

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université Abderahmane MIRA - Béjaïa

Faculté des Sciences Exactes
Département de Recherche Opérationnelle



Mémoire de Fin D'étude

En vue de l'Obtention du Diplôme Master en Recherche Opérationnelle
Option :

Modélisation Mathématique et Techniques de Décisions

THÈME

Modélisation et prévision de la pluviométrie. Apport et Transferts des Eaux cas : Barrage Béni Haroun



Présenté par : ZOGHBI Fatima & AMIRA Sarra

Devant le jury composé de :

Président	M ^{elle} N. Saadi	M.A.B.	Université de Béjaïa.
Rapporteur	M ^r D. Aissani	Professeur	Université de Béjaïa.
Rapporteur	M ^{elle} Y. Ziane	Docteur	Université de Béjaïa.
Examineur	M ^{elle} S. Amroun	Magister	Université de Béjaïa.
Examinatrice	M ^{elle} N. Belaid	Doctorante	Université de Béjaïa.
Invité	M ^r Lemanaa Azzedine	Directeur	Barrage Béni Haroun.

Juillet 2016

** Remerciements **

Nous tenons à remercier Dieu tous puissant qui nous a procuré le courage, la patience et la volonté durant la réalisation de ce travail.

Un grand merci pour nos familles, surtout nos parents qui nous ont épaulés, soutenus et suivis, qui nous ont toujours orientés vers le droit chemin.

Nous tenons à remercier M^r le professeur **D. Aissani** qui nous a fait l'honneur d'être notre promoteur, d'avoir proposé ce travail, Son soutien, sa patience et ses conseils qui nous ont beaucoup aidés dans la réalisation et la finalisation de ce travail.

Nous tenons à remercier aussi notre promotrice M^{lle} **Y.Ziane** pour son encouragement son aide , ses précieux conseil et son suivi attentif pour la réalisation de ce travail.

Nous remercions également tout le personnel de la Direction du Barrage de Béni Haroun pour leurs accueils à leur tête monsieur le directeur du barrage **Lemanaa Azzedine** ainsi notre encadreur M^r **Boudjefdjouf Hamza** pour le temps qu'il nous a consacré tout au long de notre stage.

Nous remercions vivement tous les membres de jury d'avoir accepter d'évaluer et de juger notre mémoire à sa juste valeur.

Nous remercions tous les enseignants qui ont contribué à notre formation.

Enfin, à tous ce qui nous ont soutenus de près et de loin, et tous ceux qui nous ont incitées même involontairement à faire mieux, veuillez trouver ici le témoignage de notre profonde gratitude.

※ *Dédicaces* ※

Louange a Dieu, le miséricordieux, sans lui rien de tout cela n'aurait pu être

Je dédie ce modeste travail aux êtres les plus chère de ma vie qui ont donné sens a mon existence :

Au témoignage d'affection, d'amour et de grande reconnaissances , à celle qui ma donné la vie, le symbole de tendresse, qui s'est sacrifiée pour mon bonheur, l'école de mon enfance, qui a été mon ombre durant toutes les années d'études et qui a veillé tout au long de ma vie à m'encourager et à me protéger. à ma mère. . .

À mon père, pour son soutiens ; conseils et son amour, Que dieu me les protège et les gardes.

Maman, papa, Je tiens la parole pour vous dire :
” VOILA LES FRUITS DE VOS SACRIFICES ”

Ainsi que mes frères et sœurs qui m'ont soutenu durant ce travail. Que dieu vous protège . . .Merci sans vous je ne serai rien, vous faites ma fierté et ma force de vivre.

A la mémoire de mes grands parents, mes oncles et tantes, ainsi qu'à mes adorable ami(e)s.

A tous ceux ou celles qui ont contribué de près ou de loin a la réalisation de ce modeste travail.

Merci infiniment

Fatima

※ *Dédicaces* ※

Tout au début, je tiens à remercier le Dieu le plus puissant pour m'avoir donné la force et la patience.

Je dédie ce modeste travail :

A mes deux plus chers sur terre : mes parents, merci pour les sacrifices le long de ces années car je ne pourrai oublier la tendresse, l'amour dévoué par lesquels ils m'ont toujours entourés depuis mon enfance puisse Dieu vous préserver.

A mes chers frères : Nadjib, Badr Eddine.

A mes chères sœurs : Imane, Linda et Nadjwa.

A l'esprit de mon cher frère : Bassem.

A toute ma famille oncles, tantes, cousins, coisines.

A mes amies : Ilham, Omria, Fatima, Nada, Malika, Amina.

Et tous ceux que j'aime et qui m'aiment.

Sarra

Table des matières

Table des matières	i
Table des Figures	v
Liste des tableaux	vi
Listes des abréviations	vii
Introduction Générale	1
1 Présentation du barrage Béni Haroun	4
1.1 Introduction	4
1.2 Historique	4
1.3 Localisation	5
1.4 Principales Caractéristiques du barrage BCR	7
1.5 Caractéristiques principales de l'aménagement	10
1.6 Présentation générale de l'aménagement	11
1.6.1 Barrage de Béni Haroun-Volumes Mobilisable	11
1.6.2 Le transfert de Béni Haroun	11
1.7 Caractéristiques générales du bassin	13
1.8 Présentation de la direction du barrage	15
1.8.1 Direction	15
1.8.2 La station météorologique	17
1.8.3 Salle de contrôle :	20
1.9 Position du problème	21
1.10 Conclusion	21
2 Méthodes statistiques de la prévision	23
2.1 Introduction	23

2.2	Les séries chronologiques	24
2.2.1	Définition et exemple :	24
2.2.2	Les objectifs principaux :	24
2.2.3	Les composantes d'une série chronologique	25
2.3	Quelques définitions :	25
2.4	Modélisation d'une série chronologique	28
2.5	Lissage Exponentiel	29
2.5.1	Lissage Exponentiel simple (LES) :	30
2.5.2	Le lissage exponentiel double (LED)	31
2.5.3	Méthode de Holt Winters	33
2.6	Méthode de Box et Jenkins	36
2.6.1	Fonction autocovariance et autocorrélation :	37
2.6.2	Les processus autorégressifs d'ordre p : $AR(p)$	38
2.6.3	Le processus $AR(1)$	38
2.6.4	Les processus de moyennes mobiles : $MA(q)$	40
2.6.5	Le processus $MA(1)$	41
2.6.6	Les processus autorégressifs et de moyennes mobiles : $ARMA(p, q)$	42
2.6.7	Les processus $ARIMA$	44
2.6.8	Les processus $SARIMA$	44
2.6.9	Estimation des paramètres	46
2.6.10	Validation du Modèle	47
2.6.11	Prévision	49
2.7	Conclusion	51
3	Application de la méthode de Box-Jenkins et du lissage exponentiel	52
3.1	Introduction	52
3.2	Présentation du logiciel R	52
3.3	Modélisation et prévision	54
3.4	Processus de prévision	54
3.5	Collecte des données	54
3.6	Application de la méthode du lissage exponentiel	55
3.7	Application de la méthode de Box et Jenkins	61
3.7.1	Interprétation des résultats obtenus par les deux méthodes	69
3.8	Conclusion	69

4 Objectifs des prévisions	71
4.1 Introduction	71
4.2 Importance des prévisions	71
4.3 Horizon des prévisions	72
4.4 Caractéristiques des prévisions	72
4.5 Les caractéristiques climatiques	73
4.5.1 Les précipitations	73
4.6 L'objectif du travail	75
4.7 Conclusion	79
Conclusion générale	81
Bibliographie	82

Table des figures

1.1	Plan de localisation	6
1.2	Coupe évacuateur de crues	7
1.3	Coupe plot non déversant	8
1.4	Allocation des ressources du transfert à l’horizon 2020	13
1.5	Caractéristiques générales du bassin	14
1.6	Superficies drainées aux points caractéristiques du bassin	15
1.7	Organigramme type de Barrage	16
1.8	Automatique et Manuelle.	17
1.9	Le bac Colorado	18
1.10	Diagramme d’un bac Colorado	18
1.11	Appareil pour mesurer la vitesse dU vent	19
1.12	Tableau synoptique	20
1.13	Exemple d’une feuille de calcul	21
2.1	Organigramme de la démarche Box et Jenkins	46
3.1	Graphe de la série originale	56
3.2	Graphe de la série originale ainsi que ses différentes composantes	57
3.3	La série originale et son ajustement	59
3.4	Graphe de la série originale et sa prévision	61
3.5	Corrélogramme et corrélogramme partiel de la série originale.	62
3.6	Graphe de la série après la differenciation	63
3.7	Corrélogramme et corrélogramme partiel de la série transformé.	64
3.8	Corrélogramme et corrélogramme partiel du résidu.	67
3.9	Graphe de la série originale et sa prévision	68
4.1	Variations interannuelles des précipitations 2010/2015	74
4.2	Variations interannuelles des précipitations et précipitations prévues en 2016	75

4.3 Objectifs et Défis 77

Liste des tableaux

2.1	Différentes méthodes de lissage exponentiel	30
2.2	Méthode de Holt-Winters additive et multiplicative	35
2.3	Résumé des propriétés des fonctions d'autocorrélations simples et partielles	43
3.1	La pluviométrie en millimètres	55
3.2	Les coefficients saisonniers	58
3.3	Prévisions pour l'année 2016	60
3.4	Prévisions pour l'année 2016	68
3.5	Tableau de comparaison	69
4.1	Les quantités d'AEP transférés aux cinq wilaya	76
4.2	Les quantités d'eau d'irrigation transférés	76

Listes des abréviations

BCR	: Béton Compacté au Rouleau
RN	: Retenu Normale
PHE	: Plus Haut Eau
CMP	: Crue Maximale Probale
LES	: Lissage Exponentiel Simple.
LED	: Lissage Exponentiel Double.
AR	: AutoRegressif.
MA	: Moving Average.
ARMA	: AutoRegressif Moving Average.
ARIMA	: AutoRegressif Integrated Moving Average.
SARIMA	: Seasonal AutoRegressif Integrated Moving Average.
ACF	: AutoCorrelation Function.
PACF	: Partial AutoCorrelation Function.
QDt	: Quantité D'eau transféré.
BQD	: Besoin en Quantité D'eau.
QDT	: Quantité D'eau Total.
QDB_I	: Quantité D'eau Initial dans le Barrage.
QDB_F	: Quantité D'eau Final dans le Barrage.
VTM	: Volume de la Tranche Morte.

Introduction Générale

La recherche opérationnelle est une discipline dont le but est de fournir des méthodes pour répondre à un type précis de problème, c'est-à-dire à élaborer une démarche universelle pour un type de problème qui aboutit là où les solutions sont les plus efficaces.

La particularité de la recherche opérationnelle est que les méthodes proposées sont des démarches rationnelles basées sur des concepts et outils mathématiques et/ou statistiques. Généralement, ces méthodes sont employées sur des problèmes tels que leur utilisation "manuelle" devient impossible. C'est pourquoi, du fait qu'elles sont rationnelles, les démarches proposées par la recherche opérationnelle peuvent être traduites en programmes informatiques. Dès qu'un phénomène, qu'il soit physique, biologique ou autre, est trop complexe ou encore trop bruyant pour accéder à une description analytique débouchant sur une modélisation déterministe, un ensemble d'approches ont été élaborées afin d'en décrire au mieux le comportement à partir d'une série d'observations.

Citons la reconnaissance de la parole ou de caractères manuscrits, l'imagerie médicale ou satellitaire, la prévision d'une grandeur climatique ou économique, du comportement d'un client.... La plupart des disciplines scientifiques sont concernées. Historiquement, la Statistique s'est beaucoup développée autour de ce type de problèmes et a proposé des modèles incorporant d'une part des variables explicatives et, d'autre part, une composante aléatoire ou bruit. Il s'agit alors d'estimer les paramètres du modèle à partir des observations.

La statistique est une science qui permet l'exploitation d'une masse d'informations provenant de l'observation d'un phénomène. Elle concerne un ensemble de méthodes mathématiques indispensables à l'analyse de grands ensembles de données et dont l'objectif est de fournir les éléments nécessaires à l'appréciation d'une situation et l'aide à la décision.

Il existe de nos jours un ensemble de méthodes rigoureuses, basées sur des algorithmes,

permettant de faire des prévisions, notamment sur les séries temporelles. La qualité d'une prévision dépend en grande partie du choix porté sur l'une ou l'autre de ces méthodes.

L'étude des séries temporelles correspond à l'analyse statistique d'observations régulièrement espacées dans le temps. Cette étude est appliquée dans plusieurs domaines, en météorologie, en économie, etc. Elle consiste à appliquer les méthodes de prévision. On distingue notamment, deux méthodes qui sont souvent utilisées, la méthode de Box et Jenkins (1976) et le lissage exponentiel. La première est très populaire grâce à sa simplicité et ses résultats fiables. La deuxième méthode a l'avantage de fournir une prévision "bon marché" et parfois très satisfaisante, mais elles sont parfois loin d'être les mieux adaptées.

Afin d'avoir une idée précise et claire de la manière de la mise en œuvre des outils statistiques, nous avons effectué un stage pratique au sein du Barrage de Béni Haroun, où nous avons pu mettre en pratique nos connaissances théoriques acquises en la matière. Nous nous sommes intéressées à la modélisation des changements climatiques liés aux quantités pluviométriques obtenues en millimètres par mois sur la période allant de l'année 2010 jusqu'à l'année 2015. Cette modélisation nous servira à établir des prévisions à court et à moyen terme.

Le rôle de n'importe quel barrage est de satisfaire un ou plusieurs besoins simultanés (irrigation, alimentation en eau potable, production d'énergie électrique, ...) et de protéger l'environnement des phénomènes naturels tel que les crues,

Actuellement, le barrage de Béni Haroun alimente deux régions pour l'irrigation et deux Wilaya en eau potable. Et il est prévu d'alimenter trois autres wilaya.

Notre objectif est de répondre à certaines problématiques rencontrées au niveau de ce barrage. Il s'agit notamment de vérifier si les apports pluviométriques peuvent compenser les quantités transférées. A cet effet, nous avons réalisé une étude de prévision.

Afin d'atteindre nos objectifs, nous avons structuré notre travail en quatre chapitres : Le premier chapitre est consacré à la présentation du Barrage de Béni Haroun où nous avons effectué notre stage, en retraçant son historique, sa localisation et ses principales caractéristiques ainsi que les différentes tâches des services de la direction, en particulier le regroupement du paramètre de la station climatologique.

Le deuxième chapitre portera sur le cadre général d'une série chronologique, ainsi que les éléments théoriques indispensables à l'élaboration des modèles de prévision, en exposant tous les outils statistiques qui vont nous permettre de réaliser notre étude empirique. Dans ce cadre, la méthode de Box et Jenkins et de lissage exponentiel sera présentée.

Dans le troisième chapitre, on a appliqué les méthodes citées dans le chapitre précédent sous le logiciel R ainsi que l'interprétation des résultats.

Dans le quatrième chapitre, nous avons considéré le problème des qualités transférées et renouvelées.

Notre mémoire se termine par une conclusion.

Chapitre 1

Présentation du barrage Béni Haroun

1.1 Introduction

La région de l'Algérie est caractérisée par le fait que, dans la tranche nord qui présente une largeur de 100 km environ, la demande en eau est faible alors que les ressources y sont importantes. Par contre, plus au sud, sur les hauts plateaux, les ressources sont réduites alors que la demande en eau potable est en forte croissance et qu'on y trouve de grandes superficies agricoles à irriguer.

Un grand projet de transfert a donc été élaboré au début des années 1970. Celui-ci prévoit la mobilisation de plus de 500 millions de m³ dans la retenue de Béni Haroun et leur refoulement sur les hauts plateaux du sud. La retenue est créée par la réalisation du barrage de Béni Haroun dont la hauteur est voisine de 120m.

Le barrage de Béni Haroun constitue donc un élément clé du transfert d'eau en Algérie. Sa réalisation qui a débuté dans la seconde moitié des années 1980 (1988) s'est achevée en 2001 après divers aléas.

1.2 Historique

La réalisation du barrage de Béni Haroun a été particulièrement longue est perturbée par de nombreuses péripéties et retournement de situation. Elle s'est déroulée entre 1986 et 2001.

Au cours de cette période, la réalisation a démarré sur base d'un projet de digue zonée en enrochement. Celle-ci présentait un volume total voisin de 10 millions de m³ constitué

de 95% d'énrochement et de 5% de béton. Elle s'est achevée par la réalisation d'un barrage en béton compacté au rouleau (BCR) présentant un volume global de 1.700.000 m^3 .

1.3 Localisation

Le site de barrage de Béni Haroun dans le nord-est de l'Algérie sur l'Oued Kebir à une quarantaine de km de son embouchure dans la mer méditerranée. La retenue créée par le barrage se trouve au sud de ce dernier. Elle est située à Mila et à 350 Km à l'est d'Alger.

Le barrage est situé à l'extrémité amont des gorges calcaro-marneuses de Béni Haroun et à environ 4 Km du confluent de l'Oued Rhumel et de l'Oued Endja.

Le bassin total de Oued Kebir au site du barrage couvre une superficie voisine de 7.725 Km^2 . Tenant compte qu'une partie du bassin du Rhumel est mobilisée par le barrage de Hammam Grouz, la superficie intéressant du barrage de Béni Haroun est de 6.595 Km^2 . L'apport annuel régularisé est estimé à 435 millions de m^3 .

Le volume total de la retenue est plus d'un milliard de m^3 . Le volume de la tranche utile entre les niveaux 172 et 200 est de 723 millions de m^3 .

PLAN DE LOCALISATION



FIGURE 1.1 – Plan de localisation

1.4 Principales Caractéristiques du barrage BCR

- Le barrage est de type poids rectiligne en béton compacté au rouleau s'appuyant sur une fondation rocheuse calcaire.
- Le niveau de retenue normale (RN) qui est aussi la crête de l'évacuateur à seuil libre, est à la cote 200.
- Il correspond un ouvrage d'une hauteur maximale sur fondation de 118 m et une longueur en crête de 710 m.

