

Higher National School of Hydraulic

The Library

Digital Repository of ENSH



المدرسة الوطنية العليا للري

المكتبة

المستودع الرقمي للمدرسة العليا للري



The title (العنوان):

**Le filtre de Kalman pour la prédiction multi-site des apports liquides annuels du bassin versant Cheliff.**

The paper document Shelf mark (الشفرة) : 6-0002-17

APA Citation (توثيق APA):

**Kadri, Mohammed (2017). Le filtre de Kalman pour la prédiction multi-site des apports liquides annuels du bassin versant Cheliff [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Digital Repository of ENSH aims to limit scientific production, whether published or unpublished (theses, pedagogical publications, periodical articles, books...)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 بنمّين الإنتاج العلمي لأساتذة و باحثي المدرسة.

يهدف المستودع الرقمي للمدرسة إلى حصر الإنتاج العلمي سواء كان منشورا أو غير منشور (أطروحات، مطبوعات، مبداعات، مقالات، دوريات، كتب....) و بثه 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 :**

**Le filtre de Kalman pour la prédiction multi-site  
des apports liquides annuels du bassin versant  
Cheliff**

### **Présenté par :**

M<sup>r</sup> : KADRI Mohammed

### **DEVANT LES MEMBRES DU JURY**

Nom et Prénom	Grade	Qualité
M <sup>r</sup> SALAH Boualem	Professeur	Président
M <sup>me</sup> MOKRANE Wahiba	M.A.A	Examinatrice
M <sup>r</sup> HACHEMI Abdelkader	M.A.A	Examineur
M <sup>me</sup> TAFAT Leila	M.A.A	Examinatrice
M <sup>me</sup> SADOUNE Samra	M.C.B	Promotrice

Avril 2017

## *Remerciement*

Je remercie MES PARENTS en premier lieu qui m'ont permis et facilité les choses pour accomplir mes études.



Je remercie aussi tous les autres membres de ma famille qui m'ont aidés et encouragés quel que soit les conditions, en particulier ma deuxième mère HORIA et père TAHER



Le devoir est très agréable d'exprimer ma vive gratitude à madame SADOUNE pour son aide indéfectible ainsi que ses conseils précieux tout au long de ce travail. Sans cette brave femme je ne serais fait la moitié de ce que j'avais fait.



Je tiens à remercier aussi :



Mes amis pour leur présence ;

Les membres de jury pour avoir accepté d'évaluer mon travail ;

Tous les professeurs qui ont contribué à ma formation ;

Tous ceux qui ont contribué de près ou de loin à l'élaboration et l'aboutissement de ce mémoire.



Sans oublier tout personnage de l'ENSH et de LANRH BLIDA, à qui j'ai une affection particulière.



# Dédicace

Je dédie ce travail à :

*Celle qui sont la plus chers au monde, mes parents, à qui je n'arriverai jamais à exprimer ma gratitude et ma reconnaissance, pour son amour son soutien tout au long de mes études afin de faire de moi ce que je suis aujourd'hui, que DIEU les protège toujours.*

*A mes frères  
A mes collègues et mes amis.  
A toutes la famille.  
A tous ceux qui me sont chers.*

**KADRI MOHAMMED**

ملخص:

تتدرج هذه الدراسة في اطار انشاء التصاميم الرقمية الزمنية للظواهر الهيدرولوجية من اجل التنبؤات الدقيقة. في هذا العمل إهتمنا بدراسة عامل التدفقات السائلة و تغيراتها وفق المجال الجغرافي و البعد الزمني في حوض الشلف العليا بالجزائر، إذ ان الهدف هو إثبات إمكانية تطبيق مصفاة كالمان على التصاميم الرقمية و التنبؤات متعددة المواقع وأيضا تحسين مجال الخطأ في هذه التنبؤات. تعطي هذه التقنية فائدة كبيرة تتمثل في إمكانية تحديد قيمة الخطأ في التنبؤات و التي تعد مقياسا في تحسين الدقة في مختلف التنبؤات. بالإضافة لهذا فإن خوارزمية كالمان تعمل وفق البعد الزمني بطبيعة تراجعية ، الامر الآخر هو أن هذه التقنية تستعمل كل المعطيات المتوفرة و تضيف معطيات جديدة باستمرار .

وعليه قمنا باستخدام المعطيات المتوفرة من خمس محطات هيدرمتريية موزعة بانتظام على حوض الشلف العليا وفي فترة زمنية مشتركة تمتد من 1990 الى 2011، جاءت النتائج تنطوي على النقاط التالية:

- نسبة الخطأ المقدره بمصفاة كالمان تعد جد صغيرة إذ أن الفرق بين التوقعات و المعطيات ضئيل وفق البعد المكاني و الزمني على حد سواء.
- هذه النتائج تثبت فعالية مصفاة كالمان في القيام بالتوقعات الجيدة.
- تصل نسبة متوسط الخطأ في جميع المحطات وفق كل الفترة الزمنية الى حوالي 10 % والذي يعد خطأ مقبولاً .

**الكلمات المفتاحية:** مصفاة كالمان , التدفقات السائلة , التنبؤات متعددة المواقع , حوض الشلف العليا بالجزائر

## Résumé :

L'objectif de ce travail est l'étude de l'applicabilité de la technique du filtre de Kalman (FK) discret à la modélisation et la prédiction multi-site des apports dans le bassin versant de haut Cheliff (Algérie), ainsi que l'amélioration des erreurs de prédiction, cette technique présente l'avantage de procurer l'erreur de prédiction qui présente en soit un indicateur de précision. En plus, son algorithme travaille dans le domaine temporel avec une nature récursive et dispose d'un estimateur optimal dans le sens des moindres carrés. Un autre aspect de son optimalité est qu'il incorpore toute l'information disponible sur le système, les mesures et les erreurs, dans un opérateur adaptatif qui est recalé à chaque fois qu'une nouvelle mesure devient disponible.

