

Direction des bibliothèques

AVIS

Ce document a été numérisé par la Division de la gestion des documents et des archives de l'Université de Montréal.

L'auteur a autorisé l'Université de Montréal à reproduire et diffuser, en totalité ou en partie, par quelque moyen que ce soit et sur quelque support que ce soit, et exclusivement à des fins non lucratives d'enseignement et de recherche, des copies de ce mémoire ou de cette thèse.

L'auteur et les coauteurs le cas échéant conservent la propriété du droit d'auteur et des droits moraux qui protègent ce document. Ni la thèse ou le mémoire, ni des extraits substantiels de ce document, ne doivent être imprimés ou autrement reproduits sans l'autorisation de l'auteur.

Afin de se conformer à la Loi canadienne sur la protection des renseignements personnels, quelques formulaires secondaires, coordonnées ou signatures intégrées au texte ont pu être enlevés de ce document. Bien que cela ait pu affecter la pagination, il n'y a aucun contenu manquant.

NOTICE

This document was digitized by the Records Management & Archives Division of Université de Montréal.

The author of this thesis or dissertation has granted a nonexclusive license allowing Université de Montréal to reproduce and publish the document, in part or in whole, and in any format, solely for noncommercial educational and research purposes.

The author and co-authors if applicable retain copyright ownership and moral rights in this document. Neither the whole thesis or dissertation, nor substantial extracts from it, may be printed or otherwise reproduced without the author's permission.

In compliance with the Canadian Privacy Act some supporting forms, contact information or signatures may have been removed from the document. While this may affect the document page count, it does not represent any loss of content from the document.

Université de Montréal

**Modélisation et optimisation d'un centre d'appels téléphoniques: étude
du processus d'arrivée**

par
Nabil Channouf

Département d'informatique et de recherche opérationnelle
Faculté des arts et des sciences

Thèse présentée à la Faculté des études supérieures
en vue de l'obtention du grade de Philosophiæ Doctor (Ph.D.)
en informatique

Avril, 2008

© Nabil Channouf, 2008.



Université de Montréal
Faculté des études supérieures

Cette thèse intitulée:

**Modélisation et optimisation d'un centre d'appels téléphoniques: étude
du processus d'arrivée**

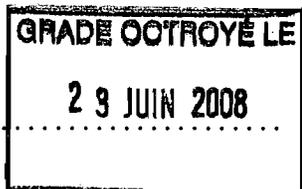
présentée par:

Nabil Channouf

a été évaluée par un jury composé des personnes suivantes:

Patrice Marcotte, président-rapporteur
Pierre L'Ecuyer, directeur de recherche
Michel Gendreau, membre du jury
Enver Yucesan, examinateur externe
Richard Duncan, représentant du doyen de la FES

Thèse acceptée le:



RÉSUMÉ

L'objectif premier de cette thèse est d'étudier le processus d'arrivée dans un centre d'appels téléphoniques et de présenter des modèles pour ce processus qui doivent bien décrire la réalité dans ces centres. Le but ultime serait de développer des méthodes d'optimisation pour la gestion des centres d'appels téléphoniques.

La gestion de ces centres est de très grande importance vu la place qu'occupent les centres d'appels dans l'activité économique des pays développés. Cette gestion est aussi complexe et elle constitue un domaine d'application intéressant pour les chercheurs dans les systèmes de files d'attente en particulier et la recherche opérationnelle en général.

Dans une première partie, nous nous intéressons à des modèles basés sur le processus de Poisson et nous exploitons une forme particulière qui permet de capter la dépendance intra-journalière. Nous proposons une amélioration de ce modèle en utilisant une estimation non paramétrique pour la fonction de taux. Ensuite, nous proposons un modèle basé sur les copules normales et nous l'adaptions au contexte des centres d'appels téléphoniques. Ces modèles, ou les copules en général, sont connus pour leur difficulté pratique dans des grandes dimensions. Dans le cas des centres d'appels, où la dimension représente le nombre de périodes de temps à considérer à l'intérieur d'une journée, nous confrontons ce même problème et nous proposons des solutions. Nous proposons aussi des méthodes avec trois algorithmes efficaces en précision et en temps de calcul pour la phase d'estimation des paramètres dans la méthode utilisée (méthode NORTA).

Nous consacrons deux chapitres pour les exemples numériques, où nous présentons une application des modèles proposés avec une analyse des données, une analyse des résultats, ainsi que des comparaisons. Nous montrons que nos modèles

peuvent fournir des meilleurs résultats que des modèles connus dans la littérature. Nous présentons aussi une simulation d'un centre de Bell Canada pour étudier la sensibilité du centre face au changement du processus d'arrivée. Nous proposons aussi une application de ces modèles aux données d'un centre d'appels d'une banque commerciale américaine.

Dans la deuxième partie de cette thèse, nous proposons des modèles de prévision pour les arrivées par jour et par intervalle de temps dans une même journée dans un centre d'appels téléphoniques. Pour les arrivées à l'intérieur d'une journée, nous proposons un modèle qui met à jour les prévisions en fonction des arrivées déjà observées au début de la journée. Nous étudions en particulier un centre d'appels au service du 911 et un des centres d'appels de Bell Canada.

Mots clés: Processus d'arrivée, centre d'appels téléphoniques, analyse de données, modélisation, simulation stochastique, étude de sensibilité, prévisions.

ABSTRACT

The main objective of this thesis is to study the arrival process to any call center and to present models for this process that describe better the reality in these centers. The ultimate goal would be to develop optimization methods for the management of call centers.

Managing call centers is of very great importance considering the place which they occupy in the economic activity of developed countries. This management is also complex and it constitutes an interesting field of application for the researchers in queuing systems in particular and operations research in general.

In a first part, we are interested in models based on the Poisson process, so we exploit a particular form of this process that captures the intra-day dependence. We propose an improvement of this model by using a nonparametric estimate for the rate function based on splines. Then, we propose a model based on normal copulas and we adapt it to the context of call centers. These kind of models or the copulas in general, are known for their practical difficulty in huge dimensions. In the case of call centers, where dimension represents the number of periods of time to consider within a day, we confront the same problem and we propose solutions. We propose also efficient methods with three accurate and fast algorithms for the step of parameters estimation of the method used (NORTA method).

We devote two chapters to numerical examples, where we present an application of the models suggested with datasets description, results analysis and also comparisons. We show that our models can provide better results than models known in the literature. We present also a simulation of a call center of Bell Canada to study the sensitivity when changing some scenarios, in particular the arrival process. We propose also an application of our models to the data from a call center

of an American commercial bank.

In the second part of this thesis, we propose forecasting models for the arrivals per day and per time intervals in the same day in a call center. For the arrivals within a day, we propose a model which allows updates of the forecasts according to the arrivals already observed at the beginning of the day. We study in particular a call center of emergency (911 calls) and a call center of Bell Canada.

Key words: Arrival process, call centers, data analysis, modeling, stochastic simulation, study of sensitivity, forecasting.

TABLE DES MATIÈRES

RÉSUMÉ	iii
ABSTRACT	v
TABLE DES MATIÈRES	vii
LISTE DES TABLEAUX	xi
LISTE DES FIGURES	xiv
LISTE DES ANNEXES	xix
LISTE DES SIGLES ET ABRÉVIATIONS	xx
REMERCIEMENTS	xxii
CHAPITRE 1 : INTRODUCTION	1
1.1 Aperçu et cadre général de la thèse	1
1.2 Le processus d'arrivée	6
1.3 Les centres d'appels téléphoniques	9
1.4 Le projet avec Bell Canada	12
1.5 Les appels téléphoniques dans une banque américaine	13
1.6 Les appels téléphoniques au service médical d'urgence	14
1.7 Logiciels et bibliothèques de simulation utilisés	14
1.8 Plan de la thèse et principales contributions	14
CHAPITRE 2 : MODÉLISATION AVEC LE PROCESSUS DE POISSON	19

2.1	Processus de Poisson et les arrivées des appels téléphoniques : quelques références	19
2.2	Estimation non paramétrique de la fonction de taux par une fonction spline	22
2.2.1	Les fonction splines	22
2.2.2	Processus de Poisson non homogène	24
2.2.3	Processus de Poisson doublement stochastique	29

CHAPITRE 3 : LES ARRIVÉES ET LA MODÉLISATION DE LA DÉPENDANCE INTRA-JOURNALIÈRE 36

3.1	Introduction	36
3.2	Méthode NORTA	40
3.2.1	Méthode NORTA avec les corrélations linéaires	41
3.2.2	Méthode NORTA avec les corrélations de rang	44
3.2.3	Estimation des paramètres des distributions marginales	45
3.3	Initialisation de la méthode NORTA	47
3.3.1	Propriétés mathématiques des corrélations	47
3.3.2	Méthode M1 : approximation en utilisant la loi normale bivariée	52
3.3.3	Méthode M2 : calcul de la dérivée	54
3.3.4	Méthode M3 : calcul de la fonction et de sa dérivée	57
3.3.5	Exemples résolus avec les 3 méthodes	58
3.3.6	Méthode M4 : Initialisation de NORTA avec un mélange de distributions discrètes et continues	63
3.4	Difficultés avec la méthode NORTA	74
3.5	Améliorations proposées pour la méthode NORTA	78

CHAPITRE 4 : LES ARRIVÉES AUX CENTRES D'APPELS DE

	BELL CANADA	82
4.1	Description des centres d'appels étudiés et analyse des données . . .	82
4.1.1	Centre C1	82
4.1.2	Centre C2	87
4.2	Application des modèles	96
4.2.1	Simulation du centre d'appels C1	96
4.2.2	Simulation des arrivées au centre d'appels C2	110
	CHAPITRE 5 : LES ARRIVÉES AU CENTRE D'APPELS D'UNE	
	BANQUE	117
5.1	Description du centre d'appels C3	117
5.2	Analyse descriptive des données	117
5.3	Application et comparaison des modèles	121
	CHAPITRE 6 : MODÈLES DE SÉRIES CHRONOLOGIQUES POUR	
	LES ARRIVÉES DANS UN CENTRE D'APPELS	127
6.1	Modèles pour les appels au service du 911	127
6.1.1	Modèles de planification pour les systèmes médicaux d'urgence dans la littérature	127
6.1.2	Description du centre d'appels étudié C4	132
6.1.3	Analyse descriptive des données	134
6.1.4	Modèles pour les arrivées par jour	136
6.1.5	Modèles pour les arrivées par intervalles de temps dans une journée	158
6.1.6	Conclusion	168
6.2	Modèles de prévision pour les arrivées au centre C2	170
6.2.1	Les arrivées par jour	170

6.2.2 Les arrivées par intervalles de temps dans une journée	176
CHAPITRE 7 : CONCLUSION	183
BIBLIOGRAPHIE	186

LISTE DES TABLEAUX

3.1	Initialisation de NORTA : résultats avec la méthode M1.	64
3.2	Initialisation de NORTA : résultats avec la méthode M2A, où $\delta = 10^{-4}$ et $2h$ est proche de 10^{-2}	65
3.3	Initialisation de NORTA : résultats avec la méthode M2B où $m = 5$, $\delta = 10^{-4}$ ($h = 0.09999$).	66
3.4	Initialisation de NORTA : résultats avec la méthode M3.	67
3.5	Initialisation de NORTA : résultats avec la méthode M4 en utilisant les corrélations de rang.	72
3.6	Initialisation de NORTA : résultats avec la méthode M4 en utilisant les corrélations linéaires.	73
3.7	Types de résolutions dans l'initialisation de NORTA selon les différents cas.	74
4.1	Nombre moyen d'appels par type de journée dans le centre C1.	84
4.2	Nombre moyen d'appels par type de journée pour l'ensemble des deux provinces au centre C2.	90
4.3	Identification des journées spéciales selon deux méthodes, pour le centre C2.	93
4.4	Estimation du taux moyen d'arrivée et des paramètres de la loi binomiale négative pour chaque demi-heure avec les données du centre C1	99
4.5	Mesures de performance obtenues par la simulation d'une journée de 8h00 à 14h00 du type mardi, mercredi ou jeudi au centre C1 avec les modèles proposés.	100

4.6	Mesures de performance obtenues par la simulation d'une journée de 8h00 à 20h30, du type mardi, mercredi ou jeudi au centre C1 avec les modèles proposés.	100
4.7	Le nombre moyen d'arrivées par demi-heure pour une journée de 8h00 à 20h30 au centre C1. Les données empiriques sont présentés avec l'écart-type "bootstrap" et sont comparées avec des données simulées (30000 jours) selon différents modèles.	105
4.8	Quelques mesures obtenues par la simulation d'une période d'une journée de 8h00 à 14h00, du type mardi, mercredi ou jeudi au centre C1 en considérant des intervalles de temps de 15 mn.	109
4.9	Estimation du taux moyen d'arrivée et des paramètres de la loi binomiale négative pour chaque demi-heure avec les données du centre C2.	113
4.10	Le nombre moyen d'arrivées par quart d'heure pour une journée de 8h00 à 20h00 au centre C2. Les données empiriques sont présentés avec l'écart-type "bootstrap" et sont comparées avec des données simulées (18000 jours) selon différents modèles.	114
5.1	Nombre moyen d'appels par type de journée dans le centre C3.	119
5.2	Estimation du taux moyen d'arrivée et des paramètres de la loi binomiale négative pour chaque intervalle de demi-heure avec les données dans le centre C3	124
5.3	Le nombre moyen d'arrivées par intervalle de temps pour une journée dans le centre C3. Les données empiriques sont comparées avec des données simulées (35750 jours) selon différents modèles.	125
6.1	Estimation des paramètres du modèle 2.	146

6.2	Estimation des paramètres du modèle 3	149
6.3	Estimation des paramètres du modèle 4	152
6.4	Estimation des paramètres du modèle 5	155
6.5	Comparaison des modèles pour les arrivées journalières	157
6.6	Erreur de prévision (RCEQM) par origine et par horizon avec les deux modèles 6 et 7 pour les arrivées par heure.	167
6.7	Résultats des estimations des paramètres et les erreurs de prévision pour le volume d'appels par jour avec le modèle où les erreurs E_t sont corrélées.	175
6.8	Erreur standard et erreurs de prévision pour le volume d'appels par jour avec le modèle où les erreurs E_t sont i.i.d.	176
6.9	Erreurs de prévisions pour le nombre d'arrivées entre 11h00 et 20h00.	180

LISTE DES FIGURES

1.1	Architecture générale d'un centre d'appels opérant en mode mixte .	10
3.1	La fonction $r_{1,2}^X(\rho)$ dans l'intervalle $[0.98, 1]$ avec le cas de la loi binomiale ($n_1 = n_2 = 3$), comparée à la fonction $r_{1,2}(\rho) = 6 \arcsin(\rho/2)/\pi$.	61
4.1	Nombre moyen d'arrivées par intervalle d'une demi-heure et par type de journée dans le centre C1.	85
4.2	Nombre moyen d'arrivées par intervalle de demi-heure pour une journée de type mardi, mercredi ou jeudi dans le centre C1.	85
4.3	Diagrammes en boîtes pour le volume d'appels par type de journée dans le centre C1.	86
4.4	Volume d'appels reçus par jour dans le centre C1.	86
4.5	Volume d'appels reçus par jour pour chaque province et pour l'ensemble dans le centre C2.	88
4.6	Nombre moyen d'arrivées par intervalle de quart d'heure et par type de journée dans le centre C2.	89
4.7	Diagrammes en boîtes pour le volume d'appels par jour pour chaque type de journée dans le centre C2.	89
4.8	Nombre moyen d'arrivées par intervalle de quart d'heure pour une journée de type mercredi, jeudi ou vendredi pour chaque province et pour l'ensemble dans le centre C2.	90
4.9	Graphique de la paire (u_1, u_2) pour tous les jours de la semaine . .	93
4.10	Diagramme de dispersion entre les arrivées avant et après 10h00 pour tous les jours de la semaine au centre C2.	95

4.11	Les résidus du modèle pour les arrivées par jour après avoir éliminé l'effet du jour de la semaine. Des observations aberrantes indiquent la présence de journées spéciales.	96
4.12	Diagnostic de la normalité des résidus pour les arrivées par jour après avoir éliminé l'effet du jour de la semaine.	97
4.13	Diagnostic de la normalité des résidus pour les arrivées par quart d'heure après avoir éliminé l'effet du jour de la semaine et de l'heure de la journée.	97
4.14	Modèles pour les corrélations $r_{i,j}^X$ dans le centre C1, (graphique pour $i = 5$ et $j = 1, \dots, 25$).	103
4.15	Fonction de taux estimée par une spline vs une fonction constante par morceaux dans le centre C1.	106
4.16	Comparaison des modèles selon les corrélations estimées entre les différentes parties de la journée dans le centre C1.	107
4.17	Comparaison des modèles selon le coefficient de variation pour chaque demi-heure dans le centre C1.	107
4.18	Modèles pour les corrélations $r_{i,j}^X$ dans le centre C2, (graphique pour $i = 10$ et $j = 1, \dots, 48$).	111
4.19	Comparaison des modèles selon le coefficient de variation pour chaque quart d'heure dans le centre C2.	115
4.20	Comparaison des modèles selon les corrélations estimées entre les différentes parties de la journée dans le centre C2.	115
5.1	Diagrammes en boîtes pour le volume d'appels par type de journée dans le centre C3.	119
5.2	Nombre moyen d'arrivées par intervalle de temps et par type de journée dans le centre C3.	120

5.3	Nombre moyen d'arrivées par intervalle de temps pour une journée de type mercredi, jeudi ou vendredi dans le centre C3.	120
5.4	Série du volume d'appels reçus par jour de mercredi à vendredi dans le centre C3.	121
5.5	Modèles pour les corrélations $r_{i,j}^X$ dans le centre C3, (graphique pour $i = 14$ et $j = 1, \dots, 35$).	123
5.6	Comparaison des modèles selon les corrélations estimées entre les différentes parties de la journée dans le centre C3.	125
5.7	Comparaison des modèles selon le coefficient de variation pour chaque intervalle de temps dans le centre C3.	126
6.1	Volume d'appels par jours pour l'année 2000 dans le centre C4, où quelques journées spéciales apparaissent.	137
6.2	Volume d'appels par mois pour toute la période dans le centre C4.	137
6.3	Volume d'appels moyen par heure pour un cycle d'une semaine dans le centre C4.	138
6.4	Diagrammes en boîtes pour le volume journalier pour chaque type de jour de la semaine dans le centre C4.	138
6.5	Diagrammes en boîtes pour le volume par type de mois de l'année dans le centre C4.	139
6.6	Les résidus E_t du modèle linéaire simple donné dans l'équation (6.2).	142
6.7	Analyse de la normalité des résidus du modèle 2	147
6.8	Analyse des corrélations entre les résidus du modèle 2	147
6.9	Analyse de la normalité des résidus du modèle 3	150
6.10	Analyse des corrélations entre les résidus du modèle 3	150
6.11	Analyse de la normalité des résidus du modèle 4	152
6.12	Analyse des corrélations entre les résidus du modèle 4	153

6.13	Erreur de prévision $RCEQM(l)$ pour les modèles de 1 à 5, avec les horizons $l = 1, \dots, 21$ jours.	158
6.14	Diagramme Quantile-Quantile pour la distribution empirique A versus la distribution empirique B pour un échantillon de $Q_{m-1,t}$ générés selon le modèle 6 sous l'hypothèse de la loi multinomiale.	162
6.15	Analyse de la normalité des résidus du modèle 7	166
6.16	Analyse des résidus pour le modèle des arrivées par jour pour la province du Québec.	172
6.17	Analyse des résidus pour le modèle des arrivées par jour pour la province de l'Ontario.	173
6.18	Analyse des résidus pour le modèle des arrivées par jour pour l'ensemble des deux provinces.	173
6.19	Prévisions et limites de confiance (95%) à un horizon d'une journée pour les arrivées par jour pour la province du Québec.	177
6.20	Prévisions et limites de confiance (95%) à un horizon d'une journée pour les arrivées par jour pour la province de l'Ontario.	177
6.21	Prévisions et limites de confiance (95%) à un horizon d'une journée pour les arrivées par jour pour l'ensemble des deux provinces.	178
6.22	Exemple de prévisions avec mise à jour et limites de confiance (95%) dans une journée du type jeudi pour les arrivées par 15 minutes entre 11h00 et 20h00, pour la province du Québec.	181
6.23	Exemple de prévisions avec mise à jour et limites de confiance (95%) dans une journée du type jeudi pour les arrivées par 15 minutes entre 11h00 et 20h00, pour la province de l'Ontario.	181

6.24 Exemple de prévisions avec mise à jour et limites de confiance (95%) dans une journée du type jeudi pour les arrivées par 15 minutes entre 11h00 et 20h00, pour l'ensemble des deux provinces.	182
---	-----

LISTE DES ANNEXES

- Annexe I :** Classes Java pour l'initialisation de NORTA avec les corrélations de rang et des distributions marginales discrètes i
- Annexe II :** Classes Java pour l'initialisation de NORTA avec les corrélations de rang et un mélange de distributions marginales continues et discrètes xxxii

LISTE DES SIGLES ET ABRÉVIATIONS

- $\lambda(t)$ fonction du taux d'arrivée des appels en fonction du temps t .
- W facteur aléatoire associé à une journée qui multiplie la fonction de taux d'arrivée.
- $s(t)$ fonction spline.
- B^p opérateur de différence d'ordre p ; $B^p X_t = X_{t-p}$.
- ∇_p^d opérateur de différence d'ordre p ;
 $\nabla_p^d X_t = (I - B^p)^d X_t = (X_t - X_{t-p})^d$.
- Δ_i $x_{i+1} - x_i$.
- R_ρ^X matrice de corrélations linéaires du vecteur X .
- R_r^X matrice de corrélations de rang du vecteur X .
- $\rho_{i,j}^X$ $\text{Corr}(X_i, X_j)$.
- $r_{i,j}^X$ $\text{Corr}(F_i(X_i), F_j(X_j))$, où F_i et F_j sont les distributions marginales de X_i et X_j , respectivement.
- V^t transposée du vecteur V .
- $\text{diag}(V)$ matrice carrée avec les éléments du vecteur V sur sa diagonale et 0 pour ses autres éléments.
- C^t transposée de la matrice C .

- AIC Critère d'Information d'Akaike.
- AR AutoRégressif.
- ARIMA "Autoregressive Integrated Moving Average".
- CRSNG Conseil de Recherches en Sciences Naturelles et en Génie.
- DVS Décomposition en Valeurs Singulières.
- EARM Erreur Absolue Relative Moyenne.
- EIQ Étendue Inter-Quartiles.
- i.i.d. indépendants et identiquement distribués.
- IVR "Integrated Voice Response".
- MC Monte Carlo.
- NORTA "NORmal To Anything".
- PPNH Processus de Poisson Non Homogène.
- PPDS Processus de Poisson Doublement Stochastique.
- QMC Quasi Monte Carlo.
- RCEQM Racine Carrée de l'Erreur Quadratique Moyenne.
- SMU Système Médical d'Urgence.
- VRU "Voice Response Unit".

REMERCIEMENTS

Je dois remercier infiniment mon directeur de recherche, Pierre L'Ecuyer. Il n'a cessé de me donner des conseils précieux et des critiques très éclairées. Il m'a aussi suggéré plusieurs idées qui ont permis de réaliser cet ouvrage. Il a montré beaucoup de patience lors des moments difficiles que j'ai vécus pendant ce long parcours. En fin, je dois le remercier beaucoup pour son soutien financier pendant toute la période de la thèse. Par la même occasion, je remercie Bell Canada qui, par le financement du projet, a permis de financer mes études.

Je remercie de tout mon cœur ma mère Mabrouka, ainsi que mes frères et sœurs Ramzi, Amel, Tarek et Sihem pour leur soutien continu. Je remercie mon épouse, Manel que j'ai rencontrée durant la période de ma thèse. Elle a montré beaucoup de patience et elle m'a fourni un soutien de valeur, ainsi que ses parents Amor et Fatma. Une fois de plus je dédie cette thèse à l'âme de mon père que j'ai perdu pendant mes études antérieures.

J'exprime toute ma gratitude à Hatem Ben-Ameur pour son grand soutien et ses encouragements. Merci aussi à mes amis Abdeljelil Farhat, Habib Korbi, Mohamed Ayadi et Samir Trabelsi.

Je remercie mes collègues du département, Alexandre Deslauriers, Amina Lamghari, Richard Simard, Eric Buist et Wyeon Chan.

Je remercie les co-auteurs des deux articles, Armann Ingolfsson et Athanasios Avramidis. Je remercie également les responsables de Bell Canada, du service médical d'urgence de la ville de Calgary, ainsi que Avi Mandelbaum pour les données qu'ils nous ont fournies sur les centres d'appels étudiés.

Finalement, je remercie les membres du jury pour avoir examiné cette thèse et pour m'avoir donné des commentaires et des suggestions pertinentes.

CHAPITRE 1

INTRODUCTION

1.1 Aperçu et cadre général de la thèse

Lorsqu'une personne compose un numéro de téléphone d'un centre d'appels pour solliciter un service quelconque, la première question qui lui passe par l'esprit est : "est ce que je vais attendre beaucoup avant d'être servi si jamais j'arrive à joindre le centre d'appels?". Le meilleur scénario serait d'avoir directement une personne au bout de la ligne au lieu d'un répondeur automatique. En pratique, aucun centre d'appels n'est capable de satisfaire les clients de la sorte. C'est de là qu'un système de file d'attente s'est imposé pour créer un certain équilibre entre des gestionnaires de centres qui veulent fournir un bon service à moindre coût d'un côté, et des clients qui veulent être servis sans avoir à attendre beaucoup de l'autre côté.

Depuis le début du 20ème siècle, les chercheurs n'ont cessé de travailler pour construire des modèles mathématiques pour les systèmes de files d'attente, y compris ceux des centres d'appels. Ces modèles ont contribué au développement de plusieurs méthodes d'optimisation de la gestion de ces centres.

Cette thèse a été réalisée dans le cadre du projet *Simulation et optimisation de la gestion des centres d'appels téléphoniques*, commandé par Bell Canada et financé conjointement avec le conseil de recherches en sciences naturelles et en génie du Canada (CRSNG). Ce projet a pour but de développer de nouvelles méthodes, ainsi que des logiciels pour la simulation et l'optimisation de la gestion des centres d'appels téléphoniques. Ces centres peuvent opérer en mode mixte (ou "blend"), c'est-à-dire un mélange d'appels entrants et d'appels sortants ou seulement avec

des appels entrants.

Ce projet touche à une branche d'activité qui ne cesse de prendre de l'ampleur en Amérique du Nord. En terme d'emplois, on estime que le nombre d'employés dans les centres d'appels aux États-Unis dépasse 3.5 millions, ce qui représente plus de 2.5% de la population active [90]. Ce taux avoisine 3.4% au Canada selon un rapport publié par le ministère canadien des ressources humaines en 2002. Ces centres injectent entre 36 et 38 millions de dollars canadiens dans le produit intérieur brut (P.I.B.) et ce secteur connaît un taux d'accroissement annuel moyen de près de 20%. Selon le même rapport, plus d'un demi million de personnes travaillent dans plus de 1400 centres à travers tout le pays. Les plus gros centres de Bell emploient jusqu'à 1500 agents. Au début des années 2000, le Canada est même devenu un marché attrayant pour les compagnies américaines à la recherche d'un emplacement pour des centre d'appels. A cette époque, ces compagnies pouvaient maintenir leurs frais de gestion bas en raison de la vigueur du dollar américain et aussi grâce à une main d'oeuvre de qualité. Cependant, on remarque ces dernières années l'émergence de nouveaux marchés pour l'emplacement des centres d'appels, comme l'Inde et la Chine pour les États-Unis et l'Afrique du Nord pour la France [70]. Ces centres deviennent aussi une source incontournable de profits pour plusieurs compagnies modernes qui fournissent des services. Ceci justifie largement l'intérêt à optimiser la gestion de ces centres.

Le fonctionnement des centres d'appels varie d'une compagnie à l'autre. Dans un des centres de Bell par exemple, du lundi au samedi, les clients appellent tout au long de la journée pour des produits et des services de Bell Canada, Bell Mobilité, Bell ExpressVu ou Bell Sympatico. Les agents peuvent à leur tour appeler des clients chez eux pour offrir les mêmes biens et services ou pour d'autres raisons comme la facturation.

Les arrivées des appels des clients, le temps alloué pour les servir ainsi que les agents disponibles dans le centre et le mécanisme d'attente constituent un système de file d'attente. La première composante a fait l'objet principal de nos travaux.

La modélisation du processus d'arrivée est une tâche assez difficile si nous voulons adopter des hypothèses réalistes. En effet, ce processus a un caractère hautement stochastique et le taux d'arrivée des appels varie substantiellement au cours d'une même journée, ainsi que d'une journée à l'autre, selon plusieurs facteurs. Nous devons donc construire des modèles avec des règles et des hypothèses qui décrivent bien la réalité des centres d'appels.

Dans la littérature, la majorité des travaux font toujours référence aux fameux processus de Poisson. Ce processus est caractérisé essentiellement par la propriété de l'indépendance entre les arrivées dans deux intervalles de temps disjoints. Cependant, nous allons voir plus loin dans ce document que si cette caractéristique n'est pas vérifiée dans un centre en particulier, ce processus peut quand même être utilisé sous une forme particulière (processus de Cox).

Dans cette section nous faisons un survol de la littérature où nous présentons quelques références qui ont étudié le processus d'arrivée des appels téléphoniques. Cela permettra de situer nos travaux et la contribution de nos modèles dans l'ensemble des modèles présentés dans ce domaine de recherche. Plus tard et au niveau de chaque chapitre, nous reviendrons plus en détails sur quelques références et des modèles en particulier qui nous permettent parfois de faire quelques comparaisons. Comme modèles descriptifs pour le processus d'arrivée, nous pouvons citer Buffa et al. [14], qui ont étudié le volume d'appels reçus selon deux unités de temps : le mois et le jour. Le modèle pour les arrivées par jour est plus utile dans la planification du nombre global d'agents nécessaire pour travailler dans le centre, qui peut varier par exemple selon la saison, alors que le modèle pour le volume d'appels

reçus par jour sert plus dans l'affectation quotidienne des agents. Dans la même catégorie, le travail de Mandelbaum et al. [63] a présenté une étude descriptive du processus d'arrivée. Ces auteurs ont étudié la variabilité du taux d'arrivée sa nature stochastique qui permettent de faire des études prédictives. Parmi les études détaillées sur le processus d'arrivée dans les centres d'appels téléphoniques figure celui de Brown et al. [13], Gans et al. [31], où on peut trouver une description du fonctionnement des centres d'appels, ainsi que des différentes composantes stochastiques du système de file d'attente que constitue tout centre d'appels (les arrivées, la durée du service, les abandons, le temps d'attente, ...). Pour le processus d'arrivée en particulier, ils ont proposé une méthode pour les prévisions du taux d'arrivée au cours d'une journée basée sur les observations d'une journée antérieure.

Dans le contexte des prévisions du volume d'appels, Andrews et Cunningham [3] et Bianchi et al. [8] ont proposé des modèles basés sur les processus ARIMA ("AutoRegressive Integrated Moving Average") pour des prévisions en tenant compte de variables exogènes comme la saison, le jour de la semaine, les jours fériés ou les jours qui suivent l'envoi des catalogues publicitaires dans le cas de centres d'appels spécialisés dans les ventes. Dans ce même contexte, nous proposons des modèles simples pour les prévisions du nombre d'arrivées dans un centre d'appels qui tiennent compte du même type de variables. Nos modèles donnent de meilleurs résultats comparativement à d'autres modèles basés par exemple sur les processus ARIMA.

Pour ce qui est des modèles basés sur le processus de Poisson, Massey et al. [64] ont proposé un modèle où la fonction de taux est estimée par une fonction linéaire par morceaux et chaque morceau correspond à un intervalle de temps dans une journée. Jongbloed et Koole [46] ont quant à eux proposé un modèle où le taux d'arrivée est stochastique dans chaque période de temps et est indépendant

des autres périodes. Ceci donne lieu au processus de Poisson doublement stochastique (ou processus de Cox). D'autres auteurs ont montré que ce processus peut être mieux exploité pour décrire la réalité dans les centres d'appels. Effectivement, Avramidis et al. [5] ont montré que l'hypothèse de l'indépendance entre les arrivées dans différentes parties de la journée n'est pas toujours vérifiée dans les centres d'appels. Parmi les modèles présentés, on retrouve celui basé sur le processus de Cox avec une forme particulière qui peut capter cette dépendance, où la fonction de taux est considérée constante par morceau, mais elle est multipliée par une composante stochastique qui varie d'une journée à l'autre.

Dans cette même catégorie, nous présentons ici des modèles qui tiennent compte de la dépendance entre les arrivées dans différentes parties de la journée et qui permettent d'avoir la quantité de corrélation désirée entre le nombre d'arrivées dans deux périodes de temps différentes dans une même journée. Nous proposons aussi d'utiliser une fonction spline pour la fonction de taux dans un processus de Poisson doublement stochastique au lieu d'une fonction constante par morceau, ce qui décrira mieux la nature variable de ce taux pendant la journée.

En plus de la grande importance du processus d'arrivée dans le système de file d'attente que constitue tout centre d'appels, la disponibilité des données sur plusieurs centres d'appels constitue une bonne motivation pour nos travaux. Ces données ont permis de tester et valider nos modèles. Effectivement, les différentes modélisations que nous proposons sont basées sur des données statistiques fournies par Bell Canada et qui concernent deux centres d'appels différents. Nous disposons aussi de données sur les arrivées dans un centre d'appels d'une grande banque commerciale américaine, ainsi que des données des arrivées des appels au service du 911 dans la ville de Calgary, en Alberta.

1.2 Le processus d'arrivée

Les appels téléphoniques arrivent durant les heures d'ouverture du centre de façon aléatoire. On identifie alors le "processus" d'arrivée qui enregistre l'instant où chaque appel aboutit. Dans la majorité de centres, les appels arrivent d'abord à un système automatisé où une bonne partie de ces appels sollicitent le service d'un préposé. L'appel est pris en charge par le premier agent libre, sinon il est placé dans une file d'attente si le système le permet.

Pour la notation, on désigne par $\{X(t), t \geq 0\}$ le processus d'arrivée, où $X(t)$ est le nombre d'arrivées dans l'intervalle $[0, t)$. On pourra noter aussi $X(t, s) = X(t + s) - X(s)$, le nombre d'arrivées dans l'intervalle $[t, t + s)$. Dans la suite du travail, on note par $X_i = X(t_{i-1}, t_i) = X(t_i) - X(t_{i-1})$, pour $i = 1, \dots, d$, le nombre d'appels reçus dans l'intervalle de temps $[t_{i-1}, t_i)$, pour $i = 1, \dots, d$, où les instants $0 = t_0 < t_1 < \dots < t_d$, représentent les débuts des périodes qui partagent la journée en d intervalles de temps.

Le processus d'arrivée peut être étudié selon plusieurs niveaux dépendamment des hypothèses adoptées et aussi du but de l'étude. On peut étudier les instants d'arrivées ou les intervalles inter-arrivées ou encore le nombre d'arrivées par unité de temps. Cette unité peut être l'année, le mois, le jour ou l'heure. Cependant pour des fins de prévisions ou des prise de décision liées au nombre et aux horaires des agents, il est plus utile d'étudier le processus des arrivées par jour ou par intervalles de temps dans une même journée.

Dans plusieurs situations, on commence par vérifier si le processus de Poisson peut bien décrire les données disponibles. Ce processus est caractérisé par sa fonction de taux (ou intensité). Si on suppose que les appels dans des intervalles disjoints de la journée arrivent indépendamment d'un intervalle à l'autre et que le taux est stationnaire sur toute la journée, on peut modéliser par le processus

de Poisson à taux homogène et la fonction de taux est donc une constante. Par ailleurs, si on remarque une variabilité considérable tout au long de la journée, on peut adopter une fonction variable pour le taux d'arrivée et donc un processus de Poisson non homogène.

Le processus de Poisson homogène est utilisé dans plusieurs systèmes de files d'attente, puisque l'hypothèse de la stationnarité permet de calculer plusieurs mesures de performance sur horizon infini. Avec un taux constant, on peut parler par exemple de l'utilité de ce processus dans les modèles d'Erlang (A, B et C). Dans le cas du processus de Poisson non homogène, la moyenne de la fonction de taux permet aussi de calculer des mesures de performance [35]. Sur un horizon fini, les formules données par les modèles d'Erlang ne permettent de fournir que des approximations pour les mesures de performance.

Dans la majorité des centres d'appels, les données recueillies sont généralement agrégées par tranches horaires (quart d'heure ou demi-heure par exemple). Pour cela, le processus de Poisson non homogène avec une fonction de taux linéaire par morceaux, ou constante pour chaque intervalle $[t_{i-1}, t_i)$, pour $i = 1, \dots, d$, s'adapte bien à une telle situation [58, 64]. La fonction de taux peut aussi prendre une forme cyclique (trigonométrique), polynomiale ou exponentielle et parfois une combinaison de ces formes [36, 59].

Dans d'autres situations, la fonction de taux peut être elle-même stochastique et on parle alors du processus de Poisson doublement stochastique ou processus de Cox. Conditionnellement à une réalisation de la fonction de taux, on retrouve le processus de Poisson homogène dans le cas où cette fonction est une variable aléatoire à valeur réelle ou le processus de Poisson non homogène si cette fonction est un processus stochastique.

Dans plusieurs situations, le processus de Poisson s'avère non réaliste, puisque

l'hypothèse que les arrivées dans des intervalles de temps disjoints sont indépendants, n'est pas toujours vérifiée. Whitt [93] a utilisé le modèle de Cox en proposant une forme particulière qui permet de capter la dépendance entre les arrivées dans différentes parties de la journée. Notre point de départ pour le modèle que nous proposons est le travail d'Avramidis et al. [5] et plus précisément le modèle qu'ils ont baptisé M1. L'idée principale pour le choix de cette forme est qu'elle permet d'introduire de la corrélation entre les arrivées d'un intervalle à un autre en multipliant la fonction de taux déterministe par une composante stochastique.

Une autre alternative poursuivie dans cette thèse est de modéliser directement les arrivées par une distribution multidimensionnelle pour le vecteur (X_1, \dots, X_d) , qui s'ajuste bien aux données et qui tient compte de la structure de dépendance entre les différentes périodes de la journée.

Pour mesurer la dépendance, nous définissons alors la matrice de corrélations R_ρ^X où chaque élément, noté $\rho_{i,j}^X$, pour $1 \leq i < j \leq d$, représente le coefficient de corrélation linéaire (ou coefficient de Pearson) entre X_i et X_j . Nous avons considéré aussi le coefficient de corrélation de rang (ou coefficient de Spearman), noté $r_{i,j}^X$, pour $1 \leq i < j \leq d$ et la matrice est notée R_r^X .

Il y a plusieurs autres approches pour capter la dépendance ; on cite par exemple les fonctions copules et les méthodes NORTA et VARTA, introduites respectivement par Cario et Nelson [18] et Deler et Nelson [24]. On cite aussi la méthode d'estimation non paramétrique des densités conjointes par une méthode de noyau, en utilisant des données empiriques [42].

Dans notre travail, nous présentons des modèles pour les arrivées qui décrivent bien la nature des centres d'appels étudiés et qui tiennent compte d'hypothèses plus réalistes que celles utilisées dans d'autres modèles qui ne tiennent pas compte de la dépendance entre les arrivées à l'intérieur d'une même journée. La disponibilité

des données pour quatre centres d'appels différents a motivé ces recherches.

1.3 Les centres d'appels téléphoniques

On peut définir un centre d'appels comme étant un groupe de personnes qui ont comme tâche principale de répondre aux appels téléphoniques de clients actuels ou potentiels [66]. En général, les appels de ces clients aboutissent à un système automatique, appelé "Interactive Voice Response" (IVR), qui les reçoit pour les acheminer par la suite vers les agents selon le type de service sollicité par les clients. Le centre dispose d'un système de files qui regroupe les appels en attente de service par un agent libre. Les appels peuvent alors être classés selon plusieurs types avec des taux d'arrivée différents. Les agents peuvent aussi être classés en différents types selon le type d'appels qu'ils peuvent prendre en charge. Dans la suite de nos travaux, nous nous intéressons seulement à l'étude du processus d'arrivée et par conséquent nous tenons compte seulement des appels entrants sans nous soucier du type d'agent. Dans la figure 1.1, nous présentons un schéma simplifié d'un centre d'appels en mode mixte.

Dans le système de file d'attente d'un centre d'appels, nous pouvons observer le phénomène d'abandon qui est dû essentiellement à une attente relativement longue et aussi à l'impatience des clients. Celle ci varie selon le comportement de chaque client.

Dans le cas d'un centre d'appels opérant en mode mixte, les appels sortants sont initiés à partir d'un composeur automatique qui agit selon des règles bien précises fixées par les responsables du centre et qui dépendent généralement de facteurs comme le flux des appels entrants, l'occupation des agents et la qualité de service des temps antérieurs.

Les responsables du centre d'appels veillent à augmenter sa productivité et à

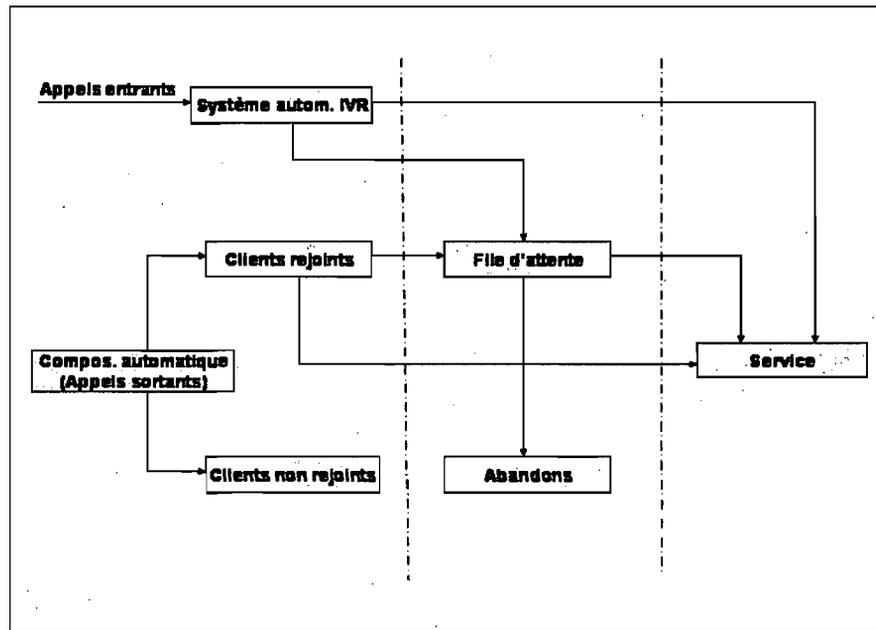


Figure 1.1 – Architecture générale d'un centre d'appels opérant en mode mixte

optimiser sa gestion tout en respectant un certain niveau de service. En général, ils veillent à minimiser un coût engendré en grande partie par le salaire des agents, en assurant une bonne occupation de ces derniers et respectant une règle pour la qualité de service. La règle (α, s) , souvent la plus utilisée, est définie comme suit : $\alpha\%$ des clients doivent attendre moins de s secondes. Des valeurs de paramètres fréquemment utilisées sont $(\alpha = 80\%, s = 20)$ et aussi $(\alpha = 70\%, s = 30)$. D'autres mesures, comme le taux d'abandon et le temps moyen d'attente, constituent aussi des indicateurs de qualité de service. On pourra alors imposer des contraintes sous forme de bornes supérieures sur le taux d'abandon ou le temps moyen d'attente.

Plusieurs modèles de gestion optimale des centres d'appels sont proposés dans la littérature. Ils cherchent à maximiser un profit ou minimiser des coûts sous des contraintes budgétaires ou de qualité de service. Les critères de décision varient selon la stratégie choisie par les gestionnaires du centre. La gestion du personnel

constitue la composante principale dans la gestion d'un centre d'appels, puisque la majeure partie des coûts est sous forme de salaires. En général, les politiques de gestion diffèrent d'un centre à un autre, mais l'objectif principal est le même, soit maximiser la performance en respectant une certaine qualité de service. Pour définir les caractéristiques d'une politique de gestion, il faut d'abord étudier le centre d'appels, puis établir un modèle mathématique qui le décrit convenablement.

Les mesures de performance d'un système de file d'attente nous permettent d'identifier et de quantifier les objectifs et les contraintes qui définissent différents programmes d'optimisation. Ces mesures peuvent être obtenues ou estimées par simulation ou analytiquement en identifiant un modèle, comme par exemple la formule d'Erlang ou les chaînes de Markov à temps continu [25, 71]. Dans ce cas, le modèle mathématique nous permet de calculer analytiquement ou numériquement, dépendamment du modèle et des hypothèses, les différentes mesures désirées. Avec la simulation, on peut estimer les différentes mesures de performance, comme on peut étudier la sensibilité du centre d'appels suite à des changements de scénarios de gestion, comme par exemple les changements du nombre d'agents par période de temps.

Dans cette thèse, nous définissons les mesures de performance suivantes : la qualité de service, l'occupation des agents, le nombre d'appels servis, le nombre d'abandons, le temps moyen d'attente dans la file et le temps moyen passé dans le système. La qualité de service sera mesurée par le pourcentage des clients qui ont attendu moins de 20 secondes avant d'être servis. L'occupation des agents est mesurée par le rapport entre la durée totale où les agents sont occupés et la durée totale de leur disponibilité. Le temps moyen d'attente dans la file concerne seulement les appels ayant attendu, alors que le temps moyen dans le système concerne tous les appels entrés dans le système y compris les abandons.

1.4 Le projet avec Bell Canada

Notre projet de recherche avec Bell Canada a démarré en 2001 et il vise à développer des méthodes pour optimiser les politiques de gestion des centres d'appels de Bell Canada. La source des données utilisées pour les différentes modélisations sont deux centres d'appels téléphoniques de Bell.

Le premier est un centre qui opère en mode mixte qui était ouvert de 8h00 à 20h30 la semaine et de 8h30 à 17h00 le samedi et fermé le dimanche. Les appels proviennent des provinces de Québec et de l'Ontario. Près de 1200 appels en moyenne arrivent les jours de la semaine et près de 350 le samedi. Le lundi présente un taux relativement plus haut par rapport aux autres journées. Le nombre d'agents disponibles dans ce centre varie selon le jour et les heures et ils sont environ 85 le lundi et 70 les autres jours. Seulement une proportion des ces agents sont du type mixte. Les appels sortants sont effectués à l'aide d'un composeur automatique, qui peut composer jusqu'à 30 numéros à la fois. La politique de composition dépend généralement du nombre d'agents libres et de la distribution du temps d'attente par clients pour les temps antérieurs. Si un client est rejoint (appel sortant) et qu'entre temps tous les agents sont devenus occupés, il est considéré comme un "mismatch" [25]. Pour ce centre, nous disposons de tous les ingrédients pour faire la simulation. On parle ici de données sur les arrivées, le temps de service, les abandons, le nombre d'agents, la politique de composition... Il est bon de rappeler ici que les données concernant ce centre ont été fournies au début du projet et que depuis un certain temps ce centre est fermé.

Un second lot de données, plus récentes, concerne un grand centre d'appels avec une plus grande capacité et qui est ouvert tous les jours de la semaine de 8h00 à 20h00. Les appels proviennent également de deux provinces, le Québec et l'Ontario. Pour ce centre, nous ne disposons que des données sur les appels entrants, avec des

détails sur l'origine de l'appel et le numéro de téléphone pour lequel le service est sollicité. Le volume d'appels diffère d'un jour à l'autre (voir section 4.1.2), samedi et dimanche enregistrant beaucoup moins d'appels que les autres jours. En gros, près de 52 000 appels arrivent par jour, dont près du tiers proviennent de la province du Québec. Dans la suite des travaux, nous appelons les deux centres d'appels de Bell, respectivement, centres C1 et C2.

1.5 Les appels téléphoniques dans une banque américaine

Les données disponibles proviennent d'un centre d'appels d'une grande banque commerciale américaine et les appels concernent des clients qui résident dans les états de New York, Pennsylvanie, Rhode Island et Massachussets. Ce centre reçoit jusqu'à 300 000 appels par jour, qui aboutissent au système automatique appelé "Voice Response Unit" (VRU) pour solliciter des services bancaires. Approximativement seulement 20% des clients qui appellent cherchent à parler à un agent. Les données disponibles concernent cette proportion des appels. Ce centre est ouvert 24 heures, les 7 jours de la semaine et il met à la disposition des clients entre 900 et 1 200 agents la semaine et entre 200 et 500 le samedi et dimanche. Les données couvrent une période allant de mars 2001 à avril 2003 [88]. Nous présentons dans le chapitre 5 une analyse descriptive des données de ce centre que nous appelons C3.

1.6 Les appels téléphoniques au service médical d'urgence

Le projet consiste à développer des modèles de prévision pour les appels téléphoniques au service médical d'urgence (911) dans la ville de Calgary, en Alberta. Les appels reçus sont analysés à deux niveaux : appels reçus toutes les heures et appels journaliers. Le département du système médical d'urgence de la ville de Calgary nous

a fourni toutes les données sur les arrivées des appels au service du 911 pour un total de 1537 jours entre 2001 et 2004 [19]. Nous appelons ce centre C4.

1.7 Logiciels et bibliothèques de simulation utilisés

Pour la simulation des centres d'appels, nous avons utilisé les deux bibliothèques logicielles de simulation en Java : *SSJ* [57] et *ContactCenters* [15]. Ces bibliothèques sont développées sous la direction du professeur Pierre L'Ecuyer au département d'informatique et de recherche opérationnelle de l'Université de Montréal. Pour l'analyse et l'exploitation des données, ainsi que les estimations et tous les calculs statistiques, nous avons utilisé les logiciels R et SAS.

1.8 Plan de la thèse et principales contributions

Nous présentons dans le deuxième chapitre la méthodologie utilisée dans la modélisation des arrivées avec le processus de Poisson non homogène et le processus de Poisson doublement stochastique (processus de Cox). Nous présentons quelques références de travaux qui se sont basés sur le processus de Poisson pour la modélisation des arrivées, particulièrement dans le contexte des centres d'appels téléphoniques. La majorité de ces travaux suppose donc l'indépendance entre le nombre d'arrivées d'un intervalle à un autre, qui est en fait une caractéristique du processus de Poisson (homogène ou non homogène). Cette hypothèse semble être de grande importance pour nous, surtout dans le contexte des centres d'appels téléphoniques. Nous verrons dans les exemples numériques que tous les centres étudiés confirment la présence de corrélations entre les arrivées dans différentes parties d'une même journée. Nous avons alors repris une forme particulière du processus de Poisson doublement stochastique, utilisée par Avramidis et al. [5], et qui permet de capter cette dépendance. Nous avons ajouté à cela un modèle non

paramétrique pour la fonction de taux par une fonction spline qui a amélioré les modèles précédents, dans lesquels la fonction de taux est estimée par une fonction constante par morceaux.

Une contribution de ce chapitre est de présenter un nouveau modèle réaliste pour la fonction de taux des arrivées des appels téléphoniques dans un processus de Poisson doublement stochastique. Cette combinaison permet de capter les variations du taux d'arrivée avec une fonction continue qui est la fonction spline, et en même temps de pouvoir capter la dépendance entre les arrivées des différentes parties de la journée en considérant une forme particulière du processus de Poisson doublement stochastique. Avec cette nouvelle modélisation, nous présentons une méthode d'estimation basée sur le maximum de vraisemblance qui permet d'estimer tous les paramètres du processus, y compris ceux de la fonction spline.

Nous abordons dans le troisième chapitre le problème de dépendance intra-journalière et nous proposons à l'occasion une revue de littérature. Nous voulions ici présenter une méthode efficace basée sur les copules normales. Cette méthode, appelée méthode NORTA, présente des avantages, comme par exemple le fait qu'elle permet de générer des valeurs des variables aléatoires étudiées avec les mêmes quantités de corrélations désirées et elle permet aussi aux distributions marginales de ces variables d'avoir n'importe quelle forme. Elles peuvent être discrètes, continues ou un mélange des deux. La principale contribution de ce chapitre est l'adaptation de la méthode NORTA au contexte particulier des centres d'appels téléphoniques où la dimension (nombre d'intervalles de temps à considérer dans une journée) est assez importante. Il est connu que cette méthode, ou les copules en général, affichent quelques difficultés dans des grandes dimensions, comme par exemple la difficulté d'obtenir une matrice de corrélations semi-définie positive. Parmi les solutions, nous proposons des modèles pour les corrélations qui permettent de réduire

considérablement le nombre de paramètres à estimer et en même temps de résoudre le problème de la matrice de corrélations.

Une autre contribution concerne l'initialisation de la méthode NORTA, qui représente l'étape d'estimation et qu'on définira en détails par la suite. La revue que nous présenterons montre que cette étape est très peu étudiée dans le cas où les distributions marginales sont discrètes. A notre connaissance, une seule méthode basée sur une heuristique est présentée par Cario et Nelson [18]. Nous proposons ici trois méthodes efficaces. Pour le cas mixte, c.-à.-d. un mélange de distributions discrètes et continue, aucune méthode n'a été proposée dans la littérature et nous présentons une méthode avec plusieurs exemples numériques.

Au chapitre 4, nous présentons les résultats numériques des applications dans les deux centres d'appels de Bell Canada. Dans le premier cas, nous effectuons la simulation du centre d'appels avec les différents modèles, puisque nous disposons de toutes les données pour modéliser les autres inputs (temps de service, temps de patience, politique de composition des appels sortants...). Bien que la simulation d'un centre d'appels ne soit pas le meilleur critère pour comparer nos modèles pour les arrivées avec d'autres modèles proposés dans la littérature, elle permet tout au moins de valider nos modèles quand les mesures de performance obtenues sont proches des mesures calculées avec les données. Pour cette raison, nous fournissons aussi une étude comparative des modèles pour le processus d'arrivée où nous analysons la corrélation et la variabilité du nombre d'arrivées dans différents intervalles de temps avec des journées simulées. Nous présentons dans ce chapitre des études descriptives des centres étudiés, ainsi que des données fournies. Nous présentons donc une validation des modèles proposés dans les chapitres 2 et 3 avec les données de deux centres d'appels téléphoniques de Bell Canada.

Le chapitre 5 concerne les données du centre d'appels de la banque américaine.

Après une description et une analyse statistique descriptive des données recueillies, nous présentons une application des modèles proposés avec une étude comparative. Ce centre présente la particularité d'être ouvert pendant les 24 heures de la journée avec des agents qui travaillent en permanence. Le taux d'arrivée durant le jour est très élevé dans ce centre et nous verrons que même dans un tel cas, nos modèles fournissent des bons résultats. Nous voulions que les données de ce centre permettent une application et une validation de plus pour nos modèles.

Au chapitre 6, nous présentons un ensemble de modèles de séries chronologiques pour les arrivées par jour et par heure d'une même journée. Nous comparons les différents modèles à l'aide de l'erreur d'estimation et des performances des prévisions. Nous présentons également des modèles de prévision à court terme à l'intérieur d'une journée. Nous voulions présenter des modèles simples et qui soient plus adaptés aux centres d'appels téléphoniques. En effet, nous avons remarqué que tous les centres étudiés présentent des caractéristiques communes. Nous avons remarqué la présence de l'effet du jour de la semaine, du mois de l'année, les journées spéciales, ainsi que de corrélations entre les arrivées durant des journées successives. Parmi nos modèles, nous avons sélectionné celui qui est parcimonieux et qui est plus performant de point de vue estimation et prévision. Parmi les contributions de ce chapitre, on présente des modèles simples adaptés aux centres d'appels. Nous avons constaté que ces modèles donnent de meilleurs résultats que d'autres modèles populaires présentés dans la littérature. Nous avons proposé aussi des modèles pour les arrivées par tranches horaires basés sur l'observation du modèle pour les arrivées par jour. Ces modèles permettent de faire des prévisions du taux d'arrivée pendant une journée, sachant le volume global et en tenant compte de l'heure de la journée. Une autre contribution est de proposer un modèle basé sur la distribution conditionnelle des arrivées par heure, sachant le volume total de la journée. Ce modèle

permet de mettre à jour les prévisions chaque fois qu'on avance dans la journée et qu'une partie de cette journée est déjà observée. Nous verrons que ce modèle fournit des bons résultats avec deux centres étudiés.

Comme application de ces modèles, nous considérons les données du centre d'appels au service médical d'urgence dans la ville de Calgary, en Alberta. Nous fournissons alors une description détaillée des données recueillis et une application des modèles de prévisions proposés pour les arrivées par jour et par heure à l'intérieur d'une même journée en se basant sur un des modèles pour les arrivées par jour. Nous présentons aussi une deuxième application des modèles proposés avec les données du centre C2 de Bell Canada.

Finalement, au chapitre 7, nous présentons les conclusions avec des perspectives de travail futur.

Il faut noter qu'une partie de ces travaux a permis la publication de deux articles [4, 19] et ils vont éventuellement permettre des extensions futures.

étudié. Comme suite à ce travail, Avramidis et al. [5] ont proposé une variété de modèles pour le processus d'arrivées en tenant compte de la dépendance entre les demi-heures de la journée.

Massey et al. [64] ont ajusté un processus de Poisson non homogène avec fonction de taux linéaire par morceaux du fait que l'estimation est faite sur la base de données agrégées. Green et Kolesar [36] ont modélisé les arrivées par un processus de Poisson avec fonction de taux sinusoïdale dans un modèle Markovien. Ils ont adopté l'hypothèse de la non stationnarité et ont introduit une méthode pour l'estimation des mesures de performance du système dans un tel contexte, qu'ils ont baptisée la "Pointwise Stationary Approximation". Jongbloed et Koole [46] ont adopté le modèle de Cox et ont proposé un processus doublement stochastique, mais les taux d'arrivées de chaque tranche horaire sont modélisées indépendamment les unes des autres. Lewis et Shedler [59] ont utilisé une fonction de taux du type exponentiel et ont estimé ses paramètres par maximum de vraisemblance en utilisant des données observées sur les instants d'arrivées dans un centre d'appels. D'autres auteurs ont utilisé des méthodes non paramétriques pour estimer la fonction de taux. Dans ce modèle, le taux d'arrivée dans chaque intervalle de temps suit une loi gamma, ce qui donne lieu à une loi binomiale négative pour la distribution inconditionnelle du nombre d'arrivées dans cet intervalle.

Kuhl et Bhairgond [50] ont utilisé les fonctions ondelettes, alors que Kuhl et Damerджи [51] ont présenté des méthodes basées sur des heuristiques pour estimer la densité et la fonction de taux cumulatif. Dans le même contexte, Henderson [38] a étudié l'estimation non paramétrique du taux d'arrivée en utilisant des heuristiques avec des données agrégées par intervalles de temps. L'estimateur utilisé est présenté par Law et Kelton [55] et il suppose que la fonction est constante par morceaux avec des noeuds (extrémités des intervalles) bien connus. Henderson [38] a étudié le cas

CHAPITRE 2

MODÉLISATION AVEC LE PROCESSUS DE POISSON

Dans ce chapitre, nous étudions l'utilisation du processus de Poisson dans la modélisation des arrivées dans le contexte des centres d'appels téléphoniques. Dans ce même contexte, nous présentons tout d'abord quelques références sur ce sujet et par la suite, nous présentons le point de départ de nos modèles et leur motivation. Nous verrons dans les chapitres 4 et 5 la validation de ces modèles et leur illustration par des exemples numériques.

2.1 Processus de Poisson et les arrivées des appels téléphoniques : quelques références

Nous citons ici différents modèles utilisés pour la fonction de taux dans les cas où le processus d'arrivée est modélisé par un processus de Poisson et nous proposons une nouvelle alternative.

Dans les travaux de Deslauriers [25], on trouve plusieurs modélisations pour le processus d'arrivée, à commencer par la modélisation simple par un processus de Poisson homogène, jusqu'à un processus non homogène ou encore un processus de Poisson doublement stochastique défini par l'équation (2.12). L'auteur a comparé les résultats des simulations des différents modèles et a établi différents tests d'ajustement et a démontré que plusieurs hypothèses simplistes, comme par exemple celle qui ne tient pas compte de la dépendance entre les arrivées dans les demi-heures de la journée, ne décrivent pas bien la réalité dans les centres d'appels. A la fin, il a proposé un modèle basé sur l'hypothèse de la dépendance entre les différentes périodes de la journée, qui semble être plus adapté à la nature du centre d'appels

particulier où les intervalles de temps sont égaux et a montré que cet estimateur est assez robuste quand les intervalles de temps diminuent pour une certaine longueur et la quantité de données disponible pour l'estimation est assez importante.

Steckley et al. [85] ont souligné l'impact de l'erreur de prévision du taux d'arrivée sur les mesures de performance dans tout système de service qui comporte un processus d'arrivée. Leur travail est motivé par le fait que beaucoup de modèles utilisent le processus de Poisson sans se soucier trop de l'erreur de prévision qui peut en découler si ce modèle ne représente pas la réalité du système. Ils ont illustré le travail en utilisant un modèle doublement stochastique pour le processus d'arrivée dans leur exemples et ont montré que l'erreur de prévision peut avoir un impact sur les modèles de planification à long terme, à court terme et même en temps réel dans une journée. Ils ont même suggéré que l'erreur de prévision soit modélisée explicitement dans l'optimisation d'un système quelconque. Dans le même contexte, Steckley et al. [86] ont aussi souligné l'importance d'un bon choix de modèle pour le processus d'arrivée dans l'approximation des mesures de performance stationnaires, alors que Robbins et al. [76] ont étudié l'impact de la variabilité et le caractère stochastique du processus d'arrivées sur les mesures de performance de tout système de file d'attente.

Nous nous basons sur le fait que la fonction de taux constante par morceaux ne représente pas la réalité dans les centres d'appels [84], pour proposer un modèle qui permet d'avoir une fonction de taux continue. Nous proposons alors d'estimer la fonction de taux par une fonction spline dans deux modèles différents : avec le processus de Poisson non homogène et le processus de Poisson doublement stochastique. Dans le deuxième cas, nous modélisons la fonction de taux selon la forme (2.12), proposée par Whitt [93], et c'est sa composante déterministe $\lambda(t)$ qui sera estimée par une fonction spline.

2.2 Estimation non paramétrique de la fonction de taux par une fonction spline

Le processus de Poisson est caractérisé essentiellement par sa fonction de taux. Par conséquent la construction d'un tel modèle repose sur l'estimation de cette fonction.

Dans les travaux qui ont précédé dans ce projet de Bell Canada, la fonction de taux a été estimée par une fonction constante par morceaux (par demi-heure pour le centre C1). Il faut noter ici que le manque de détails dans les données sur les arrivées dans ce centre a motivé cette modélisation. Or, comme ce taux varie d'une demi-heure à l'autre, il peut aussi varier à l'intérieur de chaque demi-heure. Dans plusieurs situations, il est plus approprié qu'il soit représenté par une fonction continue [84].

Ainsi, il serait utile de trouver une fonction continue qui tient compte de cette variation et qui remédie aussi au problème de sauts de la fin d'une demi-heure et le début de la demi-heure et qui s'ajuste bien aux données disponibles. C'est à partir de là qu'est venue l'idée des fonctions splines.

2.2.1 Les fonction splines

Ce type de fonction présente un outil mathématique très utile dans l'estimation statistique non paramétrique dans plusieurs domaines comme l'imagerie et le traitement des signaux. Le principe est de construire un estimateur $\hat{\lambda}(t)$ de la fonction de taux sur des intervalles départagés par des noeuds bien choisis et à chaque intervalle, on construit une fonction continue et dérivable jusqu'à un certain ordre. Ensuite les constantes ou les paramètres sont estimés de façon à lier tous les noeuds par une seule fonction continue sur toute la période étudiée. Dans notre cas, nous avons utilisé la spline cubique.

Définition 2.2.1. Une fonction spline $s(t)$ d'ordre k (ou de degré $k - 1$), avec les noeuds $r_1 < r_2 \dots < r_h$ et un domaine $[a, b]$ tel que

$$-\infty \leq a < r_1 < r_2 \dots < r_h < b \leq \infty,$$

est une fonction qui satisfait les conditions suivantes :

1. Dans chacun des intervalles $[a, r_1], \dots, [r_{j-1}, r_j], \dots, [r_h, b]$, $s(t)$ est un polynôme de degré au plus $(k - 1)$.
2. $s(t), s'(t), \dots, s^{(k-2)}(t)$, sont continues sur $[a, b]$.

L'ordre le plus utilisé est l'ordre 4, qui donne lieu à des splines cubiques.

La fonction spline cubique peut être écrite d'une façon unique sur l'intervalle $[a, b]$, nous avons

$$s(t) = \sum_{k=1}^3 \alpha_k t^k + \sum_{l=1}^h \beta_l (t - r_l)_+^3,$$

où

$$u_+ = \begin{cases} u & \text{si } u \geq 0 \\ 0 & \text{si sinon,} \end{cases}$$

et $\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \dots, \beta_h$, sont des paramètres [78].

Le théorème suivant nous assure l'unicité d'une spline cubique dans un intervalle donné.