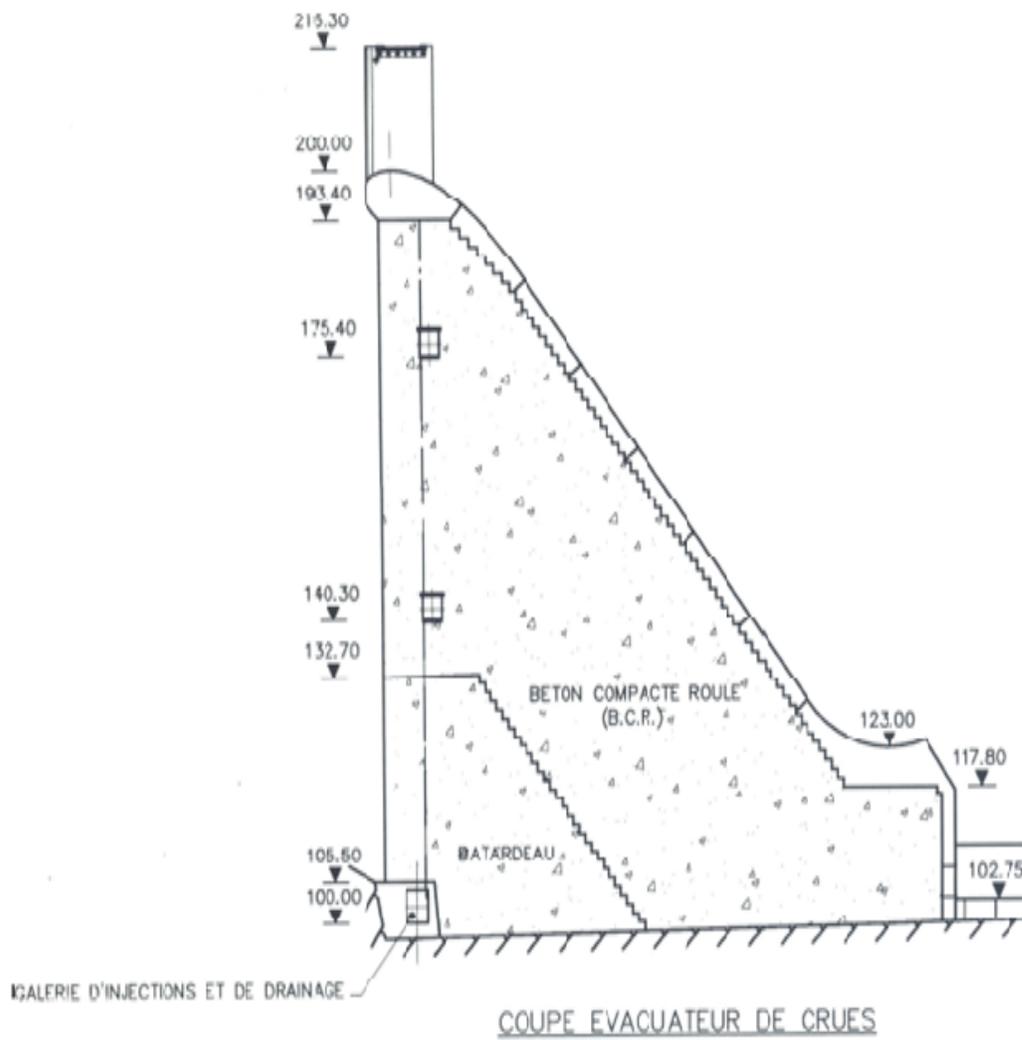


FIGURE 1.2 – Coupe évacuateur de crues

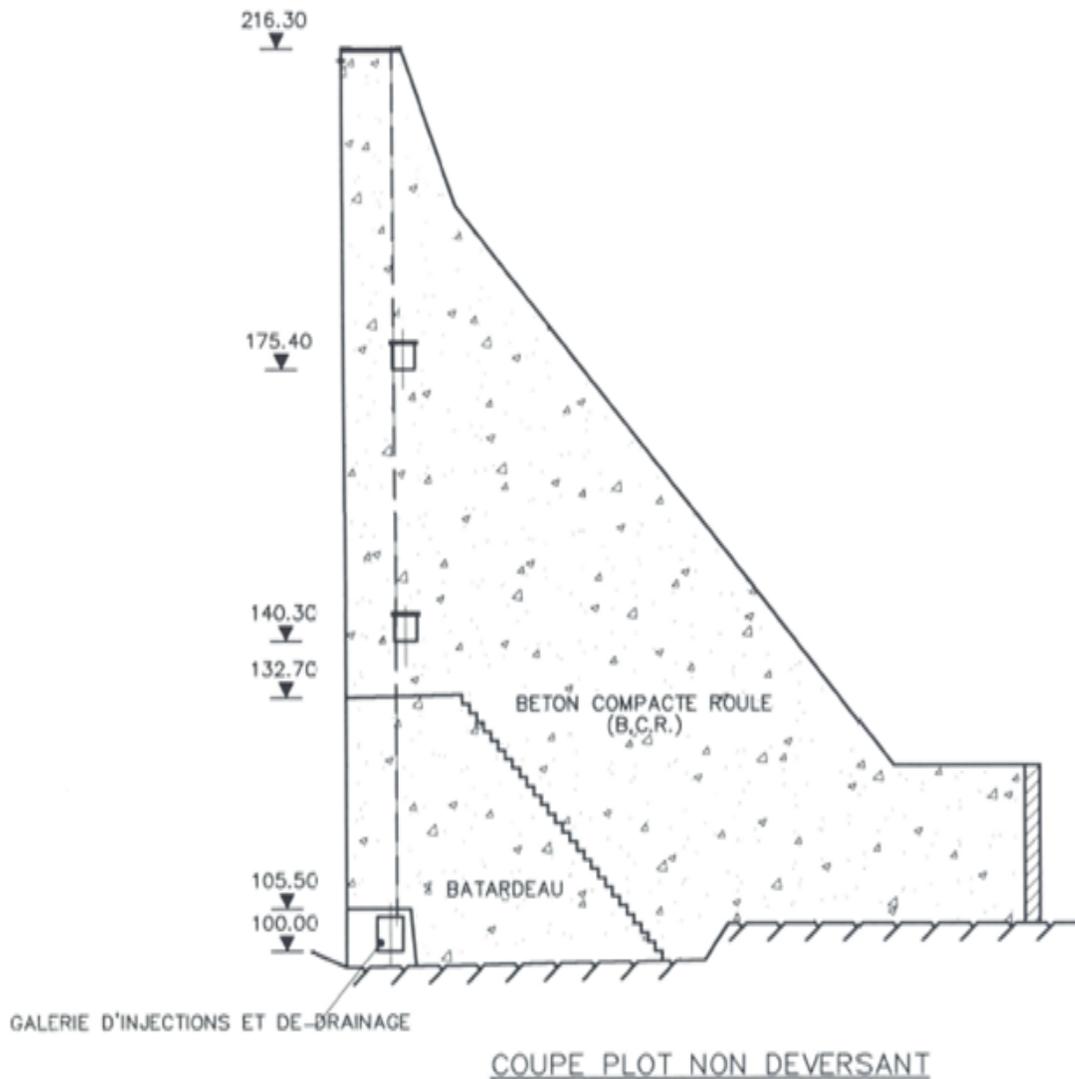


FIGURE 1.3 – Coupe plot non déversant

- Le profil de base du barrage est un triangle de caractéristiques ci-après :
 - Sommet à la cote 214.80, niveau maximum de crue (P.H.E)
 - Parement amont vertical
 - Parement aval incliné à 0.8 H pour 1 V
- La section non déversant possède une crête d'une largeur de 8m, arasée à la cote 216.30 m, ce qui laisse une revanche de 1.5 m au-dessus des plus hautes eaux.
- Le parement aval de cette section change de pente au-dessus de la cote 196, ce qui renforce la résistance aux séismes et permet l'installation d'une voie de circulation.
- L'évacuateur de surface, de type "à seuil libre", est obtenu à peu de frais, comme

dans tout barrage en béton, en aménageant en déversoir la partie centrale du barrage. Sa longueur est de 124 répartie 6 passes surmontées par un pont route.

- Le seuil de l'évacuateur est prolongé par un coursier raccorde à un saut de ski, calé au-dessus des plus hautes eaux. La capacité de l'évacuateur est de $3.230 \text{ m}^3/\text{sec}$, pour un niveau d'eau de 214.74, ce qui correspond au débit maximal restitué après laminage de la crue maximum probable (CMP) de $16.640 \text{ m}^3/\text{sec}$
- La fond de la vallée est élargi en entaillant la rive droite sur une centaine de mètre afin d'assurer une bonne assise de la zone du barrage concernée par l'évacuateur de crue et la vidange de fond.
- La vidange de fond contiguë à l'évacuateur est implantée en rive droite au niveau 140. Elle est constituée de deux pertuis dimensionnés pour évacuer un débit de $670 \text{ m}^3/\text{sec}$ (335 m^3 par pertuis) sous le niveau de retenue normale et se terminant par un saut de ski.
- Durant les travaux, la dérivation de la rivière a été assurée en saison sèche (mai - octobre) par les deux galeries de dérivation de la rive gauche prévues dans la solution de base en enrochement. Pendant la saison des pluies, ces galeries ont également été suffisantes et le passage éventuel des plus fortes crues, bien que possible par déversement par-dessus la partie centrale du barrage, n'a pas nécessité.
- Trois galeries, réparties sur 3 niveaux et situées à proximité du parement amont du barrage, permettent la réalisation et le contrôle des opérations d'injection et de drainage dans le corps du barrage et dans la fondation. Calées respectivement aux cotes 100 m, 140 m et 175,10 m, elles sont implantées à une distance du parement amont égale à 4,00 m pour la galerie inférieure et 6,50 m pour les deux supérieures. Ces galeries se prolongent dans les appuis sur des distances variant de 70 à 130 m. Une galerie en rocher complémentaire, au niveau 120.00, renforce le drainage de la rive droite.

Le système de galeries décrit ci-dessus permet de collecter les eaux de drainage provenant du corps du barrage, de la fondation et des versants. Il permet également le contrôle du comportement de l'ouvrage et les opérations de maintenance et de réparation éventuellement nécessaires.

Sous le barrage, l'écran d'étanchéité est constitué d'un rideau d'injection multifilaire (en principe 3 files) d'une profondeur de 40 à 120 m. des contrôles plus profonds ont également été effectués.

Des injections de consolidation du rocher de fondation ont été effectuées sur toute l'emprise du barrage sur une profondeur de 8 m et suivant une maille carrée de 4 m de côté.

Un rideau de drainage d'une profondeur maximum de 40 m est réalisé à partir de la galerie de pied amont.

Les agrégats pour le BCR proviennent du traitement d'alluvions de l'oued proche du barrage.

1.5 Caractéristiques principales de l'aménagement

Retenu

Niveau normale (RN) : 200,00

Niveau des plus hautes eaux (PHE) : 214 ,80

Niveau minimum d'exploitation : 172,00

Volume de la tranche morte (172-200) : 240 106 m^3

Volume utile (172-200) : 723 106 m^3

Volume totale (110-200) : 963 106 m^3

Barrage

Type : barrage poids en béton compacté au rouleau (BCR)

Hauteur maximale :

- hauteur au dessus du terrain naturel : 107 m
- hauteur au dessus des fondations : 118 m

Longueur en crête : 710 m

Largeur en crête : 8 m

Largeur maximale au niveau des fondations : 93 m

Fruit des parements amont : verticale

Fruit du parement aval : 08H/1V

Cote de la crête : 216,30

Evacuateur de crue

Type : déversoir de surface à seuil libre - coursier et saut de ski

Localisation : partie centrale du barrage

Cote de la crête : 200 m

Cote du bec aval du saut de ski : variable de 124,00 à 126,00

Longueur totale de l'évacuateur en crête : 124 m

Longueur utile de l'évacuateur en crête : 114 m répartie en 6 passes de 9 m

Longueur de la cuillère : 130 m

Débit maximale : 13230 m^3/sec pour PHE à 214, 74

Vidange de demi-fond

Type : 2 pertuis blindés dans corps du barrage avec saut de ski

Localisation : rive droite

Calage de l'entrée : 140,0

Section des pertuis : (3,0 m * 4,10 m) * 2

Vannes :

- 2 vannes wagon à l'amont
- 2 vanne segment à l'aval

Débit maximale :

- 670 m^3/sec sous RN 200
- 700 m^3/sec sous RN 200

Injection et drainage de la fondation

Voile d'injection : verticale multifilaire axé sur galerie de pied amont, profondeur de 40 à 120 m

Rideau de drainage : 1 rideau foré à partir de la galerie de pied amont

Ouvrage de dérivation

2 galeries souterraines : 8 m (galerie de la solution " digue en enrochements ")

1.6 Présentation générale de l'aménagement

1.6.1 Barrage de Béni Haroun-Volumes Mobilisable

Le barrage de Béni Haroun ouvrage clé du transfert à été terminé en 2001.

Il permet de mobiliser les eaux de l'Oued Rhumel qui est la principale ressource hydrique de la région. Suivant les estimations 435 Hm^3/an seront régularisés. Un complément de 69 Hm^3/an pourra être mobilisé et transférer à Béni Haroun par un ouvrage complémentaire à construire sur l'Oued Bou Siaba, ce qui porte à 504 Hm^3 le volume totale à transférer annuellement vers le Sud.

1.6.2 Le transfert de Béni Haroun

Les ressources transférées permettront :

- de faire face aux besoins en eau potable et industrielle des 6 Wilayas concernées, à l'horizon 2020.
- D'équiper de nouveaux périmètres agricoles (Telerghma, Chemoura, Batna-Ain Touta) d'une superficie totale de 29.523 Ha

Le transfert de Béni Haroun a pour objectif de transférer les ressources mobilisées par les barrages de Béni Haroun et de Bou Siaba (partiellement) vers les 6 wilayas de Batna, Khenchela, Mila, Oum El Bouaghi, Constantine et Jijel (région d'El Milia). Le transfert de Béni Haroun est un vaste complexe hydraulique qui comporte les ouvrages principaux suivants :

- La retenue de Béni Haroun créée par le barrage.
- Le transfert de Béni Haroun-Oued El Athmenia.
- Le réservoir de l'Oued El Athmenia.
- Le transfert du réservoir de l'Oued El Athmenia vers les réservoirs de Tallizerdane et Koudiat Medaour.
- Le réservoir Koudiat Medaour(en construction).
- Le transfert de Bou Siaba Béni Haroun.

Le fonctionnement d'ensemble de système est le suivant : (voir les Figures ci-après, le tableau 1.1 et le schéma 1.2 d'allocation des ressources)

- Les apports sont captés et régularisés sur une base pluriannuelle par des réservoirs de Béni Haroun et Bou Siaba.
- Ils sont ensuite transférés vers le réservoir intermédiaire d'Oued Athmenia puis vers les réservoirs d'extrémité de Koudiat Medaour (3) et Tallizerdane.
- A partir de ces trois réservoirs des conduites d'adduction mèneront l'eau vers les utilisateurs (agglomération et périmètres agricoles).

Afin de minimiser les coûts, les organes de transfert sont conçus pour transporter des volumes constants chaque jour de l'année. Les trois réservoirs précités ont donc une fonction de modulation saisonnière des fournitures d'eau, particulièrement indispensables pour l'irrigation.

Réservoir d'origine	Agglomérations AEPI hm ³	Quantité en	Périmètre Irrigation	Quantité en hm ³	Total AEPI+Irrigation	Quantité en hm ³
Bassin Ain Tinn	Mila	18				18
Oued Athmenia	Constantine	106				106
	Tadjenanent	11				11
	Ain Mila	11	Teleghma	40		11
Koudiat Medaour	Batna-Tazoult-Arris	84				84
	Kais-Khenchela	27				27
			Batna-Ain-Touta	50		50
			Chemoura	30		30
Tallizerdane			Chemoura	127		127
Total		257		247		504

FIGURE 1.4 – Allocation des ressources du transfert à l'horizon 2020

1.7 Caractéristiques générales du bassin

Le bassin de l'Oued El kebir est situé en Algérie Orientale. Il est bordé au sud par les Haut plateaux Tellien, à l'ouest par la montagne de petit Kabylie, au nord par la frontière de partage des eaux que forme l'Atlas Tellien, et à l'est par le bassin de Seybousse. Le bassin se divise en deux parties distinctes. La partie occidentale, constituée par le bassin de l'Oued Endja, se caractérise par des précipitations relativement élevées (700 mm/an en moyenne) et une topographie montagneuse. Dans ce bassin, les cotes atteignent 1.400 m ou davantage.

La superficie de ce bassin est de 2.169 Km². La partie orientale, constituée par le bassin de l'Oued Rhumel, se caractérise par une topographie plus douce et des précipitations modérées (500 mm/an en moyenne). La superficie de ce bassin est de 5.490 Km². Dans ce bassin, le terrain s'élève jusqu'à 1.300 m ou davantage.

Les deux rivières, l'Endja et le Rhumel, confluent près de l'extrémité amont de la gorge de Béni Haroun pour former l'Oued El kebir. La confluence se situe à environ 3 kilomètres au sud de l'extrémité amont de la gorge. Le bassin versant total au site du barrage de Béni Haroun à une superficie de 7.725 km².

Dans le bassin de l'Oued Endja situé à l'ouest et au nord-ouest de l'ensemble du bassin versant du kebir-Rhumel, la saison pluvieuse s'étend généralement du mois de décembre au mois de mai.

Dans le bassin de l'Oued Rhumel qui s'étend du sud au nord-est du bassin, la pluviométrie est distribuée de façon plus uniforme tout au long de l'année que dans le bassin de l'Oued Endja. Les propriétés d'infiltration dans le bassin de l'Oued Rhumel semblent aussi indiquer que le sol de y est plus perméable que celui de l'Oued Endja.

Les débits des Oued Endja et Rhumel à leur confluence ne sont pas proportionnels à leur surface drainantes respectives. Dus à des caractéristiques hydrologiques et géologiques différente. Les apports de l'Oued Endja sont proportionnellement plus grands que ceux de l'Oued Rhumel.

Les températures moyennes varient d'environ 5°C en janvier à environ 25°C en aout. La précipitation moyenne près du bassin varie d'environ 7 mm /mois en juillet à environ 80mm /mois en décembre. L'humidité relative varie entre 50% environ en aout à 70% environ entre novembre et mars.

Les surfaces drainées en différents points caractéristiques du bassin sont données au tableau suivant :



FIGURE 1.5 – Caractéristiques générales du bassin

Sites	Superficie (Km ²)
OUED ENDJA	
- Station de jaugeage Tassadant	937
- Confluence avec Oued Rhumel	2169
OUED RHUMEL	
- Barrage Hammam Grouz	1130
- Station de jaugeage Grarem	5400
- Confluence avec Oued Endja (avec Hammam Grouz)	5490
OUED ELKEBIR	
- Confluence Oued Rhumel et Oued Endja	7659
- Entre confluence et Béni Haroun	66
- Site de barrage :	
. avec Hammam Grouz	7725
. sans Hammam Grouz	6595
- Station de jaugeage El Ancer (avec Hammam Grouz)	8500

FIGURE 1.6 – Superficies drainées aux points caractéristiques du bassin

1.8 Présentation de la direction du barrage

1.8.1 Direction

La direction de barrage de Béni Haroun est structurée en plusieurs services chaque service a son responsable parmi eux il y a des ingénieurs spécialisés, tout ces services sont chapotées par le directeur.

1. Service d'auscultation : Prendre soin du contrôle technique du barrage en faisant toutes les mesures relatives à la stabilité et l'équilibre du barrage
2. Service contentieux : Spécialisés dans tout les cas contre l'Agence nationale des Barrages et des transferts en suivant toutes les procédures nécessaires et adoptées soit devant les tribunaux ou conseils soit avec les huissiers de justice, dans ce service il y a deux travailleurs spécialisés qui travaillent avec les tribunaux et les avocats.
3. Service maintenance : Il se compose de trois ingénieurs et deux techniciens, la tache ingénieure est presque la même (la maintenance et l'entretien, la réparation des machines et des appareils électriques, des pompes, etc...), il y a deux type de maintenance :
 - Maintenance préventive : (prévenir la panne) d'où les techniciens signale le problème et les ingénieurs le résolve ;
 - Maintenance corrective : (après la panne) c'est la tache des techniciens.

4. Service comptabilité :Pour compter tout les besoins du barrage en ce qui concerne les dépenses des autres services (électricité, internet, les petit achats, les uniforme des travailleurs, etc...).
5. Service d'exploitation :Représente les donnés journalière comme la température du barrage, le volume d'évaporation, cote, vitesse de vent, pluviométrie. . . ,
6. Service d'entretient :Pour garder la propreté interne du barrage.

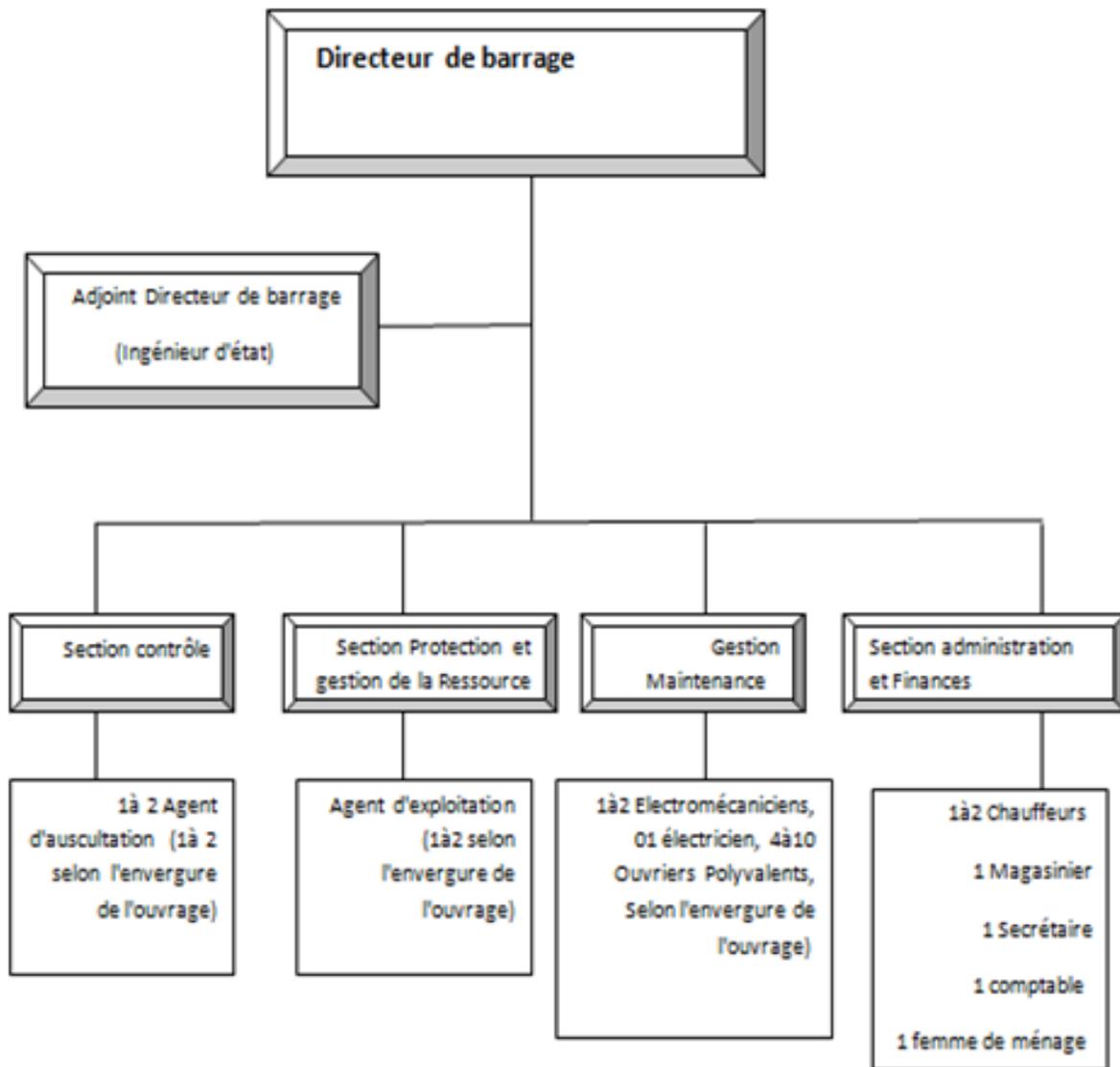


FIGURE 1.7 – Organigramme type de Barrage

1.8.2 La station météorologique

La station météorologique du barrage de Béni Haroun est un ensemble de capteurs qui enregistrent et fournissent des mesures physiques et des paramètres météorologiques liés aux variations du climat, ces capteurs étant placés dans un boîtier, abri météorologique qui réalise l'équilibre thermique du thermomètre avec l'air et le protège du rayonnement solaire. Les variables à mesurer sont la température, la vitesse du vent, la pluviométrie ainsi que l'évaporation. Les stations peuvent comporter des capteurs pour toutes ou une partie seulement de ces informations, selon leur type : manuel ou automatique.

Ces appareils sont :

1. **La pluviométrie** : Pour mesurer la quantité de pluie de chaque jour, y en a deux type d'appareils : automatique et manuel comme indique la figure suivante.



FIGURE 1.8 – Automatique et Manuelle.

2. **Le bac d'évaporation (bac Colorado)** :

Un bac d'évaporation est un évaporomètre constitué par un bassin ou un bac d'eau d'assez grande surface et assez profond où l'on mesure le changement du niveau de l'eau dû à l'évaporation. Les bassins vont de 1 à 5 mètres de diamètre et de 10 à 70 cm de profondeur. Ils sont posés sur ou dans le sol (bacs enterrés) ou encore dans l'eau (bacs flottants). Dans tous les cas, le niveau de l'eau est maintenu à faible distance au-dessous du bord du bac. Les variations du niveau d'eau du bac, mesurées à des intervalles fixes, sont le reflet de l'intensité de l'évaporation.



FIGURE 1.9 – Le bac Colorado

- **Caractéristiques** :Le bac Colorado s’agit d’une cuve ronde de 1 220 mm de diamètre et 254 mm de hauteur, rempli jusqu’à entre 50 et 75 mm de son bord. Le bac d’évaporation est surélevé du sol (en général 10 cm) sur une palette de bois ajouré afin que l’air puisse circuler sous celui-ci. Cela empêche la transmission de chaleur sol-bac qui pourrait fausser les résultats. Il permet de mesurer la vitesse d’évaporation d’un volume d’eau et d’une surface donnés exposés à l’air (exprimée en mm par jour, par mois ou par an) et l’évapotranspiration. Cette donnée dépend exclusivement de la température et de l’humidité ambiante, ainsi que des précipitations.