A cet effet, les données de cinq (5) stations hydrométriques dans le bassin versant de haut Cheliff ont été investies sur une période de 21 ans (1990 à 2011). Les résultats obtenus montrent que les prédictions multi-sites suivent de plus près les observations enregistrées et les erreurs de prédiction sont minimales. Ceci est valable aussi bien dans la dimension temporelle que spatiale, ce qui montre que le FK est un outil assez performant procurant de bonnes prédictions.

En fin L'erreur relative moyenne des prédictions est inférieure à 10% ce qui est largement acceptable.

**Mots-clés :** Filtre de Kalman, Prédiction multi-site, Précipitations, Bassin versant du Chélif, Algérie.

## Abstract:

The objective the study is to investigate the applicability of the Kalman filter discrete to modeling and the multisite prediction of the contributions in the catchment area of high Cheliff (Algeria), as well as the improvement of the errors of prediction. This technique has the advantage of getting the error of prediction which presents an indicator of precision. Moreover, its algorithm works in the temporal field with a recursive nature, and has an optimal estimator in the direction of least squares. Another aspect of its optimality is that it incorporates all information available on the system, measurements and the errors, in an adaptive operator which is readjusted with each time a new measurement becomes available. For this purpose, the data of five (5) hydrometric stations in the catchment area of high Cheliff were investigated over a period of 21 years (1990 to 2011) and then put into action. The results show that the multisite predictions more closely follow the recorded observations and the errors of prediction are minimal. This is valid as well in both dimensions, which shows that the KF is a rather powerful tool getting good predictions. In the end, the average error relative of the predictions is lower than 10% which is largely acceptable.

**Keywords:** Kalman filter, multisite prediction, liquid contributions, catchment area of high Cheliff (Algeria).

## *Liste des tableaux*

<i>Tableau II.1 : Classes de relief en fonction de dénivelée spécifique. ....</i>	<i>17</i>
<i>Tableau II.2 : Classes de relief en fonction de dénivelée spécifique. ....</i>	<i>17</i>
<i>Tableau II.3 : Caractéristiques hydrométriques du bassin versant « le haut Cheliff ». ....</i>	<i>18</i>
<i>Tableau II.4 : Les Stations pluviométriques retenues. ....</i>	<i>20</i>
<i>Tableau II.5 : Caractéristiques des stations. ....</i>	<i>21</i>
<i>Tableau II.6 : les paramètres de l'échantillon. ....</i>	<i>21</i>
<i>Tableau II.7 : test de Mood. ....</i>	<i>21</i>

## Liste des figures

<i>Figure 1.1 : Graphique de la fonction de répartition</i> .....	4
<i>Figure II.1 Représentation des cinq régions hydrographiques en Algerie</i> .....	13
<i>Figure II.2 Représentation du découpage par sous-région Du bassin Cheliff- Zahrez</i> .....	14
<i>Figure II.3: Découpage administratif de la sous-région n°1</i> .....	15
<i>Figure.II.4 : La courbe hypsométrique du bassin versant de Cheliff « le haut Cheliff</i> .....	17
<i>Figure III.1: Les étapes du filtre de Kalman</i> .....	29
<i>Figure. IV.1.1 Prédiction obtenues à la station hydrométrique EL ABABSA</i> .....	37
<i>Figure. IV.1.2 Prédiction obtenues à la station hydrométrique ARIB CHELIF</i> .....	38
<i>Figure. IV.1.3 Prédiction obtenues à la station hydrométrique BIR OULED TAHAR</i> .....	38
<i>Figure. IV.1.4 Prédiction obtenues à la station hydrométrique MARABOUT BLANC</i> .....	39
<i>Figure. IV.1.5 Prédiction obtenues à la station hydrométrique DJENANE BEN OUADAH</i> .....	39
<i>Figure.IV.2.1 Prédiction des apports annuels (1994-1995)</i> .....	40
<i>Figure.IV.2.2 Prédiction des apports annuels (1995-1996)</i> .....	40
<i>Figure.IV.2.3 Prédiction des apports annuels (1996-1997)</i> .....	41
<i>Figure.IV.2.4 Prédiction des apports annuels (2003-2004)</i> .....	41
<i>Figure.IV.2.5 Prédiction des apports annuels (2004-2005)</i> .....	42
<i>Figure.IV.2.6 Prédiction des apports annuels (2006-2007)</i> .....	42
<i>Figure.IV.2.7 Prédiction des apports annuels (2007-2008)</i> .....	43
<i>Figure.IV.2.8 Prédiction des apports annuels (2010-2011)</i> .....	43
<i>Figure.IV.3 : Gain de filtre FK</i> .....	44
<i>Figure.IV.3 Covariance d'erreur de filtre FK</i> .....	45
<i>Fig.IV.3.1 Innovations du FK à ABABSA</i> .....	46
<i>Fig.IV.3.2 Innovations du FK à ARIB CHELIF</i> .....	46
<i>Fig.IV.3.3 Innovations du FK à BIR OULED TAHAR</i> .....	47
<i>Fig.IV.3.4 Innovations du FK à DJENANE BEN OUADAH</i> .....	47
<i>Fig.IV.3.5 Innovations du FK à MARABOUT BLANC</i> .....	48



# Sommaire

**Remerciement.**

**Dédicace.**

**Résumé.**

**Liste des tableaux.**

**Liste des figures.**

**Introduction générale.**

## *Chapitre I: Notions de base pour aborder le filtre de kalman*

I.1. Introduction.....	1
I.2. Loi de probabilité.....	1
I.3. variable aléatoire réelle.....	2
I.4. Loi de probabilité à plusieurs variables .....	2
I.5. Le vecteur aléatoire .....	2
I.6. Variables aléatoires discrètes.....	3
I.7. Fonction de répartition .....	3
I.8. Rappel statistique et méthodes d'estimation.....	4
I.8.1. Paramètres de l'échantillon.....	4
a.1. Paramètre de tendance centrale : .....	4
a.2 paramètre de dispersion.....	6
b. Paramètre de forme.....	7
b.1 coefficient d'asymétrie .....	7
b.2. Coefficient d'aplatissement .....	7
I.9. Test relatif à l'échantillonnage.....	8
a. Test de stationnarité et d'homogénéité de mann et whitney (1947).....	8
b. Test de stationnarité et d'indépendance de wald et wolfowitz (1943) .....	9
c. Test de médiane (test de mood).....	9
d. Test d'anderson .....	9
e. Test du cumul des résidus .....	10
I.10. Estimation stochastique linéaire.....	10
a. Processus aléatoire.....	10
b. Bruit blanc .....	10
I.11. Modèle espace d'états ou modèle état mesure .....	11
I.11. Problème du filtrage .....	11
i.12. Conclusion : .....	12

