

Higher National School of Hydraulic

The Library

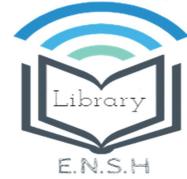
Digital Repository of ENSH



المدرسة الوطنية العليا للري

المكتبة

المستودع الرقمي للمدرسة العليا للري



The title (العنوان):

**Modélisation et prédiction multi-site des précipitations
annuelles et mensuelles dans le bassin versant du Cheliff par
le filtre de Kalman (Algérie).**

The paper document Shelf mark (الشفرة) : 6-0046-17

APA Citation (توثيق APA):

**Lounes, Adel (2017). Modélisation et prédiction multi-site des précipitations
annuelles et mensuelles dans le bassin versant du Cheliff par le filtre de Kalman
(Algérie)[Thèse de master, ENSH].**

The digital repository of the Higher National School for Hydraulics "Digital Repository of ENSH" is a platform for valuing the scientific production of the school's teachers and researchers.

Digital Repository of ENSH aims to limit scientific production, whether published or unpublished (theses, pedagogical publications, periodical articles, books...) and broadcasting it online.

Digital Repository of ENSH is built on the open DSpace software platform and is managed by the Library of the National Higher School for Hydraulics. <http://dspace.ensh.dz/jspui/>

المستودع الرقمي للمدرسة الوطنية العليا للري هو منصة خاصة بتقييم الإنتاج العلمي لأساتذة و باحثي المدرسة.

يهدف المستودع الرقمي للمدرسة إلى حصر الإنتاج العلمي سواء كان منشورا أو غير منشور (أطروحات، مطبوعات، مبداعات، مقالات، البوريات، كتب....) و بنه على الخط.

المستودع الرقمي للمدرسة مبني على المنصة المفتوحة DSpace و يتم إدارته من طرف مديرية المكتبة للمدرسة العليا للري.

كل الحقوق محفوظة للمدرسة الوطنية العليا للري.



Département Hydraulique Urbaine

MEMOIRE DE MASTER

Pour l'obtention du diplôme de Master en Hydraulique

OPTION : Assainissement

THEME :

**Modélisation et prédiction multi-site des précipitations
annuelles et mensuelles dans le bassin versant du
Cheliff par le filtre de Kalman (ALGERIE)**

Présenté par :

M^r : LOUNES Adel

DEVANT LES MEMBRES DU JURY

Nom et Prénom	Grade	Qualité
M ^{me} MEDDI Hind	M.C.A	Présidente
M ^r AMMARI Abdelhadi	M.C.B	Examineur
M ^r BOUFEKANE Abdelmadjid	M.A.A	Examineur
M ^{me} HOULI Samia	M.A.A	Examinatrice
M ^{me} SADOUNE Samra	M.C.B	Promotrice

Avril 2017

Dédicaces

À mes parents qui me sont les plus chers au monde, que Dieu les protège,

À mon frère et mes sœurs qui m'ont toujours soutenu dans mes études.

À tous mes amis ;

À tous ceux que j'aime et qui m'aiment, où qu'ils soient ;

À tous ceux qui comptent pour moi ;

À tous ceux qui m'ont aidé de loin et de près ;

Recevez ici l'expression de ma plus haute gratitude.

Lounes Adel.

Remerciements

Tout d'abord, je remercie Allah le tout puissant de m'avoir donné le courage et la patience de mener à bien mon Master.

A Madame S.SAADOUE, ma promotrice recevez Madame, mes plus sincères remerciements.

ملخص:

هدفنا في هذا العمل هو اختبار فعالية مرشح كالمان في قطاع مياه وتحديدًا في تنبؤ هطول الأمطار في مواقع متعددة في مجتمعات مياه منطقة شلف الأوسط (الجزائر)، وتحسين طاقته التنبؤ عن خطوة سنوية وشهرية، وبعدها مكاني وزماني.

لقيام بذلك، استخدمنا بيانات هطول الأمطار من 5 محطات في منطقة دراسة، وحظت على مدى 11 عامًا (2000-2010). و أظهرت النتائج أن الالاتلافات بين الملاحظات والتنبؤات التي تم الحصول عليها هي الحد الأدنى وذلك خطوات وأعاد مختلفة.

وأيرا، تم التحقق من أداء مرشح كالمان من خلال رسوم بيانية الأثر، واكتساب مرشح كالمان ومعياري الانحرافات، ناهيك عن احتساب خطأ نسبي متوسط للتنبؤات و تحصلنا على نتيجة اقل من 10%. وقد تبين أيضا أن نتائج التي تم الحصول عليها لنسبة لخطوة شهرية أكثر دقة من تلك التي تم الحصول عليها في خطوة سنوية، مما علنا نستنتج أن مرشح كالمان أكثر كفاءة في خطوة شهرية مقارنة بخطوة سنوية.

الكلمات المفتاحية: مرشح كالمان، تنبؤ في مواقع متعددة، هطول الأمطار، مجتمعات مياه، شلف الأوسط، الجزائر.

Résumé

Notre objectif dans ce travail est de tester l'efficacité du filtre de Kalman discret dans le domaine hydrologique et plus précisément dans la prédiction multi-site de la pluviométrie dans le bassin versant Moyen Cheliff (Algérie), et l'amélioration de l'erreur de prédiction pour un pas annuel et mensuel, et pour une dimension spatiale et temporelle. Pour ce faire, nous avons utilisés les données pluviométriques de 5 stations de notre région d'étude, observées sur une période de 11 ans (2000-2010). Les résultats obtenus montrent que les écarts entre les observations et les prédictions obtenues sont minimales et cela pour les différents pas et dimensions. Enfin, la performance du filtre de Kalman a été vérifiée grâce aux graphes de la trace, du gain du FK et des écarts types, sans oublier le calcul de l'erreur relative moyenne des prédictions qui est inférieure à 10%. On a aussi constaté que les résultats obtenus pour le pas mensuel sont plus précis que ceux obtenus pour le pas annuel, ce qui nous laisse dire que le filtre de Kalman est plus performant pour le pas mensuel comparé au pas annuel.

Mots-clés : Filtre de Kalman, Prédiction multi-site, Pluviométrie, Bassin versant, moyen Chéiff, Algérie.

Abstract

The objective of this work is to test the forcefulness of the discrete Kalman filter in the hydrological domain and more precisely in the multi-site prediction of rainfall in the Middle Cheliff watershed area (Algeria), the improvement of the error of prediction for an annual and monthly step, and for a spatial and temporal dimension. To do this, rainfall data from five stations are used in our study area over a period of 11 years (2000-2010). The results obtained show that the differences between the observations and the predictions are minimal and this for the different steps and dimensions.

Finally, the performance of the Kalman filter was checked using trace graphs, FK gain and standard deviations, as well as calculating the mean relative error of the predictions, which is less than 10%. It was also found that the results obtained for the monthly step are more precise than those obtained for the annual step, which allows to say that the Kalman filter is better for the monthly step compared to the annual step.

Keywords: Kalman filter, Multi-site prediction, Rainfall, Watershed, Middle Chéiff, Algeria.

Sommaire

Introduction générale

Chapitre I : Notion de base pour aborder le filtre de Kalman

Introduction	Page 2
I.1 Probabilité et variables aléatoires	Page 2
I.1.1 Loi de probabilité	Page 2
I.1.2 Variable aléatoire réelle	Page 3
I.1.3 Loi de probabilité à plusieurs variables	Page 3
I.1.4 Espérance, variance et moments d'une variable aléatoire réelle à densité de probabilité	Page 4
I.1.5 Densité de probabilité d'un vecteur aléatoire	Page 4
I.1.6 Densité de Probabilité d'un vecteur aléatoire	Page 5
I.1.7 Variable aléatoires discrètes	Page 5
I.1.8 Fonction de répartition	Page 5
I.1.9 Probabilité conditionnelle	Page 6
I.1.10 Densité conditionnelle	Page 6
I.2 Rappel statistique et méthode d'estimation	Page 6
I.2.1 Rappel statistique	Page 6
I.2.2 Test relatif à l'échantillonnage	Page 9
I.2.3 Choix du modèle fréquentiel	Page 11
I.3 Processus stochastique et séries chronologiques	Page 12
I.3.1 Estimation stochastique linéaire	Page 12
I.3.2 Modèle espace d'états ou modèle état mesure	Page 13

I.3.3 Problème du filtrage	Page 13
Conclusion	Page 13

Chapitre II : Méthode du filtre de Kalman

Introduction	Page 16
II.1 Généralité sur le filtre de Kalman	Page 16
II.1.1 Historique du Filtre de Kalman	Page 16
II.1.2 Domaine d'utilisation du Filtre de Kalman	Page 17
II.1.3 Définition du Filtre de Kalman	Page 17
II.1.4 Principe du filtre de Kalman	Page 17
II.1.5 Les limites du Filtre de Kalman	Page 18
II.2 L'estimateur optimal du filtre de Kalman	Page 18
II.2.1 L'estimateur	Page 18
II.2.2 Les estimateurs déterministes	Page 18
II.2.3 Estimateur optimale	Page 19
II.3 Mécanisme de la méthode de filtre de Kalman	Page 20
II.3.1 Variables et équations d'états	Page 20
II.4 Présentation de l'algorithme récursif du filtre de Kalman discret	Page 23
II.4.1 Les différentes étapes d'évaluation des paramètres du système de FK	Page 23
II.4.2 Stabilité du filtre vis-a-vis de la divergence	Page 24
Conclusion	Page 25

Chapitre III : Présentation de la région d'étude et traitement des données

Introduction	Page 27
III.1 Présentation de la région d'étude	Page 29

III.1.1 Situation géographique du bassin versant	Page 29
III.2 Le réseau pluviométrique	Page 30
III.2.1 Acquisition des données	Page 30
III.2.2 Caractéristiques du réseau pluviométrique	Page 30
III.3 Traitement des données hydrologiques	Page 30
III.3.1 Test d'Homogénéité	Page 30
III.3.2 Ajustement de la série pluviométrique	Page 31
Conclusion	Page 40

Chapitre IV : Application du filtre de Kalman

Introduction	Page 42
IV.1 Données de la prédiction multi-site	Page 42
IV.2 Hypothèses	Page 42
IV.3 Formulation espace-état du modèle adopté	Page 42
IV.5 Matrice de covariance des bruits du système et celle des bruits des mesures	Page 45
IV.6 Matrice de transition d'état et matrice de mesure	Page 45
IV.7 Les équations récurrentes du filtre de Kalman	Page 45
IV.7.1 Matrice gain du filtre	Page 45
IV.7.2 Utilisation de la mesure pour actualiser l'estimation	Page 46
IV.7.3 Matrice de covariance d'erreur associée à l'estimation (actualisée) corrigée par la mesure	Page 46
IV.7.4 Prédiction un pas en avant de l'estimation (actualisée) corrigée par la mesure	Page 47
IV.7.5 Matrice de covariance de l'erreur de prédiction un pas en avant	Page 47

IV.8 Résultats et interprétations	Page 47
IV.8.1 Prédictions multi-site	Page 48
IV.9 Analyse des innovations du filtre (résidus)	Page 56
IV.10 Comparaison entre le pas annuel et mensuel	Page 58
IV.10.1 Moyennes et déviation standard des prédictions multi-sites	Page 58
Conclusion	Page 60
Conclusion générale	
Références bibliographiques	

LISTE DES TABLEAUX

Tableau N°III.1 - Les Sous-régions du Bassin Cheliff-Zahrez	Page 28
Tableau N°III.2 - Coordonnées des stations hydrométriques	Page 30
Tableau N°III.3 Résultat de test d'homogénéité pour les cinq stations hydrométriques	Page 31
Tableau N°III.4- Caractéristiques empiriques de la série pluviométrique 1 (BENAIRIA)	Page 32
Tableau N°III.5 - Résultat du test de Khi-2 pour les cinq stations hydrométriques	Page 33
Tableau N°III.6 - Résultat de l'ajustement	Page 34

LISTE DE FIGURES

Figure N°I.1 : Courbe de la loi normale	Page 11
Figure N°II.1 : Schéma explicatif du mécanisme du filtre de Kalman	Page 21
Figure N°II.2 : Algorithme du Filtre de Kalman	Page 23
Figure N°III.1 : Représentation des sous-bassins composants la région de Cheliff	Page 27
Figure N°III.2 : Représentation des sous-régions du bassin Cheliff-Zahrez	Page 28
Figure N°III.3 : Situation géographique de la sous-région N°2	Page 29
Figure N° III.4 : Ajustement de la série N°1 à la loi de Gumbel	Page 35
Figure N° III.5 : Ajustement de la série N°1 à la loi de Lognormale	Page 35
Figure N° III.6 : Ajustement de la série N°2 à la loi de Gumbel	Page 36
Figure N° III.7 : Ajustement de la série N°2 à la loi de Lognormale	Page 36
Figure N° III.8 : Ajustement de la série N°3 à la loi de Gumbel	Page 37
Figure N° III.9 : Ajustement de la série N°3 à la loi de Lognormale	Page 37
Figure N° III.10 : Ajustement de la série N°4 à la loi de Gumbel	Page 38
Figure N° III.11 : Ajustement de la série N°4 à la loi de Lognormale	Page 38
Figure N° III.12 : Ajustement de la série N°5 à la loi de Gumbel	Page 39
Figure N° III.13 : Ajustement de la série N°5 à la loi de Lognormale	Page 39
Figure N° IV.1 : Gain du filtre de Kalman (Pas annuel)	Page 48
Figure N° IV.2 : Trace de la matrice de covariance d'erreur (Pas annuel)	Page 49
Figure N° IV.3 : Prédictions annuelles des précipitations aux 5 stations pluviométriques (exp : Année 2002 et 2010)	Page 50
Figure N° IV.4 : Prédictions annuelles des précipitations en fonction des années (exp : Station N° 2 et Station N°4)	Page 51
Figure N° IV.5 : Gain du filtre de Kalman (Pas mensuel)	Page 52

Figure N° IV.6 : Trace de la matrice de covariance d'erreur (Pas mensuel)	Page 53
Figure N° IV.7 : Prédictions mensuelles des précipitations aux cinq stations pluviométriques (exp : Mars 2000 et Octobre 2010)	Page 54
Figure N° IV.8 : Prédictions mensuelles des précipitations en fonction des mois (exp : Station N°3 et Station N°4)	Page 55
Figure N° IV.9 : Innovation du FK à la Station N°1 (Pas Annuel)	Page 57
Figure N° IV.10 : Innovation du FK à la Station N°1 (Pas Mensuel)	Page 57
Figure N° IV.11 : Ecart-type moyen des observations et des prédictions annuelles aux 5 stations (2000-2011)	Page 58
Figure N° IV.12 : Ecart-type moyen des observations et des prédictions mensuelles aux 5	Page 59

INTRODUCTION GENERALE

L'hydrologie est la science de l'eau. Molécule, gaz, liquide ou solide, l'eau voit son étude ressortir à la physique et à la chimie. C'est à l'étude de l'eau dans la nature, ou s'expriment évidemment ses propriétés physico-chimiques, auxquels est consacrée l'hydrologie.

Le Filtre de Kalman est un algorithme optimal et récursif de traitement de données, autrement dit, il n'exige pas de garder ou de stocker en mémoire l'intégralité des données précédentes, pour être prises en considération dans les calculs à chaque fois qu'une nouvelle mesure lui est fournie. Ceci est une propriété très importante qui rend l'implémentation du filtre très pratique. Il peut aussi être défini comme étant une approche statistique d'assimilation de données, dont le principe est de corriger la trajectoire du modèle en combinant les observations avec l'information fournie par le modèle, de façon à minimiser l'erreur entre l'état vrai et l'état filtré.

Notre application consiste en l'utilisation du Filtre de Kalman discret dans la variabilité pluviométrique dans le bassin versant du Moyen Cheliff afin d'étudier son applicabilité à la modélisation et la prédiction multi-site de la pluviométrie, ainsi que l'amélioration des erreurs de prédiction et ce pour le pas mensuel et annuel.

A cette fin, on aura besoin des précipitations mensuelles et annuelles observées dans cinq stations hydrométriques pour une période de 11 ans qu'on utilisera pour tracer des graphes spatiales et temporelles, et cela pour un pas annuel et mensuel.

Dans ce cadre, notre travail a été structuré comme suit :

Dans le premier chapitre, nous avons fait un rappel sur les lois de probabilités et de statistiques, le processus stochastique et le problème de filtrage.

Dans le deuxième chapitre, nous avons donné des généralités sur le Filtre de Kalman tout en présentant le mécanisme de la méthode du Filtre de Kalman et son algorithme.

Dans le troisième chapitre, nous avons procédé à la présentation de la zone d'étude, l'homogénéisation des données pluviométrique et leur ajustement.

Dans le quatrième et dernier chapitre, nous avons appliqué le Filtre de Kalman à nos séries pluviométriques et comparé les résultats entre les différentes phases de travail.

**LE CHAPITRE I : NOTIONS DE
BASE POUR ABORDER LE FILTRE
DE KALMAN**

Introduction

La théorie des probabilités est l'étude mathématique des phénomènes caractérisée par le hasard et l'incertitude. La statistique est, quant à elle, l'activité qui consiste à recueillir, traiter et interpréter un ensemble de données. Il existe des interconnexions entre ces domaines des sciences de l'aléatoire.

Ces domaines mathématiques sont en relation avec les autres domaines mathématiques comme l'algorithme, l'analyse, l'informatique théorique ou la logique. Les probabilités se retrouvent dans la théorie des jeux, la biologie, l'économie ou la physique, entre autres. On retrouve la statistique dans des domaines comme l'économie, la physique, la sociologie ...etc[1]

I.1 Probabilité et variables aléatoires

I.1.1 Loi de probabilité

En probabilités, nous utilisons beaucoup de mots du langage courant, mais ceux-ci revêtent un sens précis, il faut par conséquent les définir et les commenter.

Epreuve : on désigne par ce mot une expérience concrète : lancer un dé, prendre un objet au hasard pour en noter ses caractéristiques, prélever un échantillon d'une population en vue d'une observation quelconque.

Univers Ω des éventualités : c'est l'ensemble des résultats possibles d'une épreuve. Ainsi par exemple, en lançant un dé dont les faces sont numérotées de 1 à 6, a-t-on $\Omega = \{1, 2, 3, 4, 5, 6\}$. Le symbole Ω représente un ensemble fini et les éléments de cette ensemble sont appelés éventualités ou évènements élémentaires.

Évènement : C'est un type de résultat pouvant se produire lors d'une épreuve. Il peut être caractérisé par une proposition logique et doit pouvoir se ramener sans ambiguïté à une partie de l'univers Ω . Ainsi par exemple si on prend l'ensemble $\Omega = \{1, 2, 3, 4, 5, 6\}$, l'évènement décrit par la proposition p « être pair » correspond à la partie $\{2, 4, 6\}$.

Évènements certains et impossibles : Ce sont ceux pouvant se mathématiser respectivement par les parties pleine Ω et vide. Par exemple en prenant l'ensemble $\Omega = \{1, 2, 3, 4, 5, 6\}$, la phrase « obtenir 7 » caractérise l'évènement impossible.

Conjonction et disjonction de deux évènements : la conjonction « et » de deux évènements correspond à l'intersection de ces derniers, et la disjonction à leurs union.

Ainsi si nous notons A l'évènement « nombre pair » et B l'évènement « mois de 4 », nous avons :

$$A = \{2, 4, 6\} \qquad B = \{1, 2, 3\}$$

$A \cap B = \{2\}$ représente l'évènement « un résultat pair et inférieur à 4 »

$A \cup B = \{1, 2, 3, 4, 5, 6\}$ représente l'évènement « un résultat pair ou inférieur à 4 ».

Événements contraires : ce mot correspond bien évidemment à la négation logique, et à la notion ensembliste de complémentaire. Le contraire est le plus souvent par une barre de surlignement.

Événements incompatibles : deux évènements sont dits incompatibles s'ils ne peuvent se produire en même temps (ils ne sont pas obligatoirement contraire). Ce sont donc simplement des parties disjointes de Ω :

$$A, B \text{ sont incompatibles} \leftrightarrow A \cap B = \emptyset$$

Système complet (ou exhaustif) d'événements : c'est une partition de Ω , c'est-à-dire un sous-ensemble A_1, A_2, \dots, A_n de $P(\Omega)$ de parties incompatibles deux à deux, et dont la réunion constitue Ω . (**Louis Gacogne/Gérard Frugier, 1990**)

Probabilité : Par définition, une probabilité P sur l'ensemble Ω des éventualités est une application $P(\Omega) \rightarrow [0, 1]$ qui possède les deux propriétés suivantes :

1 est la probabilité maximale pour l'événement certain $P(\Omega) = 1$

Si A et B sont incompatibles : $P(A \cup B) = P(A) + P(B)$.

I.1.2 Variable aléatoire réelle

Une variables aléatoire réelle est une variable aléatoire à valeurs dans R , ou une partie de R , c'est une fonction définie depuis l'ensemble des résultats possibles d'une expérience aléatoire, dont on doit déterminer la probabilité qu'elle prenne une valeur donnée ou un ensemble donnée de valeurs. Les variables aléatoires réelles sont les variables aléatoires les plus couramment étudiées, ce qui conduit les auteurs à omettre l'adjectif réel, et à parler de variables aléatoires tout court. (**Jean-Yves DAUXOIS, 2013**)

I.1.3 Loi de probabilité à plusieurs variables

Dans certains problèmes interviennent simultanément plusieurs variables aléatoires. Mis à part les cas particulier de variables indépendantes et de variables liées fonctionnellement, cela introduit la notion de loi de probabilité à plusieurs variables autrement appelée loi jointe. La description des notions correspondantes dont certaines généralisent les notions relatives à une seule variable, est simplifiée de deux manières :

- Seules les variables continues sont considérées. Il est possible de passer aux variables discrètes en utilisant la fonction de Heaviside et la fonction de Dirac.
- Pour éviter la lourdeur des formules, l'exposé est limité à deux variables. (**Jean-Yves DAUXOIS, 2013**)

I.1.4 Espérance, variance et moments d'une variable aléatoire réelle à densité de probabilité

Soit X une variable aléatoire réelle ayant une densité de probabilité f . D'après le théorème de transfert, X possède un moment d'ordre k si et seulement si l'intégrale :

$$\int_{-\infty}^{+\infty} |t|^k f(t) dt \text{ est finie}$$

Dans ce cas :

$$E[X^k] = \int_{-\infty}^{+\infty} t^k f(t) dt \dots \text{I.1}$$

L'espérance sera :

$$E[X] = \int_{-\infty}^{+\infty} t f(t) dt \dots \text{I.2}$$

Le moment d'ordre 2 :

$$E[X^2] = \int_{-\infty}^{+\infty} t^2 f(t) dt \dots \text{I.3}$$

La variance :

$$V(X) = \int_{-\infty}^{+\infty} t^2 f(t) dt - \left(\int_{-\infty}^{+\infty} t f(t) dt \right)^2 \dots \text{I.3}$$

I.1.5 Densité de probabilité d'un vecteur aléatoire

Un vecteur aléatoire est une généralisation à n dimensions d'une variable aléatoire réelle. Alors qu'une variable aléatoire réelle est une fonction qui à chaque éventualité fait correspondre un nombre réel, le vecteur aléatoire est une fonction X qui à chaque éventualité fait correspondre un vecteur de \mathbb{R}^n (**Jean-Yves DAUXOIS, 2013**)

Les applications x_1, x_2, \dots, x_n sont des variables aléatoires réelles appelées composantes du vecteur aléatoire X alors $X = (x_1, x_2, \dots, x_n)$

Une fonction $f: \mathbb{R}^n \rightarrow [0,1]$ est une densité c'est elle satisfait

$$\int_{\mathbb{R}^n} f(x_1, x_2, x_3, \dots, x_n) dx_1 \dots dx_n = 1$$

On dit qu'un vecteur aléatoire $X = (x_1, x_2, \dots, x_n)$ admet la densité f si :

$$\mathbb{P}(X_1 \leq a_1, \dots, X_n \leq a_n) = \int_{]-\infty, a_1] \dots]-\infty, a_n]} f(x_1, x_2, x_3, \dots, x_n) dx_1 \dots dx_n = 1$$

I.1.6 Densité de Probabilité d'un vecteur aléatoire

On dit qu'un vecteur aléatoire $X = (X_1, \dots, X_n)$ est absolument continu(e) si il existe une fonction mesurable

$$f: (\mathbb{R}^n, \beta_{\mathbb{R}^n}) \rightarrow (\mathbb{R}^+, \beta_{\mathbb{R}^+})$$

Telle que, pour tout $x = (x_1, \dots, x_n)$ dans \mathbb{R}^n , on ait :

$$P_X (]-\infty, x_1] \times \dots \times]-\infty, x_n]) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} f_X(u_1, \dots, u_n) du_1 \dots du_n$$

La fonction f_X est appelée densité de probabilité conjointe du vecteur X .

Toute densité de probabilité conjointe f_X de R_n vérifie les trois assertions suivantes :

- i) f_X est positive.
- ii) f_X est mesurable.
- iii) f_X est intégrable et : $\int_{R_n} f_X(x_1, \dots, x_n) dx_1 \dots dx_n = 1$

Réciproquement toute fonction f_X dans R_n vérifiant i), ii), iii) est une densité de probabilité. **(Jean-Yves DAUXOIS, 2013)**

I.1.7 Variable aléatoires discrètes

On dit d'une variable aléatoire est discrète si elle ne peut prendre que des valeurs numériques isolées. Dans le cas contraire, on dit qu'elle est continue.

Une variable aléatoire discrète finie est donc une application

$$\Omega \xrightarrow{X} \text{Ensemble } \{x_1, \dots, x_n\} \text{ fini des valeurs de } X.$$

Ce sera le cas chaque fois que Ω est fini, par exemple la somme des numéros des faces lors du lancer de trois dés, le nombre d'enfants d'une famille ...

Une variable aléatoire discrète dénombrable est une application de Ω dans N qui est infini, mais formé de valeurs bien isolées les unes des autres. **(Louis Gacogne/Gérard Frugier, 1990)**

I.1.8 Fonction de répartition

Cette notion est surtout utile à la description des lois continues, notamment à cause de sa dérivée lorsqu'elle existe, qui détermine la densité de probabilité en un point. **(Louis Gacogne/Gérard Frugier, 1990)**

La fonction de répartition F associée à une variable aléatoire X , est la fonction définie sur R , à valeurs dans $[0,1]$ par :

$$F(x) = P(X \leq x) \dots I.4$$

I.1.9 Probabilité conditionnelle

Pour tout évènement B de probabilité non nulle, on appelle probabilité conditionnelle à B , la probabilité sur (Ω, A)

$$P^B : A \rightarrow [0, 1] \qquad A \rightarrow P^B(A) = \frac{P(A \cap B)}{P(B)}$$

$P^B(A)$ s'appelle probabilité conditionnelle à B de A (ou encore probabilité de A sachant B . On note aussi

$$P^B(A) = P(A / B)$$

I.1.10 Densité conditionnelle

Supposons que $Z = (X, Y)$ soit un vecteur dans $R^m \times R^n$ de variables aléatoires discrètes.

La fonction :

$$f_Y^X = x : R^m \rightarrow R^+$$

$$y \rightarrow \frac{f_Z(x, y)}{f_X(x)}$$

Est une densité de probabilité sur R^m et est appelée densité conditionnelle de Y sachant que $X = x$. On note $P_Y^X = x$ la loi de probabilité associée, appelée loi de Y sachant que $X = x$.

