

Titre: Génération de scénarios pour la demande en personnels durant
Title: plusieurs périodes

Auteur: Antoine Legrain
Author:

Date: 2011

Type: Mémoire ou thèse / Dissertation or Thesis

Référence: Legrain, A. (2011). Génération de scénarios pour la demande en personnels
Citation: durant plusieurs périodes [Mémoire de maîtrise, École Polytechnique de
Montréal]. PolyPublie. <https://publications.polymtl.ca/687/>

 **Document en libre accès dans PolyPublie**
Open Access document in PolyPublie

URL de PolyPublie: <https://publications.polymtl.ca/687/>
PolyPublie URL:

**Directeurs de
recherche:** François Soumis, & Richard Labib
Advisors:

Programme: Génie industriel
Program:

UNIVERSITÉ DE MONTRÉAL

GÉNÉRATION DE SCÉNARIOS POUR LA DEMANDE EN PERSONNELS
DURANT PLUSIEURS PÉRIODES

ANTOINE LEGRAIN
DÉPARTEMENT DE MATHÉMATIQUES ET GÉNIE INDUSTRIEL
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

MÉMOIRE PRÉSENTÉ EN VUE DE L'OBTENTION DU DIPLÔME DE
MAÎTRISE ÈS SCIENCES APPLIQUÉES
(MATHÉMATIQUES APPLIQUÉES)
DÉCEMBRE 2011

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Ce mémoire intitulé :

GÉNÉRATION DE SCÉNARIOS POUR LA DEMANDE EN PERSONNELS
DURANT PLUSIEURS PÉRIODES

présenté par : M. LEGRAIN Antoine, Ms.Ing.

en vue de l'obtention du diplôme de : Maîtrise ès Sciences Appliquées

a été dûment accepté par le jury constitué de :

M. DESAULNIERS Guy, Ph.D., président.

M. SOUMIS François, Ph.D., membre et directeur de recherche.

M. LABIB Richard, Ph.D., membre et co-directeur de recherche.

M. GENDREAU Michel, Ph.D., membre.

Remerciements

Je souhaite remercier tout d'abord François Soumis et Richard Labib d'avoir accepté de diriger ma maîtrise, de m'avoir aidé tout au long de mon passage à l'École Polytechnique de Montréal.

Je remercie aussi KRONOS et le CRSNG de m'avoir subventionné pendant la durée de ma maîtrise.

Je remercie le GERAD et son administration pour toute l'aide qu'ils nous fournissent au quotidien.

Je remercie tous mes camarades de laboratoire (Anthony Guillou, Quentin Lequy et Rémi Pacqueau en particulier) qui m'ont beaucoup aidé et m'ont bien conseillé.

Merci à tous les joueurs de Tarot avec qui j'ai passé de bons moments et qui animent le GERAD.

Merci à tous mes amis de Montréal avec qui je me suis bien amusé et détendu.

Merci à tout ceux que j'ai pu oublier.

Résumé

La recherche opérationnelle a permis de grandes avancées dans la fabrication d'horaires. Nous sommes capables maintenant de traiter de très grands problèmes de plusieurs millions de variables : l'être humain est incapable de résoudre de tels problèmes de manière aussi efficace que les logiciels actuels. Ces derniers prennent comme données les conventions collectives et la demande prévisionnelle en nombres d'employés. À partir de ces informations, ils construisent les jours de travail et de repos, les quarts de travail puis ils allouent les quarts et les tâches aux employés. Lorsque ces informations ne sont pas exactes, en particulier la demande, tout le processus du logiciel s'en trouve affecté et devient moins précis. Ce mémoire s'attaque à ce dernier problème et définit un modèle pour représenter l'incertitude de la demande prévisionnelle.

Cette incertitude se traduit par une variation stochastique en pourcentage de la demande. Ces variations sont représentées par plusieurs types de perturbations : on tire de manière aléatoire l'apparition de ces perturbations ainsi que leur amplitude. Nous obtenons ainsi un modèle qui est modifiable et qui tente de représenter la réalité. Chaque type de perturbation est décrit par sa longueur, sa probabilité d'apparition ainsi que les paramètres qui régissent la loi log-normale de l'amplitude.

Par la suite, une méthode d'estimation des paramètres du modèle est développée : elle calcule différents moments de la variation pour calibrer correctement les perturbations. Elle utilise des données historiques, suppose que ces dernières suivent le modèle et tente de retrouver les paramètres correspondant à cette configuration. Les résultats ainsi obtenus sont assez variables et plus ou moins bons en fonction des instances. Lorsque la variance de l'incertitude est faible ($< .002$), nous obtenons des bons résultats : le coût du planning est réduit de 1%. Pour des variances plus importantes, les résultats sont plus variables, mais il y a toujours des gains au niveau du coût.

Abstract

Operations research has really increased the performance of workforce scheduling. We are now able to process huge problems with millions of variables : a human being cannot solve this kind of problems as efficiently as current software. The inputs of those softwares are the forecasted workforce demand and collective agreements. Days off, work days and shifts are built from those pieces of information, then shifts and tasks are assigned to the employees. The operation of the software is based on this information, if this one is not accurate, the whole process is affected and becomes less efficient. This master thesis tackles this last problem and defines a model to represent the uncertainty of the demand.

This uncertainty is represented by random variations expressed in percentage of the demand. Those variations are modelled by several kinds of perturbations: we are drawing randomly the appearance of these perturbations and their amplitude. We are getting a model which can be modified and which tries to represent the reality. Each kind of perturbation is described by its length, its probability of appearing and the parameters which control the log-normal distribution of the amplitude.

Then we are developing a method to infer the parameters of the model: we are computing different moments of the variation to size correctly the perturbations. Historic data are used to estimate the original parameters of those data. So the computed results are volatile and more or less accurate according to the instances. If the variance of uncertainty is low ($< .002$), results are good and there is a real gain (1%) for the cost of the scheduling. If the variance is more important, results are more volatile but there is still a gain for the cost.

Table des matières

Remerciements	iii
Résumé	iv
Abstract	v
Table des matières	vi
Liste des tableaux	viii
Liste des figures	ix
Liste des sigles et abréviations	x
Chapitre 1 INTRODUCTION	1
Chapitre 2 REVUE DE LITTÉRATURE	3
Chapitre 3 LE MODÈLE	7
3.1 Modélisation des coefficients de variation	9
3.2 Lois Normale et Log-normale	12
3.2.1 Définitions et formules	12
3.2.2 Génération	13
3.3 Modélisation des perturbations	15
3.3.1 Modélisation de la tendance de la journée	15
3.3.2 Modélisation du bruit périodique	15
3.3.3 Modélisation des corrélations	16
3.3.4 Modèle général	18
Chapitre 4 ESTIMATION DES PARAMÈTRES	20
4.1 Estimateurs classiques	20
4.2 Estimation globale de la variance	22
4.2.1 Variance des variations Θ_p	25

4.2.2	L'annulation des effets de la tendance	27
4.2.3	Autres formules des variances	29
4.2.4	Choix des écarts	31
4.2.5	Méthodes de résolution	32
4.3	Estimation des moments	36
4.3.1	Moment d'ordre un des variations Θ_p	36
4.3.2	Moment d'ordre trois des variations Θ_p	39
4.3.3	Moment d'ordre quatre des variations Θ_p	43
4.4	Estimation de tous les paramètres	47
4.4.1	Recherche par dichotomie	47
4.4.2	Analyse de sensibilité	50
4.5	Résumé	52
Chapitre 5 RÉSULTATS DE L'ESTIMATION		54
5.1	Instances	54
5.2	Coût des instances	55
5.3	Résultats	58
Chapitre 6 CONCLUSION		61
6.1	Synthèse des travaux	61
6.2	Limites de la méthode proposée	62
6.3	Améliorations futures	63
RÉFÉRENCES		64

Liste des tableaux

TABLEAU 3.1	Pourcentage moyen de rejet de l'hypothèse H	14
TABLEAU 4.1	Résultats pour la première méthode de l'écart-type de S_j	35
TABLEAU 4.2	Résultats pour la deuxième méthode de l'écart-type de S_j	35
TABLEAU 4.3	Résultats de l'écart-type de $\mathbb{E}[\Theta_p]$	38
TABLEAU 4.4	Résultats de l'écart-type de $[\mathbb{L}_3]^{\frac{1}{3}}$	43
TABLEAU 4.5	Résultats de l'écart-type de U_j	47
TABLEAU 4.6	Résultats de l'écart-type des estimations $\hat{\beta}_j$ et $\hat{\sigma}_j$	51
TABLEAU 5.1	Instances	54
TABLEAU 5.2	Gain de l'optimisation stochastique	57
TABLEAU 5.3	Estimations des paramètres des instances	59
TABLEAU 5.4	Comparaison entre les anciens et les nouveaux gains	59

Liste des figures

FIGURE 3.1	Différentes formes du coefficient de variation θ_p	10
FIGURE 3.2	Exemple de coefficient de variation $\theta_p^{tendance}$	16
FIGURE 3.3	Exemple de coefficient de variation θ_p^{bruit}	17
FIGURE 3.4	Exemple des coefficients de variation θ_p^{1heure} et $\theta_p^{4heures}$	18
FIGURE 3.5	Exemple du coefficient de variation $\theta_p = \theta_p^{bruit} \theta_p^{1heure} \theta_p^{4heures} \theta_p^{tendance}$	19
FIGURE 4.1	Exemple de variation $\Theta_{p+8}^{4heures} - \Theta_p^{4heures}$	23
FIGURE 4.2	Exemple de variation $\Theta_{p+8}^{1heure} - \Theta_p^{1heure}$	24
FIGURE 4.3	Variation $\Theta_{p+32}^{tendance} - \Theta_p^{tendance}$	28
FIGURE 4.4	Variation $\Theta_{p+1}^{1heure} - \Theta_p^{1heure}$	31
FIGURE 4.5	Représentation d'une observation avec de petites perturbations	37
FIGURE 4.6	Représentation d'une observation avec de grandes perturbations	37
FIGURE 4.7	Représentation de f pour différentes valeurs de S_j et pour $U_j = 0, 1$	50
FIGURE 5.1	Représentation de la demande prévisionnelle	55

Liste des sigles et abréviations

E	Espérance mathématique
IP	Integer Programming
Var	Variance

Chapitre 1

INTRODUCTION

La recherche opérationnelle permet aux industriels de mieux gérer et contrôler leurs ressources. Les ressources humaines représentent un coût important pour bon nombre d'entreprises ; la fabrication des horaires se doit donc d'être optimisée de manière efficace et assez rapide pour parvenir à réduire de façon substantielle ces coûts. Le problème de génération d'horaires a déjà été bien exploré, les modèles mathématiques qui le décrivent sont très précis et de plus en plus flexibles.

Il existe deux grands types de problèmes. Le premier décrit des problèmes où les tâches à affecter sont ininterrompibles : par exemple on ne peut pas remplacer le pilote d'un avion pendant son vol : c'est la même personne qui doit effectuer la tâche dans sa totalité. Le deuxième type comprend les problèmes où les tâches sont interrompibles : un caissier dans un supermarché peut par exemple être remplacé à n'importe quel moment de la journée. Ce mémoire s'intéresse à ce dernier genre de problèmes et aux méthodes de résolution.

Ces dernières sont maintenant très efficaces : l'augmentation de la puissance de calculs des ordinateurs et l'amélioration des algorithmes d'optimisation ont réduit les temps de résolution et permis de traiter des problèmes de plus en plus complexes. Il devient maintenant difficile de baisser les coûts des ressources humaines en améliorant les algorithmes de résolution et les modèles.

Tous ces modèles prennent une demande prévisionnelle en nombre d'employés pour fabriquer les horaires. Ainsi une mauvaise prévision de la demande entraîne forcément un planning inapproprié à la réalité : l'entreprise se retrouve avec trop ou pas assez d'employés selon les moments. La qualité de la planification dépend fortement de la modélisation de la demande. Actuellement les demandes prévisionnelles sont très bien estimées, mais l'incertitude restante est ignorée dans les modèles. Il faut donc modéliser cette erreur, l'insérer dans les modèles déjà existants et développer des algorithmes de résolution différents.

L'incertitude de la demande a déjà été prise en compte dans des problèmes de gestion de production d'électricité. Dans le domaine de la génération d'horaires, des modèles stochastiques sont surtout apparus pour le secteur de l'aviation. La première difficulté est d'obtenir un modèle de l'incertitude fiable et surtout avec peu de paramètres. S'il y a trop de paramètres, il devient très difficile de calibrer le modèle. Il faut aussi pouvoir confronter ce dernier à des données réelles. Ces dernières sont confidentielles et souvent manquent de précision : on a les chiffres sur une journée ou sur chaque demi-journée. Ce texte se consacrera à l'élaboration d'un modèle de la demande et d'un générateur de données. Il est important de se fabriquer une banque de données correctement dimensionnée. Nous allons exposer le modèle et par la suite une technique de calibration.

Plan du mémoire : Un état rapide de l'art est dressé en premier lieu. Nous présentons ensuite un modèle d'incertitude de la demande prévisionnelle : les «incidents» qui proviennent de manière imprévisible durant la journée tentent d'être modélisés. Puis des idées pour estimer différents paramètres du modèle sont présentées et une technique de calibration du modèle est finalement choisie. Enfin tous les résultats obtenus sont présentés et critiqués.

Chapitre 2

REVUE DE LITTÉRATURE

La construction d'un planning horaire est toujours très délicate. Ernst *et al.* (2004) ont réalisé une bibliographie très détaillée du sujet. Dantzig (1954) a présenté une première formulation du problème : il utilise un modèle de recouvrement pour résoudre ce problème. Par la suite, plusieurs auteurs ont proposé d'autres formulations. Par exemple Aykin (1996) reprend le modèle de recouvrement et en propose une formulation implicite : celle-ci sera tout le temps reprise par la suite car elle simplifie énormément le problème. Bechtold et Jacobs (1990) réécrivent le modèle afin d'éliminer le problème de transport utilisé dans l'affectation des pauses. Les problèmes, par exemple d'affectation des pauses ou des heures supplémentaires, peuvent être résolus plus tard. Bard *et al.* (2007) considèrent un problème avec quarts de travail dans un centre de traitement de la poste américaine et montrent comment décomposer le problème de construction d'un horaire en deux problèmes qui se succèdent : la construction des quarts de travail et le recours. Ce recours contient, par exemple l'affectation des heures supplémentaires. Cette modélisation permet de prendre en compte l'arrivée d'informations entre les deux phases : c'est un modèle stochastique.

Birge et Louveaux (1997) différencient les problèmes déterministes de ceux stochastiques et présentent des techniques de résolution pour les problèmes stochastiques. Les programmes déterministes permettent de n'avoir qu'une décision à prendre en fonction de données immuables, alors que les programmes stochastiques permettent de prendre plusieurs décisions une fois un scénario choisi. Les scénarios sont des situations différentes sélectionnées afin de représenter la réalité. L'optimisation stochastique permet de construire un planning plus proche de la réalité. Pour cela, elle prend en compte deux coûts : le coût planifié de la fabrication des quarts de travail et celui de l'espérance du recours. L'espérance est calculée sur l'ensemble des scénarios choisis et le recours est optimisé pour chaque scénario. L'objectif est d'obtenir une bonne fonction du coût du recours et celle-ci dépend de la qualité des scénarios pris en compte

dans ce calcul. Ainsi une bonne représentation de la réalité permet une estimation précise du coût opérationnel du planning. La modélisation et la génération de ces scénarios sont donc deux décisions décisives par rapport à la qualité de l'horaire ainsi obtenu.

L'optimisation stochastique des horaires de personnels est un domaine assez jeune et il existe peu d'articles sur ce sujet. Quelques auteurs se sont quand même aventurés dans ce domaine.

Yen et Birge (2006) modélisent les différents retards pour une compagnie aérienne : retard de l'avion, retard de l'équipage par exemple. Ces derniers sont modélisés par des processus stochastiques corrélés. Ce point de vue est obligatoire mais il complexifie beaucoup le problème.

De la même façon, dans le problème du centre de traitement de la poste américaine, Bard *et al.* (2007) modélisent la demande en employés par un processus stochastiques. Leur modélisation diffère peu de celle de Yen et Birge (2006).

Finalement Pacqueau (2011) présente dans son mémoire de maîtrise un algorithme de résolution des problèmes d'horaire avec quarts de travail et avec une demande en employés stochastique. L'avantage de son algorithme est de pouvoir prendre un grand nombre (500) de scénarios comme données entrantes.

Ce mémoire a pour objectif de générer des scénarios représentatifs de la demande réelle. Ces derniers sont destinés à être utilisés par des algorithmes stochastiques. Une distribution classique choisie pour modéliser une incertitude est la loi normale, mais lorsque les variables sont positives, il est préférable de choisir une loi log-normale. Crow et Shimizu (1988) exposent une multitude de problèmes où cette distribution est choisie. Elle est souvent utilisée pour modéliser des problèmes complexes où les facteurs humain et naturel sont importants.

De plus, il est difficile de mesurer la demande en employés, elle n'est jamais mesurée ainsi. On l'estime en regardant les commandes entrantes, le travail à effectuer, le nombre de produits vendus, par exemple. Dans tous les cas il faut la transformer en une demande en nombre d'employés et réduire le nombre de données. Zhang et Bard (2005) nous montrent comment transformer dans un centre de traitement de courrier le flux entrant en une demande employés discrétisée sur des périodes grâce à un algorithme du type FIFO (First In First Out).

Ensuite il faut générer des scénarios et en sélectionner. Kaut et Wallace (2007) ont réalisé un état de l'art sur la génération d'arbre de scénarios. De tels arbres permettent de discrétiser la distribution de processus stochastiques et de les corrélérer. Il apparaît qu'il existe beaucoup de techniques générales très différentes.

Un premier ensemble de méthodes est basé sur le tirage des scénarios selon une distribution.

- Certains auteurs comme Loretan (1997) se servent d'une analyse par composantes principales pour réduire les paramètres décrivant la corrélation entre les états du processus. Puis ils utilisent ces vecteurs pour générer des arbres de scénarios corrélés pour ce processus stochastique.
- D'autres auteurs comme Cario et Nelson (1997) utilisent les mêmes méthodes, mais au lieu de corrélérer les scénarios avec les composantes principales, ils prennent en compte directement la distribution marginale ou la matrice de corrélation.
- Un nombre important d'auteurs développent des techniques comparables aux précédentes. Il se trouve qu'il est très souvent difficile d'obtenir les distributions marginales. Il est beaucoup plus facile d'utiliser les moments de la distribution. Vale et Maurelli (1983) se servent des quatre premiers moments pour générer des arbres simulant le score de tests psychologiques. Smith (1993) et Høyland *et al.* (2003) illustrent d'autres méthodes et exemples utilisant les moments pour choisir les scénarios.
- Dupačová *et al.* (2000) montrent comment générer des arbres de scénarios à partir d'un ensemble de scénarios, des chemins donnés.
- Pflug (2001) construit une approximation du processus stochastique en minimisant une fonction d'erreur.

Un deuxième point de cette génération est la réduction de la taille des ensembles de scénarios, des arbres de scénarios. Cet aspect est aussi très important, car la majorité des modèles d'optimisation stochastique ne peut être résolue qu'avec un nombre restreint de scénarios. En effet, le temps de résolution sera beaucoup trop grand si la taille n'est pas correctement dimensionnée.

- Dupačová *et al.* (2003) présentent une technique très intéressante pour sélectionner des scénarios. Ils introduisent une norme sur un espace de probabilités et montrent un algorithme permettant de trouver la distance optimale entre deux distributions

pour cette norme. Cet algorithme est plutôt simple à implémenter.

- Dantzig et Infanger (1992) montrent que les événements peuvent être générés au fur et à mesure de l'optimisation. L'algorithme enlève ou rajoute des scénarios en fonction de leur pertinence, c'est-à-dire par rapport à leur impact sur l'objectif.

La génération de scénarios est donc une étape primordiale pour l'obtention d'un modèle efficace et satisfaisant. Ce mémoire s'inspirera des idées générales de génération de scénarios mais d'aucune de ces méthodes en particulier.

La principale contribution de ce mémoire est l'élaboration d'un modèle pour générer des scénarios à partir d'une base de données historiques. Ce modèle est valide pour la fabrication d'horaire avec quarts de travail ; la modélisation des corrélations est assez poussée. Ce mémoire apporte aussi une méthode d'estimation des paramètres du modèle.

Chapitre 3

LE MODÈLE

Les données que l'on a actuellement pour fabriquer des horaires sont des demandes prévisionnelles estimées pour des périodes de quinze minutes. La réalité diffère forcément de celle-ci comme on ne peut pas prévoir le futur. Il existe ainsi une erreur entre les données que nous utilisons pour construire un planning et les valeurs réelles de celles-ci. L'objectif de ce modèle est d'estimer l'incertitude de la demande prévisionnelle et d'obtenir une représentation assez fiable de la réalité tout en ayant un nombre de paramètres à calibrer assez restreint. Comme les problèmes de planification d'horaires et les techniques d'optimisation utilisées sont très complexes, l'optimisation stochastique des plannings est un réel défi. Les techniques utilisent actuellement la programmation linéaire en nombres entiers et la génération de colonnes (Gamache *et al.* (1999)). Il est impératif de conserver ces outils en programmation stochastique car ce sont les plus puissants en ce moment. Nous résoudrons les problèmes d'horaires avec la programmation stochastique en nombres entiers sur un ensemble fini de scénarios. Nous représenterons donc la demande stochastique par un ensemble de scénarios que le modèle doit produire au final.

