



# Produits dérivés climatiques : aspects économétriques et financiers

Olivier Roustant

► **To cite this version:**

Olivier Roustant. Produits dérivés climatiques : aspects économétriques et financiers. Gestion et management. Université Claude Bernard - Lyon I, 2003. Français. <NNT : 2003LYO10111>. <tel-00804727>

**HAL Id: tel-00804727**

**<https://tel.archives-ouvertes.fr/tel-00804727>**

Submitted on 26 Mar 2013

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

N° d'ordre : 111 - 2003

**PRODUITS DÉRIVÉS CLIMATIQUES :  
ASPECTS ÉCONOMÉTRIQUES ET FINANCIERS**

**THÈSE**

présentée

devant **L'UNIVERSITÉ CLAUDE BERNARD – LYON I –**

pour l'obtention

du **DIPLOME DE DOCTORAT DE SCIENCES DE GESTION**

(arrêté du 25 avril 2002)

présentée et soutenue publiquement le

**lundi 7 juillet 2003**

par

**Olivier ROUSTANT**

**JURY :**

Jean-Paul LAURENT, Professeur à l'Université Claude Bernard Lyon I, Directeur de recherche

François DUFRESNE, Professeur à HEC Lausanne, Rapporteur

Olivier SCAILLET, Professeur à HEC Genève, Rapporteur

Laurent CARRARO, Professeur à l'Ecole Nationale Supérieure des Mines de Saint-Etienne

Jean-Claude AUGROS, Professeur à l'Université Claude Bernard – Lyon I

Daniel SERANT, Professeur à l'Université Claude Bernard – Lyon I



*Le domaine des produits dérivés climatiques est, par nature, à l'intersection de la finance, de l'assurance, des mathématiques appliquées, de la météorologie et de l'environnement... Je tiens par conséquent à exprimer ma profonde reconnaissance à Monsieur Jean-Paul LAURENT pour avoir eu l'audace d'accepter de lancer sur le sujet... un algébriste de formation ! Qu'il trouve ici l'expression de ma gratitude pour la confiance qu'il m'a accordée ainsi que pour ses conseils tout au long de la thèse.*

*Ce travail a été réalisé au sein du département 3MI de l'école des mines de Saint-Etienne, où d'importants moyens en statistiques ont été déployés, avec le concours de MM. Laurent CARRARO et Xavier BAY. Je les remercie vivement pour leur soutien.*

*Je tiens à remercier MM. François DUFRESNE et Olivier SCAILLET d'avoir bien voulu accepter d'être les rapporteurs de mes travaux. Je voudrais aussi remercier MM. Daniel SERANT et Jean-Claude AUGROS pour l'intérêt qu'ils ont porté à ce travail, les relectures qu'ils ont effectuées, et les discussions constructives dont ils m'ont fait profiter : leur participation à ce jury est pour moi un honneur.*

*Je remercie la société Météo-France, qui nous a fourni les données météorologiques, et la société Fininfo, qui nous a fourni les données boursières.*

*Mes remerciements vont également à ceux qui m'ont apporté leurs compétences pour les questions relatives au climat et à l'environnement. Merci à Monsieur Pierre VOIGNIER de la société Rhéa, qui m'a fait spontanément bénéficier d'informations rares et pertinentes ; merci à Monsieur Serge TABOULOT de Météo-France, et à Madame Mireille BATTON-HUBERT du centre SITE de l'école des mines, pour leur aide en météorologie ; merci à Madame Natacha GONDRAN pour les nombreuses discussions amicales sur les aspects environnementaux.*

*Merci à tous les membres du département 3MI, et donc Mmes Anca BADEA, Céline HELBERT, Laure MESSIAEN et Monsieur Eric TOUBOUL, pour l'agréable atmosphère de travail et l'ambiance cordiale qu'ils ont su faire régner autour de moi durant ces quelques années. Merci aussi à tous les membres de l'ISFA pour leur accueil chaleureux et constant malgré l'intermittence de mes visites !*

*Merci à Muriel et à Julie pour leur patience et leur soutien au quotidien.*

*Enfin, je voudrais remercier tout spécialement Madame Danièle TARRAL, sans qui cette thèse n'aurait probablement jamais débutée.*



A ma mère,  
pour sa merveilleuse dignité  
et sa présence affectueuse  
tout au long de ces années.

A mon père,  
disparu trop tôt.







## Table des matières.

<b>CHAPITRE 0. INTRODUCTION GENERALE.</b> .....	<b>1</b>
A. DES RISQUES CLIMATIQUES NON CATASTROPHIQUES AUX PRODUITS DERIVES CLIMATIQUES.....	3
I. <i>Les risques climatiques non catastrophiques.</i> .....	3
II. <i>Modalités de gestion : produits dérivés climatiques.</i> .....	5
B. QUELQUES EXEMPLES DE PRODUITS DERIVES CLIMATIQUES. ....	6
I. <i>Le cas « standard » d'un producteur d'énergie.</i> .....	6
II. <i>Un exemple de swap entre deux utilisateurs finaux.</i> .....	9
III. <i>Un exemple de contrat exotique : l'assurance de la floraison des cerisiers au Japon.</i> .....	10
C. LE CONTEXTE. ....	11
I. <i>Le développement des produits dérivés.</i> .....	11
II. <i>Le changement climatique et l'augmentation des risques climatiques.</i> .....	12
III. <i>La mondialisation et l'augmentation des besoins de couverture des risques climatiques.</i> .....	12
IV. <i>L'émergence de transferts de risques « alternatifs ».</i> .....	13
D. QUELQUES APPLICATIONS DES PRODUITS DERIVES CLIMATIQUES. ....	13
I. <i>Réduire la volatilité induite par le climat.</i> .....	13
II. <i>De nouveaux instruments de diversification.</i> .....	14
III. <i>Applications à l'environnement.</i> .....	14
E. LE MARCHE DES PRODUITS DERIVES CLIMATIQUES.....	15
I. <i>Les acteurs.</i> .....	16
II. <i>Etat du marché.</i> .....	18
III. <i>Quelques spécifications de contrats.</i> .....	22
F. PANORAMA DES PROBLEMES. ....	26
I. <i>Difficultés spécifiques.</i> .....	26
II. <i>Place des modèles statistiques.</i> .....	27
III. <i>Problèmes spécifiques.</i> .....	28
G. APPLICATION DES MODELES STATISTIQUES AUX PRODUITS DERIVES CLIMATIQUES. ....	30
I. <i>Modélisation statistique des aléas climatiques.</i> .....	30
II. <i>La pratique de l'évaluation.</i> .....	31
III. <i>Le risque de modèle.</i> .....	32
H. OBJECTIF ET ORGANISATION DE LA THESE. ....	34
REFERENCES .....	36
<b>CHAPITRE 1. MODELISATION DE LA TEMPERATURE.</b> .....	<b>41</b>
A. UNE APPLICATION DE DEUX MODELES ECONOMETRIQUES DE TEMPERATURE A LA GESTION DES RISQUES CLIMATIQUES. ....	43
I. <i>Statistique descriptive.</i> .....	47
II. <i>Ajustement d'un modèle linéaire ARMA.</i> .....	50
III. <i>Modèle linéaire à variance périodique.</i> .....	53
IV. <i>Performances des modèles. Application à la gestion des risques climatiques.</i> .....	64
V. <i>Réponse du modèle (VP) à un choc.</i> .....	72
VI. <i>Performances réelles.</i> .....	75
VII. <i>Conclusions de la section A.</i> .....	78
B. COMPLEMENTS.....	84
C. MODELISATION DE LA TEMPERATURE : ASPECT MULTIVARIE.....	97
I. <i>Quelques précisions sur le cadre de travail.</i> .....	99
II. <i>Etudes descriptives préliminaires.</i> .....	100
III. <i>Modélisation VAR.</i> .....	105
IV. <i>Comparaison des prévisions.</i> .....	111
V. <i>Conclusion de la section C.</i> .....	121
ANNEXE DU CHAPITRE 1, SECTION C .....	123
CONCLUSIONS DU CHAPITRE 1. ....	131
REFERENCES DU CHAPITRE 1. ....	133

<b>CHAPITRE 2. ASPECTS FINANCIERS.....</b>	<b>135</b>
A. L'ÉVALUATION DES PRODUITS DERIVES CLIMATIQUES.....	137
I. <i>Evaluation dans l'univers historique et bêta nul.</i> .....	137
II. <i>Indépendance entre marché et température.</i> .....	138
B. ÉVALUATION DES ACTIONS DES ENTREPRISES SENSIBLES AU CLIMAT. ....	147
I. <i>Résultats théoriques.</i> .....	147
II. <i>Etude empirique de la sensibilité des actions à la température.</i> .....	156
III. <i>Conclusions de la section B.</i> .....	175
CONCLUSIONS DU CHAPITRE 2. ....	177
REFERENCES DU CHAPITRE 2 .....	178
ANNEXE DU CHAPITRE 2 .....	179
<b>CHAPITRE 3. INCERTITUDES DE MODELISATION. REPERCUSSIONS SUR LES PRIX.....</b>	<b>181</b>
A. CALCUL DES PRIX.....	183
I. <i>Formules d'évaluation.</i> .....	183
II. <i>Méthodes de calcul.</i> .....	184
B. METHODOLOGIE.....	186
I. <i>Evaluation de l'erreur d'estimation.</i> .....	187
II. <i>Evaluation de l'incertitude des prix.</i> .....	188
C. IMPACT DES ERREURS DE MODELE SUR LA PRIME PURE D'UN CONTRAT FUTURE.....	190
I. <i>Cadre de travail.</i> .....	190
II. <i>Evaluation de l'impact de l'erreur d'estimation.</i> .....	191
III. <i>Evaluation de l'impact de mauvaises spécifications du modèle.</i> .....	193
IV. <i>Conclusions de la section C.</i> .....	200
D. CONSEQUENCES ET ORIGINES DE L'ERREUR D'ESTIMATION SUR LE PRIX DES PRODUITS DERIVES CLIMATIQUES. ....	202
I. <i>Cadre de travail. Notations.</i> .....	202
II. <i>Incertitude des prix avec la modélisation des indices de température.</i> .....	203
III. <i>Incertitude des prix avec la modélisation de la température.</i> .....	208
IV. <i>Conclusions de la section D.</i> .....	212
CONCLUSIONS DU CHAPITRE 3. ....	217
REFERENCES DU CHAPITRE 3. ....	218
ANNEXE DU CHAPITRE 3 .....	221
<b>CONCLUSION GENERALE. ....</b>	<b>229</b>
<b>BIBLIOGRAPHIE GENERALE.....</b>	<b>233</b>



## **Chapitre 0. Introduction générale.**

Apparus aux Etats-Unis en 1997, à la suite de la déréglementation du marché de l'électricité, les produits dérivés climatiques ont été introduits pour aider à se prémunir contre les risques climatiques non catastrophiques. Historiquement, ils ont permis d'apporter une protection aux producteurs d'électricité qui subissent des manques à gagner importants à chaque fois que l'hiver est doux. Cependant, il ne s'agit pas de produits d'assurance mais d'authentiques produits financiers, construits sur des « sous-jacents » d'un type particulier, qui sont des mesures de grandeurs climatiques.

La présente thèse rassemble les études de quelques aspects de ces nouveaux produits ; ses objectifs seront précisés à la fin du chapitre. Auparavant, au cours de quelques paragraphes introductifs, on se propose de familiariser le lecteur avec les produits dérivés climatiques et de répondre aux questions essentielles qu'il peut se poser à leur sujet. Nous ferons d'abord une description des risques climatiques sous-jacents et expliquerons en quoi les produits dérivés climatiques, qui sont de nature purement financière, permettent d'aider à leur gestion ; nous donnerons alors quelques exemples de transactions réalisées avec les produits dérivés climatiques ; nous présenterons ensuite le contexte, économique et géographique, qui explique pour une part leur apparition ; nous donnerons quelques applications, et observerons que leur champ va au-delà des objectifs initiaux de conception ; nous procéderons à un état des lieux des cinq premières années d'existence du marché des produits dérivés climatiques ; nous ferons ensuite un panorama des problèmes de gestion et remarquerons la place particulière allouée aux modèles statistiques dans leur résolution ; un chapitre est alors consacré à ces modèles, leur utilisation et les risques qu'elle génère. Les principaux travaux réalisés sur les produits dérivés climatiques sont présentés au cours de ces deux derniers paragraphes.

## **A. Des risques climatiques non catastrophiques aux produits dérivés climatiques.**

### **I. Les risques climatiques non catastrophiques.**

#### **1. Quelques exemples.**

Les risques climatiques non catastrophiques sont des risques économiques que supportent les entreprises dont l'activité est sensible au climat : l'impact du climat sur la consommation est alors la source de manques à gagner pouvant être importants. On entend souvent les professionnels du tourisme se plaindre, à juste titre, des « mauvaises saisons », responsables d'une diminution de l'affluence des vacanciers, mais d'autres secteurs sont également concernés comme ceux de l'énergie, de l'agroalimentaire, de l'agriculture... Par exemple, les producteurs d'électricité sont touchés par le niveau inhabituellement faible de la demande des usagers, en chauffage lorsque les hivers sont cléments, et en climatisation lorsque les étés sont doux ; les distributeurs de sodas, de sirops et autres boissons rafraîchissantes sont pénalisés par la fraîcheur relative des températures estivales, à l'origine d'une diminution de la satiété chez leurs clients ; les agriculteurs sont également sensibles à ces risques qui viennent s'ajouter aux autres types de risques de nature catastrophique, comme le gel ou la grêle, car le rendement et la qualité des récoltes dépendent des bonnes conditions climatiques durant tout le cycle de maturation.

#### **2. Analyse des risques climatiques non catastrophiques.**

Ces risques climatiques se distinguent des risques climatiques catastrophiques par plusieurs aspects.

Tout d'abord, alors que les catastrophes naturelles peuvent porter atteinte à la vie humaine, les risques étudiés ont des conséquences purement financières ; celles-ci peuvent se répercuter au plan humain, comme par des licenciements, mais la gravité des impacts est assurément moindre.

Le risque est souvent volumétrique car les impacts sont fonction de la diminution du volume du chiffre d'affaires. Mais celle-ci peut également être couplée avec une diminution du prix de vente, imposé par la médiocrité de la demande.

D'autre part, il n'y a pas ici de sinistres physiquement constatables : pas de dégâts matériels, pas de dommages visibles. Les conséquences s'observent dans les comptes des entreprises, ce qui les rend plus difficiles à évaluer.

La localisation en temps et en espace est aussi très différente. Cette différence s'apprécie en termes de concentration. Les événements catastrophiques sont toujours concentrés dans le temps : leur durée est courte, allant de quelques secondes pour un tremblement de terre à quelques heures pour une tempête ou un cyclone ; en revanche, les manques à gagner dus à de mauvaises conditions climatiques se font ressentir sur des périodes plus longues, de l'ordre de quelques mois. Les événements catastrophiques sont également souvent concentrés dans l'espace : la zone sinistrée est restreinte, de la taille d'un champ pour un orage de grêle à la surface d'une région pour un tremblement de terre ou une tempête ; au contraire, les mauvaises conditions climatiques responsables d'une baisse de l'affluence ou de la consommation touchent généralement un territoire plus étendu, à l'échelle d'un pays.

Enfin des possibilités d'échange existent entre des agents économiques exposés à des risques symétriques. Par exemple, le gérant d'un parc d'attractions et le gérant d'un complexe de cinémas ont des intérêts opposés à ce qu'il pleuve ; de même, si la clémence d'un hiver peut défavoriser un producteur d'électricité, il est certain qu'elle avantage les « gros » consommateurs, comme les hôpitaux, les PME<sup>1</sup>...

### **3. Mesure des risques. Evaluation des impacts.**

Certains risques peuvent se mesurer avant qu'ils ne se matérialisent, comme le risque de change, que l'on peut suivre en observant le mouvement des taux de change. Cette possibilité existe pour les risques climatiques non catastrophiques. On peut en effet souvent relier la qualité d'une saison à des données météorologiques, ce qui suppose en réalité deux étapes : l'identification d'un ou plusieurs facteurs climatiques influents, et le choix d'un instrument de mesure approprié.

Dans le cas de l'électricité, c'est la température, essentiellement, qui dirige la consommation. Il existait d'ailleurs, avant que les produits climatiques ne voient le jour, des indices journaliers fonction de la température, appelés degrés-jours, pour mesurer le risque climatique. Ceux-ci mesurent l'écart de la température à un niveau seuil correspondant peu ou

---

<sup>1</sup> Pour les grandes entreprises, l'implantation à l'étranger atténue la sensibilité au climat.

prou à la température ambiante à laquelle la consommation est minimale ; l'hiver par exemple, l'indice degré-jour (en anglais : *degree-day*) est défini par :

$$DD = (18 - T_t)^+ = \max(18 - T_t; 0)$$

où  $T_t$  est la température à la date  $t$ , exprimée en degrés. Une mesure du risque de température sur une saison peut alors être effectuée au moyen du cumul de ces indices :

$$DD_{\text{saison}} = \sum_{\text{saison}} (18 - T_t)^+$$

Les instruments de mesure des risques climatiques étudiés varient : hauteur de pluie, hauteur de neige, vitesse du vent ... et dépendent quelquefois de plusieurs aléas climatiques. A l'heure actuelle, c'est la température qui est majoritairement employée, mais parfois à l'aide d'indices différents de celui montré ci-dessus.

Les impacts, on l'a vu, sont difficiles à chiffrer directement car les dommages ne sont pas facilement observables. Une solution existe cependant, consistant à rechercher une fonction reliant la mesure du risque aux conséquences qu'il engendre : ce sont les *fonctions d'endommagement*, bien connues des producteurs d'électricité. Naturellement, l'inconvénient majeur est l'asymétrie d'une telle mesure, qui ne peut être faite - en réalité - que par celui qui en subit les effets ! Et ceci nous amène à étudier les modalités de gestion des risques climatiques non catastrophiques.

## II. Modalités de gestion : produits dérivés climatiques.

### 1. L'impossibilité d'une gestion de type assurance.

Le premier constat à la suite de la présentation précédente est l'impossibilité d'une gestion de type assurance des risques climatiques non catastrophiques. Il y a d'abord l'importance du volume du risque : l'épisode du cyclone Andrew, qui a déseparé jusqu'aux réassureurs, rappelle l'importance de ce facteur. D'autre part, une couverture indemnitaire est ici quasi-impossible : à la difficulté objective de mesurer les impacts économiques du climat sur le chiffre d'affaires des entreprises, s'ajoute un problème d'asymétrie puisque, même si les mesures de variables climatiques sont publiques, seul l'assuré peut connaître avec précision l'impact de ces variables sur les résultats. Par ailleurs, ce sont les entreprises les plus exposées aux risques climatiques qui veulent se prémunir. On peut craindre alors qu'un phénomène



d'antisélection conduise à les exclure du marché en raison de la concurrence... Est-il possible en effet, d'imaginer un marché d'assurances où, de façon similaire à ce qui se passe sur le marché automobile, les « mauvais conducteurs » sont les principaux clients ? Enfin l'aléa moral serait également présent : comment pourrait-on contraindre les conducteurs à se soucier de leur conduite, autrement dit, comment pourrait-on contraindre les entreprises concernées à procéder en interne à une gestion de leurs risques climatiques ?

### **2. D'une assurance climatique aux produits dérivés climatiques.**

On observe dans ce qui précède qu'une difficulté majeure est la quasi-impossibilité de chiffrer les sinistres consécutifs aux mauvaises conditions climatiques. Une solution consiste alors... à laisser ce problème aux clients ! Autrement dit, on calcule les versements en fonction de la mesure des risques climatiques et non pas selon le montant des pertes financières dues au climat. On passe donc d'une couverture indemnitaire à une couverture paramétrique, et la mesure des impacts est laissée aux entreprises.

Concrètement, il s'agit de produits financiers construits à partir d'indices climatiques d'où leur nom de produits dérivés climatiques. Ils se présentent sous la forme de futures (ou « swaps »), d'options, collars, etc. Nous en donnerons quelques exemples dans la prochaine section.

### **3. Les avantages d'une gestion financière.**

La gestion financière des risques climatiques non catastrophiques constitue une alternative viable à la gestion de type assurance. Le fait d'englober ces produits dans le marché financier permet en particulier de profiter de sa capacité d'absorption (il n'y a donc plus de problème de volume) et de sa souplesse. En outre, les risques climatiques sont relativement décorrélés du marché financier, ce qui attire les spéculateurs et les investisseurs et devrait favoriser à terme la liquidité du marché des produits dérivés climatiques.

## **B. Quelques exemples de produits dérivés climatiques.**

### **I. Le cas « standard » d'un producteur d'énergie.**

Parmi les transactions de produits dérivés climatiques, le cas le plus fréquent est celui d'un producteur d'énergie souhaitant se protéger contre un hiver doux. Il va sans dire que la douceur de l'hiver, en faisant chuter la consommation, affecte les résultats de l'industriel. La

sensibilité du résultat au climat est assez bien mesurée par un indice à base des degrés-jours introduits au paragraphe 3 de la section précédente ; dans cette situation, on utilisera l'indice HDD<sup>2</sup>, égal à la somme sur la période hivernale<sup>3</sup> des écarts de température relativement à la température de confort de 18°C :

$$HDD = \sum_{\text{période hivernale}} (18 - T_t)^+$$

où  $T_t$  est la température moyenne journalière à la date  $t$ .

A titre d'illustration, on a représenté sur la Figure 0.1 une approximation de la fonction d'endommagement d'un producteur de gaz basé à Paris<sup>4</sup> ; la température de référence est celle d'Orly et la période considérée s'étale du 15/11 au 15/3.

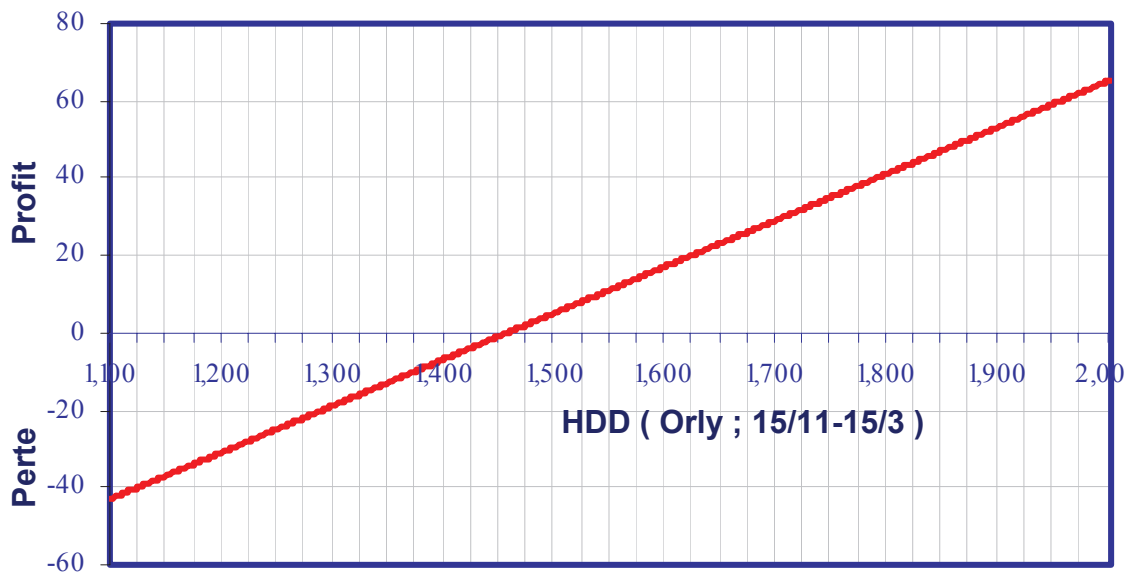


Figure 0.1. Le cas d'un producteur de gaz : sensibilité du résultat au climat.

Vu la définition de l'indice HDD, on observe qu'un hiver doux correspond à de faibles valeurs de l'indice. Pour se protéger, le producteur de gaz peut donc prendre une position courte dans un contrat future sur l'indice HDD ou une position longue dans une option de vente sur cet indice ; c'est cette dernière possibilité qui a été retenue par l'industriel. Les caractéristiques précises du contrat sont présentés dans le Tableau 0.1 . Comme une option de

<sup>2</sup> Les indices HDD et CDD sont présentés en détail au paragraphe III de la section E.

<sup>3</sup> Celle-ci dépend des contrats ; elle s'étale souvent d'octobre à avril inclus (c'est le cas au CME, voir § E.III.1) ou bien de novembre à mars. Dans l'exemple ci-dessus, la période hivernale est encore plus courte, commençant le 15 novembre et se terminant le 15 mars.

<sup>4</sup> D'après (Deboaisne, 2000).

vente standard, l'industriel en position d'acheteur va donc exercer l'option si l'indice HDD du prochain hiver est inférieur à 1500 (°C), réalisant alors un gain de  $(1500 - HDD) \times 130.000$  FF ; comme c'est souvent le cas, une valeur supérieure de paiement (*cap*) a été fixée, limitant le gain à 40.000.000 FF. Le profil de paiement est représenté sur la Figure 0.2 . Comme pour une option standard, l'industriel est protégé par rapport à de faibles valeurs du sous-jacent (voir Figure 0.3), c'est-à-dire ici par rapport à des températures hivernales douces.

☐ <b>Type</b>	: Put Option HDD
☐ <b>Acheteur du Put</b>	: Producteur de Gaz
☐ <b>Vendeur du Put</b>	: BNP Paribas
☐ <b>Période de Référence</b>	: Du 15 novembre 2000 au 15 mars 2001
☐ <b>Lieu de Référence</b>	: Station Météo France de Paris Orly
☐ <b>Indice Climatique</b>	: HDD cumulés
☐ <b>Base de Calcul</b>	: Données fournies par Météo France
☐ <b>Montant Notionnel</b>	: 40.000.000 FRF
☐ <b>Montant de la Prime</b>	: ● francs
☐ <b>Prix d 'Exercice</b>	: 1500 HDD
☐ <b>Valeur du Tick</b>	: 130.000 FRF par HDD

Tableau 0.1. Le cas d'un producteur de gaz : termes du produit dérivé climatique.

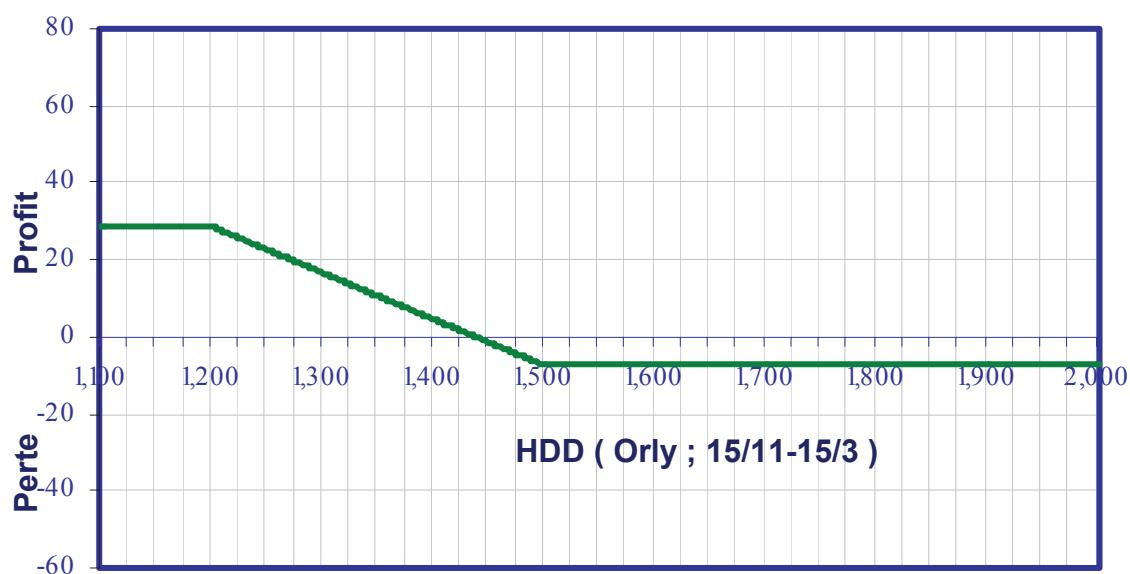


Figure 0.2. Le cas d'un producteur de gaz : profil de paiement du produit dérivé climatique.

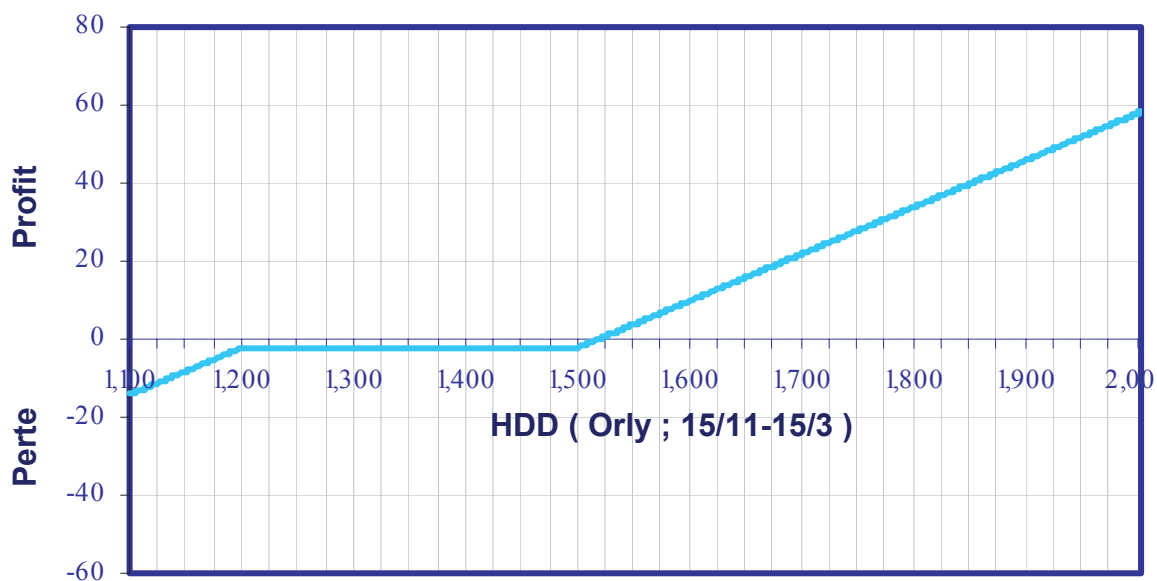


Figure 0.3. Le cas d'un producteur d'énergie : sensibilité du résultat au climat après couverture.

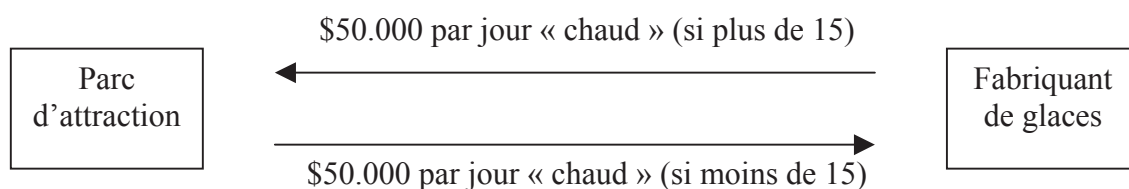
## II. Un exemple de swap entre deux utilisateurs finaux<sup>5</sup>.

Il est d'usage d'appeler « swap » un contrat entre deux parties pour lequel aucun versement initial n'a lieu : un contrat future est donc une forme de swap, et c'est d'ailleurs son appellation courante dans le monde des dérivés climatiques. Pourtant, un swap au sens fort sous-entend un échange entre deux acteurs concernés par un risque commun et non pas seulement entre un acteur et une institution financière (voir par exemple les swaps de taux d'intérêt, réalisant l'échange entre un taux variable et un taux fixe), et c'est ce type de swap dont nous voulons donner un exemple.

La transaction a été réalisée par Worldwide Weather Trading Co entre un parc d'attraction et un fabricant de glaces situés au même endroit. Les deux parties ont des risques symétriques par rapport à la température estivale : tandis que les températures caniculaires avantagent le fabricant de glaces en faisant grimper la consommation de glaces, elles nuisent au parc d'attraction en provoquant la baisse de la fréquentation. L'impact du climat a été évalué de part et d'autre, et les deux parties sont tombées d'accord pour réaliser un swap sur le nombre de jours « chauds », définis par les jours où la température excède 96°F. De façon précise, le parc d'attraction accepte la transaction lorsque l'été comprend moins de 15 jours chauds et verse alors la somme de 50.000\$/jour chaud au fabricant de glaces ; inversement, le fabricant

<sup>5</sup> En raison du manque de liquidité et de moyens de couverture avérés, les swaps s'effectuent la plupart du temps entre des utilisateurs finaux (voir § E.I.) – L'exemple est tiré de (Barrieu, 2000).

de glaces s'engage à verser 50.000\$/jour chaud à la condition qu'il y ait eu plus de 15 jours chauds dans l'été. D'autre part, il a été convenu que le montant total échangeable serait de 1 million de \$. Il s'agit donc d'un swap dont le sous-jacent est un indice conditionnel de température (voir Figure 0.4).



**Figure 0.4. Un swap de température entre un parc d'attraction et un fabricant de glaces.**

### **III. Un exemple de contrat exotique : l'assurance de la floraison des cerisiers au Japon<sup>6</sup>.**

Au Japon, la floraison des cerisiers est un événement d'une grande importance, qui amène les habitants à sortir dans les parcs et les jardins pour pique-niquer sous les frondaisons étoilées de petites fleurs roses. Quelquefois pourtant, lorsque la température est trop clémente pendant la période de maturation des arbres, la floraison se produit trop tôt, à une période où il fait encore trop frais pour en profiter, et cette précocité est alors préjudiciable aux petites et moyennes entreprises de tourisme et de loisir qui en tirent bénéfice.

L'idée est donc venue au premier assureur dommages du Japon, Tokio Marine and Fire Insurance, de proposer une protection contre ce risque. Les contrats prennent la forme d'une option à saut (*gap call option*), la compensation devenant effective lorsque la température moyenne journalière dépasse 7.5°C tous les jours d'une longue période, le versement étant alors proportionnel au nombre de jours « chauds » (voir la Figure 0.5). Les caractéristiques et la valeur des contrats dépendent des clients, et à titre d'exemple, un restaurant peut acheter un contrat pour un prix de 1.5 million de yens (11.960 €) correspondant à une valeur d'indemnisation de 450.000 yens (3.590 €) par jour.

---

<sup>6</sup> D'après une dépêche AFP datée du 17/01/03.

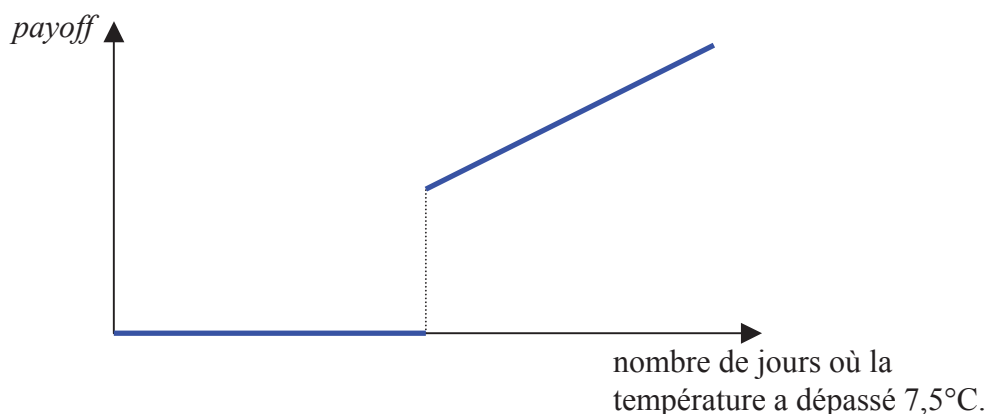


Figure 0.5. Une option à saut (*gap call option*) climatique pour la floraison des cerisiers au Japon.

### C. Le contexte.

Les produits dérivés climatiques ont vu le jour dans un contexte particulier, à l'intersection de deux changements majeurs, climatique pour le premier avec le réchauffement de la planète, économique pour le second avec la mondialisation. Ce contexte justifie pour une part leur création, rendue possible grâce aux moyens modernes de gestion. Par ailleurs, leur marché s'inscrit dans la nouvelle vague des transferts de risques alternatifs, et se place à côté des nouveaux marchés des produits dérivés d'assurance et des produits dérivés de l'électricité.

#### I. Le développement des produits dérivés.

Précisons tout de suite qu'il était difficile, de toute façon, de procéder à la gestion des risques climatiques non catastrophiques auparavant. Ces risques ont pourtant toujours existé, dès les premiers temps de l'agriculture. Mais pouvait-on faire autrement pour s'en prémunir que de constituer des réserves lors des années grasses ? Nous avons vu en effet dans la section précédente que la gestion externe des risques climatiques non catastrophiques ne peut pas se faire sous la forme d'une assurance mais sous la forme de produits dérivés. Or la gestion des produits dérivés a longtemps été délicate, et ce n'est véritablement qu'après la découverte de la formule de valorisation des options standard en 1973 par Black et Scholes que ce marché a pris de l'importance<sup>7</sup>. La gestion des risques climatiques ne pouvait donc guère se faire plus tôt. A l'heure actuelle en revanche, le développement de la finance quantitative et de l'informatique donne plus de moyens aux ingénieurs de gestion.

<sup>7</sup> La taille du marché global des produits dérivés est estimée aujourd'hui à cent mille milliards de dollars.

Mais les produits dérivés climatiques n'auraient peut-être pas vu le jour si dans le même temps, les besoins de couverture des risques climatiques et les risques eux-mêmes n'avaient pas augmenté.

## **II. Le changement climatique et l'augmentation des risques climatiques.**

Le réchauffement maintenant avéré de la planète n'a pas toujours toutes les conséquences qu'on lui attribue. Ainsi, dans l'accroissement vertigineux de l'impact économique des catastrophes naturelles observé au cours de la dernière décennie<sup>8</sup>, une part est sans doute due à la progression de l'urbanisation. Notons au passage que certains événements ont même dépassé la capacité de couverture des assureurs<sup>9</sup>. Toutefois, l'augmentation concomitante du nombre de catastrophes<sup>10</sup> renforce le sentiment que l'on peut nourrir au quotidien que le changement climatique se traduit par une augmentation des risques climatiques. Comme ce changement va se poursuivre dans les prochaines années, nous pouvons donc nous préparer à ce que cela dure.

## **III. La mondialisation et l'augmentation des besoins de couverture des risques climatiques.**

D'une façon générale, la mondialisation se traduit à l'heure actuelle par une intensification de la concurrence. Celle-ci a certainement contribué à la modification de la gestion interne des entreprises, avec un fonctionnement avec moins de fonds propres et une externalisation des services. Il n'est pas difficile alors de comprendre que les entreprises souhaitent désormais réduire au maximum la volatilité de leurs résultats, et donc en particulier la volatilité induite par le climat. En outre, dans le secteur de l'énergie, l'intensification de la concurrence a entraîné une vague de déréglementations, partie des Etats-Unis en 1997 et se déroulant à présent en Europe<sup>11</sup> : de grosses entreprises sensibles aux risques climatiques se sont donc retrouvées exposées aux aléas de l'économie. Cette conjugaison de facteurs a donc engendré un accroissement net des besoins en couverture de risques climatiques non catastrophiques à travers le monde.

---

<sup>8</sup> cf. le rapport de l'ONU pour les catastrophes naturelles : facteur 6.7 des années 60 aux dix dernières années en valeur.

<sup>9</sup> Ouragan « Andrew » en 1992 aux Etats-Unis.

<sup>10</sup> Dans le même rapport de l'ONU (cf. supra), on indique un facteur 4 des années 60 aux dix dernières années pour chiffrer l'augmentation du nombre de catastrophes naturelles.

<sup>11</sup> Le 25 novembre 2002, l'union européenne a voté à l'unanimité la déréglementation du marché de l'électricité ; celle-ci deviendra opérationnelle en 2004 pour les particuliers, en 2007 pour les entreprises.

#### **IV. L'émergence de transferts de risques « alternatifs ».**

C'est donc dans ce contexte d'accroissement des risques climatiques et économiques que sont apparus en 1997 les produits dérivés climatiques, pour répondre aux besoins de gestion des risques climatiques non catastrophiques ; c'est, pour l'essentiel, le formidable développement des produits dérivés, qui a rendu possible leur création. On conçoit alors aisément que d'autres marchés aient pu voir le jour, résultant à la fois de l'augmentation ou de l'apparition de risques, et de la possibilité de pouvoir les gérer sur un marché mondial en pleine expansion. Ainsi, deux autres marchés avaient vu le jour quelques années plus tôt<sup>12</sup> :

- les produits dérivés d'assurance (*insurance derivatives*)<sup>13</sup>, pour la gestion de produits d'assurance, limités pour l'heure aux catastrophes naturelles, et lancés par le Chicago Boarding Of Trade (CBOT) en 1993 ;
- et les produits dérivés de l'électricité (*power derivatives*), pour la gestion du prix de l'électricité, avec le lancement du 1<sup>er</sup> contrat future par la place norvégienne Nord Pool en 1995 et la 1<sup>ère</sup> option par le New York Mercantile Exchange (NYMEX) en 1996.

L'émergence de ces nouveaux marchés est un phénomène nouveau et remarquable car ils concernent des risques qui, justement, ne sont pas des risques de marché, mais qui au contraire sont relativement décorrélés des fluctuations de la bourse : en ce sens, on parle de transferts de risques alternatifs.

### **D. Quelques applications des produits dérivés climatiques.**

#### **I. Réduire la volatilité induite par le climat.**

Le but premier des produits dérivés climatiques est de pouvoir réduire la volatilité du chiffre d'affaires relativement aux risques climatiques non catastrophiques. Aujourd'hui encore, c'est la principale utilisation qui est faite de ces produits. En outre, la dégradation des conditions climatiques observée au cours des dernières décennies – et susceptible de s'aggraver dans le futur, pourrait conduire les agents économiques à recourir davantage aux produits dérivés climatiques ; ainsi dans un rapport récent<sup>14</sup>, l'ONU préconise leur emploi pour répondre à

---

<sup>12</sup> Voir (Geman, 1999), (Geman, 2000).

<sup>13</sup> Voir aussi (Geman, 1994) et (Scherer, 2000).

<sup>14</sup> Voir (ONU, 2002).



l'augmentation envisagée des pertes économiques liées au climat dans le secteur agricole. Cependant, d'autres applications ou perspectives d'applications existent.

## **II. De nouveaux instruments de diversification.**

Tout d'abord, les produits dérivés climatiques peuvent être employés comme outil de diversification dans la gestion de portefeuilles. Comme les aléas du climat sont en principe largement décorrélés des fluctuations de la bourse, c'est à juste titre qu'un investisseur peut compter sur une diminution du risque en plaçant une partie de son capital dans des produits dérivés climatiques. Les institutions bancaires l'ont bien compris et proposent maintenant des « fonds climatiques », constitués par une panoplie d'options, de swaps, d'obligations climatiques et d'obligations catastrophes.

## **III. Applications à l'environnement.**

Le renforcement des politiques environnementales, matérialisé par le protocole de Kyoto (1997), peut également suggérer des applications dans le domaine de l'environnement. Weinstein (2001), par exemple, montre que les produits dérivés climatiques sont un outil potentiel pour la gestion des permis d'émission et des certificats verts.

### **1. Des outils pour la gestion des permis d'émission.**

Les permis d'émission visent à réduire l'émission des gaz à effet de serre (CO<sub>2</sub>, CH<sub>4</sub> et autres composés organiques à base de carbone) par les industriels. Rappelons brièvement leur principe<sup>15</sup>. Chaque industriel dispose d'un quota fixé par le législateur indiquant la quantité maximale de gaz à effet de serre qu'il peut dégager dans l'atmosphère ; l'écart à ce quota est alors comptabilisé au crédit ou au débit de l'industriel. Un marché est alors possible : les industriels qui se trouvent dans l'incapacité de respecter leur quota peuvent acheter les permis d'émission dont ils ont besoin à ceux qui ne l'ont pas atteint ; les échanges de crédit sont incités par la menace de pénalités – voire d'interdictions - qui, en théorie, sont applicables aux industriels qui dépassent leur quota sans acheter de crédits à d'autres industriels. Or, dans le cas des producteurs d'énergie, on observe que le dégagement des gaz à effet de serre dépend fortement des conditions climatiques puisque celles-ci régissent l'activité industrielle. On peut donc imaginer des produits dérivés climatiques d'un nouveau type, permettant l'échange des risques climatiques contre les permis d'émission. Par exemple, un tel swap, conclu entre un producteur d'énergie et une banque, permettrait à l'industriel de recevoir les permis

---

<sup>15</sup> Pour plus de précisions, voir par exemple (Gondran, 2001) ou (Barrieu, 2002).

d'émission dont il a besoin en cas de bonne saison et le contraindrait à en céder en cas de mauvaise saison. Remarquons que les contrats de ce type ont l'avantage de ne donner lieu à aucune transaction financière.

## **2. Des outils pour la gestion des certificats verts.**

L'application des produits dérivés climatiques peut également s'envisager dans le cadre de la gestion des certificats verts<sup>16</sup>. Ce système de certificats concerne l'électricité « verte », c'est-à-dire celle produite à partir de sources renouvelables telles que l'éolien, la géothermie... Il a été mis en place à la suite de la directive européenne datant de septembre 2001, imposant aux pays des pourcentages minimums de consommation en électricité verte. S'agissant de consommation, et non de production, il est apparu nécessaire de trouver un système permettant de suivre à la trace l'électricité verte. De là les certificats verts. A la base, les certificats verts, associés à une quantité d'énergie, sont délivrés aux producteurs d'électricité verte. Ces derniers peuvent ensuite vendre leurs certificats soit directement auprès de clients finaux, dans le cadre de contrats classiques d'approvisionnement, soit par l'intermédiaire de distributeurs d'énergie.

Les produits dérivés climatiques peuvent s'avérer utiles pour faciliter la gestion des certificats. Par exemple, on peut imaginer un contrat entre un exploitant de fermes éoliennes et un institut financier échangeant des risques climatiques contre des certificats verts. En cas de conditions climatiques défavorables, l'exploitant se verrait recevoir des certificats verts qu'il n'aurait pu obtenir par sa seule production, et se verrait contraint d'en céder à l'institution financière en cas de bonnes conditions de vent<sup>17</sup>.

## **E. Le marché des produits dérivés climatiques.**

Initié en 1997 aux Etats-Unis par un contrat signé entre le groupe Enron et de petits producteurs d'énergie suite à la déréglementation du marché de l'énergie<sup>18</sup>, le marché des produits dérivés climatiques atteint aujourd'hui un volume cumulé de 11,8 milliards de dollars. Deux marchés organisés ont été créés : aux Etats-Unis, sur la place boursière du Chicago Mercantile Exchange (CME, septembre 1999), et en Europe, au London International

---

<sup>16</sup> Le système des certificats verts que nous allons présenter maintenant est décrit en détail dans (Martin, 2002).

<sup>17</sup> L'institution financière permet ainsi la mise en relation entre producteurs et utilisateurs (finaux ou intermédiaires) - Exemple tiré de Weinstein (2001).

<sup>18</sup> Voir (Barrieu, 2000).

Financial Futures and Options Exchange (Euronext-LIFFE, fin 2001). Toutefois la liquidité y est faible et la majorité des transactions s'effectue de gré à gré. Le marché est en expansion depuis sa création ; cependant, la faillite du même Enron - l'un des plus importants courtiers - en Janvier 2002, a jeté un froid ; signalons également le retrait de BNP Paribas à la même date.

Nous allons maintenant présenter le marché plus en détail, en commençant par faire un tour d'horizon des différents acteurs, puis en examinant ses caractéristiques : taille, localisation, nature des contrats..., et leur évolution. Nous précisons ensuite quelques spécifications de contrats ; ceux-ci proviennent uniquement des marchés organisés, en raison d'une part de la grande variété des transactions de gré à gré – ce qui anéantirait toute tentative de présentation exhaustive, et d'autre part de la difficulté d'accès aux informations publiques sur ces transactions. En réalité comme on le verra, plus de 80% des contrats sont du type de ceux que l'on trouve sur les marchés organisés.

## **I. Les acteurs.**

Une première catégorie d'acteurs est naturellement constituée des entreprises qui ont une demande en couverture pour les risques climatiques auxquelles elles sont exposées. En tête, semblent toujours figurer les producteurs d'énergie<sup>19</sup>, initiateurs du marché, mais d'autres secteurs sont maintenant représentés comme l'agroalimentaire, les loisirs, le tourisme... Notons que certaines entreprises susceptibles d'être concernées se montrent frileuses à entrer dans le marché, semble-t-il en raison des coûts quelquefois « dissuasifs » des transactions ; ce serait le cas, notamment, des entreprises agricoles<sup>20</sup>.

Les interlocuteurs tout désignés des entreprises exposées au climat sont les organismes bancaires et d'assurances qui forment une deuxième catégorie d'acteurs. Leur rôle, à l'heure actuelle, semble consister la plupart du temps à mettre en relation des utilisateurs finaux plutôt qu'à prendre une partie du risque ; le manque de liquidité d'une part, et l'absence de moyens de couverture avérés peuvent sans doute expliquer cet état de fait. Outre les contrats personnalisés, on trouve souvent, parmi les produits proposés, des *hedge funds* ou « fonds

---

<sup>19</sup> On dispose de peu d'informations à ce sujet ; cependant, en 2000, 90% des transactions concernaient les producteurs d'énergie (d'après Barrieu, 2000).

<sup>20</sup> Voir (Trevor, Linsley, 2002).

climatiques », regroupant un ensemble de produits dérivés attachés à des risques climatiques<sup>21</sup> (obligations catastrophes, produits dérivés climatiques...). Les souscripteurs de ces fonds endossent alors une pluralité de risques climatiques.

Parmi ces souscripteurs, on rencontre une troisième catégorie d'acteurs constituée des investisseurs ou des spéculateurs attirés par la perspective de gains importants sur un marché décorrélé des fluctuations boursières.

A titre indicatif, on pourra consulter sur le Tableau 0.2 une liste de vingt acteurs principaux, utilisés comme référence pour une étude faisant l'état des lieux du marché des produits dérivés climatiques, et dont nous présentons maintenant les résultats.

1.	Accord Energy Limited (Centrica Subsidiary)
2.	Aquila Energy
3.	AXA Corporate Solutions
4.	Bank of Tokyo-Mitsubishi Ltd.
5.	BNP Paribas
6.	Cargill Incorporated
7.	El Paso Merchant Energy
8.	Element Re Capital Products Inc.
9.	Enron
10.	Entergy-Koch (formerly Axia Energy)
11.	Hess Energy Trading Company
12.	Mirant Americas Energy Marketing LP
13.	Mitsui Sumitomo Insurance Company, Ltd. (formerly Mitsui Marine and Fire Insurance Company Ltd.)
14.	Mizuho Corporate Bank (formerly Industrial Bank of Japan)
15.	Reliant Energy Services, Inc.
16.	Societe Generale
17.	Swiss Re New Markets
18.	The Tokio Marine and Fire Insurance Co. Ltd.
19.	Tokyo Electric Power Company, Inc.
20.	TXU Europe Energy Trading Ltd.

**Tableau 0.2. Quelques intervenants du marché des produits dérivés climatiques (avril 2001 à mars 2002).**

<sup>21</sup> La Société Générale (SG) ou la Caisse des Dépôts et Consignations (CDC), par exemple, proposent ce type d'investissement.

## **II. Etat du marché.**

Nous donnons ici quelques chiffres essentiels du marché des produits dérivés climatiques. Ils proviennent d'une étude réalisée par PriceWaterhouseCoopers auprès d'une vingtaine d'intervenants majeurs<sup>22</sup> pour la période d'avril 2001 à mars 2002. L'étude mentionne également les résultats des années précédentes, ce qui permet de suivre l'évolution du marché depuis son origine. Les résultats sont présentés sur la Figure 0.6, la Figure 0.7 et la Figure 0.8.

### **1. Taille – Répartition géographique.**

On observe globalement une expansion du marché avec un doublement de la valeur des transactions depuis la 1<sup>ère</sup> étude de 1998. L'expansion est également géographique avec une récente, mais fulgurante progression du marché en Europe et en Asie<sup>23</sup> ; cependant les Etats-Unis dominant encore nettement le marché avec environ 80% des échanges.

### **2. Répartition des transactions selon la nature des risques.**

Des conclusions analogues peuvent être tirées à propos de la nature des risques climatiques échangés : la température reste le risque majoritaire mais les autres risques – la pluie, en particulier, sont de plus en plus représentés ; cette diversification se remarque aussi « à l'intérieur » du risque de température : les contrats standard de type HDD (pour l'hiver) et CDD (pour l'été) se voient concurrencés par des contrats d'autres types<sup>24</sup>. Cette diversification va sans doute de pair avec l'apparition de nouveaux acteurs dans le marché des produits dérivés climatiques : à côté des producteurs d'énergie préoccupés par le risque de température, sont venues se joindre des chaînes de parc d'attraction dont le principal risque est la pluie, des stations de sport d'hiver soucieuses de la hauteur de neige, etc. Remarquons pour finir l'inégalité entre l'hiver et l'été, les échanges relatifs à la saison hivernale étant de loin les plus importants (le phénomène a d'ailleurs tendance à s'accroître). Pour l'expliquer, on peut raisonner sur les transactions sur la température réalisées en Amérique du Nord, qui constituent une large majorité. Or on peut vérifier que l'amplitude des variations de température est plus élevée au cours de l'hiver<sup>25</sup>. Le risque de température est donc plus important dans cette saison.

---

<sup>22</sup> Voir la liste complète sur le Tableau 0.2.

<sup>23</sup> On peut remarquer, en comparant le nombre de transactions et leur valeur, que les transactions réalisées en Asie se rapportent à des nominaux très inférieurs, en moyenne, à ce qui est pratiqué dans les autres continents.

<sup>24</sup> La définition précise des indices HDD et CDD sera donnée au paragraphe suivant.

<sup>25</sup> Voir par exemple (Cao, Wei, 1998).

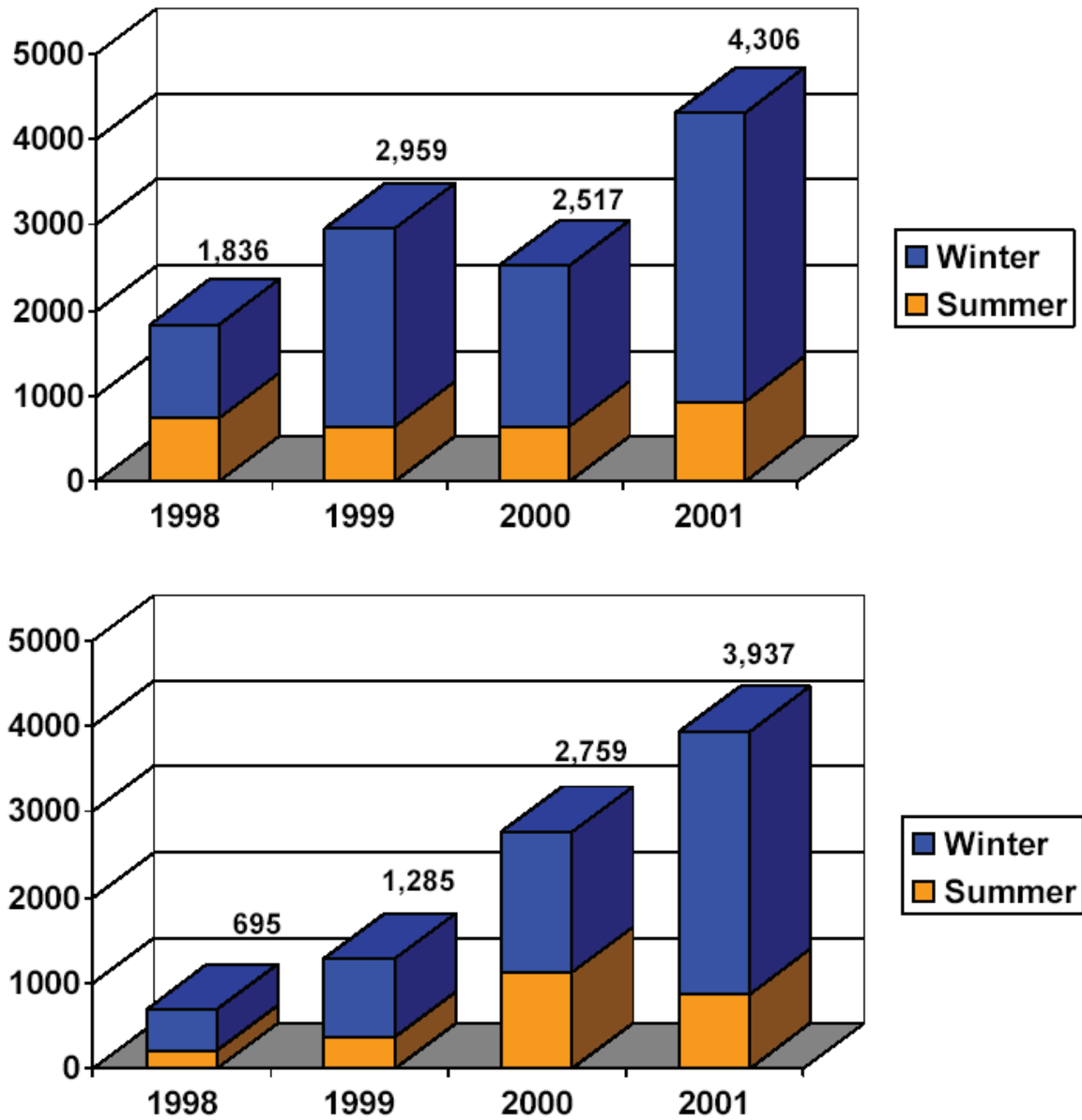


Figure 0.6. Taille du marché des produits dérivés climatiques : valeur (haut) en millions de dollars, et nombre de transactions (bas).

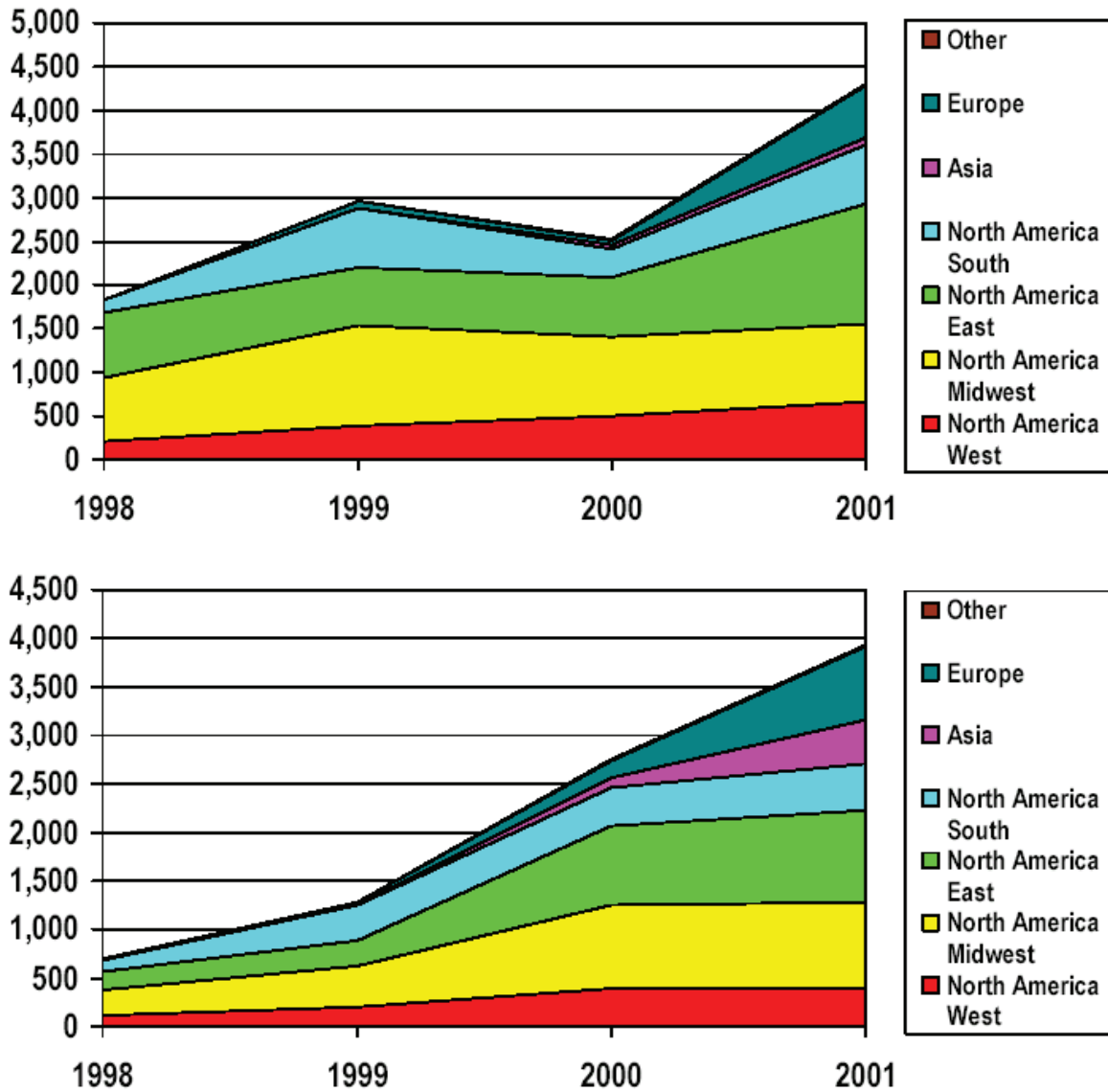


Figure 0.7. Répartition géographique du marché des produits dérivés climatiques : valeur (haut) en millions de dollars, et nombre de transactions (bas).

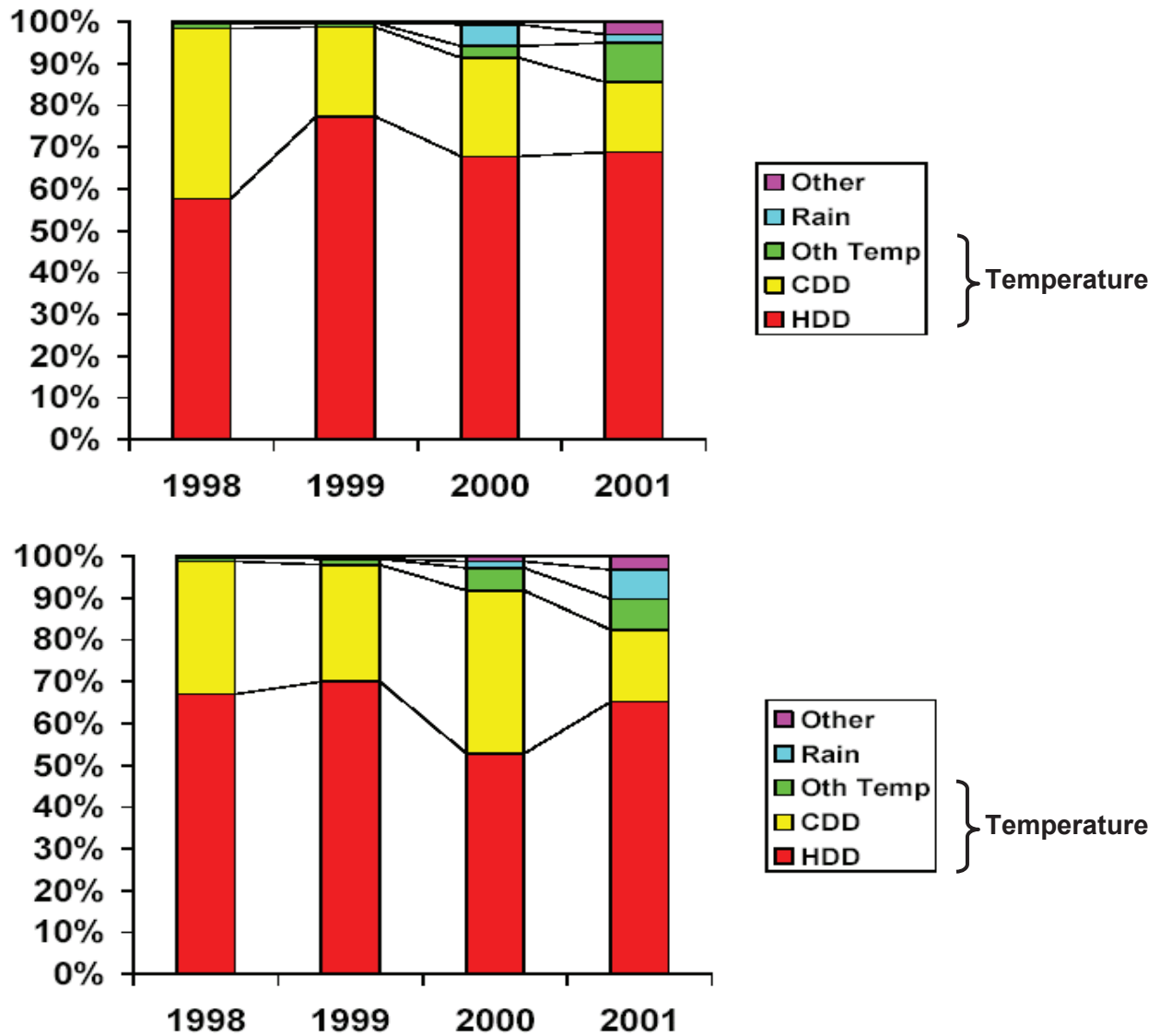


Figure 0.8. Répartition des transactions de produits dérivés climatiques selon la nature du risque : en valeur (haut) et en nombre de transactions (bas).



### III. Quelques spécifications de contrats.

#### 1. Sur les marchés organisés.

Sur le CME et le LIFFE-Euronext, seul le risque de température apparaît dans les contrats ; c'est le risque climatique pour lequel le besoin de couverture est le plus important à l'heure actuelle. Les spécifications sont résumées dans le Tableau 0.3.

Marché	CME	LIFFE-Euronext
Indice de référence	Indice HDD, indice CDD	100 + Température moyenne
Longueur de la période d'exposition au risque	1 mois	a) 1 mois b) 1 saison
Localisation des contrats	10 villes des Etats-Unis : Atlanta, Chicago, Cincinnati, Dallas, Des Moines, Las Vegas, New York, Philadelphia, Portland, Tucson.	Londres, Paris, Berlin
Nature des contrats	Futures ou options	Futures
Maturité	1 à 12 mois	a) 1 à 12 mois b) 1 à 2 saisons identiques
Nominal	1°F = 100\$	1°C = 3.000€ (ou £)

Tableau 0.3. Spécifications des contrats sur les marchés organisés, en 2002.

#### *Localisation*

Le marché se concentre autour des sites qui enregistrent la plus grande liquidité : Londres, Paris, Berlin en Europe, New York, Chicago, Dallas, ... en tout une dizaine de villes, aux Etats-Unis.

#### *Période d'exposition au risque*

Les périodes d'exposition au risque doivent être au minimum de 1 mois et peuvent aller jusqu'à une saison.

#### *Maturité*

Pour les contrats mensuels (resp. saisonniers), la maturité s'étale de 1 à 12 mois (resp. deux saisons identiques : deux hivers ou deux étés) : à une date donnée, on peut donc souscrire 12 contrats mensuels différents, un pour chacun des douze prochains mois (resp. quatre contrats saisonniers : un pour chacun des deux prochains hivers et des deux prochains étés).

« *Sous-jacent* »

Le risque de température est mesuré par le biais d'indices construits à partir de la température journalière moyenne, définie comme la moyenne des températures minimale et maximale observées. Ces indices font office de sous-jacents pour les produits dérivés de température.

Au LIFFE-Euronext, c'est simplement la moyenne mensuelle (ou sur une saison) de ces températures à laquelle on ajoute 100, qui fait référence :

$$I(\text{période}) = 100 + \frac{1}{L} \cdot \sum_{t \in \text{période}} T_t$$

où  $T_t$  désigne la température moyenne du jour  $t$  en degrés Celsius et  $L$  la longueur de la période.

Au CME, on utilise les *degrés-jours* (*degree days*) déjà bien connus des professionnels de l'énergie,

$$CDD(t) = (T_t - 65)^+ \quad HDD(t) = (65 - T_t)^+$$

- où la température  $T_t$  est donnée en degrés Fahrenheit<sup>26</sup>, et les indices de référence sont le cumul des degrés-jours sur la période concernée :

$$CDD(\text{mois}) = \sum_{t \in \text{mois}} (T_t - 65)^+ \quad HDD(\text{mois}) = \sum_{t \in \text{mois}} (65 - T_t)^+$$

Le seuil de 65°F - environ 18°C, correspond à peu de choses près à un minimum dans la consommation d'électricité : quand on s'écarte de cette température « de confort », la consommation augmente - celle des appareils de chauffage l'hiver, celle des réfrigérateurs et des climatiseurs l'été<sup>27</sup>. L'indice HDD - pour *Heating Degree Day* (« degré-jour chauffage »), qui cumule les écarts en dessous de cette température, est donc adapté à l'hiver ; en pratique, il sert de sous-jacent pour les mois d'octobre à avril inclus. Inversement, l'indice CDD - pour *Cooling Degree Day* (« degré-jour réfrigération »), convient à l'été ; il est employé d'avril à octobre inclus. Les mois d'octobre et d'avril se placent à la charnière de deux saisons<sup>28</sup> : c'est pourquoi on les retrouve dans les deux types d'indices.

On peut d'ores et déjà remarquer que pour les mois d'hiver et d'été les mesures de risque par des indices de type degré-jour (CME) donnent des résultats équivalents à la température moyenne (LIFFE-Euronext) : pour les mois d'hiver par exemple, la quasi-totalité des

<sup>26</sup> La correspondance entre degré Fahrenheit et degré Celsius s'effectue au moyen de la formule :

$$TF = 32 + 1,8 \times TC$$

où  $TF$  (resp.  $TC$ ) désigne la température en degré Fahrenheit (resp. Celsius).

<sup>27</sup> Voir par exemple (Ellithorpe, Putnam, 2000).

<sup>28</sup> ... commerciales : « l'hiver boursier » s'étale d'octobre à avril, « l'été boursier » de mai à septembre.

températures est située en dessous de 65°F si bien que l'indice HDD d'un mois d'hiver est une fonction affine de la température moyenne :

$$HDD(\text{mois d'hiver}) \approx \sum_{t \in \text{mois}} (65 - T_t) = L \times \left[ 65 - \frac{1}{L} \cdot \sum_{t \in \text{mois}} T_t \right].$$

### *Nominal*

Enfin, on donne à chaque unité d'indice (en degré) une valeur financière arbitraire par le biais du nominal. Notons que les risques unitaires associés aux contrats futures mensuels du CME (100\$) et du LIFFE-Euronext (3.000€) se correspondent à peu près à l'heure actuelle, au moins en ce qui concerne les mois d'hiver ou d'été. En effet, considérons par exemple un mois d'hiver. Un contrat future paye la différence entre la valeur de l'indice constatée et la valeur contractuelle. Or, la formule ci-dessus indique qu'une variation de 1°C de la température moyenne<sup>29</sup> est équivalente à une variation de 30°C dans l'indice HDD, ce qui correspond à une variation de 54°F (cf. note 26). Etant donnée que les devises sont - à l'heure où nous parlons, du même ordre de grandeur, on en déduit qu'un paiement de 100\$ dans un contrat future acheté au CME est équivalent à un paiement d'environ 5.400€ pour le contrat correspondant au LIFFE-Euronext.

D'autre part, avec 100\$ par degré-jour, la taille des contrats proposés au CME est inférieure à ce qui se pratique sur les marchés OTC<sup>30</sup> ; avec la remarque précédente, il en est donc de même pour les contrats du LIFFE-Euronext. Comme pour le choix des localités, cette disposition favorise la liquidité du marché.

### ■ Un exemple d'utilisation.

Considérons le cas le plus fréquent d'un producteur d'énergie confronté au risque de température. Plaçons-nous par exemple juste avant l'hiver, et considérons un petit fournisseur d'énergie alimentant la ville de Chicago, désireux de se protéger contre le risque que la température soit trop douce lors du prochain hiver. Celui-ci peut alors se couvrir en prenant des positions courtes au CME sur un certain nombre de contrats futures de type HDD pour les mois d'octobre, novembre, décembre, janvier, février, mars et avril. La protection fonctionne alors comme une couverture par des futures classique, lui donnant une compensation au cas où le risque se produit et amputant ses résultats dans le cas contraire. Supposons par exemple que les contrats HDD du mois de décembre qu'il a souscrits aient été cotés à 1115 points au

---

<sup>29</sup> Ou, ce qui revient au même, une variation de 1°C dans l'indice du LIFFE-Euronext.

<sup>30</sup> En 2000, Ellithorpe et Putnam donnaient comme « taille typique », la valeur de 5.000\$ par degré-jour.

moment de la transaction (ce qui porte à  $1115 \times 100\$ = 115.000\$$  la valeur future d'un de ces contrats, le nominal étant de 100\$ au CME). A la fin du mois de décembre, deux situations sont possibles :

- Le mois de décembre a été clément et l'indice HDD a été évalué, par exemple, à 900. Ses pertes de production sont alors compensées par un gain égal, pour chaque contrat souscrit, à :

$$(1115 - 900) \times 100\$ = 21.500\$$$

- Inversement si l'hiver a été rude et que l'indice HDD a été évalué à 1300, alors ses gains de production seront compensés par une perte de  $(1300 - 1115) \times 100\$ = 18.500\$$  par contrat. ■

## 2. Sur le marché OTC.

Les transactions de gré à gré s'effectuent sous des formes très variées, autorisant la couverture d'une gamme étendue de risques climatiques : la température bien sûr, mais aussi la pluie, la neige, le vent, l'ensoleillement, etc. Cette diversité se retrouve également :

- dans la nature des produits dérivés :
  - o en plus des futures et des options, on trouve des *collars*, *strangles* et autres montages classiques (qu'il est d'ailleurs souvent possible de synthétiser à partir d'options standard) ;
  - o il existe également des *obligations climatiques (weather-linked bonds)*, analogues aux obligations catastrophes : l'obligation procure un taux intéressant à son détenteur en échange de l'engagement à porter tout ou partie d'un risque climatique non catastrophique<sup>31</sup> ;
- dans la maturité (parfois plusieurs années) ;
- ou dans le choix des indices climatiques. Ainsi ne serait-ce que pour la température, on rencontre en plus des indices HDD, CDD et de température moyenne, les indices de température maximale, température minimale, l'indice GDD utilisé en agriculture<sup>32</sup>, l'indice EDD<sup>33</sup> ..., ainsi que des combinaisons de ces indices.