I.2 Rappel statistique et méthode d'estimation

I.2.1 Rappel statistique

A. Paramètre de l'échantillon

A.1 Paramètre à tendance centrale

A.1.1 Le mode

a) Variables quantitatives discrètes (non classées)

Le mode correspond à la valeur de la variable pour laquelle l'effectif ou la fréquence est le plus grand. Certaines variables peuvent présenter plusieurs modes.

b) Variable quantitative continue ou discrète classée

La classe modale est la classe dont la fréquence par unité d'amplitude est la plus élevée, cette classe correspond donc au rectangle le plus haut de l'histogramme des fréquences.

A.1.2 La médiane

La médiane (notée M_e) d'une variable quantitatives est la valeur de cette variable qui permet de scinder la population étudiée en deux sous-population de même effectif. Plus précisément, il y a autant d'individus pour lesquels on a observé une valeur supérieur à M_e que d'individus pour lesquels on a observé une valeur inférieure à M_e . (**Antoine Ayache & Julien Harmonier**)

A.1.3 La moyenne

On dispose d'une population de N individus et on observe x_1, x_2, \dots, x_N les valeurs d'une variable quantitative discrète X pour ces individus. (**Antoine Ayache & Julien Harmonier**)

a) Moyenne arithmétique

Elle est notée par \bar{x} et elle est définie de la manière suivante :

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_N}{N} = \frac{1}{N} \sum_{i=1}^N x_i \quad \dots \text{I.5}$$

b) Moyenne quadratique

Elle est notée par m_2 et elle est définie par la manière suivante :

$$m_2 = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2} = \sqrt{\sum_{i=1}^K f_i x_i^2} \quad \dots \text{I.6}$$

c) Moyenne harmonique

Elle est notée par m_{-1} et elle est définie de la manière suivante :

$$m_{-1} = \frac{N}{\sum_{i=1}^N \frac{1}{x_i}} = \frac{1}{\sum_{i=1}^K \frac{f_i}{x_i}} \quad \dots \text{I.7}$$

d) Moyenne géométrique

On ne peut définir cette moyenne que lorsque les observations x_1, \dots, x_N sont toutes des membres réels positifs. Si tel est le cas, la moyenne géométrique de ces observations est notée par M_g , et elle est définie par :

$$M_g = \sqrt[N]{x_1 x_2 \dots x_N} = \sqrt[N]{x_1^{n_1} \dots x_K^{n_K}} = x_1^{f_1} \dots x_K^{f_K} \quad \dots \text{I.8}$$

A.2 Paramètre de dispersion

On dispose d'une population de N individus et on observe x_1, \dots, x_N les valeurs d'une variable quantitative discrète X pour ces individus.

A.2.1 Étendue

L'étendue e_X de la variable quantitative discrète X est la différence entre la plus grande et la plus petite des valeurs observées :

$$e_X = \max_{1 \leq i \leq N} x_i - \min_{1 \leq i \leq N} x_i \dots \text{I.9}$$

A.2.2 Variance et écart-type

La variance de la variable quantitative X , notée par $\text{Var}(X)$, est, par définition, la moyenne arithmétique des carrés des écarts à la moyenne arithmétique :

$$\text{Var}(X) = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2 \dots \text{I.10}$$

Cette formule peut également se réécrire sous la forme :

$$\text{Var}(X) = \sum_{i=1}^K f_i (x_i - \bar{x})^2 \dots \text{I.11}$$

Où K désigne le nombre de valeurs distinctes de X et $f_i = n_i/N$ est la fréquence de la valeur x_i . Une autre formule importante (parfois désignée par la formule de Huygens) permettant le calcul de la variance est :

$$\text{Var}(X) = \left(\frac{1}{N} \sum_{i=1}^N x_i^2 \right) - (\bar{x})^2 = \left(\sum_{i=1}^K f_i x_i^2 \right) - (\bar{x})^2 \dots \text{I.12}$$

$$= (\text{Moyenne quadratique de } X)^2 - (\text{Moyenne arithmétique de } X)^2$$

L'écart type de la variable X , noté par σ_X , est, par définition, la racine carrée de la variance de cette variable :

$$\sigma_X = \sqrt{\text{Var}(X)} \dots \text{I.13}$$

A.2.3 L'écart absolu moyen

L'écart absolu moyen à la variable quantitative discrète X est la moyenne arithmétique des valeurs absolues des écarts à la moyenne arithmétique :

$$e_{\bar{x}} = \frac{1}{N} \sum_{i=1}^N |x_i - \bar{x}| = \sum_{i=1}^K f_i |x_i - \bar{x}| \dots \text{I.14}$$

Où K désigne le nombre de valeurs distinctes de X et f_i la fréquence de x_i . (**Antoine Ayache & Julien Harmonier**)

A.2.4 Coefficient de variation

Pour une variable statistique réelle X , on appelle le coefficient de variation le rapport

$$c = \frac{s(X)}{\bar{X}} \dots \text{I.15}$$

Le coefficient de variation est un nombre sans dimension qui permet de comparer deux variables statistiques de natures différentes. (**Antoine Ayache & Julien Harmonier**)

B. Paramètres de formes

B.1 Coefficient d'asymétrie

Il existe plusieurs coefficients d'asymétrie. Les principaux sont les suivants :

- a) Le coefficient d'asymétrie de **Pearson** fait intervenir le mode M_0 : quand il existe, il est définie par :

$$P = \frac{\bar{X} - M_0}{s(X)} \dots \text{I.16}$$

- b) Le coefficient d'asymétrie de **Yule** fait intervenir la médiane et les quartiles, il est défini par :

$$Y = \frac{Q_1 + Q_3 - 2M_e}{2(Q_3 - Q_1)} \dots \text{I.17}$$

Tout en sachant que les quartiles sont les nombres réels Q_1, Q_2, Q_3 pour lesquels les fréquences cumulées de X sont respectivement 0,25 ; 0,5 ; 0,75.

- c) Le coefficient d'asymétrie de **Fisher** fait intervenir les moments centrés, il est défini par :

$$F = \frac{\mu_3}{\frac{\mu_2^{\frac{3}{2}}}{s^3(X)}} \dots \text{I.18}$$

Lorsque le coefficient d'asymétrie est positif, la distribution est plus étalée à droite : on dit qu'il y a **oblicité à gauche**.

Lorsque le coefficient d'asymétrie est négatif, la distribution est plus étalée à gauche : on dit qu'il y a **oblicité à droite**.

B.2 Coefficient d'aplatissement

Là encore plusieurs définitions sont possibles.

Le coefficient d'aplatissement de **Pearson** est : $\beta_2 = \frac{\mu_4}{\mu_2^2}$

Le coefficient d'aplatissement de **Yule** est : $F_2 = \frac{\mu_4}{\mu_2^2} - 3$

I.2.2 Test relatif à l'échantillonnage

A) Test de stationnarité et d'homogénéité de Mann et Whitney (1947)

Le test de la médiane utilise très peu d'informations et comme le test du signe il est peu puissant. On introduit alors le test de Mann Whitney qui est une extension du test des

signes et rangs. Ce test est utilisé pour comparer deux échantillons qui ne peuvent se distinguer que par un glissement de leur position. Aussi, pour utiliser ce test, on fait l'hypothèse que la dispersion des deux échantillons est comparable.

B) Test de stationnarité et d'indépendance de Wald et Wolfowitz (1943)

Le test des suites de Wald-Wolfowitz est une alternative non-paramétrique au test t pour des échantillons indépendants. La procédure nécessite une organisation des données similaires à celles utilisées pour un test t sur des échantillons indépendants.

Ce test suppose que la variable en question est continue, et qu'elle soit mesurée sur au moins une échelle ordinale (c'est-à-dire avec des rangs). Le test des suites de Wald-Wolfowitz permet de tester l'hypothèse selon laquelle les deux échantillons indépendants sont tirés de deux populations divergeant sur plusieurs aspects, c'est-à-dire non seulement par rapport à la moyenne, mais également par rapport à la forme générale de la distribution. L'hypothèse nulle stipule que les deux échantillons appartiennent à la même population. Ainsi, ce test est différent du test paramétrique t qui ne teste strictement que les différences de positions (moyennes) entre les deux échantillons.

C) Test d'Anderson

La statistique d'Anderson-Darling (AD) détermine dans quelle mesure les données suivent une distribution spécifique. Généralement, plus la distribution s'ajuste correctement aux données, plus cette statistique est faible.

La statistique d'AD est utilisée pour calculer la valeur de p du test d'adéquation de l'ajustement, ce qui nous aide à identifier la loi qui s'ajuste le mieux à nos données. Elle est également utilisée pour vérifier si un échantillon de données provient d'une population avec une distribution spécifique.

Les hypothèses du test d'Anderson-Darling sont les suivantes :

- H_0 : les données suivent une distribution spécifiée.
- H_1 : les données ne suivent pas une distribution spécifiée.

Si la valeur de p du test d'Anderson-Darling est inférieure au seuil de signification choisi (généralement 0,05 ou 0,10), on peut en conclure que les données ne suivent pas la distribution spécifiée. Minitab n'affiche pas systématiquement une valeur de p pour le test d'Anderson-Darling car elle n'existe pas mathématiquement dans certains cas.

D) Test du cumul des résidus

Dans la méthode du cumul des résidus, on calcule d'abord la régression entre la série étudiée et la série de référence. On peut alors former la série des résidus (le résidu étant l'écart entre la valeur vraie et la valeur estimée par régression). On effectue ensuite la somme de ces

résidus que l'on pointe sur un graphe en fonction du temps. Si la série testée est homogène, la distribution de ces résidus est complètement aléatoire.

Par contre, si la série est hétérogène, la courbe présente des portions anormalement croissantes ou décroissantes. Si elle sort d'une ellipse de significativité définie pour un intervalle de confiance donné, on peut conclure qu'une hétérogénéité est détectée à la cassure de la courbe (Bois, 1986).

I.2.3 Choix du modèle fréquentiel

A- Distributions théoriques

A.1 Loi normale ou loi de Gauss

La loi normale est la loi statistique la plus répandue et la plus utile, elle représente beaucoup de phénomènes aléatoires. De plus, de nombreuses autres lois statistiques peuvent être approchées par la loi normale, tout spécialement dans le cas des grands échantillons.

Son expression mathématique est la suivante:

$$n(x) = \frac{n}{\sqrt{2\pi} \sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \dots \text{I.19}$$

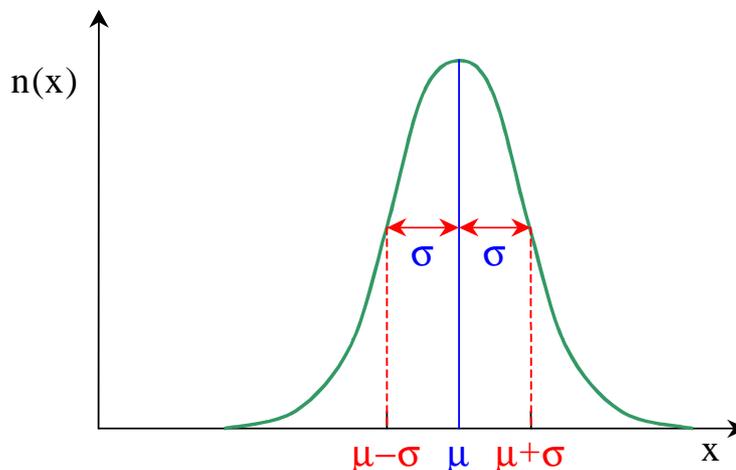


Figure N°I.1 : Courbe de la loi normale

- μ : La moyenne.
- σ : L'écart type.
- n : Le nombre total d'individus dans l'échantillon.
- $n(x)$: Le nombre d'individus pour lesquels la grandeur analysée a la valeur x .

A.2 Loi log-normale

La loi log-normale est un cas particulier de la loi normale. On dit que X suit une loi log-normale, si son logarithme, $Y = \ln(X)$, suit une loi normale. Les conditions de formation de la loi log-normale est identique à la loi normale, à la différence que les facteurs agissent de manière multiplicative.

C'est une loi définie sur R^+ , asymétrique, étalée à droite, sa densité de probabilité est :

$$f(x) = \frac{1}{\sigma x \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{\ln x - \mu}{\sigma} \right)^2} \dots \text{I.20}$$

A.3 Loi de Gumbel

La loi de Gumbel est utilisée pour modéliser les valeurs extrêmes. Par exemple, pour définir de manière adéquate la puissance d'un serveur, on s'intéresse au nombre maximal d'accès simultanés à un site web dans une journée, observé sur plusieurs jours.

La loi n'est pas paramétrée, sa fonction de densité est :

$$f(x) = e^{-x-e^{-x}} \dots \text{I.21}$$

La fonction de répartition est obtenue aisément :

$$f(x) = e^{-e^{-x}} \dots \text{I.22}$$

I.3 PROCESSUS STOCHASTIQUE ET SERIES CHRONOLOGIQUES

I.3.1 Estimation stochastique linéaire

On appelle processus stochastique ou processus aléatoire toute famille de variables aléatoires X_t . Cela signifie qu'à tout $t \in T$ est associée une variable aléatoire prenant ses valeurs dans un ensemble numérique E . On note le processus X_t . Si T est dénombrable, on dit que le processus est discret, si T est un intervalle, on dit que le processus est permanent.

Un processus aléatoire (ou stochastique) peut être défini de deux façons:

- Une application de S -- l'espace des réalisations -- dans un espace de fonctions de variable réelle (temps)

$$x(t): S \rightarrow F \\ \omega \mapsto \{x(t, \omega), t \in T\}$$

Que, à chaque événement élémentaire fait correspondre une fonction du temps;

- Une collection de variables aléatoires indexées

$$x(t) = \{x(t), t \in T\},$$

Où $x(t)$ est une variable aléatoire pour chaque t .

Les deux définitions sont équivalentes, et on utilisera une et l'autre selon le problème en étude.

I.3.2 Modèle espace d'états ou modèle état mesure

L'étude de systèmes physiques émettant au cours du temps des signaux déterminés par des états internes non observés, a conduit à développer en traitement du signal (une branche des sciences de l'ingénieur) les modèles dits espace-état. Dans un premier temps, ces modèles sont présentés en comparaison des techniques économétriques plus usuelles pour modéliser les séries temporelles. Les méthodes d'estimations de tels modèles sont ensuite expliquées en deux temps : l'estimation des variables cachées d'abord (avec le filtre de Kalman), puis celle des paramètres (avec l'algorithme EM).

Les modèles espace-état intègrent la distinction entre les variables observées (le signal) et les variables cachées (l'état interne). Les modèles espace-état peuvent s'écrire sous une forme dite développée (encadrée) qui implique, sous certaines hypothèses, que la variable de mesure s'écrit sous la forme d'une moyenne mobile. Où réside donc l'intérêt de recourir aux modèles espace-état pour décomposer ou prévoir une série temporelle relativement aux techniques classiques développées par Box et Jenkins (1976). On s'épargne ici les problèmes de stationnarité et de racine unitaire qui se posent préalablement à l'estimation d'un modèle ARMA. En particulier, les résultats du filtre de Kalman restent valides en présence de séries non-stationnaires. Par ailleurs, ce cadre permet également de relâcher l'hypothèse d'une distribution gaussienne pour les bruits.

Dans leur version élémentaire, les modèles espace-état reposent sur un certain nombre d'hypothèses principales : les équations de mesure et d'état sont linéaires ; les bruits d'observation et d'innovation sont des bruits blancs, les variables cachées suivent à un instant initial donné une loi gaussienne. À ces dernières, se sont ajoutées des hypothèses secondaires permettant de déterminer la forme canonique : l'indépendance entre les bruits d'observation et d'innovation et l'indépendance entre la variable cachée initiale et ces bruits. Toutes ces hypothèses sont destinées à simplifier les procédures d'estimation.

I.3.3 Problème du filtrage

Le filtrage est une opération fondamentale en traitement du signal et en automatique. Par filtrage, on entend généralement plusieurs types de traitements, qu'il convient de préciser : – le filtrage (filtering) proprement dit, qui consiste à extraire de l'information à l'instant t à partir de données disponibles jusqu'à l'instant t inclus ; – le lissage (smoothing) qui consiste à extraire de l'information à l'instant t , à partir de données observées au-delà de l'instant t ; le lissage exploitant plus d'informations que le filtrage conduit à des résultats plus performants ; – la prédiction (prediction) qui consiste à prédire la valeur d'une quantité d'intérêt à l'instant $t + \tau$, connaissant cette quantité jusqu'à l'instant t inclus. (**Boukharouba, 2009**)

Conclusion

Dans ce chapitre nous avons :

- Cité les différentes lois de probabilité et de statistique.
- Donn  un aper u sur le processus stochastique et le probl me de filtrage.

Le prochain chapitre sera quant   lui consacr    la pr sentation du Filtre de Kalman et de son algorithme, ainsi que le m canisme de la m thode de ce dernier.

CHAPITRE II : METHODE DU FILTRE DE KALMAN

Introduction

D'une façon générale, la fonction de filtrage consiste à estimer une information (signal) utile qui est polluée par un bruit. Alors que le filtrage fréquentiel suppose qu'il existe une séparation entre les réponses fréquentielles du signal utile et du bruit et consiste à trouver une fonction de transfert satisfaisant un gabarit sur le gain de sa réponse fréquentielle.

Le filtre de Kalman est un filtre récursif, qui utilise une approche temporelle. Il permet d'estimer l'état (cet état correspond à l'information utile) d'un système dynamique à partir d'une information, à priori, sur l'évolution de cet état (modèle) et de mesure réelles.

L'utilisation de ce filtre est particulièrement recommandée lorsque l'imprécision sur les données traitées peut être interprétée comme un bruit aléatoire, dont la distribution est connue à priori. (ALAZARD, 2006)

II.1 GENERALITE SUR LE FILTRE DE KALMAN

II.1.1 Historique du Filtre de KALMAN

Rudolf Kalman, né en 1930, poursuit des études scientifiques dans les domaines de l'électronique et de l'automatique et soutint une thèse en 1957 sur le comportement des solutions des équations récurrentes du second degré. A partir de 1955, il travailla dans différentes entreprises en tant qu'automaticien. En 1958, alors qu'il travaillait au Research Institute for Advanced Studies à Blatmore, il eut l'idée originale d'employer le formalisme des équations d'état au filtre de Wiener, ce qui lui permit de construire l'estimateur que nous connaissons aujourd'hui. (BELMAHDI, 2015)

La solution du problème d'estimation a été développée par Kalman en 1960. Les algorithmes associés à ce problème d'estimation d'état présentent un caractère séquentiel, sont linéaires et correspondent à une estimation minimisant la variance de l'erreur quadratique : ils sont appelés filtre ou estimateur de « WINER-Kalman » ou de « Bucy-Kalman » ou plus simplement, filtre de « Kalman ». (BELMAHDI, 2015)

Le filtre de Kalman et ses nombreuses dérivées ont été appliqués avec succès dans le domaine aérospatial depuis le début des années soixante à de nombreux problèmes : navigation inertielle, poursuite radar et sonar, détermination des trajectoires et d'orbites. De plus, ces algorithmes font intervenir la notion d'équation aux différences et de systèmes récurrents qui sont bien adaptée au traitement sur calculateur numérique. (DAHMANI, 2012)

L'application du filtre de Kalman (estimateur linéaire) pour l'estimation de la position d'un objet nécessite, assez souvent, la linéarisation des équations d'un système non linéaire. Cette étape supplémentaire conduit à définir ce que l'on appelle, aujourd'hui, le filtre de Kalman étendu. (DAHMANI, 2012)

II.1.2 Domaine d'utilisation du Filtre de Kalman

Le filtre de Kalman est utilisé dans une large gamme de domaines technologiques (radar, vision électronique, communication ...). C'est un thème majeur de l'automatique et du traitement du signal. Un exemple d'utilisation peut être la mise à disposition, en continu, d'informations telles que la position ou la vitesse d'un objet à partir d'une série d'observations relative à sa position, incluant éventuellement des erreurs de mesures. Par exemple, pour le cas des radars où l'on désire suivre une cible, des données sur sa position, sa vitesse et son accélération sont mesurées à chaque instant mais avec énormément de perturbations dues au bruit ou aux erreurs de mesure. Le filtre de Kalman fait appel à la dynamique de la cible qui définit son évolution dans le temps pour obtenir de meilleures données, éliminant ainsi l'effet du bruit. Ces données peuvent être calculées pour l'instant présent (filtrage), dans le passé (lissage), ou sur un horizon futur (prédiction).

Le filtrage de Kalman est aussi de plus en plus utilisé en dehors du domaine du traitement du signal, par exemple en météorologie (environnement) et en océanographie, pour l'assimilation de données dans un modèle numérique, en finance ... etc.

II.1.3 Définition du Filtre de Kalman

Le FK peut être défini d'une manière simple comme étant un algorithme optimal et récursif de traitement des données qui permet d'estimer l'état courant d'un système en se basant sur l'estimation des mesures précédentes et actuelles, c'est aussi une approche statistique, d'assimilation de données, dont le principe est de corriger la trajectoire du modèle en combinant les observations avec l'information fournie par le modèle, de façon à minimiser l'erreur entre l'état vrai et l'état filtré. Cette méthode utilise une prédiction qui s'appuie sur un modèle déterministe et un recalage qui s'appuie sur l'innovation (différence entre la mesure et la sortie prédite).

II.1.4 Principe du filtre de Kalman

Le filtre de Kalman est un estimateur récursif. Cela signifie que pour estimer l'état courant d'un système, l'estimation de l'état précédent et des mesures actuelles sont nécessaires. Le principe d'un tel filtre repose sur un modèle à variables d'état basé sur deux équations :

- L'équation récursive de modélisation du processus aléatoire x_k (vecteur d'état) discret, dite équation d'état ou équation de transition.
- L'équation d'observation (ou équation de mesure) du processus y_k (vecteur de mesure).

L'objectif est de déterminer une estimation récursive des x_k à partir des mesures y_k et du modèle de génération du processus.

II.1.5 Les limites du Filtre de Kalman

Le filtrage de Kalman est une méthode d'estimation et de prédiction puissante prenant en compte les modélisations du système. Néanmoins, ce filtre n'est pas forcément l'outil appliqué dans tous les cas. En effet, le système doit être modélisé assez précisément afin de désigner un filtre efficace. Le problème provient du fait que certains systèmes sont difficilement modélisables et encore moins linéaires. Dans le cas où la modélisation est trop approximative, le filtre n'est pas assez performant et l'erreur des estimations ne convergera pas assez rapidement provoquant alors un temps de calcul trop important, et donc pour pallier aux problèmes des systèmes non linéaires un filtre de Kalman étendu a été développé et permet de prendre en compte une modélisation non linéaire ; Néanmoins cette technique a aussi ses défauts.

Une autre limite de ce filtre est qu'il permet de prendre en compte uniquement un modèle de bruit Gaussien. Cette restriction limite l'utilisation du filtre de Kalman car dans certains cas un autre type de bruit est requis comme les bruits de poisson par exemple.

Le filtre de Kalman est donc une méthode d'estimation intéressante, mais qui n'est utilisable que lorsque le système peut être décrit précisément. S'il est impossible de trouver une modélisation correcte du système, il est préférable de se tourner vers d'autres méthodes.

II.2 L'estimateur optimal du filtre de Kalman

II.2.1 L'estimateur

L'estimation est une branche des mathématiques statistiques qui permet, à partir de mesures effectuées sur un système, d'estimer la valeur de différents paramètres de ce système.

II.2.2 Les estimateurs déterministes

Tout d'abord, un point important dans la philosophie des estimateurs déterministes est que le vecteur à estimer est justement déterministe. C'est-à-dire que l'on suppose que ce vecteur possède des valeurs non aléatoires, même si celles-ci sont inaccessibles. Ce que l'on cherche à trouver, c'est un estimateur qui estime au mieux ce vecteur.

Pour ça, il faut que l'estimateur soit sans biais et de variance minimale (pour minimiser l'erreur d'estimation). (BELMAHDI, 2015)

La qualité d'un estimateur, c'est-à-dire son aptitude à déterminer la grandeur cherchée de la manière la plus précise possible, est caractérisée par un certain nombre de paramètres qui sont les suivants :

A. Biais d'un estimateur

En statistique un biais est une démarche ou un procédé qui engendre des erreurs dans les résultats d'une étude. Un estimateur sans biais signifie que l'espérance de l'estimation doit être égale au vecteur à estimer. En gros, ça veut donc dire que l'estimateur, en moyenne, ne commet pas d'erreur. Autrement dit, si l'on fait la moyenne des estimations, on trouve la valeur réelle du paramètre. Bien sûr, le biais n'est pas calculable en réalité, car on ignore la valeur réelle du paramètre ou du vecteur à estimer. Par contre, il est très utile pour caractériser nos estimateurs. Le biais doit être égal à 0 pour avoir un estimateur parfait (dans ce cas, l'estimateur est dit donc sans biais). (BELMAHDI, 2015)

$$B(\hat{\theta}) = E(\hat{\theta}) - \theta \quad \boxed{\text{II.1}}$$

B. La variance

La variance est une grandeur statistique permettant de caractériser la dispersion des échantillons d'une variable aléatoire, c'est-à-dire de quelle manière ces échantillons se répartissent autour de la valeur moyenne, elle peut être assimilée à l'erreur. Un estimateur est dit à variance minimale lorsque l'on a réussi à minimiser les variations entre le vecteur d'état et l'estimation. On a donc minimisé l'erreur. Il faut bien faire attention, car ces deux notions sont bien distinctes. Un estimateur peut très bien avoir un biais nul, mais si sa variance est forte, cet estimateur ne sera pas performant. (BELMAHDI, 2015)

$$\text{Var}(\hat{\theta}) = E(\hat{\theta} - E(\hat{\theta}))^2 \quad \boxed{\text{II.2}}$$

II.2.3 Estimateur optimale

Dans la littérature, on trouvera qu'un estimateur dont le biais est nul aura une variance toujours supérieure ou égale à ce que l'on appelle la "**borne de Cramer-Rao**". Finalement, c'est assez logique. On ne peut pas obtenir une variance nulle et donc, on ne peut pas connaître exactement la valeur du vecteur d'état. On aura toujours une incertitude égale à la borne de Cramer-Rao pour l'estimateur optimal. Si la variance est nulle, alors ce n'est plus un problème d'estimation, car ça voudrait dire que l'on connaît la valeur à estimer dès le départ. Il se trouve que dans certains cas particuliers, il existe un estimateur sans biais à variance minimale calculable. Mais pour cela, il faut que le système respecte quelques contraintes :

- Le bruit doit être Gaussien
- Le système doit être linéaire

Cet estimateur est appelé "**Estimateur sans biais à variance minimale linéaire gaussien**", ou encore "**Estimateur optimal linéaire gaussien**" ou tout simplement "**Méthode des moindres carrés**".