Nous avons choisi de modéliser cette incertitude par un coefficient de variation par rapport à la demande prévisionnelle. Pourquoi privilégier un tel modèle par rapport, par exemple à un modèle additif du coefficient de variation ou à une modélisation directe de la demande ? Notre expertise est la recherche opérationnelle : l'optimisation de la fabrication des horaires de personnels. Nous ne savons pas comment prévoir une demande. La situation n'est pas stationnaire, elle évolue d'année en année. On ne peut pas regarder directement l'historique, il faut évaluer par exemple l'impact de l'arrivée d'un nouveau concurrent ou les tendances du marché pour fabriquer une prévision de qualité. Cette dernière comprend aussi des facteurs spécifiques à chaque entreprise et ce sont les planificateurs qui ont les connaissances suffisantes pour prédire la tendance moyenne en tenant compte des données historiques et des autres facteurs qui

influencent la journée considérée. Par exemple la stratégie commerciale, le lancement d'un nouveau produit, d'une nouvelle promotion sont des facteurs internes à chaque entreprise et influencent énormément la demande prévisionnelle. Si nous souhaitons réaliser un modèle un tant soit peu général, nous ne devons pas utiliser un modèle qui dépend de la prévision. Des modélisations additive ou directe de la réalisation sont forcément dépendantes de la demande prévisionnelle. Par exemple pour un modèle additif, une variation d'un employé n'a absolument pas la même importance si la prévision était de deux employés ou de quinze : la charge de travail supplémentaire pour chaque employé n'est pas la même. Un modèle multiplicatif nous semble a priori une bonne idée car on s'affranchit ainsi des facteurs externes à notre étude et nous nous concentrons sur la modélisation des événements totalement imprévisibles comme la météo ou le retard d'un employé. Et surtout ce modèle se transpose sur n'importe quelle autre prévision de l'entreprise afin de générer des scénarios.

Nous considérons maintenant que les données historiques, c'est-à-dire la demande prévisionnelle ainsi que sa réalisation sont connues. Les coefficients de variation peuvent être calculées à partir de cet historique. Dans ce cas, pourquoi construire un modèle multiplicatif ? Nous pourrions considérer que l'ensemble de nos coefficients de variation historiques représente la réalité de façon assez fiable, c'est-à-dire que nous avantageons ce qui s'est déjà réalisé. Comme les coefficients de variation observés n'ont pas de raison de se reproduire de la même manière dans le futur, il est plus naturel de déduire un modèle probabiliste de cet historique. De plus, de telles données sont peu archivées et ne sont pas toujours disponibles en quantités suffisantes ; il est très difficile d'obtenir un historique avec une précision de quinze minutes. Il est ainsi important de créer un modèle permettant de générer autant de demandes que nous souhaitons pour avoir suffisamment de scénarios de qualité pour l'optimisation stochastique. En effet, l'optimisation stochastique sur un grand nombre de variables avec un nombre trop petit de scénarios risque de produire une solution qui s'adapte aux scénarios mais qui n'est pas robuste face aux variations futures.

Présentons maintenant la formulation mathématique : on travaille sur une journée qu'on divise en quatre-vingt seize périodes de quinze minutes, ce qui est la décomposition habituelle. Soit D la demande en nombre d'employés. Cette demande n'est pas forcément entière. Si la demande est nulle à une certaine période, l'activité est fermée

et personne ne travaille. Nous considérons que cet évènement a été pris en compte lors du traitement de l'information, le coefficient de variation est fixé à 1 dans ce cas. La demande est donc strictement positive dans tous les autres cas.

Soit D_p la demande observée à la période p .

Soit $Dmoy_p$ la demande prévisionnelle à la période p .

Soit θ_p^0 le coefficient de variation de D_p par rapport à $Dmoy_p$.

$$\theta_p^0 = \frac{D_p}{Dmoy_p}, \forall p = 1 \dots 96$$

Nous allons présenter par la suite la modélisation de ce coefficient de variation θ_p^0 . Ces coefficients de variation permettront de générer des scénarios de la demande à partir d'une demande prévisionnelle.

3.1 Modélisation des coefficients de variation

Il est difficile de modéliser ces coefficients de variation. En effet, les coefficients de variation θ_p^0 sont dépendants entre eux : une période dépend des autres. Comment modéliser cette dépendance ? Si on essaie de trouver la loi conjointe de cette famille de variables aléatoires $(\theta_p^0)_{p=1\dots 96}$, on obtient une matrice carrée de corrélation de dimension 96×96 ; on a donc $96^2 \approx 10\,000$ paramètres pour définir ce modèle. Nous pourrions réaliser une analyse par composantes principales (Jolliffe (2002)) pour réduire ce nombre de paramètres, car nous n'aurons jamais assez de données pour estimer correctement tous ces paramètres. Une telle analyse a l'avantage de se mettre facilement en place : il existe plusieurs logiciels (matlab par exemple) gérant très bien cette méthode. Nous cherchons cependant un modèle qui peut facilement se transporter d'une prévision à une autre. Il n'est pas possible d'ajuster régulièrement les paramètres du modèle car, comme nous l'avons précisé un peu plus haut, il est difficile d'obtenir les données nécessaires au dimensionnement. De plus, une analyse par composantes principales risque d'incorporer les erreurs contenues dans l'historique dans le modèle. En effet, il peut y avoir des erreurs humaines dans la prévision (le planificateur a surestimé les effets d'une promotion), dans ce cas l'analyse par composantes principales risque de considérer cette erreur comme une variation à prendre en compte. Il en est de même pour des phénomènes observés de manière temporaire, par exemple pendant quelques jours. Nous avons choisi un modèle plus approximatif,

mais avec beaucoup moins de paramètres. Ce modèle est aussi plus régulier et ne prend pas en compte ces erreurs qui restent alors dans le bruit non capté par le modèle. Nous ne sommes pas certains que notre modèle soit meilleur, mais nous l'espérons.

Plutôt que de définir toutes les dépendances imaginables, nous avons préféré choisir quatre types de dépendances qui nous semblaient assez représentatifs de la réalité. Nous avons construit un modèle par perturbations. Nous considérons qu'en choisissant quatre types de perturbations, nous représentons assez bien les dépendances. Une perturbation peut durer une période (quinze minutes), quatre périodes (une heure), seize périodes (quatre heures) ou quatre-vingt-seize périodes (une journée). Chaque perturbation a la même amplitude pendant toute sa durée et a une certaine probabilité d'apparaître à chaque période. La forme de la perturbation est importante et doit sûrement dépendre des entreprises. Il est tout à fait possible de choisir d'autres formes que celles présentées dans la figure (3.1).

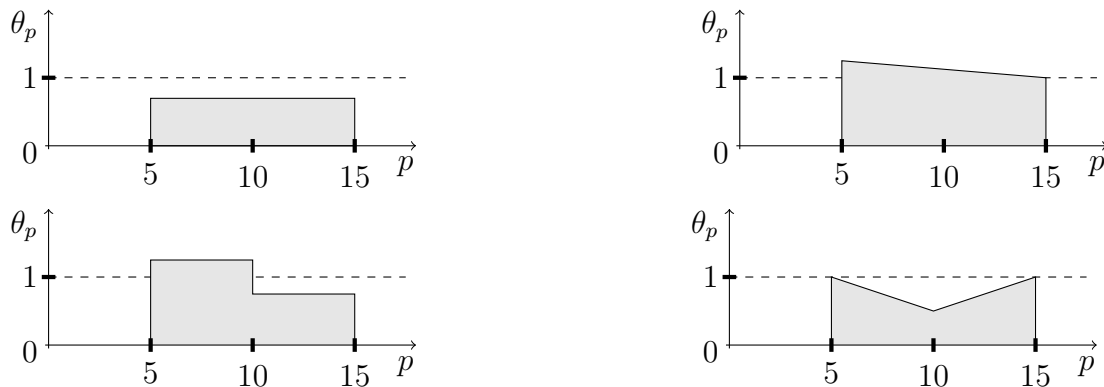


FIGURE 3.1 Différentes formes du coefficient de variation θ_p

Comment motiver un choix de patron de perturbation ? Le choix n'a pas tant d'importance. Les formules et techniques présentées par la suite pourraient être utilisées de la même façon, ce sont juste les coefficients qui seront modifiés. Par contre la forme modifie énormément les corrélations ; en fonction des données il faut tester les différents patrons, peut-être en mélanger aussi. La forme rectangulaire est la plus simple imaginable et capte forcément une partie de la corrélation. La forme avec deux états opposés (en bas à gauche) permet de modéliser de façon un peu plus fine, elle transmet l'idée suivante : s'il y a eu moins de monde pendant une période, peut-être à cause du beau temps, ces personnes viendront plus tard. De la même façon, la forme

en triangle (en bas à droite) traduit une autre corrélation réalisable : avec l'arrivée de la pluie les personnes restent chez elles et bougent moins, le pic correspond au moment où il pleut le plus, ensuite les clients recommencent doucement à arriver ; la transition de cette perturbation n'est pas brutale mais ce fait peu à peu. En tout cas, le choix de la forme ne peut pas se faire totalement a priori, il doit y avoir une analyse des résidus une fois le modèle dimensionné afin de valider les différentes hypothèses. Nous prendrons la forme rectangulaire pour la suite, elle est la plus simple à mettre en équations. Il faut noter qu'il reste a priori une erreur que nous négligerons dans ce modèle : les corrélations entre les périodes ne s'expriment pas forcément toutes comme des rectangles.

Nous avons choisi une loi log-normale pour représenter chacun de ces coefficients de variation. Nous avons hésité entre cette loi et une loi gamma ; en effet ces deux lois permettent de modéliser des grandeurs positives ce qui est le cas ici. Le coefficient de variation ne peut être négative, car la demande en nombre d'employés est forcément positive. La loi log-normale a été retenue, car les coefficients de variation que nous cherchons à modéliser semblent se multiplier entre elles et que cette loi permet facilement de décomposer plusieurs variables aléatoires qui se multiplient. Il sera donc plus facile d'estimer les paramètres de cette loi.

Nous avons finalement un modèle qui ressemble grossièrement au modèle ci-dessous.

$$\theta_p^0 = \frac{D_p}{D_{moy_p}} = \theta_p * (1 + \varepsilon_p)$$

$$\theta_p = \theta_p^{bruit} \theta_p^{1heure} \theta_p^{4heures} \theta_p^{tendance}$$

θ_p^{bruit} représente le coefficient de variation des perturbations d'une période

θ_p^{1heure} représente le coefficient de variation des perturbations de quatre périodes (3.1)

$\theta_p^{4heures}$ représente le coefficient de variation des perturbations de seize périodes

$\theta_p^{tendance}$ représente le coefficient de variation des perturbations de quatre-vingt-seize périodes

ε_p représente l'erreur commise par le modèle à chaque période

L'erreur est considérée comme négligeable devant les autres perturbations. Il faut par la suite utiliser les données pour définir des conditions où cette hypothèse est valide.

Nous travaillerons donc avec le coefficient de variation périodique θ_p dans la suite de ce mémoire. La loi log-normale a été choisie pour modéliser ces coefficients de variation. Faisons donc un bref rappel de la loi log-normale et de quelques unes de ses propriétés avant de présenter plus précisément le modèle.

3.2 Lois Normale et Log-normale

Le livre de Saporta (2006) parcourt un ensemble très important de notions en probabilités et en statistiques. Nous en présentons quelques unes par la suite.

3.2.1 Définitions et formules

Soit Y une variable aléatoire réelle à valeurs positives. On dit que Y suit une loi log-normale $\ln \mathcal{N}(\mu, \sigma^2)$ de paramètres μ et σ , si la variable aléatoire $X = \ln Y$ suit une loi normale $\mathcal{N}(\mu, \sigma^2)$ de moyenne μ et d'écart-type σ . On a ainsi les formules suivantes :

$$\begin{aligned}\mathbb{E}[Y] &= e^{\mu + \frac{\sigma^2}{2}} \\ \text{Var}[Y] &= (e^{\sigma^2} - 1)(\mathbb{E}[Y])^2\end{aligned}$$

On dit que la loi normale est additive. En effet, considérons une famille de deux variables aléatoires réelles et indépendantes X_1 et X_2 qui suivent respectivement les lois normales $\mathcal{N}(\mu_1, \sigma_1^2)$ et $\mathcal{N}(\mu_2, \sigma_2^2)$; alors la variable aléatoire $X_1 + X_2$ suit une loi normale $\mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$. De la même manière, soit une famille de deux variables aléatoires réelles et indépendantes Y_1 et Y_2 qui suivent respectivement les lois log-normales $\ln \mathcal{N}(\mu_1, \sigma_1^2)$ et $\ln \mathcal{N}(\mu_2, \sigma_2^2)$; alors la variable aléatoire $Y_1 Y_2$ suit une loi log-normale $\ln \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

Généralisons maintenant. Soit une famille de n variables aléatoires réelles et indépendantes $(X_i)_{i=1..n}$ qui suivent respectivement des lois normales $\mathcal{N}(\mu_i, \sigma_i^2)$; alors la variable aléatoire $\sum_{i=1}^n X_i$ suit une loi normale $\mathcal{N}(\sum_{i=1}^n \mu_i, \sum_{i=1}^n \sigma_i^2)$. On a immédiatement

les relations suivantes :

$$\begin{aligned}\mathbb{E}\left[\sum_{i=1}^n X_i\right] &= \sum_{i=1}^n \mathbb{E}[X_i] = \sum_{i=1}^n \mu_i \\ \text{Var}\left[\sum_{i=1}^n X_i\right] &= \sum_{i=1}^n \text{Var}[X_i] = \sum_{i=1}^n \sigma_i^2\end{aligned}$$

Les dernières formules que nous verrons ici sont celles qui génèrent les moments. Soit X une variable aléatoire réelle qui suit la loi normale $\mathcal{N}(\mu, \sigma^2)$. Sa fonction génératrice des moments $M_X(t)$ vaut $M_X(t) = \exp(\mu t + \sigma^2 \frac{t^2}{2})$ et on a :

$$\mathbb{E}[X^n] = \frac{\partial^n M_X}{\partial t^n}(0)$$

On obtient ainsi les relations suivantes :

$$\begin{aligned}\mathbb{E}[X] &= \mu \\ \mathbb{E}[X^2] &= \mu^2 + \sigma^2 \\ \mathbb{E}[X^3] &= \mu^3 + 3\mu\sigma^2 \\ \mathbb{E}[X^4] &= \mu^4 + 6\mu^2\sigma^2 + 3\sigma^4\end{aligned}$$

Ces dernières formules nous seront très utiles par la suite.

3.2.2 Génération

Tous les résultats s'appuient sur la génération de la loi log-normale. Nous allons vérifier ici que notre générateur de nombres pseudo-aléatoires n'est pas biaisé. Nous utilisons le générateur de base de Java : c'est un générateur congruentiel linéaire, il est inclus dans la classe Random. Pour contrôler sa qualité, nous vérifierons l'indépendance des valeurs produites pour la loi uniforme standard $U(0, 1)$ grâce à un test du χ^2 présenté par Plackett (1983).

Soit une famille de m classes équiprobables $(C_i)_{1..m}$. Pour la loi uniforme standard $U(0, 1)$, nous prenons les classes suivantes $C_i = [\frac{i-1}{m}, \frac{i}{m}[$; on remarque que la famille $(C_i)_{1..m}$ est bien un recouvrement disjoint de l'intervalle $[0, 1[$ et qu'elle est associée à

la probabilité $p = \frac{1}{m}$. On génère n valeurs $(x_i)_{i=1..n}$ dans $[0, 1[$ et on compte le nombre de fois où la valeur x_j appartient à la classe C_i . Appliquons donc un test du χ^2 avec les classes (C_i) construites. Notre hypothèse H est que le générateur de la classe Random de Java n'est pas biaisé, nous souhaitons avoir un risque de se tromper inférieur à une valeur α appelée tolérance.

Soient O_i le nombre de fois où x_j appartient à la classe C_i et E_i la valeur espérée de O_i : dans ce cas $E_i = \frac{n}{m}$. Si nous définissons $D^2 = \sum_{i=0}^{m-1} \frac{(O_i - E_i)^2}{E_i} = \sum_{i=0}^{m-1} \frac{(O_i - \frac{n}{m})^2}{\frac{n}{m}}$, alors D^2 suit une loi du χ^2 à $m - 1$ degrés de liberté : χ_{m-1}^2 . Soit $F(D^2, m - 1)$ la fonction de répartition de la loi χ_{m-1}^2 . Par la suite, nous tirons les n valeurs x_j de la loi uniforme, on calcule D^2 et nous évaluons $F(D^2, m - 1)$.

		Tolérance α		
		5,00 %	2,50 %	1,00 %
Nombres de valeurs n	10 000	4,80 %	2,76 %	0,92 %
	100 000	5,00 %	2,46 %	1,17 %
	1 000 000	5,25 %	2,56 %	0,85 %
	10 000 000	4,91 %	2,36 %	0,96 %

TABLEAU 3.1 Pourcentage moyen de rejet de l'hypothèse H

Le tableau (3.1) montre les résultats obtenus pour différentes valeurs de n et de la tolérance α . Pour chaque tirage, on a calculé le pourcentage de fois où $F(D^2, m - 1) > 1 - \alpha$ (c'est-à-dire que l'hypothèse H est rejetée) et nous avons pris la moyenne de ce pourcentage sur 10 000 observations. Normalement cette moyenne se doit d'être égale à la tolérance α .

Le nombre de classes m choisi pour réaliser le test a été fixé à 101 et la table de valeurs utilisée est dans le livre de Saporta (2006). Les erreurs obtenues sur le rejet moyen de l'hypothèse H sont acceptables. En vue des résultats et des tailles données, on peut considérer que le générateur de nombres pseudo-aléatoires fourni par le classe Random de Java n'est pas biaisé.

3.3 Modélisation des perturbations

3.3.1 Modélisation de la tendance de la journée

Prenons l'exemple d'un supermarché : chaque journée il y a finalement plus ou moins de clients que ce qui avait été prévu qui vont venir faire leurs courses. La somme de la demande réelle en employés sur toutes les périodes de la journée va forcément être légèrement différente de celle de la demande prévisionnelle. On modélise ce coefficient de variation de la somme par une variable d'écart moyen ϕ_{96} . Cette variable va suivre une loi log-normale de moyenne égale à 1. La demande prévisionnelle est considérée comme la moyenne de la demande dans ce modèle, le coefficient de variation a donc forcément une moyenne de 1 : $\mathbb{E}[D_p] = Dmoy_p, \forall p = 1 \dots 96$. Par conséquent d'après le modèle (3.1) : $\mathbb{E}[\theta_p] = 1$. Comme les différents types de perturbations sont indépendants et se comportent de manière symétrique (ils vont suivre les mêmes lois avec des paramètres différents), leur moyenne doit être aussi égale à 1.

On a $\mathbb{E}[\theta_p] = \mathbb{E}[\theta_p^{bruit}] \mathbb{E}[\theta_p^{1heure}] \mathbb{E}[\theta_p^{4heures}] \mathbb{E}[\theta_p^{tendance}] = 1$. Si, par exemple, $\mathbb{E}[\theta_p^{bruit}] > 1$ et $\mathbb{E}[\theta_p^{1heure}] < 1$. Ces perturbations ne sont plus symétriques et ne se comportent plus du tout de la même façon. Si on fixe à 1 la moyenne de toutes les perturbations, on n'introduit aucune corrélation et ni asymétrie entre les perturbations.

On obtient donc pour la tendance :

$$\begin{aligned} \theta_p^{tendance} &= \phi_{96}, \forall p = 1 \dots 96 \\ \phi_{96} &\longrightarrow \ln \mathcal{N}(\mu_{96}, \sigma_{96}) \\ \mathbb{E}[\phi_{96}] &= e^{\mu_{96} + \frac{\sigma_{96}^2}{2}} = 1 \end{aligned}$$

ϕ_{96} représente ainsi la tendance d'une réalisation par rapport à la demande prévisionnelle comme le montre la figure (3.2).