<sup>31</sup> La première obligation climatique aurait été lancée par Koch Industries pour une maturité de trois ans sur un nominal de 45 millions de dollars (Ellithorpe, Putnam, 2000).

<sup>32</sup> L'indice GDD (Growing Degree Day), défini par  $GDD(t) = (T_t - 50)^+$ , est donc un indice CDD dont le seuil est fixé à 50°F (environ 10°C) au lieu de la valeur usuelle 65°F. Il semble avoir été introduit en agriculture pour traduire le fait que la température est un facteur influent pour la croissance des cultures au-delà de 50°F (en en deçà de 86°F – environ 30°C, voir Trevor, Linsley, 2002). L'existence d'un deuxième seuil explique que l'on

Cependant on peut remarquer que plus de 80% des transactions se font sur la température et sont basées sur les indices CDD et HDD qu'utilise le CME<sup>34</sup>.

D'autre part, les dérivés sont souvent « cappés » c'est-à-dire que le paiement, ou *payoff*, est limité. En ce qui concerne les transactions de type futures ou options, il est aisé de remarquer qu'ajouter une clause limitant le montant des paiements revient à rajouter au portefeuille une position - longue ou courte selon le cas - dans une option d'achat. Ainsi, détenir une option d'achat cappée (*capped call* ou *call spread*) de prix d'exercice  $K$  et de paiement maximal  $M$ , est équivalent à détenir un portefeuille constitué de l'option d'achat standard de même prix d'exercice et d'une position courte dans l'option d'achat de prix d'exercice  $K + M$  (voir la Figure 0.9.).

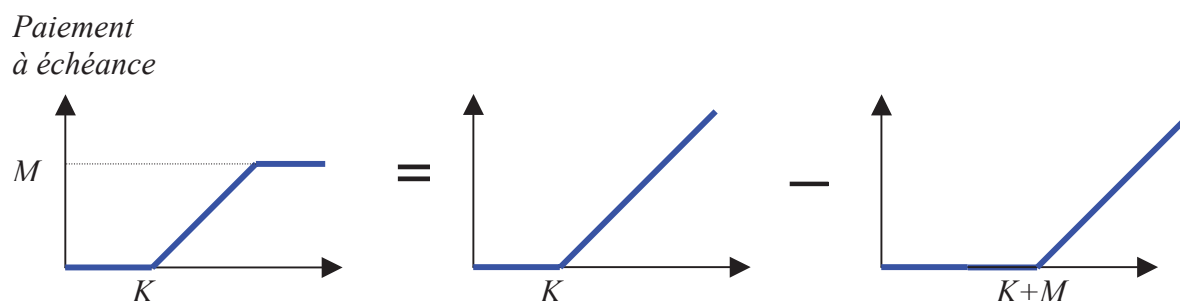


Figure 0.9. Décomposition d'une option d'achat cappée (*call spread*) en deux options d'achat standard.

## F. Panorama des problèmes.

### I. Difficultés spécifiques.

Comme pour tout nouveau produit dérivé, se posent pour les produits dérivés climatiques les questions de l'évaluation, de la couverture, de la gestion de portefeuilles, etc. Toutefois, les caractéristiques tout à fait particulières du sous-jacent rendent la tâche délicate.

La difficulté majeure provient du fait que les sous-jacents climatiques ne sont pas échangeables. Ainsi, pour les contrats Futures, la procédure d'arbitrage de type *cash-and-*

---

rencontre également la définition plus restrictive (par exemple dans Ellithorpe, Putnam, 2000) :

$$GDD(t) = \max\left((T_t - 50)^+ ; 86\right).$$

<sup>33</sup> Energy Degree Day,  $EDD = HDD + CDD$  (Ellithorpe, Putnam, 2000).

<sup>34</sup> D'après l'étude réalisée par PriceWaterhouseCoopers pour WRMA, présentée dans la section "Etat du marché" de cette introduction (voir en particulier la Figure 0.8).

*carry* est inopérante. De même, pour les options écrites directement sur des sous-jacents climatiques, la méthodologie de Black et Scholes basée sur la duplication de l'option avec le sous-jacent est hors de propos. En revanche, elle pourrait être envisagée lorsque le sous-jacent de l'option est constitué d'un contrat future climatique (comme c'est le cas sur le marché du Chicago Mercantile Exchange). Cependant, la faible liquidité sur le marché des contrats futures rend le mécanisme inapplicable. Mais, évidemment, cette faible liquidité est liée - justement - au caractère non-monnayable des sous-jacents climatiques. C'est donc ici aussi cette caractéristique qui est responsable, bien que d'une façon moins directe, de l'inefficacité de la méthodologie de Black et Scholes. Au final, la situation est donc équivalente à une situation de marchés incomplets, et des techniques *ad hoc* doivent être employées.

Par ailleurs, l'incertitude liée aux risques climatiques n'est pas seulement temporelle, de même que pour les actions ou les taux d'intérêts, mais également spatiale. Cette dimension complique la situation à plusieurs égards. D'une part, d'un point de vue technique, les modélisations sont plus complexes. D'autre part, le nombre des stations météorologiques est limité, ce qui oblige pour étudier un risque climatique en un endroit donné à recourir aux mesures de stations voisines ; le risque de base qui en résulte peut être très important (cas de la pluie).

Enfin une troisième source de difficultés est le manque de données de prix : la confidentialité sur le marché OTC, la faible liquidité sur les marchés organisés font que les prix des transactions parviennent au compte-goutte. Des pratiques courantes ne peuvent donc plus être effectuées : ainsi le calage sur les données de marché (*mark-to-market*) est-il impossible et la gestion de portefeuilles « traditionnelle », utilisant les caractéristiques financières individuelles des actifs, doit-elle s'adapter au fait que ces caractéristiques ne sont pas disponibles sur le marché.

## **II. Place des modèles statistiques.**

### **1. Du *mark-to-market* au *mark-to-model*.**

Le vide formé par l'absence de données de prix revêt la modélisation statistique d'une importance particulière. Car si, pour les marchés classiques, les sorties de modèles sont utilisées comme une aide à la décision, leur rôle pour les produits dérivés climatiques est celui de référence : n'ayant pas d'historiques de prix sur lesquels s'appuyer, les courtiers n'ont

guère d'autre choix avant d'entamer les négociations, que d'examiner les prix de modèles obtenus par des simulations du climat. Ces simulations peuvent également être d'un grand secours pour la gestion de portefeuilles. En effet, s'il est impossible, à partir d'historiques de prix, de connaître une estimation de la rentabilité espérée ou de la matrice de variance-covariance des produits constitutifs d'un *hedge fund*, on peut espérer pouvoir obtenir ces estimations en calculant des prix de modèles de ces produits.

## **2. Utilisation pour la résolution des problèmes généraux.**

La modélisation statistique est donc à la base des solutions envisageables pour les produits dérivés climatiques aux problèmes-types des produits dérivés : évaluation, couverture, gestion de portefeuilles, etc. On trouvera dans le Tableau 0.4 un panorama simplifié de la situation.

Bien souvent, la modélisation va concerner un ensemble de sous-jacents. Il peut s'agir de prendre en compte la présence de plusieurs risques comme dans le cas de la gestion des *hedge-funds*. Il peut s'agir également d'utiliser la dimension spatiale pour améliorer la modélisation relative à un lieu, étant donné que les variables climatiques sont des éléments interdépendants du phénomène complexe des déplacements de masses d'air. Il peut s'agir encore d'utiliser la dimension géographique comme moyen de couverture partielle : on peut penser qu'un bon moyen de se protéger du risque relatif à une position est de prendre la position opposée dans un lieu proche, par exemple souscrire à Lyon le même contrat que l'on a vendu à Paris. Cette idée ne peut s'appliquer qu'aux aléas climatiques pour lesquels on observe une réelle dépendance spatiale : la température rentre dans ce cadre, mais la pluie, par exemple, en sort.

Outre la couverture géographique, un autre axe de recherche pour la couverture et l'évaluation consiste à utiliser des produits dérivés corrélés, comme les produits de l'électricité. Ceci pourrait s'appliquer, notamment, à toutes les transactions – la majorité actuellement – qui font intervenir un producteur d'énergie.

## **III. Problèmes spécifiques.**

A côté des problèmes généraux, communs à tous les produits dérivés, se posent des problèmes spécifiques aux produits dérivés climatiques. Nous donnons ci-dessous une liste de quelques-uns de ces problèmes.

### **1. Que faire si aucune mesure de la variable climatique n'est disponible ?**

Les marchés organisés ont restreint les contrats qu'ils proposent à des grandes villes, pour lesquelles on dispose d'un historique conséquent de données de température. Cependant lorsqu'on sort de ces quelques villes, il est fréquent d'avoir des historiques limités, voire inexistant, et la question se pose de savoir s'il est possible, et comment il est possible, de gérer un contrat relatif à un tel site. Dans le cas de la température où des dépendances spatiales existent, un axe de recherche est d'utiliser les sites avoisinants pour lesquels des mesures sont disponibles.

### **2. Comment concevoir des produits dérivés climatiques adaptés aux besoins des acteurs ?**

Cette question se pose par exemple pour le secteur de l'agriculture. Pour fournir une protection efficace contre de mauvaises récoltes, il semble nécessaire de prendre en compte au moins deux sous-jacents, la température et la pluie, en des sites en général non référencés. Si, pour la température, le problème de l'absence de données peut être contourné, il n'en va pas de même de la pluie pour laquelle le risque de base est très élevé. En outre, il s'avère indispensable de préciser le lien entre les conditions climatiques et la quantité de récolte obtenue.

Cette question se pose également, dans une moindre mesure, sous l'angle de l'amélioration de la couverture existante. Pour l'heure, les produits dérivés climatiques offrent une compensation proportionnelle aux écarts d'une variable climatique par rapport à un niveau ou un seuil convenu ; or il est possible que la fonction d'endommagement des acteurs ne soit pas linéaire : on peut imaginer que la sévérité d'une mauvaise saison oblige l'entreprise touchée à procéder à des licenciements, à avoir recours à des emprunts dans de mauvaises conditions, etc. qui sont d'autant de surcoûts envisageables.

### **3. Gestion des risques climatiques.**

Outre son utilité directe pour les produits dérivés climatiques, la connaissance des risques climatiques peut intéresser les agents économiques dans une perspective de prévision des risques.



Problèmes-types	Difficultés spécifiques	Axes de recherche
<i>Evaluation</i>	Sous-jacent non échangeable	- Modélisation statistique du sous-jacent - Développement de méthodologies propres aux marchés incomplets (emploi de fonctions d'utilité)
<i>Couverture / Evaluation</i>	Sous-jacent non échangeable	- Modélisation statistique d'un ensemble de sous-jacents - Couverture partielle avec des produits financiers corrélés (dérivés de l'énergie) - Couverture géographique en utilisant la corrélation spatiale des sous-jacents (cas de la température par exemple)
<i>Gestion de portefeuilles</i>	Manque de données de prix	- Modélisation statistique de l'ensemble des sous-jacents du portefeuille

Tableau 0.4. Spécificités des produits dérivés climatiques pour la résolution des problèmes-types des produits dérivés.

## G. Application des modèles statistiques aux produits dérivés climatiques.

### I. Modélisation statistique des aléas climatiques.

Jusqu'à présent, les recherches ont porté principalement sur la température : c'est en effet l'aléa climatique qui intervient le plus souvent dans les contrats (voir le paragraphe « Etat du marché »). On trouve également quelques travaux sur la pluie (Moréno, 2001). Au plan statistique, la pluie paraît plus difficile à modéliser que la température, car elle suppose la connaissance de la loi jointe de *deux* variables aléatoires : l'occurrence de la pluie, et la hauteur de pluie lorsqu'il pleut ; il y a, en outre, le problème de la discontinuité spatiale : il peut pleuvoir des cordes sur un lieu... aux environs duquel le soleil brille<sup>35</sup> ! Paradoxalement pourtant, il semble que la pluie ait été davantage étudiée, notamment en raison des applications à l'agriculture, l'hydrologie et l'écologie. Des modèles élaborés, de type spatio-temporel, sont déjà développés (voir Sanso, Guenni, 2000). Cependant, l'adaptation de ces modèles à la thématique financière ne semble pas avoir été encore réalisée. Dans le cas de la température, les travaux se concentrent sur la modélisation de l'aléa en un lieu fixé ; le mouvement des masses nuageuses, ou des phénomènes climatiques exogènes comme El Niño ou son homologue de l'Atlantique, ne sont pas encore incorporés<sup>36</sup>.

<sup>35</sup> Le risque de base pour la pluie n'est-il pas surévalué ? C'est cette question que Dischel (2000) s'est posée et dont il nous invite à reconsidérer la réponse : on trouve, dans certaines situations, des corrélations nettes entre les précipitations mesurées en des lieux différents.

<sup>36</sup> On pourra néanmoins consulter (Dischel, 1998) pour une discussion autour du phénomène El Niño, et (Antoniadou et al., 2001) pour une étude de l'oscillation Atlantique Nord.

Parmi les recherches déjà effectuées sur la température, on peut distinguer essentiellement deux voies. La première consiste à adapter des modèles traditionnels de la finance au cas de la température. Ainsi, Dischel (1998), Dornier et Quéruef (2000) ou Moréno (2000) ont présenté quelques extensions du modèle de Hull et White<sup>37</sup>. On rencontre également le mouvement brownien géométrique, pour modéliser les indices de température de type HDD ou CDD (Davis, 2001). Et plus récemment, Brody, Syroka et Zervos (2001) ont introduit le processus d'Ornstein-Uhlenbeck *fractionnaire*, obtenu en remplaçant le mouvement brownien standard par un mouvement brownien fractionnaire, de façon à prendre en compte un effet de mémoire longue. Cette stratégie est également observée par (Caballero, Jewson, Brix, 2002), qui emploient un modèle de type ARFIMA. La deuxième approche est davantage de nature inférentielle, la modélisation s'effectuant exclusivement à partir de l'examen des données ; on emploie alors des méthodes et des résultats issus du domaine des séries temporelles. Initialement, Cao et Wei (1998) ont introduit un modèle linéaire ARMA avec une variance périodique. Son adéquation avec les données françaises semble relativement bonne (Roustant, 2002). Cependant, ses défauts semblent amplifiés dans le cas des villes des Etats-Unis et il devient alors nécessaire d'introduire une composante d'hétéroscédasticité conditionnelle (Campbell, Diebold, 2000).

## II. La pratique de l'évaluation.

Comment passer de la modélisation statistique à l'évaluation et la gestion des produits dérivés climatiques ? Au plan théorique, la question n'a pas une réponse immédiate comme c'est le cas des produits dérivés pour lesquels le sous-jacent est coté : il est absurde de vouloir répliquer une option climatique avec... des indices de température, de pluie, de vent ou de neige ! La méthodologie de Black et Scholes ne peut donc pas s'appliquer ici. Des approches spécifiques ont été ébauchées. Davis (2001) fait remarquer que dans le cas du marché climatique, les agents (comme les compagnies d'énergie) ne sont pas représentatifs, et utilise le concept de « valeur marginale » de l'économie ; cela l'amène à étudier la loi de probabilité jointe d'un indice HDD et du prix de l'énergie de référence. Barrieu et El Karoui (2002) proposent un cadre général pour les produits dérivés écrits sur des actifs non liquides et déterminent simultanément la structure optimale des contrats et leur prix ; selon leurs résultats, on ne devrait pas modéliser le risque non financier (ici : climatique) séparément du risque financier. Pour les produits dérivés climatiques, on peut voir ce résultat comme l'extension d'une approche actuarielle. En effet, lorsqu'on fait les hypothèses d'une faible

---

<sup>37</sup> Voir (Hull, White, 1990).

aversion au risque de l'investisseur<sup>38</sup> et de l'indépendance entre la température et le marché financier<sup>39</sup>, la formule d'évaluation qui est donnée dans la proposition 3 se réduit au principe de la variance :

$$P = E[\textit{payoff actualisé}] + \lambda \cdot \textit{var}[\textit{payoff actualisé}].$$

Plus généralement, dans l'avenir, des méthodes d'évaluation pourraient se baser sur des approches spécifiques aux marchés incomplets telles que celle développée par Carr, Geman et Madan (1999).

Dans la pratique - à défaut d'un véritable consensus, l'évaluation se fait selon une approche actuarielle. Les produits dérivés climatiques sont évalués par une formule du type (voir Geman, 2001) :

$$P = E[\textit{payoff actualisé}] + \lambda \cdot \sigma[\textit{payoff actualisé}]$$

où l'espérance et l'écart-type sont calculés dans l'univers historique, et où  $\lambda$  est choisi... à la discrétion des intéressés. C'est le « principe de l'écart-type », bien connu en assurance, où il est employé pour le calcul des primes ; sa popularité est due pour l'essentiel à son invariance par rapport à une transformation d'échelle. On pourra consulter (Bühlmann H., 1996) pour une présentation des différents principes de calcul de primes. Notons que certaines de ces règles heuristiques peuvent être justifiées par des résultats théoriques : le principe de la variance par exemple, est une approximation de méthodes d'évaluation non linéaires basées sur les fonctions d'utilité (Gerber, Parfumi, 1998) ; en outre, l'aversion pour le risque  $\lambda$  peut être reliée à la probabilité de la ruine de l'assureur (voir par exemple Gerber, 1980).

### III. Le risque de modèle.

La faible liquidité du marché des produits dérivés climatiques rend impossible toute tentative de validation à partir des données de marché (calibration) et accroît par conséquent l'importance des prix de modèles. Or les modèles sont des approximations de la réalité, toujours entachés d'erreurs. Puisque les modèles prennent de l'importance, il importe de se soucier davantage de leurs résultats.

Or, si la validation statistique d'un modèle est évidemment nécessaire pour reproduire avec réalisme des trajectoires de sous-jacents climatiques, elle n'en demeure pas moins insuffisante

---

<sup>38</sup> Ce qui est vérifié dans le cas des entreprises.

<sup>39</sup> Ce qui est vrai en première approximation : on n'observe pas, par exemple, de corrélation linéaire significative entre l'indice CAC 40 et la température à Paris.

pour assurer de l'efficacité du modèle pour l'évaluation. En effet, considérons par exemple le cas du modèle ARMA, utilisé pour modéliser la température. Il consiste, *grosso modo*, en une somme de trois termes : un terme stochastique, une composante de saisonnalité et un terme de tendance sous la forme d'une droite. On peut penser qu'une petite erreur sur la pente de la droite a un impact non négligeable sur les trajectoires de température prédites, et par voie de conséquence en une erreur significative sur les prix annoncés. Pareil désagrément ne serait pas le premier à être observé en finance. Par exemple, même si, théoriquement, l'analyse moyenne-variance de Markovitz est un outil puissant pour sélectionner un portefeuille optimal, son implémentation pose de sérieuses difficultés car l'estimation des poids du portefeuille optimal est très sensible aux erreurs d'estimation des rendements espérés et de la matrice de variance-covariance des actifs individuels (voir Bawa, Brown et Klein, 1979). La prise en compte de ce phénomène a même amené à reconsidérer certaines conclusions pourtant « bien » établies. Par exemple, Jorion (1985) a montré que l'idée selon laquelle la diversification peut à la fois augmenter le rendement et diminuer le risque est erronée ; en considérant l'erreur d'estimation, on observe que les bénéfices de la diversification se situent plus vraisemblablement dans la réduction du risque. On aura compris, en général, la nécessité de se préoccuper des erreurs de modélisation.

« Le » risque de modèle est, en réalité, multiple. Il peut s'agir, comme nous l'avons envisagé jusqu'à présent, de simples *erreurs d'estimation*. C'est le cas en gestion de portefeuilles, c'est aussi le cas dans le contexte des options standard : c'est ainsi que l'on quantifie l'incertitude des prix d'options données par la formule de Black et Scholes en évaluant les répercussions de l'erreur d'estimation de la volatilité (voir Campbell, Lo, Mac Kinlay, 1997, § 9.3.3.). Les insuffisances de cette démarche consistent dans le fait qu'elle ne considère que les erreurs intrinsèques au modèle, sans remettre en cause la validité même du modèle : il ne s'agit pas, à proprement parler, des risques de modèle, mais des risques *du* modèle. Une approche plus générale et plus réaliste consiste donc à supposer que le modèle lui-même est incertain : il devient un événement possible parmi d'autres, auquel on assigne une certaine probabilité d'occurrence. L'inférence est alors réalisée dans un cadre bayésien, et les intervalles de confiance des paramètres englobent alors l'incertitude autour du modèle. Une introduction à ce sujet est faite dans (Cairns, 2000). En pratique cependant, l'évaluation du risque de modèle est délicate et requiert une bonne connaissance de plusieurs modèles : c'est une approche sans doute prématurée pour les produits dérivés climatiques.

## **H. Objectif et organisation de la thèse.**

La présente thèse aborde les aspects économétriques et financiers de l'évaluation des produits dérivés climatiques de température.

Sa structuration en trois chapitres retrace la logique de l'évaluation présentée à la section précédente : modélisation statistique du sous-jacent, aspects financiers de l'évaluation, et estimation des incertitudes des prix dues aux risques de modèle.

Le premier chapitre traite donc la modélisation statistique de la température. Dans un premier temps, seul l'aspect temporel est pris en compte. Deux modèles linéaires univariés de type AR sont alors étudiés et comparés, permettant d'une part de dégager les caractéristiques communes de la dynamique de la température aux différents lieux, et d'autre part d'aboutir à une première conclusion quant à la pertinence de tenir compte de la situation courante dans la gestion des risques climatiques. Aussi naturelle que soit cette première approche, consistant à modéliser la température grâce à l'historique des températures, elle n'en est pas moins limitée car un grand nombre de variables météorologiques influence la température. Au moins aimerait-on prendre en compte le phénomène de circulation atmosphérique. Ainsi, dans un deuxième temps, le modèle linéaire le plus performant est généralisé à un modèle spatio-temporel de type VAR ; les progrès qu'apporte cette extension sont alors quantifiés.

Le second chapitre aborde quelques aspects financiers, en élargissant le champ de l'étude aux actions des entreprises sensibles au climat. Dans un premier temps, nous revenons sur la pratique de l'évaluation, et montrons dans un cadre simple que l'indépendance de la température avec le marché financier donne un argument en faveur de l'utilisation de l'univers historique dans l'évaluation. Dans un deuxième temps, nous examinons l'indépendance entre la température et le marché. Des anomalies ont en effet été remarquées sur le cours des actions : ainsi dans une étude empirique réalisée sur la ville de New York, Saunders (1993) observe que le rendement des actions est significativement plus élevé les jours où l'ensoleillement est fort. Ces effets, cependant relativement faibles, sont généralement attribués à des causes comportementales<sup>40</sup>. D'autre part, dans le cas des entreprises sensibles au climat, les bonnes conditions climatiques dopent les ventes et pourraient donc avoir des conséquences visibles sur les cours de leurs actions. Dans ce cas, on

---

<sup>40</sup> D'autres anomalies ont été observées, comme la possible corrélation des cycles lunaires avec les rendements des actions (Yuan, Zheng, Zhu, 2001). Voir également (Cao, Wei, 2001).

s'attend à ce que le facteur « température » soit pris en compte par le marché, ce que nous étudions.

Le troisième chapitre est consacré à l'estimation de l'incertitude autour des prix que l'on obtient en utilisant le modèle de température développé au chapitre 1 et en adoptant l'approche actuarielle pour l'évaluation. Les deux types de produits dérivés climatiques (standard) sont considérés : futures et options, et deux types d'erreurs de modèle sont envisagés : l'erreur d'estimation des paramètres, et l'écart à la normalité de la distribution des résidus. Outre les renseignements sur les prix que l'on peut en tirer, cette étude permet de porter un regard critique sur l'utilisation du modèle de type AR et d'en préciser les principaux défauts.

## Références

- Antoniadou T., Besse P., Fougères A.L., Le Gall C., Stephenson D.B., *L'oscillation Atlantique Nord et son influence sur le climat européen*, *Revue de Statistique Appliquée*, **XLIX**(3), 39-60, 2001.
- Bank for International Settlements, *Regular OTC Derivatives Market Statistics*, May 2002.
- Barrieu P., *Produits dérivés météorologiques et environnement*, thèse de doctorat en sciences de gestion, Groupe HEC, [www.hec.fr](http://www.hec.fr), 2002.
- Barrieu P., El Karoui N., *Optimal design of derivatives in illiquid markets*, *Quantitative Finance*, **2**, 181-188, 2002.
- Barrieu P., *Introduction au marché des produits dérivés météo*, séminaire ISFA « Couverture des risques naturels et climatiques », Paris, 24 et 25 octobre 2000.
- Bawa V.S., Brown S.J., Klein R.W., *Estimation Risk and Optimal Portfolio Choice*, *Studies in Bayesian Econometrics Bell Laboratories Series*, New York: North Holland, 1979.
- Brody D.C., Syroka J., Zervos M., *Dynamical Pricing of Weather Derivatives*, working paper, 2001.
- Bühlmann H., *Mathematical Methods in Risk Theory*, Springer-Verlag, 1996.
- Caballero R., Jewson S., Brix A., *Long memory in surface air temperature: detection, modeling, and application to weather derivative valuation*, *Climate Research*, **21**, 127-140, 2002.
- Cairns A.J.G., *A discussion of parameter and model uncertainty in insurance*, *Insurance: Mathematics and Economics*, **27**, 313-330, 2000.
- Campbell J.Y., Lo A.W., MacKinlay A.C., *The Econometrics of Financial Markets*, (Princeton University Press), 1997.

Campbell S., Diebold F.X., *Weather Forecasting for Weather Derivatives*, working paper, 2000.

Cao M., Wei J., *Stock Market Returns : A Temperature Anomaly*, working Paper, 2001.

Cao M., Wei J., *Pricing Weather Derivatives: an Equilibrium Approach*, working paper, Queen's University Kingston and University of Toronto, Ontario, Canada, 1998.

Carmona R., *Calibrating Degree Days Options*, working paper, 1999.

Carr P., Geman H., Madan D., *Pricing and Hedging in Incomplete Markets*, working paper, 1999.

Davis M., *Pricing weather derivatives by marginal value*, Quantitative Finance, **1**, 1-4, 2001.

Deboaisne E., *Les dérivés climatiques*, séminaire ISFA « Couverture des risques naturels et climatiques », Paris, 24 et 25 octobre 2000.

Dischel R., *Double Trouble: Hedging Rainfall and Temperature*, Risk Magazine and Energy Power Risk Management, August 2001.

Dischel R., *Seeding a Rain Market*, Environmental Finance, September 2000.

Dischel R., *Is precipitation Basis Risk Overstated?*, Risk Magazine and Energy Power Risk Management, August 2000.

Dischel R., *Warning - La Niña Volatility*, Energy and Power Risk Management, November 1998.

Dischel R., *The Fledgling Weather Market Takes Off, Part. 2.*, Applied Derivatives Trading, October 1998.

Dornier F., Quéruef M., *Caution to the wind*, Risk, Energy and Power Risk Management, August 2000.

Ellithorpe D., Putnam S., *Weather Derivatives and Their Implications for Power Markets*, The Journal of Risk Finance, 19-28, Winter 2000.

Eydeland A., Geman H., *Pricing power derivatives*, Risk, 71-73, October 1998.



- Fewings D.R., Pekoç E.A., *Robust Crop Hedging Strategy for Free Market Agriculture*, working Paper.
- Geman H., *Instruments dérivés à sous-jacent exotique : l'exemple des dérivés climatiques*, Banque & Marchés, **53**, 5-9, 2001.
- Geman H., *The Bermuda Triangle: Electricity, Weather and Insurance Derivatives*, Journal of Alternative Investments, 2000.
- Geman H., *Insurance and Weather Derivatives: from Exotic Options to Exotic Underlyings*. Risk Books, London, 1999
- Geman H., *Cat Calls*, Insurance, **7**(9), 86-89, 1994.
- Gerber H.U. and Parfumi G., *Utility Functions: From Risk Theory to Finance*, North American Actuarial Journal, **2**(3), 74-100, 1998.
- Gerber H.U., *An Introduction to Mathematical Risk Theory*, Huebner, 1980
- Gondran N. *Système de diffusion d'information pour encourager les PME/PMI à améliorer leurs performances environnementales*, thèse de doctorat en Sciences et Techniques du Déchet, Ecole Nationale Supérieure des Mines de Saint-Etienne, [www.agora21.org/entreprise/these2.html](http://www.agora21.org/entreprise/these2.html), 2001.
- Jorion P., *International Portfolio Diversification with Estimation Risk*, Journal of Business, **58**(3), 259-278, 1985.
- Martin P.-E., *Certificats verts : situation européenne et mode d'emploi*, Systèmes Solaires, **147**, 10-17, 2002.
- Moréno M., *Rain Risk*, Speedwell Weather Derivatives, [www.weatherderivs.com](http://www.weatherderivs.com), 2001.
- Moréno M., *Evaluation des Dérivés Climatiques*, Speedwell Weather Derivatives, [www.weatherderivs.com](http://www.weatherderivs.com), 2000.
- Moréno M., *Riding the Temp*, Weather Derivatives, FOW Special Supplement, December 2000.

Nott D.J., Dunsmuir W.T.M., Kohn R., Woodcock F., *Statistical Correction of a Deterministic Numerical Weather Prediction Model*, Journal of the American Statistical Association, **96**(455), 2001.

ONU, Climate Change & The Financial Services Industry, [www.unepfi.net](http://www.unepfi.net), July 2002.

Quéruef M., *Les risques climatiques et naturels : de nouvelles opportunités pour les investisseurs*, séminaire ISFA « Couverture des risques naturels et climatiques », Paris, 24 et 25 octobre 2000

Roustant O., *Une application de deux modèles économétriques de température à la gestion de risques climatiques (1<sup>ère</sup> partie)*, Banque & Marchés, **58**, 22-29, 2002.

Roustant O., *Une application de deux modèles économétriques de température à la gestion de risques climatiques (2<sup>ème</sup> partie)*, Banque & Marchés, **59**, 36-44, 2002.

Sanso B., Guenni L., *A Nonstationary Multisite Model for Rainfall*, Journal of the American Statistical Association, **95**(452), 2000.

Saunders E.M.J., *Stock Prices and Wall Street Weather*, American Economic Review, **83**, 1337-1345.

Scherer A., *Les cat-bonds : de nouvelles opportunités pour les réassureurs*, séminaire ISFA « Couverture des risques naturels et climatiques », Paris, 24 et 25 octobre 2000.

Tchapda Djamen I., *Evaluation de flux monétaires en présence d'un ou plusieurs risques de défaut*, thèse de doctorat, 2002.

Trevor C., Linsley P., *The use of weather derivatives in agriculture*, University of Northumbria at Newcastle, 2002.

Weinstein J., *Carbon-denominated weather swaps*, Environmental Finance, November 2001.

Weather Risk Management Association, *Results of Second Annual Industry-Wide Survey*, [www.wrma.org](http://www.wrma.org), 2002.

Wikle C.K., Milliff R.F., Nychka D., Berliner L.M., *Spatiotemporal Hierarchical Bayesian Modeling: Tropical Ocean Surface Winds*, Journal of the American Statistical Association, **96**(454), 2001.

Yuan K., Zheng L., Zhu Q., *Are Investors Moonstruck? Lunar Phases and Stock Returns*, University of Michigan Business School, 2001.

## **Chapitre 1. Modélisation de la température.**

Le chapitre est structuré de la façon suivante.

La première section propose une étude statistique comparative de deux modèles de température et donne une application à la gestion des risques climatiques. Elle a fait l'objet d'une publication en deux parties dans la revue « Banque & Marchés ». C'est le texte de cette publication qui est présenté ici.

La deuxième section apporte des compléments en fournissant notamment les résultats complets pour une dizaine de stations supplémentaires, réparties dans toute la France.

Elle fait donc la transition avec la troisième section qui aborde des aspects multivariés.

## A. Une application de deux modèles économétriques de température à la gestion des risques climatiques.

### Introduction de la 1<sup>ère</sup> partie.

Les conditions climatiques ont un impact important sur l'économie mondiale. Ainsi l'on estimerait à 20% la part de l'économie américaine directement affectée par le climat<sup>41</sup>. Le secteur énergétique est particulièrement exposé aux risques climatiques, notamment au risque de température : un hiver doux, par exemple, est synonyme de manques à gagner, parfois énormes, pour un producteur d'électricité. Partant de ce double constat, et parallèlement à la déréglementation du marché de l'énergie aux Etats-Unis, un marché de gré à gré de dérivés "climatiques" a vu le jour en 1997 aux Etats-Unis. Depuis le marché s'est développé avec l'apparition de marchés réglementés - Chicago Mercantile Exchange<sup>42</sup> (CME), London International Financial Futures and Options Exchange<sup>43</sup> (LIFFE), ouverture prévue au 4<sup>ème</sup> trimestre 2001 – , et diversifié : les acteurs, initialement issus du secteur énergétique, se rencontrent maintenant dans d'autres secteurs comme celui du tourisme (sports d'hivers, parcs d'attractions, etc.) , ou de la distribution (boissons, etc.).

Les dérivés climatiques procurent pendant une période prédéfinie, une compensation pour des pertes ou manques à gagner directement imputables aux aléas climatiques, à l'exclusion notable des catastrophes naturelles<sup>44</sup>. Ces aléas représentent une situation anormale, exceptionnelle, mais non catastrophique. Ils sont mesurés par des indices observables de tous, tels que le nombre de jours où la hauteur de neige est inférieure à 10 cm, le nombre de jours où la quantité de pluie est supérieure à 10 mm, les indices CDD et HDD<sup>45</sup> pour la température, etc. A la différence d'une assurance, aucune constatation de dommages n'est nécessaire : la compensation est déclenchée lorsqu'une condition sur l'indice est réalisée. A la

<sup>41</sup> Si l'on se réfère à l'information communiquée sur le site Internet du Chicago Mercantile Exchange.

<sup>42</sup> Site Internet : [www.cme.com](http://www.cme.com)

<sup>43</sup> Site Internet : [www.i-wex.com](http://www.i-wex.com)

<sup>44</sup> Certains contrats peuvent néanmoins, de façon "involontaire" et partielle, participer à la couverture d'événements catastrophiques. Les quantités inhabituelles de pluie tombées au cours du passage d'un cyclone seront comptabilisées dans les indices sur la pluie (voir ci-après). Mais les dérivés climatiques ne sont pas conçus pour procurer une indemnisation à la hauteur des préjudices subis.

<sup>45</sup> L'indice CDD (Cooling Degree Day), mesuré en degrés, est la somme des "excédents" de température par rapport à un seuil sur la période couverte :  $CDD = \sum_t \max(T(t) - \text{seuil}; 0)$ , où  $T(t)$  désigne la température moyenne du jour  $t$ . Inversement, l'indice HDD (Heating Degree Day) est donné par :  $HDD = \sum_t \max(\text{seuil} - T(t); 0)$ .

différence du marché de gré à gré, on rencontre des contrats standardisés sur la place boursière de Chicago (CME). Ceux-ci concernent le risque de température et sont constitués de contrats futures et d'options sur futures sur les indices mensuels HDD ou CDD. Les périodes d'exposition au risque sont des mois et s'étalent jusqu'à un horizon d'un an : au 15 octobre 2001 par exemple, on peut souscrire 7 contrats futures HDD différents, un pour chaque mois "d'hiver" de la période Novembre 2001 à Avril 2002 inclus, et 7 contrats futures CDD, un pour chaque mois "d'été" de la période d'Avril 2002 à Octobre 2002 inclus.

Parmi tous les aléas climatiques, c'est la température qui constitue, à l'heure actuelle, le support de la plupart des dérivés. A ce jour, les recherches autour de ces produits ont porté principalement sur la modélisation de la température. En effet, il est utile de mieux connaître le risque sous-jacent, que l'objectif soit la valorisation ou la gestion des risques climatiques comme c'est exclusivement le cas ici. Plus précisément, on fixe une période d'exposition au risque de température et on suppose connue une fonction "d'endommagement" exprimant les pertes par une fonction des températures sur toute la période. On peut alors chercher à connaître à la date courante l'espérance des pertes, la valeur à risque concernant les pertes... etc. On est donc amené à étudier la loi conditionnelle des températures sur la période d'exposition au risque.

Dans la littérature concernant la modélisation de la température, une première approche a été proposée, visant à essayer d'adapter certains modèles issus de la finance. En remarquant que la température s'écarte plus ou moins de la moyenne historique, on peut y voir une analogie avec le phénomène de régression vers la moyenne observé dans les modèles de taux d'intérêt. F. Dornier et M. Quéruel (Dornier, Quéruel, 2000) ont montré comment adapter le modèle de Hull et White à cet effet. Le modèle proposé par B. Dischel (Dischel, 1998), en est un autre exemple. Dans le même temps, le très riche domaine des séries chronologiques a été utilisé. R. Carmona a obtenu un modèle linéaire de type AR(4) pour la station de La Guardia, New York (Carmona, 1999). Dischel a mis en évidence le phénomène de saisonnalité dans la variabilité des températures. M. Cao et J. Wei (Cao, Wei, 1998) ont utilisé cette propriété pour proposer un modèle autorégressif avec une variance périodique. En utilisant les modèles ARMA, M. Moreno (Moreno, 2000, a) observe une dissymétrie dans la distribution des résidus et une saisonnalité dans leur variance. Une approche complémentaire de la modélisation des températures pratiquée par les établissements financiers met en oeuvre des

méthodes non paramétriques de Bootstrap<sup>46</sup> (rééchantillonnage de la loi historique) après correction de la saisonnalité.

Le but de cet article est double. *Primo*, il s'agit de **comparer les deux modèles de température de type ARMA** en étudiant notamment l'importance de l'hypothèse de saisonnalité de la "volatilité". *Secundo*, on se doute que le rapprochement entre la date courante et la période d'exposition au risque peut jouer un rôle. Ainsi si l'on se place au 31 octobre 2001, la connaissance de la température courante n'aura qu'un intérêt limité par rapport à un risque de température situé en avril 2002, tandis qu'elle pourra se révéler utile pour un risque localisé au mois de novembre 2001. D'une façon générale, nous cherchons à **évaluer l'impact de la situation courante sur la gestion des risques climatiques**. Supposons par exemple qu'à la date courante, la température soit inférieure à la normale saisonnière. En quoi ceci a-t-il un effet sur la gestion du risque climatique? Le problème est lié à la rapidité avec laquelle les températures "reviennent à la moyenne". Les modèles économétriques étudiés nous permettent de quantifier ces effets. Plus précisément si l'on note  $t$  la date courante,  $[t+h; t+u]$ ,  $u > h$ , la période d'exposition au risque, la température à la date  $t$ ,  $L_t(X_{t+h}, \dots, X_{t+u})$  la loi conditionnelle (à l'information courante) des températures futures et  $L(X_{t+h}, \dots, X_{t+u})$  la loi non conditionnelle, dans quelle mesure peut-on approximer  $L_t$  par la loi plus simple  $L$  ?

On aurait pu penser utiliser les prévisions de température fournies par Météo France au lieu de modèles économétriques. Il est fort possible que la prévision ainsi réalisée soit plus précise<sup>47</sup>. Néanmoins dans le cadre d'une analyse des risques climatiques, l'horizon est beaucoup plus éloigné que celui des prévisions fournies par Météo France. En outre, pour les études de risque, nous avons besoin d'indications sur la dispersion des températures futures autour de la valeur prévue (intervalles de confiance, volatilité) non fournies par Météo France.

L'étude est structurée en six sections. La première dégage des caractéristiques essentielles de la température : saisonnalité de la variance (qui apparaît comme un défaut de stationnarité), dissymétrie des distributions hivernales. Dans les deux suivantes, nous étudions la validité de

<sup>46</sup> Ces méthodes s'apparentent au calcul de la Value-at-Risk d'un portefeuille selon une méthode "historique". En ce qui concerne les températures, il peut sembler plus approprié d'utiliser des méthodes de rééchantillonnage par blocs compte-tenu des autocorrélations entre les températures (voir par ex. (Li, Maddala, 1996).

<sup>47</sup> Au sens où l'erreur de prévision c'est-à-dire l'écart entre la prévision de température et la température constatée est, en moyenne, plus faible.



deux modèles linéaires classiques. Un modèle de type ARMA d'abord dont nous constatons les insuffisances, confirmant ainsi les travaux de M. Moreno (Moreno, 2000, a). Nous proposons alors un modèle ARMA à variance périodique dans le même esprit que M. Cao et J. Wei et étudions de façon plus approfondie les améliorations qu'il apporte. Par ailleurs, la normalité des lois est discutée. Dans la quatrième section, les prévisions théoriques sont étudiées en détail. En particulier nous quantifions l'horizon à partir duquel on peut approximer la loi conditionnelle  $L_t$  par la loi non conditionnelle  $L$ . La cinquième section s'attache à regarder l'impact d'un "choc" extérieur. Une notion de réponse impulsionnelle adaptée au cadre des dérivés climatiques est introduite à cet effet. Enfin dans la sixième section, les erreurs théoriques de prévision sont comparées avec les erreurs réellement observées ("back-testing"). Cela donne un nouveau critère pour juger de la qualité de chacun des modèles. Les développements mathématiques figurent en annexe.

### **Présentation des données.**

Chaque série de températures correspond à une période s'étalant du 01/01/79 au 31/12/99, pour une fréquence journalière, soit un total de 7670 valeurs. Par "température journalière", nous entendons température journalière moyenne c'est-à-dire la moyenne de la température minimale et de la température maximale observées dans la journée. Chaque série provient d'une station météorologique d'une ville de France, et nous a été fournie par Météo France.

### **Notations.**

Nous avons convenu de retirer les 29 février : de cette façon chaque année comprend le même nombre de jours. Cette transformation représente une suppression de 5 valeurs seulement par série. Nous considérons que les températures observées sont des réalisations de variables régies par des lois de probabilité. La température à la date  $t$  sera représentée par une variable aléatoire  $X_t$ ,  $t = 1, 2, \dots$ . La date d'origine  $t = 1$  est associée au 01/01/79.

## I. Statistique descriptive.

Dans ce paragraphe, nous étudions la série de températures provenant de la station météorologique de Lyon-Bron. Comme il n'y a pas d'ambiguïté sur la localisation, nous noterons simplement  $X_t$  la variable aléatoire représentant la température à la date  $t$ , et  $x_t$  la valeur observée. L'objectif est de cerner les particularités du processus  $(X_t)_{t \geq 1}$  à partir de ces données.

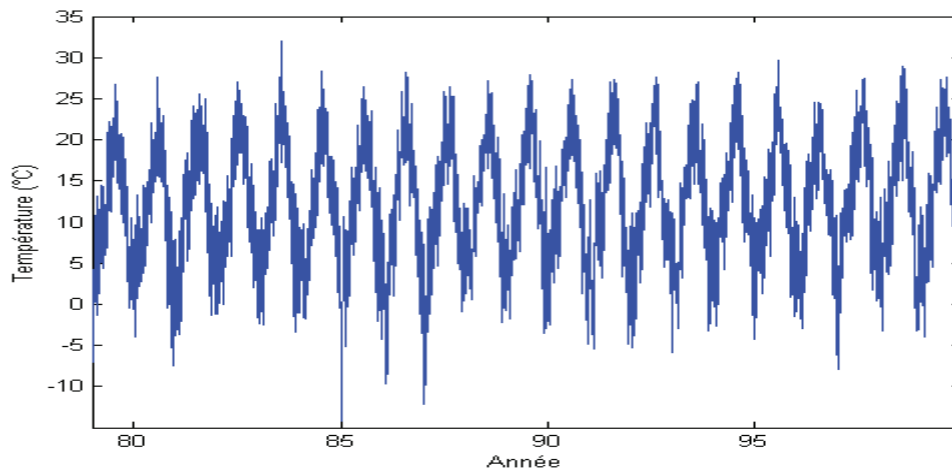


Figure 1.1. Température journalière moyenne à Lyon du 01/01/79 au 31/12/99.

Tout d'abord, la série de températures représentée sur la présente une forte composante de saisonnalité. Nous proposons pour  $X_t$  la décomposition additive classique :

$$X_t = m_t + s_t + Y_t$$

où :

1.  $m_t$  est la tendance, supposée linéaire.
2.  $s_t$  est la composante de saisonnalité; elle vérifie :  $s_{t+365} = s_t$  et  $\frac{1}{365} \sum_{t=1}^{365} s_t = 0$ .
3.  $(Y_t)$  est un processus stationnaire centré.

Il sera commode d'écrire la décomposition sous la forme

$$X_t = (m_t - m_0) + (s_t + m_0) + Y_t$$

De cette façon,  $m_t - m_0$  représente une tendance relative, et  $s_t + m_0$  s'interprète comme la température moyenne en l'absence de tendance.

### 1. Estimation de la tendance et de la saisonnalité.

On estime d'abord la saisonnalité en négligeant dans un premier temps la tendance. Puis celle-ci est estimée à partir de la série désaisonnalisée par ajustement d'une droite au sens des moindres carrés. (Voir (Brockwell, Davis, 1991), §1.4.). Les résultats de l'estimation sont montrés sur la Figure 1.2. Le terme de tendance est haussier. L'urbanisation croissante et un réchauffement global peuvent contribuer à comprendre ce résultat. La série résiduelle ne semble plus présenter de tendance et de saisonnalité. Cependant, nous allons voir qu'elle n'est pas stationnaire.

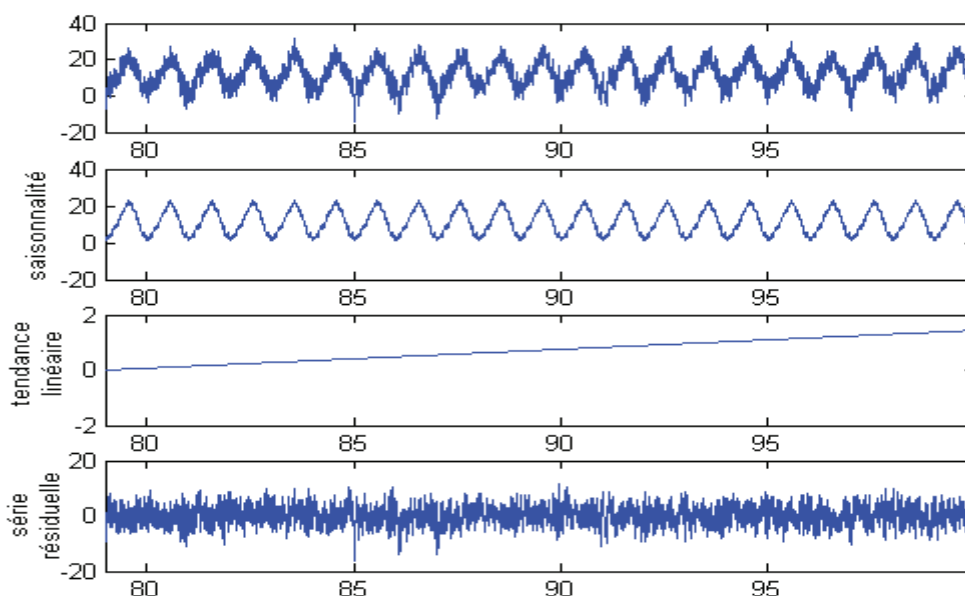


Figure 1.2 – Décomposition additive (2<sup>ème</sup> forme) de la série de température de Lyon-Bron.

### 2. Problèmes de saisonnalité.

Le tracé de la série a fait apparaître une saisonnalité évidente des températures. La décomposition précédente a permis d'éliminer le phénomène sur les valeurs ; Mais la série résiduelle présente encore une saisonnalité dans la *dispersion* de ces valeurs. Pour la mettre en évidence, nous regroupons les données par mois. Plus précisément, pour chaque mois, nous regroupons les 28 premières données du mois pour toutes les années, ce qui représente  $28 \times 21 = 588$  données. La comparaison des distributions des 12 séries mensuelles ainsi obtenues est faite au moyen de "boxplots". La Figure 1.3 présente le boxplot pour la série

résiduelle. Ce type de représentation graphique permet de visualiser d'un seul coup d'œil les principales caractéristiques des 12 distributions. Les côtés inférieurs et supérieurs de chaque rectangle représentent respectivement le premier et le dernier quartile ; le trait à l'intérieur, la médiane. Les points situés à l'extérieur représentent les valeurs extrêmes : leur répartition renseigne sur les queues de distribution.

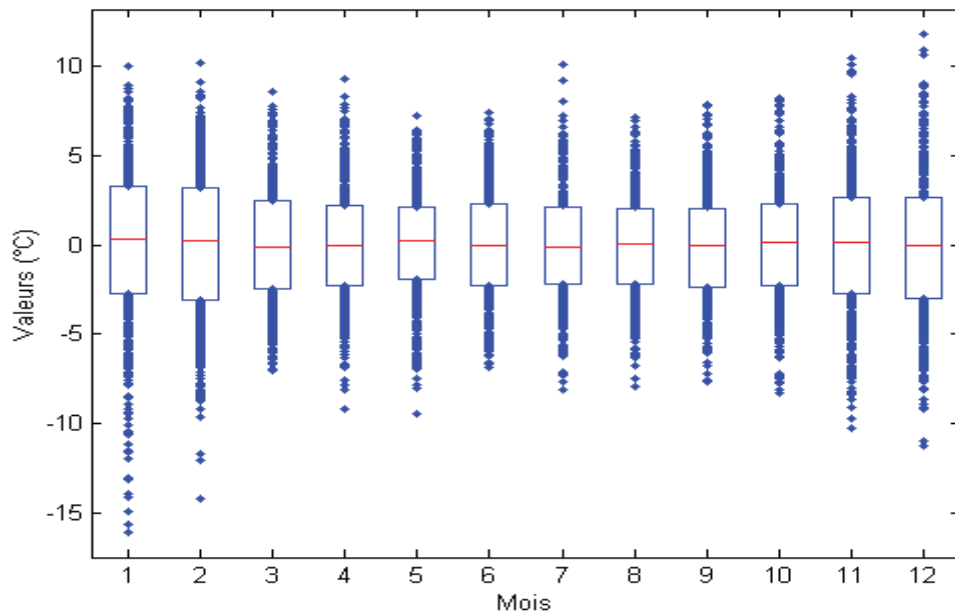


Figure 1.3. Comparaison des distributions de température mois par mois. Série résiduelle.

Comme annoncé, on constate une forte saisonnalité au niveau de la dispersion des valeurs de la température. En hiver, la dispersion est plus importante qu'en été. De plus on constate pour les mois d'hiver - janvier et février surtout - une *dissymétrie* de la distribution : la probabilité d'observer des températures exceptionnellement froides est plus élevée que celle d'observer des températures exceptionnellement douces.

Nous constatons donc que la série désaisonnalisée  $Y_t$  n'est pas stationnaire. Aussi l'ajustement "brutal" d'un modèle linéaire de type ARMA ne donnera-t-il pas des résultats de bonne qualité, comme nous allons le constater. Les modèles économétriques peuvent prendre ou non en compte la saisonnalité de la dispersion des températures. Nous mettons en œuvre ces deux approches afin de les comparer.

## II. Ajustement d'un modèle linéaire ARMA.

Au vu de la Figure 1.3 qui fait apparaître une non-stationnarité légère, on peut se demander si le modèle ARMA, n'est pas, malgré tout, suffisant pour décrire le processus  $Y_t$ . Examinons rapidement ce qu'il en est.

**Sélection des ordres  $p$  et  $q$ .** Le diagramme des autocorrélations (ACF) et des autocorrélations partielles (PACF) sont représentés sur la Figure 1.4 et la Figure 1.5. Nous constatons que les autocorrélations partielles ne sont pas significatives à partir de 4 ou peut-être 3. Cela suggère d'essayer un modèle  $AR(p)$ , avec  $p = 3$  ou 4. Pour préciser l'ordre, nous remarquons que le critère SBC<sup>48</sup> a un minimum pour  $p = 3$ . Nous faisons donc l'hypothèse que  $Y_t$  suit un modèle  $AR(3)$  : il existe des coefficients réels  $\phi_1, \phi_2, \phi_3$  tels que pour tout entier  $t \geq 0$ ,

$$Y_t = \phi_1.Y_{t-1} + \phi_2.Y_{t-2} + \phi_3.Y_{t-3} + \varepsilon_t$$

où  $(\varepsilon_t)$  est un bruit blanc gaussien de variance  $\sigma^2$ .

**Estimation, diagnostic.** Sous cette hypothèse, l'estimation des paramètres et l'étude des résidus sont rassemblés dans le Tableau 1.1. Pour valider le modèle, il faut contrôler que les résidus  $(\varepsilon_t)$  forment un bruit blanc gaussien<sup>49</sup>. Une analyse (trop!) rapide pourrait nous conduire à accepter le modèle  $AR(3)$ . En effet, les  $p$ -valeurs du test de non-corrélation de Ljung-Box<sup>50</sup> sont correctes. Cependant, la  $p$ -valeur du test d'indépendance de Kendall<sup>50</sup> de 2.1% est faible. D'autre part, l'indépendance des résidus a pour conséquence l'indépendance de leur carré. Et cette hypothèse est, elle, clairement rejetée par le test de Ljung-Box. Il y a, de plus, un doute quant au caractère gaussien des résidus. La statistique du test de Kolmogorov<sup>51</sup> d'adéquation à une loi normale est de 0.0155 ce qui est juste égal au seuil de 0.0155 correspondant au niveau 5%. Et le tracé de la droite de Henri (Figure 1.6) fait apparaître une dissymétrie. Elle traduit une queue de distribution plus épaisse dans les valeurs négatives, ce qui est conforme aux observations de la première section.

---

<sup>48</sup> Nous préférons le critère SBC au critère AIC car ce dernier a tendance à favoriser les modèles sur-paramétrés. Pour une définition de ces critères et leur comparaison, on pourra voir par exemple (Lutkepohl, 1993), §4.3.

<sup>49</sup> En réalité, les résidus formeraient un bruit blanc gaussien si le modèle étudié était le *vrai* modèle. Or ici, les coefficients du modèle sont seulement estimés à partir des données. Il est donc faux que les résidus estimés  $(\hat{\varepsilon}_t)$  suivent exactement un bruit blanc gaussien. Néanmoins, c'est vrai approximativement. (A ce sujet, on pourra consulter (Brockwell, Davis, 1991), §9.4.)

<sup>50</sup> Référence : (Brockwell, Davis, 1991), §9.4.

<sup>51</sup> Référence : (Saporta, 1990), §15.4.2.

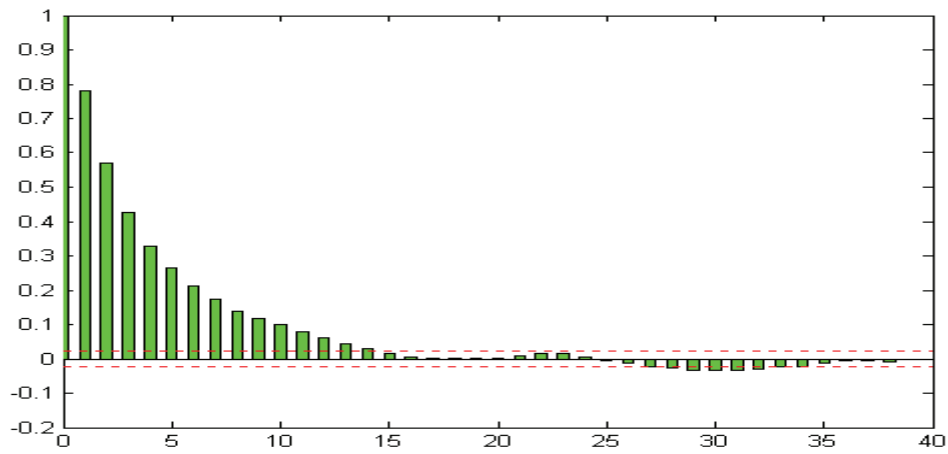


Figure 1.4. – ACF de la série sans tendance et désaisonnalisée.

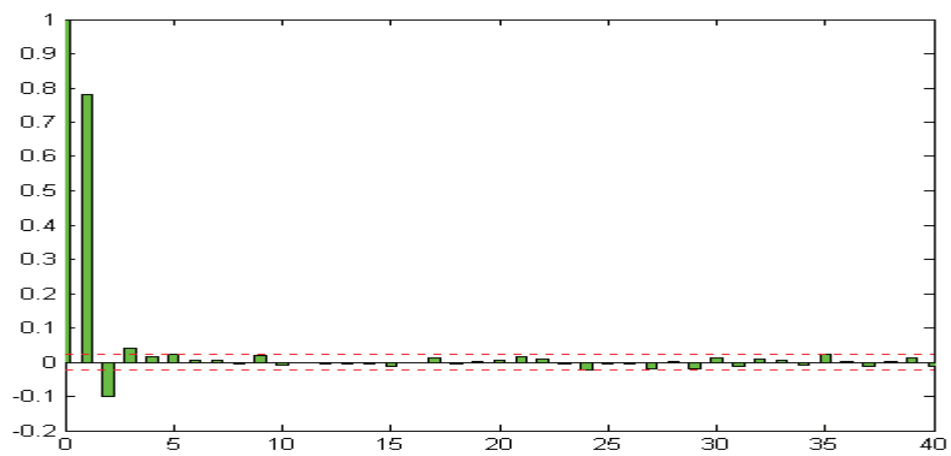
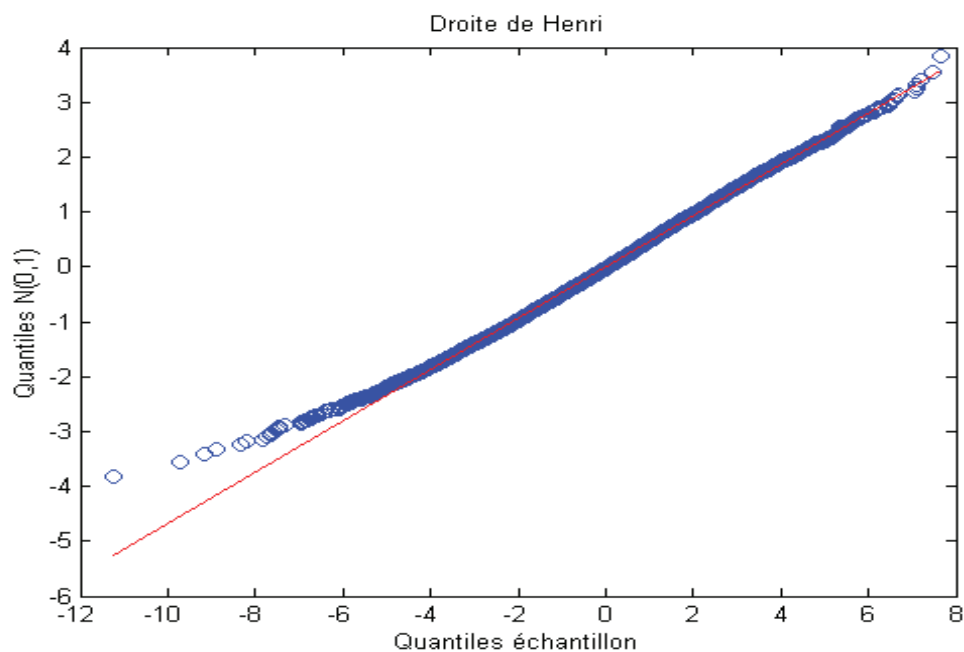


Figure 1.5. PACF de la série sans tendance et désaisonnalisée.

Paramètres estimés					
	Valeur	Ecart-type	Ratio	p-valeur	
<i>bruit blanc</i>					
$\sigma$	2.1385				
<i>partie autorégressive</i>					
$\phi_1$	0.8615	0.0114	75.4813	0.0000	
$\phi_2$	-0.1347	0.0150	-8.9867	0.0000	
$\phi_3$	0.0417	0.0114	3.6516	0.0003	
Résidus			Carrés des résidus		
Test	Largeur	Stat.	p-valeur	Stat.	p-valeur
<i>Ljung-Box</i>	21	12.7474	0.8064	64.3055	0.0000
	42	41.7092	0.3538	115.2589	0.0000
	63	63.3667	0.3585	148.4623	0.0000
	84	81.7677	0.4552	169.6804	0.0000
<i>Kendall</i>			0.0218		
<i>Kolmogorov</i>		0.0155			

Tableau 1.1. Ajustement d'un modèle AR(3). Estimation et analyse des résidus.

Figure 1.6. Droite de Henri des résidus -- Modèle (AR).



On peut alors essayer de recommencer l'étude précédente en augmentant les ordres  $p$  et  $q$ . Mais les résultats sont comparables : l'étude des résidus ne permet pas de rejeter le modèle mais l'étude de leur carré écarte l'hypothèse d'indépendance. Les nouveaux coefficients introduits sont, en plus, très mal estimés.

Ces difficultés ne sont pas surprenantes. Les trop fortes statistiques pour les carrés des résidus traduisent (au moins en partie) la saisonnalité de la variance (voir paragraphe suivant). Et le caractère douteux de la normalité traduit la dissymétrie des distributions de température.

**Conclusion.** L'ajustement d'un modèle linéaire ARMA conduit à sélectionner un processus autorégressif d'ordre 3, AR(3). Toutefois les réserves mentionnées ci-dessus à propos de la validité de ce modèle, montrent que l'ajustement n'est pas de bonne qualité.

### III. Modèle linéaire à variance périodique.

#### 1. Le modèle.

Nous voulons corriger le modèle précédent de type ARMA de façon à tenir compte de la saisonnalité de la variance. Aussi proposons-nous le modèle suivant :

#### Modèle linéaire à variance périodique (VP).

$$X_t = m_t + s_t + \rho_t \cdot Z_t$$

avec :

1.  $m_t = d \cdot t + e$ , représentant la tendance, supposée linéaire ;
2.  $s_t$ , représentant la saisonnalité, périodique de périodicité annuelle et de moyenne nulle sur une période ;
3.  $\rho_t = a + b \cdot \cos(\omega t) + c \cdot \sin(\omega t)$  avec  $\omega = \frac{2\pi}{365}$  représentant l'écart-type de la série. On ne le suppose plus constant, mais périodique de périodicité annuelle ; On impose la condition  $\rho_t > 0$ .
4.  $Z_t$  est un processus de type ARMA gaussien, avec  $\text{var } Z_t = 1$ .

On notera  $(\varepsilon_t)$  le bruit correspondant.



Ce modèle est très proche de celui proposé par M. Cao et J. Wei (Cao, Wei, 1998). Une petite différence se situe au niveau de l'expression de  $\rho_t$  qui est parfaitement lisse ici. D'autre part, nous ne supposons pas *a priori* que  $(Z_t)$  est un processus autorégressif pur. Nous allons ici plus loin dans l'étude du modèle en examinant sa *validité* et en la comparant avec celle du modèle (AR).

Tel qu'il est écrit, le modèle (VP) n'est pas complètement paramétrique : aucune forme n'est imposée pour la composante de saisonnalité  $s_t$ . Pour les objectifs que nous voulons atteindre, cela n'est pas véritablement gênant et nous nous contenterons ici de cette approche "traditionnelle". Cependant, l'absence de paramétrisation de  $s_t$  présente deux inconvénients. *Primo*, l'estimation naturelle de  $s_t$  par la moyenne historique au jour  $t$  n'est pas lisse : ainsi  $\hat{s}_2 = -9.7^\circ C$ , tandis que  $\hat{s}_3 = -10.8^\circ C$  ; autrement dit les températures historiques entre le 2 janvier et le 3 janvier diffèrent de plus d'un degré. *Secundo*, l'estimation des paramètres du modèle (VP) dépend de l'estimation faite au préalable pour la composante de saisonnalité. Observant la régularité assez exceptionnelle de la saisonnalité, il semble possible de "capter" l'essentiel du comportement saisonnier avec un petit nombre de paramètres. Une analyse spectrale fait apparaître en plus de la périodicité annuelle, une périodicité semestrielle (et, dans une moindre mesure, une périodicité d'un cinquième d'année). Il est alors naturel de proposer la paramétrisation de la composante de saisonnalité sous la forme d'une somme de sinusoides :

$$s_t = a_1 \cdot \cos(\omega t) + b_1 \cdot \sin(\omega t) + a_2 \cdot \cos(2\omega t) + b_2 \cdot \sin(2\omega t)$$

L'ajustement est, au moins à l'œil, satisfaisant. Et en reprenant l'étude qui suit avec cette forme paramétrique de  $s_t$ , on observe des résultats tout à fait comparables ce qui semble justifier *a posteriori* l'utilisation de ces quatre paramètres supplémentaires.

## 2. Estimation préliminaire des paramètres - Sélection du modèle ARMA.

Dans un premier temps, nous utilisons l'estimation de  $d$  et  $e$  obtenue précédemment. Pour estimer  $\rho_t$ , on remarque que  $\text{var}(Y_t) = \rho_t^2$ , où  $Y_t = X_t - m_t - s_t$ . On peut alors estimer  $\rho_t^2$  par une variance mobile lissée, centrée en  $t$  (voir (Jones, Wand, 1995)). La courbe  $t \rightarrow \hat{\rho}_t$  est

représentée sur la Figure 1.7. Les valeurs de  $\hat{a}$ ,  $\hat{b}$ ,  $\hat{c}$  sont alors obtenues par simple régression linéaire de  $\hat{\rho}_t$  sur  $1, \cos(\omega t), \sin(\omega t)$ . On obtient :

$$(\hat{a}, \hat{b}, \hat{c}) = (3.30, 0.51, 0.07)$$

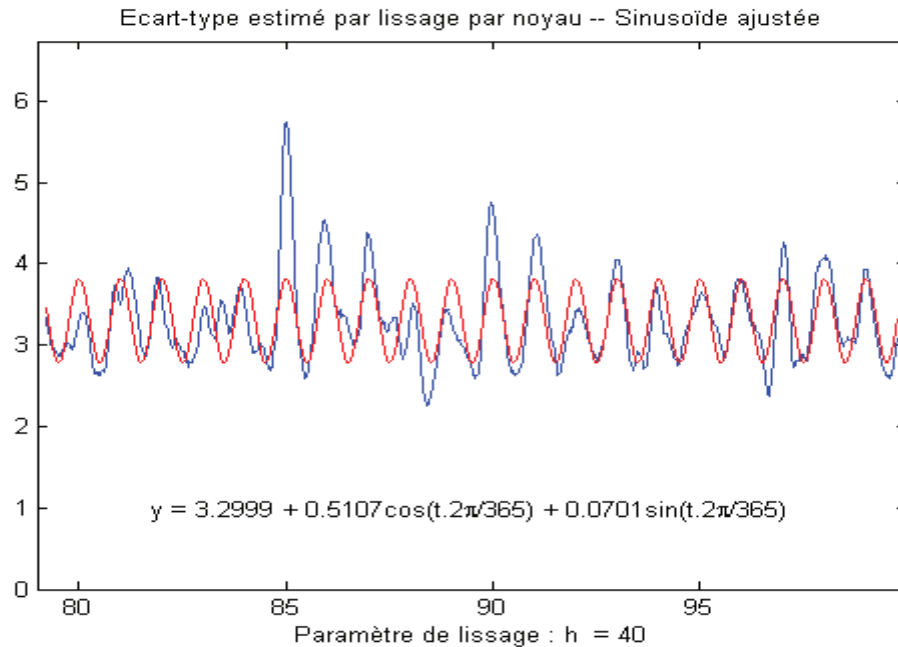


Figure 1.7. Estimation préliminaire de la volatilité de la température.

Notons que la sinusoïde ajustée garde des valeurs strictement positives. D'autre part, on constate visuellement *a posteriori* que le choix  $\omega = \frac{2\pi}{365}$  est justifié.

Connaissant les estimations de  $a, b, c, d, e$ , on peut calculer  $\hat{z}_t$ . Les techniques d'ajustement d'un modèle linéaire pour  $Z_t$  sont alors mises en œuvre de la même façon que l'on a fait pour  $Y_t$ . Elles conduisent au choix peu surprenant  $p = 3, q = 0$  et à l'estimation des paramètres du processus AR(3) ainsi déterminé. Numériquement, on obtient les estimations suivantes<sup>52</sup> :

$$(\hat{\phi}_1, \hat{\phi}_2, \hat{\phi}_3) = (0.86, -0.14, 0.04)$$

<sup>52</sup> Il n'est pas illogique de trouver quasiment les mêmes estimations que pour le processus  $(Y_t)$  du modèle précédent. En effet, à supposer que  $Z_t$  vérifie la relation  $Z_t = \phi_1.Z_{t-1} + \dots + \phi_p.Z_{t-p} + \varepsilon_t$ , alors  $Y_t$ , égal à  $\rho_t.Z_t$ , vérifiera *approximativement* la relation  $Y_t = \phi_1.Y_{t-1} + \dots + \phi_p.Y_{t-p} + \rho_t.\varepsilon_t$ , car  $\rho_t$  est *localement* quasiment constant.

### 3. Estimation plus fine des paramètres.

L'estimation des paramètres est réalisée en maximisant la vraisemblance (le développement mathématique est traité dans l'annexe A). On utilise les estimations préliminaires comme condition initiale d'un algorithme d'optimisation. A chaque étape, on contrôle que la condition de positivité de  $\rho_t$  est satisfaite.

### 4. Validité du modèle.

Le modèle à variance périodique suppose que les résidus  $\varepsilon_t$  constituent des réalisations indépendantes d'une loi normale. Peut-on mettre en défaut cette hypothèse ? L'analyse des résidus est résumée sur le Tableau 1.2. Nous constatons clairement une amélioration au niveau du test de non-corrélation linéaire de Ljung-Box : les statistiques des carrés des résidus sont correctes. Les statistiques des autres tests sont sensiblement les mêmes que pour le modèle (AR). Et le tracé de la droite de Henri (Figure 1.8) fait apparaître la même dissymétrie observée jusqu'à présent. Finalement l'amélioration par rapport au modèle linéaire AR(3) se situe au niveau de l'indépendance des résidus. Les tests effectués ne permettent pas de rejeter le modèle au niveau 3% pour la série étudiée<sup>53</sup>.

### 5. Impact du modèle sur la variance.

Au paragraphe 1.2., nous avons constaté un phénomène de saisonnalité pour la dispersion de la série désaisonnalisée  $Y_t$ . Le modèle proposé, à variance non constante, devrait permettre de diminuer l'importance du phénomène. Le boxplot obtenu à partir de  $Z_t$  (voir paragraphe : "Problèmes de saisonnalité", ci-avant) est représenté sur la Figure 1.9. On constate visuellement une nette amélioration : l'impression de "creux" que l'on avait sur la Figure 1.3 a quasiment disparu. Les intervalles interquartiles ne présentent plus une variation saisonnière. En revanche, la dissymétrie des mois de janvier et février subsiste.

---

<sup>53</sup> Les améliorations ne sont pas aussi nettes pour d'autres villes de France : les statistiques des carrés des résidus du modèle (VP) sont meilleures par rapport au modèle (AR), mais restent toutefois trop grandes pour être acceptables.

Nombre de paramètres estimés du modèle : 8

Test	Résidus			Carrés des résidus	
	Largeur	Stat.	p-valeur	Stat.	p-valeur
Ljung-Box	21	15.0286	0.3056	21.4338	0.4327
	42	43.1062	0.1361	57.2561	0.0584
	63	63.3494	0.2056	71.3749	0.2194
	84	82.8517	0.2764	88.125	0.3578
Kendall			0.0490		
Kolmogorov		0.0147			

Tableau 1.2. Analyse des résidus -- Modèle linéaire à variance périodique (VP).

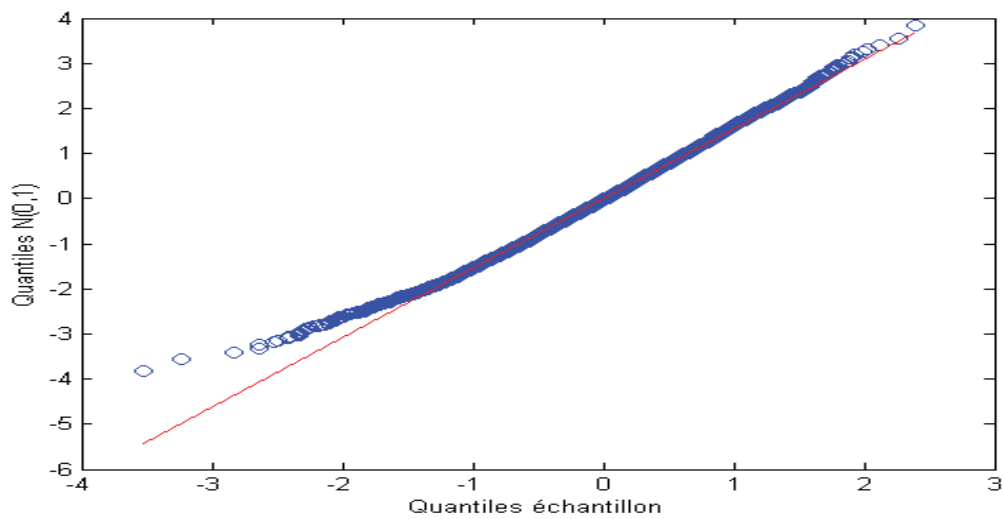


Figure 1.8. Droite de Henri des résidus -- Modèle (VP).

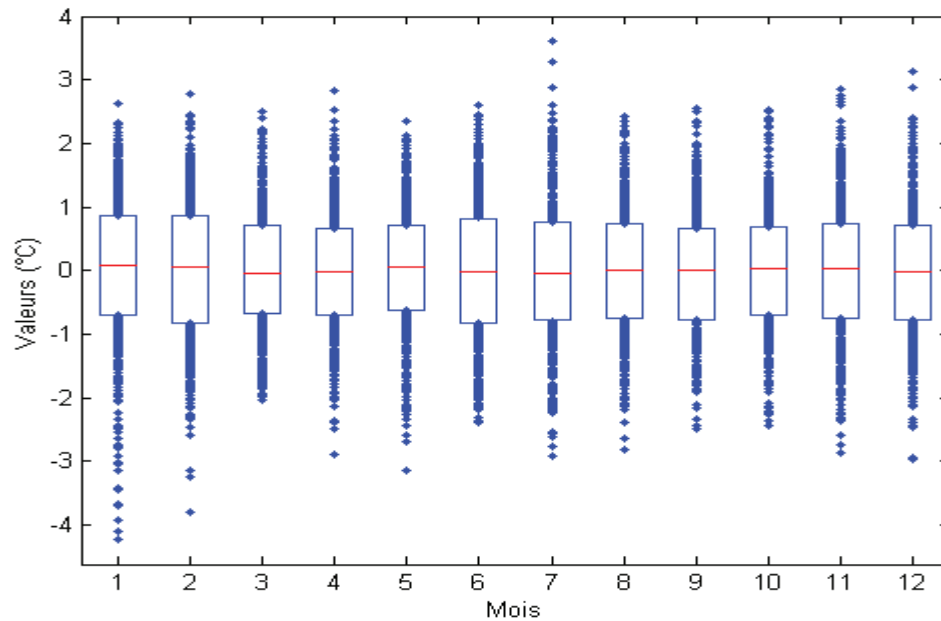


Figure 1.9. Comparaison des distributions de température mois par mois. Série sans tendance et désaisonnalisée, avec variance périodique.