II.3 MECANISME DE LA METHODE DE FILTRE DE KALMAN

II.3.1 Variables et équations d'états

Nous considérons un système dont l'état est défini par un ensemble (fini) de variables d'état $X_1(t)$, $X_2(t)$, ..., $X_n(t)$ obéissant à un système d'équations de la forme suivante :

$$\frac{dX_1}{dt} = A_{11}X_1 + A_{12}X_2 + \dots + A_{1n}X_n + B_1$$

$$\frac{dX_2}{dt} = A_{21}X_1 + A_{22}X_2 + \dots + A_{2n}X_n + B_2$$

$$\frac{dX_n}{dt} = A_{n1}X_1 + A_{n2}X_2 + \dots + A_{nn}X_n + B_n$$

II.3

Nous admettons que la variable indépendante t représente le temps, bien qu'il n'y ait là aucune limitation essentielle. Les coefficients A_{ij} peuvent être constants, ou encore être fonction régulière de t .

Les seconds membres du système d'équations sont des formes linéaires des variables d'état X_i : on dit que ce système est linéaire.

Les variables d'état X_i sont fréquemment assimilées aux composantes d'un vecteur à n dimensions, appelé vecteur d'état.

Les coefficients A_{ij} peuvent être constants ; le système est alors dit à coefficients constants. Ces coefficients peuvent aussi être fonction de la variable indépendante t ; le système est alors dit « à coefficients variables ».

Lorsque les B_i sont nuls, le système est dit homogène (il est en effet de degré 1 par rapport à l'ensemble des variables d'état).

Forme matricielle des équations d'état

Il est très commode d'écrire le système précédent sous forme matricielle (Sorenson,1985). Dans ce but, on considère :

- Une matrice colonne définissant le vecteur d'état

$$\mathbf{X}(t) = \begin{bmatrix} X_1(t) \\ X_2(t) \\ \vdots \\ X_n(t) \end{bmatrix} \quad \boxed{\text{II.4}}$$

- Une matrice carrée définissant les coefficients du système

$$\mathbf{A}(t) = |A_{ij}(t)| \quad \boxed{\text{II.5}}$$

- Une matrice colonne définissant l'excitation ou la commande du système

$$\mathbf{B}(t) = \begin{bmatrix} B_1(t) \\ B_2(t) \\ \vdots \\ B_n(t) \end{bmatrix} \quad \boxed{\text{II.6}}$$

Dans ces conditions, le système s'écrit simplement

$$\dot{\mathbf{X}}(t) = \frac{d\mathbf{X}(t)}{dt} = \mathbf{A}(t)\mathbf{X}(t) + \mathbf{B}(t) \quad \boxed{\text{II.7}}$$

Le filtrage de Kalman comporte deux phases distinctes : Prédiction et mise à jour. La phase de prédiction utilise l'état estimé de l'instant précédent pour produire une estimation courante. Dans l'étape de mise à jour, les observations de l'instant courant sont utilisées pour corriger l'état prédit dans le but d'obtenir une estimation plus précise.

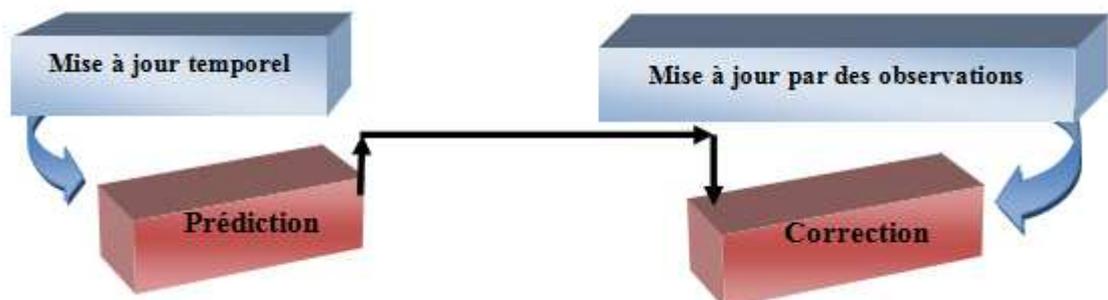


Figure N°II.1 : Schéma explicatif du mécanisme du filtre de Kalman

- **La phase de prédiction**

A l'instant k , on dispose d'une estimation appelée estimation à priori qui est fondée sur la connaissance du processus jusqu'à l'instant précédent $k-1$.

Si on notera $\hat{x}_{k/k-1}$ l'estimation d'état à priori, donc l'erreur à priori est donné par :

$$\mathbf{e}_{k/k-1} = \mathbf{x}_k - \hat{\mathbf{x}}_{k/k-1} \quad \boxed{\text{II.8}}$$

Les équations de la phase de prédiction sont :

$$\begin{cases} \hat{\mathbf{x}}_{k/k-1} = \mathbf{A}_{k-1}\hat{\mathbf{x}}_{k-1/k-1} + \mathbf{B}_{k-1}\mathbf{u}_{k-1} & \boxed{\text{(I)}} \\ \mathbf{P}_{k/k-1} = \mathbf{A}_{k-1}\mathbf{P}_{k-1/k-1}\mathbf{A}_{k-1}^T + \mathbf{Q}_{k-1} & \boxed{\text{(II)}} \end{cases}$$

- **La phase correction :**

Nous allons, maintenant, utiliser la y_k pour corriger l'estimation à priori $\hat{x}_{k/k-1}$ et obtenir l'estimation à posteriori $\hat{x}_{k/k}$.

L'erreur à posteriori :

$$\mathbf{e}_{k/k} = \mathbf{x}_k - \hat{\mathbf{x}}_{k/k} \quad \boxed{\text{II.9}}$$

Les équations de la phase de correction sont :

$$\begin{cases} \hat{\mathbf{x}}_{k/k} = \hat{\mathbf{x}}_{k/k-1} + \mathbf{K}_k(\mathbf{y}_k - \mathbf{H}_k\hat{\mathbf{x}}_{k/k-1}) & \boxed{\text{(III)}} \\ \mathbf{P}_{k/k} = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_{k/k-1} & \boxed{\text{(IV)}} \end{cases}$$

Avec :

$$\mathbf{K}_k = \mathbf{P}_{k/k-1}\mathbf{H}_k^T(\mathbf{H}_k\mathbf{P}_{k/k-1}\mathbf{H}_k^T + \mathbf{R}_k)^{-1} \quad \boxed{\text{(V)}}$$

\mathbf{K}_k : Gain de Kalman optimal, au sens de l'erreur quadratique moyenne, c'est-à-dire qui entraîne une erreur à posteriori minimale, est appelé gain de Kalman.

\mathbf{I} : est la matrice identité de même taille que $\mathbf{P}_{k/k-1}$.

II.4 PRESENTATION DE L'ALGORITHME RECURSIF DU FK DISCRET

II.4.1 Les différentes étapes d'évaluation des paramètres du système de FK

Pour obtenir l'état optimal du système, on doit combiner les observations y_k avec l'information fournie par le modèle x_k .

L'algorithme suivant indique les différentes étapes qui permettent d'obtenir un état estimé à chaque instant. (Harkat, 2014)

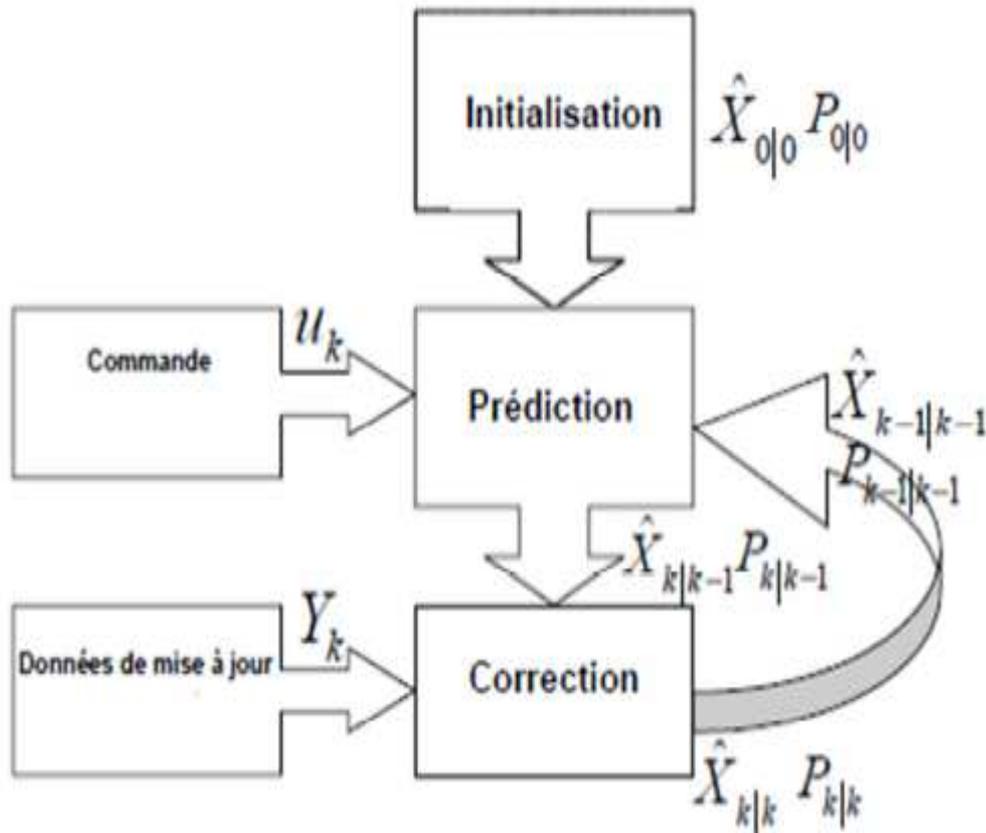


Figure N°II.2 : Algorithme du Filtre de Kalman

Les différentes étapes du calcul sont les suivantes :

1) Initialisation

Cette étape consiste à calculer les valeurs initiales des variables : vecteur d'état \hat{x}_0 , matrice de variances-covariances de l'erreur d'estimation P_0 , la matrice de variances-covariances du bruit de processus Q_0 et pour finir la matrice de variances-covariances du bruit de mesure R_0 .

$$\hat{\mathbf{x}}_0 = E[\mathbf{x}_0] \quad \boxed{\text{II.10}}$$

$$\mathbf{P}_0 = E[(\mathbf{x}_0 - \hat{\mathbf{x}}_0)(\mathbf{x}_0 - \hat{\mathbf{x}}_0)^T] \quad \boxed{\text{II.11}}$$

$$\mathbf{Q}_0 = E[(\mathbf{w} - \bar{\mathbf{w}})(\mathbf{w} - \bar{\mathbf{w}})^T] \quad \boxed{\text{II.12}}$$

$$\mathbf{R}_0 = E[(\mathbf{v} - \bar{\mathbf{v}})(\mathbf{v} - \bar{\mathbf{v}})^T] \quad \boxed{\text{II.13}}$$

2) Calcul de $\hat{\mathbf{x}}_{k/k-1}$, $\mathbf{P}_{k/k-1}$ à partir des mesures disponible à l'instant k-1 :

$$\hat{\mathbf{x}}_{k/k-1} = \mathbf{A}_{k-1}\hat{\mathbf{x}}_{k-1/k-1} + \mathbf{B}_{k-1}\mathbf{u}_{k-1} \quad \boxed{\text{II.14}}$$

$$\mathbf{P}_{k/k-1} = \mathbf{A}_{k-1}\mathbf{P}_{k-1/k-1}\mathbf{A}_{k-1}^T + \mathbf{Q}_{k-1} \quad \boxed{\text{II.15}}$$

3) Calcul du gain de Kalman :

$$\mathbf{K}_k = \mathbf{P}_{k/k-1}\mathbf{H}_k^T(\mathbf{H}_k\mathbf{P}_{k/k-1}\mathbf{H}_k^T + \mathbf{R}_k)^{-1} \quad \boxed{\text{II.16}}$$

4) Mise à jour de l'estimation :

$$\hat{\mathbf{x}}_{k/k} = \hat{\mathbf{x}}_{k/k-1} + \mathbf{K}_k(\mathbf{y}_k - \mathbf{H}_k\hat{\mathbf{x}}_{k/k-1}) \quad \boxed{\text{II.17}}$$

5) Calcul de la matrice de covariance de l'erreur :

$$\mathbf{P}_{k/k} = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_{k/k-1} \quad \boxed{\text{II.18}}$$

6) Préparation du pas suivant :

$$\hat{\mathbf{x}}_{k+1/k} = \mathbf{A}_k\hat{\mathbf{x}}_{k/k} + \mathbf{B}_k\mathbf{u}_k \quad \boxed{\text{II.19}}$$

$$\mathbf{P}_{k+1/k} = \mathbf{A}_k\mathbf{P}_{k/k}\mathbf{A}_k^T + \mathbf{Q}_k \quad \boxed{\text{II.20}}$$

II.4.2 Stabilité du filtre vis-a-vis de la divergence

Quand nous avons parlé de l'optimalité du filtre, nous n'avons pas fait allusion à sa stabilité vis-à-vis de la divergence. En effet, un filtre optimal n'est pas automatiquement stable et sa stabilité doit être prouvée.

On dit qu'il y a divergence du filtre, quand l'estimation qu'il nous fournit est entachée d'erreurs qui deviennent de plus en plus importantes.

Le filtre devient alors instable et donc insatisfaisant (**Schlee et al., 1967**). Le critère de convergence est donné par la matrice de covariance P_k (plus précisément par sa trace) qui doit être définie positive et minimale à la fin de la dernière itération (**Grewal & Andrews, 1997**). Elle doit donc converger vers le zéro au fur et à mesure que l'on avance dans le calcul sans pour autant lui être inférieure.

Les problèmes de stabilité du filtre sont reliés à la connaissance exacte des matrices et se traduisent par le fait que l'algorithme donne l'état erroné qui peut bien fonctionner à travers quelques données, mais qui n'est plus optimal pour longtemps. Ces problèmes peuvent être dus au fait que le modèle a été utilisé d'une manière erronée pouvant être engendrée par une mauvaise estimation des paramètres, comme ça peut être le résultat d'une simplification du modèle lui-même qui gère la dynamique du système.

Il ya un certain nombre de méthodes qui peuvent intervenir pour éviter l'apparition de tels problèmes. La procédure la plus utilisée est celle qui consiste à modifier les équations du filtre de sorte que les mesures récentes puissent avoir une influence plus grande sur les estimations de ce dernier et plus les mesures sont lointaines dans le temps, moins devient leur influence. (**Boukharouba, 2009**)

Conclusion

Dans ce chapitre nous avons présenté un outil de filtrage de type stochastique, qui est le Filtre de Kalman simple utilisé dans le cas des systèmes linéaires. Nous avons aussi évoqué la stabilité de ce dernier vis-à-vis de la divergence.

L'application du filtre de Kalman à un problème particulier exige le passage par les étapes suivantes :

1. Choix de la variable d'état du système étudié et détermination du modèle mathématique.
2. Estimation des conditions initiales.
3. Calcul conformément aux équations récurrentes du filtre.
4. Vérification des hypothèses d'application par l'examen des résidus et critique des résultats.

Dans le prochain chapitre on va présenter notre zone d'étude et traiter nos données pluviométriques afin de pouvoir les utiliser par la suite.

CHAPITRE III : PRÉSENTATION DE LA RÉGION D'ÉTUDE ET TRAITEMENT DES DONNÉES

Introduction

Le Bassin Hydrographique Cheliff-Zahrez couvre une superficie d'environ 56 227 km² (soit plus de 22% de la superficie de l'Algérie du Nord).

Cette région hydrographique est découpée en trois grands sous bassins versants, le bassin du Cheliff est le plus grand du point de vue superficie en Algérie, il représente plus de 77 % de la superficie totale du bassin Cheliff-Zahrez. Ce même bassin, limité au Nord par les monts du Dahra et au sud par l'Atlas saharien, comprend au Nord, la vallée du Cheliff, au sud les hauts plateaux de Saïda à Tiaret, les plaines de Nahr Oussel et Ain Oussera et au centre le massif de l'Ouarensenis.

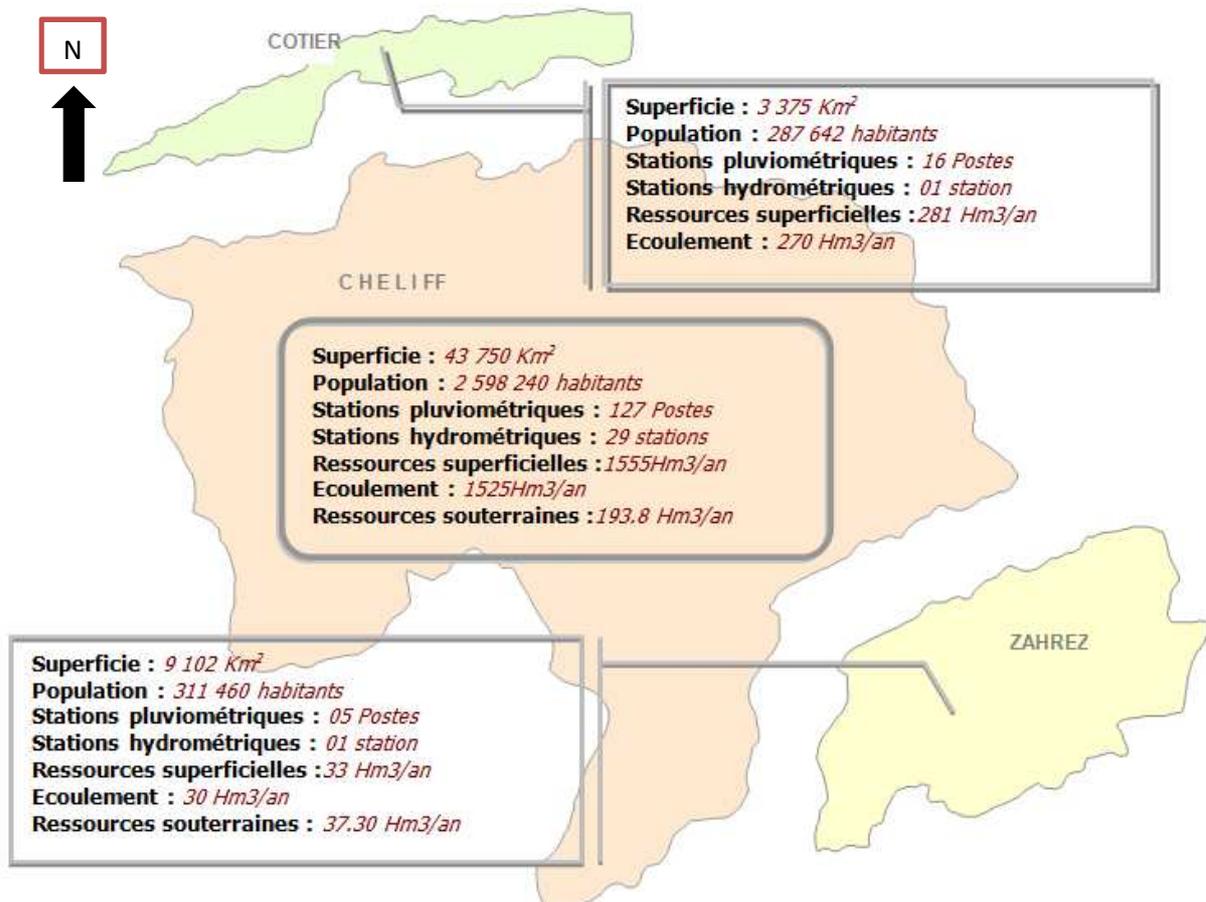


Figure N°III.1: Représentation des sous-bassins composants la région de Cheliff (ATLAS DU CHELIFF)

Vu l'étendue de la région Cheliff Zahrez, il est apparu nécessaire de définir, des sous régions pour mieux cerner l'adéquation ressources - besoins, Ce découpage est basé sur :

- ❖ Les bassins hydrographiques,
- ❖ Les infrastructures existantes ou projetées,
- ❖ Les limites administratives.

Cependant, le bassin du Cheliff qui est un bassin de superficie assez importante et à son tour découper en quatre sous-région :

Tableau N°III.1 - Les Sous-régions du Bassin Cheliff-Zahrez

Sous-Région	Bassin Hydrographique	Wilayas
Sous-Région 1	Haut Cheliff	Ain Defla, Média, Tissemsilt et Tipaza
Sous-Région 2	Moyen Cheliff	Chlef
Sous-Région 3	Mina, Bas Cheliff	Relizane, Tiaret, Mostaganem et Mascara
Sous-Région 4	Haut Plateaux	Djelfa, Laghouat et M'sila

Source : ATLAS DU CHELIFF



Figure N°III.2 : Représentation des sous-régions du bassin Cheliff-Zahrez (ATLAS DU CHELIFF)

III.1 Présentation de la région d'étude

III.1.1 Situation géographique du bassin versant

Cette région dont la superficie est près de 4154.30 Km², se situe dans le centre ouest du nord Algérie ; elle est limitée au nord par la mer Méditerranée, à l'est par les Wilayas de Tipaza et Ain Defla, à l'ouest par Mostaganem et Relizane, et au sud par Tissemsilt. Elle est composée de 35 communes, soit 132 agglomérations.



Figure N°III.3 : Situation géographique de la sous-région N°2 [1]

III.2 Le réseau pluviométrique

III.2.1 Acquisition des données

Les données nécessaires pour notre étude ont été recueillies au sein de l'Agence Nationale Des Ressources Hydriques (ANRH BLIDA) et cela pour 5 stations hydrométriques dont les caractéristiques sont les suivants :

Tableau N°III.2 - Coordonnées des stations hydrométriques

	Station	Code	Coordonnées			Période d'observation
			X (km)	Y (km)	Z (m)	
1	BENAIRIA	012205	378.15	338.4	320	1980 - 2011
2	OULED FARES	012201	368.1	326.65	116	1980 - 2011
3	MEDJADJA	012221	/	/	/	1986 - 2012
4	SOUK-EL-HAAD	012304	/	/	/	1984 - 2011
5	OULED BEN ABDELKADER	012308	/	/	/	1985 - 2013

Source : ANRH Blida

III.2.2 Caractéristiques du réseau pluviométrique

La région est équipée de 26 stations pluviométriques bien réparties dont trois (03) de longue durée, et sont gérées toutes par L'ANRH.

Les précipitations sont très irrégulières dans le temps et dans l'espace ; on distingue deux zones extrêmes ; DAHRA (côtier) et l'OUARSERNIS (Cheliff) avec une moyenne annuelle d'environ 600mm. La plaine incluse entre elles présente une moyenne annuelle d'environ 300mm. (ATLAS DU CHELIFF)

III.3 Traitement des données hydrologiques

III.3.1 Test d'Homogénéité

Avant de passer à l'ajustement de notre série il faut d'abord commencer par la vérification de l'homogénéité de cette dernière.

Pour vérifier l'homogénéité de notre série pluviométrique on va passer par le **Test de Médiane (Test de Mood)**.

-Principe du Test de la Médiane :

- ✓ On classe l'échantillon par ordre décroissant.
- ✓ On détermine la Médiane Me.
- ✓ On remplace les valeurs de la série non classée par :
 - Un signe (-) si : $X_i < Me$.
 - Un signe (+) si : $X_i > Me$.
- ✓ On calcule les valeurs de Ts et Ns tel que :

Ns : Nombre total de séries de + ou de - ;

Ts : Taille de la plus grande série de + ou de - ;

$$Ns > \frac{1}{2} (n + 1 - u_{1-\alpha/2} \sqrt{n + 1}) \dots \text{(III.1)}$$

$$Ts < 3.3(\log n + 1) \dots \text{(III.2)}$$

Le résultat du test d'homogénéité pour les cinq stations est représenté dans le tableau suivant :

Tableau N°III.3 - Résultat de test d'homogénéité pour les cinq stations hydrométriques

Station	n	Médiane	Ns	Ts	Ns Cal	Ts Cal	Résultat
1	32	39,4	14	8	10,87	8,27	Homogène
2	32	28,76	15	4	10,87	8,27	Homogène
3	27	29,49	13	5	8,81	8,02	Homogène
4	28	27,48	11	8	9,22	8,08	Homogène
5	29	26,38	11	6	9,63	8,13	Homogène

Avec : n : nombre d'observation par station.

III.3.2 Ajustement de la série pluviométrique

Le but de cette étude est de faire un ajustement de la série des précipitations maximales journalières par la loi choisie et de déterminer une intensité de pluie de durée et de période de retour donnée.

Les différentes étapes de l'étude :

- ✓ On classe la série des précipitations par ordre croissant.
- ✓ On détermine la fréquence expérimentale.
- ✓ On détermine les caractéristiques empiriques de la série de données.
- ✓ On procède à l'ajustement graphique avec la loi choisie.
- ✓ On calcule le quantile et son intervalle de confiance.

Ce travail va être effectué avec le logiciel « Hyfran ».

Tableau N°III.4 - Caractéristiques empiriques de la série pluviométrique 1 (BENAIRIA)

Caractéristiques de la série	Valeurs
Nombre D'années	33
Minimum	234
Maximum	663
Moyenne	473
Ecart-type	122
Médiane	472
Coefficient de Variation (Cv)	0.257
Coefficient d'asymétrie (Cs)	-0.26
Coefficient d'aplatissement (Ck)	1.79
Exposant climatique	0.4

a. Ajustement de la série pluviométrique par la loi de Gumbel :

La formule de répartition selon laquelle s'exprime la loi de Gumbel et la suivante :

$$F(x) = e^{-e^{-y}} \dots \text{(III.3)}$$

Tel que : $y = a(x - x_0)$;

$1/a$: la pente de la droite de Gumbel ;

y : variable réduite de Gumbel ;

x : précipitation maximale journalière (mm);

x_0 : ordonnée à l'origine en (mm).

On peut écrire : $y = -\ln(-\ln(F(x))) \dots \text{(III.4)}$

Avant de procéder à l'ajustement il faut suivre les étapes suivantes :

- Classer les valeurs des précipitations par ordre croissant avec attribution d'un rang 1, 2, 3, n.
- Calculer pour chaque valeur de précipitation la fréquence expérimentale par la formule de HAZEN.

$$F(x) = \frac{m-0.5}{n} \dots \text{(III.5)}$$

Avec : m : Rang de précipitation

n : nombre d'observations

- Calculer la variable réduite de **GUMBEL** donnée par la formule suivante : $y = -\ln(-\ln(F(x)))$
- Calculer les paramètres d'ajustement $\square a$ » et « x_0 »
- Représenter graphiquement les couples (x_i, y_i) sur papier GUMBEL.

b. Ajustement de la série pluviométrique à la loi de Galton (Log normale)

Une variable aléatoire x a une distribution log normale lorsque $(y = \ln x)$ est normale. La loi de Galton résulte de la loi normale mais rendue dissymétrique par un changement de variables. Sa fonction de répartition est donnée par :

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}u^2} du \quad \dots \text{III.6}$$

$f(x)$: Fréquence au non dépassement.