## *Chapitre II : Cadre de l'étude*

II.1. Introduction .....	13
II.2. présentation de la sous-région d'étude haut cheliff .....	14
II.3. caractéristiques morpho métriques de la région[13].....	15
II.3.1 superficie, altitude et forme .....	16
II.3.2. Rectangle équivalent .....	16
II.3.3. Relief (indice de pente) .....	17
II.3.4. Répartition et courbe hypsométrique .....	17
II.3.4. Indice de pente globale " ig".....	18

II.3.5. Densité de drainage .....	18
II.3.6 temps de concentration du bassin versant .....	18
II.3.7. Caractéristiques morphologiques et hydrologiques des sous- bassins versants .....	19
II.4. Caractéristiques du réseau pluviométrique .....	19
II.5. Répartition géographique des précipitations .....	20
I.6. Acquisition de données .....	21
II.7. Paramètre de l'échantillon .....	21
II.8. Teste d'homogénéité .....	21
II.9. Conclusion .....	22

### *Chapitre III : Présentation du filtre de kalman*

III.1. Introduction .....	23
III.2 l'estimateur optimal au filtre de kalman.....	23
III.2.1. L'estimation .....	23
III.2.2. Les estimateurs déterministes.....	23
a. Biais d'un estimateur .....	23
b. La variance .....	24
III.2.3. L'estimateur optimal .....	24
III.2.4. Pourquoi utiliser ce filtre dans une modélisation hydrologique ?.....	24
III.2.5. Historique du filtre de kalman.....	25
III.2.6. Domaines d'utilisation .....	26
III.2.7. Le filtre de kalman.....	26
III.2.7.1. Définition .....	26
III.2.7.2. Principe du filtre .....	26
III.2.7.3 avantage et inconvénient de fk .....	26
III.2.7.4. Les limites .....	27
III.3. Présentation du filtre de kalman .....	28
a. Formulation linéaire .....	28
b. Formulation matricielle .....	28
III.4. Fonctionnement du filtre de kalman.....	29
III.5. Algorithme du filtre de kalman .....	30
III.6 stabilité du filtre vis-à-vis de la divergence .....	32
III.6. Conclusion : .....	32

### *Chapitre IV : La prédiction multi-site des apports annuels*

IV.1. Introduction .....	34
IV.2. Données de la prédiction multi-site .....	34
IV.3. Formulation du modèle .....	34
IV.4. Résultats et interprétations .....	37
IV.4.1. Prédications multi-sites .....	37
IV.4.2. Prédications temporelles (selon les années) .....	37
IV.4.3. Prédications spatiales (selon les station) .....	40
IV.4.4. Optimalité des résultats.....	44
IV.5. Analyse des innovations du filtre .....	45
IV.5. Conclusion .....	48

**Conclusion et perspectives**  
**Références bibliographiques.**

## *Introduction générale*

La prédiction peut être définie comme le comportement futur probable d'un processus et la confiance qui lui est assignée dépend de la nature du processus étudié et de la qualité de l'adéquation du modèle adopté.

Le FK est une version moderne et probabiliste du concept des moindres carrés qui se base sur la description du système en se basant sur la notion d'état, il peut être défini d'une manière simple comme étant un algorithme optimal et récursif de traitement de données. Son principe est simple, il consiste à combiner deux estimations indépendantes pour former une estimation pondérée ou prédiction. La première estimation peut être une prédiction à-priori ou une estimation basée sur une information a-priori, la seconde, une prédiction basée sur la nouvelle information (nouvelle mesure). L'objectif du FK est justement de combiner ces deux pièces d'information afin d'obtenir une estimation améliorée et son grand avantage est de procurer avec précision la covariance des erreurs de prédiction qui constitue en soi un critère d'optimalité. Un autre aspect d'optimalité du filtre de Kalman c'est qu'il incorpore toute l'information disponible sur le système et ses perturbations et peut être amorcé avec le minimum d'information disponible.

En effet, depuis son introduction par Kalman (1960) et ses applications initiales par Kalman et Bucy une année après le filtre de Kalman n'a pas cessé d'intriguer les chercheurs

L'optimalité des résultats obtenus depuis sa mise en application dans les années soixante du siècle dernier jusqu'à ce jour, fait du filtre de Kalman l'outil favori des ingénieurs et chercheurs travaillant dans les domaines où la prédiction ou l'estimation linéaire sont requises. Ce puissant outil a également été appliqué à une variété de problèmes hydrologiques ou à des problèmes en relation avec l'eau car ces derniers sont souvent sujets à des influences aléatoires d'origines inconnues.

En Algérie et jusqu'à ce jour, la recherche bibliographique a montré que les modèles dynamiques basés sur la description d'état n'ont jamais eu une véritable chance d'être appliqués à la prédiction hydrologique et ce malgré le succès remarquable observé., cette technique a toujours été évitée dans les centres de recherche et de prise de décision, exception faite de quelques contributions modestes dans ce domaine (Boukharouba, 1996), (Boukharouba et Mokrane , 1997), (Boukharouba et Kettab, 2000) pour l'étude des apports , et (Boukharouba et Kettab, 2006 et (S.Harkat en 2016) pour l'étude des précipitations annuel

et mensuels mais leur travaux innovantes restées inexploitées malgré la fiabilité et la crédibilité des résultats obtenus par ces chercheurs algériens .

Notre application dans ce travail consiste en l'utilisation de ce filtre pour la prédiction multi sites des apports liquides dans le bassin de haut chelif. L'idée de base derrière cette méthode, consiste à utiliser les informations à priori (les observations), afin d'améliorer les performances des méthodes existantes et de concevoir de nouvelles techniques de filtrage.