## 6. Gaussien ou non gaussien ?

Les boxplots ont montré une dissymétrie des distributions hivernales. Nous avons vu qu'elle est conservée sur les résidus des modèles (AR) et (VP). Néanmoins le test classique d'adéquation à la loi normale ne permet pas de rejeter le caractère gaussien au niveau 5%. D'un autre côté, cela invite à penser que la dissymétrie n'est pas très importante. Il n'est donc pas impensable d'accepter l'hypothèse de normalité en première approximation. Bien qu'insatisfaisante, on peut raisonnablement penser que les résultats obtenus sous cette hypothèse ne seront pas radicalement différents. En revanche il ne semble pas indiqué de conserver cette hypothèse pour obtenir des résultats précis.

Comme notre étude cherche à donner des ordres de grandeur, nous supposons par la suite que les processus considérés sont gaussiens.

**Conclusion de la première partie.**

Motivée par le développement récent des produits dérivés climatiques, la première partie de l'étude a permis de dégager des caractéristiques essentielles de la dynamique de la température et de faire une première comparaison de deux modèles économétriques. Outre les composantes déterministes évidentes (termes de tendance et de saisonnalité), on observe une dispersion plus importante des températures en hiver qu'en été. La prise en compte de ce phénomène permet d'aboutir à une meilleure stationnarisation des données<sup>54</sup>, qui peuvent alors être modélisées par un processus autorégressif. Ces résultats ont été obtenus à partir d'une série de températures journalières mesurées à la station météorologique de Lyon-Bron, mais s'observent également pour les 10 autres stations disséminées sur l'hexagone que nous avons regardées<sup>55</sup>. Le modèle "(VP)", de type ARMA à variance périodique, semble ainsi particulièrement adapté pour décrire la dynamique de la température. A la vue des tests statistiques réalisés<sup>56</sup>, il paraît être un premier modèle assez satisfaisant. Dans la deuxième partie de l'article, nous l'utilisons pour déduire une application à la gestion des risques climatiques.

*Fin de la première partie.*



<sup>54</sup> Voir notamment les figures 3 et 11.

<sup>55</sup> Lille, Strasbourg, Brest, Rennes, Paris (Paris-Montsouris), Clermont-Ferrand, Grenoble, Bordeaux, Toulouse, Nice. La dynamique retenue est, en général, de type  $AR(p)$ , avec  $p = 2$  ou  $3$ .

<sup>56</sup> Et en dépit de la dissymétrie des distributions hivernales (voir les figures 3 et 7 ainsi que le paragraphe "Gaussien ou non gaussien").

## Annexe A. Vraisemblance du modèle (VP).

Notons  $\sigma^2$  la variance de  $\varepsilon_t$ . Le lien théorique entre la variance du bruit  $\varepsilon_t$  et la variance du processus  $Z_t$  est donné par (Brockwell, Davis), §3.3., formule 3.3.1. :

$$\text{var } Z_t = \sigma^2 \times \sum_{j=0}^{+\infty} \psi_j^2$$

où  $\sum_{j=0}^{+\infty} \psi_j z^j$  est le développement formel de  $(1 - \phi_1 z - \phi_2 z^2 - \phi_3 z^3)^{-1}$ .

$\sigma$  est donc une fonction de  $\phi_1, \phi_2, \phi_3$ , donné par :

$$\sigma = f(\phi_1, \phi_2, \phi_3) = \left( \sum_{j=0}^{+\infty} \psi_j^2 \right)^{-1/2}$$

Posons  $\Theta = (a, b, c, d, e, \phi_1, \phi_2, \phi_3)$ . D'après (Brockwell, Davis), §8.7., la vraisemblance du processus AR(3),  $Z_t$ , conditionnellement à  $Z_0$ , est donnée par :

$$L_Z(Z_1, Z_2, \dots, Z_n) = (2\pi\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2} \sum_{t=1}^n \varepsilon_t^2\right)$$

Par ailleurs, un changement de variables permet d'obtenir

$$L(X_1, X_2, \dots, X_n, \Theta) = \frac{1}{\rho_1 \rho_2 \dots \rho_n} L_Z(Z_1, Z_2, \dots, Z_n)$$

d'où finalement :

$$L(X_1, X_2, \dots, X_n, \Theta) = \frac{1}{\sigma \rho_1 \cdot \sigma \rho_2 \dots \sigma \rho_n} (2\pi)^{-n/2} \times \exp\left(-\frac{1}{2} \sum_{t=1}^n \left( \frac{X_t - (m_t + s_t)}{\sigma \rho_t} - \sum_{i=1}^3 \phi_i \cdot \frac{X_{t-i} - (m_{t-i} + s_{t-i})}{\sigma \rho_{t-i}} \right)^2\right)$$

En remarquant qu'il s'agit d'une fonction  $L'$  des variables  $a, \sigma, b, \sigma, c, \sigma, d, e, \phi_1, \phi_2, \phi_3$  avec  $\sigma = f(\phi_1, \phi_2, \phi_3)$ , on en déduit que  $\Theta$  maximise  $L$  ssi  $\Theta' := (a', b', c', d', e', \phi_1', \phi_2', \phi_3')$  maximise  $L'$ , avec la relation :

$$\Theta := \left( \frac{a'}{f(\phi_1', \phi_2', \phi_3')}, \frac{b'}{f(\phi_1', \phi_2', \phi_3')}, \frac{c'}{f(\phi_1', \phi_2', \phi_3')}, d', e', \phi_1', \phi_2', \phi_3' \right)$$

**Application numérique.**

L'optimisation de  $L'$  donne :

$$\Theta' = (2.1299, 0.2573, 0.0200, 0.0002, 11.3684, 0.8618, -0.1407, 0.0418).$$

On obtient  $f(\phi'_1, \phi'_2, \phi'_3) = 0.6267$ , d'où finalement l'estimation :

$$\Theta = (3.3985, 0.4106, 0.0319, 0.0002, 11.3684, 0.8618, -0.1407, 0.0418).$$



## Une application de deux modèles économétriques de température à la gestion des risques climatiques (2<sup>ème</sup> partie)

### Introduction de la 2<sup>ème</sup> partie.

Les dérivés climatiques sont apparus en 1997 aux Etats-Unis, parallèlement à la déréglementation du marché de l'énergie pour répondre aux besoins de protection des producteurs d'énergie face aux risques climatiques; un hiver doux, par exemple, est synonyme d'importants manques à gagner. Depuis le marché s'est largement développé avec l'apparition de marchés réglementés - Chicago Mercantile Exchange<sup>57</sup> (CME), London International Financial Futures and Options Exchange<sup>58</sup> (LIFFE) - et diversifié avec l'apparition de nouveaux acteurs issus du secteur du tourisme, de la distribution, etc. Les contrats standard du CME concernent le risque de température et sont constitués de contrats futures et d'options sur futures sur les indices mensuels CDD et HDD<sup>59</sup>. Les périodes d'exposition au risque sont des mois et s'étalent jusqu'à un horizon d'un an : au 15 octobre 2001 par exemple, on pouvait souscrire 7 contrats futures HDD différents, un pour chaque mois "d'hiver" de la période Novembre 2001 à Avril 2002 inclus, et 7 contrats futures CDD, un pour chaque mois "d'été" de la période d'Avril 2002 à Octobre 2002 inclus.

A l'heure actuelle, la température est le support le plus répandu des dérivés climatiques et sa modélisation est naturellement un des thèmes de recherche. Dans les trois premières sections de l'article (voir le précédent numéro de *Banque & Marchés*), nous avons étudié les caractéristiques essentielles de la température ainsi que la validité de deux modèles économétriques linéaires classiques, à partir d'une série de 21 années de températures journalières de la station de Lyon-Bron :

– un modèle, noté (AR), du type:

$$X_t = m_t + s_t + Y_t$$

---

<sup>57</sup> Site Internet : [www.cme.com](http://www.cme.com)

<sup>58</sup> Site Internet : [www.i-wex.com](http://www.i-wex.com)

<sup>59</sup> L'indice CDD (Cooling Degree Day), mesuré en degrés, est la somme des "excédents" de température par rapport à un seuil sur la période couverte :  $CDD = \sum_t \max(T(t) - \text{seuil}; 0)$ , où  $T(t)$  désigne la température moyenne du jour  $t$ . Inversement, l'indice HDD (Heating Degree Day) est donné par :  $HDD = \sum_t \max(\text{seuil} - T(t); 0)$ .

où  $m_t$ ,  $s_t$  désignent respectivement un terme de tendance (supposée linéaire) et de saisonnalité, et  $Y_t$  est un processus ARMA :

$$Y_t = \phi_1.Y_{t-1} + \dots + \phi_p.Y_{t-p} + \varepsilon_t + \theta_1.\varepsilon_{t-1} + \dots + \theta_q.\varepsilon_{t-q}$$

avec  $(\varepsilon_t)$  un bruit blanc gaussien de variance  $\sigma^2$ .

– un modèle, noté (VP), à "variance périodique":

$$X_t = m_t + s_t + \rho_t.Z_t$$

où  $\rho_t = a + b.\cos(\omega t) + c.\sin(\omega t)$  avec  $\omega = \frac{2\pi}{365}$  représente l'écart-type de la série, ( $\rho_t > 0$ ), et

$Z_t$  est un processus de type ARMA gaussien, avec  $\text{var } Z_t = 1$ .

La sélection du modèle ARMA a conduit à choisir un modèle autorégressif pur d'ordre 3 dans les deux cas. D'autre part, l'ajout du terme  $\rho_t$  permet de modéliser le phénomène de dispersion saisonnière (dispersion plus importante en hiver) de la température. Ce n'est donc pas une surprise que la validation statistique soit plus satisfaisante pour le modèle (VP).

On aurait pu penser utiliser les prévisions de température fournies par Météo France au lieu de modèles économétriques. Cependant que ce soit dans le cadre d'une valorisation ou d'une analyse des risques climatiques comme c'est exclusivement le cas ici, *c'est toute la dynamique de la température qui est importante*, et pas seulement les moments conditionnels d'ordre 1. En effet, supposons fixée une période d'exposition au risque de température et supposons connue une fonction "d'endommagement" exprimant les pertes par une fonction des températures sur toute la période ; on peut alors chercher à connaître à la date courante l'espérance des pertes, la valeur à risque concernant les pertes... et l'on est donc amené à étudier la *loi conditionnelle des températures* sur la période d'exposition au risque. Par ailleurs, les horizons de gestion considérés sont souvent de plusieurs mois et sont donc très éloignés de ceux des prévisions fournies par Météo France.

Le but de cet article est double. *Primo*, il s'agit de **comparer les deux modèles de température (AR) et (VP)**. *Secundo*, on se doute que le rapprochement entre la date courante et la période d'exposition au risque peut jouer un rôle. Ainsi si l'on se place au 31 octobre 2001, la connaissance de la température courante n'aura qu'un intérêt limité par rapport à un risque de température situé en avril 2002, tandis qu'elle pourra se révéler utile pour un risque localisé au mois de novembre 2001. D'une façon générale, nous cherchons à **évaluer l'impact de la**

***situation courante sur la gestion des risques climatiques.*** Le problème est lié à la rapidité avec laquelle les températures "reviennent à la moyenne". Les modèles économétriques étudiés nous permettent de quantifier ces effets. Plus précisément si l'on note  $t$  la date courante,  $[t+h; t+u]$ ,  $u > h$ , la période d'exposition au risque, la température à la date  $t$ ,  $L_t(X_{t+h}, \dots, X_{t+u})$  la loi conditionnelle (à l'information courante) des températures futures et  $L(X_{t+h}, \dots, X_{t+u})$  la loi non conditionnelle, dans quelle mesure peut-on remplacer  $L_t$  par la loi plus simple  $L$  ?

Nous présentons ici les sections 4 à 6 de l'article. La quatrième section étudie les prévisions théoriques en détail. En particulier nous quantifions l'horizon à partir duquel on peut remplacer la loi conditionnelle  $L_t$  par la loi non conditionnelle  $L$ . La cinquième section s'attache à regarder l'impact d'un "choc" extérieur. Une notion de réponse impulsionnelle adaptée au cadre des dérivés climatiques est introduite à cet effet. Enfin dans la sixième section, les erreurs théoriques de prévision sont comparées avec les erreurs réellement observées ("back-testing"). Cela donne un nouveau critère pour juger de la qualité de chacun des modèles. Les développements mathématiques figurent en annexe.

#### **IV. Performances des modèles. Application à la gestion des risques climatiques.**

Comme il a été dit en introduction, la gestion des risques climatiques requiert la connaissance de la loi conditionnelle du vecteur des températures sur la période d'exposition au risque  $(X_u, X_{u+1}, \dots, X_v)$  sachant l'information  $X_t$  connue à la date courante  $t$ , ce qui, mathématiquement, correspond à la notion de prévision. Dans le cas gaussien auquel nous nous limitons ici, il suffit d'étudier l'espérance et la variance conditionnelles. Dans ce paragraphe nous verrons que si la période d'exposition au risque débute suffisamment longtemps après la date courante, les modèles étudiés sont peu "performants" c'est-à-dire que la prévision ne fera quasiment pas appel à la situation courante. Dans ce cas la gestion des risques climatiques pourra se faire sans se soucier de la situation courante, aussi exceptionnelle soit-elle.

Avant de considérer le cas général on s'intéressera d'abord à une seule date et l'on étudiera la loi de  $X_s$  conditionnellement à  $X_t$  avec  $u \leq s \leq v$ . Outre sa simplicité technique, cette étude

permet de traiter le cas où la fonction d'endommagement qui exprime les pertes comme fonction de  $X_u, X_{u+1}, \dots, X_v$  est la somme de fonctions d'endommagement "ponctuelles", chacune d'elles ayant comme seul argument l'un des  $X_s$ ,  $u \leq s \leq v$ .

Dans ce cadre la faible performance des modèles pour un horizon de prévision suffisamment grand est assez naturelle. Si l'on s'intéresse par exemple à l'espérance conditionnelle  $E(X_u | X_t)$  et que l'on demande à une personne n'ayant aucune compétence météorologique particulière de prévoir à la date  $t$  la température dans un mois (pour la date  $u$ ), il est probable qu'elle raisonnera comme suit. Pour fixer les idées supposons que nous soyons le 1<sup>er</sup> janvier et que le défi consiste à prévoir la température pour le 1<sup>er</sup> février. Elle commencera vraisemblablement par faire l'inventaire des températures observées le 1<sup>er</sup> février au cours des années précédentes. Elle fera ensuite la moyenne, ce qui lui donnera une première approximation. Eventuellement, s'étant tenue informée des changements climatiques récents, elle prendra en compte le réchauffement observé depuis quelques années.

Cette prévision, que nous qualifierons de "naïve", peut néanmoins être une bonne prévision, voire la seule que l'on puisse faire, dans certains cas.

En revanche, imaginons qu'au lieu du 1<sup>er</sup> février, on demande une prévision pour le lendemain. Imaginons en plus qu'au 1<sup>er</sup> janvier, la température est de  $-14^\circ\text{C}$ , et que depuis une semaine la température est exceptionnellement froide. En appliquant le raisonnement précédent, on trouve que la moyenne historique des températures observées le 2 janvier est de  $2.4^\circ\text{C}$ , soit plus de  $16^\circ\text{C}$  d'écart par rapport à la veille !

Bien que possible, cette prévision n'est pas très *probable*, et d'ailleurs, la personne interrogée aurait très certainement modifié son raisonnement. Les modèles (AR) et (VP) agissent en ce sens : les prévisions qu'ils donnent sont réajustées en fonction de la situation présente et d'un passé proche dans notre cas. On conçoit donc qu'ils soient particulièrement efficaces dans deux situations :

- Lorsque la température présente et/ou des jours précédents s'éloigne très nettement de la température observée en moyenne en ce jour ;
- Lorsque la prévision est une prévision à court terme. En effet, l'impact de la température du jour et des jours passés ne se fait ressentir qu'à court terme.

Au contraire la prévision naïve sera satisfaisante dans les cas suivants :

- Lorsque la température présente et/ou des jours précédents est proche de la température observée en moyenne en ce jour ;

– Lorsque la prévision est une prévision à long ou moyen terme.

On doit donc pouvoir déterminer un nombre de jours  $h$  à partir duquel la situation présente n'a plus d'influence sur la prévision à l'horizon  $h$ , et dans cette section nous le quantifions.

La situation particulière où l'on se limite à une seule date englobe des cas où les risques climatiques se décomposent en risques ponctuels ne dépendant que de la température en une seule date. Or certains risques peuvent être liés à une variation brutale de température d'un jour à l'autre ou à la température de trois jours consécutifs, etc. Il est donc indispensable d'étudier en toute généralité la loi conditionnelle du vecteur des températures sur la période d'exposition au risque. Nous verrons que la conclusion reste la même et qu'à partir de l'horizon qui a été déterminé, on peut considérer que la loi conditionnelle est approximativement la même que la loi non conditionnelle.

Pratiquement, au-delà de cet horizon, la situation courante n'aura plus aucune influence dans la gestion des risques climatiques. Il faut toutefois prendre garde aux corrélations qui subsistent dans la loi non conditionnelle en raison de la dynamique des températures.

Les quatre paragraphes à venir concernent le cas particulier où l'on se limite à une seule date. Le cas général est traité au cinquième paragraphe.

### 1. Prévision. Erreur théorique de prévision.

#### Définitions.

1. Pour un processus  $X_t$ , la prévision en  $t$  à l'horizon  $h$ , notée  $P_t X_{t+h}$ , est l'espérance conditionnelle de  $X_{t+h}$  sachant  $X_1, X_2, \dots, X_t$  :

$$P_t X_{t+h} = E(X_{t+h} | \underline{X}_t)$$

L'erreur de prévision est  $e_t(X_{t+h}) = X_{t+h} - P_t X_{t+h}$ . La variance conditionnelle de  $e_t(X_{t+h})$  (ou de  $X_{t+h}$ ) sera notée  $V_t X_{t+h}$  :

$$V_t X_{t+h} = E((X_{t+h} - P_t X_{t+h})^2 | \underline{X}_t)$$

2. La prévision en l'absence d'observations en  $t$  à l'horizon  $h$ , notée  $PX_{t+h}$ , est l'espérance (non conditionnelle) de  $X_{t+h}$  :

$$PX_{t+h} = E(X_{t+h})$$

Sa variance sera notée  $VX_{t+h}$ .

Dans un premier temps, on s'attache à comparer  $P_t X_{t+h}$  à  $PX_{t+h}$  et  $V_t X_{t+h}$  à  $VX_{t+h}$  pour différentes valeurs de  $h$ .

Ensuite on généralise en comparant l'espérance et la variance conditionnelles avec l'espérance et la variance non conditionnelles du vecteur  $(X_u, X_{u+1}, \dots, X_v)$  avec les notations du début du paragraphe.

## 2. Prédiction naïve. Modèle "naïf" (N).

La prédiction naïve est obtenue à partir d'un modèle du type suivant :

$$X_t = m_t + s_t + Y_t$$

où l'on demande au processus  $Y_t$  d'être au moins centré et stationnaire.

Elle correspond alors à la prédiction en l'absence d'observations  $m_{t+h} + s_{t+h}$ . Il s'agit de la composante déterministe de la série de températures ; on vérifie immédiatement qu'elle est égale à la moyenne historique des températures avec un terme correctif lié à la tendance.

Pour de petites valeurs de  $h$ , cette prédiction est en général "grossière". En revanche, on conçoit que pour de plus grandes valeurs de  $h$ , la température à l'instant  $t+h$  est approximativement indépendante des températures aux instants  $t, t-1, \dots$ . Autrement dit, on a asymptotiquement :  $P_t X_{t+h} \approx PX_{t+h}$ .

La prédiction dans le modèle (N) devra être prise au sens de la prédiction en l'absence d'observations. En résumé, on a le résultat suivant :

### Prévisions et erreurs de prédiction du modèle (N).

1.  $PX_{t+h} = m_{t+h} + s_{t+h}$ .
2. La variance de l'erreur de prédiction est constante et vaut  $\text{var } Y_t$  (indépendant de  $t$  et de  $h$ ).

## 3. Cas des modèles AR et (VP).

Les calculs des prévisions et des erreurs de prédiction de ces modèles découlent directement des résultats sur les processus  $AR(p)$ . Ils figurent en annexe B. La Figure 1.10 montre les intervalles de prédiction des trois modèles, en situation hivernale en fonction de l'horizon  $h$  de prédiction ( $h=0$  correspond au jour où est faite la prédiction). La courbe correspondant à

$h < 0$  représente les températures des 20 jours précédents. En réalité seules les températures correspondantes à  $h = -2, -1, 0$  sont prises en compte dans les prévisions ; les autres sont données à titre indicatif. La Figure 1.11 est une représentation similaire en se plaçant cette fois en situation estivale.

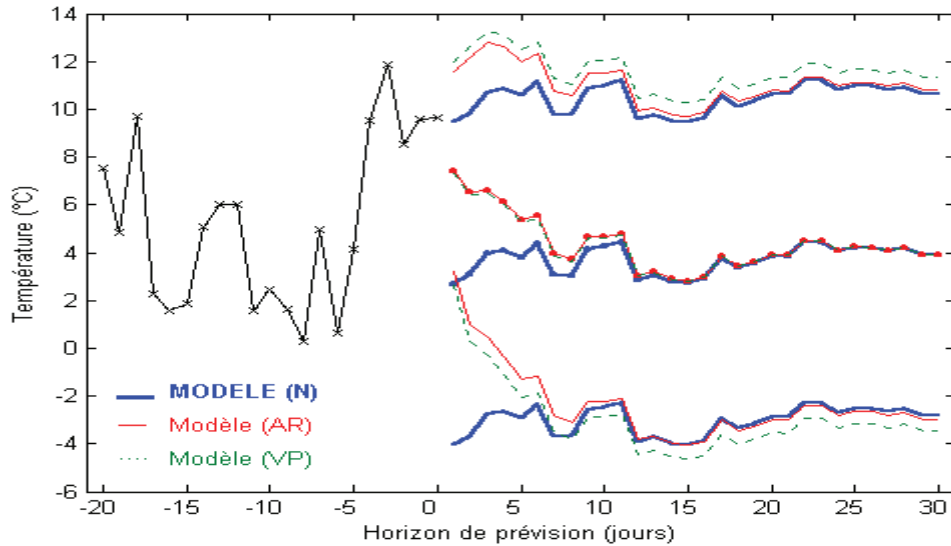


Figure 1.10. Prévisions et intervalles de prévision en situation hivernale.

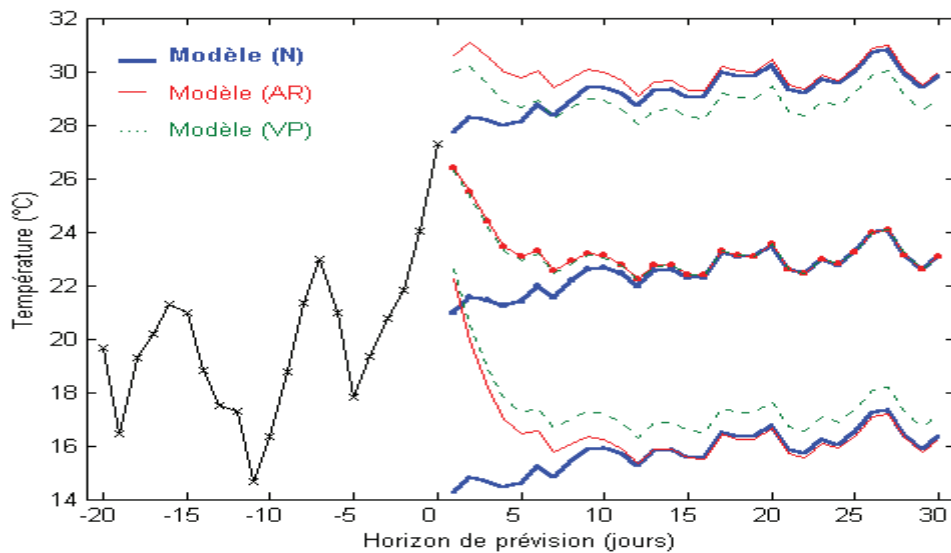


Figure 1.11. Prévisions et intervalles de prévision en situation estivale.

Nous pouvons faire les conclusions suivantes :

**Conjonction des prévisions.** On constate que les prévisions des trois modèles se confondent pour un horizon supérieur à 15 jours environ<sup>60</sup>.

**Incertitudes des prévisions.** Par ailleurs, à partir d'un tel horizon, l'utilisation du modèle (AR) ne permet pas d'améliorer l'incertitude par rapport à la prévision naïve : les intervalles de prévision sont également confondus.

Le modèle (VP) tient compte de la saisonnalité de la dispersion des températures ce qui explique la différence des intervalles de prévision selon la saison. Cela constitue une amélioration comme nous le verrons en étudiant les erreurs réelles. Toutefois, l'amélioration n'est pas meilleure au-delà de 15 jours lorsqu'on fait la comparaison avec un modèle élémentaire avec variance saisonnière : on retrouve les mêmes intervalles lorsqu'on impose seulement que l'écart-type du processus de température soit égal à  $\rho_t$  sans utiliser la forme autorégressive du processus.

En fait les données de la Figure 1.10 proviennent du 31 décembre 1998 et des 20 jours précédents, tandis que celles de la Figure 1.11 proviennent du 3 juillet 1999 et des 20 jours précédents.

On peut considérer ces exemples comme des réalisations particulières de l'espérance et de la variance conditionnelles de  $X_{t+h}$  pour  $h=1,2,\dots,30$  et pour les deux modèles. Les conclusions précédentes signifient qu'elles sont peu différentes de l'espérance et la variance (non conditionnelles) lorsque  $h > 15$ .

#### 4. Conjonction des prévisions.

Nous allons maintenant quantifier avec précision l'horizon à partir duquel l'écart entre les prévisions des modèles (AR) et (VP) d'une part, et le modèle (N) d'autre part est inférieur à une valeur fixée.

Fixons un horizon  $h$ . La quantité  $P_t Y_{t+h}$  mesure l'écart entre la prévision du modèle (AR) et la prévision naïve à l'horizon  $h$ . S'agissant d'une variable aléatoire centrée, on peut mesurer son amplitude par exemple à l'aide d'un quantile  $q_{t,h}$  tel que :

<sup>60</sup> Les prévisions des modèles (AR) et (VP) se confondent dès le lendemain. Cela tient au fait que pour de petits horizons,  $\rho_{t+h}$  est voisin de  $\rho_t$ .



$$P(|P_t Y_{t+h}| \leq q_{t,h}) = \alpha$$

Dans la suite, nous ferons l'hypothèse que  $(Y_t)$  est un processus *gaussien*. Nous avons discuté au paragraphe "Gaussien ou non gaussien ?" de la qualité de cette hypothèse : le calcul effectué ici constitue un premier pas.

Dans ce cas,  $P_t Y_{t+h}$  est une variable aléatoire normale centrée, donc le quantile  $q_{t,h}$  est une fonction simple de la variance  $\sigma_{t,h}^2 = \text{var } P_t Y_{t+h}$ .

On vérifie que

$$\begin{aligned} \sigma_{t,h}^2 &= \text{var } Y_{t+h} - \text{var}(Y_{t+h} - P_t Y_{t+h}) \\ &= \gamma(0)^2 - \sigma^2 \sum_{j=0}^{h-1} \psi_j^2 = \sigma^2 \sum_{j=h}^{+\infty} \psi_j^2 \end{aligned}$$

On constate que cette quantité décroît en fonction de  $h$  : on pourra donc, par exemple, donner une valeur de  $h$  à partir de laquelle l'intervalle  $[-0.5^\circ\text{C}; 0.5^\circ\text{C}]$  contient l'écart  $P_t Y_{t+h}$  avec une probabilité de 95%.

Cependant, la quantité  $\sigma_{t,h}^2$  est estimée à partir des données historiques et il faut tenir compte de l'erreur d'estimation. Celle-ci, sous forme d'intervalle de confiance, peut s'obtenir à partir d'une région de confiance sur les paramètres inconnus  $(\phi_1, \phi_2, \phi_3, \sigma)$  du fait d'une relation fonctionnelle entre  $\sigma_{t,h}^2$  et  $(\phi_1, \phi_2, \phi_3, \sigma)$ . La détermination d'une région de confiance est réalisée dans l'annexe C.

Les résultats numériques obtenus sont donnés dans le tableau suivant:

<i>Horizon</i>	2	4	6	8	10	12	14	16	18	20
$\sigma_h$	2.34	1.46	0.93	0.59	0.38	0.24	0.16	0.10	0.07	0.04

On peut donc en déduire que pour  $h \geq 12$  jours, l'intervalle  $[-0.5^\circ\text{C}; 0.5^\circ\text{C}]$  est un intervalle de prévision à 95% de l'écart entre la prévision du modèle dynamique (AR) et la moyenne historique. Il est obtenu à partir d'une région de confiance à 95% des vrais paramètres.

*On peut donc considérer qu'au-delà de 12 jours, le modèle dynamique (AR) n'apporte pas de plus par rapport à la moyenne historique avec  $0.5^\circ\text{C}$  de tolérance ; Au-delà de 20 jours, le modèle (AR) n'apporte pas de plus avec une tolérance de  $0.1^\circ\text{C}$ .*

### 5. Conjonction des moments conditionnels et non conditionnels.

Dans le cadre de la gestion des risques climatiques, il importe de comparer la loi du vecteur  $(X_u, X_{u+1}, \dots, X_v)$  avec et sans le conditionnement par rapport à  $X_{t_0}$  (où  $t_0$  désigne la date courante) lorsque  $h = u - t_0$  devient "grand". Bien entendu, il est difficile de répondre à cette question en toute généralité. Nous avons choisi de comparer les caractéristiques des lois (gaussiennes) c'est-à-dire les deux premiers moments conditionnels avec les deux premiers moments (non conditionnels). Le moment conditionnel d'ordre 1 est un vecteur constitué de  $v - u + 1$  variables aléatoires. Pour pouvoir le comparer nous utilisons la norme définie par :

$$\|(M_u, \dots, M_v)\|_\infty = \max_{u \leq i \leq v} \|M_i\|_{L^2}$$

où  $\|\cdot\|_{L^2}$  est la norme  $L^2$  usuelle sur l'ensemble des v.a. de carré intégrable.

Pour les moments d'ordre 2, on remarque que la matrice de covariance conditionnelle est constituée de v.a. constantes, si bien qu'on peut se contenter d'utiliser :

$$\|\Sigma\|_\infty = \max_{1 \leq i, j \leq v-u+1} |\Sigma_{i,j}|$$

où  $|\cdot|$  désigne la valeur absolue usuelle.

On a alors les résultats suivants avec les notations propres à chacun des modèles :

**Proposition.** Notons  $A_h = \sigma^2 \sum_{j=h}^{+\infty} \psi_j^2$ .

Si  $(X_t)$  suit le modèle (AR), alors  $A_h = \sigma_{t,h}$  et :

1.  $\left\| E(X_u, \dots, X_v | X_{t_0}) - E(X_u, \dots, X_v) \right\|_\infty = A_h^{1/2}$
2.  $\left\| \text{cov}(X_u, \dots, X_v | X_{t_0}) - \text{cov}(X_u, \dots, X_v) \right\|_\infty = A_h$

Si  $(X_t)$  suit le modèle (VP), et si  $M = \max_{u \leq t \leq v} \rho_t$ ,

1.  $\left\| E(X_u, \dots, X_v | X_{t_0}) - E(X_u, \dots, X_v) \right\|_\infty \leq M \cdot A_h^{1/2}$
2.  $\left\| \text{cov}(X_u, \dots, X_v | X_{t_0}) - \text{cov}(X_u, \dots, X_v) \right\|_\infty \leq M^2 \cdot A_h$

**Démonstration :** voir annexe D.

On constate que la fonction  $\sigma_{t,h}$  rencontrée au paragraphe précédent joue un rôle clef pour calculer les écarts ci-dessus pour le modèle (AR); En appliquant les résultats obtenus à partir d'une région de confiance à 95% des coefficients on en déduit que pour  $h$  supérieur à 20 jours, l'écart (au sens de  $\|\cdot\|_\infty$ ) entre les matrices de covariances conditionnelles et non conditionnelles est inférieur à 0.01. Les résultats précis sont plus durs à obtenir pour le modèle (VP). Néanmoins on peut conjecturer, à la vue des courbes de la Figure 1.10 et de la Figure 1.11, qu'ils ne seront pas très différents.

## V. Réponse du modèle (VP) à un choc.

Jusqu'à présent, nous avons étudié le temps mis par les modèles (AR) et (VP) pour "digérer" la situation courante. De façon analogue, on peut se demander quelle serait la réaction des modèles à un "choc" extérieur, c'est-à-dire par exemple une chute de pluie, l'arrivée d'un anticyclone... dont l'effet est de faire chuter ou grimper brutalement la température. Plus précisément et de façon à faire le lien avec la gestion des risques climatiques, nous étudierons l'effet d'un choc survenu après la date courante  $t_0$  où l'on souhaite effectuer une estimation des risques, sur la loi jointe des températures sur la période d'exposition au risque  $[u, u+1, \dots, v]$ . On pourra savoir par exemple les conséquences sur la loi (et donc sur la gestion des risques) de la possibilité d'anticiper un choc ayant eu lieu le lendemain ou quelques jours après  $t_0$ .

De nombreuses études ont été menées concernant les chocs et les réponses impulsionnelles, pour savoir par exemple si l'on doit considérer les chocs sur le processus lui-même ou ses innovations, et pour définir ce qu'on entend par "réponse". Ici les modèles étudiés sont linéaires et ces questions perdent une partie de leur intérêt mais nous nous sommes quand même référés aux derniers travaux sur le sujet ((Gourieroux, Jasiak, 1998), (Koop, Pesaran, Potter, 1996)) pour préciser les notions de choc et de réponse dans le cadre des dérivés climatiques. Pour la brièveté de l'exposé nous nous limitons au cas du modèle (VP), le modèle (AR) se traitant de la même façon. Le modèle étant linéaire, il est suffisant d'étudier les chocs "transitoires" ne perturbant le modèle qu'en un jour précis. Notons  $s$  la date du choc et supposons que  $t_0 < s < u$ . Les chocs que nous considérons ici affectent la partie aléatoire du

processus c'est-à-dire les innovations  $(\varepsilon_t)$ . Un choc transitoire est défini par la transformation:

$$\varepsilon_t^D = \begin{cases} \varepsilon_t & \text{si } t \neq s \\ a.\varepsilon_s + \delta & \text{si } t = s \end{cases}$$

Un tel choc affecte uniquement l'innovation  $\varepsilon_s$ . Il a été choisi de façon à modifier son espérance comme c'est souvent le cas (rôle du paramètre  $\delta$ ), mais aussi la variance, étant donné le rôle particulier de la dispersion des températures (rôle de  $a$ ). Pour tout processus  $(X_t)$ , on note alors  $(X_t^D)$  le processus perturbé ;  $X_t^D$  est la même v.a. que  $X_t$  avant la date  $s$ , puis se déduit du modèle (VP) avec le bruit  $(\varepsilon_t^D)$  pour  $t \geq s$ .

Comme il s'agit d'un modèle linéaire, il n'est pas indispensable d'utiliser la réponse impulsionnelle généralisée (Koop, Pesaran, Potter, 1996) ou son amélioration faisant appel aux "innovations non linéaires" (Gourieroux, Jasiak, 1998). La réponse "traditionnelle" constituée des v.a.  $E(X_t^D | \underline{X}_{s-1}) - E(X_t | \underline{X}_{s-1})$  et  $\text{var}(X_t^D | \underline{X}_{s-1}) - \text{var}(X_t | \underline{X}_{s-1})$  n'est pourtant pas satisfaisante puisqu'elle ne concerne qu'une loi marginale, et le conditionnement englobe toute l'information avant le choc, même celle intervenant après la date courante. Dans notre contexte il paraît plus naturel de définir la réponse impulsionnelle comme l'ensemble formé du vecteur aléatoire  $\Delta E$  et de la matrice aléatoire  $\Delta V$  avec :

$$\begin{aligned} \Delta E &= E(X_u^D, \dots, X_v^D | \underline{X}_{t_0}) - E(X_u, \dots, X_v | \underline{X}_{t_0}) \\ \Delta V &= \text{cov}(X_u^D, \dots, X_v^D | \underline{X}_{t_0}) - \text{cov}(X_u, \dots, X_v | \underline{X}_{t_0}) \end{aligned}$$

Notons que le conditionnement est inutile car l'écart entre les moments conditionnels et non conditionnels ne fait intervenir que la partie du processus antérieure à  $t_0$  (voir annexe C) et donc antérieure au choc, si bien que, par exemple :

$$E(X_u^D, \dots, X_v^D | \underline{X}_{t_0}) - E(X_u^D, \dots, X_v^D) = E(X_u, \dots, X_v | \underline{X}_{t_0}) - E(X_u, \dots, X_v)$$

On a donc également :

$$\begin{aligned} \Delta E &= E(X_u^D, \dots, X_v^D) - E(X_u, \dots, X_v) \\ \Delta V &= \text{cov}(X_u^D, \dots, X_v^D) - \text{cov}(X_u, \dots, X_v) \end{aligned}$$

En particulier, la réponse pour ce modèle linéaire ne dépend pas de l'historique (cette propriété est connue pour les modèles ARMA avec une réponse "classique").

On a alors le résultat suivant qui découle de calculs élémentaires à partir de la forme  $MA(\infty)$  du processus  $(Z_t)$  :

**Proposition.**

$$\begin{aligned}\Delta E &= (\rho_u \psi_l; \rho_{u+1} \psi_{l+1}; \dots; \rho_v \psi_{l+v-u+1}) \times \delta \\ \Delta V &= (a^2 - 1) \sigma^2 \times M\end{aligned}$$

où  $M$  est la matrice carrée d'ordre  $v - u + 1$  de terme général

$$M_{i,j} = \rho_{u+i-1} \rho_{u+j-1} \psi_{l+i-1} \psi_{l+j-1}$$

et  $l$  est le temps écoulé entre le choc et la période de couverture :  $l = u - s$ .

On observe que la réponse  $(\Delta E; \Delta V)$  dépend de termes du type  $\rho_t \cdot \psi_h$  avec  $h \geq l$ . Ce dernier est constitué du terme saisonnier borné  $\rho_t$  et du coefficient  $\psi_h$  qui tend vers 0 de façon exponentielle :

$$\exists C > 0, \exists \alpha \in ]0; 1[ \text{ t.q. } |\psi_h| \leq C \cdot \alpha^h$$

Le terme  $\rho_t \cdot \psi_h$  peut donc être considéré comme un facteur d'amortissement, et on a  $\|\Delta E\|_\infty \xrightarrow{l \rightarrow +\infty} 0$  et  $\|\Delta V\|_\infty \xrightarrow{l \rightarrow +\infty} 0$ . L'effet du terme saisonnier est vite occulté par la vitesse de convergence de  $\psi_h$  comme on le voit sur la Figure 1.12.

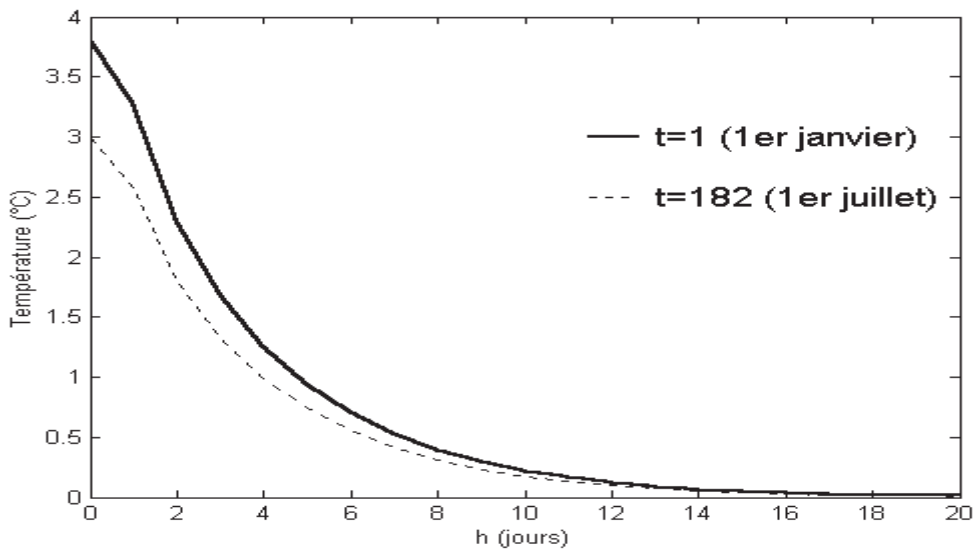


Figure 1.12. Facteur d'amortissement  $\rho_t \cdot \psi_h$  en fonction de  $h$ .

En choisissant par exemple  $\delta = 5^\circ C$  et  $a = 2$ , on a pour  $l \geq 20$  jours,  $\|\Delta E\|_\infty \leq 0.06$  et  $\|\Delta V\|_\infty \leq (0.02)^2$  avec les estimations obtenues dans l'annexe A. A l'inverse des résultats obtenus au paragraphe "Conjonction des prévisions", les écarts ci-dessus n'ont pas de valeur statistique, mais donnent une idée de la réponse en fonction du temps écoulé depuis le choc. Comme on pouvait s'y attendre on constate qu'un choc transitoire même important (comme dans l'exemple ci-dessus par exemple) ayant lieu 2 ou 3 jours après la date courante n'aura quasiment pas d'influence si la période d'exposition au risque commence plus de 20 jours après le choc. Etre capable de l'anticiper n'aura aucune incidence sur la gestion des risques climatiques.

## **VI. Performances réelles.**

L'erreur théorique ne vaut que si les données proviennent du modèle théorique. Or nous avons mis en évidence une mauvaise qualité d'ajustement du modèle (AR) et la  $p$ -valeur du test de Kendall pour le modèle (VP) n'est que de 3%. Il est donc essentiel de comparer les erreurs théoriques avec les erreurs "réelles". Pour les évaluer nous procédons de la façon suivante.

1. On commence par partager les données en deux :
  - Un paquet constitué des 15 premières années de données, à partir duquel l'ajustement du modèle sera réalisé ;
  - Un paquet constitué des données des 6 dernières années, servant de zone de test.
  
2. Une fois le modèle ajusté avec les données du premier paquet, on fixe un horizon  $h$ . On calcule alors pour  $t$  dans la zone de test, la prévision théorique à l'horizon  $h$  donnée par le modèle,  $P_t X_{t+h}$ . L'erreur de prévision réelle mesure l'écart  $x_{t+h} - P_t X_{t+h}$  entre la valeur *réellement* observée  $x_{t+h}$  et la prévision  $P_t X_{t+h}$ .
  
3. De façon à comparer les erreurs moyennes suivant les mois de l'année, on regroupe ensuite les erreurs de prévision pour les valeurs de  $t$  correspondant à un même mois. Ainsi pour  $m = 1, 2, \dots, 12$  on obtient une série d'erreurs associée au  $m^{\text{ième}}$  mois de l'année

$$A_{m,h} = \{x_{t+h} - P_t X_{t+h}, t \in M_m\}$$

(où  $M_m$  désigne l'ensemble des valeurs de  $t$  correspondant au mois  $m$ ), dont on peut calculer l'écart-type  $\sigma_{m,h}$ . Ce nombre, en dépit de la corrélation des variables de  $A_{m,h}$ , est un indicateur de la dispersion des erreurs réelles de prévision faites au mois  $m$ .

4. On peut alors comparer cet écart-type avec la moyenne des écarts-types des erreurs de prévision théorique  $X_{t+h} - P_t X_{t+h}$  lorsque  $t$  décrit  $M_m$ . Notons que ces v.a. sont non linéairement corrélées.

Dans le cas du modèle (AR), on doit donc comparer  $\sigma_{m,h}$  avec la quantité  $(\text{var } Y_{t+h})^{1/2}$ . La Figure 1.13 montre les résultats obtenus en fonction de l'horizon de prévision, entre 1 et 180 jours, pour 4 mois différents. Et pour le modèle (VP), on doit comparer  $\sigma_{m,h}$  avec

$$\left( \frac{1}{N_m} \times \sum_{t \in M_m} \rho_{t+h} \right) \times (\text{var } Z_t)^{1/2}, \text{ où } N_m \text{ désigne le nombre de jours du mois } m. \text{ Les résultats}$$

sont montrés sur la Figure 1.14. Ces graphiques font apparaître très clairement que les erreurs théoriques données par le modèle (AR) ne correspondent pas à la réalité : elles sont sous-estimées lorsque l'horizon de prévision est en hiver, et surestimées quand l'horizon de prévision est en été. En revanche les erreurs théoriques du modèle (VP) sont beaucoup plus proches des erreurs réelles et reflètent bien leur aspect saisonnier. On observe cependant qu'elles sont encore parfois sous-estimées pour l'hiver. C'est tout à fait prévisible car la loi des erreurs du modèle est gaussienne alors que la loi réelle, dissymétrique, est à queue plus épaisse dans la partie négative.

**Remarque.**

On constate que les erreurs réelles pour les modèles (AR) et (VP) sont quasi-identiques. Cela vient du fait que les prévisions des deux modèles sont très proches (voir la note 60).

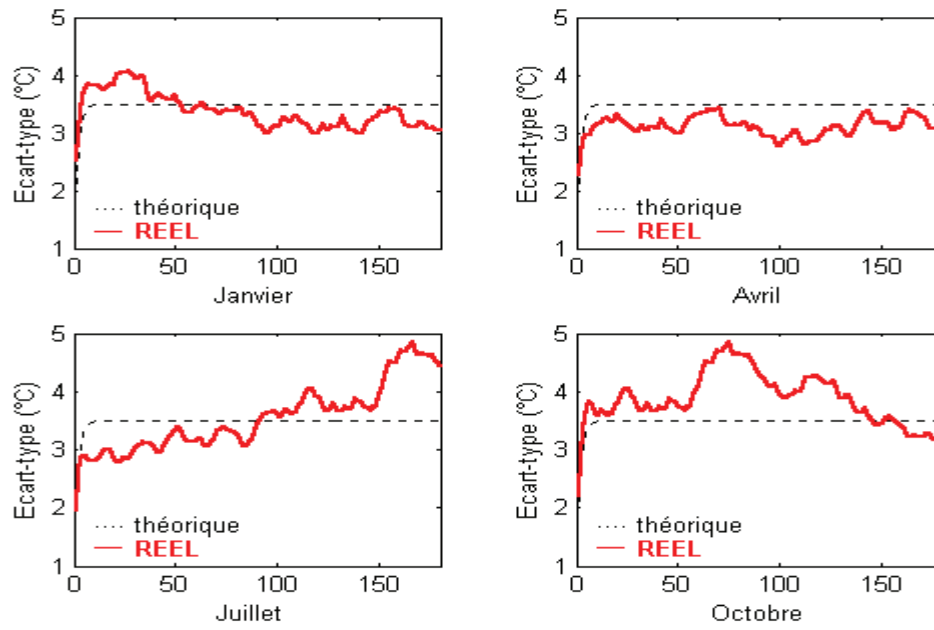


Figure 1.13. Erreurs de prévision théoriques et réelles - Modèle (AR).

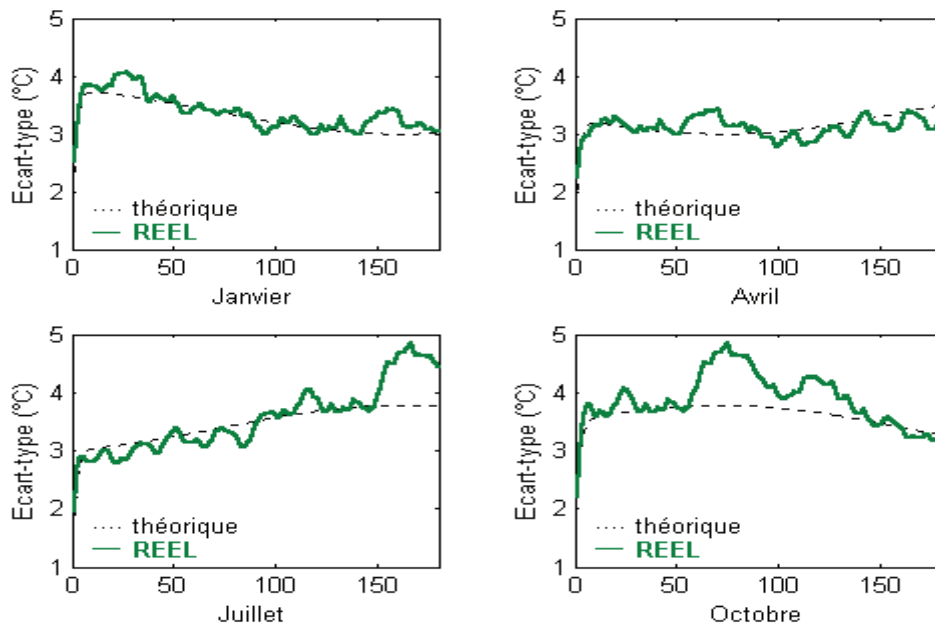


Figure 1.14. Erreurs de prévision théoriques et réelles - Modèle (VP).



## VII. Conclusions de la section A.

L'étude d'une série de 21 années de données journalières de température pour la station météorologique de Lyon-Bron nous a permis de dégager quelques caractéristiques de la température, notamment : dynamique autorégressive (après filtrage des composantes de tendance et de saisonnalité) et dispersion saisonnière. Ces propriétés se retrouvent pour la température de 10 autres stations disséminées sur l'hexagone dont nous disposions<sup>61</sup>. Des deux modèles économétriques étudiés – un modèle autorégressif pur ("AR"), et un modèle autorégressif à variance périodique ("VP") – le modèle (VP) est naturellement le plus adapté puisqu'il tient compte de la saisonnalité de la dispersion et, par conséquent, les erreurs théoriques sont davantage conformes à la réalité<sup>62</sup>.

D'autre part, lorsque l'on souhaite mener une évaluation des risques climatiques plus de 20 jours avant la période exposée au risque, on constate que les particularités de la situation courante ne sont pas significatives ; dans ce cas, le fait que les températures du jour courant et des quelques jours précédents soient exceptionnellement élevées par exemple n'apporte pas d'information pour une étude des risques. Pratiquement, dans la gestion des risques climatiques dont la période d'exposition au risque débute au-delà de 20 jours, il n'y aura pas d'inconvénient à utiliser la loi non conditionnelle au lieu de la loi conditionnelle, tout en tenant compte de la structure autorégressive (et donc des autocorrélations) de la dynamique des températures<sup>62</sup>.

Il n'est pas étonnant qu'un modèle aussi simple ne puisse donner de meilleurs résultats. Pour aller plus loin, il faut construire des modèles d'évolution beaucoup plus complexes faisant intervenir plus de variables c'est-à-dire des modèles météorologiques. Par exemple, avec de tels modèles on est parfois capable de prévoir l'impact du phénomène El Niño sur le climat des Etats-Unis jusqu'à plusieurs mois à l'avance (Dischel, 2000). Mais cela nécessite une base de données et des moyens de calculs très importants. Avec des possibilités moins grandes, on peut néanmoins tenir compte du phénomène en incorporant une ou plusieurs variables exogènes fournies par un organisme météorologique.

---

<sup>61</sup> Lille, Strasbourg, Brest, Rennes, Paris (Paris-Montsouris), Clermont-Ferrand, Grenoble, Bordeaux, Toulouse, Nice. La dynamique retenue est, en général, de type  $AR(p)$ , avec  $p = 2$  ou  $3$  (voir la section suivante).

<sup>62</sup> Ce résultat s'observe également pour les autres stations étudiées.

En ce qui concerne l'utilisation de modélisations financières classiques pour traduire la dynamique des températures, l'espoir semble assez mince au moins pour la série étudiée ici. Car autant il est possible de faire des analogies avec les séries financières, autant les comparaisons font apparaître les caractéristiques tout à fait particulières de la température. Ainsi la propriété de retour à la moyenne des températures peut-elle faire penser au premier abord à la modélisation de la dynamique des taux d'intérêts souvent décrite par le modèle de Vasicek. Cependant un tel modèle est, dans sa version discrète, de type AR(1) et serait donc trop grossier pour traduire la dynamique plus complexe de type AR(3) observée ici pour les températures. Par ailleurs, sans lien avec l'observation précédente, nous observons que les temps de retour à la moyenne sont très différents : de l'ordre d'une dizaine de jours pour la température, on passe à plusieurs années pour les taux d'intérêt. Par rapport aux rentabilités boursières, les effets ARCH, de kurtosis ou de skewness existent pour la température mais semblent moins prononcés. Enfin, la modélisation des températures journalières a fait apparaître deux types de saisonnalité "synchronisées" c'est-à-dire de même périodicité (annuelle en l'occurrence). La saisonnalité sur la moyenne d'abord, qui évoque par exemple les effets calendaires des cours boursiers (effet du lundi), mais qui est beaucoup plus forte dans le cas des températures ; Et la saisonnalité de la volatilité que l'on peut rapprocher des effets intra-day (courbe en U).

Nous avons donné une première conséquence des propriétés de la température pour la gestion des risques climatiques. Une deuxième conséquence, que nous nous contenterons cette fois de conjecturer, est que la saisonnalité de la volatilité et/ou la dissymétrie des distributions de température se traduiront par une saisonnalité dans l'évaluation des risques.



## Annexe B. Prévisions et erreurs de prévision.

### Proposition.

Soit  $W_t$  un processus  $AR(p)$  gaussien

$$W_t = \phi_1.W_{t-1} + \dots + \phi_p.W_{t-p} + \varepsilon_t$$

où  $\varepsilon_t$  est de loi  $N(0, \sigma^2)$ . Notons  $\psi_0, \psi_1, \dots$  les coefficients du développement de  $W_t$  en  $MA(\infty)$  :

$$W_t = \sum_{j=0}^{+\infty} \psi_j \varepsilon_{t-j}$$

Alors

1. Les prévisions vérifient la relation de récurrence valable pour tout  $h$

$$P_t W_{t+h} = \phi_1.P_t W_{t+h-1} + \dots + \phi_p.P_t W_{t+h-p}$$

avec  $P_t W_{t+i} = W_{t+i}$  si  $i \leq 0$ .

2. L'erreur de prévision à l'horizon  $h$  suit la loi normale  $N(0, \sigma_t^2(h))$  avec

$$\sigma_t^2(h) = \sigma^2 \cdot \sum_{j=0}^{h-1} \psi_j^2. \text{ En particulier l'intervalle } I_{t+h} = [P_t W_{t+h} - 1.96\sigma_t(h); P_t W_{t+h} + 1.96\sigma_t(h)]$$

est un intervalle de prévision à 95% :  $P(W_{t+h} \in I_{t+h}) = 0.95$

3.  $V_t X_{t+h}$  est une v.a. constante, égale à  $\sigma_t^2(h)$ . On a  $\sigma_t^2(h) \xrightarrow{h \rightarrow +\infty} \text{var } W_t$

### Preuve.

1. et 2. sont démontrés dans (Brockwell, Davis), §5.3., §5.4., et exercice 5.10.

3. résulte immédiatement de 2. et de (Brockwell, Davis), §3.3., formule 3.3.1.  $\square$

### Corollaire 1 -- Prévisions et erreurs de prévision du modèle (AR).

Si  $(X_t)$  est donné par le modèle (AR), alors

1.  $P_t X_{t+h} = m_{t+h} + s_{t+h} + P_t Y_{t+h}$

2.  $e_t(X_{t+h})$  suit la loi normale  $N(0, \sigma_t^2(h))$

3.  $\sigma_t^2(h) \xrightarrow{h \rightarrow +\infty} \text{var } Y_t$

$P_t Y_{t+h}$  et  $\sigma_t^2(h)$  se calculent en appliquant la proposition à  $Y_t$  avec  $p=3$ .

**Corollaire 2 -- Prévisions et erreurs de prévision du modèle (VP).**

Si  $(X_t)$  est donné par le modèle (VP), alors

1.  $P_t X_{t+h} = m_{t+h} + s_{t+h} + \rho_{t+h} \cdot P_t Z_{t+h}$
2.  $e_t(X_{t+h})$  suit la loi normale  $N(0, \rho_{t+h}^2 \cdot \sigma_t^2(h))$
3.  $\sigma_t(h)$  est équivalent à  $\rho_{t+h}$  lorsque  $h$  tend vers l'infini.

$P_t Z_{t+h}$  et  $\sigma_t^2(h)$  se calculent en appliquant la proposition à  $Z_t$  avec  $p=3$ .

**Annexe C. Choix d'une région de confiance des paramètres du processus (AR).**

Soit  $(X_t)$  un processus AR( $p$ ) avec

$$X_t = \phi_1 \cdot X_{t-1} + \dots + \phi_p \cdot X_{t-p} + \varepsilon_t$$

où  $(\varepsilon_t)$  est un bruit blanc de variance  $\sigma^2$ .

Le problème est de déterminer une région de confiance pour le vecteur  $(\Phi; \sigma^2)$  avec  $\Phi = (\phi_1, \dots, \phi_p)'$ . On devra pour cela utiliser les estimations obtenues, qui sont réalisées par maximum de vraisemblance. Notons  $\hat{\Phi}$  et  $\hat{\sigma}^2$  les estimateurs correspondants pour un échantillon de taille  $n$ . On a alors le résultat suivant (voir (Gourieroux, Montfort), §9.2. F et G) :

**Théorème.**

$$\sqrt{n} \begin{pmatrix} \hat{\Phi} - \Phi \\ \hat{\sigma}^2 - \sigma^2 \end{pmatrix} \xrightarrow{d} N \left( 0; \begin{pmatrix} \sigma^2 \Gamma_p^{-1} & 0 \\ 0 & 2\sigma^4 \end{pmatrix} \right)$$

avec  $\Gamma_p$  la matrice d'ordre  $p$  de terme général  $\gamma(i-j)$ ,  $1 \leq i, j \leq p$ .

En pratique, on utilise les estimateurs convergents usuels  $\hat{\gamma}(k)$  ((Brockwell, Davis) §7.2.) et on note de façon analogue  $\hat{\Gamma}_p$  la matrice de terme général  $\hat{\gamma}(i-j)$ ,  $1 \leq i, j \leq p$ . On déduit du théorème précédent que  $\hat{\Phi}$  et  $\hat{\sigma}^2$  sont asymptotiquement indépendants et que pour tous  $\alpha, \beta$  de  $]0;1[$ ,

$$\left\{ \begin{aligned} &(\Phi; \sigma^2), (\hat{\Phi} - \Phi)' \Gamma_p (\hat{\Phi} - \Phi) \leq n^{-1} \sigma^2 \chi_{1-\alpha, p}^2 \\ &\frac{\hat{\sigma}^2}{1 + \sqrt{\frac{2}{n}} \cdot N_{1-\beta/2}} \leq \sigma^2 \leq \frac{\hat{\sigma}^2}{1 - \sqrt{\frac{2}{n}} \cdot N_{1-\beta/2}} \end{aligned} \right\}$$

est approximativement une région de confiance de  $(\Phi; \sigma^2)$  au niveau  $(1-\alpha)(1-\beta)$ .

On a noté  $\chi_{1-\alpha, p}^2$  le  $(1-\alpha)$ -quantile de la loi du  $\chi^2$  à  $p$  degrés de liberté et  $N_{1-\beta/2}$  le  $(1-\beta/2)$ -quantile de la loi de Laplace-Gauss.

Pour les calculs, on a choisi  $\alpha$  égal à  $\beta$  tel que :  $(1-\alpha)(1-\beta) = 0.95$ .

## Annexe D. Ecarts entre les moments conditionnels et non conditionnels.

**Lemme.** Soit  $(W_t)$  un processus  $AR(p)$  centré. On note  $(\varepsilon_t)$  les résidus,  $\sigma^2$  leur variance et

$\psi_j$  le  $j^{\text{ème}}$  coefficient du développement de  $(W_t)$  en  $MA(\infty)$  :  $W_t = \sum_{s=0}^{+\infty} \psi_s \cdot \varepsilon_{t-s}$ . On pose

$$A_h = \sigma^2 \cdot \sum_{j=h}^{+\infty} \psi_j^2.$$

Soient deux entiers  $l$  et  $m$  tels que  $0 \leq l \leq m \leq v-u$  et  $h = u - t_0$ , avec  $u > t_0$ . Alors

1.  $\text{var } E(W_{u+l} | \underline{W}_{t_0}) = A_{h+l}$

2.  $\text{cov}(W_{u+l}; W_{u+m} | \underline{W}_{t_0}) - \text{cov}(W_{u+l}; W_{u+m}) = -\sigma^2 \cdot \sum_{j=h+l}^{+\infty} \psi_j \cdot \psi_{j+m-l}$

*Preuve.* De la décomposition du processus en  $MA(\infty)$ , on déduit  $E(W_t | \underline{W}_{t_0}) = \sum_{s=t-t_0}^{+\infty} \psi_s \cdot \varepsilon_{t-s}$ . Le

lemme en découle.

La partie de la proposition relative au processus (AR) s'obtient en appliquant le lemme à  $(Y_t)$  et en utilisant l'inégalité de Cauchy-Schwarz permettant d'écrire, par exemple,

$$\left| \sum_{j=h+l}^{+\infty} \psi_j \cdot \psi_{j+m-l} \right| \leq \sum_{j=h}^{+\infty} \psi_j^2$$

avec égalité pour  $m = l = 0$ .

En appliquant le lemme au processus  $(Z_t)$ , on obtient d'abord

$$1. \quad \text{var } E(X_{u+l} | \underline{X_{t_0}}) = \rho_{u+l}^2 \cdot \sigma^2 \cdot \sum_{j=h+l}^{+\infty} \psi_j^2$$

$$2. \quad \text{cov}(X_{u+l}; X_{u+m} | \underline{X_{t_0}}) - \text{cov}(X_{u+l}; X_{u+m}) = -\rho_{u+l} \cdot \rho_{u+m} \cdot \sigma^2 \cdot \sum_{j=h+l}^{+\infty} \psi_j \cdot \psi_{j+m-l}$$

d'où le résultat de la même façon.

## B. Compléments.

A l'issue des travaux présentés dans la section précédente, on dispose maintenant d'un modèle statistique pour la température. Celui-ci a été étudié pour une dizaine de stations réparties dans toute la France, mais seul le cas de la station de Lyon-Bron a été présenté. Nous voulons à présent combler ce manque.

Nous en profitons pour présenter des améliorations du modèle, suite aux critiques mentionnées dans le paragraphe A.III.1. Celles-ci avaient trait à la saisonnalité en pointant du doigt d'une part l'irrégularité de l'estimation (sauts d'un jour à l'autre, parfois de plus d'un degré) et, d'autre part, la dépendance des résultats de l'estimation statistique des paramètres du modèle de cette estimation préliminaire. Pour éliminer ces défauts, on peut par exemple proposer une forme paramétrique pour la saisonnalité. Celle-ci est choisie sous la forme d'une somme de sinusoides :

$$s_t = \sum_{i=1}^{N_s} (a_i \cos(i\omega t) + b_i \sin(i\omega t))$$

Notons que la seule fréquence annuelle ne suffit pas en général car les estimations de la saisonnalité font ressortir une dissymétrie nette. Le choix des fréquences à utiliser est effectué au moyen d'une analyse spectrale préliminaire, et conduit à sélectionner, en plus de la fréquence annuelle, la fréquence semestrielle ( $i = 2$ ) pour chacune des stations. La Figure 1.15 illustre le cas de Lyon. Le modèle est alors complètement paramétrique et compte  $9 + p + q$  paramètres ( $p$  et  $q$  sont les ordres du modèle ARMA). L'estimation est réalisée d'un seul coup par maximum de vraisemblance.

Ci-après figurent les résultats d'estimation de 10 villes en plus de Lyon. Pour chacune d'elles, on a également représenté les estimations de la droite de tendance, de la saisonnalité et de la courbe de volatilité, les résultats liés à la partie stochastique – sélection du modèle ARMA, tests graphiques des résidus. D'une manière générale, on retrouve les caractéristiques de la température précédemment citées : tendance à la hausse, saisonnalité dans la volatilité, persistance à quelques jours et queues de distributions épaisses pour les basses températures (d'un point de vue purement statistique, on aura également remarqué la même difficulté à propos des carrés des résidus avec des corrélations significatives pour les premiers ordres).

Cependant le climat français est le jeu de plusieurs types d'influences, océanique, méditerranéenne et même continentale à l'Est. Il n'est donc pas étonnant de repérer quelques spécificités locales. Ainsi l'influence méditerranéenne se retrouve-t-elle sur les résultats de Nice et Ajaccio (et Marseille dans une moindre mesure). La courbe de volatilité est quasiment plate indiquant une régularité du climat (dans l'amplitude des variations de température) sur toute l'année. Le niveau très faible de cette volatilité indique par ailleurs une faible amplitude de ces températures. La structure aléatoire retenue d'un ARMA(2,1) est différente du modèle AR(3) convenant (en première approximation) aux autres stations. Enfin on observe des queues de distributions épaisses pour les températures positives (c'est également le cas de Brest) ; ceci pourrait s'expliquer par l'abondance de températures douces. D'autre part, on retrouve des influences ou micro-climats de nature continentale à Strasbourg, Clermont-Ferrand ou même Lyon avec des courbes de volatilité creusées et de niveau élevé traduisant d'importantes fluctuations pendant l'année avec un maximum en hiver.

Pour finir, notons que la tendance estimée, quoique toujours haussière, est de niveau très inégal d'une ville à l'autre. Ces différences locales peuvent peut-être s'expliquer par les différences d'urbanisation des villes. Cela peut également traduire une erreur d'estimation relativement importante.



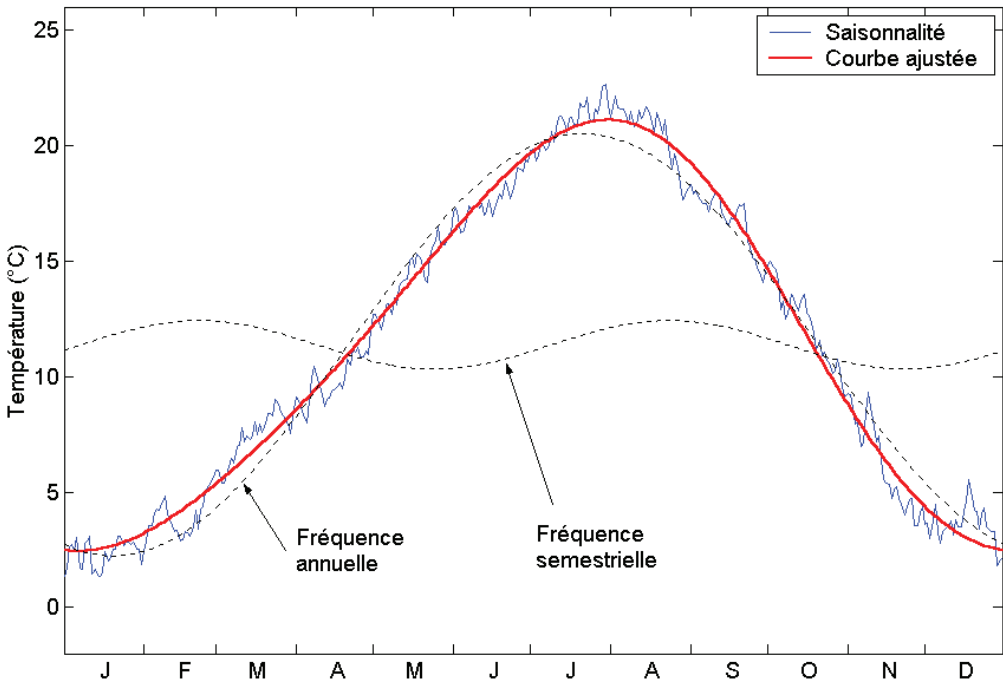


Figure 1.15. Paramétrisation de la saisonnalité. Cas de Lyon.

Tableau 1.3. Résultats d'estimation.

	Tendance		Saisonnalité			
	$d^*$	$e$	$a_1$	$b_1$	$a_2$	$b_2$
Ajaccio	1.33	14.68	-6.11	-3.79	0.04	0.80
Bordeaux	2.56	12.56	-6.80	-2.85	-0.21	0.89
Brest	1.38	10.86	-4.54	-2.59	0.06	0.59
Clermont-Ferrand	2.18	10.55	-7.62	-2.99	-0.20	1.04
Lille	2.08	9.67	-7.07	-2.79	-0.17	0.63
Lyon	1.94	11.34	-8.61	-3.05	-0.26	1.04
Marseille	1.73	14.98	-7.44	-3.25	0.06	1.02
Nice	0.85	15.44	-6.47	-3.50	0.01	1.04
Paris	1.85	11.44	-7.39	-2.67	-0.14	0.74
Strasbourg	1.61	9.94	-8.79	-2.63	-0.29	0.68
Toulouse	1.96	12.77	-7.40	-3.17	-0.18	1.10

	Volatilité			AR			MA
	a	b	c	$\phi_1$	$\phi_2$	$\phi_3$	$\theta_1$
Ajaccio	2.10	0.18	0.01	1.56	-0.59	.	-0.89
Bordeaux	3.30	0.42	0.03	0.84	-0.15	0.06	.
Brest	2.56	0.28	0.08	0.80	-0.11	0.07	.
Clermont-Ferrand	3.81	0.49	-0.04	0.85	-0.16	0.05	.
Lille	3.34	0.24	0.02	0.87	-0.15	0.06	.
Lyon	3.48	0.42	0.02	0.86	-0.14	0.04	.
Marseille	2.64	0.40	0.08	0.79	-0.05	0.03	.
Nice	1.90	0.16	0.08	1.58	-0.61	.	-0.84
Paris	3.36	0.11	0.15	0.91	-0.20	0.06	.
Strasbourg	3.58	0.49	0.05	0.89	-0.15	0.04	.
Toulouse	3.25	0.43	-0.07	0.83	-0.14	0.06	.

\*  $\times 10^{-4}$

Figure 1.16. Tendence estimée.

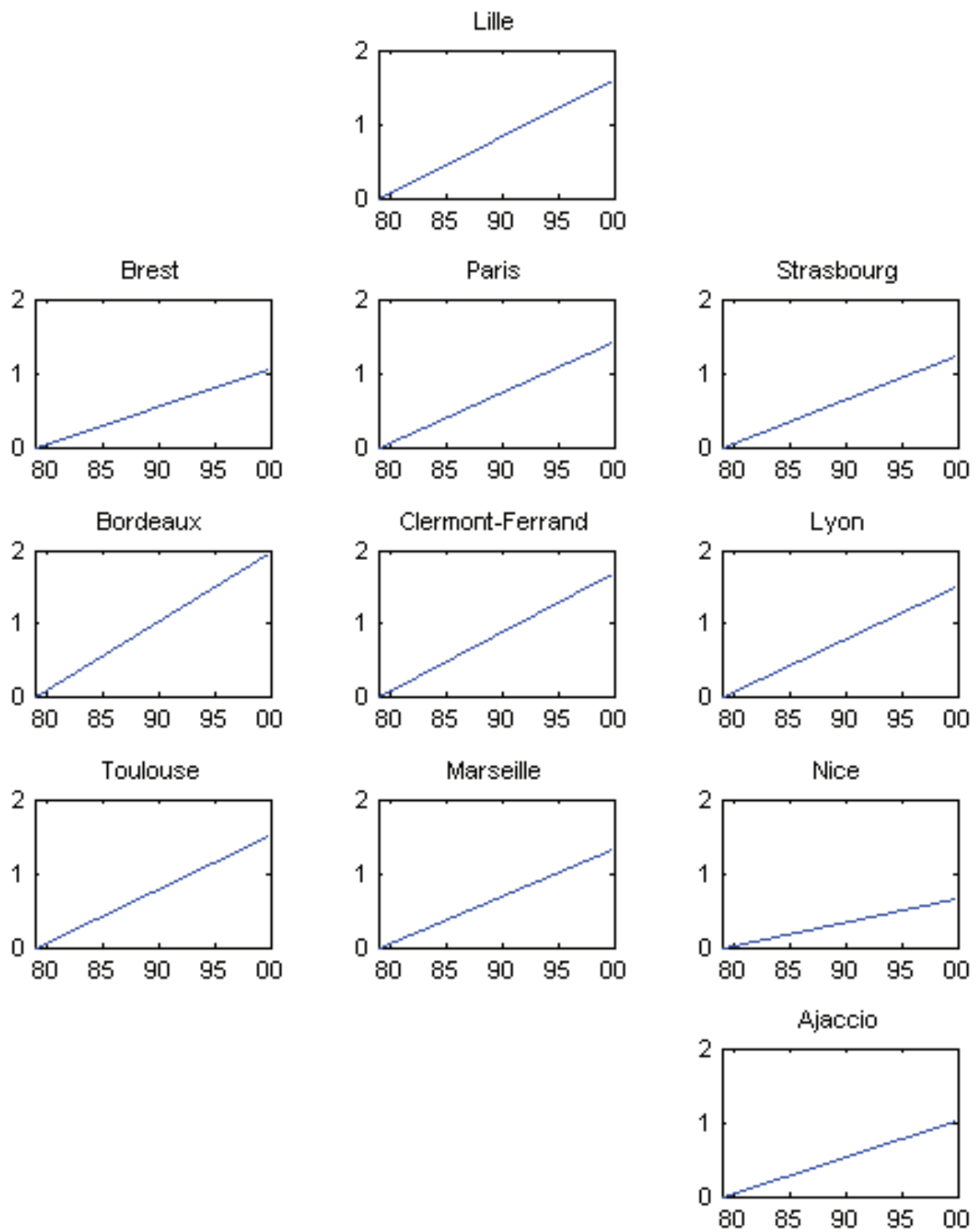


Figure 1.17. Saisonnalité historique et saisonnalité estimée.