La variable réduite est de la forme : $u = \frac{x_i - \bar{x}}{\sigma_x} \quad \dots \text{III.7}$

L'équation de la variable réduite présentée sous la forme $\overline{\text{Log}x} + u_{p\%} \sigma_{\text{Log}x}$ est l'équation d'une droite sur papier gauss-logarithmique, avec en abscisse l'échelle gaussienne et en ordonnée l'échelle logarithmique.

c. Résultats de l'ajustement et loi choisie :

Les résultats du test de Khi-2 sont représentés dans le tableau ci-dessous :

Tableau N°III.5 - Résultat du test de Khi-2 pour les cinq stations hydrométriques

Station	Loi	X ²	p-value	D.D.L	N de Classe	X ² Théo	Résultat
Station 1	Gumbel	12.5	0.0104	5	8	11.07	Non Adéquate
	Lognormale	11	0.0423	5	8	11.07	Adéquate
Station 2	Gumbel	4.5	0.3579	5	8	11.07	Adéquate
	Lognormale	3.5	0.6234	5	8	11.07	Adéquate
Station 3	Gumbel	3.85	0.2989	4	7	9.4877	Adéquate
	Lognormale	9.56	0.0486	4	7	9.4877	Non Adéquate
Station 4	Gumbel	3.00	0.4046	4	7	9.4877	Adéquate
	Lognormale	2.00	0.7358	4	7	9.4877	Adéquate
Station 5	Gumbel	5.03	0.0775	4	7	9.4877	Adéquate
	Lognormale	8.90	0.0637	4	7	9.4877	Adéquate

D'après le tableau ci-dessous, le test de Khi-2 nous permet de déduire que :

- La série N°1 s'ajuste à la loi Lognormale.
- La série N°3 s'ajuste à la loi de Gumbel.

Pour le reste des séries on doit procéder à l'analyse visuelle des graphiques d'ajustement, et les conditions sont les suivantes :

- Pas d'intersection avec axe ordonnées en des points négatifs.
- Pas d'observation en dehors des intervalles de confiances.
- Le nuage de point expérimental colle bien à la courbe théorique.

Après l'analyse visuelle des graphes, les résultats finals de l'ajustement sont illustrés dans le tableau suivant :

Tableau N°III.6 - Résultat de l'ajustement

Station	Lois d'ajustement
Station N° 1	Loi Lognormale
Station N° 2	Loi Lognormale
Station N° 3	Loi de Gumbel
Station N° 4	Loi Lognormale
Station N° 5	Loi Lognormale

- Les graphes d'ajustement :

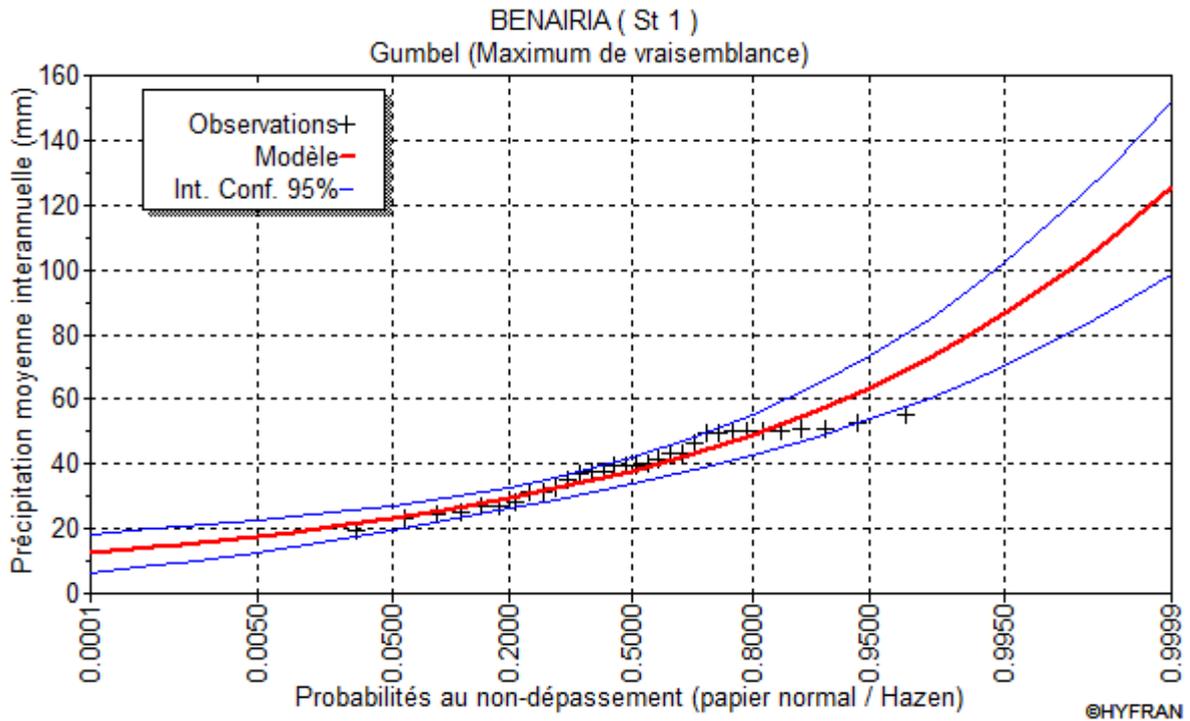


Figure N° III.4 : Ajustement de la série N°1 à la loi de Gumbel

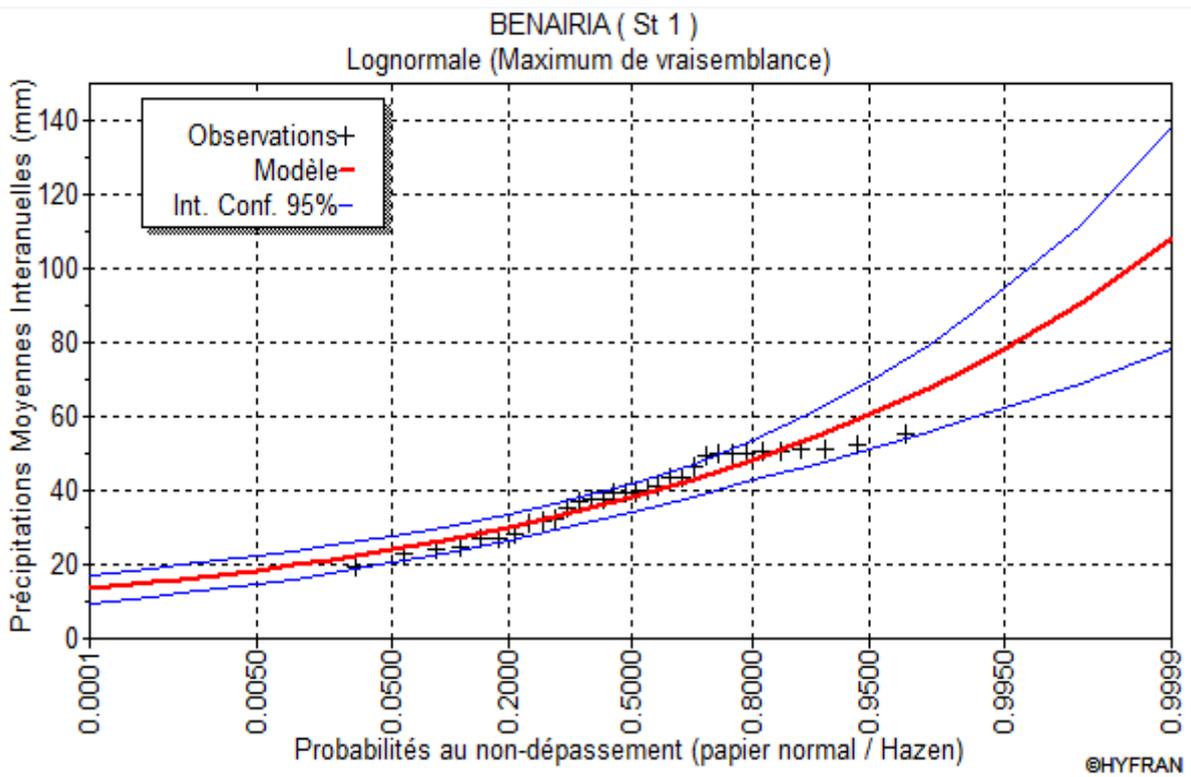


Figure N° III.5 : Ajustement de la série N°1 à la loi de Lognormale

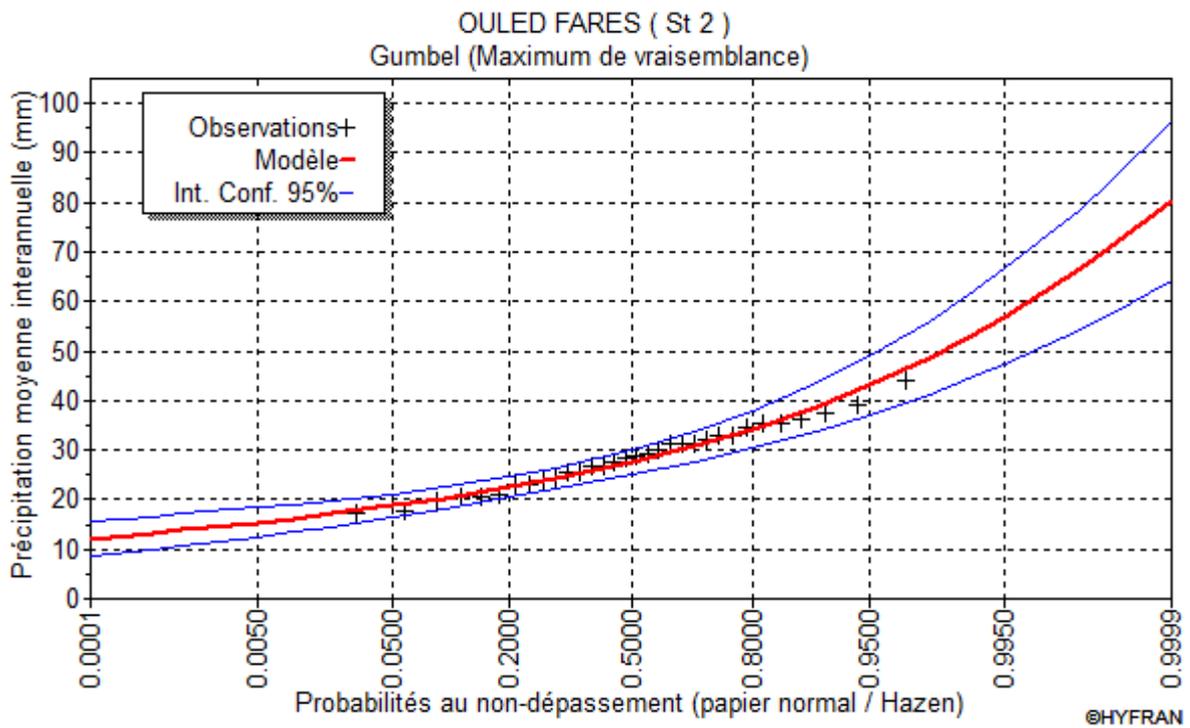


Figure N° III.6 : Ajustement de la série N°2 à la loi de Gumbel

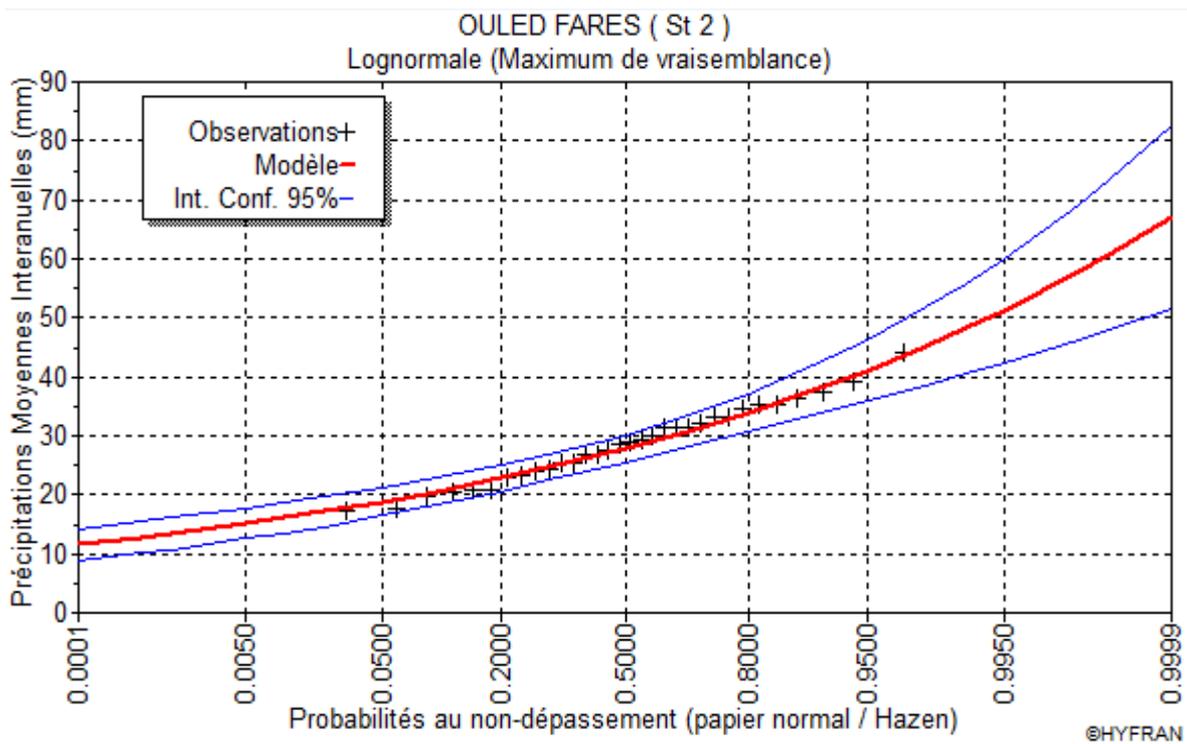


Figure N° III.7 : Ajustement de la série N°2 à la loi de Lognormale

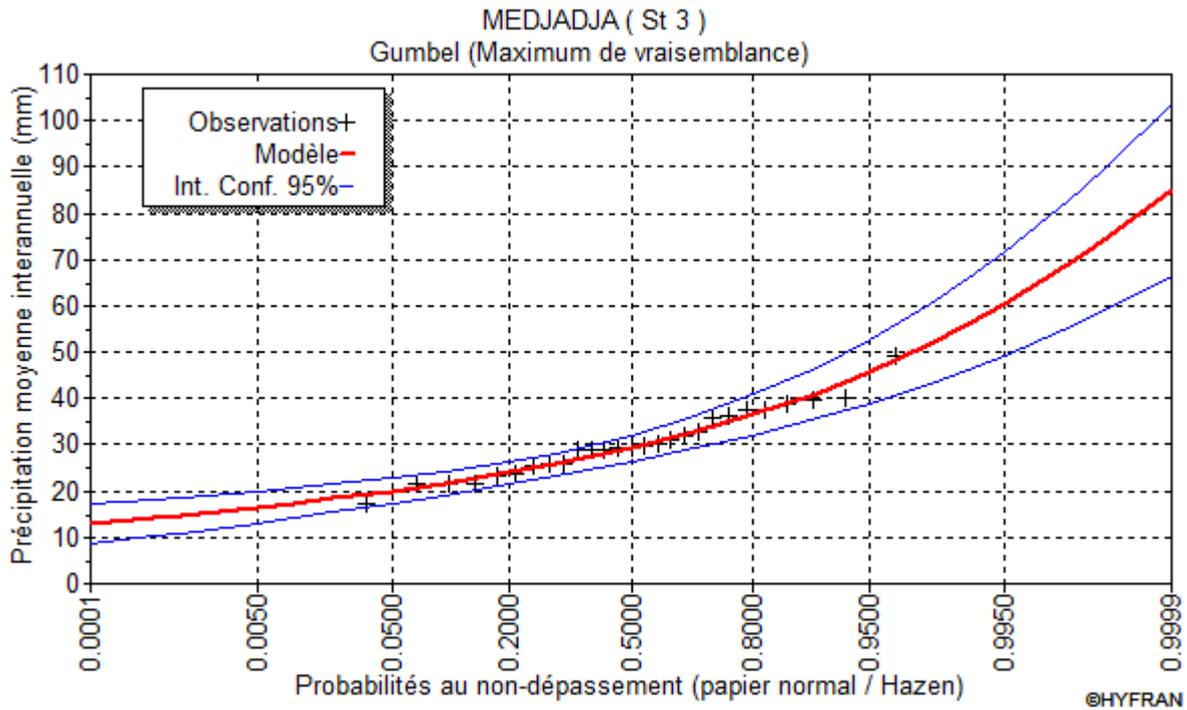


Figure N° III.8 : Ajustement de la série N°3 à la loi de Gumbel

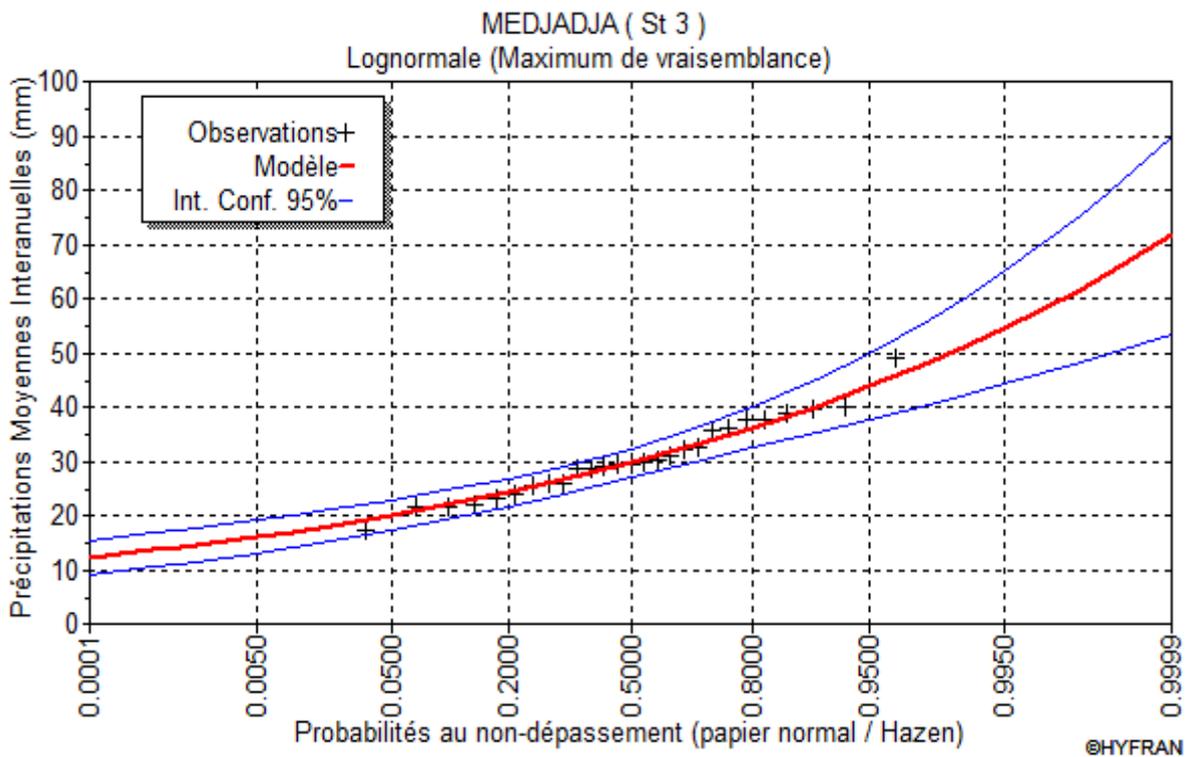


Figure N° III.9 : Ajustement de la série N°3 à la loi de Lognormale

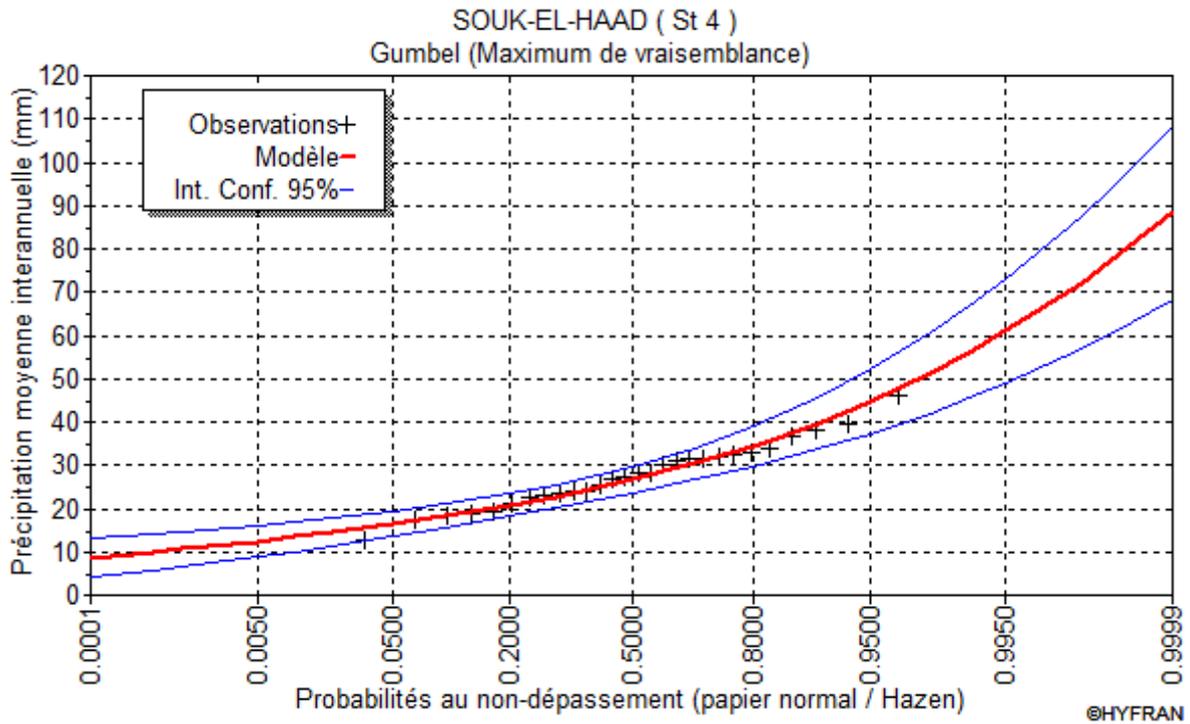


Figure N° III.10 : Ajustement de la série N°4 à la loi de Gumbel

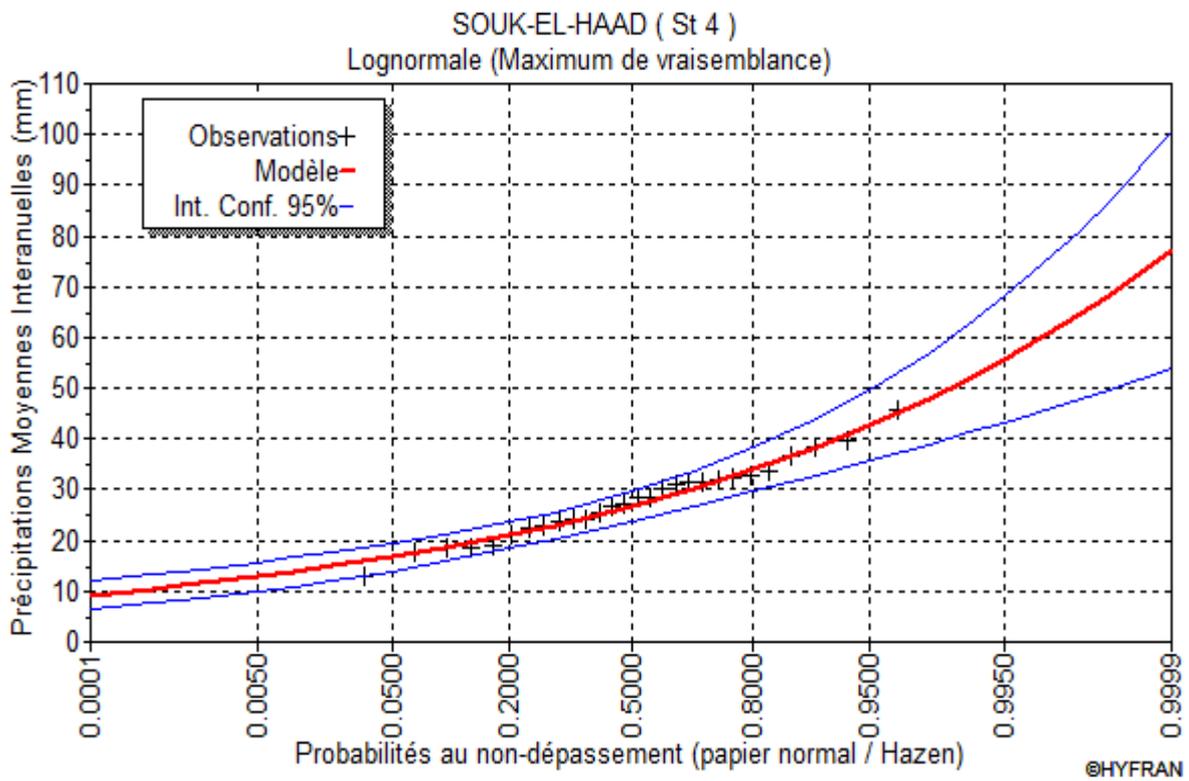


Figure N° III.11 : Ajustement de la série N°4 à la loi de Lognormale

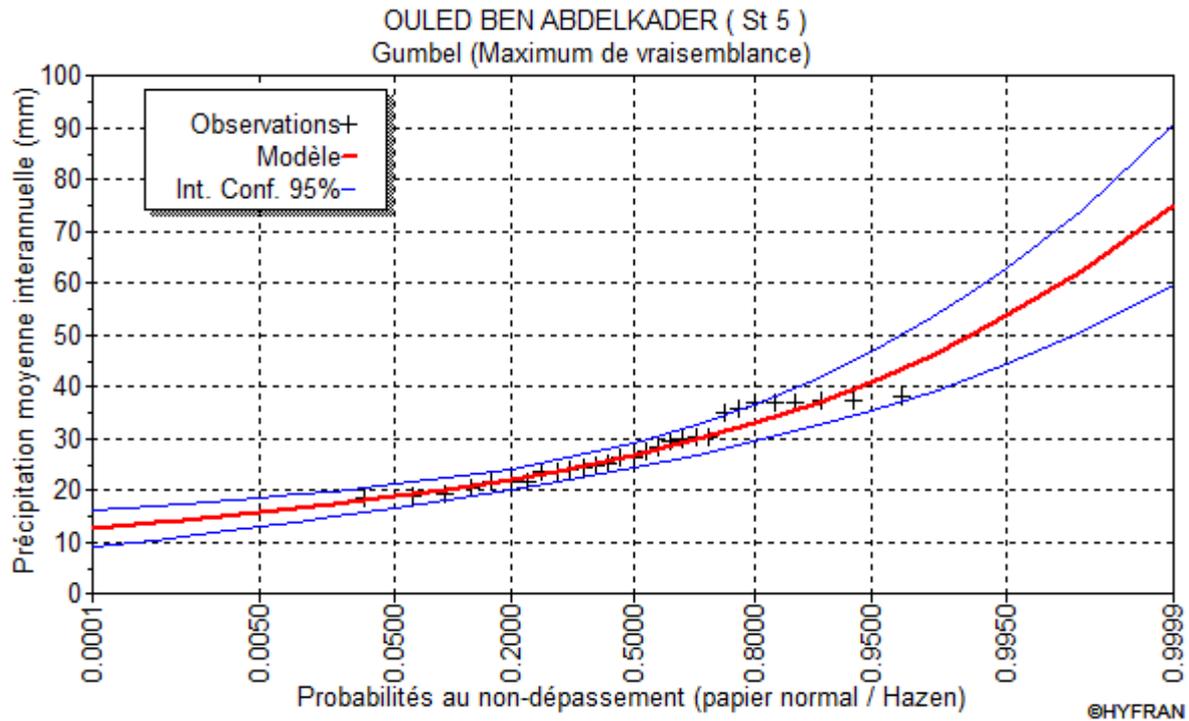


Figure N° III.12 : Ajustement de la série N°5 à la loi de Gumbel

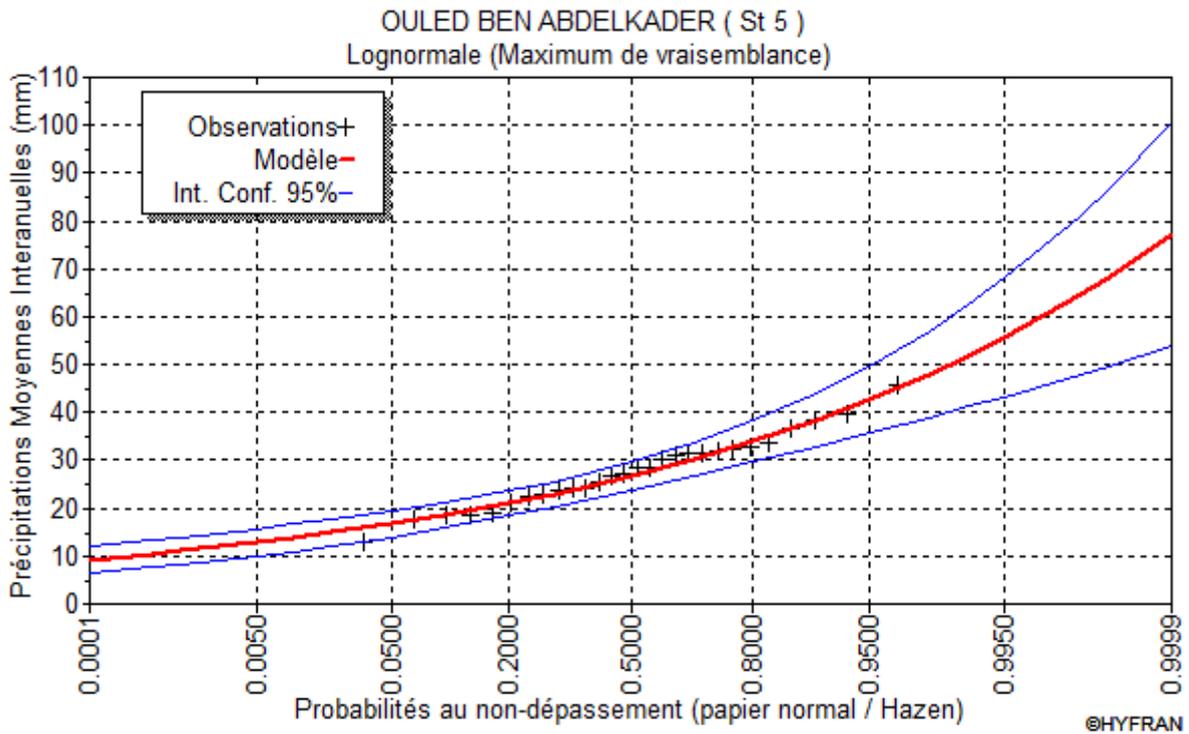


Figure N° III.13 : Ajustement de la série N°5 à la loi de Lognormale

Conclusion

Dans cette partie nous avons présenté notre zone du point de vue géographique et pluviométrique, nous avons aussi procédé à l'homogénéisation des données pluviométriques et leur ajustement.