Dans ce cadre, nous avons structuré notre travail comme suit :

Dans le premier chapitre, consacré à la notion de base sur la probabilité et statistiques qui formant le corps de n'importe quel système mathématique, ce chapitre va définir les différents variables et éléments de l'algorithme de Kalman.

Le deuxième chapitre est consacré aux généralités sur la zone d'étude le haut chelif au Nord d'Algérie

Le troisième chapitre est consacré au principe de fonctionnement du filtre de Kalman linéaire, et son algorithme

Dans le dernier chapitre, nous présentons, dans un premier temps, les étapes de l'application du filtre de Kalman linéaire pour la prédiction des apports liquides, par la suite, nous appliquons le filtre de Kalman et donnons les résultats

Nous terminons notre travail par une conclusion générale et les perspectives envisagées.

*Chapitre II :*  
*Notions de base pour*  
*aborder le filtre de*  
*Kalman*

**I.1. Introduction :**

La théorie des probabilités en mathématiques est l'étude des phénomènes caractérisés par le hasard et l'incertitude. Elle forme avec la statistique les deux sciences du hasard qui sont partie intégrante des mathématiques. Les débuts de l'étude des probabilités correspondent aux premières observations du hasard dans les jeux ou dans les phénomènes climatiques par exemple.

Bien que le calcul de probabilités sur des questions liées au hasard existe depuis longtemps, la formalisation mathématique n'est que récente. Des objets tels que les événements, les mesures de probabilité, les espaces probabilisés ou les variables aléatoires sont centraux dans la théorie. Ils permettent de traduire de manière abstraite les comportements ou des quantités mesurées qui peuvent être supposés aléatoires. En fonction du nombre de valeurs possibles pour le phénomène aléatoire étudié,

**I.2. Loi de probabilité :**

En théorie des probabilités et en statistique, une loi de probabilité décrit le comportement aléatoire d'un phénomène dépendant du hasard.

Une loi de probabilité décrit de manière théorique le caractère aléatoire d'une expérience qui est considérée comme aléatoire. La notion d'« expérience aléatoire » est dégagée pour désigner un processus réel de nature expérimentale, où le hasard intervient, avec des issues possibles bien identifiées . [3]

En théorie des probabilités, une loi de probabilité est une mesure dont la masse totale vaut 1. En particulier, cette mesure vérifie les trois axiomes des probabilités.

Pour  $(\Omega, \mathcal{A})$  un espace mesurable,  $\mathbb{P}$  est une loi de probabilité, mesure de probabilité ou plus simplement probabilité si :

- 1.  $\mathbb{P}$  est une application de  $\mathcal{A}$  dans  $[0,1]$
- 2.  $\mathbb{P}(\Omega) = 1$
- 3.  $\mathbb{P}$  est  $\sigma$  – additive , pour toute famille finie ou dénombrable d'éléments disjoints  $(\mathcal{A}_i , i \in I)$  de  $\mathcal{A}$ :

$$\mathbb{P} \left( \bigcup_{i \in I} \mathcal{A}_i \right) = \sum_{i \in I} \mathbb{P}(\mathcal{A}_i) \dots \dots \dots (I - 01)$$

Une conséquence immédiate est :  $\mathbb{P}(\emptyset) = 0$

Le triplet  $(\Omega, \mathcal{A}, \mathbb{P})$  est appelé espace probabilisé. Une loi de probabilité est également appelée distribution de probabilité pour une étude plus appliquée

Une manière usuelle d'expression d'une loi est l'utilisation d'une variable aléatoire puisque, pour toute loi de probabilité  $\mathbb{P}$  sur  $(\Omega, \mathcal{A})$  il existe<sup>14</sup> une variable aléatoire X définie sur un espace probabilisé ) et de loi  $\mathbb{P}$ .

**I.3. Variable aléatoire réelle :**

Une variable aléatoire réelle est une variable aléatoire à valeurs dans  $\mathbb{R}$  ou une partie de  $\mathbb{R}$  ; c'est une fonction définie depuis l'ensemble des résultats possibles d'une expérience aléatoire, dont on doit pouvoir déterminer la probabilité qu'elle prenne une valeur donnée ou un ensemble donné de valeurs. Les variables aléatoires réelles sont les variables aléatoires les plus couramment étudiées.

Les variables aléatoires sont très utilisées en théorie des probabilités et en statistiques. Dans les applications, les variables aléatoires sont utilisées pour modéliser le résultat d'un mécanisme non-déterministe ou encore comme le résultat d'une expérience non-déterministe qui génère un résultat aléatoire.

Selon sa définition mathématique , la loi de X est la mesure de probabilité  $\mathbb{P}_x$  définie par la relation  $\mathbb{P}_x = \mathbb{P}(X \in \mathcal{A}) = \mathbb{P}(X^{-1}(\mathcal{A})) \dots \dots \dots (I - 02)$

**I.4. Loi de probabilité à plusieurs variables :**

Dans certains problèmes interviennent simultanément plusieurs variables aléatoires. Mis à part les cas particuliers 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, certaines d'entre elles généralisant les notions relatives à une seule variable, est simplifiée de deux manières :

1. Seules les variables continues sont considérées.
2. Pour éviter la lourdeur des formules, l'exposé est limité à deux variables.