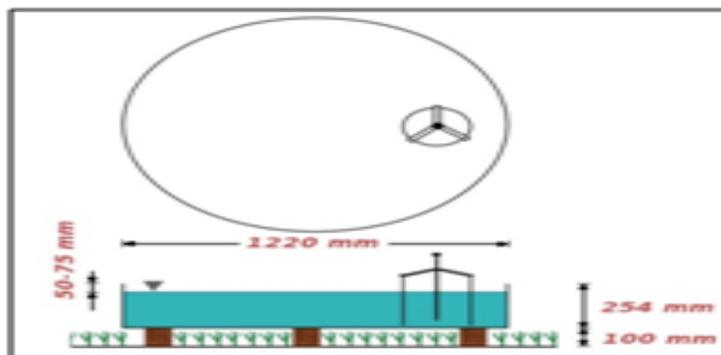


FIGURE 1.10 – Diagramme d’un bac Colorado

- **Usage :** Le niveau dans les bacs d'évaporation est relevé quotidiennement. Les relevés sont effectués en mesurant la quantité d'eau ajoutée chaque jour pour avoir un niveau constant. Le pluviomètre, généralement fixé au sol, de la station météorologique connexe permet d'évaluer les précipitations tombées dans le bac car la quantité de précipitations tombées doit être tenue en compte pour calculer l'évaporation.

Si après une pluie, le niveau du bac dépasse la pointe, il faut retirer le nombre de litres nécessaire afin de revenir au niveau de base. L'évaporation est dans ce cas égale à la pluie moins le nombre de litres d'eau retirés. Si la pluie fait déborder le bac, il est impossible de réaliser la mesure d'évaporation journalière. Elle peut être cependant estimée en minimisant les évaporations du jour précédent et du jour suivant.

3. Appareil pour mesurer la vitesse dU vent :



FIGURE 1.11 – Appareil pour mesurer la vitesse dU vent

Cet instrument nous permet de mesurer instantanément la vitesse du vent, la mesure se fait en tour par minute puis on fait la conversion Km/h.

4. **Thermomètre :** pour mesurer la température.

A l'abri se trouve un thermomètre pour mesurer la température de l'air ambiante.

1.8.3 Salle de contrôle :

Cette salle est équipée par un tableau qui s'appelle " le tableau synoptique " il est utilisé pour afficher les données nécessaires qu'elles sont calculées automatiquement. (Cote, volume, température, sous-pressions ...ect).



FIGURE 1.12 – Tableau synoptique

Dans cette salle, il existe aussi un micro ordinateur dont lequel des ingénieurs font plusieurs calculs dans un fichier Excel qui contient différentes fonctions pour calculer instantanément différentes mesures comme les affluents, les défluents, les déversoirs et le facteur K qui est une facteur de stabilité de l'ouvrage.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1														
2	DATE/HEURE	Lecture VENT	vitesse du vent	T -°C	PLUVIO-METRIE (mm)	PLUVIO-GRAPHIE (mm)	Eau ajouté (mm)	BAC + Eau ajouté + pluie (mm)	EVAPORATION au BAC A	Lecture Bac	Hauteur (mm)	Côte de la retenue	SURFACE Cuvette (m2)	VOLUME RETENUE EVAPORE m3
6	01/04/2016 08:00	3295,47	128,54	14,0					5,80	0,30	3,40	200,22	37,837	1006,229
7	01/04/2016 18:00								5,00	0,80			0,000	0,000
8	02/04/2016 08:00	3435,68	140,21	10,0	40,00		26,50	45,00	44,50	0,50	1,30	200,31	38 083 100,00	38 616,26
9	02/04/2016 18:00								26,00	0,50				
10	03/04/2016 08:00	3488,19	52,51	11,0					25,60	0,40	0,90	200,44	38 053 900,00	26 713,84
11	03/04/2016 18:00								24,00	1,60				
12	04/04/2016 08:00	3503,24	15,05	13,0					23,20	0,80	2,40	200,42	38 039 300,00	71 209,57
13	04/04/2016 18:00								21,20	2,00				
14	05/04/2016 08:00	3524,94	21,70	14,0					19,90	1,30	3,30	200,39	38 025 020,00	97 876,40
15	05/04/2016 18:00								17,40	2,50				
16	06/04/2016 08:00	3718,69	193,75	13,0					15,70	1,70	4,20	200,35	38 018 040,00	124 547,10
17	06/04/2016 18:00								15,00	0,70				
18	07/04/2016 08:00	3923,66	204,97	13,0	4,5			19,50	19,20	0,30	1,00	200,33	38 004 080,00	29 643,18
19	07/04/2016 18:00				2,5			21,70	21,00	0,70				
20	08/04/2016 08:00	4030,72	107,06	11,0					20,90	0,10	0,80	200,35	37 997 100,00	23 710,19
21	08/04/2016 18:00								18,50	2,40				
22	09/04/2016 08:00	4097,36	66,64	7,0	4,0			22,50	21,00	1,50	3,90	200,34	37 990 120,00	115 565,95
23	09/04/2016 18:00								18,50	2,50				
24	10/04/2016 08:00	4153,96	56,60	9,0					17,20	1,30	3,80	200,33	37 983 140,00	112 582,03
25	10/04/2016 18:00								15,00	2,20				
26	11/04/2016 08:00	4165,67	11,71	14,0					13,40	1,60	3,80	200,32	37 976 160,00	112 561,34
27	11/04/2016 18:00								11,00	2,40				
28	12/04/2016 08:00	4181,50	15,83	16,0					8,60	2,40	4,80	200,29	37 962 200,00	142 130,48
29	12/04/2016 18:00								5,00	3,60				

FIGURE 1.13 – Exemple d'une feuille de calcul

1.9 Position du problème

A l'heure actuelle, les eaux du Barrage de Béni Haroun sont utilisées pour alimenter deux Wilaya (Constantine, Mila). De même, les eaux servent à l'irrigation des régions de Téleghma et Chemora.

Les autorités planifient l'alimentation en eau de trois autres Wilaya : Khenchela, Batna et Oum Elbouaghi. Sur la base des besoins exprimés, la question qui se pose est de savoir ; sur la base des prévisions, est-ce-que les apports en eau (pluviométrie) seront suffisants pour compenser les quantités d'eau transférés ?

1.10 Conclusion

Le barrage de Béni Haroun est le projet du siècle du programme algérien d'aménagement des ressources en eau. Il constitue la plus grande retenue artificielle algérienne et la seconde du continent africain avec une réserve de 1 milliard de m³ d'eau atteinte en février 2012 (soit 40 millions de m³ au-delà de sa capacité d'objectif), répartis sur 3900 hectares. Situé

sur l'oued el Kébir, il est alimenté par deux bras principaux, avec les oueds Rhumel et Enndja. Dans ce chapitre, nous avons fait la présentation générale de ce barrage. Dans le chapitre suivant, nous allons citer les principales méthodes statistiques de prévision qui nous servirons pour faire la partie de l'application.

Chapitre 2

Méthodes statistiques de la prévision

2.1 Introduction

L'étude du comportement d'une ou de plusieurs séries chronologiques a depuis longtemps suscité l'intérêt des chercheurs dans des domaines très variés. Elle permet d'analyser, expliquer et décrire un phénomène au cours de temps et d'enterrer des conséquences pour des prises de décision. L'un des objectifs principaux de cette étude est la prévision qui consiste à prévoir des valeurs futures Y_{t+h} ($h = 1, \dots$) de la série chronologique à partir de ces valeurs observées jusqu'au temps Y_1, \dots, Y_t .

L'analyse des séries chronologiques permet d'étudier les situations passées et présentes et peut extrapoler l'événement dans un futur relativement proche (la prévision) [4]. La prévision se fonde donc sur la connaissance du passé et du présent. Par l'analyse des séries chronologiques, on étudie une classe particulière de processus appelés processus aléatoires stationnaires. Ces processus sont caractérisés par le fait que leurs propriétés ne changent pas au cours du temps. Au sein des processus stationnaires, on se limite au processus dit stationnaire au second ordre.

L'objectif de ce chapitre est de présenter les éléments théoriques utilisés pour l'élaboration d'un modèle de prévision, les modèles autorégressif AR(p), moyenne mobile MA(q), mixte ARMA (p, q), ARIMA, SARIMA et particulièrement, la méthode de Box et Jenkins et Lissage exponentiel.

2.2 Les séries chronologiques

2.2.1 Définition et exemple :

Une série chronologique $(Y_t, t \in T)$ est une suite d'observation d'une variable Y à différentes dates(t). Habituellement T est dénombrable, de sorte que $t = \overline{1, T}$.

Le but de l'analyse des séries temporelles (séries chronologiques) est de s'intéresser à la dynamique d'une variable. Cette dernière est importante pour au moins une raison :

D'un point de vue économétrique, on ne peut relier que deux variables qui ont des propriétés similaires, en particulier une même stabilité ou instabilité; les propriétés mathématiques permettant de relier deux variables dépendent de leur dynamique.

Une série temporelle peut concerner des données macroéconomiques (masse monétaire, inflation,) microéconomiques (nombre d'employés d'une entreprise, ventes,), politiques (nombre de votants, nombre de votes nuls,) démographiques (âge moyen des habitants d'une localité, leur taille,), financières (cours d'une action,), climatologique (pluviométrie, évaporation, température,). La périodicité de la série importe peu. Il peut s'agir de mesures annuelles, semestrielles, mensuelles etc.

- Si T est dénombrable (en générale $T \in Z$), on parle d'une série temporelle à temps discret.
- Si T n'est pas dénombrable (en général un intervalle de R on parle de série temporelle à temps continu.

2.2.2 Les objectifs principaux :

L'étude de la série peut avoir les objectifs suivants non exclusifs :

- Comprendre le passé : expliquer les variations observées.
- Prédire les valeurs futures (proches) par exemple, la température ou le cours d'une action du lendemain ou l'évolution de la population mondiale au cours du siècle prochain.
- Etudier le lien avec d'autres séries. Par exemple, l'évolution démographique de deux régions.

2.2.3 Les composantes d'une série chronologique

Dans un premier temps, on fait une représentation graphique de la série étudiée $y_t, t = \overline{1, T}$ ce qui nous permet de déterminer les composantes de cette série. ces composantes vont former les fondements de la série chronologique. On considère alors que la série est résultante de plusieurs composantes, il faut alors analyser ces composantes en les dissociant les uns des autres [1]

La Tendance (Z_t) : Cette composante représente l'évolution à long terme de la variable étudiée et traduit le comportement moyen de la série. cette tendance peut-être linéaire, polynomiale ou exponentiel.

La Composante Saisonnière (S_t) : Elle correspond à un phénomène qui se répète à un intervalle de temps régulier (périodique). En générale est un phénomène saisonnier, d'où le nom de saisonnalité. La composante saisonnière est donc périodique de période p . Il suffit de connaître les p premières valeurs S_1, \dots, S_p et donc par périodique $S_T = S_{T+p}$.

La Composante Irrégulière (ϵ_t) : Celle-ci regroupe tout ce qui n'est pas pris en compte par les deux composantes précédentes, on l'appelle Bruit Blanc (Résidu).

2.3 Quelques définitions :

a) **processus stochastique [7] :**

Un processus stochastique (Y_t) est une famille de variables aléatoires ($Y_t, \forall t \in \Theta$) indexée par un ensemble (Θ) si :

- ($t \in Z$) est dit un temps discret.
- (t) appartient à intervalle de \mathbb{R} est dit continue.

b) **Le processus stochastique du second ordre [7] :**

Soit (Y_t) un processus stochastique (i.e. une suite de variable aléatoire). Le processus y_t est dit du seconde ordre si pour tout ($t \in Z$), (Y_t) est une variable aléatoire de carré intégrable i.e. $E(|Y_t|^2) < \infty$.

c) **Processus stationnaire :**

On dit qu'un processus (Y_t) est stationnaire si les moments du premier et du second ordre sont invariants par translation dans le temps, c'est-à-dire que :

$$\begin{cases} E(Y_t) = \mu, \forall t \in Z \\ Cov(Y_t; Y_{t-h}) = E(Y_t - \mu)(Y_{t-h} - \mu), \forall t \in Z \end{cases}$$

On se limite à des processus stationnaire du second ordre, ou encore faiblement stationnaire, si le processus (Y_t) vérifie :

$$\begin{cases} E(Y_t) = E(Y_{t+h}) = \mu, \forall t \in Z \\ V(Y_t) = \sigma^2, \forall t \in Z \\ Cov(Y_t; Y_{t+h}) = \gamma(h), \forall t \in Z \end{cases}$$

d) Processus aléatoire stationnaire :

Définition :

On dit qu'un processus aléatoire temporel (Y_t) est complétement stationnaire si pour toute suite d'instantns t_1, \dots, t_n , il existe un entier h quelconque tel que :

$$f(y_{t_1}, \dots, y_{t_n}) = f(y_{t_1+h}, \dots, y_{t_n+h})$$

Aussi on dit qu'il est stationnaire en seconde ordre si :

$$\begin{cases} E(Y_t) = \mu, \forall t \in Z \\ V(Y_t) = \sigma^2, \forall t \in Z \\ Cov(Y_t; Y_{t+h}) = \gamma(h), \forall t \in Z \end{cases}$$

e) Bruit blanc [7] :

le processus stationnaire bruit blanc est une suite de variables aléatoires (ϵ_t) de même loi, de même variance et sont mutuellement indépendantes tel que :

$$\begin{cases} E(\epsilon_t) = 0, \forall t \in Z \\ V(\epsilon_t) = \sigma^2, \forall t \in Z \\ E(\epsilon_t, \epsilon_{t'}) = 0, \forall (\epsilon_t; \epsilon_{t'}) \end{cases}$$

- Bruit Blanc fort :

On dit qu'un bruit blanc est fort, si (ϵ_t) est une suite de variables aléatoires réelles indépendantes et identiquement (iid), veut dire :

$$\begin{cases} E(\epsilon_t) = 0, \forall t \in Z \\ E(\epsilon_t^2) = \sigma^2, \forall t \in Z \end{cases}$$

- Bruit Blanc faible :

On dit qu'un bruit blanc est faible, si (ϵ_t) est une suite de variables aléatoires réelles indépendantes et identiquement (iid), veut dire :

$$\begin{cases} E(\epsilon_t) = 0, \forall t \in Z \\ V(\epsilon_t^2) = \sigma^2, \forall t \in Z \\ Cov(\epsilon_t, \epsilon_{t'}) = 0, \forall (t, t') \in Z^2, t \neq t' \end{cases}$$

f) Opérateur retard :

on appelle opérateur retard l'opérateur noté B qui a tout processus (Y_t) associe le processus

$(Y_t), t \in Z$ défini par :

$$\forall t \in Z, BY_t = Y_{(t-1)}.$$

L'opérateur B décale le processus d'une unité de temps vers le passé.

Si on applique h fois cette opérateur, on décale le processus de h unité de temps :

$$B(B(\dots BY_t)) = B^h Y_t = Y_{(t-h)}$$

g) Opérateur différence :

L'opérateur ∇ fait la différence entre le processus et sa version décalée d'une unité de temps. Cette opération se construit en utilisant l'opérateur précédent :

$$\begin{aligned} \nabla Y_t &= Y_t - Y_{(t-1)} = Y_t - BY_t \\ &\iff \nabla = 1 - B; \end{aligned}$$

– Élimination de la tendance[1]

pour éliminer la tendance, on applique le filtre différence noté ∇ tel que :

$$Z_t = \nabla Y_t = Y_t - Y_{t-1}$$

si la série admet toujours une tendance on peut appliquer de nouveau le filtre ∇ :

$$\begin{aligned} \nabla Z_t &= \nabla Y_t - \nabla Y_{t-1} = Y_t - Y_{t-1} - Y_{t-1} + Y_{t-2} \\ &\iff = \nabla^2 Y_t = \nabla Z_t = Y_t - 2Y_{t-1} + Y_{t-2}. \end{aligned}$$

Remarque : le filtre ∇ peut s'écrire en fonction de l'opérateur de retard B tel que :

$$BY_t = Y_{t-1}$$

Donc :

$$\begin{aligned} \nabla Y_t &= Y_t - Y_{t-1} = Y_t - BY_t = Y_t(1 - B) \\ \iff \nabla^2 Y_t &= Y_t - 2Y_{t-1} + Y_{t-2} = Y_t - 2BY_t + B^2 Y_t \\ &= Y_t(1 - 2B + B^2) \\ &= Y_t(1 - B)^2. \end{aligned}$$

– Élimination de la saisonnalité[1]

dans le cas de présence d'une saisonnalité on applique le filtre saisonnier ∇_s , tel que :

$$S_t = \nabla_s Y_t = Y_t - Y_{t-s}$$

D'après la remarque précédente on peut écrire :

$$\begin{aligned} \nabla_s Y_t &= Y_t - Y_{t-s} = Y_t - B^s Y_t \\ \nabla_s Y_t &= Y_t(1 - B^s) \end{aligned}$$

2.4 Modélisation d'une série chronologique

Elle s'agit de déterminer des modèles décrivant la façon dont la série évolue. On distingue trois types de modèles :[1]

1. Modèles d'ajustements :

Les deux modèles les plus utilisés sont les suivants :

a) Modèles additifs et multiplicatifs :

ce sont des modèles qui permettant de combiner les deux composantes principales de la série.

- Modèles additifs : dans ce modèle l'amplitude de la composante saisonnière et du bruit reste constante au cours du temps, son expression est de la forme : $y_t = Z_t + S_t + \epsilon_t$ Avec $t = \overline{1..n}$.

- Modèle multiplicatifs : son expression est de la forme : $y_t = Z_t * S_t * \epsilon_t$ Avec $t = \overline{1..n}$, Dans ce modèle, l'amplitude de la composante saisonnière et du bruit n'est plus constant au cours du temps, elle varie proportionnellement à tendance Z_t .

b) Modèles de Regression :

c'est la méthode la plus simple pour analyser une serie chronologique. le modèle s'écrit $Y_t = g(t, \theta) + \epsilon_t, t = \overline{1..T}$. $g(t, \theta)$ est une fonction déterministe en fonction de t, et $\theta = (\theta_1, \theta_2, \dots, \theta_p) \in \mathbb{R}^p$.

le modèle le plus connu est le modèle :

$Y_t = \alpha t + \beta + S_t + \epsilon_t, t = 1, ..T$. appelé modèle de BUYS-BALLOT, la prévision est : $\hat{Y}_{t+1} = \hat{\alpha}(T + 1) + \hat{\beta} + S_{T+1}$. $\theta = (\alpha, \beta) \in \mathbb{R}^p$.

2. Modèles explicatifs :

Dans cette catégorie de modèles, la variable y_t est exprimée en fonction d'un vecteur de variable explicatives ou variables EXOGENES. c-à-d :

$$Y_t = f(X_t) + \epsilon_t$$

Où Y_t peut être déterministe ou aléatoire.

a) Les modèles explicatifs statiques : dans un modèle explicatif statique, les variables Y_t et les ϵ_t sont indépendante entre eux.

b) Les modèles explicatifs dynamiques : un modèle explicatif peut être dynamique, de fait que les ϵ_t sont auto corrélés, ou bien X_t contiennent les valeurs passées de Y_t .

3. Modèle auto-projectifs

on suppose que y_t est une fonction de ses valeurs passées et d'une perturbation

aléatoire. Le modèle s'écrit :

$$y_t = f(y_{t-1}, y_{t-2}, \dots) + \epsilon_t$$

Parmi les méthodes mises en évidence, on trouve la méthode de lissage exponentielle et les modèles de Box et Jenkins.

2.5 Lissage Exponentiel

Les techniques de lissage exponentiel ont été introduites par Holt en 1958 ainsi que par Winters en 1960, elles sont popularisées par le livre de Brown en 1963 qui utilisa pour la première fois cette méthode pour la prévision des pièces de rechange et pour la gestion des stocks.

Les méthodes de lissage constituent l'ensemble des techniques empiriques qui ont pour caractéristiques de donner un poids prépondérant aux valeurs récentes de la série chronologique, leur domaine d'utilisation privilégié est le court terme. Ces méthodes étant relativement basiques et simples de mise en œuvre, elles sont souvent utilisées dans l'industrie, notamment lorsque le nombre de prévisions à réaliser est important.

Tout lissage repose sur le choix d'une fonction de prévision, c'est-à-dire d'une fonction mathématique qui va servir à l'extrapolation. Cette fonction qui peut être linéaire, exponentielle, sinusoidale . . . , constitue le premier facteur de différenciation de ces méthodes (lissage simple, double, triple).

La fonction choisie le plus souvent empiriquement, dépend d'un certain nombre de paramètres a_0, a_1, \dots, a_k qui sont calculés à partir de l'historique disponible et mise à jour à l'arrivée de chaque nouvelle observation.

Nous présentons trois types de lissage exponentiel [5] :

- le lissage exponentiel simple qui ajuste la série si elle ne représente ni tendance ni saisonnalité.
- le lissage exponentiel double qui ajuste quant à lui une droite.
- le lissage exponentiel de Holt-Winters qui considère des fonctions plus complexes (polynomiales, périodiques, etc.). Voir le tableau 2.2

TENDANCE	SAISONNALITE	METHODE
NON	NON	LISSAGE SIMPLE
OUI	NON	LISSAGE DE HOLT
OUI	OUI	LISSAGE DE WINTERS

TABLE 2.1 – Différentes méthodes de lissage exponentiel

2.5.1 Lissage Exponentiel simple (LES) :

Le lissage exponentiel simple permet d'effectuer des prévisions pour des séries chronologiques dont la tendance est constante et sans saisonnalité. Le modèle considéré s'écrit [9] :

$$y_t = a + \epsilon_t, \epsilon_t \text{ est la composante irrégulière}$$

Soit y_t une telle série dont on a observé les T premiers instants y_1, \dots, y_t . Pour $h \in \mathbb{N}^*$, on cherche à prévoir la valeur y_{t+h} , c'est-à-dire faire une prévision de la série à l'horizon h . Etant donné un réel α tel que $0 < \alpha < 1$, comme la tendance est constante, on cherche une prévision $\hat{y}_T(h)$ sous la forme de la constante qui s'ajuste le mieux au sens des moindres carrés pondérés au voisinage de T , c'est-à-dire la solution du problème de minimisation

$$\min_{\alpha} \sum_{j=0}^{T-1} \alpha^j (y_{T-j} - a)^2$$

Définition :[10]

La prévision de la série à l'horizon h , $\hat{y}_T(h)$ fournie par la méthode de lissage exponentiel simple est donnée par :

$$\hat{y}_T(h) = (1 - \alpha) \sum_{j=0}^{T-1} \alpha^j y_{T-j}$$

Où α est la constante de lissage.