Des séries chronologiques des précipitations annuelles et mensuelles couvrant une période d'observation commune de 11 ans ont été investies. Ces dernières sont observées à 5 Stations pluviométriques dans le bassin versant du Moyen Cheliff.

CHAPITRE IV : APPLICATION DU FILTRE DE KALMAN

Introduction

Par définition, la prédiction est le comportement futur probable d'un processus. La confiance qui lui est assignée dépend de la nature du processus étudié et de la qualité de l'adéquation du modèle choisi pour mieux le représenter (**Sorenson, 1985**).

Ces prédictions, si elles sont valables, forment la base pour les décisions concernant la gestion et le contrôle des ressources hydriques. Clairement, la fiabilité de telles prédictions gouvernera celle de ces décisions, et si en plus cette validité est étendue au temps réel, les résultats peuvent alors être interpolés ou extrapolés à n'importe quel temps puisque la nouvelle donnée est tout-à-fait valable.

Dans ce chapitre, notre attention sera focalisée sur les modèles linéaires à temps discrets pouvant être représentés par des équations de différence plutôt que des équations différentielles. Cette restriction nous paraît assez raisonnable, sachant que les prédictions en hydrologie se font à des temps discrets. L'objectif étant de faire des prédictions multi-sites, notre choix est orienté vers le filtre de Kalman discret.

IV.1 Données de la prédiction multi-site

La méthode de Kalman va être appliquée aux données mensuelles et annuelles des précipitations observées par l'Agence Nationale des Ressources Hydriques.

IV.2 Hypothèses

- Le système est linéaire.
- Les résidus sont des bruits blancs gaussiens.

IV.3 Formulation espace-état du modèle adopté

L'une des étapes les plus importantes dans l'application de la technique du Filtre de Kalman est la formulation des équations d'état et de mesure conformément à la structure d'un modèle espace état. La variable état du système est un vecteur donnant l'apport liquide annuel observé simultanément aux cinq stations hydrométriques considérées. Ainsi, pour $n=5$ les vecteurs d'état du système et de la mesure ont pour dimension (5×1) , alors que les matrices de covariances, de transition d'état et celle de mesure sont de dimension (5×5) . (**HARKAT, 2014**)

L'équation d'état et celle de mesure sont respectivement formulées de la manière suivante :

- L'équation d'état :

$$\begin{bmatrix} X_1 \\ X_2 \\ \cdot \\ \cdot \\ X_5 \end{bmatrix}_k = \begin{bmatrix} \Phi_{1,1} & \cdot & \cdot & \Phi_{1,5} \\ \Phi_{2,1} & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \Phi_{5,1} & \cdot & \cdot & \Phi_{5,5} \end{bmatrix}_{k/k-1} \begin{bmatrix} X_1 \\ X_2 \\ \cdot \\ \cdot \\ X_5 \end{bmatrix}_{k-1} + \begin{bmatrix} W_1 \\ W_2 \\ \cdot \\ \cdot \\ W_5 \end{bmatrix}_{k-1}$$

Avec :

- $\begin{bmatrix} X_1 \\ X_2 \\ \cdot \\ \cdot \\ X_5 \end{bmatrix}_k$: Le vecteur état au temps k,
- $\begin{bmatrix} \Phi_{1,1} & \cdot & \cdot & \Phi_{1,5} \\ \Phi_{2,1} & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \Phi_{5,1} & \cdot & \cdot & \Phi_{5,5} \end{bmatrix}_{k/k-1}$: matrice de transition d'état
- $\begin{bmatrix} W_1 \\ W_2 \\ \cdot \\ \cdot \\ W_5 \end{bmatrix}_{k-1}$: Vecteur des bruits du système
- Equation de mesure :

$$\begin{bmatrix} Z_1 \\ Z_2 \\ \cdot \\ \cdot \\ Z_5 \end{bmatrix}_{k+1} = \begin{bmatrix} H_{1,1} & \cdot & \cdot & H_{1,5} \\ H_{2,1} & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ H_{5,1} & \cdot & \cdot & H_{5,5} \end{bmatrix}_k \begin{bmatrix} X_1 \\ X_2 \\ \cdot \\ \cdot \\ X_5 \end{bmatrix}_k + \begin{bmatrix} V_1 \\ V_2 \\ \cdot \\ \cdot \\ V_5 \end{bmatrix}_k$$

Avec :

- $\begin{bmatrix} Z_1 \\ Z_2 \\ \cdot \\ \cdot \\ Z_5 \end{bmatrix}_k$: Vecteur de mesure au temps k
- $\begin{bmatrix} H_{1,1} & \cdot & \cdot & H_{1,5} \\ H_{2,1} & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ H_{5,1} & \cdot & \cdot & H_{5,5} \end{bmatrix}_k$: matrice de mesure au temps k

- $\begin{bmatrix} X_1 \\ X_2 \\ \cdot \\ \cdot \\ X_5 \end{bmatrix}_k$: Vecteur d'état au temps k

- $\begin{bmatrix} V_1 \\ V_1 \\ \cdot \\ \cdot \\ V_5 \end{bmatrix}_k$: vecteur des bruits de mesure au temps k

IV.4 Vecteur d'état initial et matrice de covariance d'erreur associée

En présence d'une information à priori, le vecteur d'état initial $X_{k/k-1}$, pour $k = 1$ est constitué par la moyenne des précipitations moyennes mensuelles et annuelles observées aux 5 différentes stations hydrométriques en question $(\bar{X}_1, \bar{X}_2, \dots, \bar{X}_5)$ comme suit :

$$\hat{X}_{1/0} = \begin{bmatrix} \bar{X}_1 \\ \bar{X}_2 \\ \cdot \\ \cdot \\ \bar{X}_5 \end{bmatrix}$$

Nous pouvons commencer avec des éléments ayant des valeurs assez élevées dans la diagonale principale pour spécifier la matrice de covariance d'erreur qui lui est associée $P_{k/k-1}$ et qui n'est pas exactement connue. L'algorithme bénéficiera donc d'une certaine flexibilité afin de s'ajuster aux valeurs sensibles dans un temps relativement court.

Dans notre travail, nous avons choisis la matrice suivante (5x5) :

$$P_{1/0} = \begin{bmatrix} \mathbf{1000} & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{1000} \end{bmatrix}$$

Ce choix va mener à une augmentation de la matrice de covariance $P_{k/k-1}$ et celle du gain du filtre K_k , permettant ainsi au filtre adaptatif de pondérer plus lourdement la nouvelle information. Ainsi, la trace de la matrice de covariance initiale est de 5 000. Cette valeur est sensé décroître jusqu'à ce qu'elle devienne presque nulle. Cette réduction est dûe à l'adaptation automatique de l'algorithme du filtre de Kalman à la réception de nouvelles

informations objectives. Autrement dit, la performance du filtre de Kalman peut être mesurée à travers l'estimation de la trace de la matrice de covariance.

IV.5 Matrice de covariance des bruits du système et celle des bruits des mesures

La matrice de covariance du bruit du système Q et celle du bruit de mesure R , toutes les deux (5x5) sont choisies de sortes que les mesures soient moins bruitées que la dynamique du système : (HARKAT, 2014)

$$Q = \begin{bmatrix} 100 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 100 \end{bmatrix}$$

$$R = \begin{bmatrix} 50 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 50 \end{bmatrix}$$

IV.6 Matrice de transition d'état et matrice de mesure

L'estimation de la matrice de transition d'état $\Phi_{k/k-1}$ représente l'une des plus grandes difficultés dans l'application du filtre de Kalman, même si la valeur de cette matrice n'affecte pas de manière considérable les résultats de ce dernier. Nous considérons alors pour l'estimation de la matrice de transition d'état les inter-corrélations entre les observations des cinq stations pluviométriques.

En ce qui concerne la matrice de mesure H_k nous avons opté pour la matrice identité (5x5) (HARKAT, 2014)

$$H_k = \begin{bmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{bmatrix}$$

IV.7 Les équations récurrentes du filtre de Kalman

IV.7.1 Matrice gain du filtre

La matrice gain du filtre est un indicateur sur la performance de notre algorithme, elle (5x5) peut être calculée pour une prédiction à un pas en utilisant l'équation suivante :

$$K_k = P_{k/k-1} H_k^T (H_k P_{k/k-1} H_k^T + R_k)^{-1}$$

Pour $k=1$ comme suit :

$$K_1 = \begin{bmatrix} \mathbf{1000} & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{1000} \end{bmatrix} \left[\begin{bmatrix} \mathbf{1000} & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{1000} \end{bmatrix} + \begin{bmatrix} \mathbf{50} & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{50} \end{bmatrix} \right]^{-1}$$

IV.7.2 Utilisation de la mesure pour actualiser l'estimation

C'est la seconde étape de calculs pour le même pas de temps $k=1$. L'estimation est corrigée moyennant la nouvelle mesure \mathbf{Z}_k afin d'améliorer l'estimation initiale $X_{1/0}$. Ceci est fait en faisant intervenir un terme de correction comprenant la différence $(\mathbf{Z}_k - \mathbf{H}_k \hat{\mathbf{X}}_{k/k-1})$ pondérée par le facteur \mathbf{K}_1 en concordance avec l'équation II.17

Et on obtient :

$$\hat{\mathbf{X}}_{1/1} = \begin{bmatrix} \bar{X}_1 \\ \bar{X}_2 \\ \vdots \\ \bar{X}_5 \end{bmatrix}_{1/0} + \begin{bmatrix} k_{1,1} & \dots & k_{1,5} \\ k_{2,1} & \dots & \cdot \\ \cdot & \dots & \cdot \\ k_{5,1} & \dots & k_{5,5} \end{bmatrix}_1 \left(\begin{bmatrix} \mathbf{Z}_1 \\ \mathbf{Z}_2 \\ \vdots \\ \mathbf{Z}_5 \end{bmatrix}_1 - \begin{bmatrix} \bar{X}_1 \\ \bar{X}_2 \\ \vdots \\ \bar{X}_5 \end{bmatrix}_{1/0} \right)$$

IV.7.3 Matrice de covariance d'erreur associée à l'estimation (actualisée) corrigée par la mesure

La matrice de covariance d'erreur associée à l'estimation corrigée par la mesure peut être obtenue comme suit :

$$P_{1/1} = \begin{bmatrix} \mathbf{1000} & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{1000} \end{bmatrix}_{1/0} - \begin{bmatrix} k_{1,1} & \dots & k_{1,5} \\ k_{2,1} & \dots & \cdot \\ \cdot & \dots & \cdot \\ k_{5,1} & \dots & k_{5,5} \end{bmatrix}_1 \begin{bmatrix} \mathbf{1000} & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{1000} \end{bmatrix}_{1/0}$$

IV.7.4 Prédiction un pas en avant de l'estimation (actualisée) corrigée par la mesure

L'estimation corrigée $\hat{X}_{1/1}$ est projetée un pas en avant tout en ignorant le terme W_{k-1} parce que sa moyenne est nulle et parce qu'il est non corrélé avec les W précédents. Ceci nous conduit à la prédiction un pas en avant de l'estimation corrigée $\hat{X}_{2/1}$ tel que :

$$\begin{bmatrix} \bar{X}_1 \\ \bar{X}_2 \\ \cdot \\ \cdot \\ \bar{X}_5 \end{bmatrix}_{2/1} = \begin{bmatrix} \phi_{1,1} & \cdot & \cdot & \phi_{1,5} \\ \phi_{2,1} & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \phi_{5,1} & \cdot & \cdot & \phi_{5,5} \end{bmatrix}_{2/1} \begin{bmatrix} \bar{X}_1 \\ \bar{X}_2 \\ \cdot \\ \cdot \\ \bar{X}_5 \end{bmatrix}_{1/1}$$

IV.7.5 Matrice de covariance de l'erreur de prédiction un pas en avant

Calcul de la matrice de covariance d'erreur de prédiction un pas en avant $P_{2/1}$ associée à $X_{2/1}$:

$$P_{2/1} = \begin{bmatrix} \phi_{1,1} & \cdot & \cdot & \phi_{1,5} \\ \phi_{2,1} & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \phi_{5,1} & \cdot & \cdot & \phi_{5,5} \end{bmatrix}_{2/1} \begin{bmatrix} p_{1,1} & \cdot & \cdot & p_{1,5} \\ p_{2,1} & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ p_{5,1} & \cdot & \cdot & p \end{bmatrix}_{1/1} \begin{bmatrix} \phi_{1,1} & \cdot & \cdot & \phi_{1,5} \\ \phi_{2,1} & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \phi_{5,1} & \cdot & \cdot & \phi_{5,5} \end{bmatrix}_{2/1}^T + \begin{bmatrix} \mathbf{100} & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{100} \end{bmatrix}$$

$P_{2/1}$ et $X_{2/1}$ sont considérés comme conditions initiales pour les prochaines itérations, et ainsi de suite pour le restant des itérations jusqu'à la fin du calcul.

IV.8 Résultats et interprétations

Pour développer et appliquer l'approche du modèle de FK à la modélisation et la prédiction multi-site des précipitations dans le bassin versant du Moyen Chélif, les observations enregistrées aux 5 stations pluviométriques, observées au pas de temps mensuel et annuel, sur une période de 11 ans (2000 - 2010) ont été étudiées.

L'exécution consécutive des équations de Kalman sur la période d'observation représente l'application de l'approche du FK à la modélisation et à la prédiction des précipitations annuelles et mensuelles (**Modélisation de FK sous Matlab**) (HARKAT Samra, 2014). A la fin des calculs, il s'avère que certaines des valeurs prédites sont sous-estimées et d'autres sont surestimées. La différence entre les mesures et leurs prédictions conduit à une série de résidus (innovations) pour chaque station.

IV.8.1 Prédictions multi-site

A. Pas Annuel

Les prédictions pour le pas annuel ont été obtenues en créant une matrice de 5 colonnes qui représentent le nombre de stations hydrométriques et 11 lignes qui représentent le nombre d'années. Les résultats des prédictions et des erreurs relatives vont être illustrés dans les graphes suivants dans une dimension spatiale ($P \text{ (mm)} = f(\text{St})$) et temporelle ($P \text{ (mm)} = f(\text{Année})$).

Gain du filtre de Kalman :

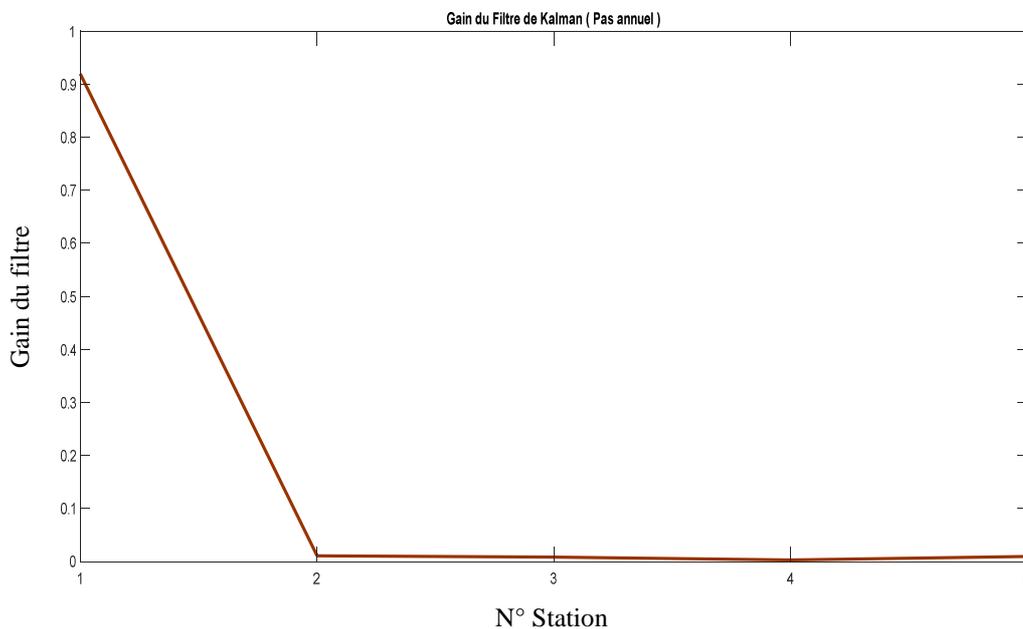
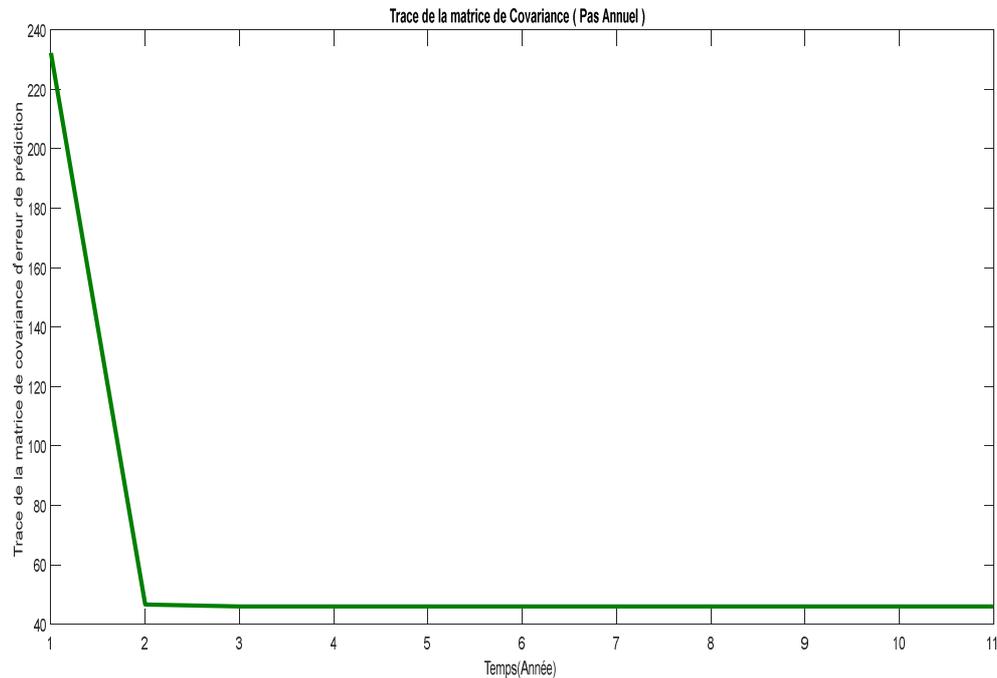


Figure N° IV.1 : Gain du filtre de Kalman (Pas annuel)

Initialement, la matrice gain du FK prend des valeurs importantes, car la mesure représente toute l'information objective et le rôle de la matrice gain est justement d'assurer que la mesure soit plus lourdement pondérée dans l'estimation des paramètres d'état, et la conséquence est une estimation automatiquement mauvaise ce qui explique des erreurs relativement importantes durant les premières itérations. Mais, avec le déroulement des calculs quand la confiance assignée à la précision des paramètres du modèle commence à augmenter, on remarque que les valeurs du gain de Kalman commencent à diminuer progressivement pour atteindre asymptotiquement une valeur proche de zéro (**Figure N° IV.1**), ce qui signifie la diminution de l'influence de la mesure dans la mise à jour de l'estimation des paramètres du modèle et les erreurs associées.

Trace de la matrice de Covariance :**Figure N° IV.2 : Trace de la matrice de covariance d'erreur (Pas annuel)**

La décroissance de la trace de la matrice de covariance de l'erreur de prédiction est un critère de convergence et d'optimalité du filtre. Comme on peut le remarquer dans la **Figure N° IV.2** la trace de cette matrice commence par des valeurs élevées au début des calculs, puis décroît rapidement pour converger vers une valeur positive stable, la plus proche possible de zéro. Cette convergence confirme l'adéquation du modèle ajusté au processus étudié et prouve l'efficacité et l'optimalité de l'algorithme.

