

Pratique R

# Séries temporelles avec R

Yves Aragon

# Séries temporelles avec $\mathbb{R}$

**Vj k' r ci g' k p v g p v k p c m ( ' i g h ' d r e p m**

Yves Aragon

# Séries temporelles avec R

 edp sciences

ISBN : 978-2-7598-1779-5

© 2016, **EDP Sciences**, 17, avenue du Hoggar, BP 112, Parc d'activités de Courtabœuf,  
91944 Les Ulis Cedex A

Imprimé en France

Tous droits de traduction, d'adaptation et de reproduction par tous procédés réservés pour tous pays. Toute reproduction ou représentation intégrale ou partielle, par quelque procédé que ce soit, des pages publiées dans le présent ouvrage, faite sans l'autorisation de l'éditeur est illicite et constitue une contrefaçon. Seules sont autorisées, d'une part, les reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective, et d'autre part, les courtes citations justifiées par le caractère scientifique ou d'information de l'œuvre dans laquelle elles sont incorporées (art. L. 122-4, L. 122-5 et L. 335-2 du Code de la propriété intellectuelle). Des photocopies payantes peuvent être réalisées avec l'accord de l'éditeur. S'adresser au : Centre français d'exploitation du droit de copie, 3, rue Hautefeuille, 75006 Paris. Tél. : 01 43 26 95 35.

## Collection Pratique R

dirigée par Pierre-André Cornillon  
et Eric Matzner-Løber

Département MASS  
Université Rennes-2-Haute-Bretagne  
France

### Comité éditorial

#### **Eva Cantoni**

Institut de recherche en statistique  
& Département d'économétrie  
Université de Genève  
Suisse

#### **Rémy Drouilhet**

Laboratoire Jean Kuntzmann  
Université Pierre Mendès France  
Grenoble, France

#### **Ana Karina Fermin Rodriguez**

Laboratoire Modal'X  
Université Paris Ouest Nanterre  
La défense, France

#### **François Husson**

Département Sciences de l'ingénieur  
Agrocampus Ouest  
France

#### **Pierre Lafaye de Micheaux**

Département de Mathématiques  
et Statistique  
Université de Montréal  
Canada

#### **Sébastien Marque**

Président Société Capionis  
Bordeaux  
France

### Déjà paru dans la même collection :

*Psychologie statistique avec R*

Yvonnick Noël, 2015

ISBN : 978-2-7598-1736-8 – EDP Sciences

*Réseaux bayésiens avec R*

Jean-Baptiste Denis, Marco Scutati, 2014

ISBN : 978-2-7598-1198-4 – EDP Sciences

*Analyse factorielle multiple avec R*

Jérôme Pagès, 2013

ISBN : 978-2-7598-0963-9 – EDP Sciences

*Régression avec R*

Pierre-André Cornillon, Eric Matzner-Løber, 2011

ISBN : 978-2-8178-0184-1 – Springer

*Méthodes de Monte-Carlo avec R*

Christian P. Robert, George Casella, 2011

ISBN : 978-2-8178-0181-0 – Springer

**Vj k' r ci g' k p v g p v k p c m ( ' i g h ' d r e p m**

## PREFACE

C'est un réel plaisir de vous inviter à entrer dans le monde des séries chronologiques en utilisant cet excellent livre, écrit par Yves Aragon, professeur émérite à l'université Toulouse 1 Capitole.

Le contenu est présenté de manière efficace et pragmatique, ce qui le rend très accessible non seulement aux chercheurs, mais aussi aux utilisateurs non universitaires. Cela est très important, parce que de tous les côtés de presque tous les océans, il y a une forte demande de praticiens en statistique appliquée (discipline également connue sous le nom d'*Analytics* dans le monde de l'entreprise).

L'esprit du volume est tout à fait celui de la revue *Case Studies in Business, Industry and Government Statistics* (CSBIGS), dont le professeur Aragon est membre de l'équipe éditoriale. Cette revue a été fondée il y a plusieurs années pour répondre à la demande des praticiens ainsi que des chercheurs, pour des cas qui privilégient les approches pratiques et comportent des données permettant de reproduire les analyses. Ce même esprit anime également le programme exceptionnel de statistiques appliquées à l'université de Toulouse 1 Capitole, dirigé par le professeur Christine Thomas-Agnan (coéditeur Europe pour CSBIGS), où le professeur Aragon a enseigné de nombreuses années.