$\hat{y}_T(h)$ est une moyenne des observations passées où le poids de chaque observation décroît de façon exponentielle avec la distance.

Formules récursives de mise à jour [5]

A partir de la définition, on obtient les formules de mise à jour suivantes :

$$\begin{aligned} \hat{y}_T(h) &= \alpha \hat{y}_{T-1}(h) + (1 - \alpha) y_T \\ &= y(h) + (1 - \alpha)(y_T - \hat{y}_{T-1}(h)) \end{aligned}$$

La première équation montre que la prévision au temps T est une moyenne pondérée de la prévision au temps $T-1$ et du résultat de la série au temps T . Elle permet de faire une mise à jour très simple. La deuxième relation définit la prévision au temps T comme la prévision au temps $T-1$ corrigée de l'erreur de prévision commise au temps $T-1$ amortie d'un facteur α

L'initialisation :

L'utilisation des formules récursives nécessite d'initialiser la récurrence, une manière habituelle d'initialiser $\hat{y}_T(h)$ est de le choisir égale à la moyenne arithmétique des observations, ou de choisir $\hat{y}_1(h) = y_1$, on remarque que pour T assez grand, la valeur initiale a peu d'influence [7].

Choix de la constante de lissage[12]

Un problème important en pratique est le choix du α qui en générale très subjectif et varie selon le contexte de l'étude et/ ou le type de prévision souhaité. En pratique, des valeurs de α comprises entre 0.7 et 0.99 produisent des prévisions qui tiennent compte du passé lointain, elles sont rigides et peu sensibles aux variations conjoncturelles. Il n'en n'est plus de même si α est compris entre 0.01 et 0.3, la prévision est souple c'est à dire influencée par les observations récentes.

On peut aussi déterminer un α adapté à la prévision d'une série donnée. Pour un horizon h on cherche la valeur de α qui minimise :

$$\sum_{t=1}^{T-h} (y_{t+h} - (1 - \alpha) \sum_{j=0}^{T-1} \alpha^j y_{T-j})^2 = \sum_{t=1}^{T-h} (y_{t+h} - \hat{y}_t(h))^2$$

C'est à dire la valeur qui en moyenne réalise la meilleure prévision à l'horizon h sur les bases $1, 2, \dots, T-h$ en utilisant le critère des moindres carrés. On choisit une grille de valeurs de α pour lesquelles on calcule le critère et on retient la valeur qui le minimise. Cette technique de choix de constante peut être reprise dans les méthodes qui vont suivre (lissage exponentiel double, ainsi que les méthodes de Holt-Winters). Les calculs sont souvent plus longs car ce n'est, en fait, plus une constante de lissage que nous sommes amenés à déterminer mais souvent deux ou trois.

2.5.2 Le lissage exponentiel double (LED)

Quand la série chronologique présente une tendance, le lissage exponentiel simple n'est plus adapté, on utilise alors le lissage exponentiel double qui généralise l'idée du lissage exponentiel simple au cas où la série peut être ajustée par une droite au voisinage de T .

On suppose alors que le modèle peut s'écrire au voisinage de T par [9] :

$$y_t = a + b(t - T) + \epsilon_t$$

on cherche donc une prévision à l'horizon h , $\hat{y}_T(h)$ de la forme.

$$\hat{y}_T(h) = \hat{a}(T) + h\hat{b}(T)$$

Où les coefficients : $\hat{a}(T)$ et $\hat{b}(T)$ sont solution de [12]

$$\sum_{j=0}^{T-1} \alpha^j (y_{T-j} - (\hat{a} - \hat{b}_j))^2 = \inf_{a \in \mathbb{R}, b \in \mathbb{R}} \sum_{j=0}^{T-1} \alpha^j (y_{T-j} - (a - b_j))^2$$

En notant la serie lissée L_1 et la serie doublement lissée L_2 définies par :

$$L_1(t) = (1 - \alpha) \sum_{j=0}^{T-1} \alpha^j y_{T-j}$$

$$L_2(T) = (1 - \alpha) \sum_{j=0}^{T-1} \alpha^j L_1(T - j)$$

On déduit la définition suivante :

Définition :

La prévision de la série à l'horizon h , $\hat{y}_T(h)$ fournie par la méthode de lissage exponentielle double est donnée par :

$$\hat{y}_T(h) = \hat{a}(T) + h\hat{b}(T)$$

le couple $(\hat{a}(T), \hat{b}(T))$ et donné par :

$$\begin{cases} \hat{a}(T) = 2L_1(T) - L_2(T), \\ \hat{b}(T) = \frac{1-\alpha}{\alpha}(L_1(T) - L_2(T)), \end{cases}$$

Formules de mise à jour[5]

Les formules de mise à jour s'obtiennent à partir de ces expressions :

$$\hat{a}(T) = \hat{a}(T-1) + \hat{b}(T-1) + (1 - \alpha^2)(y_T - \hat{y}_{T-1}(1)) \quad (2.1)$$

$$\hat{b}(T) = \hat{b}(T-1) + (1 - \alpha^2)(y_T - \hat{y}_{T-1}(1)) \quad (2.2)$$

L'initialisation

Le plus simple consiste à prendre $\hat{a}(2) = y_2$ et $b = y_2 - y_1$, mais d'autres techniques peuvent être envisagées, par exemple une régression linéaire sur les premières valeurs pour donner une estimation locale de la tendance initiale.[7]

2.5.3 Méthode de Holt Winters

a) Méthode de Holt Winters (séries non saisonnières)

Les méthodes de Brown (lissage exponentiel généralisé), ont pour origine une généralisation de la méthode du lissage exponentiel simple considérée comme un ajustement local d'une droite horizontale. Holt Winters (1960) ont adopté une approche différente.

La méthode de Holt Winters sans saisonnalité est fondée comme la méthode du lissage exponentiel double, sur l'hypothèse qu'un ajustement au voisinage de T d'une droite quelconque d'équation $y_t = a + b(t - T) + \epsilon_t$ est préférable à l'ajustement d'une droite horizontale. La prévision prend la forme [8] :

$$\hat{y}_T(h) = \hat{a}(T) + h\hat{b}(T)$$

Formules récursives de mise à jour[8]

La différence par rapport à la méthode de lissage exponentiel double est au niveau des formules de mise à jour dans l'estimation des paramètres a et b .

soient deux constantes de lissage $0 < \beta < 1$ et $0 < \gamma < 1$

les formules de mise à jour sont :

$$\hat{a}(T) = (1 - \beta)y_T + \beta[\hat{a}(T - 1) + \hat{b}(T - 1)] \quad (2.3)$$

$$\hat{b}(T) = (1 - \gamma)[\hat{a}(T) - \hat{a}(T - 1) + \gamma\hat{b}(T - 1)] \quad (2.4)$$

La formule (2.3) s'interprète comme une moyenne pondérée de deux informations sur le " niveau " a de la série à la date T : l'observation y_t et la prévision faite à $T-1$.

La formule (2.4) s'interprète comme une moyenne pondérée de deux informations sur la " pente " b de la série à la date T : la différence entre les niveaux estimés à $T-1$ et T est la pente estimée en $T-1$.

Cette méthode est plus flexible que la méthode du lissage exponentiel double dans la mesure où elle fait intervenir deux constantes (γ, β) au lieu d'une seule (α) .

$$\begin{aligned} \hat{a}(T) &= (1 - \alpha^2)y_T + \alpha^2[\hat{a}(T - 1) + \hat{b}(T - 1)] \\ \hat{b}(T) &= \hat{b}(T - 1) + \frac{(1 - \alpha)^2}{1 - \alpha^2}[\hat{a}(T) - \hat{a}(T - 1) + \hat{b}(T - 1)] \end{aligned}$$

Ces deux équations sont identiques aux équations (2.1) et (2.2), si on prend :

$$\beta = \alpha^2 \text{ et } \gamma = \frac{(1 - \alpha)^2}{1 - \alpha^2} = \frac{2\alpha}{1 + \alpha}$$

Remarque :

L'initialisation peut se faire comme dans le cas du lissage exponentiel double.

b) Méthode de Holt Winters (séries saisonnières)

Ce sont les méthodes à privilégier parmi les techniques de lissage exponentiel dans le cas de séries d'observations présentant à la fois un terme de tendance et une saisonnalité. Elles opèrent le lissage simultané de 3 termes correspondant respectivement à des estimations locales du niveau de la série désaisonnalisée a_T , de la pente de la tendance b_T et de la saisonnalité S_T qui représente la saisonnalité de période s . On peut citer au moins deux méthodes dont l'une est adaptée aux séries admettant une décomposition multiplicative et l'autre correspondant aux décompositions additives [12].

	Modèle additif	Modèle multiplicatif
Modèle	$y_t = a + (t - T)b + S_t + \varepsilon_t$	$y_t = [a + (t - T)b]S_t * \varepsilon_t$
Fonction de prévision	$\hat{y}_t(h) = \hat{a}(T) + h\hat{b}(T) + \hat{S}_{T+h-s}$ $1 \leq h \leq s$ $\hat{y}_t(h) = \hat{a}(T) + h\hat{b}(T) + \hat{S}_{T+h-2s}$ $s + 1 \leq h \leq 2s$ <p>et ainsi de suite pour $2s < h$</p>	$\hat{y}_t(h) = \hat{a}(T) + h\hat{b}(T) + \hat{S}_{T+h-s}$ $1 \leq h \leq s$ $\hat{y}_t(h) = \hat{a}(T) + h\hat{b}(T) + \hat{S}_{T+h-2s}$ $s + 1 \leq h \leq 2s$ <p>et ainsi de suite pour $2s < h$</p>
Mise à jour	$\hat{a}(T) = (1 - \beta)(y_T - \hat{S}_{T-s}) + \beta$ $[\hat{a}(T - 1) + \hat{b}(T - 1)]$ $\hat{b}(T) = (1 - \gamma)[\hat{a}(T) - \hat{a}(T - 1)]$ $+ \gamma\hat{b}(T - 1)$ $\hat{S}_T = (1 - \delta)[y_T - \hat{a}(T)] + \delta\hat{S}_{T-s}$	$\hat{a}(T) = (1 - \beta)\frac{y_T}{\hat{S}_{T-s}} + \beta$ $[\hat{a}(T - 1) + \hat{b}(T - 1)]$ $\hat{b}(T) = (1 - \gamma)[\hat{a}(T) - \hat{a}(T - 1)]$ $+ \gamma\hat{b}(T - 1)$ $\hat{S}_T = (1 - \delta)\frac{y_T}{\hat{a}(T)} + \delta\hat{S}_{T-s}$

TABLE 2.2 – Méthode de Holt-Winters additive et multiplicative

La première formule de mise à jour s’interprète comme une moyenne pondérée de la différence des niveaux estimés aux instants T et T-1 et la pente estimée à l’instant T-1.

La deuxième comme une moyenne pondérée de l’observation y_t (à laquelle on a retranché la composante saisonnière estimée à l’étape précédente) et l’estimation de la tendance faite à l’instant T-1.

La troisième comme une moyenne pondérée de l’observation y_t (à laquelle on a retranché

le niveau calculé à l'instant T) et de la composante saisonnière calculée à l'instant $T-s$.

Initialisation : [7]

L'initialisation requiert cette fois trois valeurs : $\hat{a}_T, \hat{b}_T, \hat{S}_T$. Il est naturel de choisir pour $T = 1, \dots, s$:

$$\hat{a}_T = \frac{y_1 + \dots + y_s}{s}$$

$$\hat{b}_T = \left[\frac{y_{1+s} - y_1}{s} + \dots + \frac{y_{2s} - y_s}{s} \right]$$

$$\hat{S}_T = \frac{y_T}{\hat{a}_T}$$

mais d'autres choix restent possibles, par exemple pour $s = 4$ on peut prendre :

$$\hat{a}(3) = \frac{1}{8}y_1 + \frac{1}{4}y_2 + \frac{1}{4}y_3 + \frac{1}{4}y_4 + \frac{1}{8}y_5$$

$$\hat{a}(4) = \frac{1}{8}y_2 + \frac{1}{4}y_3 + \frac{1}{4}y_4 + \frac{1}{4}y_5 + \frac{1}{8}y_6$$

$$\hat{b}_4 = \hat{a}(4) - \hat{a}(3)$$

$$\hat{S}_4 = y_4 - \hat{a}(4)$$

$$\hat{S}_3 = y_3 - \hat{a}(3)$$

$$\hat{S}_2 = y_2 - \hat{a}(3) + \hat{b}_4$$

$$\hat{S}_1 = y_1 - \hat{a}(3) + 2\hat{b}_4$$

$\hat{a}(3), \hat{a}(4), \hat{b}_4$ Même valeur que pour la méthode additive saisonnière.

$\hat{S}_4, \hat{S}_3, \hat{S}_2, \hat{S}_1$: On divise les observations par la tendance linéaire calculée à partir de $\hat{a}(4)$ et \hat{b}_4 .

2.6 Méthode de Box et Jenkins

C'est une méthode qui permet en premier étape de modéliser une série chronologique par un modèle : AR, MA, ARMA, ARIMA et SARIMA.

L'idée de Box et Jenkins est de ne tenir compte que des observations qui sont fortement liées avec le présent.

2.6.1 Fonction autocovariance et autocorrélation :

a) Fonction autocovariance :

Soit (Y_t) un processus stationnaire, on appelle fonction d'autocovariance la fonction γ définie pour tout $t \in Z$ par :

$$\begin{aligned} \forall h, t \in Z, \gamma_h &= Cov(Y_t; Y_{t-h}) \\ &= E((Y_t - E(Y_t))(Y_{t-h} - E(Y_{t-h}))) \end{aligned}$$

L'expression de γ_h est estimée par l'autocovariance empirique :

$$\hat{\gamma}_h = \frac{1}{n} \sum_{t=h+1}^n (Y_t - \bar{Y})(Y_{t-h} - \bar{Y})$$

où \bar{Y} désigne la moyenne de la série calculé sur n périodes [2, 3]

b) Fonction d'autocorrélation :

Soit (Y_t) un processus stationnaire, on appelle fonction d'autocorrélation la fonction ρ définie pour tout $t \in (Z)$ par :

$$\forall h, t \in Z, \rho_h = Corr(Y_t; X_{t-h}) = \frac{\gamma_h}{\gamma_0}.$$

– si ρ_h est proche de -1 cela signifie que Y_t et Y_{t-h} sont fortement liées.

– si ρ_h est proche de 0 cela signifie que Y_t et Y_{t-h} ne sont pas liées.

L'expression de ρ_h est estimée par l'autocovariance empirique :

$$\hat{\rho}_h = \frac{\sum_{t=h+1}^n (Y_t - \bar{Y})(Y_{t-h} - \bar{Y})}{\sum_{t=1}^n (Y_t - \bar{Y})^2}$$

où \bar{Y} désigne la moyenne de la série calculé sur n périodes [2, 3] Le graphique de la suite des ρ_h constitue le corrélogramme.

c) Définition de la fonction d'autocorrélation partielle

On appelle fonction d'autocorrélation partielle la fonction d'autocorrélation entre Y_t et Y_{t-h} en éliminant les autres variables $Y_{t-1}, Y_{t-2}, \dots, Y_{t-(h-1)}$.

$$\pi_h = \frac{cov(Y_t - Y_t^*, Y_{t-h} - Y_{t-h}^*)}{V(Y_t - Y_t^*)}$$

où Y_t^* désigne l'observation estimée de la variable à expliquer Y_t par la régression linéaire multiple $Y_t^* = \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \dots + \alpha_{h-1} Y_{t-(h-1)}$.
 et Y_{t-k}^* désigne l'observation estimée de la variable à expliquer Y_{t-h}^* par la régression linéaire multiple $Y_{t-h}^* = \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \dots + \beta_{h-1} Y_{t-(h-1)}$. Il s'agit du coefficient de corrélation entre Y_t et Y_{t-h} , l'influence des autres variables $Y_{t-1}, Y_{t-2}, \dots, Y_{t-(h-1)}$, ayant été retirée [3]. La représentation graphique des π_h est appelée corrélogramme partiel.

L'intérêt pratique des fonctions d'autocovariance, d'autocorrélation et d'autocorrélation partielle se retrouve dans l'étude des processus stationnaires tels que les processus autorégressifs (notés AR), les processus de moyenne mobile (notés MA) et les processus mixtes autorégressifs et de moyenne mobile (notés ARMA) [3].

2.6.2 Les processus autorégressifs d'ordre p : AR(p)

Un processus autorégressif (Auto Regressive model) d'ordre p retrace la façon dont est générée l'observation présente y_t à partir de ses observations passées $y_{t-1}, y_{t-2}, \dots, y_{t-p}$. Il est dit autorégressif d'ordre p, et on le note AR(p), s'il vérifie une relation de la forme :

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \epsilon_t, \text{ pour tout } t \in \mathbb{Z},$$

ou encore, En introduisant l'opérateur retard B tel que $B^j y_t = y_{t-j}$;

$$\begin{aligned} y_t - \phi_1 y_{t-1} - \phi_2 y_{t-2} - \dots - \phi_p y_{t-p} &= \epsilon_t \iff \\ y_t &= \phi_1 B y_t + \phi_2 B^2 y_t + \dots + \phi_p B^p y_t = \epsilon_t \iff \\ (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) y_t &= \epsilon_t \iff \\ \Phi(B) y_t &= \epsilon_t \end{aligned}$$

où $\phi_1, \phi_2, \dots, \phi_p$ désignent les paramètres du modèle, (ϵ_t) est un bruit blanc et $\Phi(B) = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)$. Un processus AR(p) peut aussi comporter un terme constant [2, 3, 4]

2.6.3 Le processus AR(1)

La forme générale des processus de type AR(1) est :

$$y_t = \phi_1 y_{t-1} + \epsilon_t$$

ou encore, on introduisant l'opérateur retard B tel que :

$$y_t - \phi_1 y_{t-1} = \epsilon_t \iff y_t - \phi_1 B y_t = \epsilon_t \iff (1 - \phi_1 B) y_t = \epsilon_t \iff \Phi(B) y_t = \epsilon_t$$

où ϕ_1 désigne le paramètre du modèle, (ϵ_t) est un bruit blanc et $\Phi(B) = (1 - \phi_1 B)$.

Stationnarité et inversibilité du processus AR(1)

Afin de déterminer la condition de stationnarité d'un processus AR(1), on exprime y_t sous forme d'une série de ϵ_t ; soit :

$$y_t = \phi_1 y_{t-1} + \epsilon_t = \phi_1(\phi_1 y_{t-2} + \epsilon_{t-1}) + \epsilon_t = \phi_1^2 y_{t-2} + \phi_1 \epsilon_{t-1} + \epsilon_t$$

et de proche en proche, on arrive à :

$$y_t = \phi_1^{t-1} y_1 + \phi_1^{t-2} \epsilon_2 + \dots + \phi_1^2 \epsilon_{t-2} + \phi_1 \epsilon_{t-1} + \epsilon_t$$

en utilisant la condition d'initialisation du processus $y_0 = 0$, on obtient $y_1 = \epsilon_1$ et donc :

$$y_t = \phi_1^{t-1} \epsilon_1 + \phi_1^{t-2} \epsilon_2 + \dots + \phi_1^2 \epsilon_{t-2} + \phi_1 \epsilon_{t-1} + \epsilon_t$$

il en résulte, du fait que (ϵ_t) est un bruit blanc que :

$$\begin{cases} E(y_t) = 0, \\ V(y_t) = (1 + \phi_1^2 + (\phi_1^2)^2 + \dots + (\phi_1^2)^{t-1})\sigma^2 = \left(\frac{1 - (\phi_1^2)^t}{1 - \phi_1^2}\right)\sigma^2 \\ cov(y_t, y_{t-h}) = \phi_1^h (1 + \phi_1^2 + (\phi_1^2)^2 + \dots + (\phi_1^2)^{t-h-1})\sigma^2 = \phi_1^h \left(\frac{1 - (\phi_1^2)^{t-h}}{1 - \phi_1^2}\right)\sigma^2 \end{cases}$$

avec : $\sigma^2 = V(\epsilon_t)$

on constate que, pour $|\phi_1| < 1$,

$$\begin{cases} E(y_t) = 0, \\ \lim_{t \rightarrow \infty} V(y_t) = \left(\frac{\sigma^2}{1 - \phi_1^2}\right) \\ cov(y_t, y_{t-h}) = \left(\frac{\phi_1^h \sigma^2}{1 - \phi_1^2}\right) \end{cases}$$

le processus (y_t) est donc asymptotiquement stationnaire, ainsi que $|\phi_1| < 1$ est la condition de stationnarité d'un processus AR(1).

Propriété d'inversibilité

Un processus AR(1) est stationnaire si les racines du polynôme retard $1 - \phi_1 z$ sont de module supérieur à 1.

En effet, $1 - \phi_1 z = 0 \iff z = \frac{1}{\phi_1}$.

ainsi, si $|\phi_1| < 1$, $|z| = \left|\frac{1}{\phi_1}\right| > 1$, d'où le résultat.