La probabilité pour que la variable aléatoire X prenne une valeur numérique inférieure à x alors que Y prend une valeur inférieure à y définit la fonction de répartition :

$$F(x, y) = \mathbb{P}(X < x, Y < y) \dots \dots \dots (I - 03)$$

**I.5. Le 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  $\mathbb{R}^n$ . [11] :

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

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

$$\int_{\mathbb{R}^n} f(x_1, x_2, x_3 \dots \dots x_n) dx_1 \dots dx_n = 1 \dots \dots \dots (I - 04)$$

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

$$\mathbb{P}(X_1 \leq a_1, \dots, X_n \leq a_n) =$$

$$\int_{]-\infty, a_1] \dots ]-\infty, a_n]} f(x_1, x_2, x_3, \dots, x_n) dx_1 \dots dx_n = 1 \dots \dots \dots (I - 05)$$

**I.6. Variables aléatoires discrètes :**

Une variable aléatoire est dite discrète lorsque l'ensemble des valeurs qu'elle peut prendre est fini ou infini dénombrable

Soit  $(\Omega, \mathcal{F}, \mathbb{P})$  un espace probabilisé et  $(E, \mathcal{E})$  un espace mesurable. Une variable aléatoire de  $\Omega$  vers  $E$  est dite discrète s'il existe un ensemble fini ou dénombrable  $S \subset E$  tel que  $\mathbb{P}(X \in S) = 1$ .

**I.7. Fonction de répartition :**

la fonction de répartition d'une variable aléatoire réelle caractérise la loi de probabilité de cette variable aléatoire réelle. La fonction de répartition de la variable aléatoire réelle  $X$  est la fonction  $F_X$  qui à tout réel  $x$  associe .

$$F(x) = \mathbb{P}(X < x) \dots \dots \dots (I - 06)$$

La probabilité que  $x \in ]a, b]$  est donc :

$$\text{si } a < b : \mathbb{P}(a < X \leq b) = F_x(b) - F_x(a) \dots \dots \dots (I - 07)$$



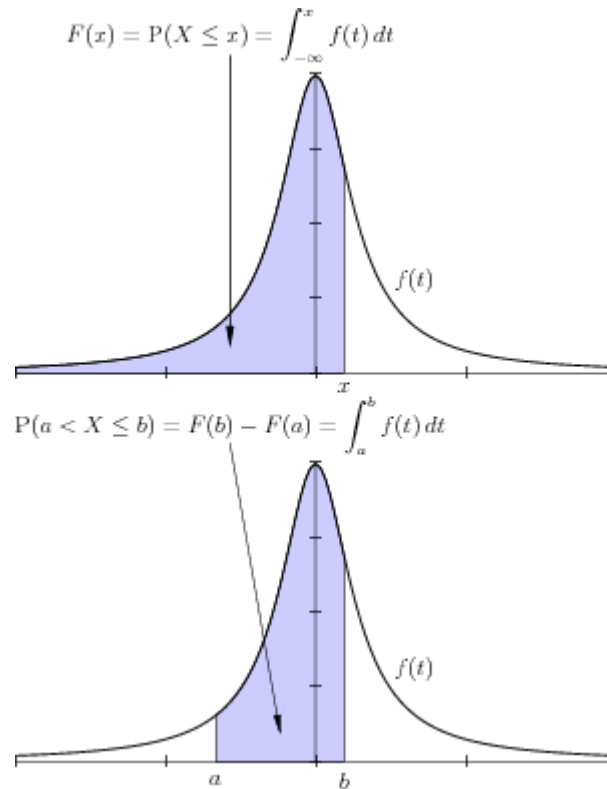


Fig I.1 : Graphique de la fonction de répartition.

**I.8. Rappel statistique et méthodes d'estimation :**

**I.8.1. Paramètres de l'échantillon :**

**A. Paramètre de tendance centrale :**

En statistiques, la tendance centrale désigne le point autour duquel se regroupent à un certain degré les différentes valeurs d'une variable quantitative mesurées dans un échantillon ou issues d'une population, c'est-à-dire le « centre » autour duquel ces valeurs « tendent » à se concentrer. On parle aussi de la position de cet ensemble de valeurs par opposition à la dispersion statistique, qui en reflète la variabilité. Il existe plusieurs mesures de la tendance centrale dites aussi critères de position ; les plus connues sont la moyenne, la médiane ou le mode.

- La moyenne

La moyenne représente la mesure la plus courante de tendance centrale des observations. Elle se calcule en additionnant les valeurs observées de chaque individu divisé par le nombre de sujets observés  $n$  :

$$\bar{X} = \frac{\sum X}{n} \dots \dots \dots (I - 08)$$

Moyenne quadratique :

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

$$m_2 = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2} = \sqrt{\sum_{i=1}^K f_i x_i^2} \dots \dots \dots (I - 09)$$

Moyenne harmonique :

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

$$m_{-1} = \frac{N}{\sum_{i=1}^N \frac{1}{x_i}} = \frac{1}{\sum_{i=1}^K \frac{f_i}{x_i}} \dots \dots \dots (I - 10)$$

Moyenne géométrique :

On ne peut définir cette moyenne que lorsque les observations  $x_1, \dots, x_N$  sont tous 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 :

$$M_g = \sqrt[N]{x_1 x_2 \dots x_N} = \sqrt[N]{x_1^{n_1} \dots x_K^{n_K}} = x_1^{f_1} \dots x_K^{f_K} \dots \dots \dots (II - 11)$$

- La médiane :

La médiane correspond à l'observation du milieu, c'est-à-dire la valeur de part et d'autre de laquelle se situe la moitié des observations. Pour la mesurer, on établit la liste des observations individuelles par ordre croissant ou décroissant. La position de la médiane se calcule de manière suivante :

$n$  = nombre de sujets observés

Si  $n$  impair :  $\left(\frac{n+1}{2}\right)$  *ème valeur* ,

Si  $n$  pair : moyenne arithmétique des  $\left(\frac{n}{2}\right)$  *et*  $\left(\frac{n}{2} + 1\right)$  *ème valeurs*

- Le mode :

Le mode d'une série d'observations est la valeur la plus fréquente d'un ensemble de données. Le mode est rarement employé seul pour mesurer la tendance centrale, parce qu'avec un petit nombre d'observations, chaque valeur est unique. Dans ce cas, il n'y a pas de mode.

**A.2 Paramètre de dispersion :**

- L'étendue :

L'étendue  $e_X$  mesure l'écart entre la valeur la plus élevée et la plus petite.

$$e_X = \max_{1 \leq i \leq N} x_i - \min_{1 \leq i \leq N} x_i \dots \dots \dots (I - 12)$$

Bien qu'elle soit facile à calculer et à comprendre, cette mesure de dispersion présente un certain nombre de désavantages.