Le choix du logiciel de statistique R, et la fourniture de code R et de données permettant aux lecteurs de s'entraîner, ajoutent encore à l'intérêt de ce livre. R, déjà largement utilisé dans les milieux universitaires, est un outil de plus en plus important pour les praticiens dans le monde de l'entreprise. Dans la formation des étudiants à l'université de Bentley en *Analytics*, l'accent est mis sur le trio SAS<sup>®</sup>, IBM<sup>®</sup> SPSS<sup>®</sup> et R. Les étudiants sont encouragés à prendre un cours pratique de séries chronologiques basé sur R, non seulement parce que l'étude des séries chronologiques est importante, mais parce que les employeurs veulent embaucher des analystes maîtrisant R ; ce langage doit donc être ajouté au *curriculum vitae*. L'ouvrage aborde l'étude de séries temporelles avec une approche claire et logique. Il commence par les moindres carrés ordinaires, en soulignant les limites de la méthode (car les erreurs sont corrélées dans de nombreux cas). Puis il conduit le lecteur vers le modèle ARIMA et ses extensions, y compris les modèles avec hétéroscédasticité conditionnelle. Un aspect particulièrement intéressant du livre est l'utilisation de la simulation comme outil de validation pour les modèles de séries chronologiques.

Le volume présente plusieurs cas fascinants, y compris une étude du trafic de passagers à l'aéroport de Toulouse-Blagnac avant et après les attaques terroristes du 11 septembre 2001.

J'espère qu'il y aura bientôt une traduction en anglais de cet ouvrage, afin que les lecteurs qui n'ont pas de compétences en français, puissent en bénéficier. Ce serait la cerise sur le gâteau.

*Professeur Dominique Haughton,  
Bentley University.*

**Vj k' r ci g' k p v g p v k p c m ( ' i g h ' d r e p m**

## REMERCIEMENTS

Un certain nombre de collègues m'ont apporté une aide décisive dans l'élaboration de cet ouvrage.

Thibault Laurent a prêté attention de manière spontanée et désintéressée à mon travail, et relu minutieusement de larges pans du manuscrit. Il m'a fait mesurer toute la puissance de Sweave et exploiter son automatiser par Make. Plus largement, j'ai bénéficié de sa connaissance étendue et précise de R. Il a amélioré très sensiblement le code des exemples et levé toutes les difficultés de programmation en expert. Il s'est chargé du package et du site du livre. Pour la réédition de 2015, j'ai eu de nouveau recours à son solide savoir et à sa disponibilité pour mettre à jour mes connaissances de l'environnement R.

Nadine Galy, professeur à l'ESC de Toulouse, m'a suggéré des séries financières avec leurs problématiques. Michel Simioni, directeur de recherche à l'INRA, et Anne Vanhems, professeur à l'ESC de Toulouse, ont relu attentivement certains chapitres. Leurs questions m'ont amené à clarifier plusieurs points.

Cet ouvrage est issu d'un cours en master Statistique et économétrie, appuyé sur R pendant quelques années ; des étudiants de ce master, aussi bien en face à face qu'à distance, ainsi que des étudiants inconnus qui ont consulté des parties du cours sur Internet, m'ont signalé des points épineux. J'ai pu ainsi améliorer l'exposé.

Ce travail utilise tantôt des données classiques, ce qui permet au lecteur de comparer notre approche à d'autres démarches, tantôt des données nouvelles. Ces dernières, originales sans être confidentielles, ne sont pas faciles à obtenir. Le SRISE-DRAAF Champagne-Ardenne m'a facilité l'accès à la série sur le vin de Champagne. Le service de la statistique et de la prospective du ministère de l'Alimentation, de l'Agriculture et de la Pêche a mis à ma disposition ses données sur la collecte mensuelle de lait. La chambre régionale de commerce et d'industrie Midi-Pyrénées m'a fourni les données de trafic de l'aéroport de Toulouse-Blagnac. Les relectures et conseils de Pierre-André Cornillon et Eric Matzner-Løber ont largement contribué à rendre le livre plus clair. Charles Ruelle de Springer-Verlag, d'emblée intéressé par cet ouvrage, présent à chaque étape du travail, m'a piloté de bonne grâce.