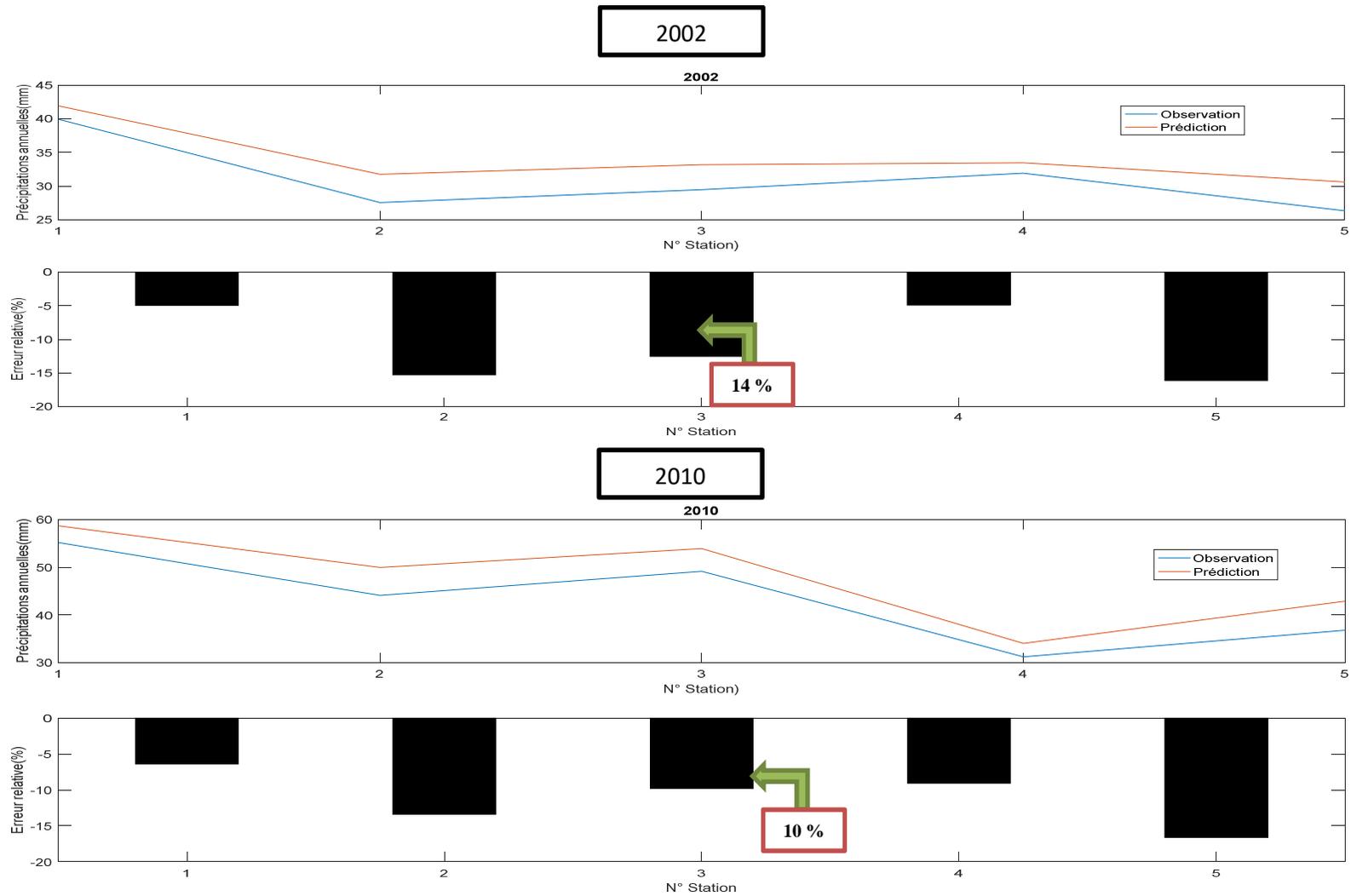


Figure N° IV.3 : Prédictions annuelles des précipitations aux 5 stations pluviométriques (exp : Année 2002 et 2010)

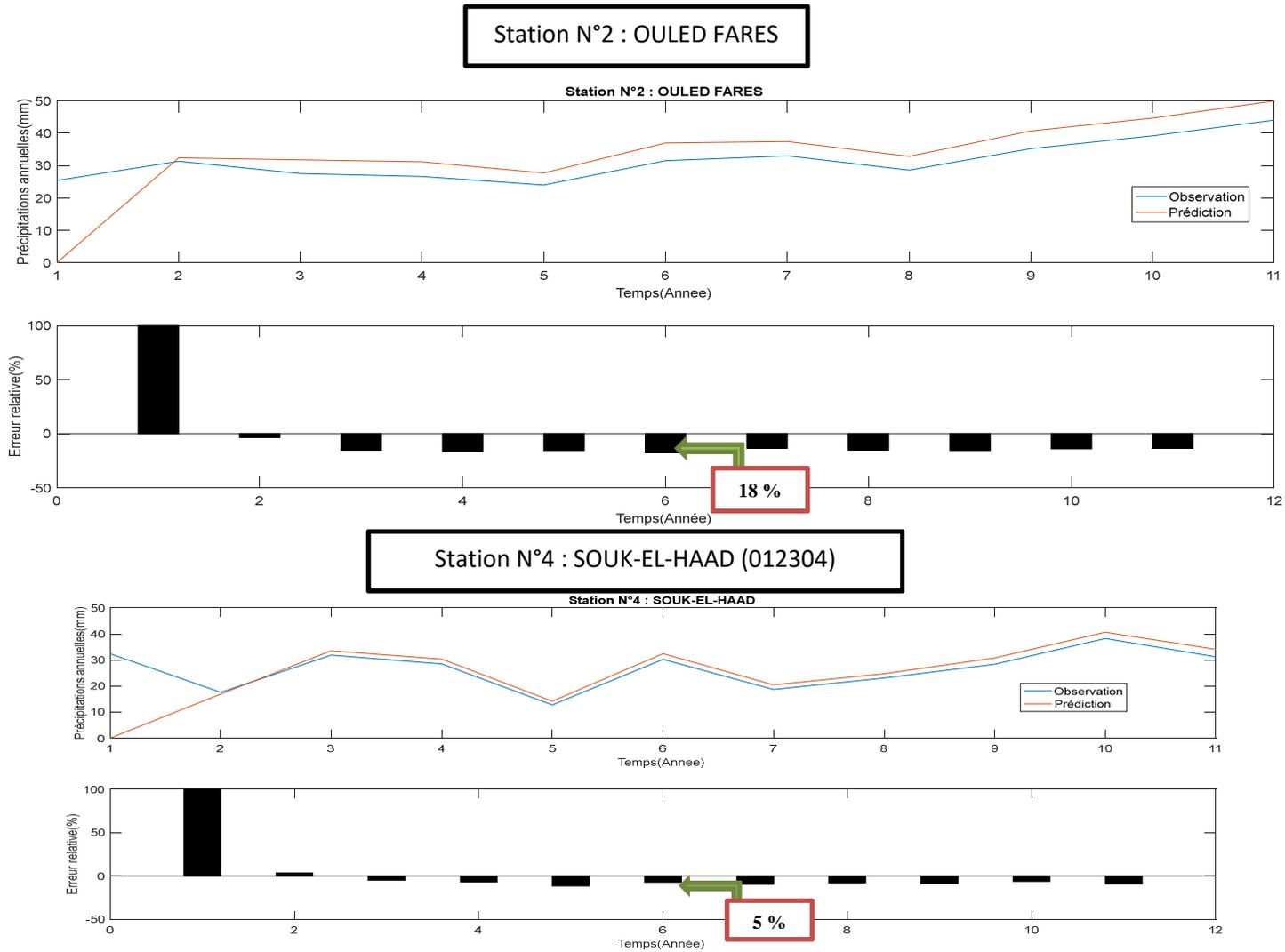


Figure N° IV.4 : Prédictions annuelles des précipitations en fonction des années (exp : Station N° 2 et Station N°4)

Interprétation :

La troisième figure représente les prédictions annuelles des précipitations en fonction des cinq stations hydrométriques (Spatial) avec l'erreur relative observée en chaque station et année, nous remarquons que l'erreur dans les premières années de calcul est importante, cet écart considérable entre l'observation et la prédiction est dû principalement à l'adaptation de l'algorithme du filtre qui fait que durant les premières itérations l'on assigne plus de confiance à la mesure en tant qu'information objective qu'au modèle ajusté. Alors que dans les dernières années de calcul, nous remarquons une baisse remarquable de l'erreur relative, chose qui prouve que l'algorithme s'est adapté aux données qu'il est entrain de traiter.

Même chose pour la quatrième figure ou l'on remarque que l'erreur est en baisse dans la station N°4 par rapport à celle dans la station N°2, et cela pour les mêmes raisons citées auparavant.

B. Pas Mensuel

Les prédictions pour le pas mensuel ont été obtenues en créant une matrice de 5 colonnes qui représentent le nombre de stations hydrométriques et 132 lignes qui représentent le nombre de mois. Les résultats des prédictions et des erreurs relatives vont être illustrés dans les graphes suivants dans une dimension spatiale ($P \text{ (mm)} = f(\text{St})$) et temporelle ($P \text{ (mm)} = f(\text{Mois})$).

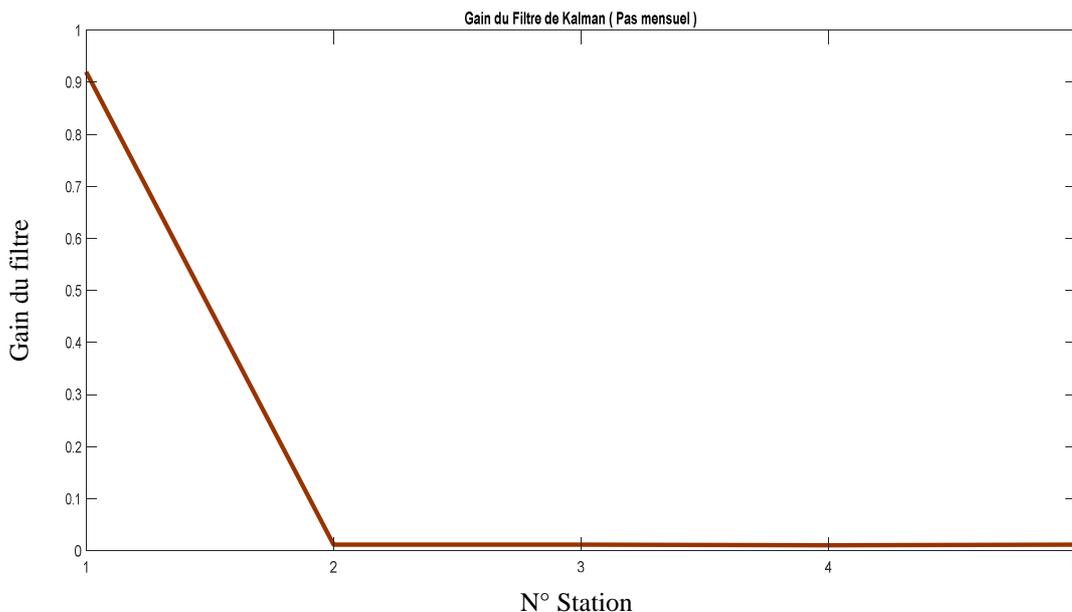
Gain du filtre de Kalman :

Figure N° IV.5 : Gain du filtre de Kalman (Pas mensuel)

Pour le pas mensuel, on remarque la même chose concernant le graphe du Gain du filtre de Kalman, ou ce dernier est en décroissance jusqu'à atteindre des valeurs très proche de zéro.

Trace de la matrice de Covariance :

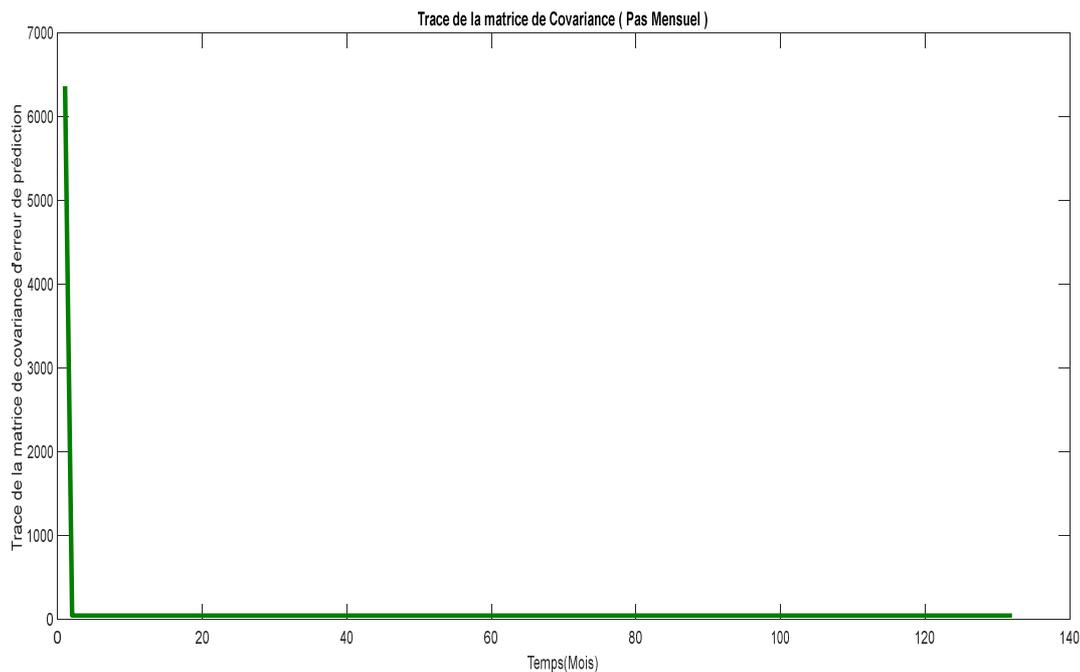


Figure N° IV.6 : Trace de la matrice de covariance d'erreur (Pas mensuel)

On remarque d'après la **Figure N° IV.6** que la trace de la matrice de covariance prend des valeurs importantes, et puis après un certain nombre d'itérations, elle décroît rapidement jusqu'à atteindre des valeurs qui se rapprochent de zéro à la fin des calculs, tout en restant positive.

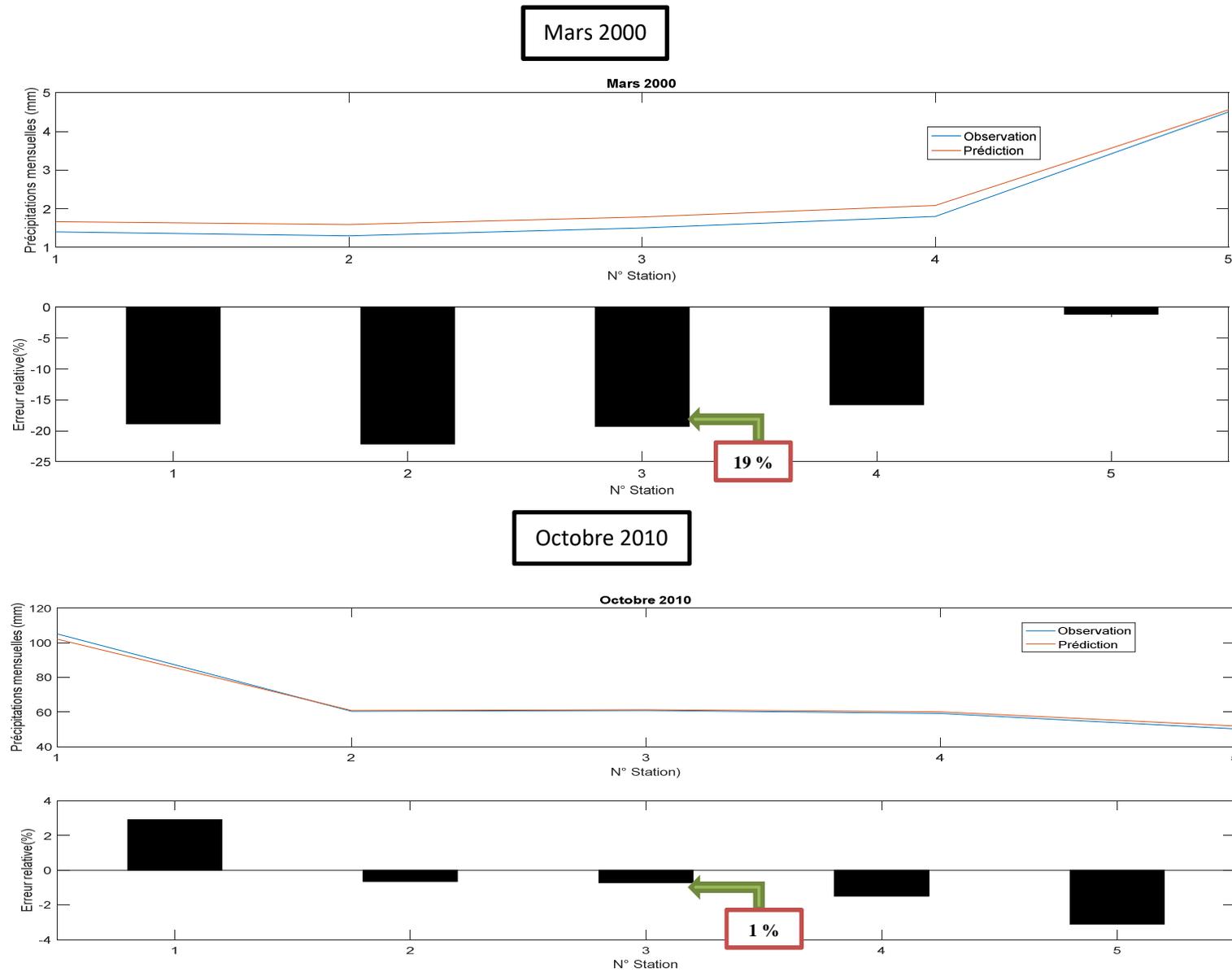


Figure N° IV.7 : Prédictions mensuelles des précipitations aux cinq stations pluviométriques (exp : Mars 2000 et Octobre 2010)

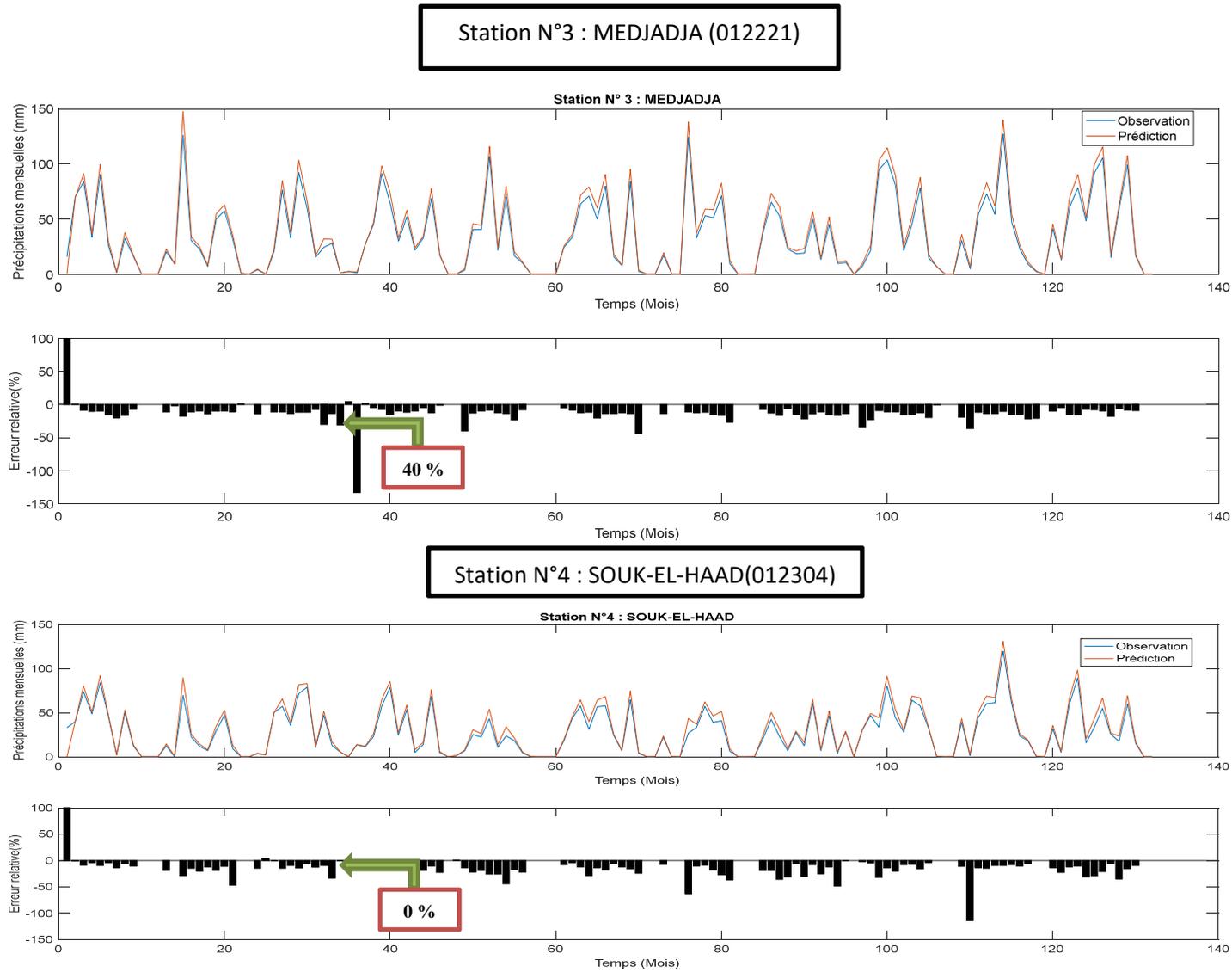


Figure N° IV.8 : Prédictions mensuelles des précipitations en fonction des mois (exp : Station N°3 et Station N°4)

Interprétation :

La 3^{ème} figure représente les prédictions mensuelles des précipitations en fonction des cinq stations hydrométriques avec l'erreur relative observées en chaque station et mois, nous remarquons que l'erreur lors des premiers mois de calcul est de 20%, alors que pour les derniers mois elle varie entre 3% et -3%, chose qui prouve l'amélioration de l'algorithme et son adaptation avec les données au fil des itérations.

La même chose pour la 4^{ème} figure ou on remarque la baisse de l'erreur relative entre la station N°3 et la station N°4.

IV.9 Analyse des innovations du filtre (résidus)

Les innovations du filtre dans le cas des 5 stations pluviométriques, et pour le pas annuel et mensuel, présente des valeurs élevées au cours des premières années ou des premiers mois par rapport aux restes. Cette différence est tout à fait normale et est principalement due à l'adaptation de l'algorithme du filtre qui fait que durant les premières itérations l'on assigne plus de confiance à la mesure en tant qu'information objective qu'au modèle ajusté. L'écart entre l'observation et la prédiction est par conséquent important au début, mais une fois que ce dernier s'adapte on obtient des innovations beaucoup plus petites vers la fin du calcul, et cela est valable pour toutes les stations étudiées comme on peut le constater dans les figures suivantes :

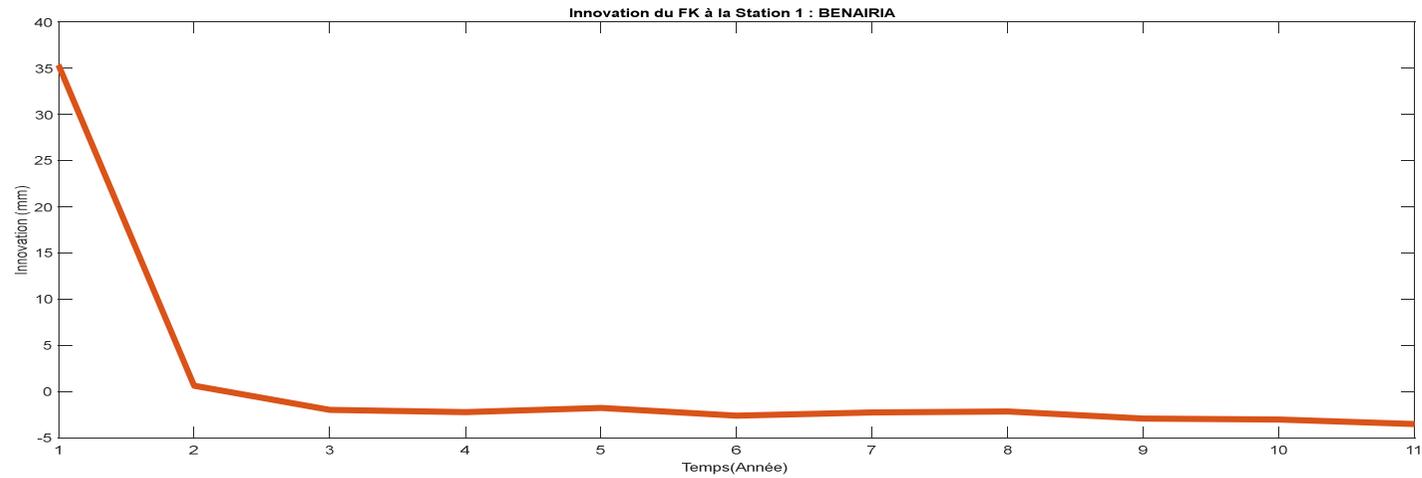


Figure N° IV.9 : Innovation du FK à la Station N°1 (Pas Annuel)

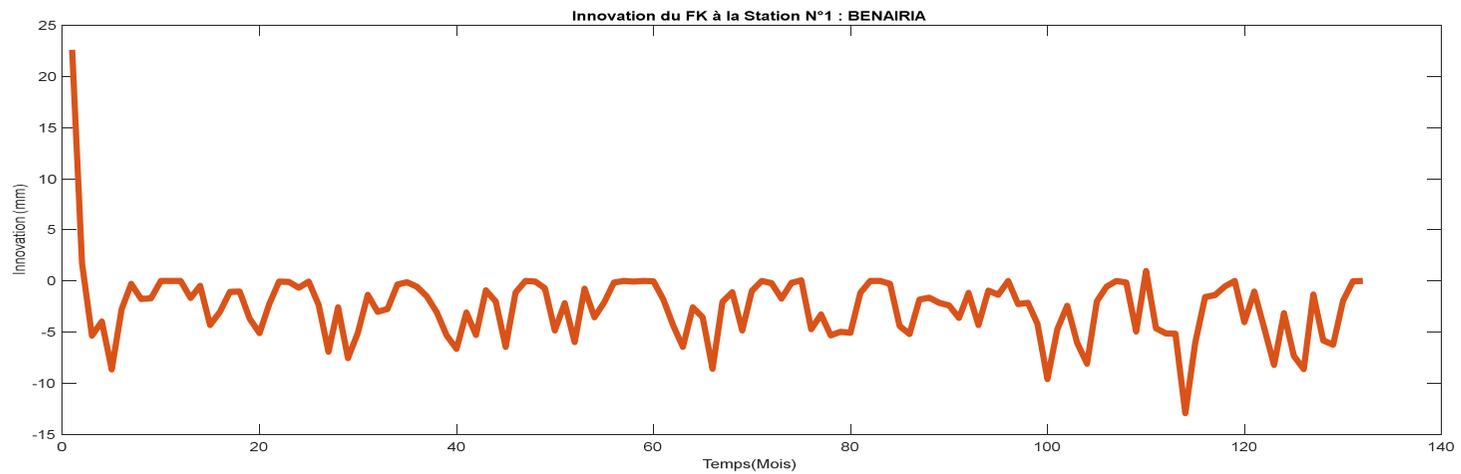


Figure N° IV.10 : Innovation du FK à la Station N°1 (Pas Mensuel)

Remarque : Le reste des résultats des innovations du filtre de Kalman sont donnés dans l'Annexe N° 1

IV.10 Comparaison entre le pas annuel et mensuel

En comparant les graphes avec un pas mensuel et ceux avec un pas annuel, on remarque que pour le pas annuel l'erreur lors des dernières itérations varie entre -10 et -20%. Tandis que pour le pas mensuel, cette dernière varie entre 3% et -3%. Chose qui prouve que le filtre est plus performant au pas mensuel.

On remarque aussi en comparant la **Figure N° IV.2** et la **Figure N° IV.6** que la trace de la matrice de covariance de l'erreur dans le cas du pas annuel converge vers une valeur positive estimée à 45 alors que pour le pas mensuel, elle converge vers une valeur très proche du zéro, ce qui prouve encore une fois la précision du pas mensuel par rapport au pas annuel.

IV.10.1 Moyennes et déviation standard des prédictions multi-sites

Les figures suivantes montrent la variation de l'écart moyen des observations et prédictions obtenues aux 5 stations au cours de la période d'étude, pour le pas annuel et mensuel. On remarque que les valeurs des observations sont constamment supérieures à celles des prédictions.