Théorème 2.2.1. Soit f une fonction définie aux points $a = r_0 < r_1 < \dots < r_h < b = r_{h+1}$, de l'intervalle $[a, b]$. Si f est continue et dérivable jusqu'à l'ordre 2 sur $[a, b]$, alors il existe une spline cubique unique s qui approxime f sur $[a, b]$ en l'interpolant aux points (noeuds) r_0, r_1, \dots, r_{h+1} .

Si s vérifie $s^{(2)}(r_0) = s^{(2)}(r_{h+1}) = 0$, alors elle est appelée *spline cubique naturelle*.

Démonstration. Voir Burden et Faires [16, page 135], par exemple. □

2.2.2 Processus de Poisson non homogène

Avant d'expliquer le modèle basé sur une fonction spline pour la fonction de taux, nous devons tout d'abord nous assurer du choix du processus de Poisson comme modèle pour les arrivées dans les centres d'appels étudiés. Nous avons commencé par éliminer le processus de Poisson homogène comme candidat puisque le taux d'arrivée varie tout au long de la journée. Dans un premier temps, nous allons vérifier si le processus de Poisson non homogène peut représenter un bon modèle ou non pour les arrivées au cours d'une journée. Parmi les possibilités, il suffit de considérer des intervalles de temps de la journée où le taux est relativement stationnaire à l'intérieur de chaque intervalle et vérifier si le nombre d'arrivées à l'intérieur de chaque intervalle est un processus de Poisson avec un taux λ constant. Il est équivalent aussi à tester si les inter-arrivées dans cet intervalle sont i.i.d. (indépendants et identiquement distribués) de loi exponentielle de moyenne $1/\lambda$, ce qui permettra de conclure si le nombre d'arrivées par intervalles de temps disjoints sont indépendants ou non. Avec les données disponibles, les instants d'arrivées sont agrégés et enregistrés en secondes, ce qui fait que les instants et les inter-arrivées prennent des valeurs entières. Alors, nous utilisons plutôt le test de chi-deux de Pearson pour tester l'hypothèse nulle suivante :

$$H_0 : \quad \text{Les observations } X_i \text{ sont des variables i.i.d. provenant} \\ \text{de la distribution de la loi de Poisson avec un taux } \lambda, \quad (2.1)$$

où X_i est le nombre d'arrivées dans un intervalle bien choisi au jour $i = 1, \dots, n$.

Pour réaliser le test, nous devons diviser la plage des nombres observés des arrivées en k intervalles adjacents $[a_0, a_1), [a_1, a_2), \dots, [a_{k-1}, a_k]$, et nous devons calculer

$O_j =$ le nombre des X_i qui tombent dans l'intervalle $[a_{j-1}, a_j)$

pour $i = 1, 2, \dots, n$; et $j = 1, 2, \dots, k$,

et $p_j = Pr \{X \in [a_{j-1}, a_j)\}$, qui est la probabilité que le nombre d'arrivées soit dans l'intervalle $[a_{j-1}, a_j)$. Pour la validité du test, il faut s'assurer que les np_j sont assez grands et pour cela nous utiliserons la règle $np_j \geq 5$ pour $j = 1, \dots, k = 11$. Cette règle, qui est assez populaire avec ce genre de test, est suggérée par plusieurs auteurs comme Yarnold [94], par exemple.

Alors, sous H_0 , la statistique

$$X^2 = \sum_{j=1}^k \frac{(O_j - np_j)^2}{np_j} \quad (2.2)$$

suit approximativement une loi chi-deux avec $(k - 1)$ degrés de liberté.

Nous verrons plus loin avec l'application de ce test aux données du centre C2 que le processus de Poisson non homogène peut être considéré parmi les modèles pour les arrivées des appels téléphoniques. Nous verrons en même temps que la modélisation par un taux d'arrivée constant par morceaux n'est pas un très bon choix pour cet exemple, et que même pour des petits intervalles, ce taux n'est pas vraiment stationnaire. Ceci constituera une bonne motivation pour estimer la fonction de taux par une fonction continue qui décrit mieux la réalité dans les centres d'appels, au lieu d'une fonction constante par morceaux. Nous avons opté

pour une fonction spline cubique.

2.2.2.1 Estimation des paramètres

Pour l'estimation de la fonction spline, nous disposons donc d'une fonction, notée f , évaluée seulement aux points $r_1 < r_2 \dots < r_h$. Dans le cas où a et b sont finis, quelques auteurs considèrent aussi des noeuds ($a = r_0$ et $b = r_{h+1}$), et la fonction f est aussi évaluée aux points a et b . Pour tout le domaine $[a, b]$, cette fonction est estimée par :

$$\begin{aligned} s(t) &= P_i(t) \\ &= c_{1,i} + c_{2,i}(t - r_i) + c_{3,i}(t - r_i)^2 + c_{4,i}(t - r_i)^3, \quad (2.3) \\ &\text{pour } t \in [r_i, r_{i+1}], \quad i = 0, \dots, h, \end{aligned}$$

qui n'est autre que la formule d'interpolation polynomiale de Newton (voir, par exemple, Conte et de Boor [21]) qui utilise la formule générale de récurrence suivante : pour une fonction f définie aux points distincts x_0, \dots, x_n , avec respectivement ses valeurs, $f(x_0), \dots, f(x_n)$, nous avons la relation de récurrence suivante :

$$\begin{aligned} f[x_j] &= f(x_j), \quad j = 0, \dots, n, \\ f[x_j, \dots, x_{j+k}] &= \frac{f[x_{j+1}, \dots, x_{j+k}] - f[x_j, \dots, x_{j+k-1}]}{x_{j+k} - x_j}, \quad (2.4) \\ &k = 1, \dots, n, \\ &j = 0, \dots, n - k. \end{aligned}$$

Les coefficients de l'équation (2.3) sont donnés par :

$$\begin{cases} c_{1,i} = f(r_i) \\ c_{2,i} = f'(r_i) \\ c_{3,i} = f[r_i, r_i, r_{i+1}] - f[r_i, r_i, r_{i+1}, r_{i+1}](r_{i+1} - r_i) \\ c_{4,i} = f[r_i, r_i, r_{i+1}, r_{i+1}], \\ i = 0, \dots, h. \end{cases} \quad (2.5)$$

Les conditions citées dans la définition 2.2.1 peuvent être écrites sous-forme d'équations comme suit :

$$\begin{cases} P_i(r_i) = f(r_i), \quad P_i(r_{i+1}) = f(r_{i+1}), \quad i = 0, \dots, h; \\ P_{i-1}^{(k)}(r_i) = P_i^{(k)}(r_i), \quad k = 0, 1, 2, \quad i = 1, \dots, h. \end{cases} \quad (2.6)$$

Dans (2.3), nous nous retrouvons avec $(h + 1)$ polynômes avec 4 paramètres chacun, $c_{1,i}, c_{2,i}, c_{3,i}$ et $c_{4,i}$; $i = 0, \dots, h$. Pour estimer ces paramètres, on pose

$$s_i = P'_{i-1}(r_i) = P'_i(r_i), \quad \text{pour } i = 1, \dots, h,$$

et (2.7)

$$\Delta_i = r_{i+1} - r_i, \quad \text{pour } i = 0, \dots, h.$$

Si nous écrivons les paramètres $c_{1,i}, c_{2,i}, c_{3,i}$ et $c_{4,i}$ en fonction des s_i , alors les coefficients dans (2.5) deviennent :

$$\begin{cases} c_{1,i} = f(r_i) \\ c_{2,i} = s_i \\ c_{3,i} = \frac{f[r_i, r_{i+1}] - s_i}{\Delta_i} - c_{4,i} \Delta_i \\ c_{4,i} = \frac{s_{i+1} + s_i - 2f[r_i, r_{i+1}]}{\Delta_i^2}, \end{cases} \quad (2.8)$$

pour $i = 0, \dots, h$.

Nous reprenons la deuxième condition dans (2.6) pour $k = 2$ et en remplaçant les polynômes P_i par leur expression dans (2.3), nous obtenons un système linéaire constitué des équations suivantes :

$$2c_{3,i-1} + 6c_{4,i-1}\Delta_{i-1} = 2c_{3,i}, \quad i = 1, \dots, h, \quad (2.9)$$

et en remplaçant les coefficients $c_{i,i}$ par leurs valeurs dans (2.8), nous obtenons finalement le système suivant :

$$\begin{aligned} &\Delta_i s_{i-1} + 2(\Delta_{i-1} + \Delta_i) s_i + \Delta_{i-1} s_{i+1} \\ &= 3(f[r_{i-1}, r_i] \Delta_i + f[r_i, r_{i+1}] \Delta_{i-1}), \quad i = 1, \dots, h. \end{aligned} \quad (2.10)$$

Maintenant, nous nous retrouvons avec un système linéaire réduit de h équations à h inconnues ($s_i, i = 1, \dots, h$) qu'on résoudra avec la méthode d'élimination de Gauss. Effectivement, la matrice du système linéaire (2.10) étant triangulaire supérieure avec ses éléments de la diagonale non nuls, on obtient directement une solution unique avec cette méthode [23]. Les deux paramètres s_0 et s_{h+1} peuvent être calculés de plusieurs façons. Parmi les solutions, l'une consiste à utiliser la condition $f^{(2)}(r_0) = f^{(2)}(r_{h+1}) = 0$, ce qui donne lieu à la spline cubique naturelle. En pratique, on peut utiliser des informations a priori sur la courbe aux deux extrémités.

Le théorème suivant va nous permettre de borner l'erreur de l'estimation par une fonction spline cubique.

Théorème 2.2.2. *Soit f une fonction dérivable jusqu'à l'ordre 4 sur $[a, b]$ et soit les noeuds d'interpolation $a = r_0 < r_1 < r_2 \dots < r_h < b = r_{h+1}$. Si s est la fonction*

spline cubique naturelle qui interpole f aux points r_0, r_1, \dots, r_{h+1} , alors l'erreur d'approximation satisfait :

$$|f(t) - s(t)| \leq \frac{5 \max_{a \leq \xi \leq b} |f^{(4)}(\xi)| (\max_{0 \leq i \leq h} \Delta_i)^4}{384}, \quad \text{pour } a \leq t \leq b. \quad (2.11)$$

Démonstration. Voir Schultz [77, Théorème 7.4, page 57], par exemple. \square

Par exemple, avec h noeuds équidistants sur $[a, b]$, nous avons $\max_{0 \leq i \leq h} \Delta_i = (b - a)/(h + 1)$. Nous pouvons donc établir que si $\max_{a \leq \xi \leq b} |f^{(4)}(\xi)| < \infty$, alors la norme uniforme $|f - s| = \max_{a \leq t \leq b} |f(t) - s(t)| = O(h^{-4})$ [23].

Dans le contexte du centre d'appels, nous avons opté pour une fonction spline cubique et les noeuds correspondent au début de chaque période de temps étudiée, quart d'heure ou demi-heure par exemple, dépendamment du cas étudié. Nous avons aussi opté en particulier pour la spline cubique naturelle, puisqu'on peut par exemple établir que la fonction de taux est nulle jusqu'à l'ouverture du centre, c.-à-d. $s(t) = 0$, pour $t \leq r_0$, et par conséquent $f^{(2)}(r_0) = s^{(2)}(r_0) = 0$. Pour la fin de la période de temps étudiée, nous allons aussi supposer que la fonction de taux est positive jusqu'à l'instant de fermeture, c.-à-d. que $s(t) = 0$, pour $t \geq r_{h+1}$, et par conséquent, $f^{(2)}(r_{h+1}) = s^{(2)}(r_{h+1}) = 0$.

Ceci est expliqué par le fait qu'en pratique, même si après l'heure de fermeture du centre un certain nombre d'appels continuent d'arriver, les agents ne prennent en charge que les appels qui sont arrivés avant la fermeture et qui sont encore en attente dans la file.

2.2.3 Processus de Poisson doublement stochastique

Une autre façon d'exploiter le processus de Poisson doublement stochastique ou le processus de Cox est de l'utiliser sous la formulation particulière pour la fonction

de taux, noté $\Lambda(t)$ et qui est donnée par :

$$\Lambda(t) = W\lambda(t), \quad t \geq 0, \quad (2.12)$$

où W est une variable aléatoire indépendante du temps et qui varie par exemple d'une journée à l'autre, et $\lambda(t)$ est une fonction qui représente la variation du taux d'arrivée dans le temps. Pour une journée j et conditionnellement à une réalisation de $W = w$, on a un processus de Poisson non homogène de taux $w\lambda(t)$. Le travail d'Avramidis et al. [5] illustre bien la validité et le réalisme de cette approche qui tient compte de la dépendance. Ces mêmes auteurs ont aussi proposé des modèles paramétriques pour le nombre d'arrivées pour une journée sous forme de vecteur de dimension $d = 25$, où chaque élément représente une demi-heure dans le centre.

Il faut noter ici que Jongbloed et Koole [46] ont utilisé aussi le processus de Poisson doublement stochastique dans le contexte des centres d'appels, où chaque fonction de taux dans chaque intervalle de temps est une variable aléatoire qui est indépendante de chaque variable pour les autres intervalles.

Avec le succès enregistré par le modèle présenté par Avramidis et al. [5], nous avons utilisé ce même modèle, mais avec une fonction de taux $\lambda(t)$ qui est estimée par la fonction spline $s(t)$ au lieu d'une fonction constante par morceaux.

Nous supposons que W suit une loi gamma de paramètres α et β , et nous notons sa densité $f_{\alpha,\beta}$. Ce choix repose essentiellement sur le fait que cette loi permet d'avoir une loi paramétrique pour la distribution conjointe inconditionnelle du nombre d'arrivées dans chaque intervalle de temps. Cette loi n'est autre que la loi multinomiale négative.

Alors, sous l'hypothèse d'un processus de Poisson, nous avons que le nombre d'arrivées X_i dans chaque intervalle de temps $[r_i, r_{i+1}]$, $i = 0, \dots, h$, suit la loi de

Poisson définie par la fonction de masse suivante :

$$P(X_i = x|w) = \frac{e^{-w\lambda_i}(w\lambda_i)^x}{x!}, \quad \text{pour } x = 0, 1, \dots \quad \text{et } i = 0, \dots, h, \quad (2.13)$$

avec

$$\begin{aligned} \lambda_i &= \int_{r_i}^{r_{i+1}} s(t) dt, \\ &= c_{1,i}\Delta_i + c_{2,i}\Delta_i^2/2 + c_{3,i}\Delta_i^3/3 + c_{4,i}\Delta_i^4/4, \end{aligned} \quad (2.14)$$

pour $i = 0, \dots, h$. Ces λ_i sont obtenus en utilisant la fonction $s(t)$ définie dans (2.3).

2.2.3.1 Estimation des paramètres

Pour estimer les paramètres α et β de la loi W et en même temps les paramètres de la spline cubique s , nous avons opté pour la méthode de maximum de vraisemblance, puisque nous pouvons avoir la fonction de masse conjointe des variables $X_i, i = 0, \dots, h$, qui contient tous ces paramètres en même temps. Nous devons tout d'abord établir la distribution inconditionnelle de chaque variable $X_i, i = 0, \dots, h$.

Proposition 2.2.1. *Sous l'hypothèse que W suit une loi gamma de paramètres α et β , le vecteur de variables aléatoires $X = (X_0, \dots, X_h)$, pour le nombre d'arrivées par intervalle $[r_i, r_{i+1}], i = 0, \dots, h$, suit une loi multinomiale négative de paramètres $(\alpha; \beta\lambda_0, \dots, \beta\lambda_h)$. Chaque variable aléatoire X_i suit une loi binomiale négative de paramètres $(\alpha, \beta\lambda_i), i = 0, \dots, h$.*

Démonstration. Nous savons déjà que chaque variable aléatoire $X_i|W$ suit une loi de Poisson de paramètre $(w\lambda_i), i = 0, \dots, h$, pour une réalisation w de W , tel que donné dans (2.13). Par la propriété d'indépendance du processus de Poisson, nous

avons que la fonction de masse conjointe de $X|W$ est donnée par :

$$f_{X|W}(x_0, \dots, x_h|w) = \prod_{i=0}^h \left(\frac{(\lambda_i w)^{x_i} e^{-\lambda_i w}}{x_i!} \right).$$

La fonction de masse conjointe inconditionnelle de X est donc donnée par :

$$\begin{aligned} f_X(x_0, \dots, x_h) &= \int_0^\infty f_{X|W}(x_0, \dots, x_h|w) f_{\alpha, \beta}(w) dw \\ &= \int_0^\infty \prod_{i=0}^h \left(\frac{(\lambda_i w)^{x_i} e^{-\lambda_i w}}{x_i!} \right) \left(\frac{\beta^{-\alpha} w^{\alpha-1} e^{-w/\beta}}{\Gamma(\alpha)} \right) dw \\ &= \frac{\Gamma(\sum_{i=0}^h x_i + \alpha)}{\Gamma(\alpha) \prod_{i=0}^h x_i!} \left(\frac{1}{\sum_{k=0}^h \beta \lambda_k + 1} \right)^\alpha \\ &\quad \prod_{i=0}^h \left(\frac{\beta \lambda_i}{\sum_{k=0}^h \lambda_k + 1} \right)^{x_i}, \end{aligned} \tag{2.15}$$

qui n'est autre que la densité de la loi multinomiale négative de paramètres $(\alpha; \beta \lambda_0, \dots, \beta \lambda_h)$ [5, 25], où chaque $\lambda_i, i = 0, \dots, h$, est donné par (2.14). Il en découle directement des propriétés de la loi multinomiale négative que chaque élément $X_i, i = 0, \dots, h$, du vecteur X suit la loi binomiale négative [45, page 294]. \square

Si nous disposons de données observées pour les arrivées par intervalles de temps pour un nombre de journées n , nous appelons $X_{i,j}$ le nombre d'arrivées dans l'intervalle $[r_i, r_{i+1}], i = 0, \dots, h$, au jour $j = 1, \dots, n$. Nous appelons aussi $f_\theta(x_{0,1}, \dots, x_{h,1}; x_{0,2}, \dots, x_{h,2}; \dots; x_{0,n}, \dots, x_{h,n})$, la fonction de masse conjointe de ces variables aléatoires, où $\theta = (\alpha, \beta; c_{1,0}, \dots, c_{1,h}, \dots, c_{4,h})$ représente l'ensemble des paramètres du processus. En supposant que les arrivées sont indépendantes d'une

journée à l'autre, nous avons :

$$f_{\theta}(x_{0,1}, \dots, x_{h,1}; x_{0,2}, \dots, x_{h,2}; \dots; x_{0,n}, \dots, x_{h,n}) = \prod_{j=1}^n f_X(x_{0,j}, \dots, x_{h,j}), \quad (2.16)$$

où $f_X(x_{0,j}, \dots, x_{h,j})$ est la fonction de masse conjointe pour le nombre d'arrivées dans les intervalles $[r_i, r_{i+1}]$, $i = 0, \dots, h$, du jour $j = 1, \dots, n$. Selon l'équation (2.15), cette densité est donnée par :

$$f_X(x_{0,j}, \dots, x_{h,j}) = \frac{\Gamma(\sum_{i=0}^h x_{i,j} + \alpha)}{\Gamma(\alpha) \prod_{i=0}^h x_{i,j}!} \left(\frac{1}{\sum_{k=0}^h \beta \lambda_k + 1} \right)^{\alpha} \prod_{i=0}^h \left(\frac{\beta \lambda_i}{\sum_{k=0}^h \lambda_k + 1} \right)^{x_{i,j}}$$

Pour trouver les estimateurs de maximum de vraisemblance de θ , nous avons utilisé la formule (2.14) pour remplacer les λ_i , $i = 0, \dots, h$, pour les exprimer en fonction des paramètres $c_{1,i}$, $c_{2,i}$, $c_{3,i}$ et $c_{4,i}$. Pour les n journées, la fonction de log-vraisemblance est donnée par :

$$\begin{aligned} \phi(\theta) &= \log(f_{\theta}(x_{0,1}, \dots, x_{h,1}; x_{0,2}, \dots, x_{h,2}; \dots; x_{0,n}, \dots, x_{h,n})) \\ &= \sum_{j=1}^n \log(f_X(x_{0,j}, \dots, x_{h,j})) \\ &= \sum_{j=1}^n \log \left(\frac{\Gamma(\alpha + \sum_{i=0}^h x_{i,j})}{\Gamma(\alpha)} \right) - \sum_{j=1}^n \log \left(\prod_{i=0}^h x_{i,j}! \right) \\ &\quad + n\alpha \log \left(\frac{1}{1 + \beta \sum_{k=0}^h \sum_{l=1}^4 c_{l,k} \Delta_k^l / l} \right) \\ &\quad + \sum_{j=1}^n \sum_{i=0}^h x_{i,j} \log \left(\frac{\beta \sum_{l=1}^4 c_{l,i} \Delta_i^l / l}{1 + \beta \sum_{k=0}^h \sum_{l=1}^4 c_{l,k} \Delta_k^l / l} \right). \end{aligned} \quad (2.17)$$

Si nous fixons $\beta = 1/\alpha$, ce qui garantit aussi que la variable W aura une

moyenne de 1 [25], nous obtenons

$$\begin{aligned} \phi(\theta) = & \sum_{j=1}^n \sum_{m=0}^{x_j-1} \log(\alpha + m) - \sum_{j=1}^n \sum_{i=0}^h \log(x_{i,j}!) \\ & + n\alpha \log \left(\frac{\alpha}{\alpha + \sum_{k=0}^h \sum_{l=1}^4 c_{l,k} \Delta_k^l / l} \right) \\ & + \sum_{j=1}^n \sum_{i=0}^h x_{i,j} \log \left(\frac{\sum_{l=1}^4 c_{l,i} \Delta_i^l / l}{\alpha + \sum_{k=0}^h \sum_{l=1}^4 c_{l,k} \Delta_k^l / l} \right), \end{aligned} \quad (2.18)$$

avec $x_j = \sum_{i=0}^h x_{i,j}$, le volume total des arrivées enregistrées le jour $j = 1, \dots, n$. Les dérivées de ϕ par rapport aux paramètres α et $c_{1,i}, c_{2,i}, c_{3,i}, c_{4,i}, i = 0, \dots, h$, sont données par

$$\begin{aligned} \frac{\partial \phi(\theta)}{\partial \alpha} = & \sum_{j=1}^n \sum_{m=0}^{x_j-1} \frac{1}{(\alpha + m)} + n \log \left(\frac{\alpha}{\alpha + \sum_{k=0}^h \sum_{l=1}^4 c_{l,k} \Delta_k^l / l} \right) \\ & - \frac{\sum_{j=1}^n \sum_{i=0}^h x_{i,j} - n \sum_{k=0}^h \sum_{l=1}^4 c_{l,k} \Delta_k^l / l}{\alpha + \sum_{k=0}^h \sum_{l=1}^4 c_{l,k} \Delta_k^l / l} \end{aligned} \quad (2.19)$$

$$\frac{\partial \phi(\theta)}{\partial c_{l,i}} = \sum_{j=1}^n \sum_{i=0}^h x_{i,j} \left(\frac{\Delta_i^l / l}{\sum_{l=1}^4 c_{l,i} \Delta_i^l / l} \right) - \frac{\Delta_i^l / l (\sum_{j=1}^n \sum_{i=0}^h x_{i,j} + n\alpha)}{\alpha + \sum_{k=0}^h \sum_{l=1}^4 c_{l,k} \Delta_k^l / l}, \quad (2.20)$$

pour $l = 1, \dots, 4$ et $i = 0, \dots, h$.

Pour calculer un estimateur pour θ qui maximise la fonction ϕ en annulant ses dérivées partielles, nous avons utilisé la fonction *nlnmb* du logiciel R. Cette fonction est utilisée pour minimiser une fonction à plusieurs variables avec ou sans contraintes linéaires. Elle nécessite des valeurs initiales et utilise les dérivées premières et les dérivées secondes comme option. Nous l'avons donc utilisée plutôt pour minimiser $(-\phi)$. Nous n'avons pas pu fournir les dérivées secondes à cause de la complexité de leur calcul et aussi l'ordre de la matrice hessienne correspondante

est assez élevé ($4 \times h + 1$). Par contre, pour accélérer et assurer la convergence de la maximisation, nous avons utilisé comme valeurs initiales les estimations des paramètres $c_{l,i}$, $l = 1, \dots, d$, $i = 0, \dots, h$, obtenues avec la procédure décrite dans la section 2.2.1, c.-à.-d. avec une spline cubique sans la composante W , et nous avons ajouté la contrainte $\alpha > 0$.

Dans les chapitres 4 et 5, nous présentons les résultats numériques des estimations pour chaque centre, et l'application des modèles avec une étude comparative.

CHAPITRE 3

LES ARRIVÉES ET LA MODÉLISATION DE LA DÉPENDANCE INTRA-JOURNALIÈRE

3.1 Introduction

Dans ce chapitre, nous étudions la dépendance qui peut exister entre les arrivées dans différents intervalles d'une journée. En d'autres termes nous présentons des modèles qui tiennent compte de la corrélation qui peut exister entre le nombre d'arrivées dans un intervalle quelconque et celui d'un autre intervalle dans une même journée. Pour les modèles que nous proposons ici, nous supposons que les volumes d'arrivées d'une journée à une autre sont indépendants. Citons tout d'abord quelques travaux où la modélisation du processus d'arrivée tient compte de la dépendance entre les arrivées de différentes périodes dans une même journée ou même entre des journées successives.

Brown et al. [13] ont évoqué le problème de dépendance dans un contexte de prévision, où ils ont présenté une technique pour prédire le taux d'arrivée dans un processus de Poisson dans une journée avec une structure autorégressive qui tient compte du volume d'appels reçus dans une journée antérieure. Dans le modèle proposé, la variable représentant le nombre d'appels reçus dans chaque intervalle de temps suit une loi de Poisson avec une fonction de taux sous forme de produit de deux composantes, une déterministe et une stochastique qui lui donne une nature prédictive, plus un terme d'erreur.

Dans le cadre du projet de Bell Canada, Avramidis et al. [5] ont présenté un ensemble de modèles qui tiennent compte de la dépendance intra-journalière. Ils ont modélisé les arrivées à l'aide de plusieurs distributions paramétriques comme, par

exemple, la loi multinomiale négative et la loi de Dirichlet. Ces modèles présentent un vrai gain en réalisme par rapport à ceux basés sur des hypothèses plus simplistes, comme par exemple ne pas tenir compte de la corrélation qui peut exister entre les volumes d'appels dans différentes parties de la journée. Bien que ces modèles permettent d'introduire de la corrélation positive entre les arrivées des différents intervalles, cette corrélation s'avère assez élevée par rapport aux corrélations réelles. Selon ces modèles, les corrélations sont toujours positives, alors que dans la réalité, nous pouvons avoir des corrélations nulles ou très faibles entre les arrivées des heures éloignées d'une même journée, comme par exemple les premières heures de la matinée avec celles de la fin de la journée. Nous verrons ceci avec les exemples numériques quand nous ferons les comparaisons sur la base des corrélations.

Il est donc logique de continuer dans ce sens et de proposer des modèles qui tiennent compte de cette dépendance et qui peuvent garantir les quantités de corrélations désirées. En effet, nous verrons plus loin dans les exemples numériques qu'au niveau de chaque centre étudié nous avons détecté la présence de corrélation entre les volume d'appels dans différents intervalles de temps d'une journée. Nous avons opté pour une méthode basé sur les copules normales appelé méthode NORTA ("NORmal To Anything"). Cette méthode permet de générer un vecteur aléatoire, avec des distributions marginales et des corrélations connues, à l'aide d'une matrice de corrélations spécifiques d'un vecteur aléatoire de distribution normale standard. L'étape initiale qui consiste à trouver cette matrice est appelée "initialisation de NORTA". Nous reviendrons en détails sur cette méthode plus loin dans ce chapitre, mais nous allons tout d'abord citer quelques références qui présentent des modèles basés sur des méthodes qui tiennent compte de la dépendance dans un contexte plus général que les centres d'appels. Pour une description des différents concepts de dépendance dans les modèles multivariés, nous citons le livre de Joe [44]. Parmi

ces méthodes on retrouve les copules et pour une bonne référence sur ce sujet, nous citons le livre de Nelsen [69].

Plusieurs auteurs ont utilisé les copules normales, comme par exemple Lurie et Goldberg [61] qui ont présenté une modélisation et un algorithme pour générer un vecteur avec des distributions marginales et une structure de corrélations donnée, mais leur méthode s'applique seulement dans le cas où les distributions marginales sont continues et strictement croissantes. Cario et Nelson [18] ont présenté la méthode ARTA ("AutoRegressive To Anything") pour manipuler l'autocorrélation dans une série temporelle avec distribution marginale donnée et ont généralisé le résultat en 2002, pour une série chronologique vectorielle VARTA. Les mêmes auteurs ont proposé une méthode basée aussi sur les copules normales, qu'ils ont baptisé (NORTA), pour générer un vecteur avec des distributions marginales discrètes ou continues et une structure de corrélations donnée, sans se soucier de la distribution conjointe. Ghosh et Henderson [32] ont évoqué la difficulté de la méthode NORTA pour des grandes dimensions, où il faut transformer la matrice de corrélations obtenue en une qui soit semi-définie positive et ils ont proposé des solutions.

Chen [20] a présenté une approche générale pour l'initialisation de la méthode NORTA. Elle spécifie les différentes procédures à utiliser selon la nature des distributions marginales. Elle propose d'utiliser la simulation pour résoudre le problème d'intégration, donnée en (3.6) et qu'on verra par la suite, dans le cas où les distributions sont discrètes, ou même continues mais que la corrélation est proche de 1 ou -1 . Ce dernier cas particulier a été ignoré dans la littérature. Pour la phase d'initialisation de la méthode NORTA, l'algorithme qu'elle propose est basé sur une méthode stochastique plus adaptée à la simulation que tout autre résolution numérique. Dans la simulation, elle utilise une variable de contrôle pour la réduction

de la variance. Elle a prouvé l'efficacité de la méthode par rapport à la méthode numérique dans des cas particuliers de distributions marginales. Cependant, la méthode risque d'être coûteuse pour des grandes dimensions d du vecteur X . Dans la partie empirique, elle présente une variété d'exemples numériques pour illustrer la méthode, avec des distributions continues et même des cas mixtes (différentes distributions marginales, discrètes ou continues), sauf qu'aucun exemple ne traite le cas où les distributions sont discrètes. En gros, le travail montre l'efficacité de la simulation utilisée dans l'initialisation de NORTA. Au début, l'auteur insiste sur son efficacité pour les distributions discrètes ou distributions continues avec corrélations proche de 1 ou -1 . Cependant, les exemples empiriques ne traitent que des distributions continues.

Nous avons exploité les données disponibles pour modéliser le processus d'arrivée dans un centre d'appels téléphoniques en tenant compte des structures de corrélations à l'intérieur d'une même journée. Nous avons opté pour la méthode NORTA, mais dans le contexte particulier des centres d'appels. Cette méthode s'avère flexible et elle nous permet de générer des arrivées en préservant la structure de corrélation avec des distributions marginales quelconques sans se soucier de la distribution conjointe. Nous avons proposé trois méthodes efficaces pour l'initialisation de NORTA dans le cas où les distributions marginales sont discrètes. Nous avons aussi proposé une méthode de résolution dans le cas où nous disposons d'un mélange de distributions discrètes et continues. Il faut noter que ce dernier cas a été très peu soulevé dans la littérature et qu'aucune méthode n'a été proposée. Nous avons aussi proposé des solutions aux difficultés de la méthode en grandes dimensions. Cette méthode est expliquée en détails dans la section suivante.

3.2 Méthode NORTA

Le principe est de modéliser directement les arrivées par une distribution multidimensionnelle pour le vecteur $X = (X_1, \dots, X_d)$ qui s'ajuste bien aux données et qui tient compte de la structure de dépendance entre les différentes périodes de la journée. Chaque X_i représente le nombre d'appels reçus pendant l'intervalle de temps i , pour $i = 1, \dots, d$.

Le travail de Avramidis et al. [5] illustre bien la validité de l'approche de modélisation qui tient compte de la dépendance. Ils ont proposé des modèles paramétriques pour le nombre d'arrivées pour une journée sous forme de vecteur de dimension $d = 25$ pour le centre d'appels C1, où chaque élément représente une demi-heure dans le centre. Le modèle, qu'ils ont baptisé M1, propose la multinomiale négative pour estimer la distribution conjointe de (X_1, \dots, X_d) , avec distribution marginale la binomiale négative pour chaque variable $X_i, i = 1, \dots, d$. Il s'agit en fait du processus de Poisson doublement stochastique qu'on a présenté dans la section 2.2.3, mais dans leur modèle la partie déterministe de la fonction de taux est constante par morceaux et la composante stochastique W suit une loi gamma. Nous appelons ce modèle PPDS-constante.

Avec la méthode NORTA, on veut générer $X = (X_1, \dots, X_d)$, un vecteur aléatoire dont on connaît la matrice de corrélations linéaires ou de rang, ainsi que les distributions marginales. Pour chaque $X_i, i = 1, \dots, d$, on note la distribution conjointe

$$F(x_1, \dots, x_d) = Pr\{X_1 \leq x_1, \dots, X_n \leq x_d\} \quad (3.1)$$

et les fonctions de distributions marginales

$$F_i(x) = Pr\{X_i \leq x\}, \quad 1 \leq i \leq d. \quad (3.2)$$

Le vecteur X est obtenu avec la méthode NORTA si chacun de ses éléments est calculé de la façon suivante :

$$X_i = F_i^{-1}(\Phi(Z_i)), \quad i = 1, \dots, d, \quad (3.3)$$

où Φ est la distribution cumulative de la normale standard et Z_i est la $i^{\text{ème}}$ composante d'une normale multivariée avec les corrélations linéaires suivantes :

$$\rho_{i,j}^Z = \text{Corr}(Z_i, Z_j), \quad 1 \leq i < j \leq d. \quad (3.4)$$

Une fois que nous disposons de la matrice de corrélation de Z , notée R_ρ^Z , où chaque élément représente $\rho_{i,j}^Z$, $1 \leq i < j \leq d$, nous générons au début un vecteur $N = (N_1, \dots, N_d)$, i.i.d. d'une variable aléatoire normale standard (moyenne zéro et variance 1). Par la suite, nous obtenons le vecteur $Z = (Z_1, \dots, Z_d)$ d'une loi normale multivariée de moyenne zéro et de matrice de covariance R_ρ^Z en multipliant N par L , où L provient d'une décomposition de la matrice $R_\rho^Z = LL^t$. Par la suite chaque X_i du vecteur X est obtenue selon la formule (3.3).

3.2.1 Méthode NORTA avec les corrélations linéaires

Pour mesurer la dépendance, nous définissons la matrice de corrélations R_ρ^X où chaque élément, noté $\rho_{i,j}^X$, pour $1 \leq i < j \leq d$, représente la corrélation de Pearson entre X_i et X_j .

Puisque la matrice R_ρ^Z est symétrique et tous ses éléments de la diagonale sont égaux à 1, le problème se ramène à calculer les $d(d-1)/2$ corrélations $\rho_{i,j}^Z$ dans des

équations du type :

$$\begin{aligned}
 \rho_{i,j}^X &= \text{Corr}(X_i, X_j) \\
 &= \text{Corr}(F_i^{-1}(\Phi(Z_i)), F_j^{-1}(\Phi(Z_j))) \\
 &= \frac{E[F_i^{-1}(\Phi(Z_i))F_j^{-1}(\Phi(Z_j))] - \mu_i\mu_j}{\sigma_i\sigma_j}, \quad 1 \leq i < j \leq d, \quad (3.5)
 \end{aligned}$$

où

$$\begin{aligned}
 E[F_i^{-1}(\Phi(Z_i))F_j^{-1}(\Phi(Z_j))] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_i^{-1}(\Phi(z_i))F_j^{-1}(\Phi(z_j))\phi(z_i, z_j, \rho_{i,j}^Z)dz_idz_j \\
 &= g_i(\rho_{i,j}^Z), \quad 1 \leq i < j \leq d, \quad (3.6)
 \end{aligned}$$

avec μ_i et σ_i^2 représentent respectivement la moyenne et la variance de X_i , pour $i = 1, \dots, d$, la fonction inverse $F^{-1}(u) = \inf\{x : F(x) \geq u; 0 \leq u \leq 1\}$ et ϕ est la densité de la loi normale standard bivariable avec corrélation $\rho_{i,j}^Z$.

L'estimation des paramètres repose alors sur le calcul de la matrice R_ρ^Z et la résolution de la double intégrale pour chaque paire (i, j) , où $1 \leq i < j \leq d$. En général, la double intégrale peut être calculée analytiquement ou approximée numériquement ou en utilisant la simulation *Monte Carlo (MC)* ou *Quasi Monte Carlo (QMC)*, voir L'Ecuyer [56], par exemple.

Analytiquement, elle est résolue seulement pour des cas particuliers. Par exemple, dans le cas où les distributions marginales sont continues et uniformes, la relation dans l'équation (3.5) se simplifie et la résolution devient :

$$\rho_{i,j}^Z = 2 \sin\left(\frac{\pi\rho_{i,j}^X}{6}\right), \quad 1 \leq i < j \leq d, \quad (3.7)$$

Numériquement, il y a plusieurs procédures pour résoudre l'intégrale double

comme par exemple la quadrature gaussienne [21], qui est définie comme suit :

$$g_l(\rho_{i,j}^Z) \approx \sum_{k=1}^d \sum_{l=1}^d w_k w_l F_i^{-1}(z_i) F_j^{-1}(z_j) \phi(z_i, z_j, \rho_{i,j}^Z), \quad 1 \leq i < j \leq d, \quad (3.8)$$

avec les z_i et z_j bien choisis dans les axes de Z_i et Z_j et où w_l et w_k sont des poids. Cette méthode est implantée dans la majorité des logiciels et elle est efficace seulement dans le cas où la fonction à intégrer est assez lisse [20].

Avec la simulation, nous devons effectuer des transformations sur la double intégrale dans (3.6) pour pouvoir intégrer sur le plan $[0, 1]^2$. Nous avons

$$\begin{aligned} g_l(\rho_{i,j}^Z) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_i^{-1}(\Phi(z_i)) F_j^{-1}(\Phi(z_j)) \phi(z_i, z_j, \rho_{i,j}^Z) dz_i dz_j \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{F_i^{-1}(\Phi(z_i)) F_j^{-1}(\Phi(z_j)) \phi(z_i) \phi\left(\frac{z_j - \rho_{i,j}^Z z_i}{\sqrt{1 - (\rho_{i,j}^Z)^2}}\right)}{\sqrt{1 - (\rho_{i,j}^Z)^2}} dz_i dz_j, \end{aligned} \quad (3.9)$$

où ϕ représente la densité de la loi normale standard. Si on pose

$$\left\{ \begin{array}{l} u_i = \Phi(z_i) \\ \text{et} \\ u_j = \Phi\left(\frac{z_j - \rho_{i,j}^Z z_i}{\sqrt{1 - (\rho_{i,j}^Z)^2}}\right). \end{array} \right. \quad (3.10)$$

on aura finalement :

$$g_l(\rho_{i,j}^Z) = \int_0^1 \int_0^1 F_i^{-1}(u_i) F_j^{-1}(\Phi(\rho_{i,j}^Z \Phi^{-1}(u_i) + \sqrt{1 - \rho_{i,j}^Z{}^2} \Phi^{-1}(u_j))) du_i du_j. \quad (3.11)$$

Maintenant que le domaine original est transformé sur $[0, 1]^2$, nous pouvons utiliser l'intégration Monte Carlo pour estimer la double intégrale par simulation.

3.2.2 Méthode NORTA avec les corrélations de rang

La méthode NORTA peut être utilisée aussi avec les corrélations de rang ou corrélations de Spearman, qu'on notera $r_{i,j}^X$, $1 \leq i < j \leq d$, et la matrice correspondante R_r^X . Les corrélations de rang ont l'avantage d'être plus stables, surtout face aux variations monotones des variables aléatoires, et elles ne dépendent ni des paramètres, ni des des moments d'ordre 1 et 2 des distributions. Nous avons

$$\begin{aligned} r_{i,j}^X &= \text{Corr}(F_i(X_i), F_j(X_j)) \\ &= \frac{E [F_i(F_i^{-1}(\Phi(Z_i)))F_j(F_j^{-1}(\Phi(Z_j)))] - \mu_{F_i}\mu_{F_j}}{\sigma_{F_i}\sigma_{F_j}}, \end{aligned} \quad (3.12)$$

$$1 \leq i < j \leq d,$$

où μ_{F_i} et $\sigma_{F_i}^2$ représentent la moyenne et la variance de $F_i(X_i)$, $i = 1, \dots, d$ ($1/2$ et $1/12$ si la variable aléatoire X_i est continue), et

$$\begin{aligned} E [F_i(F_i^{-1}(\Phi(Z_i)))F_j(F_j^{-1}(\Phi(Z_j)))] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_i(F_i^{-1}(\Phi(z_i)))F_j(F_j^{-1}(\Phi(z_j))) \\ &\quad \phi(z_i, z_j, \rho_{i,j}^Z) dz_i dz_j \\ &= g_r(\rho_{i,j}^Z), \quad 1 \leq i < j \leq d. \end{aligned} \quad (3.13)$$

Si chaque variable aléatoire X_i est continue et que chaque $F_i(X_i) = \Phi(Z_i) = U_i$, pour $i = 1, \dots, d$, est uniforme et continue sur $[0, 1]$, alors on dispose d'une formule analytique semblable à celle donnée en (3.7), soit :

$$\rho_{i,j}^Z = 2 \sin \left(\frac{\pi r_{i,j}^X}{6} \right), \quad 1 \leq i < j \leq d, \quad (3.14)$$

Dans le cas où les distributions marginales sont discrètes, il faut donc résoudre la double intégrale, donnée en (3.13). La méthode numérique est inefficace puisque

la fonction à intégrer n'est pas lisse.

Avec la simulation, il y a une procédure semblable à celle des corrélations linéaires pour calculer $g_r(\rho_{i,j}^Z)$, soit :

$$g_r(\rho_{i,j}^Z) = \int_0^1 \int_0^1 F_i(F_i^{-1}(u_i)) F_j(F_j^{-1}(\Phi(\rho_{i,j}^Z \Phi^{-1}(u_i) + \sqrt{1 - \rho_{i,j}^Z{}^2} \Phi^{-1}(u_j)))) du_i du_j, \quad (3.15)$$

et par la suite, nous pouvons utiliser l'intégration Monte Carlo pour estimer la double intégrale par simulation.

Nous proposons des méthodes efficaces pour l'estimation des paramètres dans l'étape d'initialisation de la méthode NORTA, en résolvant la double intégrale en (3.6) ou (3.13), pour les corrélations linéaires ou de rang, respectivement.

Avant d'étudier en détails cette étape et ces méthodes, regardons tout d'abord l'estimation des paramètres des distributions marginales.

3.2.3 Estimation des paramètres des distributions marginales

Comme suite au travail d'Avramidis et al. [5], nous présentons d'autres modèles pour le processus d'arrivées, avec l'aspect de dépendance entre les différents intervalles de temps de la journée. Nous utilisons donc la méthode NORTA et nous supposons que la distribution marginale de chaque $X_i, i = 1, \dots, d$, est une binomiale négative. Cette loi s'ajuste bien aux données des centres d'appels étudiés.

La fonction de masse de la binomiale négative est donnée par :

$$\begin{aligned} P_{s_i, r_i}(X_i = x_i) &= \frac{\Gamma(x_i + s_i)}{x_i! \Gamma(s_i)} \left(\frac{1}{1 + r_i} \right)^{s_i} \left(1 - \frac{1}{1 + r_i} \right)^{x_i} \\ &= \binom{x_i + s_i - 1}{x_i} \left(\frac{1}{1 + r_i} \right)^{s_i} \left(1 - \frac{1}{1 + r_i} \right)^{x_i}, \end{aligned} \quad (3.16)$$

avec $s_i \in \mathbb{N}$ et $0 < 1/(1 + r_i) < 1$. Cette loi correspond au nombre d'échecs avant d'avoir s_i succès dans une expérience de Bernoulli avec probabilité de succès $p_i = 1/(1 + r_i)$. Pour estimer les paramètres s_i et p_i de chaque distribution marginale $F_i, i = 1, \dots, d$, nous utilisons une méthode basée sur le maximum de vraisemblance. Nous disposons des observations $x_{i,1}, \dots, x_{i,n}$, où chaque $x_{i,j}$ représente le nombre d'arrivées observé dans l'intervalle $i = 1, \dots, d$, du jour $j = 1, \dots, n$, des n journées observées. Nous supposons aussi que les journées sont indépendantes. Le logarithme de la fonction de vraisemblance, exprimée en fonction de s_i et p_i , est donné par :

$$\begin{aligned} \phi(s_i, p_i) &= \ln \left(\prod_{j=1}^n P_{s_i, p_i}(X_i = x_{i,j}) \right) \\ &= \sum_{j=1}^n \ln \left(\frac{\Gamma(x_{i,j} + s_i)}{x_{i,j}! \Gamma(s_i)} \right) + n s_i \ln p_i + \ln(1 - p_i) \sum_{j=1}^n x_{i,j}. \end{aligned} \quad (3.17)$$

Pour s_i fixé, on peut facilement montrer que cette fonction est maximale pour $\hat{p}_i = s_i/(s_i + \bar{x}_i)$, où \bar{x}_i désigne le nombre moyen d'appels reçus dans l'intervalle de temps $i = 1, \dots, d$. Alors, en gardant seulement les composantes qui contiennent la variable s_i , la fonction devient :

$$\begin{aligned} \phi(s_i) &= n \bar{x}_i \ln \bar{x}_i + n s_i \ln s_i - n(s_i + \bar{x}_i) \ln(s_i + \bar{x}_i) \\ &\quad + \sum_{k=1}^{m_i} f_i(k) \ln(s_i + k - 1), \end{aligned} \quad (3.18)$$

avec $m_i = \max_{1 \leq j \leq n} x_{i,j}$ et $f_i(k) = \text{nombre des } x_{i,j} \geq k$ [55].

3.3 Initialisation de la méthode NORTA

La phase d'initialisation de la méthode NORTA consiste donc à trouver la matrice R_ρ^Z sachant que nous disposons de la matrice R_r^X (ou R_ρ^X pour les corrélations linéaires), ainsi que les distributions marginales. Pour cela il faut résoudre les $d(d-1)/2$ équations et pour chaque corrélation donnée $r_{i,j}^X$ (ou $\rho_{i,j}^X$), $1 \leq i < j \leq d$, nous devons trouver le $\rho_{i,j}^Z$ correspondant dans des équations du type

$$f_r(\rho) = g_r(\rho) - r_{i,j}^X \sigma_{F_i} \sigma_{F_j} - \mu_{F_i} \mu_{F_j} = 0, \quad (3.19)$$

pour les corrélations de rang, et

$$f_l(\rho) = g_l(\rho) - \rho_{i,j}^X \sigma_i \sigma_j - \mu_i \mu_j = 0, \quad (3.20)$$

pour les corrélations linéaires.

3.3.1 Propriétés mathématiques des corrélations

Le théorème suivant va résumer les propriétés que nous utiliserons pour les deux types de corrélations. Ces propriétés sont aussi développées en détails dans Avramidis et al. [4]. Nous avons

$$r_{i,j}^X(\rho_{i,j}^Z) = \frac{g_r(\rho_{i,j}^Z) - \mu_{F_i} \mu_{F_j}}{\sigma_{F_i} \sigma_{F_j}}, \quad 1 \leq i < j \leq d. \quad (3.21)$$

$$\rho_{i,j}^X(\rho_{i,j}^Z) = \frac{g_l(\rho_{i,j}^Z) - \mu_i \mu_j}{\sigma_i \sigma_j}, \quad 1 \leq i < j \leq d. \quad (3.22)$$

Théorème 3.3.1. [4] Soit F_i et F_j deux distributions marginales associées à X_i et X_j , respectivement. Soit $\rho_{i,j}^X(\rho_{i,j}^Z)$, $r_{i,j}^X(\rho_{i,j}^Z)$ et $\rho_{i,j}^Z$ les corrélations définies dans (3.22), (3.21) et (3.4), respectivement. Nous avons

1. Les fonctions $\rho_{i,j}^X(\rho_{i,j}^Z)$ et $r_{i,j}^X(\rho_{i,j}^Z)$ sont non-décroissantes sur $[-1, 1]$, avec $\rho_{i,j}^X(0) = 0$ et $r_{i,j}^X(0) = 0$.
2. S'il existe $\delta > 0$ qui vérifie $E[|X_i X_j|^{1+\delta}] < \infty$ pour $\rho_{i,j}^Z \in [-1, 1]$. Alors, $\rho_{i,j}^X$ et $r_{i,j}^X$ sont continues sur $[-1, 1]$.
3. Si X_i et X_j sont deux variables continues, alors

$$r_{i,j}^X = 12g_r(\rho_{i,j}^Z) - 3 = \frac{6}{\pi} \arcsin\left(\frac{\rho_{i,j}^Z}{2}\right), \quad (3.23)$$

où

$$g_r(\rho_{i,j}^Z) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Phi(z_i) \Phi(z_j) \phi(z_i, z_j, \rho_{i,j}^Z) dz_i dz_j.$$

Démonstration. Pour les corrélations linéaires $\rho_{i,j}^X$, les parties 1 et 2 sont prouvées par Cario et Nelson [17]. Pour les corrélations de rang $r_{i,j}^X$, la même preuve est valable en remplaçant $F_l(\Phi(Z_l))$ par $F_l^{-1}(F_l(\Phi(Z_l)))$, pour $l = i, j$. Dans la partie 3, la relation en (3.23) a été démontrée en 1907 par Karl Pearson, selon Kurowicka et Cooke [52] et par d'autres auteurs par la suite, comme par exemple Kruskal [49]. □

A partir de ce théorème, nous pouvons établir le corollaire suivant :

Corollaire 3.3.1. [4] Si les distributions F_i et F_j sont non-dégénérées, alors la fonction $r_{i,j}^X(\rho_{i,j}^Z)$ est strictement croissante sur $[-1, 1]$ et elle possède une fonction inverse.

Ce corollaire nous garantit l'existence d'une solution unique $\rho_{i,j}^Z$ associée à chaque $r_{i,j}^X \in [r_{i,j}^X(-1), r_{i,j}^X(1)]$.

Dans la suite des travaux, nous comparons les différentes méthodes en nous basant essentiellement sur les corrélations de rang au lieu des corrélations linéaires. Pour la résolution des équations du type (3.19), nous utilisons des algorithmes qui tiennent compte de la propriété suivante : la fonction g_r (et par conséquent f_r aussi) est croissante sur $[-1, 1]$ et elle s'annule à zéro [4, 18, 20].

Pour une paire (i, j) de variables aléatoires X_i et X_j , on note par $\{x_{i,1}, x_{i,2}, \dots\}$ et $\{x_{j,1}, x_{j,2}, \dots\}$, les points de masse des deux variables aléatoires, respectivement. On note par $\{p_{i,1}, p_{i,2}, \dots\}$ et $\{p_{j,1}, p_{j,2}, \dots\}$, les probabilités de masse correspondantes, et par $\{f_{i,1}, f_{i,2}, \dots\}$ et $\{f_{j,1}, f_{j,2}, \dots\}$, les probabilités cumulatives.

La double intégrale dans (3.13) peut être écrite comme suit :

$$\begin{aligned}
 g_r(\rho) &= \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} f_{i,k} f_{j,l} \left(\int_{z_{i,k-1}}^{z_{i,k}} \int_{z_{j,l-1}}^{z_{j,l}} \phi(x_i, x_j, \rho) dx_i dx_j \right) \\
 &= \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} (f_{i,k+1} - f_{i,k})(f_{j,l+1} - f_{j,l}) \bar{\Phi}(z_{i,k}, z_{j,l}, \rho) \\
 &= \sum_{k=-\infty}^{\infty} p_{i,k+1} \sum_{l=-\infty}^{\infty} p_{j,l+1} \bar{\Phi}(z_{i,k}, z_{j,l}, \rho),
 \end{aligned} \tag{3.24}$$

qui fait apparaître la distribution de la loi normale bivariée

$\bar{\Phi}(x, y, \rho) = \int_x^{\infty} \int_y^{\infty} \phi(z_1, z_2, \rho) dz_1 dz_2$ [4]. De la même manière et pour les corrélations linéaires, nous pouvons facilement développer les formules semblables pour avoir à la fin :

$$g_l(\rho) = \sum_{k=-\infty}^{\infty} (x_{i,k+1} - x_{i,k}) \sum_{l=-\infty}^{\infty} (x_{j,l+1} - x_{j,l}) \bar{\Phi}(z_{i,k}, z_{j,l}, \rho). \tag{3.25}$$

Dans les méthodes que nous proposons pour l'initialisation de NORTA, nous utilisons la fonction dérivée de g_r . Pour cela, il fallait démontrer que cette fonction

est différentiable. Nous présentons seulement la différentiabilité à l'ordre 1, puisque dans [4], cette propriété a été généralisée pour un tout ordre supérieur à 1.

Proposition 3.3.1. *La fonction g_r est infiniment différentiable sur $[-1, 1]$ et sa dérivée première est donnée par :*

$$g_r'(\rho) = \sum_{k=-\infty}^{\infty} p_{i,k} \sum_{l=-\infty}^{\infty} p_{i,l} \phi(z_{i,k}, z_{j,l}, \rho). \quad (3.26)$$

Démonstration. Pour prouver l'existence de la dérivée première, nous exploitons la propriété de la densité de la loi normale standard bivariée où pour $-1 < \rho < 1$,

$$\frac{d}{d\rho} \phi_\rho(x, y) = \frac{\partial^2}{\partial x \partial y} \phi_\rho(x, y) \quad \text{pour tout } x, y \quad (3.27)$$

[48, p. 393]. Nous avons

$$\begin{aligned} \frac{d}{d\rho} \bar{\Phi}_\rho(x, y) &= \int_x^\infty \int_y^\infty \frac{d}{d\rho} \phi_\rho(z_i, z_j) dz_j dz_i \\ &= \int_x^\infty \frac{d}{dz_i} \left[\int_y^\infty \frac{\partial}{\partial z_j} \phi_\rho(z_i, z_j) dz_j \right] dz_i \\ &= \int_x^\infty \frac{d}{dz_i} [-\phi_\rho(z_i, y)] dz_i \\ &= \phi_\rho(x, y). \end{aligned} \quad (3.28)$$

Pour prouver la continuité de g_r' , nous utilisons la propriété de la convergence uniforme (Rudin, 1976, le théorème 7.9) et montrer qu'il existe une suite de fonctions qu'on notera $g'_{r,n}$, qui sont continues et qui convergent vers g'_r en tout point de l'intervalle $[-1, 1]$ et particulièrement dans le voisinage de chaque point de l'intervalle. Alors, pour tout $\rho_0 \in [-1, 1]$, il existe un voisinage de ρ_0 , noté $N_\epsilon(\rho_0) = [\rho_0 - \epsilon, \rho_0 + \epsilon] \subset [-1, 1]$, et la convergence uniforme tient si on considère

la suite de fonctions de la forme

$$g'_{r,n}(\rho) = \sum_{(k,l): \max(|k|,|l|) \leq n} p_{i,k+1} p_{j,l+1} \phi_{\rho}(z_{i,k}, z_{j,l}). \quad (3.29)$$

Il est facile de vérifier que $\lim_{n \rightarrow \infty} g'_{r,n}(\rho) = g'_r(\rho)$, puisque nous avons :

$$\begin{aligned} \sup_{\rho \in N_{\epsilon}(\rho_0)} |g'_{r,n}(\rho) - g'_r(\rho)| &= \sup_{\rho \in N_{\epsilon}(\rho_0)} \sum_{(k,l): \max(|k|,|l|) > n} p_{i,k+1} p_{j,l+1} \phi_{\rho}(z_{i,k}, z_{j,l}) \\ &\leq \frac{1}{2\pi \sqrt{1 - \rho_*^2}} \left[\sum_{k: |k| > n} p_{i,k+1} + \sum_{l: |l| > n} p_{j,l+1} \right], \end{aligned} \quad (3.30)$$

où $\rho_* = \max(|\rho - \epsilon|, |\rho + \epsilon|)$ [4]. La convergence uniforme de $g'_{r,n}$ vers g'_r est vérifiée par le fait que :

$$\lim_{n \rightarrow \infty} \sup_{\rho \in N_{\epsilon}(\rho_0)} \sum_{(k,l): \max(|k|,|l|) > n} p_{i,k+1} p_{j,l+1} \phi_{\rho}(z_{i,k}, z_{j,l}) = 0, \quad (3.31)$$

ce qui veut dire que g'_r est définie et continue sur tout voisinage de ϵ_0 dans $[-1, 1]$ et par conséquent sur tout l'intervalle. Pour la différentiabilité à tout ordre supérieur à 1, la preuve est donnée par Avramidis et al. [4]. \square

En ce qui concerne les corrélations linéaires, des propriétés analogues peuvent être établies. Nous avons :

Proposition 3.3.2. *La fonction g_l est différentiable sur le domaine $[-1, 1]$ et sa dérivée première est donnée par :*

$$g'_l(\rho) = \sum_{k=-\infty}^{\infty} (x_{i,k+1} - x_{i,k}) \sum_{l=-\infty}^{\infty} (x_{j,l+1} - x_{j,l}) \phi_{\rho}(z_{i,k}, z_{j,l}). \quad (3.32)$$

Démonstration. La preuve est semblable à celle de la proposition 3.3.1, il suffit de

changer les $p_{i,k+1}$ et les $p_{j,l+1}$ par les $x_{i,k+1}$ et les $x_{j,l+1}$, respectivement. \square

Corollaire 3.3.2. *Si les distributions F_i et F_j sont non-dégénérées, alors la fonction $\rho_{i,j}^X(\rho_{i,j}^Z)$ est strictement croissante sur $[-1, 1]$ et elle possède une fonction inverse.*

Ce corollaire nous garantit l'existence d'une solution unique $\rho_{i,j}^Z$ associée à chaque $\rho_{i,j}^X \in \rho_{i,j}^X[(-1), \rho_{i,j}^X(1)]$.

3.3.2 Méthode M1 : approximation en utilisant la loi normale bivariée

La première méthode est basée sur les travaux de Cario et Nelson [18] pour calculer la fonction $g_r(x)$, donnée dans (3.13), et résoudre l'équation (3.19). Dans le cas où les deux variables sont discrètes, le calcul se réduit donc à une sommation sur les domaines des distributions et à une double intégration sur le domaine tronqué de la densité de la binormale standard et l'équation (3.24) devient :

$$\begin{aligned} g_r(\rho) &\approx \sum_{k=1}^{m_i} \sum_{l=1}^{m_j} f_{i,k} f_{j,l} \left(\int_{z_{i,k-1}}^{z_{i,k}} \int_{z_{j,l-1}}^{z_{j,l}} \phi(x_i, x_j, \rho) dx_i dx_j \right) \\ &= \sum_{k=0}^{m_i-1} p_{i,k+1} \sum_{l=0}^{m_j-1} p_{j,l+1} \bar{\Phi}(z_{i,k}, z_{j,l}, \rho), \end{aligned} \quad (3.33)$$

où $z_{i,0} = z_{j,0} = -10$, $z_{i,k} = \Phi^{-1}(f_{i,k})$, pour $k = 1, \dots, m_i - 1$, $z_{j,l} = \Phi^{-1}(f_{j,l})$, pour $l = 1, \dots, m_j - 1$ et $z_{i,m_i} = z_{j,m_j} = 10$. Ici nous remarquons que le domaine de l'intégration est tronqué entre $[-10, 10]^2$ au lieu de $[-\infty, \infty]^2$. Avec le logiciel MATLAB par exemple, cette double intégrale (intégration numérique) converge à partir de $[-6, 6]^2$. Dans notre cas, pour des distributions discrètes avec domaine infini, nous procédons à la troncation dans les domaines de F_i et F_j jusqu'aux points $m_i = F_i^{-1}(1 - 10^{-6})$ et $m_j = F_j^{-1}(1 - 10^{-6})$; nous avons par conséquent

$f_{i,m_i} = f_{j,m_j} = 1$. Pour $z_{i,0}$, $z_{j,0}$ et z_{i,m_i} , z_{j,m_j} , respectivement, au lieu de $\Phi^{-1}(0)$ et $\Phi^{-1}(1)$ (autrement dit $-\infty$ et ∞), nous avons considéré les valeurs utilisées dans la librairie de simulation en Java *SSJ* [57], soit $z_{i,0} = z_{j,0} = \Phi^{-1}(2.2 \times 10^{-308}) = -37.5197$ et $z_{i,m_i} = z_{j,m_j} = \Phi^{-1}(1 - 2.2 \times 10^{-16}) = 8.1259$. Pour chaque corrélation, pour le calcul de cette double intégrale, nous nous sommes basés sur l'algorithme 462 de Donnelly [26].

Cario et Nelson [18] ont utilisé la méthode NORTA en considérant les corrélations linéaires $\rho_{i,j}^X$ du vecteur X pour trouver les corrélations correspondantes $\rho_{i,j}^Z$ du vecteur Z , pour $1 \leq i < j \leq d$. Ils ont utilisé un algorithme numérique qui répète le calcul de la fonction $g_l(x)$ sur une grille de valeurs candidates dans l'intervalle $[-1, 1]$ et ensuite l'algorithme affine la grille sur des itérations jusqu'à ce que la valeur $\rho_{i,j}^Z$ soit atteinte avec un seuil de tolérance bien défini.

Dans cette méthode, aucune information n'est spécifiée sur le coût computationnel, sur l'erreur d'approximation ou le seuil de tolérance dans les calculs des doubles intégrales pour le cas discret. Il n'est pas spécifié clairement aussi si dans le cas d'une distribution discrète avec domaine infini, on utilise la même procédure en fixant des bornes (troncation), ou bien si on utilise la méthode numérique générale qu'ils proposent pour tout autre type de distributions.

Dans le premier cas, il s'agit de deux troncations à la fois et on risque de perdre la précision dans l'approximation. Dans la deuxième méthode, si la fonction à intégrer n'est pas lisse, l'intégration numérique devient inefficace et il n'est pas approprié de l'utiliser pour des distributions discrètes [20]. On note aussi qu'aucun des exemple numériques proposés dans les références ne traite le cas des distributions discrètes avec la méthode NORTA.

Nous avons donc utilisé un autre algorithme plus efficace, connu sous le nom de la méthode de Brent [10], présenté par Press et al. [73] et qui combine la bipartition

et la réduction de l'intervalle avec l'interpolation quadratique inverse. Avec cet algorithme, on cherche à résoudre l'équation (3.19) pour chaque couple (i, j) , pour $1 \leq i < j \leq d$. On distingue alors deux cas pour $-1 \leq r_{i,j}^X < 0$ ou $0 < r_{i,j}^X \leq 1$ et dans chaque intervalle et à chaque itération i de l'algorithme on dispose de trois points a , b et c , où

- le point a représente la solution de l'itération $(i - 1)$ ou la valeur de b dans cette itération.
- le point b est la solution la plus proche de zéro après les $(i - 1)$ itérations.
- le point c est la nouvelle valeur après avoir coupé l'intervalle par la bisection, tel que $f_r(b)f_r(c) \leq 0$, $|f_r(b)| \leq |f_r(c)|$.

Initialement $a = c$, et à chaque itération l'intervalle $[b, c]$ contient la solution donnée par l'interpolation quadratique inverse et l'algorithme s'arrête dès que $f_r(b) = 0$ ou $|b - c| \leq \epsilon$, où ϵ est un seuil de tolérance fixé et la solution est b .

La même procédure est valable aussi si on utilise les corrélations linéaires et on trouvera la solution de l'équation (3.20).

3.3.3 Méthode M2 : calcul de la dérivée

La deuxième méthode est basée sur la fonction dérivée de g_r [4]. A partir de (3.33) et de la proposition 3.3.1 on calcule la fonction dérivée suivante :

$$g'_r(\rho) \approx \sum_{k=0}^{m_i-1} p_{i,k} \sum_{l=0}^{m_j-1} p_{j,l} \phi(z_{i,k}, z_{j,l}, \rho). \quad (3.34)$$

Le principe de la méthode est de commencer par choisir une valeur initiale, qu'on note ρ_0 . Nous proposons deux valeurs possibles pour ρ_0 , ce qui nous amène à distinguer deux variantes de la méthode M2. Nous avons la variante M2A pour $\rho_0 = 2 \sin(\pi r_{i,j}^X/6)$, et la variante M2B pour $\rho_0 = 0$. Par la suite, la fonction f_r ,

donnée en (3.19), est évaluée à ρ_0 et sur une séquence $\rho_k, k = 0, 1, \dots$, qui est croissante ou décroissante dépendamment du signe de $r_{i,j}^X$ et de $f_r(\rho_0)$.