- la valeur de l'étendue a tendance à augmenter avec la taille de l'échantillon.
- Ensuite, puisqu'elle est calculée à partir des valeurs extrêmes, sa valeur peut être étendue même si les valeurs de la majorité des observations sont proches les unes des autres.
- De plus, elle est inutilisable pour les tests statistiques. On préférera donc une mesure qui tienne compte de la totalité des observations.

- La variance et la déviation standard (l'écart type)

Ces deux mesures représentent les mesures de variabilité les plus courantes et les plus instructives. Elles mesurent la dispersion de chaque observation autour de la moyenne.

$$Var(X) = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2 \dots \dots \dots (I - 13)$$

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

$$\sigma_X = \sqrt{Var(X)} \dots \dots \dots (I - 14)$$

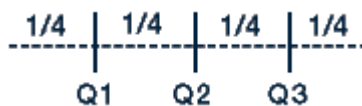
- Coefficient de variation :

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

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

Le coefficient de variation est un nombre sans dimension qui permet de comparer deux variables statistiques de natures différentes.

L'idée est de diviser les observations, classées en ordre croissant, en 4 groupes égaux et de mesurer la distance entre les quartiles extrêmes (Q1 et Q3).



Q1 = observation telle que 25% des sujets lui sont inférieurs.

Q2 = médiane.

Q3 = observation telle que 75% des sujets lui sont inférieurs.

$$EIQ = Q3 - Q1 \dots \dots \dots (I - 16)$$

**B. Paramètre de forme :**

Un paramètre de forme est un paramètre d'une loi de probabilité qui n'est ni un paramètre de position ni un paramètre d'échelle. Un tel paramètre régit uniquement la forme de la distribution.

**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 :

$$P = \frac{\bar{X} - M_0}{s(X)} \dots \dots \dots (II - 17)$$

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

$$Y = \frac{Q_1 + Q_3 - 2M_e}{2(Q_3 - Q_1)} \dots \dots \dots (I - 18)$$

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 :

$$F = \frac{\mu_3}{\frac{\mu_2^2}{3}} = \frac{\mu_3}{s^3(X)} \dots \dots \dots (I - 19)$$

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.9. Test relatif à l'échantillonnage :**

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

Le test de Wilcoxon-Mann-Whitney (ou test U de Mann-Whitney ou encore test de la somme des rangs de Wilcoxon) est un test statistique non paramétrique qui permet de tester l'hypothèse selon laquelle la distribution des données est la même dans deux groupes. Il a été proposé par Frank Wilcoxon en 1945 et par Henry Mann et Donald Ransom Whitney en 1947.

L'énorme avantage de ce test est sa simplicité, même si de ce fait son utilisation est limitée. Comme tous les tests statistiques, il consiste à partir de ce qui est observé à mettre en évidence un évènement dont on connaît la loi de probabilité (au moins sa forme asymptotique). La valeur obtenue, si elle est peu probable selon cette loi, suggèrera de rejeter l'hypothèse nulle.

On considère deux populations X et Y de taille respective  $n_x, n_y$  n suppose les observations indépendantes et disposant d'une relation d'ordre. On souhaite tester l'hypothèse suivante :

H0 : la probabilité qu'une observation de la population X soit supérieure à une observation de la population Y est égale à la probabilité qu'une observation de la population Y soit supérieure à une observation de la population Y X :  $P(X > Y) = P(Y > X)$ .

En général l'hypothèse plus forte « les deux distributions sont égales » est utilisée.

Si nous ordonnons les  $n_x + n_y$  éléments de  $X \cup Y$  par ordre croissant, nous pouvons définir, pour chaque individu, son rang dans la séquence ainsi formée. Soit  $S_x$  la somme des  $n_x$  rangs des éléments de X.

On montre que, sous H0, l'évènement  $S_x = t$  suit une distribution connue, tabulée pour de petits échantillons et qui peut être approchée par une loi de probabilité gaussienne de moyenne

$$E = \frac{n_x(n_x + n_y + 1)}{2} \dots \dots \dots (I - 20)$$

$$V = \frac{n_x n_y (n_x + n_y + 1)}{12} \dots \dots \dots (I - 21)$$

pour des échantillons de taille supérieure à 20 environ.

Le test est construit en confrontant la valeur effectivement obtenue à cette moyenne et cet écart type : on peut ainsi estimer la probabilité de cette valeur sous l'hypothèse nulle et ainsi décider ou non de rejeter cette hypothèse nulle.

On calculera la valeur  $:\epsilon = \left| \frac{S_x - E}{\sqrt{V}} \right|$  qui, si elle est supérieure à 1,96 (risque de 5 %), permet de rejeter l'hypothèse nulle H0 d'égalité des deux échantillons.

**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 tests pour des échantillons indépendants. La procédure nécessite une organisation des données similaire à celle utilisées pour un test 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 position (moyennes) entre les deux échantillons.

**C. Test de Médiane (Test de Mood).**

Principe :

- ✓ 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  $T_s$  et  $N_s$  tel que :

$N_s$  : Nombre total de séries de + ou de - ;

$T_s$  : Taille de la plus grande série de + ou de - ;

$$N_s > \frac{1}{2} (n + 1 - u_{1-\alpha/2} \sqrt{n + 1}) \dots\dots (I-22)$$

$$T_s < 3.3(\log n + 1) \dots\dots\dots (I-23)$$

**D. 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.

### E. 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 porte 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.

## I.10. Estimation stochastique linéaire :

### A. Processus aléatoire :

De nombreux domaines utilisent des observations en fonction du temps

Un processus aléatoire généralise la notion de variable aléatoire utilisée en statistiques élémentaires. On le définit comme une famille de variables aléatoires  $X(t)$ , associées à toutes les valeurs  $t \in T$ . L'ensemble des observations disponibles  $x(t)$ , constitue une réalisation du processus.