Identification d'un processus AR(1)

L'identification d'un processus AR(1) peut se faire par sa fonction d'autocorrélation et sa fonction d'autocorrélation partielle. En effet, un processus AR(1) est caractérisé par une décroissance géométrique des termes de sa fonction d'autocorrélation de type :

$$\rho_h = \phi_1^h$$

et par le fait que sa fonction d'autocorrélation partielle est tronquée au-delà de l'ordre 1 :

$$\pi_h = \begin{cases} \rho_1 & \text{si } h = 1 \\ 0 & \text{si } h > 1 \end{cases}$$

Stationnarité et inversibilité du processus AR(p)

Le polynôme ϕ du processus AR(p) peut s'écrire

$$\Phi(B) = \prod_{i=1}^p (1 - \lambda_i B)$$

comme pour le processus AR(1), afin d'obtenir la stationnarité, il est nécessaire que $|\lambda_i| < 1$. cette condition de stationnarité correspond à des racines du polynome $\Phi(B) = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)$ supérieures à 1 en module (propriété d'inversibilité) [4]

2.6.4 Les processus de moyennes mobiles : MA(q)

On appelle processus moyenne mobile (moving average) d'ordre q, noté MA(q), un processus vérifiant la relation :

$$y_t = \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \dots - \theta_q \epsilon_{t-q}, \text{ pour tout } t \in \mathbb{Z},$$

ou encore en introduisant l'opérateur retard B tel que $B^j \epsilon_t = \epsilon_{t-j}$,

$$y_t = \Theta(B) \epsilon_t$$

où $\theta_1, \theta_2, \dots, \theta_q$ désignent les paramètres du modèle, (ϵ_t) est un bruit blanc et $\Theta(B) = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q)$.

Contrairement aux processus AR(p), les processus MA(q) sont toujours des processus stationnaires puisque y_t s'écrit comme combinaison linéaire des ϵ_t [2, 3, 4]

2.6.5 Le processus MA(1)

La forme générale des processus de type MA(1) est $y_t = \epsilon_t - \theta_1 \epsilon_{t-1}$. Ou encore, en introduisant l'opérateur retard B , $y_t = \epsilon_t - \theta_1 \epsilon_{t-1} = (1 - \theta_1 B) \epsilon_t = \Theta(B) \epsilon_t$, où θ_1 désigne le paramètre du modèle, (ϵ_t) est un bruit blanc et $\Theta(B) = (1 - \theta_1 B)$

Condition d'inversibilité d'un processus MA(1)

Afin de déterminer la condition d'inversibilité d'un processus MA(1), on procède comme suit :

$$\begin{aligned} y_t &= \epsilon_t - \theta_1 \epsilon_{t-1} = -\theta_1 \epsilon_{t-1} + \epsilon_t \\ y_t &= -\theta_1 (y_{t-1} + \theta_1 \epsilon_{t-2}) + \epsilon_t = -\theta_1 y_{t-1} - \theta_1^2 \epsilon_{t-2} + \epsilon_t \\ &= \dots = -\theta_1 y_{t-1} - \theta_1^2 y_{t-2} - \theta_1^3 y_{t-3} - \dots + \epsilon_t \end{aligned}$$

Si $\theta_1 > 1$ ou $\theta_1 < -1$ le poids du passé va en grandissant, ce qui est absurde. Alors, les seules valeurs acceptables pour θ_1 soient comprises entre -1 et +1.

Ainsi, si $|\theta_1| < 1$ est la condition d'inversibilité d'un processus MA(1). Un processus MA(1) est inversible si les racines du polynôme $(1 - \theta_1 z)$ sont à l'extérieur du disque unité. En effet,

$$1 - \theta_1 z = 0 \iff z = \frac{1}{\theta_1}$$

Ainsi, si $|\theta_1| < 1$, $|z| = \left|\frac{1}{\theta_1}\right| > 1$, d'où le résultat.

Condition d'inversibilité d'un processus MA(q)

Le polynôme Θ du processus MA(q) peut s'écrire

$$\theta(z) = \prod_{i=1}^q (1 - \lambda_i z)$$

comme pour le processus MA(1), afin d'obtenir la condition d'inversibilité, il faut que $|\lambda_i| < 1$. Cette condition d'inversibilité correspond à des racines du polynôme

$\Theta(B) = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q)$ supérieur à 1 en module.

La fonction d'autocorrélation d'un processus MA(q)

La fonction d'autocorrélation d'un processus MA(q) est de la forme générale

$$\rho_h = \begin{cases} \frac{(-\theta_h + \theta_1 \theta_{h+1} + \dots + \theta_{q-h} \theta_q)}{1 + \theta_1^2 + \dots + \theta_q^2}, & h = 1 \dots q; \\ 0, & h > q. \end{cases}$$

La fonction d'autocorrélation est donc tronquée au-delà du retard q .

La fonction d'autocorrélation partielle d'un processus MA(q)

Contrairement aux processus AR(p), on montre que la fonction d'autocorrélation partielle d'un processus MA(q) n'est pas tronquée au-delà de l'ordre q.

En faisant varier h de 1 à q et en remplaçant ρ_h par $\hat{\rho}_h$, on obtient un système non linéaire de q inconnues à q équations suivant :

$$\left\{ \begin{array}{l} \hat{\rho}_1 = \frac{(-\theta_1 + \theta_1 \theta_2 + \dots + \theta_{q-1} \theta_q)}{1 + \theta_1^2 + \dots + \theta_q^2}, \\ \hat{\rho}_2 = \frac{(-\theta_2 + \theta_1 \theta_3 + \dots + \theta_{q-2} \theta_q)}{1 + \theta_1^2 + \dots + \theta_q^2}, \\ \cdot \\ \cdot \\ \cdot \\ \hat{\rho}_q = \frac{(-\theta_q)}{1 + \theta_1^2 + \dots + \theta_q^2}. \end{array} \right.$$

L'estimation préliminaire des paramètres $\theta_1, \theta_2, \dots, \theta_q$ d'un processus MA(q) s'obtient en résolvant ce système [3].

2.6.6 Les processus autorégressifs et de moyennes mobiles : ARMA (p, q)

Un processus y_t ; $t \in \mathbb{Z}$ est dit autorégressif d'ordre p et de moyennes mobiles d'ordre q, s'il peut être représenté sous la forme :

$$\begin{aligned} y_t - \phi_1 y_{t-1} - \phi_2 y_{t-2} - \dots - \phi_p y_{t-p} &= \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \dots - \theta_q \epsilon_{t-q} \\ \iff \Phi(B)y_t &= \Theta(B)\epsilon_t \end{aligned}$$

où :

$$\Phi(B) = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) \text{ et } \Theta(B) = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q)$$

les coefficients $\phi_1, \phi_2, \dots, \phi_p, \theta_1, \theta_2, \dots, \theta_q$ sont les paramètres du modèle et ϵ_t est un bruit blanc.

Pour obtenir un processus ARMA (p, q) stationnaire, il est nécessaire que les racines de $\Phi(B)y_t = 0$ soient en module supérieure à 1.

Pour que le processus ARMA (p, q) soit inversible, il faut que les racines de $\Theta(B)\epsilon_t = 0$ soient en module supérieure à 1.

Les corrélogrammes et les corrélogrammes partiels sont, par voie de conséquence, un mélange des deux corrélogrammes des processus AR et MA purs [2, 3, 4]

Dans le tableau 2.1, Bourbonnais et Terraza (2004) proposent un récapitulatif sur les formes des fonctions d'autocorrélation et d'autocorrélation partielles théoriques des processus AR(p), MA(q) et ARMA (p, q).

L'estimation des paramètres dans un modèle ARMA (p, q), si l'hypothèse de normalité des erreurs est réalisée, se fait par la méthode du maximum de vraisemblance.

modèle	autocorrélation	autocorrélation partielle
Bruit blanc	Nulle pour $k < 0$	Nulle pour $k > 0$
AR(p) : $\Phi(B)X_t = \epsilon_t$	Décroissance exponentielle et/ou sinusoidale	Pics significatifs pour les p premiers retards, les autres coefficients sont nuls pour des retards $> p$
MA(q) : $X_t = \theta(B)\epsilon_t$	Pics significatifs pour les q premiers retards, les autres coefficients sont nuls pour des retards $> q$	Décroissance exponentielle et/ou sinusoidale
ARMA(p,q) : $\Phi(B)X_t = \theta(B)\epsilon_t$	Décroissance exponentielle ou sinusoidale amortie	Décroissance exponentielle ou sinusoidale amortie

TABLE 2.3 – Résumé des propriétés des fonctions d'autocorrélations simples et partielles

Lorsque les paramètres du modèle sont estimés, il faut vérifier si les conditions de stationnarité et d'inversibilité sont satisfaites. Les coefficients peuvent être testés individuellement à l'aide du test de student et le modèle peut être testé globalement à l'aide du test de Box et Pierce ou à partir du test de Ljung et Box.

2.6.7 Les processus ARIMA

Les modèles AR, MA et ARMA ne sont représentatifs que des processus stationnaires en tendance et corrigés des variations saisonnières. Si le processus étudié n'est pas stationnaire en tendance, il convient de le stationnariser par passage aux différences selon l'ordre d'intégration d (c'est-à-dire le nombre de fois qu'il faut différencier le processus pour le rendre stationnaire) par la transformation [4] :

$$\Phi(B)\nabla^d y_t = \Theta(B)\epsilon_t$$

avec :

$$\begin{aligned}\nabla y_t &= y_t - y_{t-1} = (1 - B)y_t \\ \nabla^2 y_t &= \nabla(\nabla y_t) = \nabla(y_t - y_{t-1}) = (y_t - y_{t-1}) - (y_{t-1} - y_{t-2}) \\ &= y_t - 2y_{t-1} + y_{t-2} = (1 - 2B + B^2)y_t = (1 - B)^2 y_t \\ &\dots\dots\dots \\ \nabla^d y_t &= (1 - B)^d y_t\end{aligned}$$

∇^d est appelé l'opérateur de différenciation à l'ordre d (le but de cet opérateur est donc d'éliminer la tendance).

Par définition, on appelle modèle ARIMA (autoregressive integrated moving average model) de paramètres p, d, q , un modèle de la forme :

$$\Phi(B)\nabla^d y_t = \Theta(B)\epsilon_t$$

où $\Phi(B) = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)$ et $\Theta(B) = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q)$

2.6.8 Les processus SARIMA

Si le processus étudié comporte une saisonnalité (une valeur élevée du coefficient de corrélation ρ_s entre y_t et y_{t-s}) alors il s'agit d'un processus à variation saisonnières (donc non stationnaire), il convient de le stationnariser (élimination de la saisonnalité) [4] en utilisant l'opérateur de différenciation saisonnière ∇_s défini par :

$$\nabla_s y_t = y_t - y_{t-s} = y_t - B^s y_t = (1 - B^s) y_t$$

on se ramène alors à un modèle

$$\Phi(B) \nabla_s y_t = \Theta(B) \epsilon_t \iff \Phi(B) (1 - B^s) y_t = \Theta(B) \epsilon_t$$

Lorsque l'opérateur de différenciation saisonnière est d'ordre S (c'est-à-dire que y_t est différencié S fois d'une période s), on utilise l'opérateur de différenciation d'ordre S défini par :

$$\nabla_s^S = (1 - B^s)^S y_t$$

par exemple, pour $S = 2$, on vérifie bien que

$$\nabla_s^2 = (1 - B^s)^2 y_t$$

En effet

$$\begin{aligned} \nabla y_t &= y_t - y_{t-1} = (1 - B) y_t \\ \nabla_s^2 y_t &= \nabla_s (\nabla_s y_t) = \nabla_s (y_t - y_{t-s}) = (y_t - y_{t-s}) - (y_{t-s} - y_{t-s-s}) = (y_t - y_{t-s}) - (y_{t-s} - y_{t-2s}) \\ &= y_t - 2y_{t-s} + y_{t-2s} = (1 - 2B^s + B^{2s}) y_t = (1 - B^s)^2 y_t \end{aligned}$$

Dans ce cas, le modèle transformé par l'opérateur de différenciation saisonnière d'ordre S s'écrit :

$$\Phi(B) \nabla_s^S y_t = \Theta(B) \epsilon_t \iff \Phi(B) (1 - B^s)^S y_t = \Theta(B) \epsilon_t$$

Il s'agit d'un cas particulier d'un modèle général mixte multiplicatif saisonnier SARIMA qui s'écrit sous la forme :

$$\Phi(B) \Psi(B^s) \nabla^d y_t \nabla_s^S y_t = \Theta(B) \Lambda(B^s) \epsilon_t$$

avec

$$\Psi(B^s) = (1 - \alpha_s B^s - \alpha_{2s} B^{2s} - \dots - \alpha_{P_s} B^{P_s s})$$

et

$$\Lambda(B^s) = (1 - \beta_s B^s - \beta_{2s} B^{2s} - \dots - \beta_{Q_s} B^{Q_s s})$$

Où (ϵ_t) est un bruit blanc, Φ et Ψ des opérateurs autorégressifs de degrés respectifs p et P , Φ et Λ des opérateurs de moyennes mobiles de degrés respectifs q et Q .

Un tel modèle est un modèle SARIMA $(p, d, q) (P, S, Q)_s$

Nous pouvons résumer les différentes étapes de la méthodologie de Box et Jenkins Selon l'organigramme 2.1, [5] :

Etape 1 : Analyse des caractéristiques de stationnarité du processus : élimination de la tendance et de la saisonnalité,

Etape 2 : Identification du modèle par les fonctions d'autocorrélation et d'autocorrélation partielle,

Etape 3 : Estimation des paramètres du modèle,

Etape 4 : Validation du modèle en utilisant le test de Ljung-Box (ou le test de Box-Pierce),

Etape 5 :Prévision.

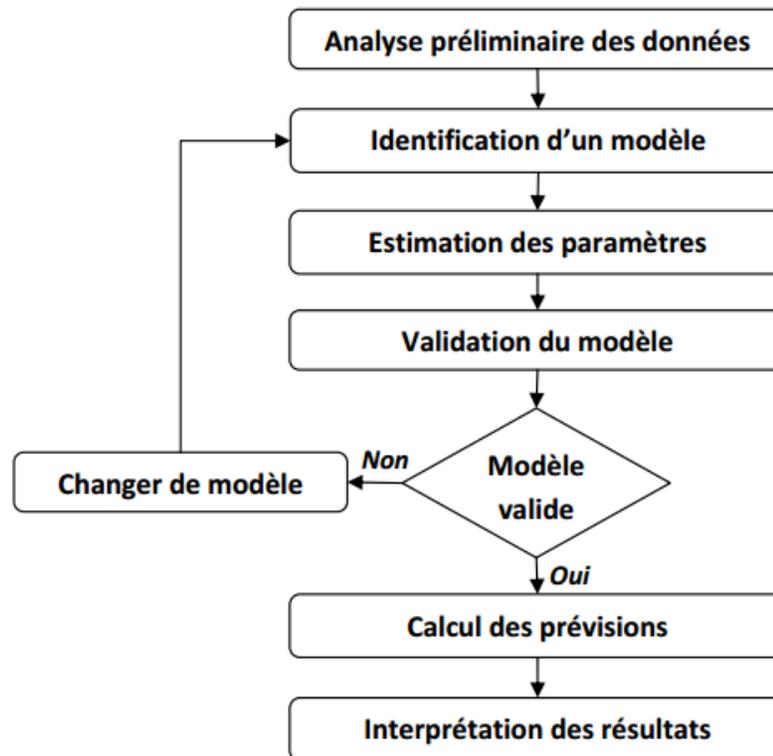


FIGURE 2.1 – Organigramme de la démarche Box et Jenkins

2.6.9 Estimation des paramètres

Après avoir vérifié tous les cas possibles et opté pour un modèle vérifiant toute conditions nécessaires à l'application de la méthode de Box et Jenkins, on procède à l'estimation des paramètres afin de déterminer le modèle. Dans les cas simples on peut évaluer les au-

tocorrélations du processus en nombre égal au nombre des paramètres à estimer. Cette méthode appelée en statistique méthode des moments est lourde et souvent peu précise, on préfère les méthodes les plus générales de minimisation d'un critère tels que l'erreur moyenne carrée, ou l'erreur moyenne absolue MAE. La méthode de Box et Jenkins est associée au critère MSE et au principe des moindres carrés.

2.6.10 Validation du Modèle

1. Vérifier que les coefficients du modèle satisfont aux conditions de stationnarité et d'inversibilité.
2. On teste si les paramètres du modèle sont significatifs.

Une fois les paramètres du modèle estimés, nous examinons les résultats d'estimation. Cet examen se fait par plusieurs tests qu'on classera en deux catégories :

- Tests concernant les paramètres du modèle.
- Tests concernant les résidus.

• Tests concernant les paramètres du modèle

- Test de Student [11]

$H_0 : \theta_i = 0$ contre $H_1 : \theta_i \neq 0$

Statistique de décision

Ce test est basé sur la statistique :

$$T_i = \frac{|\hat{\theta}_i|}{\hat{\sigma}(\hat{\theta}_i)}$$

où $\hat{\theta}_i$ est l'estimateur de θ_i , $\hat{\sigma}(\hat{\theta}_i)$, et l'écart type associé à $\hat{\theta}_i$.

On montre que sous H_0 , T suit une loi de Student à $(n - k - 1)$ degré de liberté.

La région critique est :

$$RC = [T_i > T_{tab}]$$

où $T_{tab} = T_{(n-k-1, \frac{\alpha}{2})}$ est la valeur théorique lue sur la table de Student à $(n - k - 1)$ degré de liberté, au seuil de signification α .

tel que k est le nombre de paramètres du modèle.

n est la taille de l'échantillon.

Règle de décision

Si $T_i > T_{(n-k-1, \frac{\alpha}{2})}$ on rejette H_0 , donc le paramètre θ_i est significatif.

Si $T_i < T_{(n-k-1, \frac{\alpha}{2})}$ on accepte H_0 , donc le paramètre θ_i n'est pas significatif.

Remarque

On peut aussi procéder à l'estimation par intervalle de confiance. Si les bornes de l'intervalle sont de même signe alors le paramètre ne peut pas prendre la valeur zéro, dans ce cas ce paramètre peut être considéré significativement différent de zéro au seuil de signification α donné. Le résultat obtenu par le test de student est toujours confirmé par l'estimation par intervalle de confiance.

● **Tests concernant les résidus** Ces tests ont pour objet de vérifier que les résidus obtenus suivent bien un processus de Bruit Blanc. Cela se fait à l'aide des tests suivants :

- **Test de Box-Pierce [5]** Soit ρ_k l'autocorrélation d'ordre k du processus $(\varepsilon_t, t \in \mathbb{Z})$.

Les hypothèses de test de Box et Pierce sont les suivantes :

$$\begin{cases} H_0 : \rho_1 = \rho_2 = \dots = \rho_k, \\ \text{contre} \\ \exists i = \overline{1, k}, \rho_i \neq 0, \end{cases}$$

La statistique du test

$$Q = n \sum_{i=1}^k \rho_i^2 \rightsquigarrow \chi_k^2$$

où n est le nombre d'observation

La règle de décision

-Si $Q \leq \chi_k^2(\alpha)$ on accepte H_0 (les résidus forment un bruit blanc). Alors le modèle est bon.

-Si $Q > \chi_k^2(\alpha)$, alors on rejette l'hypothèse H_0 de bruit blanc au seuil de signification α et on constate que le modèle estimé est mauvais. La valeur $\chi_k^2(\alpha)$ est lue dans la table de Khi-deux à k degré de liberté au seuil de signification α .

- **Test de Ljung et Box [11]**

On peut également utiliser la statistique de Ljung et Box (dont les propriétés sont meilleures). Ce test est à appliquer de préférence au test de Box-Pierce, lorsque l'échantillon est de petite taille. La distribution de la statistique du test de Ljung-Box est en effet plus proche de celle du Khideux en petit échantillon que ne l'est celle du test de Box -pierce.

La statistique du test

$$Q = n(n+2) \sum_{j=1}^M \frac{\rho_j^2}{n-j} \rightsquigarrow \chi_{M-k}^2$$

Où n désigne la taille de l'échantillon, ρ_j désigne les coefficients d'autocorrélation théoriques, M est la partie entière de $\min(\frac{n}{2}, 3\sqrt{n})$.

Cette statistique est aussi distribuée selon une χ^2 à $(M - k)$ degrés de liberté.

Règle de décision

Pour $\alpha = 0.05$, Q^* est la réalisation de la statistique Q , $\chi_\alpha^2(M - k)$ est la valeur théorique lue sur la table de khi-deux à $(M - k)$ degré de liberté, on aura :

- Si $Q^* > \chi_\alpha^2(M - k)$ alors on rejette l'hypothèse H_0 , de bruit blanc au seuil de signification α , donc le modèle estimé n'est pas bon.
-
- Si $Q^* < \chi_\alpha^2(M - k)$ alors on accepte l'hypothèse H_0 , donc les résidus forment un bruit blanc, le modèle est adéquat.

- Test de normalité des résidus "Test de Shapiro-Wilk" [6]

Très populaire, le test de Shapiro-Wilk est basé sur la statistique W . En comparaison des autres tests, il est particulièrement puissant pour les petits effectifs ($n < 50$). La statistique du test s'écrit :

$$W = \frac{[\sum_{i=1}^{n/2} a_i(y_{(n-i+1)} - y_i)]^2}{\sum_i (y_i - \bar{y})}$$

Où : y_i correspond à la série des données triées.

$[n/2]$ est la partie entière du rapport $n/2$.

a_i sont des constantes générées à partir de la moyenne et de la matrice de variance covariance des quantiles d'un échantillon de taille n suivant la loi normale. Ces constantes sont fournies dans des tables spécifiques. La statistique W peut donc être interprété comme le coefficient de détermination (le carré du coefficient de corrélation) entre la série des quantiles générés à partir de la loi normale et les quantiles empiriques obtenus à partir des données. Plus W est élevé, plus la compatibilité avec la loi normale est crédible. La région critique s'écrit :

$$R.C : W < W_{crit}$$

Les valeurs seuils W_{crit} pour différents risque α et les effectifs n sont lues dans la table de Shapiro-Wilk.

2.6.11 Prévision

La prévision qui minimise l'erreur à l'instant $n + h$ est donnée par l'espérance conditionnelle de y_{n+h} quand y_n, y_{n-1}, \dots, y_1 sont connu :

$$\hat{y}_{n+h} = E(y_{n+h} | y_n, y_{n-1}, \dots, y_1).$$

Cette espérance conditionnelle s'obtient en ne tenant compte que des termes d'indice inférieur ou égal à n pour ce qui concerne les erreurs.