Maintenant, le calcul de la double intégrale dans (3.13) se réduit à une intégration simple en utilisant le fait que :

$$g_r(\rho) = \int_{\rho_0}^{\rho} g'_r(t) dt + g_r(\rho_0), \quad \text{pour } \rho \in [0, 1], \quad (3.35)$$

et

$$g_r(\rho) = - \int_{\rho}^{\rho_0} g'_r(t) dt + g_r(\rho_0), \quad \text{pour } \rho \in [-1, 0]. \quad (3.36)$$

A cette étape, on arrive à réduire le calcul d'une intégrale double à une simple. Pour résoudre l'équation (3.19) et trouver les éléments de la matrice R_{ρ}^Z , nous proposons l'intégration numérique par la règle de Simpson [21]. Avec l'algorithme que nous utilisons, nous distinguons deux cas pour $r_{i,j}^X$, soit $-1 < r_{i,j}^X < 0$, ou $0 < r_{i,j}^X < 1$. Si par exemple $r_{i,j}^X$ est dans l'intervalle $[0, 1]$, on sait déjà que la solution $\rho_{i,j}^Z$ s'y trouve aussi et nous savons aussi que $\rho_0 = 2 \sin(\pi r_{i,j}^X / 6) < \rho_{i,j}^Z$. Pour cela, on fait l'intégration seulement dans l'intervalle $[\rho_0, 1 - \delta]$, avec $\delta > 0$ et assez petit, puisque la fonction dérivée g'_r n'est pas définie en 1 ou -1 . Sinon, pour $r_{i,j}^X$ dans l'intervalle $[-1, 0]$, l'intégration est faite sur $[-1 + \delta, \rho_0]$.

Nous allons expliquer l'algorithme seulement pour le premier cas, c.-à.-d. $r_{i,j}^X \in [0, 1]$, puisque les deux cas sont similaires et c'est seulement le domaine d'intégration qui change. La fonction f_r , définie en (3.19), que nous utilisons dans l'algorithme devient :

$$f_r(\rho) = \int_{\rho_0}^{\rho} g'_r(t) dt + g_r(\rho_0) - r_{i,j}^X \sigma_{F_i} \sigma_{F_j} - \mu_{F_i} \mu_{F_j}, \quad (3.37)$$

et pour une paire (i, j) et le $r_{i,j}^X$, nous cherchons le $\rho_{i,j}^Z$ (ρ pour simplifier) qui vérifie $f_r(\rho) = 0$. Voici l'algorithme utilisé

1. L'intervalle $[\rho_0, \rho_m = 1 - \delta]$ est départagé sur la grille $\rho_k = \rho_0 + 2kh$, $k = 0, 1, \dots, m$, c.-à-d., que $|1 - \delta - \rho_0| = 2hm$, où $h > 0$ est le pas utilisé dans la grille de façon à avoir $1 - 2h < \rho_m < 1$.
2. L'intégrale $\int_{\rho_0}^{\rho} g'_r(t) dt$ est alors calculée en petites parties sur la grille de calcul, soit $\int_{\rho_0}^{\rho_0+2kh} g'_r(t) dt$, $k = 0, 1, \dots, m$. Chaque intégrale est approximée par intégration numérique en utilisant la règle de Simpson, comme suit :

$$\int_{\rho_0}^{\rho_0+2kh} g'_r(t) dt \approx \frac{kh}{3} \left(g'_r(\rho_0) + 4g'_r(\rho_0 + kh) + g'_r(\rho_0 + 2kh) \right). \quad (3.38)$$

Ce qui nous donne la relation de récurrence suivante :

$$I_k = I_{k-1} + \frac{h}{3} \left(g'_r(\rho_0 + 2kh - 2h) + 4g'_r(\rho_0 + 2kh - h) + g'_r(\rho_0 + 2kh) \right), \quad (3.39)$$

pour $k = 0, 1, \dots, m$, où $I_0 = 0$.

L'algorithme s'arrête à une itération k où la solution est entre ρ_{k-1} et ρ_k , c.-à-d. la fonction f_r change de signe.

3. Par la suite, nous appliquons une interpolation quadratique pour trouver la solution cherchée $r_{i,j}^Z$. Si l'algorithme s'arrête à la première itération, nous appliquons l'interpolation linéaire, puisque nous ne disposons que des deux points ρ_0 et ρ_1 .

Dans la variante M2A, où l'intégration se fait à partir de ρ_0 jusqu'à $\rho_m = 1 - \delta$ et nécessite le calcul de $f_r(\rho)$ via la formule (3.33), on doit fixer le pas h . Dans la deuxième variante M2B, on fixe plutôt m et l'intégration se fait sur l'intervalle

$[0, \rho_m]$. Ici nous avons $f_r(0) = -r_{i,j}^X \sigma_{F_i} \sigma_{F_j}$ et qui ne nécessite pas de calcul. Le coût de cette méthode dépend essentiellement de la précision et du choix de m dans la grille de calcul. Nous présenterons dans les exemples numériques les résolutions avec les deux méthodes pour les comparaisons.

Cette méthode de résolution proposée pour le cas discret permet de contourner le problème et simplifier la procédure d'initialisation de NORTA et de réduire le problème à une intégration simple, où plusieurs méthodes numériques permettent de résoudre le problème avec une bonne exactitude, comme la règle de Simpson que nous avons utilisé dans notre cas. Dans le cas où les distributions ont des domaines finis, la résolution devient très précise. Tandis que dans le cas où le domaine est infini, la méthode présente une borne sur l'erreur d'approximation de la dérivée tout en tronquant le domaine [4].

Cette méthode présente aussi l'avantage d'être beaucoup moins coûteuse que les méthodes d'intégrations numériques qui pourraient être utilisées pour résoudre directement la double intégrale définie en (3.24). Notre méthode va nous éviter l'inversion des distributions dans la phase d'initialisation, et en plus elle ne nécessite qu'une intégration simple. Les exemples numériques qui ont illustré la méthode ont prouvé sa supériorité par rapport à l'intégration numérique [4]. En plus des corrélations de rang, la même méthode a été adaptée aux corrélations linéaires.

3.3.4 Méthode M3 : calcul de la fonction et de sa dérivée

Dans cette méthode, nous profitons du fait que nous pouvons évaluer la fonction f_r , ainsi que sa fonction dérivée ($f_r'(\rho) = g_r'(\rho)$), pour utiliser l'algorithme de Newton-Raphson pour résoudre l'équation (3.19).

A chaque itération k de l'algorithme, nous utilisons (3.33) pour calculer $f_r(\rho_k)$, et (3.26) pour calculer $f_r'(\rho_k)$ et la formule de récurrence définissant ρ_{k+1} est donnée

par :

$$\rho_{k+1} = \rho_k - \frac{f_r(\rho_k)}{f_r'(\rho_k)}, \quad (3.40)$$

ou en d'autres termes, l'algorithme choisit le point ρ_{k+1} de telle sorte que la droite qui passe par les points $(\rho_k, f_r(\rho_k))$ et $(\rho_{k+1}, 0)$ a pour pente celle de f_r au point ρ_k . L'algorithme s'arrête à une itération k par exemple, quand $|\rho_{k-1} - \rho_k| \leq \epsilon$, où ϵ est un seuil de tolérance bien choisi.

Cependant, il est connu que cet algorithme peut ne pas converger dans le cas où le point initial, ou la solution à chaque itération, n'est pas dans le voisinage de la racine recherchée. Nous avons donc utilisé une version modifiée présentée par Press et al. [73] qui garantit la convergence de l'algorithme en ajoutant la bipartition. Si par exemple lors d'une itération, la solution tombe à l'extérieur de l'intervalle de recherche de la racine, l'algorithme remplace cette solution par le milieu de l'intervalle. Pour garantir aussi une convergence rapide de l'algorithme nous avons utilisé comme valeur initiale $\rho_0 = 2 \sin(\pi r_{i,j}^X/6)$, au lieu du milieu de l'intervalle tel que suggéré dans l'algorithme original.

3.3.5 Exemples résolus avec les 3 méthodes

Nous avons testé nos 3 méthodes sur deux ensembles d'exemples, le premier avec des distributions marginales de domaine fini et le deuxième concerne le cas d'une distribution marginale avec domaine infini. Dans le premier ensemble d'exemples, les trois distributions marginales sont celles d'une loi binomiale, notée $\text{Bin}(n, p)$, avec la probabilité de succès $p = 1/2$ et nous avons choisi trois valeurs pour le nombre d'expériences n afin de vérifier les méthodes avec des domaines de différentes grosseurs. Nous avons donc pris de façon arbitraire $n = 3, 100$ et 1000 .

Le deuxième ensemble d'exemples est inspiré des données disponibles, où le nombre d'appels reçus par période de temps suit une loi binomiale négative de paramètres s et p , notée $\text{BinNeg}(s, p)$. Nous avons considéré le couple de variable (X_1, X_2) , pour les arrivées des périodes de temps (8h00-8h30) et (8h30-9h00), respectivement, dans le centre C1 de Bell. La distribution binomiale négative s'ajuste bien aux données de ce centre. Les paramètres (s, p) des deux variables sont estimés à partir des données et sont $s_1 = 15.68$, $s_2 = 60.21$, $p_1 = 0.3861$ et $p_2 = 0.6211$. La corrélation de rang entre X_1 et X_2 , estimée par les données, est de 0.43. Le domaine étant infini, nous utilisons la troncation. Nous faisons varier s_1 et s_2 ainsi que la corrélation $r_{1,2}^X$ pour couvrir plusieurs cas possibles. Pour les méthodes M1 et M3, nous avons utilisé deux seuils de tolérance ϵ , soit 10^{-2} et 10^{-4} .

Les calculs préliminaires nous ont révélé que dans le cas où la racine est proche de 1, il faut affiner la recherche. Avec la méthode M2A, nous avons considéré $\rho_m = 1 - \delta$, avec $\delta = 10^{-4}$, et $2h = 10^{-2}$. Pour la méthode M2B, nous avons fixé m à 5, ce qui donne $h = (1 - \delta)/10 \approx 0.2$.

Les tableaux 3.1, 3.2, 3.3 et 3.4 récapitulent les résultats pour les méthodes M1, M2A, M2B et M3, respectivement. Chacune des six parties de chaque tableau correspond à un exemple de distribution marginale. Dans chacun des cas et dans la première colonne, nous donnons les paramètres définissant chaque distribution avec les corrélations de rang extrêmes, c.-à-d., $r_{1,2}^X(-1)$ et $r_{1,2}^X(1)$, ainsi que le nombre de points $n = n_1 n_2$, où n_1 et n_2 représentent le cardinal du domaine de chaque distribution. Dans la deuxième colonne, nous présentons la corrélation donnée $r_{1,2}^X$. Pour chaque exemple de méthode et chaque cas traité, nous rapportons dans la troisième colonne la corrélation $\rho_{1,2}^Z$ trouvée, et le temps de calcul "CPU" dans la quatrième colonne. Pour chacune des méthodes, nous rapportons aussi les mesures suivantes : pour la méthode M1, N_1 désigne le nombre d'itérations de l'algorithme

de recherche de la racine et qui est aussi le nombre d'évaluations de la fonction g_r . Pour M2A, nous calculons $N_{g'}$ le nombre d'évaluations de g'_r ; pour NI3, le nombre N_3 désigne le nombre d'itérations de l'algorithme qui est aussi le nombre d'évaluations de g_r et de g'_r . Dans la sixième et la septième colonne, nous rapportons la corrélation entre X_1 et X_2 recalculée avec $\rho_{1,2}^Z$, ainsi que l'erreur relative. Toutes les expériences ont été faites sur une machine "Linux" avec un processeur 64-bit de vitesse 2.4 ghz.

Avec un seuil de tolérance $\epsilon = 10^{-4}$, les méthodes M1 et M3 donnent les meilleurs résultats dans tous les cas et nécessitent un petit nombre d'itérations. Par contre, avec un plus grand seuil où ϵ passe à 10^{-2} , le nombre d'itérations dans M3 n'augmente pas beaucoup, en général une itération de plus à l'exception du cas avec une loi binomiale $\text{Bin}(3, 1/2)$. Donc plus de précision n'exige pas un grand coût de calcul additionnel. Pour toutes les méthodes, les plus grandes erreurs se produisent dans l'exemple de la loi binomiale $\text{Bin}(3, 0.5)$, que nous examinons plus en détails par la suite. Excepté pour cet exemple avec $r_{1,2}^X = 0.98$ par exemple, M3 nécessite toujours moins de travail que M1, soit environ 30% en moyenne et généralement entre 20% et 45%. D'ailleurs, à deux exceptions dans le même exemple, M3 est plus précise que M1. Avec un seuil de 10^{-2} , avec la méthode M1 l'erreur relative est de 4% à 5% environ quand $r_{1,2}^X = 0.05$, mais l'erreur absolue est peut-être plus appropriée dans ce cas vu que la corrélation elle-même est assez petite. De point de vue rapidité d'exécution, M2A est généralement plus rapide que les autres méthodes; elle est également précise, à une exception. Cette méthode devient très efficace et rapide quand la distance entre ρ_0 et $\rho_{1,2}^Z$ est assez petite. Effectivement, dans le cas de la loi binomiale avec $n_1 = n_2 = 1000$, cette distance est très petite, et M2A est aussi précise que M1 ou M3 et elle est généralement plus rapide. La plus grande valeur observée de cette distance est de environ 0.09 ob-

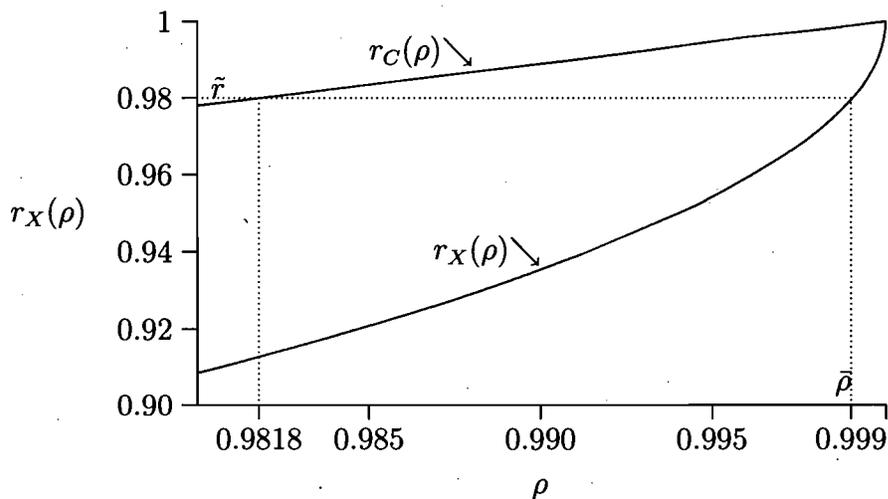


Figure 3.1 – La fonction $r_{1,2}^X(\rho)$ dans l'intervalle $[0.98, 1]$ avec le cas de la loi binomiale ($n_1 = n_2 = 3$), comparée à la fonction $r_{1,2}(\rho) = 6 \arcsin(\rho/2)/\pi$.

servée dans le cas de la loi binomiale avec $n_1 = n_2 = 3$ et $r_{1,2}^X = -0.5$. Évidemment, la méthode M2B ne bénéficie aucunement de cette distance, puisqu'on n'utilise pas $\rho_0 = 2 \sin(\pi r_{1,2}^X/6)$, mais l'intégration se fait sur l'intervalle $[0, \rho_m]$. Cette méthode affiche fréquemment de grandes erreurs qui tendent à augmenter avec l'augmentation du nombre de points n_1 et n_2 et avec une corrélation $r_{1,2}^X$ proche de 0 ou de 1.

Pour le cas de la binomiale avec $n_1 = n_2 = 3$ et $r_{1,2}^X = 0.98$, la racine calculée est $\rho_{1,2}^Z = 0.999041$ et son approximation est $2 \sin(0.98\pi/6) \approx 0.981808$. Le graphique 3.1 montre la fonction $r_{1,2}^X(\rho)$ sur l'intervalle $[0.98, 1]$. Dans un tel cas, la méthode M3 pratique la bipartition à toutes les itérations, puisque la valeur initiale n'est pas dans le voisinage de la solution. Par contre, la méthode M1 exige moins d'itérations que M3 dans ce cas-ci. Les approximations polynomiales d'ordre réduit de g_r font que la méthode M2 avec ces deux variantes souffre plus dans ce cas à cause de l'erreur d'intégration relativement grande. Avec la méthode M2B, nous ne gagnons

en exactitude qu'à partir de $m = 128$ où nous avons obtenu $\rho_{1,2}^Z = 0.9999$, et une erreur relative d'environ 1.7%; elle est due à l'erreur d'intégration et non pas à l'erreur d'interpolation. Par contre la variante M2A fait mieux que M2B; par exemple, avec $m = 16$, l'erreur relative est de 0.06%. Comme cité précédemment, le fait de prendre δ très petit ne règle pas mieux le problème : pour M2A, avec $\delta = 10^{-12}$ et $m = 2$, l'erreur relative grimpe jusqu'à 4.2%.

En résumé, avec un bon programme pour calculer la distribution normale bivariée (et ainsi f_r), nous recommandons M3; les deux variantes de M2 ne garantissent pas toujours une bonne précision selon le cas. Si par contre un tel programme n'est pas disponible, alors M2B est une solution plus facile et moins coûteuse, puisqu'elle nécessite seulement le calcul de f'_r .

Nous avons étudié donc le problème de corrélation avec la méthode NORTA pour les distributions marginales discrètes et nous avons proposé des méthodes efficaces et peu coûteuses en temps d'exécution. Nous avons pu obtenir une formule pour la fonction dérivée et nous avons pu développer et analyser de nouveaux algorithmes qui exploitent cette fonction dérivée. Pour les distributions marginales avec domaine infini, nos méthodes sont adaptées à cette situation avec des approximations et nous avons fourni des limites sur l'erreur résultante. Le lecteur peut se référer à Avramidis et al. [4] pour plus de détails sur ce sujet.

Pour récapituler les résultats des expériences numériques, nous pouvons conclure qu'au début, nous croyions que le fait d'utiliser la fonction dérivée pourrait nous faire gagner beaucoup par rapport à l'évaluation de fonction elle-même, ce qui favorisera la méthode M2. A notre surprise, les programmes qui calculent l'intégrale de la distribution de la loi normale bivariée (et ainsi f_r) sont assez efficaces et l'erreur est négligeable avec un coût de calcul assez modéré. Nous avons pu améliorer ces programmes et nous avons établi des programmes assez efficaces pour le calcul de

fonction dérivée. En résumé, si un bon code de programme est disponible pour calculer l'intégrale de la distribution de la loi normale bivariée, nous recommandons la méthode M3. Autrement, M2B est une solution facile et moins coûteuse.

Nous avons également exploité la convergence de la distance entre la fonction $r_{1,2}^X(\rho)$, et la fonction $(6/\pi) \arcsin(\rho/2)$ dans le cas des distributions marginales continues, quand nous avons pris comme valeur initiale $\rho_0 = 2 \sin(\pi r_{i,j}^X/6)$. Plus les dimensions des domaines ou supports de X_1 et X_2 (n_1 et n_2) augmentent, plus la distance diminue. Avec ce résultat nous avons constaté que la valeur $2 \sin(\pi r_{1,2}^X/6)$ est une approximation à la solution cherchée on l'a utilisée comme point initial dans les méthodes de résolution proposées par l'initialisation de NORTA.

3.3.6 Méthode M4 : Initialisation de NORTA avec un mélange de distributions discrètes et continues

Les méthodes que nous avons présentées dans ce chapitre et dans l'article [4] traitent du problème d'initialisation de NORTA dans le cas où les distributions marginales sont discrètes. Dans le cas où ces distributions sont continues, nous avons une résolution analytique directe si on utilise les corrélations de rang ou encore avec les corrélations linéaires et deux variables aléatoires qui sont uniformément distribuées. Dans le cas où on a un couple de variables aléatoires mixte, c.-à.-d. une discrète et l'autre continue, nous proposons une méthode de résolution qui utilise le même algorithme de recherche de la racine utilisé avec la méthode M1. Nous baptisons cette méthode M4.

Supposons que nous avons un couple de variables aléatoires (X_i, X_j) , avec X_i une variable continue et X_j une variable discrète. Supposons aussi que nous disposons des corrélations de rang $r_{i,j}^X$. Nous utilisons le fait que la variable X_j est discrète pour transformer la double intégrale dans la formule de g_r . Nous reprenons

Tableau 3.1 – Initialisation de NORTA : résultats avec la méthode M1.

	ϵ	$r_{1,2}^X$	$\rho_{1,2}^Z$	CPU (s)	N_1	$r_{1,2}^X(\rho_{1,2}^Z)$	Erreur rel. (%)
Binomiale $n_1 = n_2 = 3$ $p_1 = p_2 = 0.5$ $n = 16$ $r_{1,2}^X(-1) = -0.9241$ $r_{1,2}^X(1) = 1$	10^{-2}	-0.50	-0.6076	0.084×10^{-3}	5	-0.4999	0.049
		0.05	0.0603	0.063×10^{-3}	5	0.0499	0.115
		0.20	0.2387	0.063×10^{-3}	5	0.1991	0.475
		0.90	0.9760	0.102×10^{-3}	6	0.8999	0.009
		0.98	0.9999	0.050×10^{-3}	3	0.9935	1.382
	10^{-4}	-0.50	-0.6079	0.103×10^{-3}	6	-0.5000	< 0.001
		0.05	0.0604	0.082×10^{-3}	5	0.0500	< 0.001
		0.20	0.2399	0.084×10^{-3}	5	0.2000	< 0.001
		0.90	0.9760	0.118×10^{-3}	6	0.9000	< 0.001
		0.98	0.9990	0.130×10^{-3}	7	0.9800	< 0.001
Binomiale $n_1 = n_2 = 100$ $p_1 = p_2 = 0.5$ $n = 10201$ $r_{1,2}^X(-1) = -0.9971$ $r_{1,2}^X(1) = 1$	10^{-2}	-0.50	-0.5194	0.093	4	-0.4991	0.175
		0.05	0.0551	0.069	3	0.0524	4.703
		0.20	0.2099	0.092	4	0.1999	0.036
		0.90	0.9107	0.085	4	0.8996	0.043
		0.98	0.9861	0.061	3	0.9811	0.114
	10^{-4}	-0.50	-0.5203	0.119	5	-0.5000	< 0.001
		0.05	0.0526	0.094	4	0.0500	< 0.001
		0.20	0.2099	0.117	5	0.2000	< 0.001
		0.90	0.9111	0.113	5	0.9000	< 0.001
		0.98	0.9851	0.098	5	0.9800	< 0.001
Binomiale $n_1 = n_2 = 1000$ $p_1 = p_2 = 0.5$ $n = 1002001$ $r_{1,2}^X(-1) = -0.9997$ $r_{1,2}^X(1) = 1$	10^{-2}	-0.50	-0.5171	5.503	4	-0.4992	0.159
		0.05	0.0500	4.410	3	0.0477	4.511
		0.20	0.2091	5.858	4	0.1999	0.031
		0.90	0.9082	4.703	4	0.8999	0.014
		0.98	0.9803	3.153	3	0.9780	0.207
	10^{-4}	-0.50	-0.5179	6.995	5	-0.5000	< 0.001
		0.05	0.0524	6.193	4	0.0500	< 0.001
		0.20	0.2091	7.509	5	0.2000	< 0.001
		0.90	0.9083	5.848	5	0.9000	< 0.001
		0.98	0.9821	4.289	4	0.9800	< 0.001
Binomiale négative $r_1 = 1.568$ $p_1 = 0.3861$ $r_2 = 6.021$ $p_2 = 0.6211$ $n = 768$ $r_{1,2}^X(-1) = -0.9738$ $r_{1,2}^X(1) = 0.9652$	10^{-2}	-0.50	-0.5330	6.57×10^{-3}	4	-0.4960	0.209
		0.05	0.0518	4.33×10^{-3}	3	0.0478	4.412
		0.43	0.4614	5.49×10^{-3}	4	0.4299	0.030
		0.90	0.9323	4.46×10^{-3}	3	0.8986	0.151
		0.96	0.9895	5.88×10^{-3}	4	0.9593	0.076
	10^{-4}	-0.50	-0.5341	8.31×10^{-3}	5	-0.5000	< 0.001
		0.05	0.0542	5.78×10^{-3}	4	0.0500	< 0.001
		0.43	0.4616	7.14×10^{-3}	5	0.4300	< 0.001
		0.90	0.9336	5.97×10^{-3}	4	0.9000	< 0.001
		0.96	0.9903	9.59×10^{-3}	6	0.9600	< 0.001
Binomiale négative $r_1 = 15.68$ $p_1 = 0.3861$ $r_2 = 60.21$ $p_2 = 0.6211$ $n = 6560$ $r_{1,2}^X(-1) = -0.9971$ $r_{1,2}^X(1) = 0.9989$	10^{-2}	-0.50	-0.5177	0.053	4	-0.4993	0.148
		0.05	0.0501	0.041	3	0.0478	4.481
		0.43	0.4467	0.056	4	0.4298	0.042
		0.90	0.9090	0.053	4	0.8999	0.017
		0.98	0.9811	0.037	3	0.9778	0.229
	10^{-4}	-0.50	-0.5184	0.070	5	-0.5000	< 0.001
		0.05	0.0524	0.052	4	0.0500	< 0.001
		0.43	0.4469	0.066	5	0.4300	< 0.001
		0.90	0.9092	0.067	5	0.9000	< 0.001
		0.98	0.9832	0.054	4	0.9800	< 0.001
Binomiale négative $r_1 = 156.7$ $p_1 = 0.3861$ $r_2 = 602.1$ $p_2 = 0.6211$ $n = 189912$ $r_{1,2}^X(-1) = -0.9997$ $r_{1,2}^X(1) = 0.9999$	10^{-2}	-0.50	-0.5169	1.301	4	-0.4992	0.159
		0.05	0.0500	0.871	3	0.0478	4.491
		0.43	0.4464	1.273	4	0.4298	0.042
		0.90	0.9080	1.255	4	0.8999	0.013
		0.98	0.9802	0.877	3	0.9780	0.199
	10^{-4}	-0.50	-0.5177	1.639	5	-0.5000	< 0.001
		0.05	0.0524	1.236	4	0.0500	< 0.001
		0.43	0.4465	1.616	5	0.4300	< 0.001
		0.90	0.9081	1.611	5	0.9000	< 0.001
		0.98	0.9819	1.190	4	0.9800	< 0.001

Tableau 3.2 – Initialisation de NORTA : résultats avec la méthode M2A, où $\delta = 10^{-4}$ et $2h$ est proche de 10^{-2} .

	$r_{1,2}^X$	$\rho_{1,2}^Z$	CPU (s)	$N_{g'}$	$r_{1,2}^X(\rho_{1,2}^Z)$	Erreur rel. (%)
Binomiale	-0.50	-0.6079	0.062×10^{-3}	21	-0.5000	< 0.001
$n_1 = n_2 = 3$	0.05	0.0604	0.032×10^{-3}	5	0.0500	< 0.001
$p_1 = p_2 = 0.5$	0.20	0.2399	0.045×10^{-3}	9	0.2000	< 0.001
$n = 16$	0.90	0.9760	0.067×10^{-3}	17	0.8999	0.011
$r_{1,2}^X(-1) = -0.9241$	0.98	0.9962	0.031×10^{-3}	5	0.9602	2.024
$r_{1,2}^X(1) = 1$						
Binomiale	-0.50	-0.5203	0.032	5	-0.5000	< 0.001
$n_1 = n_2 = 100$	0.05	0.0526	0.032	5	0.0500	< 0.001
$p_1 = p_2 = 0.5$	0.20	0.2099	0.032	5	0.2000	< 0.001
$n = 10201$	0.90	0.9111	0.032	5	0.9000	< 0.001
$r_{1,2}^X(-1) = -0.9971$	0.98	0.9851	0.028	5	0.9810	0.099
$r_{1,2}^X(1) = 1$						
Binomiale	-0.50	-0.5179	2.17	5	-0.5000	< 0.001
$n_1 = n_2 = 1000$	0.05	0.0524	2.48	5	0.0500	< 0.001
$p_1 = p_2 = 0.5$	0.20	0.2091	2.40	5	0.2000	< 0.001
$n = 1002001$	0.90	0.9083	2.01	5	0.9000	< 0.001
$r_{1,2}^X(-1) = -0.9997$	0.98	0.9821	1.89	5	0.9800	< 0.001
$r_{1,2}^X(1) = 1$						
Binomiale négative	-0.50	-0.5341	2.27×10^{-3}	5	-0.5000	< 0.001
$r_1 = 1.568$	0.05	0.0542	2.09×10^{-3}	5	0.0500	< 0.001
$p_1 = 0.3861$	0.43	0.4616	1.95×10^{-3}	5	0.4300	< 0.001
$r_2 = 6.021$	0.90	0.9336	2.29×10^{-3}	7	0.9000	< 0.001
$p_2 = 0.6211$	0.96	0.9903	2.33×10^{-3}	7	0.9600	< 0.001
$n = 768$						
$r_{1,2}^X(-1) = -0.9738$						
$r_{1,2}^X(1) = 0.9652$						
Binomiale négative	-0.50	-0.5184	0.019	5	-0.5000	< 0.001
$r_1 = 15.68$	0.05	0.0524	0.019	5	0.0500	< 0.001
$p_1 = 0.3861$	0.43	0.4469	0.018	5	0.4300	< 0.001
$r_2 = 60.21$	0.90	0.9092	0.019	5	0.9000	< 0.001
$p_2 = 0.6211$	0.98	0.9832	0.018	5	0.9800	< 0.001
$n = 6560$						
$r_{1,2}^X(-1) = -0.9971$						
$r_{1,2}^X(1) = 0.9989$						
Binomiale négative	-0.50	-0.5177	0.50	5	-0.5000	< 0.001
$r_1 = 156.7$	0.05	0.0524	0.46	5	0.0500	< 0.001
$p_1 = 0.3861$	0.43	0.4465	0.47	5	0.4300	< 0.001
$r_2 = 602.1$	0.90	0.9081	0.47	5	0.9000	< 0.001
$p_2 = 0.6211$	0.98	0.9819	0.45	5	0.9800	< 0.001
$n = 189912$						
$r_{1,2}^X(-1) = -0.9997$						
$r_{1,2}^X(1) = 0.9999$						

Tableau 3.3 – Initialisation de NORTA : résultats avec la méthode M2B où $m = 5$, $\delta = 10^{-4}$ ($h = 0.09999$).

	$r_{1,2}^X$	$\rho_{1,2}^Z$	CPU (s)	$N_{g'}$	$r_{1,2}^X(\rho_{1,2}^Z)$	Erreur rel. (%)
Binomiale	-0.50	-0.6078	0.024×10^{-3}	9	-0.5000	< 0.001
$n_1 = n_2 = 3$	0.05	0.0601	0.010×10^{-3}	3	0.0497	0.514
$p_1 = p_2 = 0.5$	0.20	0.2389	0.015×10^{-3}	5	0.1999	0.026
$n = 16$	0.90	0.8485	0.029×10^{-3}	11	0.7409	17.676
$r_{1,2}^X(-1) = -0.9241$	0.98	0.8642	0.029×10^{-3}	11	0.7565	22.805
$r_{1,2}^X(1) = 1$						
Binomiale	-0.50	-0.5202	9.65×10^{-3}	7	-0.4999	0.029
$n_1 = n_2 = 100$	0.05	0.0525	4.20×10^{-3}	3	0.0499	0.172
$p_1 = p_2 = 0.5$	0.20	0.2099	6.95×10^{-3}	5	0.2000	< 0.001
$n = 10201$	0.90	0.8718	15.60×10^{-3}	11	0.8582	4.641
$r_{1,2}^X(-1) = -0.9971$	0.98	0.9142	15.80×10^{-3}	11	0.9034	7.821
$r_{1,2}^X(1) = 1$						
Binomiale	-0.50	-0.5178	1.023	7	-0.4998	0.030
$n_1 = n_2 = 1000$	0.05	0.0523	0.44	3	0.0499	0.157
$p_1 = p_2 = 0.5$	0.20	0.2091	0.73	5	0.2000	< 0.001
$n = 1002001$	0.90	0.8982	1.64	11	0.8892	1.202
$r_{1,2}^X(-1) = -0.9997$	0.98	0.9625	1.62	11	0.9586	2.187
$r_{1,2}^X(1) = 1$						
Binomiale négative	-0.50	-0.5340	0.76×10^{-3}	7	-0.4998	0.031
$r_1 = 1.568$	0.05	0.0541	0.34×10^{-3}	3	0.0499	0.157
$p_1 = 0.3861$	0.43	0.4614	0.75×10^{-3}	7	0.4299	0.028
$r_2 = 6.021$	0.90	0.9567	1.19×10^{-3}	11	0.9246	2.736
$p_2 = 0.6211$	0.96	1.00	1.13×10^{-3}	11	0.9652	0.545
$n = 768$						
$r_{1,2}^X(-1) = -0.9738$						
$r_{1,2}^X(1) = 0.9652$						
Binomiale négative	-0.50	-0.5183	6.38×10^{-3}	7	-0.4999	0.029
$r_1 = 15.68$	0.05	0.0523	2.72×10^{-3}	3	0.0499	0.160
$p_1 = 0.3861$	0.43	0.4468	6.44×10^{-3}	7	0.4299	0.019
$r_2 = 60.21$	0.90	0.8995	9.77×10^{-3}	11	0.8896	1.150
$p_2 = 0.6211$	0.98	0.9644	9.74×10^{-3}	11	0.9595	2.091
$n = 6560$						
$r_{1,2}^X(-1) = -0.9971$						
$r_{1,2}^X(1) = 0.9989$						
Binomiale négative	-0.50	-0.5176	0.18	7	-0.4998	0.030
$r_1 = 156.7$	0.05	0.0523	0.078	3	0.0499	0.155
$p_1 = 0.3861$	0.43	0.4465	0.18	7	0.4296	0.020
$r_2 = 602.1$	0.90	0.9077	0.29	11	0.8996	0.041
$p_2 = 0.6211$	0.98	0.9816	0.28	11	0.9797	0.035
$n = 189912$						
$r_{1,2}^X(-1) = -0.9997$						
$r_{1,2}^X(1) = 0.9999$						

Tableau 3.4 – Initialisation de NORTA : résultats avec la méthode M3.

	ϵ	$r_{1,2}^X$	$\rho_{1,2}^2$	CPU (s)	N_1	$r_{1,2}^X(\rho_{1,2}^2)$	Erreur rel. (%)
Binomiale $n_1 = n_2 = 3$ $p_1 = p_2 = 0.5$ $n = 16$ $r_{1,2}^X(-1) = -0.9241$ $r_{1,2}^X(1) = 1$	10^{-2}	-0.50	-0.6079	0.038×10^{-3}	2	-0.5000	< 0.001
		0.05	0.0604	0.020×10^{-3}	1	0.0500	0.004
		0.20	0.2399	0.040×10^{-3}	2	0.2000	< 0.001
		0.90	0.9767	0.060×10^{-3}	3	0.9013	0.142
		0.98	0.9922	0.137×10^{-3}	7	0.9429	3.783
	10^{-4}	-0.50	-0.6079	0.064×10^{-3}	3	-0.5000	< 0.001
		0.05	0.0604	0.053×10^{-3}	2	0.0500	< 0.001
		0.20	0.2399	0.054×10^{-3}	2	0.2000	< 0.001
		0.90	0.9760	0.106×10^{-3}	5	0.9000	< 0.001
		0.98	0.9990	0.222×10^{-3}	12	0.9800	< 0.001
Binomiale $n_1 = n_2 = 100$ $p_1 = p_2 = 0.5$ $n = 10201$ $r_{1,2}^X(-1) = -0.9971$ $r_{1,2}^X(1) = 1$	10^{-2}	-0.50	-0.5203	0.028	1	-0.5000	< 0.001
		0.05	0.0526	0.029	1	0.0500	< 0.001
		0.20	0.2099	0.028	1	0.2000	< 0.001
		0.90	0.9111	0.025	1	0.9000	< 0.001
		0.98	0.9851	0.022	1	0.9800	< 0.001
	10^{-4}	-0.50	-0.5203	0.054	2	-0.5000	< 0.001
		0.05	0.0526	0.056	2	0.0500	< 0.001
		0.20	0.2099	0.053	2	0.2000	< 0.001
		0.90	0.9111	0.047	2	0.9000	< 0.001
		0.98	0.9851	0.044	2	0.9800	< 0.001
Binomiale $n_1 = n_2 = 1000$ $p_1 = p_2 = 0.5$ $n = 1002001$ $r_{1,2}^X(-1) = -0.9997$ $r_{1,2}^X(1) = 1$	10^{-2}	-0.50	-0.5179	1.642	1	-0.5000	< 0.001
		0.05	0.0524	1.924	1	0.0500	< 0.001
		0.20	0.2091	1.831	1	0.2000	< 0.001
		0.90	0.9083	1.387	1	0.9000	< 0.001
		0.98	0.9821	1.214	1	0.9800	< 0.001
	10^{-4}	-0.50	-0.5179	3.360	2	-0.5000	< 0.001
		0.05	0.0524	1.918	1	0.0500	< 0.001
		0.20	0.2091	1.771	1	0.2000	< 0.001
		0.90	0.9083	2.796	2	0.9000	< 0.001
		0.98	0.9821	2.550	2	0.9800	< 0.001
Binomiale négative $r_1 = 1.568$ $p_1 = 0.3861$ $r_2 = 6.021$ $p_2 = 0.6211$ $n = 768$ $r_{1,2}^X(-1) = -0.9738$ $r_{1,2}^X(1) = 0.9652$	10^{-2}	-0.50	-0.5341	3.79×10^{-3}	2	-0.5000	< 0.001
		0.05	0.0542	1.78×10^{-3}	1	0.0500	< 0.001
		0.43	0.4616	3.15×10^{-3}	2	0.4300	< 0.001
		0.90	0.9336	3.45×10^{-3}	2	0.9000	< 0.001
		0.96	0.9902	3.46×10^{-3}	2	0.9616	< 0.001
	10^{-4}	-0.50	-0.5341	3.77×10^{-3}	2	-0.5000	< 0.001
		0.05	0.0542	3.42×10^{-3}	2	0.0500	< 0.001
		0.43	0.4616	3.16×10^{-3}	2	0.4300	< 0.001
		0.90	0.9336	5.13×10^{-3}	3	0.9000	< 0.001
		0.96	0.9903	5.10×10^{-3}	3	0.9600	< 0.001
Binomiale négative $r_1 = 15.68$ $p_1 = 0.3861$ $r_2 = 60.21$ $p_2 = 0.6211$ $n = 6560$ $r_{1,2}^X(-1) = -0.9971$ $r_{1,2}^X(1) = 0.9989$	10^{-2}	-0.50	-0.5184	0.017	1	-0.5000	< 0.001
		0.05	0.0524	0.016	1	0.0500	< 0.001
		0.43	0.4469	0.016	1	0.4300	< 0.001
		0.90	0.9092	0.016	1	0.9000	< 0.001
		0.98	0.9832	0.015	1	0.9800	< 0.001
	10^{-4}	-0.50	-0.5184	0.031	2	-0.5000	< 0.001
		0.05	0.0524	0.016	1	0.0500	< 0.001
		0.43	0.4469	0.031	2	0.4300	< 0.001
		0.90	0.9092	0.031	2	0.9000	< 0.001
		0.98	0.9832	0.028	2	0.9800	< 0.001
Binomiale négative $r_1 = 156.7$ $p_1 = 0.3861$ $r_2 = 602.1$ $p_2 = 0.6211$ $n = 189912$ $r_{1,2}^X(-1) = -0.9997$ $r_{1,2}^X(1) = 0.9999$	10^{-2}	-0.50	-0.5177	0.393	1	-0.5000	< 0.001
		0.05	0.0524	0.370	1	0.0500	< 0.001
		0.43	0.4465	0.378	1	0.4300	< 0.001
		0.90	0.9081	0.383	1	0.9000	< 0.001
		0.98	0.9819	0.332	1	0.9800	< 0.001
	10^{-4}	-0.50	-0.5177	0.394	1	-0.5000	< 0.001
		0.05	0.0524	0.369	1	0.0500	< 0.001
		0.43	0.4465	0.366	1	0.4300	< 0.001
		0.90	0.9081	0.740	2	0.9000	< 0.001
		0.98	0.9819	0.689	2	0.9800	< 0.001

les équations (3.12) et (3.13) et nous avons :

$$\begin{aligned} g_r(\rho_{i,j}^Z) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_i(F_i^{-1}(\Phi(z_i))) F_j(F_j^{-1}(\Phi(z_j))) \phi(z_i, z_j, \rho_{i,j}^Z) dz_i dz_j \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\Phi(z_i) F_j(F_j^{-1}(\Phi(z_j))) \phi(z_i) \phi\left(\frac{z_j - \rho_{i,j}^Z z_i}{\sqrt{1 - (\rho_{i,j}^Z)^2}}\right)}{\sqrt{1 - (\rho_{i,j}^Z)^2}} dz_i dz_j, \end{aligned}$$

en utilisant le fait que

$$\phi(z_i, z_j, \rho_{i,j}^Z) = \frac{\phi(z_i) \phi\left(\frac{z_j - \rho_{i,j}^Z z_i}{\sqrt{1 - (\rho_{i,j}^Z)^2}}\right)}{\sqrt{1 - (\rho_{i,j}^Z)^2}},$$

où ϕ est la densité de la loi normale standard univariée. Nous avons donc :

$$\begin{aligned} g_r(\rho_{i,j}^Z) &= \int_{-\infty}^{\infty} \frac{\Phi(z_i) \phi(z_i)}{\sqrt{1 - (\rho_{i,j}^Z)^2}} \left[\int_{-\infty}^{\infty} F_j(F_j^{-1}(\Phi(z_j))) \phi\left(\frac{z_j - \rho_{i,j}^Z z_i}{\sqrt{1 - (\rho_{i,j}^Z)^2}}\right) dz_j \right] dz_i, \\ &= \int_{-\infty}^{\infty} \frac{\Phi(z_i) \phi(z_i)}{\sqrt{1 - (\rho_{i,j}^Z)^2}} \left[\sum_{l=-\infty}^{\infty} f_{j,l} \int_{z_{j,l-1}}^{z_{j,l}} \phi\left(\frac{z_j - \rho_{i,j}^Z z_i}{\sqrt{1 - (\rho_{i,j}^Z)^2}}\right) dz_j \right] dz_i \\ &= \sum_{l=-\infty}^{\infty} f_{j,l} \left[\int_{-\infty}^{\infty} \Phi(z_i) \phi(z_i) \left(\Phi\left(\frac{z_{j,l} - \rho_{i,j}^Z z_i}{\sqrt{1 - (\rho_{i,j}^Z)^2}}\right) \right. \right. \\ &\quad \left. \left. - \Phi\left(\frac{z_{j,l-1} - \rho_{i,j}^Z z_i}{\sqrt{1 - (\rho_{i,j}^Z)^2}}\right) \right) dz_i \right]. \end{aligned}$$

Nous pouvons effectuer un changement de variable pour simplifier l'intégration

et la ramener sur le domaine $[0, 1]$ en posant $u_i = \Phi(z_i)$. nous aurons alors :

$$g_r(\rho_{i,j}^Z) = \sum_{l=-\infty}^{\infty} f_{j,l} \left[\int_0^1 u_i \left(\Phi \left(\frac{z_{j,l} - \rho_{i,j}^Z \Phi^{-1}(u_i)}{\sqrt{1 - (\rho_{i,j}^Z)^2}} \right) - \Phi \left(\frac{z_{j,l-1} - \rho_{i,j}^Z \Phi^{-1}(u_i)}{\sqrt{1 - (\rho_{i,j}^Z)^2}} \right) \right) du_i \right]. \quad (3.41)$$

La résolution repose donc sur une simple intégration et une sommation sur le domaine de distribution F_j qui est discrète. La fonction à intégrer étant assez lisse, nous pouvons résoudre l'intégrale numériquement. Nous pouvons facilement remarquer que la loi de X_i ne paraît pas dans la formule de $g_r(\rho_{i,j}^Z)$, il suffit donc de spécifier seulement la corrélation entre X_i et X_j et la loi discrète de X_j . Il faut noter que ceci est vrai seulement dans la phase d'initialisation, puisque nous n'avons pas à calculer la moyenne et l'écart-type de $F_i(X_i)$ comme dans le cas d'une distribution discrète, et nous avons $\mu_{F_i} = 0.5$ et $\sigma_{F_i} = \sqrt{1/12}$. Par contre il faut spécifier la distribution de F_i pour générer des valeurs aléatoires de X_i en utilisant la relation (3.3).

En ce qui concerne la même méthode avec les corrélations linéaires $\rho_{i,j}^X$, nous avons une formule semblable où il faut spécifier la distribution de la variable aléatoire continue X_i . Nous avons :

$$g_l(\rho_{i,j}^Z) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_i^{-1}(\Phi(z_i)) F_j^{-1}(\Phi(z_j)) \phi(z_i, z_j, \rho_{i,j}^Z) dz_i dz_j,$$

et avec les mêmes transformations effectuées dans le cas de $r_{i,j}^X$, nous obtenons à la

fin une formulation avec une simple intégration sur le domaine $[0, 1]$:

$$g_l(\rho_{i,j}^Z) = \sum_{l=-\infty}^{\infty} y_{j,l} \left[\int_0^1 F_i^{-1}(u_i) \left(\Phi \left(\frac{z_{j,l} - \rho_{i,j}^Z \Phi^{-1}(u_i)}{\sqrt{1 - (\rho_{i,j}^Z)^2}} \right) - \Phi \left(\frac{z_{j,l-1} - \rho_{i,j}^Z \Phi^{-1}(u_i)}{\sqrt{1 - (\rho_{i,j}^Z)^2}} \right) \right) du_i \right]. \quad (3.42)$$

Pour essayer la méthodes avec des exemples numériques, nous avons considéré 3 cas pour des couples de variables aléatoires mixtes (X_1 continue et X_2 discrète), soit :

Cas 1 : X_1 suit une loi exponentielle de paramètre $\lambda_1 = 25$ et X_2 suit une loi binomiale de paramètres $n_2 = 100$ et $p_2 = 0.5$.

Cas 2 : X_1 suit une loi gamma de paramètres $\alpha_1 = 20$ et $\beta_1 = 30$ et X_2 suit une loi binomiale négative de paramètres $s_2 = 60.21$ et $p_2 = 0.611$.

Cas 3 : X_1 suit une loi normale de paramètres $\mu_1 = 25$ $\sigma_1 = 10$ et X_2 suit une loi de Poisson de paramètre $\lambda_2 = 25$.

Nous avons utilisé la quadrature Gauss-Legendre pour résoudre l'intégrale donnée en (3.41) ou en (3.42) et la librairie en Java de Flanagan [30]. Nous avons testé la méthode d'initialisation avec les corrélations de rang et les corrélations linéaires. Nous avons considéré des valeurs de $r_{1,2}^X$ et $\rho_{1,2}^X$ négatives, proches de zéro, positives et proche de 1 pour couvrir le maximum de valeurs possibles des corrélations. Le seuil de tolérance pour l'algorithme de recherche étant de deux valeurs possibles $\epsilon = 10^{-1}, 10^{-4}$.

Nous avons calculé le temps moyen d'exécution en CPU pour chaque cas, le nombre d'itérations qui est aussi le nombre de fois que la fonction g_r ou g_l est évaluée en utilisant l'intégration numérique, ainsi que l'erreur relative.

La méthode semble être rapide en temps d'exécution avec les deux types de

corrélations, puisque le temps moyen calculé est généralement en bas d'un centième de seconde. Le nombre d'itérations n'augmente que de 1 ou 2 si on passe d'un seuil de tolérance de $\epsilon = 10^{-2}$ à $\epsilon = 10^{-4}$. L'erreur relative est aussi faible, à l'exception du cas où la corrélation est égale à 0.05 et dans ce cas-ci l'erreur absolue serait plus appropriée puisque c'est une valeur proche de zéro. Les tableaux 3.5 et 3.6 montrent que les méthodes fournissent des bons résultats pour les deux types de corrélations surtout avec un seuil de tolérance $\epsilon = 10^{-4}$.

Nous pouvons résumer tous les cas possibles comme suit : dans le cas où les deux distributions sont discrètes, les méthodes de résolution M1, M2A, M2B et M3 sont valables pour les deux types de corrélations, avec une recommandation pour les méthodes M1 et M2B, selon ce qu'on veut comme précision, par exemple. Dans le cas où nous disposons seulement des corrélations linéaires et que les distributions marginales sont continues mais pas uniformes, alors l'initialisation de NORTA est effectuée en résolvant numériquement la double intégrale de l'équation (3.6) et en utilisant un algorithme de recherche de la racine, comme celui utilisé avec la méthode M1. Cario et Nelson [18] ont utilisé la quadrature gaussienne pour résoudre la double intégrale. Dans le cas où les distributions marginales sont continues et nous disposons des corrélations de rang, ou que ces distributions sont continues et uniformément distribuées et nous disposons des corrélations linéaires, alors la résolution devient analytique avec la formulation (3.7). Dans le cas où nous avons un mélange de distributions continues et de distributions discrètes, la méthode M4 sera suffisante pour les deux types de corrélations. Toutes ces situations sont résumées dans le tableau 3.7.

Tableau 3.5 – Initialisation de NORTA : résultats avec la méthode M4 en utilisant les corrélations de rang.

	ϵ	$r_{1,2}^X$	$\rho_{1,2}^Z$	CPU (s)	N_1	$r_{1,2}^X(\rho_{1,2}^Z)$	Erreur rel. (%)
$-X_1$: Exponentielle $\lambda_1 = 25$ $-X_2$: Binomiale $n_2 = 100$ $p_2 = 0.65$ $-r_{1,2}^X(-1) = -0.9918$ $-r_{1,2}^X(1) = 0.9934$	10^{-2}	-0.50	-0.5189	0.0033	4	-0.4995	0.106
		0.05	0.0503	0.0023	3	0.0479	4.293
		0.20	0.2099	0.0032	4	0.1999	0.025
		0.90	0.9110	0.0023	3	0.9010	0.109
		0.98	0.9866	0.0023	3	0.9827	0.275
	10^{-4}	-0.50	-0.5194	0.0042	5	-0.5000	< 0.001
		0.05	0.0526	0.0032	4	0.0500	< 0.001
		0.20	0.2099	0.0041	5	0.2000	< 0.001
		0.90	0.9101	0.0041	5	0.9000	< 0.001
		0.98	0.9841	0.0041	5	0.9800	< 0.001
$-X_1$: Gamma $\alpha_1 = 20$ $\beta_1 = 30$ $-X_2$: Bin. négative $s_2 = 60.21$ $p_2 = 0.611$ $-r_{1,2}^X(-1) = -1$ $-r_{1,2}^X(1) = 0.9746$	10^{-2}	-0.50	-0.5183	0.0036	4	-0.4997	0.053
		0.05	0.0513	0.0026	3	0.0489	2.300
		0.20	0.2102	0.0026	3	0.2006	0.280
		0.90	0.9088	0.0047	5	0.9000	< 0.001
		0.96	0.9633	0.0047	5	0.9587	0.135
	10^{-4}	-0.50	-0.5186	0.0046	5	-0.5000	< 0.001
		0.05	0.0525	0.0036	4	0.0500	< 0.001
		0.20	0.2096	0.0036	4	0.2000	< 0.001
		0.90	0.9088	0.0047	5	0.9000	< 0.001
		0.96	0.9644	0.0058	6	0.9600	< 0.001
$-X_1$: Normale $\mu_1 = 25$ $\sigma_1 = 10$ $-X_2$: Poisson $\lambda = 25$ $-r_{1,2}^X(-1) = -0.9975$ $-r_{1,2}^X(1) = 0.9871$	10^{-2}	-0.50	-0.5185	0.0024	4	-0.4992	0.163
		0.05	0.0507	0.0017	3	0.0482	3.657
		0.20	0.2098	0.0024	4	0.2000	< 0.001
		0.90	0.9118	0.0018	3	0.9018	0.201
		0.98	0.9836	0.0024	4	0.9794	0.058
	10^{-4}	-0.50	-0.5193	0.0030	5	-0.5000	< 0.001
		0.05	0.0526	0.0024	4	0.0500	< 0.001
		0.20	0.2099	0.0024	4	0.2000	< 0.001
		0.90	0.9101	0.0031	5	0.9000	< 0.001
		0.98	0.9841	0.0039	6	0.9800	< 0.001

Tableau 3.6 – Initialisation de NORTA : résultats avec la méthode M4 en utilisant les corrélations linéaires.

	ϵ	$\rho_{1,2}^X$	$\rho_{1,2}^Z$	CPU (s)	N_1	$r_{1,2}^X(\rho_{1,2}^Z)$	Erreur rel. (%)
$-X_1$: Exponentielle $\lambda_1 = 25$ $-X_2$: Binomiale $n_2 = 100$ $p_2 = 0.5$ $-r_{1,2}^X(-1) = -0.9134$ $-r_{1,2}^X(1) = 0.8670$	10^{-2}	-0.50	-0.5365	0.0038	4	-0.5002	0.044
		0.05	0.0824	0.0049	5	0.0500	< 0.001
		0.20	0.2531	0.0038	4	0.2018	0.886
		0.50	0.5884	0.0038	4	0.5000	< 0.001
		0.80	0.9277	0.0028	3	0.8017	0.211
	10^{-4}	-0.50	-0.5362	0.0049	5	-0.5000	< 0.001
		0.05	0.0824	0.0049	5	0.0500	< 0.001
		0.20	0.2511	0.0049	5	0.2000	< 0.001
		0.50	0.5885	0.0038	4	0.5000	< 0.001
		0.80	0.9258	0.0039	4	0.8000	< 0.001
$-X_1$: Gamma $\alpha_1 = 20$ $\beta_1 = 30$ $-X_2$: Bin. négative $s_2 = 60.21$ $p_2 = 0.611$ $-r_{1,2}^X(-1) = -0.9994$ $-r_{1,2}^X(1) = 0.9469$	10^{-2}	-0.50	-0.5067	0.0343	4	-0.4998	0.044
		0.05	0.0528	0.0253	3	0.0489	2.264
		0.20	0.2062	0.0255	3	0.1998	0.101
		0.50	0.5098	0.0341	4	0.4992	0.167
		0.90	0.9149	0.0430	5	0.9000	< 0.001
	10^{-4}	-0.50	-0.5070	0.0424	5	-0.5000	< 0.001
		0.05	0.0540	0.0342	4	0.0500	< 0.001
		0.20	0.2064	0.0428	5	0.2000	< 0.001
		0.50	0.5106	0.0428	5	0.5000	< 0.001
		0.90	0.9149	0.0428	5	0.9000	< 0.001
$-X_1$: Normale $\mu_1 = 25$ $\sigma_1 = 10$ $-X_2$: Poisson $\lambda = 25$ $-r_{1,2}^X(-1) = -0.9842$ $-r_{1,2}^X(1) = 0.9954$	10^{-2}	-0.50	-0.5080	0.0019	3	-0.5024	0.487
		0.05	0.0502	0.0019	3	0.0496	0.724
		0.20	0.2009	0.0019	3	0.1986	0.704
		0.90	0.9104	0.0027	4	0.8995	0.053
		0.98	0.9896	0.0027	4	0.9777	0.236
	10^{-4}	-0.50	-0.5055	0.0026	4	-0.5000	< 0.001
		0.05	0.0506	0.0027	4	0.0500	< 0.001
		0.20	0.2024	0.0026	4	0.2000	< 0.001
		0.90	0.9109	0.0034	5	0.9000	< 0.001
		0.98	0.9919	0.0041	6	0.9800	< 0.001

	Corrélation de rang $r_{i,j}^X$	Corrélation linéaire $\rho_{i,j}^X$
X_i, X_j continues uniformes	analytique : équation (3.7)	analytique : équation (3.7)
X_i, X_j continues (sauf loi uniforme)	analytique : équation (3.7)	numérique : méthode de Cario et Nelson [18] par ex.
X_i, X_j discrètes	méthodes : M1, M2A, M2B, M3	méthodes : M1, M2A, M2B, M3
X_i continue X_j discrète	méthode : M4	méthode : M4

Tableau 3.7 – Types de résolutions dans l’initialisation de NORTA selon les différents cas.

3.4 Difficultés avec la méthode NORTA

Les difficultés computationnelles de la méthode se manifestent en grande partie dans l’initialisation pour calculer les $d(d-1)/2$ éléments de la matrice R_ρ^Z . Par la suite, un coût est engendré par la génération du vecteur X , où la complexité de l’inversion varie d’une distribution à une autre. Éventuellement, il y a un coût additionnel pour corriger la matrice de corrélations obtenue dans le cas où elle n’est pas semi-définie positive.

Rappelons tout d’abord les propriétés d’une matrice de corrélations. Toute matrice C qui correspond aux corrélations (linéaires ou de rang) entre d variables aléatoires d’un même vecteur doit vérifier les propriétés suivantes :

- P1 : tous les éléments de la matrice doivent être dans l’intervalle $[-1, 1]$.
- P2 : les éléments de la diagonale doivent être égaux à 1.
- P3 : la matrice doit être symétrique.
- P4 : la matrice doit être semi-définie positive, c.-à.-d. que pour tout vecteur $x \in \mathbb{R}^d$, nous avons $x^t C x \geq 0$.

Dans notre cas, le problème se pose avec la matrice R_ρ^Z obtenue après la phase

d'initialisation de la méthode NORTA. Les propriétés P1, P2 et P3 sont vérifiées, sauf pour la propriété P4. Avec des grandes dimensions, les chances d'obtenir une matrice semi-définie positive sont très faibles. Ghosh et Henderson [32] ont même montré que la probabilité que cette matrice ne soit pas semi-définie positive est presque égale à 1 pour $d > 15$. Comme solution, les mêmes auteurs ont proposé un programme pour chercher la matrice semi-définie positive la «plus proche» de R_ρ^Z , soit

$$\begin{aligned} \text{Min} \quad & d(R_\rho^Z, LL^t) \\ \text{s. à} \quad & LL^t \geq 0 \\ & (LL^t)_{i,i} = 1, \quad i = 1, \dots, d, \end{aligned} \quad (3.43)$$

avec LL^t la matrice recherchée, qu'on écrit sous cette forme puisqu'elle doit être décomposable et $d(R_\rho^Z, LL^t)$ est la distance entre les deux matrices et qui peut prendre l'une des formes suivantes :

$$L_1 = \sum_{i>j} |\rho_{i,j}^Z - (LL^t)_{i,j}|$$

ou

$$L_\infty = \max_{i>j} |\rho_{i,j}^Z - (LL^t)_{i,j}|.$$

Lurie et Goldberg [61] ont proposé le programme suivant :

$$\begin{aligned} \text{Min}_{l_{i,j}} \quad & \frac{1}{2} \|R_\rho^Z - LL^t\| = \frac{1}{2} \sum_{i=1}^k \sum_{j=1}^k w_{i,j} (\rho_{i,j}^Z - s_{i,j})^2 \\ \text{s. à} \quad & s_{i,j} = \sum_{h=1}^m l_{ih} l_{jh}, \quad i, j = 1, \dots, d \\ & s_{i,i} = 1, \quad i = 1, \dots, d, \end{aligned} \quad (3.44)$$

où $m = \min(i, j) \leq d$, les $l_{i,j}$ sont les éléments de L et les $w_{i,j}$, pour $i, j = 1, \dots, d$, sont des poids bien choisis.

Davenport et Iman [22] ont proposé aussi un algorithme pour construire une matrice semi-définie positive à partir d'une matrice quelconque avec des transformations légères sur quelques valeurs propres. Il n'y a aucune base théorique qui supporte la méthode, mais les auteurs soutiennent que dans tous les cas empiriques étudiés, l'algorithme a convergé et a donné des bons résultats dans le cas où toutes les conditions sont satisfaites. Ils ont donc montré la validité de leur algorithme par des exemples d'applications.

La procédure transforme tout simplement la matrice en une avec des valeurs propres non négatives. Cette méthode garantit que les éléments de la matrice transformée sont proches de celle de départ sous certaines conditions. Si les conditions ne sont pas remplies, l'algorithme converge quand même, mais ne donne pas de bons résultats. Pour avoir un bon résultat, il faut avoir les conditions suivantes satisfaites :

- Il faut que le nombre de valeurs propres négatives de la matrice initiale soit relativement petit par rapport à la dimension de la matrice et par rapport aux autres valeurs non négatives.
- Les valeurs propres doivent aussi être proches de zéro (petites en valeur absolue).
- Il faut que quelques unes des valeurs non négatives soient proches de zéro aussi.

Voici l'algorithme utilisé : on définit la matrice initiale C non singulière, symétrique et vérifiant $c_{i,i} = 1$, pour $i = 1, \dots, d$, et $-1 < c_{i,j} < 1$, pour $1 \leq i < j \leq d$. On suppose, sans perte de généralité, que les d valeurs propres de C sont ordonnées, $\lambda_1 < \dots < \lambda_d$, et que les $k < d$ premières valeurs propres $\lambda_1 < \dots < \lambda_k$ sont

négatives. On applique alors l'algorithme suivant :

1. On décompose la matrice C :

$$C = ZDZ^t = \sum_{i=1}^d \lambda_i Z_i Z_i^t,$$

où, $D = \text{diag}(\lambda_i)$, c.-à.-d. que les λ_i , $i = 1, \dots, d$, constituent les éléments de la diagonale de D et ses autres éléments sont nuls, et les Z_i , $i = 1, \dots, d$, sont les vecteurs propres associés aux valeurs propres et qui constituent les colonnes de Z .

2. Mettre $\lambda_i = \epsilon$, pour $i = 1, \dots, k$, où ϵ a une valeur proche de zéro (0.001 par exemple).
3. Augmenter les valeurs $\lambda_{k+1} < \dots < \lambda_{2k}$ à ϵ , comme suit :

$$\lambda_i^* = \begin{cases} \epsilon & \text{si } \lambda_i \leq \epsilon \\ \lambda_i & \text{si } \lambda_i > \epsilon. \end{cases}$$

4. Garder le reste des valeurs propres $\lambda_{2k+1} < \dots < \lambda_d$, et prendre la nouvelle matrice :

$$C^* = \sum_{i=1}^k \epsilon Z_i Z_i^t + \sum_{i=k+1}^{2k} \lambda_i^* Z_i Z_i^t + \sum_{i=2k+1}^d \lambda_i Z_i Z_i^t.$$

5. Mettre des 1 dans la diagonale $c_{i,i} = 1$, $i = 1, \dots, d$, si nécessaire.
6. Si toutes les valeurs propres λ_i , $i = 1, \dots, d$ de C^* sont non négatives, la matrice recherchée sera C^* et l'algorithme se termine. Sinon, refaire une autre itération.

3.5 Améliorations proposées pour la méthode NORTA

Pour améliorer notre modèle pour les arrivées, nous proposons au début de réduire le nombre de paramètres estimés. Pour chercher la matrice R_ρ^Z , nous avons à résoudre, $d(d-1)/2$, équations du type (3.6). Or les corrélations $r_{i,j}^X$, pour $1 \leq i < j \leq d$, ne sont pas connues, mais estimées par les corrélations échantillonales. Nous avons aussi à estimer les paramètres des distributions marginales F_i , pour $i = 1, \dots, d$. Ce grand nombre d'estimations devient un inconvénient en grande dimension, comme dans l'exemple traité dans le centre d'appels C1 et C2 de Bell, par exemple, où les dimensions sont de 25 et 48, respectivement.

Pour régler ce problème, une des solutions possibles serait de réduire le nombre de paramètres à estimer, $d(d-1)/2$, à un nombre beaucoup plus petit en trouvant un modèle qui paramétrise la matrice en question. En plus de réduire le nombre de paramètres, nous voulons en même temps garantir la propriété P4 de toute matrice de corrélation, étant donné que c'est difficile d'avoir cette propriété satisfaite et d'avoir une matrice semi-définie positive dans des grandes dimensions, comme dans notre cas [32].

Il y a plusieurs modèles proposés dans la littérature pour paramétriser la matrice de corrélation et ces modèles sont classés en deux groupes : les modèles qui garantissent une matrice estimée avec le même rang que la matrice originale et les modèles qui fournissent une matrice de moindre rang. Pour plus de détails sur la paramétrisation d'une matrice de corrélations, le lecteur peut se référer, par exemple à Brigo [11].

L'idée est de trouver un modèle pour la matrice de départ R_r^X , avec un ensemble θ de paramètres qui minimise la norme de la différence entre la matrice originale

et celle qu'on cherche, soit :

$$\min_{\theta} \sum_{i=1}^d \sum_{\substack{j=1 \\ j \neq i}}^d (r_{i,j}^X - r_{i,j}^X(\theta))^2, \quad (3.45)$$

dans un modèle de la forme :

$$r_{i,j}^X = r_{i,j}^X(\theta) + \epsilon_{i,j}, \quad i, j = 1, \dots, d, \quad j \neq i, \quad (3.46)$$

avec $\epsilon_{i,j}$ des résidus qui sont i.i.d. selon la loi normale de moyenne 0 et de variance σ_{ϵ}^2 .