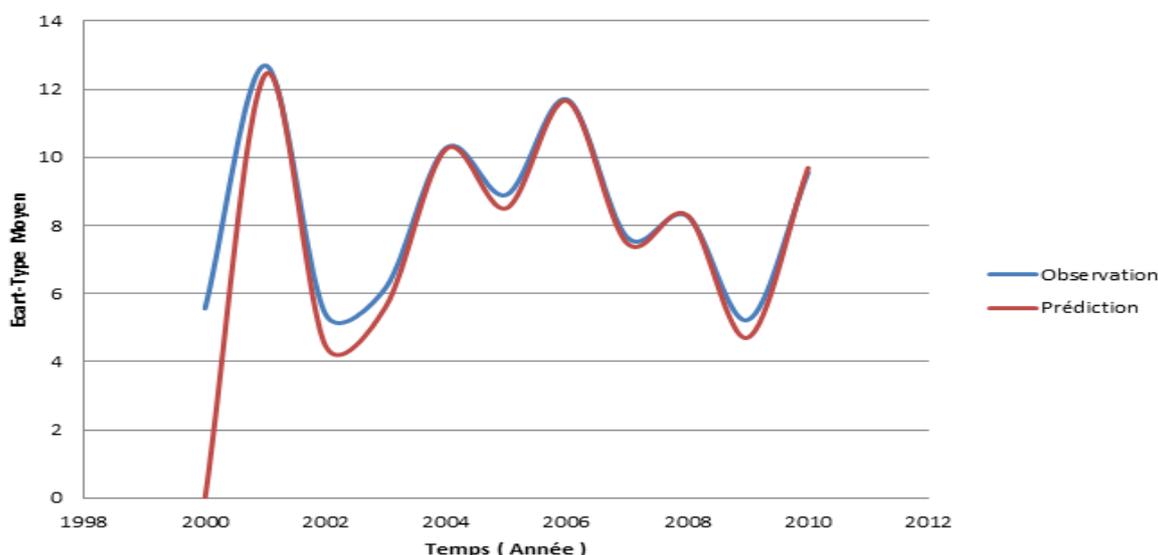


Figure N° IV.11 : Ecart-type moyen des observations et des prédictions annuelles aux 5 stations (2000-2011)

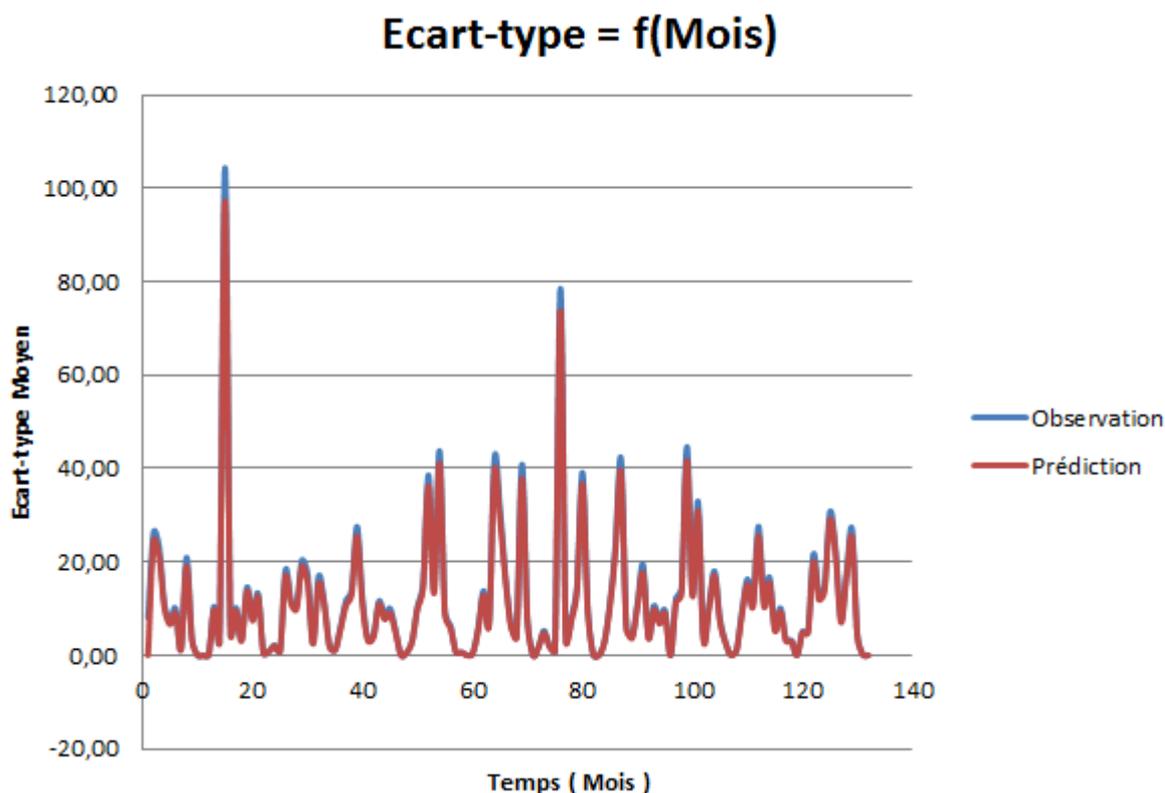


Figure N° IV.12 : Ecart-type moyen des observations et des prédictions mensuelles aux 5

Les résultats montrent que les prédictions multi-sites obtenues par le FK suivent de plus près les observations enregistrées avec une erreur relative moyenne en valeur absolue de l'ordre de 10,11% pour le pas annuel et de 7,33% pour le pas mensuel. Ces erreurs de prédictions sont minimales vu que leurs moyenne est inférieure ou égale à 10%, ce qui montre l'efficacité du filtre de Kalman dans la modélisation et la prédiction des précipitations au bassin versant Moyen-Cheliff.

On remarque aussi en comparant la **Figure N° IV.11** et la **Figure N° IV.12** que la différence entre les écarts type moyen pour le pas mensuel est moindre par rapport à celle observée au pas annuel. Chose qui confirme que le filtre de Kalman à une meilleure adaptation avec le pas mensuel, vu que l'écart type moyen est l'un des outils qui prouve la performance du filtre de Kalman.

Conclusion

Au cours de ce chapitre, nous avons utilisé un modèle de type filtre de Kalman pour la prédiction multi-site des précipitations. A cet effet, les séries chronologiques des précipitations annuelles et mensuelles, couvrant une période d'observation commune de 11 ans (2000 à 2010) ont été investies. Ces dernières sont observées à 5 stations pluviométriques dans le bassin versant du Moyen Cheliff, au Nord d'Algérie, et Les prédictions annuelles et mensuelles obtenues dans la dimension temporelle tout comme la dimension spatiale, sont proches des valeurs observées sur la même période et surtout en ce qui concerne le pas mensuel, ce qui indique que le filtre de Kalman multi-site est un outil efficace pour la modélisation et la prédiction des précipitations dans le Moyen Chélif.

La performance du modèle utilisé a été mise en épreuve par le calcul du pourcentage relative des prédictions multi-sites. Ce pourcentage moyen sur toute la période d'observation est inférieur à 20 % pour le pas annuel et à 5% pour le pas mensuel, ce qui est très acceptable.

Le tracé des graphes du gain du filtre de Kalman et de la covariance des erreurs ainsi que des écarts type moyens a montré la performance du filtre de Kalman, et prouvés que la prédiction pour le pas mensuel est plus précise que celle pour le pas annuel.

Les valeurs élevées de la prédiction par rapport à celles de l'observation nous ont aussi laissé déduire que le modèle à parfois tendance à la surestimation.

En fin, nous pouvons dire que le filtre de Kalman possède une propriété très importante qui est l'adaptation rapide au fur et à mesure que les données hydrologiques sont injectées. Un autre avantage de la technique est que l'application est faite dans le domaine temporel, cette caractéristique joue un rôle important dans la prévision en temps réel des séries chronologiques en hydrologie.

CONCLUSION GENERALE

Dans la présente étude, nous avons développé un modèle stochastique linéaire multi-site, formulé en espace-état, pour satisfaire la structure d'un filtre de Kalman discret.

Ce modèle est composé de deux équations. La première exprime l'état du système hydrologique concerné, dans notre cas, c'est un vecteur donnant les précipitations annuelles (respectivement mensuelles) à plusieurs sites à la fois. La deuxième est une équation de mesure, qui est dans ce cas un vecteur de même dimension que l'état, obtenu par observation de précipitations annuelles (respectivement mensuelles) aux sites précédents.

Nous avons utilisé le filtre de Kalman discret pour la prédiction spatiale et temporelle des précipitations dans 5 stations hydrométriques sur une période d'observation de 11 ans (2000-2010)

Les résultats de cette étude montrent l'efficacité du filtre de Kalman et son adaptation rapide aux données qui lui sont injectées au fil des itérations, ils montrent aussi la différence entre la prédiction avec un pas annuel et celle avec un pas mensuel, ou on a remarqué que les prédictions avec le pas mensuel sont beaucoup plus proches de la mesure réelle. Ce qui nous permet de dire que la précision des valeurs données par notre algorithme augmente avec la diminution du pas utilisé.

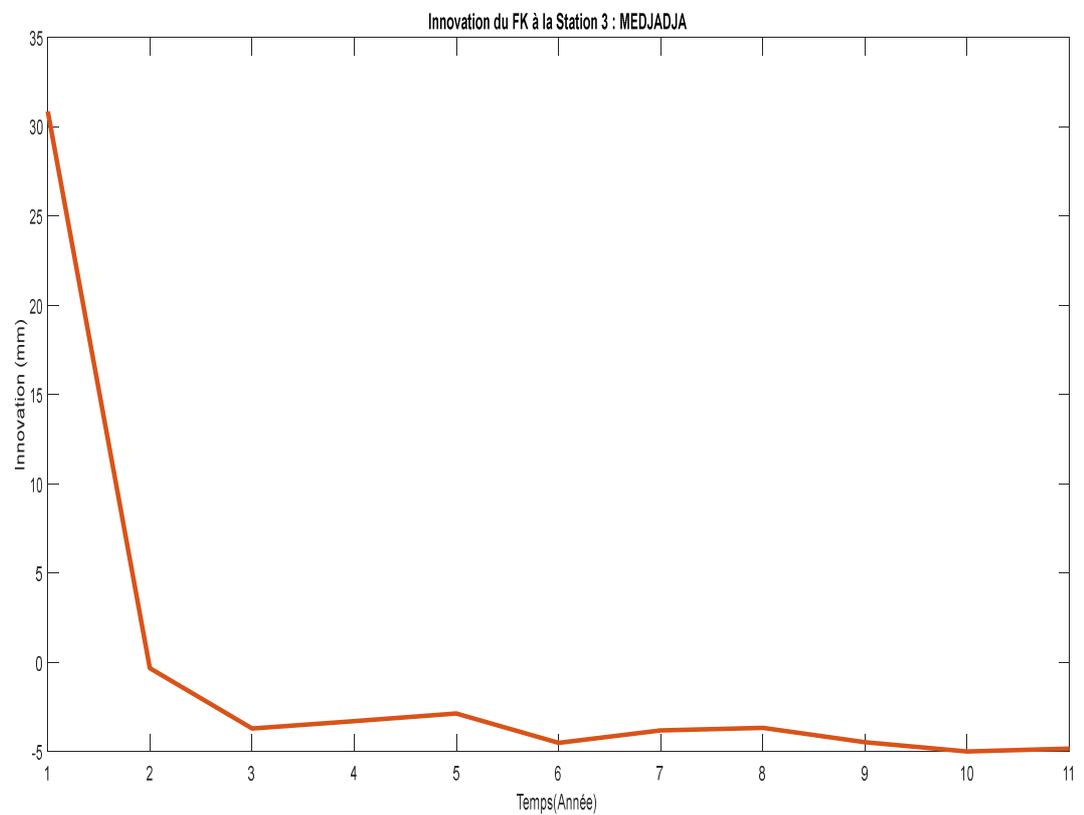
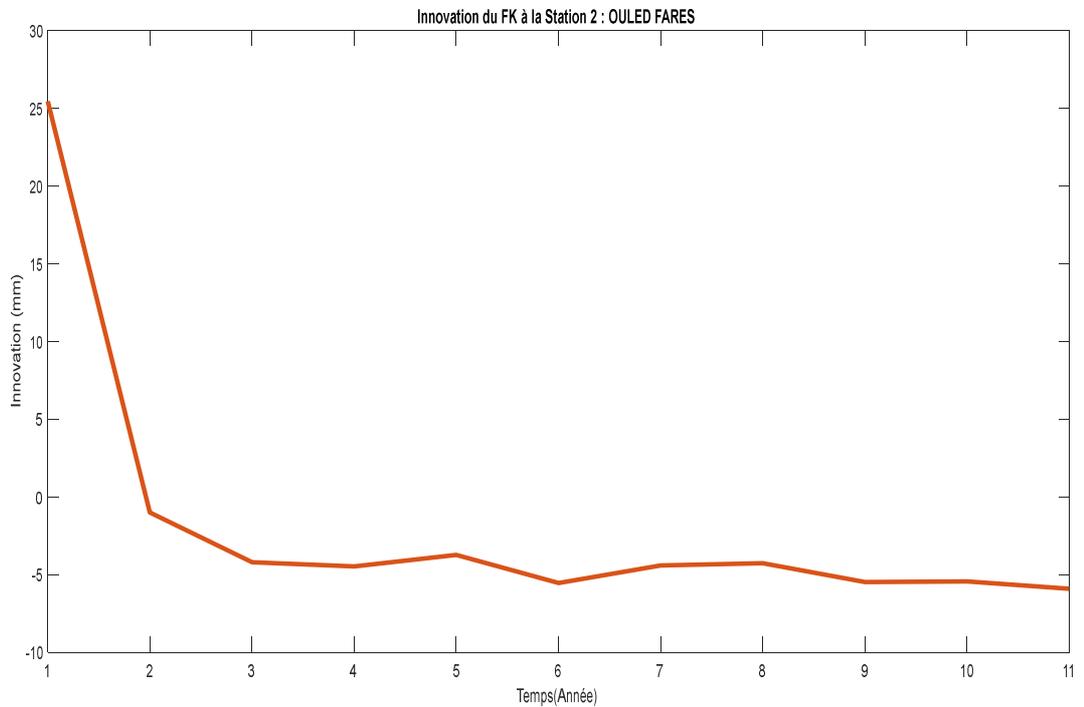
En outre, le calcul des moyennes et des déviations standards pour les données observées et leurs valeurs prédites a montré que les déviations standards des valeurs observées sont plus grandes que celles des valeurs prédites, ce qui pourrait traduire une éventuelle tendance du (FK) multi-site à une surestimation.

Le filtre de Kalman est alors un algorithme très efficace qui peut être utilisé dans le domaine de l'hydrologie. Et cela grâce aux facultés qu'il a de s'adapter à n'importe quel type de données et de fournir des résultats optimaux.

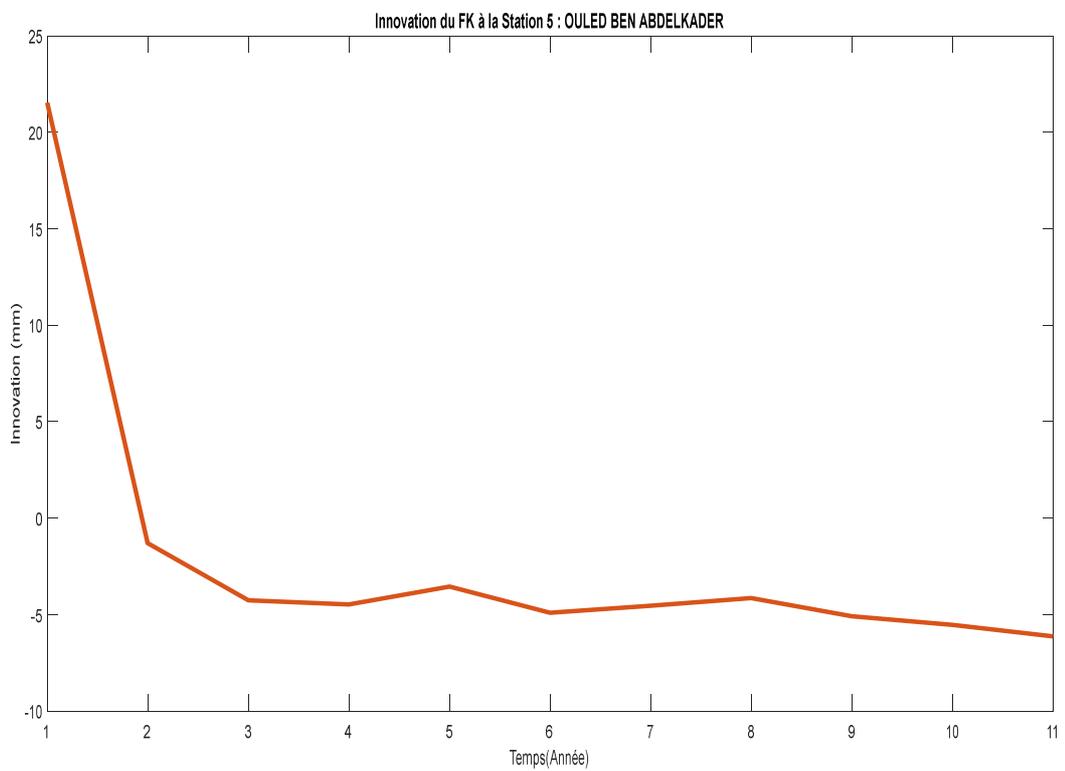
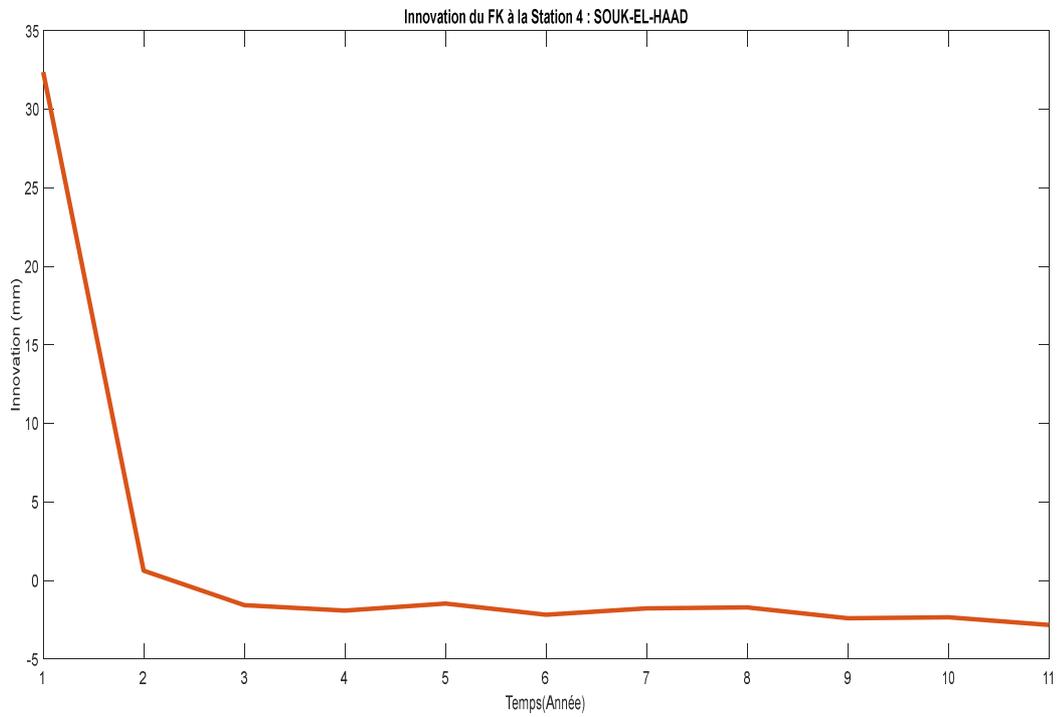
Références bibliographiques

- ❖ [1] Agence De Bassin Hydrographique Cheliff-Zahrez, ATLAS DU CHELIFF.
- ❖ [2] ALAZARD,D . Octobre 2006. Introduction au filtre de KALMAN.
- ❖ [3] AYACHE, A., HARMONIER.J. Cours de statistique descriptive.
- ❖ [4] BELMAHDI, F.2015. Application du filtre de Kalman pour le débruitage des signaux. Mémoire de Magister.
- ❖ [5] BOUKHAROUBA, A. Kettab, Kalman filter technique for mutli-site modeling and stream flow prediction in Algeria, J. Food Agric. Environ. 7(2) (2009) 671–677.
- ❖ [6] DAHMANI, M.Avril 2012. Application du filtre de Kalman linéaire et non linéaire et ses dérivées à la poursuite d'une cible manœuvrant. Thèse de Doctorat en Science.
- ❖ [7] DAUXOIS, JY. Septembre 2013. Cours de probabilité.
- ❖ [8] GACOGNE, L., FRUGIER, G. 1990. Probabilités et statistiques. cours de mathématiques. TOME 4.
- ❖ [9] HARKAT, S. Application du Filtre de Kalman sur la variabilité pluviométrique dans le bassin versant de Cheliff » Algerie. Université de Chlef, 2014.

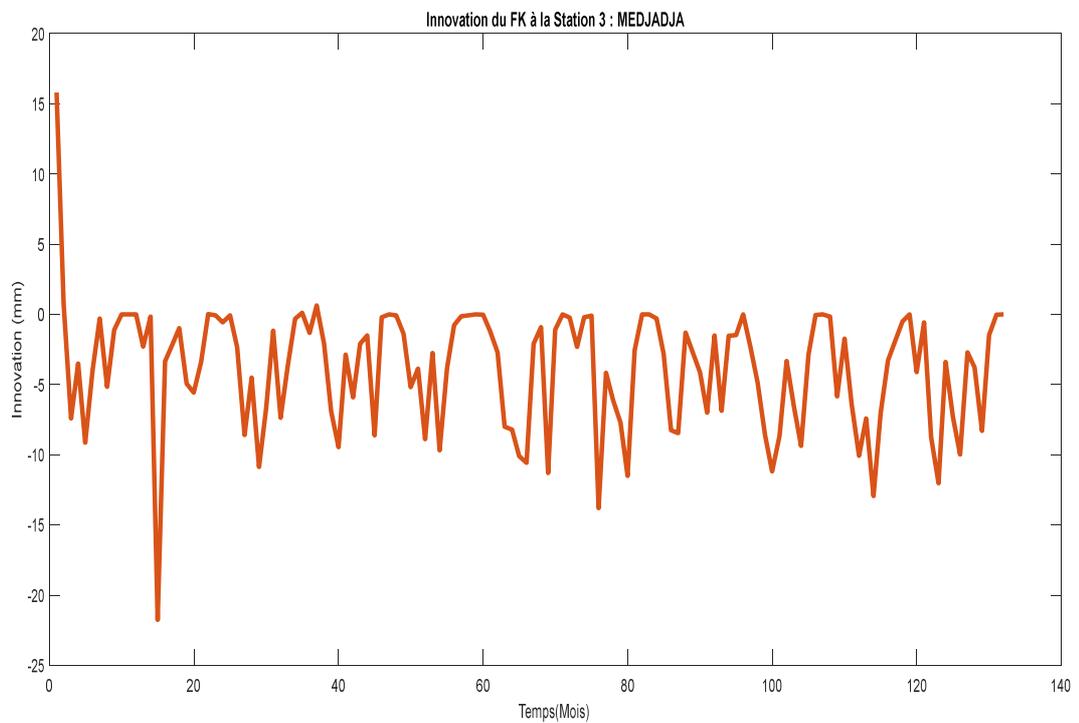
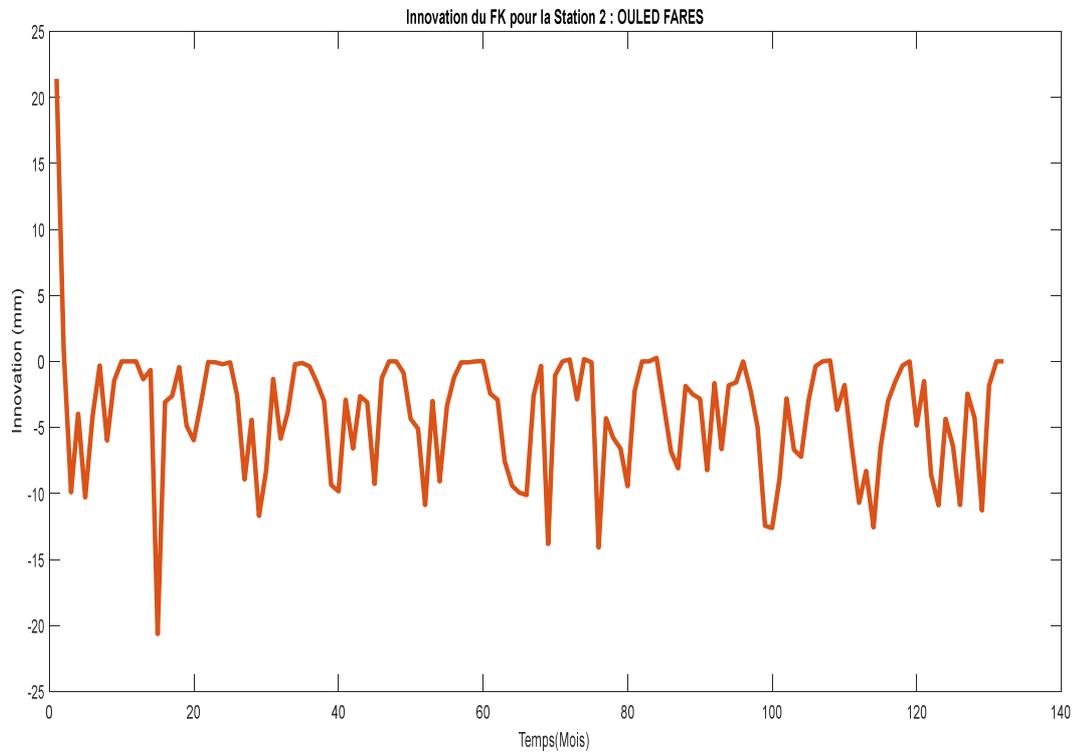
Annexe N°1 : Innovation du Filtre de Kalman (Pas Annuel)



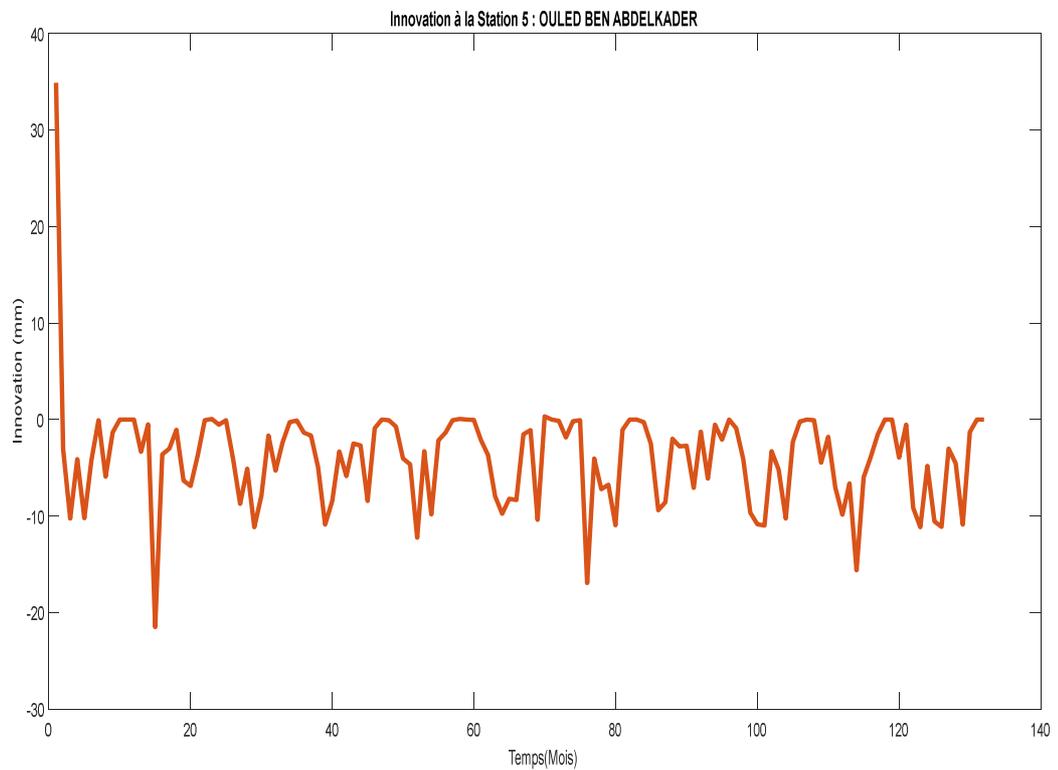
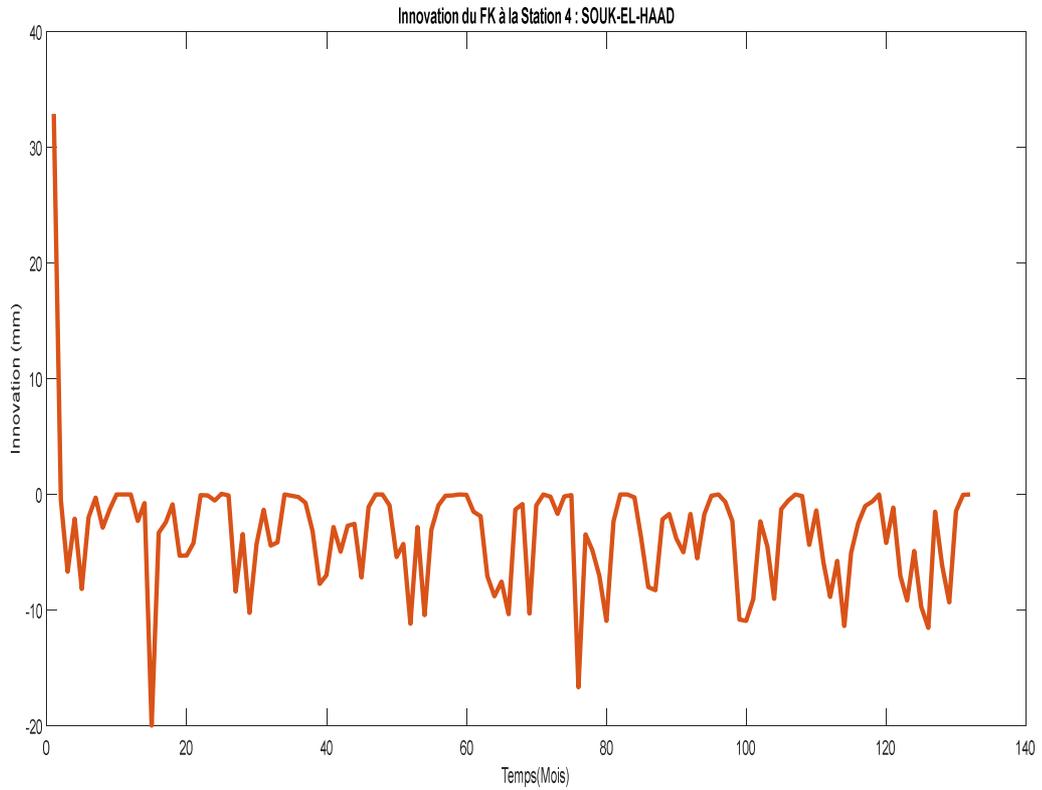
Annexe N°1 : (Suite)