Exemple : Considérons un modèle ARMA(1,1) :

$$y_t - \phi_1 y_{t-1} = \epsilon_t - \theta_1 \epsilon_{t-1}, t = 1, \dots, n$$

la variable à l'horizon h s'écrit :

$$y_{n+h} - \phi_1 y_{n+h-1} = \epsilon_{n+h} - \theta_1 \epsilon_{n+h-1}.$$

pour $h = 1$, on aura :

$$y_{n+1} = \phi_1 y_n + \epsilon_{n+1} - \theta_1 \epsilon_n.$$

la prévision qui minimise l'erreur à l'instant $n + 1$ est :

$$\begin{aligned} \hat{y}_{n+1} &= E(y_{n+1} | y_n, y_{n-1}, \dots, y_1) \\ &= \phi_1 E(y_n | y_n, y_{n-1}, \dots, y_1) + E(\epsilon_{n+1} | y_n, y_{n-1}, \dots, y_1) - \theta_1 E(\epsilon_n | y_n, y_{n-1}, \dots, y_1) \\ &= \phi_1 y_n - \theta_1 \epsilon_n. \end{aligned}$$

car $E(\epsilon_{n+1} | y_n, y_{n-1}, \dots, y_1) = 0$ (on ne tient pas compte de ϵ_{n+1}).

Si $h = 2$, on aura :

$$y_{n+2} = \phi_1 y_{n+1} + \epsilon_{n+2} - \theta_1 \epsilon_{n+1}.$$

la prévision qui minimise l'erreur à l'instant $n + 2$ est :

$$\begin{aligned} \hat{y}_{n+2} &= E(y_{n+2} | y_n, y_{n-1}, \dots, y_1) \\ &= \phi_1 E(y_{n+1} | y_n, y_{n-1}, \dots, y_1) + E(\epsilon_{n+2} | y_n, y_{n-1}, \dots, y_1) - \theta_1 E(\epsilon_{n+1} | y_n, y_{n-1}, \dots, y_1) \\ &= \phi_1 E(y_{n+1} | y_n, y_{n-1}, \dots, y_1). \\ &= \phi_1 \hat{y}_{n+1} \end{aligned}$$

car $E(\epsilon_{n+1} | y_n, y_{n-1}, \dots, y_1) = E(\epsilon_{n+2} | y_n, y_{n-1}, \dots, y_1) = 0$ (on ne tient pas compte de ϵ_{n+1} et de ϵ_{n+2}).

de proche en proche, on montre que :

$$\hat{y}_{n+3} = \phi_1^2 \hat{y}_{n+1}, \dots, \hat{y}_{n+h} = \phi_1^{h-1} \hat{y}_{n+1}$$

avec :

$$\hat{y}_{n+1} = \phi_1 y_n - \theta_1 \epsilon_n$$

Cette prévision dépend des vraies valeurs des paramètres ϕ_1 et θ_1 : Dans la pratique, on utilise les valeurs estimées $\hat{\phi}_1$ et $\hat{\theta}_1$.

2.7 Conclusion

Dans ce chapitre, nous avons présenté différentes méthodes statistiques de prévision à savoir la méthode de Lissage exponentiel ainsi que la méthode de Box & Jenkins qui seront appliquées sur des données réelles (la pluviométrie) dans le chapitre suivant.

Chapitre 3

Application de la méthode de Box-Jenkins et du lissage exponentiel

3.1 Introduction

Cette partie concerne la modélisation statistique des données climatiques obtenues au sein du barrage de Béni Haroun. Pour ce faire, nous appliquerons deux méthodes, les modèles de Box et Jenkins et les techniques du lissage exponentiel. En utilisant le logiciel R, nous allons déterminer les prévisions des quantités pluviométriques en millimètres sur une durée de douze mois et de six ans.

3.2 Présentation du logiciel R



R est un langage de programmation et un environnement mathématique utilisés pour le traitement de données et l'analyse statistique. C'est un projet GNU (General Public

Licence) fondé sur le langage S et sur l'environnement développé dans les laboratoires Bell par John Chambers et ses collègues.

Le logiciel R a été développé dans les années 90 par Robert Gentleman et Ross Ihaka (Département de Statistique, Université d'Auckland, Nouvelle-Zélande), c'est un logiciel libre (avec code source) et peut être distribué librement.

Il est le plus populaire dans la communauté des statisticiens-chercheurs car :

- de nombreuses techniques statistiques modernes et classiques ont été implémentées.
- L'utilisateur n'est pas limitée par les procédures préprogrammées par un package.
- Il comporte de nombreuses fonctions pour les analyses statistiques et les graphiques, ceux-ci sont visualisé immédiatement dans une fenêtre propre et peuvent être exportés sous divers formats jpg, png, bmp, ps, pdf, emf, pictex et xfig. Les résultats des analyses statistiques sont affichés à l'écran, certains résultats partiels (valeurs de P, coefficients de régression, résidus,...) peuvent être sauvés à part, exportées dans un fichier ou utilisés dans des analyses ultérieures..

D'après une enquête menée par Rexer Analytics auprès de 1300 analystes, R est le logiciel le plus utilisé aussi bien chez les analystes qui travaillent en entreprise, ceux qui travaillent dans le monde académique, ceux qui travaillent pour des organismes publics ou des ONG ou encore ceux qui travaillent comme consultants.

R comporte des moyens qui rendent possibles la manipulation des données et les calculs. R a aussi la possibilité d'exécuter des programmes stockés dans des fichiers textes.

En effet R possède :

- un système efficace de manipulation et de stockage des données,
- différents opérateurs pour le calcul sur tableaux, en particulier les matrices,
- un grand nombre d'outils pour l'analyse des données et les méthodes statistiques,
- des moyens graphiques pour visualiser les analyses,
- un langage de programmation simple et performant comportant : conditions, boucles.

Le logiciel R est un logiciel performant en termes de calculs et de représentations graphiques. C'est pourquoi, dans le cadre de notre études statistiques, nous l'avons utilisé afin d'estimer les modèles statistiques et de calculer les prévisions.

3.3 Modélisation et prévision

La modélisation est une représentation d'une série chronologique par un modèle qui décrit son évolution dans le temps. Plusieurs méthodes de modélisations ont été établies, citons en particulier, les méthodes du lissage exponentiel et la méthode de Box et Jenkins. La prévision aide à élaborer des plans de production (prévision à court terme) et peut être d'une grande utilité dans la planification de l'acquisition de ressources humaines, financières ou matérielles (prévision à long terme). La prévision d'une série consiste à déterminer son évolution sur les périodes non observées.

3.4 Processus de prévision

- Etablir le but des prévisions.
- Définir l'horizon de prévision.
- Recueillir les données pertinentes.
- Modélisation : choisir une technique (modèle) de prévision.
- Etablir les prévisions.
- Validation : vérifier si les erreurs forment un bruit blanc en se basant sur les données disponibles.
- Procéder à l'implémentation et suivre l'évolution des prévisions.

3.5 Collecte des données

Les données dont nous disposons proviennent des annuaires statistiques établis par le service d'exploitation de la Direction du barrage du Béni Haroun, qui regroupe les données journalière comme la température du barrage, le volume d'évaporation, cote, vitesse de vent, pluviométrie. . . . Nous nous sommes intéressées aux données concernant la climatologie :

- La pluviométrie en millimètres par mois

Les données de cette série sont cumulées de janvier 2010 jusqu'à décembre 2015.

Après la collecte de ces données, nous les avons regroupées dans un tableau de 12 mois sur 6 ans.

Dans le but de prévoir la pluviométrie en millimètres pour l'année 2016, nous appliquerons la méthode du lissage exponentiel et celle de Box et Jenkins.

Le tableau 3.1 représente nos données :

mois \ anné	2010	2011	2012	2013	2014	2015
Jan	96	21,4	52,5	112	69	195
Fév	47,5	160	197,5	180	59	192
Mar	56,5	38	40,5	69	253	98
Avr	34,6	60	93,5	30	10	4,5
Mai	81	1,3	6,5	39	24,5	26,6
Jui	22	0	10,5	3,5	21	3
Juil	2,5	0	2	3,5	0	0
Aou	3	0	4	26,5	0	17,5
Sept	33	25	49	27	0,5	38
Oct	74	89	9,5	54,5	5,5	82,5
Nov	100	57	35	248,5	63	82,5
Déc	61,1	99	55,5	67,5	202	0,5

TABLE 3.1 – La pluviométrie en millimètres

3.6 Application de la méthode du lissage exponentiel

Afin de tracer le graphe de la série originale, il faut introduire les données sous forme d'un vecteur dans le logiciel R puis appliquer les commandes suivantes qui permettent de le rendre sous forme d'une série chronologique.

```
> y=c(96,47.5,56.5,34.6,81,22,2.5,3,33,74,100,61.1,21.4,
160,38,60,1.3,0,0,0,25,89,57,99,52.5,197.5,40.5,93.5,6.5,
10.5,2,4,49,9.5,35,55.5,112,180,69,30,39,3.5,3.5,26.5,27,
54.5,248.5,67.5,69,59,253,10,24.5,21,0,0,0.5,5.5,63,202,
195,192,98,4.5,26.6,3,0,17.5,38,82.5,82.5,0.5)
> serie=ts(y,start=c(2010,1),end=c(2015,12),frequency=12)
```

La commande qui nous permet de visualiser le graphe de la série originale est :

```
> plot(serie)
```

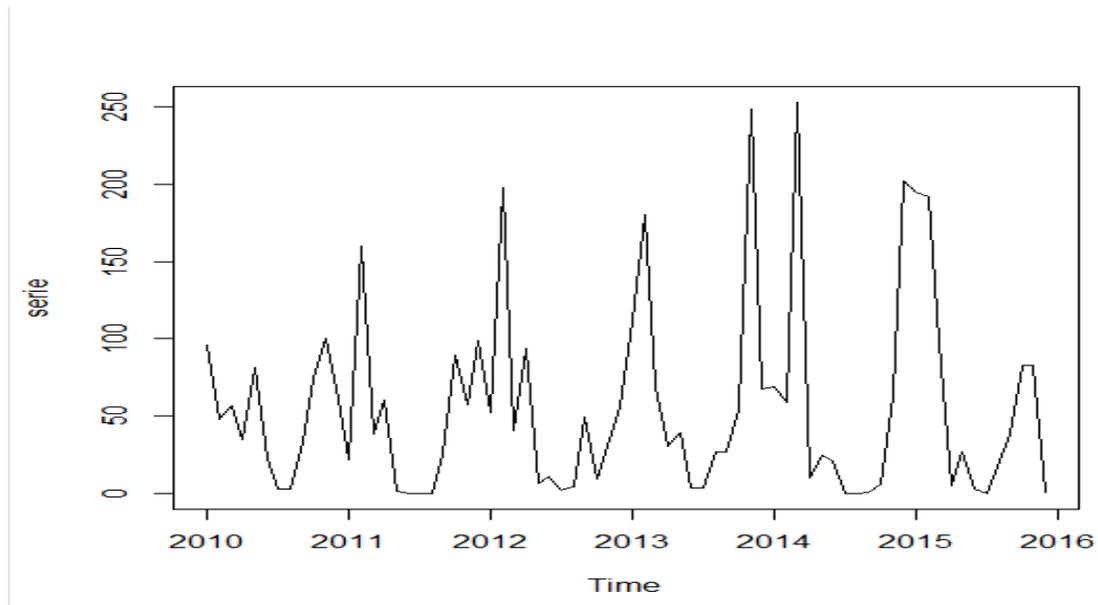


FIGURE 3.1 – Graphe de la série originale

Le graphe de la série originale figure 3.2 nous montre que celle-ci comporte une tendance et une composante saisonnière. Pour illustrer cela, nous allons décomposer notre série originale en ses différentes composantes, pour ce faire nous utilisons les commandes suivantes :

```
> decompose=decompose(serie,type=c("additive","multiplicative"),filter=NULL)
> plot(decompose)
```

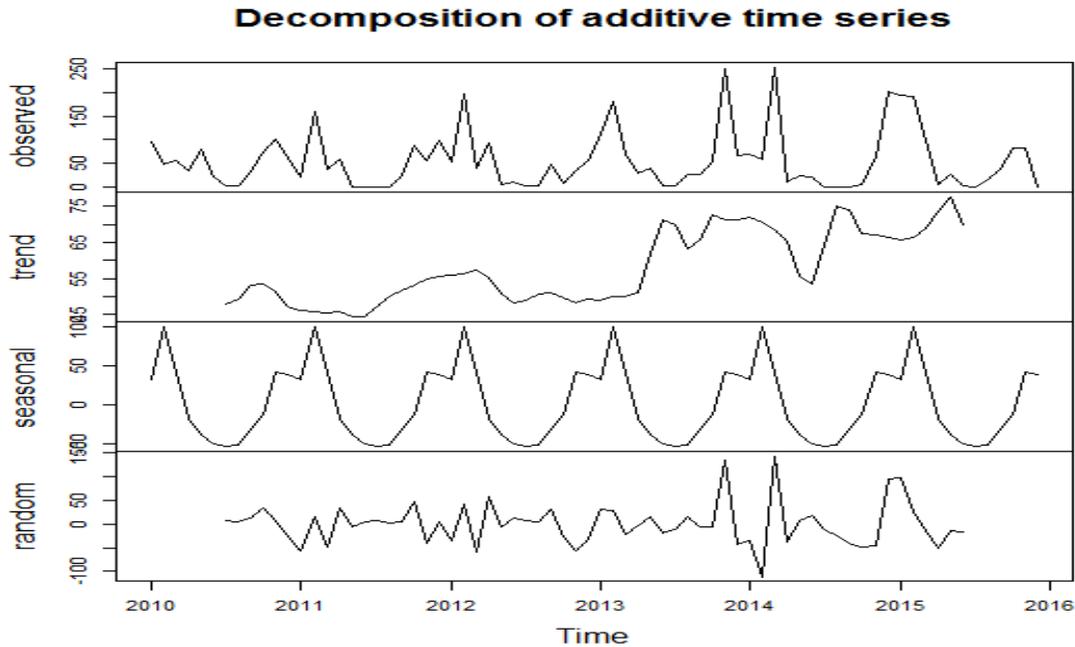


FIGURE 3.2 – Graphe de la série originale ainsi que ses différentes composantes

D'après la figure 3.3 nous confirmons que notre série comporte une composante saisonnière et une tendance croissante. Le modèle du lissage exponentiel adéquat est alors celui de Holt Winters, donc nous appliquons le lissage exponentiel de Winters additif. La commande qui permet d'effectuer le lissage de Winters additif est :

```
> lisse=HoltWinters(serie,seasonal="add")
```

Cette commande affiche les résultats suivants :

les valeurs des paramètres α , β et γ qui minimisent la somme des carrés de l'erreur :

$$\Theta(\alpha, \beta, \gamma) = \sum_{i=1}^{T-1} (Y_{t+1} - \hat{Y}_T)$$

$$\alpha = 0.02993478$$

$$\beta = 0.09332182$$

$$\gamma = 0.1393696$$

Les coefficients \hat{a} et \hat{b} sont respectivement :

$$\hat{a} = 61.6134389$$

$$\hat{b} = 0.3190487$$

Les coefficients saisonniers sont :

\hat{S}_1	16.2981025
\hat{S}_2	110.8423529
\hat{S}_3	27.9746336
\hat{S}_4	-1.1937900
\hat{S}_5	-35.0451208
\hat{S}_6	-42.7340387
\hat{S}_7	-46.8813967
\hat{S}_8	-42.5045239
\hat{S}_9	-21.2861531
\hat{S}_{10}	9.2078982
\hat{S}_{11}	49.2985336
\hat{S}_{12}	24.3733509

TABLE 3.2 – Les coefficients saisonniers

Le modèle de prévision s'écrit sous la forme :

$$\hat{y}_t = 61.6134389 + (t - T)0.3190487 + \hat{S}_t$$

Pour représenter la série originale et son ajustement par la méthode de Holt-Winters additive, nous utilisons la commande suivante :

```
> plot(lisse,xlab="Mois:01/2005-12/2013",ylab="La pluviometrie en millimetres")
```

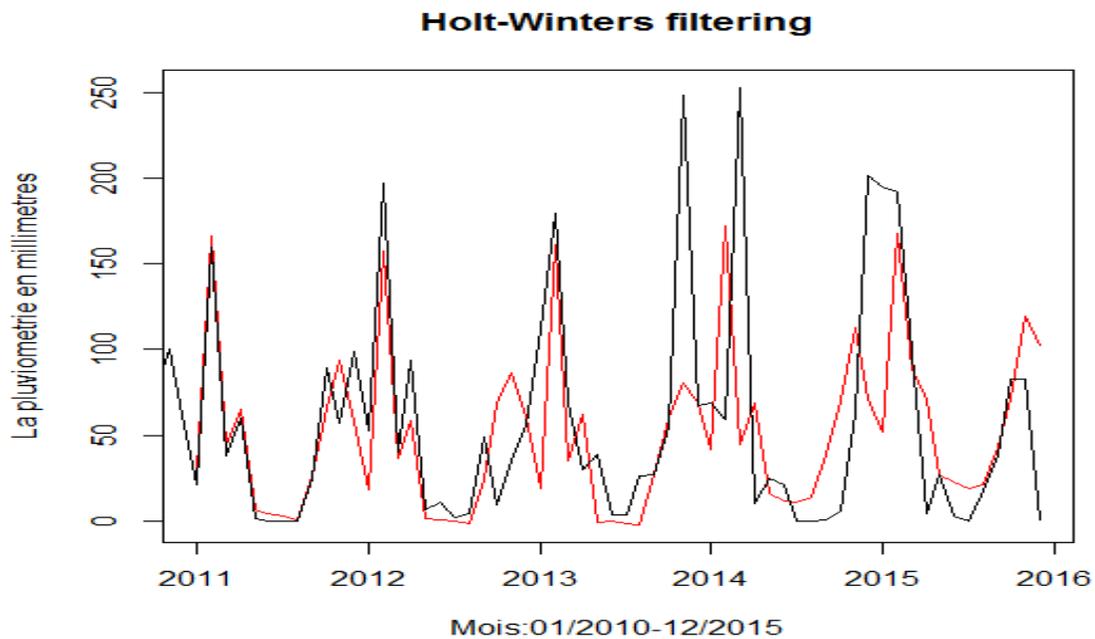


FIGURE 3.3 – La série originale et son ajustement

Nous avons un assez bon ajustement, nous pouvons ainsi faire des prévisions. Pour ce faire, nous utilisons la commande suivante :

```
> prevision=predict(lisse,n.ahead=12)
```

Le tableau 3.3 et la figure 3.5 représentent les prévisions obtenues par la méthode de Holt Winters additive pour l'année 2016 :

Mois	Valeur-prévue
Janvier	78.23059
Février	173.09389
Mars	90.54522
Avril	61.69584
Mai	28.16356
Juin	20.79369
juillet	16.96538
<i>Août</i>	21.66130
Septembre	43.19872
Octobre	74.01182
Novembre	114.42151
Décembre	89.81537

TABLE 3.3 – Prévisions pour l'année 2016

Afin d'obtenir le graphe suivant nous avons utilisé les commandes suivantes :

```
> serie_prevision=ts(c(serie,prevision),start=2005,frequency=12)
> plot(serie_prevision,xlab="Mois:01/2010-12/2016",ylab="Pluviometrie en millimetres")
> l=lines(prevision,col="red")
```

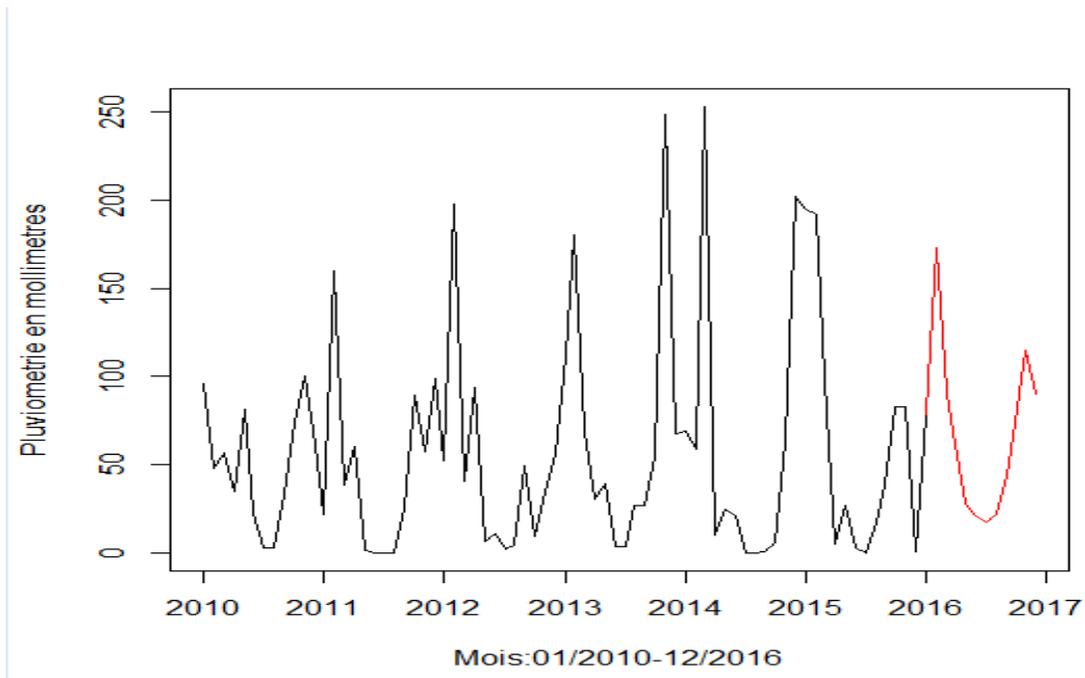


FIGURE 3.4 – Graphe de la série originale et sa prévision

3.7 Application de la méthode de Box et Jenkins

• Analyse préliminaire

Les commandes qui nous permettent de visualiser respectivement le corrélogramme et le corrélogramme partiel de la série originale sont :

- > acf(serie)
- > pacf(serie)

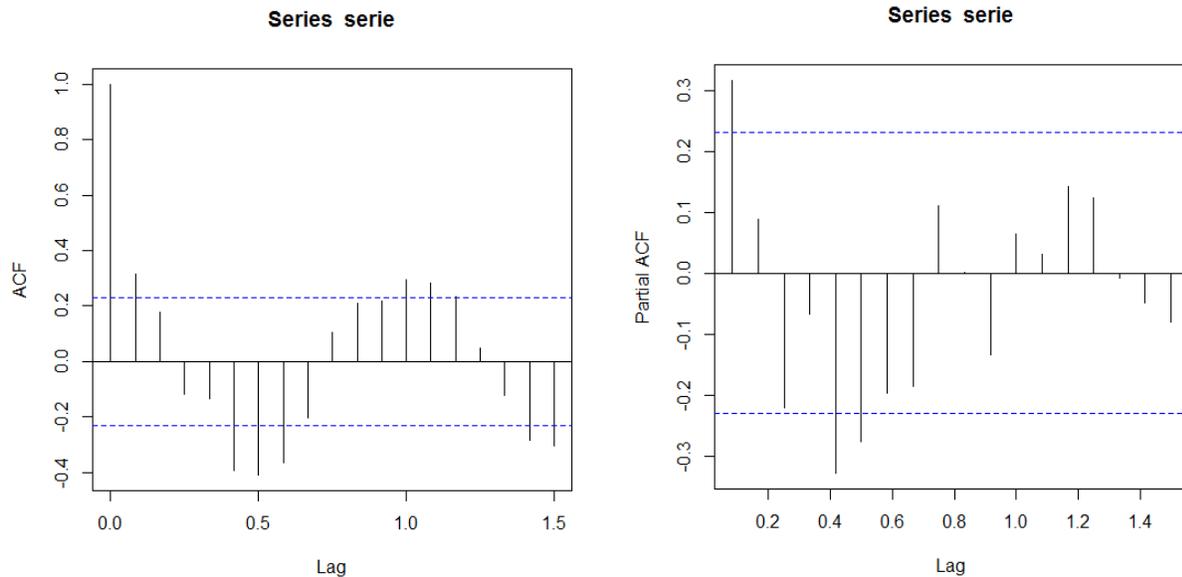


FIGURE 3.5 – Corrélogramme et corrélogramme partiel de la série originale.