Dans notre cas, le but principal était de trouver des modèles qui décrivent bien les corrélations dans le contexte des centres d'appels, avec peu de paramètres et qui garantissent d'avoir une matrice qui satisfait la propriété P4. Nous avons donc examiné la structure des corrélations entre les arrivées dans chaque intervalle de temps dans une journée au niveau de chaque centre d'appels étudié et nous avons remarqué que généralement ces corrélations décroissent rapidement quand la différence entre ces intervalles augmente. Il fallait donc trouver un modèle qui permet de capter cette propriété, sans trop de paramètres. Une idée qui vient tout de suite à l'esprit est de permettre aux corrélations $r_{i,j}$ de varier en fonction de $|j - i|$, pour $i, j = 1, \dots, d, j \neq i$. Au début, nous avons pensé au modèle simple où chaque corrélation $r_{i,j}$ dépend directement de $|j - i|$, $i, j = 1, \dots, d, j \neq i$, en ayant $(d - 1)$ paramètres, c.-à.-d. $r_{i,j} = p_{|j-i|}$, $i, j = 1, \dots, d, j \neq i$. Bien que le nombre de paramètres est réduit par rapport à $d(d - 1)/2$, nous avons constaté que le seul moyen d'estimer les paramètres $p_{|j-i|}$, $i, j = 1, \dots, d, j \neq i$, est de prendre la moyenne arithmétique des corrélations échantillonnales, c.-à.-d. avoir $\hat{p}_{|j-i|} = \frac{\sum_{k=1}^{d-|j-i|} r_{k,k+|j-i|}}{d-|j-i|}$, $i, j = 1, \dots, d, j \neq i$. Nous avons aussi constaté que

plus la différence $|j - i|$ augmente, moins nous avons de données pour l'estimation des paramètres. Par exemple, pour $|j - i| = d - 1$, le paramètre p_{d-1} est estimé en utilisant seulement $r_{1,d}$. Bien entendu, il y a plusieurs autres possibilités pour modéliser les corrélations, mais nous avons insisté sur les modèles où les corrélations décroissent en fonction de la différence entre les intervalles de temps.

Nous avons analysé la tendance dans les corrélations de rang échantillonales dans les trois centres étudiés C1, C2 et C3, et nous avons constaté qu'il y a plusieurs modèles possibles qui sont basés, par exemple, sur des fonctions polynomiales, exponentielles ou rationnelles. Nous avons sélectionné les trois modèles qui s'ajustent mieux aux données et qui fournissent des résultats semblables du point de vue de l'erreur standard d'estimation. Il faut noter que ces modèles permettent d'obtenir une matrice avec le même rang que celui de la matrice originale, contrairement à la méthode que nous avons présentée à la section précédente. Voici les trois modèles proposés :

– Modèle 1 : $\theta = (a, b)$

$$r_{i,j}^X(\theta) = a^{|j-i|} + b, \quad i, j = 1, \dots, d, \quad j \neq i. \quad (3.47)$$

– Modèle 2 : polynomial, $\theta = (p_1, p_2, p_3, p_4)$

$$r_{i,j}^X(\theta) = p_1|j - i|^3 + p_2|j - i|^2 + p_3|j - i| + p_4, \quad (3.48)$$

$$i, j = 1, \dots, d, \quad j \neq i.$$

– Modèle 3 : rapport de deux polynômes, $\theta = (p_1, p_2, q_1, q_2, q_3)$.

$$r_{i,j}^X(\theta) = \frac{p_1|j - i| + p_2}{|j - i|^3 + q_1|j - i|^2 + q_2|j - i| + q_3}, \quad (3.49)$$

$$i, j = 1, \dots, d, \quad j \neq i.$$

Pour la sélection du modèle adéquat dans chaque centre d'appels étudié, nous allons utiliser le *critère d'information d'Akaike* (AIC). Le lecteur peut se référer à Akaike [2] et Wei [91, page 153]. Ce critère possède l'avantage qu'il prend en compte l'erreur quadratique moyenne des résidus et le nombre des paramètres estimés dans le modèle. Il représente une estimation approximative de la *distance de Kullback-Leibler* entre le vrai modèle et le modèle qui l'estime. En supposant que les résidus sont i.i.d. de distribution normale de moyenne 0 et de variance σ^2 , ce critère est défini par :

$$\text{AIC}(s) = n \log(\hat{\sigma}^2) + 2s, \quad (3.50)$$

où n est le nombre d'observations, s est le nombre de paramètres estimés dans le modèle, et $\hat{\sigma}^2$ l'estimateur de maximum de vraisemblance de la variance des résidus. D'autres variantes de ce critère, connues sous le nom d'AICC, sont présentées, par exemple, dans Brockwell et Davis [12, pages 301–304].

Nous verrons dans les chapitres 4 et 5, les résultats numériques obtenus avec ces modèles. Il faut mentionner ici qu'avec les trois modèles utilisés pour la matrice R_r^X , la matrice correspondante R_ρ^Z obtenue est toujours semi-définie positive malgré des dimensions importantes.

CHAPITRE 4

LES ARRIVÉES AUX CENTRES D'APPELS DE BELL CANADA

4.1 Description des centres d'appels étudiés et analyse des données

Dans ce chapitre nous présentons une étude de cas avec des données observées dans deux centres d'appels de Bell Canada. Nous étudions la sensibilité du centre par rapport au changement du processus d'arrivée en faisant des simulations avec les différents modèles présentés. Nous commençons tout d'abord par une étude préliminaire descriptive des données recueillies dans ces centres pour ensuite appliquer les modèles et observer et comparer leur performances.

Il est important de signaler que les données fournies par Bell Canada sont insuffisantes pour une bonne validation des modèles. Dans plusieurs cas, ces données sont agrégées ou incomplètes, et le plus important c'est l'absence de données sur de longues périodes de temps. L'absence de données pour le centre C2 nous a empêché de construire un modèle mathématique complet pour effectuer des simulations sur le centre afin d'estimer certaines mesures de performance avec les modèles proposés et d'observer aussi la sensibilité du centre face aux changements de différents scénarios possibles.

4.1.1 Centre C1

Le premier centre de Bell Canada étudié est ouvert du lundi au samedi et il est fermé le dimanche. Les appels proviennent de clients résidant au Québec et en Ontario qui sollicitent des services sur Bell Canada, Bell Mobilité, Bell Express Vu et Bell Internet Sympatico. C'est un centre qui opère en mode "mixte", où les clients peuvent aussi être appelés.

Dans ce centre, le nombre d'agents employés varie selon des horaires journaliers fixés par les gestionnaires du centre. En gros, chaque agent travaille pendant 7 heures et demie. Il commence au début par répondre seulement aux appels entrants et ceci pendant 4 heures et demie, ensuite pour les 3 heures qui restent, en plus de ce type d'appels, il répond aussi aux appels sortants des clients rejoints par le composeur automatique. En moyenne, 27 agents travaillent du lundi au vendredi et 13 le samedi. Les préposés répondent aux appels des clients de 8h00 à 20h30 la semaine et de 8h30 à 17h00 le samedi. Pour la politique de composition des appels sortants, elle dépend du nombre d'agents libres et de la qualité de service des dernières 10 minutes. Les gestionnaires du centre peuvent aussi intervenir pour contrôler le composeur. S'ils remarquent une baisse de qualité de service au cours des dernières demi-heures, due par exemple à un taux d'arrivée élevé inhabituel ou à une mauvaise occupation des agents, ils peuvent changer les paramètres du composeur pour diminuer son agressivité et permettre de servir plus d'appels entrants.

Ce premier lot de données fourni par Bell concerne 253 journées du 2 mai 2001 au 3 mars 2002. Pour les arrivées, les données sont agrégées par demi-heure et aucune information n'est fournie sur l'instant de chaque appel, son type ou son origine. Avant de procéder à l'application des modèles mathématiques pour le processus d'arrivée à ces données, nous les avons analysées afin de distinguer les jours de la semaine qui se ressemblent ou en d'autres termes identifier l'effet du jour de la semaine. Le samedi est traité à part, puisque le centre est ouvert de 8h30 à 17h00 seulement. Le graphique 4.1 montre le nombre moyen d'arrivées par demi-heure pour chaque type de journée du lundi au vendredi. Il est clair que le lundi est différent des autres journées. Le vendredi semble aussi être différent, et ceci est confirmé en regardant le tableau 4.1, où on montre le volume moyen d'appels par type de jour, ainsi que les diagrammes en boîtes dans la figure 4.3.

Jour	Estimation
Lundi	1419±86
Mardi	1108±53
Mercredi	1132±58
Jeudi	1153±60
Vendredi	1228±67

Tableau 4.1 – Nombre moyen d'appels par type de journée dans le centre C1.

Dans chaque diagramme, on retrouve la médiane, le premier et le troisième quartiles (le fond et le haut de la boîte centrale), l'étendue d'inter-quartiles EIQ (la hauteur de la boîte) et de deux barres trouvés à une distance de 1.5 fois l'EIQ au-dessous de premier quartile et au-dessus de quatrième quartile, respectivement. Les petits cercles dénotent les observations qui dépassent les deux barres qu'on considère ici comme des données extrêmes ou journées spéciales pour notre cas.

En observant les graphiques 4.1 et 4.3, ainsi que le tableau 4.1, nous concluons alors qu'il y a quatre types de jours : lundi, mardi+mercredi+jeudi, vendredi et samedi. Pour la modélisation et dans la suite de nos travaux, nous considérons seulement les jours du type mardi, mercredi et jeudi, soit un total de 120 jours. Ce nombre n'inclut pas les 5 journées spéciales où les horaires et le nombre d'agents sont réduits, par exemple, la veille du jour de l'an et la veille de Noël. La figure 4.4 montre l'évolution de la série du volume d'arrivées par jour pour ce type de journée. Le graphique 4.2 nous donne une idée sur la variation du taux d'arrivée pendant la journée, où on remarque la présence de deux sommets. Les appels affluent en grand nombre dans la période de la matinée entre 10h et 13h, un relâchement est ensuite observé pour la période du dîner et une autre période avec un grand taux d'arrivée entre 16h00 et 18h00.

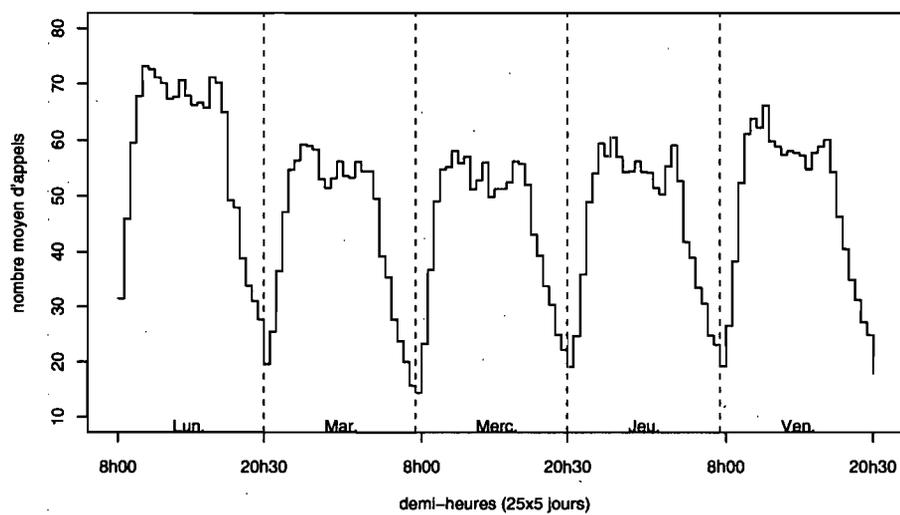


Figure 4.1 – Nombre moyen d'arrivées par intervalle d'une demi-heure et par type de journée dans le centre C1.

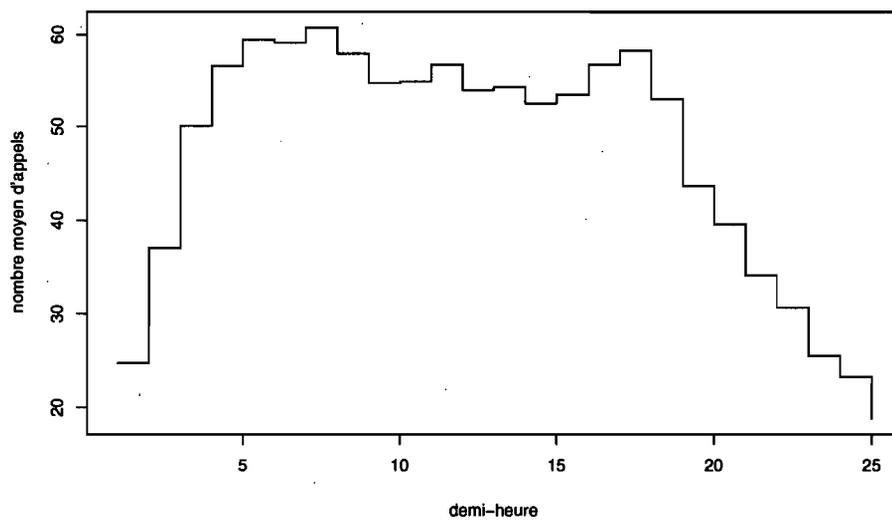


Figure 4.2 – Nombre moyen d'arrivées par intervalle de demi-heure pour une journée de type mardi, mercredi ou jeudi dans le centre C1.

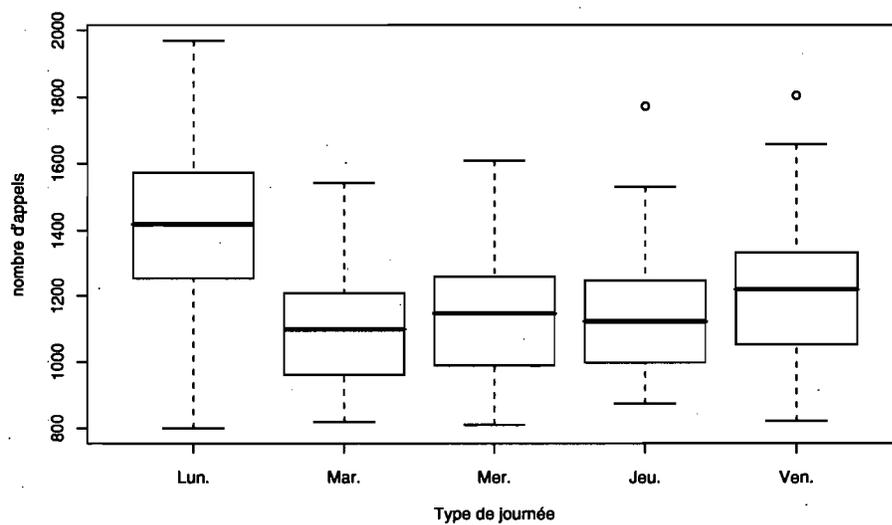


Figure 4.3 – Diagrammes en boîtes pour le volume d'appels par type de journée dans le centre C1.

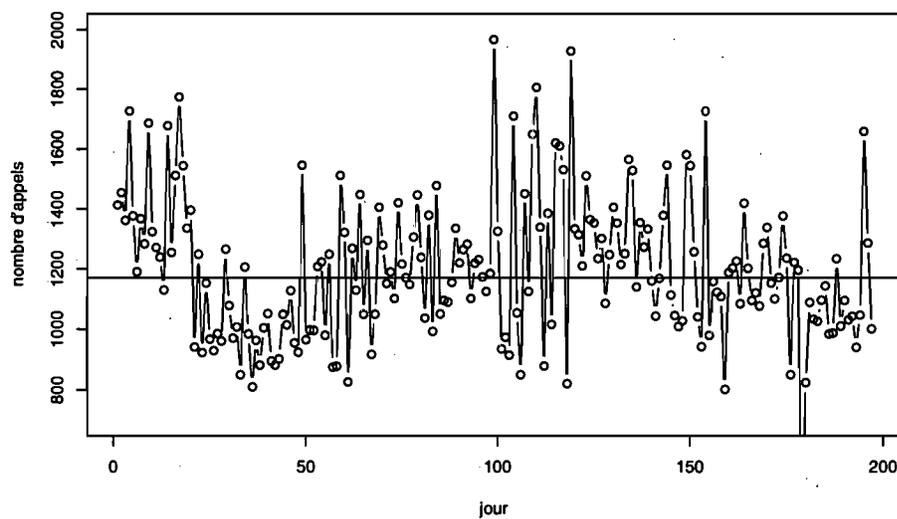


Figure 4.4 – Volume d'appels reçus par jour dans le centre C1.

4.1.2 Centre C2

Ce centre est plus grand que le premier et il est spécialisé essentiellement dans les ventes et les services de produits de télécommunication principalement filaires, mais qui incluent aussi la vente de produits non filaires et de satellite. Le centre opère avec environ 1500 à 2000 agents à travers les deux provinces du Québec et de l'Ontario. Les agents répondent aux appels des clients seulement du lundi au vendredi de 8h00 à 20h00. Le reste du temps les clients appellent un service automatisé pour des informations ou des services et parfois ils laissent des messages dans le but d'être rappelés durant les heures d'ouverture dans la semaine.

Dans ce centre, les données couvrent une période plus récente de 173 jours, du premier janvier au 22 juin 2005. Durant 24 heures, les appels aboutissent au système automatique IVR ("Integrated Voice Response") et sont enregistrés par leurs instant d'arrivée, avec des détails sur la localisation de l'appel par province (Québec ou Ontario), l'information indiquant si l'appel a été transféré ou non vers un agent durant les heures de travail, le numéro du centre d'appel approprié par région, le numéro d'où l'appel est composé, ainsi que le numéro concerné par l'appel. Les trois dernières informations ne sont pas fournies pour tous les appels enregistrés. L'information sur le numéro concerné par l'appel pourrait être utile dans l'identification des recompositions d'un même client alors qu'elles sont comptabilisées comme des arrivées distinctes par le système automatisé. Nous ne tenons pas compte de ces recompositions dans nos modèles et elles sont toutes considérées comme des arrivées distinctes.

Contrairement aux données du premier centre, celles du deuxième peuvent permettre de faire des analyses à plusieurs niveaux, dépendamment du besoin. On peut par exemple étudier la distribution de la durée inter-arrivées ou la distribution des instants d'arrivées. On peut aussi estimer une fonction de taux si l'on

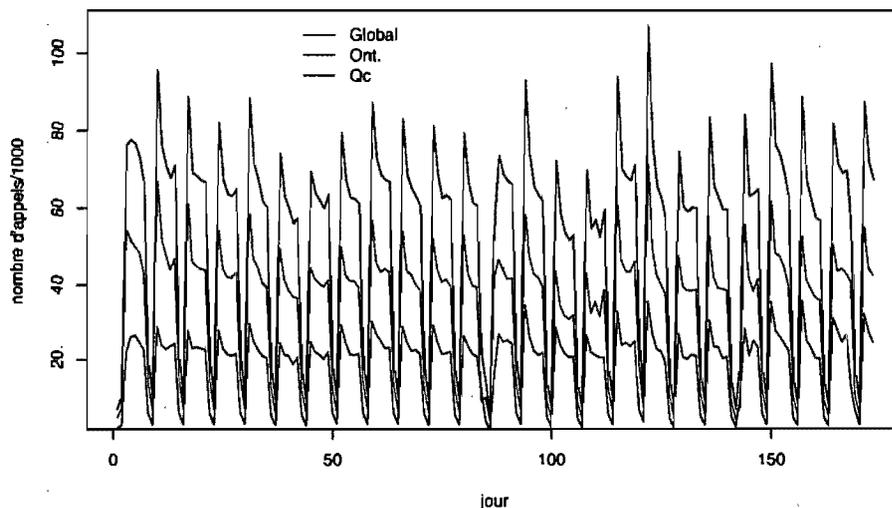


Figure 4.5 – Volume d'appels reçus par jour pour chaque province et pour l'ensemble dans le centre C2.

adopte l'hypothèse que le processus de Poisson décrit bien les arrivées. Dans notre cas, nous considérons le processus d'arrivée et nous nous intéressons au nombre d'appels reçus par intervalle de temps durant la journée.

Comme première analyse, nous étudions le nombre d'arrivées durant les différentes périodes ou intervalles de la journée selon le type de journée de la semaine. Nous essayons tout d'abord de distinguer le type de journée de la semaine et par la suite identifier et éliminer l'effet de quelques journées spéciales considérées comme des données aberrantes.

Dans la figure 4.5, nous présentons la série du volume d'appels reçus par jour pour la période couverte, par province et pour l'ensemble. Si nous regardons les graphiques 4.6 et 4.7, nous pouvons déjà constater que les arrivées de mercredi, jeudi et vendredi se ressemblent et diffèrent des autres journées. A priori, nous pouvons considérer 5 types de journées : lundi, mardi, mercredi+jeudi+vendredi, samedi et dimanche. Cette hypothèse est renforcée si on regarde le tableau 4.2, où

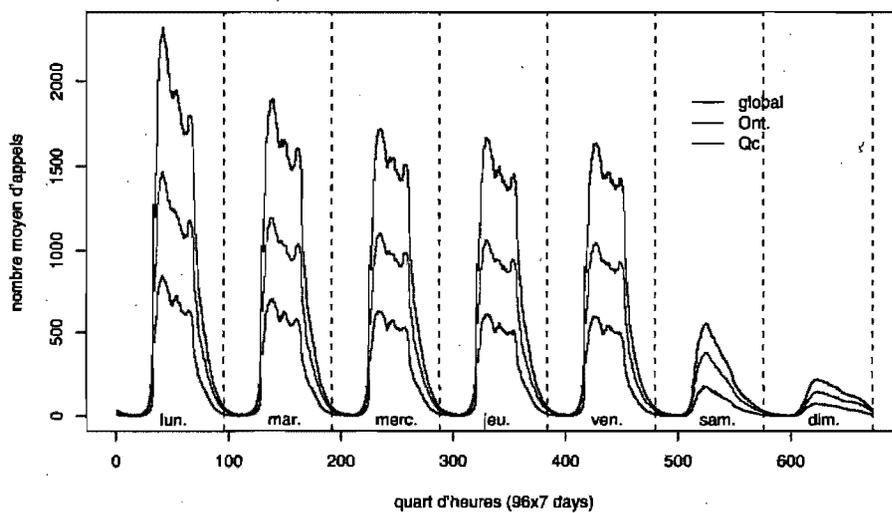


Figure 4.6 – Nombre moyen d'arrivées par intervalle de quart d'heure et par type de journée dans le centre C2.

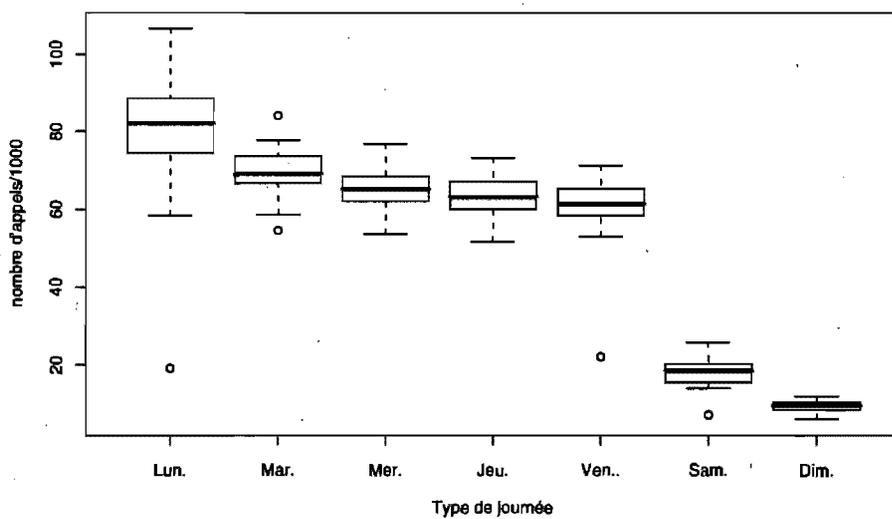


Figure 4.7 – Diagrammes en boîtes pour le volume d'appels par jour pour chaque type de journée dans le centre C2.

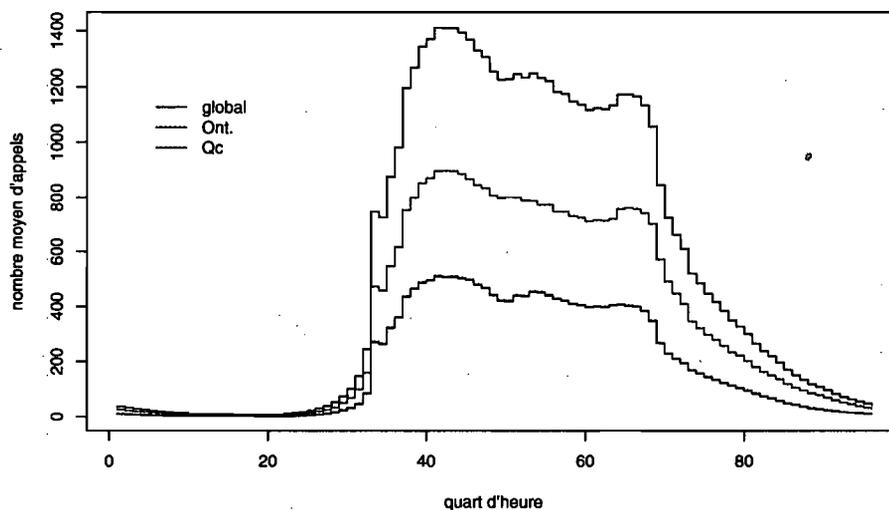


Figure 4.8 – Nombre moyen d’arrivées par intervalle de quart d’heure pour une journée de type mercredi, jeudi ou vendredi pour chaque province et pour l’ensemble dans le centre C2.

Jour	Estimation
Lundi	80445±6450
Mardi	69392±2565
Mercredi	64911±2005
Jeudi	62796±2093
Vendredi	60268±3747
Samedi	17881±1400
Dimanche	9396±565

Tableau 4.2 – Nombre moyen d’appels par type de journée pour l’ensemble des deux provinces au centre C2.

on estime le volume d'appels moyen par jour pour chaque type de journée. Dans le graphique 4.8, nous présentons le nombre moyen d'appels reçus par tranches de 15 minutes pour illustrer la variabilité du taux d'arrivée dans une même journée.

Pour les journées spéciales, Shen et Huang [81] ont utilisé une méthode basée sur la décomposition en valeurs singulières (DVS) des éléments de la matrice des données pour analyser les arrivées dans un centre d'appels. Cette méthode permet non seulement d'identifier les données aberrantes, mais aussi de détecter l'effet du jour de la semaine. Elle permet entre autre de résumer les données et permet d'identifier la partie qui contient la majorité de l'information et aussi une bonne visualisation des données. Nous avons utilisé cette méthode pour identifier les journées spéciales dans ce centre.

Pour appliquer la méthode, nous considérons la matrice $Z = (Z_{i,j})$ de dimension (n, m) et de rang $r \leq \min(n, m)$, où chaque $Z_{i,j}$ correspond au nombre d'appels reçus durant le quart d'heure $j = 1, \dots, m$, du jour $i = 1, \dots, n$. Chaque ligne, notée $r_i, i = 1, \dots, n$, représente les arrivées de la journée i réparties sur les m quart d'heures et toutes ces lignes $r_i, i = 1, \dots, n$, vont capter l'effet de l'heure de la journée ou l'effet intra-journalier. Chaque colonne de Z , notée $c_j, j = 1, \dots, m$, correspond au nombre d'appels reçus pendant le quart d'heure j de chacune des n journées et par conséquent, ces colonnes $c_j, j = 1, \dots, m$, vont capter l'effet du jour ou l'effet inter-journalier. On décompose la matrice Z comme suit :

$$Z = USV^t, \quad (4.1)$$

où U est une matrice de dimension (n, m) et ses éléments (colonnes) proviennent des $c_j, j = 1, \dots, m$ et représentent l'effet inter-journalier. Les élément de la matrice V^t de dimension (m, m) (les lignes) proviennent des $r_i, i = 1, \dots, n$ et représentent

l'effet intra-journalier. Quant à la matrice S qui est de dimension (m, m) , elle est diagonale et ses éléments $s_1 \geq s_2 \geq \dots \geq s_r$ ($s_k = 0$ pour $r + 1 \leq k \leq m$), appelés valeurs singulières, indiquent le poids de chaque paire de vecteurs qui expliquent les données.

Nous nous intéressons plutôt à l'effet inter-journalier ou l'effet du jour de la semaine, qui va permettre par conséquent de détecter les journées spéciales par type de jour. Selon Shen et Huang [81], la quantité relative $s_k^2 / (\sum_{i=1}^r s_i^2)$, $k = 1, \dots, r$, permet d'identifier l'importance relative de chaque vecteur u_k , $k = 1, \dots, m$, de la matrice U dans l'explication des données.

Nous avons appliqué cette méthode aux données du centre C2, nous disposons donc de 96 quart d'heures par jour et de 173 journées et notre matrice Z est donc de dimension $(173, 96)$. Nous avons décomposé la matrice, qui est de rang 96, et nous avons calculé les quantités $s_k^2 / (\sum_{i=1}^{96} s_i^2)$, $k = 1, \dots, 96$, et nous avons constaté que $(s_1^2 + s_2^2) / (\sum_i s_i^2) = 0.995$, ce qui veut dire que 99.5% de l'information est expliquée par la paire des vecteurs singuliers (u_1, u_2) de la matrice U , qui rappelle le représente l'effet inter-journalier. Nous pouvons donc analyser l'effet du jour de la semaine avec ces deux vecteurs, puisqu'ils contiennent presque la totalité de l'information [81]. Nous présentons le diagramme de dispersion de ces deux vecteurs dans le graphique 4.9, où on peut remarquer qu'il y a des points qui s'éloignent de l'ensemble. On pourra alors suspecter ces points pour des données aberrantes ou des journées spéciales. Dans le tableau 4.3, nous comparons le résultat de la méthode décrite ci-dessus avec une méthode classique qui identifie les données aberrantes comme étant les valeurs qui dépassent de $1.5 \times \text{EIQ}$ en dessus du 3ème quartile ou en dessous du premier quartile, où le EIQ représente l'étendue inter-quartiles.

Dans notre analyse, nous tenons compte de toutes les journées identifiées par la méthode DVS, qui sont au nombre de 8. Parmi ces journées nous retrouvons le Jour

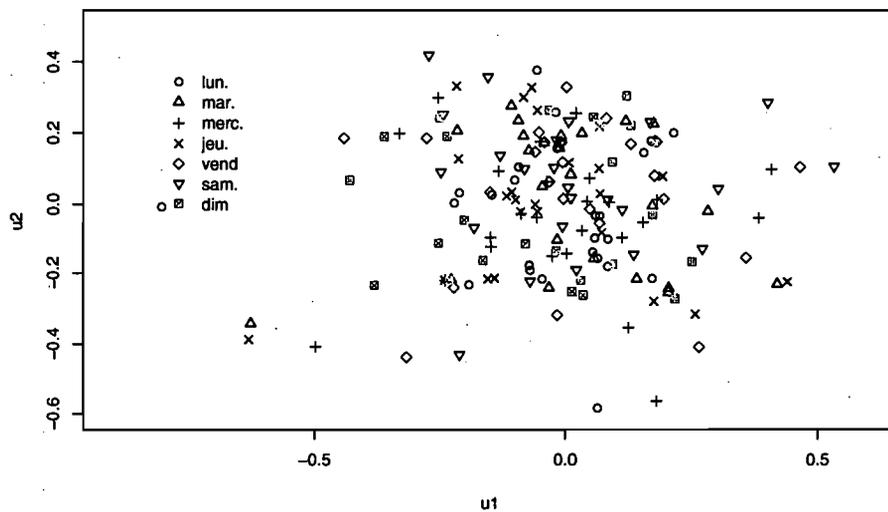


Figure 4.9 – Graphique de la paire (u_1, u_2) pour tous les jours de la semaine

Jour	avec DVS	avec EIQ
1 janvier	×	×
2 janvier	×	
25 mars	×	×
13 avril	×	
14 avril	×	
19 avril	×	×
2 mai	×	
23 mai	×	×
24 mai		×

Tableau 4.3 – Identification des journées spéciales selon deux méthodes, pour le centre C2.

de l'An, Vendredi Saint et la Fête de la Reine. Dans la modélisation du processus d'arrivée à l'intérieur d'une journée, nous éliminons tout simplement ces journées et nous considérons seulement les journées qui se ressemblent (mardi, mercredi et jeudi).

Dans le chapitre 2, en équation (2.1), nous avons présenté un test qui permet de justifier le choix d'un processus de Poisson pour les arrivées des appels à l'intérieur d'une journée. Nous avons donc appliqué ce test aux données du centre C2, où nous disposons de plus de détails sur les instants d'arrivées. Nous avons considéré au début un intervalle de longueur 5 minutes pour tous les jours entre 10h00 et 10h05 pour les journées semblables du type mercredi, jeudi et vendredi pour éliminer l'effet du jour de la semaine, excluant les journées spéciales identifiées auparavant, soit un total de $n = 70$ jours. Nous avons obtenu $X^2 = 84.528$ et une p-valeur $p = 2.03 \times 10^{-14}$, ce qui nous amène à rejeter l'hypothèse nulle. Nous avons repris le même test avec un intervalle de longueur une minute, nous avons obtenu les résultats suivants : $X^2 = 7.236$ et $p = 0.204$. Dans ce cas-ci, l'hypothèse n'est pas rejetée jusqu'à un niveau de 20%. Ce résultat permet de considérer le processus de Poisson non homogène parmi nos modèles. Nous ajoutons à ce modèle une estimation de la fonction de taux par une fonction spline cubique, puisque le test montre aussi que même dans des intervalles de 5 minutes, le taux n'est pas vraiment stationnaire.

La question qu'on s'est posée : est ce que ce processus décrit bien la réalité dans ce centre d'appels ? Pour cela, nous devons étudier le problème de dépendance entre les arrivées successives dans une même journée et même entre les journées. Dans le graphique 4.10, nous présentons un diagramme de dispersion pour les arrivées de tous les jours avant et après 10h00. Nous remarquons bien qu'il y a une dépendance entre les arrivées des deux parties de la journée. Parmi les solutions on

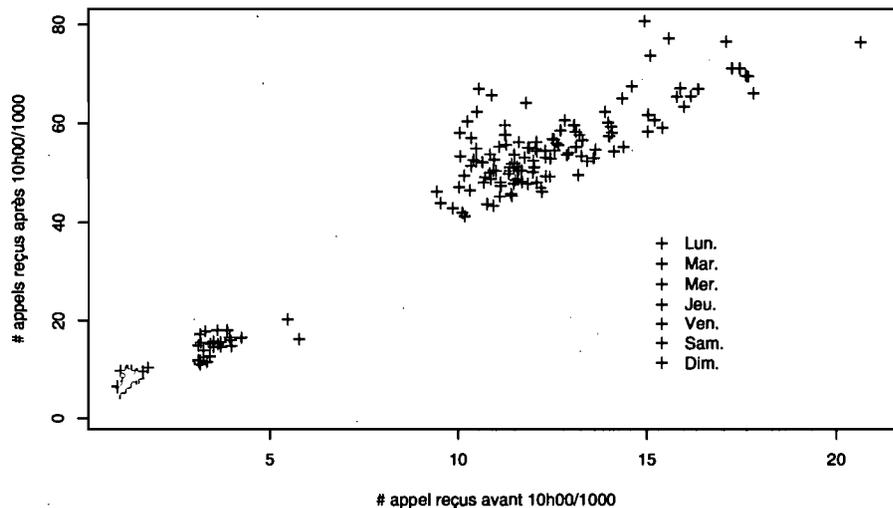


Figure 4.10 – Diagramme de dispersion entre les arrivées avant et après 10h00 pour tous les jours de la semaine au centre C2.

pourrait considérer le processus de Poisson doublement stochastique ou encore le processus NORTA, puisque dorénavant nous sommes convaincus qu'un modèle qui tient compte de la dépendance entre les arrivées dans les différentes parties d'une même journée décrit mieux le processus d'arrivée des appels. Nous verrons plus loin que le processus NORTA a l'avantage de générer des arrivées avec exactement les quantités de corrélations désirées.

La deuxième expérience consiste à étudier le comportement des arrivées si on enlève l'effet de la journée de la semaine et l'effet de l'heure de la journée. Les graphiques 4.11, 4.12 et 4.13 montrent l'analyse des résidus quand on enlève les deux effets en question. Nous pouvons constater que dans les deux cas les résidus ne sont pas i.i.d. de distribution normale. Pour les arrivées par période de temps dans la journée, le test de Box-Ljung ($Q = 45.8$, $p\text{-valeur} = 7.534 \times 10^{-6}$) révèle aussi la présence de corrélations entre les résidus, ce qui nous motive à chercher des modèles qui vont capter ces corrélations. Ceci renforce l'hypothèse de la dépendance entre

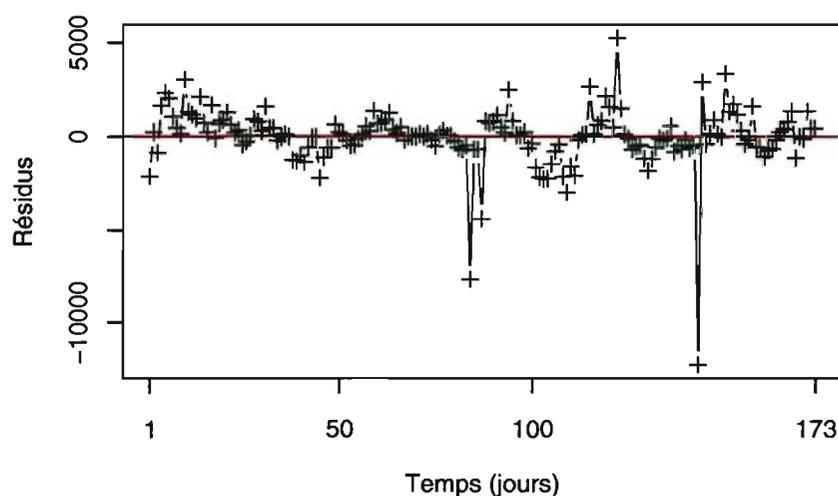


Figure 4.11 – Les résidus du modèle pour les arrivées par jour après avoir éliminé l’effet du jour de la semaine. Des observations aberrantes indiquent la présence de journées spéciales.

les volumes d’arrivées dans différentes périodes de la journée. Pour l’instant, nous allons supposer que les arrivées d’une journée à l’autre sont indépendantes.

4.2 Application des modèles

4.2.1 Simulation du centre d’appels C1

Rappelons que la simulation nous permet d’effectuer plusieurs expérimentations sur le centre étudié et permet essentiellement d’estimer des mesures de performance utiles pour la bonne gestion du centre et permet aussi d’observer la sensibilité du centre d’appels face au changement de scénarios dans toute composante du système de file d’attente que constitue le centre d’appels en général et dans le processus d’arrivée en particulier. Une fois que tous ces objectifs sont fixés, l’étape suivante consiste à construire un modèle mathématique à partir des données observées qui

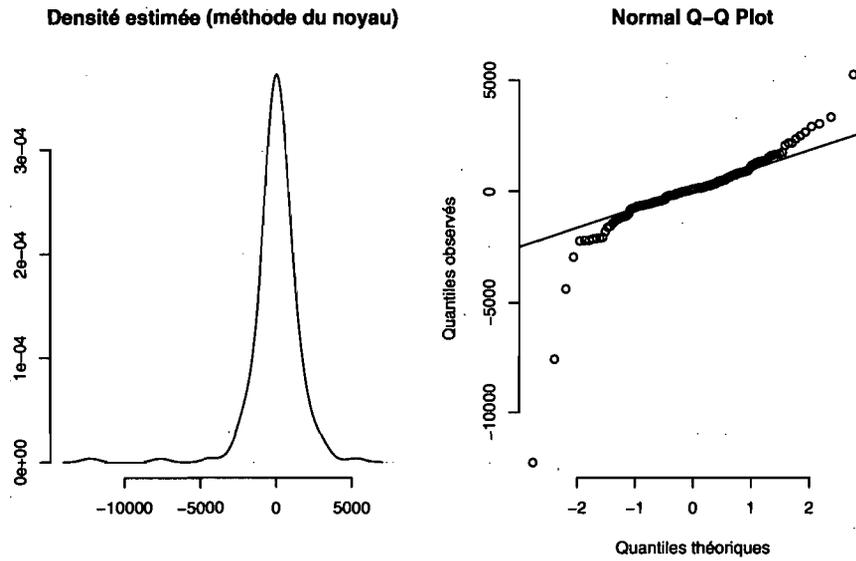


Figure 4.12 – Diagnostic de la normalité des résidus pour les arrivées par jour après avoir éliminé l'effet du jour de la semaine.

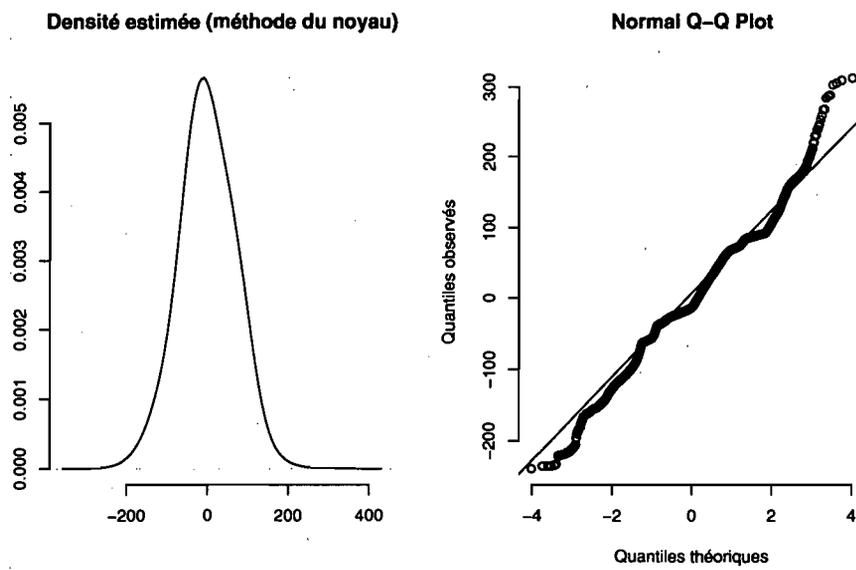


Figure 4.13 – Diagnostic de la normalité des résidus pour les arrivées par quart d'heure après avoir éliminé l'effet du jour de la semaine et de l'heure de la journée.

soit en même temps le plus réaliste possible et qu'on peut manipuler avec la simulation.

Dans notre cas, nous avons utilisé le modèle présenté par Deslauriers [25], puisque nous avons utilisé les mêmes données, sauf pour le processus d'arrivée où nous utilisons les modèles que nous avons proposés dans cette thèse.

Dans un premier temps, nous avons commencé par estimer les paramètres de la loi binomiale négative pour chaque distribution marginale du nombre d'arrivées par demi-heure. Les estimations, basées sur la méthode présentée dans la section 3.2.3, sont présentés dans le tableau 4.4.

Pour tester la validité de nos modèles pour les arrivées, nous avons procédé à une simulation d'un centre d'appels en mode mixte, c.-à-d., appels sortants et appels entrants. Nous avons utilisé les mêmes données et distributions que celles utilisés par Deslauriers [25] et Avramidis et al. [5]. Nous avons simulé deux types de journées, une journée en mode mixte entre 8h00 et 20h30, avec des appels composés à partir de 14h00, et une journée en appels entrants seulement de 8h00 à 14h00. Pour le processus de Poisson, nous présentons les résultats avec deux formes de la fonction de taux : constante par morceaux et une fonction spline cubique. Les résultats sont présentés dans les tableaux 4.5 et 4.6, sous forme de moyennes par jour avec les intervalles de confiance à 95% suite à la simulation de 30000 journées indépendantes. Les valeurs inférieures à 0.05 sont remplacées par ϵ . La ligne correspondant aux "Données" dans les tableaux des résultats de la simulation est obtenue avec un rééchantillonnage de 250 fois pour chaque journée des 120 observées pour obtenir des données pour 30000 journées. Ceci va nous permettre de calculer les mesures de performance pour les données avec le même nombre que celui des journées simulées.

Dans la simulation, nous nous sommes intéressés aux mesures de performance

Demi-heure	$\hat{\mu}_i$	\hat{s}_i	\hat{p}_i
8h00-8h30	24.7± 1.4	15.7	0.388
8h30-9h00	37.0± 1.4	60.2	0.619
9h00-9h30	50.1± 2.0	36.9	0.424
9h30-10h00	56.6±2.2	36.7	0.393
10h00-10h30	59.5±2.3	36.4	0.380
10h30-11h00	59.1±2.2	40.9	0.409
11h00-11h30	60.8±2.7	25.1	0.292
11h30-12h00	58.0±2.5	24.6	0.298
12h00-12h30	54.8±2.4	24.8	0.312
12h30-13h00	54.9±2.3	25.3	0.315
13h00-13h30	56.7± 2.6	21.8	0.278
13h30-14h00	54.0±2.3	25.1	0.317
14h00-14h30	54.3±2.4	22.8	0.296
14h30-15h00	52.5± 2.3	23.3	0.308
15h00-15h30	53.5±2.0	42.9	0.445
15h30-16h00	56.8± 2.2	36.2	0.389
16h00-16h30	58.3± 2.0	47.7	0.450
16h30-17h00	53.0±2.2	27.7	0.343
17h00-17h30	43.7±2.0	21.1	0.325
17h30-18h00	39.6± 1.8	26.5	0.401
18h00-18h30	34.1±1.7	21.3	0.385
18h30-19h00	30.7± 1.7	14.9	0.327
19h00-19h30	25.5±1.4	16.8	0.397
19h30-20h00	23.3± 1.3	19.2	0.451
20h00-20h30	18.8± 1.2	10.9	0.368

Tableau 4.4 – Estimation du taux moyen d'arrivée et des paramètres de la loi binomiale négative pour chaque demi-heure avec les données du centre C1

	Qualité de service %	Nbr. d'abandons	Nbr. d'appels servis	Occup. des agents %	Temps d'attente moy. appels reçus	Temps d'attente moy. dans la file
Données	83.8±.3	20.1±6.1	606±15	65.6±1.6	16.2±3.7	83.9±6.3
PPNH-cte	89.8±ε	10.3±.1	615.7±.2	66.7±ε	6.7±.1	51.3±.3
PPNH-spline	88.8±ε	11.6±.1	612.7±.2	66.3±ε	9.3±.1	63.8±.1
PPDS-cte	84.5±ε	18.1±.2	607.6±.7	65.8±.1	14.5±.1	77.5±.3
PPDS-spline	84.5±ε	18±.2	606.2±.6	65.6±.1	14.3 ±.2	74.4±.2
NORTA	82.7±ε	21.5±.2	605.9±.6	65.4 ±ε	16.8±.1	79.3±.3
NORTA réduit	82.9±ε	21.1±.2	606.2±.7	65.4±.1	16.4±.2	78.6±.3

Tableau 4.5 – Mesures de performance obtenues par la simulation d'une journée de 8h00 à 14h00 du type mardi, mercredi ou jeudi au centre C1 avec les modèles proposés.

	Qualité de service %	Nbr. d'abandons	Nbr. d'appels servis	Occup. des agents %	Temps d'attente moy. appels reçus	Temps d'attente moy. dans la file
Données	88.1±.2	26.9±6.8	1143.1±30.6	70.5±.1	10.6±2.3	68±5.2
PPNH-cte	92±ε	15.6±.1	1154.1±.3	71.2±ε	5.1±ε	44.9±.2
PPNH-spline	90.8±ε	19±ε	1145.4±.3	72±ε	7.1±ε	54.2±.1
PPDS-cte	88.7±ε	24.4±.2	1144.6±1.4	70.6±ε	9.6±.1	63.2±.3
PPDS-spline	88±ε	26.8±.3	1137.5±1.2	71.5±.1	10.2±.1	62.8±.1
NORTA	87.4±.1	28.4±.2	1142.5±1.3	70.8±ε	11 ±.1	64.5±.3
NORTA réduit	87.6±.1	27.9±.2	1142.3±1.4	70.7±.1	10.8±.1	63.9±.5

Tableau 4.6 – Mesures de performance obtenues par la simulation d'une journée de 8h00 à 20h30, du type mardi, mercredi ou jeudi au centre C1 avec les modèles proposés.

suivantes : (1) la qualité de service qui représente le pourcentage des appels ayant attendu moins de 20 secondes avant d'être servis; (2) le nombre d'abandons; (3) le nombre d'appels servis; (4) le taux d'occupation des agents qui représente le rapport de la durée de service pour tous les appels sur la disponibilité totale des agents; (5) le taux moyen passé dans le système pour tous les appels; (6) le taux moyen d'attente dans la file.

Nous désignons par PPNH, le processus de Poisson non homogène et par PPDS le processus de Poisson doublement stochastique. Les termes "constante" et "spline" sont associés au type d'estimation de la fonction du taux. Pour la génération des arrivées selon le processus de Poisson avec une fonction de taux spline, nous

avons utilisé la méthode de découpage pour générer les instants d'arrivées. Cette méthode, proposée par Lewis et Shedler [60], est une variante de la méthode d'acceptation/rejet qui est utilisée quand le taux cumulé d'un processus de Poisson est difficile à inverser. Pour un processus de Poisson de taux, par exemple $\lambda(t)$, l'algorithme consiste à générer des pseudo-arrivées selon un processus de Poisson de taux constant $\bar{\lambda} = \max_{t \geq 0} \lambda(t)$, et chaque pseudo-arrivée t est acceptée avec probabilité $\lambda(t)/\bar{\lambda}$.

Pour le processus PPDS, au début de chaque journée nous générons une variable W suivant une loi gamma avec les paramètres estimés $\hat{\alpha}$ et $\hat{\beta}$ pour la multiplier par la fonction de taux et utiliser par la suite la même méthode (découpage). Le nombre d'agents disponibles, selon les deux types, est fixe pour chaque demi-heure. La variable durée de service des appels entrants suit une loi gamma, alors que pour les appels sortants, la densité de la variable durée de service est estimée par une fonction noyau à partir des données observées. Le lecteur peut se référer à Deslauriers [25] pour plus de détails sur les différents inputs et modélisations utilisés dans la simulation.

Pour utiliser les résultats de l'estimation avec la fonction spline s , définie dans la section 2.2.1, avec le premier centre d'appels étudié C1, nous avons considéré l'intervalle $[a, b]$, la période de temps où ce centre est ouvert, c.-à-d, de 8h00 à 20h30. Nous avons donc 25 intervalles de 30 minutes chacun, et les noeuds ont été choisis de la façon suivante : r_0 correspond à 8h00, r_1 à 8h30, ..., r_{24} à 20h00, et r_{25} à 20h30. Nous disposons de $n = 120$ jours du type mardi, mercredi et jeudi. Si on considère que le centre est fermé avant 8h00, on peut établir que $s(r_0) = 0$, par contre pour la fermeture, on considère que le taux d'arrivée s'annule à la fermeture du centre à 20h30 malgré qu'il y a encore des appels qui continuent d'arriver même après ce temps, mais que les agents ne les prennent pas en charge. On peut donc supposer

que la fonction s devient nulle après r_{25} . Dans la figure 4.15, nous présentons une fonction spline cubique avec $h = 24$ noeuds comparée à une estimation par morceaux.

En plus du modèle NORTA, nous appelons NORTA réduit, le modèle où les corrélations sont modélisées selon l'une des formulations (3.47), (3.48) et (3.49) proposées pour les trois modèles 1, 2 et 3 dans la section 3.5. Nous utilisons ici le critère d'AIC, donné en (3.50) dans la même section pour le choix du modèle adéquat. Selon ce critère, le modèle 1, donné en (3.47), fait légèrement mieux que les deux autres modèles. Nous avons donc calculé le AIC pour les trois modèles et nous avons obtenu -1464.57 pour le modèle 1, -1462.30 pour le modèle 2 et -1459.47 pour le modèle 3. Nous présentons dans le graphique 4.14 les corrélations de rang échantillonales, $r_{i,j}^X$, pour $1 \leq i < j \leq d$, comparées à celles obtenues avec les trois modèles. Pour l'estimation des paramètres du modèle 1 choisi, nous avons obtenu, avec la méthode des moindres carrés non linéaires, les résultats suivants :

$$\hat{a} = 0.990 \pm 0.002$$

$$\hat{b} = -0.467 \pm 0.019,$$

avec une erreur standard $\hat{\sigma}_\epsilon$ (écart-type des résidus) de 0.0865. Nous avons aussi obtenu une erreur standard de 0.0863 et de 0.0864 pour les deux modèles 2 et 3, respectivement. Il faut mentionner ici que nous avons pu obtenir une matrice semi-définie positive en utilisant ce modèle, malgré une dimension de 25.

Pour pouvoir analyser la sensibilité de la simulation d'un centre d'appels au changement du processus d'arrivée, nous présentons dans les tableaux 4.5 et 4.6 les résultats des simulations avec les différents modèles pour le processus d'arrivée. Bien que nous sommes convaincus de l'infériorité du processus de Poisson non

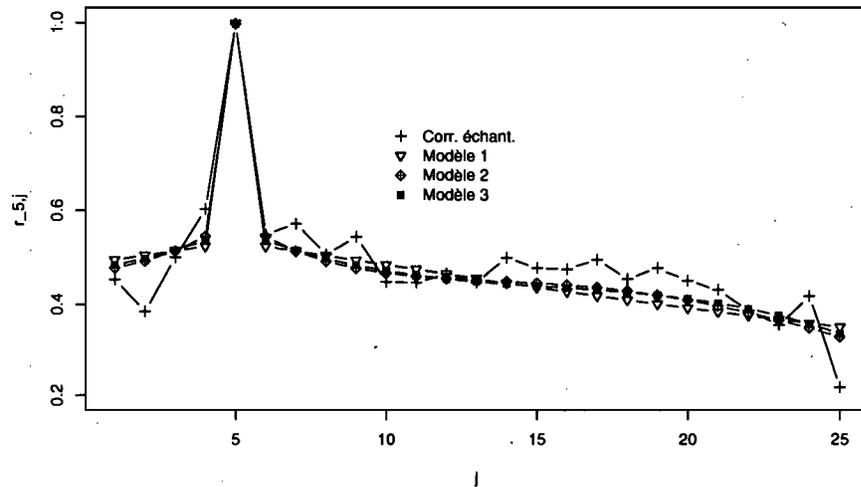


Figure 4.14 – Modèles pour les corrélations $r_{i,j}^X$ dans le centre C1, (graphique pour $i = 5$ et $j = 1, \dots, 25$).

homogène par rapport aux autres modèles dans notre cas, nous présentons quand même les résultats des simulations avec tous les modèles. Cela permet en même temps d'illustrer l'avantage des modèles qui tiennent compte de la dépendance intra-journalière.

Si nous examinons les résultats de la simulation avec un processus de Poisson non homogène, nous remarquons qu'une fonction de taux spline fait légèrement mieux qu'une fonction à taux constant par morceaux. Nous n'avons pas obtenu de bons résultats avec le processus de Poisson non homogène en général. Avec le processus de Poisson doublement stochastique, les résultats sont meilleurs et avec une spline cubique pour la fonction de taux les résultats s'améliorent un peu plus qu'avec un taux fixe par morceaux. Le modèle NORTA donne des bons résultats surtout dans le cas de la mi-journée de travail (de 8h00 à 14h00). Ce modèle et celui qui a moins de paramètres (NORTA réduit), où les corrélations sont définies selon

l'équation (3.47), deviennent presque équivalents et les deux sont plus réalistes et plus proches des données empiriques. Il faut noter que c'est particulier à ce centre où le flux des appels entrants n'est pas si important et on verra la différence dans le centre d'appels C2.

Il faut rappeler ici que la simulation ne constitue pas le principal critère pour comparer ou classer les modèles, puisque d'autres composantes stochastiques de la simulation peuvent être une source d'erreur. Cependant, elle permet quand même de valider nos modèles quand les mesures de performance obtenues sont proches de celles calculées avec les données. La simulation permet aussi d'observer la sensibilité des mesures de performance face au changement du processus d'arrivée. Dans le tableau 4.7, nous présentons les arrivées par demi-heure obtenus dans la simulation des 30000 journées indépendantes. En comparant les résultats, une fois de plus nous remarquons que les modèles qui tiennent compte de la dépendance intra-journalière donnent des meilleurs résultats, notamment le processus de Poisson doublement stochastique avec une fonction spline et le processus NORTA.

Pour pouvoir comparer d'avantage les résultats obtenus selon les différentes méthodes nous avons calculé

$$\rho_m = \text{Corr} \left(\sum_{i=1}^m X_i, \sum_{i=m+1}^{25} X_i \right), \quad m = 1, \dots, 24, \quad (4.2)$$

ce qui permet de capter les corrélations entre différentes parties de la journée [5]. Nous avons aussi calculé le coefficient de variation, qui est une mesure relative de dispersion des arrivées dans chaque intervalle de temps. Bien que ce coefficient ne dépend que de la loi marginale, il permet tout de même de faire des comparaisons si on considère les coefficients de variation pour les arrivées dans tous les intervalles obtenus avec les différents modèles et ceux obtenus avec les données. Ces coefficients

Demi-heure	Données	PPNH (spline)	PPDS (constante)	PPDS (spline)	NORTA	NORTA (réduit)
8h00-8h30	24.71 (0.71)	25.86±0.04	24.33±0.05	25.87±0.05	24.79±0.06	24.81±0.07
8h30-9h00	36.97 (0.67)	37.27±0.04	36.63±0.06	37.28±0.07	37.08±0.06	37.02±0.06
9h00-9h30	50.16 (1.01)	49.63±0.06	48.90±0.08	49.56±0.08	50.09±0.09	50.10±0.09
9h30-10h00	56.65 (1.16)	56.53±0.06	55.81±0.08	56.52±0.09	56.62±0.10	56.63±0.10
10h00-10h30	59.45 (1.13)	59.16±0.06	58.48±0.09	59.15±0.09	59.51±0.10	59.47±0.10
10h30-11h00	59.14 (1.11)	59.26±0.06	57.87±0.09	59.33±0.09	59.14±0.10	59.19±0.10
11h00-11h30	60.79 (1.32)	60.05±0.06	58.52±0.09	59.96±0.09	60.84±0.12	60.78±0.12
11h30-12h00	57.87 (1.33)	57.85±0.06	56.08±0.08	57.96±0.09	58.00±0.11	58.03±0.11
12h00-12h30	54.71 (1.39)	54.80±0.06	53.13±0.08	54.78±0.09	54.78±0.11	54.79±0.11
12h30-13h00	54.90 (1.16)	54.80±0.05	54.38±0.08	54.94±0.09	54.90±0.11	54.87±0.11
13h00-13h30	56.71 (1.32)	56.38±0.06	55.20±0.08	56.36±0.09	56.67±0.12	56.67±0.11
13h30-14h00	53.93 (1.20)	54.19±0.06	53.48±0.08	54.20±0.09	53.99±0.10	54.05±0.11
14h00-14h30	54.19 (1.32)	54.01±0.06	52.47±0.08	54.05±0.09	54.30±0.11	54.24±0.11
14h30-15h00	52.59 (1.11)	52.68±0.06	52.15±0.08	52.70±0.08	52.58±0.10	52.55±0.10
15h00-15h30	53.46 (0.99)	53.48±0.06	52.53±0.08	53.51±0.09	53.57±0.09	53.55±0.09
15h30-16h00	56.80 (1.12)	56.57±0.06	56.15±0.08	56.66±0.09	56.77±0.10	56.72±0.10
16h00-16h30	58.23 (1.05)	57.23±0.06	57.91±0.09	57.90±0.09	58.30±0.09	58.25±0.09
16h30-17h00	53.10 (1.06)	52.78±0.06	52.22±0.08	52.74±0.08	53.02±0.10	52.96±0.10
17h00-17h30	43.66 (0.99)	43.98±0.05	43.14±0.07	44.00±0.07	43.71±0.09	43.66±0.09
17h30-18h00	39.57 (0.92)	39.30±0.05	39.03±0.06	39.27±0.07	39.61±0.08	39.53±0.08
18h00-18h30	34.16 (0.79)	34.15±0.05	33.86±0.06	34.21±0.06	34.14±0.08	34.05±0.08
18h30-19h00	30.63 (0.94)	30.45±0.04	29.99±0.05	30.42±0.06	30.64±0.08	30.61±0.08
19h00-19h30	25.50 (0.73)	25.74±0.04	25.48±0.05	25.71±0.05	25.44±0.06	25.52±0.06
19h30-20h00	23.31 (0.62)	23.09±0.04	23.42±0.05	23.07±0.05	23.29±0.06	23.27±0.06
20h00-20h30	18.81 (0.65)	18.46±0.03	18.58±0.04	18.43±0.04	18.76±0.06	18.75±0.06

Tableau 4.7 – Le nombre moyen d’arrivées par demi-heure pour une journée de 8h00 à 20h30 au centre C1. Les données empiriques sont présentés avec l’écart-type “bootstrap” et sont comparées avec des données simulées (30000 jours) selon différents modèles.

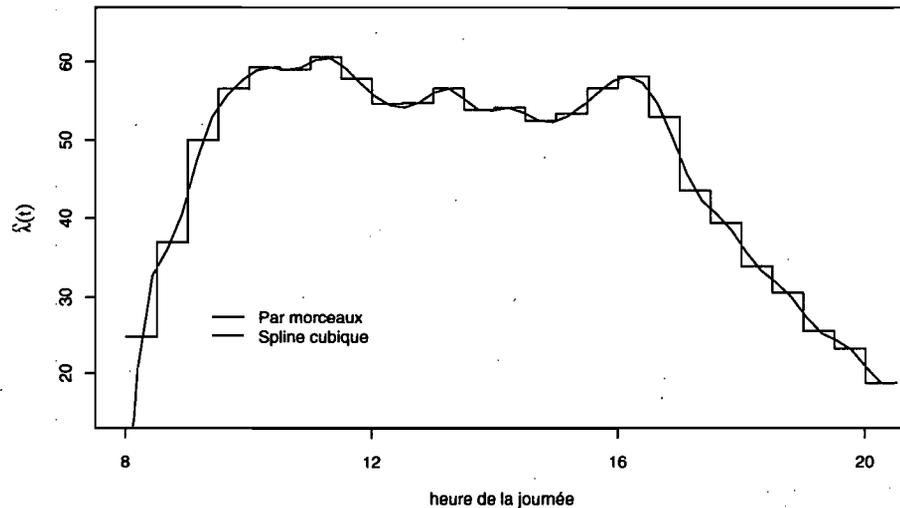


Figure 4.15 – Fonction de taux estimée par une spline vs une fonction constante par morceaux dans le centre C1.

sont définis comme suit :

$$CV(X_i) = \frac{\sqrt{\text{Var}(X_i)}}{E[X_i]}, \quad \text{pour } i = 0, \dots, d.$$

L'examen du graphique 4.16 permet de constater que le modèle NORTA semble être le plus adéquat par rapport aux autres modèles. Il fournit la meilleure estimation de point de vue des corrélations. La version réduite de ce modèle donne des résultats moins bons, alors que le processus de Poisson doublement stochastique tend à surestimer les corrélations. En résumé, le modèle NORTA est meilleur, suivi de NORTA réduit et enfin le processus de Poisson doublement stochastique. Il ne faut pas oublier que la structure des corrélations a été légèrement modifiée avec la modélisation de ces corrélations selon la formulation (3.47). Il est donc normal que les corrélations obtenues avec le modèle NORTA réduit s'éloignent un peu de celles

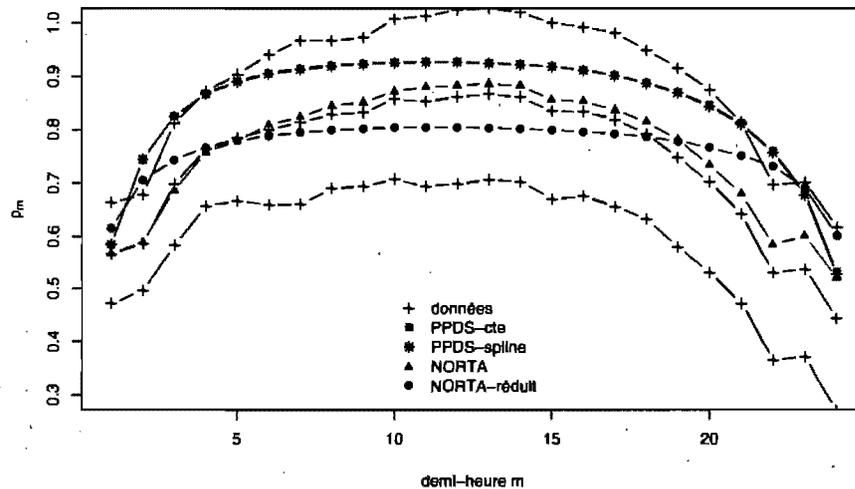


Figure 4.16 – Comparaison des modèles selon les corrélations estimées entre les différentes parties de la journée dans le centre C1.

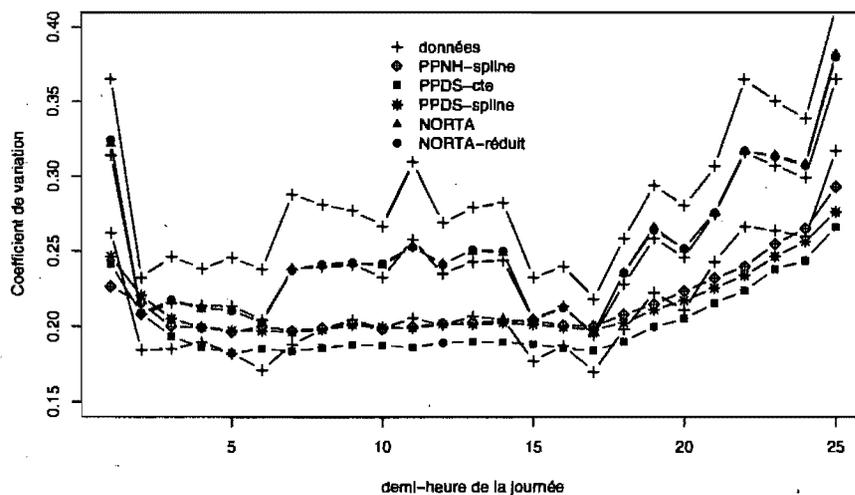


Figure 4.17 – Comparaison des modèles selon le coefficient de variation pour chaque demi-heure dans le centre C1.

des données. Par contre, ce modèle fait toujours mieux que le processus de Poisson doublement stochastique qui tend à surestimer les corrélations. Le modèle NORTA réduit présente un compromis entre les deux autres modèles, il permet de réduire le nombre de paramètres par rapport au modèle NORTA et d'avoir des corrélations plus proches des données.