Que tous ces collaborateurs, connus ou inconnus, qui ont permis de transformer un travail assez solitaire en un travail d'équipe, veuillent bien accepter mes très vifs remerciements.

Bien entendu, je suis seul responsable des fautes et imprécisions qui subsisteraient dans la version finale.

**Vj k' r ci g' k p v g p v k p c m ( ' i g h ' d r e p m**

## AVANT-PROPOS

### **Panorama**

Ce livre est bâti autour de l'étude de quelques séries temporelles régulières, c'est-à-dire de suites d'observations d'un phénomène à des dates régulièrement espacées. Avant d'étudier ces séries, une première partie, les chapitres 1 à 7, est consacrée à quelques rappels sur les méthodes, en particulier sur leur utilisation concrète, avec, souvent, des exemples de mise en pratique dans R.

Les outils de visualisation de série, nombreux dans R, permettent de comprendre la structure d'une série avant toute modélisation ; ils sont présentés dès le chapitre 1 et utilisés systématiquement dans l'étude des séries. Les graphiques sont importants à toutes les étapes du traitement. Avant modélisation, ils aident à saisir la structure de la série ; après modélisation, ils offrent une vision globale de l'ajustement, vision que ne peut donner un niveau de signification empirique considéré isolément.

Quelques éléments sur R pour les séries temporelles sont donnés au chapitre 2, qui concernent principalement les dates et les structures de séries. Mais la lecture de ce chapitre demande une connaissance préalable de R.

Le chapitre 3 est consacré à la régression linéaire. Elle est illustrée par la régression par moindres carrés ordinaires, d'une consommation d'électricité sur des variables de température, sans considération de la nature temporelle des données. Or dans une régression sur séries temporelles, les erreurs sont habituellement autocorrélées...

Pour modéliser les erreurs autocorrélées ou toute série présentant une dynamique, il est indispensable d'avoir des notions sur les modèles ARMA, ARIMA et leurs versions saisonnières. Le chapitre 4 est précisément consacré à des rappels sur les modèles stationnaires, ARMA en particulier. Nous ne traitons que marginalement les méthodes d'estimation ou de prévision, mais insistons sur les concepts qui sont souvent source de confusion. Par exemple, régularité ne veut pas dire stationnarité, comme le montre la série des températures à Nottingham Castle. Nous présentons le test du portmanteau et les principes d'identification du modèle d'une série temporelle. Nous définissons les modèles ARMAX, modèles de régression linéaire d'une série temporelle sur des séries prédéterminées où l'erreur présente une dynamique. Plusieurs séries étudiées dans l'ouvrage peuvent relever d'un tel modèle : le niveau du lac Huron, la température moyenne à Nottingham, la collecte de lait, la consommation d'électricité.

Les modèles non stationnaires, pour cause de tendance déterministe ou pour cause de racine unitaire, sont examinés au chapitre 5, et les tests classiques de l'une ou l'autre situation sont mis en pratique sur des séries classiques ou sur des séries simulées. Dans ces méthodes également, l'exploration de la série doit orienter le champ des questions théoriques qu'elle soulève.

Le lissage exponentiel, chapitre 6, qui s'intéresse à la prévision d'une série plus qu'à sa modélisation, n'est considéré que dans ses modèles les plus simples, mais le traitement retenu passe par le filtre d'innovation ; le lissage échappe ainsi au traitement habituel par bricolage et intuition pure, et gagne une estimation par

maximum de vraisemblance : donc une précision d'estimation.