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)  $F$  :

$$x(t): S \rightarrow F$$

$$\omega \mapsto \{x(t, \omega), t \in T\}$$

à chaque événement élémentaire fait correspondre une fonction du temps;

- Une collection de variables aléatoires :

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

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

### B. Bruit blanc :

Par analogie avec la lumière blanche qui mélange toutes les fréquences lumineuses, un bruit blanc est un processus stochastique qui possède la même densité spectrale de puissance à toutes les fréquences. Ceci correspond à une autocorrélation nulle en tout point sauf à l'origine : le processus est décorréolé. S'il est gaussien, cette décorrélation entraîne l'indépendance.

**I.11. 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. [11]

Les modèles espace-état intègrent la distinction entre les variables observées (le signal) et les variables cachées (l'état interne). Ils sont constitués :

- D'une ou plusieurs équation(s) de mesure (équation (1) dans I) décrivant la manière dont les variables observées sont générées par les variables cachées et les résidus.
- D'une ou plusieurs équation(s) d'état (équation (2) dans I) décrivant la manière dont les variables cachées sont générées à partir de leur retard et d'innovations.

La terminologie suivante est utilisée :

- $Y_t$  est appelé observation ou variable de mesure ;
- $Z_t$  est la variable d'état à la date  $t$  ;
- $\varepsilon_t$  est le vecteur des innovations à la date  $t$  ;
- $\eta_t$  est le vecteur des erreurs de mesures à la date  $t$  ;
- $A_t$  est la matrice de transition ;
- $C_t$  est la matrice de mesure ;
- $X_{1,t}, X_{2,t}$  sont des variables exogènes, prédéterminées ;
- $C_t Z_t$  est le signal à la date  $t$ .

Soit un processus multidimensionnel  $Y_t$ , on appelle **modèle espace-état** de ce processus, le système (I) décrit par les équations matricielles (1) et (2) de système **(I)** :

$$\begin{cases} (1): Z_{t+1} = A_t Z_t + B_t X_{1,t} + \varepsilon_t \\ (2): Y_t = C_t Z_t + D_t X_{2,t} + \eta_t \end{cases} \text{ où } \begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix} \sim NID \left( 0, \begin{pmatrix} Q_t & S_t \\ S_t & R_t \end{pmatrix} \right) \quad (I) \quad \dots \dots \dots (I - 01)$$

Où les matrices  $A_t, C_t$  sont de taille  $K \times K$  et  $n \times K$ ,  $B_t$  et  $D_t$  sont des matrices déterministes de taille  $K_1 \times K$  et  $K_2 \times K$  et  $Z_0$  est un vecteur aléatoire .

**I.11. Problème du filtrage :**

Le filtrage est une opération fondamentale, 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'information 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.

Dans ce travail, on se restreindra à des filtres linéaires, c'est-à-dire dont l'information extraite est une combinaison linéaire des observations exploitées. Ceci entraîne en particulier des modèles de filtres très simples, avec des propriétés séduisantes, et permet d'utiliser l'arsenal puissant de l'algèbre linéaire pour formaliser ces modèles.

Le terme optimal doit se référer à un critère. Dans ce cadre du filtrage linéaire optimal, le critère usuel est l'erreur quadratique moyenne. Par conséquent, le filtrage linéaire optimal a pour objectif de rechercher le "meilleur" filtre linéaire, c'est-à-dire celui qui fournit une approximation telle que l'erreur quadratique moyenne est minimale. [6]

### **I.12. Conclusion :**

Dans ce chapitre, nous avons introduits les notions mathématiques élémentaires fondamentales couramment utilisées en descriptions des modèles numériques requises, en précision pour la compréhension du filtrage de Kalman ce dernier faisant l'objet principal de notre travail dans le cadre de la présente thèse.

A cet effet, nous avons présenté quelques concepts sur lesquels reposent les fondements mathématiques du FK, ainsi que quelques outils mathématiques indispensables pour analyser les processus stochastiques et leur transmission dans les systèmes dynamiques linéaires. Nous avons également introduit la notion de bruit blanc, qui est entièrement caractérisé par sa fonction d'autocorrélation, et qui présente l'une des hypothèses de base du filtre de Kalman. Par la même occasion, nous avons évoqué la notion des modèles espace-état, et filtrage qui sont à la base de la formulation du filtre de Kalman.