De point de vue variabilité, selon le graphique 4.17, nous remarquons que les modèles NORTA et NORTA réduit donnent les mêmes coefficients de variation. Sur ce graphique, les lignes qui tracent ces coefficients sont confondues pour ces deux modèles et ils sont très proches de celles obtenues avec les données. Les résultats obtenus avec les modèles NORTA et NORTA réduit sont meilleurs que ceux obtenus avec les modèles basés sur le processus de Poisson (homogène ou doublement stochastique). Ces derniers tendent à sous-estimer les coefficients de variation réels.

4.2.1.1 Fonction spline pour le taux versus fonction constante par morceaux

Pour justifier le choix de la fonction spline pour le taux d'arrivée dans le cas du processus de Poisson doublement stochastique, nous avons procédé à la simulation du même centre C1 pendant les trois premières heures de la journée en considérant des intervalles de temps plus courts, soit 15 minutes. Avec le processus PPDS-constante, nous n'avons pas le choix de garder le taux constant pour les deux périodes de 15 minutes de chaque demi-heure. Alors que pour le PPDS-spline, la fonction spline permet d'évaluer la fonction de taux à n'importe quelle minute de la demi-heure. Nous avons donc simulé 30000 journées indépendantes dans le centre C1 et nous avons recueilli les quatre mesures suivantes : nombre d'appels servis, nombre d'abandons, qualité de service (pourcentage des clients ayant attendu moins de 20 secondes) et le taux d'occupation des agents. Nous avons comparé ces mesures

	Qualité de service %	Nbr. d'abandons	Nbr. d'appels servis	Occup. des agents %	Temps d'attente moy. appels reçus	Temps d'attente moy. dans la file
Données	83.8±.3	20.1±6.1	606±15	65.6±1.6	16.2±3.7	83.9±6.3
PPNH-cte	89.8±ε	10.3±.1	615.7±.2	66.7±ε	6.7±.1	51.3±.3
PPNH-spline	84.0±ε	19.8±ε	606±.2	67.3 ±.1	13.5±.1	66.7±.2

Tableau 4.8 – Quelques mesures obtenues par la simulation d'une période d'une journée de 8h00 à 14h00, du type mardi, mercredi ou jeudi au centre C1 en considérant des intervalles de temps de 15 mn.

avec celles des données pour la même période de temps. Les résultats présentés dans le tableau 4.7, montrent l'avantage d'utiliser une fonction spline. Les résultats avec le modèle PPDS-spline s'approchent plus de la réalité. Avec cette fonction le fait d'avoir des données agrégées ne constitue pas un obstacle pour modéliser les arrivées dans des intervalles de temps plus courts. Si l'on veut assurer une certaine flexibilité dans la simulation quant au choix des intervalles de temps, ou pour pouvoir connaître le taux d'arrivée à n'importe quel instant de la journée, il faudra bien se doter d'une fonction de taux continue sur tout l'intervalle de temps, comme la fonction spline qui semble être un bon choix dans notre cas. Il serait donc intéressant de trouver un moyen pour combiner le modèle NORTA et la fonction spline pour tirer avantage des deux en même temps. Nous avons remarqué que le processus NORTA permet de générer le nombre d'arrivées par intervalle de temps avec les quantités de corrélations désirées, alors que la fonction spline permet de générer les instants de ces arrivées de façon adéquate, puisqu'elle capte la variabilité du taux sur toutes les périodes de temps étudiées. Ceci constituera une bonne avenue de recherche future.

4.2.2 Simulation des arrivées au centre d'appels C2

Dans ce centre, nous considérons les arrivées par intervalles de 15 min chacun, de 8h00 à 20h00 et nous disposons de données de $n = 72$ jours. Nous avons donc 48 intervalles $[r_i, r_{i+1}]$, $i = 0, \dots, h = 47$. Il faut encore noter que pour ce centre nous ne disposons que des données sur les arrivées et que pour pouvoir faire une comparaison entre les différents modèles, nous devons simuler avec chaque modèle le nombre d'arrivées pour chaque période de temps pour obtenir le total de la journée et répéter l'opération pendant un certain nombre de journées indépendantes et ensuite les comparer avec les données observées dans le centre d'appels étudié.

Pour le modèle NORTA réduit, nous avons utilisé le critère AIC pour sélectionner un modèle pour les corrélations. Le modèle 3, donné en (3.49), présente un plus petit AIC parmi les trois modèles. Nous avons obtenu -3069.98 pour le modèle 1, -3087.23 pour le modèle 2 et -3090.57 pour le modèle 3. Nous remarquons qu'il y a une légère différence entre les trois valeurs du AIC. L'estimation des paramètres avec la méthode des moindres carrés non linéaires a donné les résultats suivants :

$$\hat{p}_1 = -701.4 \pm 912.6$$

$$\hat{p}_2 = 29221 \pm 37626$$

$$\hat{q}_1 = -66.68 \pm 31.32$$

$$\hat{q}_2 = 1079 \pm 789.6$$

$$\hat{q}_3 = 30570 \pm 40950,$$

avec une erreur standard $\hat{\sigma}_\epsilon$ (écart-type des résidus) de 0.0640, alors que l'erreur est de 0.0654 et 0.0643 pour les modèles 1 et 2, respectivement.

Nous avons dû simuler 18000 journées indépendantes pour ensuite comparer les arrivées selon les corrélations entre les différentes parties de la journée ou les

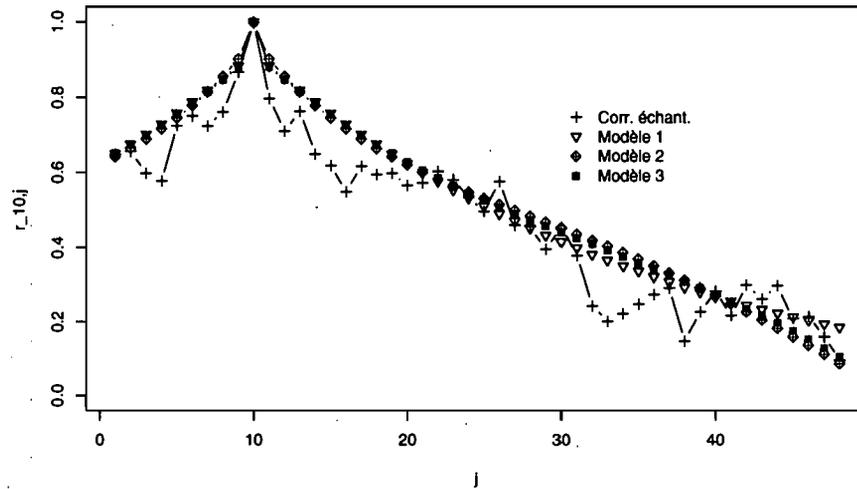


Figure 4.18 – Modèles pour les corrélations $r_{i,j}^X$ dans le centre C2, (graphique pour $i = 10$ et $j = 1, \dots, 48$).

coefficients de variation, ou encore les arrivées moyennes par période de temps.

Dans le tableau 4.9 nous présentons le résultats des estimations, à partir des données, des paramètres de chaque distribution marginale (binomiale négative) pour le nombre d'arrivées à chaque quart d'heure.

Pour les données observées, nous avons utilisé 250 répétitions “bootstrap” pour obtenir les intervalles de confiances et les 18000 journées (250×72), c.-à.-d. que les données de chaque jours sont rééchantillonnées 250 fois (avec remise) pour pouvoir obtenir 18000 journées. Cela permettra de comparer les mesures de performance calculées avec les données et celles obtenues avec la simulation avec le même nombre de jours.

Dans le tableau 4.10, nous pouvons remarquer que les résultats obtenus sont presque tous aussi bons que les données. Cependant, nous remarquons qu'avec le modèle PPDS et une fonction spline, il y a un léger gain par rapport au modèle où la

fonction de taux est constante par morceaux. Mais, les deux modèles surestiment les coefficients de corrélation. Pour cela, le modèle NORTA serait préférable si nous examinons le graphique 4.20, où ce modèle nous permet d'avoir des corrélations très proches de celles des données. Si nous comparons les résultats obtenus dans ce centre avec ceux obtenus dans le centre C1, nous pouvons facilement observer le succès obtenu avec le modèle NORTA et à un degré moindre le modèle NORTA réduit, dans le cas où le centre est plus important de point de vue nombre d'appels reçus par jour et de point de vue variabilité à l'intérieur de la journée.

Avec le coefficient de variation (figure 4.19) et sur la base d'une comparaison entre les coefficients obtenus avec ce modèle et celles estimées par l'échantillon, on peut déjà faire une première constatation : il y a plus de variabilité dans les arrivées durant les premiers quarts d'heures et dans les dernières demi-heures de la journée. C'est durant ces moments de la journée que tous les modèles, à part le modèle NORTA, sous-estiment les coefficients de variation théoriques. Nous remarquons aussi qu'il n'y a pas d'écart entre les coefficients obtenus à partir des données observées et celles obtenues par le modèle NORTA réduit.

Pour résumer, nous pouvons constater que le processus de Poisson non homogène ne donne pas de bons résultats de point de vue des corrélations. Les résultats obtenus avec le processus NORTA sont aussi bons, sinon meilleurs, que ceux obtenus avec le processus doublement stochastique. En ce qui concerne le modèle NORTA réduit, où les corrélations sont modélisées selon le modèle 3 donné en (3.49), nous remarquons qu'il y a peu de perte d'information et les résultats de la simulation sont semblables qu'avec un modèle qui utilise tous les paramètres.

Il faut mentionner ici que les corrélations échantillonnales sont très petites et tendent à s'annuler entre des quarts d'heures éloignés dans la journée, comme par exemple entre celle du début de la matinée et de la fin de la journée. Le proces-

Quart d'heure	$\hat{\mu}_i$	\hat{s}_i	\hat{p}_i
8h00-8h15	94.9±5.5	17.9	0.158
8h15-8h30	127.1±7.1	19.5	0.133
8h30-8h45	182.9± 9.2	24.1	0.116
8h45-9h00	301.7±12.3	35.8	0.106
9h00-9h15	917.8±37.6	32.6	0.034
9h15-9h30	887.1±31.6	45.5	0.049
9h30-9h45	1050.1±32.0	62.8	0.056
9h45-10h00	1178.4±29.1	98.0	0.077
10h00-10h15	1420.2±32.8	111.9	0.073
10h15-10h30	1512.7±33.0	126.3	0.077
10h30-10h45	1606.1±29.1	180.1	0.101
10h45-11h00	1640.5±32.2	149.7	0.084
11h00-11h15	1685.3±31.4	170.1	0.092
11h15-11h30	1685.2±35.7	125.5	0.069
11h30-11h45	1689.1±36.1	126.6	0.070
11h45-12h00	1676.1±35.0	134.0	0.074
12h00-12h15	1644.9±37.1	115.0	0.065
12h15-12h30	1604.8±36.0	119.8	0.069
12h30-12h45	1573.7±34.6	123.6	0.073
12h45-13h00	1507.2±34.0	116.3	0.072
13h00-13h15	1468.3±34.3	107.9	0.068
13h15-13h30	1484.3±34.4	107.4	0.067
13h30-13h45	1506.5±35.0	109.2	0.068
13h45-14h00	1491.8±36.0	103.5	0.065
14h00-14h15	1517.6±35.3	109.7	0.067
14h15-14h30	1501.0±37.2	96.7	0.061
14h30-14h45	1499.5±38.2	91.1	0.057
14h45-15h00	1447.0±36.7	91.5	0.059
15h00-15h15	1454.9±37.3	90.8	0.059
15h15-15h30	1427.7±35.3	96.8	0.064
15h30-15h45	1422.8 ±36.0	91.6	0.060
15h45-16h00	1394.0±36.2	86.3	0.058
16h00-16h15	1407.0±32.0	113.3	0.075
16h15-16h30	1404.0±29.9	130.9	0.085
16h30-16h45	1419.5±31.1	118.9	0.077
16h45-17h00	1472.7±33.1	114.9	0.072
17h00-17h15	1469.8±35.6	95.2	0.061
17h15-17h30	1464.6±32.6	116.7	0.074
17h30-17h45	1427.7±31.0	123.4	0.080
17h45-18h00	1322.4±29.9	109.2	0.076
18h00-18h15	1055.9±27.6	83.4	0.073
18h15-18h30	902.2±27.8	54.3	0.057
18h30-18h45	818.5±23.7	68.1	0.077
18h45-19h00	751.2±21.7	71.6	0.087
19h00-19h15	634.2±18.7	67.4	0.096
19h15-19h30	594.0±17.8	65.1	0.098
19h30-19h45	544.6±15.4	75.9	0.122
19h45-20h00	503.4±16.9	50.7	0.091

Tableau 4.9 – Estimation du taux moyen d'arrivée et des paramètres de la loi binomiale négative pour chaque demi-heure avec les données du centre C2.

Quart d'heure	Données	PPNH (spline)	PPDS (constante)	PPDS (spline)	NORTA	NORTA (réduit)
8h00-8h15	94.84 (2.78)	95.10±0.10	95.01±0.25	95.20±0.24	94.76±0.48	94.94±0.48
8h15-8h30	127.01 (3.41)	131.15±0.12	127.08±0.30	131.22±0.31	126.88±0.61	126.99±0.60
8h30-8h45	183.27 (4.65)	177.91±0.14	182.85±0.39	178.12±0.39	182.70±0.78	183.00±0.77
8h45-9h00	301.80 (6.32)	339.76±0.19	301.98±0.59	339.49±0.66	301.61±1.04	302.29±1.04
9h00-9h15	918.96 (19.65)	869.98±0.30	919.00±1.59	870.23±1.51	916.75±3.20	92.147±3.22
9h15-9h30	888.16 (16.74)	909.43±0.31	887.13±1.55	909.35±1.58	886.37±2.65	888.87±2.65
9h30-9h45	1051.27 (15.88)	1041.14±0.33	1049.82±1.79	1041.95±1.78	1049.42±2.68	1050.33±2.66
9h45-10h00	1178.47 (15.24)	1188.86±0.36	1178.70±2.01	1189.23±2.03	1177.00±2.42	1178.55±2.41
10h00-10h15	1420.59 (16.96)	1408.81±0.39	1421.13±2.41	1409.58±2.37	1418.71±2.74	1421.31±2.74
10h15-10h30	1511.34 (15.21)	1515.91±0.40	1511.84±2.54	1516.57±2.54	1510.52±2.77	1513.67±2.75
10h30-10h45	1606.43 (15.67)	1601.66±0.41	1606.36±2.69	1602.20±2.68	1605.03±2.48	1606.06±2.47
10h45-11h00	1640.75 (16.63)	1642.79±0.42	1640.12±2.74	1642.34±2.75	1639.34±2.78	1641.28±2.75
11h00-11h15	1686.77 (15.85)	1677.45±0.42	1685.66±2.82	1677.84±2.80	1682.73±2.68	1685.38±2.65
11h15-11h30	1684.52 (19.28)	1680.10±0.42	1685.16±2.82	1680.69±2.83	1683.48±3.09	1685.57±3.06
11h30-11h45	1688.83 (18.10)	1680.19±0.42	1688.61±2.85	1679.90±2.82	1686.53±3.08	1688.59±3.06
11h45-12h00	1678.15 (17.19)	1674.50±0.42	1676.60±2.81	1674.16±2.82	1674.07±2.98	1675.25±2.95
12h00-12h15	1641.98 (18.27)	1644.26±0.42	1644.53±2.75	1644.47±2.76	1644.48±3.15	1645.12±3.12
12h15-12h30	1606.03 (17.92)	1605.68±0.42	1604.90±2.69	1605.76±2.69	1603.64±3.02	1604.60±2.97
12h30-12h45	1573.34 (18.10)	1571.30±0.41	1574.05±2.63	1571.63±2.65	1573.96±2.93	1574.42±2.89
12h45-13h00	1508.40 (17.86)	1508.80±0.40	1507.48±2.52	1509.09±2.54	1506.67±2.87	1506.68±2.86
13h00-13h15	1466.90 (16.97)	1470.92±0.40	1468.43±2.48	1470.83±2.48	1468.23±2.89	1468.16±2.88
13h15-13h30	1486.48 (16.98)	1485.13±0.40	1484.13±2.50	1484.94±2.49	1483.29±2.94	1484.17±2.94
13h30-13h45	1507.72 (16.71)	1503.60±0.40	1507.35±2.53	1503.79±2.52	1505.70±2.95	1506.10±2.95
13h45-14h00	1492.07 (17.26)	1495.47±0.40	1492.41±2.50	1495.31±2.52	1491.62±2.99	1492.06±2.98
14h00-14h15	1516.52 (16.96)	1514.20±0.40	1518.26±2.57	1514.33±2.55	1517.35±2.98	1516.83±2.96
14h15-14h30	1497.66 (20.17)	1503.24±0.40	1501.14±2.53	1503.21±2.54	1499.72±3.13	1500.66±3.10
14h30-14h45	1498.20 (18.97)	1495.29±0.40	1499.06±2.52	1495.81±2.53	1498.74±3.22	1498.87±3.19
14h45-15h00	1446.48 (19.40)	1451.70±0.39	1447.33±2.44	1451.51±2.46	1446.86±3.10	1447.18±3.07
15h00-15h15	1454.90 (20.45)	1451.66±0.40	1455.31±2.46	1451.78±2.46	1455.22±3.12	1454.89±3.09
15h15-15h30	1427.24 (16.94)	1430.08±0.39	1427.76±2.40	1430.50±2.40	1426.93±2.99	1427.05±2.94
15h30-15h45	1422.53 (18.91)	1420.35±0.39	1422.82±2.40	1420.00±2.40	1421.97±3.04	1421.66±3.02
15h45-16h00	1393.32 (18.59)	1397.04±0.39	1393.94±2.36	1397.55±2.36	1393.93±3.08	1393.50±3.04
16h00-16h15	1407.55 (14.95)	1405.06±0.39	1407.71±2.38	1405.80±2.37	1406.51±2.73	1405.82±2.70
16h15-16h30	1403.72 (14.97)	1405.03±0.39	1404.32±2.37	1405.24±2.37	1403.91±2.55	1403.15±2.52
16h30-16h45	1420.17 (16.28)	1422.10±0.39	1419.41±2.40	1422.31±2.40	1419.02±2.68	1419.04±2.67
16h45-17h00	1474.37 (16.04)	1469.05±0.40	1473.64±2.48	1469.17±2.49	1471.30±2.83	1470.45±2.80
17h00-17h15	1469.39 (18.44)	1470.99±0.39	1469.24±2.48	1471.23±2.49	1468.34±3.08	1467.81±3.04
17h15-17h30	1464.41 (17.61)	1462.72±0.40	1464.56±2.48	1463.38±2.47	1464.72±2.78	1462.94±2.76
17h30-17h45	1428.02 (15.64)	1426.44±0.39	1427.91±2.41	1426.95±2.41	1428.10±2.67	1426.44±2.64
17h45-18h00	1323.27 (16.03)	1312.18±0.38	1323.04±2.23	1312.13±2.23	1322.48±2.61	1321.21±2.60
18h00-18h15	1055.59 (14.87)	1064.02±0.33	1056.10±1.80	1063.65±1.82	1055.85±2.38	1054.56±2.36
18h15-18h30	903.75 (13.78)	904.23±0.31	902.19±1.56	904.00±1.57	902.15±2.49	901.07±2.46
18h30-18h45	816.60 (12.70)	820.02±0.30	818.66±1.43	819.98±1.44	818.41±2.03	817.57±2.02
18h45-19h00	750.53 (11.16)	746.89±0.28	751.32±1.32	747.16±1.31	751.04±1.82	750.04±1.81
19h00-19h15	634.84 (9.33)	640.08±0.26	633.69±1.13	640.21±1.14	635.15±1.61	632.97±1.59
19h15-19h30	593.79 (9.52)	590.70±0.25	594.11±1.07	591.36±1.07	594.39±1.53	593.52±1.52
19h30-19h45	543.40 (7.36)	550.60±0.24	544.14±0.99	551.00±0.99	545.47±1.32	543.86±1.31
19h45-20h00	504.09 (8.26)	476.75±0.23	503.54±0.91	476.61±0.87	504.34±1.47	503.18±1.45

Tableau 4.10 – Le nombre moyen d'arrivées par quart d'heure pour une journée de 8h00 à 20h00 au centre C2. Les données empiriques sont présentés avec l'écart-type "bootstrap" et sont comparées avec des données simulées (18000 jours) selon différents modèles.

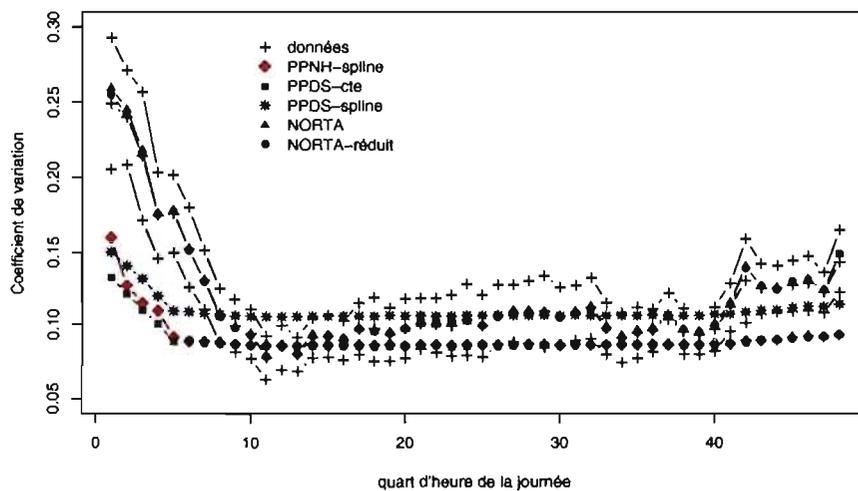


Figure 4.19 – Comparaison des modèles selon le coefficient de variation pour chaque quart d’heure dans le centre C2.

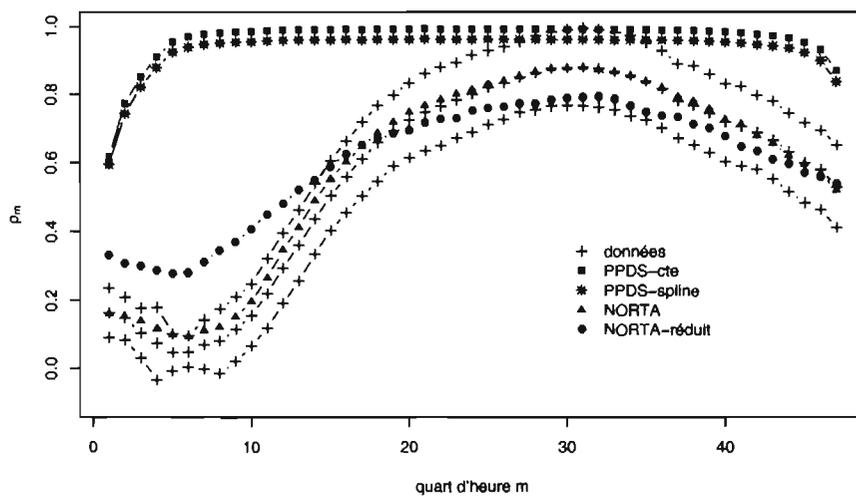


Figure 4.20 – Comparaison des modèles selon les corrélations estimées entre les différentes parties de la journée dans le centre C2.

Le modèle sus-NORTA permet de capturer ces variations au niveau des corrélations, puisqu'il permet de générer des arrivées avec les quantités de corrélations désirées (celles des données), alors que le processus de Poisson doublement stochastique permet de générer des arrivées avec des quantités élevées de corrélations. Ceci explique l'écart entre les corrélations dans le graphique 4.20 et le succès obtenu avec le modèle NORTA et à un degré moindre le modèle NORTA réduit. Il faut rappeler ici qu'avec le processus de Poisson doublement stochastique que nous présentons le vecteur aléatoire X , qui représente le nombre d'arrivées par intervalles de temps, suit une loi multinomiale négative et que la corrélation entre chaque couple d'éléments de X est toujours positive et elle tend à surestimer les corrélations réelles.

CHAPITRE 5

LES ARRIVÉES AU CENTRE D'APPELS D'UNE BANQUE

5.1 Description du centre d'appels C3

Ce centre reçoit des appels en provenance de 4 états américains : New York, Pennsylvanie, Rhode Island et Massachussetts. Il est ouvert pendant 24 heures, les 7 jours de la semaine. Pendant les jours fériés le centre est ouvert, mais avec un personnel réduit. La grande majorité des appels arrivent au début au système automatique VRU ("Voice Response Unit") et une autre partie des appels sont transférés directement à partir des lignes d'informations et quelques appels aboutissent directement à des agents. Près de 20% de ces appels continuent à travers ce système dans le but de solliciter la communication directe avec un agent. La banque de données mise à notre disposition contient les détails sur cette proportions des appels qui attendent pour parler à un agent libre.

Ce centre emploie entre 900 et 1 200 agents les jours de semaine et entre 200 et 500 pour le samedi et dimanche. Le nombre total des appels qui aboutissent au système VRU, ainsi que les autres types, se situe entre 200 000 et 270 000 la semaine, entre 120 000 et 140 000 le samedi et entre 60 000 et 100 000 le dimanche.

5.2 Analyse descriptive des données

Au départ, nous disposons de données s'étalant sur près de deux ans. Après avoir identifié et éliminé les journées spéciales et les journées où il y avait des problèmes et des difficultés pour collecter les données, nous avons pu garder un total de 389 jours.

En deuxième étape, nous avons procédé à l'élimination de l'effet du jour de la

semaine. Il paraît évident dès le début que le samedi et dimanche ne ressemblent pas aux autres jours. Dans le tableau 5.1, nous présentons le volume moyen des appels reçus pour chaque type de jour avec la marge d'erreur, en supposant bien sûr que les journées sont indépendantes. Les volumes d'appels sont plus élevés pendant la semaine que le samedi et dimanche. Dans le graphique 5.1 des diagrammes en boîtes, on peut aussi constater une ressemblance entre les arrivées de mercredi, jeudi et vendredi. Les autres jours présentent des différences.

A l'intérieur d'une journée, on enregistre le niveau d'appels le plus élevé entre 10h00 et 12h00 et le niveau baisse un peu jusqu'à 18h00. Par la suite, le nombre d'appels continue de baisser et de minuit jusqu'à 7h du matin on observe le niveau le plus bas. Nous présentons dans le graphique 5.3 le nombre moyen d'appels par périodes de demi-heures dans une journée du type mercredi, jeudi ou vendredi et dans le graphique 5.2 celui du volume moyen d'appels reçus pour tous les jours de la semaine. Dans le graphique 5.4, nous présentons la série du volume total des arrivées pour ces journées.

Nous commençons tout d'abord par éliminer l'effet du jour de la semaine et avoir des journées semblables. Avec les graphiques 5.1 et 5.2, nous remarquons déjà que mercredi, jeudi et vendredi se ressemblent, lundi et mardi sont différents, alors que dimanche et samedi sont complètement différents des autres journées. Nous avons enlevé les journées où le système d'enregistrement des données a mal fonctionné, tel que mentionné par Trofimov et al. [88]. A cet effet, nous avons utilisé la méthode basée sur l'étendue inter quartiles (EIQ). Dans la suite du travail et pour l'application des modèles, nous considérons donc seulement les jours du type mercredi, jeudi et vendredi, soit un total de 143 jours.

En ce qui concerne les arrivées par période de temps dans la journée, nous avons considéré seulement 34 périodes de 30 minutes chacune de 7h00 à minuit. Pendant

Jour	Estimation
Lundi	59977±3278
Mardi	58809±2084
Mercredi	56328±1788
Jeudi	55274±2081
Vendredi	56726±1530
Samedi	23821±657
Dimanche	11329±576

Tableau 5.1 – Nombre moyen d'appels par type de journée dans le centre C3.

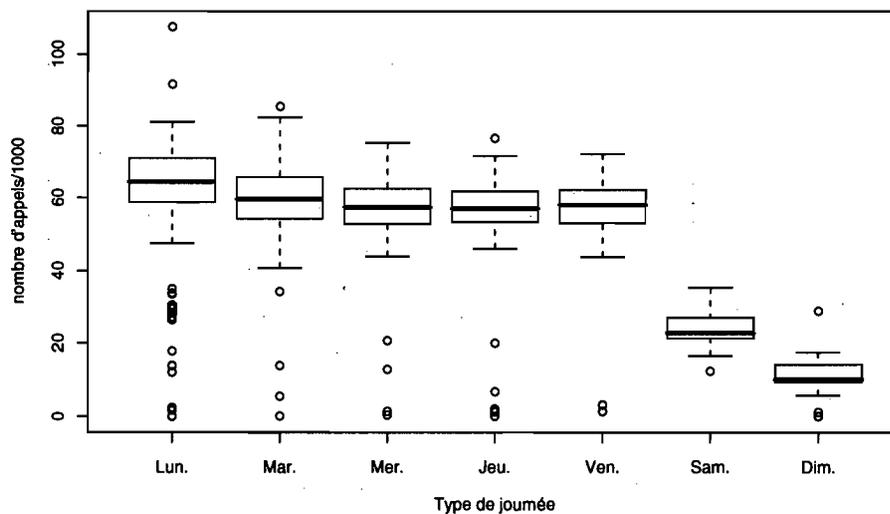


Figure 5.1 – Diagrammes en boîtes pour le volume d'appels par type de journée dans le centre C3.

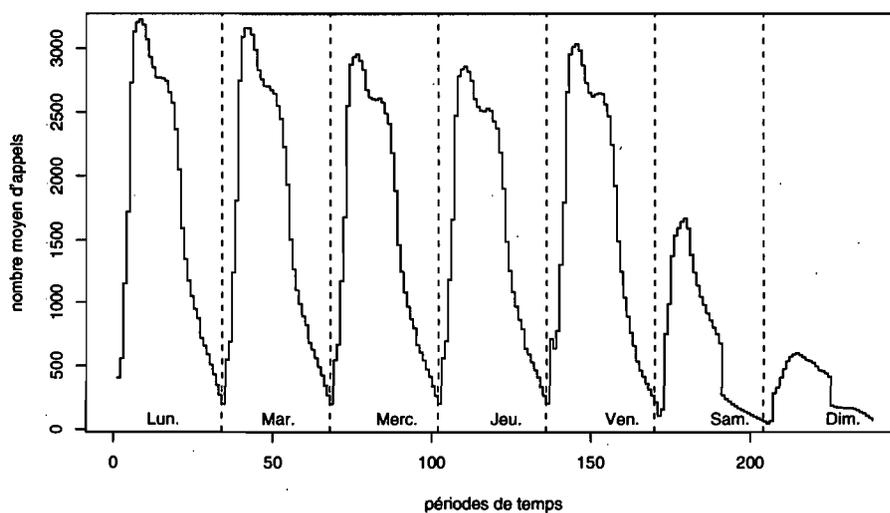


Figure 5.2 – Nombre moyen d’arrivées par intervalle de temps et par type de journée dans le centre C3.

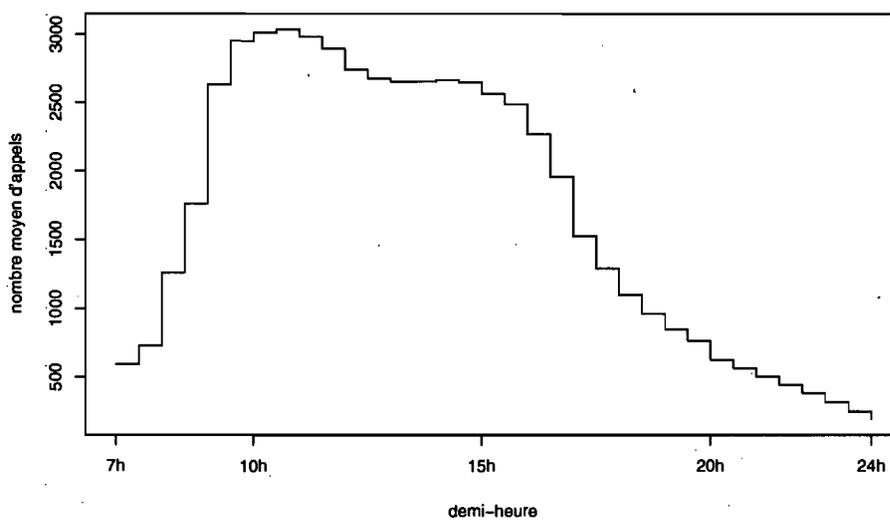


Figure 5.3 – Nombre moyen d’arrivées par intervalle de temps pour une journée de type mercredi, jeudi ou vendredi dans le centre C3.

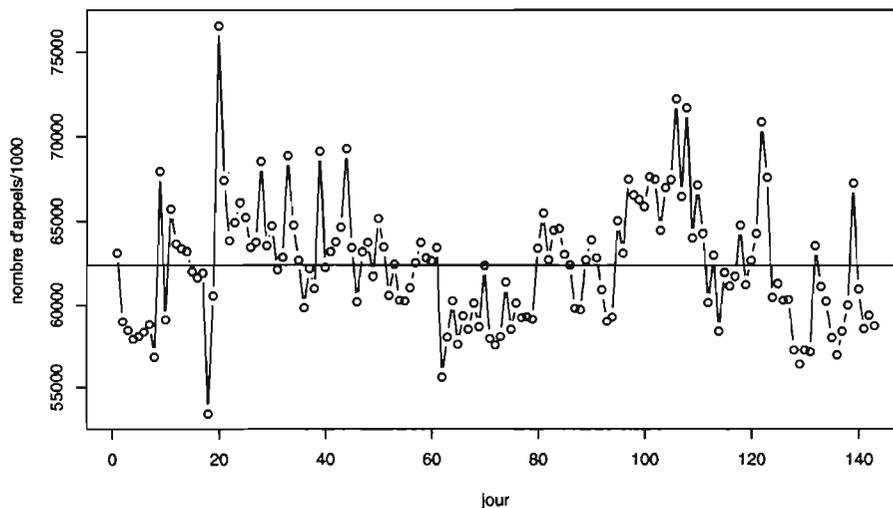


Figure 5.4 – Série du volume d’appels reçus par jour de mercredi à vendredi dans le centre C3.

la période de temps de minuit à 7h00, il y a très peu d’appels par rapport aux autres périodes de la journée et pour cela nous n’avons pas considéré cette période. Nous allons donc appliquer les modèles NORTA, NORTA réduit et le processus de Poisson doublement stochastique pour capter la dépendance intra-journalière. Nous supposons donc que les arrivées d’une journée à l’autre sont indépendantes.

5.3 Application et comparaison des modèles

La présente section fournit un résumé des résultats obtenus avec la simulation des arrivées au centre d’appels C3 selon les différents modèles proposés. Nous avons procédé à la simulation des arrivées par intervalle de temps sur 35 750 journées indépendantes. Nous ne disposons pas d’autres données pour la simulation du centre au complet. Nous avons fait intervenir la méthode “bootstrap” pour les intervalles de confiance pour les données empiriques avec 250 répliques des 143 journées

observées pour donner un total de 35750 journées pour les données.

Pour le modèle NORTA réduit, nous avons encore considéré le critère AIC pour sélectionner le modèle adéquat pour les corrélations. Nous avons obtenu les résultats suivants : -2295.59 pour le modèle 1, -2315.63 pour le modèle 2 et -2312.80 pour le modèle 3. Nous avons donc opté pour le deuxième modèle, donné en (3.48), et les estimations des paramètres, par la méthode des moindres carrés non linéaires, a donné les résultats suivants :

$$\hat{p}_1 = -1.60 \times 10^{-5} \pm 1.96 \times 10^{-5}$$

$$\hat{p}_2 = 1.44 \times 10^{-3} \pm 9.16 \times 10^{-4}$$

$$\hat{p}_3 = -0.0622 \pm 0.012$$

$$\hat{p}_4 = 0.900 \pm 0.041,$$

avec une erreur standard $\hat{\sigma}_\epsilon$ (écart-type des résidus) de 0.0201 pour ce modèle, et de 0.0216 et 0.0202 pour les modèle 1 et 3, respectivement.

En premier lieu, nous présentons dans le tableau 5.2 le résultats des estimations des paramètres de la loi binomiale négative pour chaque distribution marginale du nombre d'arrivées par période de temps, ainsi que le nombre moyen d'appels reçus à chaque période. Dans le tableau 5.3, nous présentons les résultats de la simulation des arrivées par intervalle de temps et par type de modèle. Nous remarquons que le processus NORTA, ainsi que sa version réduite donne des résultats très proches de ce qui est observé. Rappelons que dans ce dernier modèle les corrélations de rang échantillonales satisfont la relation (3.48). Dans les graphique 5.6 et 5.7, nous pouvons aussi remarquer l'avantage du processus de NORTA par rapport aux autres modèles. De point de vue des corrélations, les modèles basés sur le processus de Poisson doublement stochastique avec une spline ou une fonction de

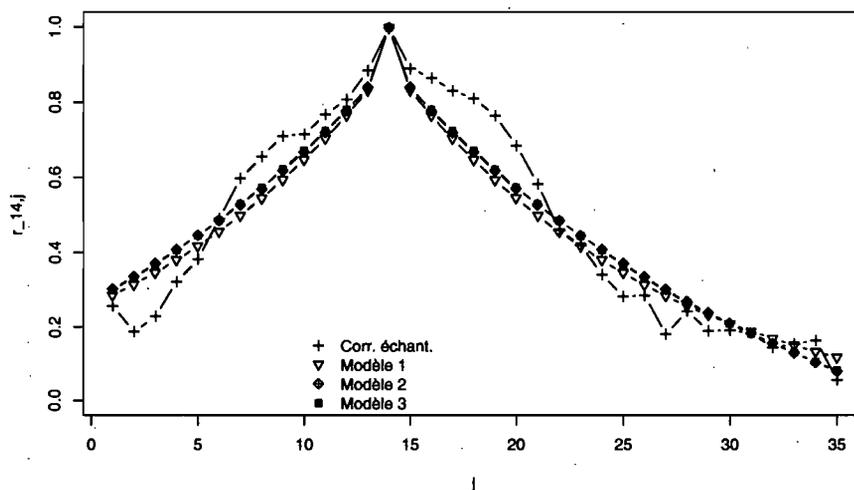


Figure 5.5 – Modèles pour les corrélations $r_{i,j}^X$ dans le centre C3, (graphique pour $i = 14$ et $j = 1, \dots, 35$).

taux constante par morceau pour le taux, tendent à surestimer les corrélations théoriques. Les résultats sont meilleurs avec le modèle NORTA, mais un peu moins bons avec le modèle NORTA réduit. Même constatation de point de vue coefficient de variation, les modèles NORTA et NORTA réduit donnent des résultats très proches de celles obtenus avec les données, alors que les autres modèles tendent à sous-estimer ces coefficients.

De façon générale, le processus de Poisson non homogène donne des résultats moins satisfaisants, du moment qu'il ne modélise pas la dépendance entre les différentes parties de la journée. Le processus de Poisson doublement stochastique donne de meilleurs résultats, mais une fois de plus nous remarquons que ce modèle introduit des corrélations positives entre les arrivées de toutes les périodes de temps dans la journée, alors qu'en réalité cette corrélation s'approche de zéro entre des intervalles de temps éloignés dans la journée. Ceci explique l'écart sur le graphique

Période	$\hat{\mu}_i$	\hat{s}_i	\hat{p}_i
7h00-7h30	615.7±15.8	45.7	0.069
7h30-8h00	713.4±14.6	70.1	0.090
8h00-8h30	1292.1±20.1	122.5	0.087
8h30-9h00	1863.5±24.4	169.7	0.083
9h00-9h30	2847.7±33.4	207.2	0.068
9h30-10h00	3275.9±40.3	192.8	0.056
10h00-10h30	3357.0±38.5	223.2	0.062
10h30-11h00	3376.0±39.3	219.9	0.061
11h00-11h30	3311.4±38.2	222.6	0.063
11h30-12h00	3197.6±37.0	224.0	0.065
12h00-12h30	3039.2±33.6	244.0	0.074
12h30-13h00	2968.8±35.1	212.9	0.067
13h00-13h30	2949.8±34.4	219.4	0.069
13h30-14h00	2954.2±32.5	248.6	0.078
14h00-14h30	2957.8±33.0	239.4	0.075
14h30-15h00	2960.2±32.6	244.2	0.076
15h00-15h30	2879.7±32.2	234.7	0.075
15h30-16h00	2817.6±32.0	225.2	0.074
16h00-16h30	2539.5±29.4	216.2	0.078
16h30-17h00	2174.8±28.9	159.2	0.068
17h00-17h30	1657.1±24.9	124.5	0.070
17h30-18h00	1382.9±19.5	146.9	0.096
18h00-18h30	1158.5±17.5	128.9	0.100
18h30-19h00	1001.1±16.8	102.7	0.093
19h00-19h30	873.8±15.7	91.4	0.095
19h30-20h00	783.3±16.5	64.6	0.076
20h00-20h30	621.9±13.9	58.9	0.086
20h30-21h00	560.2±16.2	37.7	0.063
21h00-21h30	507.0±13.9	39.0	0.071
21h30-22h00	442.5±11.7	41.3	0.085
22h00-22h30	386.6±10.7	38.6	0.091
22h30-23h00	323.8±9.0	39.0	0.107
23h00-23h30	254.7±6.6	48.2	0.159
23h30-24h00	196.3±4.7	61.9	0.240

Tableau 5.2 – Estimation du taux moyen d'arrivée et des paramètres de la loi binomiale négative pour chaque intervalle de demi-heure avec les données dans le centre C3

Période	Données	PPNH (spline)	PPDS (constante)	PPDS (spline)	NORTA	NORTA (réduit)
7h00-7h30	615.54 (7.91)	632.77±0.26	615.90±0.84	632.88±0.46	616.17±1.85	614.10±1.84
7h30-8h00	713.71 (7.34)	750.04±0.28	713.49±0.96	750.46±0.53	713.77±1.75	711.61±1.75
8h00-8h30	1292.23 (10.27)	1277.54±0.37	1291.72±1.62	1278.38±0.85	1292.71±2.41	1290.24±2.40
8h30-9h00	1862.56 (12.50)	1896.30±0.46	1863.51±2.30	1896.92±1.21	1864.57±2.95	1860.94±2.94
9h00-9h30	2848.55 (18.30)	2812.32±0.55	2846.97±3.38	2814.11±1.76	2850.99±4.04	2845.16±4.00
9h30-10h00	3275.19 (19.06)	3264.07±0.60	3277.03±3.89	3264.75±2.03	3280.96±4.79	3273.07±4.75
10h00-10h30	3355.41 (18.30)	3355.35±0.60	3357.79±3.98	3357.54±2.08	3362.08±4.56	3354.89±4.54
10h30-11h00	3376.32 (18.77)	3371.55±0.60	3375.41±3.99	3372.38±2.08	3380.51±4.64	3373.15±4.59
11h00-11h30	3311.58 (18.63)	3309.99±0.60	3311.72±3.92	3310.95±2.06	3315.88±4.53	3308.44±4.45
11h30-12h00	3196.71 (18.73)	3194.33±0.59	3197.89±3.79	3195.78±1.98	3201.60±4.39	3194.60±4.33
12h00-12h30	3038.29 (16.82)	3044.91±0.57	3039.14±3.60	3045.96±1.89	3043.34±3.98	3034.53±3.96
12h30-13h00	2969.13 (19.14)	2970.96±0.56	2969.88±3.55	2971.08±1.85	2972.86±4.18	2965.83±4.15
13h00-13h30	2949.62 (17.40)	2951.12±0.56	2950.23±3.51	2951.84±1.84	2954.28±4.11	2947.64±4.05
13h30-14h00	2954.38 (16.46)	2953.98±0.56	2955.03±3.51	2955.09±1.84	2957.88±3.88	2950.70±3.82
14h00-14h30	2956.94 (17.46)	2959.30±0.56	2958.42±3.51	2960.23±1.84	2961.57±3.94	2954.21±3.88
14h30-15h00	2960.56 (15.63)	2954.76±0.57	2960.38±3.52	2955.31±1.84	2963.90±3.93	2956.94±3.85
15h00-15h30	2879.12 (16.79)	2883.76±0.56	2880.05±3.43	2885.25±1.80	2883.74±3.88	2876.62±3.83
15h30-16h00	2818.38 (17.20)	2804.61±0.55	2818.06±3.35	2805.00±1.75	2821.00±3.87	2813.89±3.83
16h00-16h30	2538.79 (15.12)	2539.73±0.52	2539.73±3.07	2540.43±1.60	2540.38±3.56	2536.91±3.53
16h30-17h00	2174.84 (14.57)	2162.82±0.48	2175.00±2.64	2163.90±1.37	2174.83±3.53	2173.13±3.51
17h00-17h30	1656.85 (12.25)	1673.19±0.42	1657.37±2.03	1673.80±1.08	1656.69±3.00	1655.15±3.03
17h30-18h00	1382.73 (9.88)	1381.46±0.39	1383.38±1.73	1382.17±0.91	1382.32±2.35	1382.06±2.37
18h00-18h30	1158.04 (7.86)	1162.34±0.36	1158.36±1.46	1163.04±0.78	1157.53±2.13	1157.24±2.12
18h30-19h00	1001.17 (8.15)	1001.16±0.33	1000.14±1.29	1001.26±0.68	1000.33±2.05	999.34±2.03
19h00-19h30	873.99 (7.28)	877.33±0.31	873.34±1.15	877.52±0.60	873.71±1.88	872.84±1.88
19h30-20h00	783.19 (7.68)	777.40±0.29	783.30±1.05	777.82±0.54	777.82±1.88	772.84±1.88
20h00-20h30	621.82 (6.73)	628.82±0.26	621.88±0.86	628.80±0.45	621.98±1.67	621.82±1.66
20h30-21h00	559.88 (8.41)	559.31±0.25	560.14±0.79	559.75±0.41	560.44±1.86	560.31±1.84
21h00-21h30	506.95 (7.41)	506.38±0.23	506.90±0.73	506.63±0.38	507.33±1.65	506.21±1.64
21h30-22h00	442.33 (5.99)	443.16±0.22	442.65±0.64	443.33±0.34	442.78±1.41	442.40±1.41
22h00-22h30	386.58 (5.52)	386.27±0.21	386.74±0.58	386.13±0.31	386.61±1.29	386.37±1.28
22h30-23h00	323.53 (4.34)	323.43±0.19	323.84±0.51	323.63±0.27	323.88±1.09	323.47±1.08
23h00-23h30	254.93 (3.65)	255.47±0.17	254.68±0.43	255.38±0.22	254.89±0.79	254.82±0.79
23h30-24h00	196.26 (2.50)	196.33±0.16	196.19±0.35	196.38±0.19	196.40±0.56	196.16±0.57

Tableau 5.3 – Le nombre moyen d’arrivées par intervalle de temps pour une journée dans le centre C3. Les données empiriques sont comparées avec des données simulées (35750 jours) selon différents modèles.

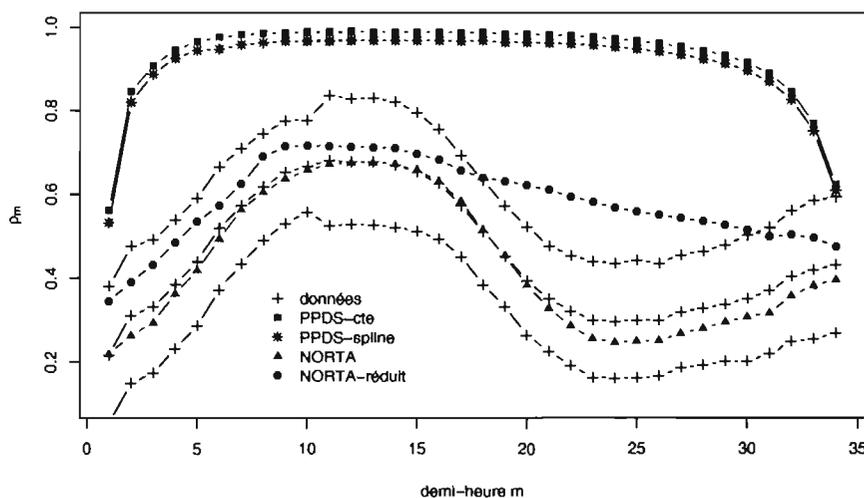


Figure 5.6 – Comparaison des modèles selon les corrélations estimées entre les différentes parties de la journée dans le centre C3.

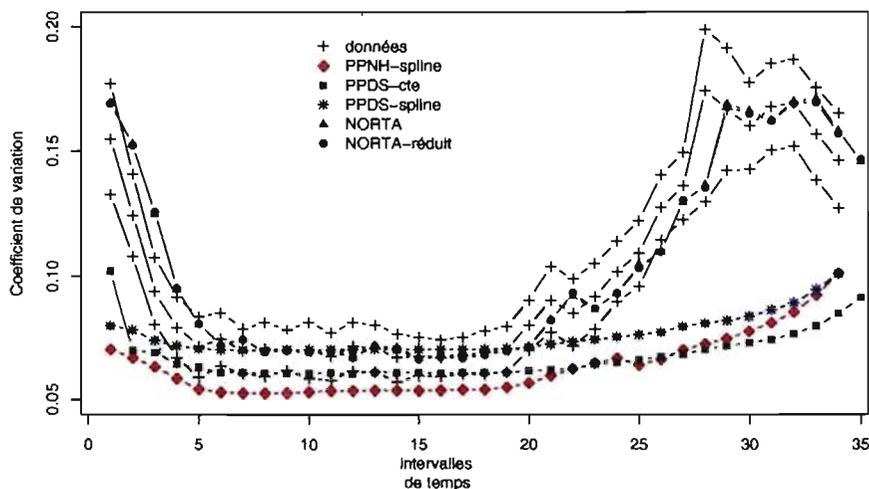


Figure 5.7 – Comparaison des modèles selon le coefficient de variation pour chaque intervalle de temps dans le centre C3.

5.6. Rappelons que les corrélations ρ_m , $m = 1, \dots, 33$, dans ce graphique sont calculées avec la même formule donnée en (4.2), avec 34 périodes pour le centre C3. Nous remarquons que c'est le modèle NORTA qui affiche sa supériorité et qui permet de donner des résultats très proche de la réalité. Le modèle NORTA réduit semble être moins bon de point de vue des corrélations. Mais nous remarquons qu'il fait quand même mieux que le processus de Poisson doublement stochastique. Ce modèle se présente comme un bon compromis entre les deux autres modèles, puisqu'il permet de réduire le nombre de paramètres à estimer tout en respectant la dépendance, mais en s'éloignant un peu de la structure de corrélation initiale.

CHAPITRE 6

MODÈLES DE SÉRIES CHRONOLOGIQUES POUR LES ARRIVÉES DANS UN CENTRE D'APPELS

6.1 Modèles pour les appels au service du 911

6.1.1 Modèles de planification pour les systèmes médicaux d'urgence dans la littérature

Les modèles que nous présentons dans ce chapitre sont motivés par la disponibilité des données observées sur les appels téléphoniques au service 911 de la ville de Calgary, en Alberta.

Depuis les années 70, plusieurs études ont été réalisées sur des modèles de planification pour les Systèmes Médicaux d'Urgence (SMU), aussi bien que le service d'incendies et le service de police. Green et Kolesar [34] ont récemment fourni une perspective sur l'impact de ces modèles de planification sur l'aide aux décisions et sur le bon fonctionnement de ces systèmes. Ces mêmes auteurs ont aussi fourni une description des modèles proposés dans la littérature sur ce sujet. Swersey [87] a présenté une revue de la littérature théorique sur ce thème et a étudié le problème d'optimisation pour la localisation des services d'urgence, alors que Goldberg [33] a présenté une revue détaillée plutôt sur le côté pratique de ce sujet et les travaux réalisés pendant une période de 35 ans.

Des exemples récents sur des modèles de planification et des modèles de simulation ont été présentés par Henderson et Mason [40], par exemple. Leur modèles d'optimisation concernent l'affectation des ambulances (nombre nécessaire, emplacements...) dans la ville d'Auckland, en Nouvelle-Zélande. Dans le même contexte, Ingolfsson et al. [43] ont étudié l'impact de l'utilisation d'une station unique, pour

les départs et les arrivées des ambulances qui ont la même période de travail, sur l'optimisation du service des ambulances dans la ville d'Edmonton, en Alberta. Il y a eu aussi des modèles analytiques de files d'attente, notamment le modèle de files d'attente hypercube dans le contexte d'affectation et de déploiement des ambulances [53, 54], ainsi que des modèles d'optimisation pour la localisation des unités de soins.

En général, les modèles de files d'attente supposent que les arrivées des appels suivent un processus de Poisson, une hypothèse qui est soutenue par des arguments aussi bien théoriques [39] qu'empiriques [37, 95]. Des études empiriques tant pour un système médical d'urgence (SMU) que pour d'autres services montrent qu'un taux d'arrivée variable dans le temps, ou même stochastique, d'un processus de Poisson décrit bien la réalité dans les centres d'appels [4, 13]. Si le taux d'arrivée varie beaucoup pendant une heure donnée, il est donc inapproprié de l'approximer par un taux constant et il est plus prudent de considérer des périodes de temps plus courtes, par exemple des demi-heures. Ainsi, ce modèle basé sur le processus de Poisson non homogène devient plus plausible dans des études analytiques ou des simulations. Pour illustration, supposons que les appels arrivent pendant l'heure h suivant un processus de Poisson avec un taux qui reste constant pendant l'heure. Alors, conditionnellement au nombre d'appels reçus pendant toute l'heure, les instants d'arrivées individuels sont indépendants et uniformément distribués [75, Sections 4.5-4.6]. Cette propriété est aussi valable pour un processus de Poisson avec un taux d'arrivée stochastique.

Dans notre travail, nous présentons des modèles pour le taux d'arrivée pendant une partie ou durant les heures de la journée. Nous fournissons les détails dans les sections 6.1.4 et 6.1.5 et nous montrons l'utilité de ces modèles dans toute simulation et étude analytique où les arrivées suivent un processus de Poisson.

Goldberg [33] a montré la grande utilité des prévisions des arrivées des appels dans les centres, et a déploré le peu d'études systématiques réalisés dans ce sens. Dans son travail, il a classifié les modèles utilisés dans littérature en deux catégories : (1) des modèles pour la distribution du volume d'appels en fonction de variables démographiques et (2) des modèles pour le volume d'appels évoluant dans le temps.

Dans la première catégorie, Kamenetsky et al. [47] ont présenté la littérature avant 1982 et ont proposé des modèles de régression pour les prévisions du volume d'appels au SMU en fonction de la population, de l'emploi et deux autres variables démographiques. Dans leur modèle les variables utilisées ont expliqué la plupart de la variation ($R^2 = 92\%$) et les données ont couvert 200 unités de soins le sud-ouest de la Pennsylvanie, aux États-Unis. Dans un article plus récent et dans la même catégorie, McConnell et Wilson [65] ont mis le point sur une autre variable importante dans l'explication du volume d'appels. Il s'agit de la distribution par âge de la population.

Notre travail tombe dans la deuxième catégorie, du fait que nous présentons des modèles de séries chronologiques pour les arrivées journalières et par heure. Ces modèles permettront de faire des prévisions dans les centres d'appels au service 911 dans la ville de Calgary, en Alberta. En général, le volume d'appels qui arrivent à ce type de centre varie beaucoup entre les jours de semaine et les jours de la fin de semaine [37, 95]. Parmi les travaux réalisés dans le passé sur les prévisions des appels au service SMU, il y a celui de Mabert [62] qui a analysé les arrivées des appels d'urgence au département de police de la ville d'Indianapolis, en Indiana, aux États-Unis. Il a considéré plusieurs méthodes simples basées sur les modèles saisonniers et a constaté qu'un modèle ARIMA simple [9] donnait de meilleurs résultats dans son cas. Baker et Fitzpatrick [6] ont utilisé le modèle de lissage

exponentiel de "Winter" pour les prévisions des appels journaliers reçus dans le centre d'appels d'urgence au SMU.

Pour les prévisions du volume d'appels journaliers qui arrivent au centre d'appel spécialisé dans la télé-vente de détails, Andrews et Cunningham [3] ont utilisé un modèle ARIMA avec une fonction de transfert pour les journées spéciales ; ils ont utilisé des variables indicatrices pour les jours où les catalogues sont envoyés aux clients actuels ou potentiels. Bianchi et al. [8] ont aussi utilisé un modèle ARIMA pour les prévisions des appels reçus quotidiennement dans un centre de télé-vente. Ils ont comparé ce modèle avec celui de "Holt-Winters" et ont aussi montré l'avantage de prendre en considération les journées spéciales. Tych et al. [89] ont présenté des modèles de prévision pour les arrivées par heure dans un centre d'appels d'une banque en utilisant un modèle relativement complexe avec les composantes non observées appelé *la régression harmonique dynamique*, qui a donné de meilleurs résultats qu'un modèle ARIMA saisonnier dans leur cas. Brown et al. [13] ont développé une méthode pour les prévisions du taux d'arrivée dans des périodes de temps au cours d'une journée, notamment avec un modèle autorégressif basé sur le volume d'appels du jour précédent.

Tout récemment Shen et Huang [80] ont présenté des modèles de prévision pour le volume d'appels par jour et par intervalles de temps dans une même journée. Ils ont considéré un vecteur qui représente le profil du jour pour les arrivées par intervalles de temps et l'ont traité comme une série chronologique multivariée. Ils ont procédé à la réduction de la dimension de ce vecteur en utilisant la méthode de décomposition en valeurs singulières de la matrice qui représente les données observées dans un centre d'appels. Par la suite ils ont utilisé un modèle de série chronologique univariée pour le nombre d'arrivées par jour pour ensuite combiner les deux modèles afin d'obtenir des prévisions des arrivées à l'intérieur d'une journée

en supposant que le profil du jour ne varie pas à un horizon proche. Pour mettre à jour les prévisions dans le cas où une partie de la journée est déjà observée, ils ont proposé une méthode basée sur les moindres carrées pénalisées qui combine les l'information observée avec les prévisions faites sur la même période.

Weinberg et al. [92] ont présenté un modèle multiplicatif pour les prévisions du taux d'arrivée dans un centre d'appels téléphoniques dans le cas où ces arrivées suivent un processus de Poisson non homogène. Leur modèle est en fait basé sur celui de Brown et al. [13]. L'estimation des paramètres est faite sur la base de données observées dans un centre d'appels d'une banque commerciale américaine et elle est faite à l'aide des méthodes Monte Carlo par chaînes de Markov. Ces méthodes permettent d'obtenir, par simulation, un échantillon d'une loi quelconque sans simuler directement celle-ci, mais plutôt une chaîne de Markov. Ils ont montré que leur modèle donne de meilleurs résultats pour le cas du centre étudié que des modèles additifs qui tiennent compte de l'heure de la journée, du jour de la semaine et un effet croisé.

Les normes de niveau de service pour les systèmes SMU sont des mandataires incomplets comparés au rôle très important que jouent tels systèmes, à savoir sauver des vies et prévenir la souffrance, voir, par exemple, Erdogan et al. [29] pour une discussion de modèles qui essaient de quantifier de tels buts plus explicitement. La satisfaction de ces normes de niveau de service n'est certes pas une tâche facile, mais c'est un problème d'intérêt économique et social substantiel de diriger des systèmes SMU efficacement. L'incertitude et l'environnement stochastique pour le volume d'appels complique le processus, surtout pour déterminer les niveaux de recrutement du personnel SMU et de l'équipement. Il est donc très important de modéliser correctement la nature stochastique du volume d'appels et particulièrement faire des prévisions dans le futur. Ceci explique largement la motivation pour construire

des modèles de prévisions pour les appels à court, à moyen et à long terme de façon la plus efficace possible.

Parmi les mesures de performance pour un système SMU et d'une manière caractéristique, nous pouvons citer la fraction d'appels qui sont satisfaits au bout d'un temps bien défini, à partir de l'instant que l'appel a été fait. Bien qu'il n'existe pas des normes universelles [67], on considère que le temps de réponse commence quand l'évaluation d'appel débute et il se termine lorsque ambulance atteint l'adresse de l'appel. On peut aussi citer d'autres mesures de performance comme le temps d'attente au téléphone avant qu'un opérateur du 911 réponde ; par exemple, 90% dans 10 secondes ou moins, ou 95% dans 5 secondes ou moins [68]. On peut aussi regarder le temps moyen passé dans l'évaluation de l'appel, le temps moyen mis avant d'être dans le lieu, ainsi que le temps moyen passé à l'hôpital.

Les décisions principales dans la gestion du personnel et des ambulances reposent essentiellement sur les modèles de prévision du volume d'appels à court et à moyen terme (quelques jours à quelques semaines dans l'avenir). Les prévisions à long terme sont nécessaires pour le plan stratégique d'expansion du système ou de réorganisation. Pour le terme plus court (l'intérieur d'une journée), les prévisions sont généralement utiles pour des décisions immédiates comme les réunions ou la réorganisation des pauses-café des agents, ainsi que pour l'appel de ressources supplémentaires.

6.1.2 Description du centre d'appels étudié C4

La majorité des grandes villes dans les pays industrialisés disposent d'un système de services médicaux d'urgence (SMU) qui fournissent les soins médicaux primaires à des personnes malades ou blessées, et assurent le transport de ces personnes vers des établissements de soins médicaux. La demande pour de tels services augmente

constamment dans ces pays à cause de l'accroissement et surtout le vieillissement de la population.

Presque toute la demande aux systèmes SMU arrive par téléphone en composant un numéro d'urgence (911 en Amérique du Nord). Les appels qui arrivent à ce numéro sont au départ acheminés à trois services différents, soit le service SMU, le service des incendies ou le service de police dépendamment de l'intervention demandée. Nous nous concentrons seulement sur les appels acheminés au service SMU et nous étudierons spécifiquement le centre d'appels au 911 dans la ville de Calgary, en Alberta. Les données observées dont nous disposons concernent ce centre pour une période s'étalant sur 4 années.

Chaque appel qui arrive au centre d'appels 911 est directement pris en charge par un agent qui évalue la nature et l'importance de l'incident. Au cours de cette première étape, l'agent peut aussi fournir des instructions à une personne présente sur la scène sur les procédures des premiers soins en attendant l'arrivée d'une ambulance. Une fois que l'adresse est connue, ainsi que la nature et l'importance de l'incident, une ambulance est dépêchée sur les lieux. Cette étape est généralement débutée au cours de la première, c.-à.-d. qu'en même temps l'agent traite l'appel, l'ambulance et l'équipage du véhicule sont en route vers la scène de l'appel, où ils évaluent la situation et fournissent des soins médicaux sur place et déterminent par la suite si le transport à un établissements de soins médicaux est nécessaire ou non. En général, près de 75% des cas nécessitent le transport à un centre médical. Le personnel accompagne le patient jusqu'à ce que la responsabilité soit transférée à un infirmier ou à un médecin. Par la suite, l'équipage peut alors finir les formalités nécessaires avant d'être disponible pour répondre à d'autres demandes.

En général, la première étape nécessite seulement quelques minutes pour l'évaluation de l'appel et la dépêche de l'ambulance. Pour la deuxième étape, on compte un

temps de l'ordre d'une heure pour l'intervention de l'ambulance et de son équipage sur la scène. Cependant, la durée de cette étape est en train d'augmenter dans beaucoup d'endroits à cause des temps d'attente élevés qui sont enregistrés dans les urgences des hôpitaux [27, 28, 79, 83].

Dans notre cas, nous présentons des modèles pour le volume d'appels journaliers et par heure au SMU et nous validons ces modèles en utilisant des données observées dans le centre d'appels 911 de la ville de Calgary, en Alberta. Bien que nous nous concentrons sur ces données bien spécifiques, nos modèles pourraient être utiles et valables pour tout autre SMU dans d'autres villes. Nous voulions que ces modèles soient plutôt simples et efficaces et puissent être utilisés pour la simulation du système du centre d'appels.

Les séries de données observées et analysées ont révélé une tendance à la hausse dans le temps, des cycles saisonniers de jour, de semaine et d'année, avec un effet de journées spéciales et une structure d'autocorrélation positive. Nous utilisons deux approches : (1) les modèles autorégressifs après avoir éliminé la tendance et toutes les saisonnalités et l'effet des journées spéciales; et (2) ARIMA doublement saisonnier avec les effets des journées spéciales. Nous comparons les modèles du point de vue de performance des estimations et des prévisions. Nous présentons aussi deux modèles pour les arrivées par heure : (3) une distribution multinomiale pour le vecteur du nombre d'appels à chaque heure conditionnellement au volume total d'appels pendant le jour et (4) un modèle de séries chronologiques. Dans notre cas et pour nos données, (1) et (3) semblent être meilleurs.