Le chapitre 7 est consacré à la simulation. Son importance n'est pas à démontrer. D'abord, elle reproduit un mécanisme aléatoire autant de fois qu'on le souhaite. La transposition informatique du modèle d'une série, même si elle se limite à l'utilisation d'une fonction de R, est déjà une façon de vérifier que l'on a compris ce modèle, et sait lire les résultats d'estimation fournis par le logiciel. Ensuite, la simulation pousse un modèle dans ses retranchements : simuler de nombreuses trajectoires d'une série permet de voir comment il se comporte. Appliquée sur le résultat d'une estimation, elle permet de vérifier que des estimations apparemment raisonnables conviennent bien à la série étudiée. Etant donné une série simulée suivant un mécanisme particulier, il est toujours instructif d'estimer sur la série le modèle qui a servi à la simuler pour voir comment se retrouve le modèle initial après simulation et estimation. Autre démarche éclairante : estimer un modèle incorrect, afin de repérer quels mécanismes d'alerte offrent les méthodes, quand on les applique à des séries de manière inappropriée. La simulation est parfois indispensable, par exemple si la série a subi une transformation non linéaire vers une série plus normalement distribuée que l'original. L'effet du modèle doit s'examiner sur la série initiale. Mais la transformation réciproque est une opération le plus souvent très compliquée du point de vue théorique. La simulation, plus simple et plus sûre, permet de la contourner. On a la chance, en séries temporelles, de simuler facilement des modèles très variés sans grand effort d'imagination : cela par le caractère limité des modèles et par la variété des fonctions disponibles pour la simulation.

La deuxième partie de l'ouvrage est consacrée à des études de séries. Pour chaque série étudiée, on se pose un certain nombre de questions et l'on essaie d'y répondre, d'abord par l'exploration graphique puis par la modélisation. Le trafic passager à l'aéroport de Toulouse-Blagnac, chapitre 8, avant et après le 11 septembre, se prête à différentes comparaisons : dynamique de la série, comparaison de l'évolution attendue du trafic annuel en l'absence d'attentats, avec sa réalisation après le 11 septembre. Une série peut relever de plusieurs traitements. La série classique des températures à Nottingham Castle, objet du chapitre 9, en est un exemple : traditionnellement, elle est modélisée par un SARIMA, mais en réalité son exploration suggère de la régresser sur des fonctions trigonométriques et de modéliser la dynamique de l'erreur. La prévision de la consommation d'électricité, chapitre 10, expliquée par certaines fonctions de la température, passe par une modélisation ARMAX et requiert le concours de plusieurs méthodes, chacune présentée dans la première partie. La collecte mensuelle de lait, chapitre 11, a subi l'introduction de la politique des quotas. On essaie d'apprécier les conséquences de cette politique et on compare deux modélisations de la série.

Le traitement des séries présentant une hétéroscédasticité conditionnelle, méthodes et exemples, fait l'objet du chapitre 12. La compréhension de la structure de ces séries permet incidemment de lever des confusions sur la notion de stationnarité. L'estimation et la prévision portent principalement sur la variance de la série et non sur la moyenne, au contraire des séries sans hétéroscédasticité conditionnelle.

L'interprétation des résultats est donc plus délicate que lorsque c'est la moyenne qui est modélisée.

Au cours de l'ouvrage, on pourra constater qu'une série peut recevoir plusieurs modélisations, toutes satisfaisantes et pourtant contradictoires d'un point de vue théorique : tendance stochastique ou déterministe par exemple. Cela peut s'expliquer par au moins deux raisons. D'une part, nous disposons d'un nombre limité de classes de modèles et il n'est pas certain qu'une série donnée ne soit pas mieux modélisée par une autre classe de modèles. D'autre part, nous étudions des séries finies, alors que les modèles concernent souvent des séries infinies. Observons enfin que la façon dont une série est étudiée dépend de nombreux paramètres : la compétence du statisticien, les outils théoriques et pratiques dont il dispose, le temps qu'il peut y consacrer, l'objectif de l'étude, académique, pratique...

### Public

Ce livre pratique entre dans les détails concrets de l'analyse des séries temporelles. Il suppose un minimum de connaissances en statistique mathématique ou en économétrie. Il nécessite également quelques connaissances de base du logiciel R. Cet ouvrage s'adresse à tous les étudiants de masters statistique, économétrie, d'écoles d'ingénieurs mais aussi aux ingénieurs et chercheurs désirant s'initier à l'analyse des séries temporelles.

### Convention

Une fonction est écrite en style *télétype* et avec parenthèses (`mean()`), un package est écrit en **gras** (**forecast**), un objet est écrit en style *télétype* sans parenthèses. Du code R apparaît dans le texte. S'il s'agit de code après exécution, il commence par un chevron `>` et est en style *télétype italique*. Le résultat de ce code, s'il figure dans le livre, apparaît en style *télétype*. Dans ce même style, on rencontrera également du code non exécuté, mais donné comme exemple de syntaxe. On utilise le point et non la virgule dans la représentation des nombres décimaux.