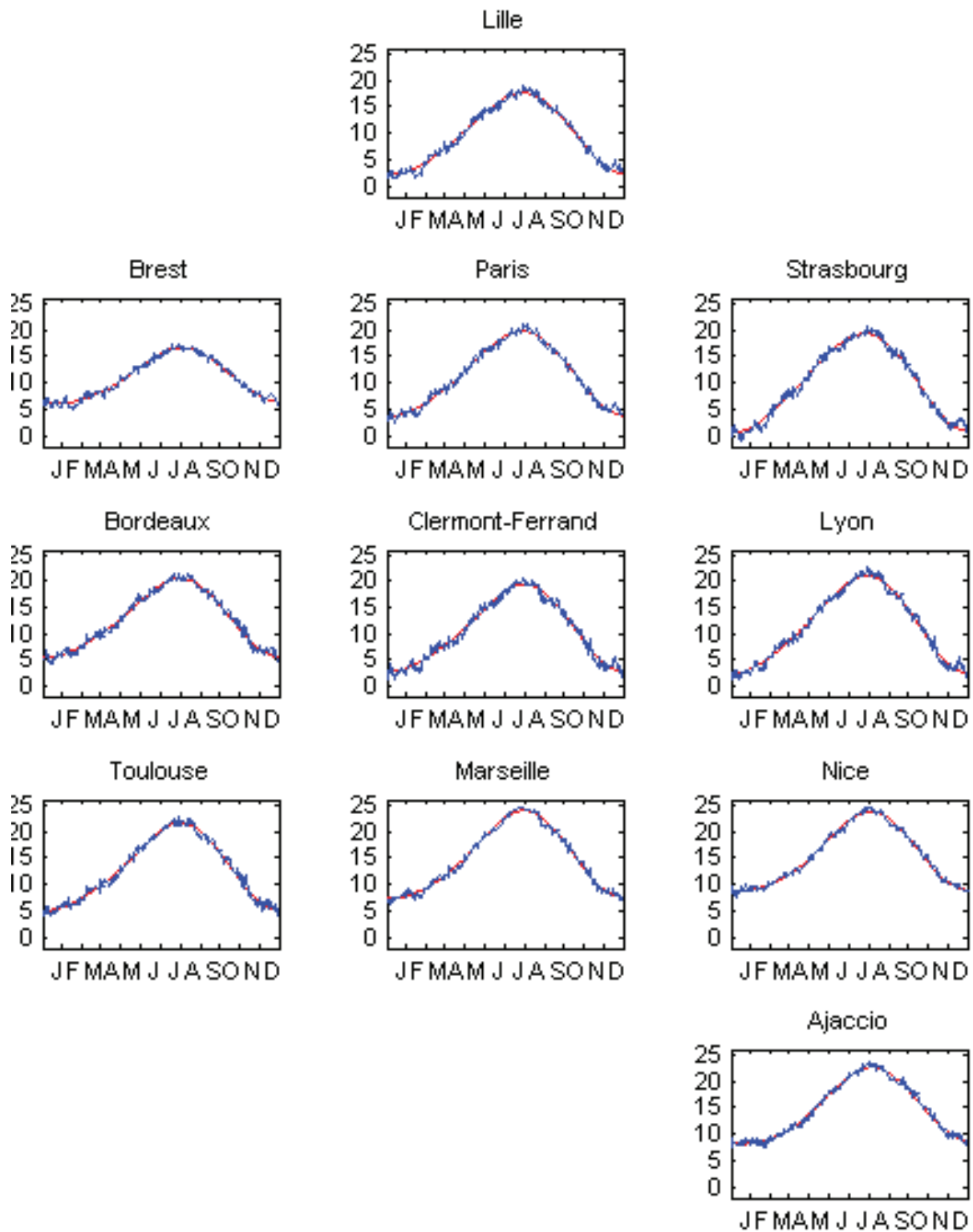


Figure 1.18. Volatilité estimée.

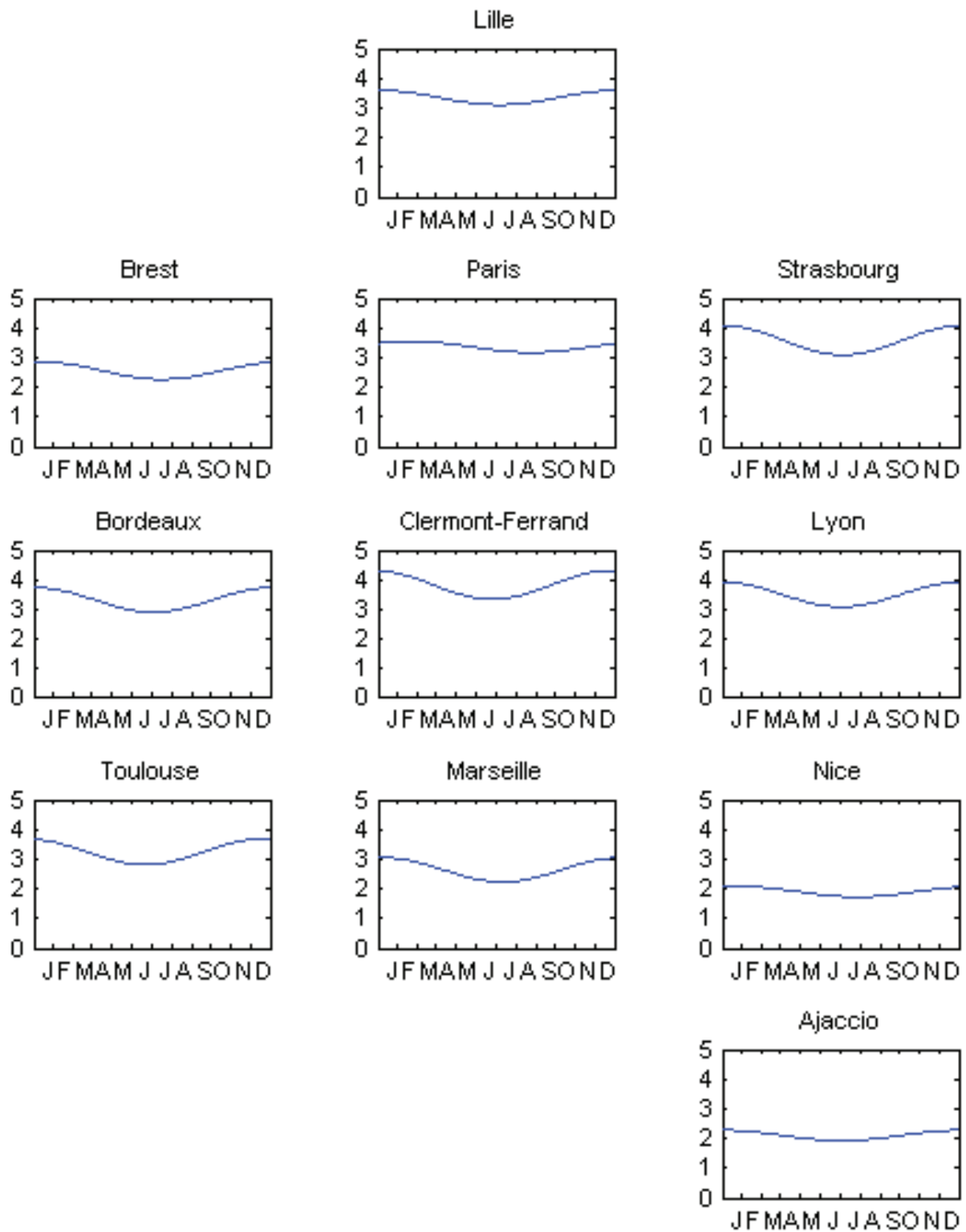


Figure 1.19. Résultat de la sélection du modèle ARMA.

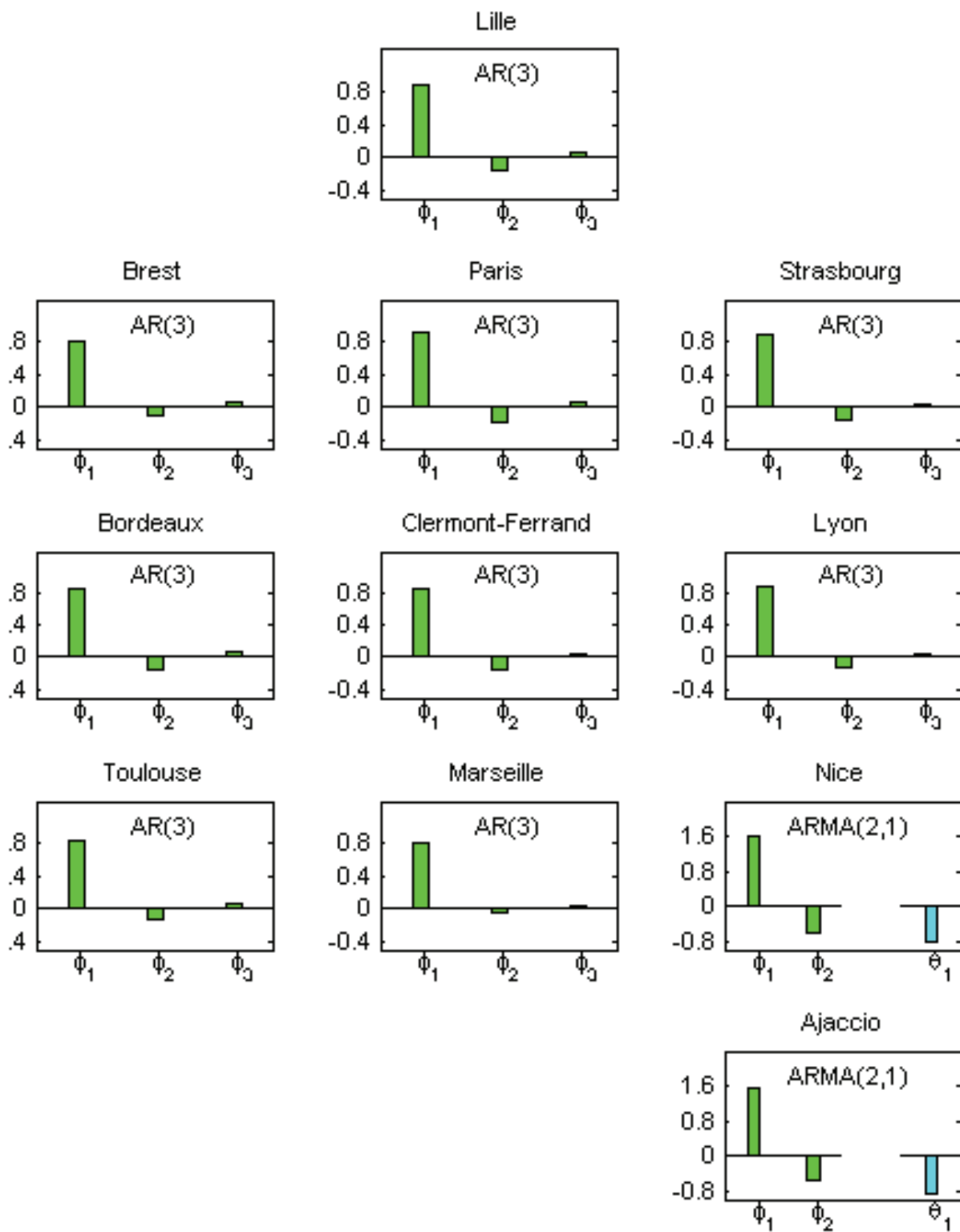


Figure 1.20. Autocorrélogramme des résidus estimés (1).

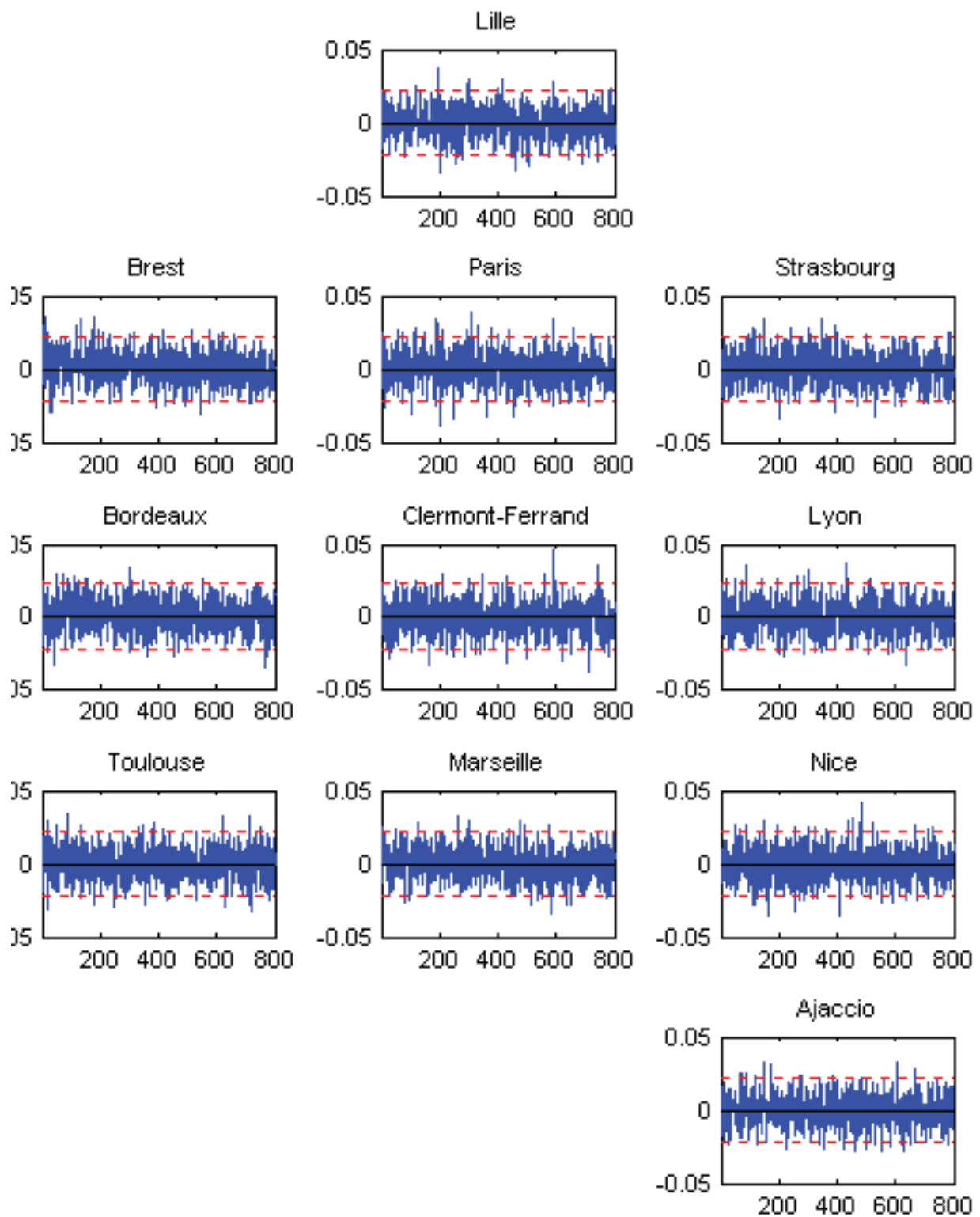


Figure 1.21. Autocorrélogramme des carrés des résidus estimés (1).

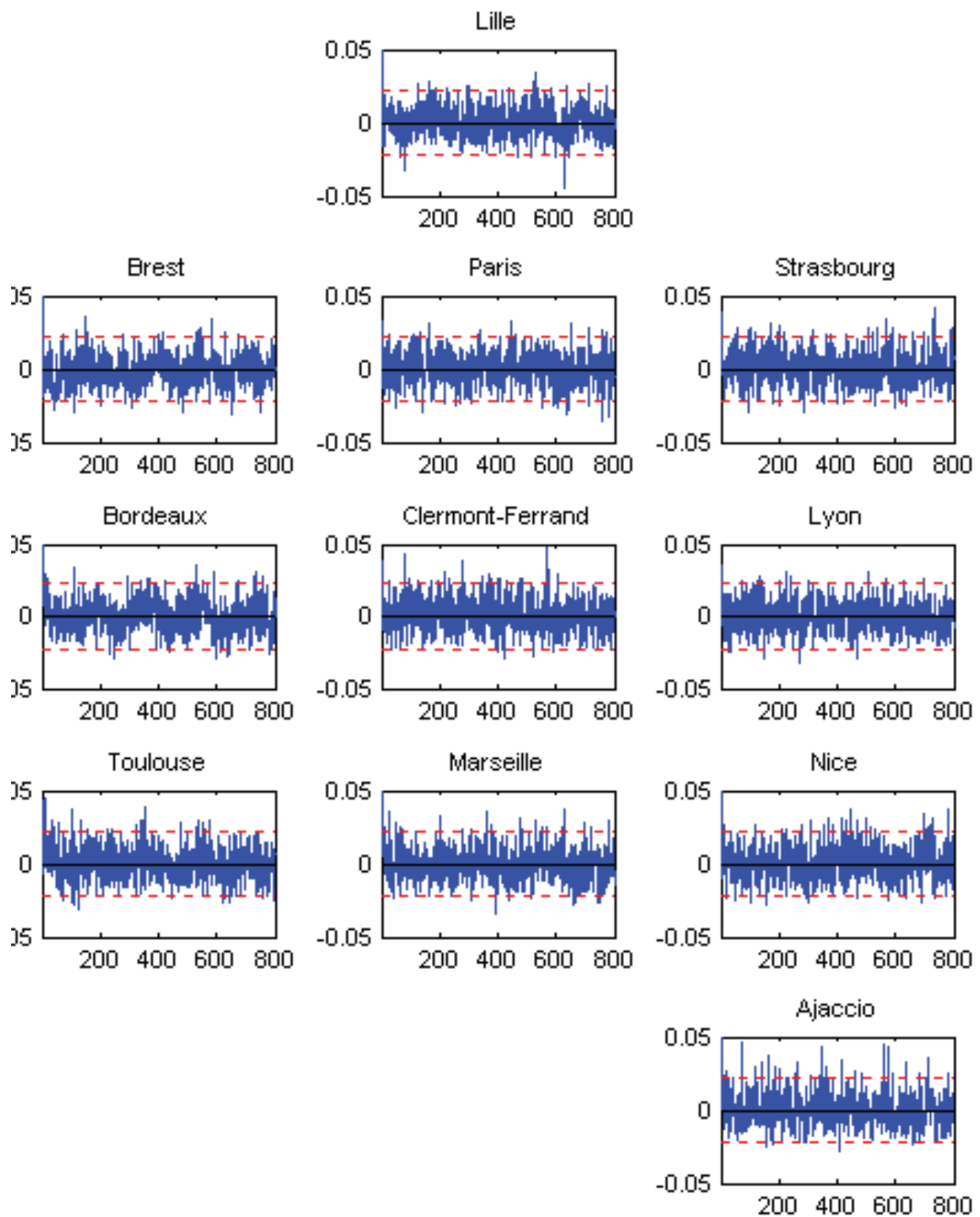




Figure 1.22. Autocorrélogramme des résidus estimés (2).

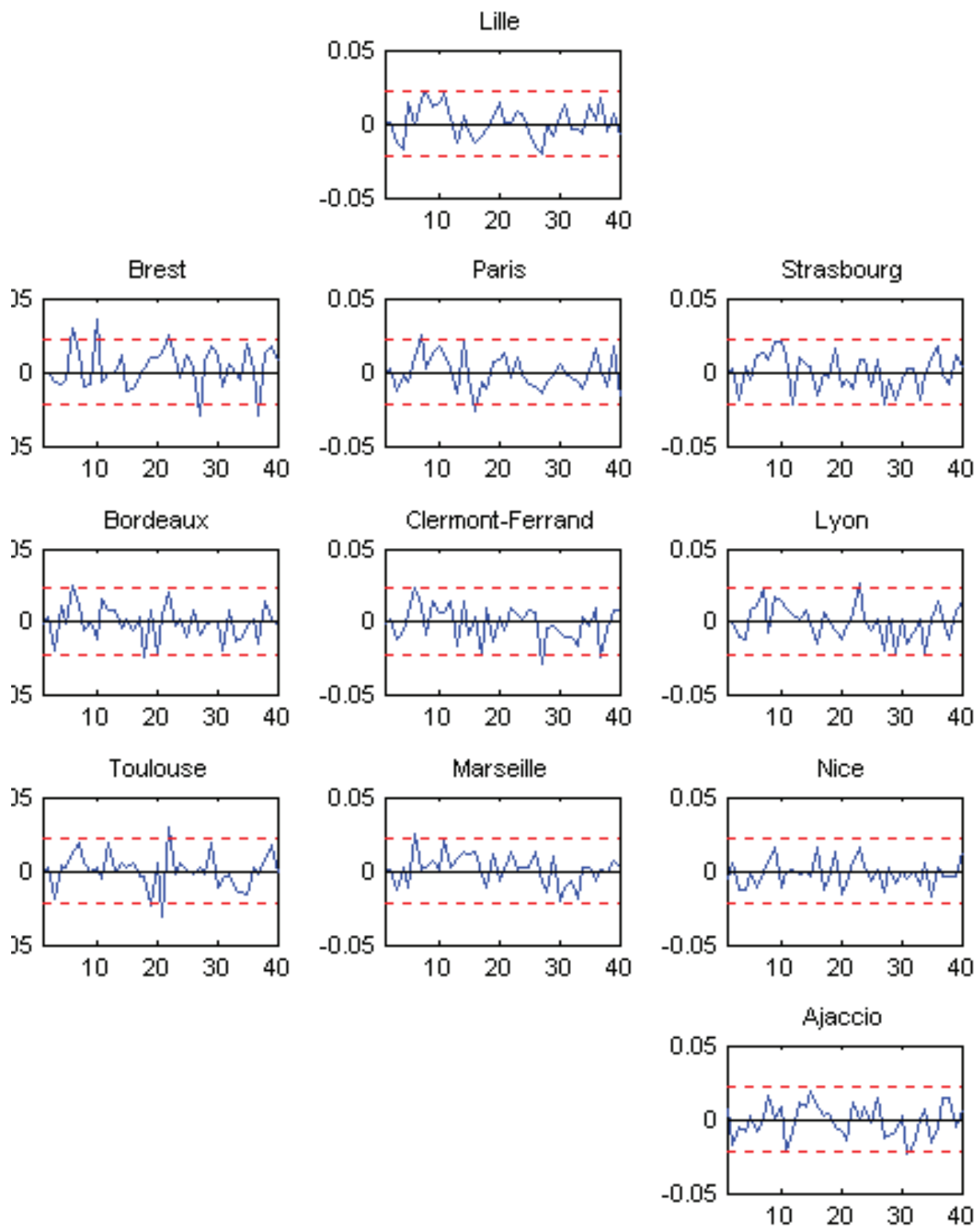


Figure 1.23. Autocorrélogramme des carrés des résidus estimés (2).

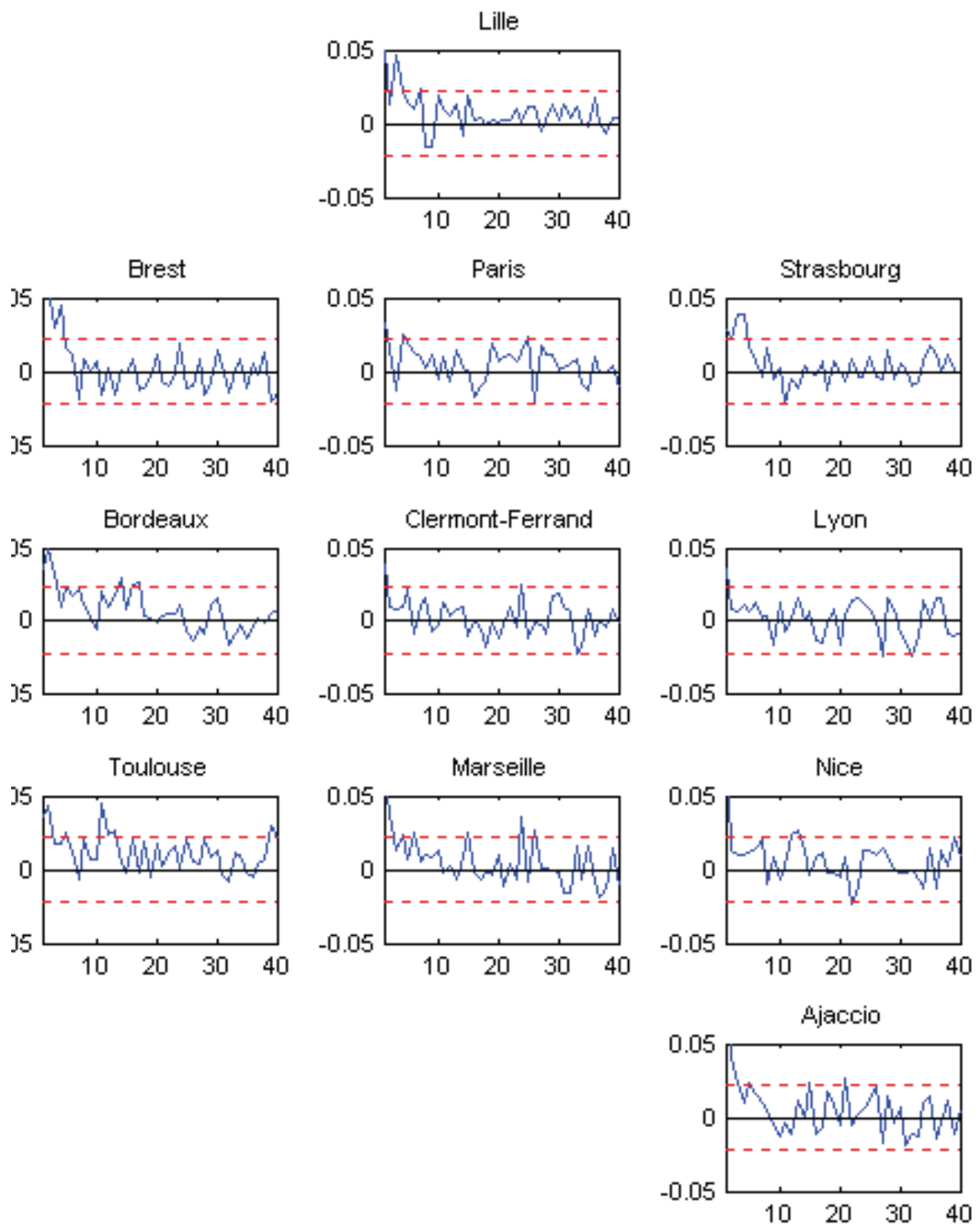
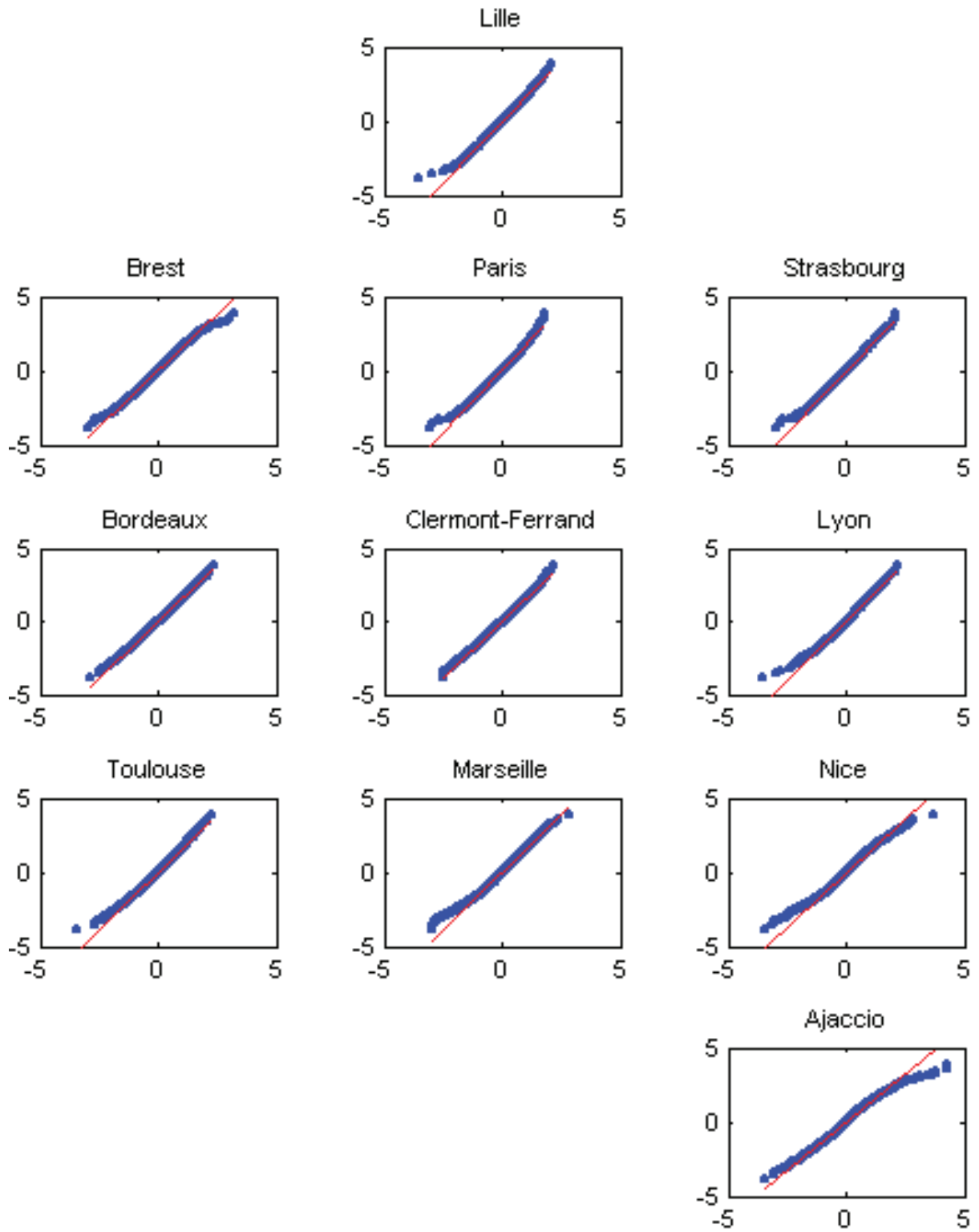


Figure 1.24. Droite de Henri des résidus estimés.



### C. Modélisation de la température : aspect multivarié.

Dans la section précédente, on a proposé une première modélisation de la température en un lieu donné. On a ainsi extrait des caractéristiques essentielles de la température, comme la saisonnalité de la dispersion, et mis en évidence une dynamique autorégressive dominante. Cependant le modèle présente le défaut important d'être basé sur une information plutôt pauvre, à savoir l'historique de température du lieu. Or la température en un lieu donné dépend de plusieurs variables météorologiques comme la pression atmosphérique, la vitesse du vent... ainsi que de la circulation atmosphérique. Pour modéliser la température avec précision, il faudrait donc en théorie disposer des historiques de ces variables en un grand nombre de lieux. Même si en pratique une telle stratégie serait trop coûteuse, on aimerait tenir compte de ces aspects pour affiner la modélisation. L'incorporation d'autres variables requiert des études préliminaires importantes comme le choix de ces variables, ainsi que leur modélisation. En revanche la prise en compte de la dimension spatiale est applicable à peu de frais pour la seule variable de température ; pour cette raison nous avons privilégié cette voie comme première extension du modèle autorégressif AR.

L'objectif de cette section est d'apprécier l'amélioration de la dynamique de la température lorsqu'on adjoint l'aspect spatial dans la modélisation. Si l'on reprend le cas de Lyon, la question est de savoir si l'on augmente « beaucoup » la qualité de la dynamique de la température à Lyon lorsqu'on dispose des températures passées à Lyon, Brest, Clermont-Ferrand, Paris, Marseille, Grenoble, etc. Pour répondre, on pourrait utiliser un modèle de régression linéaire où la réponse serait la température (stationnarisée) à Lyon et les facteurs la température (stationnarisée) des autres villes au jour courant et aux jours précédents. Cependant il n'y a pas à proprement parler de « facteurs » qui sont à l'origine d'une réponse mais plutôt des influences entre les températures des différentes villes traduisant non pas une relation de cause à effet, mais le phénomène global de circulation de masses d'air. Il semble donc plus naturel de considérer le système formé par l'ensemble des températures des différents lieux. L'extension naturelle du modèle local AR est alors le modèle multivarié VAR (Vector AutoRegressive). Son intérêt est multiple. *Primo*, il permet de rendre compte de la dynamique spatiale dans son ensemble et non seulement de déplacements vers un lieu fixé. *Secundo*, il donne donc d'un seul coup la dynamique de la température en chaque lieu (contrairement à un modèle de régression). *Tertio*, il est particulièrement adapté à la gestion

de portefeuilles de dérivés climatiques puisqu'il donne la loi jointe du processus vectoriel des températures en différents lieux.

La modélisation de la dynamique spatio-temporelle de la température par un processus de type VAR est séduisante, mais quelles améliorations peut-on réellement attendre par rapport au modèle temporel AR ? Naturellement, on espère prendre en compte le déplacement de masses d'air, et plus particulièrement l'influence océanique qui est dominante dans le climat de la France. Une première limitation vient de la localisation des stations météorologiques. Ne disposant que de stations terrestres, il paraît difficile de tirer profit du déplacement Ouest-Est pour la température de... Brest ! Une deuxième limitation vient de la fréquence d'échantillonnage des données. Celle-ci étant journalière, les circulations de durée inférieure à la journée seront peu accessibles. Toutefois, le temps de traversée d'une perturbation océanique varie de 24h à 72h. On peut donc être assez optimiste de ce côté-là. Une troisième limitation vient de la simplicité du modèle VAR. Celui-ci ne permet absolument pas de tenir compte des déplacements saisonniers ou ponctuels car les coefficients du modèle sont constants au cours du temps. On percevra au contraire une dynamique générale, sorte de « moyenne » de la dynamique sur l'année. Enfin, et c'est peut-être le point le plus sérieux, rien ne dit que ce soit la variable température qui permette de rendre compte de la circulation atmosphérique... Cependant, même si la variable « pression » est sans doute mieux adaptée pour annoncer une perturbation, il est vrai que celle-ci se ressentira au niveau des températures et que, inversement, une chute brutale des températures est souvent le signe de la traversée d'une perturbation. En résumé, le modèle devrait permettre de percevoir des déplacements de masses d'air dominants comme l'influence océanique. En revanche, les déplacements rapides ou saisonniers risquent d'être atténués, voire éliminés.

On est donc en droit de penser que le modèle VAR intégrera au moins l'influence dominante Ouest-Est (et nous verrons - en effet - qu'elle l'est) et que, partant, la dynamique de la température au lieu considéré sera améliorée. Cependant la mise en place d'une modélisation VAR a un coût. Le modèle AR(3) local comporte déjà 12 paramètres ; avec un modèle VAR(3) faisant intervenir  $S$  stations, le nombre de paramètres est multiplié par  $S$ . Or nous verrons qu'une des difficultés posées par la modélisation AR est la répercussion quelquefois désastreuse de l'erreur d'estimation sur le prix de produits dérivés climatiques. La multiplication du nombre des paramètres est donc à prendre avec précaution (il y a, en outre, peu de chances que ces paramètres soient mieux estimés en raison de la simplicité du modèle

linéaire VAR pour décrire la dynamique spatiale des températures). Se pose donc le problème de la *rentabilité* du modèle multivarié. Afin d'apprécier l'importance des améliorations qu'il apporte par rapport au modèle local, on a choisi de comparer les prévisions post-sample à quelques jours des deux modèles avec celles de Météo-France. Les prévisions de Météo-France sont le fruit d'une modélisation très poussée du climat dans son ensemble ; on peut les considérer comme des prévisions étalon d'un super-modèle multivarié. La comparaison des prévisions du modèle spatio-temporel de type VAR avec les prévisions du modèle temporel AR d'une part, et celles de Météo-France d'autre part, fourniront un élément important d'appréciation du gain réalisé par la prise en compte de l'aspect spatial dans la modélisation statistique de la température.

Cette section est structurée comme suit. Dans une première partie, on précise le cadre de travail. Dans la seconde partie, des études descriptives sont menées visant à retrouver dans les données de température les influences climatiques. Elles sont utilisées pour faire une sélection des stations pour un modèle VAR. Une troisième partie présente le modèle et donne des résultats d'estimation. La comparaison avec ceux du modèle AR fournit une première indication de l'amélioration réalisée. Enfin dans la dernière partie, les prévisions des modèles AR et VAR sont comparées avec celles de Météo-France dans le cadre d'une analyse post-sample.

## I. Quelques précisions sur le cadre de travail.

Nous étudions l'amélioration de la dynamique de la température dans deux cas : Lyon et Paris. Le premier est présenté en détail ; seuls les résultats essentiels sont fournis pour Paris. Les conclusions, les résultats intermédiaires sont comparables pour les deux villes. Nous disposons des prévisions quotidiennes de température de Météo-France à Lyon et Paris pour les horizons  $h = 1, 2, \dots, 5$  jours, sur la période allant du 01/01/96 au 31/12/99. Pour mettre en œuvre la modélisation VAR, nous disposons des historiques de température de 13 villes de France ; cependant de façon à ne pas avoir à recourir à un modèle de type VARMA, nous n'utiliserons pas les températures relevées à Nice et Ajaccio pour lesquelles le modèle local retenu est de type ARMA(2,1). De façon à comparer les erreurs *réelles* de prévision, nous utilisons seulement les 17 premières années de chaque historique (du 01/01/79 au 31/12/95), réservant les 4 dernières années de données (du 01/01/96 au 31/12/99) pour l'analyse post-

sample. Pour chacune des 11 autres stations, la température est modélisée (localement) par le modèle AR(3) sous sa forme paramétrique (voir B). Les paramètres des modèles locaux sont réestimés en conséquence par maximum de vraisemblance. La modélisation VAR est effectuée à partir des séries stationnarisées : (avec les notations du paragraphe A.III.1)

$$Z_t = \frac{X_t - (m_t + s_t)}{\sigma_t}$$

Enfin concernant les notations, on accole les trois premières lettres de la station considérée à chaque notation. On notera donc  $Z_{Lyo,t}$ ,  $Z_{Bor,t}$  ... en référence à Lyon, Bordeaux...

## II. Etudes descriptives préliminaires.

Avant de se lancer dans une modélisation VAR, il est utile de rechercher les corrélations entre les différentes variables de façon à accroître la compréhension des données d'une part, et d'autre part à proposer une sélection pertinente d'un petit nombre de variables. Dans notre cas, on aimerait savoir notamment s'il est possible de retrouver dans les données de température des influences climatiques dominantes comme l'influence océanique. En existe-t-il d'autres ?

### 1. Autocorrélations croisées.

Le premier réflexe consiste à regarder les corrélations croisées entre les températures (stationnarisées) des différentes villes. Si l'on note  $Z_{1,t}$  et  $Z_{2,t}$  les températures en deux stations  $s^1$  et  $s^2$ , elles sont définies par (Brockwell, Davis, § 11.2)

$$\rho(s_t^1, s_{t-h}^2) = \text{corr}(Z_{1,t}, Z_{2,t-h}) := \frac{\text{cov}(Z_{1,t}, Z_{2,t-h})}{\sqrt{\text{var}(Z_{1,t})} \sqrt{\text{var}(Z_{2,t-h})}}$$

et estimées au moyen de l'estimateur usuel

$$\hat{\rho}(s_t^1, s_{t-h}^2) = \frac{\sum_{t=h+1}^N (Z_{1,t} - \bar{Z}_1)(Z_{2,t-h} - \bar{Z}_2)}{\sqrt{\sum_{t=1}^N (Z_{1,t} - \bar{Z}_1)^2} \sqrt{\sum_{t=1}^N (Z_{2,t} - \bar{Z}_2)^2}}$$

où  $\bar{Z}_i = \frac{1}{N} \sum_{t=1}^N Z_{i,t}$ ,  $i=1,2$ , et  $N$  le nombre de données.

Les propriétés asymptotiques de  $\hat{\rho}(s_t^1, s_{t-h}^2)$  sont rappelées dans (Brockwell, Davis, § 11.2).

En pratique, comme dans le cas univarié, on se donne comme bornes de confiance les valeurs  $\pm \frac{2}{\sqrt{n}}$  pour juger de la signification des autocorrélations croisées.

La Figure 1.25 représente les autocorrélations croisées entre les températures de Lyon, Bordeaux et Paris. On remarque d'importantes autocorrélations entre chaque station et, qui plus est... dans tous les sens ! Sur les cadrans diagonaux, on retrouve les autocorrélations (temporelles) des modèles locaux avec l'allure caractéristique d'un processus autorégressif. C'est également cette allure qui est observée pour les autocorrélations « spatio-temporelles » entre deux stations différentes. En réalité, on ne détecte ici rien d'autre que la structure autorégressive dominante de la dynamique des températures. On peut penser que la variable utilisée - c'est-à-dire l'écart à la normale saisonnière (stationnarisé), est trop « grossière » pour rendre compte d'une dynamique spatiale.

Il est alors normal de considérer les résidus des modèles de température comme nouvelle variable. Pour chaque station, on s'intéresse donc au processus  $(\varepsilon_t)$  tel que

$$Z_t = \phi_1.Z_{t-1} + \dots + \phi_p.Z_{t-p} + \varepsilon_t$$

On peut aisément se convaincre que  $\varepsilon_t$  est bien plus approprié pour détecter le passage d'une perturbation. Imaginons en effet que ladite perturbation arrive à la date  $t$  sur la façade océanique. A Bordeaux, par exemple, la température chute et cette baisse se retrouve très précisément dans le résidu estimé à partir des données  $\hat{\varepsilon}_{Bor,t}$ . C'est en effet la partie de  $Z_t$  qui rend compte des « surprises » instantanées. L'autre partie,  $\phi_1.Z_{t-1} + \dots + \phi_p.Z_{t-p}$ , relève du passé ; d'ailleurs elle est égale à la prévision que l'on ferait à la date  $t-1$  avec l'information passée (relative à Bordeaux uniquement). Supposons que la perturbation mette 1 jour pour arriver à Lyon, elle sera « détectée » par le résidu  $\hat{\varepsilon}_{Lyo,t+1}$ . Si ce déplacement de masses d'air est suffisamment régulier dans l'année, on s'attend donc à trouver une corrélation significative entre  $\hat{\varepsilon}_{Bor,t}$  et  $\hat{\varepsilon}_{Lyo,t+1}$ . En revanche, il n'y a pas de raison, a priori, d'observer une corrélation dans l'autre sens, entre  $\hat{\varepsilon}_{Lyo,t}$  et  $\hat{\varepsilon}_{Bor,t+1}$ .



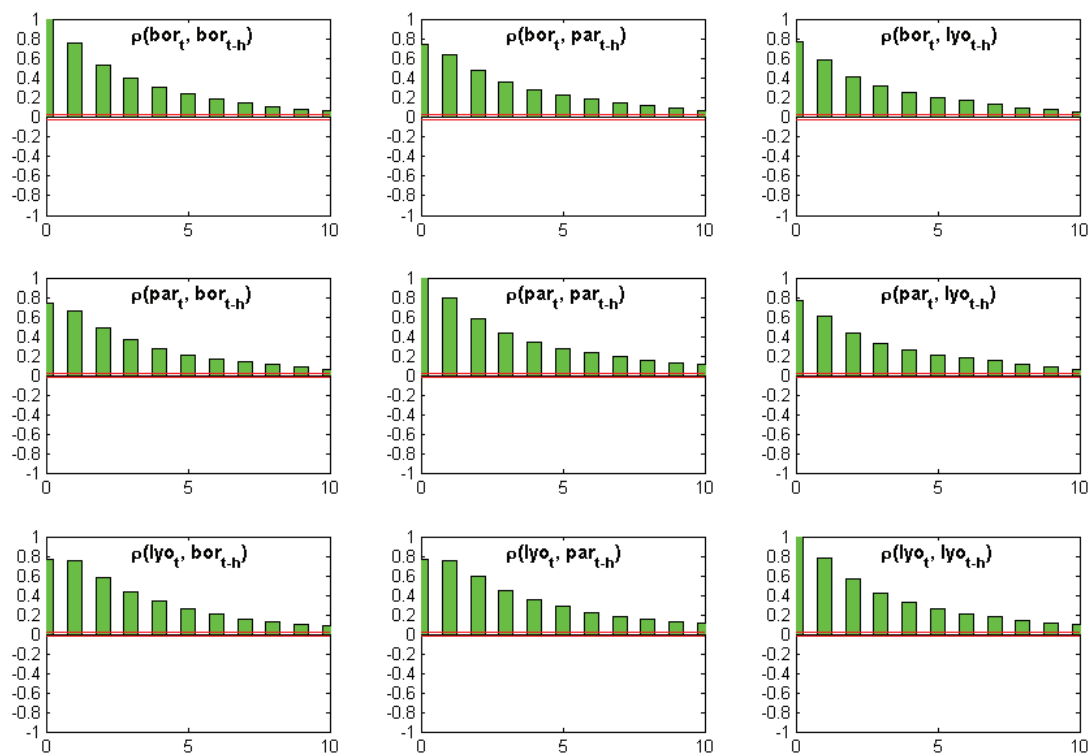
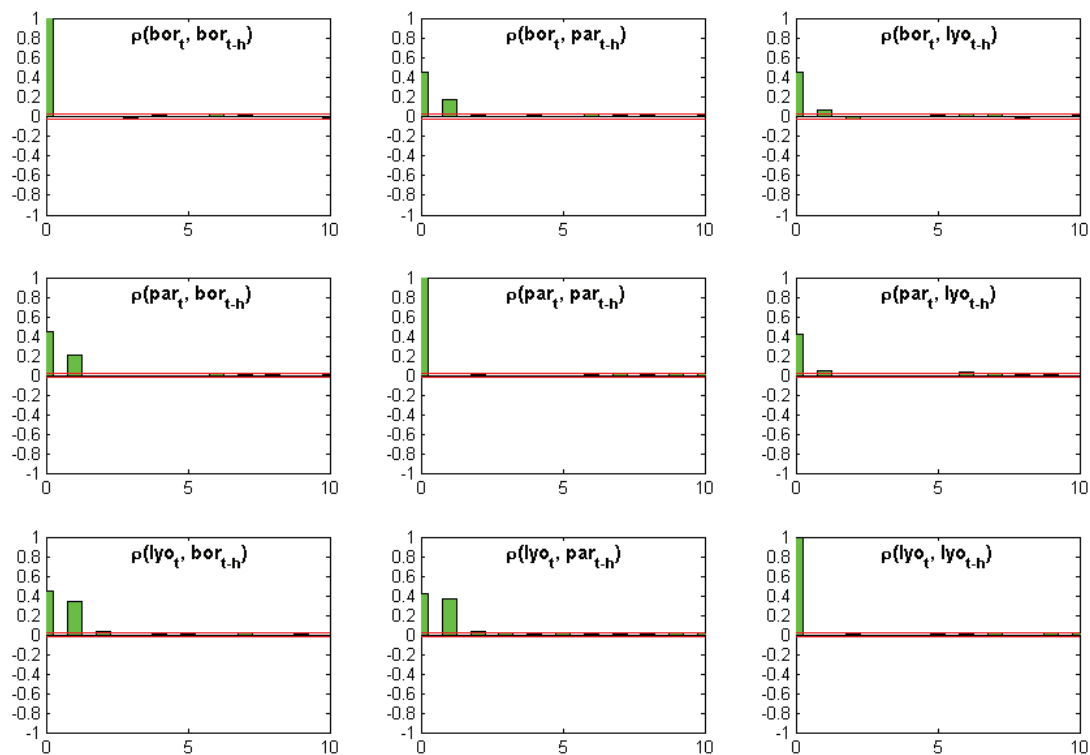


Figure 1.25. Autocorrélations croisées pour Bordeaux, Paris, Lyon.

Figure 1.26. Autocorrélations croisées des résidus pour Bordeaux, Paris, Lyon.



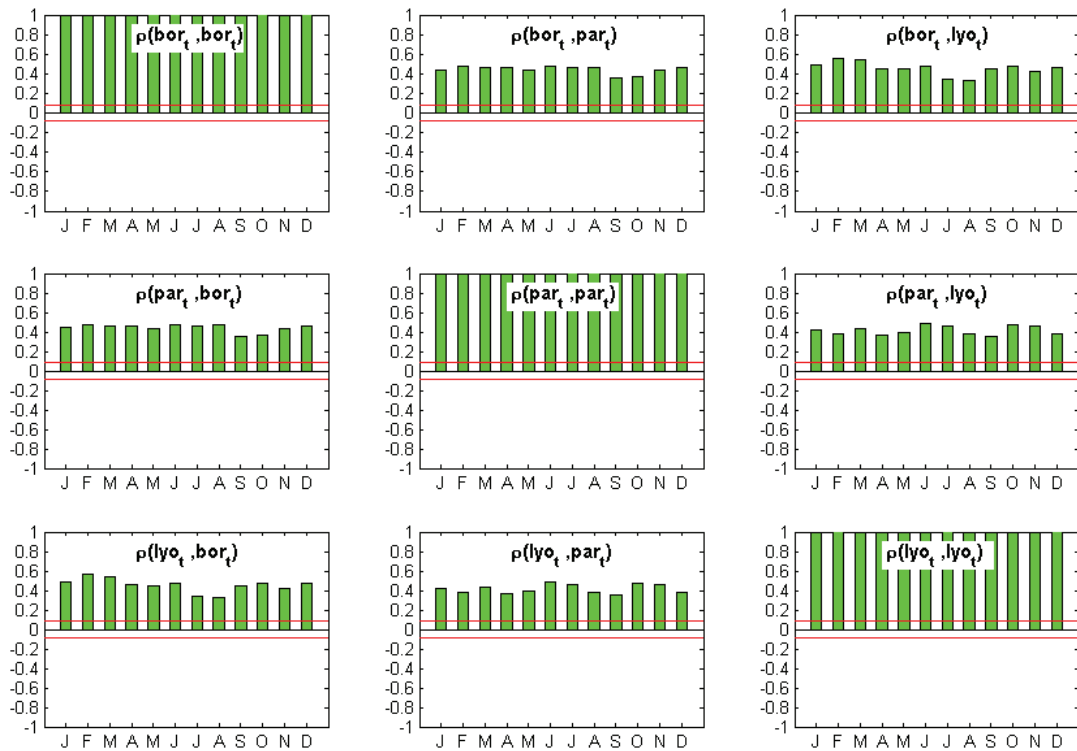
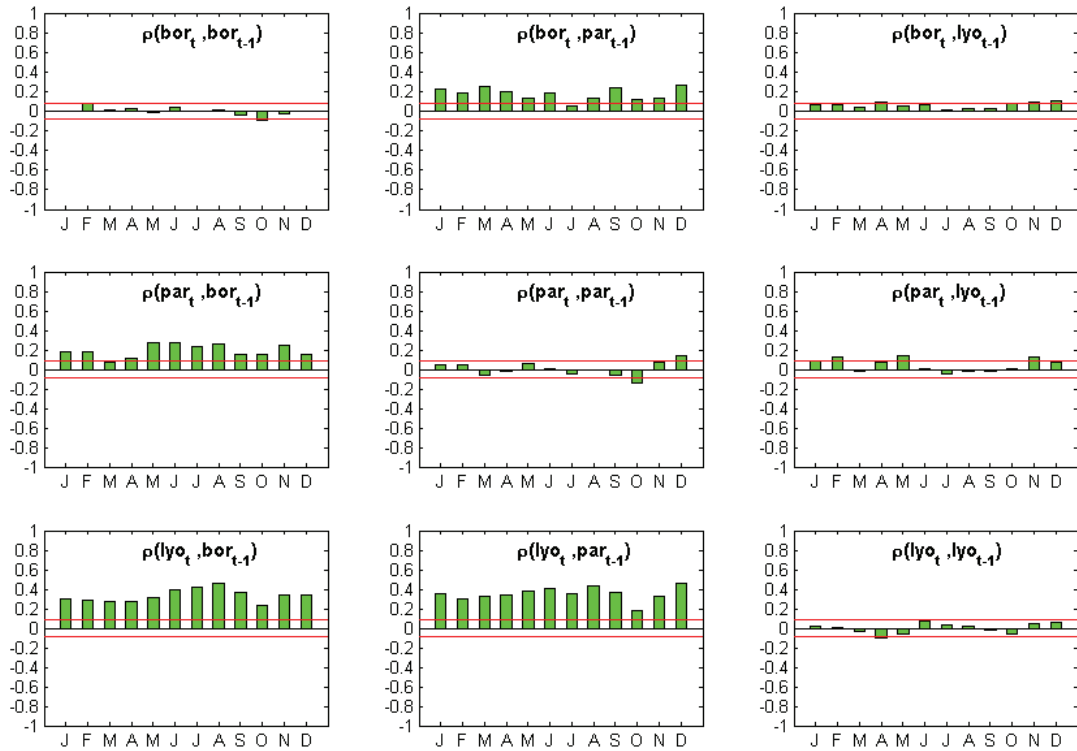


Figure 1.27. Fluctuations saisonnières des autocorrélations croisées instantanées des résidus des modèles locaux. Cas de Lyon, Bordeaux, Paris.

Figure 1.28. Fluctuations saisonnières des autocorrélations croisées de décalage 1 des résidus des modèles locaux. Cas de Lyon, Bordeaux, Paris.



On a donc représenté les autocorrélations croisées des résidus des modèles locaux de température des stations de Lyon, Bordeaux et Paris (Figure 1.26). On observe en effet une corrélation significative de Bordeaux vers Lyon, mais pas de Lyon vers Bordeaux. D'autre part, on observe une corrélation de Bordeaux vers Paris ce qui peut également s'expliquer par une influence océanique. Il semble exister également des corrélations de Paris vers Bordeaux et Lyon, mais pas de Lyon vers Paris. Au niveau local, il est connu que le temps qu'il fait à Paris se retrouve le lendemain à Lyon ; le résultat que l'on trouve ici semble être en accord avec cette information « de terrain ». On peut penser qu'on détecte ici une influence Nord-Sud. Les figures Figure 1.34, Figure 1.35 et Figure 1.36 (pages 124 à 125) reproduisent le même type de graphiques pour d'autres groupes de villes. On remarque d'une façon générale des corrélations dans le sens Ouest-Est et Nord-Sud. En revanche, on ne détecte pas, sur ces graphiques, d'influence méditerranéenne claire : il n'y a pas, par exemple, de corrélations de Marseille (ou Nice, ou Ajaccio) vers Lyon.

## **2. Des fluctuations saisonnières dans les autocorrélations croisées ?**

On peut se demander si les corrélations observées dépendent ou non de la saison. L'influence Nord-Sud est-elle plus importante l'hiver que l'été, l'influence océanique est-elle constante au cours du temps ? Nous avons donc représenté les autocorrélations croisées des résidus des modèles locaux de température, mois par mois. Nous nous sommes limités aux corrélations instantanées et aux corrélations du jour au lendemain qui sont les seules qui soient significatives. Le cas de Lyon, Bordeaux, Paris est illustré sur la Figure 1.27 et la Figure 1.28. Les graphiques concernant d'autres stations sont regroupés en fin de chapitre (Figure 1.37 à Figure 1.42, pages 126 à 128). Les résultats semblent différer un peu selon que l'on regarde les autocorrélations instantanées ou retardées. Dans le premier cas on observe, en général, une légère saisonnalité avec un maximum pour l'hiver ; dans le second, il semble y avoir une faible augmentation en été suivie d'un creux en octobre. Cependant, ces conclusions doivent être appréciées avec beaucoup de prudence du fait que l'on se base sur des corrélations estimées : il n'est pas clair que les différences que l'on observe soient significatives. D'autre part, on doit être prudent également par rapport aux interprétations que l'on peut en tirer. On pourrait penser, de façon un peu hâtive, à utiliser un modèle VAR à coefficients non constants (dépendant de la saison) pour traduire ces légères fluctuations. Toutefois, il est plus vraisemblable que celles-ci traduisent un défaut dans la stationnarisation des températures : on

ne doit pas oublier, en effet, que la variance des modèles locaux est modélisée assez sommairement par un processus déterministe sinusoïdal...

### 3. Conclusion. Sélection de stations.

Dans cette section, on a montré comment la dynamique spatiale des températures (conséquence de déplacement de masses d'air) pouvait - en théorie, s'observer sur les autocorrélations croisées des résidus des modèles locaux de température. On retrouve alors effectivement l'influence océanique dominante du climat de la France, ainsi, semble-t-il, qu'une influence Nord-Sud. Ces influences paraissent réparties d'une façon relativement uniforme sur l'année, aux quelques variations près indiquées au paragraphe précédent.

Pour améliorer la dynamique de température à Lyon, il semble donc pertinent, d'utiliser une station située à l'Ouest (tenant ainsi compte de l'influence océanique) par exemple Bordeaux, mais aussi une ville située au nord comme Paris. En revanche, il semble parfaitement inutile d'adjoindre une ville située au sud comme Marseille. Nous aurons la confirmation de ceci à la fois par le résultat d'estimation du modèle VAR, et par la comparaison des prévisions de différents modèles VAR. Dans la suite, nous avons sélectionné le système {Lyon, Bordeaux, Paris} pour le modèle multivarié.

## III. Modélisation VAR.

### 1. Présentation du modèle. Amélioration par rapport au modèle local AR.

Le modèle VAR( $p$ ) (Vector AutoRegressive) s'écrit

$$\mathbf{Z}_t = A_1 \mathbf{Z}_{t-1} + \dots + A_p \mathbf{Z}_{t-p} + \mathbf{U}_t$$

où  $\mathbf{Z}_t = (Z_{1,t}, \dots, Z_{K,t})'$  est un vecteur  $K \times 1$ , les  $A_i$ ,  $i = 1, \dots, p$  sont des matrices à coefficients constants de taille  $K \times K$  et  $\mathbf{U}_t = (U_{1,t}, \dots, U_{K,t})'$  est un bruit blanc  $K$ - dimensionnel, autrement dit tel que  $E(\mathbf{U}_t) = 0$ ,  $E(\mathbf{U}_t \mathbf{U}_s') = 0$  si  $t \neq s$  et  $E(\mathbf{U}_t \mathbf{U}_t') = \Sigma_U$ . La matrice de variances-

covariances  $\Sigma_U$  est supposée inversible. On suppose de plus que le processus  $(U_t)$  est gaussien.

Dans notre situation,  $Z_{1,t}$  représente la température (stationnarisée) à Lyon ou Paris et  $Z_{2,t}, \dots, Z_{K,t}$  la température d'autres stations. Dans la suite, la présentation est axée sur le cas où  $Z_{1,t} = Z_{Lyo,t}$ ,  $Z_{2,t} = Z_{Bor,t}$  et  $Z_{3,t} = Z_{Par,t}$  comme annoncé à la fin du paragraphe précédent mais nous examinons également d'autres situations.

Choisissons donc, par exemple, le cas de {Lyon, Bordeaux, Paris} et imaginons que le modèle VAR( $p$ ) soit connu. On peut alors juger de l'importance de l'amélioration par rapport au modèle local AR, en étudiant la pertinence du système constitué des variables supplémentaires. Ceci est lié à la notion de causalité au sens de Granger, définie dans (Lütkepohl, 1993). On dit que {Bordeaux, Paris} est causal pour Lyon si l'erreur de prévision augmente strictement lorsque l'on retire l'information relative à Bordeaux et Paris, c'est-à-dire si pour au moins un horizon  $h$ ,

$$\text{var}\left(Z_{Lyo,t+h} - E\left[Z_{Lyo,t+h} \mid \underline{Z_{Lyo,t}}, \underline{Z_{Bor,t}}, \underline{Z_{Par,t}}\right]\right) < \text{var}\left(Z_{Lyo,t+h} - E\left[Z_{Lyo,t+h} \mid \underline{Z_{Lyo,t}}\right]\right)$$

où  $\underline{Z_{s,t}}$  désigne l'information passée relative à la station  $s$ .

Dans le cas d'un processus VAR( $p$ ), la causalité se lit sur les coefficients des matrices  $A_t$ , et dire que {Bordeaux, Paris} est causal pour Lyon est équivalent à dire que les coefficients  $A_{i,12}$  et  $A_{i,13}$  ne sont pas tous nuls (Lütkepohl, 1993, § 2.3.1b). En particulier, si les nombres  $A_{i,12}$  et  $A_{i,13}$  sont tous nuls, on ne peut espérer réduire l'erreur de prévision de la température à Lyon, ce qui est clair car dans ce cas l'expression de  $Z_{Lyo,t}$  ne fait plus intervenir ni la température de Bordeaux ni celle de Paris. En revanche, il n'est pas vrai que si les  $A_{i,12}$  seulement sont nuls, alors l'erreur de prévision ne puisse pas diminuer lorsque l'on retire l'information relative à Bordeaux ; en effet, l'expression de  $Z_{Lyo,t}$  fait encore intervenir l'une des variables  $Z_{Par,t-1}, \dots, Z_{Par,t-p}$  et donc, de proche en proche, l'une des variables  $Z_{Bor,t-2}, \dots, Z_{Bor,t-p-1}$  ce qui fait que la prévision pour un certain horizon  $h \geq 2$  fait intervenir la température de Bordeaux. Ainsi la causalité au sens de Granger ne permet pas d'apprécier la pertinence de l'ajout d'une seule variable mais du *système* constitué des

variables supplémentaires. Remarquons en outre qu'elle ne renseigne en aucune façon sur le degré de pertinence du système. Pour cela on peut regarder l'importance de la diminution des erreurs de prévision aux horizons  $h = 1, 2, \dots$

En pratique, le modèle VAR est estimé et il existe des tests statistiques de causalité (voir par exemple Lütkepohl, 1993, § 3.6.1). Cependant, l'appréciation que nous souhaitons porter est davantage de nature qualitative que décisionnelle. On jugera de l'amélioration du modèle VAR par rapport au modèle AR selon deux aspects :

1. La signification des coefficients  $A_{i,1j}$ ,  $j > 1$ ,  $i = 1, 2, \dots, p$
2. La diminution de l'erreur de prévision à l'horizon 1 lorsqu'on passe du modèle local AR au modèle multivarié VAR ; on compare donc  $\text{var}(U_{Ly\sigma,t})$  à  $\text{var}(\varepsilon_{Ly\sigma,t})$ .

Ceci donne un avant-goût de l'étude de comparaison des prévisions – beaucoup plus complète, qui sera réalisée au paragraphe suivant. Cependant, il s'agit plutôt ici de confirmer la sélection faite au moyen des techniques descriptives présentées précédemment que de comparer des prévisions. En outre, celles-ci sont, dans ce paragraphe, in-sample, c'est-à-dire que l'on compare les prévisions avec des températures de modèle (et non les températures réellement observées). Le résultat est donc dépendant de la validité des modèles considérés.

## **2. Résultats d'estimation.**

On donne maintenant des résultats d'estimation du modèle VAR( $p$ ) pour différents systèmes. L'estimation relative à {Lyon, Bordeaux, Paris} est détaillée ci-dessous ; les autres résultats figurent en annexe. L'estimation est réalisée de façon standard par moindres carrés (Lütkepohl, 1993, § 3.2.). Elle prend en compte l'existence d'une moyenne non nulle dans le processus  $(Z_t)$  ; toutefois comme ce processus est centré, la moyenne estimée est quasi-nulle et nous ne l'avons pas mentionnée dans les résultats. La sélection de l'ordre du modèle,  $p$ , aurait pu se faire au moyen d'un critère AIC ou SBC. Néanmoins, nous avons choisi la valeur de  $p$  en nous basant sur le résultat de la sélection des modèles locaux donnant  $p = 3$  ; c'est cette valeur qui a été retenue. On a vérifié qu'au-delà, les coefficients estimés sont, en majorité, non significatifs.

L'estimation du modèle VAR(3) pour {Lyon, Bordeaux, Paris} conduit aux résultats donnés ci-dessous. On a également mentionné les résultats d'estimation des modèles individuels. Par souci de clarté, l'erreur d'estimation n'est pas donnée mais les coefficients non significatifs sont signalés et mis entre parenthèses. Pour chaque coefficient, elle est comprise entre 0.01 et 0.02 que ce soit pour le modèle VAR(3) ou chacun des modèles AR(3). En particulier, on a confirmation du fait que le passage au modèle VAR conduit à multiplier le nombre de paramètres, avec une erreur d'estimation du même ordre pour chacun d'eux. On observe que le système {Bordeaux, Paris} est « largement » causal pour Lyon ; en outre, la variance de l'erreur de prévision à 1 pas chute de 0.3936 pour le modèle univarié AR(3) à 0.2866 pour le modèle VAR(3). En revanche, on remarque que si l'objectif avait été d'améliorer la modélisation de la température à Bordeaux, la prise en compte des températures de Paris et Lyon aurait été quelque peu inefficace puisque la variance de l'erreur passe de 0.4187 à 0.4006 seulement en rajoutant cette information.

**Tableau 1.4. Résultats d'estimation du modèle VAR(3) pour le système {Lyon, Bordeaux, Paris}.**

$$\begin{bmatrix} Z_{Lyo,t} \\ Z_{Bor,t} \\ Z_{Par,t} \end{bmatrix} = \begin{bmatrix} 0.4392 & 0.3403 & 0.3888 \\ (0.0031) & 0.7135 & 0.2556 \\ -0.0426 & 0.2756 & 0.7796 \end{bmatrix} \begin{bmatrix} Z_{Lyo,t-1} \\ Z_{Bor,t-1} \\ Z_{Par,t-1} \end{bmatrix} + \begin{bmatrix} (0.0059) & -0.1775 & -0.1866 \\ -0.0628 & -0.1119 & -0.1380 \\ (-0.0333) & -0.1362 & -0.1314 \end{bmatrix} \begin{bmatrix} Z_{Lyo,t-2} \\ Z_{Bor,t-2} \\ Z_{Par,t-2} \end{bmatrix} \\
 + \begin{bmatrix} 0.0648 & (0.0073) & (0.0068) \\ 0.0438 & 0.0679 & (0.0151) \\ 0.0371 & (-0.0101) & 0.0734 \end{bmatrix} \begin{bmatrix} Z_{Lyo,t-3} \\ Z_{Bor,t-3} \\ Z_{Par,t-3} \end{bmatrix} + \begin{bmatrix} U_{Lyo,t} \\ U_{Bor,t} \\ U_{Par,t} \end{bmatrix}$$

avec  $\Sigma_U = \begin{bmatrix} 0.2866 & 0.1650 & 0.1381 \\ 0.1650 & 0.4001 & 0.1831 \\ 0.1381 & 0.1831 & 0.3437 \end{bmatrix}$ .

Comparaison avec les modèles individuels :

$$\begin{bmatrix} Z_{Lyo,t} \\ Z_{Bor,t} \\ Z_{Par,t} \end{bmatrix} = \begin{bmatrix} 0.8615 & 0 & 0 \\ 0 & 0.8353 & 0 \\ 0 & 0 & 0.9107 \end{bmatrix} \begin{bmatrix} Z_{Lyo,t-1} \\ Z_{Bor,t-1} \\ Z_{Par,t-1} \end{bmatrix} + \begin{bmatrix} -0.1507 & 0 & 0 \\ 0 & -0.1436 & 0 \\ 0 & 0 & -0.1992 \end{bmatrix} \begin{bmatrix} Z_{Lyo,t-2} \\ Z_{Bor,t-2} \\ Z_{Par,t-2} \end{bmatrix} \\
 + \begin{bmatrix} 0.0545 & 0 & 0 \\ 0 & 0.0610 & 0 \\ 0 & 0 & 0.0658 \end{bmatrix} \begin{bmatrix} Z_{Lyo,t-3} \\ Z_{Bor,t-3} \\ Z_{Par,t-3} \end{bmatrix} + \begin{bmatrix} \varepsilon_{Lyo,t} \\ \varepsilon_{Bor,t} \\ \varepsilon_{Par,t} \end{bmatrix}$$

avec  $\text{var}(\varepsilon_{Lyo,t}) = 0.3936$ ,  $\text{var}(\varepsilon_{Bor,t}) = 0.4187$  et  $\text{var}(\varepsilon_{Par,t}) = 0.3656$ .

D'autre part, dans l'annexe de fin de chapitre figurent les résultats d'estimation pour les systèmes {Lyon, Bordeaux}, {Lyon, Paris} et {Lyon, Marseille}. Les conclusions tirées des autocorrélations croisées des résidus sont pleinement confirmées. Lyon ou Paris font ainsi chuter l'erreur de prévision tandis que l'ajout de Marseille (seule) ne la modifie quasiment pas ; toutefois, remarquons que le coefficient  $A_{12}$  est significatif, semblant indiquer une (faible) causalité au sens de Granger de Marseille vers Lyon. Enfin remarquons que dans la réduction de variance obtenue avec l'incorporation de Bordeaux et Paris, égale à  $-0.11$ , la part la plus importante revient à l'ajout d'une seule ville, Bordeaux ( $-0.7$ ) ou Paris ( $-0.8$ ), la diminution réalisée en passant d'une ville à deux villes n'étant que de  $-0.4$  environ, soit deux fois moindre. Cette remarque se généralise à d'autres exemples comme on le verra plus tard, et va dans le sens d'une économie de variables : le passage à 4, 5, 6 stations ou plus est coûteux (en terme de qualité d'estimation) et conduit à une amélioration somme toute modeste par rapport à un système de 3 stations.

### 3. Validation.

Dans ce paragraphe, nous examinons la validité du modèle VAR. En fait, pour l'objectif poursuivi ici, cette étape n'est pas strictement nécessaire. En revanche, elle serait indispensable dans le cadre de la gestion de portefeuilles de dérivés climatiques. En outre, il est utile de savoir si un modèle linéaire aussi simple que celui-ci peut suffire à décrire la dynamique spatio-temporelle des températures. On s'attend bien évidemment à des difficultés puisque le résultat de la validation des modèles locaux avait déjà conduit à émettre des réserves.

On doit vérifier que les résidus  $(U_t)$  du modèle VAR forment un bruit blanc gaussien multidimensionnel. Cette hypothèse implique en particulier les deux propriétés de non-corrélation et de normalité :

- « non corrélation » Pour tous  $1 \leq i, j \leq K$ , et  $h > 0$ ,  $\text{corr}(U_{i,t}, U_{j,t-h}) = 0$
- « normalité » Pour tout  $t$ ,  $U_t$  suit une loi normale (multidimensionnelle).

La propriété de non-corrélation signifie la nullité des autocorrélations temporelles et des autocorrélations spatiales retardées. Seules peuvent subsister des autocorrélations spatiales instantanées. La dynamique spatiale est donc intégrée par le modèle VAR, on ne doit plus l'observer dans ses résidus comme ce fut le cas avec les modèles locaux (paragraphe 1, ci-



dessus). La propriété de normalité spécifie qu'à un instant donné, le vecteur des résidus du modèle VAR relatifs à l'ensemble des stations est de loi normale. Il s'agit d'une propriété plus forte que la normalité marginale des résidus correspondant à une station donnée.

Comme dans le cas univarié, les tests statistiques sont employés pour éprouver les deux propriétés mentionnées ci-dessus de non-corrélation et de normalité des résidus. Ils sont présentés en détail dans (Lütkepohl, 1993, § 4.4.). Le fait qu'ils s'appuient sur des propriétés asymptotiques n'est pas gênant ici car la taille de l'échantillon,  $N = 7665$ , est relativement « grande ». On présente ci-dessous les résultats pour {Lyon, Bordeaux, Paris}, la conclusion étant identique dans les autres cas.

- **Non-corrélation**

Un premier test consiste à vérifier que les autocorrélations croisées des résidus estimés se trouvent « majoritairement » dans les bandes de confiance à 95%,  $\pm \frac{2}{\sqrt{N}}$  (Lütkepohl, 1993, § 4.4.2.). Ceci est assez bien vérifié (Figure 1.29) ; en particulier, selon ce test qui traite les autocorrélations de façon « individuelle », les autocorrélations spatiales retardées ne sont pas significatives. En revanche, le test de Portmanteau (Lütkepohl, 1993, § 4.4.3.), qui s'intéresse à des paquets d'autocorrélations, donne des  $p$ -valeurs faibles : de l'ordre de 1% lorsque la largeur des paquets est de 20, 50 ou 100. L'indépendance des résidus est donc rejetée au niveau 5%, et tolérée au niveau 1%.

- **Normalité**

Le tracé des droites de Henri des lois marginales fait ressortir les mêmes observations que dans le cas univarié, à savoir principalement des queues de distributions trop épaisses pour une loi normale du côté des températures basses. Le test de Kolmogorov au niveau 5% conduit au rejet de la normalité. Par ailleurs, la normalité des lois marginales (à supposer qu'elle soit validée) ne suffit pas à garantir la normalité multidimensionnelle. On a donc utilisé le test basé sur *skewness* et *kurtosis* pour une loi multidimensionnelle (Lütkepohl, 1993, § 4.5.2). Le résultat est sans appel : les  $p$ -valeurs sont à peu près... nulles. La normalité est donc rejetée de façon claire.

En conclusion, même si la Figure 1.29 montre que les résidus sont relativement « contrôlés », les tests statistiques utilisés conduisent indiscutablement au rejet du modèle VAR.