6.1.3 Analyse descriptive des données

Nous disposons de données de 50 mois (de 2000 à 2004) qui couvrent les appels au SMU dans la ville de Calgary, en Alberta. L'analyse préliminaire de ces données

révèle une tendance à la hausse, un cycle saisonnier chaque jour, un cycle saisonnier chaque semaine, un cycle saisonnier chaque année, les effets de journées spéciales et également une structure de corrélation. En vue de cela, nous considérons les deux approches principales suivantes : (1) des modèles autorégressifs pour les résidus d'un modèle avec une tendance positive, les saisonnalités et les effets des journées spéciales ; et (2) un modèle ARIMA doublement saisonnier pour le résidu d'un modèle qui capte seulement des effets de journées spéciales. Dans l'approche (1), nous explorons des modèles dont les effets sont le jour de la semaine et le mois de l'année. Nous considérons aussi un modèle avec les effets croisés (les termes d'interaction) et un modèle plus parcimonieux pour réduire le nombre de paramètres, avec aussi les effets croisés, mais seulement les effets statistiquement significatifs sont retenus. Nous verrons plus loin que ce dernier se révèle être le meilleur tant du point de vue de performance des prévisions que du point de vue estimation. Tous les modèles sont estimés avec les 36 premiers mois de données et l'erreur de prévisions est mesurée avec les données à partir des 14 derniers mois.

Nous disposons des données sur les arrivées entre le premier janvier 2000 et le 16 mars 2004. Chaque appel est enregistré par son instant d'arrivée, et il est classé selon sa priorité et la zone géographique d'où il est composé. Pour nos modèles, nous procédons avec le nombre d'appels par heure et par jour pour l'application de modèles de séries chronologiques. Le nombre moyen d'arrivées est environ 174 par jour, et environ 7 par heure. La figure 6.1 montre le volume d'appels journaliers pour l'année 2000, où on peut observer une tendance à la hausse. Ceci peut être expliqué par une croissance démographique et un vieillissement de la population dans la ville en question. Des volumes élevés sont enregistrés en juillet et en décembre, par exemple, le 1 janvier, le 8 juillet, le 11 novembre, le 1 décembre ; et des volumes bas, comme par exemple le 26 janvier, le 9 septembre. Le graphique 6.2, montre le

volume mensuel pendant la période entière, qui révèle aussi la tendance à la hausse. Le graphique 6.3 affiche le volume moyen des appels par heure pour chaque jour de la semaine. Il révèle le cycle saisonnier de 24 heures qu'on qualifie d'effet de l'heure de la journée. On remarque aussi que le niveau le plus haut du volume d'appels s'affiche entre 10h00 et 20h00; les volumes les plus bas sont observés au cours de la nuit. On observe aussi l'effet du jour de la semaine, où l'activité augmente vendredi et samedi soir et au début de la nuit. Pour le volume quotidien, les plus grandes valeurs sont observées vendredi et samedi par rapport aux autres jours de la semaine.

Les graphiques 6.4 et 6.5 montrent le diagramme en boîtes du volume d'appels journaliers par type de jour de la semaine, et celui du volume mensuel pour chaque mois de l'année, respectivement. Nous remarquons clairement que vendredi et samedi affichent plus d'appels que la moyenne. Juillet, décembre et novembre sont les mois les plus importants en volume, alors que le mois d'avril affiche moins de volumes d'appels.

6.1.4 Modèles pour les arrivées par jour

Nous considérons maintenant cinq modèles différents de séries chronologiques pour le volume d'appels reçus au cours des jours successifs. Bien qu'à la fin nous concluons qu'un de ces modèles s'ajuste mieux aux données de Calgary, nous analysons tous les modèles que nous présentons, car ils peuvent être appropriés selon la ville étudiée et le but de l'analyse. Dans la Section 6.1.4.6, nous comparons ces modèles du point de vue de la qualité de l'estimation et des performances des prévisions.

Dans la suite du travail, t désigne l'unité de temps utilisée pour désigner le jour et le nombre d'arrivées au jour t est noté Y_t , pour $t = 1, 2, \dots, n$, où $n = 1537$

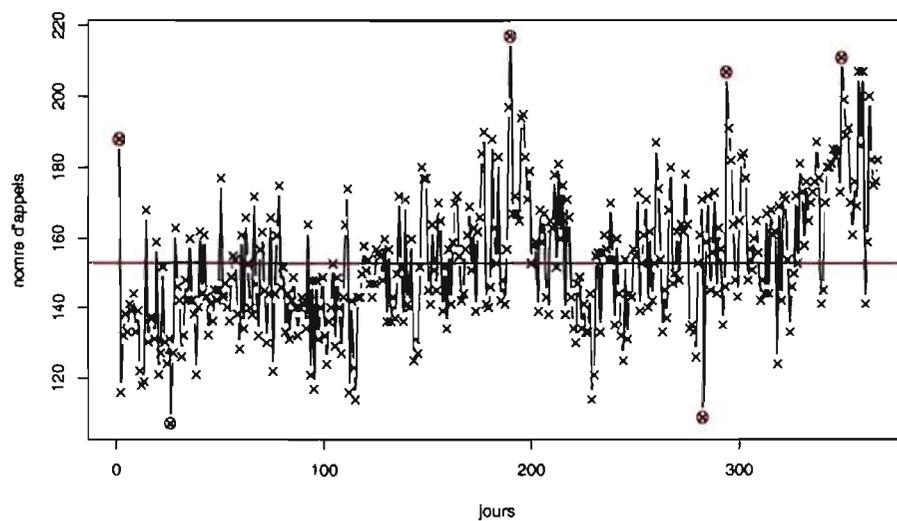


Figure 6.1 – Volume d'appels par jours pour l'année 2000 dans le centre C4, où quelques journées spéciales apparaissent.

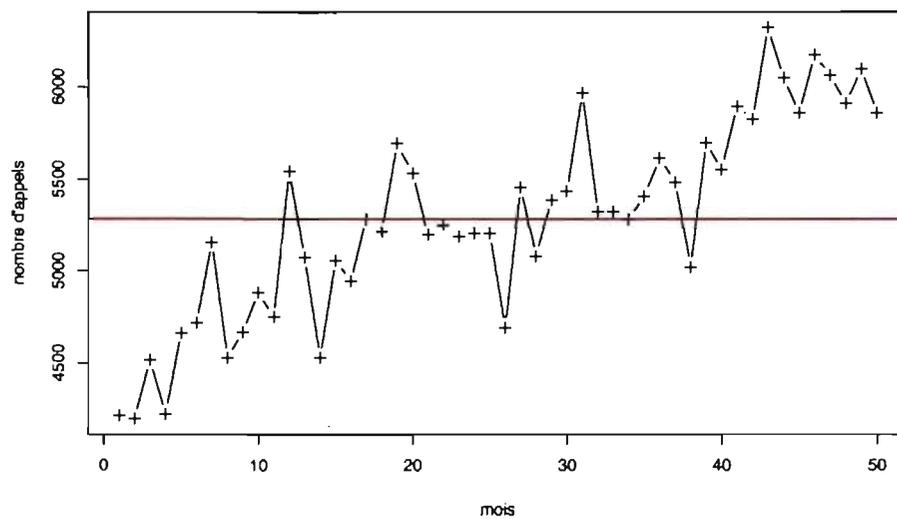


Figure 6.2 – Volume d'appels par mois pour toute la période dans le centre C4.

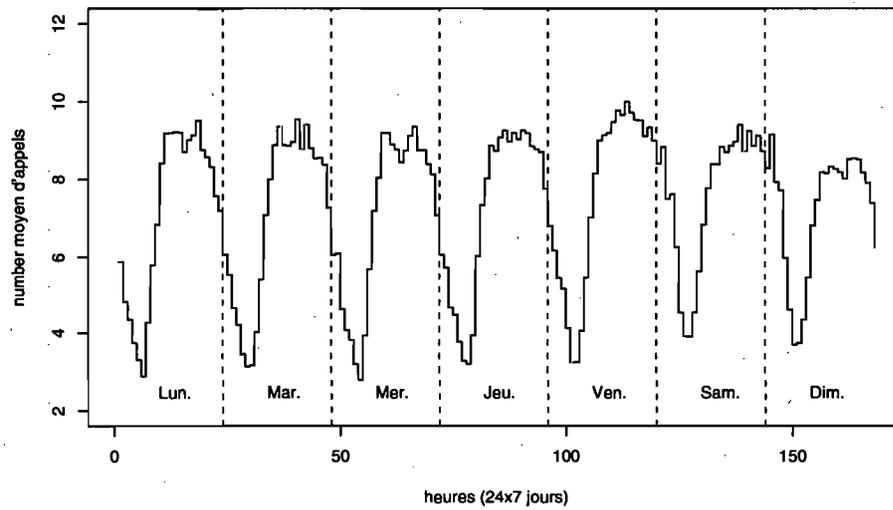


Figure 6.3 – Volume d'appels moyen par heure pour un cycle d'une semaine dans le centre C4.

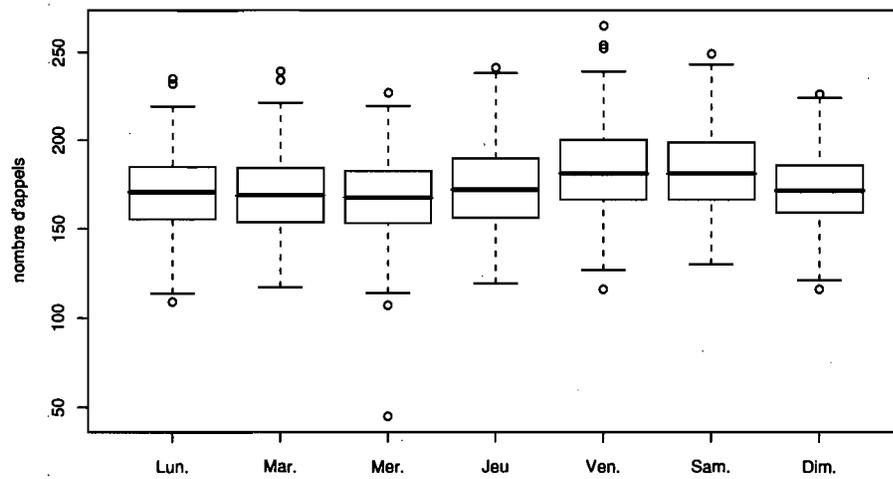


Figure 6.4 – Diagrammes en boîtes pour le volume journalier pour chaque type de jour de la semaine dans le centre C4.

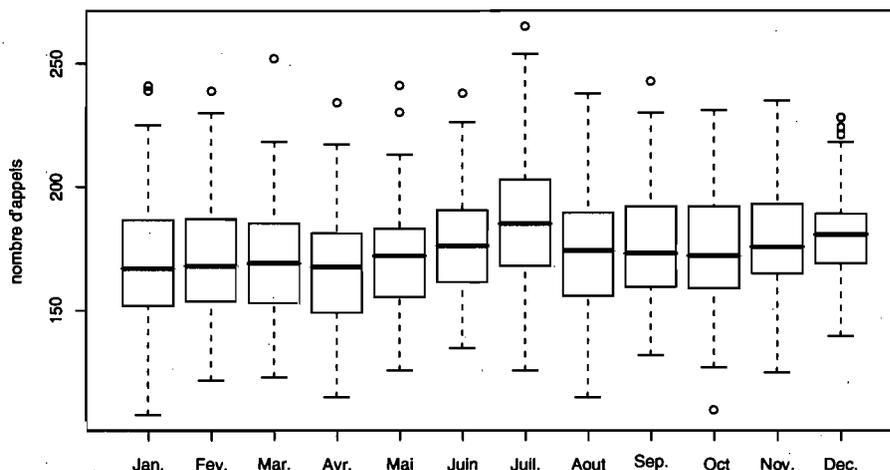


Figure 6.5 – Diagrammes en boîtes pour le volume par type de mois de l'année dans le centre C4.

représente le nombre de jours observés dans le centre. La phase d'estimation est faite sur la base des 1096 premières observations (du premier janvier 2000 au 31 décembre 2002) et les 441 observations (du premier janvier 2003 au 16 mars 2004) sont utilisées pour la validation.

Dans nos modèles pour le nombre d'arrivées par jour Y_t et par heure Z_h , on peut utiliser la procédure suivante pour simuler les arrivées au jour t :

1. Simuler le volume journalier Y_t et nous verrons dans cette section que cela implique de simuler des résidus selon un processus autorégressif standard.
2. Connaissant le volume journalier Y_t , nous pouvons facilement générer un vecteur $Z_t = (Z_{24(t-1)+1}, \dots, Z_{24t})$, pour le nombre d'arrivées par heure et nous verrons plus loin dans ce chapitre que cela implique de simuler un vecteur selon une distribution multinomiale.
3. Ayant obtenu le nombre généré Z_h des appels reçus pendant une heure

donnée, nous pouvons utiliser la propriété des statistiques d'ordre pour simuler les instants d'arrivées à l'intérieur de l'heure donnée.

6.1.4.1 Modèle 1 : effets fixes et résidus indépendants

Le premier modèle simple qu'on a proposé ici tient compte des effets du jour de la semaine et du mois de l'année pour le volume d'appels par jour au SMU. Ceci peut être valable pour la plupart des grandes villes. L'analyse préliminaire des données confirme la présence de ces effets et indique aussi une tendance positive. Cela suggère le modèle linéaire suivant comme une première approximation :

$$Y_{j,k,l} = a + \tilde{\beta}_j + \tilde{\gamma}_k + \tilde{\alpha}_l + \tilde{\epsilon}_{j,k,l}, \quad (6.1)$$

où $Y_{j,k,l}$ est le nombre d'appels par jour du type j pendant le mois k de l'année l , les paramètres a , $\tilde{\beta}_j$, $\tilde{\gamma}_k$, et $\tilde{\alpha}_l$, sont des constantes réelles, et les résidus $\tilde{\epsilon}_{j,k,l}$ sont i.i.d. selon la loi normale de moyenne 0. Les premières analyses montrent une tendance linéaire de l'effet de l'année, en fonction de l , et qui nous permet de réécrire le modèle sous la forme d'une série temporelle, comme suit :

$$Y_t = a + bt + \sum_{j=1}^7 \beta_j C_{t,j} + \sum_{k=1}^{12} \gamma_k S_{t,k} + E_t, \quad (6.2)$$

où a , b , les β_j , et les γ_k sont des constantes réelles, l'indicateur $C_{t,j}$ vaut 1 si l'observation t est dans le jour j de la semaine et 0 sinon ; l'indicateur $S_{t,k}$ vaut 1 si l'observation t est dans le mois k de l'année et 0 sinon. Dans d'autres villes et avec d'autres données, il peut être plus approprié de modéliser avec un effet non linéaire pour l'année en fonction de t . Nous supposons que les résidus E_t sont i.i.d. de distribution normale de moyenne 0 et de variance σ_E^2 , c.-à-d., un bruit blanc gaussien. Avec la présence du paramètre a , nous imposons les contraintes

d'identifiabilité suivantes :

$$\sum_{j=1}^7 \beta_j = \sum_{k=1}^{12} \gamma_k = 0. \quad (6.3)$$

Ceci évitera la redondance dans le modèle; par exemple, en ajoutant une constante κ à tous les β_j 's en la soustrayant de a le modèle ne changera pas. Nous estimons alors le modèle de régression, donné en (6.2) par moindres carrés ordinaires et nous obtenons les résidus montrés dans la figure 6.6, où nous avons encerclé les points à une distance supérieure à $3\hat{\sigma}_E$, où $\hat{\sigma}_E^2$ représente la variance empirique des résidus. Nous notons la présence d'un seul point qui dépasse $4\hat{\sigma}_E$ et qui correspond premier janvier 2002, et 7 autres qui dépassent $3\hat{\sigma}_E$: le premier décembre 2000 ; le premier janvier 2001 ; le 27 mai 2001 ; le 2 août 2001 ; le 8 septembre 2001 ; le 27 juin 2002 ; le 12 juillet 2002. La valeur qui est inférieure à $-3\hat{\sigma}_E$ coïncide avec le 30 juillet 2001. Le premier janvier de chaque année parait comme une journée spéciale avec un volume supérieur à la moyenne. On constate aussi que le volume d'appels des jours du mois de juillet sont supérieurs à ceux des autres mois. Ceci, nous a motivé pour considérer l'effet du "Calgary Stampede", qui a lieu chaque année au mois de juillet. Il représente un des plus grands "rodéos" dans le monde et le plus grand festival de l'année dans la ville Calgary, en Alberta (<http://calgarystampede.com>). Pour la période qui couvre nos données, il a eu lieu entre le 7 et le 16 juillet 2000 ; du 6 au 15 juillet 2001 ; du 5 au 14 juillet 2002 ; et du 4 au 13 juillet 2003.

Pour évaluer les journées spéciales, nous ajoutons deux variables indicatrices $H_{t,1}$ et $H_{t,2}$, où $H_{t,1}$ vaut 1 si l'observation t coïncide avec le premier janvier et 0 sinon; alors que $H_{t,2}$ vaut 1 si l'observation t coïncide avec un des 40 jours du

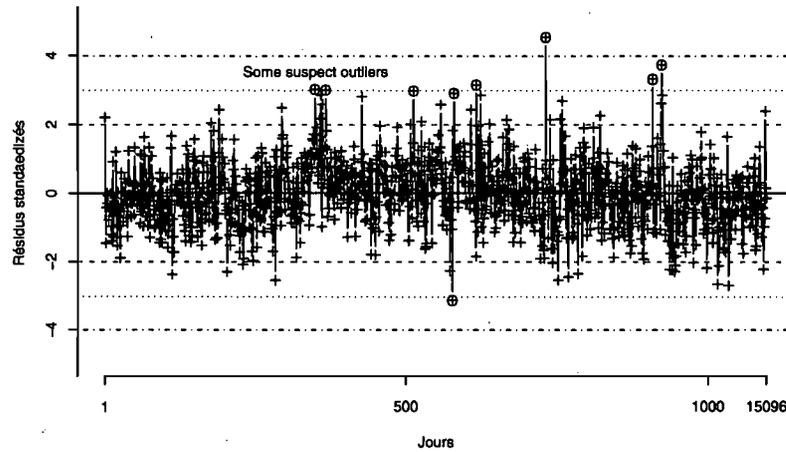


Figure 6.6 – Les résidus E_t du modèle linéaire simple donné dans l'équation (6.2).

“Stampede” mentionnés auparavant, et 0 sinon. Le modèle devient :

$$Y_t = a + bt + \sum_{j=1}^7 \beta_j C_{t,j} + \sum_{k=1}^{12} \gamma_k S_{t,k} + \omega_1 H_{t,1} + \omega_2 H_{t,2} + E_t, \quad (6.4)$$

qui contient deux nouveaux paramètres ω_1 and ω_2 . Nous procédons alors à l'estimation des paramètres du modèle de régression par moindres carrés ordinaires en utilisant les $n = 1096$ premières observations. Si nous dénotons les estimations de ces paramètres par \hat{a} , \hat{b} , $\hat{\beta}_j$, $\hat{\gamma}_k$, $\hat{\omega}_1$ et $\hat{\omega}_2$, alors les estimations de Y_t et E_t sont données par :

$$\hat{Y}_t = \hat{a} + \hat{b}t + \sum_{j=1}^7 \hat{\beta}_j C_{t,j} + \sum_{k=1}^{12} \hat{\gamma}_k S_{t,k} + \hat{\omega}_1 H_{t,1} + \hat{\omega}_2 H_{t,2} \quad (6.5)$$

et

$$\hat{E}_t = Y_t - \hat{Y}_t. \quad (6.6)$$

Un estimateur naïf de σ_E^2 serait la variance empirique, donnée par :

$$\hat{\sigma}_E^2 = \frac{1}{n-s} \sum_{t=1}^n \hat{E}_t^2, \quad (6.7)$$

où $s = 21$ représente le nombre de paramètres estimés dans le modèle. Il est évident que cet estimateur est biaisé si les résidus sont corrélés [7]. Pour cela, nous devons vérifier si les résidus forment un bruit blanc ou non, c.-à-d., non corrélés et distribués selon une loi normale standard de moyenne 0 et de variance constante. La stationnarité peut être observée dans la figure 6.6. Pour la corrélation entre les résidus, nous utilisons le test de "Ljung-Box test", où la statistique est donnée par :

$$Q = n(n+2) \sum_{i=1}^l \frac{\hat{r}_i^2}{n-i},$$

avec n le nombre de résidus, \hat{r}_i est l'autocorrélation échantillonnale d'ordre i et l est l'ordre maximum qu'on utilise pour tester les autocorrélations. Sous l'hypothèse nulle que les résidus sont non corrélés et $n \gg s$, Q est approximativement distribuée selon la loi chi-deux de l degrés de liberté. Nous avons $n = 1096$ et $s = 21$. Nous appliquons ce test avec $l = 30$ et nous obtenons $Q = 154.8$, avec une p-valeur inférieure à 2.2×10^{-16} , ce qui nous amène à rejeter l'hypothèse nulle. La présence de corrélation entre les résidus nous incite à considérer le modèle 2.

6.1.4.2 Modèle 2 : un modèle autorégressif pour les erreurs

Nous améliorons le modèle 1 en considérant un modèle de séries chronologiques pour les résidus E_t . Ce processus étant gaussien et stationnaire, et il suffit donc de capter la structure d'autocorrélation. Nous le faisons avec un processus autorégressif

d'ordre p , noté $AR(p)$, défini par

$$E_t = \phi_1 E_{t-1} + \dots + \phi_p E_{t-p} + a_t, \quad (6.8)$$

où a_t sont i.i.d. distribués selon une loi normale de moyenne 0 et de variance σ_a^2 . En se basant sur les résidus dans (6.6), et en utilisant les outils standards pour l'identification d'un modèle ARIMA adéquat [9, 91], nous obtenons un modèle avec $p = 3$. Il faut noter que cet ordre pourrait varier pour d'autres données et d'autres villes. Avec un modèle où $p > 3$, les coefficients ϕ_l pour $l > 3$ sont non significatifs à un niveau de 5%. Par exemple pour ϕ_4 , nous avons une p-valeur de 0.153.

Le modèle obtenu en combinant (6.4) et (6.8) avec $p = 3$ est donné par :

$$\phi(B) \left[Y_t - a - bt - \sum_{j=1}^7 \beta_j C_{t,j} - \sum_{k=1}^{12} \gamma_k S_{t,k} - \omega_1 H_{t,1} - \omega_2 H_{t,2} \right] = a_t, \quad (6.9)$$

avec $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3$, B est défini par $B^p E_t = E_{t-p}$, et ϕ_1, ϕ_2, ϕ_3 sont des paramètres autorégressifs.

Les paramètres $(a, b, \beta_1, \dots, \beta_7, \gamma_1, \dots, \gamma_{12}, \omega_1, \omega_2, \phi_1, \phi_2, \phi_3)$ sont estimés par moindres carrés non linéaires [1, page 67], en utilisant les observations Y_t pour $t = 4, \dots, n = 1096$. Les résultats des estimations sont données dans le tableau 6.1, avec l'écart-type et la p -valeur du test pour l'hypothèse que le paramètre est non significatif (égal à zéro). Les résidus du modèle sont alors $\hat{a}_t = \hat{\phi}(B)(Y_t - \hat{Y}_t)$ et leur variance σ_a^2 est estimée par

$$\hat{\sigma}_a^2 = \frac{1}{n-s} \sum_{t=4}^n \hat{a}_t^2, \quad (6.10)$$

avec $n = 1096$ et $s = 24$. Nous obtenons $\hat{\sigma}_a^2 = 250.1$. La figure 6.7 montre l'analyse

de la normalité des résidus obtenus : estimation de la densité et un diagramme Quantile-Quantile, c.-à-d., les quantiles empiriques versus les quantiles de la loi normale standard. La densité est estimée par la méthode du noyau en utilisant un bruit blanc gaussien comme noyau et un facteur de lissage suggéré par [82, page 48]. Dans la figure 6.8, nous vérifions les autocorrélations échantillonnales jusqu'à l'ordre 30, et la p -valeur du test de "Ljung-Box" pour chaque ordre. Nous constatons que l'hypothèse d'indépendance des résidus n'est pas rejetée, ce qui veut dire que ce modèle estime mieux que le modèle 1.

Les paramètres, présentés dans le tableau 6.1, sont a (la moyenne), b (la tendance positive), ω_1 (l'effet positif du premier janvier), ϕ_1 , ϕ_2 , et ϕ_3 (les paramètres autorégressifs), γ_7 (l'effet positif du mois de juillet), et β_5 et β_6 (l'effet positif de vendredi et samedi). Les paramètres qui sont significatifs à un niveau de 10% sont β_1 et β_3 (l'effet négatif du lundi et de mercredi) et γ_1 (effet négatif du mois de janvier). Puisque le mois d'avril possède le niveau moyen le plus bas (figure 6.5), il paraîtrait étrange de voir l'effet du mois de janvier négatif. Ceci est dû au fait qu'on a retranché l'effet du premier de ce mois. Il faut ajouter aussi que les estimations sont faites avec seulement les premiers 1096 jours des données, alors que dans la figure 6.5 on représente les 1537 jours. Nous avons un total de 13 paramètres significatifs. Dans la section 6.1.4.3, nous éliminerons les paramètres non significatifs. L'effet du "Stampede" ω_2 n'étant pas significatif et ceci est expliqué par le fait que cet effet est capturé par l'effet du mois de juillet. Le volume moyen par jour durant ces événements est de 186, et il est de 180 durant le reste des journées de juillet, alors que pour tous les jours de l'année, il est de 174 en moyenne.

Pour ce modèle ci, la variance des résidus est estimée autrement, puisqu'ils sont corrélés. Dans le modèle 1, la variance échantillonnale (6.7) "surestime" $\sigma_E^2 = \text{Var}[E_t]$. Nous multiplions alors l'équation (6.8) par E_t et nous appliquons l'opérateur

Paramètre	<i>a</i>	<i>b</i>	ω_1	ω_2			
	constante	tendance/mois	Jan. 1	Stampede			
Estimation	149.3	0.031	60.5	2.7			
Erreur std.	3.3	0.003	11.0	4.5			
<i>p</i> -val. du t-test	< 0.001	< 0.001	< 0.001	0.544			
Paramètre	β_1	β_2	β_3	β_4	β_5	β_6	β_7
	Lun.	Mar.	Mer.	Jeu.	Ven.	Sam.	Dim.
Est.	-4.2	-5.0	-5.3	-0.9	8.5	7.6	-0.8
Erreur std.	2.5	2.5	2.5	2.5	2.5	2.5	2.5
<i>p</i> -val. du t-test	0.095	0.046	0.034	0.719	0.001	0.002	0.747
Paramètre	γ_1	γ_2	γ_3	γ_4	γ_5	γ_6	γ_7
	Jan.	Fev.	Mar.	Avr.	Mai	Juin	Juil.
Estimation	-5.6	-4.1	-2.5	-4.0	0.5	4.1	13.2
Erreur std.	2.9	2.8	2.8	2.8	2.7	2.8	2.9
<i>p</i> -val. du t-test	0.048	0.152	0.358	0.149	0.849	0.138	< 0.001
Paramètre	γ_8	γ_9	γ_{10}	γ_{11}	γ_{12}		
	Aout	Sep.	Oct.	Nov.	Déc.		
Estimation	-1.3	-0.3	-4.3	0.3	4.1		
Erreur std.	2.8	2.8	2.8	2.8	2.8		
<i>p</i> -val du t-test	0.627	0.928	0.121	0.916	0.150		
Paramètre	ϕ_1	ϕ_2	ϕ_3	σ_a^2			
Estimation	0.192	0.108	0.083	250.1			
Erreur std.	0.030	0.031	0.030	-			
<i>p</i> -val. du t-test	< 0.001	< 0.001	0.006				

Tableau 6.1 – Estimation des paramètres du modèle 2.

d'espérance. Nous obtenons

$$\sigma_E^2 = \mathbb{E}[E_t^2] = \phi_1\gamma_1 + \phi_2\gamma_2 + \phi_3\gamma_3 + \sigma_a^2 = (\phi_1\rho_1 + \phi_2\rho_2 + \phi_3\rho_3)\sigma_E^2 + \sigma_a^2,$$

où $\gamma_i = \text{Cov}(E_t, E_{t-i})$ et $\rho_i = \text{Corr}(E_t, E_{t-i})$ pour chaque i . Nous remplaçons les quantités dans la dernière expression par leur estimations pour σ_E^2 , nous obtenons $\hat{\sigma}_E^2 = 291.8$ comme estimation de σ_E^2 du modèle 1. Si on le compare avec l'estimation $\sigma_a^2 = 250.1$, nous constatons que le modèle 1 a 17% de plus en variance que le modèle 2.

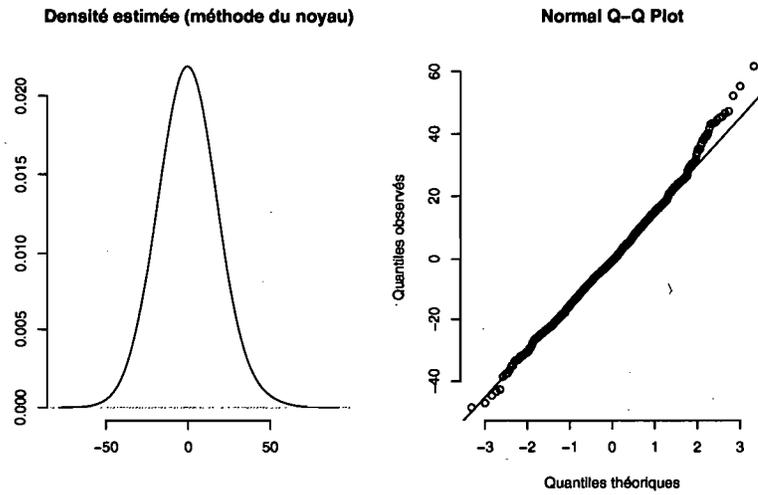


Figure 6.7 – Analyse de la normalité des résidus du modèle 2

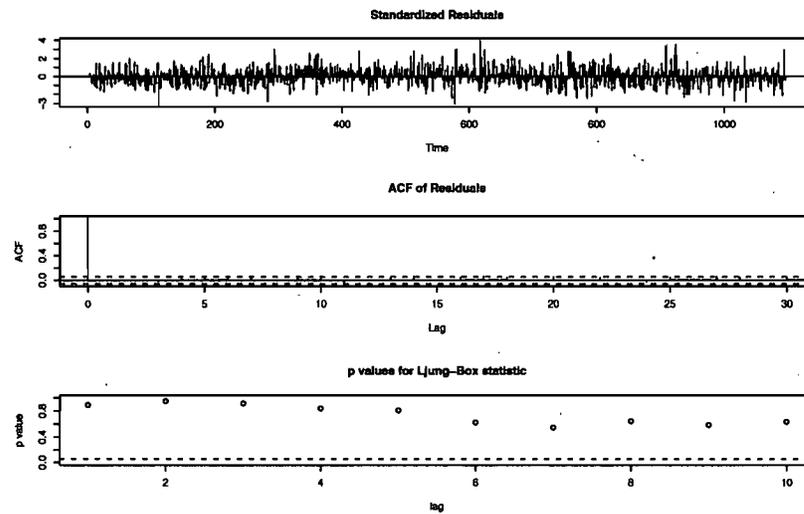


Figure 6.8 – Analyse des corrélations entre les résidus du modèle 2

6.1.4.3 Modèle 3 : ajout des effets croisés

Ce modèle 3 est une extension du modèle 2 en ajoutant les effets croisés ou plus exactement l'interaction entre les jours et les mois. Nous ajoutons la composante suivante :

$$\sum_{j=1}^7 \sum_{k=1}^{12} \delta_{j,k} M_{t,j,k} \quad (6.11)$$

à la partie de droite de l'équation (6.4) et nous la soustrayons entre les crochets en (6.9). La variable indicatrice $M_{t,j,k}$ vaut 1 si l'observation t coïncide avec le jour j du mois k . Nous avons donc les nouveaux paramètres $\delta_{j,k}$, qui satisferont aussi les contraintes d'identifiabilité $\sum_{k=1}^{12} \delta_{j,k} = 0$, pour chaque j et $\sum_{j=1}^7 \delta_{j,k} = 0$, pour chaque k .

Les estimations obtenues pour les paramètres β_j , γ_k , et ω_i dans ce modèle demeurent inchangées par rapport au modèle 2. Dans le tableaux 6.2, nous présentons le reste des paramètres du modèle 3 avec les p -valeurs du test d'hypothèse par la signification des paramètres (égale ou non à zéro). La variance des résidus est de $\sigma_a^2 = 241.5$, qui représente autour de 4% de moins que le modèle 2. Le diagnostic des résidus dans la figure 6.9 et le test de "Ljung-Box" ne rejettent pas l'hypothèse que ces résidus constituent un bruit blanc ($n = 1096$, $Q = 5.394$, et la p -valeur du test est de 0.944), ce qui veut dire que notre modèle est adéquat.

La première remarque concernant ce modèle, est qu'il améliore un peu le modèle 2, mais avec un peu plus de paramètres à estimer. Ceci nous incite à considérer seulement les paramètres significatifs.

	Lun.	Mar.	Mer.	Jeu.	Ven.	Sam.	Dim.	
Param.	$\delta_{1,1}$	$\delta_{2,1}$	$\delta_{3,1}$	$\delta_{4,1}$	$\delta_{5,1}$	$\delta_{6,1}$	$\delta_{7,1}$	Jan.
Est.	-0.3	5.0	4.6	-6.9	4.0	-6.3	-0.2	
Erreur std.	3.9	4.0	4.0	4.2	4.2	4.2	4.2	
p-v. du t-test	0.942	0.210	0.256	0.102	0.339	0.137	0.965	
Param.	$\delta_{1,2}$	$\delta_{2,2}$	$\delta_{3,2}$	$\delta_{4,2}$	$\delta_{5,2}$	$\delta_{6,2}$	$\delta_{7,2}$	Fev.
Est.	-5.4	5.6	3.7	-3.6	1.0	2.9	-4.2	
Erreur std.	4.2	4.1	4.2	4.2	4.2	4.2	4.2	
p-v. du t-test	0.193	0.173	0.382	0.386	0.805	0.486	0.321	
Param.	$\delta_{1,3}$	$\delta_{2,3}$	$\delta_{3,3}$	$\delta_{4,3}$	$\delta_{5,3}$	$\delta_{6,3}$	$\delta_{7,3}$	Mar.
Est.	8.1	6.2	-3.7	6.2	0.5	-12.3	-4.9	
Erreur std.	4.1	4.2	4.1	3.9	3.9	4.2	4.2	
p-v. du t-test	0.052	0.137	0.364	0.113	0.891	0.003	0.240	
Param.	$\delta_{1,4}$	$\delta_{2,4}$	$\delta_{3,4}$	$\delta_{4,4}$	$\delta_{5,4}$	$\delta_{6,4}$	$\delta_{7,4}$	Avr.
Est.	1.7	-3.7	0.4	4.2	-3.3	4.1	-3.4	
Erreur std.	4.2	4.2	4.2	4.2	4.2	3.9	3.9	
p-v. du t-test	0.686	0.369	0.927	0.314	0.434	0.287	0.393	
Param.	$\delta_{1,5}$	$\delta_{2,5}$	$\delta_{3,5}$	$\delta_{4,5}$	$\delta_{5,5}$	$\delta_{6,5}$	$\delta_{7,5}$	Mai
Est.	5.6	-0.8	-0.5	0.1	-6.6	-1.8	4.0	
St.err	3.8	3.9	3.9	4.2	4.2	4.2	4.2	
p-v. du t-test	0.143	0.840	0.908	0.986	0.113	0.663	0.345	
Param.	$\delta_{1,6}$	$\delta_{2,6}$	$\delta_{3,6}$	$\delta_{4,6}$	$\delta_{5,6}$	$\delta_{6,6}$	$\delta_{7,6}$	Juin
Est.	-2.6	-4.5	-1.3	11.5	-6.7	3.0	0.7	
Erreur std.	4.1	4.1	4.2	3.8	3.9	4.2	4.2	
p-v. du t-test	0.525	0.276	0.747	0.003	0.083	0.468	0.875	
Param.	$\delta_{1,7}$	$\delta_{2,7}$	$\delta_{3,7}$	$\delta_{4,7}$	$\delta_{5,7}$	$\delta_{6,7}$	$\delta_{7,7}$	Juil.
Est.	-3.9	-4.6	1.5	0.2	6.6	1.8	-1.6	
Erreur std.	3.9	4.2	4.2	4.2	4.2	3.8	3.9	
p-v. du t-test	0.324	0.269	0.728	0.954	0.112	0.649	0.685	
Param.	$\delta_{1,8}$	$\delta_{2,8}$	$\delta_{3,8}$	$\delta_{4,8}$	$\delta_{5,8}$	$\delta_{6,8}$	$\delta_{7,8}$	Aout
Est.	6.1	4.7	-1.4	-0.8	-3.1	0.1	-5.4	
Erreur std.	4.2	3.9	3.9	3.9	4.2	4.2	4.2	
p-v. du t-test	0.142	0.227	0.711	0.833	0.452	0.986	0.195	
Param.	$\delta_{1,9}$	$\delta_{2,9}$	$\delta_{3,9}$	$\delta_{4,9}$	$\delta_{5,9}$	$\delta_{6,9}$	$\delta_{7,9}$	Sep.
Est.	-1.6	-5.3	-4.5	-2.3	2.9	4.9	5.9	
Erreur std.	7.1	6.8	6.7	6.7	5.4	5.8	7.9	
p-v. du t-test	0.819	0.441	0.502	0.732	0.594	0.399	0.457	
Param.	$\delta_{1,10}$	$\delta_{2,10}$	$\delta_{3,10}$	$\delta_{4,10}$	$\delta_{5,10}$	$\delta_{6,10}$	$\delta_{7,10}$	Oct.
Est.	0.9	-1.5	-3.7	-6.8	3.2	0.8	7.1	
Erreur std.	4.0	4.0	4.2	4.2	4.2	4.2	3.9	
p-v. du t-test	0.822	0.702	0.371	0.105	0.436	0.839	0.073	
Param.	$\delta_{1,11}$	$\delta_{2,11}$	$\delta_{3,11}$	$\delta_{4,11}$	$\delta_{5,11}$	$\delta_{6,11}$	$\delta_{7,11}$	Nov.
Est.	-1.2	0.1	7.8	-2.2	4.5	-5.4	-3.5	
Erreur std.	4.3	3.9	3.9	3.9	4.1	4.1	4.6	
p-v. du t-test	0.775	0.982	0.044	0.578	0.280	0.187	0.444	
Param.	$\delta_{1,12}$	$\delta_{2,12}$	$\delta_{3,12}$	$\delta_{4,12}$	$\delta_{5,12}$	$\delta_{6,12}$	$\delta_{7,12}$	Déc.
Est.	-7.3	-1.1	-2.7	0.4	-3.1	8.2	5.7	
Erreur std.	4.3	4.3	4.2	4.2	3.9	3.9	4.0	
p-v. du t-test	0.092	0.797	0.518	0.924	0.417	0.036	0.154	
Param.	ϕ_1	ϕ_2	ϕ_3	σ_a^2				
Est.	0.213	0.126	0.085	241.5				
Erreur std.	0.030	0.031	0.031	-				
p-v. du t-test	< 0.001	< 0.001	0.006					

Tableau 6.2 – Estimation des paramètres du modèle 3

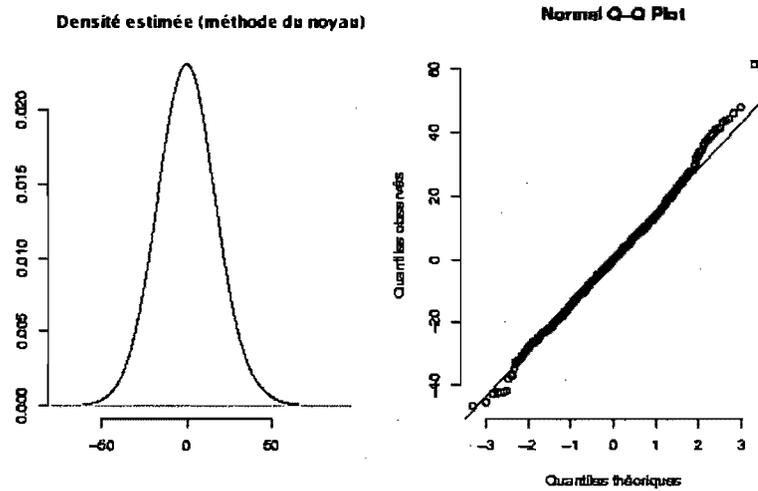


Figure 6.9 – Analyse de la normalité des résidus du modèle 3

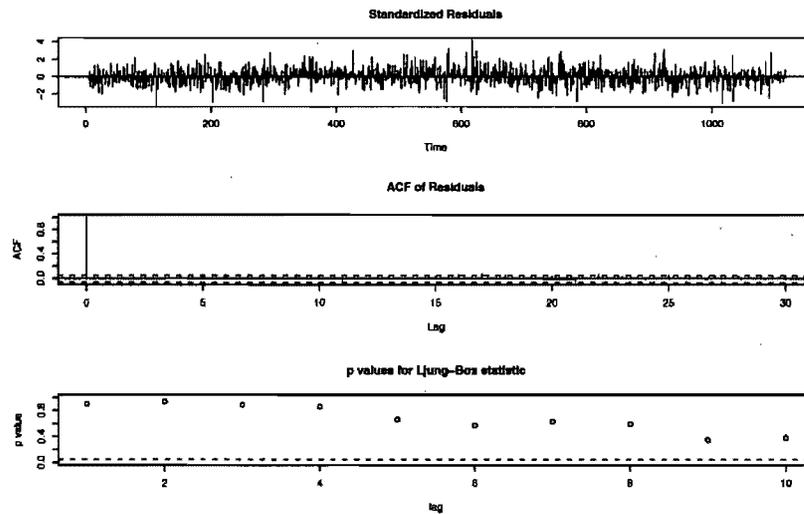


Figure 6.10 – Analyse des corrélations entre les résidus du modèle 3

6.1.4.4 Modèle 4 : seulement des paramètres significatifs

Dans ce modèle, nous retenons seulement les paramètres significatifs du modèle 3. Nous choisirons un niveau de 10% (c.-à-d., seulement les paramètres qui ont une p -valeur inférieure à 0.10 dans les tableaux 6.1 et 6.2). Dans le tableaux 6.2, 8 paramètres $\delta_{j,k}$ et 3 paramètres ϕ_i sont significatifs à un niveau de 90%. Nous avons aussi 10 autres paramètres significatifs dans le tableau 6.1, pour un total de 20. En considérant les contraintes d'identifiabilité, seulement $s = 15$ paramètres indépendants restent parmi les 20. Il faut bien sûr noter que ceci est valable seulement pour la ville en question. Pour d'autres villes et d'autres données les variables significatives changent.

Nous recommençons alors l'estimation en considérant les paramètres non significatifs comme nuls et les résultats sont dans le tableau 6.3. Les paramètres estimés sont tous significatifs et nous pouvons aussi remarquer que les paramètres $\delta_{j,k}$ pour les samedis du mois de mars ont un effet négatif, alors que les mercredis de juin ont un effet positif. A un niveau 10%, nous avons aussi les samedis de décembre, les jeudis de novembre, les lundis du mois de mars, les dimanches du mois d'octobre, les vendredis du mois de juin et les lundis du mois de décembre.

La variance des résidus est estimée à $\sigma_a^2 = 241.9$. L'analyse des résidus est représentée dans les graphiques 6.11 et 6.12. Le test de "Ljung-Box" ne détecte pas de corrélations entre les résidus (avec $n = 1096$, $Q = 11.581$, et une p -valeur de 0.48) et ne rejette pas l'hypothèse d'indépendance de ces résidus, ce qui veut dire que ce modèle est adéquat pour nos données.

Paramètre	a	b	ω_1					
	constante	tendance/mois	Jan. 1					
Estimation	149.6	0.032	57.8					
Erreur std.	1.9	0.003	10.7					
p-val. du t-test	< 0.001	< 0.001	< 0.001					
Paramètre	β_1	β_2	β_3	β_5	β_6			
	Lun.	Mar.	Mer.	Ven.	Sam.			
Estimation	-4.3	-5.3	-6.2	7.8	7.9			
Erreur std.	1.5	1.4	1.4	1.4	1.5			
p-val. du t-test	0.004	< 0.001	< 0.001	< 0.001	< 0.001			
Paramètre	γ_1	γ_7						
	Jan.	Juil.						
Estimation	-5.7	12.4						
Erreur std.	2.9	2.7						
p-val. du t-test	0.050	< 0.001						
Paramètre	$\delta_{1,3}$	$\delta_{6,3}$	$\delta_{4,6}$	$\delta_{5,6}$	$\delta_{7,10}$	$\delta_{3,11}$	$\delta_{1,12}$	$\delta_{6,12}$
	Lun.	Sam.	Jeu.	Ven.	Dim.	mer.	Lun.	Sam.
	Mar.	Mar.	Juin	Juin	Oct.	Nov.	Déc.	Déc.
Estimation	7.9	-11.1	12.0	-7.1	8.3	8.5	-7.8	8.2
Erreur std.	4.2	4.3	3.9	3.9	3.8	4.0	4.5	4.1
p-v. du t-test	0.060	0.010	0.002	0.069	0.029	0.034	0.083	0.046
Paramètre	ϕ_1	ϕ_2	ϕ_3	σ_a^2				
Estimation	0.212	0.132	0.094	241.6				
Erreur std.	0.030	0.031	0.031	-				
p-v. du t-test	< 0.001	< 0.001	0.002	-				

Tableau 6.3 – Estimation des paramètres du modèle 4

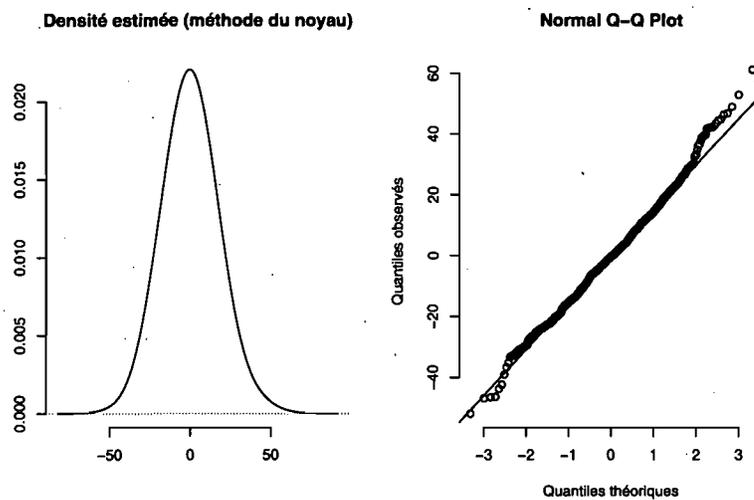


Figure 6.11 – Analyse de la normalité des résidus du modèle 4

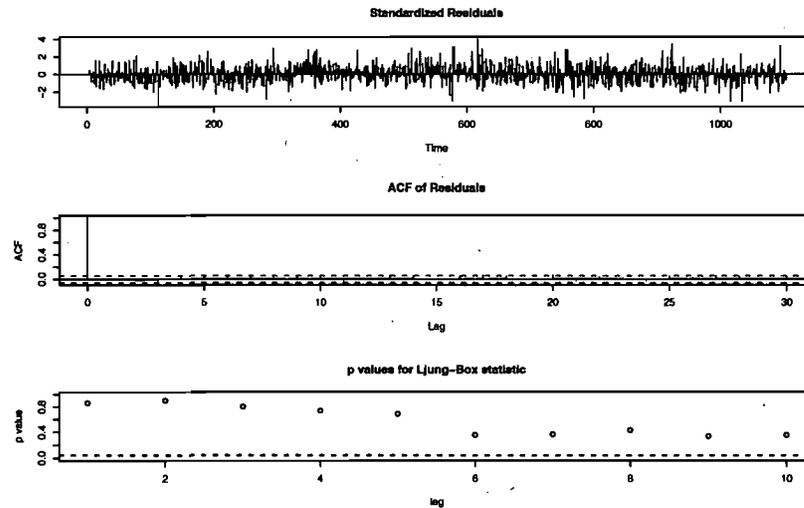


Figure 6.12 – Analyse des corrélations entre les résidus du modèle 4

6.1.4.5 Modèle 5 : ARIMA doublement saisonnier

Nous considérons ici un modèle différent : le modèle ARIMA doublement saisonnier. La série est décomposée comme suit :

$$Y_t = N_t + \omega_1 H_{t,1} + \omega_2 H_{t,2}, \quad (6.12)$$

où $\{N_t\}$ est modélisé par un processus ARIMA doublement saisonnier et la deuxième composante capte l'effet des journées spéciales (le premier janvier et les journées du "Stampede"). Les deux cycles qu'on considère ici sont ceux de la semaine et de l'année : le premier cycle de la semaine avec une période $s_1 = 7$, et un cycle annuel approximatif avec une période de $s_2 = 365$. Ceci implique que chaque jour de l'année affiche le même niveau que celui de l'année passée plus un bruit. Autrement dit, la moyenne conditionnelle de N_t est régressée sur N_{t-365} ; par exemple le premier janvier 2004 est régressé sur le premier janvier 2003, en veillant à enlever le 29 février pour avoir la périodicité de 365 jours.

La forme générale du modèle ARIMA doublement saisonnier [9, 12, 91] avec des périodes s_1 et s_2 est donnée par :

$$\phi(B)\Phi_{s_1}(B^{s_1})\Phi_{s_2}(B^{s_2})\nabla^d\nabla_{s_1}^{d_1}\nabla_{s_2}^{d_2}N_t = \theta(B)\Theta_{s_1}(B^{s_1})\Theta_{s_2}(B^{s_2})a_t, \quad (6.13)$$

où $\nabla_s^d = (1 - B^s)^d$, ϕ , Φ_{s_1} , Φ_{s_2} , θ , Θ_{s_1} , et Θ_{s_2} , sont des fonctions polynomiales d'ordre p , p_1 , p_2 , q , q_1 , et q_2 , respectivement, et $\{a_t\}$ un bruit blanc gaussien. Ce processus est noté $\text{ARIMA}(p, d, q) \times (p_1, d_1, q_1)_{s_1} \times (p_2, d_2, q_2)_{s_2}$.

Nous utilisons les outils nécessaire pour l'identification du modèle adéquat (choix de l'ordre des polynômes et les composantes d , d_1 , et d_2), l'estimation des paramètres (ω_1 , ω_2 , et les coefficients des polynômes), ainsi que le diagnostic et la validation du modèle [9, 91].

En général, les modèles ARIMA avec plus qu'un cycle saisonnier sont difficiles à identifier et aussi à estimer. Ceci est dû à la nature multiplicative du modèle avec ses multiples opérateurs dans (6.13). Pour cela, un critère pour l'identification et la sélection du meilleur modèle s'impose. Nous avons utilisé le *critère d'information d'Akaike* (AIC), présenté dans la section 3.5. Dans ce cas-ci, $\hat{\sigma}_a^2$ est l'estimateur de maximum de vraisemblance de la variance des résidus qui est approximativement la même que la variance échantillonnale, donnée en (6.10), sans l'hypothèse que les résidus sont i.i.d. de distribution normale [72].

Le modèle retenu présente un AIC minimum, tout en respectant la condition que les résidus présentent un bruit blanc [9]. Voici donc le modèle pour N_t

$$(1 - \phi_7 B^7 - \phi_{14} B^{14} - \phi_{28} B^{28})(1 - \phi_{365} B^{365})(1 - B)N_t = (1 - \theta_1 B)a_t. \quad (6.14)$$

Les paramètres sont estimés avec moindres carrés en se basant sur (6.12) et (6.14), c.-à-d., les paramètres qui minimisent la somme des carrés des résidus es-

Paramètre	ω_1	ω_2				
	Jan. 1	“Stampede”				
Estimation	43.8	16.6				
Erreur std.	12.0	3.8				
p-v. du t-test	< 0.001	< 0.001				
Paramètre	ϕ_7	ϕ_{14}	ϕ_{28}	ϕ_{365}	θ_1	σ_a^2
Estimation	0.064	0.103	0.082	0.128	0.905	251.7
Erreur std.	0.03	0.03	0.03	0.04	0.01	-
p-v. du t-test	0.038	0.001	0.007	0.001	< 0.0001	

Tableau 6.4 – Estimation des paramètres du modèle 5

timés. Les estimations sont présentées dans le tableau 6.4, avec les p -valeurs. On peut facilement remarquer ici que ce modèle présente peu de paramètres par rapport aux modèles précédents. Aussi, on peut noter que ce modèle ne capte pas l'effet du mois de janvier, mais en ajoutant l'effet des journées spéciales il a pu capter l'effet du “Stampede”, qui est très significatif (paramètre ω_2) pour ce modèle-ci.

6.1.4.6 Comparaison des modèles pour les arrivées par jour

Dans cette section, nous comparons la performance des 5 modèles proposés en terme de la qualité de l'estimation et la performance des prévisions. Les résultats sont dans le tableau 6.5.

Pour la qualité de l'estimation, nous reportons l'erreur standard pour les résidus de chaque modèle, $\hat{\sigma}_a$, le nombre s de paramètres estimés, et le *critère d'information d'Akaike* (AIC).

Les modèles présentés sont estimés en utilisant les 1096 premiers jours des données. Ces modèles sont alors utilisés pour les prévisions des 441 jours restants des données (pour $t = 1097, \dots, 1537$), avec un *horizon* allant de 1 à 21 jours dans

le futur. L'erreur de prévision d'ordre ℓ au jour t est définie par :

$$e_t(\ell) = Y_{t+\ell} - \hat{Y}_t(\ell),$$

où $\hat{Y}_t(\ell)$ représente les prévisions de $Y_{t+\ell}$ basée sur l'information observée au jour t . Les formules explicites des prévisions par le modèle 5 (ARIMA doublement saisonnier) s'avèrent assez compliquées [12, pages 175-182], cependant les logiciels utilisés les fournissent.

Les mesures habituellement utilisées pour mesurer les performances des prévisions sont : *racine carrée de l'erreur quadratique moyenne* (RCEQM) et l'*erreur absolue relative moyenne* (EARM) à différents ordre de prévision, qu'on définit comme suit :

$$\text{RCEQM}(\ell) = \sqrt{\frac{1}{442 - \ell} \sum_{t=1097}^{1538-\ell} e_t^2(\ell)}$$

et

$$\text{EARM}(\ell) = \frac{1}{442 - \ell} \sum_{t=1097}^{1538-\ell} \frac{|e_t(\ell)|}{Y_{t+\ell}},$$

pour l'ordre ℓ .

Avec la EARM, l'erreur est relativisée par rapport au volume d'appels correspondant, $Y_{t+\ell}$, ce qui reflète l'idée qu'un volume élevé engendre une petite erreur. D'un autre coté aussi un volume très bas $|Y_{t+\ell}|$ va engendrer une erreur importante. Ce genre de mesure doit être considéré avec plus de précaution.

Dans le tableau 6.5, nous présentons un sommaire des comparaisons effectuées. La première partie concerne l'estimation des modèles avec les 1096 premières observations. La variance estimée des résidus $\hat{\sigma}_a^2$, et sa racine carrée $\hat{\sigma}_a$ (erreur standard d'estimation), le nombre s de paramètres indépendants estimés, le degrés de liberté

	Modèle 1	Modèle 2	Modèle 3	Modèle 4	Modèle 5
$\hat{\sigma}_a^2$	291.8	250.1	241.5	241.6	251.7
Erreur std. $\hat{\sigma}_a$	17.08	15.81	15.54	15.55	15.87
s	21	24	90	15	7
Degrés de liberté	1075	1072	1006	1081	1088
AIC(s)		6099	6194	6045	6068
RCEQM(1)	17.82	15.38	15.31	13.91	15.68
EARM(1) (en %)	7.58	6.14	6.14	5.72	6.91

Tableau 6.5 – Comparaison des modèles pour les arrivées journalières

$n - s$, le critère AIC(s) (celui-la n'est pas défini pour le modèle 1, $\hat{\sigma}_a^2$ étant remplacé par $\hat{\sigma}_E^2$, défini à la section 6.1.4.1). Selon le critère AIC, le modèle 4 semble être le meilleur, suivi du modèle 5, et puis le modèle 2. Cependant, il faut noter que le choix du modèle 5 parmi tous les modèles "possibles" ARIMA a été basé sur le critère AIC, ce qui explique son rang.

Dans la deuxième partie du tableau, nous présentons les RCEQM et EARM pour les prévisions à un horizon d'une journée. Tandis que les RCEQM à des horizons 1 à 21 sont dans la figure 6.13. Pour des petits horizons, le modèle 4 donne de meilleurs résultats en prévisions, suivi des modèles 2 et 3. Pour des horizons $l \geq 13$, les RCEQM(l) sont approximativement égaux pour tous les modèles, sauf pour le modèle 5. Ceci nous dit que le critère de sélection selon le AIC est efficace pour notre cas.

Pour l'interprétation de l'erreur standard d'estimation et du RCEQM, il est utile de faire le lien avec le nombre moyen d'appels par jour, soit 174 appels. Si on considère que les appels sont générés à partir d'un processus de Poisson avec taux constant de 174, alors l'écart-type du nombre d'appels par jour serait de $\sqrt{174} = 13.2$. Avec le modèle 4 et un horizon d'une journée, nous avons obtenu un chiffre semblable. Ceci nous amène à dire que si nous observons le volume d'appels jusqu'à un certain temps, le taux d'arrivée pour les prochaines 24 heures peut

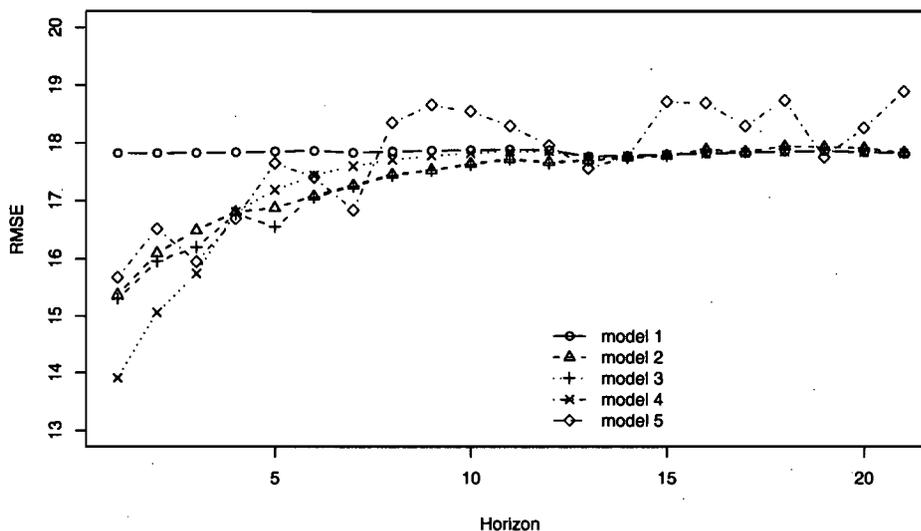


Figure 6.13 – Erreur de prévision $RCEQM(l)$ pour les modèles de 1 à 5, avec les horizons $l = 1, \dots, 21$ jours.

être alors presque déterministe. Par contre, avec un RCEQM assez élevé pour des horizons plus longs qu'une journée, on peut voir les arrivées comme étant générés selon un processus de Poisson avec un taux variable.

6.1.5 Modèles pour les arrivées par intervalles de temps dans une journée

Les modèles pour les arrivées des appels par heures sont inspirés par le succès des modèles des arrivées journaliers. Nous notons par Z_h le nombre d'appels reçus durant l'heure h de la journée ; $h = 1, \dots, 24n$ et $n = 1537$. Nous présentons deux modèles pour les heures construits sur la base des modèles pour le volume par jour Y_t en ajoutant une composant qui départage la journée en 24 heures. La première approche est basée sur la *distribution conditionnelle* du vecteur du nombre d'appels reçus par heure, sachant le volume total de la journée. La seconde approche utilise

un modèle de séries basé sur les observations des données par heure.

6.1.5.1 Modèle 6 : distribution conditionnelle

Nous utilisons ici le modèle 4 pour les arrivées journalières, puis nous supposons que pour un jour quelconque t , la distribution conditionnelle du vecteur $\mathbf{Z}_t = (Z_{24(t-1)+1}, \dots, Z_{24t})$, sachant Y_t , est indépendante des autres jours. Nous considérons alors la distribution multinomiale, de paramètres (N, p_1, \dots, p_{24}) , pour la distribution conditionnelle du vecteur donné, où $N = Y_t$. Ici, chaque p_i représente la probabilité qu'un appel quelconque parmi ceux reçus dans la journée arrive exactement dans l'heure i . Le vecteur (p_1, \dots, p_{24}) est appelé *profil du jour*. Alors, conditionnellement au volume du jour Y_t et selon la propriété de la loi multinomiale, les arrivées durant les différentes heures de la journée sont indépendants.

Dans l'analyse des données du centre C4, à la section 6.1.3, nous avons constaté qu'il est plus approprié d'utiliser différents vecteurs (p_1, \dots, p_{24}) selon le type de journée. Nous avons établi 4 types de journées : (1) du lundi au mercredi, (2) jeudi et dimanche, (3) vendredi, et (4) samedi. Pour chaque catégorie c , pour $c = 1, 2, 3, 4$, nous estimons cette probabilité par la fraction des appels de la catégorie c qui arrivent pendant l'heure i , c.-à-d.,

$$\hat{p}_{c,i} = \frac{\sum_{t=1}^n Z_{24(t-1)+i} P_{t,c}}{\sum_{t=1}^n \sum_{h=1}^{24} Z_{24(t-1)+h} P_{t,c}}, \quad (6.15)$$

pour $i = 1, 2, \dots, 24$, où la variable indicatrice $P_{t,c}$ vaut 1 si le jour t est de la catégorie c et 0 sinon. L'avantage avec ce modèle est que le modèle pour les arrivées par jour demeure inchangé. Évidemment d'autres distributions que la distribution multinomiale, pour la distribution conditionnelle de \mathbf{Z}_t , peuvent être suggérées.

Pour tester la validité de notre modèle, nous utilisons le test suivant : sous

l'hypothèse que la distribution conditionnelle à Y_t est multinomiale, et si le jour t tombe dans la catégorie c , alors la statistique

$$Q_t = \sum_{i=1}^{24} \frac{(Z_{24(t-1)+i} - Y_t p_{c,i})^2}{Y_t p_{c,i}}$$

devrait avoir une distribution approximative chi-deux avec 23 degrés de liberté, sous la condition que les $Y_t p_{c,i}$ soient assez élevés (par exemple plus grands que 5) pour chaque $i = 1, \dots, 24$ [74]. Alors, pour chaque type de jour, nous avons calculé le Q_t et nous avons comparé leurs distributions empiriques avec celle de chi-deux. Pour garantir que toutes les $Y_t p_{c,i}$ sont supérieures à 5, nous avons agrégé les heures de la nuit et tôt le matin. Nous avons regroupé les heures de 3h00 à 7h00 en un seul groupe, ce qui fait que nous avons plutôt $m = 21$ périodes et à chaque période l'espérance est au moins égale à 5. La probabilité qu'un appel de la journée arrive pendant la période i en une journée de type c est donnée par

$$\tilde{p}_{c,i} = \begin{cases} p_{c,i} & \text{for } i = 1, 2, 3, \\ p_{c,4} + p_{c,5} + p_{c,6} + p_{c,7} & \text{for } i = 4, \\ p_{c,i+3} & \text{for } i = 5, \dots, 21. \end{cases} \quad (6.16)$$

Dénotons par $\tilde{Z}_{t,i}$ le nombre d'appels dans la période i du jour t . Nous avons que $\sum_{i=1}^m \tilde{Z}_{t,i} = Y_t$. Pour une journée de la catégorie c , conditionnellement à Y_t , sous l'hypothèse nulle (pour la loi multinomiale), nous avons que $\tilde{\mathbf{Z}}_t = (\tilde{Z}_{t,1}, \dots, \tilde{Z}_{t,m})$ est distribué selon la loi multinomiale de paramètres $(Y_t, \tilde{p}_{c,1}, \dots, \tilde{p}_{c,m})$. Comme estimateurs pour les paramètres $\tilde{p}_{c,i}$, nous utilisons ceux données dans (6.16). Alors, pour n assez grand, la statistique du test de chi-deux de Pearson

$$Q_{m-1,t} = \sum_{i=1}^m \frac{(\tilde{Z}_{t,i} - Y_t \hat{\tilde{p}}_{c,i})^2}{Y_t \hat{\tilde{p}}_{c,i}}, \quad (6.17)$$

est approximativement distribuée selon la loi chi-deux avec ($m - 1 = 20$) degrés de liberté sous l'hypothèse de la loi multinomiale.