Nous avons utilisé les fonctions `Stangle()` et `Sweave()` pour récupérer le code et les sorties de R ; les erreurs de copier-coller entre le logiciel et le livre devraient ainsi être limitées.

### Ressources

Un package, **caschrono**, installé depuis R par `install.packages("caschrono")`, regroupe les données manipulées ici, ainsi que quelques fonctions souvent appelées dans l'ouvrage. Son chargement provoque celui des packages dont une ou plusieurs fonctions, ou simplement des classes d'objets, peuvent être utilisées par **caschrono**. Les corrigés des exercices ainsi que des compléments sont contenus dans des vignettes associées à certains chapitres ; ainsi `vignette("Anx2")` provoque l'affichage du fichier `.pdf` des compléments du chapitre 2. La liste des vignettes figure en tête de l'aide en ligne de **caschrono**. Le code R nécessaire à la rédaction de l'ouvrage, en particulier le code des graphiques, est disponible sur le site de l'ouvrage<sup>1</sup>. Les appels à la fonction `par()`, nécessaires pour la mise en page de l'ouvrage, ne

---

1. <http://seriestemporelles.com/>

sont pas reproduits dans le livre mais figurent sur le site. On trouve également sur le site du code R complémentaire signalé dans ce travail par le sigle SiteST. Les packages appelés hors de **caschrono** sont : **car**, **chron**, **dse**, **dynlm**, **expsmooth**, **fBasics**, **fGarch**, **FinTS**, **FitARMA**, **nlme**, **polynom**, **TSA**, **urca**, ainsi que **xtable** qui ne sert qu'à l'édition en LaTeX. Il faut évidemment les installer pour traiter les exemples.

Ceux chargés par **caschrono** sont : **forecast**, **Hmisc**, **its**, **timeSeries**.

### Orientations bibliographiques

Ruppert (2004) présente de façon très accessible et soignée des rappels de statistique mathématique et la plupart des méthodes manipulées dans ce livre, avec une orientation financière. Brocklebank & Dickey (2003) ont une approche concrète et fine des séries temporelles et n'utilisent qu'un minimum de bases mathématiques; cet ouvrage illustre l'emploi du logiciel SAS/ETS. Le présent livre ne dépasse pas leur niveau théorique. Zivot & Wang (2006), pour les séries financières, mêlent efficacement considérations théoriques et pratiques. L'approche de Tsay (2005) pour les modèles ARMA, surtout intuitive, correspond aussi à l'esprit du présent ouvrage. Pour les bases théoriques, on peut se reporter à de nombreux ouvrages, notamment : Bourbonnais & Terraza (2008), Brockwell & Davis (2002), Franses (1998), Hamilton (1994), Wei (2006). Les travaux de Cryer & Chan (2008), Pankratz (1991) et Wei, déjà cité, contiennent, en plus d'un exposé théorique, des exemples numériques fouillés. Shumway & Stoffer (2006) donnent un solide exposé théorique contemporain accompagné d'exemples réalistes tirés de domaines variés. Le manuel en ligne de méthodes statistiques NIST/SEMATECH (2006), pragmatique, s'avère utile pour réviser les bases de la statistique et même pour approfondir certaines questions. L'ensemble des packages de R consacrés aux séries temporelles est présenté par classe de problèmes dans le panorama <http://cran.r-project.org/web/views/TimeSeries.html> régulièrement mis à jour.

Les méthodes manipulées ici sont classiques et largement pratiquées; des travaux de recherche ont montré les limites de certaines d'entre elles. Le lecteur qui souhaite prendre du recul sur ces méthodes consultera avec profit Kennedy (2003) et sa bibliographie.

### Comment utiliser ce livre ?

Il faut d'abord installer **caschrono**, les packages utilisés hors **caschrono** et récupérer le code sur le site. Cela fait, la meilleure façon de travailler est d'exécuter le code pas à pas, et même ligne à ligne pour les graphiques, en vue de s'approprier les commentaires et conclusions, quitte à les contester ! Il est très formateur de simuler des séries puis de leur appliquer les techniques d'exploration et de modélisation.