3.3.2 Modélisation du bruit périodique

La demande prévisionnelle ne peut capter l'incertitude de chaque période, la réalisation va se différencier de la prévision de manière aléatoire. À chaque période il existe un bruit inidentifiable, il est impossible de prévoir de manière exacte le nombre de clients

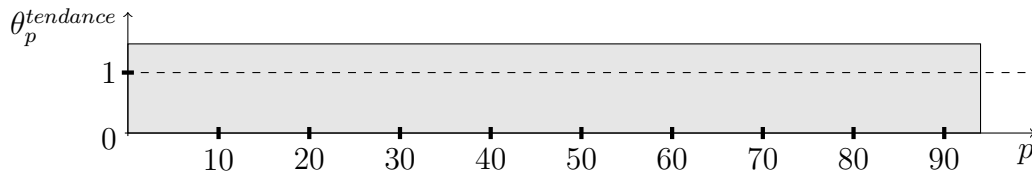


FIGURE 3.2 Exemple de coefficient de variation $\theta_p^{tendance}$

qui vont arriver dans le supermarché par exemple. Ce bruit inhérent à la nature de la réalisation va être modélisé par des coefficients de variation $\phi_{1,p}$ indépendantes sur chaque période. De la même manière que pour la tendance, le bruit périodique va suivre une loi log-normale de moyenne 1, on obtient donc :

$$\begin{aligned}\theta_p^{bruit} &= \phi_{1,p}, \forall p = 1 \dots 96 \\ \phi_{1,p} &\longrightarrow \ln \mathcal{N}(\mu_1, \sigma_1) \\ \mathbb{E}[\phi_{1,p}] &= e^{\mu_1 + \frac{\sigma_1^2}{2}} = 1\end{aligned}$$

Ces variables permettent donc de bruiteur la réalisation par rapport à la prévision comme le montre la figure (3.3). Il faut noter qu'un choix important a été fait. La même distribution a été choisie pour toutes les périodes. Avec cette hypothèse, on perd de la précision dans le modèle ; le bruit périodique est peut-être plus important à midi qu'à neuf heure. Il semble donc logique de choisir des distributions différentes, il faudrait ainsi 96 distributions différentes ce qui est trop. Ce choix permet de réduire fortement le nombre de paramètres, d'avoir un modèle plus générique et plus résistant au bruit. Un modèle avec une distribution pour la journée, une autre pour la soirée et une dernière pour la nuit peut être envisagé si ce choix apporte une amélioration significative du modèle.

3.3.3 Modélisation des corrélations

Il existe certains évènements comme de la pluie dans l'exemple du supermarché, qui vont modifier la demande sur plusieurs périodes. En effet, la pluie va pousser les

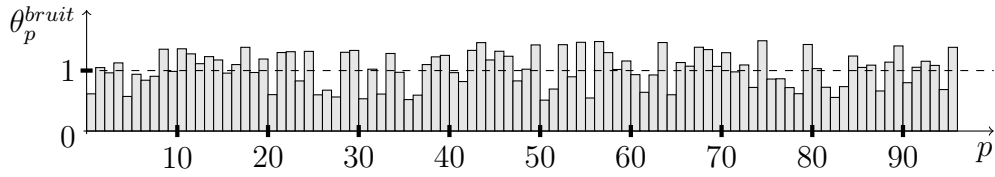


FIGURE 3.3 Exemple de coefficient de variation θ_p^{bruit}

clients à rester chez eux, plutôt que de sortir faire leurs courses. On considère que ces perturbations peuvent commencer à n'importe quelle période comme on le voit sur la figure (3.4). Par contre, afin de simplifier le modèle, on va choisir des longueurs prédéfinies pour la durée des perturbations, soient quatre périodes et seize périodes dans ce modèle. Ces durées peuvent être modifiées si elles semblent ne pas décrire correctement la réalité du marché. Les longueurs des perturbations (une heure et quatre heures) ont été décidées afin d'obtenir une représentation de toutes les longueurs imaginables sur une journée. Encore une fois ce sont des hypothèses, elles doivent donc être confrontées à des données pour être approuvées.

Soient $\delta_{4,p}$ et $\delta_{16,p}$ les variables aléatoires indiquant s'il y a une perturbation ou non commençant à la période p . Ces variables suivent une loi de Bernoulli $\mathcal{B}(\beta_4)$ et $\mathcal{B}(\beta_{16})$. Soient $\phi_{4,p}$ et $\phi_{16,p}$ les coefficients de variation des perturbations de durée quatre et seize commençant à la période p . S'il y a une variation à la période p , les coefficients de variation suivent une loi log-normale, sinon elles sont égales à 1. On a donc :

$$\theta_p^{1heure} = \prod_{k=\max\{1,p-3\}}^p \phi_{4,k}, \forall p = 1 \dots 96$$

$$\theta_p^{4heures} = \prod_{k=\max\{1,p-15\}}^p \phi_{16,k}, \forall p = 1 \dots 96$$

$$\delta_{j,p} \longrightarrow \mathcal{B}(\beta_j), j = 4, 16$$

$$\text{Si } \delta_{j,p} = 0, \text{ alors } \phi_{j,p} = 1, j = 4, 16$$

$$\text{Si } \delta_{j,p} = 1, \text{ alors } \phi_{j,p} \longrightarrow \ln \mathcal{N}(\mu_j, \sigma_j), j = 4, 16$$

$$\mathbb{E}[\phi_{j,p}] = e^{\mu_j + \frac{\sigma_j^2}{2}} = 1, j = 4, 16$$

Comme pour la modélisation du bruit périodique, on a choisi une même distribution

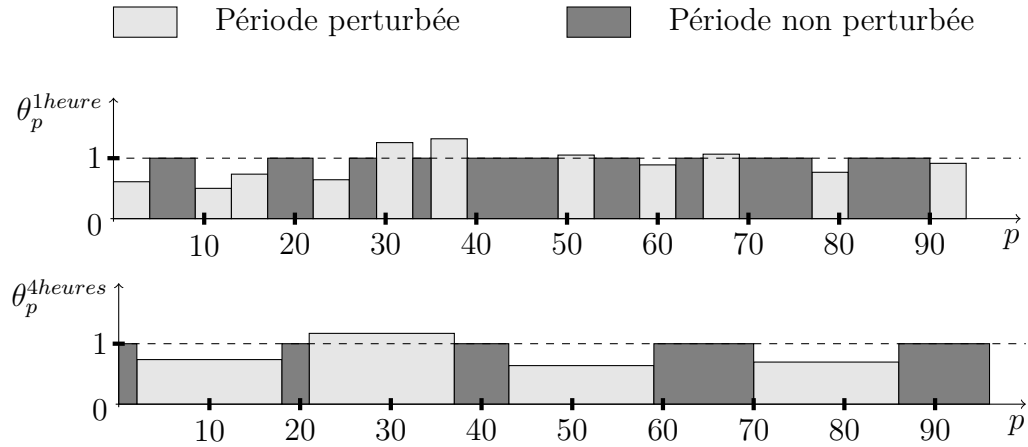


FIGURE 3.4 Exemple des coefficients de variation θ_p^{1heure} et $\theta_p^{4heures}$

pour chaque type de perturbations. Ce choix est encore une fois soumis à l'étude de la réalité afin d'être validé. Nous allons maintenant présenter le modèle général qui regroupe tous les modèles précédents.

3.3.4 Modèle général

Afin de réaliser un modèle global de l'incertitude de la demande prévisionnelle, nous allons combiner les trois modèles précédents en multipliant les variables agissant à une même période p . On a donc :

$$\forall p = 1 \dots 96, \theta_p = \theta_p^{bruit} \theta_p^{1heure} \theta_p^{4heures} \theta_p^{tendance} \quad (3.2)$$

$$= \phi_{96} \cdot \phi_{1,p} \cdot \prod_{k=\max(1,p-3)}^p \phi_{4,k} \cdot \prod_{k=\max(1,p-15)}^p \phi_{16,k} \quad (3.3)$$

Les variables ϕ caractérisent ainsi en entier le modèle; la figure (3.5) le montre bien. Elles sont définies en utilisant une ou deux lois différentes. La première loi décrit l'incertitude sur une période : c'est la loi log-normale. La deuxième loi décrit l'apparition ou non d'une perturbation : c'est la loi de Bernoulli. Le modèle permet donc de décrire de manière probabiliste la réalité. La génération des coefficients de variation selon ce modèle permet de recréer une demande corrélée entre périodes avec un nombre restreint de paramètres à dimensionner : les deux β_j et les quatre σ_j . Les

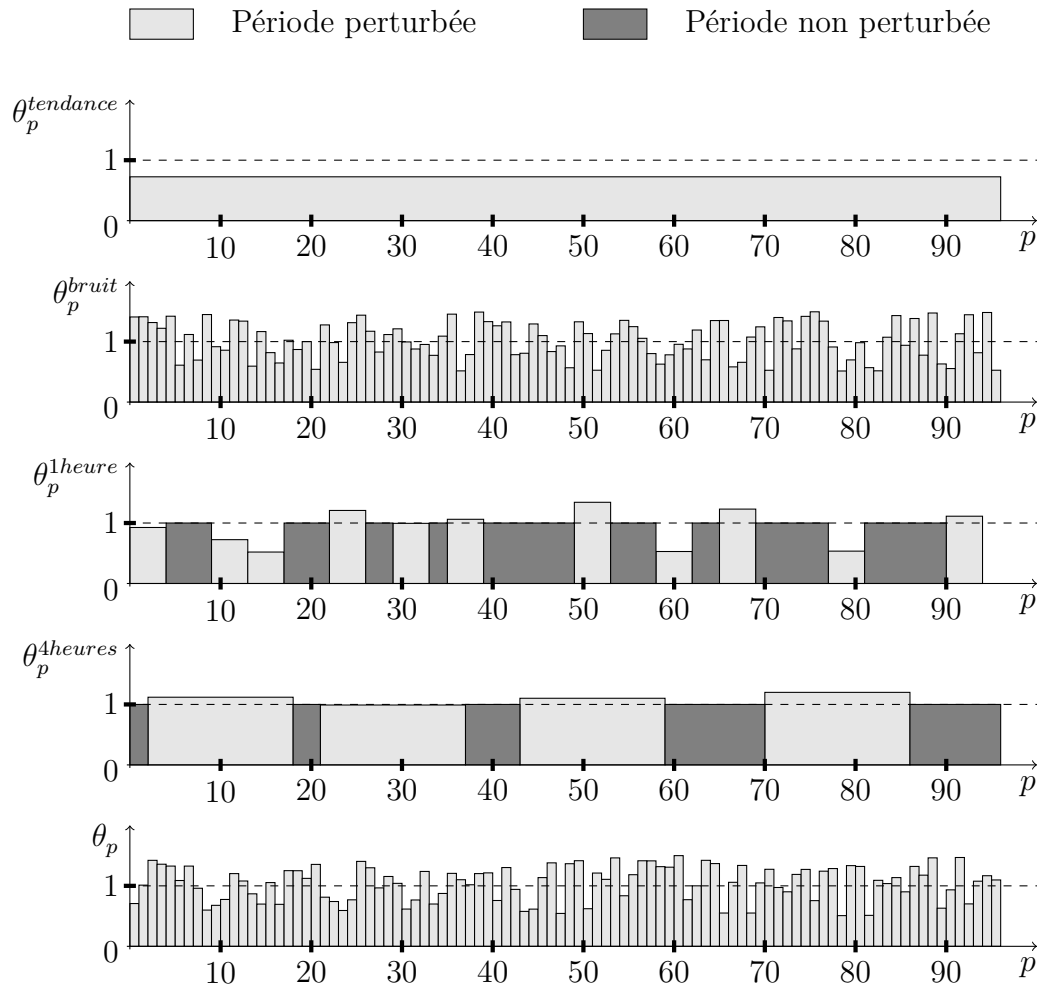


FIGURE 3.5 Exemple du coefficient de variation $\theta_p = \theta_p^{bruit} \theta_p^{1heure} \theta_p^{4heures} \theta_p^{tendance}$

paramètres μ_j n'ont pas besoin d'être considérés comme inconnus puisque $\mu_j = \frac{-1}{2} \sigma_j^2$. Nous allons présenter des techniques d'ajustement de ces paramètres dans le prochain chapitre.

Chapitre 4

ESTIMATION DES PARAMÈTRES

L'objectif de cette section est de présenter une méthode de calibration du modèle. Comme nous n'avons pas de données réelles en notre possession pour calibrer celui-ci, nous allons commencer par essayer de retrouver les paramètres à partir de données générées par le modèle. Il est sûr que l'historique ainsi créé a la forme voulue, mais si les estimateurs n'arrivent même pas à obtenir une bonne approximation des paramètres de ces données, c'est que forcément ils ne sont pas bons et ne seront pas plus efficaces a priori dans la réalité. De plus, certains paramètres du modèle sont très difficiles à obtenir, il nous faut donc développer des techniques spécifiques à ce problème. Les estimateurs classiques de la variance nous permettent de récupérer une partie de l'information, mais par la suite il faut construire des estimateurs plus précis afin de retrouver tous les paramètres du modèle.

4.1 Estimateurs classiques

En général, les statisticiens essayent de construire de bons estimateurs à partir du maximum de vraisemblance. Présentons rapidement cet outil : soient X une variable aléatoire réelle, α un paramètre à estimer et $f(x, \alpha)$ la densité de X pour un paramètre α donné. (x_1, x_2, \dots, x_n) est un échantillon de la variable aléatoire X . On note L la vraisemblance de α au vu des observations (x_1, x_2, \dots, x_n) :

$$L(x_1, \dots, x_n, \alpha) = \prod_{i=1}^n f(x_i, \alpha)$$

L'objectif est de maximiser L en fonction de la variable α . On cherche donc $\tilde{\alpha}$ tel que $\frac{\partial L(x_1, x_2, \dots, x_n, \tilde{\alpha})}{\partial \alpha} = 0$. Comme le fonction L est un produit, il est habituel de prendre le logarithme pour faciliter les calculs. Comme le logarithme est une fonction monotone

croissante, on obtient le même $\tilde{\alpha}$ en maximisant L ou $\ln(L)$. On a donc :

$$\begin{aligned}\ln(L(x_1, x_2, \dots, x_n, \alpha)) &= \sum_{i=1}^n \ln(f(x_i, \alpha)) \\ \frac{\partial \ln(L(x_1, x_2, \dots, x_n, \alpha))}{\partial \alpha} &= \sum_{i=1}^n \frac{\partial \ln(f(x_i, \alpha))}{\partial \alpha}\end{aligned}$$

D'où :

$$\sum_{i=1}^n \frac{f'_\alpha(x_i, \tilde{\alpha})}{f(x_i, \tilde{\alpha})} = 0 \quad (4.1)$$

Le problème est que dans la situation actuelle il est impossible d'appliquer cette idée à notre modèle. Les variables aléatoires $(\phi_{i,j})_{i=4,16}$ sont le mélange d'une variable aléatoire continue et d'une variable aléatoire discrète. Il faut donc introduire un nouveau concept : le maximum de vraisemblance conditionnel.

Soient deux variables aléatoires réelles X, Y et α un paramètre à déterminer. Soit $((y_1, x_1), \dots, (y_n, x_n))$ un échantillon apparié. On définit maintenant la vraisemblance L de cette façon :

$$L(y_1, \dots, y_n, x_1, \dots, x_n, \alpha) = \prod_{i=1}^n f(y_i | x_i, \alpha)$$

où $f(y_i | x_i, \alpha)$ est la densité de Y étant donné X pour un paramètre α donné. De la même façon, on peut définir le maximum de vraisemblance, mais arrêtons-nous ici pour montrer que, dans cette étude, le maximum de vraisemblance risque d'être très compliqué à mettre en place. En effet, pour l'utiliser il faut avoir un jeu de données des variables aléatoires X et Y . Dans le modèle, Y est la variable $\phi_{i,j}$ et X la loi de Bernoulli $\delta_{i,p}$. Tout d'abord les données que nous avons sont les coefficients de variation $\theta_p = \theta_p^{bruit} \theta_p^{1heure} \theta_p^{4heures} \theta_p^{tendance}$. Il semble imaginable d'arriver à récupérer les informations sur les variables $\theta_p^{tendance} = \phi_{96}$, $\theta_p^{bruit} = \phi_{1,p}$, $\theta_p^{1heure} = \prod_{k=\max(1,p-3)}^p \phi_{4,k}$

et $\theta_p^{4heures} = \prod_{k=\max(1,p-15)}^p \phi_{16,k}$, ce qui déjà est difficile. Il est sûrement faisable d'y arriver, car chacune de ces variables se comporte de manière différente. Elles n'ont pas

la même distribution et leur perturbation ne dure pas le même nombre de périodes. Cependant comme nous n'avons à notre disposition que les produits de ces quatre variables, il est impossible d'obtenir de manière exacte la réalisation de chacune. Nous ne pourrons avoir qu'une estimation de celles-ci. Par la suite, il faut aussi séparer les variables $\phi_{4,p}$ et $\phi_{16,p}$ afin de déterminer les variables $\delta_{j,p}$. Cette dernière étape est encore plus délicate, car il faut extraire la valeur d'une variable au milieu de quatre ou seize variables identiques. Il n'y a aucune possibilité de différencier les variables intrinsèques aux variations $\phi_{4,p}$ ou $\phi_{16,p}$ par construction du modèle. Arriver alors à obtenir toutes les variables $\delta_{j,p}$ est un rêve qui semble actuellement inatteignable. Nous ne voyons pas comment arriver à isoler ces variables à partir des données en notre possession. C'est pourquoi nous tentons une autre approche : la méthode des moments.

Soient une variable aléatoire réelle X et un échantillon indépendant (x_1, x_2, \dots, x_n) . Nous cherchons toujours à obtenir un paramètre α qui caractérise la densité de la variable X , mais nous allons développer une autre méthode. Pour cela, considérons les s premiers moments de X , $M = [m_1(\alpha), m_2(\alpha), \dots, m_s(\alpha)]$ où $m_i(\alpha) = \mathbb{E}[X^i]$. Un estimateur non biaisé de $m_i(\alpha)$ est $\tilde{m}_i = \frac{1}{n} \sum_{j=1}^n x_j^i$. Il suffit alors de résoudre le système d'équation $m_i(\tilde{\alpha}) = \tilde{m}_i(\tilde{\alpha})$ pour avoir un bon estimateur $\tilde{\alpha}$ de α . Dans le cas de ce modèle, nous allons nous inspirer de cette méthode pour construire les estimateurs des probabilités β_j et des écarts-type σ_j . Tout d'abord, présentons les équations utilisées puis la construction des estimateurs.

4.2 Estimation globale de la variance

Nous allons travailler avec les lois normales associées aux variables afin de simplifier les calculs. On a donc : $\Theta_p \leftarrow \ln(\theta_p)$ et $\rho_{j,p} \leftarrow \ln(\phi_{j,p})$. Il est plus facile de travailler avec une somme de lois normales qu'avec le produit de lois log-normales. Les variables $\ln(\phi_p)$ suivent maintenant une loi normale $\mathcal{N}(\mu_j, \sigma_j)$.

Les différentes durées de perturbations permettent d'isoler la variance de chacune. On va évaluer les variances et utiliser les corrélations afin de simplifier ces formules. L'étude de deux variations espacées de plusieurs périodes Θ_p et Θ_k permet de modifier les formes des perturbations. En faisant varier l'écart entre les deux variations, nous captions une partie ou non des perturbations ; en jouant sur cette différence on arrive à

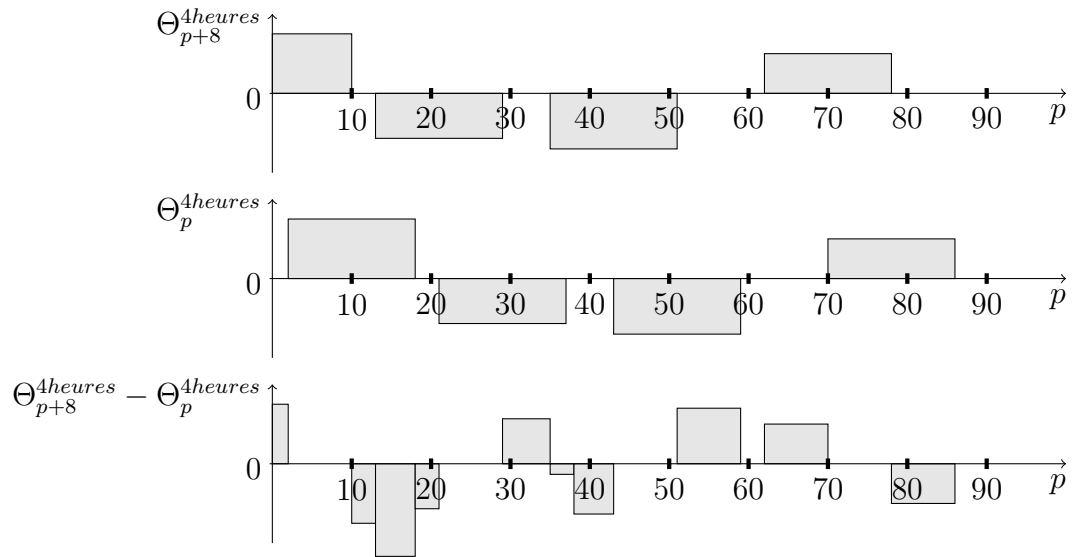


FIGURE 4.1 Exemple de variation $\Theta_{p+8}^{4heures} - \Theta_p^{4heures}$

isoler chaque type de perturbations. Étudions par exemple les perturbations de quatre heures.

En regardant la variation $\Theta_{p+8}^{4heures} - \Theta_p^{4heures}$, on annule une partie des perturbations. Par exemple, sur la figure (4.1) la perturbation comprise entre les périodes 2 et 10 disparaît. De cette façon, on a diminué l'influence des perturbations de quatre heures. Celles d'une heure, comme nous pouvons le voir sur la figure (4.2), ne s'annulent pas entre elles, car elles sont trop petites.