Annexe N°2 : Innovation du Filtre de Kalman (Pas Mensuel)

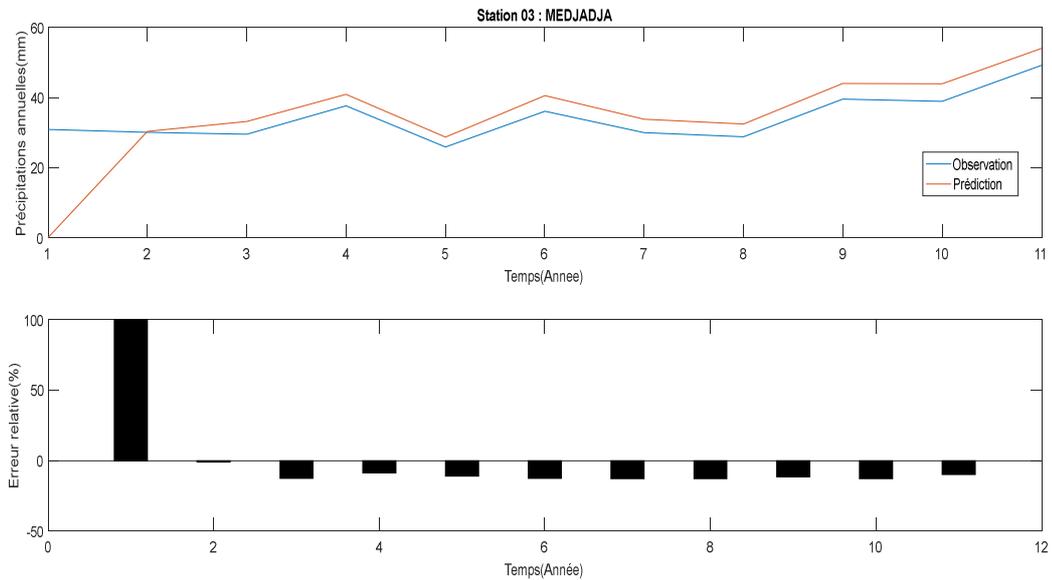


Annexe N°2 : (Suite)

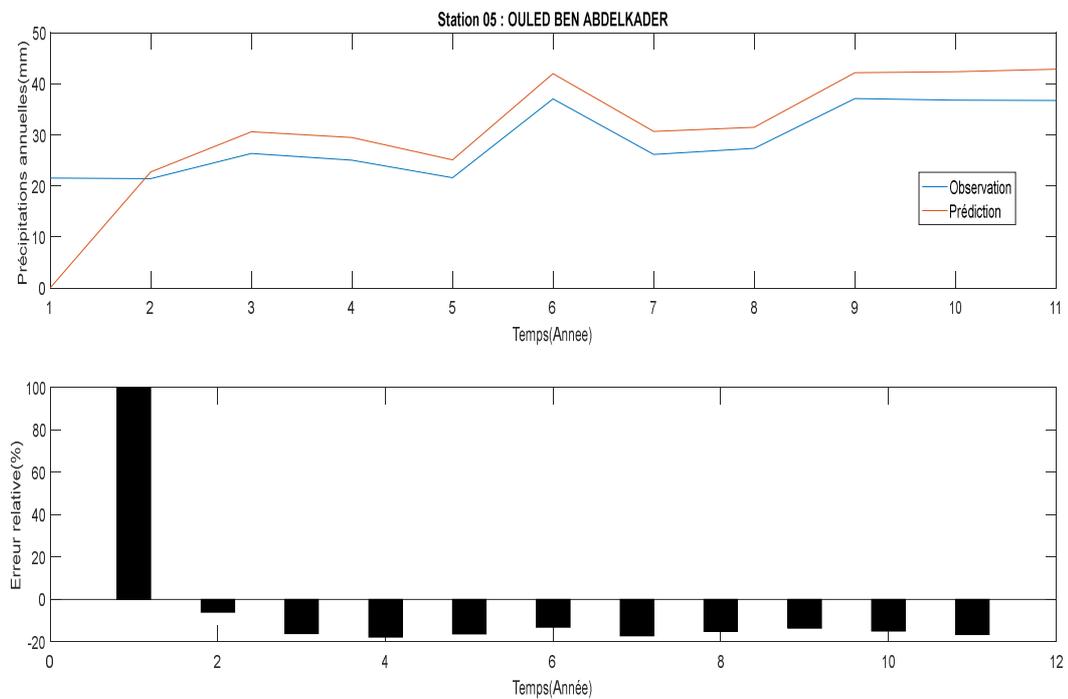


Annexe N°3 : Prédiction et erreurs relative dimension temporelle (Pas Annuel)

Station 03 : MEDJADJA

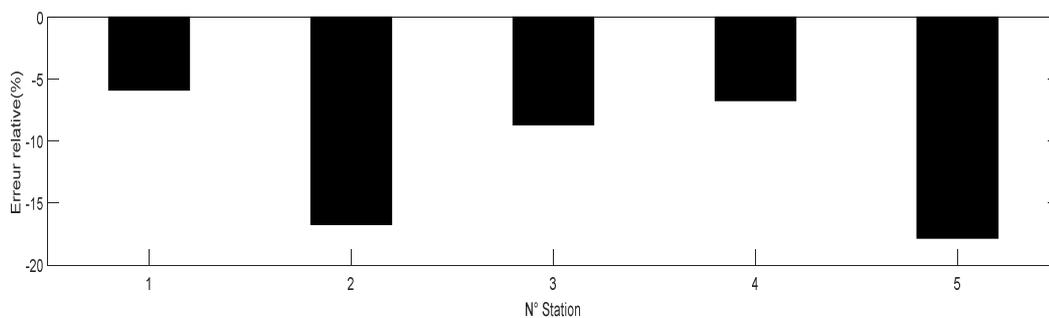
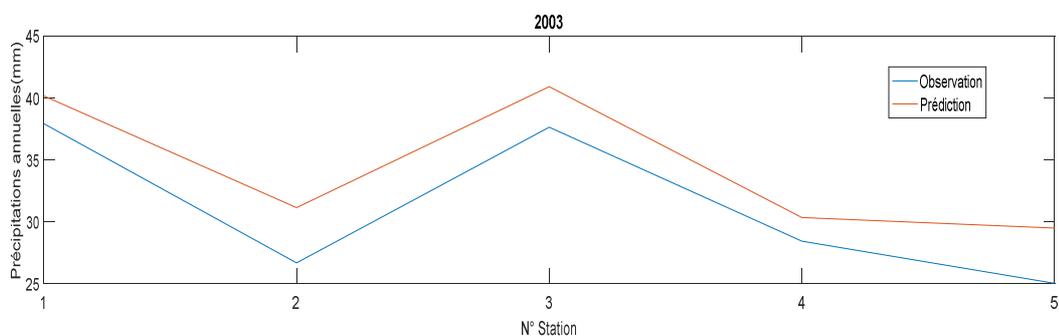


Station 05 : OULED BEN

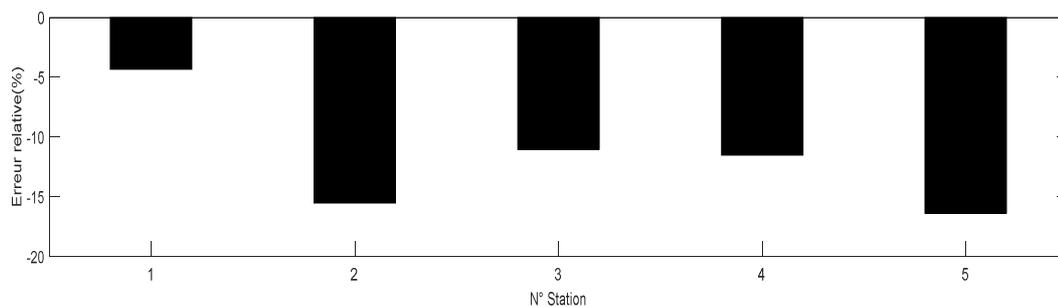
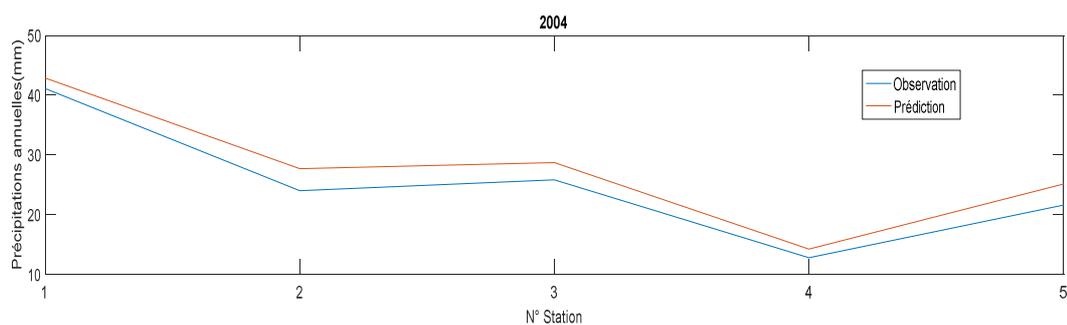


Annexe N°4 : Prédiction et erreurs relative dimension Spatiale (Pas Annuel)

2003

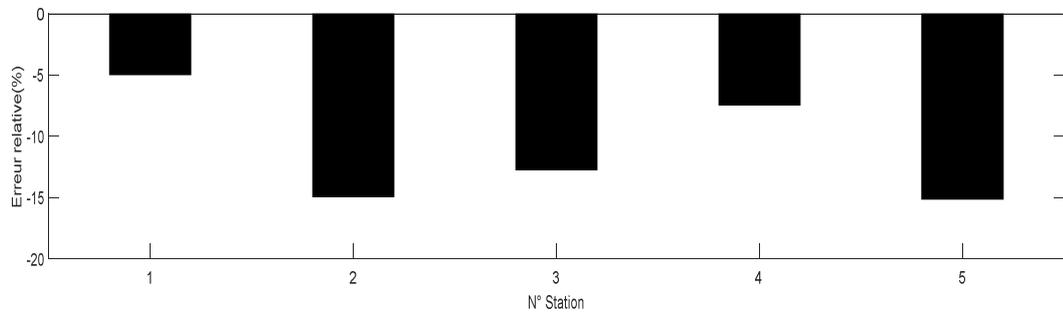
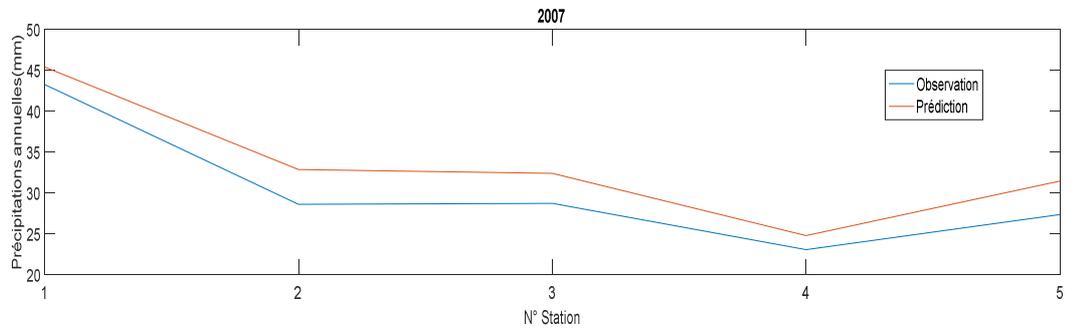


2004

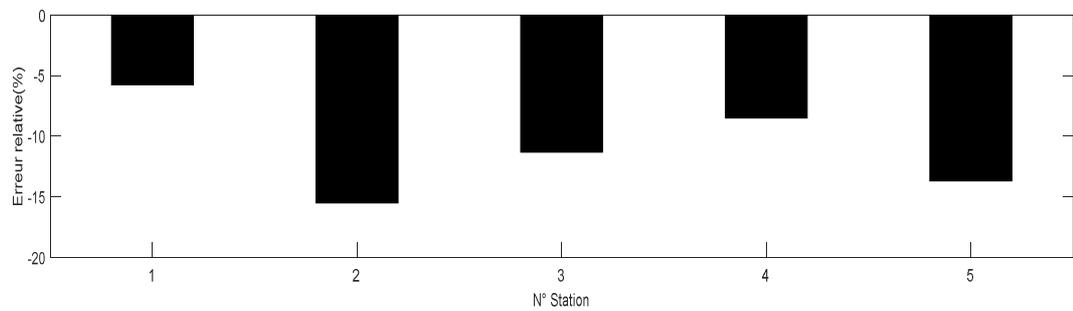
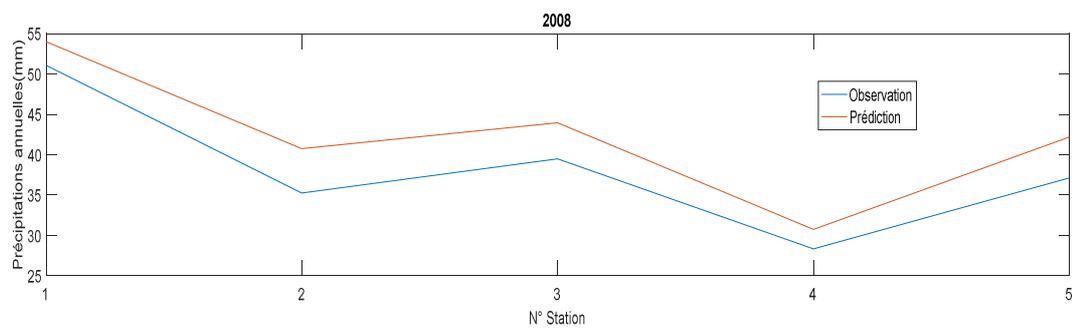


Annexe N°4 : (Suite)

2007

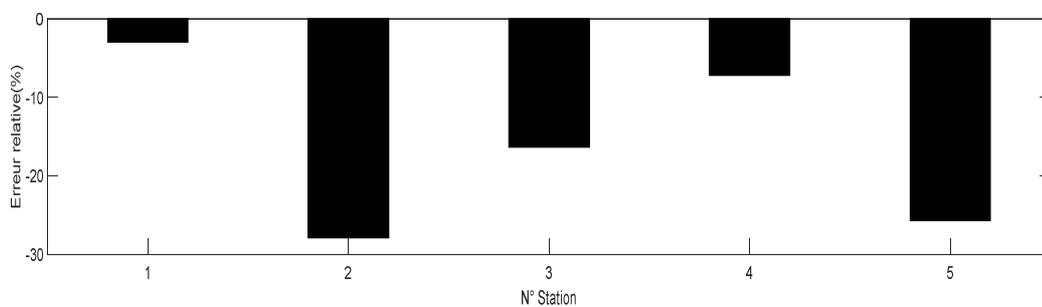
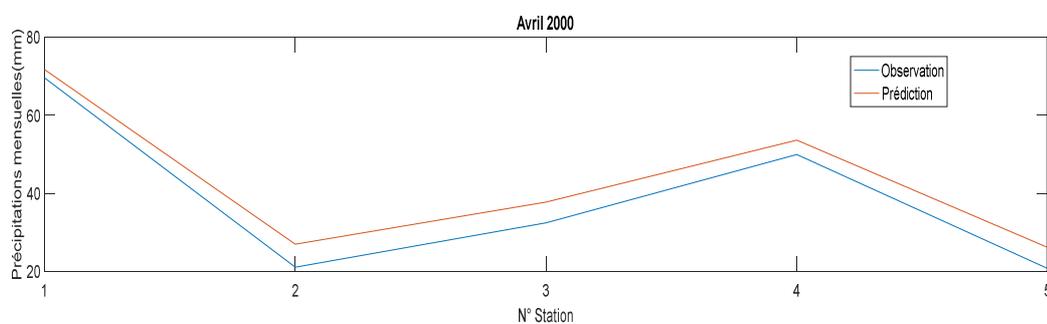


2008

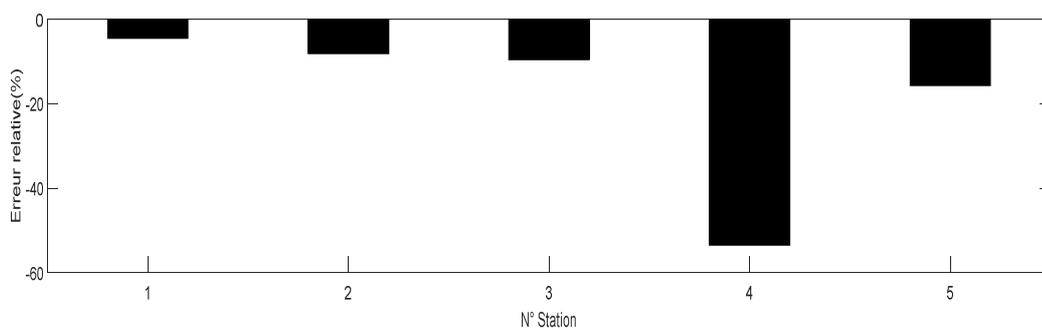
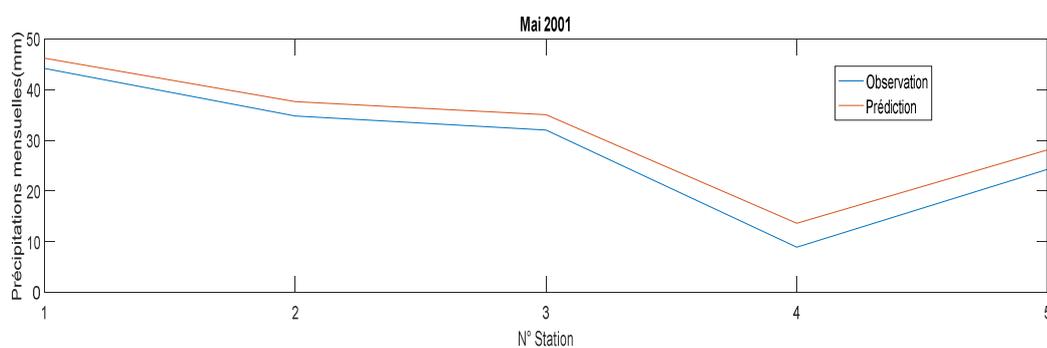


Annexe N°5 : Prédiction et erreurs relative dimension Spatiale (Pas Mensuel)

Avril 2000

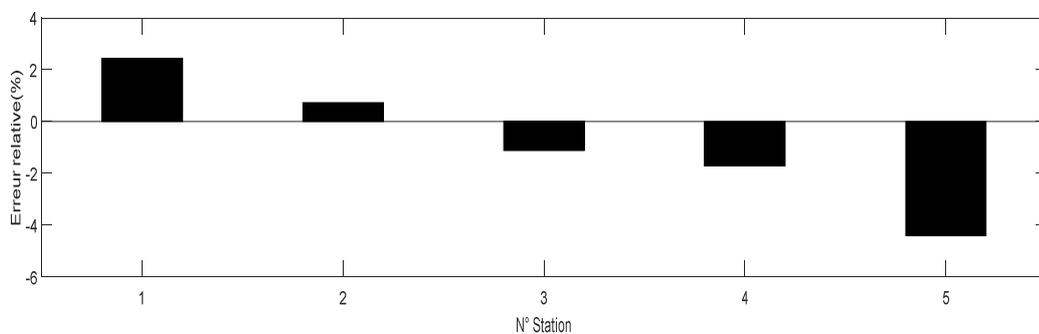
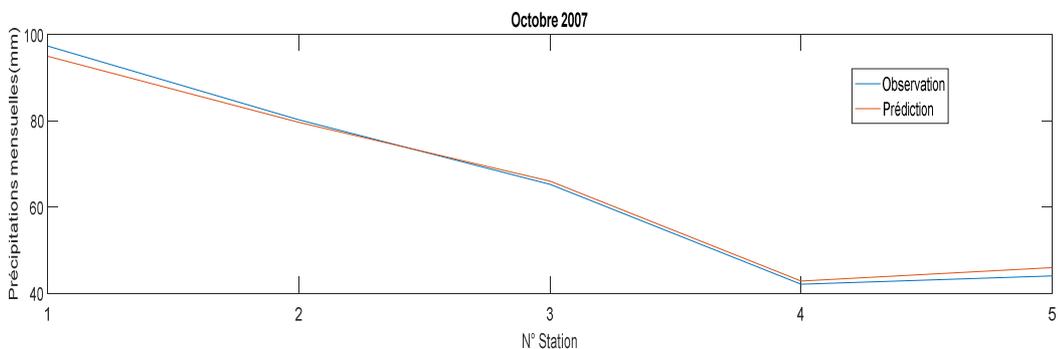


Mai 2001

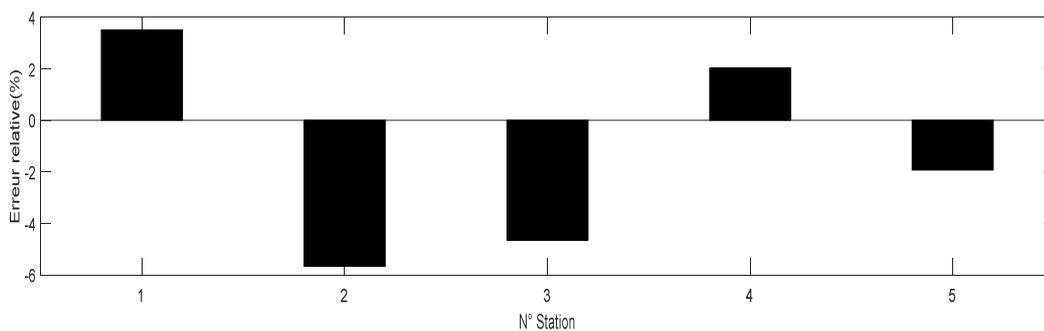
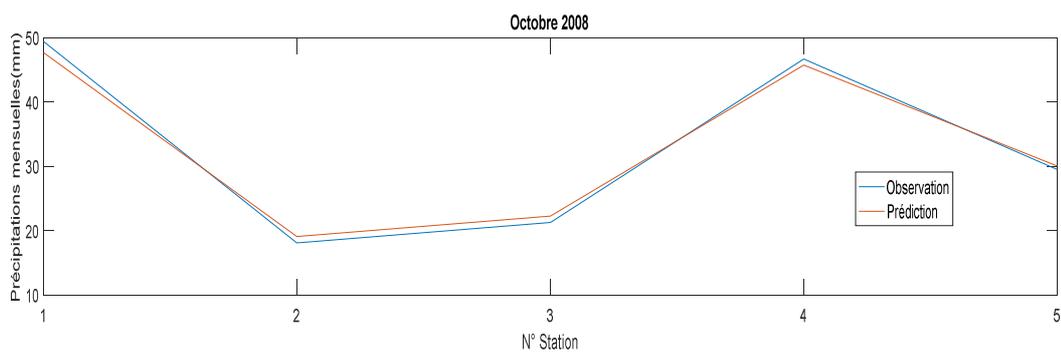


Annexe N°5 : (Suites)

Octobre 2007

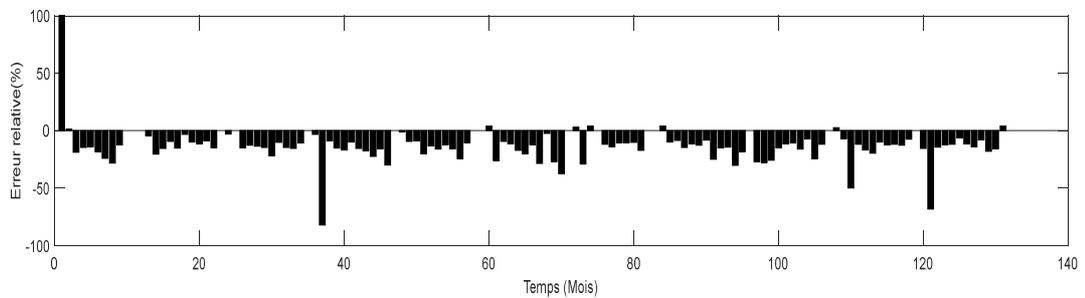
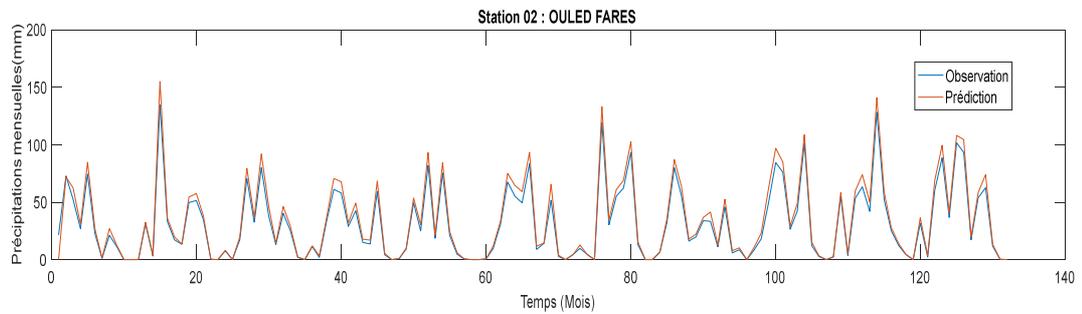


Octobre 2008



Annexe N°6 : Prédiction et erreurs relative dimension Temporelle (Pas Mensuel)

Station 02 : OULED FARES



Station 05 : OULED BEN ABDELKADER