### A propos de la présente édition 2015

Les principales différences avec l'édition parue en 2011 chez Springer proviennent des changements dans la fonction `auto.arima()` de **forecast**. En 2011, les calculs ont été effectués avec la version 2.16 de **forecast**, dans l'édition actuelle, avec la version 6.1. Si pour un modèle et un jeu de données, les estimations des coeffi-

cients coïncident dans les deux éditions, les estimations de variance peuvent différer. Ces variations sont dues aux changements (améliorations) des algorithmes d'optimisation utilisés dans la fonction `Arima()` de **forecast**. D'autre part, la log-vraisemblance est souvent numériquement mal conditionnée et la recherche de son maximum est alors délicate, comme l'illustre le tableau 4.1. Ces variations incitent donc à arrondir les estimations des coefficients, opération qui se traite comme un test d'hypothèse sur des coefficients d'une régression, voir l'entrée *test sur les coefficients d'une régression* dans l'index. L'algorithme de recherche de modèle de `auto.arima()` a été amélioré, donnant parfois des résultats plus convaincants que ceux de 2011.

**Vj k' r ci g' k p v g p v k p c m ( ' i g h ' d r e p m**

# Table des matières

Préface	vii
Remerciements	ix
Avant-propos	xi
<b>1 Démarche de base en séries temporelles</b>	<b>1</b>
1.1 Exemples de séries temporelles	1
1.2 Graphiques pour les séries temporelles	10
1.2.1 Chronogramme	10
1.2.2 Lag plot	10
1.2.3 Month plot	14
1.3 Tendance, saisonnalité, résidus	16
1.4 Etapes et objectifs de l'analyse d'une série temporelle	17
1.5 Opérateur retard	19
<b>2 R pour les séries temporelles</b>	<b>21</b>
2.1 Dates et heures	22
2.1.1 Considérations générales	22
2.1.2 Dates et heures en R	23
2.2 Les structures de séries temporelles dans R	28
2.2.1 La fonction <code>ts()</code>	28
2.2.2 Récupération de données boursières et séries de classe <code>its</code>	32
2.3 Série de classe <code>ts</code> : retard, différence, union, intersection, sous-série	33
2.4 Traitement des manquants	34
<b>3 Régression linéaire par la méthode des moindres carrés</b>	<b>39</b>
3.1 Principe de la régression linéaire	39
3.2 Significativité de la régression	42
3.3 Comparaison de modèles et critères d'information	44
3.4 Intervalle de confiance (IC)	45
3.5 Prédiction	45
3.6 Exemple : consommation d'électricité	47

<b>4</b>	<b>Modèles de base en séries temporelles</b>	<b>59</b>
4.1	Stationnarité . . . . .	59
4.1.1	Fonction d'autocorrélation d'une série stationnaire . . . . .	60
4.1.2	Bruit blanc . . . . .	62
4.2	Série linéaire . . . . .	66
4.3	Fonctions d'autocorrélation . . . . .	71
4.3.1	Fonction d'autocorrélation d'un AR . . . . .	71
4.3.2	Fonction d'autocorrélation d'un MA . . . . .	72
4.4	Prévision . . . . .	75
4.4.1	Principe . . . . .	75
4.4.2	Fonction d'autocorrélation partielle . . . . .	76
4.4.3	Prévision d'un modèle autorégressif . . . . .	78
4.4.4	Prévision d'un MA( $q$ ) . . . . .	79
4.5	Estimation . . . . .	80
4.5.1	Exemples . . . . .	82
4.5.2	Modèle ARMA saisonnier (modèle SARMA) . . . . .	85
4.5.3	Modèle ARMAX . . . . .	88
4.6	Construction d'un ARMA ou d'un SARMA . . . . .	92
4.6.1	Identification d'un ARMA . . . . .	92
4.6.2	La méthode MINIC . . . . .	95
4.7	Exercices . . . . .	96
<b>5</b>	<b>Séries temporelles non stationnaires</b>	<b>99</b>
5.1	Séries intégrées – Modèles ARIMA et SARIMA . . . . .	100
5.2	Construction d'un modèle SARIMA . . . . .	105
5.3	Non-stationnarité stochastique ou déterministe . . . . .	106
5.3.1	Test de non-stationnarité : introduction et pratique . . . . .	106
5.3.2	Test de stationnarité à une tendance déterministe près . . . . .	113
5.4	Significativité illusoire en régression . . . . .	118
<b>6</b>	<b>Lissage exponentiel</b>	<b>123</b>
6.1	Lissage exponentiel simple . . . . .	123
6.2	Lissage exponentiel double . . . . .	129
6.3	Méthode de Holt-Winters et modèle de lissage correspondant . . . . .	132
<b>7</b>	<b>Simulation</b>	<b>135</b>
7.1	Principe . . . . .	135
7.2	Simulation de séries temporelles . . . . .	136
7.2.1	Principe . . . . .	136
7.2.2	Illustration numérique . . . . .	137
7.3	Construction de séries autorégressives . . . . .	142
7.4	Construction de séries subissant une intervention . . . . .	143
7.4.1	Réponses typiques à une intervention . . . . .	143
7.4.2	Simulation d'une intervention . . . . .	145