Nous avons calculé les $Q_{m-1,t}$ pour les 1537 valeurs observées, et nous avons comparé leur distribution empirique (distribution A) avec celle de la loi chi-deux (distribution C) en utilisant le diagramme Quantile-Quantile. Au début, nous avons constaté qu'il y a beaucoup de différences entre les distributions, surtout dans la partie supérieure de la courbe. Il se peut que cette distribution ne soit pas la plus appropriée. Pour cela, nous avons plutôt utilisé la simulation pour générer une série de 1537 réalisations successives de Y_t sous le modèle 4, puis un échantillon de vecteurs \tilde{Z}_t conditionnellement à Y_t sous l'hypothèse de la loi multinomiale, et avec ces données nous avons calculé les valeurs correspondantes $Q_{m-1,t}$, pour $t = 1, \dots, 1537$. La distribution empirique de cet échantillon est appelée distribution B, et dans la figure 6.14, nous montrons le diagramme Quantile-Quantile de la distribution A versus la distribution B. L'estimation semble être excellente cette fois-ci, à part un peu dans la partie supérieure de la courbe. Ceci est dû au fait que les observations à droite qui causent cette légère différence correspondent à des journées spéciales où le profil ne correspond à aucun des profils choisis selon les 4 catégories de jours établis auparavant. Ceci peut être aussi expliqué par le fait que des événements particuliers ont eu lieu à ces dates. Les journées observées sont celles du dimanche 29 octobre 2000, lundi 19 mars 2001, vendredi 13 avril 2001, lundi 30 septembre 2001, dimanche 28 octobre 2001, samedi 28 octobre 2002, samedi 21 mai 2003, et le premier janvier de chaque année. Pour la journée du premier janvier, le profil peut être établi ou prédit d'avance vu que ça se répète chaque année. Nous pouvons conclure que le choix de la distribution multinomiale est un bon choix, et il nous permettra d'effectuer des prévisions avec ce modèle pour les arrivées pour les heures de journée, sachant le volume journalier total.

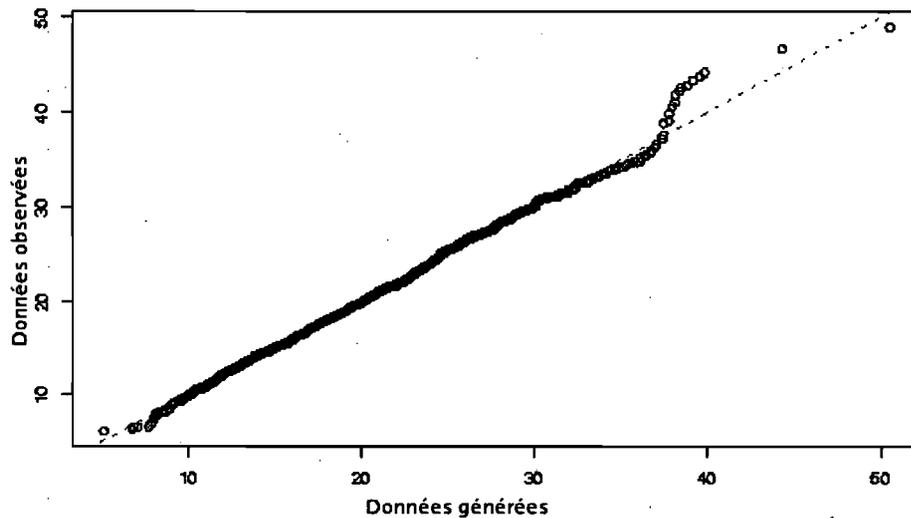


Figure 6.14 – Diagramme Quantile-Quantile pour la distribution empirique A versus la distribution empirique B pour un échantillon de $Q_{m-1,t}$ générés selon le modèle 6 sous l'hypothèse de la loi multinomiale.

Maintenant, nous entamons les prévisions pour les heures en utilisant ce modèle. Si h est la h ième heure du jour t , une prévision de Z_h faite ℓ jours avant le jour t (à la fin du jour $t - \ell$, alors $\ell = 1$ correspond au début de la journée t) est tout simplement $\hat{p}_{c,i}\hat{Y}_{t-\ell}(\ell)$, où $\hat{Y}_{t-\ell}(\ell)$ est la prévision de Y_t tel que définie dans la section 6.1.4 et c est la catégorie du jour t .

Pour les prévisions de Z_h pour l'heure h , nous utilisons $\hat{p}_{c,i}\hat{Y}_{t-1}(1)$, bien que l'on ferait mieux si on prenait en compte l'information déjà disponible sur le volume quotidien des jours précédents. Par exemple, celui du jour t , jusqu'à l'heure $h - 1$. Supposons qu'à 11h00 nous voulons calculer les prévisions des arrivées pour les prochaines 13 heures. Nous tenons donc compte de la quantité disponible, qu'on appellera $W_{t,11}$. Alors, nous calculons les prévisions du volume d'appels pour le

restant de la journée t , $\hat{Y}_{t-1}(1) - W_{t,11}$, comme suit :

$$\frac{\hat{p}_{c,i}(\hat{Y}_{t-1}(1) - W_{t,11})}{\hat{p}_{c,12} + \dots + \hat{p}_{c,24}} \quad (6.18)$$

pour la i ème heure, et $i > 11$. Un inconvénient ici, c'est que si nous enregistrons un $W_{t,11}$ assez grand, alors nous aurons aussi les prévisions assez basses pour le restant de la journée, qui expliqué par une corrélation négative entre le volume d'appels durant les heures de la journée, alors qu'elle est en fait positive. Nous pensons donc qu'il est plus approprié d'exploiter l'information déjà existante ($W_{t,11}$) pour une mise à jour des prévisions des arrivées dans la journée en cours \tilde{Y}_t . Nous utilisons donc la même formule pour les prévisions des heures qui restent, $Z_h, = 12, \dots, 24$, c.-à.-d. $\hat{p}_{c,i}\hat{Y}_{t-1}(1)$, mais en veillant à mettre à jour la prévision $\hat{Y}_{t-1}(1)$ en se servant de $W_{t,11}$ et de la formule de Bayes.

Pour un cas plus général, où nous observons les arrivées jusqu'à l'heure i de la journée, notons par $W_{t,i} = Z_{24(t-1)+1} + \dots + Z_{24(t-1)+i}$, le nombre d'appels durant les i premières heures du jour t . Sous notre modèle, $\tilde{Y}_t = \hat{Y}_{t-1}(1)$ est une statistique suffisante pour l'information des jours précédents. Avec la formule de Bayes, la distribution conditionnelle de Y_t sachant \tilde{Y}_t et $W_{t,i}$ est

$$\mathbb{P}[Y_t = y \mid \tilde{Y}_t, W_{t,i} = w] = \frac{\mathbb{P}[W_{t,i} = w \mid Y_t = y] \mathbb{P}[Y_t = y \mid \tilde{Y}_t]}{\mathbb{P}[W_{t,i} = w \mid \tilde{Y}_t]}, \quad (6.19)$$

pour tous $0 \leq w \leq y$. La distribution conditionnelle de $W_{t,i}$ sachant $Y_t = y$ est binomiale de paramètres $(y, p_{c,1:i})$, où $p_{c,1:i} = \sum_{\ell=1}^i p_{c,\ell}$. Bien que Y_t prend seulement des valeur entières, le modèle 4 approxime sa distribution conditionnelle à \tilde{Y}_t par la loi normale de moyenne \tilde{Y}_t et de variance $\sigma_Y^2 = \text{Var}[\phi^{-1}(B)(a_t)]$. Ceci peut être utilisé pour spécifier l'expression pour les probabilités dans (6.19) et les calculer numériquement. Pour Y_t , la probabilité pour un entier y peut être approximée en

intégrant la densité de la loi normale sur l'intervalle $[y - 1/2, y + 1/2]$. Il faut noter que conditionnellement à $W_{t,i}$ et Y_t , le vecteur $(Z_{t,i+1}, \dots, Z_{t,24})$ est distribué selon la loi multinomiale de paramètres $(Y_t - W_{t,i}, p_{c,i+1}/(1 - p_{c,1:i}), \dots, p_{c,24}/(1 - p_{c,1:i}))$.

Pour les prévisions, nous utilisons alors l'espérance conditionnelle $\mathbb{E}[Y_t \mid \tilde{Y}_t, W_{t,i}]$ au lieu de la distribution entière (6.19). Si nous supposons que la paire $(Y_t, W_{t,i})$ est approximativement distribuée selon la loi normale bivariée, qui est assez proche de la vraie loi sous notre modèle quand $p_{c,1:i}$ n'est pas proche de 0 ou 1 et Y_t a une espérance assez large, alors nous avons [41, page 93] :

$$\mathbb{E}[Y_t \mid \tilde{Y}_t, W_{t,i} = w] = \tilde{Y}_t + \frac{\text{Cov}[Y_t, W_{t,i} \mid \tilde{Y}_t]}{\text{Var}[W_{t,i} \mid \tilde{Y}_t]}(w - \mathbb{E}[W_{t,i} \mid \tilde{Y}_t]). \quad (6.20)$$

Or, $\mathbb{E}[W_{t,i} \mid \tilde{Y}_t] = p_{c,1:i}\tilde{Y}_t$, $\text{Cov}[Y_t, W_{t,i} \mid \tilde{Y}_t] = p_{c,1:i}\sigma_a^2$, et

$$\begin{aligned} \text{Var}[W_{t,i} \mid \tilde{Y}_t] &= \mathbb{E}_{Y_t}[\text{Var}[W_{t,i} \mid \tilde{Y}_t, Y_t]] + \text{Var}_{Y_t}[\mathbb{E}[W_{t,i} \mid \tilde{Y}_t, Y_t]] \\ &= \tilde{Y}_t p_{c,1:i}(1 - p_{c,1:i}) + p_{c,1:i}^2 \sigma_a^2. \end{aligned}$$

Si on combine ceci avec (6.20), nous obtenons

$$\mathbb{E}[Y_t \mid \tilde{Y}_t, W_{t,i} = w] = \tilde{Y}_t + \frac{w - p_{c,1:i}\tilde{Y}_t}{(1 - p_{c,1:i})\tilde{Y}_t/\sigma_a^2 + p_{c,1:i}}. \quad (6.21)$$

Nous verrons plus loin l'application de cette formule.

6.1.5.2 Modèle 7 : extension du modèle 4 et ajout de l'effet d'heure de la journée

Nous écrivons ce modèle comme suit :

$$Z_h = p_{c,i}Y_t + W_h$$

si h est la i ème heure du jour t de catégorie c , où le processus Y_t provient du modèle 4 et le processus W_h est AR(q) pour un certain ordre q .

Si Y_t provient du modèle 4, alors la variance σ_a^2 des résidus devrait être réduite d'avantage à cause de la variance additionnelle captée par les W_h . Ceci nous donne

$$Z_h = p_{c,i} \left[\hat{Y}_t + E_t \right] + \sum_{l=1}^{24} \alpha_l D_{h,l} + W_h, \quad (6.22)$$

où $D_{h,l} = 1$ si h est dans la l ème heure du jour et 0 sinon, en plus d'avoir $\{E_t\}$ et $\{W_h\}$ des processus AR.

La majeure différence entre les modèles 6 et 7 est la suivante : avec le modèle 6, nous avons une corrélation positive entre les arrivées dans différentes heures de la même journée, et la seule corrélation entre les heures de deux journées successives est captée par les autocorrélations au niveau du processus E_t . Avec le modèle 7, la corrélation entre les arrivées de deux heures d'une même journée diminue avec la distance entre ces deux heures, et la corrélation entre les heures de deux journées successives est enregistrée seulement pour des heures proches (par exemple, vendredi soir et samedi tôt le matin).

Lors de l'estimation du modèle, nous avons ajouté les deux contraintes suivantes : $\sum_{l=1}^{24} \alpha_l = 0$ et $\sum_{h=24(t-1)+1}^{24t} W_h = 0$, pour $t = 1, \dots, n$, et nous avons forcé $\{E_t\}$ d'être un processus AR(3) comme dans le cas du modèle pour les journées. Nous avons donc estimé le modèle $\{W_h\}$. Le plus grand ordre de différence (observé) pour lequel le paramètre autorégressif est significatif à un niveau 5% est l'ordre 44, cependant très peu de paramètres autorégressifs d'ordre plus grand que 25 sont significatifs à un niveau 1%. Pour cela, nous avons décidé d'adopter le modèle AR(25) pour W_h et le modèle AR(3) pour E_t . Nous avons donc ré-estimé le modèle en (6.22) avec ces contraintes et c'est notre modèle 7. Avec ce modèle

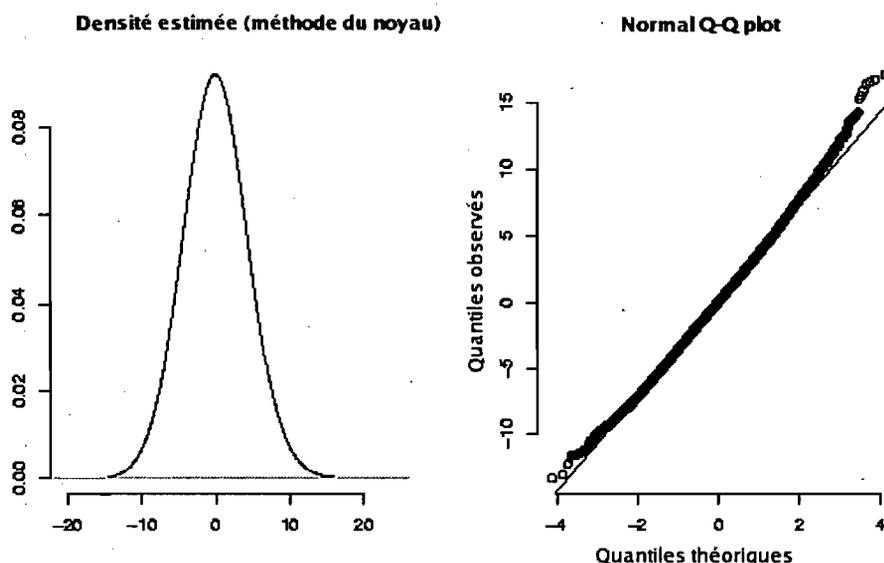


Figure 6.15 – Analyse de la normalité des résidus du modèle 7

nous avons obtenu une erreur standard de 3.6. Le diagnostic des résidus, présenté dans la figure 6.15, nous confirme la normalité de ces résidus.

Pour comparer les deux modèles 6 et 7 de point de vue performance de prévisions, nous utilisons la *racine carrée de l'erreur quadratique moyenne* (RCEQM) à différents ordres pour les 441 jours, $t = 1097, \dots, 1537$. Pour $h = 24(t - 1) + 1, \dots, 24t$, nous définissons, à l'ordre r et l'heure h , l'erreur $e_h(r) = Z_{h+r} - \hat{Z}_h(r)$, où $\hat{Z}_h(r)$ désigne la prévision de Z_{h+r} à un horizon r selon un modèle donné. Nous considérons deux cas :

(1) Pour les prévisions des prochaines 24 heures du jour t en étant au début du jour t et nous avons \tilde{Y}_t la prévision de Y_t selon le modèle 4, nous mesurons l'erreur par

$$\text{RCEQM}(r) = \sqrt{\frac{1}{441} \sum_{t=1096}^{1536} e_{24(t-1)}^2(r)};$$

Tableau 6.6 – Erreur de prévision (RCEQM) par origine et par horizon avec les deux modèles 6 et 7 pour les arrivées par heure.

	Horizon r	Modèle 6	Modèle 7
Cas (1) : Prévisions pour $Z_{24(t-1)+r}$ à Z_{24t} , au temps $h = 24(t-1)$	12 (0 :00-12 :00)	3.1	3.5
	14 (0 :00-14 :00)	3.1	3.9
	17 (0 :00-17 :00)	3.5	3.9
	23 (0 :00-23 :00)	3.1	3.2
	24 (0 :00-24 :00)	3.2	3.3
Cas (2) : Prévisions pour $Z_{24(t-1)+11+r}$ à Z_{24t} , au temps $h = 24(t-1) + 11$	1 (11 :00-12 :00)	3.1	3.1
	3 (11 :00-14 :00)	2.9	3.5
	6 (11 :00-17 :00)	2.3	3.7
	12 (11 :00-23 :00)	3.1	3.2
	13 (11 :00-24 :00)	3.1	3.1

(2) Pour les prévisions des heures $i = 12, \dots, 24$, du jour t après avoir observé les 11 premières heures en utilisant la formule (6.21) pour la mise à jour de la prévision de Y_t pour le modèle 6, nous mesurons l'erreur par

$$\text{RCEQM}(r) = \sqrt{\frac{1}{441} \sum_{t=1097}^{1537} e_{24(t-1)+11}^2(r)}.$$

Le tableau 6.6 donne une partie des résultats obtenus pour la comparaison. Généralement, le modèle 6 donne des meilleurs résultats dans les deux cas (1) et (2); la RCEQM est toujours plus petite. Pour un horizon très proche (une heure), ce n'est pas surprenant que les deux modèles ont des performances équivalentes, puisqu'avec les deux modèles on arrive à capter convenablement la corrélation entre les heures successives. Pour des horizons plus lointains, les deux modèles semblent aussi donner des résultats semblables, puisque la vraie corrélation n'est pas aussi grande dans ce cas-là. Par contre, pour les valeurs de r entre les deux (horizons de quelques heures seulement), le modèle 6 donne clairement des meilleurs résultats.

Par exemple, si nous sommes à 11h00 (donc $W_{t,11}$ étant observé) et nous voulons des prévisions pour le volume d'appels pour les heures entre 16h00 et 17h00 du même jour, le RCEQM est de 3.7 pour le modèle 7, alors qu'elle est de 2.3 pour le modèle 6. Nous avons constaté aussi l'importance d'utiliser l'information disponible pour le volume des premières heures de la journée dans les prévisions pour le volume des prochaines heures de la même journée. Par exemple, avec le modèle 6 à 11h00, pour les prévisions des volumes d'appels entre 16h00 et 17h00, le RCEQM est de 3.5 si nous ignorons les volumes reçus de minuit jusqu'à 11h00, et l'erreur se réduit à 2.3 en si on tient compte de cette information.

6.1.6 Conclusion

Nous avons considéré une variété de modèles de séries chronologiques pour estimer et calculer les prévisions pour le volume d'appels par jour et par tranche horaire au service SMU. Notons qu'en général, le volume d'appels dépend en grande partie du type d'activité des gens au cours de la journée. Ces gens peuvent être au travail, en train de faire la navette, en train de dormir ou de faire la fête et nos modèles essaient de capter ces influences en considérant l'effet de l'heure de la journée, le jour de la semaine, le mois de l'année, ainsi qu'un traitement spécial pour le jour de l'an et des journées du festival le plus important dans la ville que nous avons étudiée. Nous avons utilisé trois approches fondamentales pour le volume d'appel journaliers : la régression standard en ignorant les dépendances, la régression avec des résidus corrélés et une troisième approche (ARIMA doublement-saisonnier) qui tient compte d'une structure de dépendance et des effets saisonniers. Comme interprétation ordinaire de ces modèles : c'est que la première partie déterministe capte les composantes saisonnières et non-saisonnières et la deuxième partie stochastique (les erreurs) capte l'effet d'effets omis ou non-observables comme les

corrélations. Nous constatons qu'un modèle de la deuxième catégorie, qui inclut un sous-ensemble choisi des effets principaux et statistiquement significatifs pour les effets principaux et effets d'interaction, donne des meilleurs résultats pour des prévisions 1 ou 2 jours dans l'avenir. L'avantage de ce modèle sur le modèle de régression standard diminue quand on augmente l'horizon des prévisions et il disparaît à peu près à deux semaines d'horizon. Le modèle ARIMA doublement-saisonnier ne donne pas de bons résultats avec un horizon dépassant une semaine dans l'avenir (voir figure 6.13).

Pour le volume d'appels par heure, nous avons utilisé deux approches : une qui est construite autour de la distribution conditionnelle du volume par heure, connaissant le volume par jour et une autre approche basée sur un modèle de séries chronologiques. Les deux approches peuvent être combinées avec n'importe quel modèle que nous avons présenté pour le volume d'appels par jour. Nous avons aussi montré que le fait de prendre en considération les mises à jour des prévisions à partir du volume déjà observé au début de la journée pourrait être très utile pour les décisions d'affectation du personnel en temps réel. Nous avons donc proposé l'approche de distribution conditionnelle et nous avons constaté qu'elle donne des bons résultats et que le fait de mettre à jour les prévisions toutes les heures avec le volume observé dans la première partie de la journée peut améliorer considérablement les prévisions.

Les modèles que nous avons présentés sont simples, pratiques et pourraient être utilisés pour les prévisions de routine pour n'importe quel système SMU, aussi bien que dans les modèles de simulation de tels systèmes. Nos modèles qui combinent la régression et les processus ARIMA ont montré une supériorité sur le modèle ARIMA saisonnier. Nous avons aussi montré l'importance de modéliser les effets des journées spéciales, de l'heure de la journée, du jour de la semaine et du mois

de l'année. Bien que nous nous attendions à ce que l'approche générale décrite ici soit applicable dans d'autres villes et d'autres systèmes SMU, il est important de vérifier les spécificités et les caractéristiques de chaque ville. Il peut être utile d'étudier la possibilité d'ajouter d'autres variables au modèle de régression comme des variables démographiques par exemple.

6.2 Modèles de prévision pour les arrivées au centre C2

6.2.1 Les arrivées par jour

Nous procédons ici à une application de la classe de modèles présentés dans la section précédente avec les données du centre C2 de Bell Canada. Le but est de fournir un outil efficace et utile pour la planification des ressources humaines en agents pour une journée dans l'avenir et à l'intérieur même d'une journée.

Une première analyse des données à la section 4.1.2, nous a révélé une forte présence de l'effet du jour de la semaine. En tenant compte de ces considérations et de la présence d'une dépendance entre les journées successives, nous avons utilisé un modèle de régression à erreurs AR (modèle autorégressif) semblable au modèle donné à l'équation (6.9). Il faut noter que nous aurions espéré avoir une période couverte par les données assez longue pour pouvoir développer des modèles assez consistants pour les données journalières. Pour cela, nous ne pouvons pas considérer l'effet du mois de l'année, ainsi que les journées spéciales. Nous n'avons pas pu aussi constater aussi s'il y a une tendance à la hausse ou non dans cette série au fil des jours. Nous avons considéré des journées comme des données aberrantes et nous avons enlevé les trois jours qui correspondent au premier janvier, le vendredi saint (23 mars) et la fête de la reine (23 mai). Les données disponibles concernent donc un total de 170 jours. Nous avons utilisé les 17 premières semaines (117 jours) pour l'estimation des paramètres et nous avons gardé les 8 autres semaines (53 jours)

pour la validation des prévisions. Notre modèle se présente comme suit :

$$Y_t = \beta_0 + \sum_{j=1}^p \beta_j c_{t,j} + E_t, \quad (6.23)$$

où l'indicateur $c_{t,j}$ vaut 1 si l'observation t est dans le jour j de la semaine et 0 sinon ; les constantes réelles β_j représentent l'effet du jour de la semaine, vérifient la condition d'identifiabilité $\sum_{j=1}^p \beta_j = 0$. Le modèle peut aussi s'écrire comme suit :

$$Y_t = C_t' \beta + E_t, \quad (6.24)$$

où $C_t = (1, c_{t,1}, \dots, c_{t,p})'$ et $\beta = (\beta_0, \dots, \beta_p)'$.

Nous avons pu identifier un modèle AR(1) pour le processus E_t , donné par :

$$E_t = \phi_1 E_{t-1} + a_t, \quad (6.25)$$

où a_t est un bruit blanc de variance σ_a^2 . Le modèle en (6.23) devient :

$$(1 - \phi_1 B)(Y_t - \beta_0 + \sum_{j=1}^p \beta_j c_{t,j}) = a_t, \quad (6.26)$$

avec $BE_t = E_{t-1}$.

Avec une notation matricielle pour le modèle en (6.24), pour $t = 1, \dots, n$, nous avons :

$$Y = C\beta + E, \quad (6.27)$$

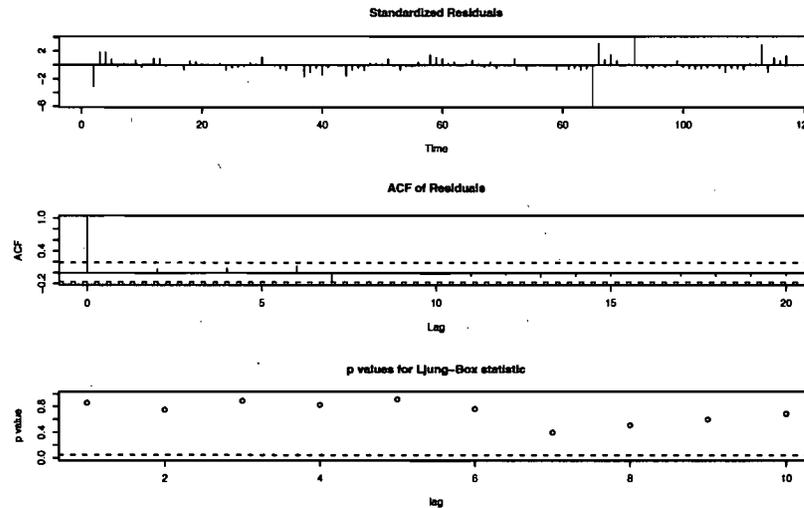


Figure 6.16 – Analyse des résidus pour le modèle des arrivées par jour pour la province du Québec.

où $Y = (y_1, \dots, y_n)'$, $E = (E_1, \dots, E_n)'$ et la matrice

$$C = (C_1, \dots, C_n)' = \begin{pmatrix} 1 & c_{1,1} & \cdots & c_{1,p} \\ 1 & c_{2,1} & \cdots & c_{2,p} \\ \vdots & \vdots & & \vdots \\ 1 & c_{n,1} & \cdots & c_{n,p} \end{pmatrix}$$

de dimension $(n, p + 1)$. Dans notre cas nous avons $n = 117$ et $p = 7$.

L'analyse des résidus dans les graphiques 6.16, 6.17 et 6.18 nous permettent de conclure que le modèle semble adéquat et valide pour les données disponibles. Dans la première partie du tableau 6.7, nous présentons les résultats des estimations sous le modèle donné en (6.26), pour les deux provinces du Québec et de l'Ontario, ainsi que pour l'ensemble.

La prévision pour le nombre d'appels à une origine t et à un horizon l , notée Y_{t+l} , est donnée par $\hat{Y}_t(l)$ et l'erreur de prévision est donnée par $e_t(l) = Y_{t+l} - \hat{Y}_t(l)$.

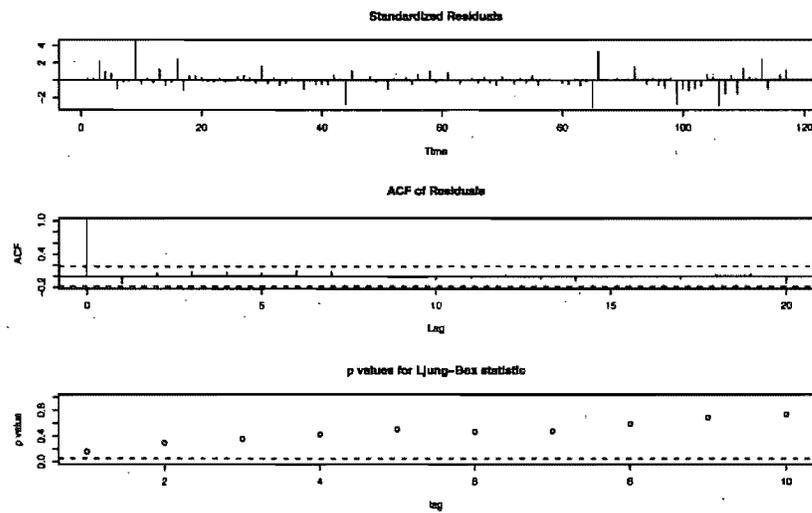


Figure 6.17 – Analyse des résidus pour le modèle des arrivées par jour pour la province de l'Ontario.

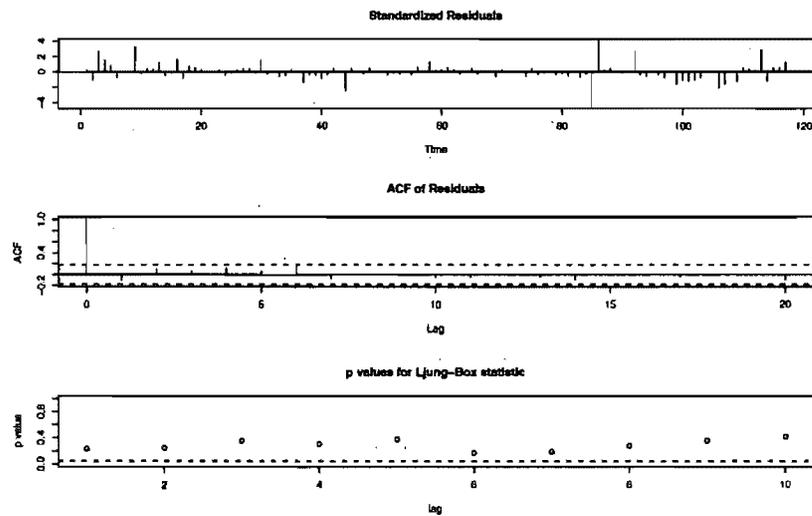


Figure 6.18 – Analyse des résidus pour le modèle des arrivées par jour pour l'ensemble des deux provinces.

La variance de l'erreur de prévision, notée $V(l)$, qui est la même que la variance de Y_{t+l} , est donnée par :

$$\begin{aligned}
 V(l) &= \text{Var}[e_t^2(l)] = \text{Var}(C'_{t+l}(\beta - \hat{\beta}) + E_{t+l}) \\
 &= \text{Var}(C'_{t+l}(\beta - \hat{\beta})) + \sigma_E^2 \\
 &= \sigma_E^2(C'_{t+l}(C'C)^{-1}C_{t+l} + 1) \\
 &= \frac{\sigma_a^2}{1 - \phi_1\rho_1}(C'_{t+l}(C'C)^{-1}C_{t+l} + 1), \quad (6.28)
 \end{aligned}$$

où $\rho_1 = \text{Corr}(E_t, E_{t-1})$. L'estimateur de cette variance est donné par :

$$\hat{V}(l) = \frac{\hat{\sigma}_a^2(C'_{t+l}(C'C)^{-1}C_{t+l} + 1)}{1 - \hat{\phi}_1\hat{\rho}_1}. \quad (6.29)$$

L'intervalle de prévision à un niveau $100(1 - \alpha)\%$ est donné par :

$$\hat{Y}_t(l) \pm t_{\alpha/2}(n - p - 1)\hat{\sigma}_a \left(\frac{C'_{t+l}(C'C)^{-1}C_{t+l} + 1}{1 - \hat{\phi}_1\hat{\rho}_1} \right)^{1/2}, \quad (6.30)$$

avec $t_{\alpha/2}(n - p - 1)$ le quantile d'ordre $(1 - \alpha/2)$ de la loi de Student avec $(n - p - 1)$ degrés de liberté.

Nous avons considéré les prévisions à un horizon d'un jour ($l = 1$) et nous avons calculé les mêmes mesures utilisées dans la section 6.1.4.6, pour évaluer les performances des prévisions, soit la *racine carrée de l'erreur quadratique moyenne* (RCEQM) et l'*erreur absolue relative moyenne* (EARM), définies comme suit :

$$\text{RCEQM}(1) = \sqrt{\frac{1}{53} \sum_{t=117}^{169} e_t^2(1)}$$

et

$$\text{EARM}(1) = \frac{1}{53} \sum_{t=117}^{169} \frac{|e_t(1)|}{Y_{t+1}}.$$

Tableau 6.7 – Résultats des estimations des paramètres et les erreurs de prévision pour le volume d'appels par jour avec le modèle où les erreurs E_t sont corrélées.

		Québec	Ontario	l'ensemble
Estimation (Erreur std.)	$\hat{\beta}_0$	18259.89 (596.07)	34451.06 (2151.14)	52688.64 (2419.38)
	$\hat{\beta}_1$	9416.49 (478.32)	18738.62 (1077.53)	28154.89 (1376.79)
	$\hat{\beta}_2$	5969.45 (478.31)	9785.91 (1078.72)	15750.90 (1377.43)
	$\hat{\beta}_3$	4467.93 (478.27)	7686.27 (1078.03)	12150.89 (1376.60)
	$\hat{\beta}_4$	3753.26 (477.75)	6361.27 (1075.15)	10108.82 (1373.34)
	$\hat{\beta}_5$	3947.99 (491.31)	6401.74 (1086.10)	10330.69 (1396.72)
	$\hat{\beta}_6$	-12412.72 (490.49)	-21251.14 (1081.75)	-33642.75 (1390.12)
	$\hat{\beta}_7$	-15142.40 (478.28)	-27722.66 (1073.64)	-42853.44 (1373.47)
	$\hat{\phi}_1$	0.274 (0.089)	0.680 (0.067)	0.573 (0.076)
	$\hat{\sigma}_a$	1896.28	3059.28	4563.53
	AIC	2116.26	2228.72	2322.08
	Prévisions	RCEQM(1)	2956.51	3837.03
EARM(1) (%)		9.84	6.97	7.13

Dans le bas du tableau 6.7, nous présentons les erreurs de prévisions RCEQM et EARM calculées à un horizon d'une seule journée.

Nous avons comparé ces mesures avec celles obtenues avec un modèle qui ne tient pas compte des corrélations entre le volume d'appels dans des journées successives, c.-à.-d., le modèle en (6.24) en supposant que les résidus E_t sont i.i.d. de distribution normale de moyenne 0 et de variance σ_E^2 . L'erreur standard ($\hat{\sigma}_E$) et les erreurs de prévision sont données dans le tableau 6.8.

Tableau 6.8 – Erreur standard et erreurs de prévision pour le volume d'appels par jour avec le modèle où les erreurs E_t sont i.i.d.

		Québec	Ontario	l'ensemble
Estimation	$\hat{\sigma}_E$	2033.44	4306.62	5739.81
Prévisions	RCEQM(1)	3017.90	4094.01	6365.34
	EARM(1) (%)	11.39	11.90	10.78

Nous pouvons remarquer facilement que le modèle qui capte la corrélation entre les volumes d'appels dans deux journées successives donne de meilleurs résultats. La RCEQM, par exemple, est réduite de 2.1% pour le Québec, de 6.3% pour l'Ontario et de 5.2% pour l'ensemble. Nous remarquons aussi une réduction au niveau de la EARM, ainsi qu'une réduction au niveau de l'estimation quand on compare l'erreur standard d'estimation ($\hat{\sigma}_a$ et $\hat{\sigma}_E$). Rappelons qu'en moyenne le centre reçoit près de 18455 appels en provenance du Québec, près de 34226 de l'Ontario et 52681 de l'ensemble, sachant que ces moyennes prennent en considération les sept jours de la semaine.

Dans les graphiques 6.19, 6.20 et 6.21, nous présentons les prévisions de Y_{t+1} , $t = 117, \dots, 169$, qui sont données par $\hat{Y}_t(1)$, $t = 117, \dots, 169$, c.-à.-d. les prévisions d'une seule journée à la fois, dans le futur, à 53 origines différentes. En examinant ces graphiques, nous ne remarquons rien d'anormal dans les prévisions, à part 2 ou 3 observations réelles qui dépassent les limites de prévision et qui représentent moins de 5%. Nous remarquons aussi que le modèle fonctionne bien pour les arrivées dans chaque province et pour l'ensemble des deux.

6.2.2 Les arrivées par intervalles de temps dans une journée

Pour les arrivées par période de temps dans la journée, nous avons considéré des périodes de 15 minutes et nous avons utilisé le modèle présenté dans la section

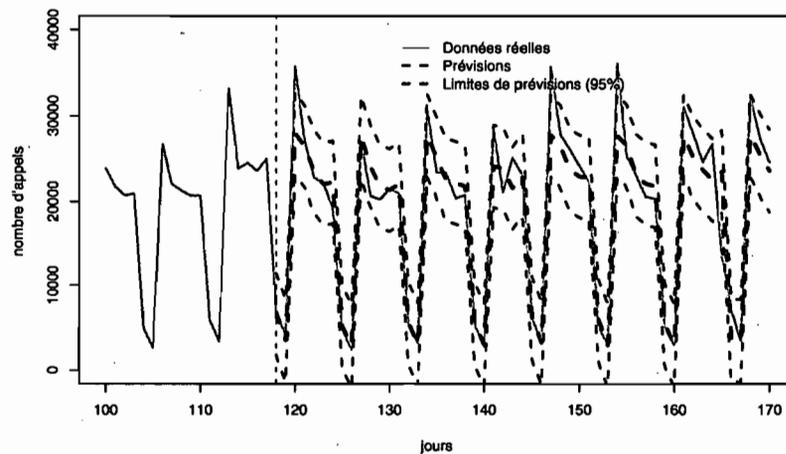


Figure 6.19 – Prévisions et limites de confiance (95%) à un horizon d'une journée pour les arrivées par jour pour la province du Québec.

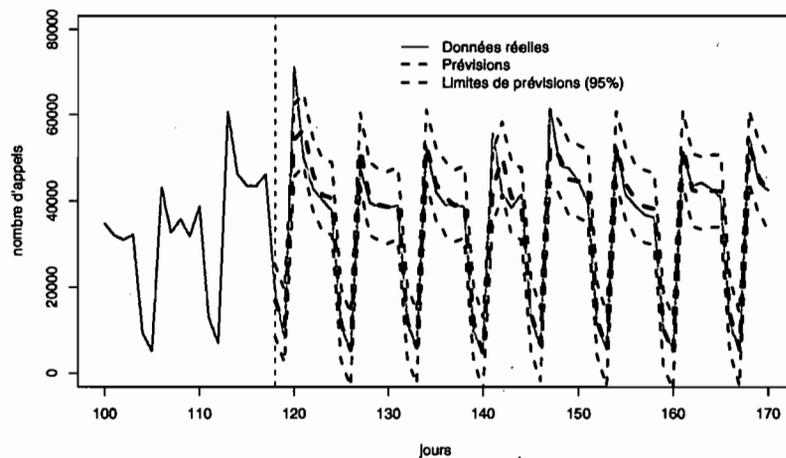


Figure 6.20 – Prévisions et limites de confiance (95%) à un horizon d'une journée pour les arrivées par jour pour la province de l'Ontario.

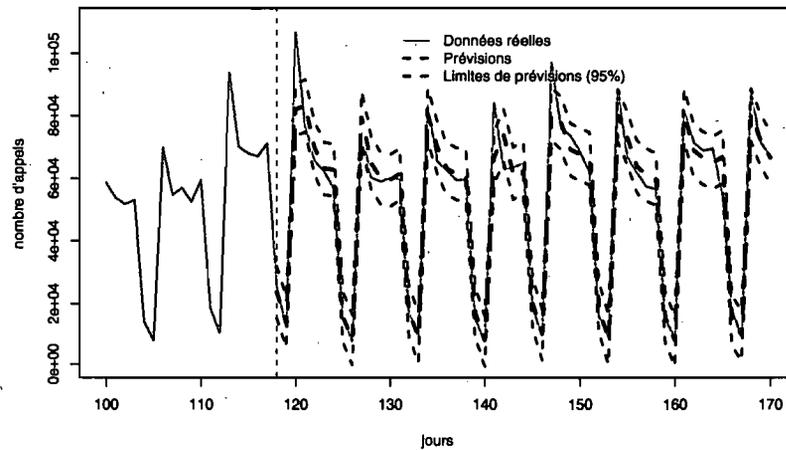


Figure 6.21 – Prévisions et limites de confiance (95%) à un horizon d'une journée pour les arrivées par jour pour l'ensemble des deux provinces.

6.1.5.1, où le vecteur $\mathbf{Z}_t = (Z_{48(t-1)+1}, \dots, Z_{48t})$, sachant $Y_t = N$, suit une distribution multinomiale de paramètres (N, p_1, \dots, p_{48}) , où le vecteur (p_1, \dots, p_{48}) , qui représente le profil du jour, vérifie l'équation (6.15). Chaque Z_h représente donc le nombre d'appels reçus durant le quart d'heure $h = 48(t-1) + 1, \dots, 48t$.

Pour les prévisions, nous appelons $e_h(r) = Z_{h+r} - \hat{Z}_h(r)$, l'erreur de prévision, où $\hat{Z}_h(r)$ désigne la prévision de Z_{h+r} selon notre modèle. Nous utilisons le fait que N est assez grand et avec le théorème de la limite centrale, nous pouvons calculer un intervalle de prévision approximatif pour Z_{h+r} . Alors l'intervalle à un niveau $100(1 - \alpha)\%$ est donné par :

$$\hat{Z}_h(r) \pm Z_{\alpha/2}(Np_h(1 - p_h))^{1/2}, \quad (6.31)$$

avec $Z_{\alpha/2}$ le quantile d'ordre $(1 - \alpha/2)$ de la loi normale standard.

Nous avons exploité le modèle 6 avec la distribution conditionnelle des arrivées connaissant le nombre total des arrivées dans la journée. Nous avons considéré donc

le cas où nous observons une partie de la journée et les prévisions des heures qui restent sont calculées en utilisant la prévision mise à jour de la journée calculée à partir de la formule de Bayes, présentée en (6.21). Supposons que nous disposons de la prévision $\hat{Y}_{t-1}(1)$ du volume d'appels, obtenues avec le modèle en (6.26), durant la journée entière t , faite à la veille, c.-à.-d., à la fin de la journée $(t-1)$, et que nous avons observé le nombre d'arrivées jusqu'à 11h00, noté $W_{t,11} = \sum_{i=1}^{12} Z_{48(t-1)+i}$. Nous devons donc calculer la prévision du volume d'appels pour la partie restante de la journée de 11h00 à 20h00, c.-à.-d. la quantité $Y_t - W_{t,11} = \sum_{i=13}^{48} Z_{48(t-1)+i}$, en utilisant $W_{t,11}$. Nous considérons cette fois-ci l'erreur de prévisions comme suit :

$$e_{t,w} = \sum_{i=13}^{48} Z_{48(t-1)+i} - \hat{p}_i \hat{Y}_{t-1}(1), \quad (6.32)$$

où les $\hat{p}_i, i = 1, \dots, 48$, représentent une estimation du profil du jour à partir des données selon la formule donnée en (6.15). Pour le modèle qui tient compte de l'information du matin ($W_{t,11}$), la prévision $\hat{Y}_{t-1}(1)$ est remplacée par sa valeur mise à jour en utilisant la formule donnée en (6.21). Pour mesurer l'erreur de prévision, nous avons utilisé la *racine carrée de l'erreur quadratique moyenne* (RCEQM) et l'*erreur absolue relative moyenne* (EARM), définies comme suit :

$$\text{RCEQM} = \sqrt{\frac{1}{53} \sum_{t=118}^{170} e_{t,w}^2}$$

et

$$\text{EARM} = \frac{1}{53} \sum_{t=118}^{170} \frac{|e_{t,w}|}{Y_t - W_{t,11}}$$

Dans le tableau 6.9 nous présentons les erreurs de prévision RCEQM et EARM calculées pour le nombre d'arrivées pour toute la période de 11h00 à 20h00 sur les 53 journées utilisées pour les prévisions. Nous comparons ces erreurs avec le modèle

Tableau 6.9 – Erreurs de prévisions pour le nombre d'arrivées entre 11h00 et 20h00.

		Québec	Ontario	l'ensemble
Modèle avec mise à jour de $\hat{Y}_{t-1}(1)$	RCEQM	1407.69	1932.46	3012.91
	EARM (%)	8.36	7.53	6.97
Modèle sans mise à jour	RCEQM	1959.05	2594.68	3905.11
	EARM (%)	11.25	8.67	8.45

qui ne tient pas compte de l'information du matin, c.-à.-d. sans mise à jour de la prévision du volume journalier faite à minuit. De point de vue de la RCEQM, par exemple, nous avons observé une réduction de près de 28.1% pour la province du Québec, 26.6% pour l'Ontario et 22.8% pour l'ensemble. Rappelons qu'en moyenne le centre reçoit près de 12318 appels en provenance du Québec entre 11h00 et 20h00, 22927 de l'Ontario et 35245 pour l'ensemble pour la même période de temps. Le calcul de ces moyennes tient compte des sept jours de la semaine.

Dans les graphiques 6.22, 6.23 et 6.24, nous présentons le résultat des prévisions pour une seule journée choisie pour illustration seulement. Nous avons considéré une journée du type jeudi et nous avons calculé les prévisions du nombre d'appels reçus durant la période de 11h00 à 20h00 en utilisant une mise à jour de la prévision du volume journalier en fonction du volume déjà observé de 8h00 à 11h00. Nous remarquons que le modèle fournit des bons résultats pour les arrivées par quarts d'heure en utilisant l'information de la matinée. Les résultats des prévisions montrent aussi des erreurs minimales et presque toutes les observations réelles sont à l'intérieur des limites de prévision.

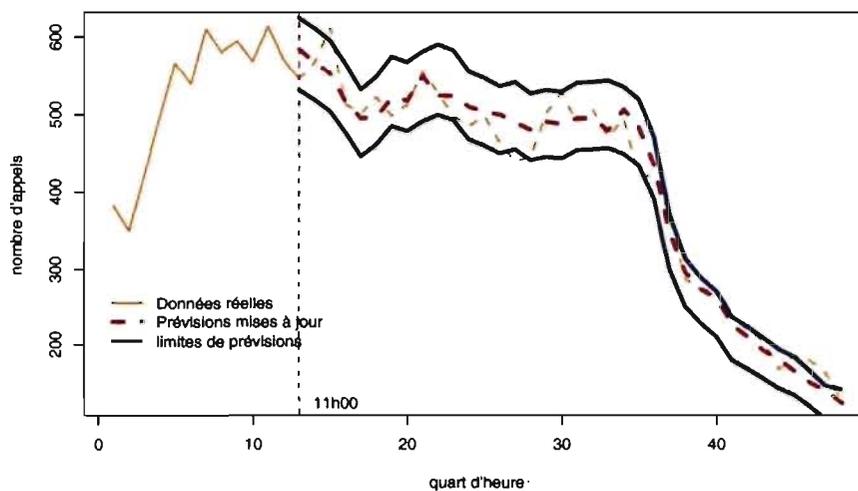


Figure 6.22 – Exemple de prévisions avec mise à jour et limites de confiance (95%) dans une journée du type jeudi pour les arrivées par 15 minutes entre 11h00 et 20h00, pour la province du Québec.

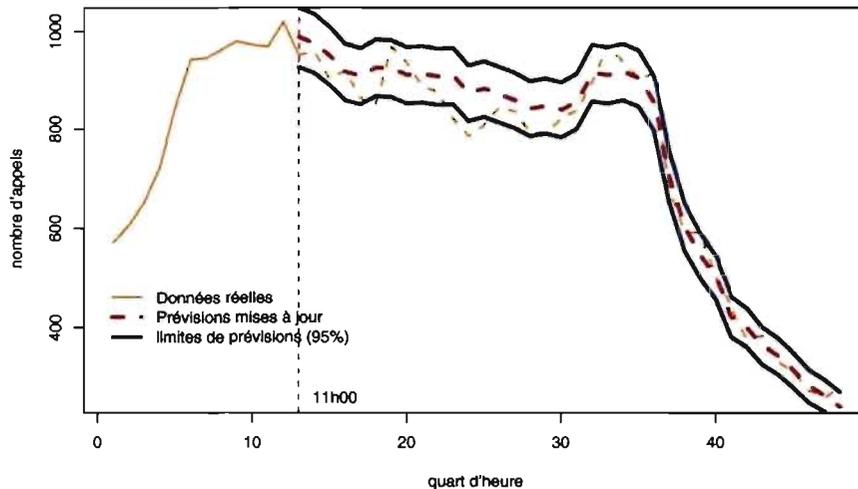


Figure 6.23 – Exemple de prévisions avec mise à jour et limites de confiance (95%) dans une journée du type jeudi pour les arrivées par 15 minutes entre 11h00 et 20h00, pour la province de l'Ontario.

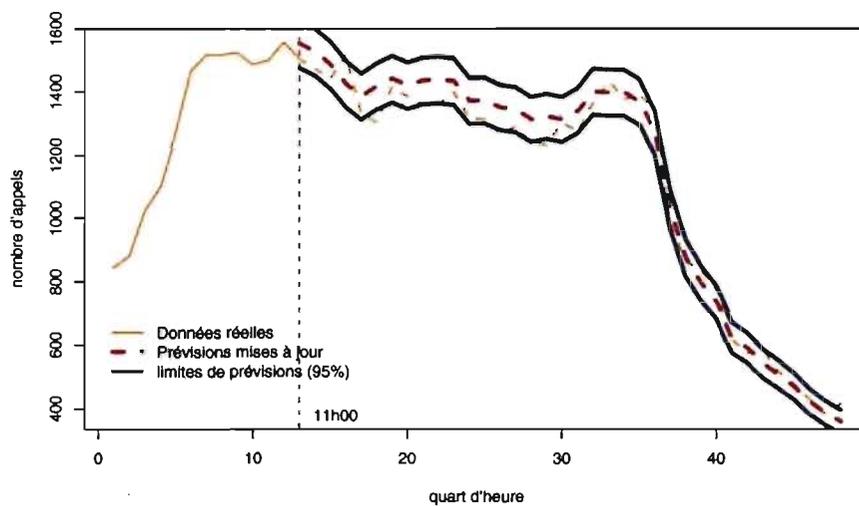


Figure 6.24 – Exemple de prévisions avec mise à jour et limites de confiance (95%) dans une journée du type jeudi pour les arrivées par 15 minutes entre 11h00 et 20h00, pour l'ensemble des deux provinces.

CHAPITRE 7

CONCLUSION

Dans cette thèse nous nous sommes intéressés à une composante très importante du système de file d'attente que constitue tout centre d'appels et qui représente le processus d'arrivée. Nous avons pu présenter une variété de modèles pour le processus d'arrivée avec des hypothèses assez réalistes et qui décrivent bien la réalité dans les centres d'appels téléphoniques étudiés. Ces modèles ont été validés et testés avec 4 jeux de données qui concernent 4 centres d'appels et nous avons obtenu des bons résultats.

L'analyse descriptive des données dans les centres étudiés nous ont révélé des effets similaires. Les arrivées diffèrent selon la période de la journée, selon le type du jour de la semaine et le mois de l'année aussi. Nous avons aussi observé la présence de corrélations entre le nombre d'arrivées durant différentes périodes d'une même journée.

Nous avons pu utiliser le processus de Poisson avec une forme particulière qui permet de capter la dépendance à l'intérieur d'une journée. A ce modèle, nous avons ajouté une amélioration pour la fonction de taux avec une estimation non paramétrique à l'aide d'une fonction spline. Nous avons obtenus des bons résultats d'estimation avec ces modèles.

Avec le modèle NORTA, nous avons eu beaucoup de succès et nous avons pu faire des bons ajustements. Ce type de modèle permet de faire des simulation en générant un vecteur d'arrivées dans une journée avec exactement la même structure de corrélation désirée observée dans la réalité. Nous avons pu apporter des améliorations à ce type de modèle qui présente quelques difficultés pratiques,

comme la réduction du nombre de paramètres à estimer, ainsi que la difficulté d'obtenir une matrice semi-définie positive. Nous avons pu développer de nouvelles méthodes précises et très rapides en temps d'exécution pour l'initialisation de la méthode NORTA. Ces méthodes sont valables avec les corrélations linéaires et les corrélations de rang et dans le cas où les distributions marginales sont discrètes. Nous avons pu proposer une méthode semblable dans le cas où nous disposons d'un mélange de distributions continues et de distributions discrètes. Lorsque les distributions marginales sont continues, il existe une résolution analytique simple en utilisant les corrélations de rang. Nous avons testé ces méthodes dans le contexte des arrivées dans les centres d'appels téléphoniques. Ces travaux ont permis la publication d'un article.

Nous avons aussi pu renforcer les bibliothèques de simulation déjà en place avec ces nouveaux modèles. Désormais, avec les bibliothèques de simulation en Java *SSJ* [57] et *ContactCenters* [15], on peut facilement générer les arrivées selon le processus NORTA ou selon le processus de Poisson non homogène ou doublement stochastique avec une fonction spline pour le taux, en utilisant la méthode de découpage. Toutes les méthodes que nous avons proposées dans cette thèse pour l'initialisation de NORTA sont aussi incorporées sous forme de classes dans la bibliothèque *SSJ* (voir annexes).

Une autre contribution qui a permis aussi de publier un article, est celle des modèles de séries chronologiques pour les arrivées journalières et par tranche horaire. Nous avons pu présenter des modèles de prévisions dans le centre d'appels au 911 dans la ville de Calgary, en Alberta, pour les journées et les heures dans une même journée. Ces modèles peuvent être adaptés à d'autres centres semblables.

Plusieurs alternatives de travaux futurs méritent d'être étudiés. Nous croyons qu'il est intéressant par exemple d'étudier d'avantage les méthodes qui permettront

d'améliorer la méthode NORTA et de diminuer ses difficultés pratiques dans des grandes dimensions. Il faudrait chercher d'autres modèles qui s'ajustent mieux avec la nature des corrélations entre les arrivées des parties d'une journée qui permettent de réduire le nombre de paramètres à estimer, et en même temps garantissent une matrice semi-définie positive. Une autre avenue de recherche qui consiste à combiner le processus NORTA avec la fonction spline. Le premier modèle s'avère efficace quand à la génération du nombre d'arrivées avec les quantités désirées de corrélations, alors qu'une fonction spline pour la fonction de taux permet générer les instants d'arrivées de façon efficace.

Un sujet très important qui a été laissé à côté dans ce travail et qui pourrait être privilégié dans des travaux futurs, est celui de l'étude du cas où il y a plusieurs types d'appels à distinguer, ce qui veut dire considérer plusieurs processus d'arrivée. Dans la réalité, ces processus sont généralement dépendants, non stationnaires et même doublement stochastiques, ce qui complique la tâche de trouver des modèles adéquats et réalistes pour ces processus. Ces propriétés ne sont pas toujours prises en compte, par exemple, dans les modèles de simulation des centres d'appels téléphoniques. Nous croyons donc qu'il serait très intéressant d'étudier les possibilités de modèles et de voir en même temps les aspects estimation et prévision dans de telles situations.

BIBLIOGRAPHIE

- [1] B. Abraham et J. Ledolter. *Statistical Methods for Forecasting*. Wiley, Toronto, 1983.
- [2] H. Akaike. Information theory and an extension of the maximum likelihood principle. Dans B. N. Petrov et F. Csaki, éditeurs, *Proc. 2nd International Symposium on Information Theory, Akademiai Kiado, Budapest*, pages 267–281, 1973.
- [3] B. Andrews et S. M. Cunningham. L.L. Bean improves call-center forecasting. *Interfaces*, 25, pages 1–13, 1995.
- [4] A. N. Avramidis, N. Channouf et P. L'Ecuyer. Efficient correlation matching for fitting discrete multivariate distributions with arbitrary marginals and normal copula dependence. *INFORMS Journal on Computing*, 2008. À paraître.
- [5] A. N. Avramidis, A. Deslauriers et P. L'Ecuyer. Modeling daily arrivals to a telephone call center. *Management Science*, 50(7), pages 896–908, 2004.
- [6] J. R. Baker et K. E. Fitzpatrick. Determination of an optimal forecast model for ambulance demand using goal programming. *Journal of the Operational Research Society*, 37(11), pages 1047–1059, 1986.
- [7] W. R. Bell et S. C. Hillmer. Modeling time series with calendar variation. *Journal of the American Statistical Association*, 78(383), pages 526–534, 1983.
- [8] L. Bianchi, J. Jarrett et R. Ch. Hanumara. Improving forecasting for telemarketing centers by ARIMA modeling with intervention. *International Journal of Forecasting*, 14, pages 497–504, 1998.

- [9] G. E. P. Box, G. M. Jenkins et G. C. Reinsel. *Time Series Analysis, Forecasting, and Control*. Prentice-Hall, Englewood Cliffs, New Jersey, troisième édition, 1994.
- [10] R. P. Brent. *Algorithms for Minimization without Derivatives*. Prentice-Hall, Englewood Cliffs, N.J., 1976.
- [11] D. Brigo. A note on correlation and rank reduction. Rapport technique, Product and Business Development Group, Banca IMI, San Paolo-IMI Group, Milan, Italie, 2002.
- [12] P. J. Brockwell et R. A. Davis. *Time Series : Theory and Methods*. Springer-Verlag, New York, deuxième édition, 1991.
- [13] L. Brown, N. Gans, A. Mandelbaum, A. Sakov, H. Shen, S. Zeltyn et L. Zhao. Statistical analysis of a telephone call center : A queueing-science perspective. *Journal of the American Statistical Association*, 100, pages 36–50, 2005.
- [14] E. Buffa, M. Cosgrove et B. Luce. An integrated work shift scheduling system. *Decision Sciences*, 7, pages 620–630, 1976.
- [15] E. Buist et P. L'Ecuyer. *ContactCenters : A Java Library for Simulating Contact Centers*, 2005. Software user's guide, disponible à <http://www.ericbuist.com/contactcenters>.
- [16] R. L. Burden et J. D. Faires. *Numerical Analysis*. Brooks/Cole Publishing Company, California, sixième édition, 1997.
- [17] M. C. Cario et B. L. Nelson. Autoregressive to anything : Time series input processes for simulation. *Operations Research Letters*, 19, pages 51–58, 1996.

- [18] M. C. Cario et B. L. Nelson. Numerical methods for fitting and simulating autoregressive-to-anything processes. *INFORMS Journal on Computing*, 10, pages 72–81, 1998.
- [19] N. Channouf, P. L'Ecuyer, A. Ingolfsson et A. N. Avramidis. The application of forecasting techniques to modeling emergency medical system calls in Calgary, Alberta. *Health Care Management Science*, 10(1), pages 25–45, 2007.
- [20] H. Chen. Initialisation for NORTA : Generation of random vectors with specified marginals and correlations. *INFORMS Journal on Computing*, 13, pages 312–331, 2000.
- [21] S. D. Conte et C. de Boor. *Elementary Numerical Analysis : An Algorithmic Approach*. McGraw-Hill, deuxième édition, 1972.
- [22] J. M. Davenport et R. L. Iman. An iterative algorithm to produce a positive definite correlation matrix from an "approximate correlation matrix". Rapport technique, Sandia National Laboratories, Albuquerque, Mexico, 1982.
- [23] C. de Boor. *A Practical Guide to Splines*. Numéro 27 dans Applied Mathematical Sciences Series. Springer-Verlag, New York, 1978.
- [24] B. Deler et B. L. Nelson. Modeling and generating multivariate time series with arbitrary marginals using a vector autoregressive technique. Rapport technique, Northwestern University, 2000.
- [25] A. Deslauriers. Modélisation et simulation d'un centre d'appels téléphoniques dans un environnement mixte. Mémoire de maîtrise, Département d'Informatique et de Recherche Opérationnelle, Université de Montréal, Montréal, Canada, 2003.

- [26] T. G. Donnelly. Algorithm 462 : Bivariate normal distribution. *Communications of the ACM*, 16(10), page 638, 1973.
- [27] M. Eckstein et L. S. Chan. The effect of emergency department crowding on paramedic ambulance availability. *Annals of Emergency Medicine*, 43, pages 100–105, 2004.
- [28] M. Eckstein, S. M. Isaacs, C. M. Slovis, B. J. Kaufman, J. R. Loflin, R. E. O'Connor et P. E. Pepe. Facilitating ems turnaround intervals at hospitals in the face of receiving facility overcrowding. *Prehospital Emergency Care*, 9, pages 267–275, 2005.
- [29] G. Erdogan, E. Erkut et A. Ingolfsson. Ambulance deployment for maximum survival. Rapport technique, School of Business, University of Alberta, 2006. Disponible à <http://www.bus.ualberta.ca/aingolfsson/publications>.
- [30] M. T. Flanagan. *Michael Thomas Flanagan's Scientific Java Library*, 2007. Integration Class : numerical integration. Disponible à <http://www.ee.ucl.ac.uk/~mflanaga>.
- [31] N. Gans, G. Koole et A. Mandelbaum. Telephone call centers : Tutorial, review, and research prospects. *Manufacturing and Service Operations Management*, 5, pages 79–141, 2003.
- [32] S. Ghosh et S. G. Henderson. Behavior of the norta method for correlated random vector generation as the dimension increases. *ACM Transactions on Modeling and Computer Simulation*, 13(3), pages 276–294, 2001.
- [33] J. B. Goldberg. Operations research models for the deployment of emergency service vehicles. *EMS Management Journal*, 1, pages 20–39, 2004.

- [34] L. V. Green et P. J. Kolesar. Improving emergency responsiveness with management science. *Management Science*, 50, pages 1001–1014, 2004.
- [35] L. V. Green et P.J. Kolesar. The pointwise stationary approximation for queues with nonstationary arrivals. *Management Science*, 37(1), pages 84–97, 1991.
- [36] L. V. Green et P.J. Kolesar. The lagged PSA for estimating peak congestion in multiserver Markovian queues with periodic arrival rates. *Management Science*, 43, pages 80–87, 1997.
- [37] E. Gunes et R. Szechtman. A simulation model of a helicopter ambulance service. Dans *Proceedings of the 2005 Winter Simulation Conference*, pages 951–957. IEEE Press, 2005.
- [38] S. G. Henderson. Estimation of nonhomogeneous Poisson processes from aggregated data. *Operations Research Letters*, 31, pages 375–382, 2003.
- [39] S. G. Henderson. Should we model dependence and nonstationarity, and if so, how? Dans *Proceedings of the 2005 Winter Simulation Conference*, pages 120–129. IEEE Press, 2005.
- [40] S. G. Henderson et A. J. Mason. Ambulance service planning : simulation and data visualization. Dans F. Sainfort, M. L. Brandeau et W. P. Pierskalla, éditeurs, *Handbook of Operations Research and Health Care Methods and Applications*, pages 77–102. Kluwer Academic, Boston, 2004.
- [41] R. V. Hogg et A. F. Craig. *Introduction to Mathematical Statistics*. Prentice-Hall, cinquième édition, 1995.

- [42] W. Hörmann, J. Leydold et G. Derflinger. *Automatic Nonuniform Random Variate Generation*. Springer-Verlag, Berlin, 2004.
- [43] A. Ingolfsson, E. Erkut et S. Budge. Simulation of single start station for Edmonton EMS. *Journal of the Operational Research Society*, 54, pages 736–746, 2003.
- [44] H. Joe. *Multivariate Models and Dependence Concepts*. Chapman and Hall, London, 1997.
- [45] N. L. Johnson et S. Kotz. *Distributions in Statistics : Discrete Distributions*. Houghton Mifflin, Boston, 1969.
- [46] G. Jongbloed et G. Koole. Managing uncertainty in call centers using Poisson mixtures. *Applied Stochastic Models in Business and Industry*, 17, pages 307–318, 2001.
- [47] R. Kamenetsky, L. Shuman et H. Wolfe. Estimating need and demand for prehospital care. *Operations Research*, 30, pages 1148–1167, 1982.
- [48] M. Kendall et A. Stuart. *The advanced theory of statistics*, volume 1 : Distribution theory. Macmillan, New York, NY, 4ème édition, 1977.
- [49] W. Kruskal. Ordinal measures of association. *Journal of the American Statistical Association*, 53, pages 814–861, 1958.
- [50] M. E. Kuhl et P. S. Bhairgond. Nonparametric estimation of nonhomogeneous poisson process using wavelets. Dans J.A. Joines, R.R. Barton, K. Kang et P.A. Fishwick, éditeurs, *Proceedings of the 2000 Winter Simulation Conference*, pages 562–571. IEEE Press, 2000.

- [51] M. E. Kuhl et H. Damerджи. Estimating and simulating poisson processes with trends or asymmetric effects. Dans S. Andradóttir, K. J. Healy, D. H. Withers et B. L. Nelson, éditeurs, *Proceedings of the 1997 Winter Simulation Conference*, pages 287–295. IEEE Press, 1997.
- [52] D. Kurowicka et R. M. Cooke. Conditional, partial and rank correlation for the elliptical copula ; dependence modelling in uncertainty analysis. Dans E. Zio, M. Demichela et N. Piccinini, éditeurs, *Safety and Reliability, 3rd volume*, pages 1795–1802. 2001.
- [53] R. C. Larson. A hypercube queuing model for facility location and redistricting in urban emergency services. *Computers & Operations Research*, 1(1), pages 67–95, 1974.
- [54] R. C. Larson. Approximating the performance of urban emergency service systems. *Operations Research*, 23(5), pages 845–868, 1975.
- [55] A. M. Law et W. D. Kelton. *Simulation Modeling and Analysis*. McGraw-Hill, New York, NY, troisième édition, 2000.
- [56] P. L'Ecuyer. Random number generation and quasi-Monte Carlo. Dans J. Teugels et B. Sundt, éditeurs, *Encyclopedia of Actuarial Science*, volume 3. John Wiley, Chichester, UK, 2004.
- [57] P. L'Ecuyer. *SSJ : A Java Library for Stochastic Simulation*, 2004. Software user's guide. Disponible à <http://www.iro.umontreal.ca/~lecuyer>.
- [58] L. M. Leemis. Nonparametric estimation of the cumulative intensity function for a nonhomogeneous poisson process. *Management Science*, 37(7), pages 886–900, July 1991.