On ne considère que les données au milieu d'une journée afin d'éviter une partie des effets de bord. En effet il en existe deux sortes. Tout d'abord, toutes les périodes n'ont pas la même probabilité d'être affectées par les perturbations d'une heure et de quatre heures. En effet, si nous considérons la première période de la journée $p = 1$, elle ne peut être touchée que par les perturbations générées à cette période, alors qu'une période en milieu de journée ($p = 45$) peut être touchée par des perturbations générées à partir de la période $p = 30$ jusqu'à la période considérée $p = 45$. C'est la première raison pour laquelle on ne considère que les périodes au milieu de la journée. Deuxièmement, lorsque que nous considérons des périodes espacées (figure (4.1)), une partie des perturbations peuvent disparaître : la première perturbation

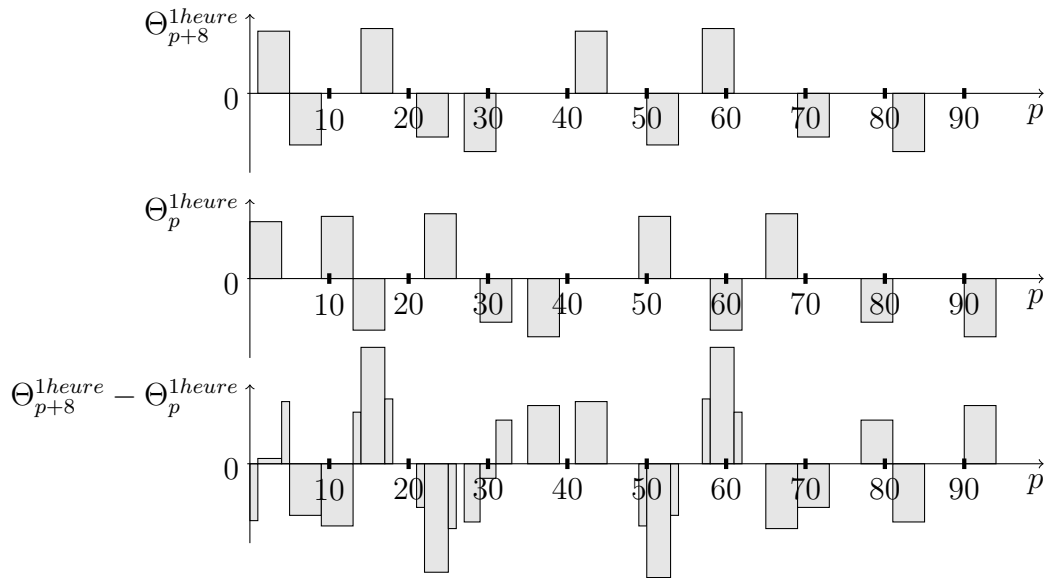


FIGURE 4.2 Exemple de variation $\Theta_{p+8}^{1heure} - \Theta_p^{1heure}$

n'est plus en entier pour la variation $\Theta_{p+8}^{4heures}$ et il n'y en a plus aucune sur les huit dernières périodes. C'est pourquoi nous ne prendrons en considération que les périodes p comprises entre 16 et 64 inclus. Le début à la période $p = 16$ est dû à la perturbation de quatre heures : comme il a été expliqué en premier lieu, avant cette période, la probabilité d'être affectée par ce type de perturbation est plus petite. La fin à la période $p = 64$ ($= 96 - 32$) vient des variations Θ_{p+32} . À partir de ce moment, il n'y a plus aucune perturbation générée

Nous allons d'abord calculer la variance théorique des variations Θ_p puis illustrer les techniques utilisées dans le cas précis où on cherche à s'affranchir des effets de la tendance, enfin nous montrerons l'ensemble des formules.

4.2.1 Variance des variations Θ_p

Retrouvons d'abord l'expression de Θ_p à partir des expressions (3.2) et (3.3) de θ_p .

$$\begin{aligned}
\forall p = 16 \dots 64, \Theta_p &= \ln(\theta_p) \\
&= \ln(\theta_p^{bruit} \theta_p^{1heure} \theta_p^{4heures} \theta_p^{tendance}) \\
&= \ln(\theta_p^{bruit}) + \ln(\theta_p^{1heure}) + \ln(\theta_p^{4heures}) + \ln(\theta_p^{tendance}) \\
&= \ln(\phi_{1,p}) + \ln\left(\prod_{k=p-3}^p \phi_{4,k}\right) + \ln\left(\prod_{k=p-15}^p \phi_{16,k}\right) + \ln(\phi_{96}) \\
&= \ln(\phi_{1,p}) + \sum_{k=p-3}^p \ln(\phi_{4,k}) + \sum_{k=p-15}^p \ln(\phi_{16,k}) + \ln(\phi_{96}) \\
&= \rho_{1,p} + \sum_{k=p-3}^p \rho_{4,k} + \sum_{k=p-15}^p \rho_{16,k} + \rho_{96} \\
&= \Theta_p^{bruit} + \Theta_p^{1heure} + \Theta_p^{4heures} + \Theta_p^{tendance}
\end{aligned}$$

où :

$$\begin{aligned}
\Theta_p^{bruit} &= \rho_{1,p} \\
\Theta_p^{1heure} &= \sum_{k=p-3}^p \rho_{4,k} \\
\Theta_p^{4heures} &= \sum_{k=p-15}^p \rho_{16,k} \\
\Theta_p^{tendance} &= \rho_{96}
\end{aligned}$$

étudions maintenant l'expression de la variance de Θ_p . Comme les variables $\Theta_p^{bruit}, \Theta_p^{1heure}, \Theta_p^{4heures}$ et $\Theta_p^{tendance}$ sont indépendantes, on a :

$$\begin{aligned}
\text{Var}[\Theta_p] &= \text{Var}[\Theta_p^{bruit}] + \text{Var}[\Theta_p^{1heure}] + \text{Var}[\Theta_p^{4heures}] + \text{Var}[\Theta_p^{tendance}] \\
&= \text{Var}[\rho_{1,p}] + \text{Var}\left[\sum_{k=p-3}^p \rho_{4,k}\right] + \text{Var}\left[\sum_{k=p-15}^p \rho_{16,k}\right] + \text{Var}[\rho_{96}]
\end{aligned}$$

Comme les variables $\rho_{j,k}$ sont aussi indépendantes, on obtient :

$$\begin{aligned}\text{Var}[\Theta_p] &= \text{Var}[\rho_{1,p}] + \sum_{k=p-3}^p \text{Var}[\rho_{4,k}] + \sum_{k=p-15}^p \text{Var}[\rho_{16,k}] + \text{Var}[\rho_{96}] \\ &= \text{Var}[\rho_{1,p}] + 4\text{Var}[\rho_{4,p}] + 16\text{Var}[\rho_{16,p}] + \text{Var}[\rho_{96}]\end{aligned}$$

D'après les formules de la loi normale, on a les relations suivantes :

$$\begin{aligned}\text{Var}[\rho_{1,p}] &= \sigma_1^2 \\ \text{Var}[\rho_{96}] &= \sigma_{96}^2\end{aligned}$$

Pour les variations $\rho_{4,k}$ et $\rho_{16,k}$, il est plus difficile d'obtenir les variances, car ce sont des variables aléatoires conditionnelles. Calculons la variance de $\rho_{j,p}$ pour $j=4$ et 16 , on a :

$$\text{Var}[\rho_{j,p}] = \mathbb{E}[\rho_{j,p}^2] - \mathbb{E}[\rho_{j,p}]^2$$

or

$$\begin{aligned}\mathbb{E}[\rho_{j,p}] &= \mathbb{E}[\rho_{j,p} \mid \delta_{j,p} = 1] * \mathbb{P}[\delta_{j,p} = 1] + \mathbb{E}[\rho_{j,p} \mid \delta_{j,p} = 0] * \mathbb{P}[\delta_{j,p} = 0] \\ &= \beta_j \mu_j + 0 \\ &= \frac{-\beta_j \sigma_j^2}{2}\end{aligned}$$

En effet, lorsque $\delta_{j,p} = 0$, il n'y a aucune perturbation et, par conséquent, $\rho_{j,p} = 0$.

De plus on a aussi :

$$\begin{aligned}\mathbb{E}[\rho_{j,p}^2] &= \mathbb{E}[\rho_{j,p}^2 \mid \delta_{j,p} = 1] * \mathbb{P}[\delta_{j,p} = 1] + \mathbb{E}[\rho_{j,p}^2 \mid \delta_{j,p} = 0] * \mathbb{P}[\delta_{j,p} = 0] \\ &= (\sigma_j^2 + \mu_j^2) \beta_j + 0 \\ &= \beta_j \left(\sigma_j^2 + \frac{\sigma_j^4}{4} \right)\end{aligned}$$

On en déduit que :

$$\begin{aligned}\text{Var}[\rho_{j,p}] &= \beta_j \left(\sigma_j^2 + \frac{\sigma_j^4}{4} \right) - \left(\frac{-\beta_j \sigma_j^2}{2} \right)^2 \\ &= \beta_j \left(\sigma_j^2 + \frac{\sigma_j^4}{4} \right) - \frac{\beta_j^2 \sigma_j^4}{4} \\ &= \beta_j \sigma_j^2 + (1 - \beta_j) \beta_j \frac{\sigma_j^4}{4}\end{aligned}$$

Résumons toutes les formules obtenues :

$$\text{Var}[\rho_{1,p}] = S_1^2 = \sigma_1^2 \quad (4.2)$$

$$\text{Var}[\rho_{4,p}] = S_4^2 = \beta_4 \sigma_4^2 + (1 - \beta_4) \beta_4 \frac{\sigma_4^4}{4} \quad (4.3)$$

$$\text{Var}[\rho_{16,p}] = S_{16}^2 = \beta_{16} \sigma_{16}^2 + (1 - \beta_{16}) \beta_{16} \frac{\sigma_{16}^4}{4} \quad (4.4)$$

$$\text{Var}[\rho_{96}] = S_{96}^2 = \sigma_{96}^2 \quad (4.5)$$

Les variables S_j^2 représentent la variance globale des perturbations de longueur j à une période donnée. La variance globale vaut ainsi :

$$\text{Var}[\Theta_p] = S_1^2 + 4S_4^2 + 16S_{16}^2 + S_{96}^2 \quad (4.6)$$

4.2.2 L'annulation des effets de la tendance

Nous allons montrer ici comment obtenir une formule sans le terme ρ_{96} qui correspond à la tendance. En considérant $\Theta_{p+32} - \Theta_p$, on annule l'effet de la tendance, comme la figure (4.3) le montre, mais on ne modifie pas les autres perturbations. L'écart entre les deux variations est trop important pour que les autres types de perturbations puissent interagir. En effet, comme leurs durées (une, quatre et seize périodes) sont inférieures à seize, elles ne peuvent pas se mélanger pour un écart de 32. La figure (4.2) illustre parfaitement cette idée pour un écart de huit.

Exprimons maintenant la variation $\Theta_{p+32} - \Theta_p$:

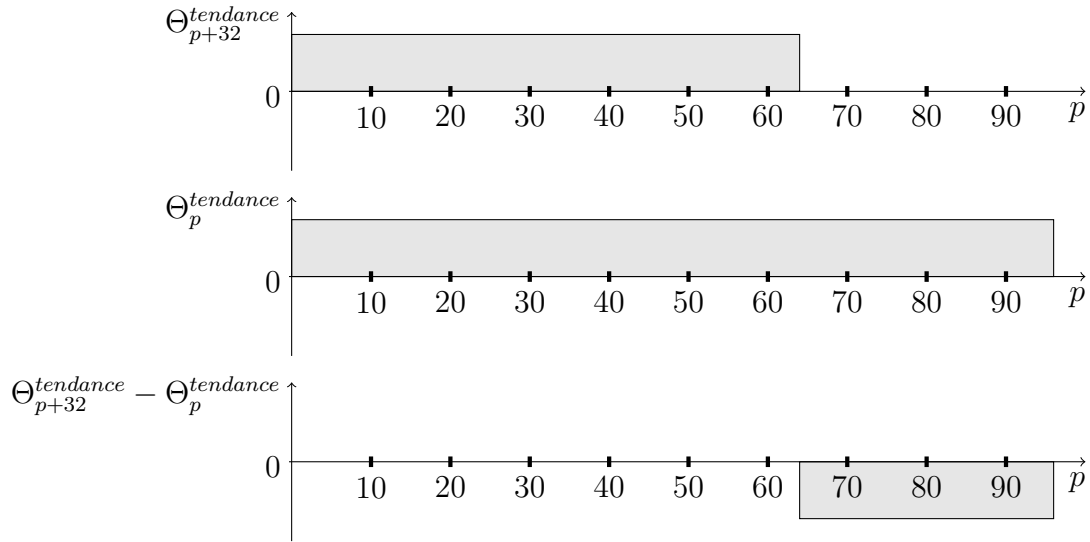


FIGURE 4.3 Variation $\Theta_{p+32}^{tendance} - \Theta_p^{tendance}$

$$\begin{aligned} \Theta_{p+32} - \Theta_p &= \Theta_{p+32}^{bruit} + \Theta_{p+32}^{1heure} + \Theta_{p+32}^{4heures} + \Theta_{p+32}^{tendance} \\ &\quad - [\Theta_p^{bruit} + \Theta_p^{1heure} + \Theta_p^{4heures} + \Theta_p^{tendance}] \\ \Theta_{p+32} - \Theta_p &= [\Theta_{p+32}^{bruit} - \Theta_p^{bruit}] + [\Theta_{p+32}^{1heure} - \Theta_p^{1heure}] \\ &\quad + [\Theta_{p+32}^{4heures} - \Theta_p^{4heures}] + [\Theta_{p+32}^{tendance} - \Theta_p^{tendance}] \end{aligned}$$

Comme $\forall p = 16 \dots 64$, $\Theta_{p+32}^{tendance} = \Theta_p^{tendance} = \rho_{96}$, on a :

$$\begin{aligned} \Theta_{p+32} - \Theta_p &= [\Theta_{p+32}^{bruit} - \Theta_p^{bruit}] + [\Theta_{p+32}^{1heure} - \Theta_p^{1heure}] + [\Theta_{p+32}^{4heures} - \Theta_p^{4heures}] \\ \text{Var}[\Theta_{p+32} - \Theta_p] &= \text{Var}[\Theta_{p+32}^{bruit} - \Theta_p^{bruit}] + \text{Var}[\Theta_{p+32}^{1heure} - \Theta_p^{1heure}] \\ &\quad + \text{Var}[\Theta_{p+32}^{4heures} - \Theta_p^{4heures}] \end{aligned}$$

Comme il a été dit juste avant, les seules perturbations qui interagissent sont celles de la tendance. Ainsi Θ_{p+32}^{bruit} et Θ_p^{bruit} , Θ_{p+32}^{1heure} et Θ_p^{1heure} , $\Theta_{p+32}^{4heures}$ et $\Theta_p^{4heures}$ sont des

variables aléatoires indépendantes. Nous obtenons donc les équations suivantes :

$$\begin{aligned}\text{Var}[\Theta_{p+32} - \Theta_p] &= \text{Var}[\Theta_{p+32}^{bruit}] + \text{Var}[\Theta_p^{bruit}] + \text{Var}[\Theta_{p+32}^{1heure}] + \text{Var}[\Theta_p^{1heure}] \\ &\quad + \text{Var}[\Theta_{p+32}^{4heures}] + \text{Var}[\Theta_p^{4heures}] \\ &= 2\text{Var}[\Theta_p^{bruit}] + 2\text{Var}[\Theta_p^{1heure}] + 2\text{Var}[\Theta_p^{4heures}]\end{aligned}$$

Comme pour la formule (4.6), l'équation précédente se réécrit pour donner la formule suivante :

$$\text{Var}[\Theta_{p+32} - \Theta_p] = 2S_1^2 + 8S_4^2 + 32S_{16}^2 \quad (4.7)$$

Le premier terme est la variance des perturbations quotidiennes, le deuxième et le troisième représentent les variances des corrélations. Ils décrivent presque entièrement la corrélation dans ce modèle. Nous allons maintenant montrer d'autres formules du même type mais avec des écarts de périodes plus petits.

4.2.3 Autres formules des variances

Si au lieu de considérer un écart de 32 périodes, nous considérons un écart de 8 périodes, il y a plus de perturbations qui interagissent. L'idée générale reste la même, mais les coefficients de la formule (4.7) changent un peu. Avec un écart de 8 périodes, l'effet de la tendance est toujours annulé, mais il y a les perturbations de durée quatre heures qui interagissent aussi. En effet comme le montre la figure (4.1), il y a une partie de la perturbation qui disparaît et l'autre qui reste toujours présente. C'est pourquoi les coefficients de la formule (4.7) sont modifiés. Nous obtenons donc pour la variation $\Theta_{p+8}^{4heures} - \Theta_p^{4heures}$ pour $p = 16 \dots 64$:

$$\begin{aligned}\Theta_{p+8}^{4heures} - \Theta_p^{4heures} &= \sum_{k=(p+8)-15}^{p+8} \rho_{16,k} - \sum_{k=p-15}^p \rho_{16,k} \\ &= \left[\sum_{k=p-7}^p \rho_{16,k} + \sum_{k=p+1}^{p+8} \rho_{16,k} \right] - \left[\sum_{k=p-7}^p \rho_{16,k} + \sum_{k=p-15}^{p-8} \rho_{16,k} \right] \\ &= \sum_{k=p+1}^{p+8} \rho_{16,k} - \sum_{k=p-15}^{p-8} \rho_{16,k}\end{aligned}$$

Nous évaluons la variance de cette variation. Comme toutes les variables aléatoires restantes sont indépendantes, on obtient :

$$\begin{aligned}
\text{Var}[\Theta_{p+8}^{4heures} - \Theta_p^{4heures}] &= \sum_{k=p+1}^{p+8} \text{Var}[\rho_{16,k}] + \sum_{k=p-15}^{p-8} \text{Var}[\rho_{16,k}] \\
&= \sum_{k=p+1}^{p+8} S_{16}^2 + \sum_{k=p-15}^{p-8} S_{16}^2 \\
&= 16S_{16}^2
\end{aligned}$$

Maintenant que le terme de la perturbation de quatre heures est calculé, nous pouvons regarder la variance totale, on obtient ainsi :

$$\text{Var}[\Theta_{p+8} - \Theta_p] = 2\text{Var}[\Theta_p^{bruit}] + 2\text{Var}[\Theta_p^{1heure}] + \text{Var}[\Theta_{p+8}^{4heures} - \Theta_p^{4heures}]$$

d'où la formule suivante :

$$\text{Var}[\Theta_{p+8} - \Theta_p] = 2S_1^2 + 8S_4^2 + 16S_{16}^2 \quad (4.8)$$

Nous obtenons de la même façon en considérant des écarts d'une période la formule suivante :

$$\text{Var}[\Theta_{p+1} - \Theta_p] = 2S_1^2 + 2S_4^2 + 2S_{16}^2 \quad (4.9)$$

La figure (4.4) explique comment les perturbations de longueur quatre interagissent ensemble ; une grande partie de la corrélation est enlevée par cet écart d'une période.