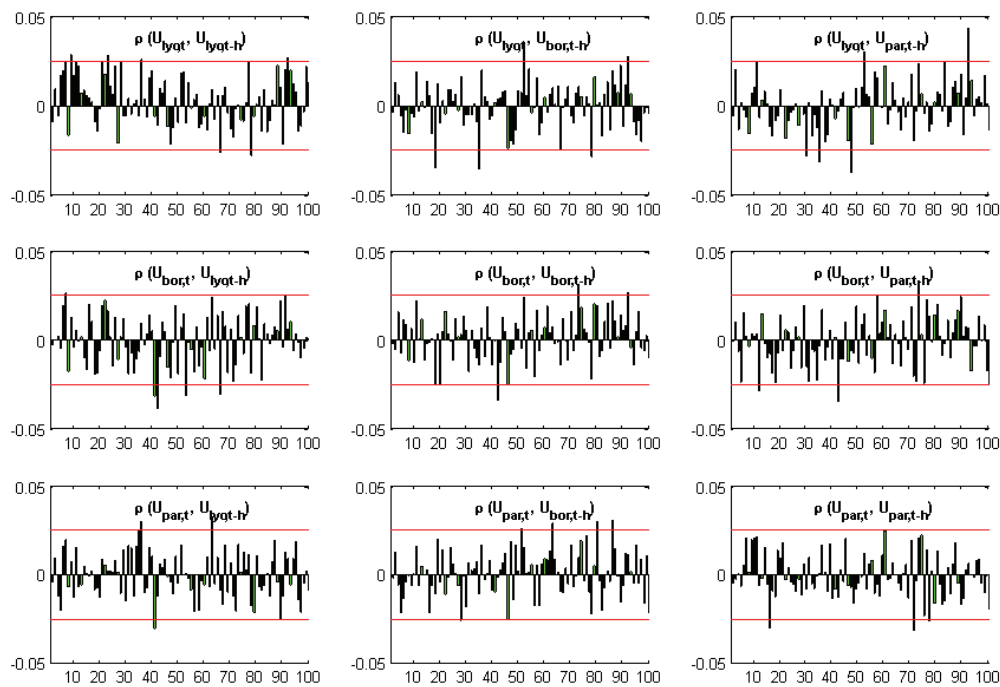


Figure 1.29. Autocorrélations croisées des résidus du modèle VAR pour le système {Lyon, Bordeaux, Paris}

#### IV. Comparaison des prévisions.

Pour apprécier l'amélioration de la modélisation de la température lorsqu'on intègre la dimension spatiale, nous comparons les prévisions de température à quelques jours des modèles AR et VAR. De façon à mieux relativiser les différences, nous les comparons avec les prévisions - normalement plus performantes - de Météo-France, et avec la normale saisonnière (par rapport à laquelle on espère évidemment faire mieux !). Il s'agit ici de la normale saisonnière de modèle, c'est-à-dire correspondant à la partie déterministe du modèle de température local :  $NS_t = m_t + s_t$  ; avec ce choix, on observera ainsi directement l'amélioration due à la dynamique aléatoire pure des processus AR et VAR. Dans un premier temps, la simple visualisation des prévisions nous permet de comprendre le « comportement » des modèles et nous faire un premier avis. Ensuite, nous comparons de façon précise les prévisions des différents modèles avec les températures réellement observées dans une analyse post-sample détaillée (sur la problématique de comparaison des prévisions et l'analyse post-sample, on pourra consulter Makridakis, Whellwright et Hyndman, 1998,

§7.8.2). A l'exception du tout dernier paragraphe, les résultats présentés ici concernent le cas de Lyon, avec Bordeaux et Paris comme stations supplémentaires.

### 1. Visualisation des prévisions.

Donnons tout d'abord quelques précisions sur le calcul des prévisions. Si  $t$  désigne le jour courant et  $h$  l'horizon de prévision, on note  $p_t(AR, h)$ ,  $p_t(VAR, h)$ ,  $p_t(MF, h)$  et  $p_t(NS, h)$  la prévision faite le jour  $t$  à l'horizon  $h$  (c'est-à-dire pour le jour  $t+h$ ) obtenues avec les modèles AR, VAR, le modèle de Météo-France et avec la normale saisonnière définie ci-dessus. Avec les autres notations de ce chapitre, on a alors :

- $p_t(NS, h) = m_{t+h} + s_{t+h}$
- $p_t(AR, h) = m_{t+h} + s_{t+h} + \sigma_{t+h} \times E \left[ Z_{Lyo, t+h} \mid \underline{Z_{Lyo, t}} \right]$
- $p_t(VAR, h) = m_{t+h} + s_{t+h} + \sigma_{t+h} \times E \left[ Z_{Lyo, t+h} \mid \underline{Z_{Lyo, t}}, \underline{Z_{Bor, t}}, \underline{Z_{Par, t}} \right]$

où  $(Z_{Lyo, t}, Z_{Bor, t}, Z_{Par, t})'$  est un processus VAR(3) modélisant le vecteur des températures stationnarisées à Lyon, Bordeaux, Paris. Pour ne pas alourdir les notations, on a également écrit  $Z_{Lyo, t}$  dans le cas AR, bien qu'il s'agisse de deux processus différents. L'espérance conditionnelle est calculée de proche en proche en utilisant une relation de récurrence linéaire (voir Lütkepohl, § 2.2.) à partir de la température actuelle, et des deux jours passés.

On a représenté sur la Figure 1.30 les prévisions obtenues sur la période du 14/07/99 au 24/07/99, pour les horizons  $h = 1, 2, \dots, 5$  jours. On constate immédiatement que les prévisions des modèles AR et VAR « fonctionnent » de façon très semblable, en convergeant vers la normale saisonnière – la *moyenne* - d'autant plus rapidement que la température courante en est éloignée. Sur cet exemple, cela se traduit par une sous-estimation des variations futures de la température. Au contraire les prévisions de Météo-France sont beaucoup plus « réactives » : lorsque la température s'écarte de la normale saisonnière, on anticipe un écart dans l'autre sens ; les prévisions traversent la moyenne. Ici, les variations de température sont même surestimées. Elles paraissent, à vue d'œil, beaucoup plus précises. A partir de ce graphique, nous voyons déjà que les améliorations apportées par le modèle VAR risquent d'être faibles. Le gain principal semble se situer à l'horizon 1. On constate en effet que les prévisions correspondantes suivent les variations de la température en s'écartant de la

moyenne lorsque la température s'en écarte. Ce phénomène, qui ne pourrait pas être observé avec un modèle univarié de type AR, traduit l'incorporation d'une information supplémentaire pertinente.

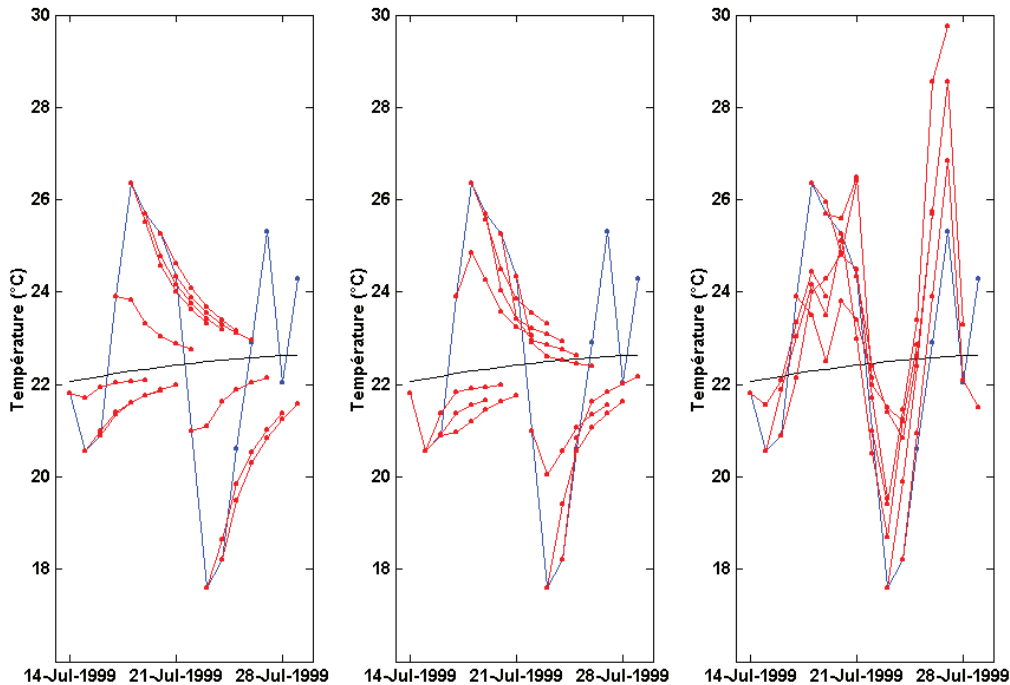
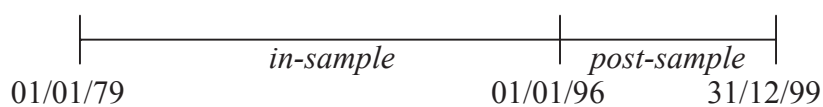


Figure 1.30 . Prévisions de température jusqu'à 5 jours pour les modèles AR, VAR et Météo France

## 2. Analyse post-sample.

L'examen de la Figure 1.30 a permis de faire une première comparaison des prévisions des modèles AR et VAR. Pour aller plus loin, il est indispensable de réaliser une étude systématique sur une période de temps plus longue. C'est ce qui est fait dans ce paragraphe. De façon précise, on compare les erreurs réelles de prévision sur une période post-sample de 4 ans (du 01/01/96 au 31/12/99) postérieure à la période utilisée – in-sample- pour estimer les modèles (du 01/01/79 au 31/12/95).



Une comparaison globale des erreurs de prévision est d'abord réalisée. Ensuite, afin de comprendre d'où proviennent les améliorations, on distingue selon la saison puis selon l'écart par rapport à la normale saisonnière. L'étude est conduite pour le cas de Lyon avec le rajout de Bordeaux et Paris pour le modèle VAR. Pour finir, on présente quelques résultats globaux obtenus avec d'autres choix de stations.

a) Comparaison globale des prévisions.

Notons  $M$  la période post-sample, allant du 01/01/96 au 31/12/99, et désignons par  $x_t$ ,  $t \in M$ , la température réellement observée au jour  $t$ . L'erreur réelle de prévision à l'horizon  $h$  faite au jour  $t$ ,  $t \in M$ , est alors donnée par

$$p_t(\cdot, h) - x_{t+h}$$

où le point symbolise l'un des 4 modèles présentés auparavant. Pour mesurer l'erreur globale de prévision sur toute la période  $M$ , nous utilisons les deux indicateurs usuels d'écart-type et de biais :

$$\text{RMSE}(\cdot, h) = \sqrt{\frac{1}{\#M} \sum_{t \in M} (p_t(\cdot, h) - x_{t+h})^2}, \quad h = 1, 2, \dots, 5$$

$$\text{Biais}(\cdot, h) = \frac{1}{\#M} \sum_{t \in M} (p_t(\cdot, h) - x_{t+h}), \quad h = 1, 2, \dots, 5$$

Les résultats sont donnés sur la Figure 1.31 et le Tableau 1.5. On constate d'abord que le biais de prévision est négligeable pour tous les modèles. D'autre part, il y a une amélioration nette du modèle VAR par rapport au modèle AR à l'horizon 1. En revanche, l'amélioration est beaucoup moins sensible dès l'horizon 2 lorsqu'on effectue la comparaison avec les prévisions de Météo-France et la normale saisonnière. De l'ordre de 8% pour  $h = 2$ , elle est inférieure à 5% pour  $h \geq 3$  (voir Tableau 1.9). Ces premières conclusions peu enthousiasmantes vis-à-vis de l'utilisation d'un modèle VAR généralisent les constats effectués sur l'exemple de la Figure 1.30.

b) Comparaison des prévisions selon le mois de l'année.

On souhaite voir ici si l'efficacité des prévisions dépend du mois où elles sont réalisées. On considère donc 12 sous-ensembles  $M$  de dates dans la période *post-sample*, notés  $M_{\text{janvier}}$ ,

$M_{\text{février}}, \dots, M_{\text{décembre}}$ , obtenus en collectant les dates de  $M$  relatives à un mois donné. On utilise les indicateurs :

$$\text{RMSE}(., \text{mois}) = \sqrt{\frac{1}{\#M_{\text{mois}}} \frac{1}{5} \sum_{t \in M_{\text{mois}}} \sum_{h=1}^5 (p_t(., h) - x_{t+h})^2}$$

$$\text{Biais}(., \text{mois}) = \frac{1}{\#M_{\text{mois}}} \frac{1}{5} \sum_{t \in M_{\text{mois}}} \sum_{h=1}^5 (p_t(., h) - x_{t+h})$$

Les résultats sont donnés sur la Figure 1.32 et le Tableau 1.6. On observe que l'erreur de prévision est plus faible en été qu'en hiver. Ceci traduit simplement la saisonnalité de la volatilité des températures. On ne constate pas, par ailleurs, d'amélioration du modèle VAR par rapport au modèle AR un mois plutôt qu'un autre. En outre, il est bien difficile d'établir une hiérarchie satisfaisante de l'efficacité mensuelle des prévisions des modèles AR et VAR par rapport à la normale saisonnière. Car si les cas de janvier et février incitent à penser que l'erreur relative est plus faible en hiver, l'exemple du mois de décembre constitue un contre-exemple flagrant.

Nous constatons d'autre part un biais négatif quasi-systématique pour l'hiver et le printemps, et positif pour l'été et l'automne. Ce biais est peu important pour Météo-France, mais semble amplifié pour les autres modèles. Or ces 3 modèles ont de commun la partie déterministe modélisant la normale saisonnière. Il se pourrait donc que l'on observe ici un défaut dans cet aspect de la modélisation. Pour illustrer cette hypothèse, on a comparé les biais de prévision pour la normale saisonnière « de modèle » utilisée jusqu'à présent, et la normale saisonnière « historique » obtenue en effectuant la moyenne des températures sur les 17 années d'historique. Les résultats, donnés dans le Tableau 1.7 montrent qu'avec la normale saisonnière historique, le biais de prévision est décalé négativement. Il est donc, en valeur absolue, plus important les six premiers mois et diminué les six derniers. Cela peut s'expliquer par le fait que la normale saisonnière historique (telle qu'elle est calculée ici) ne tient pas compte de la tendance au réchauffement. Les différences observées semblent indiquer, en tout cas, que les défauts de modélisation peuvent se répercuter de façon assez importante sur le biais des prévisions.

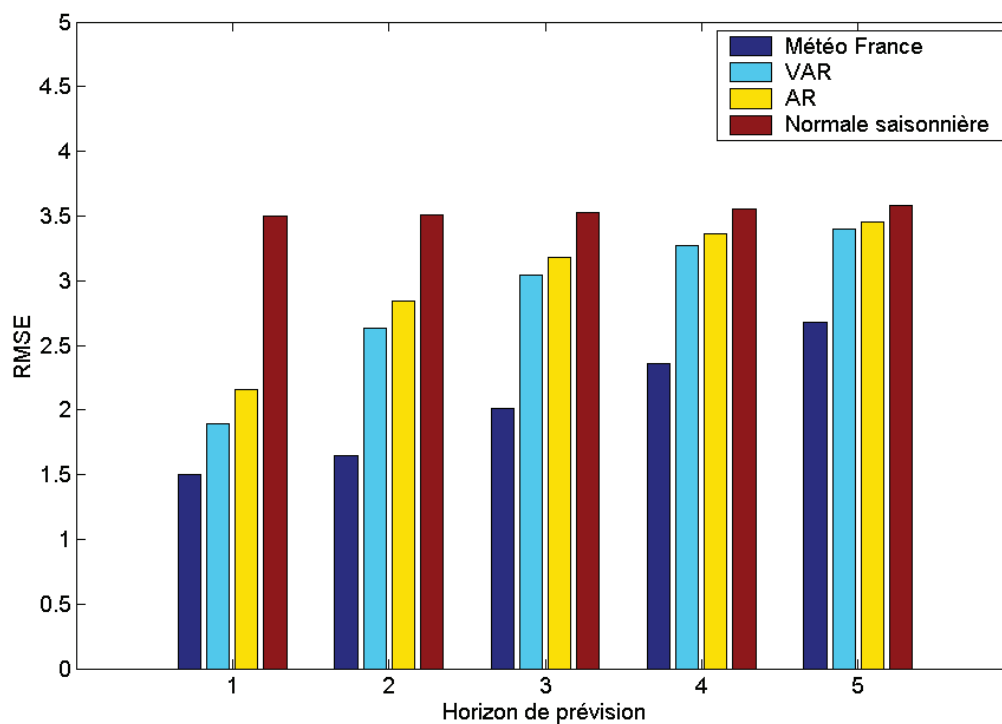


Figure 1.31. Comparaison des prévisions journalières de température de différents modèles

Tableau 1.5. Comparaison des prévisions journalières de température de différents modèles

Modèle	RMSE relatifs des prévisions					Moyenne des erreurs de prévision				
	Horiz. 1	Horiz. 2	Horiz. 3	Horiz. 4	Horiz. 5	Horiz. 1	Horiz. 2	Horiz. 3	Horiz. 4	Horiz. 5
Normale saisonnière	1.00	1.00	1.00	1.00	1.00	.13	.13	.13	.14	.14
AR	.62	.81	.90	.95	.97	.03	.06	.08	.10	.11
VAR	.54	.75	.86	.92	.95	-.09	-.08	-.04	-.01	-.01
Météo France	.43	.47	.57	.66	.75	-.18	-.08	-.03	-.01	-.03
RMSE Normale saisonnière	3.49	3.51	3.53	3.55	3.58					

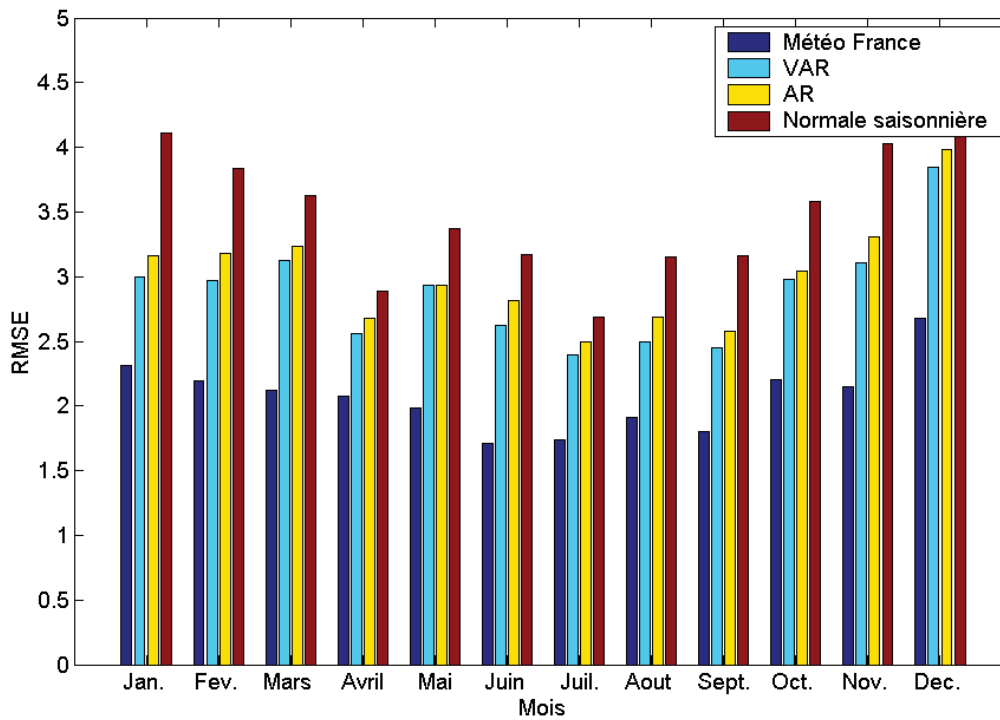


Figure 1.32. Comparaison des prévisions journalières de température de différents modèles selon le mois d'origine de la prévision

Tableau 1.6. Comparaison des prévisions journalières de température de différents modèles selon le mois d'origine des prévisions

Modèle	RMSE relatifs des prévisions aux horizons 1, 2, ..., 5 réalisées en un mois donné											
	Janvier	Février	Mars	Avril	Mai	Juin	Juillet	Août	Sept.	Oct.	Nov.	Déc.
Normale saisonnière	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
AR	.77	.83	.89	.93	.87	.89	.93	.85	.81	.85	.82	.91
VAR	.73	.77	.86	.89	.87	.83	.89	.79	.78	.83	.77	.88
Météo France	.56	.57	.58	.72	.59	.54	.65	.61	.57	.62	.53	.61
RMSE Normale Saisonnière	4.11	3.84	3.62	2.88	3.37	3.17	2.69	3.15	3.16	3.58	4.02	4.39
Modèle	Biais des prévisions aux horizons 1, 2, ..., 5 réalisées en un mois donné											
	Janvier	Février	Mars	Avril	Mai	Juin	Juillet	Août	Sept.	Oct.	Nov.	Déc.
Normale saisonnière	-.11	.07	-.90	-.15	1.52	.00	.46	-.03	1.26	.63	1.78	.19
AR	.13	-.31	-.54	-.22	-.91	-.03	.19	.02	.82	.51	1.13	.11
VAR	-.06	-.34	-.48	-.29	1.03	-.14	.01	.08	.77	.32	.80	-.15
Météo France	-.03	-.59	-.02	-.42	-.20	.11	-.16	-.22	.20	.25	.43	-.13

Tableau 1.7. Biais de prévision de deux normales saisonnières

Modèle	Biais des prévisions aux horizons 1, 2, ..., 5 réalisées en un mois donné											
	Janvier	Février	Mars	Avril	Mai	Juin	Juillet	Août	Sept.	Oct.	Nov.	Déc.
Norm. saisonnière	-.11	.07	-.90	-.15	1.52	.00	.46	-.03	1.26	.63	1.78	.19
Norm. saisonnière historique	-1.42	-.92	-1.08	-1.22	-2.21	-1.10	.18	-.91	-.00	.33	.85	-.21



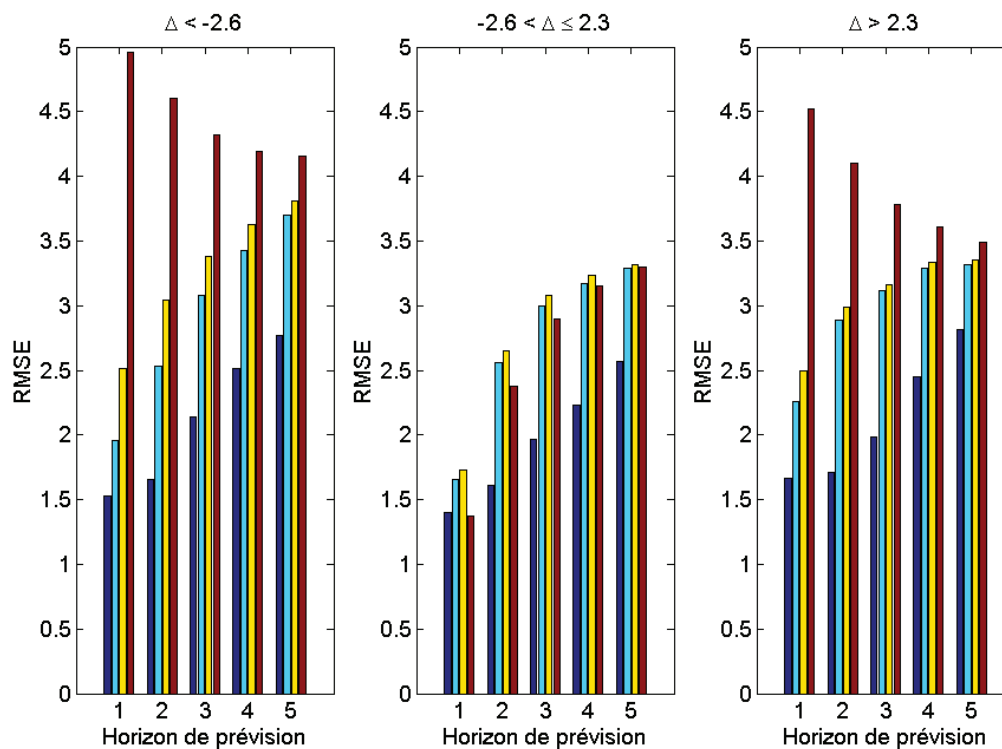


Figure 1.33. Comparaison des prévisions journalières de température de différents modèles selon l'écart à la normale saisonnière

Tableau 1.8. Comparaison des prévisions journalières de température des modèles AR et VAR suivant l'écart des températures par rapport à la normale saisonnière

		<i>RMSE relatifs des prévisions</i>				
		<i>Horizon 1</i>	<i>Horizon 2</i>	<i>Horizon 3</i>	<i>Horizon 4</i>	<i>Horizon 5</i>
AR		1.00	1.00	1.00	1.00	1.00
VAR	global	.88	.93	.96	.97	.98
	$\Delta < -2.6$	.78	.83	.91	.95	.97
	$-2.6 < \Delta \leq 2.3$	.96	.96	.97	.98	.99
	$\Delta > 2.3$	.90	.97	.99	.99	.99

c) Comparaison des prévisions selon l'écart à la normale saisonnière.

On a vu sur la Figure 1.30 que les prévisions des modèles AR et VAR tendent à converger vers la moyenne lorsque l'horizon de prévision croît, et que pour l'horizon 1 des différences assez nettes apparaissent. Nous allons montrer qu'on peut effectivement retrouver ce résultat en distinguant les prévisions selon l'éloignement de la température courante par rapport à la normale saisonnière. Pour ce faire, on a fractionné la période *post-sample*  $M$  en trois sous-périodes  $M_-$ ,  $M_0$ ,  $M_+$  selon l'importance de l'écart  $x_t - (m_t + s_t)$  de la température à la normale saisonnière :  $M_-$  est la première région interquartile,  $M_+$  la dernière, et  $M_0$  la réunion des 2<sup>ème</sup> et 3<sup>ème</sup> régions intermédiaires. On a ensuite calculé les indicateurs d'écart-type et de biais sur chacune de ces sous-périodes.

Les résultats sont donnés sur la Figure 1.33 et le Tableau 1.8. On observe tout d'abord que l'utilisation des modèles AR et VAR n'est véritablement utile pour la prévision que lorsque la température courante s'écarte suffisamment de la normale ; dans le cas contraire, l'erreur de prévision de ces modèles est en effet supérieure. La Figure 1.33 fait également apparaître une dissymétrie dans l'efficacité des prévisions selon le signe de l'écart : le modèle VAR apporte une amélioration importante par rapport au modèle AR lorsque la température courante est exceptionnellement basse. Il semble donc que l'information apportée par Paris et Bordeaux pour prévoir la température à Lyon soit particulièrement pertinente dans ce cas. Cela n'est pas tellement surprenant car une chute de température est souvent un signe traduisant le passage d'une perturbation ; les données d'autres stations permettent de tenir compte de cette dynamique spatiale dans la prévision.

d) Comparaison des prévisions selon le choix des stations.

Nous revenons ici sur les problèmes de la sélection des stations et du nombre de stations à utiliser dans le cadre de la modélisation VAR. Les études descriptives réalisées auparavant ont apporté une première réponse (II.1), qui a été complétée par les résultats d'estimation du modèle VAR (voir III.2). Une des façons d'apprécier l'importance de l'amélioration réalisée par l'ajout d'un système de stations a été de regarder l'importance de la diminution de l'erreur de prévision à l'horizon 1 lorsqu'on passe du modèle local AR au modèle multivarié VAR (III.1) III.1 Cependant, la non-validation du modèle VAR limite la portée de ces conclusions. Aussi reprenons-nous cette idée dans le cadre de l'analyse *post-sample*, en comparant les erreurs réelles de prévision des modèles AR et VAR. Notons qu'il s'agit ici de prévisions

« complètes », incorporant les composantes déterministes de la modélisation. En outre, en plus de l'horizon 1, on considère les horizons 2, 3, 4 et 5.

Les résultats figurent dans le Tableau 1.9 et le Tableau 1.10. Les conclusions sont identiques à celles du paragraphe III.2. D'une part, il ressort que la sélection préliminaire réalisée au moyen d'autocorrélations croisées, conduit effectivement à un choix pertinent de stations. D'autre part, le surenchérissement n'est pas souhaitable. Le gain réalisé en passant de 3 à 6 stations est quasiment négligeable. Enfin nous constatons que la situation est plus défavorable pour Paris (Tableau 1.10) en ce qui concerne l'utilisation d'un modèle VAR : les gains réalisés sont, de toute façon, plus faibles.

**Tableau 1.9. Comparaison de l'efficacité des prévisions journalières de température du modèle VAR suivant les stations utilisées. Cas de Lyon.**

<i>Stations utilisées</i>		<i>RMSE relatifs des prévisions</i>				
		<i>Horizon 1</i>	<i>Horizon 2</i>	<i>Horizon 3</i>	<i>Horizon 4</i>	<i>Horizon 5</i>
AR	Lyon	1.00	1.00	1.00	1.00	1.00
VAR	Lyon – Marseille	.99	.99	.99	1.00	1.00
	Lyon – Bordeaux	.92	.96	.98	.99	.99
	Lyon – Paris	.91	.94	.96	.97	.98
	Lyon – Bordeaux – Paris	.88	.92	.95	.97	.98
	Lyon – Bordeaux – Paris – Clermont-Ferrand – Brest – Lille	.86	.90	.94	.96	.98
Météo – France	.69	.58	.63	.70	.77	

**Tableau 1.10. Comparaison de l'efficacité des prévisions journalières de température du modèle VAR suivant les stations utilisées. Cas de Paris.**

<i>Stations utilisées</i>		<i>RMSE relatifs des prévisions</i>				
		<i>Horizon 1</i>	<i>Horizon 2</i>	<i>Horizon 3</i>	<i>Horizon 4</i>	<i>Horizon 5</i>
AR	Paris	1.00	1.00	1.00	1.00	1.00
VAR	Paris – Brest	0.95	0.97	0.98	0.99	0.99
	Paris – Brest – Bordeaux – Clermont-Ferrand – Strasbourg – Lille	0.92	0.96	0.98	0.98	0.99
Météo – France		0.68	0.57	0.65	0.73	0.83

## V. Conclusion de la section C.

Cette section, consacrée aux aspects multivariés, avait pour but premier d'étudier l'extension du modèle temporel de température à la dimension spatio-temporelle. Cette idée portait l'espoir de pouvoir nettement améliorer la dynamique de température en un lieu donné grâce à l'anticipation des déplacements de masses d'air, rendu possible par l'information apportée par d'autres stations. Nous avons alors considéré l'extension naturelle du modèle autorégressif utilisé pour la température locale, donc un modèle de type VAR, et examiné le gain obtenu grâce à l'ajout de l'aspect spatial dans la modélisation. Pour ce faire, nous avons comparé les prévisions *post-sample* des deux modèles statistiques au regard des prévisions météorologiques de Météo-France.

Il ressort d'une façon générale que l'amélioration attendue est en réalité peu importante même avec l'information de 5 ou 6 stations. La dynamique VAR incorpore effectivement les influences climatiques dominantes comme les déplacements Ouest-Est, mais celles-ci, pour les exemples étudiés, ne semblent avoir de conséquence sur les prévisions que lorsque la température courante est largement inférieure à la normale saisonnière. En outre, l'effet n'est sensible qu'à l'horizon 1. Finalement, à la vue du surcoût que crée le passage au spatial (une dizaine de paramètres par stations), il ne paraît pas rentable d'utiliser un modèle multivarié si l'objectif est d'améliorer la modélisation locale. On peut penser à sélectionner judicieusement les stations en limitant leur nombre ; nous avons vu comment le tracé des autocorrélations croisées des résidus des modèles locaux pouvait permettre une telle sélection. Cependant, le nombre de paramètres ajoutés reste élevé, et leur erreur d'estimation est du même ordre.

La modélisation VAR ne paraît donc pas pertinente pour l'évaluation d'un seul produit dérivé climatique. En revanche, elle peut constituer un premier pas pour la gestion de portefeuilles de dérivés climatiques. Dans ce cas, l'étude réalisée montre que des efforts devront être dirigés vers la nature de la loi du processus multivarié qu'on ne peut plus considérer comme normale. Par ailleurs, on pourrait penser à tenir compte de l'information spatiale dans la modélisation des composantes déterministes (celles-ci ont en effet été éliminées ici pour ne s'intéresser qu'à la partie stochastique). On a vu par exemple que celles-ci peuvent se traduire par un biais non négligeable dans les prévisions (Tableau 1.7). D'autre part, on verra par la suite qu'un défaut important du modèle AR réside dans l'importance de l'erreur d'estimation de ses composantes de tendance et de saisonnalité. Or l'information collective peut permettre

de réduire, globalement, l'erreur d'estimation individuelle. C'est l'idée des estimateurs de James-Stein (voir par exemple Efron, Morris, 1977, et une application à la finance dans Jorion, 1985). Peut-être peut-on penser à appliquer cette idée au problème.

**Annexe du chapitre 1, section C**

**Etudes descriptives – Résultats d'estimation**

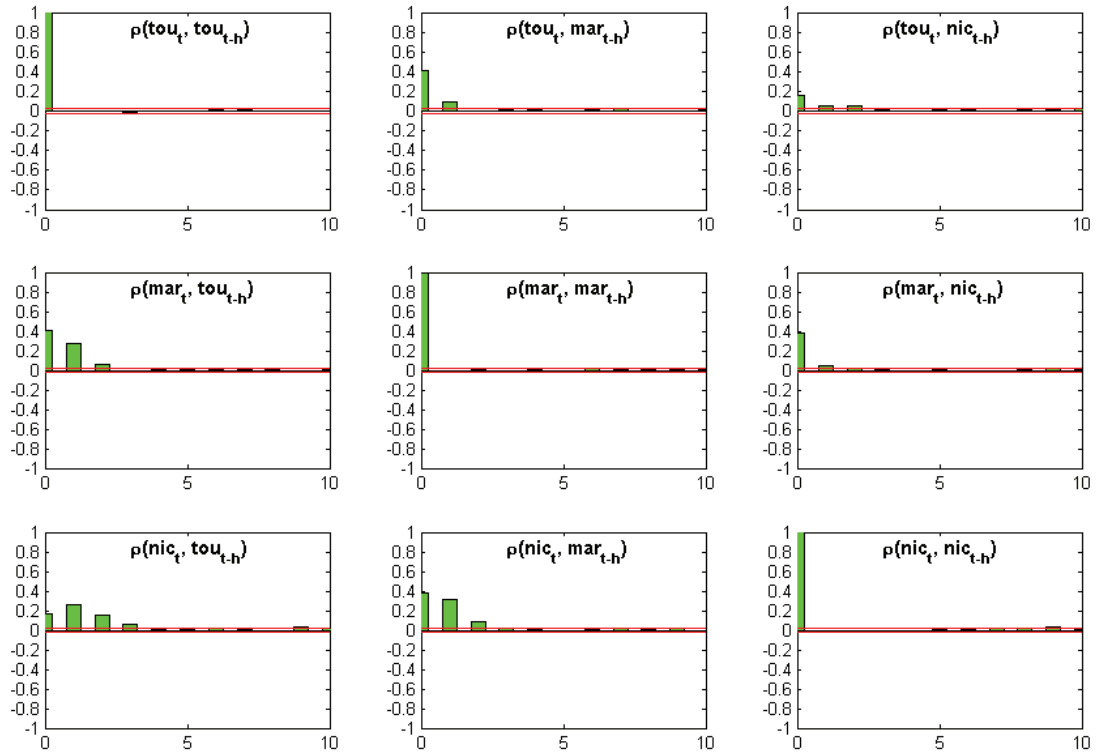
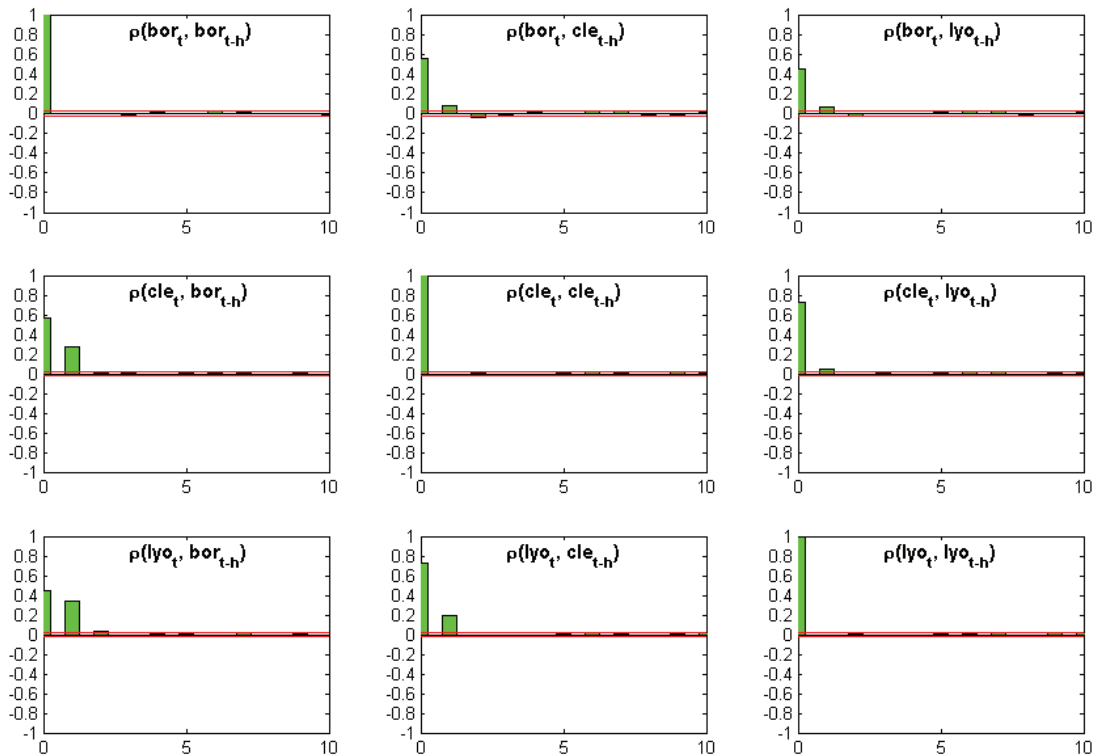


Figure 1.34. Autocorrélations croisées des résidus pour Toulouse, Marseille, Nice.

Figure 1.35. Autocorrélations croisées des résidus pour Bordeaux, Clermont-Ferrand, Lyon.



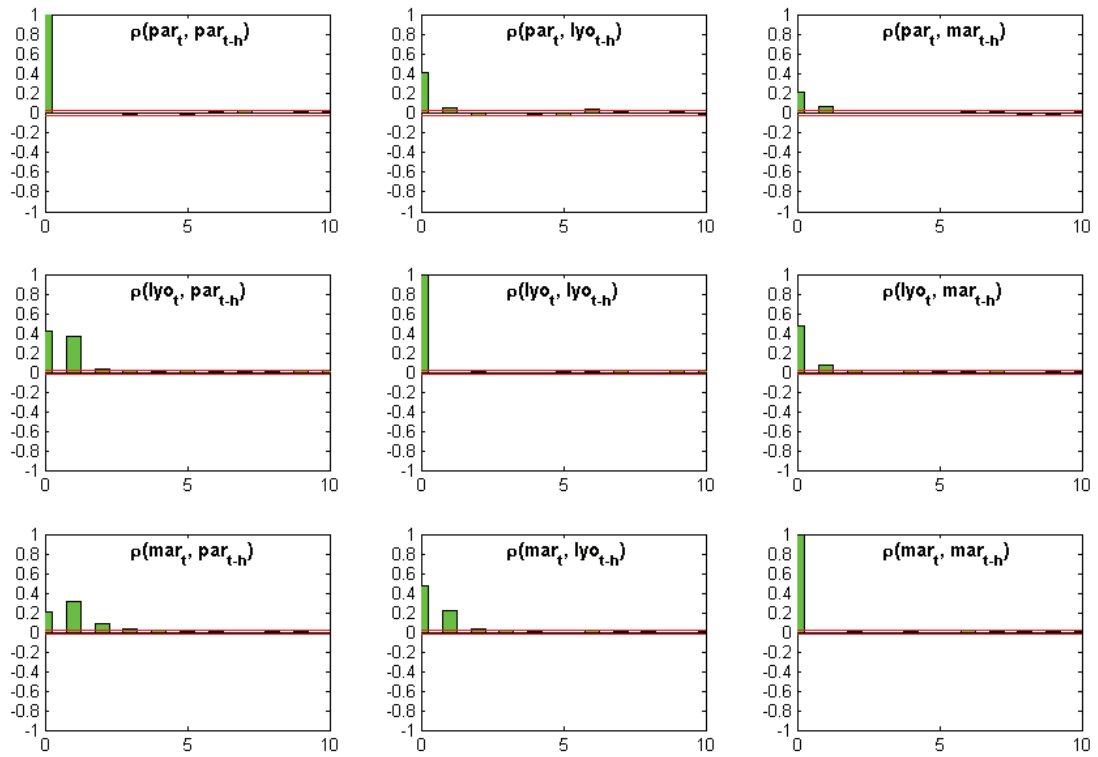


Figure 1.36. Autocorrélations croisées des résidus pour Paris, Lyon, Marseille.



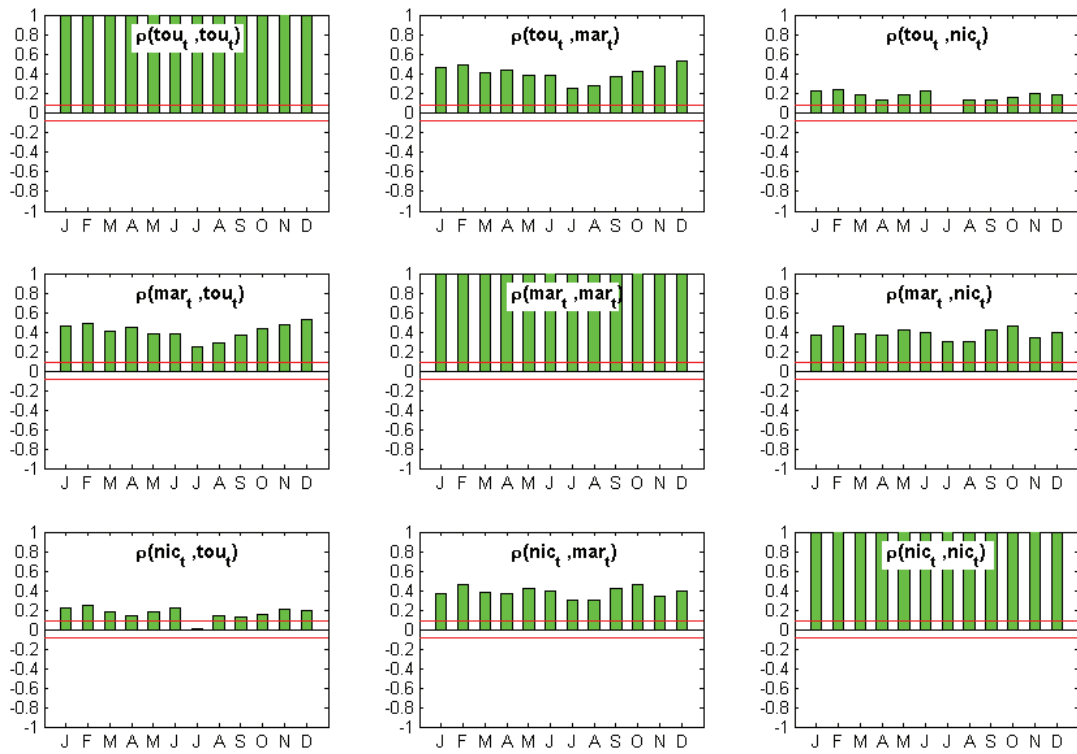
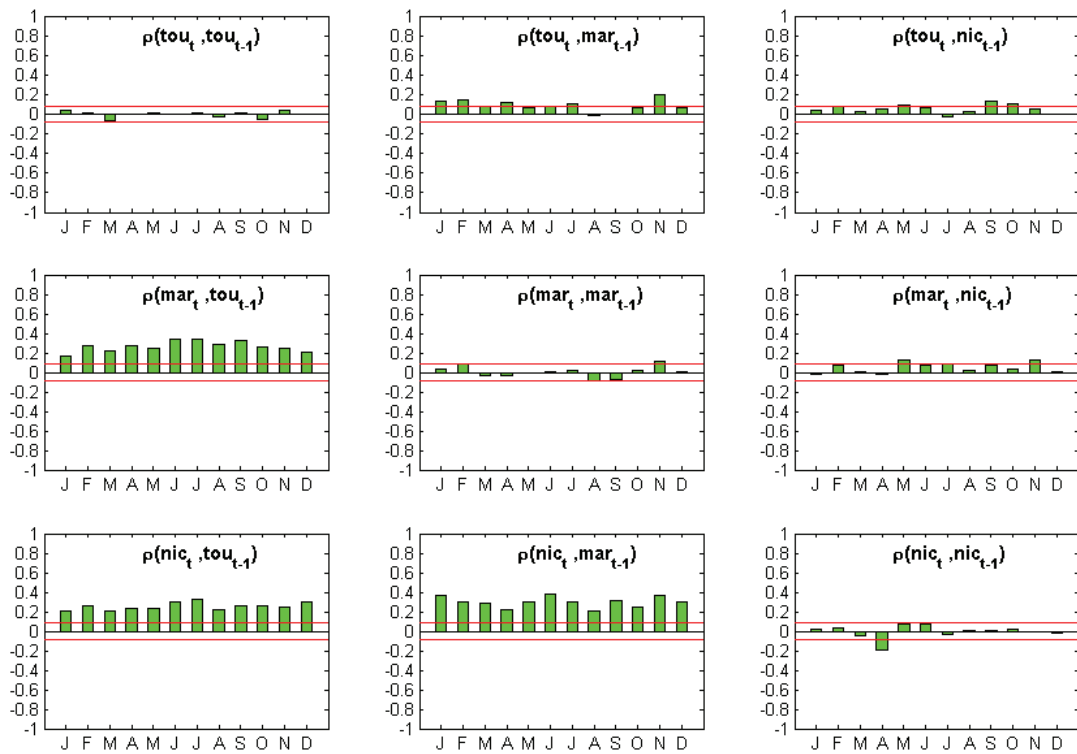


Figure 1.37. Fluctuations saisonnières des autocorrélations croisées instantanées des résidus des modèles locaux. Cas de Toulouse, Marseille, Nice.

Figure 1.38. Fluctuations saisonnières des autocorrélations croisées de décalage 1 des résidus des modèles locaux. Cas de Toulouse, Marseille, Nice.



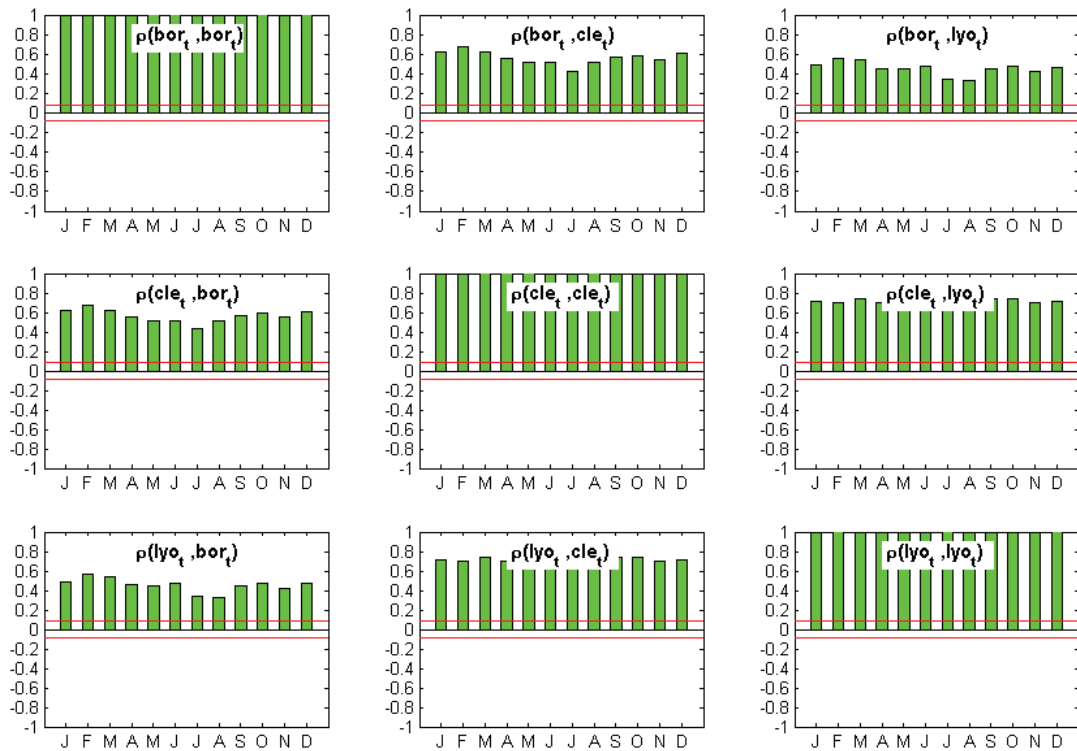
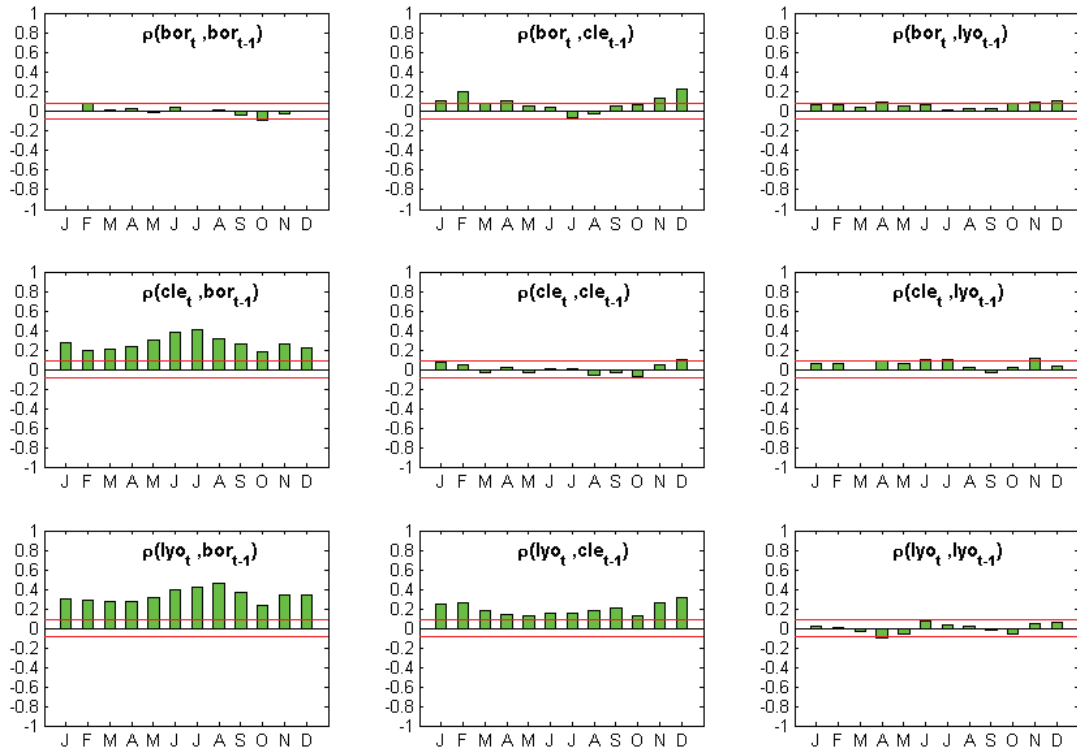


Figure 1.39. Fluctuations saisonnières des autocorrélations croisées instantanées des résidus des modèles locaux. Cas de Bordeaux, Clermont-Ferrand, Lyon.

Figure 1.40. Fluctuations saisonnières des autocorrélations croisées de décalage 1 des résidus des modèles locaux. Cas de Bordeaux, Clermont-Ferrand, Lyon.



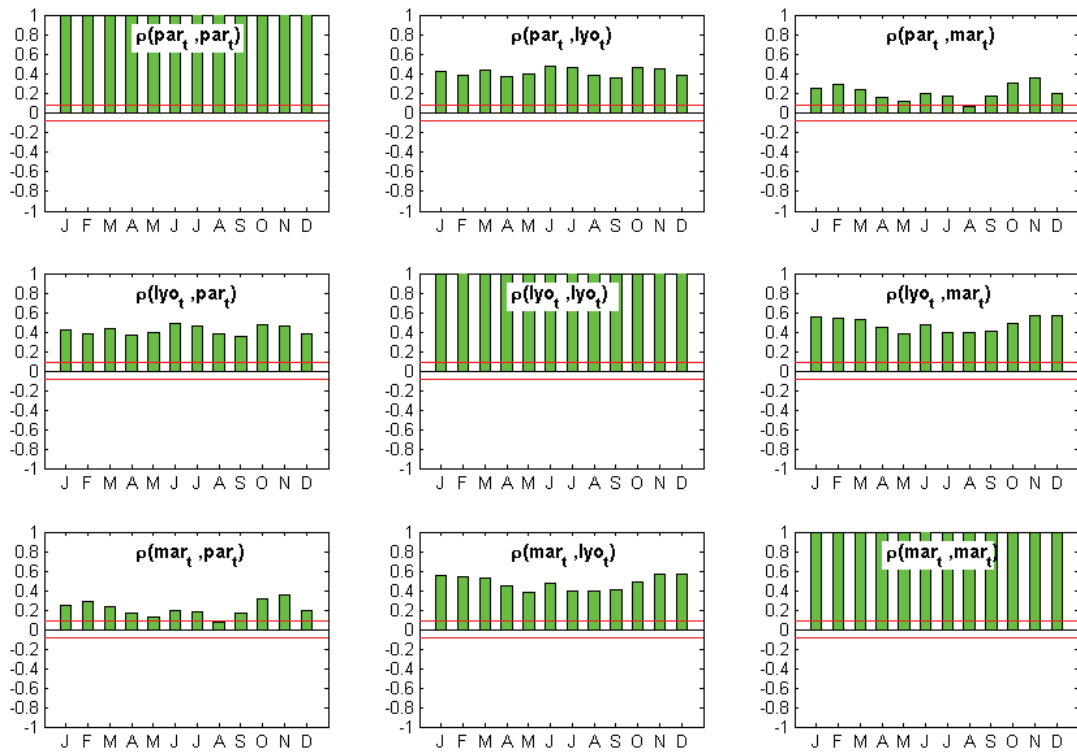
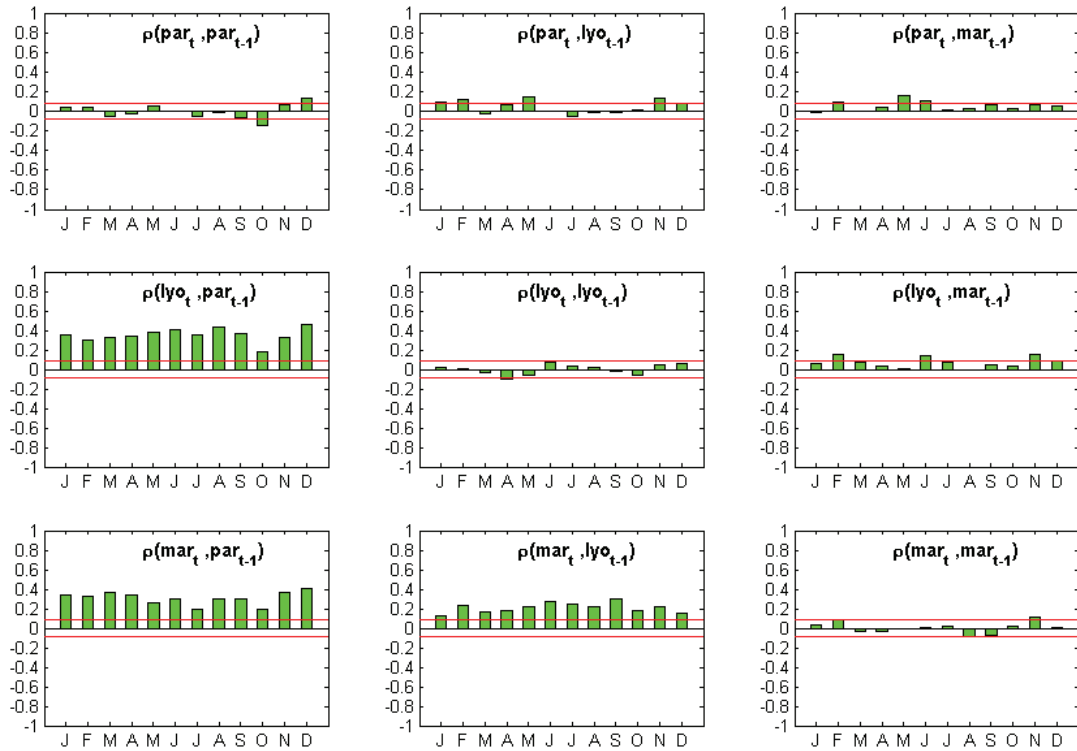


Figure 1.41. Fluctuations saisonnières des autocorrélations croisées instantanées des résidus des modèles locaux. Cas de Paris, Lyon, Marseille.

Figure 1.42. Fluctuations saisonnières des autocorrélations croisées de décalage 1 des résidus des modèles locaux. Cas de Paris, Lyon, Marseille.



## Résultats d'estimation de quelques modèles multivariés.

Les parenthèses indiquent les coefficients non significatifs au niveau 5%.

### 1. Système {Lyon, Bordeaux}

$$\begin{bmatrix} Z_{Lyo,t} \\ Z_{Bor,t} \end{bmatrix} = \begin{bmatrix} 0.5637 & 0.4731 \\ 0.0807 & 0.8006 \end{bmatrix} \begin{bmatrix} Z_{Lyo,t-1} \\ Z_{Bor,t-1} \end{bmatrix} + \begin{bmatrix} -0.0487 & -0.1958 \\ -0.1005 & -0.1269 \end{bmatrix} \begin{bmatrix} Z_{Lyo,t-2} \\ Z_{Bor,t-2} \end{bmatrix} \\ + \begin{bmatrix} 0.0693 & (-0.0114) \\ 0.0505 & 0.0567 \end{bmatrix} \begin{bmatrix} Z_{Lyo,t-3} \\ Z_{Bor,t-3} \end{bmatrix} + \begin{bmatrix} U_{Lyo,t} \\ U_{Bor,t} \end{bmatrix}$$

avec  $\Sigma_U = \begin{bmatrix} 0.3240 & 0.1895 \\ 0.1895 & 0.4162 \end{bmatrix}$ .

Comparaison avec les modèle individuels :

$$\begin{bmatrix} Z_{Lyo,t} \\ Z_{Bor,t} \end{bmatrix} = \begin{bmatrix} 0.8615 & 0 \\ 0 & 0.8353 \end{bmatrix} \begin{bmatrix} Z_{Lyo,t-1} \\ Z_{Bor,t-1} \end{bmatrix} + \begin{bmatrix} -0.1507 & 0 \\ 0 & -0.1436 \end{bmatrix} \begin{bmatrix} Z_{Lyo,t-2} \\ Z_{Bor,t-2} \end{bmatrix} \\ + \begin{bmatrix} 0.0545 & 0 \\ 0 & 0.0610 \end{bmatrix} \begin{bmatrix} Z_{Lyo,t-3} \\ Z_{Bor,t-3} \end{bmatrix} + \begin{bmatrix} \varepsilon_{Lyo,t} \\ \varepsilon_{Par,t} \end{bmatrix}$$

avec  $\text{var}(\varepsilon_{Lyo,t}) = 0.3936$ , et  $\text{var}(\varepsilon_{Bor,t}) = 0.4187$ .

### 2. Système {Lyon, Paris}

$$\begin{bmatrix} Z_{Lyo,t} \\ Z_{Par,t} \end{bmatrix} = \begin{bmatrix} 0.5712 & 0.5147 \\ 0.0662 & 0.8823 \end{bmatrix} \begin{bmatrix} Z_{Lyo,t-1} \\ Z_{Par,t-1} \end{bmatrix} + \begin{bmatrix} -0.0518 & -0.2548 \\ -0.0810 & -0.1865 \end{bmatrix} \begin{bmatrix} Z_{Lyo,t-2} \\ Z_{Par,t-2} \end{bmatrix} \\ + \begin{bmatrix} 0.0644 & (0.0074) \\ 0.0309 & 0.0704 \end{bmatrix} \begin{bmatrix} Z_{Lyo,t-3} \\ Z_{Par,t-3} \end{bmatrix} + \begin{bmatrix} U_{Lyo,t} \\ U_{Par,t} \end{bmatrix}$$

avec  $\Sigma_U = \begin{bmatrix} 0.3177 & 0.1632 \\ 0.1632 & 0.3640 \end{bmatrix}$ .

Comparaison avec les modèle individuels :

$$\begin{bmatrix} Z_{Lyo,t} \\ Z_{Par,t} \end{bmatrix} = \begin{bmatrix} 0.8615 & 0 \\ 0 & 0.9107 \end{bmatrix} \begin{bmatrix} Z_{Lyo,t-1} \\ Z_{Par,t-1} \end{bmatrix} + \begin{bmatrix} -0.1507 & 0 \\ 0 & -0.1992 \end{bmatrix} \begin{bmatrix} Z_{Lyo,t-2} \\ Z_{Par,t-2} \end{bmatrix} \\ + \begin{bmatrix} 0.0545 & 0 \\ 0 & 0.0658 \end{bmatrix} \begin{bmatrix} Z_{Lyo,t-3} \\ Z_{Par,t-3} \end{bmatrix} + \begin{bmatrix} \varepsilon_{Lyo,t} \\ \varepsilon_{Par,t} \end{bmatrix}$$

avec  $\text{var}(\varepsilon_{Lyo,t}) = 0.3936$ , et  $\text{var}(\varepsilon_{Par,t}) = 0.3656$ .

### 3. Système {Lyon, Marseille}.

$$\begin{bmatrix} Z_{Lyo,t} \\ Z_{Mar,t} \end{bmatrix} = \begin{bmatrix} 0.7899 & 0.1262 \\ 0.3250 & 0.5987 \end{bmatrix} \begin{bmatrix} Z_{Lyo,t-1} \\ Z_{Mar,t-1} \end{bmatrix} + \begin{bmatrix} -0.1608 & (-0.0220) \\ -0.1560 & (-0.0091) \end{bmatrix} \begin{bmatrix} Z_{Lyo,t-2} \\ Z_{Mar,t-2} \end{bmatrix} \\ + \begin{bmatrix} 0.0530 & (-0.0046) \\ (0.0156) & 0.0437 \end{bmatrix} \begin{bmatrix} Z_{Lyo,t-3} \\ Z_{Mar,t-3} \end{bmatrix} + \begin{bmatrix} U_{Lyo,t} \\ U_{Mar,t} \end{bmatrix}$$

avec  $\Sigma_U = \begin{bmatrix} 0.3885 & 0.2044 \\ 0.2044 & 0.3812 \end{bmatrix}$ .

Comparaison avec les modèle individuels :

$$\begin{bmatrix} Z_{Lyo,t} \\ Z_{Mar,t} \end{bmatrix} = \begin{bmatrix} 0.8615 & 0 \\ 0 & 0.7837 \end{bmatrix} \begin{bmatrix} Z_{Lyo,t-1} \\ Z_{Mar,t-1} \end{bmatrix} + \begin{bmatrix} -0.1507 & 0 \\ 0 & -0.0524 \end{bmatrix} \begin{bmatrix} Z_{Lyo,t-2} \\ Z_{Mar,t-2} \end{bmatrix} \\ + \begin{bmatrix} 0.0545 & 0 \\ 0 & 0.0399 \end{bmatrix} \begin{bmatrix} Z_{Lyo,t-3} \\ Z_{Mar,t-3} \end{bmatrix} + \begin{bmatrix} \varepsilon_{Lyo,t} \\ \varepsilon_{Mar,t} \end{bmatrix}$$

avec  $\text{var}(\varepsilon_{Lyo,t}) = 0.3936$ , et  $\text{var}(\varepsilon_{Mar,t}) = 0.4113$ .

## Conclusions du chapitre 1.

Ce premier chapitre était consacré à l'étude du risque climatique sous-jacent le plus fréquent parmi les produits dérivés climatiques : la température.

Dans un premier temps, nous avons considéré uniquement l'aspect temporel (sections A et B) et étudié la température mesurée dans une dizaine de lieux fixés, situés en France. Il ressort de cette étude qu'un modèle de type autorégressif est assez satisfaisant, à condition d'ajouter une saisonnalité dans la volatilité de manière à tenir compte du fait que les températures en France sont plus dispersées en hiver qu'en été. Le retour rapide à la moyenne de ce type de processus a alors pour conséquence qu'il n'est pas nécessaire de prendre en compte la spécificité de la situation courante dans l'évaluation ou la gestion des risques climatiques lorsque la période d'exposition au risque est située plus d'une vingtaine de jours plus tard.

Au niveau de la modélisation, deux difficultés sont apparues. D'une part, la loi de probabilité des résidus n'est pas toujours normale : elle présente souvent une dissymétrie, qui semble provenir de celle que l'on observe sur la distribution brute des températures, principalement lors des mois d'hiver, qui présente une queue épaisse dans les basses températures. D'autre part, les autocorrélogrammes des carrés des résidus font apparaître des valeurs parfois élevées, ce qui traduit une imperfection du modèle ; tandis que pour les villes françaises, cette imperfection semble acceptable, Campbell et Diebold (2000) ont montré que la situation était moins favorable pour les villes américaines, et ont abouti à une amélioration du modèle en intégrant un phénomène d'hétéroscédasticité conditionnelle.

Dans un deuxième temps, nous avons adjoint la dimension spatiale à la modélisation, en nous intéressant principalement aux améliorations que peut apporter cette nouvelle dimension à la dynamique temporelle de la température (section C). Comme on pouvait s'y attendre, l'ajustement d'un modèle multivarié de type autorégressif est cette fois moins satisfaisant, les difficultés rencontrées au niveau local se cumulant au niveau global ; toutefois, certains résultats sont encourageants comme la possibilité de retrouver les déplacements de masses d'air sur les autocorrélogrammes des résidus des modèles locaux de température. Cependant, et contrairement à l'intuition, la connaissance des températures de lieux voisins n'apporte que peu d'améliorations sur la modélisation de la température d'un lieu fixé.

Dans la suite, nous avons convenu de nous limiter - comme premier travail - à l'évaluation des produits dérivés climatiques relatifs à un lieu unique. D'après ce qui précède, il n'est alors pas utile de tenir compte de la dimension spatiale (ce qui aurait été nécessaire si l'objectif avait été, par exemple, d'étudier la gestion d'un portefeuille de produits dérivés climatiques) et on a retenu le modèle de type autorégressif avec variance périodique. Les répercussions sur les prix des défauts du modèle mentionnés ci-dessus, ainsi que l'impact de l'erreur d'estimation seront évalués dans le chapitre 3. Auparavant, il convient de mieux comprendre le lien entre la température et les prix, et plus généralement, de s'intéresser aux aspects financiers dans la gestion des risques climatiques non catastrophiques.

## Références du chapitre 1.

- Barrieu P., *Introduction au marché des produits dérivés météo*, Séminaire ISFA : Couverture des risques naturels et climatiques, Paris, 24 et 25 octobre 2000.
- Brockwell P.J., Davis R.A., *Time Series : Theory and Methods*, Springer Verlag, 1991.
- Campbell S., Diebold F.X., *Weather Forecasting for Weather Derivatives*. working Paper, University of Pennsylvania, Philadelphia, 2000.
- Cao M., Wei J., *Pricing Weather Derivatives : an Equilibrium Approach*, octobre 1998.
- Carmona R., *Calibrating Degree Days Options*, 1999.
- Dischel R., *The Fledgling Weather Market Takes Off, Part. 2.*, Applied Derivatives Trading, octobre 1998.
- Dischel R., *Seasonnal Weather Forecasts and Derivative Valuation*, Risk, Energy and Power Risk Management, août 2000.
- Dornier F., Quéruef M., *Caution to the wind*, Risk, Energy and Power Risk Management, août 2000.
- Efron B., Morris C., *Le paradoxe de Stein*. Scientific American, **1**, nov. 1977.
- Gourieroux C., Jasiak J., *Nonlinear Innovations and Impulse Responses*, octobre 1998.
- Gourieroux C., Montfort A., *Séries temporelles et modèles dynamiques*, Economica, 2<sup>ème</sup> édition, 1995.
- Jones C., Wand R., *Kernel Smoothing*, Chapman and Hall, 1995.
- Jorion P., *International Portfolio Diversification with Estimation Risk*, Journal of Business, 1985, **58**(3).



Koop G., Pesaran M.H., Potter S.M., *Impulse response analysis in nonlinear multivariate models*, Journal of Econometrics, **74**, 119-147, 1996.

Li H., Maddala G.S., *Bootstrapping Time Series Models*, Econometric Reviews, **15**(2), 115-158, 1996.

Lütkepohl H., *Introduction to Multiple Time Series Analysis*, Springer Verlag, 1993.

Makridakis S., Wheelwright S.C., Hyndman R.J., *Forecasting. Methods and Applications*, Wiley, 1998.

Moreno M., *Evaluation des Dérivés Climatiques*, Speedwell Weather Derivatives, [www.weatherderivs.com](http://www.weatherderivs.com), 2000.

Moreno M., *Riding the Temp*, Weather Derivatives, FOW Special Supplement, Décembre 2000.

Sankaran K., *Finding a value*, Risk, Energy and Power Risk Management, août 2000.

Saporta G., *Probabilités, Analyse des Données et Statistique*, Technip, 1990.

Wilks D.S., *Statistical Methods in the Atmospheric Sciences*, Academic Press, 1995.

## **Chapitre 2. Aspects financiers.**

Dans ce chapitre on considère les produits climatiques (toujours relativement au risque de température) au sens large : dérivés climatiques bien sûr, mais aussi actions d'entreprises sensibles au climat ; et on s'intéresse au problème de leur évaluation. Peut-on considérer le risque de température comme un risque financier à rémunérer ? Autrement dit, y a-t-il un « bêta » spécifique à la température ? Si, comme on peut le penser, la température est décorrélée des mouvements du marché financier, alors l'évaluation des produits dérivés climatiques se fera dans l'univers historique (§ A.I.). Cependant, l'indépendance n'est pas parfaite : on a détecté un certain nombre d'« anomalies » à l'efficacité imputables à des variables météorologiques, dont la température (voir Cao et Wei, 2001). On relie généralement ces anomalies à l'effet de phénomènes naturels sur le comportement des agents économiques dans leur globalité de sorte qu'il n'y a pas de raison de croire, a priori, que ces effets soient plus importants lorsqu'ils se rapportent à des produits climatiques (§ B.II.2.). Toutefois, la température a bel et bien un impact sur le chiffre d'affaires des entreprises sensibles au climat. Cet impact se ressent-il sur le cours de leurs actions ? Cette question est étudiée en détail, d'abord théoriquement (§ B.I.2.), puis de façon empirique (§ B.II.3.).

## A. L'évaluation des produits dérivés climatiques.

Dans cette partie nous donnons quelques résultats tenant à l'évaluation des produits dérivés climatiques. Dans un premier temps, se basant sur le modèle d'évaluation des actifs financiers (MEDAF), nous montrons que l'évaluation se fait dans l'univers historique lorsque le bêta du produit dérivé climatique est nul. La vérification empirique étant impossible à réaliser actuellement en raison de l'absence de données de prix, nous étudions l'indépendance de la température et du marché.

### I. Evaluation dans l'univers historique et bêta nul.

Nous nous plaçons ici dans le cadre du MEDAF monopériodique. La rentabilité de tout actif  $c$  vérifie donc :

$$E[R_c] = r + \beta_c (E[R_m] - r)$$

où  $R_c$  et  $R_m$  désignent les rentabilités de l'actif  $c$  et du marché respectivement,  $r$  le taux sans risque (dont on suppose l'existence) et  $\beta_c = \text{cov}(R_c, R_m) / \text{var}(R_m)$ . En notant :

-  $P_c^t$ ,  $t = 0, 1$  son prix à l'instant  $t$  ;

$$- \beta_c^* = \beta_c \times P_c^0 = \frac{\text{cov}(P_c^1, R_m)}{\text{var}(R_m)}$$

on observe que la relation du MEDAF conduit à la relation suivante, connue sous le nom de « certainty equivalent valuation formula » (voir Copeland et Weston, 1988, § 7.E.) :

$$P_c^0 = \frac{E[P_c^1] - \beta_c^* (E[R_m] - r)}{1 + r} \quad (*)$$

On peut remarquer qu'avec le changement de probabilité :

$$\frac{dQ}{dP} = 1 - \lambda_m (R_m - E[R_m]),$$

où  $\lambda_m$  est la prime de risque de marché -  $\lambda_m = \frac{E[R_m] - r}{\text{var}[R_m]}$ , la relation (\*) précédente s'écrit :

$$P_c^0 = \frac{E^Q[P_c^1]}{1 + r}$$

On retrouve donc de ce fait l'interprétation usuelle du prix comme l'espérance actualisée sous une nouvelle probabilité « corrigée du risque »,  $Q$ .

Supposons maintenant que  $c$  soit un actif particulier, à bêta nul. Alors on a également  $\beta_c^* = 0$  et la relation (\*) conduit à :

$$P_c^0 = \frac{E[P_c^1]}{1+r}$$

Ainsi, l'évaluation se fait en réalité dans l'univers historique. Il y a toujours une probabilité corrigée du risque, différente de la probabilité historique, mais le premier moment coïncide dans les deux univers.

Or on peut penser que les produits dérivés climatiques font effectivement partie de ces actifs à bêta nul. En effet, ils sont exclusivement créés à partir de la température dont on a que peu de raison de penser qu'elle ait un lien majeur avec le marché financier. Il eût été préférable de tester directement la non-corrélation entre le marché et le produit climatique lui-même, mais l'illiquidité du marché climatique rend la démarche stérile. On se rabat donc à tester l'indépendance entre le marché et la température.

## **II. Indépendance entre marché et température.**

### **1. Relations a priori. Choix de la fréquence des données.**

Pour aborder ce problème, il importe d'avoir une idée des relations qui pourraient exister entre un indice de marché et la température. La littérature sur le sujet mentionne des effets des températures extrêmes sur la rentabilité des indices de marché ; ainsi Cao et Wei (2001) ont mis en évidence, en utilisant des données journalières, une rentabilité significativement plus élevée les jours de basses températures. L'explication de ce phénomène serait de nature comportementale : l'augmentation de l'agressivité humaine constatée (et prouvée) lors de conditions extrêmes de température occasionnerait une prise de risque plus importante des agents économiques. A côté de ces conséquences irrationnelles, on peut rechercher de véritables relations économiques : retrouve-t-on dans le cours des indices de marché les impacts économiques dus à la température ? Il y a vraisemblablement des effets à long terme,

avec notamment l'impact du réchauffement climatique observé ces dernières années. Toutefois le petit nombre de données rend l'étude statistique difficile. A l'opposé, il est peu vraisemblable que les cours intègrent une variation thermique ponctuelle, fût-elle extraordinaire (à moins qu'elle ne soit couplée avec un autre événement climatique, de nature catastrophique) : les investisseurs s'inquiéteront-ils de la santé d'une entreprise à cause d'une chute brutale de température, ne durant qu'un jour ou deux ? En revanche, on peut penser qu'un mauvais mois – ou une mauvaise saison, peut ralentir l'activité économique et pourrait être pris en compte dans l'évaluation financière.

A la lumière de ces réflexions, il ressort que des données mensuelles ou trimestrielles sont sans doute les mieux appropriées pour déceler un impact du climat sur le marché. Les contraintes de l'étude statistique, imposant un nombre suffisant de données, nous font pencher pour une fréquence mensuelle. Nous donnons un argument quantitatif confortant ce choix dans le cadre des entreprises sensibles au climat<sup>63</sup> (paragraphe B.I.2). Dans la suite, nous présentons donc les résultats obtenus avec des données mensuelles<sup>64</sup>.