Le graphe de l'autocorrélation présente une décroissance lente quant à l'autocorrélation partielle apparaît tronquée, cela confirme la non stationnarité de notre série.

L'examen des deux graphes 3.2 et 3.3 nous suggèrent l'existence d'une tendance croissante donc il sera nécessaire de la différencier ainsi qu'une composante saisonnière, donc une différenciation d'ordre 12 sera nécessaire pour enlever cette saisonnalité.

Les commandes qui nous permettent d'obtenir la série chronologique différencier sont les suivantes :

```
> dift=diff(serie)
> difs=diff(dift,frequency=12)
> plot(difs)
```

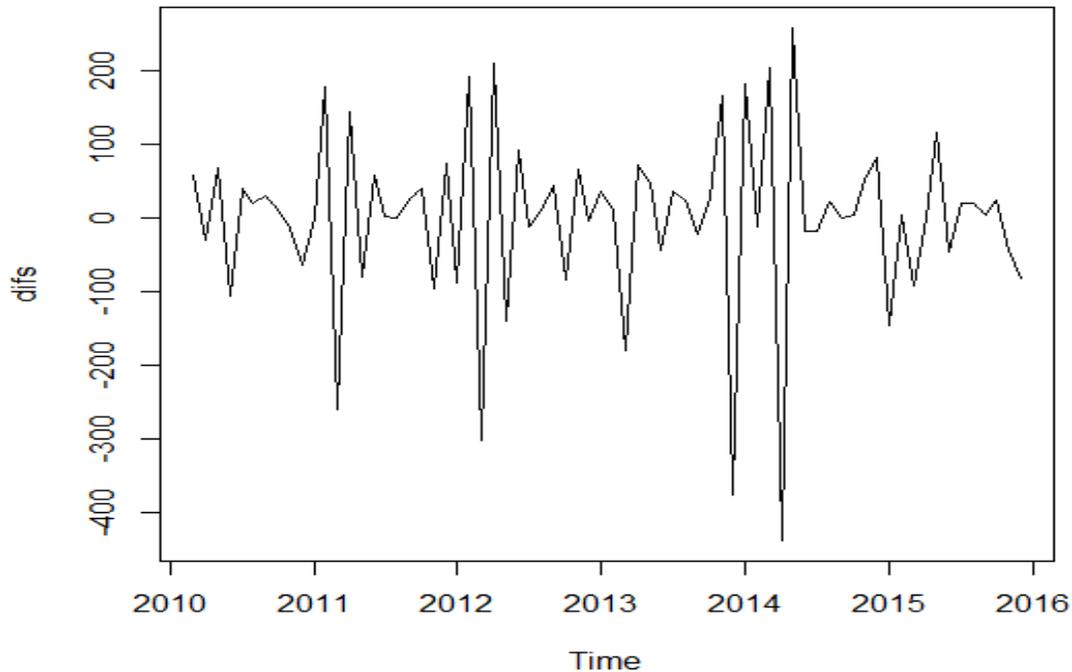


FIGURE 3.6 – Graphe de la série après la differenciation

● Identification du modèle

D'après le graphe de la série originale (Fig 3.2) on constate que cette série est sujet à une saisonnalité d'ordre 12 (12 mois) et une tendance croissante linéairement. Donc le modèle adéquat est $SARIMA(p, d, q)(P, D, Q)_{s=12}$.

En appliquant à la série originale un filtre ordinaire d'ordre 1 ($d=1$) et un filtre saisonnier d'ordre 1 ($D=1$) on a obtenu une série sans tendance ni saisonnalité.

Après avoir tester différents modèles, nous avons opté pour le modèle $SARIMA(5, 1, 2)(0, 1, 1)_{12}$. Les commandes qui nous permettent de visualiser respectivement le corrélogramme et le corrélogramme partiel de la série transformée sont :

```
> acf(differencier)
> pacf(differencier)
```

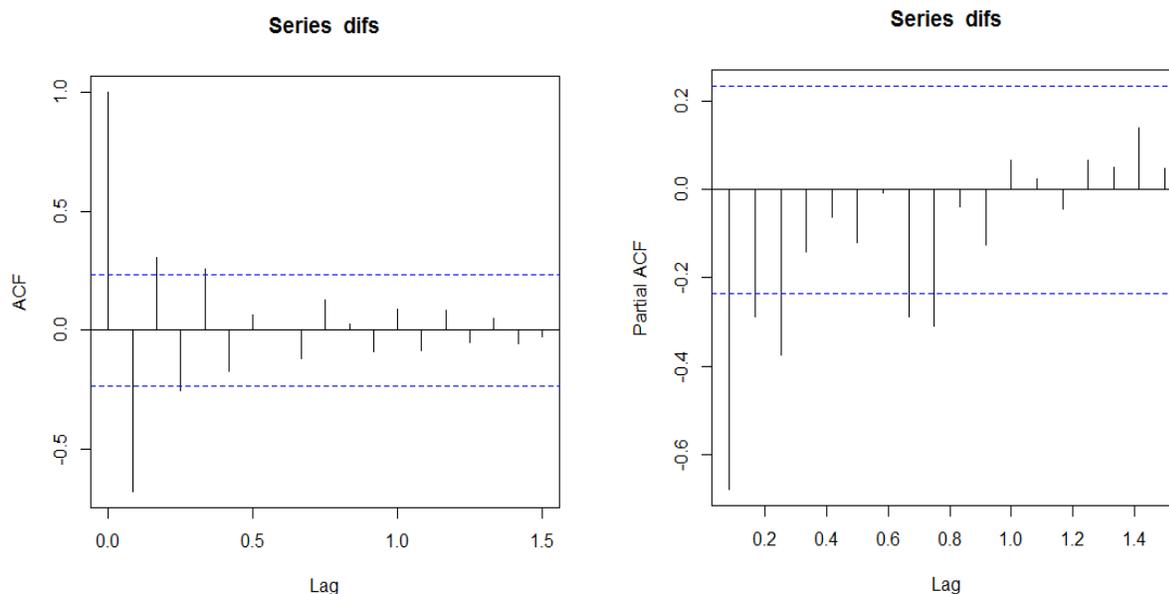


FIGURE 3.7 – Corrélogramme et corrélogramme partiel de la série transformé.

• Estimation des paramètres du modèle

Après avoir opté pour un modèle $SARIMA(5, 1, 1)(0, 1, 1)_{12}$ la commande :
`> sarima=arima(s, order=c(5, 1, 2), seasonal=list(order=c(0, 1, 1), period=12))`,
 permet d'effectuer l'ajustement de la série et d'obtenir les résultats suivants :

Les estimateurs des paramètres du modèle sont :

$$\hat{\phi}_1 = 0,8081, \hat{\phi}_2 = -0,4970, \hat{\phi}_3 = -0,6295, \hat{\phi}_4 = 0,4411, \hat{\phi}_5 = -0,9210, \hat{\theta}_1 = 0,9982, \\ \hat{\theta}_s = -1,0000$$

Les écarts type des paramètres estimés sont :

$$\hat{\sigma}(\hat{\phi}_1) = 0,1813, \hat{\sigma}(\hat{\phi}_2) = 0,1668, \hat{\sigma}(\hat{\phi}_3) = 0,1649, \hat{\sigma}(\hat{\phi}_4) = 0,1738, \hat{\sigma}(\hat{\phi}_5) = 0,1538, \\ \hat{\sigma}(\hat{\theta}) = 0,0933, \hat{\sigma}(\hat{\theta}_s) = 0,1648.$$

Nous remarquons que tous les estimateurs vérifient les conditions de stationnarité : $|\hat{\phi}| < 1$, $|\hat{\theta}| < 1$, $|\hat{\theta}_s| < 1$ et d'inversibilité.

Le modèle s'écrit :

$$(1 - \hat{\phi}_1 B - \hat{\phi}_2 B^2 - \hat{\phi}_3 B^3 - \hat{\phi}_4 B^4 - \hat{\phi}_5 B^5)(1 - B)(1 - B^{12})y_t = (1 - \hat{\theta} B)(1 - \hat{\theta}_s B^{12})\epsilon_t$$

alors :

$$(1 - 0,8081B + 0,3970B^2 + 0,6295B^3 - 0,4411B^4 + 0,9210B^5)(1 - B)(1 - B^{12})y_t = (1 - 0,9982B)(1 + B^{12})\epsilon_t.$$

● **Validation du modèle**

Tests sur les coefficients du modèle

La valeur critique lue sur la table de student est $T(72, 0.025) = 2,39$

1. **Test de $H_0 : \phi_1 = 0$** contre $H_1 : \phi_1 \neq 0$

Ce test est basé sur la statistique : $T_1 = \frac{|\widehat{\phi}_1|}{\widehat{\sigma}(\phi_1)} \rightarrow t(n - 7, \frac{\alpha}{2})$

sa réalisation $t_1 = \frac{0,8081}{0,1813} = 4,45 > 2,39$, Alors on rejette l'hypothèse H_0 et par conséquent on accepte l'hypothèse H_1 donc le paramètre ϕ_1 est significativement différent de zéro.

2. **Test de $H_0 : \phi_2 = 0$** contre $H_1 : \phi_2 \neq 0$

Ce test est basé sur la statistique : $T_2 = \frac{|\widehat{\phi}_2|}{\widehat{\sigma}(\phi_2)} \rightarrow t(n - 7, \frac{\alpha}{2})$

sa réalisation $t_2 = \frac{0,4970}{0,1668} = 2,97 > 2,39$, Alors on rejette l'hypothèse H_0 et par conséquent on accepte l'hypothèse H_1 donc le paramètre ϕ_2 est significativement différent de zéro.

3. **Test de $H_0 : \phi_3 = 0$** contre $H_1 : \phi_3 \neq 0$

Ce test est basé sur la statistique : $T_3 = \frac{|\widehat{\phi}_3|}{\widehat{\sigma}(\phi_3)} \rightarrow t(n - 7, \frac{\alpha}{2})$

sa réalisation $t_3 = \frac{0,6295}{0,1649} = 3,81 > 2,39$, Alors on rejette l'hypothèse H_0 et par conséquent on accepte l'hypothèse H_1 donc le paramètre ϕ_3 est significativement différent de zéro.

4. **Test de $H_0 : \phi_4 = 0$** contre $H_1 : \phi_4 \neq 0$

Ce test est basé sur la statistique : $T_4 = \frac{|\widehat{\phi}_4|}{\widehat{\sigma}(\phi_4)} \rightarrow t(n - 7, \frac{\alpha}{2})$

sa réalisation $t_4 = \frac{0,4411}{0,1738} = 2,53 > 2,39$, Alors on rejette l'hypothèse H_0 et par conséquent on accepte l'hypothèse H_1 donc le paramètre ϕ_4 est significativement différent de zéro.

5. **Test de $H_0 : \phi_5 = 0$** contre $H_1 : \phi_5 \neq 0$

Ce test est basé sur la statistique : $T_5 = \frac{|\widehat{\phi}_5|}{\widehat{\sigma}(\phi_5)} \rightarrow t(n - 7, \frac{\alpha}{2})$

sa réalisation $t_5 = \frac{0,9210}{0,1538} = 5,98 > 2,39$, Alors on rejette l'hypothèse H_0 et par conséquent on accepte l'hypothèse H_1 donc le paramètre ϕ_5 est significativement différent de zéro.

6. **Test de $H_0 : \theta = 0$ contre $H_1 : \theta \neq 0$**

Ce test est basé sur la statistique : $T_6 = \frac{|\hat{\theta}|}{\hat{\sigma}(\hat{\theta})} \longrightarrow t(n-7, \frac{\alpha}{2})$

sa réalisation $t_6 = \frac{0,9982}{0,0933} = 10,69 > 2,39$, Alors on rejette l'hypothèse H_0 et par conséquent on accepte l'hypothèse H_1 donc le paramètre θ est significativement différent de zéro.

7. **Test de $H_0 : \theta_s = 0$ contre $H_1 : \theta_s \neq 0$**

Ce test est basé sur la statistique : $T_7 = \frac{|\hat{\theta}_s|}{\hat{\sigma}(\hat{\theta}_s)} \longrightarrow t(n-7, \frac{\alpha}{2})$

sa réalisation $t_7 = \frac{1,0000}{0,1648} = 6,06 > 2,39$, Alors on rejette l'hypothèse H_0 et par conséquent on accepte l'hypothèse H_1 donc le paramètre θ_s est significativement différent de zéro.

- **Examen des résidus**

Test sur les résidus

Ce test est basé sur la statistique suivante :

$$Q = n(n+2) \sum_{i=1}^M \frac{\rho_i^2}{n-i} \longrightarrow \chi_{M-k}^2$$

Test : $H_0 : \rho_1 = \rho_2 = \dots = 0, i = 1, \dots, M$ contre $H_1 : \exists i \in 1, \dots, M \text{ tq } \rho_i \neq 0$, où :

- ρ_i désignent les coefficients d'autocorrélations théoriques,

- M est la partie entière de : $Min(\frac{n}{2}, 3\sqrt{n})$ donc M=25.

L'instruction :

```
> Box.test(residu1, lag=18, type="Ljung")
```

nous donne la valeur empirique de la statistique de Box et Ljung, $Q^* = 13,6550$.

Au niveau de signification $\alpha = 0.05$ la valeur critique lue sur la table de khi-deux est :

$$\chi_{M-k}^2(\alpha) = \chi_{25-7}^2(0,05) = \chi_{18}^2(0,05) = 28,8693$$

Comme $Q^* = 13,6550 < \chi_{18}^2(0,05) = 28,8693$, alors on accepte l'hypothèse H_0 , par conséquent les résidus forment un bruit blanc, le corrélogramme et le corrélogramme partiel des résidus nous montrent clairement que tous les pics sont à l'intérieure de l'enveloppe.

Les commandes sous R qui nous permet de les représenter sont :

```
> residu=residuals(sarima)
```

```
> residu1=ts(residu,start=2005,frequency=12)
> acf(residu1)
> pacf(residu1)
```

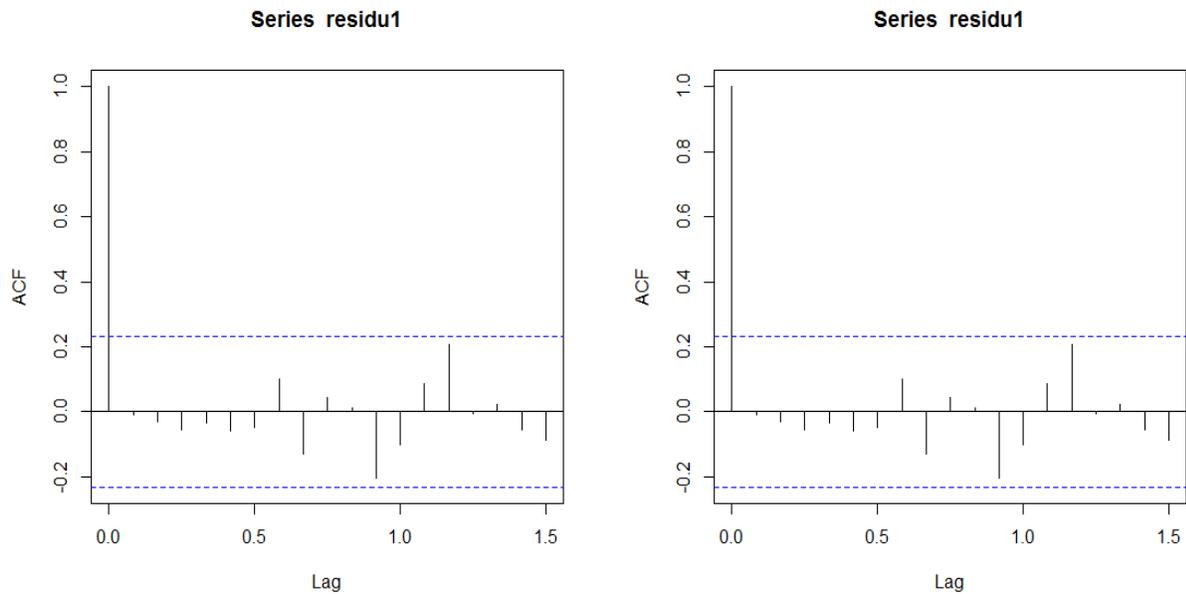


FIGURE 3.8 – Corrélogramme et corrélogramme partiel du résidu.

● Prédiction

La commande qui nous permet d'obtenir les prévisions est :

```
> prevoir=predict(sarima,12).
```

Les prévisions obtenues selon le modèle $SARIMA(5, 1, 1)(0, 1, 1)_{12}$ sont représentées dans le graphe 3.13 et le tableau 3.4 :

Mois	Valeur-prévue
Janvier	103.50
Février	173.01
Mars	134.82
Avril	42.81
Mai	42.77
Juin	14.26
juillet	17.09
<i>Août</i>	16.86
Septembre	44.47
Octobre	61.58
Novembre	115.42
Décembre	93.48

TABLE 3.4 – Prévisions pour l'année 2016

Afin d'obtenir le graphe suivant nous avons utilisé les commandes suivantes :

```
> serie_prevoir=ts(c(serie,prevoir),start=2005,frequency=12)
> plot(serie_prevoir,xlab="Mois:01/2010-12/2015",ylab="Pluviométrie en millimetres")
> l=lines(prevoir,col="red")
```

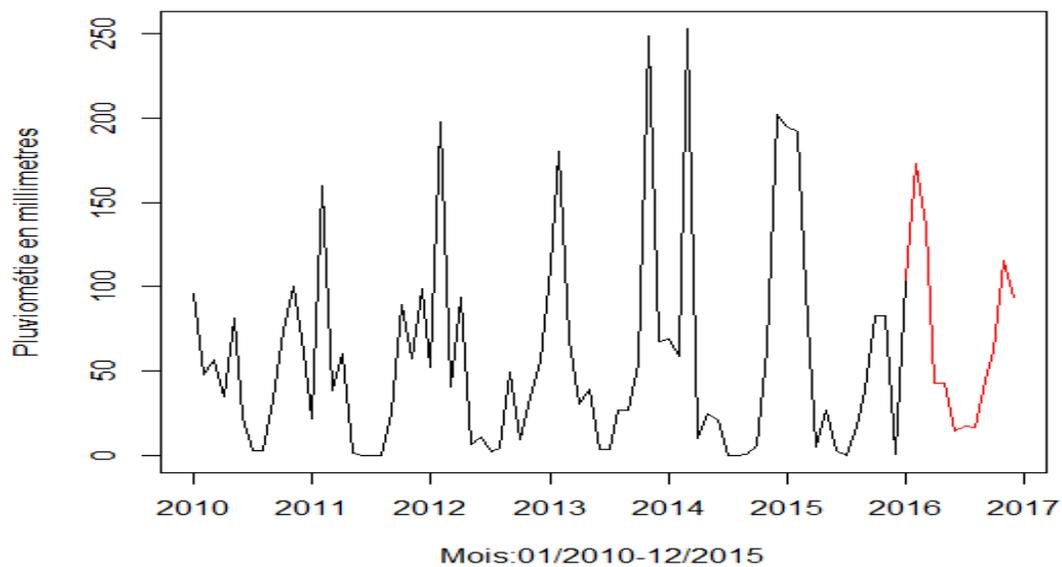


FIGURE 3.9 – Graphe de la série originale et sa prévision

3.7.1 Interprétation des résultats obtenus par les deux méthodes

Mois	Prévision avec lissage exponentiel	Prévision avec Box et Jenkins	Les valeurs réelles
Janvier	78.23	103.50	68
Février	173.09	173.01	41
Mars	90.54	134.82	170.5
Avril	61.69	42.81	51.1
Mai	28.16	42.77	1.1

TABLE 3.5 – Tableau de comparaison

D'après le tableau 3.5, les prévisions données par Box-Jenkins et le lissage exponentiel, concernant les quantités pluviométriques, montrent qu'il y a un écart important avec la réalité, en particulier pour les deux mois de février et mai, cela s'explique par le fait qu'il y a plusieurs facteurs qui influencent sur les quantités pluviométrique, parmi ces facteurs, le réchauffement climatique qui provoque un changement au niveau de la répartition saisonnière des pluies.

Sachant que notre climat est méditerranéen donc la distribution mensuelle des pluies dans nos prévisions obtenues par les deux méthodes apparaissent raisonnables, c'est-à-dire : En novembre, décembre et janvier nous remarquons un maximum de quantités pluviométriques. Au printemps- été les quantités pluviométriques diminuent et elles reviennent à la hausse en automne-hiver.

3.8 Conclusion

Dans ce chapitre, on a abordé une méthode très utilisée dans la statistique décisionnelle la modélisation des série chronologique.