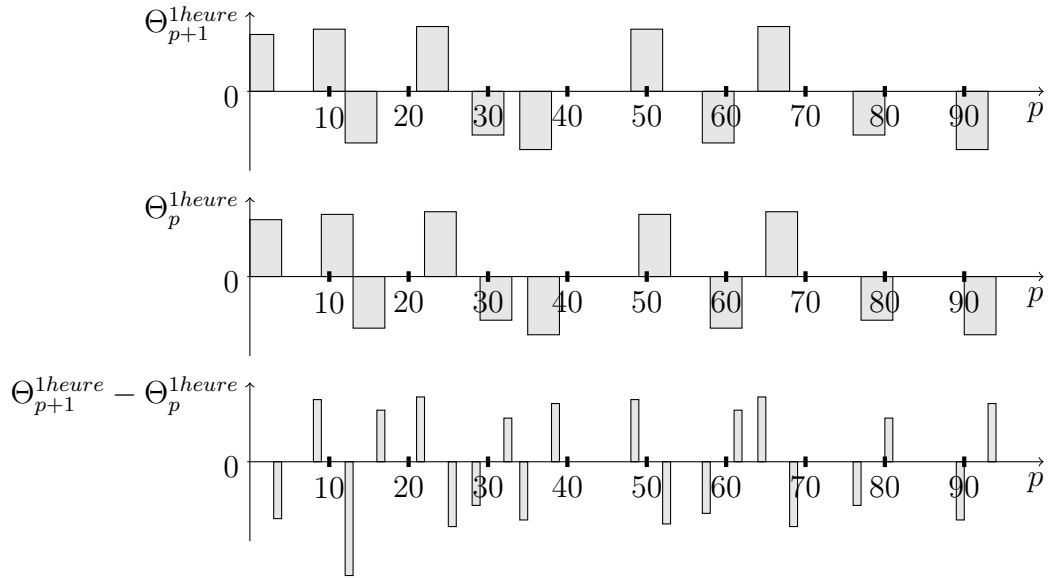


FIGURE 4.4 Variation $\Theta_{p+1}^{1heure} - \Theta_p^{1heure}$

4.2.4 Choix des écarts

Résumons les formules (4.6), (4.7), (4.8) et (4.9) qui ont été calculées dans les derniers paragraphes :

$$\begin{aligned}
 \text{Var}[\Theta_p] &= S_1^2 + 4S_4^2 + 16S_{16}^2 + S_{96}^2 \\
 \text{Var}[\Theta_{p+1} - \Theta_p] &= 2S_1^2 + 2S_4^2 + 2S_{16}^2 \\
 \text{Var}[\Theta_{p+8} - \Theta_p] &= 2S_1^2 + 8S_4^2 + 16S_{16}^2 \\
 \text{Var}[\Theta_{p+32} - \Theta_p] &= 2S_1^2 + 8S_4^2 + 32S_{16}^2
 \end{aligned}
 \tag{4.10}$$

Il apparaît qu'un choix arbitraire de un, huit et trente-deux a été fait pour l'écart entre les périodes. Pourquoi n'avons-nous pas pris en compte d'autres écarts ? Pourquoi avons-nous choisi ceux-là ? Si nous considérons un écart de trente-et-un, nous obtenons la même formule (4.7) que pour un écart de périodes de trente-deux. Le problème vient du fait que nous avons considéré un modèle qui ne correspond pas à la réalité : il a fallu choisir des longueurs pour les perturbations, alors qu'il existe sûrement d'autres longueurs dans la réalité. Par exemple il peut exister des perturbations de durée de huit heures (32 périodes). Ces dernières interagissent lorsque l'écart est trente-et-un mais plus du tout pour un écart de trente-deux. Les formules, par contre restent identiques, alors qu'elles ne contiennent plus les mêmes éléments. Choisir un écart

de trente-deux implique beaucoup plus de choses que ce qu'on peut penser au début. Cet intervalle de périodes contient d'autres perturbations qui seront assimilées à une perturbation de durée quatre heures ou une journée. Ainsi le choix des écarts (une, huit et trente deux périodes) implique que nous imposons un représentant pour une classe de perturbations : la perturbation de longueur une ne représente qu'elle-même, celle de quatre représente les perturbations de longueur deux à sept, celle de seize représente les perturbations de longueur huit à trente-et-une et celle de quatre-vingt-seize représente toutes les perturbations de longueur supérieure à trente-et-une. Il est donc obligatoire de se restreindre à ces quatre équations (4.10) pour estimer tous les paramètres.

Nous allons voir dans les prochaines sections les résultats de l'estimation pour les variances globales $(S_j^2)_{j=1,4,16,96}$, puis l'élaboration d'autres équations afin d'estimer totalement les paramètres. En effet, il y a actuellement quatre équations et six paramètres à estimer.

4.2.5 Méthodes de résolution

Pour le moment, nous n'avons pas de données industrielles afin de confronter notre modèle à la réalité. On ne peut utiliser qu'un échantillon que nous avons construit afin de tester les équations (4.10). On génère un échantillon avec notre modèle et une demande prévisionnelle arbitraire, puis nous allons essayer de retrouver les variances S_j^2 de chaque type de perturbations et nous les confronterons aux valeurs théoriques. De toutes façons, même si nous avons des données industrielles, nous serions obligés de quand même simuler notre modèle. En effet, il faut les valeurs théoriques des paramètres pour juger si une méthode de résolution est bonne ou non. Les données industrielles ne contiennent pas ces valeurs, alors que nous les connaissons dans le cadre d'une simulation du modèle.

Les méthodes ne sont pas faciles à mettre en place, car il est possible d'utiliser beaucoup de techniques de régression différentes pour trouver nos paramètres. Nous n'arrivons pas à obtenir des très bons résultats, car on utilise trop d'observations pour avoir des résultats à peu près acceptables. Au début nous avons essayé de minimiser les écarts entre les valeurs théoriques et observées. Soit $(\Theta_p^i)_{i=1\dots n}$ l'échantillon considéré. Nous

essayons de résoudre le problème suivant :

$$\min_{V, S^2} |V_1 - \text{Var}[\Theta_p^i]| + |V_2 - \text{Var}[\Theta_{p+1}^i - \Theta_p^i]| + |V_3 - \text{Var}[\Theta_{p+8}^i - \Theta_p^i]| \\ + |V_4 - \text{Var}[\Theta_{p+32}^i - \Theta_p^i]|$$

soumis aux contraintes :

$$\begin{aligned} V_1 &= S_1^2 + 4S_4^2 + 16S_{16}^2 + S_{96}^2 \\ V_2 &= 2S_1^2 + 2S_4^2 + 2S_{16}^2 \\ V_3 &= 2S_1^2 + 8S_4^2 + 16S_{16}^2 \\ V_4 &= 2S_1^2 + 8S_4^2 + 32S_{16}^2 \end{aligned} \quad (4.11)$$

$\text{Var}[\Theta_p^i]$ représente l'estimateur de la variance de l'échantillon Θ_p^i où i est l'indice de l'observation. On a donc : $\text{Var}[\Theta_p^i] = \frac{1}{n-1} \sum_{i=1}^n [\Theta_p^i - (\frac{1}{n} \sum_{i=1}^n \Theta_p^i)]^2$. Les variables S_j^2 représentent donc les variances théoriques que nous espérons retrouver.

Cette première méthode n'est pas très efficace comme nous le verrons par la suite. En effet, on mélange toutes les périodes et les variances, alors qu'il faudrait déjà obtenir les variances pour chaque période puis travailler sur ces variables. Pour cela, il suffit de résoudre le petit système d'équations (4.10). On obtient alors :

$$\begin{aligned} \text{Var}[\Theta_p] &= S_{1,p}^2 + 4S_{4,p}^2 + 16S_{16,p}^2 + S_{96,p}^2 \\ \text{Var}[\Theta_{p+1} - \Theta_p] &= 2S_{1,p}^2 + 2S_{4,p}^2 + 2S_{16,p}^2 \\ \text{Var}[\Theta_{p+8} - \Theta_p] &= 2S_{1,p}^2 + 8S_{4,p}^2 + 16S_{16,p}^2 \\ \text{Var}[\Theta_{p+32} - \Theta_p] &= 2S_{1,p}^2 + 8S_{4,p}^2 + 32S_{16,p}^2 \end{aligned}$$

$$\forall p = 16 \dots 64,$$

$$\begin{aligned} S_{1,p}^2 &= \frac{4\text{Var}[\Theta_{p+32} - \Theta_p] - 12\text{Var}[\Theta_{p+8} - \Theta_p] + 32\text{Var}[\Theta_{p+1} - \Theta_p]}{24} \\ S_{4,p}^2 &= \frac{-7\text{Var}[\Theta_{p+32} - \Theta_p] + 15\text{Var}[\Theta_{p+8} - \Theta_p] - 8\text{Var}[\Theta_{p+1} - \Theta_p]}{48} \\ S_{16,p}^2 &= \frac{\text{Var}[\Theta_{p+32} - \Theta_p] - \text{Var}[\Theta_{p+8} - \Theta_p]}{16} \\ S_{96,p}^2 &= \frac{2\text{Var}[\Theta_p] - \text{Var}[\Theta_{p+32} - \Theta_p]}{2} \end{aligned} \quad (4.12)$$

Par la suite nous résolvons le même type de problème que (4.11), mais nous incorporons les nouvelles valeurs $(S_{j,p}^2)$ comme observations :

$$\begin{aligned} \min_{V, S^2} & |V_1 - \text{Var}_i[\Theta_p^i]| + |V_2 - \text{Var}_i[\Theta_{p+1}^i - \Theta_p^i]| + |V_3 - \text{Var}_i[\Theta_{p+8}^i - \Theta_p^i]| \\ & + |V_4 - \text{Var}_i[\Theta_{p+32}^i - \Theta_p^i]| + c \sum_{j=1,4,16,96} |S_j^2 - S_{j,p}^2| \end{aligned}$$

soumis aux contraintes :

$$\begin{aligned} V_1 &= S_1^2 + 4S_4^2 + 16S_{16}^2 + S_{96}^2 \\ V_2 &= 2S_1^2 + 2S_4^2 + 2S_{16}^2 \\ V_3 &= 2S_1^2 + 8S_4^2 + 16S_{16}^2 \\ V_4 &= 2S_1^2 + 8S_4^2 + 32S_{16}^2 \end{aligned} \tag{4.13}$$

c est une constante qu'il faut ajuster afin d'obtenir les meilleurs résultats possible. Dans notre cas nous avons pris $c = 2$. Les résultats de ces deux méthodes sont présentés dans les tableaux (4.1) et (4.2). Nous avons résolu les deux problèmes d'optimisation avec les mêmes paramètres ; ils ont été résolus 1000 fois et c'est l'écart-type des solutions S_j qui est présenté. Le nombre de scénarios augmente afin d'observer la vitesse de convergence de ces estimateurs. Les simulations se sont arrêtées à 1000 scénarios, car il est improbable actuellement d'arriver à en obtenir plus dans la réalité. En effet, 1000 scénarios correspond environ à trois années de données. Il est déjà très difficile d'obtenir des données aussi précises sur quelques mois, alors sur trois années cela semble impossible.

Il est d'abord très important d'insister sur le fait que ces valeurs sont très difficiles à interpréter en tant que telles. Les variances S_j^2 représentent les variances de chaque type de perturbation. Par contre, il n'est pas possible de conclure si l'estimation qui en suit est bonne ; en effet peut-être qu'un écart de 0,1% de la valeur théorique peut impliquer des grandes différences de coût dans le planning.

Par contre, il est tout à fait naturel de vouloir comparer ces valeurs entre elles. C'est ce que nous faisons et ferons pour les prochains tableaux de résultats. Tout d'abord, les écarts-type pour les deux méthodes diminuent lorsque le nombre de scénarios augmente, ce qui est normal heureusement. Il apparaît rapidement que les résultats de la deuxième méthode sont meilleurs ; en effet, dans tous les cas, la deuxième méthode obtient de plus petits écarts-type. Les deux méthodes ont l'air de converger, mais

Nombre de scénarios	Type de perturbations	1	4	16	96
	Probabilité β_j	1	0,2	0,2	1
	Écart-type σ_j	0,01	0,2225	0,2225	0,01
	Écart-type				
100	S_j	0,1689	0,2398	0,0076	0,0643
300	S_j	0,0091	0,0133	0,0045	0,0463
500	S_j	0,0069	0,0098	0,0036	0,0354
700	S_j	0,0059	0,0084	0,0030	0,0298
1000	S_j	0,0050	0,0070	0,0025	0,0246

TABLEAU 4.1 Résultats pour la première méthode de l'écart-type de S_j

Nombre de scénarios	Type de perturbations	1	4	16	96
	Probabilité β_j	1	0,2	0,2	1
	Écart-type σ_j	0,01	0,2225	0,2225	0,01
	Écart-type				
100	S_j	0,0166	0,0200	0,0091	0,0983
300	S_j	0,0084	0,0113	0,0045	0,0384
500	S_j	0,0065	0,0089	0,0033	0,0281
700	S_j	0,0055	0,0076	0,0028	0,0249
1000	S_j	0,0046	0,0064	0,0023	0,0209

TABLEAU 4.2 Résultats pour la deuxième méthode de l'écart-type de S_j

la deuxième le fait un peu plus rapidement. Par conséquent, elle est un meilleur estimateur que la première et c'est elle qui sera utilisée par la suite. Enfin, l'estimation de la variance de la tendance a un écart-type plus important que les autres, ses valeurs risquent d'être moins précises par la suite.

La connaissance de S_j ne permet pas de retrouver σ_j et β_j pour $j = 4$ ou 16 . Les deux figures (4.5) et (4.6) représentent deux simulations du modèle pour des variances S_j^2 constantes entre les deux réalisations. Sur les deux figures, une des courbes est la réalisation et l'autre la demande prévisionnelle. La figure (4.5) a des perturbations d'amplitude plus faibles que celles de la figure (4.6). On aperçoit bien à l'oeil nu que les paramètres σ_j sont sûrement plus importants pour la deuxième figure ; normalement les paramètres β_j sont censés être plus importants pour la première figure afin de conserver des variances S_j^2 égales, mais ce fait est moins évident. Les deux réalisations obtenues sont donc bien différentes et il nous faut donc trouver d'autres estimateurs pour séparer les paramètres β_j et σ_j .

4.3 Estimation des moments

Dans cette section, nous présentons l'élaboration d'autres estimateurs et les résultats obtenus pour le dimensionnement. La méthode des moments, présentée précédemment, est utilisée pour construire d'autres équations. Le calcul des variances réalisées est en fait l'utilisation des moments deux de la variation Θ_p . Nous essayons d'obtenir le même type d'équations que celles présentées dans la formule (4.10) en considérant d'autres moments des variations.

4.3.1 Moment d'ordre un des variations Θ_p

Le moment un correspond en réalité à la moyenne, il est difficile de l'utiliser comme tel. En effet, la moyenne est peu sensible aux paramètres car les perturbations peuvent s'annuler entre elles. De plus, il est impossible de considérer la différence entre deux variations Θ_k et Θ_p car elles s'annulent totalement comme le montre la formule

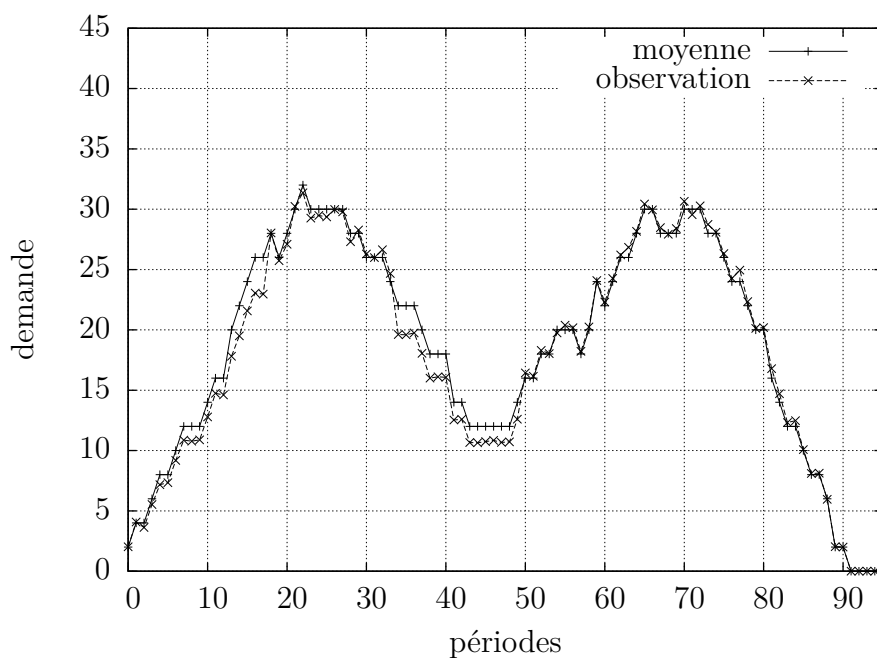


FIGURE 4.5 Représentation d'une observation avec de petites perturbations

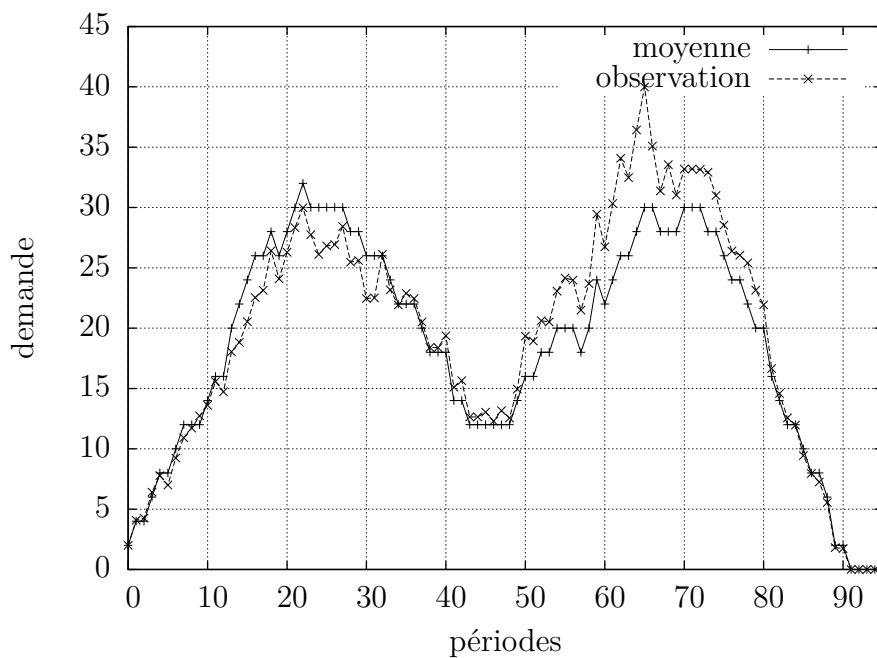


FIGURE 4.6 Représentation d'une observation avec de grandes perturbations

suivante :

$$\begin{aligned}\forall p, k = 16 \dots 64, \mathbb{E}[\Theta_k - \Theta_p] &= \mathbb{E}[\Theta_k] - \mathbb{E}[\Theta_p] \\ &= \mathbb{E}[\Theta_p] - \mathbb{E}[\Theta_p] \\ &= 0\end{aligned}$$

De plus pour la moyenne, on obtient les équations suivantes :

$$\begin{aligned}\forall p = 16 \dots 64, \text{ on a :} \\ \mathbb{E}[\Theta_p] &= \mathbb{E}[\Theta_p^{bruit}] + \mathbb{E}[\Theta_p^{1heure}] + \mathbb{E}[\Theta_p^{4heures}] + \mathbb{E}[\Theta_p^{tendance}] \\ &= \mathbb{E}[\rho_{1,p}] + \mathbb{E}\left[\sum_{k=\max(1,p-3)}^p \rho_{4,k}\right] + \mathbb{E}\left[\sum_{k=\max(1,p-15)}^p \rho_{16,k}\right] + \mathbb{E}[\rho_{96}]\end{aligned}$$

D'où :

$$\mathbb{E}[\Theta_p] = -\frac{\sigma_1^2}{2} - 4\frac{\beta_4\sigma_4^2}{2} - 16\frac{\beta_{16}\sigma_{16}^2}{2} - \frac{\sigma_{96}^2}{2}$$

Type de perturbations	1	4	16	96
Probabilité β_j	1	0,2	0,2	1
Écart-type σ_j	0,01	0,2225	0,2225	0,01

Nombre de scénarios	Écart-type de $\mathbb{E}[\Theta_p]$
100	0,0259
300	0,0141
500	0,0114
700	0,0091
1000	0,0078

TABLEAU 4.3 Résultats de l'écart-type de $\mathbb{E}[\Theta_p]$

L'estimateur classique de la moyenne $\mathbb{E}[\Theta_p^i] = \frac{1}{n} \sum_{i=1}^n \Theta_{i,p}$ est un bon estimateur et converge rapidement. C'est pourquoi une simple moyenne sur les périodes doit suffire. La même instance que celle prise pour les résultats de la variance a été utilisée. De la même manière que pour la variance, on calcule l'écart-type de la moyenne à partir

de 1000 observations. Les résultats sont présentés dans le tableau (4.3). On observe à nouveau que l'écart-type de la moyenne diminue lorsque le nombre de scénarios augmente. Ces résultats sont tout à fait acceptables au vu des erreurs commises pour la variance.

Nous avons maintenant cinq équations, il en manque une dernière.