## 2. Données.

L'étude est menée avec 5 indices du marché français (CAC 40, SBF 80, SBF 250, MidCac et Second Marché) et un indice américain (Dow Jones). Nous disposons des cours quotidiens de clôture. La longueur des historiques est précisée dans le Tableau 2.1.

Pour le marché français nous avons utilisé les températures de Paris et Marseille ; pour le marché américain, celles de New York et Chicago. Les résultats sont assez peu sensibles au choix d'une station particulière ni même d'un groupe de stations. La raison à cela est la forte corrélation spatiale des températures à l'intérieur d'une même « région climatique »<sup>65</sup>.

---

<sup>63</sup> Notons que l'anomalie comportementale de température devrait également s'observer dans le cas de données mensuelles, car les mois extrêmes correspondent à une accumulation de jours extrêmes.

<sup>64</sup> Nous avons également étudié le cas de données journalières : les conclusions sont inchangées.

<sup>65</sup> Voir paragraphe Chapitre 2.B.II.3.a).

Tableau 2.1. Longueur des historiques

Valeur	Historique
CAC 40	01/01/1988 au 31/12/1999
SBF 80	01/01/1991 au 31/12/1999
SBF 250	01/01/1991 au 31/12/1999
MidCac	01/01/1991 au 31/12/1999
Second Marché	01/01/1991 au 31/12/1999
Dow Jones	01/01/1986 au 31/12/1999
Electricité de Strasbourg	01/01/1990 au 31/12/1999
Eurodisney	01/01/1990 au 31/12/1999
Pernod-Ricard	01/07/1992 au 31/12/1999
Teisseire-France	01/01/1990 au 31/12/1999

### 3. Recherche de corrélations. Indépendance.

Dans ce paragraphe, nous étudions l'existence de corrélations « globales » entre la température et le marché, au sens où l'on n'isole pas les températures extrêmes. Nous recherchons d'abord la présence de corrélations linéaires, puis non linéaires. Pour cela, il est essentiel de travailler avec des données stationnarisées, sans quoi on peut être confronté à des problèmes d'inconsistance des estimateurs (surtout en présence de termes déterministes). Nous utilisons donc les variables suivantes :

- Pour la température, le résidu mensuel de température, noté  $\varepsilon_{\text{temp},t}$ , égal à la moyenne des résidus quotidiens du mois  $t$  :

$$\varepsilon_{\text{temp},t} = \frac{1}{\#\text{mois } t} \sum_{\text{mois } t} \varepsilon_s$$

Ce terme correspond à la partie aléatoire de la température moyenne mensuelle<sup>66</sup>.

- Pour le marché, on utilise la rentabilité logarithmique mensuelle :

<sup>66</sup> Il est à peu près indifférent de prendre en compte ou non le coefficient déterministe de dispersion saisonnière  $\rho_s$ . Dans un même mois, les variations de  $\rho_s$  sont minimes, si bien que les deux quantités  $\sum_{\text{mois } t} \varepsilon_s$

et  $\sum_{\text{mois } t} \rho_s \varepsilon_s$  sont à peu près proportionnelles.

$$R_{m,t} = \log(S_{t+1}^- / S_t)$$

où  $S_t$  et  $S_{t+1}^-$  désignent respectivement le cours du premier et du dernier jour du mois  $t$ <sup>67</sup>.

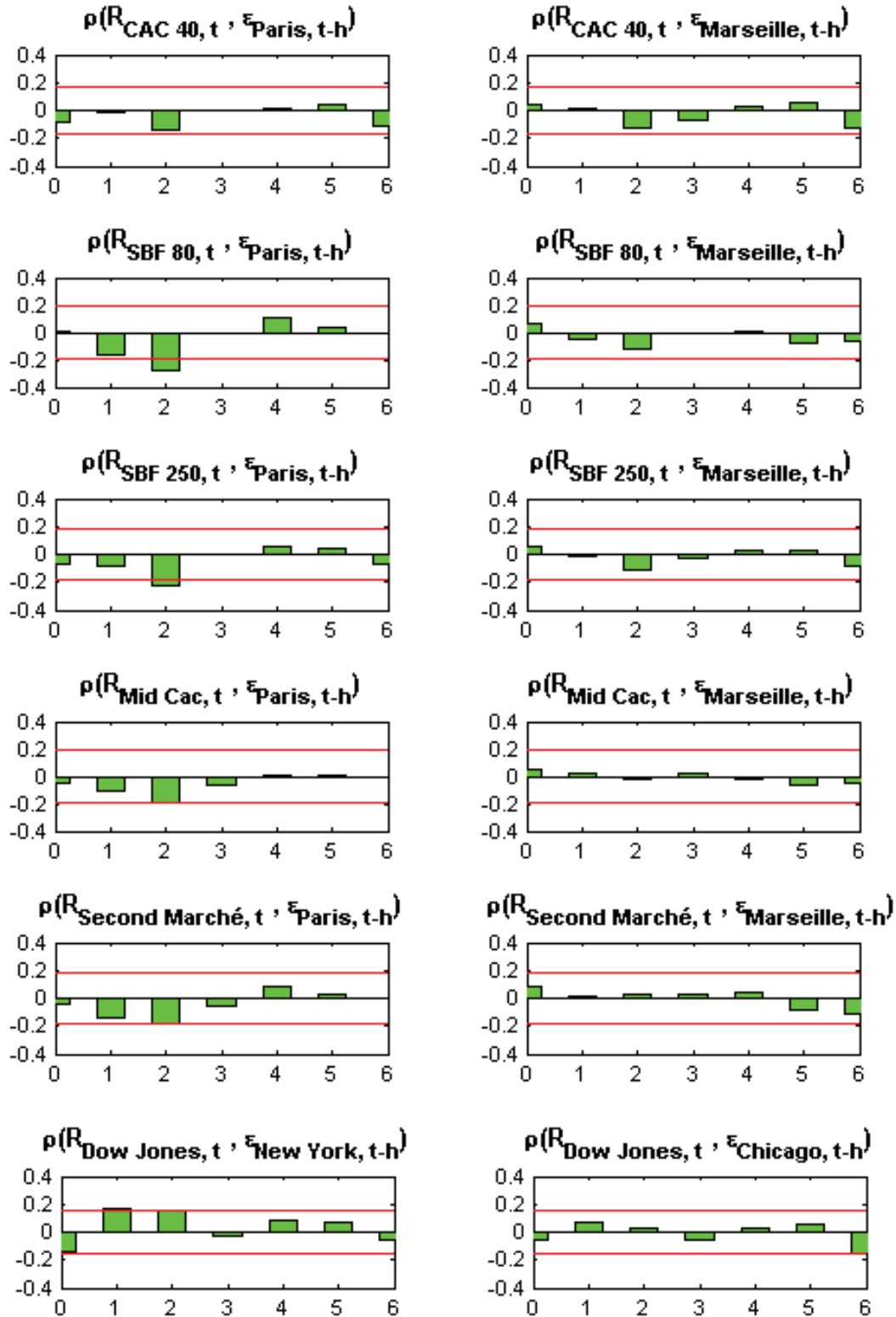
Dans un premier temps, nous recherchons des corrélations linéaires entre la température et le marché. Ainsi, et de façon à tenir compte d'éventuels retardements dans l'impact de la température, nous traçons les autocorrélations estimées entre  $R_{m,t}$  et  $\varepsilon_{\text{temp},t-h}$ ,  $h = 0, 1, \dots, 6$  avec les bandes de signification à 95% (voir Lütkepohl, 1993, § 4.4.2.). La Figure 2.1 représente le résultat obtenu pour les données de marché et de température précisées ci-avant. D'une façon générale, il ressort que les corrélations entre température et marché ne sont pas - ou peu, significatives. Cependant, quelques précisions doivent être apportées. Pour les données françaises tout d'abord, on remarque des corrélations négatives légèrement significatives entre la rentabilité des indices boursiers (à l'exception de l'indice CAC 40) et la température mensuelle observée deux mois auparavant à Paris. En revanche aucune corrélation n'est significative quand on considère la température de Marseille. Pour les données américaines on remarque de légères corrélations – positives cette fois – entre la rentabilité et la température à New York des deux mois précédents, mais rien avec Chicago.

Que conclure de tout cela ? Pour commencer, notons que l'anomalie comportementale de température observée par Cao et Wei ne devrait pas intervenir ici car elle est intégrée dans les cours le jour même. Le fait d'observer un temps de réaction pèse donc en faveur d'une cause économique : un redressement aurait lieu à la fin du trimestre. Dans cette hypothèse, il n'est pas absurde que les régions de Paris et New York soient davantage « impliquées » que ne le sont celles de Marseille et Chicago ; elles correspondent en effet à deux pôles économiques majeurs en France et aux Etats-Unis. Toutefois les fortes corrélations entre les températures mensuelles de Paris et Marseille (estimée à 0.65) et celles de New York et Chicago (estimée à 0.70) viennent nuancer ce discours. Par ailleurs, comment expliquer que le signe des corrélations change lorsqu'on traverse l'Atlantique ? Tout ceci invite à la prudence ; au mieux pourrait-il exister une corrélation, à peine significative, et retardée, entre la température et le marché.

<sup>67</sup> On a envisagé de tenir compte de l'hétéroscédasticité en prenant comme variable d'étude le résidu de régression linéaire de  $R_{m,t}$  sur  $R_{m,t-1}$ . Cependant les résultats sont très peu modifiés.



Figure 2.1. Autocorrélations estimées entre rentabilité boursière et température.



Dans un deuxième temps, à la recherche d'éventuelles non-linéarités, nous mettons en œuvre le test non paramétrique d'indépendance de Kendall (voir Hollander, Wolfe, 1999, §8.1.). Le Tableau 2.2 donne les  $p$ -valeurs obtenues en testant l'indépendance entre  $R_{m,t}$  et  $\varepsilon_{\text{temp},t-h}$ ,  $h = 0,1,2$ . Les résultats confirment les conclusions précédentes sans apporter d'informations supplémentaires.

**Tableau 2.2. Indépendance entre marché et température :  $p$ -valeur du test d'indépendance de Kendall**

	Paris			Marseille		
	retard = 0	retard = 1	retard = 2	retard = 0	retard = 1	retard = 2
CAC 40	0.5876	0.6238	0.0994	0.3098	0.8036	0.3164
SBF 80	0.7866	0.0727	0.0086	0.4382	0.7181	0.1826
SBF 250	0.6252	0.3055	0.0195	0.4670	0.9196	0.2514
MidCac	0.8776	0.3611	0.0553	0.4900	0.5954	0.8860
Second Marché	0.8567	0.1658	0.0520	0.2811	0.7460	0.8734

	New York			Chicago		
	retard = 0	retard = 1	retard = 2	retard = 0	retard = 1	retard = 2
Dow Jones	0.0428	0.1997	0.0202	0.3281	0.5748	0.7019

#### 4. Influence des températures extrêmes. Anomalie de température.

Dans ce paragraphe, on étudie l'influence des températures extrêmes sur le marché boursier. Ce travail vient donc en complément de l'étude réalisée par Cao et Wei avec des données journalières ; nous utilisons d'ailleurs leur méthodologie. Le principal apport de l'extension aux données mensuelles est de pouvoir tenir compte de causes purement économiques (*cf.* la discussion du paragraphe 1 de cette section). Celles-ci peuvent modifier *a priori* les résultats obtenus avec une fréquence quotidienne. Rappelons que la principale conclusion à laquelle aboutissent Cao et Wei est que les rentabilités d'indices boursiers sont significativement plus élevées les jours de basses températures ; la cause invoquée est de nature comportementale. On devrait retrouver cet effet en examinant les rentabilités mensuelles, les mois froids correspondant à une accumulation de jours froids. Si seule cette cause comportementale intervient, on s'attend donc à observer une rentabilité supérieure lors des mois froids, et un effet quasi-nul lors des mois chauds. Mais on pourrait également observer d'autres effets dus

à des causes économiques. Si ces effets ne sont pas les mêmes, cette étude devrait permettre de les mettre en évidence.

La méthodologie suivie est la même que Cao et Wei ; nous résumons ici les étapes essentielles, et renvoyons à leur article (2001) pour les points de détails. On commence par faire des classes de largeur identique à partir des données d'indices de températures standard HDD et CDD. On considère ensuite les rentabilités des mois correspondants, et on teste l'hypothèse nulle que les rentabilités moyennes des 1<sup>ère</sup> et dernière classes sont égales<sup>68</sup>. On emploie ici le test non paramétrique de Wilcoxon (voir Hollander, Wolfe, 1999, § 4.1.) ; le test est unilatéral pour l'indice HDD, opposant  $H_0 : "E[R_1] = E[R_d]"$  à  $H_1 : "E[R_1] < E[R_d]"$  ( $R_1$  et  $R_d$  représentent les rentabilités des 1<sup>ère</sup> et dernière classes), et bilatéral pour l'indice CDD.

Les résultats sont montrés dans le Tableau 2.3 et le Tableau 2.4. Ils indiquent les caractéristiques de translation et de dispersion des rentabilités dans chaque classe, ainsi que la  $p$ -valeur du test de Wilcoxon. On observe que les rentabilités des indices boursiers sont significativement plus élevées lors des mois froids (valeurs importantes de l'indice HDD) et inchangées lors des mois chauds (valeurs importantes de l'indice CDD). En toute rigueur, on devrait seulement conclure au rejet de l'indépendance entre le marché et la température, car le test de Wilcoxon est réalisé sous l'hypothèse  $H$  que les échantillons des rentabilités des première et dernière classes sont indépendants entre eux ; le rejet dans le test signifie donc le rejet de l'une ou l'autre des hypothèses  $H$  ou  $H_0$ . Cependant, on constate que la rentabilité moyenne est supérieure dans la dernière classe de façon quasi-systématique<sup>69</sup> ; on peut donc conjecturer que  $H_0$  doit être rejetée.

On obtient ainsi avec une fréquence mensuelle des résultats analogues à ceux obtenus par Cao et Wei avec une fréquence journalière. Puisque aucune autre constatation nouvelle n'est apparue, nous sommes amenés à penser, à la lumière de la discussion précédente, que

---

<sup>68</sup> Notons qu'avec cette méthodologie, il n'est pas possible de considérer des effets retardés comme dans le paragraphe Chapitre 2.A.II.3, en raison de la saisonnalité marquée des indices de température.

<sup>69</sup> On a envisagé plusieurs seuils pour ces indices, différents du seuil standard égal à 18°C ; les résultats obtenus sont similaires. D'autre part avec des données mensuelles, le problème des indices nuls ne se pose quasiment plus ; là encore, comme dans (Cao, Wei, 2001), les résultats sont similaires qu'on écarte ou non ces données de la 1<sup>ère</sup> classe.

l'influence des températures extrêmes sur le marché financier est de nature comportementale plutôt que de source économique.

**Tableau 2.3. Détection de l'anomalie de température avec des données mensuelles : impact des basses températures.**

Action \ Indice HDD		Nombre de classes = 2			Nombre de classes = 3			
		classe 1	classe 2	p-valeur	classe 1	classe 2	Classe 3	p-valeur
CAC 40	Moyenne	0.0035	0.0228	0.0450	-0.0056	0.0225	0.0306	0.0040
	Ecart-type	0.0551	0.0598		0.0560	0.0561	0.0534	
SBF 80	Moyenne	-0.0017	0.0343	0.0000	-0.0081	0.0215	0.0407	0.0001
	Ecart-type	0.0402	0.0451		0.0401	0.0387	0.0480	
SBF 250	Moyenne	-0.0013	0.0312	0.0005	-0.0096	0.0226	0.0378	0.0002
	Ecart-type	0.0467	0.0447		0.0473	0.0403	0.0451	
MidCac	Moyenne	-0.0078	0.0327	0.0002	-0.0144	0.0167	0.0412	0.0003
	Ecart-type	0.0455	0.0547		0.0461	0.0412	0.0625	
Second Marché	Moyenne	-0.0040	0.0274	0.0001	-0.0089	0.0177	0.0285	0.0013
	Ecart-type	0.0342	0.0424		0.0356	0.0320	0.0483	
<i>Taille approx. des classes*</i>		<i>65%</i>	<i>35%</i>		<i>45%</i>	<i>35%</i>	<i>20%</i>	

**Tableau 2.4. Détection de l'anomalie de température avec des données mensuelles : impact des hautes températures.**

Action \ Indice CDD		Nombre de classes = 2		
		classe 1	classe 2	p-valeur
CAC 40	Moyenne	0.0007	-0.0182	0.4372
	Ecart-type	0.0538	0.0751	
SBF 80	Moyenne	-0.0075	0.0015	0.7220
	Ecart-type	0.0405	0.0421	
SBF 250	Moyenne	-0.0080	-0.0062	0.8695
	Ecart-type	0.0465	0.0573	
MidCac	Moyenne	-0.0135	-0.0069	0.7633
	Ecart-type	0.0483	0.0361	
Second Marché	Moyenne	-0.0070	-0.0103	0.4767
	Ecart-type	0.0371	0.0309	
<i>Taille approx. des classes*</i>		<i>86%</i>	<i>14%</i>	

\* les différences de largeur des classes proviennent des longueurs variables des historiques d'indices de marché

## **5. Conclusion sur l'indépendance entre marché et température.**

Nous avons analysé l'indépendance entre température et marché à partir de données mensuelles d'indices des marchés français et américains. A l'exception d'un léger effet retardé, nous n'avons pas décelé de corrélations globales significatives ; en revanche, les rentabilités boursières d'indices sont significativement supérieures lors des mois froids ce qui confirme les conclusions des travaux de Cao et Wei, obtenues avec une fréquence quotidienne, et qui s'expliquerait par le rôle joué par la température sur le comportement humain. Notons que les deux conclusions obtenues ne sont pas incompatibles : dans la recherche de corrélations globales les températures extrêmes sont diluées dans un flot de données, ce qui peut expliquer que leur impact ne puisse être mis en évidence. Néanmoins, la proportion de ces températures « extrêmes » atteint parfois 35% (*cf.* Tableau 2.3), ce qui tendrait plutôt à montrer que l'impact des températures sur le marché est limité.

Pour en tirer des conséquences sur les produits dérivés climatiques, il est bon de remarquer que les produits standard ne font pas jouer un rôle particulier aux extrêmes de température. D'une part les indices non linéaires de température, HDD ou CDD, mesurent les écarts au-delà d'un seuil dont le niveau ( $18^{\circ}\text{C}$ ) est bien trop moyen pour faire intervenir les températures extrêmes ; d'autre part, même dans le cas des options où le *payoff* est non linéaire, le prix d'exercice n'est pas très éloigné de la moyenne historique des indices de température. Etant donné que la dépendance légère observée au point 4 du paragraphe précédent semble liée aux températures extrêmes, on peut conjecturer qu'elle n'interviendra pas dans l'évaluation des produits dérivés climatiques. Toutefois il conviendra, pour s'en assurer, de réaliser des études empiriques directement sur les prix des produits.

## B. Evaluation des actions des entreprises sensibles au climat.

A la section précédente, nous avons vu que le bêta d'un indice de marché par rapport à la température n'est pas significatif<sup>70</sup>. Nous voulons maintenant raffiner l'étude en considérant le cas particulier des actions d'entreprises sensibles au climat. Naturellement toutes les entreprises sont exposées aux risques climatiques, mais la spécialisation dans un secteur d'activité « à risque », la localisation géographique font de certaines d'entre elles un terrain propice aux aléas climatiques. C'est le cas par exemple de la société Teisseire-France, dont 80% de l'activité concerne la commercialisation de jus de fruits en France (voir Tableau 2.5, page 158). Pour ces entreprises, on peut concevoir que la température puisse avoir un rôle particulier dans l'évaluation des actions. En particulier, il est possible que le bêta action/température soit significatif.

La première partie est consacrée à des calculs théoriques dont le but principal est de fournir des résultats quantitatifs de l'impact de la température sur les actions. La seconde partie est une étude empirique réalisée sur les actions de 4 entreprises françaises sensibles à la température, où nous examinons la question de l'influence du facteur « température » sous un angle statistique.

### I. Résultats théoriques.

#### 1. Evaluation des actions à dividendes périodiques.

Les entreprises sensibles au climat ont des revenus saisonniers. Imaginons, pour simplifier, que cette saisonnalité soit répercutée intégralement dans les dividendes de leurs actions ; peut-on alors calculer sa contribution à la valeur des actions ?

Pour répondre à cette question, nous nous plaçons en temps continu et ajoutons une composante périodique au modèle de Gordon-Shapiro (voir Campbell, Lo, MacKinlay, § 7.1.). Nous supposons donc que les dividendes vérifient :

$$d_t = De^{gt} (1 + \lambda \sin(\omega t))$$

où  $\omega = 2\pi$  (périodicité annuelle) ; les autres quantités s'interprètent comme suit :

<sup>70</sup> en prenant comme variable de température la température elle-même.

- $D$  : dividende initial ;
- $g$  : taux de croissance des dividendes en l'absence de saisonnalité<sup>71</sup> ;
- $\lambda$  : proportion de saisonnalité dans les dividendes.

Avec ce modèle, la valeur actuelle de l'action à la date  $u$  est l'espérance actualisée de ses flux futurs :

$$V_\lambda(t) = \int_u^{+\infty} De^{gt} (1 + \lambda \sin(\omega t)) e^{-r(t-u)} dt$$

où  $r$  est le taux d'actualisation. Un calcul élémentaire donne alors :

$$V_\lambda(u) = \frac{De^{gu}}{r-g} \left\{ 1 + \lambda \cdot \frac{r-g}{\sqrt{(r-g)^2 + \omega^2}} \sin(\omega u + \phi_{r,g,\omega}) \right\}$$

avec  $(\cos \phi_{r,g,\omega}; \sin \phi_{r,g,\omega}) = (r-g; \omega) / \sqrt{(r-g)^2 + \omega^2}$ .

On constate que la valeur de l'action se décompose en une partie constante, égale à sa valeur en l'absence d'effet saisonnier, et un terme cyclique de même périodicité que les dividendes, mais présentant un déphasage et – surtout – de proportion modifiée, égale à :

$$\lambda \cdot \frac{r-g}{\sqrt{(r-g)^2 + \omega^2}}$$

En notant  $V(u) = V_0(u)$  la valeur de l'action en l'absence d'effet saisonnier, la proportion de valeur due à la saisonnalité est ainsi :

$$\frac{V_\lambda(u) - V(u)}{V(u)} = \lambda \cdot \frac{r-g}{\sqrt{(r-g)^2 + \omega^2}} \sin(\omega u + \phi_{r,g,\omega})$$

- Dans le cas de basses fréquences ( $\omega \rightarrow 0$ ),

$$\frac{V_\lambda(u) - V(u)}{V(u)} \rightarrow \lambda$$

La saisonnalité des dividendes est répercutée intégralement, en proportion, dans la valeur de l'action.

- A l'inverse, dans le cas de « hautes » fréquences, c'est-à-dire dans un des cas  $\omega \rightarrow +\infty$ , ou  $\omega \gg r-g$ ,

$$\frac{V_\lambda(u) - V(u)}{V(u)} \approx \lambda \frac{r-g}{\omega} \cos(\omega u)$$

<sup>71</sup> Remarquons que  $De^{gt}$  est le dividende moyen annuel, de sorte que  $g$  s'interprète également comme le taux de croissance instantané moyen des dividendes.

La saisonnalité est donc très atténuée. (En outre, le déphasage est total : la valeur de l'action est maximale au moment où le dividende est minimal).

Dans la pratique,  $\omega = 2\pi$  est très supérieur à  $r - g$ , de l'ordre de  $10\% = 0.1$ . On se trouve donc dans le dernier cas. Avec les données numériques précédentes, on a :

$$\max \left| \frac{V_\lambda(u) - V(u)}{V(u)} \right| \approx 0.016 \times \lambda$$

La proportion de saisonnalité dans la valeur de l'action représente 1.6% de la proportion de saisonnalité des dividendes,  $\lambda$ . Dans la pratique, on peut penser que la valeur de  $\lambda$  est de toute façon très inférieure à 1, à moins que l'entreprise ait une activité purement saisonnière. L'impact du climat sur le cours de l'action s'avère alors négligeable.

## 2. Evaluation théorique de la répercussion d'une bonne saison sur le prix des actions.

Une saison favorable se ressent directement sur le chiffre d'affaires des entreprises sensibles au climat. Ainsi, dans un article paru dans la revue Capital (Solen, 2002), on apprend qu'une augmentation de  $1^\circ\text{C}$  au-dessus de  $24^\circ\text{C}$  occasionne une augmentation de 8% des ventes de Teisseire-France. Dans ce paragraphe, nous avons souhaité quantifier l'ordre de grandeur de l'impact maximal de cette variation sur le cours de l'action. Pour ce faire, on suppose que les bénéfices exceptionnels réalisés sont entièrement reversés sous forme de dividendes<sup>72</sup>. Dans un premier temps, nous établissons alors le lien théorique entre la variation de la rentabilité de l'action et la variation de la marge de l'entreprise. Ensuite, à l'aide du bilan de chaque entreprise, nous quantifions approximativement la relation entre marge et chiffre d'affaires. Enfin, nous remontons jusqu'à la température en calculant l'impact d'une bonne saison sur le chiffre d'affaires.

Nous avons effectué l'étude pour Teisseire-France pour laquelle nous disposons d'un ordre de grandeur de l'impact de la température ; nous l'avons étendue à Pernod-Ricard, qui est dans le même secteur d'activité.

<sup>72</sup> Il s'agit d'une hypothèse commode permettant un calcul simplifié de la revalorisation de l'action ; en pratique, une partie des bénéfices est réinvestie, mais l'augmentation de valeur de l'action dépend du montant des bénéfices et non de la façon dont ils sont utilisés.



a) Lien théorique entre la variation du cours de l'action et la variation de la marge.

On suppose ici que l'on a réussi à évaluer l'impact de bonnes conditions climatiques au cours d'une année sur la marge de l'entreprise. L'augmentation de la marge va se répercuter sur le cours de l'action, car, par hypothèse, l'entreprise reverse intégralement ces bénéfices aux actionnaires sous forme de dividendes. Nous voulons maintenant quantifier l'importance de cette répercussion.

Pour cela, nous nous plaçons dans le cadre du modèle de Gordon-Shapiro (voir Campbell, Lo, MacKinlay, § 7.1.) où le taux de croissance des dividendes et le taux d'actualisation sont supposés constants. L'unité de temps est l'année. Et, pour simplifier, on fait également l'hypothèse que les dividendes sont versés intégralement à la fin de l'année.

Nous utilisons les notations suivantes :

$\Delta M, M^+, M$  : Variation de la marge due au climat, marge réalisée avec l'impact du climat ( $M^+$ ), ou sans ( $M$ ) - on a donc  $\Delta M = M^+ - M$  ;

$\Delta S, S^+, S$  : Variation du cours de l'action due au climat, valeur théorique de l'action au début de l'année courante avec l'impact du climat ( $S^+$ ), ou sans ( $S$ ) ;

$N$  : Nombre d'actions à la fin de l'année courante ;

$CB$  : Capitalisation boursière de l'entreprise à la fin de l'année courante :  
 $CB = N \times S$  ;

$D$  : Dividende par action versé à la fin de l'année courante ;

$\Delta g, g^+, g$  : Variation du taux de croissance des dividendes due au climat, valeur de ce taux avec ou sans impact du climat.

$r$  : Taux d'actualisation

D'après l'hypothèse de reversement intégral des bénéfices, on doit d'abord avoir la relation :

$$\Delta M = N \times D \times \Delta g$$

Car, pour chaque action, la part de dividende due au climat est  $D \times \Delta g$ .

D'autre part, il est possible de relier la variation du cours de l'action à  $\Delta g$ .

En effet, en l'absence d'impact du climat, les dividendes futurs successifs sont :

$$D(1+g), D(1+g)^2, \dots, D(1+g)^n, \dots$$

Or le « choc » de température modifie le dividende de l'année courante qui devient  $D(1+g^+)$ , les autres restant inchangés. Par conséquent, la valeur théorique de l'action est :

$$S = \frac{D(1+g)}{1+r} + \frac{D(1+g)^2}{(1+r)^2} + \dots + \frac{D(1+g)^n}{(1+r)^n} + \dots$$

dans une année normale, et :

$$S^+ = \frac{D(1+g^+)}{1+r} + \frac{D(1+g)^2}{(1+r)^2} + \dots + \frac{D(1+g)^n}{(1+r)^n} + \dots$$

après une saison exceptionnelle, si bien que la différence de rentabilité de l'action due à l'impact du climat est de :

$$\frac{\Delta S}{S} = \frac{1}{1+r} \left[ \frac{D}{S} \times \Delta g \right]$$

Des deux relations ci-dessus, on en déduit alors :

$$\frac{\Delta S}{S} = \frac{1}{1+r} \left[ \frac{\Delta M}{N \times S} \right]$$

ou encore :

$$\boxed{\frac{\Delta S}{S} = \frac{1}{1+r} \left[ \frac{\Delta M}{CB} \right]}$$

b) Impact d'une variation du chiffre d'affaires sur la marge.

Nous venons de mesurer la répercussion de la variation de la marge d'une entreprise due au climat sur le cours de l'action. Mais l'information dont nous disposons (voir l'introduction de ce paragraphe) permet de quantifier l'impact de bonnes conditions climatiques sur le chiffre d'affaires... Il nous reste donc à relier la variation du chiffre d'affaires à celle de la marge.

Pour cela, nous faisons l'hypothèse simplificatrice que les frais se répartissent en deux catégories : frais fixes et frais variables à proportion du chiffre d'affaires.

Introduisons les notations suivantes :

$\Delta CA, CA$  : Variation du chiffre d'affaires due au climat, ce chiffre d'affaires avant l'impact du climat ;

$FF$  : Frais fixes ;

$FV, y$  : Frais variables, taux de frais :  $FV = y \times CA$

Nous avons alors :

$$M = CA - FF - FV = (1 - y)CA - FF$$

d'où

$$\boxed{\Delta M = (1 - y)\Delta CA}$$

c) Impact d'une variation du chiffre d'affaires sur le cours de l'action.

D'après les points a) et b), nous constatons que la variation du chiffre d'affaires due à de bonnes conditions climatiques se répercute sur le cours de l'action selon la relation :

$$\boxed{\frac{\Delta S}{S} = \frac{1}{1+r} \left[ (1-y) \frac{CA}{CB} \right] \frac{\Delta CA}{CA}}$$

Application Numérique :

Les calculs qui suivent sont effectués sur la base du compte consolidé de l'année 2001 (voir les comptes des résultats de Pernod-Ricard en annexe, et le site internet [www.teisseire.fr](http://www.teisseire.fr)). Le taux d'actualisation utilisé est fixé arbitrairement à 10%, valeur couramment rencontrée en pratique. Les capitalisations boursières ont été calculées pour la fin de l'année 2001 ; elles sont recensées dans le Tableau 2.17 (voir en fin de chapitre).

Nous en déduisons les résultats suivants :

Données ou variables financières	Pernod-Ricard	Teisseire-France
CA (millions d'euros)	4 555	147
CB (millions d'euros)	4 906	26
CA/CB	0.93	5.67
$1 - y$	0.47 <sup>(73)</sup>	0.51 <sup>(74)</sup>
$\frac{\Delta S}{S}$	$0.4 \times \frac{\Delta CA}{CA}$	$2.6 \times \frac{\Delta CA}{CA}$

Il nous reste maintenant à connaître l'impact d'une bonne saison sur le chiffre d'affaires.

d) Impact d'une bonne saison sur le chiffre d'affaires.

L'information révélée par la revue Capital (Solen, 2002) donne un ordre de grandeur<sup>75</sup> de la variation relative du chiffre d'affaires par unité de température - à savoir : 8%, la relation linéaire n'étant valable qu'au-delà du seuil 24°C. Avec ce qui précède, on observe que toute la question se réduit maintenant à quantifier les chocs de température.

La Figure 2.2 montre l'excédent mensuel de température du sud de la France au-delà de 24°C sur la période 1979-1999. Il est obtenu en effectuant la moyenne des excédents journaliers de température par rapport à 24°C :

$$\text{Excédent mensuel} = \frac{1}{L} \sum_{t \in \text{mois}} (T_t - 24)^+$$

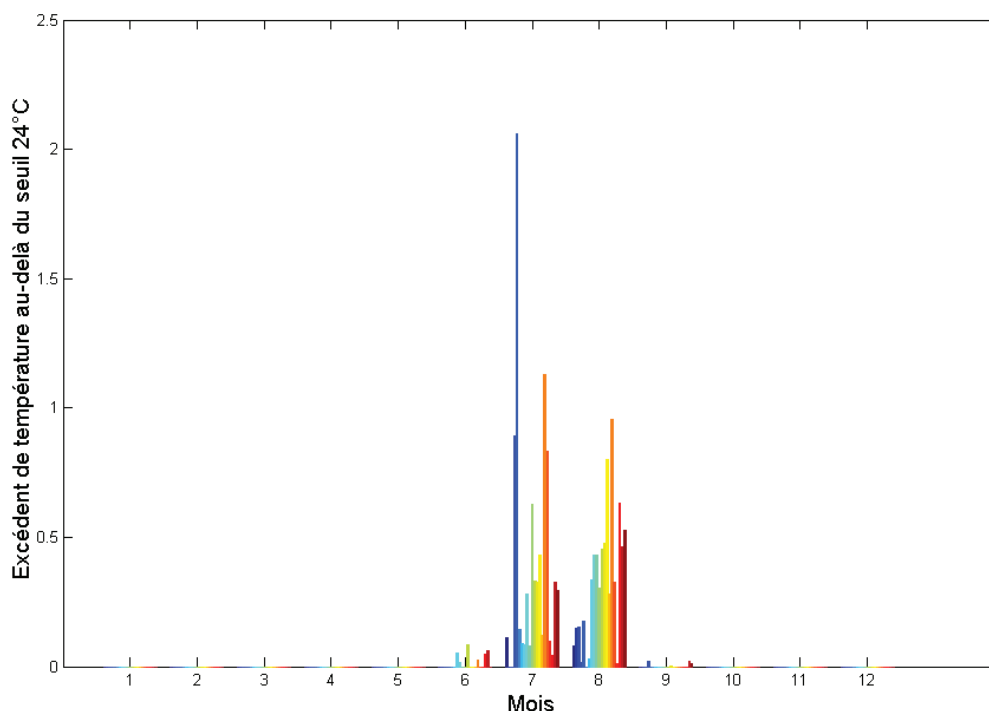
où  $L$  est la durée d'un mois et  $T_t$  la température journalière du jour  $t$ , en degrés celsius<sup>76</sup>. Celle-ci est calculée en effectuant la moyenne des température d'un ensemble de villes du sud de la France (Lyon, Toulouse, Marseille, Nice, Ajaccio). On observe que pour un mois exceptionnel, l'excédent moyen de température a représenté environ 2°C.

<sup>73</sup> Calcul de  $y$  dans le cas de Pernod-Ricard : on a, en millions d'euros :  $FV = 2412.2$  (achats consommés), et  $FF = 1031.3 + 661.2 + 39.5 = 1732$  (frais de commercialisation et de distribution, frais de production et frais généraux, frais financiers nets). On en déduit :  $y = 0.5295$ .

<sup>74</sup> Calcul de  $y$  dans le cas de Teisseire-France : en milliers d'euros cette fois, on a :  $FV = 71\,979$  (consommations), et  $FF = 147\,416 - 5\,422 - 71\,979 = 70\,015$  (différence du chiffre d'affaires, des frais variables et du résultat courant). D'où :  $y = 0.4883$ .

<sup>75</sup> Nous avons compris que si, sur la saison estivale, la moyenne des excédents de température (au-delà de 24°C) est chiffrée à 1°C, alors le chiffre d'affaires annuel augmente de 8%. Cependant, le texte de l'article n'est pas suffisamment explicite pour affirmer que cette interprétation est la bonne, et nos demandes de précisions auprès de l'auteur de l'article et de la société Teisseire sont restées infructueuses.

<sup>76</sup> Il s'agit donc de l'indice CDD de seuil 24°C, rapporté à la longueur de la période.



**Figure 2.2. Moyenne mensuelle des excédents de la température du sud de la France\* au-delà de 24°C, de 1979 à 1999.**

\* moyenne de la température mensuelle de Lyon, Toulouse, Marseille, Nice et Ajaccio.

Le même calcul peut être effectué sur l'ensemble d'une saison. L'augmentation attendue de rentabilité de l'action s'en déduit alors immédiatement, d'abord en évaluant l'augmentation du chiffre d'affaires grâce à l'information comptable ci-dessus (voir la note 75), puis en appliquant la formule du paragraphe précédent. Les résultats sont donnés sur la Figure 2.3 dans le cas de Teisseire-France où l'impact du climat est le plus important. Sur l'axe de rentabilité, l'origine a été placée de façon à correspondre avec la moyenne des excédents de température sur l'ensemble des saisons. En effet, on peut aisément imaginer que les agents économiques anticipent l'impact moyen du climat, et que seules les variations exceptionnelles sont intégrées dans le cours de l'action. On constate alors qu'à l'exception de deux saisons extraordinaires où l'excédent de rentabilité attendu est de l'ordre de 15%, celui-ci se situe en général autour de 5% pour une bonne saison « ordinaire ». Dans le cas de Pernod-Ricard, l'impact attendu est environ 6 fois moins important.

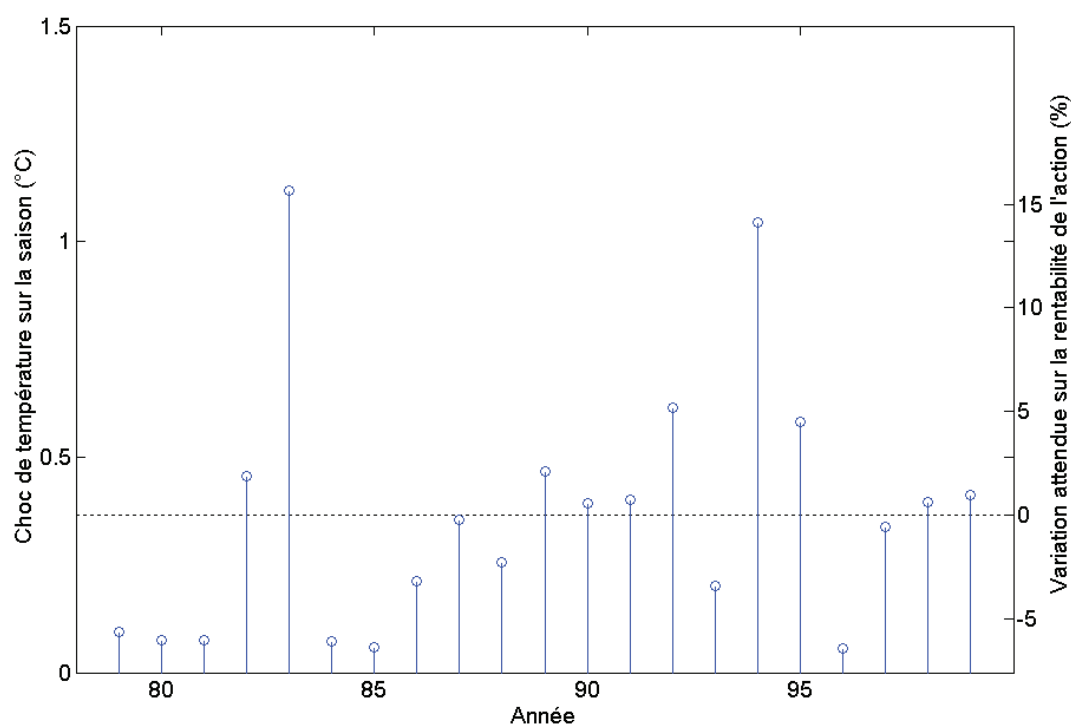


Figure 2.3. Augmentation de rentabilité attendue sur l'action Teisseire-France en fonction de la saison.

### 3. Conclusions : influence du climat sur les actions prévue par la théorie.

Dans cette première partie, nous avons essayé de déterminer l'impact théorique du climat sur les actions des entreprises qui y sont sensibles. Deux types d'influence ont été étudiés. D'une part, l'influence « normale » des variations saisonnières. D'autre part, l'influence de saisons exceptionnelles.

Dans le premier cas, même en supposant que les agents économiques n'anticipent pas l'effet du climat, on observe que celui-ci serait négligeable sur le cours des actions.

Dans le second cas, en revanche, les bénéfices dus à des conditions climatiques exceptionnellement favorables peuvent se ressentir nettement : ainsi dans le cas de Teisseire-France, l'impact a été chiffré autour de 5%, en atteignant même 15% par deux fois (sur un historique de vingt ans). L'importance de l'impact dépend directement du rapport  $CA/CB$ , entre la capitalisation boursière et le chiffre d'affaires. Si celui-ci est faible, les bénéfices réalisés sont à répartir sur un ensemble d'actions qui, en valeur, est peu important comparativement à l'ampleur des résultats. Ainsi ce rapport agit comme un véritable levier, en démultipliant l'impact des bénéfices d'autant que  $CA$  est grand devant  $CB$ .

On peut s'interroger sur la possibilité de retrouver cette augmentation par une étude statistique. Tout va dépendre, naturellement, du nombre de données à disposition. Supposons par exemple que nous ayons une centaine de données mensuelles (10 ans), et posons pour simplifier que les rentabilités mensuelles (spécifiques) sont gaussiennes de moyenne nulle et d'écart-type 5% (c'est l'ordre de grandeur de l'estimation obtenue dans la suite). L'erreur d'estimation autour de la moyenne mensuelle est donc de

$$5\% / \sqrt{100} = 0.5\%$$

Un intervalle de confiance autour de la rentabilité moyenne espérée est donc environ :

$$[-1\% ; 1\%]$$

Tout va dépendre alors de la société, bien sûr, mais aussi de la façon dont les bénéfices saisonniers sont intégrés dans les actions. En effet, si ceux-ci sont intégrés au fur et à mesure, alors il suffirait d'une augmentation annuelle de 2% (répartie en juillet et en août) pour pouvoir déceler un effet. En revanche, si les bénéfices annuels sont réinvestis mensuellement, (la décision pourrait par exemple être prise en fin d'année), alors seuls les impacts de plus de 12% seraient décelables ! Pour certaines sociétés, l'impact serait trop faible dans les deux cas. C'est le cas, par exemple, pour une société comme Pernod-Ricard pour lequel le rapport CA/CB est 6 fois moins important que celui de Teisseire-France : l'impact du climat est en général inférieur à 1%, à l'exception de deux années où il a été chiffré aux alentours de 3%.

Nous allons maintenant éprouver ces résultats théoriques et ces hypothèses avec une étude empirique, réalisée pour le cas de Teisseire-France, Pernod-Ricard et de deux autres actions.

## **II. Etude empirique de la sensibilité des actions à la température.**

Nous étudions ici d'un point de vue statistique la sensibilité des actions à la température en examinant le cas de 4 entreprises françaises « sensibles ». De façon précise, nous étudions les deux questions suivantes :

**Q1.** Le bêta action / température est-il significatif ?

**Q2.** La rentabilité de l'action est-elle liée aux températures extrêmes ?

La première est liée à la recherche de corrélations linéaires ; la seconde aux non-linéarités. Cette dernière mérite quelques précisions. Nous savons que les températures extrêmes ont un rôle sur le marché dans son ensemble ; il ne s'agit pas ici de rechercher cet effet sur les actions étudiées. Au contraire on souhaite savoir si la rentabilité spécifique de ces actions est sensible aux températures extrêmes : on veut donc éliminer le plus possible la cause comportementale (responsable, semble-t-il, de la réponse du marché à la température) pour se concentrer sur les causes économiques. Dans l'idéal, on aimerait répondre à la question :

**Q2bis.** La rentabilité de l'action est-elle liée aux ventes exceptionnelles dues aux températures extrêmes ?

Les deux questions **Q2** et **Q2bis** ne sont pas sans lien, car le climat a un impact réel sur les ventes des entreprises étudiées<sup>77</sup> ; cependant, les records de vente peuvent aussi provenir d'une décision commerciale (promotion). Il semble ainsi utopique de vouloir distinguer dans l'effet des températures extrêmes la contribution des ventes.

En résumé, en posant la question **Q2**, on souhaite répondre avec des outils statistiques au problème soulevé au point 2 du paragraphe précédent de savoir si les ventes exceptionnelles dues au climat sont intégrées dans le cours des actions.

Le paragraphe est structuré comme suit. Dans le point 1, on présente les entreprises étudiées et les données d'action. Dans le point 2 on examine la possibilité de distinguer les causes comportementales des causes économiques. Dans le point 3, une première étude empirique est menée à bien. Dans le point 4 enfin, on tient compte de façon plus précise de l'anomalie comportementale de température.

### **1. Présentation des entreprises étudiées. Données d'action.**

Les entreprises méritant l'appellation « sensible au climat » doivent satisfaire au moins deux critères :

- secteur d'activité spécifique, sensibilité au climat
- concentration géographique de la destination des ventes

---

<sup>77</sup> On pourra consulter l'étude de Garcia et Sturzenegger, 2001, pour des résultats quantitatifs à propos de l'impact des températures sur les ventes de jus de fruits.



**Tableau 2.5. Sélection d'actions sensibles au climat.**

Code Sicovam	Désignation des valeurs et marché de cotation	Secteur d'activité principal – Répartition sectorielle du CA	Répartition géographique du CA	Risque de température
3102	Electricité de Strasbourg (premier marché)	Energie	Alsace du Nord	Eté frais Hiver doux
12587	Euro-Disney (premier marché ▲)	Tourisme - parc à thème : 48% - hôtels et Disney Village : 38.5% - autres : 13.5%	Paris	Chutes de température [Pluie]
12069	Pernod-Ricard (premier marché ▲)	Boisson - spiritueux et vins : 42% - fruits transformés : 27% - distribution : 31%	France : 28.5% Europe : 45.5% Autres : 26%	Eté frais
3953	Teisseire-France (second marché)	Boisson - sirops : 71% - jus de fruits et prêt-à-boire : 19.5% - autres : 9.5%	France : 80.5%	Eté frais

▲ Valeur du SBF 80.

(\*) Les chiffres donnés sont relatifs à l'année 2001 (source : rapport annuel Teisseire-France 2001, [www.teisseire.fr](http://www.teisseire.fr) ; rapport annuel Pernod-Ricard 2001, [www.pernod-ricard.fr](http://www.pernod-ricard.fr), rapport annuel Euro Disney, 2001, [www.eurodisney.com](http://www.eurodisney.com)).

D'autre part, pour mener à bien l'étude statistique, on a dû imposer une longueur minimale aux historiques de cours d'actions.

Finalement, notre choix s'est porté sur les 4 actions dont quelques caractéristiques sont présentées dans le Tableau 2.5. La longueur des historiques est précisée dans le Tableau 2.1 (page 140). Nous avons retenu Euro-Disney bien que le risque climatique direct soit la pluie, en raison de la corrélation importante avec la température : les précipitations sont souvent accompagnées d'une chute de température. Parmi le choix de ces valeurs, la sélection de Pernod-Ricard est la plus discutable car un faible pourcentage des ventes concernent la France.

## **2. Problème de l'anomalie comportementale de température.**

Pour l'étude de la question **Q2**, on souhaiterait pouvoir éliminer l'effet global de la température sur le marché, d'origine comportementale, et les effets d'origine purement économiques. Sans quoi, les conclusions pourraient être faussées. Ainsi, dans le cas d'une entreprise fournisseur d'énergie comme Electricité de Strasbourg, les causes économique et comportementale devraient occasionner toutes deux une augmentation de la rentabilité des cours les jours de basse température. Comment discerner alors l'impact des ventes dues à la bonne saison ? Pour ce cas, il sera d'ailleurs difficile de distinguer totalement les deux effets, vu qu'ils ont la même origine. En revanche, la situation des entreprises du secteur de la boisson est plus avantageuse puisque le risque « commercial » de température est présent en été. Il est alors envisageable d'éliminer l'anomalie comportementale de température des causes possibles. Dans le cas général, on peut espérer réduire l'impact de l'anomalie de température sur les actions en considérant les rentabilités spécifiques de ces actions ; en effet comme cette anomalie est ressentie par l'ensemble du marché, on peut s'attendre à ce que la part spécifique des actions n'y soit pas sensible.

Nous proposons maintenant de vérifier empiriquement cette hypothèse. Pour cela, il est préférable de sélectionner les entreprises pour lesquelles l'anomalie est clairement identifiable : nous avons donc écarté Electricité de Strasbourg et Euro-Disney pour retenir les valeurs Pernod-Ricard et Teisseire-France pour lesquelles une hausse du cours lors des mois froids ne peut vraisemblablement pas être due à un surcroît de ventes. Pour chacune de ces

actions, nous extrayons la rentabilité spécifique en effectuant la régression linéaire de la rentabilité sur la rentabilité d'un indice de marché :

$$R_{a,t} = \alpha + \beta_{a|m} \cdot R_{m,t} + R_{a,t}^{\text{spec}}$$

avec des notations évidentes. Le choix de l'indice de marché est indiqué dans le Tableau 2.7, page 162, et sera justifié au paragraphe suivant (point a). Il n'a pas, cependant, d'incidence notable sur la suite. Nous reproduisons alors l'étude réalisée au A.II.4 pour la détection de l'anomalie de température en remplaçant les rentabilités par les rentabilités spécifiques.

Les résultats obtenus sont montrés dans le Tableau 2.6. Dans le cas de Pernod-Ricard, nous n'observons plus une rentabilité supérieure pour les mois froids ; pour Teisseire-France, le phénomène est atténué mais encore visible ; toutefois, il n'est plus significatif. Au plan statistique, on a donc éliminé l'anomalie de température pour ces actions. On peut conjecturer que ce résultat s'étend aux autres actions, pour lesquelles la vérification empirique semble impossible.

**Tableau 2.6. Résolution de l'anomalie (basses températures) pour les actions**

Action \ Indice HDD		Nombre de classes = 2			Nombre de classes = 3			
		classe 1	classe 2	p-valeur	classe 1	classe 2	classe 3	p-valeur
<u>Rentabilités brutes</u>								
Pernod-Ricard	Moyenne	-0.0120	0.0128	0.0655	-0.0205	0.0133	0.0134	0.0665
	Ecart-type	0.0646	0.0702		0.0654	0.0587	0.0750	
Teisseire-France	Moyenne	-0.0211	0.0183	0.0091	-0.0188	-0.0182	0.0451	0.0045
	Ecart-type	0.0722	0.0881		0.0747	0.0724	0.0907	
<u>Rentabilités spécifiques</u>								
Pernod-Ricard / SBF 80	Moyenne	-0.0032	-0.0040	0.4908	-0.0089	0.0040	-0.0034	0.3589
	Ecart-type	0.0617	0.0629		0.0660	0.0472	0.0695	
Teisseire-France / Second Marché	Moyenne	-0.0081	0.0088	0.1203	-0.0013	-0.0167	0.0167	0.1837
	Ecart-type	0.0662	0.0685		0.0687	0.0633	0.0680	

La conclusion de cette petite étude est remarquable à plusieurs points de vue. Tout d'abord, elle illustre sur ces exemples que l'anomalie de température mise en évidence par Cao et Wei est bien un phénomène global, touchant l'ensemble du marché. Ensuite et par conséquent, en dégageant la part spécifique des actions, on fait d'une pierre deux coups : on élimine les contributions du marché et de l'anomalie de température.

### 3. Etude empirique.

Nous essayons maintenant d'apporter une réponse aux questions **Q1** et **Q2**. La méthodologie suivie pour traiter la première est éprouvée : il s'agit de tester la nullité du coefficient de la régression linéaire de la rentabilité spécifique de chaque action sur un facteur (en l'occurrence, la température). Cependant, un soin particulier est accordé au choix des variables de travail de façon à réduire au maximum l'incertitude statistique. C'est l'objet du travail préliminaire. On étudie ensuite la question **Q2** grâce à un test d'égalité des moyennes portant sur les rentabilités spécifiques dans deux classes déterminées par un seuil de température.

a) Travail préliminaire.

Il reste encore des degrés de liberté sur les variables d'étude :

- l'indice de marché utilisé pour extraire la rentabilité spécifique ;
- la localisation des stations sur lesquelles se fait la mesure des températures.

Nous précisons ici nos choix.

#### • *Choix des indices de marché - Réduction de variance.*

On sait que dans une régression linéaire, la plus grande réduction de variance correspond au plus grand  $R^2$  ; en outre, la variance est divisée par  $1 - R^2$  (voir par exemple Campbell, Lo, MacKinlay, 1997, § 4.4.4.). Pour chaque action, nous avons donc choisi l'indice de marché qui donne le plus grand  $R^2$  dans la régression de la rentabilité de l'action sur celle de l'indice :

$$R_{a,t} = \alpha + \beta_{a|m} \cdot R_{m,t} + R_{a,t}^{\text{spec}}$$

Les résultats sont indiqués dans le Tableau 2.7. Il est rassurant de retrouver par ce biais les indices principaux auxquels appartiennent deux des actions étudiées : Pernod-Ricard et Euro-Disney sont ainsi des valeurs de l'indice SBF 80. D'autre part, Teisseire-France est coté au second marché, ce qui correspond au résultat de notre sélection statistique. Pour conclure sur

ce point, notons que l'ampleur de la réduction de variance est limitée ; le choix d'autres indices ne modifie pas les conclusions de la présente étude.

Tableau 2.7. Choix de l'indice de marché pour chaque station

Action\Indice	CAC 40	SBF 80	SBF 250	MidCac	Sec. Marché
Electricité de S.	0.0182	0.0632	0.0399	0.0683	<b>0.0740</b>
Eurodisney	0.0850	<b>0.0977</b>	0.0821	0.0718	0.0723
Pernod-Ricard	0.2040	<b>0.2313</b>	0.2064	0.1547	0.1400
Teisseire-France	0.0595	0.1467	0.0672	0.1642	<b>0.1811</b>

Le tableau mentionne les valeurs de  $R^2$  estimées dans la régression de la rentabilité d'une action sur la rentabilité de l'indice de marché. L'indice choisi est celui correspondant au plus grand  $R^2$ .

Les valeurs correspondantes des bêta action/marché sont indiquées ci-dessous.

Tableau 2.8. Bêta action/marché estimés

Action\Indice	CAC 40	SBF 80	SBF 250	MidCac	Second Marché
Electricité de S.	0.1415 (0.0956)	0.3239 (0.1211)	0.2402 (0.1144)	0.2914 (0.1045)	<b>0.3966</b> <b>(0.1362)</b>
Euro-Disney	0.5861 (0.1778)	<b>0.7995</b> <b>(0.2372)</b>	0.6841 (0.2233)	0.5932 (0.2081)	0.7844 (0.2741)
Pernod-Ricard	0.5511 (0.1161)	<b>0.7123</b> <b>(0.1384)</b>	0.6120 (0.1279)	0.5071 (0.1264)	0.6089 (0.1609)
Teisseire-France	0.3469 (0.1270)	0.6473 (0.1516)	0.4087 (0.1479)	0.5926 (0.1298)	<b>0.8139</b> <b>(0.1681)</b>

L'écart-type d'estimation figure entre parenthèses. Les valeurs indiquées en gras sont relatives à l'indice précédemment choisi pour chaque action.

• *Choix de la localisation des stations*

Dans certains cas, comme pour Electricité de Strasbourg ou Euro-Disney, la destination géographique des ventes des entreprises est suffisamment précise pour éliminer toute ambiguïté sur les stations météorologiques à utiliser. Une ambiguïté bien réelle, en revanche, subsiste pour Teisseire-France et Pernod-Ricard qui distribuent leurs boissons à l'échelle nationale (voire mondiale pour le second). Pour ce secteur d'activité, les températures du sud de la France sont certainement les plus explicatives, et une variable naturelle est donnée par la

température moyenne d'un groupe de villes situées au sud de la France. La façon de constituer ce groupe n'a, en pratique, aucune incidence sur les résultats, ce qui s'explique par la forte corrélation spatiale des températures. Dans la suite, nous avons opté pour le groupe de 5 stations situées au sud de Lyon : {Lyon, Toulouse, Marseille, Nice, Ajaccio}, que nous avons baptisé « Sud ». Pour la question Q1, nous avons également indiqué les résultats correspondants à chaque station individuelle pour illustrer les différences de pouvoir explicatif.

b) Etude de la question **Q1**

Maintenant que les choix d'indice de marché et de stations météorologiques ont été explicités, nous étudions la signification du bêta action/température. Nous utilisons une procédure en deux étapes :

- obtention de la rentabilité spécifique par régression linéaire de la rentabilité de l'action sur la rentabilité de l'indice de marché correspondant

$$R_{a,t} = \alpha + \beta_{a|m} \cdot R_{m,t} + R_{a,t}^{\text{spec}}$$

- estimation du bêta action/température par régression de la rentabilité spécifique sur la température stationnarisée (voir A.II.3) :

$$R_{a,t}^{\text{spec}} = \gamma + \beta_{a|\text{temp}} \cdot \varepsilon_{\text{temp},t} + u_t$$

Nous avons également envisagé la possibilité de retards en remplaçant  $\varepsilon_{\text{temp},t}$  par

$$\varepsilon_{\text{temp},t-h}, \quad h = 1, 2, 3, \dots$$

On pouvait naturellement employer une procédure en une seule étape en régressant d'un seul coup la rentabilité de l'action sur les deux facteurs « marché » et « température »; les résultats sont très proches.

Les valeurs des bêtas action/température estimées avec la procédure en 2 étapes sont données dans le Tableau 2.9 et le Tableau 2.10 ; les valeurs significatives aux niveaux 10%, 5% et 1% sont également indiquées. Dans les cas les plus significatifs, le tracé de la rentabilité de l'action contre la température est effectué (Figure 2.4 et Figure 2.5).

Tableau 2.9. Bêtas action/température estimés.

Station\Action	Electricité de Strasbourg	Eurodisney	Pernod -Ricard	Teisseire -France
<b>Lille</b>	0.0470 (0.0468)	0.1092 (0.0914)	0.0389 (0.0533)	0.0734 (0.0575)
<b>Brest</b>	-0.0019 (0.0465)	0.1361 (0.0902)	0.0208 (0.0532)	0.0484 (0.0572)
<b>Paris</b>	0.0306 (0.0488)	* <b>0.1552</b> <b>(0.0945)</b>	0.0823 (0.0546)	0.0628 (0.0600)
<b>Strasbourg</b>	<b>0.0341</b> <b>(0.0496)</b>	0.1395 (0.0967)	0.0515 (0.0561)	0.0586 (0.0611)
<b>Bordeaux</b>	0.0084 (0.0483)	* 0.1613 (0.0933)	* 0.1060 (0.0558)	0.0446 (0.0594)
<b>Clermont-F.</b>	0.0007 (0.0473)	0.1448 (0.0917)	* 0.0970 (0.0543)	0.0599 (0.0581)
<b>Lyon</b>	0.0249 (0.0480)	* 0.1540 (0.0929)	** 0.1090 (0.0539)	0.0466 (0.0591)
<b>Toulouse</b>	0.0103 (0.0488)	* 0.1596 (0.0944)	** 0.1426 (0.0565)	0.0473 (0.0601)
<b>Marseille</b>	0.0161 (0.0485)	0.1073 (0.0944)	*** 0.1620 (0.0540)	0.0291 (0.0598)
<b>Nice</b>	0.0222 (0.0470)	0.1370 (0.0912)	*** 0.1541 (0.0534)	0.0912 (0.0574)
<b>Ajaccio</b>	-0.0105 (0.0394)	0.0889 (0.0767)	** 0.1105 (0.0444)	0.0410 (0.0484)
<b>Nord</b>	- -	0.1496 (0.0981)	- -	- -
<b>Est</b>	0.0388 (0.0512)	- -	- -	- -
<b>Sud</b>	- -	- -	*** <b>0.1617</b> <b>(0.0568)</b>	<b>0.0609</b> <b>(0.0619)</b>

## Notes.

1. En dessous de chaque estimation, entre parenthèses, figure l'erreur d'estimation. Les symboles \*, \*\*, \*\*\* indiquent les valeurs significatives aux niveaux respectifs 10%, 5% et 1%. Les valeurs surlignées en gras sont relatives à notre choix de stations.

2. Les mentions Nord, Est, Sud correspondent à un groupement de stations (voir Chapitre 2.B.II.3.a) :

Nord : Lille, Brest, Paris, Strasbourg

Est : Lille, Paris, Strasbourg, Lyon

Sud : Lyon, Toulouse, Marseille, Nice, Ajaccio

Tableau 2.10. Bêtas action/température estimés, avec possibilité de retards.

	retard = 0	retard = 1	retard = 2	retard = 3
<b>Electricité de S. / Strasbourg</b>	0.0341 (0.0496)	-0.0135 (0.0497)	-0.0642 (0.0497)	-0.0396 (0.0499)
<b>Euro-Disney / Paris</b>	* 0.1552 (0.0945)	0.1512 (0.0946)	0.1060 (0.0957)	0.0274 (0.0963)
<b>Pernod-Ricard / Sud</b>	*** 0.1617 (0.0568)	0.0286 (0.0586)	-0.0120 (0.0582)	0.0940 (0.0575)
<b>Teisseire-France / Sud</b>	0.0609 (0.0619)	-0.0181 (0.0615)	-0.0946 (0.0612)	-0.0022 (0.0612)

*Notes.*

1. En dessous de chaque estimation, entre parenthèses, figure l'erreur d'estimation. Les symboles \*, \*\*, \*\*\* indiquent les valeurs significatives aux niveaux respectifs 10%, 5% et 1%.
2. La mention Sud correspond au groupement de stations : Lyon, Toulouse, Marseille, Nice, Ajaccio.

Tableau 2.11. Bêtas action/température estimés, avec possibilité de retards et en l'absence de rentabilités exceptionnelles.

	retard = 0	retard = 1	retard = 2	retard = 3
<b>Electricité de S. (1) / Strasbourg</b>	0.0124 (0.0441)	-0.0097 (0.0440)	* -0.0839 (0.0437)	-0.0612 (0.0440)
<b>Euro-Disney (1) / Paris</b>	0.0105 (0.0865)	0.0992 (0.0829)	0.1187 (0.0835)	0.0106 (0.0835)
<b>Pernod-Ricard / Sud</b>	*** 0.1617 (0.0568)	0.0286 (0.0586)	-0.0120 (0.0582)	0.0940 (0.0575)
<b>Teisseire-France (1) / Sud</b>	0.0592 (0.0592)	-0.0401 (0.0591)	-0.0742 (0.0592)	0.0035 (0.0586)

*Notes.*

1. On a enlevé ici quelques rentabilités exceptionnelles : pour Electricité de Strasbourg, mai 1998 (+29.6%) ; pour Euro-Disney, novembre 1993 (-41.1%), novembre 1994 (+25.9%) et novembre 1999 (-35.7%) ; pour Teisseire-France, janvier 1994 (+27.7%).
2. En dessous de chaque estimation, entre parenthèses, figure l'erreur d'estimation. Les symboles \*, \*\*, \*\*\* indiquent les valeurs significatives aux niveaux respectifs 10%, 5% et 1%.
3. La mention Sud correspond au groupement de stations : Lyon, Toulouse, Marseille, Nice, Ajaccio.



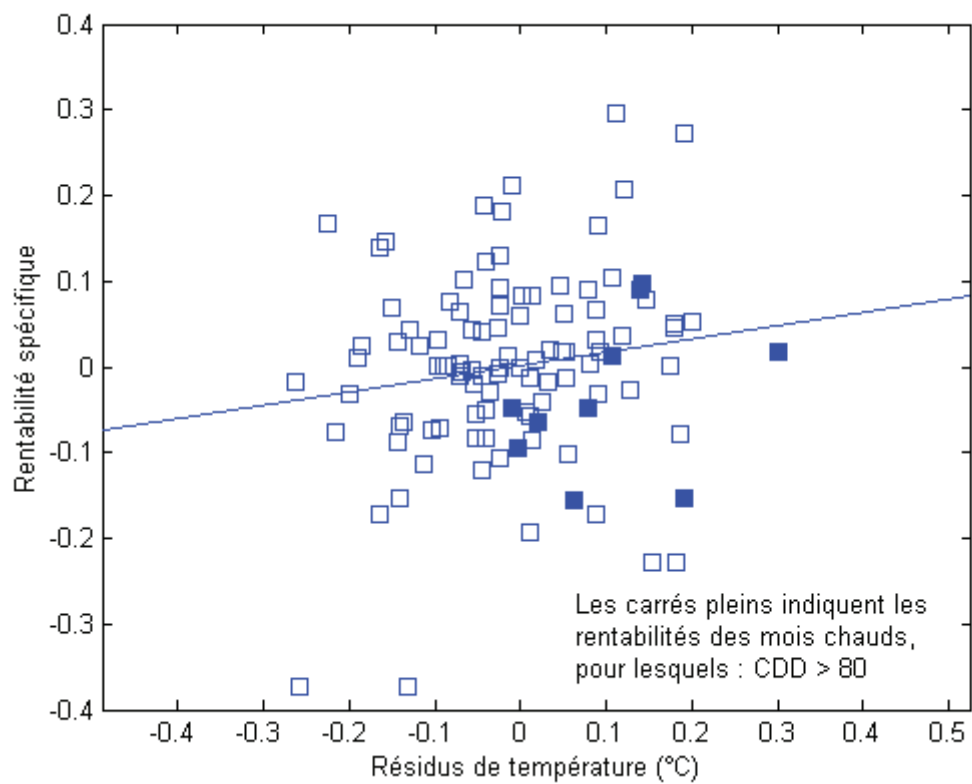


Figure 2.4. Tracé de la rentabilité de l'action Euro-Disney contre la température à Paris.

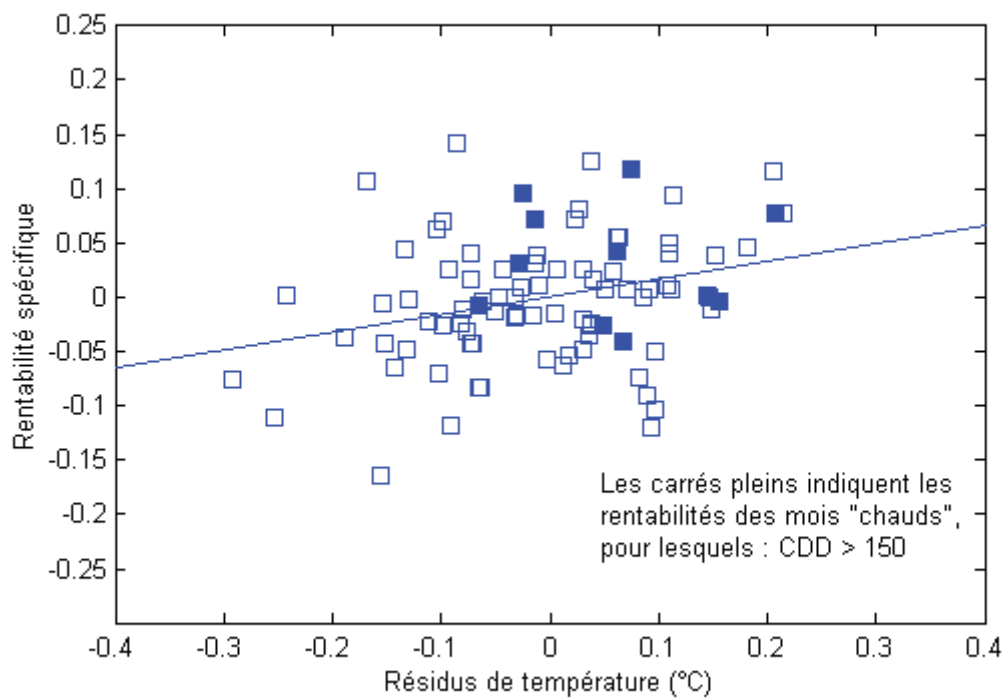


Figure 2.5. Tracé de la rentabilité de l'action Pernod-Ricard contre la température du sud de la France.

D'une façon générale nous constatons que les valeurs estimées sont faibles (relativement, par exemple, aux bêtas action/marché) et peu ou pas significatifs ; d'autre part l'importance de certains bêtas s'explique par la présence de quelques rentabilités extrêmes, comme pour Euro-Disney (voir Tableau 2.11.). En se plaçant au niveau usuel de confiance de 5%, seuls les bêtas estimés relatifs à Pernod-Ricard sont significatifs. Ce dernier résultat est quelque peu surprenant. En effet, Il est contraire aux anticipations que l'on pouvait faire à la vue des informations économiques des différentes sociétés (*cf.* Tableau 2.5) : la diversification géographique assez étendue fait de Pernod-Ricard l'entreprise *a priori* la moins exposée au risque de température parmi les quatre étudiées ici. En outre, la société Teisseire-France – qui est pourtant dans le même secteur d'activité, ne semble présenter aucune sensibilité à la température spécifique. Remarquons cependant qu'il y a une certaine cohérence des résultats au niveau géographique par rapport à l'impact économique attendu : ainsi dans le cas de Pernod-Ricard, la sensibilité est plus forte pour les stations situées au sud de la France (on peut également observer, dans l'ensemble, une certaine continuité spatiale dans les valeurs estimées).

Aussi, dans le seul cas où il est significatif, nous nous trouvons dans l'embarras pour donner une justification sensée du bêta action/température : la corrélation entre température et action est acquise ; le lien de cause à effet est, lui, douteux.

c) Etude de la question Q2.

Pour savoir si les températures extrêmes ont un lien avec la rentabilité des actions, nous employons une méthodologie analogue à celle que Cao et Wei ont utilisé pour détecter l'anomalie de température. Rappelons à ce propos que celle-ci a été en grande partie éliminée dans les rentabilités spécifiques des actions (paragraphe 2 de cette section, et Tableau 2.6). Rappelons également que – dans l'idéal -, nous cherchons à exhiber d'éventuelles répercussions sur les cours des impacts économiques causés par les températures extrêmes. « L'extrême » ici va donc dans le sens des risques : hautes températures pour Pernod-Ricard et Teisseire (augmentation de la consommation), hautes températures pour Euro-Disney (en raison de l'augmentation prévisible du nombre de visiteurs), basses ou hautes températures pour Electricité de Strasbourg (augmentation de la consommation : chauffage l'hiver, climatisation l'été). Ce sont ces situations qui nous intéressent particulièrement ici. Dans un souci d'exhaustivité, nous présentons néanmoins tous les cas de figure.

Nous procédons comme suit. Nous raisonnons sur les rentabilités spécifiques et, pour chaque action, nous constituons deux classes selon que la température mensuelle est jugée extrême ou non. Les extrêmes sont mesurés au moyen des indices HDD ou CDD. Pour choisir le seuil de sélection, un compromis doit être trouvé entre le souhait d'opter pour le niveau le plus élevé possible (de façon à sélectionner des températures réellement exceptionnelles) et la nécessité de conserver un nombre suffisant de données pour l'étude statistique. Les seuils ont alors été déterminés au cas par cas, en réservant au moins 10 données dans la classe « extrême ». Enfin, nous mettons en œuvre le test de Wilcoxon portant sur les moyennes opposant l'hypothèse nulle  $H_0 : \{ER_2 = ER_1\}$  à l'hypothèse alternative  $H_1 : \{ER_2 > ER_1\}$  où  $R_2$  représente la rentabilité spécifique de la classe « extrême ».

Les résultats sont montrés dans le Tableau 2.12 et le Tableau 2.13.

Commençons tout d'abord par le second tableau qui s'intéresse aux situations où les extrêmes de température ne sont pas de nature à avoir un impact économique sur les sociétés étudiées. Parmi les trois cas présentés, un seul est significatif correspondant aux basses températures et Teisseire-France. On peut sans doute invoquer ici un effet « parasite » de l'anomalie de température, le même qui semblait visible (sans être significatif pour autant) sur la rentabilité spécifique de Teisseire-France (voir § 2 et Tableau 2.6) ; le niveau des températures extrêmes, plus élevé dans notre cas (la classe « extrême » contient environ 13% des données contre 35% dans le Tableau 2.6.), peut contribuer à accroître l'importance de l'effet. Nous aurons la confirmation de cette hypothèse au prochain paragraphe. Enfin, on peut remarquer qu'aucun autre lien entre les basses températures et les rentabilités spécifiques n'est à signaler.

Examinons ensuite le premier tableau qui vise davantage à apporter une réponse à la question **Q2bis**. On constate que les températures extrêmes n'ont pas un impact significatif en général, à l'exception de l'action Pernod-Ricard. Les résultats de l'étude de la question **Q1** sont tout à fait cohérents avec ce dernier résultat, et nous renvoyons le lecteur à la conclusion du paragraphe précédent en ce qui concerne le cas particulier de Pernod-Ricard.

Tableau 2.12. Impact des températures extrêmes sur les rentabilités spécifiques des actions (1).

	Echantillon 1	Echantillon 2	p-valeur
<b>Electricité de S. / Strasbourg</b>	$HDD \leq 480$ (98 données) $m = 0.0015$ $s = 0.0575$	$HDD > 480$ (10 données) $m = -0.0152$ $s = 0.0425$	0.8105
	$CDD \leq 60$ (96 données) $m = -0.0018$ $s = 0.0581$	$CDD > 60$ (12 données) $m = 0.0141$ $s = 0.0393$	0.0825
<b>Eurodisney / Paris</b>	$CDD \leq 80$ (97 données) $m = 0.0035$ $s = 0.1121$	$CDD > 80$ (10 données) $m = -0.0343$ $s = 0.0889$	0.8805
<b>Pernod-Ricard / Sud</b>	$CDD \leq 150$ (79 données) $m = -0.0046$ $s = 0.0588$	$CDD > 150$ (11 données) $m = 0.0328$ $s = 0.0525$	0.0310
<b>Teisseire-France / Sud</b>	$CDD \leq 150$ (95 données) $m = -0.0012$ $s = 0.0672$	$CDD > 150$ (13 données) $m = 0.0089$ $s = 0.0871$	0.3687

Note.

$m$  et  $s$  désignent respectivement la moyenne et l'écart-type.

**Tableau 2.13. Impact des températures extrêmes sur les rentabilités spécifiques des actions (2).**

	<b>Echantillon 1</b>	<b>Echantillon 2</b>	<b>p-valeur</b>
<b>Eurodisney / Paris</b>	$HDD \leq 400$ (97 données) $m = 0.0006$ $s = 0.1051$	$HDD > 400$ (10 données) $m = -0.0057$ $s = 0.1596$	0.3986
<b>Pernod-Ricard / Sud</b>	$HDD \leq 300$ (80 données) $m = 0.0014$ $s = 0.0586$	$HDD > 300$ (10 données) $m = -0.0115$ $s = 0.0650$	0.7870
<b>Teisseire-France / Sud</b>	$HDD \leq 300$ (94 données) $m = -0.0068$ $s = 0.0647$	$HDD > 300$ (14 données) $m = 0.0459$ $s = 0.0850$	0.0141

*Note.*

$m$  et  $s$  désignent respectivement la moyenne et l'écart-type.