On a appliqué les modèles de prévisions sur un cas pratique. La série a été modélisée par deux méthodes le lissage exponentielle qui est une méthode empirique et les modèles de Box et Jenkins reconnus pour leur rigueur mathématique. La difficulté dans la modélisation avec les modèles de Box et Jenkins réside dans le choix des paramètres.

En effet, on a modélisé la quantité en mm de pluviométrie de 2010 jusqu'à 2015 par les deux méthodes citées. La méthode lissage exponentiel de Halt-Winter appliquée sur le

problème considéré permet de trouver un modèle qui permet de lissée la série et de dégager un modèle additif à base duquel on calculera les prévisions. Lors de la modélisation par Box & Jenkins, on a testé plusieurs modèles de type SARIMA. Le modèle choisi ajuste mieux la série, ce dernier nous a servi aux calculs des prévisions pour l'année 2016, et de les comparées aux données réelles .

Chapitre 4

Objectifs des prévisions

4.1 Introduction

La prévision d'une série consiste à déterminer son évolution sur les périodes non observées. Le terme "prévision" s'utilise dans le cadre d'un phénomène quantifiable, par exemple comme dans notre cas "la pluviométrie".

Cette procédure suppose donc que l'on dispose d'une série d'observations et d'un corps d'hypothèses conduisant à une modélisation du phénomène étudié. On a opéré deux étapes : construction d'un modèle à partir des données et d'hypothèses à priori, puis projection dans la partie non observée .

Après avoir obtenu les prévisions dans le chapitre précédent et après les avoir comparé avec les valeurs réelles, Ce chapitre est consacré à répondre a la questions posé précédement.

4.2 Importance des prévisions

Les prévisions sont un des éléments de base servant à coordonner les différents départements d'une compagnie. Lorsque tous les départements utilisent la même prévision dans la planification de leur travail, ils se préparent au même futur et leurs efforts sont coordonnés.

Les prévisions sont donc un élément vital à la bonne coordination des activités des différents départements d'une compagnie. Il est important que ces derniers fassent leur planification en se basant sur les mêmes suppositions en ce qui concerne l'avenir.

4.3 Horizon des prévisions

On distingue généralement trois types de prévision :

Court Terme : Généralement mesuré en jours ou en semaines. Peut aller jusqu'à un an.

Moyen Terme : Généralement mesuré en semaines ou en mois. Peut aller jusqu'à deux ans.

Long Terme : Peut aller jusqu'à cinq ans ou plus

4.4 Caractéristiques des prévisions

Voici quatre caractéristiques des prévisions. Ces caractéristiques doivent être constamment gardées à l'esprit lorsqu'on utilise les résultats de prévisions dans les décisions de gestion.

1. Elles sont généralement fausses

Une fois les prévisions déterminées, elles sont souvent utilisées comme s'il s'agissait d'informations certaines. Les transfères d'eau vers les cinq wilaya peuvent être modifiés si les prévisions de la pluviométrie ne sont pas exactes. Le système de planification doit être suffisamment robuste pour réagir en cas d'erreur de prévision non anticipée.

2. Une bonne prévision est plus qu'une valeur numérique

Puisque les prévisions sont généralement fausses, une bonne prévision doit également inclure une mesure de l'erreur anticipée pour la prévision. Ceci peut prendre la forme d'une étendue ou d'une mesure de l'erreur telle que la variance de la distribution de l'erreur de prévision.

3. Les prévisions à long terme sont moins précises

Cette caractéristique est assez intuitive. On peut prévoir plus exactement les pluviométries de la prochaine année que celles à survenir dans cinq ans.

4. Les prévisions ne remplacent pas la vraie information

Une technique de prévision donnée peut donner de bonnes prévisions dans la plupart des cas. Cependant, il peut y avoir un changement climatique non prévu comme ce qu'on a vu dans le chapitre précédent.

4.5 Les caractéristiques climatiques

Les caractéristiques les plus originales de toute la région sont réunies dans le bassin de Beni Haroun. Le nord s'avère être une zone humide et pluvieuse et le sud avec un climat semiaride à aride.

Cette opposition trouve son origine dans de multiples causes :

- Le nord influencé par un climat méditerranéen, le sud par un climat de type continental.
- La répartition des ensembles montagneux, les contrastes renforçant cette disparité géographique.

Dans cet setion, on s'intéressera à évaluer le paramètre climatique jugé important : les précipitations.

4.5.1 Les précipitations

Les précipitations constituent évidemment, le paramètre climatique essentiel. Leur intensité, leur continuité et leur périodicité sont l'origine même de l'écoulement, de sa localisation et de sa violence. Dans certaines régions, où le régime des écoulements est fortement lié aux précipitations.

Les variations interannuelles

L'interprétation de l'histogramme (Fig 4.1) fait sortir, que d'une manière générale les précipitations se caractérisent par leur irrégularité interannuelle.

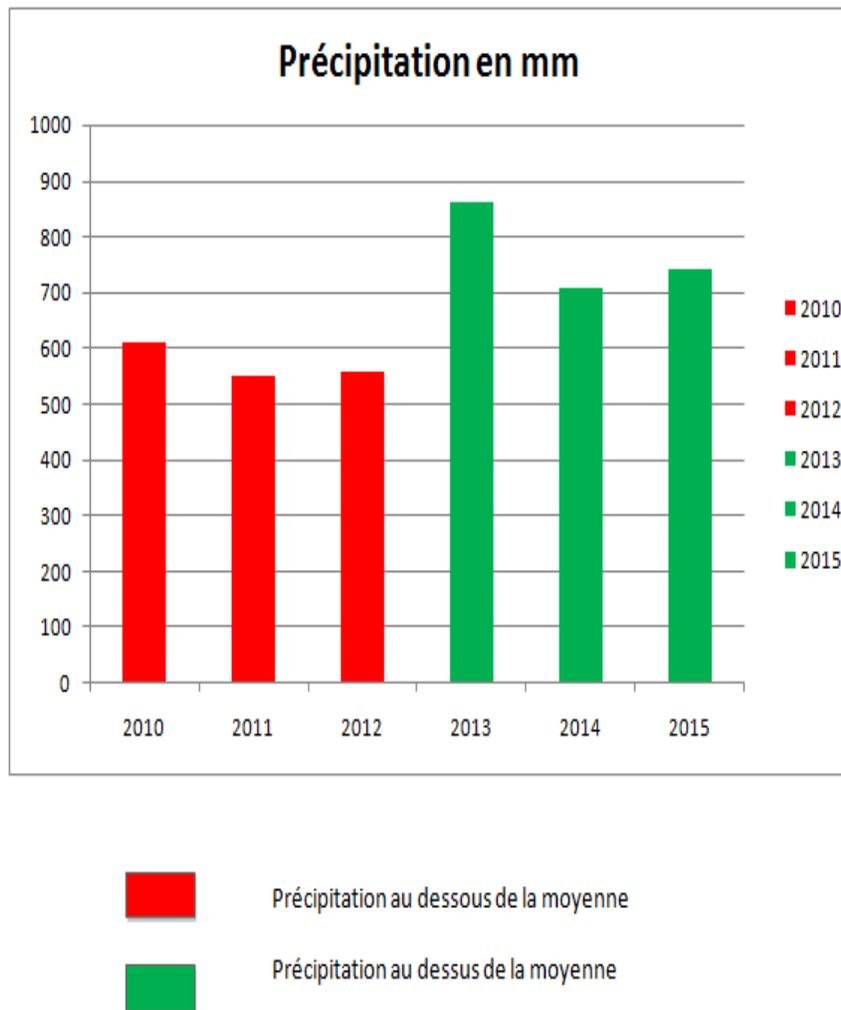


FIGURE 4.1 – Variations interannuelles des précipitations 2010/2015

L'année 2013, est l'année la plus abondante en pluie, elle a connu un excédent de 189,92 mm. A l'inverse, l'année 2011, a connu un déficit de 120.38 mm, soit (Tableau N°04), c'est l'année la plus sèche. La moyenne annuelle des pluies précipitées pour la période 2010/2015, est de 671.08 mm. A l'aide des prévisions calculées dans le chapitre précédent on peut tracer l'histogramme suivant :

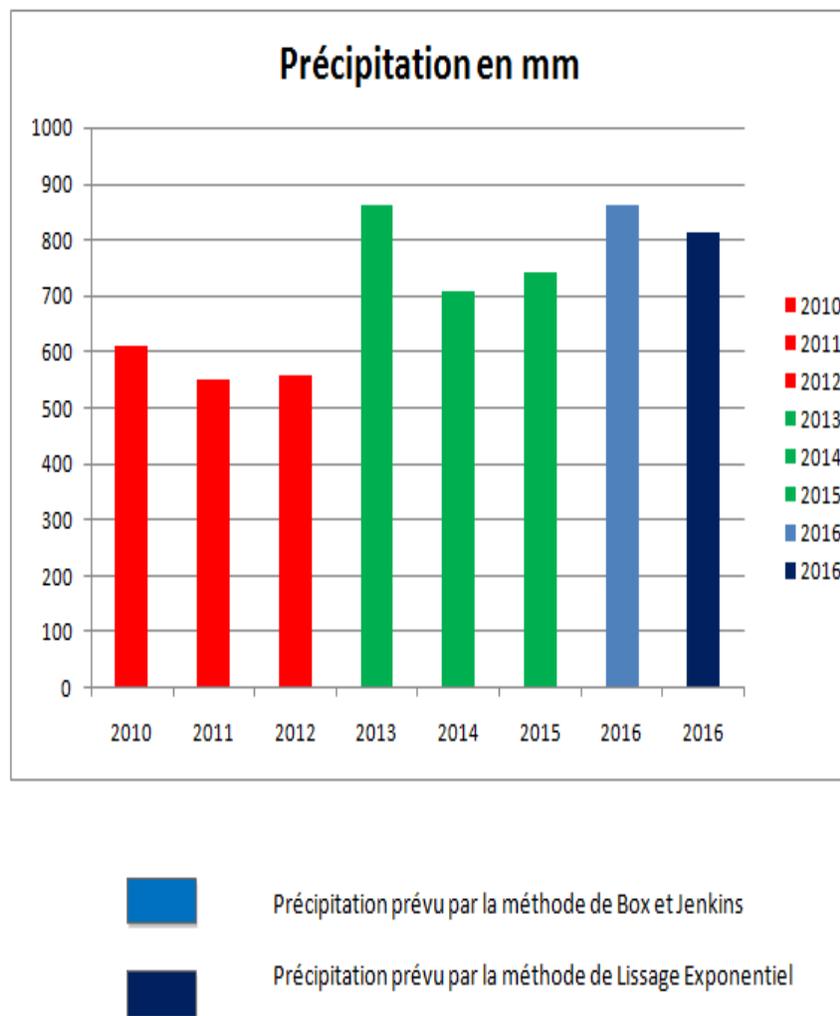


FIGURE 4.2 – Variations interannuelles des précipitations et précipitations prévues en 2016

On remarque que les précipitations dans le bassin de Beni Haroun varient entre 500-900 mm. Elles s'accroît du nord au sud, se caractérise par des variations inter et intra-annuelle, mensuelles et saisonnières et se concentre sur une part assez courte de l'année.

4.6 L'objectif du travail

Dans cette section, nous avons essayé de répondre à certaines questions qui se pose au niveau du barrage de Béni Haroun, parmi celles-ci :

- Déterminer la quantité d'eau dans l'année suivante (2016), afin d'éviter des dégâts (catastrophes) comme par exemple des crues en cas d'excé d'eau, manque d'eau ce

qui cause une sécheresse,

- L'objectif principal est de savoir combien de Wilaya le barrage peut alimenter selon les besoins des Wilaya. Sachant que le barrage alimente déjà deux (2) Wilaya (Constantine, Mila), et il prévoit d'alimenter trois(3) autres Wilaya c-à-d (Constantine, Mila, Batna, Khenchela, Oum ElBouaghi).

Notre but est de déterminer combien d'autres Wilaya peuvent être alimentées aussi.

D'après l'étude réalisée dans le chapitre précédent (les prévisions), on a déterminé quelques chiffres sur la quantité d'eau pour chaque mois de l'année suivante qui est donnée dans le tableau 3.5, En prenant en compte ces résultats et d'autre paramètres comme l'évaporation d'eau, on a déterminé la quantité d'eau actuelle dans le barrage pour l'année suivante qui est 1 milliard m^3 .

a partir de ça, on peut dire que cette quantité d'eau ne cause aucune catastrophe ni crues, ni sécheresse.

Pour répondre à l'objectif principal qu'est : combien de Wilaya peuvent être alimentées, on a besoin de certaines informations.

- La quantité d'eau actuelle dans le barrage de Béni Haroun qu'est 1 milliard m^3 ,
- La quantité d'eau qu'il faut garder dans le barrage, afin d'éviter la sécheresse sur tout durant les mois d'été,
- le besoin de chaque Wilaya en quantité d'eau annuelle.

Le barrage doit alimenter cinq Wilaya en eau potable (AEP) et deux régions pour l'irrigation pour cette année. Alors les besoins en eau de ces wilaya ont été déjà récupéré elles sont présentées par les valeurs données dans les tableaux suivantes :

Qte AEP \ Wilaya	Constantine	Mila	Batna	Khenchla	Oum Elbouaghi
AEP	108.24	47.5	93.10	26	21.3

TABLE 4.1 – Les quantités d'AEP transférés aux cinq wilaya

Qte irrigation \ Wilaya	Télegma	chemora
eau d'irrigation	44	89.50

TABLE 4.2 – Les quantités d'eau d'irrigation transférés

Ainsi la figure ci-après nous montre toutes les opérations des transferts d'eau.



FIGURE 4.3 – Objectifs et Défis

Alors la quantité d'eau qui sera transférée est :

$$\begin{aligned}
 QDt_{AEP} &= \sum_{j=1}^5 BQD_j = \\
 BQD_{Constantine} + BQD_{Mila} + BQD_{Batna} + BQD_{Khenchla} + BQD_{Oumelbouaghi} &= \\
 108.24 + 47.5 + 93.10 + 26 + 21.3 &= 296.14 Hm^3
 \end{aligned}$$

et

$$\begin{aligned}
 QDt_{irri} &= \sum_{j=1}^2 BQD_j = \\
 BQD_{Teleghma} + BQD_{chemora} &= \\
 44 + 89.5 &= 133.5 Hm^3
 \end{aligned}$$

Alors la quantité d'eau totale transféré est :

$$QDT = QDt_{AEP} + QDt_{irri} = 296.14 + 133.5 = 429.64 Hm^3$$

et la quantité d'eau qui doit toujours être présenté au barrage est le volume équivalent à la cote 172 m VTM (volume de la tranche morte) est : $242000000 m^3 = 242 Hm^3$, sachant que la quantité initial (QDB_I) dans le barrage est : 1 milliard $m^3 = 1000 Hm^3$.

Cependant le volume d'eau restant dans le barrage sera :

$$\begin{aligned}
 QDB_F &= QDB_I - VTM - QDT \iff \\
 QDB_F &= 1000 - 242 - 429.64 \iff \\
 QDB_F &= 328.36 Hm^3
 \end{aligned}$$

On remarque que le barrage contient une quantité qui permet d'alimenter un nombre de Wilaya inférieure au nombre des Wilaya déjà alimenté ou bien prévu d'être alimenté que se soit en eau potable ou pour l'irrigation des régions selon leurs besoins, mais en prenant toujours en compte la quantité d'eau qui s'évapore, ainsi que le niveau minimal d'eau dans le barrage.

4.7 Conclusion

L'eau a comme propriété particulière d'être renouvelable et en transit permanent. Comme ressource renouvelable, elle suit un cycle naturel, cheminement qui s'accompagne d'un ensemble de transformations. De l'atmosphère, ou se transforment les nuages, les précipitations arrivent au sol où elles favorisent le développement des plantes et alimentent les cours d'eaux après avoir ruisselé ou s'être infiltré. Les cours d'eaux à leur tour aboutissent à la mer ou par évaporation assure le retour d'une partie à l'atmosphère.

Il s'agit d'un mouvement perpétuel, mais non uniforme. Le cycle peut être parcouru rapidement ou lentement. En quelques jours une goutte d'eau évaporée de la méditerranée peut tomber sous forme de pluie et revenir à celle-ci.

Les prévisions sur la pluviométrie de l'année 2016 nous montre qu'elles seront suffisantes pour compenser les quantités d'eau transférées.

Conclusion Générale

A la lumière de notre recherche, nous avons présenté dans ce mémoire une revue statistique d'une série chronologique qui représente dans notre cas la pluviométrie en millimètre dans le barrage de Béni Haroun par mois sur six ans.

Comme indiqué dans l'introduction, le but de cette étude sur cette série chronologique est la prévision, qui consiste à utiliser le meilleur des modèles estimés pour donner les valeurs estimées à des instants futurs.

En premier lieu, nous avons présenté les méthodes statistiques de la prévision telle que la méthode de Box & Jenkins et de Lissage exponentiel.

En deuxième lieu, grâce au logiciel R, nous avons appliqué le lissage exponentiel de Holt Winters additif (série saisonnière) pour prévoir les quantités pluviométriques pour l'année 2016. Ensuite, nous avons adopté la méthode de Box et Jenkins pour modéliser la même série. Etant donné que la série n'est pas stationnaire (elle comporte une tendance et une saisonnalité), il a été question d'une différenciation d'ordre 1 et saisonnière d'ordre 12. Après la stationnarisation de la série, et à partir des graphiques des fonctions d'autocorrélation et d'autocorrélation partielle, nous avons sélectionné le modèle $SARIMA(5, 1, 1)(0, 1, 1)_{12}$. La validation du modèle s'est faite grâce au test de student sur les coefficients et le test de Ljung-Box sur les résidus. Ainsi, nous avons pu obtenir des prévisions pour l'année 2016.

Notre objectif de l'étude de cette série et de faire les prévisions est pour savoir combien de Wilaya le barrage peut alimenter, soit en eau potable soit pour l'irrigation.

Les résultats obtenus montrent que le barrage peut alimenter les 5 Wilaya prévues, et à partir de la quantité d'eau restante dans le barrage, nous pouvons déduire qu'il peut alimenter d'autres Wilaya à condition qu'il ne dépasse pas 328 Hm^3 .

Comme perspective :

- Réalisation d'une étude sur la pollution du Barrage
- Optimisation de la gestion du Barrage,...

Bibliographie

- [1] S. Adjabi, Cours 3^{eme} année, méthode statistique de la prévision RO,2013/2014
- [2] S. Adjabi, Notes de cours des méthodes statistiques de la prévision, 5^{ème} année R.O. Université de Béjaïa, 2010/2011.
- [3] R. Bourbonnais and M. Terraza. Analyse des séries temporelles : Application à l'économie et à la gestion. Dunod, 2^{ème} édition, 2008.
- [4] G. Bresson and A. Pirotte. Économétrie des séries temporelles : Théorie et applications. Presses Universitaires de France, 1^{ère} édition, 1995.
- [5] B. Derasse Chouck, S. Berkani. Prévision du parc des comptes d'une banque : Cas de la BADR 365. Mémoire de fin de cycle en STID. Université A. Mira de Bejaia, 2013.
- [6] A. Harradj, I. Allegui. Etude de l'évolution des ventes du sucre de CEVITAL. Mémoire de fin de cycle en STID. Université A. Mira de Bejaia, 2011.
- [7] A. Lagnoux. Séries Chronologiques. Université de Toulouse, MASTER 1, 2010-2011.
- [8] A. Louar. TP de méthodes statistiques de la prévision. Université A. Mira de Bejaia, 2011.
- [9] V. Monbet. Modélisation de séries temporelles. Université de Rennes, 2011.
- [10] M. Nana, Yakam Ander. Modèle de prévision du parc des comptes d'une banque, cas AFRILAND First Bank. Mémoire de master en statistique appliquée, université Yaoundé de Cameroun, 2005.
- [11] F. Teghbit, Y. Bouiche. Modélisation et Prévision des quantités totales de pétrole (brût et condensât) transportées par l'oléoduc OB1 reliant Hassi Messaoud et Bejaia. Mémoire de fin de cycle en STID. Université A. Mira de Bejaia, 2013.
- [12] M.C. Viano, A. Philippe. Maîtrise d'Économétrie Cours de Séries Temporelles. Université des Sciences et Technologies de Lille U.F.R. de Mathématiques Pures et Appliquées, 1999 - 2004.

RÉSUMÉ

Le barrage de Béni Haroun, situé à une quarantaine de km au Nord de Constantine (Mila), contrôle les eaux du bassin de l'Oued Kébir-Rhumel. Il est le plus grand barrage du pays, érigé dans la région bien arrosée du Tell oriental, au Sud de laquelle se situent de grands centres urbains (Constantine, Batna, Khenchela...) et de vastes terres irrigables (Hautes Plaines semi-arides) qui ne peuvent être satisfaits par les ressources d'eau locales.

Le barrage de Béni Haroun alimente actuellement deux wilaya (Constantine, Mila) et il est prévu d'alimenter trois autres Wilaya (Khenchela, Batna, Oum ElBouaghi). Dans ce travail, nous avons fait une modélisation et une prévision sur la pluviométrie par la méthode de Box et Jenkins et une prévision par la méthode du Lissage exponentiel de l'année 2016. Les résultats obtenus montrent que les prévisions de pluviométrie permettront de compenser les quantités d'eau à transférer.

Mots clés : Barrage, Modélisation, Prévision, série chronologique, Lissage Exponentiel, Méthode de Box et Jenkins.

Abstract

The barrage of Beni Haroun, located at forty km in the North of Constantine (Mila), control water of the basin of the Kébir-Rhumel wadi. It is most barrage of the country, set up in the well sprinkled area of Eastern Tell, in the South of which great urban centers are (Constantine, Batna, Khenchela...) and of vast irrigable grounds (High semi-arid Plains) which cannot be satisfied by the local resources.

The barrage of Beni Haroun currently feeds two wilaya (Constantine, Mila) and it is envisaged to feed three another Wilaya (Khenchela, Batna, Oum Elbouaghi). In this work, we made a modeling and a forecast on pluviometry by the method of Box and Jenkins and a forecast by the method of exponential Smoothing of the year 2016. The results obtained show that the forecasts of pluviometry will make it possible to compensate for the quantities of water with transférer .

Keywords : Barrage, Modeling, Forecasting, time series, Exponential Smoothing, Method of Box and Jenkins.