4.3.2 Moment d'ordre trois des variations Θ_p

Nous allons calculer le moment trois de la variation Θ_p . Il y a un problème avec cette méthode, le moment d'ordre trois n'est pas linéaire. Puisque les variations Θ_p sont la somme de vingt-deux perturbations, les moments d'ordre trois sont compliqués à exprimer. Afin de simplifier les calculs par la suite, on va construire une mesure L_3 inspirées du moment d'ordre trois afin que deux variables aléatoires réelles et indépendantes X et Y vérifient $\mathbb{L}_3[X + Y] = \mathbb{L}_3[X] + \mathbb{L}_3[Y]$. Nous avons déjà :

$$\begin{aligned}\mathbb{E}[(X + Y)^3] &= \mathbb{E}[X^3 + 3X^2Y + 3XY^2 + Y^3] \\ &= \mathbb{E}[X^3] + 3\mathbb{E}[X^2Y] + 3\mathbb{E}[XY^2] + \mathbb{E}[Y^3]\end{aligned}$$

Or les variables X et Y sont indépendantes,

$$\mathbb{E}[(X + Y)^3] = \mathbb{E}[X^3] + 3\mathbb{E}[X^2]\mathbb{E}[Y] + 3\mathbb{E}[X]\mathbb{E}[Y^2] + \mathbb{E}[Y^3]$$

Il faut trouver un moyen de supprimer les deux termes $\mathbb{E}[X^2]\mathbb{E}[Y]$ et $\mathbb{E}[X]\mathbb{E}[Y^2]$ afin de trouver la mesure \mathbb{L}_3 cherchée. Considérons le moment d'ordre deux :

$$\begin{aligned}\mathbb{E}[(X + Y)^2] &= \mathbb{E}[X^2 + 2XY + Y^2] \\ &= \mathbb{E}[X^2] + 2\mathbb{E}[XY] + \mathbb{E}[Y^2] \\ &= \mathbb{E}[X^2] + 2\mathbb{E}[X]\mathbb{E}[Y] + \mathbb{E}[Y^2]\end{aligned}$$

On obtient ainsi :

$$\begin{aligned}\mathbb{E}[(X + Y)^2]\mathbb{E}[X + Y] &= \mathbb{E}[X^2]\mathbb{E}[X] + 2(\mathbb{E}[X])^2\mathbb{E}[Y] + \mathbb{E}[X]\mathbb{E}[Y^2] \\ &\quad + \mathbb{E}[X^2]\mathbb{E}[Y] + 2\mathbb{E}[X](\mathbb{E}[Y])^2 + \mathbb{E}[Y^2]\mathbb{E}[Y]\end{aligned}$$

D'où :

$$\begin{aligned}\mathbb{E}[(X + Y)^3] - 3\mathbb{E}[(X + Y)^2]\mathbb{E}[X + Y] &= \mathbb{E}[X^3] + 3\mathbb{E}[X^2]\mathbb{E}[Y] + 3\mathbb{E}[X]\mathbb{E}[Y^2] + \mathbb{E}[Y^3] \\ &\quad - 3\mathbb{E}[X^2]\mathbb{E}[X] - 6(\mathbb{E}[X])^2\mathbb{E}[Y] - 3\mathbb{E}[X]\mathbb{E}[Y^2] \\ &\quad - 3\mathbb{E}[X^2]\mathbb{E}[Y] - 6\mathbb{E}[X](\mathbb{E}[Y])^2 - 3\mathbb{E}[Y^2]\mathbb{E}[Y] \\ \mathbb{E}[(X + Y)^3] - 3\mathbb{E}[(X + Y)^2]\mathbb{E}[X + Y] &= \mathbb{E}[X^3] - 3\mathbb{E}[X^2]\mathbb{E}[X] + \mathbb{E}[Y^3] - 3\mathbb{E}[Y^2]\mathbb{E}[Y] \\ &\quad - 6(\mathbb{E}[X])^2\mathbb{E}[Y] - 6\mathbb{E}[X](\mathbb{E}[Y])^2\end{aligned}$$

Enfin afin de supprimer les deux derniers termes $(\mathbb{E}[X])^2\mathbb{E}[Y]$ et $\mathbb{E}[X](\mathbb{E}[Y])^2$, considérons la formule suivante :

$$(\mathbb{E}[X + Y])^3 = (\mathbb{E}[X])^3 + 3(\mathbb{E}[X])^2\mathbb{E}[Y] + 3(\mathbb{E}[Y])^2\mathbb{E}[X] + (\mathbb{E}[Y])^3$$

Ainsi, nous pouvons définir \mathbb{L}_3 ainsi :

$$\mathbb{L}_3[X] = \mathbb{E}[X^3] - 3\mathbb{E}[X^2]\mathbb{E}[X] + 2(\mathbb{E}[X])^3 \quad (4.14)$$

En effet :

$$\begin{aligned}\mathbb{L}_3[X + Y] &= \mathbb{E}[(X + Y)^3] - 3\mathbb{E}[(X + Y)^2]\mathbb{E}[X + Y] + 2(\mathbb{E}[X + Y])^3 \\ &= (\mathbb{E}[X^3] - 3\mathbb{E}[X^2]\mathbb{E}[X]) + (\mathbb{E}[Y^3] - 3\mathbb{E}[Y^2]\mathbb{E}[Y]) \\ &\quad - 6(\mathbb{E}[X])^2\mathbb{E}[Y] - 6\mathbb{E}[X](\mathbb{E}[Y])^2 + 6(\mathbb{E}[X])^2\mathbb{E}[Y] + 6(\mathbb{E}[Y])^2\mathbb{E}[X] \\ &\quad + 2(\mathbb{E}[X])^3 + 2(\mathbb{E}[Y])^3 \\ \mathbb{L}_3[X + Y] &= \mathbb{L}_3[X] + \mathbb{L}_3[Y]\end{aligned}$$

D'après la formule (4.14), nous avons pour la mesure d'ordre trois de la variation Θ_p :

$$\mathbb{L}_3[\Theta_p] = \mathbb{L}_3[\Theta_p^{bruit}] + \mathbb{L}_3[\Theta_p^{1heure}] + \mathbb{L}_3[\Theta_p^{4heures}] + \mathbb{L}_3[\Theta_p^{tendance}]$$

A partir de maintenant et pour le reste du texte nous utiliserons la notation suivante : $\rho_{96,p} = \rho_{96}$. En effet cette notation est justifiée, l'indice p a été omis jusqu'ici car la perturbation de longueur une journée ne dépend pas de la période. Rajouter un indice p permet donc d'harmoniser toutes les notations. Utilisons les formules (3.2.1) des moments pour calculer la mesure d'ordre trois. Pour les perturbations de longueur 1 et 96, $j = 1, 96$, on a :

$$\begin{aligned} \mathbb{E}[\rho_{j,p}] &= \mu_j \\ &= \frac{-\sigma_j^2}{2} \end{aligned}$$

$$\begin{aligned} \mathbb{E}[\rho_{j,p}^2] &= \mu_j^2 + \sigma_j^2 \\ &= \left(\frac{-\sigma_j^2}{2}\right)^2 + \sigma_j^2 \\ &= \frac{\sigma_j^4}{4} + \sigma_j^2 \end{aligned}$$

$$\begin{aligned} \mathbb{E}[\rho_{j,p}^3] &= \mu_j^3 + 3\mu_j\sigma_j^2 \\ &= \left(\frac{-\sigma_j^2}{2}\right)^3 + 3\frac{-\sigma_j^2}{2}\sigma_j^2 \\ &= \frac{-\sigma_j^6}{8} - 3\frac{\sigma_j^4}{2} \end{aligned}$$

$$\begin{aligned} \mathbb{L}_3[\rho_{j,p}] &= \frac{-\sigma_j^6}{8} - 3\frac{\sigma_j^4}{2} - 3\left(\frac{\sigma_j^4}{4} + \sigma_j^2\right)\frac{-\sigma_j^2}{2} + 2\left(\frac{-\sigma_j^2}{2}\right)^3 \\ &= \frac{-\sigma_j^6}{8} - 3\frac{\sigma_j^4}{2} + 3\frac{\sigma_j^6}{8} + 3\frac{\sigma_j^4}{2} - 2\frac{\sigma_j^6}{8} = 0 \end{aligned}$$

De même, on a pour les variables conditionnelles $j = 4, 16$:

$$\begin{aligned}
\mathbb{E}[\rho_{j,p}] &= \beta_j \frac{-\sigma_j^2}{2} \\
\mathbb{E}[\rho_{j,p}^2] &= \beta_j \left(\sigma_j^2 + \frac{\sigma_j^4}{4} \right) \\
\mathbb{E}[\rho_{j,p}^3] &= \mathbb{E}[\rho_{j,p}^3 | \delta_{j,p} = 1] \mathbb{P}[\delta_{j,p} = 1] + \mathbb{E}[\rho_{j,p}^3 | \delta_{j,p} = 0] \mathbb{P}[\delta_{j,p} = 0] \\
&= \beta_j (\mu_j^3 + 3\mu_j \sigma_j^2) + 0 \\
&= \beta_j \frac{-\sigma_j^6}{8} - 3\beta_j \frac{\sigma_j^4}{2} \\
\mathbb{L}_3[\rho_{j,p}] &= \beta_j \frac{-\sigma_j^6}{8} - 3\beta_j \frac{\sigma_j^4}{2} - 3\beta_j^2 \left(\frac{\sigma_j^4}{4} + \sigma_j^2 \right) \frac{-\sigma_j^2}{2} + 2\beta_j^3 \left(\frac{-\sigma_j^2}{2} \right)^3 \\
&= \frac{\sigma_j^6}{8} (-\beta_j + 3\beta_j^2 - 2\beta_j^3) + 3\beta_j \frac{\sigma_j^4}{2} (\beta_j - 1)
\end{aligned}$$

Ainsi, on obtient :

$$\begin{aligned}
\mathbb{L}_3[\Theta_p] &= \frac{\sigma_4^6}{2} (-\beta_4 + 3\beta_4^2 - 2\beta_4^3) + 6\beta_4 \sigma_4^4 (\beta_4 - 1) \\
&\quad + 2\sigma_{16}^6 (-\beta_{16} + 3\beta_{16}^2 - 2\beta_{16}^3) + 24\beta_{16} \sigma_{16}^4 (\beta_{16} - 1)
\end{aligned} \tag{4.15}$$

On peut déjà noter qu'on vient en fait de retrouver un concept statistique : les cumulants. Lukacs (1970) expose différentes propriétés de ces outils dont la linéarité. La mesure d'ordre quatre (4.16) que nous donnons par la suite n'est en fait que le quatrième cumulant. La moyenne et la variance sont le premier et le deuxième cumulant : nous travaillons en fait sur les quatre premiers cumulants. Il est aussi intéressant de remarquer que cette mesure est nulle, si la variable n'est pas conditionnée. On peut dire qu'elle est vraiment adaptée à l'estimation des probabilités β_j . Il n'est pas possible à nouveau d'observer la différence entre deux variations Θ_k et Θ_p , car la mesure \mathbb{L}_3 est nulle :

$$\forall p, k = 16 \dots 64, \mathbb{L}_3[\Theta_k - \Theta_p] = \mathbb{L}_3[\Theta_k] - \mathbb{L}_3[\Theta_p] = 0$$

De la même manière que pour la moyenne, on calcule l'écart-type de la mesure d'ordre trois à partir de 1000 observations. Les résultats sont présentés dans le tableau (4.4).

Type de perturbations	1	4	16	96
Probabilité β_j	1	0,2	0,2	1
Écart-type σ_j	0,01	0,2225	0,2225	0,01

Nombre de scénarios	Écart-type de $[\mathbb{L}_3]^{\frac{1}{3}}$
100	0,0649
300	0,0502
500	0,0397
700	0,0338
1000	0,0268

TABLEAU 4.4 Résultats de l'écart-type de $[\mathbb{L}_3]^{\frac{1}{3}}$

Comme auparavant l'écart-type de la mesure d'ordre trois diminue lorsque le nombre de scénarios augmente.

Nous avons obtenu ainsi une sixième équation. Nous allons quand même chercher d'autres formules avec les moments d'ordre quatre. En effet, si nous modifions le modèle et que les perturbations de longueur une deviennent conditionnées, il faut estimer dorénavant sept paramètres. Le fait de calculer les quatrièmes cumulants permet d'obtenir quelque soit la configuration une équation par paramètre.

4.3.3 Moment d'ordre quatre des variations Θ_p

Construisons maintenant une mesure d'ordre quatre \mathbb{L}_4 qui a la même propriété que la mesure \mathbb{L}_3 et qui prend en compte le moment d'ordre quatre. En suivant le raisonnement décrit dans le paragraphe précédent, nous obtenons les étapes suivantes :

$$\begin{aligned}\mathbb{E}[(X + Y)^4] &= \mathbb{E}[X^4] + 4\mathbb{E}[X^3]\mathbb{E}[Y] + 6\mathbb{E}[X^2]\mathbb{E}[Y^2] \\ &\quad + 4\mathbb{E}[X]\mathbb{E}[Y^3] + \mathbb{E}[Y^4]\end{aligned}$$

$$\begin{aligned}\mathbb{E}[(X + Y)^3]\mathbb{E}[(X + Y)] &= \\ &\mathbb{E}[X^3]\mathbb{E}[X] + 3\mathbb{E}[X^2]\mathbb{E}[X]\mathbb{E}[Y] + 3(\mathbb{E}[X])^2\mathbb{E}[Y^2] + \mathbb{E}[X]\mathbb{E}[Y^3] \\ &\quad + \mathbb{E}[X^3]\mathbb{E}[Y] + 3\mathbb{E}[X^2](\mathbb{E}[Y])^2 + 3\mathbb{E}[X]\mathbb{E}[Y^2]\mathbb{E}[Y] + \mathbb{E}[Y]\mathbb{E}[Y^3]\end{aligned}$$

$$\begin{aligned}
(\mathbb{E}[(X + Y)^2])^2 &= (\mathbb{E}[X^2])^2 + 2\mathbb{E}[X^2]\mathbb{E}[X]\mathbb{E}[Y] + \mathbb{E}[X^2]\mathbb{E}[Y^2] \\
&\quad + 2\mathbb{E}[X^2]\mathbb{E}[X]\mathbb{E}[Y] + 4(\mathbb{E}[X])^2(\mathbb{E}[Y])^2 + 2\mathbb{E}[X]\mathbb{E}[Y]\mathbb{E}[Y^2] \\
&\quad + \mathbb{E}[X^2]\mathbb{E}[Y^2] + 2\mathbb{E}[Y^2]\mathbb{E}[X]\mathbb{E}[Y] + (\mathbb{E}[Y^2])^2 \\
(\mathbb{E}[(X + Y)^2])^2 &= (\mathbb{E}[X^2])^2 + 4\mathbb{E}[X^2]\mathbb{E}[X]\mathbb{E}[Y] + 2\mathbb{E}[X^2]\mathbb{E}[Y^2] \\
&\quad + 4(\mathbb{E}[X])^2(\mathbb{E}[Y])^2 + 4\mathbb{E}[X]\mathbb{E}[Y]\mathbb{E}[Y^2] + (\mathbb{E}[Y^2])^2
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}[(X + Y)^2](\mathbb{E}[(X + Y)])^2 &= \mathbb{E}[X^2](\mathbb{E}[X])^2 + 2(\mathbb{E}[X])^3\mathbb{E}[Y] + (\mathbb{E}[X])^2\mathbb{E}[Y^2] \\
&\quad + 2\mathbb{E}[X^2]\mathbb{E}[X]\mathbb{E}[Y] + 4(\mathbb{E}[X])^2(\mathbb{E}[Y])^2 + 2\mathbb{E}[X]\mathbb{E}[Y]\mathbb{E}[Y^2] \\
&\quad + \mathbb{E}[X^2](\mathbb{E}[Y])^2 + 2(\mathbb{E}[Y])^3\mathbb{E}[X] + \mathbb{E}[Y^2](\mathbb{E}[Y])^2
\end{aligned}$$

$$\begin{aligned}
(\mathbb{E}[(X + Y)])^4 &= (\mathbb{E}[X])^4 + 4(\mathbb{E}[X])^3\mathbb{E}[Y] + 6(\mathbb{E}[X])^2(\mathbb{E}[Y])^2 \\
&\quad + 4\mathbb{E}[X](\mathbb{E}[Y])^3 + (\mathbb{E}[Y])^4
\end{aligned}$$

Posons donc :

$$\mathbb{L}_4[X] = \mathbb{E}[X^4] - 4\mathbb{E}[X^3]\mathbb{E}[X] - 3(\mathbb{E}[X^2])^2 + 12\mathbb{E}[X^2](\mathbb{E}[X])^2 - 6(\mathbb{E}[X])^4 \quad (4.16)$$

Avec les formules ci-dessus, on peut vérifier que $\mathbb{L}_4[X + Y] = \mathbb{L}_4[X] + \mathbb{L}_4[Y]$. Calculons maintenant la mesure d'ordre quatre d'une variation. Pour cela, nous avons besoin du moment d'ordre quatre d'une perturbation. D'après les formules (3.2.1), on obtient :

$$\forall j = 1, 4, 16, 96, \forall p = 1 \dots 96,$$

$$\begin{aligned}
\mathbb{E}[\rho_{j,p}^4] &= \mathbb{E}[\rho_{j,p}^4 | \delta_{j,p} = 1] \mathbb{P}[\delta_{j,p} = 1] + \mathbb{E}[\rho_{j,p}^4 | \delta_{j,p} = 0] \mathbb{P}[\delta_{j,p} = 0] \\
&= \beta_j (\mu_j^4 + 6\mu_j^2 \sigma_j^2 + 3\sigma_j^4) \\
&= \beta_j \left(\frac{\sigma_j^8}{16} + 6 \frac{\sigma_j^6}{4} + 3\sigma_j^4 \right)
\end{aligned}$$

Ainsi, on a pour la mesure \mathbb{L}_4 :

$$\forall j = 1, 4, 16, 96, \forall p = 1 \dots 96,$$

$$\begin{aligned} \mathbb{L}_4[\rho_{j,p}] &= \beta_j \left(\frac{\sigma_j^8}{16} + 6 \frac{\sigma_j^6}{4} + 3\sigma_j^4 \right) - 4\beta_j^2 \left(\frac{-\sigma_j^6}{8} - 3 \frac{\sigma_j^4}{2} \right) \left(\frac{-\sigma_j^2}{2} \right) \\ &\quad - 3\beta_j^2 \left(\frac{\sigma_j^4}{4} + \sigma_j^2 \right)^2 + 12\beta_j^3 \left(\frac{\sigma_j^4}{4} + \sigma_j^2 \right) \left(\frac{-\sigma_j^2}{2} \right)^2 - 6\beta_j^4 \left(\frac{-\sigma_j^2}{2} \right)^4 \\ \mathbb{L}_4[\rho_{j,p}] &= \beta_j \left(\frac{\sigma_j^8}{16} + 6 \frac{\sigma_j^6}{4} + 3\sigma_j^4 \right) + 12\beta_j^3 \left(\frac{\sigma_j^8}{16} + \frac{\sigma_j^6}{4} \right) \\ &\quad - \beta_j^2 \left(\frac{\sigma_j^8}{4} + 3\sigma_j^6 + 3 \frac{\sigma_j^8}{16} + 3 \frac{\sigma_j^6}{2} + 3\sigma_j^4 \right) - 6\beta_j^4 \left(\frac{\sigma_j^8}{16} \right) \\ \mathbb{L}_4[\rho_{j,p}] &= \beta_j \left(\frac{\sigma_j^8}{16} + 6 \frac{\sigma_j^6}{4} + 3\sigma_j^4 \right) + 12\beta_j^3 \left(\frac{\sigma_j^8}{16} + \frac{\sigma_j^6}{4} \right) \\ &\quad - \beta_j^2 \left(7 \frac{\sigma_j^8}{16} + 9 \frac{\sigma_j^6}{2} + 3\sigma_j^4 \right) - 6\beta_j^4 \left(\frac{\sigma_j^8}{16} \right) \end{aligned} \tag{4.17}$$

Remarquons que pour $j = 1, 96$, l'expression (4.17) se simplifie grandement. En effet, dans ce cas, $\beta_j = 1$. On a alors :

$$\begin{aligned} \mathbb{L}_4[\rho_{j,p}] &= \frac{\sigma_j^8}{16} + 6 \frac{\sigma_j^6}{4} + 3\sigma_j^4 + 12 \left(\frac{\sigma_j^8}{16} + \frac{\sigma_j^6}{4} \right) \\ &\quad - \left(7 \frac{\sigma_j^8}{16} + 9 \frac{\sigma_j^6}{2} + 3\sigma_j^4 \right) - 6 \left(\frac{\sigma_j^8}{16} \right) \end{aligned} \tag{4.18}$$

On obtient donc : $\mathbb{L}_4[\rho_{j,p}] = 0$. De la même manière que pour la mesure d'ordre 3, la mesure d'ordre 4 est non nulle pour les perturbations conditionnelles. Ces deux mesures devraient normalement capter le paramètre β_j de chaque type de perturbations. L'avantage de cette mesure par rapport à la précédente est qu'elle est toujours positive. Il est donc possible é nouveau d'observer la différence entre deux variations Θ_k et Θ_p à travers cette mesure \mathbb{L}_4 . En effet, de la même manière que pour les équations (4.10) sur les variances, on a :

$$\begin{aligned}
\mathbb{L}_4[\Theta_p] &= \mathbb{L}_4[\rho_{1,p}] + 4\mathbb{L}_4[\rho_{4,p}] + 16\mathbb{L}_4[\rho_{16,p}] + \mathbb{L}_4[\rho_{96}] \\
\mathbb{L}_4[\Theta_{p+1} - \Theta_p] &= 2\mathbb{L}_4[\rho_{1,p}] + 2\mathbb{L}_4[\rho_{4,p}] + 2\mathbb{L}_4[\rho_{16,p}] \\
\mathbb{L}_4[\Theta_{p+8} - \Theta_p] &= 2\mathbb{L}_4[\rho_{1,p}] + 8\mathbb{L}_4[\rho_{4,p}] + 16\mathbb{L}_4[\rho_{16,p}] \\
\mathbb{L}_4[\Theta_{p+32} - \Theta_p] &= 2\mathbb{L}_4[\rho_{1,p}] + 8\mathbb{L}_4[\rho_{4,p}] + 32\mathbb{L}_4[\rho_{16,p}]
\end{aligned}$$

Comme les perturbations de longueur 1 et 96 ont une mesure d'ordre 4 nulle, il est possible de simplifier les formules ci-dessus. On obtient donc :

$$\begin{aligned}
\forall p &= 16 \dots 64, \\
\mathbb{L}_4[\Theta_p] &= 4\mathbb{L}_4[\rho_{4,p}] + 16\mathbb{L}_4[\rho_{16,p}] \\
\mathbb{L}_4[\Theta_{p+1} - \Theta_p] &= 2\mathbb{L}_4[\rho_{4,p}] + 2\mathbb{L}_4[\rho_{16,p}] \\
\mathbb{L}_4[\Theta_{p+8} - \Theta_p] &= 8\mathbb{L}_4[\rho_{4,p}] + 16\mathbb{L}_4[\rho_{16,p}] \\
\mathbb{L}_4[\Theta_{p+32} - \Theta_p] &= 8\mathbb{L}_4[\rho_{4,p}] + 32\mathbb{L}_4[\rho_{16,p}]
\end{aligned}$$

Pour les variances (4.12), le système a quatre équations et quatre inconnues. Dans ce cas il y a quatre équations et deux inconnues, la méthode utilisée dans le cas des variances n'est plus réalisable. En effet, si cette méthode est utilisée, une partie de l'information sera perdue. Il suffit d'utiliser deux équations pour trouver les inconnues, il y a donc deux équations inutilisées. Pour régler ce problème, il faut utiliser le premier modèle (4.11) vue dans le cas des variances ce qui donne ici le problème de minimisation suivant :

$$\begin{aligned}
\min_{V,U} & |V_1 - \mathbb{L}_{4,i}[\Theta_p^i]| + |V_2 - \mathbb{L}_{4,i}[\Theta_{p+1}^i - \Theta_p^i]| + |V_3 - \mathbb{L}_{4,i}[\Theta_{p+8}^i - \Theta_p^i]| \\
& + |V_4 - \mathbb{L}_{4,i}[\Theta_{p+32}^i - \Theta_p^i]|
\end{aligned}$$

soumis aux contraintes :

$$\begin{aligned}
V_1 &= 4U_4^4 + 16U_{16}^4 \\
V_2 &= 2U_4^4 + 2U_{16}^4 \\
V_3 &= 8U_4^4 + 16U_{16}^4 \\
V_4 &= 8U_4^4 + 32U_{16}^4
\end{aligned} \tag{4.19}$$

Les variable U_j représentent les mesures d'ordre quatre des perturbations. En effet U_j est l'estimation de $\mathbb{L}_4[\rho_{j,p}]$ pour $j = 4, 16$. Le tableau (4.5) présente les résultats obtenus avec la même instance que ceux présentés dans les tableaux (4.1) et (4.2). On considère la racine quatrième de la mesure \mathbb{L}_4 afin d'avoir des erreurs d'ordre un et comparable aux autres erreurs calculées auparavant.