#### 4. Prise en compte plus fine de l'anomalie de température.

Nous avons vu que la considération des rentabilités spécifiques permettait, dans une large mesure, d'écartier la cause comportementale dans l'explication de liens éventuels entre action et température (§ 2.). Toutefois, nous suspectons que cette cause ne peut être écartée totalement : il peut subsister une part de l'effet de l'anomalie dans les résidus de la régression de la rentabilité de l'action sur la rentabilité de marché (même paragraphe, cas de Teisseire-France). Or il est possible d'aller plus loin dans l'élimination de l'anomalie lorsque les basses températures (qui lui sont, pour l'essentiel, associées) ne peuvent être jugées responsables d'augmentation des ventes (voir la discussion du paragraphe 2), autrement dit lorsqu'il y a *disjonction* entre causes comportementales et causes économiques. Cette situation se présente notamment pour les entreprises du secteur de la boisson, représentées ici par Teisseire-France et Pernod-Ricard.

Pour chacune d'elles, nous allons raffiner la procédure d'extraction de la rentabilité spécifique en ajoutant une nouvelle variable, sensée absorber la part résiduelle de l'anomalie de température. Nous proposons :

$$1_{\{HDD > x\}}$$

où  $x$  est un seuil déterminant le « niveau d'extrême ». En effet, les valeurs élevées de l'indice HDD correspondent aux basses températures ; comme celles-ci ne peuvent avoir d'impact sur les ventes, nous pouvons interpréter cette variable comme une variable « comportementale » de température. Dans la suite, nous choisissons pour  $x$  le quantile des données d'indice HDD d'ordre 85%. Les conclusions sont robustes vis-à-vis d'autres valeurs de  $x$  supérieures à 75%. La rentabilité spécifique s'obtient donc maintenant par régression de la rentabilité de l'action sur la rentabilité de l'indice de marché et la variable comportementale de température :

$$R_{a,t} = \alpha + \beta_{a|m} \cdot R_{m,t} + \beta_{a|temp}^{comp} \cdot 1_{\{HDD > x\}} + R_{a,t}^{spec}$$

Partant, nous recommandons l'étude empirique avec cette nouvelle base en suivant les mêmes méthodologies (voir les points 3.b) et 3.c)).

• *Résultats.*

Les résultats sont donnés dans le Tableau 2.14, le Tableau 2.15 et le Tableau 2.16. Pour Pernod-Ricard, nous ne constatons aucune modification : les bêtas action/marché, action/température sont inchangés. Pour Teisseire-France en revanche, la variable comportementale semble faire pleinement office du rôle qu'on lui prêtait : le coefficient estimé correspondant est positif de façon significative, ce qui incline à penser que la part résiduelle de l'effet de l'anomalie de température est bien capturée par la nouvelle variable. Le résultat du Tableau 2.16 va dans le même sens : il n'y a plus d'impact des basses températures sur les nouvelles rentabilités spécifiques (on rappelle que pour les deux actions étudiées, basses températures et anomalie de température peuvent être confondues). (Au passage, ceci confirme l'explication avancée au 3.c)). Par voie de conséquence, l'estimation des autres coefficients est modifiée – ainsi que leur interprétation : la température n'intervient plus que par ses conséquences économiques « physiques ». Le bêta action/marché diminue de 0.81 à 0.72, le bêta action/température augmente de 0.06 à 0.09. La sensibilité de l'action à la température est donc plus importante – elle est même significative au niveau 10% lorsqu'on quantifie les températures extrêmes par un quantile à 90%.

**Tableau 2.14. Bêtas estimés de l'action avec le marché et la température**

	<b>Ancienne procédure</b>	<b>Nouvelle procédure</b>
<b>Pernod-Ricard</b>	$R_{a,t} = -0.0069 + 0.7123 R_{m,t} + R_{a,t}^{\text{spec}}$ <p style="text-align: center;">(0.0065) (0.1384)</p> $R_{a,t}^{\text{spec}} = 0.0008 + 0.1617 \varepsilon_{\text{temp},t} + u_t$ <p style="text-align: center;">(0.0060) (0.0568)</p>	$R_{a,t} = -0.0057 + 0.7290 R_{m,t} - 0.0095 X_t(x) + R_{a,t}^{\text{spec}}$ <p style="text-align: center;">(0.0069) (0.1430) (0.0184)</p> $R_{a,t}^{\text{spec}} = 0.0007 + 0.1573 \varepsilon_{\text{temp},t} + v_t$ <p style="text-align: center;">(0.0060) (0.0569)</p>
<b>Teisseire-France</b>	$R_{a,t} = -0.0114 + 0.8139 R_{m,t} + R_{a,t}^{\text{spec}}$ <p style="text-align: center;">(0.0068) (0.1681)</p> $R_{a,t}^{\text{spec}} = 0.0006 + 0.0609 \varepsilon_{\text{temp},t} + v_t$ <p style="text-align: center;">(0.0067) (0.0619)</p>	$R_{a,t} = -0.0178 + 0.7190 R_{m,t} + 0.0477 X_t(x) + R_{a,t}^{\text{spec}}$ <p style="text-align: center;">(0.0071) (0.1683) (0.0189)</p> $R_{a,t}^{\text{spec}} = 0.0009 + 0.0903 \varepsilon_{\text{temp},t} + v_t$ <p style="text-align: center;">(0.0065) (0.0597)</p>

*Note.*

$X_t(x) = 1_{\{HDD > x\}}$ ,  $x$  est le quantile à 85% des données d'indice HDD.

Tableau 2.15. Impact des températures extrêmes sur les nouvelles rentabilités spécifiques des actions (1).

	Echantillon 1	Echantillon 2	p-valeur
<b>Pernod-Ricard / Sud</b>	$CDD \leq 150$ (79 données) $m = -0.0044$ $s = 0.0589$	$CDD > 150$ (11 données) $m = 0.0316$ $s = 0.0521$	0.0346
<b>Teisseire-France / Sud</b>	$CDD \leq 150$ (95 données) $m = -0.0020$ $s = 0.0647$	$CDD > 150$ (13 données) $m = 0.0149$ $s = 0.0869$	0.2528

Note.

$m$  et  $s$  désignent respectivement la moyenne et l'écart-type.

Tableau 2.16. Impact des températures extrêmes sur les nouvelles rentabilités spécifiques des actions (2).

	Echantillon 1	Echantillon 2	p-valeur
<b>Pernod-Ricard / Sud</b>	$HDD \leq 300$ (80 données) $m = 0.0005$ $s = 0.0587$	$HDD > 300$ (10 données) $m = -0.0039$ $s = 0.0651$	0.6780
<b>Teisseire-France / Sud</b>	$HDD \leq 300$ (94 données) $m = -0.0011$ $s = 0.0649$	$HDD > 300$ (14 données) $m = 0.0074$ $s = 0.0854$	0.5182

Note.

$m$  et  $s$  désignent respectivement la moyenne et l'écart-type.



• *Remarques.*

Terminons le paragraphe par deux remarques. Signalons d'abord que la procédure en une seule étape (régression simultanée sur toutes les variables « marché », comportementale, « température ») conduit à des résultats très proches et à des conclusions identiques. D'autre part, on pourrait s'inquiéter que les précédents résultats puissent être faussés en raison de la corrélation entre les deux variables explicatives dépendant de la température  $1_{\{HDD>x\}}$  et  $\varepsilon_{temp,t}$ . Toutefois, cette corrélation est estimée à  $-0.13$  dans le cas de Pernod-Ricard et  $-0.17$  pour Teisseire-France (estimation obtenue avec la procédure en une étape), ce qui est acceptable ; le fait que l'introduction de la nouvelle variable ne modifie pas l'estimation dans le cas de Pernod-Ricard est également un signe réconfortant.

**5. Remarque sur l'effet janvier.**

Une rentabilité anormalement élevée est souvent observée sur les actions au cours du mois de janvier, c'est « l'effet janvier » (voir par exemple Hamon, Jacquillat, 1992). S'agissant d'un phénomène saisonnier, il eût été préférable d'en tenir compte dans notre discussion. Cependant, la variable discrète correspondante (valant 1 pour le mois de janvier et 0 sinon) est fortement corrélée avec la variable comportementale  $1_{\{HDD>x\}}$  introduite au paragraphe précédent. Cela signifie deux choses : d'une part qu'il est difficile d'isoler cet effet de l'anomalie de température et, *ipso facto*, que l'on tient compte de l'effet janvier pour une bonne part chaque fois que l'on utilise la variable comportementale. La distinction des deux anomalies est possible car l'effet janvier touche principalement les entreprises qui ont vu leur cours chuter en fin d'année précédente (même référence), il y a donc possibilité de le repérer au moyen d'une variable économique. Cependant cette distinction est du second ordre, nous ne l'avons pas abordée ici.

**6. Conclusions de l'étude empirique.**

Nous avons étudié la sensibilité de quelques actions à la température à partir de quatre entreprises « sensibles au climat » (voir § 1. pour les critères de sélection et la présentation détaillée de ces entreprises). Nous avons étudié l'impact des températures extrêmes sur la rentabilité des actions. Un effet résiduel dû à l'anomalie de température, observé par Cao et

Wei (2001), peut subsister bien que cette anomalie touche le marché dans sa globalité et qu'elle soit en grande partie absente dans les rentabilités spécifiques. Une fois que l'on a tenu compte de ce phénomène global, plus aucun impact des températures extrêmes n'est décelable, à l'exception de Pernod-Ricard. Nous avons également examiné la question de l'existence d'un bêta spécifique à la température, en éliminant le plus possible l'effet de l'anomalie de température. Les résultats sont mitigés. Inexistant pour Electricité de Strasbourg, le bêta semble positif, mais significatif seulement au niveau 10%, pour Euro-Disney et Teisseire-France, il est enfin positif de façon nette pour Pernod-Ricard ; dans tous les cas, les valeurs estimées sont de toute façon très inférieures aux bêtas « traditionnels » action/marché.

Finalement, les conséquences économiques de la température sur les ventes de ces entreprises se répercutent ou bien pas du tout, ou bien de façon très atténuée sur le cours des actions. L'entreprise Pernod-Ricard fait office d'exception : objectivement la moins exposée au risque climatique, son action est pourtant la plus (voire la seule) réactive (*cf.* § 3.b)). Peut-être faut-il songer à un argument d'ordre psychologique ou culturel, les agents économiques amplifiant la répercussion réelle des températures sur les ventes ?

### **III. Conclusions de la section B.**

Cette section était consacrée à l'étude de l'influence de la température sur le cours des actions d'entreprises sensibles au climat.

La première partie, de nature théorique, a mis en évidence la possibilité de répercussions significatives des bonnes saisons sur le cours des actions. L'importance de ces répercussions dépend de facteurs climatiques extérieurs à l'entreprise, mais aussi de facteurs économiques internes : on a ainsi montré que la variation de rentabilité attendue est proportionnelle au rapport chiffre d'affaires/capitalisation boursière ( $CA/CB$ ).

La seconde partie a consisté en une vaste étude empirique, réalisée à partir des actions des sociétés Electricité de Strasbourg, Eurodisney, Pernod-Ricard et Teisseire-France. A l'exception du cas de Pernod-Ricard, elle n'a pas permis de valider de façon nette une dépendance de nature économique du climat sur le cours des actions : les effets obtenus ont quasiment disparu après avoir tenu compte de l'impact général du climat sur le marché (qui

s'observe sur les indices d'actions, voir section A). En particulier, on est en droit de douter de la significativité du bêta action/température pour ces entreprises.

Ces résultats semblent en contradiction avec les conclusions de la partie théorique, d'autant plus que l'entreprise Pernod-Ricard est celle qui semble<sup>78</sup> avoir le rapport  $CA/CB$  le plus faible et une envergure internationale, qui la rendent a priori peu sensible aux aléas climatiques. Cependant, nous avons montré qu'une hypothèse importante pour pouvoir déceler l'impact théorique du climat tenait à la *fréquence* à laquelle les bons résultats sont intégrés dans le cours des actions : si la revalorisation de l'action a lieu d'un coup - après la saison favorable par exemple, celle-ci a toutes les chances d'être détectable ; en revanche une revalorisation mensuelle et régulière a pour effet de diluer l'effet des bénéfices de l'entreprise sur le cours de l'action. Les résultats obtenus semblent donc indiquer que les agents procèdent à un rééquilibrage progressif des bons résultats sur plusieurs mois.

---

<sup>78</sup> Cette affirmation est basée sur les chiffres de 2001. Elle mériterait d'être étendue à la période de l'étude.

## Conclusions du chapitre 2.

Dans ce chapitre, nous nous sommes intéressés au risque de température sous un angle financier, en étudiant la question de sa rémunération. Nous avons alors été amenés à rechercher d'éventuelles dépendances avec le risque de marché, et de façon plus spécifique, avec la rentabilité des actions d'entreprises exposées au climat.

Dans l'ensemble, nous n'observons pas de dépendance franche, même dans le cas des entreprises a priori plus exposées aux risques climatiques. De faibles dépendances existent, mais semblent uniquement dues au phénomène d'impact comportemental du climat qui avait été observé sur l'ensemble du marché. En particulier, elles ne semblent pas être de nature économique, ni spécifiques à certaines entreprises.

Ces résultats indiquent que le risque de température n'est pas ou peu rémunéré dans le cas des actions. Pour expliquer ce constat pour les entreprises des secteurs « à risque », on peut avancer l'hypothèse que les bénéfices exceptionnels dus aux climats sont intégrés de façon continue et régulière dans le cours des actions, de sorte que leur impact est difficilement décelable par des méthodes statistiques.

La situation est plus complexe pour les produits dérivés climatiques, car on ne dispose pas à l'heure actuelle de données de prix. Toutefois, en se plaçant dans le cadre du MEDAF monopériodique, et en supposant que la non-dépendance vérifiée pour les indices de température sous-jacents s'étend aux produits dérivés climatiques, on obtient que leur évaluation doit se faire sous la probabilité historique. C'est un argument en faveur de l'évaluation dans le cadre actuariel, ainsi qu'elle est couramment pratiquée.

## Références du chapitre 2

- Campbell J.Y., Lo A.W., MacKinlay A.C., *The Econometrics of Financial Markets*, Princeton University Press, 1997.
- Cao M., Wei J., *Stock Market Returns : A Temperature Anomaly*, working Paper, Queen's University Kingston and University of Toronto, Ontario, Canada, 2001.
- Copeland T.E., Weston J.F., *Financial Theory and Corporate Policy*, Addison Wesley, 1988.
- Garcia A.F., Sturzenegger F., *Hedging Corporate Revenues with Weather Derivatives : A Case Study*, Master of Science in Banking and Finance – MBF Master's Thesis, Université de Lausanne, HEC, Suisse, 2001.
- Hamon J., Jacquillat B., *Le marché français des actions - Etudes empiriques 1977-1991*, Presses Universitaires de France, 1992.
- Hollander M., Wolfe D.A., *Nonparametric Statistical Methods*, Wiley Series, 1999.
- Laurent J.P., Roustant O., *Weather derivatives and the stock market: a risk assessment*, working paper, [www.affi.asso.fr](http://www.affi.asso.fr), 2003.
- Lütkepohl H., *Introduction to Multiple Time Series Analysis*, Springer Verlag, 1993.
- Saunders E.M.J., *Stock Prices and Wall Street Weather*, American Economic Review, **83**, 1337-1345, 1993.
- Solen D., *Six actions à acheter maintenant pour miser sur « l'effet canicule »*, Capital, **129**, juin 2002.

## Annexe du chapitre 2

<b>ANNEXE.</b>				
<b>PERNOD RICARD - Compte de résultat consolidé</b>				
<i>En millions d'euros</i>	<b>2001</b>	<b>2000</b>	<b>1999</b>	<b>2001 / 2000 %</b>
Chiffre d'affaires hors droits et taxes (H.D.T.)	4 555,2	4 382,0	3 590,3	4,0%
Achats consommés	(2 412,2)	(2 293,3)	(1 792,8)	5,2%
<b>Marge brute</b>	<b>2 143,0</b>	<b>2 088,7</b>	<b>1 797,5</b>	<b>2,6%</b>
Frais de commercialisation et de distribution	(1 031,3)	(996,9)	(863,8)	3,5%
Frais de production et frais généraux	(661,2)	(670,7)	(561,4)	-1,4%
<b>Résultat opérationnel</b>	<b>450,5</b>	<b>421,1</b>	<b>372,3</b>	<b>7,0%</b>
Frais financiers nets	(39,5)	(52,2)	(32,1)	-24,3%
<b>Résultat courant</b>	<b>411,0</b>	<b>368,9</b>	<b>340,2</b>	<b>11,4%</b>
Résultat exceptionnel	220,0	(53,4)	(8,2)	NS
Impôts sur les sociétés	(235,7)	(97,2)	(92,8)	142,6%
Résultat des sociétés mises en équivalence	0,7	2,1	2,3	-66,9%
<b>Résultat net avant amortissement des survaleurs</b>	<b>396,0</b>	<b>220,4</b>	<b>241,5</b>	<b>79,7%</b>
Amortissement des survaleurs	(31,4)	(20,4)	(21,1)	53,8%
<b>Résultat net</b>	<b>364,6</b>	<b>200,0</b>	<b>220,3</b>	<b>82,3%</b>
Droit des tiers dans le résultat	6,4	5,0	7,1	29,2%
<b>Droit du Groupe dans le résultat</b>	<b>358,2</b>	<b>195,0</b>	<b>213,2</b>	<b>83,7%</b>
<b>Résultat par action*</b>	<b>Euros</b>	<b>Euros</b>	<b>Euros</b>	<b>2001 / 2000</b>
Résultat courant	7,29	6,54	6,03	11,4%
Résultat net part du Groupe	6,35	3,46	3,78	83,7%
<b>Résultat net avant impôt</b>	<b>Euros</b>	<b>Euros</b>	<b>Euros</b>	<b>2001 / 2000</b>
Global (en millions)	600,3	297,1	313,1	102,1%
Par action	10,65	5,27	5,55	102,0%
*Nombre d'actions en 1999, 2000 et 2001 : 56 386 660				

Tableau 2.17. Capitalisation boursière et chiffre d'affaires des quatre entreprises sélectionnées.

Données financières au 31/12/01*	Electricité de Strasbourg	Eurodisney	Pernod-Ricard	Teisseire-France
Nombre de titres (millions)	7.1	1 055.8	56.4	1.2
Valeur de l'action (euros)	33.10	0.83	87	21.63
<i>CB</i> (millions d'euros)	234	876	4 906	26
<i>CA</i> (millions d'euros)	455	1 005	4 555	147
<b><i>CA/CB</i> au 31/12/01*</b>				
<i>CA/CB</i>	1.95	1.15	0.93	5.67

\* ou le cas échéant au dernier jour boursier avant cette date, et à l'exception de Eurodisney, pour lequel la date de référence est le 30/09/01.

**Chapitre 3. Incertitudes de modélisation.  
Répercussions sur les prix.**



Dans ce chapitre, nous étudions les modalités pratiques de l'évaluation des produits dérivés climatiques. Le schéma classique : « modélisation du sous-jacent + formule d'évaluation » s'entoure de nombreuses incertitudes que nous voulons quantifier. A l'heure actuelle cependant, aucun consensus ne s'est dégagé au sujet de la technique d'évaluation à employer ; nous en avons choisi une et concentré notre étude sur l'impact des erreurs de la modélisation de la température sur les prix des produits dérivés climatiques.

Notre choix de cadre d'évaluation s'est porté sur l'approche actuarielle dans laquelle les produits sont évalués par une formule du type :

$$P = E[\textit{payoff actualisé}] + \lambda \cdot \sigma[\textit{payoff actualisé}] \quad (*)$$

où les espérances sont calculées dans l'univers historique, et  $\lambda$  est un facteur lié au risque. C'est l'approche utilisée par les praticiens (voir Geman, 2001). Dans le chapitre 2 de ce travail, nous avons donné dans le cadre du MEDAF des arguments en faveur d'une évaluation linéaire dans l'univers historique : cela revient à choisir  $\lambda$  nul dans la formule (\*) ; de ce point de vue, on peut voir  $\lambda$  comme un facteur englobant les risques non pris en compte par le MEDAF. Outre son emploi courant en assurance, l'approche actuarielle peut être vue comme une approximation d'approches plus complexes proposées pour les produits dérivés climatiques (nous renvoyons le lecteur au paragraphe « La pratique de l'évaluation » de l'introduction).

Dans ce cadre, le modèle de température développé au chapitre 1 donne des prix pour les produits dérivés climatiques. Cependant, l'adéquation de ce modèle linéaire n'est pas parfaite : on se souvient des écarts à l'indépendance et à la normalité. D'autres erreurs proviennent simplement de l'estimation du modèle. Comment ces incertitudes se traduisent-elles sur les prix ? Nous répondons à cette question selon l'organisation suivante. Après un paragraphe consacré au calcul des prix, nous détaillons la méthodologie d'évaluation de l'incertitude des prix. Ensuite, nous quantifions l'impact des erreurs de modélisation sur un cas particulier. Enfin, nous étudions les conséquences et les origines de l'erreur d'estimation sur l'ensemble des produits ; nous en déduisons les principaux défauts du modèle de température relativement à son utilisation pour l'évaluation des produits dérivés climatiques.

## A. Calcul des prix.

Avant d'étudier l'incertitude des prix, il est utile de connaître les formules qui les donnent et les mécanismes qui permettent de les calculer. Les produits dérivés de température d'étude sont des Futures (ou swaps) et des options sur un indice de température, noté  $TI$ . Nous nous limitons ici à l'un des deux indices standard :

$$HDD = \sum_{t=u}^v (X_{\text{ref}} - X_t)^+ \quad CDD = \sum_{t=u}^v (X_t - X_{\text{ref}})^+$$

$[u; v]$  est la période d'exposition au risque,  $X_t$  la température journalière moyenne et  $X_{\text{ref}}$  est une température de référence (rappelons que la valeur standard est 18°C). Donnons d'abord les formules d'évaluation.

### I. Formules d'évaluation.

Dans le cadre actuariel retenu ici, la formule générale d'évaluation est (voir l'introduction ci-avant) :

$$P = E[\textit{payoff actualisé}] + \lambda \cdot \sigma[\textit{payoff actualisé}]$$

où le calcul s'effectue dans l'univers historique. Pour l'objectif poursuivi – l'évaluation des incertitudes – il n'y a pas de réel inconvénient à « oublier » l'actualisation : même dans le cas d'un taux d'intérêt stochastique, celui-ci est – au moins en première approximation, indépendant de la température et donc du *payoff* des produits dérivés climatiques, et l'actualisation se fait en divisant par un facteur constant (indépendant du modèle de température). Avec cette simplification, les formules d'évaluation sont les suivantes.

#### 1. Prix des Futures.

Le prix des Futures est calculé de sorte qu'aucun versement initial n'est à effectuer. Dans notre cadre, on peut considérer que le *payoff* est la température elle-même puisque :

$$\begin{aligned} 0 &= E[TI - F] + \lambda \cdot \sigma[TI - F] \\ \Leftrightarrow F &= E[TI] + \lambda \cdot \sigma[TI] \end{aligned}$$

( $F$  désigne le prix Future). Dans la suite, nous adoptons ce point de vue.

#### 2. Prix des options.

Notons  $K$  le prix d'exercice. Le prix d'une option d'achat est donc :

$$C = E \left[ (TI - K)^+ \right] + \lambda \cdot \sigma \left[ (TI - K)^+ \right]$$

et celui d'une option de vente :

$$P = E \left[ (K - TI)^+ \right] + \lambda \cdot \sigma \left[ (K - TI)^+ \right]$$

## II. Méthodes de calcul.

Chacune des expressions précédentes peut se calculer numériquement sans connaître la loi exacte de l'indice de température  $TI$  grâce aux techniques de bootstrap (voir Davinson, Hinkley, 1997). Cette méthode est utilisée pour évaluer les incertitudes des prix dues aux écarts de normalité et d'indépendance du modèle de température. Cependant son coût élevé en temps de calcul pose un véritable problème lorsqu'on doit y faire appel un grand nombre de fois. Pour cette raison, elle s'avèrera inappropriée pour évaluer l'impact de l'erreur d'estimation (on aurait dans ce cas deux niveaux de simulation d'où la multiplication du temps de calcul, voir paragraphe suivant).

Cette difficulté disparaît lorsqu'on dispose de formules fermées pour les prix des Futures et des options. Or il est possible d'obtenir des expressions analytiques lorsque la température est un processus gaussien, avec les restrictions suivantes :

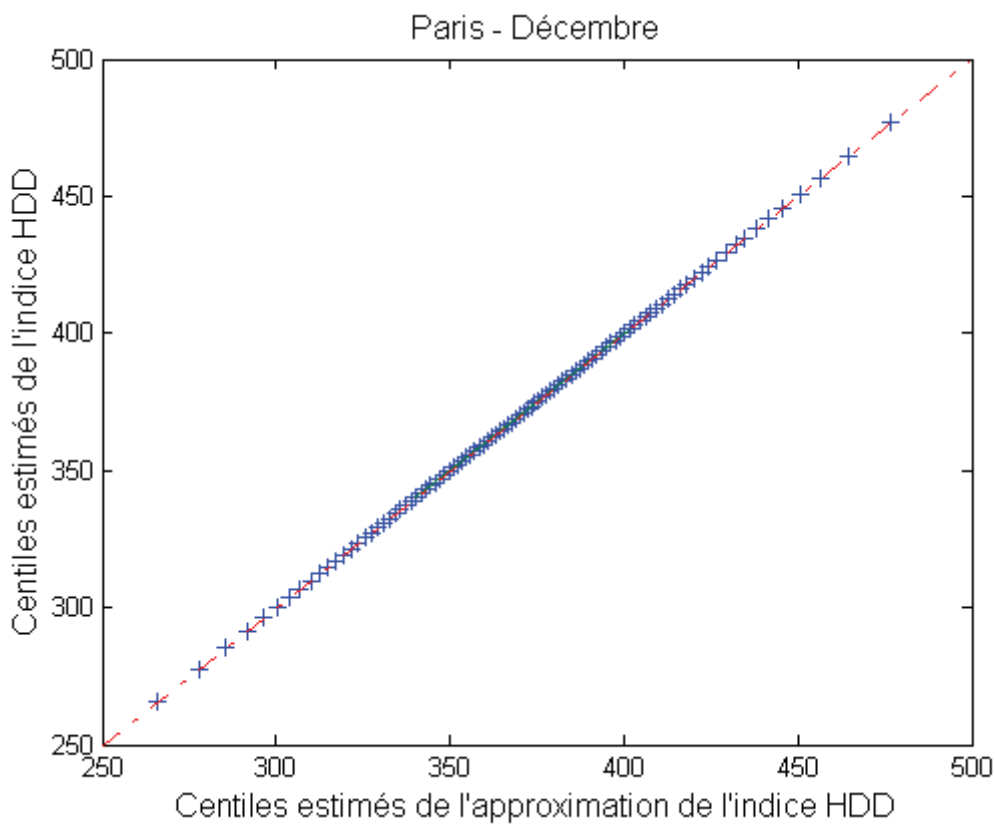
- *Prime pure* ( $E[TI]$ ) pour un contrat *Future* ; le calcul se fait alors au moyen du lemme présenté dans l'annexe A.1.
- *Période d'exposition au risque contenue dans l'hiver ou dans l'été*. Dans ce cas, on peut remarquer que les approximations suivantes sont valides :

$$HDD \approx \sum_{t=u}^v (X_{ref} - X_t) \qquad CDD \approx \sum_{t=u}^v (X_t - X_{ref})$$

( $X_{ref}$  désigne ici la valeur standard des indices HDD et CDD : 18°C). Par conséquent, les indices ci-dessus sont une fonction linéaire de la température et suivent une loi normale. Des formules fermées s'en déduisent immédiatement grâce au même lemme (voir annexe A.1.). Les simplifications précédentes sont naturelles car en hiver (en été) les températures sont au-dessous (au-dessus) de 18°C (nous parlons ici des stations de Paris et Chicago sur lesquelles nous avons réalisé l'étude qui suit). Nous nous sommes assurés de leur validité en comparant la distribution de chacun des indices avec celle de leur approximation. Nous avons donc simulé un grand nombre de trajectoires de températures et représenté les quantiles estimés

contre leur approximation. Un exemple est montré sur la Figure 3.1, correspondant à la station de Paris-Montsouris et au mois de décembre ; l’alignement parfait que l’on observe demeure lorsqu’un choc est effectué sur les paramètres du modèle de température.

Le calcul pratique des prix a guidé l’organisation du chapitre : la section C est restreinte à l’étude de la prime pure d’un contrat Future mais s’applique à tous les mois de l’année (ainsi qu’à des indices HDD et CDD non standard) ; la section D s’applique à tous les produits dérivés climatiques mais se limite aux mois d’hiver et d’été.



**Figure 3.1. Validité de l’approximation linéaire pour les indices de températures.**

*QQ-plot des centiles estimés de l’indice HDD standard de décembre pour la station de Paris-Montsouris contre ceux de son approximation linéaire, obtenue avec  $10^6$  simulations.*

## B. Méthodologie.

Nous présentons maintenant la méthodologie d'évaluation des incertitudes des prix dues à l'erreur d'estimation (les conséquences des autres erreurs de modèle s'obtiennent par bootstrap ; les techniques correspondantes sont détaillées au cours de leur utilisation (§ C.III.1 et § C.III.2). Il s'agit d'une adaptation de la démarche suivie – par exemple - pour quantifier la précision des estimateurs du prix d'options classiques par la formule de Black et Scholes. Après l'avoir rappelée, nous évaluons l'erreur d'estimation et discutons de deux voies permettant de déduire l'incertitude des prix.

■ **Un exemple.** (Campbell, Lo, Mac Kinlay, 1997, § 9.3.3.).

Nous sortons ici momentanément du cadre des produits dérivés de température.

Considérons une option d'achat sur une action, et plaçons-nous dans le cadre d'application du modèle de Black et Scholes où l'actif sous-jacent suit un mouvement brownien géométrique :

$$dS_t = \mu S_t dt + \sigma S_t dB_t$$

La valeur de l'option est alors une fonction connue, notons-là  $G$ , de la volatilité  $\sigma$ . L'incertitude du prix de l'option due à l'erreur d'estimation de la volatilité se fait en deux étapes :

- l'incertitude de la volatilité est d'abord évaluée au moyen de la distribution asymptotique :

$$\sqrt{n}(\hat{\sigma}^2 - \sigma^2) \sim N(0, V(\sigma)) \quad \text{avec} \quad V(\sigma) = 2\sigma^4.$$

- l'incertitude du prix se déduit ensuite par un développement de Taylor à l'ordre 1 (*méthode delta*) :

$$\sqrt{n}(G(\hat{\sigma}) - G(\sigma)) \sim N(0, V_G(\sigma)) \quad \text{avec} \quad V_G \equiv \frac{\partial G'}{\partial \sigma} \cdot V \cdot \frac{\partial G}{\partial \sigma}.$$

Pour une option d'achat, on obtient ainsi  $V_G(\sigma) = \frac{T}{2} S^2(0) \sigma^2 \phi^2(d_1)$ , avec

$$d_1 = \frac{\ln(S_0 / K) + (r + \sigma^2 / 2)T}{\sigma \sqrt{T}}, \quad \text{où } K \text{ est le prix d'exercice, } T \text{ la maturité et } r \text{ le taux d'intérêt}$$

sans risque. ■

Nous allons adapter cette démarche aux produits dérivés de température. Dans cette situation, le modèle de température remplace le modèle de Black et Scholes, et un vecteur d'une dizaine

de paramètres  $\Theta$  remplace la volatilité  $\sigma$  ; en outre, il sera souvent possible d'obtenir une expression explicite du prix des produits (voir § A.II.). Dans un premier temps, comme dans l'exemple, nous obtenons un résultat asymptotique pour  $\hat{\Theta}$ , l'estimateur du maximum de vraisemblance des paramètres ; cependant il s'agit cette fois d'un résultat approché et la variance asymptotique doit être calculée numériquement. Nous discutons alors de deux méthodologies pour l'évaluation de l'incertitude des prix  $P(\Theta)$ , l'une basée sur la méthode delta (comme ci-dessus) et l'autre sur l'évaluation de la distribution exacte de  $P(\hat{\Theta})$  par bootstrap.

### **I. Evaluation de l'erreur d'estimation.**

Le modèle de température est *grosso modo* un modèle ARMA augmenté de termes déterministes de tendance, de saisonnalité, et de dispersion. Pour un modèle ARMA « simple », nous avons le résultat suivant :

**Théorème.** (Gouriéroux et Montfort, 1995, § 9.2.F.). *Soit  $(Y_t)$  un processus ARMA,  $\Theta$  le vecteur des paramètres,  $\hat{\Theta}$  l'estimateur du maximum de vraisemblance de  $\Theta$ . Alors  $\hat{\Theta}$  est asymptotiquement gaussien :*

$$\sqrt{n}(\hat{\Theta} - \Theta) \xrightarrow{n \rightarrow +\infty} N(0; \Gamma(\Theta))$$

Nous faisons l'hypothèse que ce résultat s'applique pour le modèle de température. Ceci n'est pas déraisonnable car les termes additionnels  $m_t$ ,  $s_t$  et  $\sigma_t$  de tendance, de saisonnalité et de volatilité sont tous déterministes ; cependant il convient d'aller plus loin que cette intuition. Enfin, une deuxième difficulté vient du fait que  $\Gamma(\Theta)$  n'est pas donné explicitement. Pour apporter une réponse à ces questions, nous utilisons une technique de bootstrap, présentée en annexe (A.3.), qui donne des réalisations indépendantes de  $\hat{\Theta}$ . Nous pouvons alors en déduire une approximation de la matrice de covariance  $\Gamma(\Theta)$  (les résultats pour les stations de Paris et Chicago figurent dans le Tableau 3.12, page 226 et le Tableau 3.13, page 227) et effectuer des tests statistiques pour vérifier que cet échantillon provient d'une loi normale : le nombre élevé -  $n = 7665$ , de données de température laisse à penser que la distribution exacte de  $\hat{\Theta}$  n'est pas trop éloigné de la distribution asymptotique (normale). Les tests effectués sont de deux types : Kolmogorov pour les lois marginales (voir par exemple Saporta, 1990) et le test basé

sur *skewness* et *kurtosis* pour la loi multidimensionnelle (déjà utilisé au cours du chapitre 2, voir Lütkepohl, 1993, § 4.5.1.). Les résultats obtenus pour les stations de Chicago et Paris (Tableau 3.14, page 228) sont satisfaisants.

## II. Evaluation de l'incertitude des prix.

Nous présentons deux méthodologies différentes, donnant des intervalles de confiance approximatifs des prix ; lorsque le *payoff* est linéaire, les deux approches coïncident.

### 1. « Méthodologie asymptotique ».

A partir de la distribution asymptotique de  $\Theta$ , on peut obtenir une statistique asymptotiquement pivotale de  $P(\Theta)$  en utilisant la méthode delta (exemple ci-dessus ; voir aussi Davinson, Hinkley, 1997). On a alors :

$$\sqrt{n} \left( P(\hat{\Theta}) - P(\Theta) \right) \xrightarrow{n \rightarrow +\infty} N(0; v_p(\Theta))$$

avec  $v_p \equiv \frac{\partial P}{\partial \Theta'} \Gamma \frac{\partial P}{\partial \Theta}$ .

Supposons que l'on sache calculer (numériquement) l'expression ci-dessus, alors on peut estimer  $v_p(\Theta)$  de façon naturelle en remplaçant  $\Theta$  par son estimation, et en déduire un intervalle de confiance pour  $P(\Theta)$ .

L'inconvénient majeur de cette méthodologie est qu'elle repose sur une approximation de la distribution exacte du prix dont on connaît mal la précision. Il est donc utile d'estimer la distribution exacte et de préciser les situations où les deux distributions coïncident.

### 2. « Méthodologie bootstrap ».

La distribution exacte de  $P(\hat{\Theta})$  peut être approchée au moyen d'une technique de bootstrap :

- simuler de façon indépendante  $\hat{\Theta}^1, \dots, \hat{\Theta}^R$  selon la distribution normale de matrice de covariance  $\Gamma(\Theta)$  précédemment estimée ;
- calculer  $P(\hat{\Theta}^1), \dots, P(\hat{\Theta}^R)$ .

L'incertitude du prix est évaluée par deux quantiles de  $P(\hat{\Theta})$  estimés avec l'échantillon "bootstrap"  $P(\hat{\Theta}^1), \dots, P(\hat{\Theta}^R)$  ainsi obtenu.

### **3. Discussion.**

Chacune des méthodologies donne des intervalles de confiance approchés. Cependant, l'incertitude évaluée par bootstrap est sans doute plus réaliste lorsque la distribution exacte des prix est éloignée de la distribution asymptotique. En fait, même dans le cas où les deux sont proches – autrement dit lorsque la distribution exacte est approximativement normale, la méthodologie bootstrap reste efficace : en effet, le résultat asymptotique est alors exact si bien que les deux méthodologies donnent les mêmes incertitudes. En outre, l'expression analytique de la variance asymptotique n'est pas immédiate à obtenir alors qu'elle peut facilement être estimée par la variance de l'échantillon bootstrap. Ainsi, la méthodologie bootstrap est appropriée dans tous les cas. Pour finir, remarquons que dans le cas gaussien, les intervalles de confiance sont exacts. Ceci arrive, par exemple, lorsque le *payoff* est linéaire (dans l'hypothèse où le processus de température est gaussien) ; comme on l'a vu précédemment (§ A.II.), cela correspond au cas de la prime pure des contrats Futures de température pour les mois d'hiver ou d'été.



## C. Impact des erreurs de modèle sur la prime pure d'un contrat Future.

Dans cette section, nous considérons le cas particulier des contrats Futures de température et examinons l'incertitude autour de la prime pure. Pour des raisons pratiques précisées auparavant (§ A.II), cette restriction nous permet d'élargir le champ de l'étude à toutes les périodes d'exposition au risque et à des indices de température non standard. Nous étudions alors l'impact de deux types d'erreur de modèle : l'erreur d'estimation et l'erreur de spécification du bruit résiduel de modèle ; cette dernière est décomposée en deux types : mauvaise spécification de la distribution (défauts dans la normalité) ou de la dynamique (défauts dans l'indépendance).

### I. Cadre de travail.

Les contrats Futures considérés sont construits sur les indices de température les plus courants (à la date de l'étude) nommément les indices HDD et CDD :

$$HDD \approx \sum_{t=u}^v (X_{ref} - X_t) \quad CDD \approx \sum_{t=u}^v (X_t - X_{ref})$$

où  $X_{ref}$  est un seuil et  $[u;v]$  la période d'exposition au risque. Nous parlerons d'indices  $X_{ref}$ -HDD et  $X_{ref}$ -CDD, et des indices HDD et CDD, simplement, pour la valeur standard  $X_{ref} = 18^\circ C$ . Rappelons que la saison HDD s'étale généralement d'octobre à avril et la saison CDD de mai à septembre.

De façon à faire une étude à la fois fine et complète, nous avons retenu des contrats mensuels (comme c'est le cas, du reste, au Chicago Mercantile Exchange) et utilisé un ou plusieurs seuils d'indice « extrêmes » en plus de la valeur standard. Deux stations ont servi de référence : l'une en France à Paris (Paris-Montsouris), l'autre aux Etats-Unis à Chicago (O'Hare Airport). Les spécifications sont résumées dans le Tableau 3.1.

Les données de température sont celles du chapitre 1, et sont relatives à la période du 1<sup>er</sup> janvier 1979 au 31 décembre 1999. La date courante est choisie au moins un mois avant le mois d'exposition défini dans le contrat, de sorte que la situation courante n'intervient pas dans l'évaluation (voir la section A du chapitre 1).

**Tableau 3.1. Spécifications d'indices.**

Station	Paris	Chicago
<b>Indice HDD</b>		
<i>Mois d'exposition au risque</i>	Octobre à Avril	Octobre à Avril
<i>Seuil d'indice</i>	18°C (standard); 0°C	65°F (stand.); 20°F; 10°F
<b>Indice CDD</b>		
<i>Mois d'exposition au risque</i>	Mai à Septembre	Mai à Septembre
<i>Seuil d'indice</i>	18°C (standard); 26°C	65°F (standard); 80°F

## II. Evaluation de l'impact de l'erreur d'estimation.

Dans ce paragraphe, le modèle de température est supposé valide, et la seule source d'incertitude provient de l'erreur d'estimation des paramètres. La variabilité de la prime pure des contrats Futures qui en résulte est mesurée en suivant la méthodologie détaillée à la section B précédente.

Nous donnons d'abord les résultats relatifs aux contrats Futures standard (voir le précédent paragraphe) dans le Tableau 3.2; ils concernent la station sise à Chicago mais les résultats sont similaires à Paris. Nous avons indiqué la prime pure théorique obtenue avec les paramètres estimés, les deux premiers moments, l'incertitude sous forme de deux quantiles et l'incertitude relative correspondante par rapport à la prime pure moyenne. (En fait, sauf pour les mois de Mai et Septembre, la distribution estimée peut être considérée comme normale – ainsi qu'en attestent les résultats du test de Kolmogorov, et nous aurions pu évaluer l'incertitude par les deux premiers moments).

Nous observons une différence nette entre le cas des Futures HDD et CDD : si pour les contrats HDD l'incertitude ne dépasse guère 5%, elle s'étale de 10 à 20% pour les contrats CDD. Cette dissymétrie n'est pas très surprenante puisque le seuil de 18°C ne joue pas le même rôle en hiver ou en été : alors qu'en hiver presque toutes les températures sont en dessous de 18°C, elles sont de part et d'autre de cette valeur en été.

**Tableau 3.2. Impact de l'erreur d'estimation sur la prime pure des contrats Futures standard.**Station : *O'Hare Airport*

Prime pure :	Valeur théorique	Moyenne	Ecart-type	Incertitude	Incertitude relative (%)
<i>Saison HDD</i>					
Octobre	397.87	398.04	15.38	[368.0 ; 428.1]	± 7.5
Novembre	752.31	752.71	17.19	[719.1; 768.3]	± 4.4
Décembre	1079.00	1079.08	18.48	[1042.9; 1115.3]	± 3.4
Janvier	1215.41	1216.05	17.96	[1180.04; 1251.4]	± 2.9
Février	1036.39	1036.76	16.42	[1004.3; 1069.2]	± 3.1
Mars	901.86	902.12	17.94	[866.9; 938.1]	± 4.0
Avril	518.69	518.70	16.09	[487.1; 550.1]	[-6.1; 6.0]
<i>Saison CDD</i>					
Mai	53.56	53.70	5.34	[43.4; 64.7]	[-19.0; 20.4]
Juin	195.22	195.50	11.94	[172.5; 219.5]	[-11.7; 12.3]
Juillet	314.90	315.31	14.51	[287.1; 343.9]	[-8.9; 9.1]
Août	256.13	256.09	13.68	[229.2; 283.9]	[-10.4; 10.9]
Septembre	90.77	90.73	7.70	[76.2; 106.5]	[-16.0; 17.4]

**Tableau 3.3. Impact de l'erreur d'estimation sur la prime pure des contrats Futures non standard.**Station : *Paris-Montsouris*

Contrat Future	HDD Décembre		CDD Août	
	Standard	0-HDD	Standard	26-CDD
Indice de température (°C)				
<u>Prime pure :</u>				
Valeur théorique	369.97	1.74	93.52	1.97
Moyenne	370.37	1.78	93.54	1.98
Ecart-type	9.55	0.42	7.23	0.48
Incertitude	[351.1; 389.1]	[1.06 ; 2.72]	[79.3; 107.8]	[1.17; 3.05]
Incertitude relative (%)	± 5.1	[-38.8; 56.4]	± 15.1	[-39.2; 57.3]

Station : *O'Hare Airport*

Contrat Future	HDD Décembre			CDD Août	
	Standard	20-HDD	10-HDD	Standard	80-CDD
Indice de température (°F)					
<u>Prime pure :</u>					
Valeur théorique	1079.00	30.53	3.94	256.14	17.51
Moyenne	1079.08	30.68	3.97	256.48	17.67
Ecart-type	18.48	3.44	0.67	13.64	2.62
Incertitude	[1042.9; 1115.3]	[24.3; 37.9]	[2.79; 5.40]	[229.7; 283.3]	[12.9; 23.3]
Incertitude relative (%)	± 3.4	[-20.1; 24.1]	[-28.8; 37.6]	± 10.5	[-25.8; 32.7]

Nous présentons ensuite les résultats généraux dans le Tableau 3.3. Nous retrouvons la même dissymétrie entre les contrats HDD et CDD. Par ailleurs, on peut remarquer que les estimations les plus précises sont relatives aux contrats standard, l'incertitude de la prime pure des Futures non standard pouvant être énorme. D'une façon générale, on observe sur les tableaux précédents que la qualité de l'estimation décroît avec la valeur de la prime pure.

### **III. Evaluation de l'impact de mauvaises spécifications du modèle.**

Reprenons le processus de température ( $X_t$ ) :

$$X_t = m_t + s_t + \sigma_t Z_t$$

avec  $Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \phi_3 Z_{t-3} + \varepsilon_t$ , où l'on suppose (H) : « $(\varepsilon_t)$  est un bruit blanc gaussien».

Nous considérons maintenant que les paramètres estimés sont les vrais paramètres et évaluons les répercussions sur les prix des écarts à l'hypothèse (H). Nous nous intéressons d'abord aux écarts à la normalité puis aux écarts à l'indépendance.

#### **1. Mauvaise spécification de la loi marginale.**

Au cours du chapitre 1, nous avons observé que les résidus présentent des anomalies dans les queues de distribution par rapport à une loi normale. Pour évaluer leurs conséquences sur les prix, nous considérons deux distributions alternatives pour les résidus : la loi normale et la loi empirique, comparons les lois des indices de température correspondants, puis les primes pures des contrats Futures qui s'en déduisent.

##### a) Comparaison des indices de température.

Comme les primes pures des produits dérivés climatiques étudiés dépendent de la température via des indices, la première étape pour atteindre l'objectif fixé est de comparer les lois de ces indices, avec ou sans l'hypothèse de normalité sur le processus résiduel. Nous avons donc deux modèles : le premier est celui que nous avons considéré jusqu'à présent, tandis que dans

le second, nous utilisons la distribution empirique  $\hat{\mu}_n = \sum_{i=1}^n \frac{1}{n} \delta_{\varepsilon_i}$  où  $n$  est le nombre de

données (on commet une erreur statistique en remplaçant la distribution exacte  $\mu$  par la loi empirique  $\hat{\mu}_n$  ; cependant,  $\hat{\mu}_n$  converge en loi vers  $\mu$  et, comme la valeur de  $n$  est « grande » - égale à 7665, on peut espérer que cette erreur soit faible). Dans ce modèle, les indices sont obtenus avec le processus ( $\hat{X}_t$ ) :

$$\hat{X}_t = m_t + s_t + \sigma_t \cdot \hat{Z}_t$$

où  $\hat{Z}_t = \phi_1 \hat{Z}_{t-1} + \phi_2 \hat{Z}_{t-2} + \phi_3 \hat{Z}_{t-3} + \hat{\varepsilon}_t$ , et les  $\hat{\varepsilon}_k$  sont des variables aléatoires indépendantes, de même loi  $\hat{\mu}_n$ . Considérons le cas des indices HDD. Nous voulons donc comparer les lois de

$$\sum_{t=u}^v (X_{\text{ref}} - X_t)^+ \quad \text{et} \quad \sum_{t=u}^v (X_{\text{ref}} - \hat{X}_t)^+ .$$

Comme elles sont toutes deux inconnues (les densités théoriques sont difficiles à obtenir), on doit se contenter de comparer les distributions approchées. Une façon de procéder est d'examiner les quantiles de deux grands échantillons provenant de chacune des distributions.

Les échantillons sont obtenus avec des simulations de Monte Carlo. Notons  $t_0$  la date

courante. On obtient un échantillon de  $\sum_{t=u}^v (X_{\text{ref}} - X_t)^+$  en effectuant les opérations

suivantes ; pour  $r = 1, \dots, R$  :

- simuler  $\varepsilon_{t_0+1}^{*r}, \dots, \varepsilon_u^{*r}, \dots, \varepsilon_v^{*r}$  selon la loi normale  $\hat{\mu}_{norm}$  ajustée aux résidus estimés ;
- calculer la trajectoire  $X_u^{*r}, \dots, X_v^{*r}$  avec le modèle de température, puis

$$\sum_{t=u}^v (X_{\text{ref}} - X_t^{*r})^+ .$$

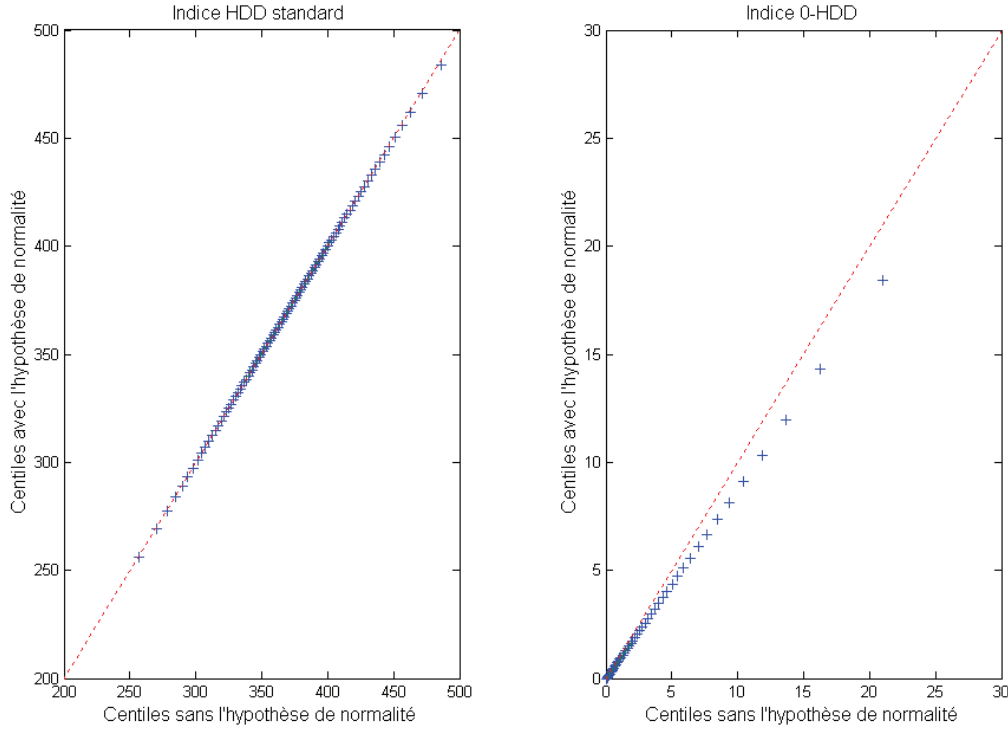
Les mêmes opérations réalisées en remplaçant  $\hat{\mu}_{norm}$  par  $\hat{\mu}_n$  permettent d'obtenir un

échantillon de  $\sum_{t=u}^v (X_{\text{ref}} - \hat{X}_t)^+ .$

La taille  $R$  doit être grande afin de limiter l'erreur d'estimation des quantiles ; nous avons choisi  $R = 10^6$  : pour cette valeur, l'erreur de simulation de la moyenne de la distribution peut être négligée, comme nous le verrons au paragraphe suivant.

Nous montrons les résultats obtenus dans le cas de Paris-Montsouris et le mois de décembre avec deux indices de température (Figure 3.2.):

- l'indice HDD standard, correspondant à  $K = 18^\circ\text{C}$ ;
- l'indice 0-HDD, correspondant à  $K = 0^\circ\text{C}$ .



**Figure 3.2. Comparaison des distributions d'indices de température. Cas standard et non standard.**

*Centiles de l'indice HDD standard (resp. 0-HDD) de décembre pour Paris-Montsouris obtenus en supposant que la distribution marginale des résidus est normale, contre ceux obtenus sans cette hypothèse (10<sup>6</sup> simulations).*

Dans chaque cas, nous avons représenté les quantiles de  $\sum_{t=u}^v (X_{\text{ref}} - X_t)^+$  contre ceux de

$\sum_{t=u}^v (X_{\text{ref}} - \hat{X}_t)^+$  ainsi que la droite  $y = x$  où les points devraient se trouver si les

distributions étaient égales. Dans le cas standard, on n'observe pas de différence nette ; dans

le second en revanche, la queue de distribution de  $\sum_{t=u}^v (X_{\text{ref}} - \hat{X}_t)^+$  est plus épaisse que celle

de  $\sum_{t=u}^v (X_{\text{ref}} - X_t)^+$ . Une explication peut être avancée en examinant les queues de

distribution des résidus qui, pour la station de Paris-Montsouris, sont épaisses pour les basses températures, et fines pour les hautes températures. En simulant avec la distribution empirique, davantage de températures vont se trouver en dessous de la normale saisonnière qu'en utilisant la loi normale ajustée. Par conséquent une fonction décroissante de  $X_t$  - et

donc un indice HDD (somme sur  $t$  de fonctions décroissantes de  $X_t$ ), devrait donner une valeur plus élevée sous la loi empirique. En outre, ce phénomène devrait être amplifié lorsque cette fonction n'implique que des températures extrêmes comme l'indice 0-HDD. Bien entendu, l'inverse devrait être vrai : un indice CDD (par exemple) devrait exhiber une queue de distribution plus fine lorsqu'on rejette la normalité. C'est en effet ce que l'on observe.

b) Comparaison des primes pures.

Nous comparons maintenant les primes pures – avec ou sans hypothèse de normalité, afin d'évaluer l'incertitude autour des prix due à une mauvaise spécification du processus de température sous-jacent.

Rappelons que les primes pures  $P(\Theta)$  sont données par des expressions de la forme

$$E \left[ \sum_{t=u}^v (X_{\text{ref}} - X_t)^+ \right] \text{ pour les contrats HDD et } E \left[ \sum_{t=u}^v (X_t - X_{\text{ref}})^+ \right] \text{ pour les contrats CDD.}$$

Avec l'hypothèse de normalité,  $P$  peut se calculer analytiquement (voir annexe A.1.). Ce n'est plus le cas lorsqu'on rejette la normalité et le calcul de la prime pure correspondante  $\hat{P}$  doit se faire numériquement. Par exemple, considérons le cas HDD. Nous avons :

$$\hat{P} = E_{\hat{\mu}_n} \left[ \sum_{t=u}^v (X_{\text{ref}} - \hat{X}_t)^+ \right]$$

En utilisant les indices  $HDD_r := \sum_{t=u}^v (X_{\text{ref}} - X_t^{*r})^+$ ,  $r = 1, \dots, R$ , précédemment obtenus par

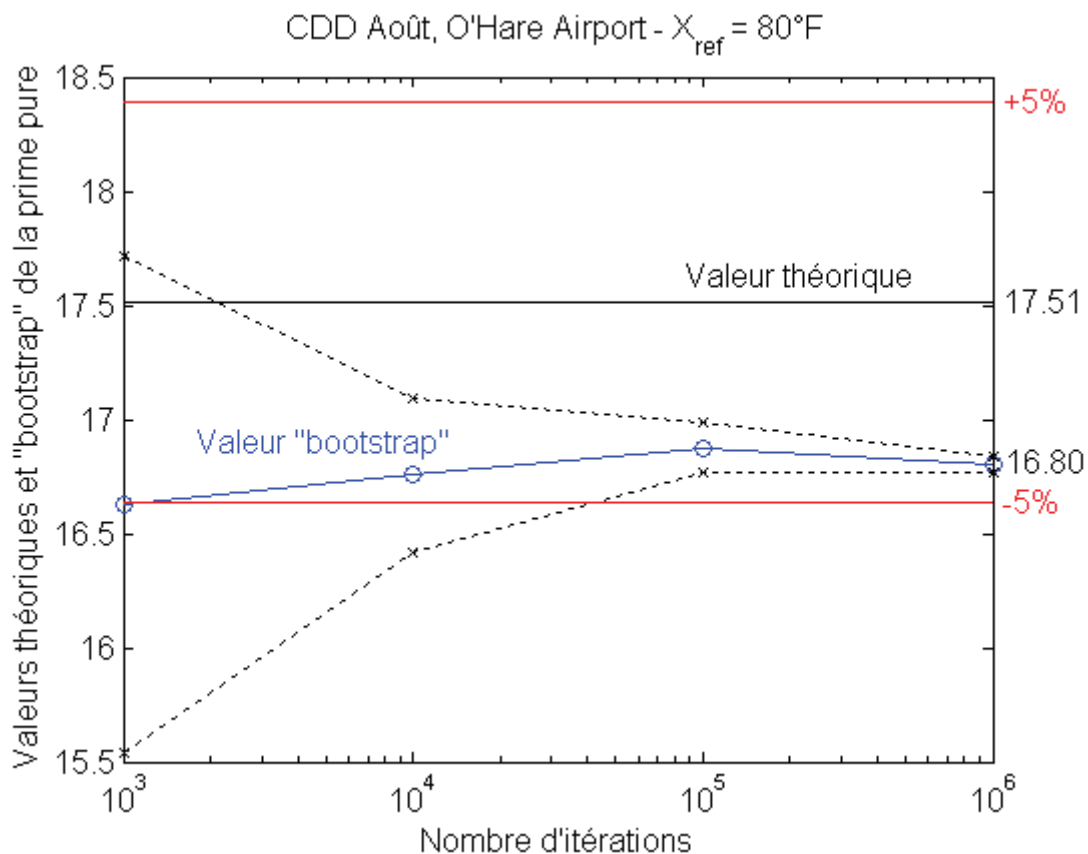
bootstrap à partir de la distribution empirique  $\hat{\mu}_n$  (voir le paragraphe précédent), on peut approcher  $\hat{P}$  par la moyenne de l'échantillon :

$$\hat{P}_R = \frac{1}{R} \sum_{r=1}^R HDD_r$$

Cette méthode est une technique de bootstrap et, pour cette raison, nous parlerons pour  $\hat{P}_R$  de « prime pure bootstrap » (voir Davinson, Hinkley, 1997 ; notons que le bootstrap s'applique aux données « éclaircies », c'est-à-dire les plus indépendantes possibles, comme cela a été fait ici).

L'approximation de  $\hat{P}$  par  $\hat{P}_R$  induit une erreur de simulation qui peut se révéler problématique. A titre d'illustration, considérons le contrat Future sur l'indice 80-CDD

relativement au mois d'août à Chicago. La Figure 3.3 montre la valeur théorique de la prime pure et quelques valeurs « bootstrap » correspondant à différents nombres d'itérations  $R$ . Les limites de confiance, résultant des propriétés standard de l'estimateur de la moyenne, sont indiquées par des lignes pointillées (extrapolées). Pour  $R=10^3$  ou  $R=10^4$ , l'erreur de simulation est encore importante et, par exemple, il est impossible de décider si l'erreur relative entre  $\hat{P}$  et  $\hat{P}_R$  est supérieure ou inférieure à 5%. En revanche pour  $R=10^6$ , celle-ci représente moins de 1% (autour de la valeur estimée 16.80) et peut donc être négligée.



**Figure 3.3. Comparaison des primes pures : problème de l'erreur de simulation.**

Pour cette raison, les calculs ont été effectués avec  $R=10^6$  simulations. Les résultats sont donnés dans le Tableau 3.4. Nous indiquons la prime pure théorique et la prime pure bootstrap  $\hat{P}_R$  obtenues respectivement avec et sans l'hypothèse de normalité ; le diamètre de confiance  $d_R$ , largeur d'un intervalle de confiance approché de  $\hat{P}$  : c'est un indicateur de l'erreur de simulation ; l'erreur relative  $e_R$  entre  $P$  et  $\hat{P}_R$  ; le signe de la différence observée  $\hat{P}_R - P$ , lorsqu'elle est supérieure à 1%.



Pour les seuils standard d'indices de température, nous constatons que les valeurs obtenues sont très proches si l'on remplace la distribution normale par la distribution empirique ; pour ces exemples, il aurait fallu imposer une fourchette de moins de 0.2% autour de la valeur théorique pour distinguer entre les deux valeurs. En revanche, la différence *relative* est nette pour des seuils extrêmes et semble aller en augmentant lorsqu'on s'écarte du seuil standard. Le signe de la différence observée  $\hat{P}_R - P$  s'explique de la même façon que pour les indices de température (voir le paragraphe précédent).

**Tableau 3.4. Impact des écarts à la normalité sur la prime pure des contrats Futures.**

Station : *Paris-Montsouris*

Contrat Future	HDD décembre		CDD août	
	Standard	0-HDD	Standard	26-CDD
Indice de température (°C)				
<u>Prime pure :</u>				
Valeur théorique ( $P$ )	369.97	1.74	93.52	1.97
Valeur bootstrap ( $\hat{P}_R$ )	369.93	2.02	93.86	1.72
Diamètre de confiance ( $d_R$ )	0.19	0.02	0.14	0.01
Erreur relative ( $e_R$ ) (%)	0.01	15.79	0.36	12.44
Signe de la différence $\hat{P}_R - P$		+		-

Station : *O'Hare Airport*

Contrat Future	HDD décembre			CDD août	
	Standard	20-HDD	10-HDD	Standard	80-CDD
Indice de température (°F)					
<u>Prime pure :</u>					
Valeur théorique ( $P$ )	1079.00	30.53	3.94	256.14	17.51
Valeur bootstrap ( $\hat{P}_R$ )	1078.92	31.71	4.72	257.04	16.80
Diamètre de confiance ( $d_R$ )	0.49	0.12	0.04	0.28	0.07
Erreur relative ( $e_R$ ) (%)	0.01	3.86	19.76	0.35	4.07
Signe de la différence $\hat{P}_R - P$		+	+		-

## 2. Mauvaise spécification de la dynamique.

Nous considérons maintenant les écarts à l'indépendance des résidus  $\varepsilon_t$ . Pour évaluer les incertitudes qui en résultent, nous employons la technique de bootstrap par blocs glissants (« moving block bootstrap », voir Bühlmann, 2000 ou Davinson, Hinkley, 1997). On considère d'abord les  $n-l+1$  blocs de longueur  $l$  de valeurs consécutives à partir de la série résiduelle  $\varepsilon_1, \dots, \varepsilon_n$  :

$$(\varepsilon_1, \dots, \varepsilon_l), (\varepsilon_2, \dots, \varepsilon_{l+1}), \dots, (\varepsilon_{n-l+1}, \dots, \varepsilon_n)$$

Puis on rééchantillonne  $k+1$  blocs avec  $k = \lceil n/l \rceil$ , en ne conservant que les premiers  $n-kl$  éléments dans le dernier bloc de façon à obtenir  $n$  données au total. En agrégeant les blocs obtenus, on obtient la série :

$$\varepsilon_{S_1+1}, \dots, \varepsilon_{S_1+l}, \varepsilon_{S_2+1}, \dots, \varepsilon_{S_2+l}, \dots, \varepsilon_{S_k+1}, \dots, \varepsilon_{S_k+n-kl}$$

où  $S_1, S_2, \dots$  sont indépendantes, de loi uniforme sur  $\{1, \dots, n-l+1\}$ . En rééchantillonnant des blocs, on conserve une partie de la dépendance des résidus mais chaque jointure crée un blanc dans la dynamique. Par conséquent on ne peut pas prendre une valeur trop petite pour la taille des blocs  $l$  ; une valeur trop grande n'est pas non plus recommandée car le nombre de séries reconstruites est alors faible, ce qui limite la portée statistique des résultats. Il ne semble pas exister de règles générales et nous avons préféré suivre la suggestion de Künsch de se baser sur des « jugements subjectifs à partir des corrélations de l'échantillon » (Li et Maddala, 1996, page 139). Ici on observe sur l'autocorrélogramme des résidus et de leurs carrés que les anomalies les plus évidentes concernent les 5 premières autocorrélations, d'où  $l \geq 5$  ; d'autre part, on peut vouloir au moins 10 blocs, ce qui impose  $l \leq 21$  pour les calculs relatifs au mois d'août et  $l \leq 33$  pour le mois de décembre (en plaçant la date courante au 31/12/99). Finalement, nous avons adopté la valeur  $l = 20$ .

Les résultats sont donnés dans le Tableau 3.5. Ils sont beaucoup plus difficiles à interpréter et par exemple, nous ne pouvons pas expliquer ici le signe de la différence  $P - \hat{P}_R$ . Toutefois, les remarques générales du paragraphe précédent restent valables : pour les seuils d'indice standard, les erreurs dues à une mauvaise spécification de la dynamique ne sont pas significatives mais deviennent importantes pour des seuils d'indice extrêmes.

Tableau 3.5. Impact des écarts à l'indépendance sur la prime pure des contrats Futures.

Station: *Paris-Montsouris*

Contrat Future	HDD décembre		CDD août	
	Standard	0-HDD	Standard	26-CDD
Indice de température (°C)				
<u>Prime pure :</u>				
Valeur théorique ( $P$ )	369.97	1.74	93.52	1.97
Valeur bootstrap ( $\hat{P}_R$ )	369.56	1.67	93.26	1.71
Diamètre de confiance ( $d_R$ )	0.19	0.02	0.14	0.01
Erreur relative ( $e_R$ ) (%)	0.11	4.56	0.28	13.01
Signe de la différence $\hat{P}_R - P$		-		-

Station : *O'Hare Airport*

Contrat Future	HDD décembre			CDD août	
	Standard	20-HDD	10-HDD	Standard	80-CDD
Indice de température (°F)					
<u>Prime pure :</u>					
Valeur théorique ( $P$ )	1079.00	30.53	3.94	256.14	17.51
Valeur bootstrap ( $\hat{P}_R$ )	1078.02	30.49	4.10	256.77	17.31
Diamètre de confiance ( $d_R$ )	0.50	0.13	0.04	0.28	0.07
Erreur relative ( $e_R$ ) (%)	0.09	0.15	3.91	0.25	1.13
Signe de la différence $\hat{P}_R - P$			+		-

#### IV. Conclusions de la section C.

Nos premières remarques concernent le modèle de température lui-même : nous avons en effet été amené à calculer une approximation de la matrice de covariance des paramètres du modèle. Bien que la tendance soit significativement croissante (ce qu'on peut expliquer par l'urbanisation et/ou un réchauffement global) l'erreur d'estimation des paramètres correspondants est importante. Cela suggère que la modélisation de la tendance est un point clé pour l'évaluation des produits dérivés climatiques.

Revenons maintenant à l'incertitude des prix. Nous avons étudié l'impact sur les prix de l'erreur d'estimation et de mauvaises spécifications du modèle. Les écarts à la normalité et à l'hypothèse d'indépendance n'affectent pas beaucoup la prime pure. En revanche, l'incertitude des prix résultant de l'erreur d'estimation ne peut être négligée. Pour les contrats Futures standard, elle est de l'ordre de 5% pour la saison HDD, et de 10 à 20% pour la saison CDD. L'importance de ces incertitudes peut se comprendre en regardant l'expression des

indices HDD (ou CDD) : il s'agit d'une somme d'écarts positifs par rapport à un seuil de sorte que peu de compensations peuvent avoir lieu.

Nous étendons maintenant l'étude de l'impact de l'erreur d'estimation à la prime de risque des Futures et au cas des options.

## D. Conséquences et origines de l'erreur d'estimation sur le prix des produits dérivés climatiques.

Nous voulons étudier l'incertitude de la prime pure  $E[\textit{payoff}]$  et de la prime de risque  $\sigma[\textit{payoff}]$  autour des contrats Futures et des options sur la température. La source d'incertitude sera limitée, dans cette section, à l'erreur d'estimation du modèle ARMA de température ; nous avons porté notre attention, dans cette première étude, sur l'erreur qui paraît la plus importante d'après les conclusions de la section précédente.

Les incertitudes des prix vont permettre d'apprécier la qualité du modèle de température pour ce qui est de son applicabilité à l'évaluation des produits dérivés climatiques. Deux éclairages différents vont permettre de préciser les choses. D'abord, nous avons considéré une modélisation des indices de températures eux-mêmes ; on ne peut pas attendre grand chose de cette voie car l'historique de données est très mince (21 données...) et les incertitudes résultant de l'erreur d'estimation de ce modèle « macroscopique » peuvent donc être vues comme une borne supérieure. Ensuite, nous avons étudié le rôle de chaque paramètre – ou groupe de paramètres, de façon à détecter l'origine des incertitudes de prix ; nous en déduisons les améliorations à apporter au modèle.

La section est organisée comme suit. Le premier paragraphe fixe le cadre de travail. L'incertitude des prix est ensuite évaluée, d'abord avec la modélisation des indices de températures puis avec le modèle ARMA.

### I. Cadre de travail. Notations.

Les indices de température de référence sont cette fois les indices HDD et CDD standard

$$HDD = \sum_{t=u}^v (X_{ref} - X_t)^+ \quad CDD = \sum_{t=u}^v (X_t - X_{ref})^+$$

avec  $X_{ref} = 18^\circ C$ . Pour des questions de faisabilité (ou, au moins, dans un but de simplification) évoquées au paragraphe A.II, la période d'exposition au risque  $[u;v]$  est placée en hiver ou en été, sa durée est d'un mois comme pour les contrats standard du CME ; les résultats précédents concernent décembre et juillet. Les indices sont alors des fonctions affines de la température :

$$HDD \approx \sum_{t=u}^v (X_{ref} - X_t) \quad CDD \approx \sum_{t=u}^v (X_t - X_{ref})$$

Par voie de conséquence, la méthodologie présentée ici s'applique aussi pour l'indice utilisé sur le LIFFE, égal à la moyenne mensuelle des températures journalières, et les conclusions sont identiques.

L'étude est réalisée à partir des données de température à Paris (station de Paris-Montsouris). Des difficultés sont apparues avec la station de Chicago : l'introduction de la modélisation des indices de température a fait apparaître un problème de calibration au niveau des options ; en se plaçant dans une situation favorable – et donc en se limitant au cas de Paris, les défauts du modèle n'en seront que plus évidents. Les données de température sont celles du chapitre 1.

Pour les options, plusieurs prix d'exercices ont été retenus. Tout d'abord nous avons considéré la moyenne historique des indices de température, noté  $K_0$ , également retenu par Cao et Wei (1998) ; cette valeur se révélera intéressante pour des raisons statistiques (voir II.2.b), proposition 2.). Bien entendu, il n'y a pas de raison de se limiter à ce choix particulier et en fin de compte, nous avons opté pour cinq prix d'exercices différents

$$K_0 - 60, K_0 - 30, K_0, K_0 + 30, K_0 + 60$$

de telle sorte que la différence entre deux prix d'exercices consécutifs correspond à une différence d'un degré par jour.

### **Rappel des notations.**

$TI$	indice de température.
$X_t$	température journalière moyenne.
$PP$	prime pure (espérance du <i>payoff</i> ).
$PR$	« prime de risque » (volatilité du <i>payoff</i> ).
$K_0$	pour un indice de température donné, moyenne historique des valeurs d'indice.

## **II. Incertitude des prix avec la modélisation des indices de température.**

Afin de fournir un élément de comparaison avec les incertitudes associées au modèle ARMA, nous modélisons dans ce paragraphe les indices de température, sous-jacents des produits

dérivés climatiques étudiés ici. Le manque de données laisse une marge importante dans le choix d'un modèle ; nous en choisissons un, compatible avec la modélisation de la température. Comme on s'y attend, l'erreur de modélisation – et, par contrecoup, l'incertitude des prix – est importante.

### **1. Modélisation des indices de température.**

Bien qu'une modélisation précise soit impossible, on peut proposer un modèle « raisonnable » pour un indice de température. On peut procéder de deux façons, selon que l'on travaille avec l'échantillon des valeurs d'indice lui-même ou à partir de la modélisation « microscopique » de la température.

Rappelons que nous considérons seulement deux saisons : l'hiver et l'été. Pour un mois donné, les valeurs d'indice ne présentent pas de dépendance significative. Par conséquent, on peut envisager de modéliser l'indice par un bruit blanc, et même un bruit blanc gaussien car le test de Kolmogorov ne permet pas de rejeter la normalité.

Un tel modèle est compatible avec la modélisation microscopique. En effet, avec le modèle autorégressif gaussien classiquement proposé pour la température journalière, il n'y a pas de dépendance de long terme et donc pas de dépendance entre les indices mensuels de température d'une année à l'autre. En outre, même si les écarts à la normalité sont réels, ceux-ci ne semblent pas induire des différences importantes sur les prix (C.III.1). Il n'est donc pas déraisonnable de maintenir l'hypothèse de normalité. Or, pour les mois d'hiver et d'été, les indices de température sont approximativement une combinaison affine de la température journalière. L'hypothèse de normalité peut donc être maintenue sur les indices.

Pour toutes ces raisons, nous proposons de modéliser la température par un bruit blanc gaussien,

$$(TI)_i \sim WN(\mu, \sigma^2)$$

### **2. Incertitude des prix.**

#### a) Cas des contrats Futures.

Pour un contrat Future de température, le *payoff* est l'indice de température lui-même. Avec la modélisation retenue, nous en déduisons immédiatement les expressions correspondantes de la prime pure et de la prime de risque :

$$PP = \mu$$

$$PR = \sigma$$

On peut remarquer, en particulier, que chaque prime dépend d'un seul paramètre. Par conséquent, les origines de l'incertitude sont aisément identifiables : l'incertitude autour de la prime pure provient entièrement de l'erreur d'estimation de la moyenne du processus de l'indice de température, alors que l'incertitude autour de la prime de risque dépend seulement de la qualité de l'estimation de la volatilité de l'indice. Dans le cas présent, on peut évaluer l'incertitude en donnant un intervalle de confiance pour chaque prix (ou paramètre). Les estimateurs usuels de  $\mu$  et  $\sigma^2$  sont :

$$\hat{\mu} = \overline{TI} := \frac{TI_1 + TI_2 + \dots + TI_n}{n}$$

$$\hat{\sigma}^2 = S_{TI}^2 := \frac{1}{n-1} \sum_{i=1}^n (TI_i - \overline{TI})^2$$

On a alors, avec des résultats élémentaires de statistique :

**Proposition 1.** *Des intervalles de confiance de niveau  $1-\alpha$  pour la prime pure  $PP$  et la prime de risque  $PR$  sont donnés par :*

$$PP: \left[ \overline{TI} - t_{n-1}(1-\alpha/2) \frac{S_{TI}}{\sqrt{n}} ; \overline{TI} + t_{n-1}(1-\alpha/2) \frac{S_{TI}}{\sqrt{n}} \right]$$

$$PR: \left[ \frac{\sqrt{n-1}}{\sqrt{\chi_{n-1}^2(1-\alpha/2)}} S_{TI} ; \frac{\sqrt{n-1}}{\sqrt{\chi_{n-1}^2(\alpha/2)}} S_{TI} \right]$$

où  $t_{n-1}(x)$  (resp.  $\chi_{n-1}^2(x)$ ) désigne le quantile d'ordre  $x$  de la loi de Student (resp. chi-deux) à  $n-1$  degrés de liberté.