- [59] P. A. W. Lewis et G. S. Shedler. Statistical analysis of non-stationary series of events un data base system. *IBM Journal of research and development*, 20, pages 465–482, 1976.
- [60] P. A. W. Lewis et G. S. Shedler. Simulation of nonhomogeneous Poisson processes by thinning. *Nav. Res. Logistics Quart.*, 26, pages 403–414, 1979.
- [61] Ph. M. Lurie et M. S. Goldberg. Approximate method for sampling correlated random variables from partially specified distributions. *Management Science*, 44, pages 203–218, 2000.
- [62] V. A. Mabert. Short interval forecasting of emergency phone call (911) work loads. *Journal of Operations Management*, 5(3), pages 259–271, 1985.
- [63] A. Mandelbaum, A. Sakov et S. Zeltyn. Empirical analysis of call center, 2000. Disponible à <http://iew3.technion.ac.il/serveng/References/references.html>.
- [64] W. A. Massey, G. A. Parker et W. Whitt. Estimating the parameters of a non-homogeneous Poisson process with linear rate. *Telecommunication Systems—Modeling, Analysis, Design and Management*, 5(4), pages 361–388, 1995.
- [65] C. E. McConnell et R. W. Wilson. The demand for prehospital emergency services in an aging society. *Social Science & Medicine*, 46(8), pages 1027–1031, 1998.
- [66] V. Mehrotra. Ringing up big business. *ORMS Today*, 24(4), pages 18–24, Octobre 1997.
- [67] A. Moeller. Obstacles to measuring emergency medical service performance. *EMS Management Journal*, 1, pages 8–15, 2004.

- [68] National Fire Protection Association. NFPA 1221 : Standard for the installation, maintenance, and use of emergency service communication systems, 2002.
- [69] R. B. Nelsen. *An Introduction to Copulas*, volume 139 de *Lecture Notes in Statistics*. Springer-Verlag, New York, 1999.
- [70] D. Nora. Délocalisation après les usines, les services? *Le Nouvel Observateur*, Février 2005.
- [71] J. Pichitlamken, A. Deslauriers, P. L'Ecuyer et A. N. Avramidis. Modeling and simulation of a telephone call center. Dans S. Chick, P. J. Sánchez, D. Ferrin et D. J. Morrice, éditeurs, *Proceedings of the 2003 Winter Simulation Conference*, pages 1805–1812. IEEE Press, 2003.
- [72] D. A. Pierce. Least squares estimation in the regression model with autoregressive-moving average errors. *Biometrika*, 58(2), pages 299–312, 1971.
- [73] W. H. Press, S. A. Teukolsky, W. T. Vetterling et B. P. Flannery. *Numerical Recipes in C : The Art of Scientific Computing*. Cambridge University Press, Cambridge, troisième édition, 1992.
- [74] T. R. C. Read et N. A. C. Cressie. *Goodness-of-Fit Statistics for Discrete Multivariate Data*. Springer Series in Statistics. Springer-Verlag, New York, 1988.
- [75] S. Resnick. *Adventures in Stochastic Processes*. Boston, 1992.
- [76] T. R. Robbins, D. J. Medeiros et P. Dum. Evaluating arrival rate uncertainty in call centers. Dans L. F. Perrone, F. P. Wieland, J. Liu, B. G. Lawson, D. M.

- Nicol et R. M. Fujimoto, éditeurs, *Proceedings of the 2006 Winter Simulation Conference*, pages 2180–2187. IEEE Press, 2006.
- [77] M. H. Schultz. *Spline analysis*. Prentice-Hall, Englewood Cliffs, New Jersey, 1973.
- [78] G. A. F. Seber. *Linear Regression Analysis*. John Wiley, New York, 1977.
- [79] W. Segal, V. Verter, A. Colacone et M. Afilalo. The in-hospital interval : A description of emt time spent in the emergency department. *Prehospital Emergency Care*, 10, pages 378–382, 2006.
- [80] H. Shen et J. Z. Huang. Interday forecasting and intraday updating of call center arrivals. *Manufacturing & Service Operations Management*, 2007. À paraître.
- [81] H. Shen et J. Z. Huang. Analysis of call center arrival data using singular value decomposition. *Applied Stochastic Models in Business and Industry*, 21, pages 251–263, 2005.
- [82] B. Silverman. *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, London, 1986.
- [83] P. Sprivulis et B. Gerrard. Internet-accessible emergency department workload information reduces ambulance diversion. *Prehospital Emergency Care*, 9, pages 285–291, 2005.
- [84] S. G. Steckley et S. G. Henderson. The error in steady-state approximations for the time-dependent waiting time distribution. *Stochastic Models*, 23(2), pages 307–332, 2007.

- [85] S. G. Steckley, S. G. Henderson et V. Mehrotra. Forecast errors in service systems. *Probability in the Engineering and Informational Sciences*, 2007. À paraître.
- [86] S. G. Steckley, S. G. Henderson et V. Mehrotra. Performance measures for service systems with a random arrival rate. Dans M. E. Kuhl, N. M. Steiger, F. B. Armstrong et J. A. Joines, éditeurs, *Proceedings of the 2005 Winter Simulation Conference*; pages 2180–2187. IEEE Press, 2005.
- [87] A. J. Swersey. The deployment of police, fire and emergency medical units. Dans S. M. Pollock, M. Rothkopf et A. Barnett, éditeurs, *Handbooks in Operations Research and Management Science*, volume 6, pages 151–190. North-Holland, New York, 1994.
- [88] V. Trofimov, P. D. Feigin, A. Mandelbaum, E. Ishay et E. Nadjharov. Data model for call center analysis. Volume 1 : Model description and introduction to user interface. Rapport technique, Technion-Israel Institute of Technology, Israel, 2004.
- [89] W. Tych, D. J. Pedregal, P. C. Young et J. Davies. An unobserved component model for multi-rate forecasting of telephone call demand : The design of a forecasting support system. *International Journal of Forecasting*, 18, pages 673–695, 2002.
- [90] L. Uchitelle. Answering '800' calls, extra income but not security. *The New York Times*, Mars 27, page 1, 2002.
- [91] W. W-S. Wei. *Time Series Analysis : Univariate and Multivariate Methods*. Addison-Wesley, New York, 1990.

- [92] J. Weinberg, L. D. Brown et J. R. Stroud. Bayesian forecasting of an inhomogeneous Poisson process with applications to call center data. *Journal of the American Statistical Association*, 102(480), pages 1185–1198, 2007.
- [93] W. Whitt. Dynamic staffing in a telephone call center aiming to immediately answer all calls. *Operations Research Letters*, 24, pages 205–212, 1999.
- [94] J. K. Yarnold. The minimum expectation in χ^2 goodness-of-fit tests and the accuracy of approximations for the null distribution. *Journal of the American Statistical Association*, 65, pages 864–886, 1970.
- [95] Z. Zhu, M. A. McKnew et J. Lee. Effects of time-varied arrival rates : an investigation in emergency ambulance service systems. Dans *Proceedings of the 1992 Winter Simulation Conference*, pages 1180–1186. IEEE Press, 1992.

Annexe I

Classes Java pour l'initialisation de NORTA avec les corrélations de rang et des distributions marginales discrètes

I.1

**Java classes for NORTA initialization with rank correlations when the
marginal distributions are discrete**

User's guide with an example
(version of 07 March 2008)

Nabil Channouf

Introduction

This document describes a Java implementation of correlation matching algorithms proposed in [1] for the situation where one wants to use the NORTA method to fit a multivariate distribution with discrete marginals. The four different algorithms discussed in [1] are implemented in four subclasses of an abstract class named `NortaInitDisc`. This software makes use of the SSJ library [2]. An example of how to use it is given at the end of this document.

The NORTA method is an approach for modeling dependence in a finite-dimensional random vector $X = (X_1, \dots, X_d)$ with given univariate marginals via normal copula that fits the rank or the linear correlation between each pair of coordinates of X . The standard normal distribution function is applied to each coordinate of a vector $Z = (Z_1, \dots, Z_d)$ of correlated standard normals to produce a vector $U = (U_1, \dots, U_d)$ of correlated uniforms over $[0, 1]$. Then X is obtained by applying the inverse of each marginal distribution function to each coordinate of U . The fitting requires finding the correlation between the coordinates of each pair of Z that would yield the correlation between the coordinates of the corresponding pair of X . The step of finding the correlation matrix of Z , given the correlation matrix of X and the marginal distributions, constitutes the NORTA initialization step. In [1], we present a detailed analysis of the NORTA method and root-finding problem when the marginal distributions are discrete.

With the NORTA method, we have the following representation :

$$X_l = F_l^{-1}(\Phi(Z_l)), \quad l = 1, \dots, d,$$

where Φ is the standard normal distribution function and $F_l^{-1}(u) = \inf\{x : F_l(x) \geq u\}$ for $0 \leq u \leq 1$, which is the quantile function of the marginal distribution

$F_l, l = 1, \dots, d.$

For the bivariate case ($d = 2$), we have a vector $X = (X_1, X_2)$ and the two marginal distributions F_1 and F_2 with means and standard deviations $\mu_{F_1} = E[F_1(X_1)]$, $\mu_{F_2} = E[F_2(X_2)]$, $\sigma_{F_1} = \text{Var}(F_1(X_1))^{1/2}$ and $\sigma_{F_2} = \text{Var}(F_2(X_2))^{1/2}$, respectively. For this case, NORTA initialization is reduced to the problem of finding the correlation $\rho_Z = \text{Corr}(Z_1, Z_2)$.

In this document, we present a set of Java classes for NORTA initialization in the bivariate case given the rank correlation and two discrete marginal distributions.

We have :

$$r_X(\rho) = \text{Corr}(F_1(X_1), F_2(X_2)) = \frac{g_r(\rho) - \mu_{F_1}\mu_{F_2}}{\sigma_{F_1}\sigma_{F_2}}, \quad (\text{I.1})$$

where :

$$\begin{aligned} g_r(\rho) &= E[F_1(X_1)F_2(X_2)] \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_1\{F_1^{-1}[\Phi(x_1)]\} F_2\{F_2^{-1}[\Phi(x_2)]\} \phi_\rho(x_1, x_2) dx_1 dx_2, \end{aligned} \quad (\text{I.2})$$

where ϕ_ρ is the bivariate standard normal density. Then, for a given correlation r_X , we use an algorithm of root-finder to find the corresponding correlation ρ_Z that verifies

$$f_r(\rho_Z) = g_r(\rho_Z) - r_X \sigma_{F_1} \sigma_{F_2} - \mu_{F_1} \mu_{F_2} = 0. \quad (\text{I.3})$$

When the marginal distributions are continuous, the root-finding problem is easy to solve when we use the rank correlation. We have an analytic solution for

(I.2) and the relation in (I.1) becomes :

$$r_X(\rho) = (6/\pi) \arcsin(\rho/2).$$

NortaInitDisc

This abstract class defines the algorithms used for NORTA initialization when the marginal distributions are discrete. Four algorithms are supported for now, and they are defined as subclasses of the class `NortaInitDisc`.

For two random variables X_1 and X_2 , and their two marginal distributions F_1 and F_2 , respectively, we specify the rank correlation $r_X = \text{Corr}(F_1(X_1), F_2(X_2))$, the parameters of the marginal distributions and a parameter for truncation tr . For the correlation matching, we must have finite supports for the two distributions. Then if the support of each marginal is infinite, we have to upper-bound it at the quantile of order tr . For example, if the marginals have their support points in $[0, +\infty)$, the software will truncate to $[0, F_l^{-1}(tr)]$, for $l = 1, 2$. The parameter tr must to be given by the user, depending on the type of the two distributions. If the marginals have finite supports, one can simply give $tr = 1$.

Each algorithm `NI1`, `NI2a`, `NI2b` and `NI3` can be used to calculate the corresponding correlation $\rho_Z = \text{Corr}(Z_1, Z_2)$, where Z_1 and Z_2 are standard normal random variables. These subclasses implement the specific methods for NORTA initialization presented in [1].

Each type of algorithm should be defined as a subclass of `NortaInitDisc`. Each subclass must implement the method `computeCorr` which returns the solution ρ_Z . When executing this method, the subclass may call the methods `integ` and `deriv`, depending on the type of algorithm. For example, the subclass `NI1` calls only the method `integ`, since the algorithm do not use the derivative [1]. Each subclass must also call the method `computeParams`, which is executed immediately before the beginning of the root-finder algorithm.

When creating a class representing an algorithm, the `toString` method can be called to display information about the inputs.

```
public abstract class NortainitDisc
```

Constructor

```
public NortainitDisc (double rX,
                     DiscreteDistributionInt dist1,
                     DiscreteDistributionInt dist2,
                     double tr)
```

Constructor with the target rank correlation rX , the two discrete marginals $dist1$ and $dist2$ and the parameter for the truncation tr . This constructor can be called only by the constructors of the subclasses.

Methods

```
public abstract double computeCorr();
```

Computes and returns the correlation ρ_Z . Every subclass of `NortainitDisc` must implement this method.

```
public void computeParams ()
```

Computes the following inputs of each marginal distribution :

- The number of support points m_1 and m_2 for the two distributions.
- The means and standard deviations of $F_1(X_1)$ and $F_2(X_2)$, respectively.
- The vectors $p_1[i]$, $p_2[j]$, $z_1[i] = \Phi^{-1}(f_1[i])$ and $z_2[j] = \Phi^{-1}(f_2[j])$, where $f_1[i]$

and $f_2[j]$, for $i = 0, \dots, m_1 - 1$; $j = 0, \dots, m_2 - 1$, are the cumulative probability functions, and Φ is the standard normal distribution function.

Every subclass of `NortaInitDisc` must call this method.

```
public double integ (double r)
```

Computes the function

$$g_r(r) = \sum_{i=0}^{m_1-2} p_{1,i+1} \sum_{j=0}^{m_2-2} p_{2,j+1} \bar{\Phi}_r(z_{1,i}, z_{2,j}), \quad (\text{I.4})$$

which involves the bivariate normal integral $\bar{\Phi}_r(x, y) = \int_x^\infty \int_y^\infty \phi_r(z_1, z_2) dz_1 dz_2$. Method `barF` of class `BiNormalDonnellyDist` (from package `probdistmulti` of SSJ [2]) is used to compute $\bar{\Phi}_r(x, y)$, with m_1 , m_2 , and the vectors $p_1[i], i = 1, \dots, m_1 - 1$; $z_1[i], i = 0, \dots, m_1 - 2$; $p_2[j], j = 1, \dots, m_2 - 1$; $z_2[j], j = 0, \dots, m_2 - 2$. The correlation parameter r must be in $[-1, 1]$. This method may be called by subclasses of `NortaInitDisc`.

```
public double deriv (double r)
```

Computes the derivative of g_r , given by

$$g'_r(r) = \sum_{i=0}^{m_1-2} p_{1,i+1} \sum_{j=0}^{m_2-2} p_{2,j+1} \phi_r(z_{1,i}, z_{2,j}), \quad (\text{I.5})$$

where ϕ_r is the bivariate standard binormal density. The method uses m_1 , m_2 , and the vectors $p_1[i], i = 1, \dots, m_1 - 1$; $z_1[i], i = 0, \dots, m_1 - 2$; $p_2[j], j = 1, \dots, m_2 - 1$; $z_2[j], j = 0, \dots, m_2 - 2$. The correlation parameter r must be in $[-1, 1]$. This method

may be called by subclasses of `NortaInitDisc`.

NI1

Extends the class `NortaInitDisc` and implements the algorithm NI1. It uses an algorithm based on Brent method for root-finding, which combines root-bracketing, bisection and inverse quadratic interpolation. It calls the method `integ` to compute the function g_r given in (I.4). The search should be done in the interval $[-1, 0]$ if $r_X \in [-1, 0]$, or $[0, 1]$ if $r_X \in [0, 1]$. At each iteration, the algorithm halves the interval length and uses an accuracy ϵ to find the root ρ_Z of equation (I.3).

```
public class NI1 extends NortaInitDisc
```

Constructor

```
public NI1 (double rX,
            DiscreteDistributionInt dist1,
            DiscreteDistributionInt dist2,
            double tr,
            double tolerance)
```

Constructor with the target rank correlation `rX`, the two discrete marginals `dist1` and `dist2`, the parameter for truncation `tr` (see the constructor of class `NortaInitDisc`) and the specific parameter $\epsilon = \text{tolerance}$ defined above for the algorithm NI1.

Methods

```
public double computeCorr ()
```

Computes and returns the correlation ρ_Z using the algorithm NI1.

NI2a

Extends the class `NortaInitDisc` and implements the algorithm NI2a. It uses the derivative, so it calls the method `deriv` to compute the function g'_r given in (I.5). The double integration in (I.2) is simplified and only a simple integration is used. The algorithm uses numerical integration with Simpson's rules over subintervals given by the finite sequence $\rho_k = \rho_0 + 2kh$, for $k = 0, 1, \dots, m$, where h is a fixed step size and m is such that $1 - 2h < \rho_m < 1$. The initial point is chosen as $\rho_0 = 2 \sin(\pi r_X / 6)$. The integration is done between ρ_0 and $\rho_m = \pm(1 - \delta)$, or between ρ_0 and 0, depending on the sign of r_X and on whether the root is to the left, or to the right of ρ_0 . So depending on the case, the worst-case integration distance will be set to $d = |1 - \delta - \rho_0|$ or $d = |\rho_0|$. Then, the step size is readjusted to $h^* = d/(2m)$, where d is the maximum number of steps (iterations) calculated based on the pre-defined step size h , so $m = \lceil d/(2h) \rceil$. The algorithm stops at iteration k if the root is in a subinterval $[\rho_{k-1}, \rho_k]$, and a quadratic interpolation is used to compute the solution. For this, the method `interpol` of class `Misc` (from package `util` of SSJ [2]) is used.

```
public class NI2a extends NortaInitDisc
```

Constructor

```
public NI2a (double rX,
             DiscreteDistributionInt dist1,
             DiscreteDistributionInt dist2,
             double tr,
             double h,
```

double delta)

Constructor with the target rank correlation r_X , the two discrete marginals `dist1` and `dist2`, the parameter for the truncation `tr` (see the constructor of class `NortaInitDisc`), and the specific parameters `h` and $\delta = \text{delta}$ for the algorithm NI2a, as described above.

Methods

```
public double computeCorr ()
```

Computes and returns the correlation ρ_Z using the algorithm NI2a.

NI2b

Extends the class `NortaInitDisc` and implements the algorithm NI2b. It is a variant of NI2a. It uses the derivative, so it calls the method `deriv` to compute the function g'_r given in (I.5) and uses numerical integration with Simpson's rules as well. But the integration grid is either in the interval $[0, 1 - \delta]$ or $[-1 + \delta, 0]$, depending on the sign of r_X . Here the number of subintervals of integration is fixed to m and the algorithm stops at iteration k if the root is in subinterval $[\rho_{k-1}, \rho_k]$, and a quadratic interpolation is used to compute the solution. For this, the method `interpol` of class `Misc` (from package `util` of SSJ [2]) is used.

```
public class NI2b extends NortaInitDisc
```

Constructor

```
public NI2b (double rX,
            DiscreteDistributionInt dist1,
            DiscreteDistributionInt dist2,
            double tr,
            int m,
            double delta)
```

Constructor with the target rank correlation `rX`, the two discrete marginals `dist1` and `dist2`, the parameter for the truncation `tr` (see the constructor of class `NortaInitDisc`), and the specific parameters `m` and $\delta = \text{delta}$ for the algorithm NI2b, as described above.

Methods

```
public double computeCorr ()
```

Computes and returns the correlation ρ_Z using the algorithm NI2b.

NI3

Extends the class `NortaInitDisc` and implements the algorithm NI3. It uses the function g_r and its derivative g'_r , so it calls the methods `integ` and `deriv`, given in (I.4) and (I.5) and uses an adapted version of the Newton-Raphson algorithm. At any iteration, if the solution falls outside the search interval, the algorithm uses bisection and halves the interval length to guarantee convergence. The initial solution is taken as $\rho_0 = 2 \sin(\pi r_X / 6)$, and then at each iteration k , $f_r(\rho_k)$ and $f'_r(\rho_k)$ are calculated and a solution is computed by the recurrence formula :

$$\rho_{k+1} = \rho_k - \frac{f_r(\rho_k)}{f'_r(\rho_k)}.$$

The algorithm stops at iteration k if $|\rho_{k-1} - \rho_k| \leq \epsilon$. The function f_r is the one given in (I.3).

```
public class NI3 extends NortaInitDisc
```

Constructor

```
public NI3 (double rX,
           DiscreteDistributionInt dist1,
           DiscreteDistributionInt dist2,
           double tr,
           double tolerance)
```

Constructor with the target rank correlation `rX`, the two discrete marginals `dist1` and `dist2`, the parameter for the truncation `tr` (see the constructor of class `NortaInitDisc`), and the specific parameter $\epsilon = \text{tolerance}$ for the algorithm NI3, as defined above.

Methods

```
public double computeCorr ()
```

Computes and returns the correlation ρ_Z using algorithm N13.

Example

In this example, we consider two random variables X_1 and X_2 with negative binomial marginals, denoted by $\text{NegBin}(s, p)$. In our example, the parameters (s, p) for X_1 and X_2 , respectively, are : $s_1 = 15.68$, $p_1 = 0.3861$, $s_2 = 60.21$ and $p_2 = 0.6211$. We want to calculate the correlation ρ_Z for a target rank correlation $r_X = 0.43$. Since the negative binomial has an unbounded support, we set the upper bound points of each support at the quantile of order $tr = 1 - 10^{-6}$, so that the number of support points $m_l = F_l^{-1}(1 - 10^{-6}) + 1$, for $l = 1, 2$.

The Java program uses the class `DiscreteDistributionInt` of package `probdist` from `SSJ`, to specify the two discrete marginal distributions. Each of the four subclasses `NI1`, `NI2a`, `NI2b` and `NI3` are called for each algorithm to compute the correlation ρ_Z , so we can compare the results.

Listing I.1 – Example with correlated negative binomial distribution.

```
import umontreal.iro.lecuyer.probdist.*;
public class ExampleNortaInitDisc
{
    public static void main (String[] args) {
        final double rX = 0.43;           // Corrélation de rang donnée rX
        final double tr = 1.0 - 1.0e-6;  // Paramètre pour tronquer le domaine
        // Définie les deux distributions marginales
        DiscreteDistributionInt dist1 = new
            NegativeBinomialDist(15.68, 0.3861);
        DiscreteDistributionInt dist2 = new
            NegativeBinomialDist(60.21, 0.6211);
        NI1 ni1Obj = new NI1(rX, dist1, dist2, tr, 1.0e-4);
        System.out.println("Résultat avec la méthode M1: rho_Z = "
            + ni1Obj.computeCorr());
        NI2a ni2aObj = new NI2a(rX, dist1, dist2, tr, 0.005, 1.0e-4);
        System.out.println("Résultat avec la méthode M2B: rho_Z = "
            + ni2aObj.computeCorr());
        NI2b ni2bObj = new NI2b(rX, dist1, dist2, tr, 5, 1.0e-4);
        System.out.println("Résultat avec la méthode M2A: rho_Z = "
            + ni2bObj.computeCorr());
        NI3 ni3Obj = new NI3(rX, dist1, dist2, tr, 1.0e-4);
        System.out.println("Résultat avec la méthode M3: rho_Z = "
            + ni3Obj.computeCorr());
    }
}
```

Listing I.2 – Results of the program ExampleNortaInitDisc.java

```
Results with method NI1: rho_Z = 0.44691
Results with method NI2b: rho_Z = 0.44685
Results with method NI2a: rho_Z = 0.44683
Results with method NI3: rho_Z = 0.44691
```

BIBLIOGRAPHY

- [1] A. N. Avramidis, N. Channouf et P. L'Ecuyer. Efficient correlation matching for fitting discrete multivariate distributions with arbitrary marginals and normal-copula dependence. *INFORMS Journal on Computing*, 2008. To appear.
- [2] P. L'Ecuyer. *SSJ: A Java Library for Stochastic Simulation*, 2004. Software user's guide. Available at <http://www.iro.umontreal.ca/~lecuyer>.

I.2

Code source des classes Java pour l'initialisation de NORTA avec les corrélations de rang et des distributions marginales discrètes

Nous présentons ici le code source des classes Java, données dans la première partie de l'annexe I, pour l'initialisation de la méthode NORTA avec des corrélations de rang et des distributions marginales discrètes. Les quatre sous-classes NI1, NI2a, NI2b et NI3 correspondent aux quatre méthodes M1, M2A, M2B et M3, respectivement, que nous avons présentées dans les sections 3.3.1, 3.3.2 et 3.3.3.

Nous faisons appels aux classes suivantes de la librairie de simulation en Java SSJ :

- Classes `DiscreteDistributionInt` et `NormalDist` du paquetage `probdist`.
- Classe `BiNormalDonnellyDist` du paquetage `probdistmulti`.
- Classe `Misc` du paquetage `util`.

Listing I.3 – La classe abstraite `NortaInitDisc`.

```
import umontreal.iro.lecuyer.probdist.*;
import umontreal.iro.lecuyer.probdistmulti.*;
public abstract class NortaInitDisc
{
    protected double rX; // Corrélation de rang donnée
    protected DiscreteDistributionInt dist1; // Dist. marginale de X_1
    protected DiscreteDistributionInt dist2; // Dist. marginale de X_2
    protected double tr; // Paramètre pour la tronquation
    protected double mu1, mu2, sd1, sd2; /* Moyennes et écart-types de
                                         F_1(X_1) et F_2(X_2). */

    // Nombre de points dans les domaines des deux distributions.
    private int m1, m2;
    // Fonctions de masses pour les deux distributions.
    private double[] p1;
    private double[] p2;
    /* Quantiles des probabilités cumulatives selon la fonc. de répart.
       de la loi normale standard pour les deux distributions. */
    private double[] z1;
    private double[] z2;

    /* Constructeur de la classe abstraite. Peut être appelé
       seulement par le constructeur de la sous-classe de la
```

```

NortaInitMix. */
public NortaInitDisc (double rX,
                     DiscreteDistributionInt dist1,
                     DiscreteDistributionInt dist2,
                     double tr)
{
    this.rX = rX;
    this.dist1 = dist1;
    this.dist2 = dist2;
    this.tr = tr;
}

public abstract double computeCorr();

/* Calcule les moyennes et les écart-types de F_1(X_1) et F_2(X_2),
   le nombre de points des domaines et les vecteurs p_1, p_2, z_1, z_2. */
public void computeParams ()
{
    m1 = dist1.inverseFInt (tr) + 1;
    m2 = dist2.inverseFInt (tr) + 1;
    int[] y1 = new int[m1];
    int[] y2 = new int[m2];
    p1 = new double[m1];
    p2 = new double[m2];
    double[] f1 = new double[m1];
    double[] f2 = new double[m2];
    z1 = new double[m1];
    z2 = new double[m2];
    double u11 = 0.0, u22 = 0.0;

    // Calcule mu1, sd1, p1 and z1 of X_1
    for (int i = 0; i < m1; i++) {
        y1[i] = i;
        p1[i] = dist1.prob (y1[i]);
        f1[i] = dist1.cdf (y1[i]);
        z1[i] = NormalDist.inverseF01 (f1[i]);
        if (z1[i] == Double.NEGATIVE_INFINITY)
            z1[i] = NormalDist.inverseF01 (2.2e-308);
        if (z1[i] == Double.POSITIVE_INFINITY)
            z1[i] = NormalDist.inverseF01 (1.0 - Math.ulp (1.0));
        mu1 += f1[i] * p1[i];
        u11 += f1[i] * f1[i] * p1[i];
    }
    sd1 = Math.sqrt (u11 - mu1 * mu1);

    // Calcule mu2, sd2, p2 and z2 of X_2
    for (int i = 0; i < m2; i++) {
        y2[i] = i;
        p2[i] = dist2.prob (y2[i]);
        f2[i] = dist2.cdf (y2[i]);
        z2[i] = NormalDist.inverseF01 (f2[i]);
        if (z2[i] == Double.NEGATIVE_INFINITY)
            z2[i] = NormalDist.inverseF01 (2.2e-308);
        if (z2[i] == Double.POSITIVE_INFINITY)
            z2[i] = NormalDist.inverseF01 (1.0 - Math.ulp (1.0));
        mu2 += f2[i] * p2[i];
        u22 += f2[i] * f2[i] * p2[i];
    }
    sd2 = Math.sqrt (u22 - mu2 * mu2);
}

```

```

}
// Calcule la fonction g_r pour chaque corrélation
public double integ (double r)
{
    double gr = 0.0, sum;
    for (int i = 0; i < m1 - 1; i++) {
        sum = 0.0;
        for (int j = 0; j < m2 - 1; j++) {
            sum += p2[j+1] * BiNormalDonnellyDist.barF (z1[i], z2[j], r);
        }
        gr += p1[i + 1] * sum;
    }
    return gr;
}
// Calcule la fonction dérivée g'_r pour chaque corrélation
public double deriv (double r)
{
    double c = Math.sqrt (1.0 - r * r);
    double c1 = 2 * c * c;
    double gp = 0.0, sum;
    for (int i = 0; i < m1 - 1; i++) {
        double z1sq = z1[i] * z1[i];
        double t1 = 2 * r * z1[i];
        sum = 0.0;
        for (int j = 0; j < m2 - 1; j++) {
            sum += p2[j+1] * Math.exp ((t1 * z2[j] - z1sq - z2[j]
                * z2[j]) / c1);
        }
        gp += p1[i + 1] * sum;
    }
    gp = gp / (2.0 * Math.PI * c);
    return gp;
}
// Pour afficher les inputs
private String tabToString(double[] tab , String message)
{
    String desc = message + "\n [";
    for (int i = 0; i < tab.length; i++)
        if (i == tab.length - 1)
            desc += "]\n";
        else
            desc += tab[i] + ",";
    return desc;
}
public String toString()
{
    String desc = "";
    desc += "rX = " + rX + "\n";
    desc += "tr = " + tr + "\n";
    desc += "m1 = " + m1 + "\n";
    desc += "m2 = " + m2 + "\n";
    desc += "mu1 = " + mu1 + "\n";
    desc += "mu2 = " + mu2 + "\n";
    desc += "sd1 = " + sd1 + "\n";
    desc += "sd2 = " + sd2 + "\n";
    desc += tabToString(p1 , "Table p1 : ");
    desc += tabToString(z1 , "Table z1 : ");
}

```

```

    desc += tabToString(p2 , "Table p2 : ");
    desc += tabToString( z2, "Table z2 : ");
    return desc;
}
}

```

Listing I.4 – La sous-classe NI1 pour la méthode M1.

```

import umontreal.iro.lecuyer.probdist.*;
import umontreal.iro.lecuyer.util.*;
public class NI1 extends NortainitDisc
{
    private double tolerance;// Seuil de tolérance
    public NI1 (double rX,
                DiscreteDistributionInt dist1,
                DiscreteDistributionInt dist2,
                double tr,
                double tolerance)
    {
        super(rX, dist1, dist2, tr);
        this.tolerance = tolerance;
        computeParams();// Appelle la méthode computeParams
    }
    public double computeCorr ()
    {
        final double ITMAX = 100;// Nombre maximum d'itérations
        final double EPS = 1.0e-15;// Précision de la machine
        double b;// La solution finale
        double a, c, mini, min2, fa, fb, fc, pp, q, rrr, s,
                tolerancel, xm,x1,x2;
        double e = 0.0, d = 0.0;
        // Constantes
        double cc = rX * sd1 * sd2;
        double ccc = cc + mu1 * mu2;

        if (rX == 0)
            return 0.0;
        if (rX > 0) { // Oriente la recherche et initialise a, b, c
            x1 = 0.0;
            x2 = 1.0;
            a = x1;
            b = x2;
            c = x2;
            fa = - cc;
            fb = integ (b) - ccc;
        }
        else {
            x1 = -1.0;
            x2 = 0.0;
            a = x1;
            b = x2;
            c = x2;
            fa = integ (a) - ccc;
            fb = - cc;
        }
    }
}

```

```

fc = fb;
for (int i = 1; i <= ITMAX; i++) { // Commencer la recherche
    if ((fb > 0.0 && fc > 0.0) || (fb < 0.0 && fc < 0.0)) {
        // Renommer a, b, c et ajuster l'intervalle d.
        c = a;
        fc = fa;
        e = d = b - a;
    }
    if (Math.abs (fc) < Math.abs (fb)) {
        a = b;
        b = c;
        c = a;
        fa = fb;
        fb = fc;
        fc = fa;
    }
    // Seuil de tolérance final
    tolerance1 = 2.0 * EPS * Math.abs (b) + 0.5 * tolerance;
    xm = 0.5 * (c - b);
    if (Math.abs (xm) <= tolerance1 || fb == 0.0)
        return b;
    if (Math.abs (e) >= tolerance1 && Math.abs (fa) > Math.abs (fb)) {
        s = fb / fa; // Utiliser l'interpolation quadratique inverse
        if (a == c) {
            pp = 2.0 * xm * s;
            q = 1.0 - s;
        }
        else {
            q = fa / fc;
            rrr = fb / fc;
            pp = s * (2.0 * xm * q * (q - rrr) - (b - a) * (rrr - 1.0));
            q = (q - 1.0) * (rrr - 1.0) * (s - 1.0);
        }
        if (pp > 0.0)
            q = -q; // Vérifier les limites de l'intervalle
        pp = Math.abs (pp);
        min1 = 3.0 * xm * q - Math.abs (tolerance1 * q);
        min2 = Math.abs (e * q);
        if (2.0 * pp < (min1 < min2 ? min1 : min2)) {
            e = d; // Accepter l'interpolation
            d = pp / q;
        }
        else { // Échec de l'interpolation, utiliser bisection
            d = xm;
            e = d;
        }
    }
    else { // Limites convergent lentement, alors utiliser la bisection
        d = xm;
        e = d;
    }
    a = b; // a devient la meilleure solution
    fa = fb;
    if (Math.abs (d) > tolerance1)
        b += d; // Évalue la nouvelle solution
    else {
        if (xm > 0) b += Math.abs (tolerance1);
        else
            b += -Math.abs (tolerance1);
    }
}

```

```

        fb = integ (b) - ccc;
    }
    return b;
}
// Pour afficher les inputs
public String toString()
{
    String desc = super.toString();
    desc += "tolerance : " + tolerance + "\n";
    return desc;
}
}

```

Listing I.5 – La sous-classe NI2a pour la méthode M2A.

```

import umontreal.iro.lecuyer.probdist.*;
import umontreal.iro.lecuyer.util.*;

public class NI2a extends NortainitDisc
{
    private double h; // Pas utilisé dans l'intégration numérique.
    private double delta; // Paramètre positif pour avoir rho_m = 1-delta
                          // (ou -1+delta), tout près de 1 (ou -1). */

    public NI2a(double rX,
                DiscreteDistributionInt dist1,
                DiscreteDistributionInt dist2,
                double tr,
                double h,
                double delta)
    {
        super(rX, dist1, dist2, tr);
        this.h = h;
        this.delta = delta;
        computeParams(); // Appelle la méthode computeParams
    }

    public double computeCorr ()
    {
        // Paramètres pour l'interpolation quadratique.
        double xtemp = 0.0, temp = 0.0;
        double[] x = new double[3];
        double[] y = new double[3];
        double[] c = new double[3];

        double b = 0.0; // La solution finale
        double xold, xnew, dold, dnew, dmid, iold, inew;
        double d = 0.0, h2 = 0.0, hd3 = 0.0;
        int m; // Nombre d'itérations

        // Corrélation min. et max.
        double lrx = (integ (-1) - mu1 * mu2) / sd1 * sd2;
        double urx = (integ (1) - mu1 * mu2) / sd1 * sd2;
        // Solution initiale
        double rho1 = 2 * Math.sin (Math.PI * rX / 6);
        double intg1 = integ (rho1); // Calcule g_r(rho1)
    }
}

```

```

double gr = rX * sd1 * sd2 + mu1 * mu2; /* integ(\rho) = gr
                                         équivalent à
                                         \rho = la solution. */

if (intg1 == gr)
    return rho1;
if (intg1 < gr) { // Orienter la recherche de gauche à droite
    if (0 < rX && rX < 1) /* Faire la recherche entre rh_0 et
                           rho_m=1 - delta. */
        d = 1 - delta - rho1;
    if (-1 < rX && rX < 0) // Faire la recherche entre rh_0 et 0
        d = -rho1;
    m = (int) Math.ceil(d / (2 * h));
    h = d / (2 * m); // Réajuster h
    hd3 = h / 3; // Constante
    h2 = 2 * h; // Constante
    xold = rho1;
    dold = deriv (xold);
    iold = intg1;
    for (int i = 1; i <= m; i++) { // Commencer la recherche
        dmid = deriv (xold + h);
        xnew = xold + h2;
        dnew = deriv (xnew);
        inew = iold + hd3 * (dold + 4 * dmid + dnew);
        if (inew >= gr) { // La solution est dans l'actuel intervalle
            // Calcule les paramètres pou l'interpolation quadratique
            x[0] = xtemp;
            x[1] = xold;
            x[2] = xnew;
            y[0] = temp;
            y[1] = iold;
            y[2] = inew;
            Misc.interpol (2, x, y, c);
            b = (c[2] * (xtemp + xold) - c[1] + Math.sqrt ((c[1]
                - c[2] * (xtemp + xold)) * (c[1] - c[2] * (xtemp
                + xold)) - 4 * c[2] * (c[0] - c[1] * xtemp + c[2]
                * xtemp * xold - gr))) / (2 * c[2]));
            return b;
        }
        xtemp = xold;
        temp = iold;
        xold = xnew;
        dold = dnew;
        iold = inew;
    }
    b = 1.0 - delta / 2; // Quand la solution est à droite de rho_m
}
if (intg1 > gr) { // Orienter la recherche de gauche à droite.
    if (0 < rX && rX < 1) // Faire la recherche entre 0 et rh_0
        d = rho1;
    if (-1 < rX && rX < 0) /* Faire la recherche entre
                           rho_m=-1+delta et rho_0. */
        d = rho1 + 1 - delta;
    m = (int) Math.ceil (d / (2 * h));
    h = d / (2 * m); // Réajuster h
    hd3 = h / 3; // Constante
    h2 = 2 * h; // Constante
    xold = rho1;
    dold = deriv (xold);
}

```

```

iold = intg1;
for (int i = 1; i <= m; i++) { // Commencer la recherche
    dmid = deriv (xold - h);
    xnew = xold - h2;
    dnew = deriv (xnew);
    inew = iold - hd3 * (dold + 4 * dmid + dnew);
    if (inew <= gr) { // La solution est dans l'actuel intervalle
        // Calcule les paramètres pou l'interpolation quadratique
        x[0] = xnew;
        x[1] = xold;
        x[2] = xtemp;
        y[0] = inew;
        y[1] = iold;
        y[2] = temp;
        Misc.interpol (2, x, y, c);
        b = (c[2] * (xnew + xold) - c[1] + Math.sqrt ((c[1]
            - c[2] * (xnew + xold)) * (c[1] - c[2] * (xnew
            + xold)) - 4 * c[2] * (c[0] - c[1] * xnew + c[2]
            * xnew * xold - gr))) / (2 * c[2]));
        return b;
    }
    xtemp = xold;
    temp = iold;
    xold = xnew;
    dold = dnew;
    iold = inew;
}
b = -1 + delta / 2; // Quand la solution est à gauche de rho_m
}
return b;
}
// Pour afficher les inputs
public String toString()
{
    String desc = super.toString();
    desc += "h : " + h + "\n";
    desc += " delta : " + delta + "\n";
    return desc;
}
}

```

Listing I.6 – La sous-classe NI2b pour la méthode M2B.

```

import umontreal.iro.lecuyer.probdist.*;
import umontreal.iro.lecuyer.util.*;
public class NI2b extends NortainitDisc
{
    private int m; // Nombre max. d'itérations pour l'intégration.
    private double delta; /* Paramètre positif pour avoir rho_m =1-delta
        (ou -1+delta), tout près de 1 (ou -1). */

    public NI2b(double rX,
        DiscreteDistributionInt dist1,
        DiscreteDistributionInt dist2,

```

```

        double tr,
        int m,
        double delta)
{
    super(rX, dist1, dist2, tr);
    this.m = m;
    this.delta = delta;
    computeParams();// Appelle la méthode computeParams
}
public double computeCorr ()
{
    // Paramètres pour l'interpolation quadratique.
    double xtemp = 0.0, temp = 0.0;
    double[] x = new double[3];
    double[] y = new double[3];
    double[] c = new double[3];

    double b = 0.0; // La solution finale
    double xold, xnew, dold, dnew, dmid, iold, inew;

    double rho1 = 0.0; // Solution initiale
    double intg1 = mu1 * mu2; // Calcule g_r(rho_1)
    double gr = rX * sd1 * sd2 + mu1 * mu2; /* integ(\rho) = gr
                                                équivalent à
                                                \rho = la solution. */

    double h = (1 - delta) / (2 * m);
    double hd3 = h / 3; // Constante
    double h2 = 2 * h; // Constante

    if (intg1 == gr)
        return rho1;
    if (0 < rX && rX < 1) { /* Faire la recherche entre rh_0 et
                                rho_m=1 - delta. */
        xold = rho1;
        dold = deriv(xold);
        iold = intg1;
        for (int i = 1; i <= m ;i++) { // Commencer la recherche
            dmid = deriv(xold + h);
            xnew = xold + h2;
            dnew = deriv(xnew);
            inew = iold + hd3 * (dold + 4 * dmid + dnew);
            if (inew >= gr) { // La solution est dans l'actuel intervalle
                // Calcule les paramètres pou l'interpolation quadratique
                x[0] = xtemp;
                x[1] = xold;
                x[2] = xnew;
                y[0] = temp;
                y[1] = iold;
                y[2] = inew;
                Misc.interpol(2, x, y, c);
                b = (c[2] * (xtemp + xold) - c[1] + Math.sqrt((c[1]
                    - c[2] * (xtemp + xold)) * (c[1] - c[2] * (xtemp
                    + xold)) - 4 * c[2] * (c[0] - c[1] * xtemp + c[2]
                    * xtemp * xold - gr))) / (2 * c[2]));
                return b;
            }
            xtemp = xold;
            temp = iold;
            xold = xnew;
        }
    }
}

```

```

        dold = dnew ;
        iold = inew;
    }
    b = 1 - delta / 2; // Quand la solution est à droite de rho_m
}
}
if ( -1 < rX && rX < 0) { /* Faire la recherche entre
                            rho_m=-1+delta et rho_0. */
    xold = rho1;
    dold = deriv(xold);
    iold = intg1;
    for (int i = 1; i <= m ;i++) { // Commencer la recherche
        dmid = deriv(xold - h);
        xnew = xold - h2;
        dnew = deriv(xnew);
        inew = iold - hd3 * (dold + 4 * dmid + dnew);
        if (inew <= gr) { // La solution est dans l'actuel intervalle
            // Calcule les paramètres pou l'interpolation quadratique
            x[0] = xnew;
            x[1] = xold;
            x[2] = xtemp;
            y[0] = inew;
            y[1] = iold;
            y[2] = temp;
            Misc.interpol(2, x, y, c);
            b = (c[2] * (xnew + xold) - c[1] + Math.sqrt((c[1]
                - c[2] * (xnew + xold)) * (c[1] - c[2] * (xnew
                + xold)) - 4 * c[2] * (c[0] - c[1] * xnew + c[2]
                * xnew * xold - gr))) / (2 * c[2]));
            return b;
        }
        xtemp = xold;
        temp = iold;
        xold = xnew;
        dold = dnew ;
        iold = inew;
    }
    b = -1 + delta / 2; // Quand la solution est à gauche de rho_m.
    return b;
}
// Affiche les inputs
public String toString()
{
    String desc = super.toString();
    desc += "m : " + m + "\n";
    desc += "delta : " + delta + "\n";
    return desc;
}
}

```

Listing I.7 – La sous-classe NI3 pour la méthode M3.

```

import umontreal.iro.lecuyer.probdist.*;
public class NI3 extends NortaInitDisc {

```

```

private double tolerance;// Seuil de tolérance
public NI3 (double rX,
            DiscreteDistributionInt dist1,
            DiscreteDistributionInt dist2,
            double tr,
            double tolerance)
{
    super(rX, dist1, dist2, tr);
    this.tolerance = tolerance;
    computeParams();// Appelle la méthode computeParams
}
public double computeCorr ()
{
    final double ITMAX = 100;// Nombre maximum d'itérations
    double b = 0.0; // La solution finale
    double f, df, dx, xl, xh, dxold, temp;
    double ccc = rX * sd1 * sd2 + mu1 * mu2;// Constante
    if (rX == 0.0)
        return 0.0;
    if (rX > 0.0) { // Orienter la recherche
        xl = 0.0;
        xh = 1.0;
    }
    else {
        xl = -1.0;
        xh = 0.0;
    }
    b = 2 * Math.sin (Math.PI * rX / 6); // Solution initiale
    dxold = xh - xl;
    dx = dxold;
    f = integ (b) - ccc;
    df = deriv (b);
    for (int i = 1; i <= ITMAX; i++) { // Commencer la recherche
        if (((b - xh) * df - f) * ((b - xl) * df - f) > 0.0)
            || (Math.abs (2.0 * f) > Math.abs (dxold * df)) {
                /* Faire la bisection si la solution sort de l'intervalle
                ou converge rapidement. */
                dxold = dx;
                dx = 0.5 * (xh - xl);
                b = xl + dx;
                if (xl == b) // S'il y a convergence
                    return b; // Accepter la solution.
            }
        else {
                dxold = dx;
                dx = f / df;
                temp = b;
                b -= dx;
                if (temp == b)
                    return b;
            }
        if (Math.abs (dx) < tolerance)
            return b; // Test pour la convergence
        f = integ (b) - ccc;
        df = deriv (b);
        if (f < 0.0)
            xl = b; // Garder la solution à l'interieur de l'intervalle
    }
}

```

```
        else
            xh = b;
    }
    return b;
}
// Pour afficher les inputs
public String toString()
{
    String desc = super.toString();
    desc += "tolerance : " + tolerance + "\n";
    return desc;
}
}
```

Annexe II

Classes Java pour l'initialisation de NORTA avec les corrélations de rang et un mélange de distributions marginales continues et discrètes

II.1

**Java classes for NORTA initialization with rank correlations and a mix
of continuous and discrete marginal distributions**

User's guide with an example

(version of 11 April 2008)

Nabil Channouf

Introduction

This document describes a Java implementation of correlation matching algorithm for the NORTA method to fit a multivariate distribution with a mix of continuous and discrete marginals. Only one algorithm, called NI4, is implemented as a subclass of an abstract class named `NortaInitMix`. The software makes use of the SSJ library [1] and the scientific Flanagan's library [2]. An example of how to use it is given at the end of this document.

The NORTA method is an approach for modeling dependence in a finite-dimensional random vector $X = (X_1, \dots, X_d)$ with given univariate marginals via normal copula that fits the rank or the linear correlation between each pair of coordinates of X . The standard normal distribution function is applied to each coordinate of a vector $Z = (Z_1, \dots, Z_d)$ of correlated standard normals to produce a vector $U = (U_1, \dots, U_d)$ of correlated uniforms over $[0, 1]$. Then X is obtained by applying the inverse of each marginal distribution function to each coordinate of U . The fitting requires finding the correlation between the coordinates of each pair of Z that would yield the correlation between the coordinates of the corresponding pair of X . The step of finding the correlation matrix of Z , given the correlation matrix of X and the marginal distributions, constitutes the NORTA initialization step.

With the NORTA method, we have the following representation :

$$X_l = F_l^{-1}(\Phi(Z_l)), \quad l = 1, \dots, d,$$

where Φ is the standard normal distribution function and $F_l^{-1}(u) = \inf\{x : F_l(x) \geq u\}$ for $0 \leq u \leq 1$, which is the quantile function of the marginal distribution $F_l, l = 1, \dots, d$.

For the bivariate case ($d = 2$), we have a vector $X = (X_1, X_2)$, with X_1 being continuous and X_2 discrete. We have their two marginal distributions F_1 and F_2 with means and standard deviations $\mu_{F_1} = E[F_1(X_1)]$, $\mu_{F_2} = E[F_2(X_2)]$, $\sigma_{F_1} = \text{Var}(F_1(X_1))^{1/2}$ and $\sigma_{F_2} = \text{Var}(F_2(X_2))^{1/2}$, respectively. Note that $F_1(X_1)$ is uniformly distributed over $[0, 1]$ since X_2 is continuous, so we have $\mu_{F_1} = 0.5$ and $\sigma_{F_1} = \sqrt{1/12}$. The NORTA initialization is reduced to the problem of finding the correlation $\rho_Z = \text{Corr}(Z_1, Z_2)$.

We present here Java classes for NORTA initialization with rank correlation, one continuous and one discrete marginal distributions. We have :

$$\tau_X(\rho) = \text{Corr}(F_1(X_1), F_2(X_2)) = \frac{g_r(\rho) - \mu_{F_1}\mu_{F_2}}{\sigma_{F_1}\sigma_{F_2}}, \quad (\text{II.1})$$

where :

$$\begin{aligned} g_r(\rho) &= E[F_1(X_1)F_2(X_2)] \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Phi(x_1)F_2\{F_2^{-1}[\Phi(x_2)]\}\phi_\rho(x_1, x_2)dx_1dx_2, \end{aligned} \quad (\text{II.2})$$

where ϕ_ρ is the bivariate standard normal density. Then, for a given correlation τ_X , we use an algorithm of root-finder to find the corresponding correlation ρ_Z that verifies

$$f_r(\rho_Z) = g_r(\rho_Z) - \tau_X\sigma_{F_1}\sigma_{F_2} - \mu_{F_1}\mu_{F_2} = 0. \quad (\text{II.3})$$

NortaInitMix

This abstract class defines the algorithm used for NORTA initialization when we have a mix of continuous and discrete marginal distributions. One algorithm is supported for now, and it is defined as subclass of the class `NortaInitMix`.

For a continuous random variable X_1 and a discrete random variable X_2 , and their two marginal distributions F_1 and F_2 , respectively, we specify the rank correlation $r_X = \text{Corr}(F_1(X_1), F_2(X_2))$, the parameters of the discrete marginal distribution, a parameter for truncation tr and a parameter for the number of points for the Gaussian quadrature $nPoints$. We do not need to specify the parameters of the continuous distribution, since $F_1(X_1)$ is uniformly distributed over $[0, 1]$. For the correlation matching, we must have finite support for the discrete distribution. Then if the support is infinite, we have to upper-bound it at the quantile of order tr . For example, if the marginal have its support points in $[0, +\infty)$, the software will truncate to $[0, F_2^{-1}(tr)]$. The parameter tr must to be given by the user, depending on the type of the two distributions. If the marginals have finite supports, one can simply give $tr = 1$.

The algorithm NI4 can be used to calculate the corresponding correlation $\rho_Z = \text{Corr}(Z_1, Z_2)$, where Z_1 and Z_2 are standard normal random variables.

The algorithm should be defined as a subclass of `NortaInitMix`. This subclass must implement the method `computeCorr` which returns the solution ρ_Z . When executing this method, the subclass may call the method `integ` and this method uses the class `IntegralFunction` and the its subclass `Funcnt` for the numerical integration by the Gauss-Legendre quadrature. The subclass must also call the method `computeParams`, which is executed immediately before the beginning of the root-finder algorithm.

When creating a class representing the algorithm, the `toString` method can be

called to display information about the inputs.

```
public abstract class NortainitMix
```

Constructor

```
public NortainitMix (double rX,
                    DiscreteDistributionInt dist2,
                    double tr,
                    int nPoints)
```

Constructor with the target rank correlation rX , the discrete marginal $dist2$, the parameter for the truncation tr and the number of points for the Gaussian quadrature $nPoints$. This constructor can be called only by the constructor of the subclass.

Methods

```
public abstract double computeCorr();
```

Computes and returns the correlation ρ_Z . The subclass of `NortainitMix` must implement this method.

```
public void computeParams ()
```

Computes the following inputs for the discrete marginal distribution :

- The number of support points m_2 .
- The mean and standard deviation of $F_2(X_2)$.

- The vectors $p_2[j]$ and $z_2[j] = \Phi^{-1}(f_2[j])$, where $f_2[j]$, $j = 0, \dots, m_2 - 1$, is the cumulative probability function, and Φ is the standard normal distribution function.

The subclass of `NortaInitMix` must call this method.

```
public double integ (double r)
```

Computes the function

$$g_r(\rho_{i,j}^Z) = \sum_{j=0}^{m_2-1} f_{2,j} \left[\int_0^1 u_i \left(\Phi \left(\frac{z_{2,j} - \rho_{i,j}^Z \Phi^{-1}(u_i)}{\sqrt{1 - (\rho_{i,j}^Z)^2}} \right) - \Phi \left(\frac{z_{2,j-1} - \rho_{i,j}^Z \Phi^{-1}(u_i)}{\sqrt{1 - (\rho_{i,j}^Z)^2}} \right) \right) du_i \right], \quad (\text{II.4})$$

which involves the normal integral $\Phi_r(x) = \int_x^\infty \phi_r(z) dz$. The method `gaussQuad` of class `Integration` (from package `Integration` of Flanagan's library [30]) is used to compute the integral in (II.4) over $[0, 1]$ with m_2 and the vectors $f_2[j]$, $j = 1, \dots, m_2 - 1$ and $z_2[j]$, $j = 0, \dots, m_2 - 2$. The correlation parameter r must be in $[-1, 1]$. This method may be called by the subclasses of `NortaInitMix`.

NI4

Extends the class `NortaInitMix` and implements the algorithm NI4. It uses an algorithm based on Brent method for root-finding, which combines root-bracketing, bisection and inverse quadratic interpolation. It calls the method `integ` to compute the function g_r , given in (II.4), using numerical integration by the Gauss-Legendre quadrature. The search should be done in the interval $[-1, 0]$ if $r_X \in [-1, 0]$, or $[0, 1]$ if $r_X \in [0, 1]$. At each iteration, the algorithm halves the interval length and uses an accuracy ϵ to find the root ρ_Z of equation (II.3).

```
public class NI4 extends NortaInitMix
```

Constructor

```
public NI4 (double rX,
           DiscreteDistributionInt dist2,
           double tr,
           int nPoints,
           double tolerance)
```

Constructor with the target rank correlation `rX`, the discrete marginal `dist2`, the parameter for truncation `tr` (see the constructor of class `NortaInitMix`), the parameter `nPoints` for the number of points for the Gauss-Legendre quadrature (numerical integration) and the specific parameter $\epsilon = \text{tolerance}$ defined above for the algorithm NI4.

Methods

```
public double computeCorr ()
```

Computes and returns the correlation ρ_Z using the algorithm NI4.

Example

In this example, we consider a continuous random variable X_1 (we do not to specify its parameters), and a discrete random variable X_2 having a Poisson distribution with mean 25. We want to calculate the correlation ρ_Z for a target rank correlation $r_X = 0.2$. Since the Poisson distribution has an unbounded support, we set the upper bound points of each support at the quantile of order $tr = 1 - 10^{-6}$, so that the number of support points $m_2 = F_2^{-1}(1 - 10^{-6}) + 1$.

The Java program uses the class `DiscreteDistributionInt` of package `probdist` from `SSJ`, to specify the discrete marginal distribution. The subclass `NI4` is called to compute the correlation ρ_Z .

Listing II.1 – Example with correlated Normal and Poisson distributions.

```
import umontreal.iro.lecuyer.probdist.*;

public class ExampleNortaInitMix
{
    public static void main (String[] args) {
        final double rX = 0.2;           // Target rank correlation rX
        final double tr = 1.0 - 1.0e-6;  // Quantile upper limit
        final int nPoints = 16;         // Nbr. of points for num. integration

        // Define the discrete marginal distribution
        DiscreteDistributionInt dist2 = new PoissonDist(25.0);

        NI4 ni4Obj = new NI4(rX, dist2, tr, nPoints, 1.0e-4);
        System.out.println("Result with method NI4: rho_Z = "
            + ni4Obj.computeCorr());
    }
}
```

Listing II.2 – Results of the program `ExampleNortaInitMix.java`

Result with method NI4: rho_Z = 0.2099

BIBLIOGRAPHY

- [1] P. L'Ecuyer. *SSJ: A Java Library for Stochastic Simulation*, 2004. Software user's guide. Available at <http://www.iro.umontreal.ca/~lecuyer>.
- [2] M. T. Flanagan. *Michael Thomas Flanagan's Scientific Java Library*, 2007. Integration Class: numerical integration. Available at <http://www.ee.ucl.ac.uk/~mflanaga>.

II.2

Code source des classes Java pour l'initialisation de NORTA avec les corrélations de rang et un mélange de distributions marginales continues et discrètes.

Nous présentons ici le code source des classes Java, données dans la première partie de l'annexe II, pour l'initialisation de NORTA avec des corrélations de rang et un mélange de distributions marginales continues et discrètes. La classe NI4 correspond à la méthode M4 présentée à la section 3.3.6.

Nous utilisons la librairie de simulation en Java SSJ en faisant appel aux classes suivantes :

- Classes `DiscreteDistributionInt` et `NormalDist` du paquetage `probdist`.
- Classe `createForSingleThread` du paquetage `util`.

De la librairie scientifique de Flanagan, nous faisons appel aux classes suivantes :

- Classe `Funct` du paquetage `IntegralFunction`.
- Classe `gaussQuad` du paquetage `Integration`.

Listing II.3 – La classe abstraite `NortaInitMix`.

```
import flanagan.integration.Integration;
import flanagan.integration.IntegralFunction;
import umontreal.iro.lecuyer.probdistmulti.*;
import umontreal.iro.lecuyer.probdist.*;

public abstract class NortaInitMix
{
    protected double rX; // Corrélation de rang donnée
    protected DiscreteDistributionInt dist2; // Dist. marginale de X_2
    protected double tr; // Paramètre pour la tronquation
    protected int nPoints; // Num. of intervals for numerical integ.
    protected double mu2, sd2; // Moyenne et écart-type de F_2(X_2).

    private int m2; // Nombre de points du domaine de F_2.
    private double[] p2; // Fonction de masses de F_2.
    private double[] f2; // Fonction cum. de F_2.
    /* Quantiles des probabilités cumulatives selon la fonc. de répart.
       de la loi normale standard pour F_2. */
    normal C.D.F for the disc. marginal */
}
```

```

private double[] z2;

/* Constructeur de la classe abstraite. Peut être appelé
seulement par les constructeurs des sous-classes de la
NortaInitDis.*/
public NortaInitMix (double rX,
                    DiscreteDistributionInt dist2,
                    double tr,
                    int nPoints)

{
    this.rX = rX;
    this.dist2 = dist2;
    this.nPoints = nPoints;
    this.tr = tr;
}

// Moyenne et écart-type de F_2(X_2)
final double mu1 = 0.5;
final double sd1 = Math.sqrt(1.0/12);

public abstract double computeCorr();

/* Calcule la moyennes et l'écart-types de F_2(X_2),
le nombre de points du domaines et les vecteurs p_2 et z_2. */
public void computeParams ()
{
    m2 = dist2.inverseFInt (tr) + 1;
    int[] y2 = new int[m2];
    p2 = new double[m2];
    f2 = new double[m2];
    z2 = new double[m2];
    double u22 = 0.0;

    // Calcule mu2, sd2, p2 and z2 of X_2
    for (int i = 0; i < m2; i++) {
        y2[i] = i;
        p2[i] = dist2.prob (y2[i]);
        f2[i] = dist2.cdf (y2[i]);
        z2[i] = NormalDist.inverseF01 (f2[i]);
        if (z2[i] == Double.NEGATIVE_INFINITY)
            z2[i] = NormalDist.inverseF01 (2.2e-308);
        if (z2[i] == Double.POSITIVE_INFINITY)
            z2[i] = NormalDist.inverseF01 (1.0 - Math.ulp (1.0));
        mu2 += f2[i] * p2[i];
        u22 += f2[i] * f2[i] * p2[i];
    }
    sd2 = Math.sqrt (u22 - mu2 * mu2);
}

// Définit la fonction à intégrer .
class Funct implements IntegralFunction{
    private double rho ,z21,z22;// Corrélacion et bornes d'intégration.
    public double function(double x){
        double y = x * (NormalDist.cdf01((z21-rho *
            NormalDist.inverseF01(x))/Math.sqrt(1-rho*rho))
            - NormalDist.cdf01((z22-rho*NormalDist.inverseF01(x))
            / Math.sqrt(1-rho*rho)));
        return y;
    }
}

```

```

    public void setA(double rho){
        this.rho = rho;
    }
    public void setB(double z21){
        this.z21 = z21;
    }
    public void setC(double z22){
        this.z22 = z22;
    }
}

// Calcule la fonction g_r pour chaque corrélation.
public double integ (double r)
{
    double gr = 0.0;
    Funct f = new Funct();
    f.setA(r);
    for (int j = 1; j < m2 ; j++) {
        f.setB(z2 [j]);
        f.setC(z2 [j-1]);
        gr+= f2[j] * Integration.gaussQuad(f, 0.0, 1.0, nPoints);
    }
    return gr;
}

// Pour afficher les inputs
private String tabToString(double[] tab , String message)
{
    String desc = message + "\n [";
    for (int i = 0; i < tab.length; i++)
        if (i == tab.length - 1)
            desc += "]\n";
        else
            desc += tab[i] + ", ";
    return desc;
}

public String toString()
{
    String desc = "";
    desc += "rX = " + rX + "\n";
    desc += "tr = " + tr + "\n";
    desc += "nPoints = " + nPoints + "\n";
    desc += "m2 = " + m2 + "\n";
    desc += "mu2 = " + mu2 + "\n";
    desc += "sd2 = " + sd2 + "\n";
    desc += tabToString(p2 , "Table p2 : ");
    desc += tabToString( z2, "Table z2 : ");
    return desc;
}
}

```

Listing II.4 – La sous-classe NI4 pour la méthode M4.

```

import umontreal.iro.lecuyer.probdist.*;
import umontreal.iro.lecuyer.util.*;

```

```

public class NI4 extends NortaInitMix
{
    private double tolerance; // Seuil de tolérance
    public NI4 (double rX,
                DiscreteDistributionInt dist2,
                double tr,
                int nPoints,
                double tolerance)
    {
        super(rX, dist2, tr, nPoints);
        this.tolerance = tolerance;
        computeParams(); // Appelle la méthode computeParams
    }
    public double computeCorr ()
    {
        final double ITMAX = 100; // Nombre maximum d'itérations
        final double EPS = 1.0e-15; // Précision de la machine
        double b; // La solution finale
        double a, c, min1, min2, fa, fb, fc, pp, q, rrr, s,
                tolerancel, xm, x1, x2;
        double e = 0.0, d = 0.0;
        // Constantes
        double cc = rX * sd1 * sd2;
        double ccc = cc + mu1 * mu2;
        if (rX == 0)
            return 0.0;
        if (rX > 0) { // Oriente la recherche et initialise a, b, c
            x1 = 0.0;
            x2 = 1.0;
            a = x1;
            b = x2;
            c = x2;
            fa = - cc;
            fb = integ (b) - ccc;
        }
        else {
            x1 = -1.0;
            x2 = 0.0;
            a = x1;
            b = x2;
            c = x2;
            fa = integ (a) - ccc;
            fb = - cc;
        }
        fc = fb;
        for (int i = 1; i <= ITMAX; i++) { // Commencer la recherche
            if ((fb > 0.0 && fc > 0.0) || (fb < 0.0 && fc < 0.0)) {
                // Renommer a, b, c et ajuster l'intervalle d.
                c = a;
                fc = fa;
                e = d = b - a;
            }
            if (Math.abs (fc) < Math.abs (fb)) {
                a = b;
                b = c;
                c = a;
                fa = fb;
                fb = fc;
                fc = fa;
            }
        }
    }
}

```

```

// Seuil de tolérance final
tolerance1 = 2.0 * EPS * Math.abs (b) + 0.5 * tolerance;
xm = 0.5 * (c - b);
if (Math.abs (xm) <= tolerance1 || fb == 0.0)
    return b;
if (Math.abs (e) >= tolerance1 && Math.abs (fa) > Math.abs (fb)) {
    s = fb / fa; // Utiliser l'interpolation quadratique inverse
    if (a == c) {
        pp = 2.0 * xm * s;
        q = 1.0 - s;
    }
    else {
        q = fa / fc;
        rrr = fb / fc;
        pp = s * (2.0 * xm * q * (q - rrr) - (b - a) * (rrr - 1.0));
        q = (q - 1.0) * (rrr - 1.0) * (s - 1.0);
    }
    if (pp > 0.0)
        q = -q; // Vérifier les limites de l'intervalle
    pp = Math.abs (pp);
    min1 = 3.0 * xm * q - Math.abs (tolerance1 * q);
    min2 = Math.abs (e * q);
    if (2.0 * pp < (min1 < min2 ? min1 : min2)) {
        e = d; // Accepter l'interpolation
        d = pp / q;
    }
    else { // Échec de l'interpolation, utiliser bisection
        d = xm;
        e = d;
    }
}
else { // Limites convergent lentement, alors utiliser la bisection
    d = xm;
    e = d;
}
a = b; // a devient la meilleure solution
fa = fb;
if (Math.abs (d) > tolerance1)
    b += d; // Évalue la nouvelle solution
else {
    if (xm > 0) b += Math.abs (tolerance1);
    else
        b += -Math.abs (tolerance1);
}
fb = integ (b) - ccc;
}
return b;
}
// Pour afficher les inputs
public String toString()
{
    String desc = super.toString();
    desc += "tolerance : " + tolerance + "\n";
    return desc;
}
}

```