Nombre de scénarios	Type de perturbations	1	4	16	96
	Probabilité β_j	1	0,2	0,2	1
	Écart-type σ_j	0,01	0,2225	0,2225	0,01
	Écart-type				
100	U_j	0	0,0921	0,1282	0
300	U_j	0	0,0502	0,0807	0
500	U_j	0	0,0646	0,0427	0
700	U_j	0	0,0505	0,0294	0
1000	U_j	0	0,0454	0,0229	0

TABLEAU 4.5 Résultats de l'écart-type de U_j

L'écart-type a été calculé sur la base de 1000 tirages. Comme pour les résultats des autres mesures, on s'aperçoit dans la table (4.5) que heureusement l'écart-type de U_j diminue lorsque que l'on augmente le nombre de scénarios. Toutefois on observe que l'écart-type est plus important que pour les autres mesures. Ce résultat laisse présager des erreurs dues à la mesure d'ordre quatre. Par contre on peut remarquer une différence de vitesse de convergence entre les deux perturbations. Les perturbations de longueur seize ont une moins bonne estimation au début, mais elle rattrape finalement assez rapidement, puis dépasse l'estimation des perturbations de longueur quatre. Nous allons voir par la suite comment utiliser tous les résultats obtenus pour trouver les paramètres recherchés β_j et σ_j .

4.4 Estimation de tous les paramètres

4.4.1 Recherche par dichotomie

Les formules des variances (4.10) et des mesures d'ordre quatre (4.16) nous fournissent assez d'équations pour estimer les paramètres. Le modèle (4.13) permet d'estimer les variances S_j^2 de chaque type de perturbation et le modèle (4.19) donne accès

aux mesures d'ordre quatre U_j^4 de chaque type de perturbations qui suivent une loi conditionnée. Ainsi pour les paramètres σ_1 et σ_{96} , on obtient :

$$\begin{aligned}\sigma_1 &= S_1 \\ \sigma_{96} &= S_{96}\end{aligned}$$

Pour les autres ($j = 4, 16$), la situation est un peu plus délicate. Avec les équations (4.3), (4.4) et (4.17), il faut maintenant résoudre le système d'équations suivant :

$$S_j^2 = \beta_j \sigma_j^2 + (1 - \beta_j) \beta_j \frac{\sigma_j^4}{4} \quad (4.20)$$

$$\begin{aligned}U_j^4 &= \beta_j \left(\frac{\sigma_j^8}{16} + 6 \frac{\sigma_j^6}{4} + 3\sigma_j^4 \right) + 12\beta_j^3 \left(\frac{\sigma_j^8}{16} + \frac{\sigma_j^6}{4} \right) \\ &\quad - \beta_j^2 \left(7 \frac{\sigma_j^8}{16} + 9 \frac{\sigma_j^6}{2} + 3\sigma_j^4 \right) - 6\beta_j^4 \left(\frac{\sigma_j^8}{16} \right)\end{aligned} \quad (4.21)$$

Calculons en premier le discriminant Δ_j de l'équation (4.20) du second degré en σ_j^2 :

$$\begin{aligned}\Delta_j &= \beta_j^2 - 4 \frac{(1 - \beta_j) \beta_j}{4} (-S_j^2) \\ &= \beta_j^2 + (1 - \beta_j) \beta_j S_j^2 \\ \sqrt{\Delta_j} &= \sqrt{\beta_j^2 + (1 - \beta_j) \beta_j S_j^2}\end{aligned}$$

Comme la solution se doit d'être positive, il ne reste qu'une seule racine :

$$\begin{aligned}\sigma_j^2 &= \frac{4(-\beta_j + \sqrt{\beta_j^2 + (1 - \beta_j) \beta_j S_j^2})}{2((1 - \beta_j) \beta_j)} \\ &= \frac{2(\sqrt{1 + (\frac{1}{\beta_j} - 1) S_j^2} - 1)}{(1 - \beta_j)}\end{aligned} \quad (4.22)$$

La simplification par β_j est réalisable, car β_j est différent de zéro. Si $\beta_j = 0$, il n'y a pas de perturbation de type j . Toutes les mesures sont donc censées être nulles, ce qui ne sera jamais le cas dans la réalité. Il suffit de réinjecter la nouvelle valeur de σ_j^2 dans la formule (4.21) afin d'obtenir une équation d'une seule variable β_j . Cette dernière

équation est assez compliquée. Définissons la fonction $f(\beta_j)$ de la manière suivante :

$$f(\beta_j) = [\beta_j(\frac{\sigma_j^8}{16} + 6\frac{\sigma_j^6}{4} + 3\sigma_j^4) + 12\beta_j^3(\frac{\sigma_j^8}{16} + \frac{\sigma_j^6}{4}) - \beta_j^2(7\frac{\sigma_j^8}{16} + 9\frac{\sigma_j^6}{2} + 3\sigma_j^4) - 6\beta_j^4(\frac{\sigma_j^8}{16})]^{\frac{1}{4}} - U_j \quad (4.23)$$

telle que : $\sigma_j^2 = \frac{2(\sqrt{1 + (\frac{1}{\beta_j} - 1)S_j^2} - 1)}{(1 - \beta_j)}$

Nous pouvons déjà remarquer que le terme sous la racine quatrième est toujours positif, il n'y a donc pas de problème de définition. Nous prenons la racine quatrième de la mesure d'ordre quatre afin de comparer des valeurs qui sont du même ordre de grandeur que les paramètres qui sont étudiés, c'est-à-dire dans ce cas les écarts-type σ_j .

étudions le signe de la fonction f afin de déterminer quand elle s'annule dans l'intervalle $[0, 1]$. Commençons par l'évaluation de f aux bornes 0 et 1.

Quand β_j tend vers 0, le terme dominant est $\beta_j(\frac{\sigma_j^8}{16})$. En effet, σ_j^8 est équivalent à $\frac{2S_j}{\beta_j^4}$. La fonction f tend donc positivement vers l'infini quand β_j tend vers 0.

Par contre quand β_j tend vers 1, σ_j^2 tend vers S_j^2 . Il a été montré avec la formule (4.18) que la mesure d'ordre quatre est nulle dans cette situation. La fonction f tend ainsi vers $-U_j^4$ quand β_j tend vers 1.

Comme la fonction f est continue sur $]0, 1[$, d'après le théorème des valeurs intermédiaires elle s'annule forcément au moins une fois dans l'intervalle $]0, 1[$ si $U_j^4 > 0$. De plus, si elle s'annule deux fois dans $]0, 1[$, il existe alors plusieurs paires (β_j, σ_j) qui sont solutions et qui représentent les mêmes variations. Cette dernière idée semble impossible. Si les probabilités β_j et les écarts-type σ_j sont différents, les modèles le sont forcément. Espérons donc que l'équation $f = 0$ n'a qu'une unique solution β_j dans l'intervalle $]0, 1[$. Cette dernière question sera étudiée dans la prochaine section.

Il suffit maintenant de trouver le zéro de cette fonction par dichotomie grâce à l'algorithme (1). On obtient avec cet algorithme l'estimation $\hat{\beta}_j$ de la probabilité puis celle de l'écart-type $\hat{\sigma}_j$ grâce à la formule (4.22). Tous les paramètres sont maintenant estimés.

```

Définir le pas
min = 0, max = 1
tantque  $|f(max) - f(min)| > pas$  faire
    milieu =  $\frac{min+max}{2}$ 
    si  $f(milieu) > 0$  alors
        min = milieu
    sinon
        max = milieu
    fin si
fin tantque

```

Algorithme 1 Recherche d'un zéro dans l'intervalle $]0; 1[$ de la fonction f

4.4.2 Analyse de sensibilité

Nous allons rapidement voir comment la fonction f se comporte en fonction de ses paramètres. Heureusement, la fonction f semble ne s'annuler qu'une fois dans cet intervalle si $U_j^4 > 0$ et ne s'annule pas si $U_j^4 = 0$. La figure (4.7) illustre cette idée.

Il est sûrement possible de montrer que f' est toujours négative, mais l'expression de

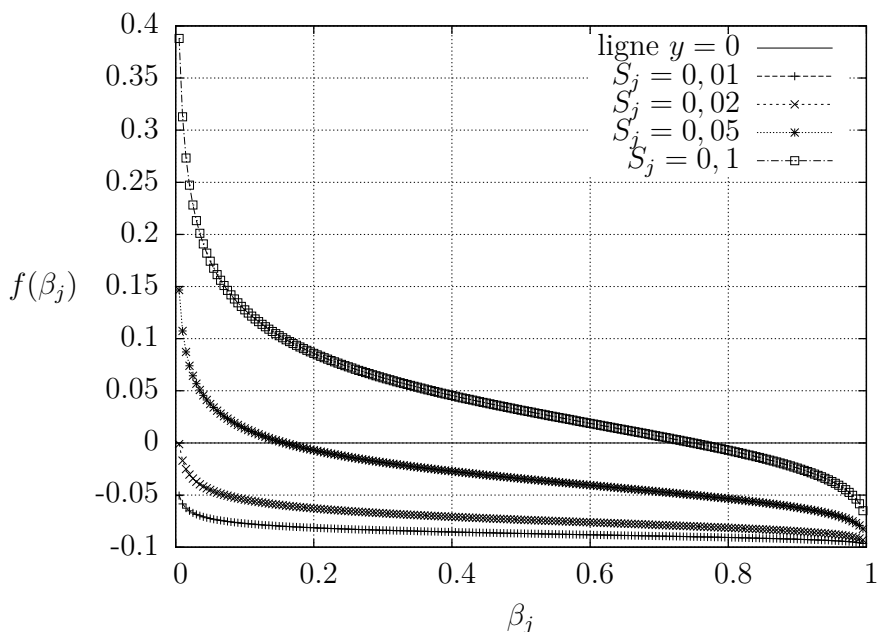


FIGURE 4.7 Représentation de f pour différentes valeurs de S_j et pour $U_j = 0, 1$

la fonction f est déjà compliquée et sa dérivée l'est encore plus. Même si l'on considère

la fonction f sans la racine quatrième, les expressions restent toujours très longues. Il apparaît aussi que la solution $\hat{\beta}_j$ de l'équation $f(\beta_j) = 0$ soit très sensible au paramètre S_j .

De plus, en fonction de l'endroit où la solution se situe, cette dernière peut être très sensible à U_j . En effet, si elle se trouve au milieu de l'intervalle, la pente de la fonction f est assez faible, une simple translation de f selon l'axe des ordonnées peut beaucoup modifier la solution.

Nombre de scénarios	Type de perturbations	1	4	16	96
	Probabilité β_j	1	0,2	0,2	1
	Écart-type σ_j	0,01	0,2225	0,2225	0,01
	Écart-type				
100	$\hat{\beta}_j$	-	0,4368	0,5638	-
	$\hat{\sigma}_j$	0,0166	0,0749	0,0989	0,0983
300	$\hat{\beta}_j$	-	0,3202	0,3385	-
	$\hat{\sigma}_j$	0,0084	0,0658	0,0671	0,0384
500	$\hat{\beta}_j$	-	0,2890	0,2049	-
	$\hat{\sigma}_j$	0,0065	0,0615	0,0485	0,0281
700	$\hat{\beta}_j$	-	0,2326	0,1390	-
	$\hat{\sigma}_j$	0,0055	0,0544	0,0395	0,0249
1000	$\hat{\beta}_j$	-	0,2035	0,1046	-
	$\hat{\sigma}_j$	0,0046	0,0489	0,0341	0,0209

TABLEAU 4.6 Résultats de l'écart-type des estimations $\hat{\beta}_j$ et $\hat{\sigma}_j$

De la même manière que pour les autres méthodes, le tableau (4.6) représente l'écart-type des paramètres β_j et σ_j sur 1000 tirages.

Comme pour tous les autres résultats, les écarts-type diminuent si le nombre de scénarios augmente. Les estimations de σ_j convergent é peu près à la même vitesse. Par contre, nous pouvons observer que les écarts-type des estimateurs des probabilités β_j sont très importants et laissent présager des mauvaises estimations de ces paramètres. Ceci peut être expliqué par le fait que la fonction f (4.23) est peu sensible pour β_j . Une petite erreur pour U_j ou S_j peut impliquer une grosse erreur pour β_j .

4.5 Résumé

Nous présentons ici un résumé de la méthode d'estimation développée.

Estimation des deuxièmes (Var) et quatrièmes (\mathbb{L}_4) cumulants

1.

$$\begin{aligned}\text{Var}[\Theta_p] &= S_1^2 + 4S_4^2 + 16S_{16}^2 + S_{96}^2 \\ \text{Var}[\Theta_{p+1} - \Theta_p] &= 2S_1^2 + 2S_4^2 + 2S_{16}^2 \\ \text{Var}[\Theta_{p+8} - \Theta_p] &= 2S_1^2 + 8S_4^2 + 16S_{16}^2 \\ \text{Var}[\Theta_{p+32} - \Theta_p] &= 2S_1^2 + 8S_4^2 + 32S_{16}^2\end{aligned}$$

où :

$$\begin{aligned}S_1^2 &= \sigma_1^2 \\ S_4^2 &= \beta_4\sigma_4^2 + (1 - \beta_4)\beta_4\frac{\sigma_4^4}{4} \\ S_{16}^2 &= \beta_{16}\sigma_{16}^2 + (1 - \beta_{16})\beta_{16}\frac{\sigma_{16}^4}{4} \\ S_{96}^2 &= \sigma_{96}^2\end{aligned}$$

2.

$$\begin{aligned}\mathbb{L}_4[\Theta_p] &= 4U_4^4 + 16U_{16}^4 \\ \mathbb{L}_4[\Theta_{p+1} - \Theta_p] &= 2U_4^4 + 2U_{16}^4 \\ \mathbb{L}_4[\Theta_{p+8} - \Theta_p] &= 8U_4^4 + 16U_{16}^4 \\ \mathbb{L}_4[\Theta_{p+32} - \Theta_p] &= 8U_4^4 + 32U_{16}^4\end{aligned}$$

où :

$$\begin{aligned}U_j^4 &= \beta_j\left(\frac{\sigma_j^8}{16} + 6\frac{\sigma_j^6}{4} + 3\sigma_j^4\right) + 12\beta_j^3\left(\frac{\sigma_j^8}{16} + \frac{\sigma_j^6}{4}\right) \\ &\quad - \beta_j^2\left(7\frac{\sigma_j^8}{16} + 9\frac{\sigma_j^6}{2} + 3\sigma_j^4\right) - 6\beta_j^4\left(\frac{\sigma_j^8}{16}\right)\end{aligned}$$

\Rightarrow On obtient S_j et U_j

Estimation des probabilités β_j et des écarts-type σ_j

Résolution du système :

$$\begin{cases} S_j^2 &= \beta_j \sigma_j^2 + (1 - \beta_j) \beta_j \frac{\sigma_j^4}{4} \\ U_j^4 &= \beta_j \left(\frac{\sigma_j^8}{16} + 6 \frac{\sigma_j^6}{4} + 3 \sigma_j^4 \right) + 12 \beta_j^3 \left(\frac{\sigma_j^8}{16} + \frac{\sigma_j^6}{4} \right) \\ &- \beta_j^2 \left(7 \frac{\sigma_j^8}{16} + 9 \frac{\sigma_j^6}{2} + 3 \sigma_j^4 \right) - 6 \beta_j^4 \left(\frac{\sigma_j^8}{16} \right) \end{cases}$$

1. Expression de σ_j en fonction de β_j et S_j

$$\sigma_4^2 = \frac{2(\sqrt{1 + (\frac{1}{\beta_4} - 1)S_4^2} - 1)}{(1 - \beta_4)}$$

$$\sigma_{16}^2 = \frac{2(\sqrt{1 + (\frac{1}{\beta_{16}} - 1)S_{16}^2} - 1)}{(1 - \beta_{16})}$$

2. On remplace dans l'expression de U_j^4 , σ_j par $\frac{2(\sqrt{1 + (\frac{1}{\beta_j} - 1)S_j^2} - 1)}{(1 - \beta_j)}$. Puis on recherche par dichotomie la seule dernière variable $\beta_j \in [0, 1]$.

Obtention de tous les paramètres β_j et σ_j .

Nous présentons par la suite les résultats obtenus sur diverses instances.

Chapitre 5

RÉSULTATS DE L'ESTIMATION

5.1 Instances

Le tableau (5.1) présente les paramètres β_j et σ_j retenus afin de construire des échantillons qui semblent correctement dimensionnés et assez représentatifs de la réalité. Toutes ces instances sont équilibrées car la variance globale de chacun des types de perturbations est à peu près constante : c'est le terme $\text{Var}[\Theta_p]$.

Instances	Numéro	β_4	β_{16}	σ_1	σ_4	σ_{16}	σ_{96}	$\sqrt{(\text{Var}[\Theta_p])}$
petites perturbations								
petit β_j , grand σ_j	1	0,025	0,025	0,01	0,06	0,06	0,01	0,02
moyen β_j , moyen σ_j	2	0,05	0,05	0,01	0,05	0,06	0,01	0,021
grand β_j , petit σ_j	3	0,1	0,1	0,01	0,03	0,06	0,01	0,02
moyennes perturbations								
petit β_j , grand σ_j	4	0,05	0,05	0,03	0,15	0,15	0,03	0,064
moyen β_j , moyen σ_j	5	0,1	0,1	0,03	0,1	0,1	0,03	0,062
grand β_j , petit σ_j	6	0,2	0,2	0,03	0,07	0,07	0,03	0,061
grandes perturbations								
petit β_j , grand σ_j	7	0,025	0,025	0,05	0,3	0,3	0,05	0,099
moyen β_j , moyen σ_j	8	0,05	0,05	0,05	0,2	0,2	0,05	0,096
grand β_j , petit σ_j	9	0,1	0,1	0,05	0,15	0,15	0,05	0,098
très grandes perturbations								
petit β_j , grand σ_j	10	0,05	0,05	0,1	0,4	0,4	0,1	0,2
moyen β_j , moyen σ_j	11	0,1	0,1	0,1	0,3	0,3	0,1	0,2
grand β_j , petit σ_j	12	0,2	0,2	0,1	0,2	0,2	0,1	0,19
géantes perturbations								
petit β_j , grand σ_j	13	0,025	0,025	0,3	1,2	1,2	0,05	0,59
moyen β_j , moyen σ_j	14	0,05	0,05	0,3	1	1	0,05	0,61
grand β_j , petit σ_j	15	0,1	0,1	0,3	0,8	0,8	0,05	0,62

TABLEAU 5.1 Instances

Ces instances serviront par la suite afin de tester le modèle et ses estimateurs. Les paramètres de chaque instance sont utilisés pour générer des scénarios à partir de la

demande prévisionnelle présentée sur la figure (5.1).