7.4.3	Estimation d'une intervention . . . . .	148
<b>8</b>	<b>Trafic mensuel de l'aéroport de Toulouse-Blagnac</b>	<b>149</b>
8.1	Préparation des données . . . . .	149
8.2	Exploration . . . . .	151
8.2.1	Décomposition de la série en tendance, saisonnalité et erreur	151
8.2.2	Month plot . . . . .	152
8.2.3	Lag plot . . . . .	153
8.3	Modélisation avant septembre 2001 . . . . .	155
8.3.1	Modélisation manuelle . . . . .	157
8.3.2	Modélisation automatique . . . . .	160
8.4	Impact sur le volume du trafic . . . . .	163
8.4.1	Prévision ponctuelle . . . . .	163
8.4.2	Simulation de trajectoires . . . . .	164
8.5	Etude après le 9/11 — Lissage exponentiel . . . . .	167
8.6	Estimation d'un SARIMA dans R — Vérification . . . . .	171
<b>9</b>	<b>Température mensuelle moyenne à Nottingham</b>	<b>173</b>
9.1	Exploration . . . . .	173
9.2	Modélisation . . . . .	174
9.2.1	Modèle SARIMA . . . . .	174
9.2.2	Régression sur fonctions trigonométriques . . . . .	175
9.3	Prévision . . . . .	181
9.4	Comparaison . . . . .	182
9.5	Analyse spectrale . . . . .	186
<b>10</b>	<b>Consommation d'électricité</b>	<b>189</b>
10.1	Identification de la série des résidus obtenus par MCO . . . . .	189
10.2	Estimation du modèle ARMAX . . . . .	193
10.3	Estimation d'un modèle à erreur non stationnaire — Modèle ARIMAX	197
10.4	Prévision de l'année 1984 . . . . .	200
10.5	Prédiction sur la série non transformée . . . . .	203
<b>11</b>	<b>Production de lait</b>	<b>205</b>
11.1	Analyse exploratoire . . . . .	205
11.2	Modélisation avant 1984 . . . . .	210
11.3	Modélisation après l'introduction des quotas . . . . .	214
11.4	Modélisation SARIMA de toute la série . . . . .	214
11.5	Modélisation ARMAX de la collecte . . . . .	217
11.5.1	Modélisation MCO . . . . .	217
11.5.2	Identification des résidus de l'ajustement MCO . . . . .	218
11.5.3	Modélisation simultanée de la moyenne et de l'erreur . . . . .	221

**Vj k' r ci g' k p v g p v k p c m ( ' i g h ' d r e p m**

# Notations

$BB(0, \sigma^2)$  Bruit blanc de variance  $\sigma^2$ , page 62

$BBN(0, \sigma^2)$  Bruit blanc gaussien de variance  $\sigma^2$ , page 62

$\chi_k^2$  Loi de  $\chi^2$  à  $k$  ddl, page 63

ddl Degrés de liberté, page 43

$\mathcal{F}(p, q)$  Loi de Fisher à  $p$  et  $q$  ddl, page 43

i.i.d. (Variables aléatoires) indépendantes identiquement distribuées, page 9

$\mathcal{N}(\mu, \sigma^2)$  Loi normale d'espérance  $\mu$ , de variance  $\sigma^2$ , page 40

B Opérateur retard, page 19

L Opérateur retard, page 19

$\mathcal{T}(k)$  Loi de Student à  $k$  ddl, page 43

$X'$  Transposée de la matrice  $X$ , page 40