En outre, l'incertitude relative autour de  $PR$ , mesurée par la variable aléatoire  $\frac{\widehat{PR} - PR}{PR}$  (où  $\widehat{PR} = \hat{\sigma}$  est l'estimateur de la prime de risque) ne dépend pas des spécifications du contrat : lieu, mois, nature de l'indice de température (HDD ou CDD).

b) Cas des options.

Les expressions de  $PP$  et  $PR$  s'écrivent sous la forme

$$PP = f_1(\mu, \sigma)$$

$$PR = f_2(\mu, \sigma)$$



où  $f_1$  et  $f_2$  peuvent être explicitées (voir l'annexe A.1.). Comme le *payoff* de l'option est une fonction non linéaire de l'indice de température, les paramètres  $\mu$  et  $\sigma$  interviennent tous deux dans ces expressions et donc, contrairement au cas des contrats Futures, l'incertitude de chaque partie du prix dépend de l'erreur d'estimation des deux paramètres. Cependant, on a le résultat suivant (la preuve est donnée dans l'annexe A.2.) :

**Proposition 2.** *Ce résultat étant valable autant pour la prime pure que pour la prime de risque, nous noterons  $P$  pour désigner l'une quelconque de ces primes –  $P$  est une fonction des caractéristiques  $\mu$  et  $\sigma$  de l'indice de température ; notons  $\hat{P}$  l'estimateur correspondant, obtenu comme fonction de leurs estimateurs naturels  $\hat{\mu}$  et  $\hat{\sigma}$  .*

*Considérons une option dont le prix d'exercice est égal à la moyenne des valeurs historiques de l'indice de température. Alors l'incertitude relative autour de  $P$ , mesurée par la variable aléatoire  $\frac{\hat{P}-P}{P}$ , ne dépend pas des spécifications de l'option : nature (option d'achat ou option de vente), lieu, mois, nature de l'indice de température (HDD ou CDD). D'un point de vue statistique, sa distribution est pivotale (indépendante de  $\mu$  et  $\sigma$ ) ; en notant  $q(x)$  le quantile d'ordre  $x$  correspondant, on en déduit un intervalle de confiance de niveau  $1-\alpha$  pour  $P$  :*

$$P: \left[ \frac{\hat{P}}{1+q(1-\alpha/2)}; \frac{\hat{P}}{1+q(\alpha/2)} \right]$$

*Les quantiles doivent être estimés par simulation.*

Dans le cas général, l'incertitude doit être évaluée d'une autre façon car les distributions exactes de  $PP$  et  $PR$ , vues comme des variables aléatoires fonction de  $\hat{\mu}$  et  $\hat{\sigma}$ , ne sont pas disponibles. Celle-ci peut néanmoins être approchée avec précision en simulant avec la loi théorique de  $\hat{\mu}$  et  $\hat{\sigma}^2$ . L'incertitude des prix est alors évaluée par deux quantiles.

c) Résultats.

Quelques résultats représentatifs figurent dans le Tableau 3.6, le Tableau 3.7 et le Tableau 3.8. D'une façon générale, nous observons une grande incertitude autour des prix, plus de 20%, et ce particulièrement dans le cas des options : elle représente environ 50% lorsque le prix d'exercice est la moyenne des valeurs historiques d'indice, et peut aller au-delà de 100% pour d'autres prix d'exercice (en fait, la qualité de l'estimation semble décroître avec la valeur estimée). La plus petite incertitude est relative à la prime pure des contrats Futures et pourrait même être acceptable pour les contrats HDD des mois d'hiver (environ 5%).

**Tableau 3.6. Incertitude des prix de contrats Futures avec la modélisation des indices de température.**

Future HDD décembre. Paris		
	<b>Prime pure</b>	<b>Prime de risque</b>
Valeur estimée	381.1	38.9
Incertitude	[363.3; 398.8]	[29.7; 56.2]
Incertitude relative (%)	[-4.7; 4.7]	[-23.5; 44.5]

Future CDD juillet. Paris		
	<b>Prime pure</b>	<b>Prime de risque</b>
Valeur estimée	80.3	41.1
Incertitude	[61.5; 99.0]	[31.4; 59.4]
Incertitude relative (%)	[-23.4; 23.4]	[-23.5; 44.5]

**Tableau 3.7. Incertitude des prix des options avec la modélisation des indices de température.**

Option de prix d'exercice $K_0$		
	<b>Prime pure</b>	<b>Prime de risque</b>
Incertitude relative (%)	[-54.7; 68.1]	[-40.3; 38.6]

*Exemple.* Option d'achat sur l'indice HDD de décembre de prix d'exercice  $K_0$  (valeur correspondante : 381.1). Paris.

	<b>Prime pure</b>	<b>Prime de risque</b>
Valeur estimée	15.5	22.7
Incertitude	[9.2; 34.3]	[16.3; 38.0]
Incertitude relative (%)	[-54.7; 68.1]	[-40.3; 38.6]

*Notes.*

1. L'incertitude relative est mesurée autour de la valeur estimée.
2.  $K_0$  est la moyenne des valeurs historiques de l'indice de température.

**Tableau 3.8. Incertitude du prix des options en fonction du prix d'exercice avec la modélisation des indices de température.**

Prime pure de quelques options sur l'indice HDD de décembre. Paris.

Prix d'exercice	Valeur estimée	Incertitude	Incertitude relative (%)
$K_0 - 60$	61.0	[45.8; 77.1]	[-25.0; 26.4]
$K_0 - 30$	34.9	[22.2; 49.2]	[-36.4; 40.8]
$K_0$	15.5	[9.2; 34.3]	[-54.7; 68.1]
$K_0 + 30$	4.9	[1.1; 11.2]	[-76.8; 126.9]
$K_0 + 60$	1.0	[0.1; 3.8]	[-92.7; 265.7]

Prime de risque de quelques options sur l'indice HDD de décembre. Paris.

Prix d'exercice	Valeur estimée	Incertitude	Incertitude relative (%)
$K_0 - 60$	36.8	[26.5; 45.9]	[-28.1; 24.7]
$K_0 - 30$	31.7	[22.8; 40.1]	[-28.2; 26.2]
$K_0$	22.7	[16.3; 38.0]	[-40.3; 38.6]
$K_0 + 30$	12.7	[5.0; 21.4]	[-60.2; 67.8]
$K_0 + 60$	5.5	[1.1; 12.3]	[-79.0; 123.3]

*Notes.*

1. L'incertitude relative est mesurée autour de la valeur estimée.
2.  $K_0$  est la moyenne des valeurs historiques de l'indice de température, ici 381,1.

### III. Incertitude des prix avec la modélisation de la température.

Nous évaluons d'abord l'impact global de l'erreur de modélisation sur les prix, puis les répercussions d'erreurs locales provenant d'un paramètre ou d'un groupe de paramètres.

#### 1. Incertitude globale.

La variabilité des prix due à l'erreur d'estimation du modèle de température est mesurée en suivant la méthodologie détaillée dans la section B.

Les résultats sont indiqués sur le Tableau 3.9, le Tableau 3.10 et le Tableau 3.11. D'une façon générale, nous observons le même phénomène que pour la modélisation de l'indice de température, à savoir une énorme incertitude autour du prix des options. En ce sens, les résultats sont quelque peu décevants. En fait, la modélisation microscopique de la température, qui était supposée réduire l'incertitude des prix, conduit à des intervalles de

confiance qui sont presque aussi larges que ceux que l'on obtient avec la modélisation macroscopique de l'indice de température (voir aussi Figure 3.4 et Figure 3.5). Une exception est fournie par l'estimation de la prime de risque des contrats Futures qui est cinq fois plus précise avec la modélisation de la température. Cependant, elle révèle un autre problème majeur : avec une prime de risque annoncée entre 45.5 et 52.2 (Tableau 3.9), la prime de risque *historique* du contrat Future sur l'indice HDD de décembre, égale à 38.9 (Tableau 3.6), devient hautement improbable ! Ainsi, dans ce cas, les résultats ne sont pas compatibles avec les données historiques. C'est le problème général de la calibration : les paramètres estimés sont les plus vraisemblables au regard des données journalières (estimation statistique) mais pas en comparaison des prix historiques (calibration).

**Tableau 3.9. Incertitude des prix de contrats Futures avec la modélisation de la température.**

Future HDD décembre. Paris		
	<b>Prime pure</b>	<b>Prime de risque</b>
Valeur estimée	370.0	49.1
Incertitude	[351.6; 388.9]	[45.5; 52.2]
Incertitude relative (%)	[-5.2; 5.0]	[-7.3; 6.3]

Future CDD juillet. Paris		
	<b>Prime pure</b>	<b>Prime de risque</b>
Valeur estimée	88.3	46.1
Incertitude	[70.1; 106.2]	[43.0; 48.9]
Incertitude relative (%)	[-20.5; 20.4]	[-6.6; 6.0]

**Tableau 3.10. Incertitude des prix des options avec la modélisation de la température.**

Option d'achat sur Future HDD décembre, de prix d'exercice  $K_0$ . Paris.

	<b>Prime pure</b>	<b>Prime de risque</b>
Valeur estimée	14.5	24.9
Incertitude	[8.1; 23.6]	[18.3; 31.3]
Incertitude relative (%)	[-44.2; 62.3]	[-26.2; 25.8]

Option d'achat sur Future CDD juillet, de prix d'exercice  $K_0$ . Paris.

	<b>Prime pure</b>	<b>Prime de risque</b>
Valeur estimée	22.6	29.6
Incertitude	[13.7; 34.2]	[23.2; 35.3]
Incertitude relative (%)	[-39.2; 51.0]	[-21.6; 19.2]

*Notes.*

1. L'incertitude relative est mesurée autour de la valeur estimée.
2.  $K_0$  est la moyenne des valeurs historiques de l'indice de température.

**Tableau 3.11. Incertitude du prix des options en fonction du prix d'exercice avec la modélisation de la température.**

Prime pure de quelques options sur l'indice HDD de décembre. Paris.

Prix d'exercice	Valeur estimée	Incertitude	Incertitude relative (%)
$K_0 - 60$	53.0	[38.4; 69.6]	[-27.5; 31.2]
$K_0 - 30$	30.4	[19.5; 44.1]	[-35.9; 44.5]
$K_0$	14.5	[8.1; 23.6]	[-44.2; 62.3]
$K_0 + 30$	5.5	[2.5; 10.5]	[-54.6; 89.6]
$K_0 + 60$	1.6	[0.6; 3.7]	[-62.6; 126.8]

Prime de risque de quelques options d'achat sur l'indice HDD de décembre. Paris.

Prix d'exercice	Valeur estimée	Incertitude	Incertitude relative (%)
$K_0 - 60$	42.5	[37.6; 46.2]	[-11.5; 8.7]
$K_0 - 30$	34.8	[28.3; 40.3]	[-18.8; 15.5]
$K_0$	24.9	[18.3; 31.3]	[-26.2; 25.8]
$K_0 + 30$	15.1	[9.7; 21.2]	[-35.2; 40.4]
$K_0 + 60$	7.8	[4.4; 12.2]	[-42.4; 56.7]

*Notes.*

1. L'incertitude relative est mesurée autour de la valeur estimée.
2.  $K_0$  est la moyenne des valeurs historiques de l'indice de température, ici 381,1.

Pour aller plus loin dans la compréhension des résultats, nous examinons l'influence locale des paramètres sur les prix.

## 2. Incertitude locale.

Jusqu'à présent, nous avons évalué l'incertitude des prix due à l'erreur d'estimation de tous les paramètres. Mais peut-on quantifier l'impact de l'erreur d'estimation de la tendance seulement (par exemple) sur le prix des produits dérivés climatiques ? Nous parlerons d'incertitude locale par opposition à l'incertitude (globale) des prix étudiée précédemment. Notons  $\Theta_1$  le vecteur formé de l'ensemble des paramètres dont nous voulons évaluer l'influence (dans notre exemple,  $\Theta_1$  est constitué des deux paramètres modélisant la tendance) et  $\Theta_2$  le vecteur contenant les paramètres restants ; soient  $\hat{\Theta}_1$  et  $\hat{\Theta}_2$  les estimateurs du maximum de vraisemblance correspondants. Alors une approche possible pour répondre au problème est d'estimer la loi conditionnelle :

$$P(\hat{\Theta}_1) \Big| \hat{\Theta}_2 = \Theta_2$$

L'incertitude locale des prix peut donc être évaluée exactement de la même façon que pour l'incertitude globale en suivant la méthodologie détaillée à la section B. En effet, supposons que :

$$\sqrt{n}(\hat{\Theta} - \Theta) \xrightarrow{n \rightarrow +\infty} N(0; \Gamma(\Theta))$$

et introduisons la partition de  $\Gamma(\Theta)$  associée à celle de  $\Theta = (\Theta_1; \Theta_2)$  :  $\Gamma = \begin{bmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{bmatrix}$  ;  $\Gamma_{ij}$

représente donc la matrice des covariances (asymptotiques) entre  $\sqrt{n}(\hat{\Theta}_i - \Theta_i)$  et  $\sqrt{n}(\hat{\Theta}_j - \Theta_j)$ . On a alors (voir Saporta, § 4.2.4) :

$$\sqrt{n}(\hat{\Theta}_1 - \Theta_1 \Big| \hat{\Theta}_2 = \Theta_2) \xrightarrow{n \rightarrow +\infty} N(0; \Gamma_{11} - \Gamma_{12} \Gamma_{22}^{-1} \Gamma_{21})$$

et en utilisant la méthode delta (voir §B.II.1), on en déduit :

$$\sqrt{n}(P(\hat{\Theta}_1) - P(\Theta_1) \Big| \hat{\Theta}_2 = \Theta_2) \xrightarrow{n \rightarrow +\infty} N(0; v_{P|\Theta_2}(\Theta_1))$$

avec  $v_{P|\Theta_2}(\Theta_1) = \frac{\partial P}{\partial \Theta_1} (\Gamma_{11} - \Gamma_{12} \Gamma_{22}^{-1} \Gamma_{21}) \frac{\partial P}{\partial \Theta_1}$ .

En pratique, comme nous l'avons expliqué au paragraphe B.II.3, nous n'utilisons pas ce dernier résultat mais préférons rééchantillonner dans la distribution asymptotique conditionnelle de  $\hat{\Theta}_1$  pour obtenir une approximation de la loi exacte de  $P(\hat{\Theta}_1) \Big| \hat{\Theta}_2 = \Theta_2$ .

### 3. Résultats.

La Figure 3.6 et la Figure 3.8 montrent l'influence marginale des groupes de paramètres du modèle de température. Elles représentent des échantillons obtenus par bootstrap (avec  $R = 10^4$  simulations) de la distribution (non conditionnelle) des prix et des distributions conditionnelles des prix relativement à la tendance, la saisonnalité, la volatilité ou la partie autorégressive. Avant de commenter les résultats, remarquons que dans le cas Future, seulement deux groupes de paramètres interviennent dans les expressions de la prime pure et de la prime de risque (voir l'annexe A.1.) – la tendance et la saisonnalité pour la prime pure, la dispersion et la partie stochastique pour la prime de risque. Ceci est analogue à ce que nous avons observé avec la modélisation des indices de température (II.2.a) puisque la tendance et la saisonnalité caractérisent la moyenne du processus de température tandis que dispersion et

partie autorégressive sont caractéristiques de sa volatilité. Maintenant, nous observons de façon nette que les plus grandes incertitudes proviennent essentiellement des erreurs d'estimation de la tendance et de la saisonnalité. Nous pouvons même ajouter que toutes les deux sont responsables de l'incertitude des prix puisque les paramètres de tendance sont approximativement indépendants des paramètres de saisonnalité (ainsi que l'atteste la matrice de corrélation estimée, voir l'annexe A.3 et le Tableau 3.13).

Comparons ensuite les distributions conditionnelles relativement à un seul paramètre (Figure 3.7 et Figure 3.9). Nous voyons que la source d'incertitude la plus grande est relative à la pente de la tendance. Pour obtenir ce résultat, un changement de paramètres a été nécessaire. En effet, avec la forme initiale, les deux paramètres de tendance et de saisonnalité sont fortement négativement corrélés et il aurait été impossible de distinguer l'influence spécifique d'un seul d'entre eux ; avec la nouvelle paramétrisation, en revanche, la pente de la tendance est approximativement indépendante de tous les autres paramètres (voir annexe A.3).

#### **IV. Conclusions de la section D.**

Deux objectifs principaux ont motivé les travaux regroupés au sein de cette section. D'une part, évaluer les conséquences de l'erreur d'estimation du modèle de température du chapitre 1 sur les prix de produits dérivés climatiques de type Future et option. D'autre part, connaître la source de ces erreurs, en identifiant les points faibles du modèle relativement au calcul des prix.

Pour le premier point, il apparaît que les incertitudes des prix de contrats Futures se situent autour de 10%, ce qui, sans être négligeable, peut être acceptable. En revanche, le cas des options est problématique puisque les incertitudes autour de la valeur annoncée dépassent facilement 50%. En outre, nous avons effectué le même genre de calculs avec un modèle censé, en principe, être plus grossier, dans le but de disposer d'un élément de comparaison. Pour cela, nous avons considéré les indices de température utilisés comme sous-jacents des produits étudiés, et avons remarqué que ces indices pouvaient être modélisés (en première approximation) par un bruit blanc. Avec ce modèle, les incertitudes autour des prix des options ne sont même pas deux fois supérieures, alors que seulement 21 données ont été utilisées pour l'estimation de ses paramètres, contre 7.665 pour le modèle ARMA...

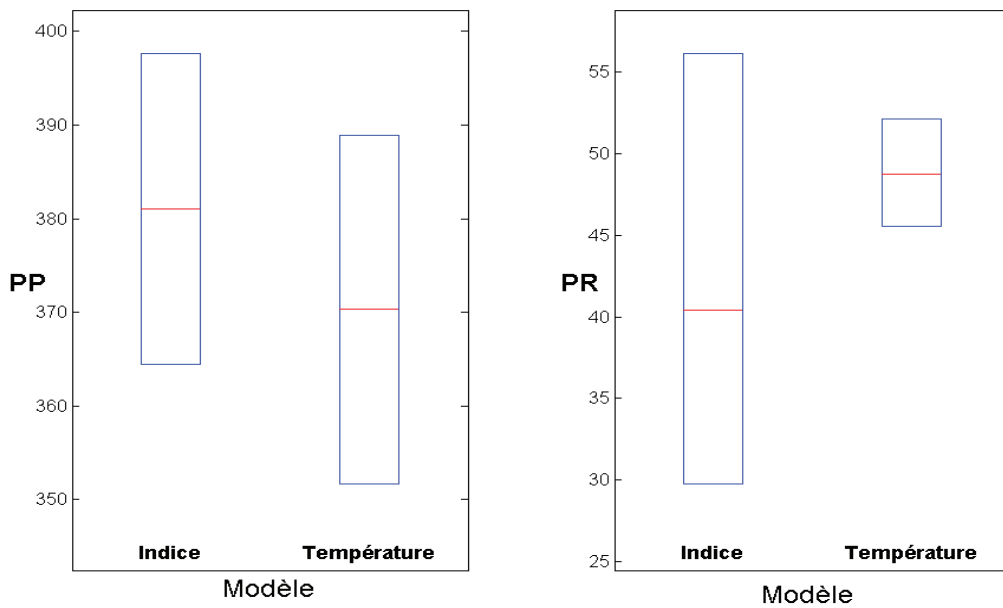
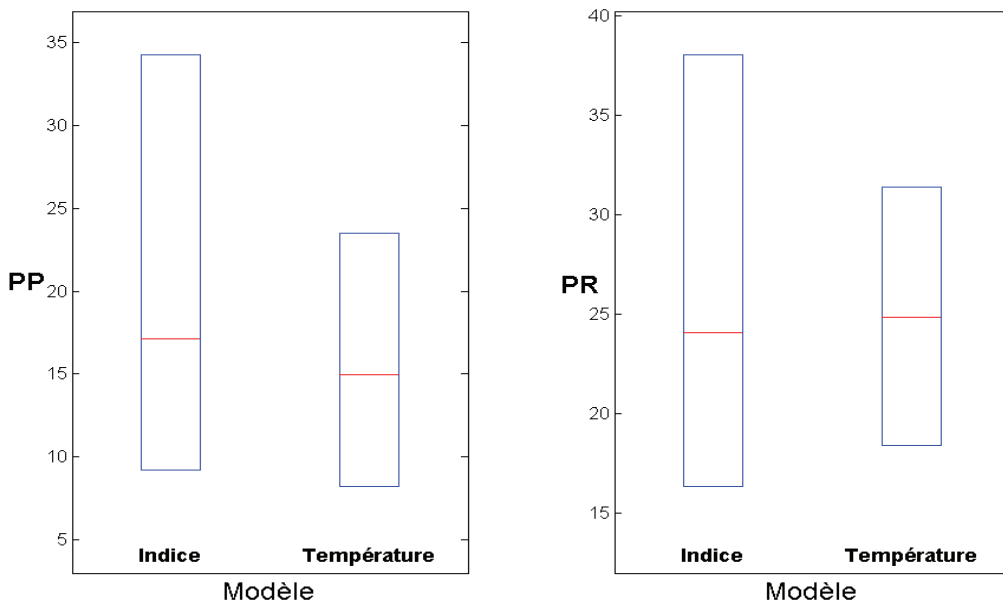


Figure 3.4. Comparaison de l'incertitude des prix des contrats Futures avec les deux modélisations.

Figure 3.5. Comparaison de l'incertitude des prix des options avec les deux modélisations.



*Notes.*

1. Les figures donnent des intervalles de confiance obtenus avec la modélisation « macroscopique » de l'indice de température ou avec la modélisation « microscopique » de la température. Les produits considérés sont le contrat Future et l'option d'achat de prix d'exercice  $K_0$ , sur l'indice HDD de décembre, à Paris.
2.  $K_0$  est la moyenne des valeurs historiques d'indice.
3.  $PP$  : prime pure –  $PR$  : prime de risque.



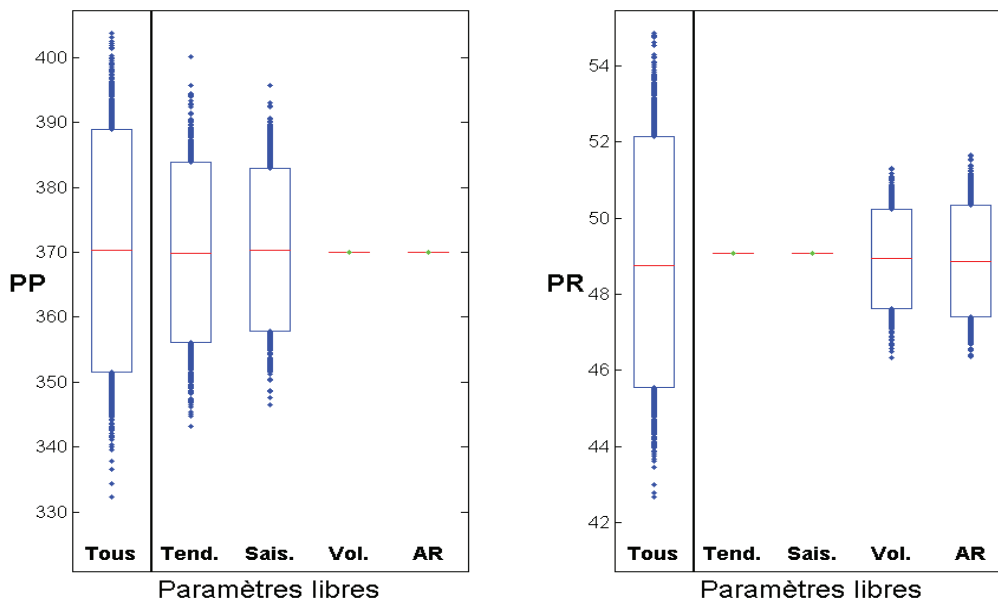
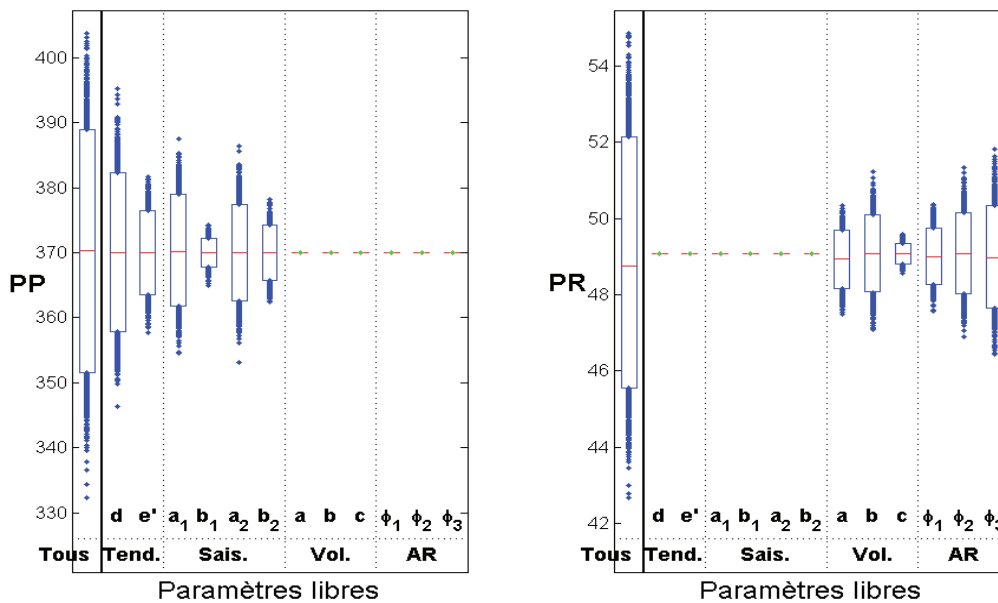


Figure 3.6. Impact de l'erreur d'estimation des composantes du modèle de température sur le prix des contrats Futures.

Figure 3.7. Impact de l'erreur d'estimation des paramètres du modèle de température sur le prix des contrats Futures.



*Notes.*

1. Les figures ci-dessus représentent des échantillons de prix (loi non conditionnelle et lois conditionnelles) obtenus par bootstrap, pour le contrat Future sur l'indice *HDD* du mois de décembre, Paris. Dans le conditionnement, seuls les paramètres (ou groupes de paramètres) indiqués ne sont pas fixés.
2. *PP* : prime pure – *PR* : prime de risque.

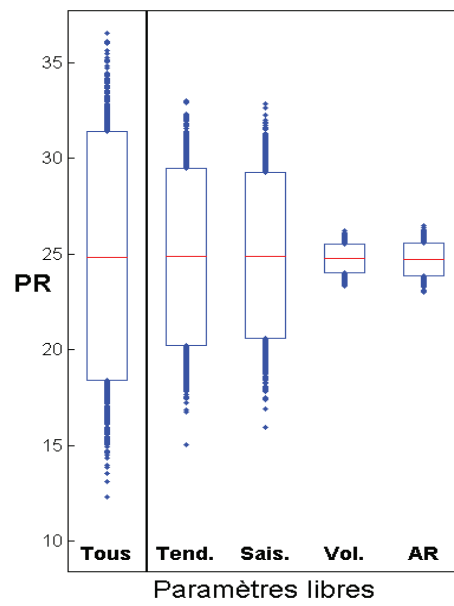
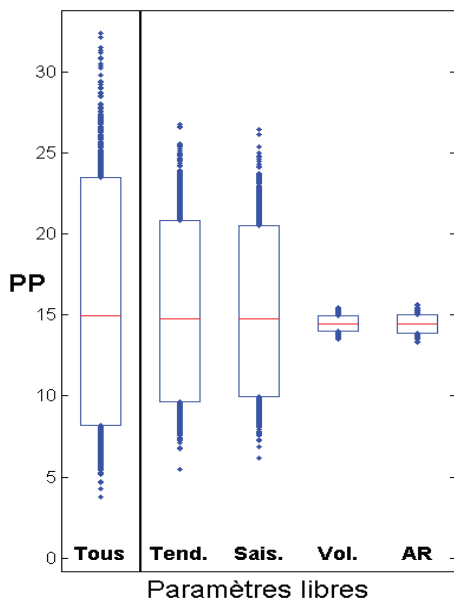
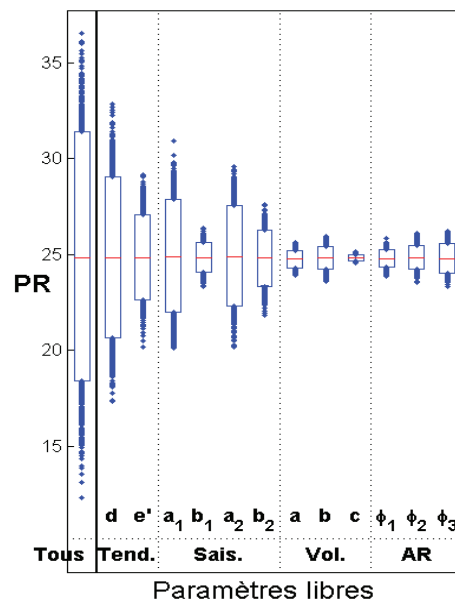
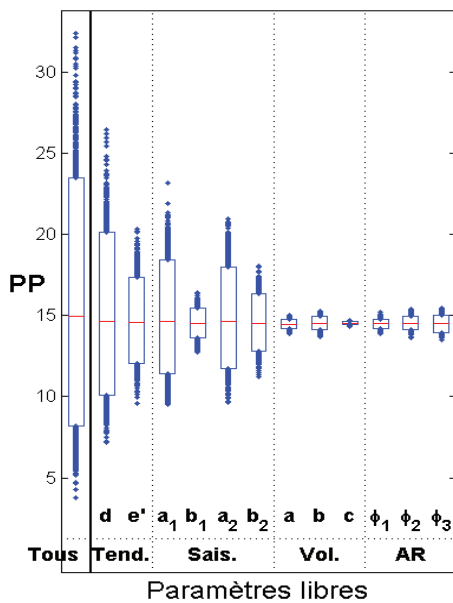


Figure 3.8. Impact de l'erreur d'estimation des composantes du modèle de température sur le prix des options

Figure 3.9. Impact de l'erreur d'estimation des paramètres du modèle de température sur le prix des options



*Notes.*

1. Les figures ci-dessus représentent des échantillons de prix (loi non conditionnelle et lois conditionnelles) obtenus par bootstrap, pour l'option d'achat sur l'indice HDD du mois de décembre de prix d'exercice  $K_0$ , Paris. Dans le conditionnement, seuls les paramètres (ou groupes de paramètres) indiqués ne sont pas fixés.
2.  $K_0$  est la moyenne des valeurs historiques d'indice.
3.  $PP$  : prime pure –  $PR$  : prime de risque.

Nous avons ensuite recherché dans le modèle de température la source des incertitudes évaluées sur les prix. Nous avons alors observé que ce sont les composantes déterministes caractéristiques de la moyenne du processus, la tendance et la saisonnalité, qui sont en grande partie – et toutes deux - responsables de ces incertitudes. Une analyse plus fine a également montré que le paramètre le plus nuisible est la pente de la tendance (bien que le manque de précision ne soit pas dû uniquement à ce seul paramètre) ; cela confirme l'idée que la modélisation de la tendance par une droite est trop grossière et, en particulier, qu'une petite variation de la pente engendre une grande variation du niveau de température, et donc des prix de modèle.

Pour finir, notons que nous avons choisi ici de réaliser cette étude pour des contrats spécifiant principalement une ville française comme lieu de référence. Il s'agit a priori de la situation la plus favorable, car on a vu (chapitre 1) que le modèle de température est plus « juste » dans ce cas. Les résultats obtenus avec des villes américaines conduisent à des conclusions comparables ; cependant, le problème de calibration soulevé au paragraphe III.1, est largement accentué.

### **Conclusions du chapitre 3.**

L'objectif de ce chapitre était de quantifier les risques de modèle associés à l'utilisation du modèle statistique dans l'évaluation des produits dérivés climatiques. Pour ce faire, nous avons supposé que l'évaluation est réalisée dans le cadre actuariel utilisé en pratique. Nous avons alors évalué les incertitudes autour des prix causées par l'erreur d'estimation dans le modèle de température, et, dans le cas des contrats Futures, par deux types d'erreurs de spécification de ce modèle.

Les résultats sont mitigés. Pour les contrats Futures, l'incertitude la plus forte, qui est celle provenant de l'erreur d'estimation, n'est pas négligeable mais reste modérée, de l'ordre de 10%. En revanche, l'incertitude autour du prix des options, qui va souvent au-delà de 50%, est très importante.

Nous avons alors cherché à connaître l'origine de ces incertitudes (section D). Il en ressort que ce sont les composantes déterministes de tendance et de saisonnalité du modèle de température qui sont les plus préjudiciables pour l'évaluation des produits dérivés climatiques. D'une façon générale, ceci pourrait s'expliquer par la forme des indices de température à partir desquels sont calculés les prix : constitués d'une somme d'écart de température sur toute la période d'exposition au risque, ils amplifieraient particulièrement les erreurs d'estimation relatives à la moyenne du processus de température (davantage, par exemple, que les erreurs d'estimation des corrélations temporelles).

### Références du chapitre 3.

- Bawa V.S., Brown S.J., and Klein R.W., *Estimation Risk and Optimal Portfolio Choice*, Studies in Bayesian Econometrics Bell Laboratories Series, New York: North Holland, 1979.
- Bouchaud J.P., Potters M., *Theory of Financial Risks. From statistical Physics to Risk Management*. Cambridge University Press, 2000.
- Brockwell P.J., Davis R.A., *Time Series : Theory and Methods*, Springer Verlag, 1991.
- Brody D.C., Syroka J., Zervos M., *Dynamical Pricing of Weather Derivatives*. working Paper, Imperial College, London. Submitted to : Quantitative Finance, 2001.
- Bühlmann P., *Bootstrap for Time Series*. Seminar für Statistik, ETH Zürich. Revised Version, 2000.
- Cairns A.J.G., *A discussion of parameter and model uncertainty in insurance*. Insurance: Mathematics and Economics **27**, 313-330, 2000.
- Campbell J.Y., Lo A.W., MacKinlay A.C., *The Econometrics of Financial Markets*. Princeton University Press, 1997.
- Campbell S., Diebold F.X., *Weather Forecasting for Weather Derivatives*. working Paper, University of Pennsylvania, Philadelphia, 2000.
- Cao M., Wei J., *Pricing Weather Derivatives: an Equilibrium Approach*, working paper, Queen's University Kingston and University of Toronto, Ontario, Canada ,1998.
- Davinson A.C., Hinkley D.V., *Bootstrap Methods and their Applications*. Cambridge University Press, 1997.
- Davis M., *Pricing weather derivatives by marginal value*, Quantitative Finance, **1**, 1-4, 2001.

- Dornier F., Quéruef M., *Caution to the wind*, Risk, Energy and Power Risk Management, August 2000.
- Efron B., Tibshirani R., *Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Accuracy*, Journal of Statistical Science **1**(1), 54-77, 1986.
- Geman H., *Instruments dérivés à sous-jacent exotique : l'exemple des dérivés climatiques*, Banque & Marchés, **53**, 5-9, 2001.
- Gourieroux C., Montfort A., *Séries temporelles et modèles dynamiques*, Economica, 2<sup>ème</sup> édition, 1995.
- Hull J., White A., *Pricing Interest rate securities*, Review of Financial Studies, **3**, 4, 573-92, 1990.
- Jorion P., *International Portfolio Diversification with Estimation Risk*, Journal of Business, **58**(3), 259-278, 1985.
- Li H., Maddala G.S., *Bootstrapping Time Series Models*, Econometric Reviews, **15**(2), 115-158, 1996.
- Moréno M., *Evaluation des Dérivés Climatiques*, Speedwell Weather Derivatives, [www.weatherderivs.com](http://www.weatherderivs.com), 2000.
- Moréno M., *Riding the Temp*, Weather Derivatives, FOW Special Supplement, December 2000.
- Moréno M., Roustant O., *Modélisation des températures*, in Blondeau J., Partrat C., *La réassurance : approche technique*, Economica, 2003.
- Roustant O., Laurent J.P., Bay X. and Carraro L., *Estimation Risk and the Pricing of Weather Derivatives*, working paper, [www.affi.asso.fr](http://www.affi.asso.fr), 2002.
- Roustant O., Laurent J.P., Bay X. and Carraro L., *An empirical study of the price uncertainty of some weather derivatives*, working paper, <http://laurent.jeanpaul.free.fr/>, 2002.
- Roustant O. *Une application de deux modèles économétriques de température à la gestion de risques climatiques (1<sup>ère</sup> partie)*, Banque & Marchés, **58**, 22-29, 2002.

Roustant O. *Une application de deux modèles économétriques de température à la gestion de risques climatiques (2ème partie)*, Banque & Marchés, **59**, 36-44, 2002.

Saporta G., *Probabilités, Analyse des Données et Statistique*, Technip, 1990.

Tassi P. *Méthodes Statistiques*. Economica, 1989.

## Annexe du chapitre 3

### A.1. Expressions analytiques des prix des produits dérivés climatiques.

Lorsque le processus de température est gaussien, les prix de produits dérivés climatiques peuvent se calculer explicitement. Le lemme suivant est alors utile :

**Lemme.** Soit un réel  $K$ , et  $X$  une variable aléatoire gaussienne,  $X \sim N(\mu, \sigma^2)$ . Alors avec

$Y = (K - X)^+$  (resp.  $Y = (X - K)^+$ ), on a :

$$E[Y^+] = \sigma \left( L.N(L) + \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{L^2}{2}\right) \right)$$

$$E[(Y^+)^2] = \sigma^2 \left( (1 + L^2).N(L) + \frac{L}{\sqrt{2\pi}} \exp\left(-\frac{L^2}{2}\right) \right)$$

où  $L = \frac{K - \mu}{\sigma}$  (resp.  $L = -\frac{K - \mu}{\sigma}$ ), et  $N(\cdot)$  désigne la fonction de répartition de la loi normale centrée réduite.

### Applications

Considérons par exemple le cas de contrats basés sur l'indice  $HDD$ , notons  $[u; v]$  la période d'exposition au risque et  $(X_t)$  le processus de température. Rappelons que l'on se place dans le cadre actuariel (voir l'introduction de ce chapitre) et que la date courante est suffisamment éloignée de la période d'exposition de sorte qu'il n'est pas nécessaire de prendre en compte la situation courante dans l'évaluation.

□ **Prime pure d'un contrat Future (cf. section C).**

La prime pure d'un contrat Future sur indice HDD est donné par

$$PP = E[HDD] = \sum_{t=u}^v E\left[(X_{\text{ref}} - X_t)^+\right]$$

Il suffit donc d'utiliser le premier résultat du lemme avec  $K = X_{\text{ref}}$ .



□ **Prime pure et prime de risque des Futures et des options relatifs aux mois d'hiver et d'été (cf. section D).**

• 1er cas : avec la modélisation de la température (section D, § III).

Pour les mois considérés, l'indice HDD est approximativement linéaire :

$$HDD \approx \sum_{t=u}^v (X_{\text{ref}} - X_t)$$

où  $X_{\text{ref}}$  désigne ici le seuil standard de 18°C. Rappelons que

$$X_t = m_t + s_t + \sigma_t \cdot Z_t,$$

avec  $Z_t = \phi_1 \cdot Z_{t-1} + \dots + \phi_p \cdot Z_{t-p} + \varepsilon_t$ , où  $(\varepsilon_t)$  est un bruit blanc gaussien. Le processus linéaire  $(X_t)$  est alors gaussien et on en déduit :

$$HDD \sim N \left( l \cdot X_{\text{ref}} - \sum_{t=u}^v (m_t + s_t); \rho' \cdot \Gamma_Z \cdot \rho \right)$$

où  $l = v - u + 1$  (longueur de la période d'exposition),  $\rho$  est le vecteur  $l \times 1$ ,  $\rho = (\rho_u; \rho_{u+1}; \dots; \rho_v)$  et  $\Gamma_Z$  la matrice d'autocorrélations  $l \times l$  du processus autorégressif  $Z$ ,  $(\Gamma_Z)_{i,j} = \text{cov}(Z_0; Z_{|i-j|})$ ,  $1 \leq i, j \leq l$ . Dans le cas de contrats Futures HDD, la prime pure et la prime de risque, données respectivement par la moyenne et l'écart-type de l'indice HDD, s'en déduisent immédiatement ; pour les options, il suffit d'appliquer le lemme ci-dessus en prenant cette fois pour  $K$  le prix d'exercice des options.

• 2ème cas : avec la modélisation de l'indice de température (section D, § II).

Dans ce cas, l'indice est supposé être un bruit blanc gaussien. La situation est donc la même qu'au paragraphe précédent, les calculs s'effectuent de la même façon.

**A.2. Preuve de la proposition 2 (section D, § II.2.b).**

Nous supposons ici qu'un indice de température est modélisé par un bruit blanc gaussien :

$$TI \sim WN(\mu, \sigma^2)$$

En pratique,  $\mu$  et  $\sigma$  sont remplacés par leur estimation naturelle et, en particulier,  $\mu$  est égal à la moyenne des valeurs historiques d'indice, ce que nous avons noté  $K_0$ . Considérons alors une option sur un indice donné de prix d'exercice  $K$ . Notons  $P$  pour désigner l'une quelconque des primes attachées à l'option : prime pure ou prime de risque –  $P$  est une

fonction de  $\mu$  et  $\sigma$  ; notons  $\hat{P}$  l'estimateur correspondant, obtenu comme fonction des estimateurs naturels  $\hat{\mu}$  et  $\hat{\sigma}$  de  $\mu$  et  $\sigma$  .

Nous allons montrer que dans le cas où  $K = K_0 (= \mu)$ , la distribution de l'incertitude de la prime  $\frac{\hat{P} - P}{P}$  est indépendante de  $\mu$  et  $\sigma$  (on parle de loi « pivotale »), et des spécifications de l'option autres que le prix d'exercice : nature (option d'achat ou option de vente), lieu, mois, nature de l'indice (HDD ou CDD).

Observons déjà qu'on peut se ramener à  $\frac{\hat{P}}{P} = \frac{\hat{P} - P}{P} + 1$ , et qu'une fois établie l'indépendance de  $\frac{\hat{P}}{P}$ , il restera seulement à s'assurer que la distribution est la même pour une option d'achat

ou une option de vente. Maintenant, remarquons que la distribution de  $\frac{TI - \mu}{\sigma}$  est pivotale. Il

en résulte que  $P$  peut s'écrire sous la forme  $P = \sigma \cdot f\left(\frac{K - \mu}{\sigma}\right)$  (par exemple, pour la prime

pure d'une option d'achat, on a  $P = E[(K - TI)^+] = \sigma \cdot E\left[\left(\frac{K - \mu}{\sigma} - \frac{TI - \mu}{\sigma}\right)^+\right]$ ). L'estimateur

associé s'écrit  $\hat{P} = \hat{\sigma} \cdot f\left(\frac{K - \hat{\mu}}{\hat{\sigma}}\right)$  et dans le cas où  $K = K_0 = \mu$ , on en déduit :

$$\frac{\hat{P}}{P} = \frac{1}{f(0)} X \cdot f\left(\frac{Y}{X}\right) \quad (*)$$

avec  $X = \frac{\hat{\sigma}}{\sigma}$  et  $Y = \frac{\hat{\mu} - \mu}{\sigma}$ . Or  $X$  et  $Y$  sont indépendantes et pivotales d'après le théorème de

Fisher (voir Tassi, IV.4.). Par conséquent, la distribution de  $\frac{\hat{P}}{P}$  est pivotale. Que cette loi ne

dépende pas de la nature de l'option résulte simplement du fait que changer la nature revient à changer  $Y$  en  $-Y$  dans (\*); le résultat en découle puisque la distribution de  $Y$  est symétrique.

### **A.3. Estimation de la variance asymptotique des paramètres du modèle de température.**

Nous considérons le modèle de température du chapitre 1,

$$X_t = m_t + s_t + \sigma_t \cdot Z_t,$$

avec  $Z_t = \phi_1 \cdot Z_{t-1} + \dots + \phi_p \cdot Z_{t-p} + \varepsilon_t$ , où le processus des résidus  $(\varepsilon_t)$  est un bruit blanc gaussien. Notons  $\Theta$  le vecteur des paramètres, et  $\hat{\Theta}$  l'estimateur du maximum de

vraisemblance. En supposant que  $\sqrt{n}(\hat{\Theta} - \Theta) \xrightarrow[n \rightarrow +\infty]{} N(0, \Gamma(\Theta))$ , comment peut-on estimer  $\Gamma(\Theta)$  ?

Si nous disposons de  $R$  réalisations indépendantes de  $\hat{\Theta} : \Theta^{*1}, \dots, \Theta^{*R}$ , nous pourrions obtenir une approximation de  $\Gamma(\Theta)$  avec la matrice de covariance calculée à partir des  $\Theta^{*r}$  :

$$\hat{\Gamma}_{i,j} = \frac{1}{n} \sum_{r=1}^R (\Theta_i^{*r} - \bar{\Theta}_i)(\Theta_j^{*r} - \bar{\Theta}_j) \quad (**)$$

où  $\bar{\Theta}_i = \frac{1}{R} \sum_{r=1}^R \Theta_i^{*r}$  ( $\Theta_i^{*r}$  désigne la  $i^{\text{ème}}$  coordonnée de  $\Theta^{*r}$ ).

Or l'échantillon  $\Theta^{*1}, \dots, \Theta^{*R}$  peut être obtenu à partir des données par la méthode standard de bootstrap (Efron et Tibshirani, 1986, §6 ou Davinson, Hinkley, 1997). Pratiquement, elle consiste à réaliser les opérations suivantes. Pour  $r = 1, \dots, R$  :

- simuler indépendamment  $\varepsilon_1^*, \dots, \varepsilon_n^*$  selon la loi normale estimée des résidus.
- reconstituer la trajectoire de température correspondante  $x_1^*, \dots, x_n^*$  avec le modèle de température et l'estimation *initiale* des paramètres :

$$x_t^* = m_t + s_t + \sigma_t \cdot z_t^*,$$

$$\text{avec } z_t^* = \hat{\phi}_1 \cdot z_{t-1}^* + \dots + \hat{\phi}_p \cdot z_{t-p}^* + \varepsilon_t^* \text{ et } z_0^* = z_{-1}^* = \dots = z_{1-p}^* = 0;$$

- calculer l'estimation du maximum de vraisemblance  $\Theta^{*r}$  pour les « nouvelles » données  $x_1^*, \dots, x_n^*$ .

En fait, on fait une erreur statistique en remplaçant la vraie distribution des résidus par la loi estimée et les vrais paramètres par leur estimation. Cependant, on peut penser que cette erreur est faible en raison du nombre assez élevé de données,  $n = 7665$ .

Finalement, on calcule l'approximation  $\hat{\Gamma}(\Theta)$  avec (\*\*).

### Application numérique

Nous donnons dans le Tableau 3.12 l'écart-type d'estimation des paramètres et la matrice de corrélation estimée obtenue par la technique de bootstrap ci-dessus dans le cas des stations météorologiques de Paris-Montsouris et O'Hare Airport (Chicago), avec  $R = 2.000$  simulations. Les paramètres ont été regroupés selon les caractéristiques du modèle : tendance, saisonnalité, dispersion, et partie autorégressive. Pour des raisons de lisibilité, seule la moitié de la matrice est présentée ; l'autre partie s'en déduit par symétrie.

Voici quelques remarques, permettant de mieux comprendre les résultats.

➤ Commençons par l'estimation de la tendance. Nous constatons que la pente estimée est significativement positive, bien que l'erreur d'estimation soit importante ; la tendance est donc croissante d'une façon significative. Cette conclusion est cohérente par rapport au phénomène global de réchauffement et à l'urbanisation observés au cours de ces dernières décades.

D'autre part, les estimateurs de la pente  $d$  et de l'ordonnée à l'origine  $e$  sont fortement négativement corrélés. En effet, surestimer la pente conduit obligatoirement, pour « coller » aux données, à sous-estimer l'ordonnée à l'origine comme dans un mouvement de balançoire. On peut éliminer cette corrélation en remplaçant la constante  $e$  par  $e' = d\bar{t} + e$  où  $\bar{t}$  est le temps moyen – l'axe de la balançoire. Le résultat est patent, comme on peut le voir dans le Tableau 3.13. Ce changement de paramétrage ne modifie pas les résultats d'incertitudes de prix; en revanche, il est absolument nécessaire pour distinguer le rôle que joue la pente dans l'impact de l'erreur d'estimation de la tendance sur les prix.

➤ Les estimateurs de  $a$  et  $\phi_1$  sont positivement corrélés. Pour comprendre pourquoi, commençons par simplifier le problème en ne gardant dans  $\sigma_t$  et  $Z_t$  que les termes les plus influents : approchons  $\sigma_t$  par la constante  $a$ , et  $Z_t$  par un processus AR(1) de telle sorte que  $X_t \approx m_t + s_t + a.Z_t$  avec  $Z_t \approx \phi_1.Z_{t-1} + \varepsilon_t$ . Alors, intuitivement, on peut remarquer que surestimer  $\phi_1$  conduit à diminuer la dispersion de  $Z_t$ , ce qui amène à surestimer  $a$  pour coller aux données.

➤ Enfin, les valeurs estimées des coefficients de la partie autorégressive sont très proches de celles qui avaient été obtenues en estimant les coefficients du modèle AR(3) après avoir retiré la tendance et la saisonnalité, et avec une dispersion constante. L'estimation par maximum de vraisemblance d'un tel modèle est en fait simplement une optimisation par moindres carrés, tandis que l'estimation du modèle de température, qui se fait également par maximum de vraisemblance, est quasiment équivalente à des moindres carrés pondérés par  $\sigma_t$ . Le fait que la pondération ne soit pas très marquée explique probablement le peu de différences des deux estimations. L'introduction d'une variance périodique est utile essentiellement pour obtenir des erreurs de prévisions réalistes.

**Tableau 3.12. Ecart-type d'estimation et matrice de corrélation des paramètres du modèle de température.**

*Paris-Montsouris*

$\theta$	$\hat{\theta}_{emv}$	$\sigma$	(Tendance)		(Saisonnalité)				(Dispersion)			(partie AR)					
			$d$	$e$	$a_1$	$b_1$	$a_2$	$b_2$	$a$	$b$	$c$	$\phi_1$	$\phi_2$	$\phi_3$			
$d$	1.85*	.48*	1														
$e$	11.43	.21	-.87	1													
$a_1$	-7.39	.15	.01	.03	1												
$b_1$	-2.67	.15	.06	-.03	-.01	1											
$a_2$	-0.14	.15	-.01	-.00	.03	-.05	1										
$b_2$	0.74	.15	.03	-.01	.04	.01	.00	1									
$a$	3.35	.05	-.00	.01	.01	.00	.02	.03	1								
$b$	0.11	.04	-.01	.01	.03	.01	-.02	.01	.02	1							
$c$	0.15	.04	-.02	.03	-.02	-.02	.02	.00	.10	-.03	1						
$\phi_1$	0.92	.01	-.01	.02	-.01	.00	.01	.02	.57	-.01	.03	1					
$\phi_2$	-0.19	.01	.01	.00	.03	.01	-.00	-.01	-.08	.01	.03	-.68	1				
$\phi_3$	0.06	.01	.02	-.03	-.02	-.01	.00	.01	.06	.02	-.01	.16	-.68	1			

\* :  $\times 10^{-4}$

*O'Hare Airport*

$\theta$	$\hat{\theta}_{emv}$	$\sigma$	(Tendance)		(Saisonnalité)		(Dispersion)			(partie AR)							
			$d$	$e$	$a_1$	$b_1$	$a$	$b$	$c$	$\phi_1$	$\phi_2$	$\phi_3$					
$d$	2.36*	.98*	1														
$e$	48.56	.44	-.85	1													
$a_1$	-23.44	.32	.00	.17	1												
$b_1$	-8.50	.31	.04	.03	.00	1											
$a$	8.88	.12	-.01	.02	.01	-.00	1										
$b$	1.91	.10	-.02	.01	.03	.01	.29	1									
$c$	.89	.10	-.02	.02	-.02	-.02	.16	-.01	1								
$\phi_1$	.87	.01	-.02	.03	-.00	-.00	.62	.12	.06	1							
$\phi_2$	-.27	.01	.01	.00	.04	.02	-.13	-.02	.02	-.66	1						
$\phi_3$	.10	.01	.02	-.03	-.01	-.02	.09	.03	.00	.20	-.66	1					

\* :  $\times 10^{-4}$

*Note.*

Les résultats ont été obtenus par bootstrap avec 2.000 simulations.

**Tableau 3.13. Ecart-type d'estimation et matrice de corrélation des paramètres du modèle de température, avec la nouvelle paramétrisation de la tendance.**

*Paris-Montsouris*

$\theta$	$\hat{\theta}_{emv}$	$\sigma$	(Tendance)		(Saisonnalité)				(Dispersion)			(partie AR)				
			$d$	$e'$	$a_1$	$b_1$	$A_2$	$b_2$	$a$	$b$	$c$	$\phi_1$	$\phi_2$	$\phi_3$		
$d$	1.85*	.48*	1													
$e'$	12.1	.11	.00	1												
$a_1$	-7.39	.15	.01	.07	1											
$b_1$	-2.67	.15	.06	.04	-.01	1										
$a_2$	-0.14	.15	-.01	-.01	.03	-.05	1									
$b_2$	0.74	.15	.03	.02	.04	.01	.00	1								
$a$	3.35	.05	-.00	.01	.01	.00	.02	.03	1							
$b$	0.11	.04	-.01	.00	.03	.01	-.02	.01	.02	1						
$c$	0.15	.04	-.02	.02	-.02	-.02	.02	.00	.10	-.03	1					
$\phi_1$	0.92	.01	-.01	.02	-.01	.00	.01	.02	.57	-.01	.03	1				
$\phi_2$	-0.19	.01	.01	.01	.03	.01	-.00	-.01	-.08	.01	.03	-.68	1			
$\phi_3$	0.06	.01	.02	.03	-.02	-.01	.00	.01	.06	.02	-.01	.16	-.68	1		

\* :  $\times 10^{-4}$

*Note.*

Les résultats ont été obtenus par bootstrap avec 2.000 simulations.

Tableau 3.14. Résultats de tests de la normalité de l'estimateur des paramètres du modèle de température.

	Paris	Chicago
<i>Lois marginales</i>	Statistique de Kolmogorov	
$\hat{d}$	* 0.0248	0.0160
$\hat{e}$	* 0.0228	0.0167
$\hat{a}_1$	0.0101	0.0152
$\hat{b}_1$	0.0169	0.0129
$\hat{a}_2$	0.0171	
$\hat{b}_2$	0.0137	
$\hat{a}$	0.0165	0.0155
$\hat{b}$	0.0105	0.0146
$\hat{c}$	0.0095	0.0094
$\hat{\phi}_1$	0.0158	0.0193
$\hat{\phi}_2$	0.0133	0.0135
$\hat{\phi}_3$	0.0141	0.0162
<i>Loi multidimensionnelle</i>	<i>p-valeur</i>	
Test <i>skewness</i>	0.5955	0.6258
Test <i>kurtosis</i>	0.0517	0.0879
Test joint	0.1497	0.2227

*Notes.*

1. Le tableau donne les résultats des tests de normalité pour des échantillons de taille  $R = 2.000$  obtenus par bootstrap des estimateurs de maximum de vraisemblance des paramètres du modèle de température pour les stations de Paris-Montsouris et O'Hare Airport (Chicago).
2. Les valeurs repérées par un astérisque conduisent à un rejet de l'hypothèse de normalité au niveau 95% avec le test de Kolmogorov. Le seuil correspondant est égal à 0.0200.
3. Les tests basés sur *skewness* et *kurtosis* peuvent être trouvés dans (Lütkepohl, 1993, § 4.5., formules 4.5.4, 4.5.5 et 4.5.8.).

## **Conclusion générale.**



La présente thèse a abordé quelques aspects économétriques et financiers des produits dérivés climatiques de température. Au plan de la gestion, les produits dérivés climatiques font partie de ces nouveaux produits financiers pour lesquels la méthodologie de Black et Scholes est inopérante : dans leur cas, la raison est simplement que les indices de température sous-jacents ne sont pas des biens échangeables. Pour la même raison, le marché qu'ils constituent se comporte comme un marché incomplet. Une autre particularité, à l'heure actuelle, provient du manque de données de prix.

Cette absence de repères, théoriques et historiques, renforce le rôle joué par les modèles statistiques dans la gestion financière, et, par voie de conséquence, accroît l'importance des interrogations liées à leur utilisation, à la fois au plan financier au sujet de la pratique de l'évaluation ou au plan statistique à propos des risques de modèle. De là, la structuration de la thèse en trois chapitres : modélisation statistique de la température, aspects financiers, incertitudes de modélisation et répercussions sur les prix.

Les travaux sur la modélisation de la température, que nous avons réalisés sur un ensemble de stations météorologiques françaises, ont permis de voir qu'au strict plan statistique un modèle temporel linéaire de type autorégressif convient dans la majorité des cas ; le modèle doit néanmoins intégrer une saisonnalité dans la volatilité afin de rendre compte de la saisonnalité de la dispersion des températures. Par ailleurs, l'adjonction de l'aspect spatial dans la modélisation ne semble pas se justifier, car les améliorations qu'elle apporte semblent dérisoires relativement à la complexification qu'elle engendre.

Muni de ce modèle, nous avons pu alors étudier quelques aspects financiers du risque de température. La question de son indépendance vis-à-vis du risque de marché a été examinée, en considérant d'abord le marché financier dans son ensemble, puis en se restreignant aux actions d'entreprises exposées au climat. Seul dans le premier cas, paradoxalement, une faible dépendance se laisse observer, que l'on attribue généralement à l'impact du climat sur le comportement humain. Par conséquent, l'impact de la température sur les valeurs financières semble être du second ordre, ce qui donnerait raison à la publicité qui est faite autour des produits dérivés climatiques, vantant la décorrélation des risques climatiques<sup>79</sup> par rapport au marché financier. En ce qui concerne la pratique de l'évaluation des produits dérivés

---

<sup>79</sup> La température est le risque de température le plus fréquent sur le marché des produits dérivés climatiques.

climatiques, ces résultats nous ont permis - sans toutefois la justifier - de donner un argument en faveur de l'approche actuarielle.

L'utilisation des modèles statistiques n'est pourtant pas sans risques. En se plaçant dans un cadre actuariel, et en utilisant le modèle autorégressif du chapitre 1, nous avons mis en évidence de très importantes incertitudes autour des prix d'options climatiques. Ainsi les critères statistiques intrinsèques qui ont suffi à valider le modèle de température ne sont pas assez restrictifs pour prévenir des risques de son utilisation pour l'évaluation des produits dérivés climatiques. La source des erreurs a été identifiée : il s'agit des composantes déterministes de tendance et de saisonnalité relatives à la moyenne du processus.

Une controverse existe actuellement quant à la voie à suivre pour modéliser la température : certains se suffisent de modèles à mémoire courte comme le modèle autorégressif étudié ici, d'autres en revanche prônent l'incorporation de mémoire longue dans la modélisation<sup>80</sup>. Il nous semble que les résultats que nous obtenons placent le débat à un autre niveau. Car, dans ces deux types d'approches, la première étape, consistant à se débarrasser des composantes déterministes pour se consacrer à l'étude de la partie aléatoire pure, est souvent considérée comme une procédure automatique, et reléguée au second plan. Pourtant, il apparaît que ce sont ces composantes qui ont l'influence la plus néfaste sur l'évaluation des produits dérivés climatiques. Pour ces produits en effet, la structure des indices climatiques semble démultiplier les erreurs de modélisation de la moyenne du processus.

Parmi les voies de recherches futures, la recherche d'un modèle permettant une modélisation plus précise de la normale saisonnière paraît inévitable. A cette fin, il est vraisemblable que celle-ci doive être considérée non plus comme un terme déterministe de second ordre, mais comme un processus stochastique à part entière. Le développement des modèles non linéaires pourrait permettre de trouver une modélisation adéquate. Une deuxième voie de recherche consiste à l'extension de l'étude de la rémunération de la température à un cadre plus large que le MEDAF. Une étude économétrique reste ainsi à faire à propos de la significativité des bêtas action/température avec un modèle multifacteurs.

---

<sup>80</sup> Voir par exemple (Brody, Syroka, Zervos, 2001) ou (Caballero, Jewson, Brix, 2002).



## Bibliographie générale.

Antoniadou T., Besse P., Fougères A.L., Le Gall C., Stephenson D.B., *L'oscillation Atlantique Nord et son influence sur le climat européen*, Revue de Statistique Appliquée, **XLIX**(3), 39-60, 2001.

Bank for International Settlements, *Regular OTC Derivatives Market Statistics*, May 2002.

Barrieu P., *Produits dérivés météorologiques et environnement*, thèse de doctorat en sciences de gestion, Groupe HEC, [www.hec.fr](http://www.hec.fr), 2002.

Barrieu P., and El Karoui N., *Optimal design of derivatives in illiquid markets*, Quantitative Finance, **2**, 181-188, 2002.

Barrieu P., *Introduction au marché des produits dérivés météo*, séminaire ISFA « Couverture des risques naturels et climatiques », Paris, 24 et 25 octobre 2000.

Bawa V.S., Brown S.J., and Klein R.W., *Estimation Risk and Optimal Portfolio Choice*, Studies in Bayesian Econometrics Bell Laboratories Series, New York: North Holland, 1979.

Bouchaud J.P., Potters M., *Theory of Financial Risks. From statistical Physics to Risk Management*. Cambridge University Press, 2000.

Brockwell P.J., Davis R.A., *Time Series : Theory and Methods*, Springer-Verlag, 1991.

Brody D.C., Syroka J., Zervos M., *Dynamical Pricing of Weather Derivatives*. working Paper, Imperial College, London. Submitted to : Quantitative Finance, 2001.

Bühlmann H., *Mathematical Methods in Risk Theory*, Springer-Verlag, 1996.

Bühlmann P., *Bootstrap for Time Series*. Seminar für Statistik, ETH Zürich. Revised Version, 2000.

- Caballero R., Jewson S., Brix A., *Long memory in surface air temperature: detection, modeling, and application to weather derivative valuation*, *Climate Research*, **21**, 127-140, 2002.
- Cairns A.J.G., *A discussion of parameter and model uncertainty in insurance*. *Insurance: Mathematics and Economics* **27**, 313-330, 2000.
- Campbell J.Y., Lo A.W., MacKinlay A.C., *The Econometrics of Financial Markets*. Princeton University Press, 1997.
- Campbell S., Diebold F.X., *Weather Forecasting for Weather Derivatives*. working Paper, University of Pennsylvania, Philadelphia, 2000.
- Cao M., Wei J., *Stock Market Returns : A Temperature Anomaly*, working Paper, 2001.
- Cao M., Wei J., *Pricing Weather Derivatives: an Equilibrium Approach*, working paper, Queen's University Kingston and University of Toronto, Ontario, Canada, 1998.
- Carmona R., *Calibrating Degree Days Options*, working paper, 1999.
- Carr P., Geman H. and Madan D., *Pricing and Hedging in Incomplete Markets*, working paper, 1999.
- Copeland T.E., Weston J.F., *Financial Theory and Corporate Policy*, Addison Wesley, 1988.
- Davinson A.C., Hinkley D.V., *Bootstrap Methods and their Applications*. Cambridge University Press, 1997.
- Davis M., *Pricing weather derivatives by marginal value*, *Quantitative Finance*, **1**, 1-4, 2001.
- Deboaisne E., *Les dérivés climatiques*, séminaire ISFA « Couverture des risques naturels et climatiques », Paris, 24 et 25 octobre 2000.
- Dischel R., *Double Trouble: Hedging Rainfall and Temperature*, *Risk Magazine and Energy Power Risk Management*, August 2001.

Dischel R., *Seeding a Rain Market*, Environmental Finance, September 2000.

Dischel R., *Is precipitation Basis Risk Overstated?*, Risk Magazine and Energy Power Risk Management, August 2000.

Dischel R., *Seasonnal Weather Forecasts and Derivative Valuation*, Risk, Energy and Power Risk Management, août 2000.

Dischel R., *Warning - La Niña Volatility*, Energy and Power Risk Management, November 1998.

Dischel R., *The Fledgling Weather Market Takes Off, Part. 2.*, Applied Derivatives Trading, October 1998.

Dornier F., Quérueu M., *Caution to the wind*, Risk, Energy and Power Risk Management, August 2000.

Efron B., Tibshirani R., *Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Accuracy*, Journal of Statistical Science **1**(1), 54-77, 1986.

Efron B., Morris C., *Le paradoxe de Stein*, Scientific American, n°1, nov. 1977.

Ellithorpe D., Putnam S., *Weather Derivatives and Their Implications for Power Markets*, The Journal of Risk Finance, 19-28, Winter 2000.

Eydeland A., Geman H., *Pricing power derivatives*, Risk, 71-73, October 1998.

Fewings D.R., Pekoz E.A., *Robust Crop Hedging Strategy for Free Market Agriculture*, working Paper.

Garcia A.F., Sturzenegger F., *Hedging Corporate Revenues with Weather Derivatives : A Case Study*, Master of Science in Banking and Finance – MBF Master's Thesis, Université de Lausanne, HEC, Suisse, 2001.

Geman H., *Instruments dérivés à sous-jacent exotique : l'exemple des dérivés climatiques*, Banque & Marchés, **53**, 5-9, 2001.

Geman H., *The Bermuda Triangle: Electricity, Weather and Insurance Derivatives*, Journal of Alternative Investments, 2000.

- Geman H., *Insurance and Weather Derivatives: from Exotic Options to Exotic Underlyings*.  
Risk Books, London, 1999
- Geman H., *Cat Calls*, *Insurance*, **7**(9) 86-89, 1994.
- Gerber H.U. and Parfumi G., *Utility Functions: From Risk Theory to Finance*, *North American Actuarial Journal*, **2**(3), 74-100, 1998.
- Gerber H.U., *An Introduction to Mathematical Risk Theory*, Huebner, 1980
- Gondran N. *Système de diffusion d'information pour encourager les PME/PMI à améliorer leurs performances environnementales*, thèse de doctorat en Sciences et Techniques du Déchet, Ecole Nationale Supérieure des Mines de Saint-Etienne, [www.agora21.org/entreprise/these2.html](http://www.agora21.org/entreprise/these2.html), 2001.
- Gourieroux C., Jasiak J., *Nonlinear Innovations and Impulse Responses*, octobre 1998.
- Gourieroux C., Montfort A., *Séries temporelles et modèles dynamiques*, *Economica*, 2<sup>ème</sup> édition, 1995.
- Hamon J., Jacquillat B., *Le marché français des actions - Etudes empiriques 1977-1991*, Presses Universitaires de France, 1992.
- Hollander M., Wolfe D.A., *Nonparametric Statistical Methods*, Wiley Series, 1999.
- Hull J., White A., *Pricing Interest rate securities*, *Review of Financial Studies*, **3**, 4, 573-92, 1990.
- Jones C., Wand R., *Kernel Smoothing*, Chapman and Hall, 1995.
- Jorion P., *International Portfolio Diversification with Estimation Risk*, *Journal of Business*, **58**(3), 259-278, 1985.
- Koop G., Pesaran M.H., Potter S.M., *Impulse response analysis in nonlinear multivariate models*, *Journal of Econometrics*, **74**, 119-147, 1996.

- Laurent J.P., Roustant O., Weather derivatives and the stock market: a risk assessment, working paper, [www.affi.asso.fr](http://www.affi.asso.fr), 2003.
- Li H., Maddala G.S., *Bootstrapping Time Series Models*, *Econometric Reviews*, **15**(2), 115-158, 1996.
- Lütkepohl H., *Introduction to Multiple Time Series Analysis*, Springer Verlag, 1993.
- Makridakis S., Wheelwright S.C., Hyndman R.J., *Forecasting. Methods and Applications*, Wiley, 1998.
- Martin P.-E., *Certificats verts : situation européenne et mode d'emploi*, *Systèmes Solaires*, **147**, 10-17, 2002.
- Moréno M., Roustant O., *Modélisation des températures*, in Blondeau J., Partrat C., *La réassurance : approche technique*, Economica, 2003.
- Moréno M., *Rain Risk*, Speedwell Weather Derivatives, [www.weatherderivs.com](http://www.weatherderivs.com), 2001.
- Moréno M., *Evaluation des Dérivés Climatiques*, Speedwell Weather Derivatives, [www.weatherderivs.com](http://www.weatherderivs.com), 2000.
- Moréno M., *Riding the Temp*, Weather Derivatives, FOW Special Supplement, December 2000.
- Nott D.J., Dunsmuir W.T.M., Kohn R., Woodcock F., *Statistical Correction of a Deterministic Numerical Weather Prediction Model*, *Journal of the American Statistical Association*, **96**(455), 2001.
- ONU, Climate Change & The Financial Services Industry, [www.unepfi.net](http://www.unepfi.net), July 2002.
- Quéruel M., *Les risques climatiques et naturels : de nouvelles opportunités pour les investisseurs*, séminaire ISFA « Couverture des risques naturels et climatiques », Paris, 24 et 25 octobre 2000.
- Roustant O., Laurent J.P., Bay X. and Carraro L., *Estimation Risk and the Pricing of Weather Derivatives*, working paper, [www.affi.asso.fr](http://www.affi.asso.fr), 2002.



- Roustant O., Laurent J.P., Bay X. and Carraro L., *An empirical study of the price uncertainty of some weather derivatives*, working paper, <http://laurent.jeanpaul.free.fr/>, 2002.
- Roustant O. *Une application de deux modèles économétriques de température à la gestion de risques climatiques (1<sup>ère</sup> partie)*, Banque & Marchés, **58**, 22-29, 2002.
- Roustant O. *Une application de deux modèles économétriques de température à la gestion de risques climatiques (2<sup>ème</sup> partie)*, Banque & Marchés, **59**, 36-44, 2002.
- Sankaran K., *Finding a value*, Risk, Energy and Power Risk Management, août 2000.
- Sanso B., Guenni L., *A Nonstationary Multisite Model for Rainfall*, Journal of the American Statistical Association, **95**(452), 2000.
- Saporta G., *Probabilités, Analyse des Données et Statistique*, Technip, 1990.
- Saunders E.M.J., *Stock Prices and Wall Street Weather*, American Economic Review, **83**, 1337-1345, 1993.
- Scherer A., *Les cat-bonds : de nouvelles opportunités pour les réassureurs*, séminaire ISFA « Couverture des risques naturels et climatiques », Paris, 24 et 25 octobre 2000.
- Solen D, *Six actions à acheter maintenant pour miser sur « l'effet canicule »*, Capital, **129**, juin 2002.
- Tassi P. *Méthodes Statistiques*. Economica, 1989.
- Tchapda Djamen I., *Evaluation de flux monétaires en présence d'un ou plusieurs risques de défaut*, thèse de doctorat, 2002.
- Trevor C., Linsley P., *The use of weather derivatives in agriculture*, University of Northumbria at Newcastle, 2002.
- Weinstein J., *Carbon-denominated weather swaps*, Environmental Finance, November 2001.

Weather Risk Management Association, *Results of Second Annual Industry-Wide Survey*, [www.wrma.org](http://www.wrma.org) , 2002.

Wikle C.K., Milliff R.F., Nychka D., Berliner L.M., *Spatiotemporal Hierarchical Bayesian Modeling: Tropical Ocean Surface Winds*, Journal of the American Statistical Association, **96**(454), 2001.

Wilks D.S., *Statistical Methods in the Atmospheric Sciences*, Academic Press, 1995.

Yuan K., Zheng L., Zhu Q., *Are Investors Moonstruck? Lunar Phases and Stock Returns*, University of Michigan Business School, 2001.





# **PRODUITS DÉRIVÉS CLIMATIQUES : ASPECTS ÉCONOMÉTRIQUES ET FINANCIERS.**

## **Résumé.**

Cette thèse constitue l'une des premières études des produits dérivés climatiques. Elle examine d'abord la modélisation de la température, qui est la variable climatique la plus fréquente (chapitre 1). Un modèle univarié de type autorégressif à volatilité périodique est en général approprié pour décrire sa dynamique, notamment pour les données françaises. Elle aborde alors quelques aspects financiers des risques climatiques. En particulier, la quasi-indépendance du marché à la température semble justifier la pratique de l'évaluation selon une approche actuarielle (chapitre 2). Enfin, elle quantifie le risque de modèle lié à cette pratique (chapitre 3). Tandis que les prix des contrats Futures sont robustes par rapport aux erreurs de modélisation, d'importantes incertitudes sont mises en évidence autour du prix des options. La source des erreurs a été identifiée : il s'agit des composantes déterministes de tendance et de saisonnalité relatives à la moyenne du processus de température.

## **Mots-clés :**

Produits dérivés climatiques, température, modèle ARMA, modèle VAR, bêta action/température, erreur d'estimation, bootstrap, risque de modèle.

# **WEATHER DERIVATIVES: ECONOMETRICS AND FINANCIAL ASPECTS.**

## **Abstract.**

This thesis is one of the first research studies about weather derivatives. Firstly, it focuses on the modelling of temperature, which is the most frequent weather variable (chapter 1). An autoregressive univariate model with periodic volatility is shown to be appropriate to describe its dynamics, especially with French data. Then, it addresses some financial aspects of weather risks. In particular, the fact that market and temperature are quasi independent seems to justify the common use of an actuarial framework for valuation (chapter 2). Finally, the model risk corresponding to this practise is quantified in chapter 3; whereas Futures prices appear to be robust to modelling errors, large uncertainties are shown around option prices. This source of uncertainty has been identified: it concerns trend and seasonality, the deterministic components of the temperature process mean.

## **Keywords:**

Weather derivatives, temperature, ARMA model, VAR model, stock/temperature beta, parameter uncertainty, bootstrap, model risk.

## **Laboratoire de recherche :**

Ecole ISFA - Campus de la Doua - 21, av. Claude Bernard  
Bât. Doyen Jean Braconnier - 69 622 Villeurbanne cedex.

**Discipline :** Sciences de Gestion.