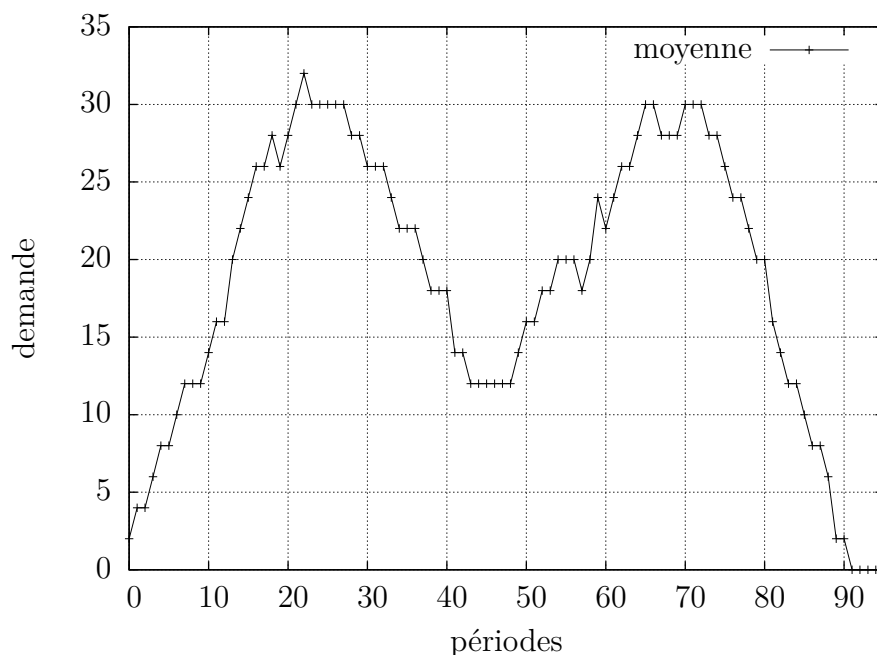


FIGURE 5.1 Représentation de la demande prévisionnelle

Les prochaines sections évaluent la qualité des instances ainsi que celle des estimateurs.

5.2 Coût des instances

Le tableau (5.2) reprend les instances représentées dans le tableau (5.1) présentées dans la section précédente. Ces dernières ont été utilisées pour faire tourner un algorithme stochastique. Le modèle de Pacqueau (2011) a été utilisé et ses programmes ont testé ces différentes données. Le tableau (5.2) présente donc le gain entre la solution IP moyenne et la solution IP stochastique. IP signifie que les problèmes ont été résolus de manière entière et que la solution est forcément entière. La solution IP moyenne est le résultat que l'on obtient en résolvant le problème de fabrication d'horaires avec la demande prévisionnelle. On évalue ensuite cette solution sur 10000 scénarios générés avec les paramètres de l'instance. Ces scénarios sont censés représenter la réalité.

Au contraire, la solution IP stochastique est obtenue à partir d'un problème stochastique. On résout un problème maître où on assigne les quarts aux employés puis on

résout le sous-problème stochastique qui minimise le recours. Ce dernier comprend l'affectation des heures supplémentaires et des pauses et l'appel d'employés à temps partiel pour la journée. Ce sous-problème prend comme données 500 scénarios de l'instance. Le coût de la solution stochastique est aussi évalué sur les 10000 scénarios. Le gain G présenté dans le tableau suit donc la formule suivante :

$$G = \frac{Sol_m - Sol_s}{Sol_m}$$

où : Sol_m est le coût de la solution moyenne et Sol_s est le coût de la solution stochastique.

La valeur du gain permet donc d'évaluer l'apport de l'optimisation stochastique par rapport à l'optimisation déterministe. Le nombre d'employés à plein temps que la solution utilise permet aussi de juger du poids du recours. C'est ce nombre qui donne le coût du problème de première décision : le reste du coût est le recours.

Le tableau (5.2) montre en premier lieu que si la variance totale S_j^2 augmente, l'optimisation stochastique est plus efficace : le gain et le coût du recours augmentent. De plus au sein d'un même type d'instances, moyennes perturbations par exemple, le gain augmente aussi si les perturbations sont moins rares mais de plus petites amplitudes. Ces deux premiers résultats sont rassurants car ils confirment notre intuition :

- plus la variance des scénarios est importante, plus l'optimisation stochastique est efficace.

- le coût du recours est moins important s'il y a beaucoup de petites perturbations.

De plus, il faut remarquer que la demande prévisionnelle est la même pour toutes les instances. Ainsi, pour un même type d'instances, le nombre d'employés à plein temps est à peu près constant. L'horaire doit donc être à peu près le même pour ces instances. La solution du problème maître dépend ainsi assez fortement de la variance de la demande. En effet, un groupe d'instances a des paramètres β_j et σ_j différents mais les mêmes variances. Les écarts de coûts au sein d'un même groupe viendraient donc principalement du coût de recours. Il est quand même utile de ne pas limiter la modélisation des perturbations à leur variance. En effet, il suffit de regarder les gains obtenus pour se rendre compte que les instances au sein d'un même groupe représentent bien des réalités différentes.

Instances	Numéro	Gain G en %	Nombre d'employés à plein temps
petites perturbations			
petit β_j , grand σ_j	1	-0,15	55
moyen β_j , moyen σ_j	2	0,07	56
grand β_j , petit σ_j	3	-0,04	55
moyennes perturbations			
petit β_j , grand σ_j	4	0,69	52
moyen β_j , moyen σ_j	5	0,86	52
grand β_j , petit σ_j	6	0,82	52
grandes perturbations			
petit β_j , grand σ_j	7	0,73	53
moyen β_j , moyen σ_j	8	1,11	52
grand β_j , petit σ_j	9	1,56	51
très grandes perturbations			
petit β_j , grand σ_j	10	2,30	40
moyen β_j , moyen σ_j	11	3,45	38
grand β_j , petit σ_j	12	3,74	38
perturbations géantes			
petit β_j , grand σ_j	13	5,82	25
moyen β_j , moyen σ_j	14	8,48	26
grand β_j , petit σ_j	15	11,57	26

TABLEAU 5.2 Gain de l'optimisation stochastique

On peut rajouter que le dernier jeu d'instances (les géantes perturbations) semblent peu réalistes ou, en tout cas, semblent représenter une situation rare. En effet, ces dernières instances ont plus de deux fois moins d'employés que les premières, c'est-à-dire que plus de la moitié de la demande est en général couverte par les heures supplémentaires et les employés à temps partiels. De telles situations semblent donc peu probables dans des contextes comme ceux des supermarchés, des usines ou des hôpitaux par exemple.

Enfin on peut noter que le gain des instances numéro 1 et 3 est négatif; ce résultat n'est pas forcément aberrant. La solution stochastique n'est pas optimale, une heuristique est utilisée dans l'algorithme d'énumération implicite qui calcule le coût du recours. Comme ces gains sont proches de zéro, il est possible, quand on ajoute l'erreur due à l'heuristique, d'obtenir une solution stochastique moins bonne que la solution moyenne.

5.3 Résultats

Pour chaque instance, 500 scénarios ont été générés. Les résultats présentés dans le tableau (5.3) ont été calculés à partir de ces données. La procédure d'estimation construite dans le chapitre précédent a été utilisée pour retrouver les paramètres $\hat{\beta}_j$ et $\hat{\sigma}_j$ de chacune des instances.

Pour critiquer la qualité des résultats, les paramètres $\hat{\beta}_j$ et $\hat{\sigma}_j$ ont été retenus pour générer à nouveau 500 scénarios. Ces derniers ont été réinsérés dans l'algorithme d'optimisation stochastique afin d'obtenir une solution. Enfin le coût de cette solution est évalué sur les mêmes 10000 scénarios qui ont été construits pour le calcul du coût et qui représentent la réalité de l'instance. Toute cette procédure permet d'évaluer la qualité de l'estimation : si le (nouveau) gain reste positif, les paramètres continuent d'apporter une amélioration. De plus, il est possible de comparer ce gain à celui (l'ancien) présenté dans le tableau (5.4). A priori il est moins bon, mais c'est surtout l'écart entre les deux qui est intéressant.

Il est difficile d'interpréter le tableau (5.3) car les paramètres ne signifient pas grand chose si le gain de l'instance associée n'est pas étudié. Il y a toutefois quelques remarques qui peuvent être faites. Tout d'abord, les résultats de l'estimation sont très variables d'une instance à une autre. Pour les instances 1 à 5, les résultats semblent être bons et assez proches des paramètres originaux ; c'est aussi le cas de l'instance 9. Par contre, pour les autres données, on observe des grandes erreurs dans l'estimation de la probabilité β_j . Les deux extrêmes de cette valeur (0,001 et 1) sont plusieurs fois obtenus. Nous sommes tentés de dire que l'estimation est mauvaise dans ce cas. Cette grande variance de la probabilité a déjà été aperçue dans les résultats du tableau (4.6). De plus, les paramètres σ_j pour $j = 4, 16$ explosent à partir de l'instance 10. Ce phénomène peut être expliqué grâce à la mesure d'ordre quatre (4.16). En effet, quand les écarts augmentent entre les variations Θ_p et Θ_k , l'exposant quatre de la mesure fait exploser les différences. C'est pourquoi les paramètres σ_j sont assez grands dans le cas des perturbations conditionnées.

Enfin, on peut remarquer que l'estimation se détériore de manière générale quand la variance $\text{Var}[\Theta_p]$ augmente. Elle devient encore plus approximative pour les perturbations de types 4 et 16 à cause des deux raisons énoncées ci-dessus. De toute façon, c'est l'étude du tableau (5.4) qui permettra de trancher sur la qualité des estimateurs.

Instances	Numéro	$\hat{\beta}_4$	$\hat{\beta}_{16}$	$\hat{\sigma}_1$	$\hat{\sigma}_4$	$\hat{\sigma}_{16}$	$\hat{\sigma}_{96}$
petites perturbations							
petit β_j , grand σ_j	1	0,011	0,035	0,011	0,078	0,053	0,008
moyen β_j , moyen σ_j	2	0,095	0,043	0,009	0,037	0,054	0,011
grand β_j , petit σ_j	3	0,075	0,112	0,009	0,037	0,028	0,014
moyennes perturbations							
petit β_j , grand σ_j	4	0,107	0,021	0,025	0,126	0,220	0,06
moyen β_j , moyen σ_j	5	0,12	0,063	0,029	0,95	0,125	0,032
grand β_j , petit σ_j	6	1	0,079	0,029	0,033	0,11	0,026
grandes perturbations							
petit β_j , grand σ_j	7	1	0,007	0,048	0,05	0,565	0,046
moyen β_j , moyen σ_j	8	1	0,018	0,048	0,046	0,334	0,064
grand β_j , petit σ_j	9	0,068	0,056	0,049	0,202	0,194	0,051
très grandes perturbations							
petit β_j , grand σ_j	10	0,163	0,005	0,104	0,299	1,132	0,178
moyen β_j , moyen σ_j	11	1	0,009	0,097	0,125	0,913	0,095
grand β_j , petit σ_j	12	1	0,013	0,105	0,111	0,772	0,123
géantes perturbations							
petit β_j , grand σ_j	13	0,002	0,001	0,453	4,626	6,635	0,738
moyen β_j , moyen σ_j	14	0,002	0,001	0,483	4,813	7,413	0,927
grand β_j , petit σ_j	15	1	0,001	0,708	0,378	6,678	0,949

TABLEAU 5.3 Estimations des paramètres des instances

Instance numéro	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Ancien gain(%)	-0,15	0,06	-0,04	0,69	0,86	0,88	0,73	1,11	1,55	2,30	3,45	3,74	5,82	8,48	11,57
Nouveau gain(%)	0,03	-0,25	-0,13	0,73	0,86	1,01	0,59	1,05	1,56	2,43	3,38	3,63	5,41	7,49	11,11

TABLEAU 5.4 Comparaison entre les anciens et les nouveaux gains

Les résultats sont très intéressants. Tout d'abord, nous pouvons remarquer que pour les trois instances constituant les petites perturbations, les gains sont comparables. De toute manière, pour les instances de petites variances (< 0.0004), le gain de l'optimisation est très petit et pas forcément intéressant ; des gains négatifs peuvent être observés.

Par contre, les autres résultats sont très encourageants. Tant que les paramètres σ_j n'explorent pas, l'ancien et le nouveau gains sont similaires et positifs. On peut même observer une amélioration du coût pour certaines instances. Ce phénomène s'explique

à nouveau par le fait que l'algorithme d'optimisation stochastique n'est pas exact. À partir de l'instance 13, il y a une perte de gain assez importante entre l'estimation et la réalité. Le nouveau gain reste heureusement toujours positif et améliore nettement la qualité de la solution.

Ces résultats poussent à conclure que la qualité des scénarios n'est pas si primordiale. Il y a certes une détérioration de la solution pour les instances qui ont une grande variance ($> 0,5$), mais les gains sont tout le temps positifs ou presque nuls. Néanmoins, le modèle a permis de réduire de 10 000 paramètres à 6 la modélisation de la demande. Il est sûrement faisable de réduire encore le nombre de paramètres, comme la qualité n'est pas si importante, mais des très bons résultats sont déjà obtenus ainsi. Le fait d'utiliser un algorithme stochastique permet donc d'obtenir des solutions qui sont beaucoup plus robustes et qui s'adaptent bien aux changements. Toutefois, il est difficile de juger de la qualité de la modélisation par perturbations, il faudrait tester le modèle sur des données réelles pour pouvoir vraiment conclure sur l'efficacité de cette méthode de génération de scénarios.

Chapitre 6

CONCLUSION

Ce mémoire a présenté un modèle de génération de scénarios à partir de données historiques. Les points les plus importants sont rappelés ci-dessous. S'il ne faut retenir qu'une seule chose de ce mémoire, c'est que l'optimisation stochastique rend les solutions plus robustes et cela en tenant peu compte de la qualité de l'échantillon des coefficients de variation.

6.1 Synthèse des travaux

Nos travaux peuvent être résumés en trois étapes importantes :

1. Nous avons développé un modèle par perturbations pour représenter l'incertitude de la demande prévisionnelle ;
2. Une méthode d'estimation a été proposée afin de retrouver les paramètres régissant le modèle ;
3. Nous avons étudié la qualité de l'estimation et son impact sur la solution.

La modélisation de la demande prévisionnelle est une méthode novatrice : nous n'avons pas trouvé d'approche similaire réalisée auparavant. Une telle approche permet de conserver tout le travail déjà accompli pour obtenir une demande prévisionnelle de qualité. Le modèle proposé est de plus très ajustable : la forme et la longueur des perturbations peuvent être modifiées, les paramètres β_j et σ_j permettent de décrire beaucoup de situations différentes. Nous sommes aussi passés d'une matrice de corrélation avec environ 10 000 paramètres à un modèle avec seulement 6 paramètres.

La méthode d'estimation est assez compliquée et comprend plusieurs étapes intermédiaires :

1. Estimation des variances ;
2. Estimation des mesures d'ordre quatre ;
3. Recherche des paramètres originaux par dichotomie.

À chaque étape, une erreur s'introduit dans l'estimation. Le résultat peut être par conséquent assez variable. Comme pour les perturbations non conditionnées, l'estimation se résume à l'étape 1 : elle est en général de meilleure qualité que celle des perturbations conditionnées. Il reste difficile de juger de la qualité de ses résultats sans calculer l'impact sur le coût du planning.

Le test du modèle avec des données historiques fabriquées donne de bons résultats. Les gains observés sont très satisfaisants et permettent en général peu de pertes : l'estimation donne des solutions équivalentes à la réalité dans la plupart des cas testés. Ces résultats sont très encourageants et donnent envie de tester le modèle sur des données réelles.

6.2 Limites de la méthode proposée

La principale limite du modèle est le fait que nous n'ayons pas pu le tester dans un cadre réel. Nous n'avons pas de données réelles qui nous permettent d'évaluer notre méthode d'estimation. Tant que cela ne sera pas fait, il est très difficile de pouvoir vraiment conclure quant à l'efficacité de l'estimation. Même si tous les résultats tendent à dire que les solutions obtenues sont meilleures avec un algorithme d'optimisation stochastique par scénarios, il nous faut attendre ces données réelles. Celles-ci sont difficiles à obtenir, car elles demandent beaucoup de travail pour être construites. Elles existent chez les entreprises, mais de manières brutes. Les estimations en nombres d'employés doivent être déduites du nombre de clients ou patients qui sont arrivés pendant la tranche de quinze minutes, le volume de leur commande ou la gravité de la maladie par exemple. Il est donc vraiment important d'obtenir ces données dans le futur.

6.3 Améliorations futures

Comme il a été dit dans la section précédente, il faut obtenir des données. Une fois celles-ci obtenues, nous pourrions réaliser beaucoup de tests différents :

- Refaire la même étude que celle qui a été présentée dans la partie résultat ;
- Analyser l'incertitude oubliée par le modèle ;
- Essayer de modifier la forme et la durée des perturbations.

Ce ne sont que quelques idées parmi d'autres qui sont présentées. Il faudra réaliser autant d'essais différents que possible afin de pousser au maximum l'analyse du modèle.

RÉFÉRENCES

- AYKIN, T. (1996). Optimal shift scheduling with multiple break windows. *Management Science*, 42, 591–602.
- BARD, J. F., MORTON, D. P. et WANG, Y. M. (2007). Workforce planning at USPS mail processing and distribution centers using stochastic optimization. *Ann. Oper. Res.*, 155, 51–78.
- BECHTOLD, S. et JACOBS, L. (1990). Implicit modeling of flexible break assignment in optimal shift scheduling. *Management Science*, 36, 1339–1351.
- BIRGE, J.-R. et LOUVEAUX, F. (1997). *Introduction to stochastic programming*. Springer.
- CARIO, M. C. et NELSON, B. L. (1997). Modeling and generating random vectors with arbitrary marginal distributions and correlation matrix. Rapport technique, Northwestern University.
- CROW, E. L. et SHIMIZU, K. (1988). *Lognormal Distributions : Theory and Application*. Dekker, New York.
- DANTZIG, G. (1954). A comment on Edie’s traffic delay at toll booths. *Operations Research*, 2, 339–341.
- DANTZIG, G. et INFANGER, G. (1992). Large-scale stochastic linear programs - importance sampling and benders decomposition. *Computational and applied mathematics*, 111–120.
- DUPAČOVÁ, J., CONSIGLI, G. et WALLACE, S. W. (2000). Scenarios for multistage stochastic programs. *Ann. Oper. Res.*, 25–53.
- DUPAČOVÁ, J., GRÖWE-KUSKA, N. et RÖMISH, W. (2003). Scenario reduction in stochastic programming. *Mathematical Programming*, 95, 493–511.
- ERNST, A., JIANG, H., KRISHNAMOORTHY, M. et SIER, D. (2004). Staff scheduling and rostering : A review of applications, methods and models. *European Journal of Operational Research*, 153, 3–27.
- GAMACHE, M., SOUMIS, F., MARQUIS, G. et DESROSIERS, J. (1999). A column generation approach for large scale aircrew rostering problems. *Operations Research*, 47, 247–263.

- HØYLAND, K., KAUT, M. et WALLACE, S. W. (2003). A heuristic for moment-matching scenario generation. *Computational Optimization and Applications*, 169–185.
- JOLLIFFE, L. T. (2002). *Principal Component Analysis*. Springer Series in Statistics. Springer, seconde édition.
- KAUT, M. et WALLACE, S. W. (2007). Evaluation of scenario-generation methods for stochastic programming. *Pacific Journal of Optimization*, 3, 257–271.
- LORETAN, M. (1997). Generating market risk scenarios using principal components analysis : methodological and practical considerations. *The Measurement of Aggregate Market Risk*, 23–60.
- LUKACS, E. (1970). *Characteristic Functions*. Griffin, London, seconde édition.
- PACQUEAU, R. (2011). *Optimisation stochastique d’horaires de personnel*. Mémoire de maîtrise, École Polytechnique de Montréal.
- PFLUG, G. C. (2001). Scenario tree generation for multiperiod financial optimization by optimal discretization. *Mathematical Programming*, 251–271.
- PLACKETT, R. (1983). Karl pearson and the chi-squared test. *International Statistical Review*, 51, 59–72.
- SAPORTA, G. (2006). *Probabilités, Analyse des données et Statistique*. TechNip.
- SMITH, J. E. (1993). Moment methods for decision analysis. *Management Science*, 340–358.
- VALE, C. D. et MAURELLI, V. A. (1983). Simulating multivariate nonnormal distributions. *Psychometrika*, 465–471.
- YEN, W. J. et BIRGE, J.-R. (2006). A stochastic programming approach to the airline crew scheduling problem. *Transportation Science*, 40, 3–14.
- ZHANG, X. et BARD, J. F. (2005). Equipment scheduling at mail processing and distribution centers. *IIE Trans Schedul Logist*, 37, 175–187.