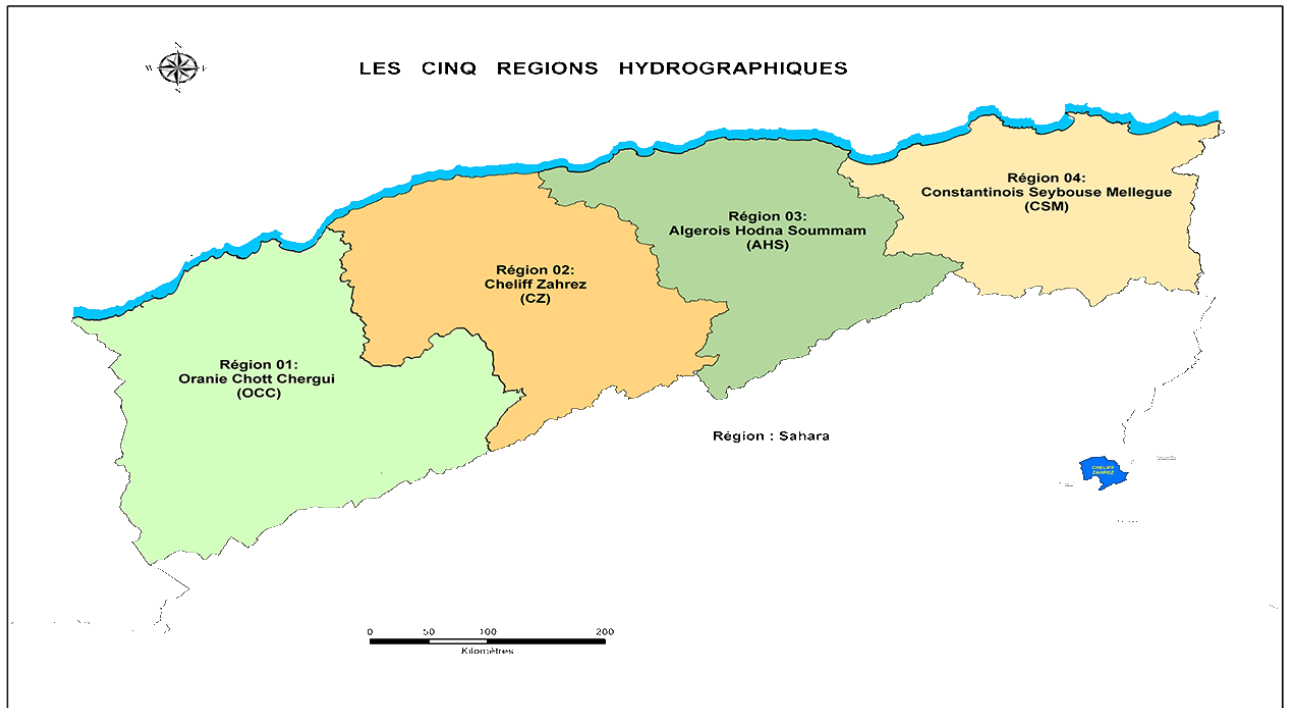
## *Chapitre II :*

### *Cadre de l'étude*

### II.1. Introduction :

Le Bassin Hydrographique Cheliff-Zahrez couvre une superficie d'environ 56 227 km<sup>2</sup> (soit plus de 22% de la superficie de l'Algérie du Nord).

Une région hydrographique limitée naturellement au Nord par la Mer Méditerranée, à l'Ouest par la région Oranie – Chott Chergui, à l'Est par la région Algéroise – Hodna – Soummam et au sud par le Sahara. [1]



*Figure II.1 Représentation des cinq régions hydrographiques en Algérie*

Toutes les infrastructures de mobilisation existantes sont situées dans le bassin du Cheliff qui est le bassin principal de la région avec une superficie de 43750 Km<sup>2</sup> de la région et renferme, presque la totalité des ressources en eau de surface.

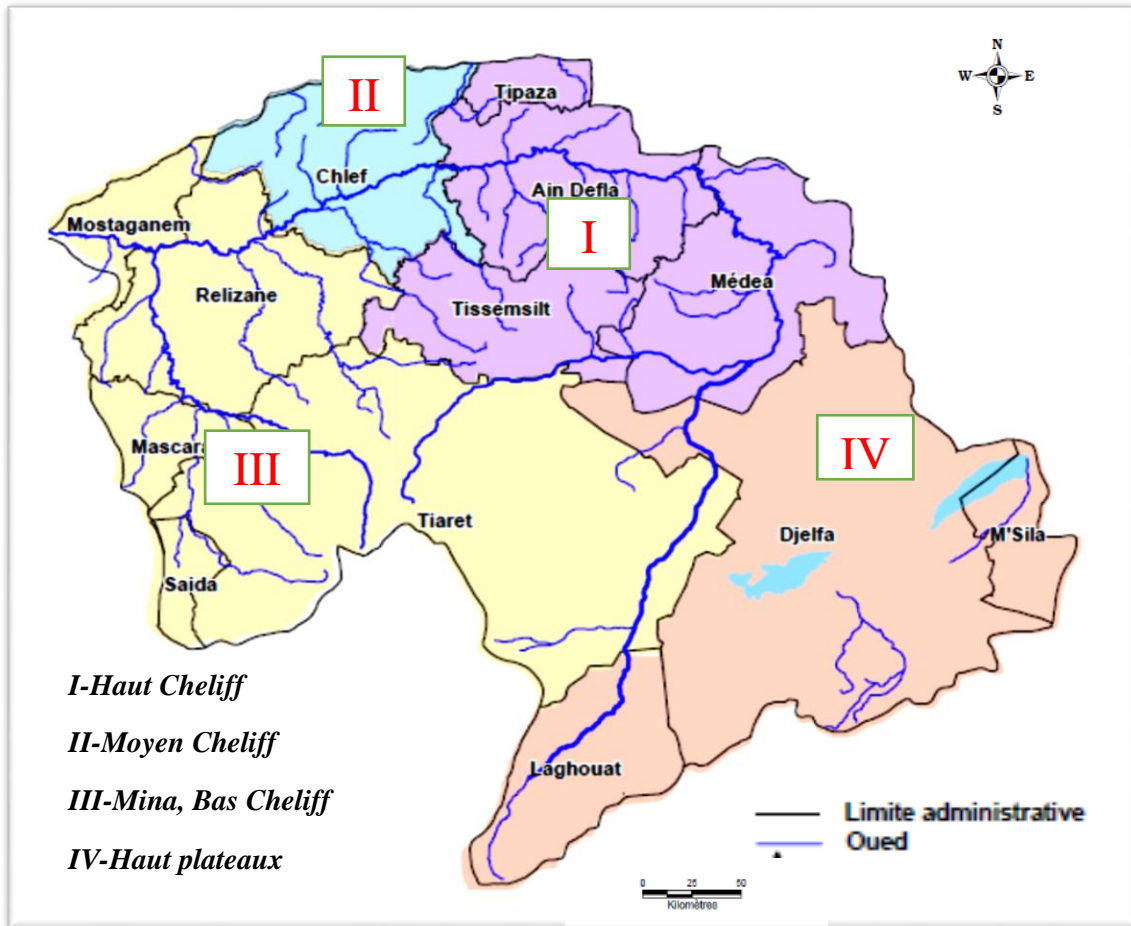


Figure II.2 Représentation du découpage par sous-région Du bassin Cheliff [1]

**II.2.Présentation de la sous-région d'étude Haut Cheliff :**

Cette région se situe dans le centre Ouest du Nord Algérien ; elle est limitée au Nord par la mer Méditerranée, à l'Est par le bassin hydrographique Algerois-Hodna-Soumam, à l'Ouest par les wilayas de Chlef, Relizane et Tiaret , et au Sud par Djelfa . Elle est composée de 04 wilaya (Ain Defla, Tipaza, Medea et Tissemsilt), soit 90 communes.

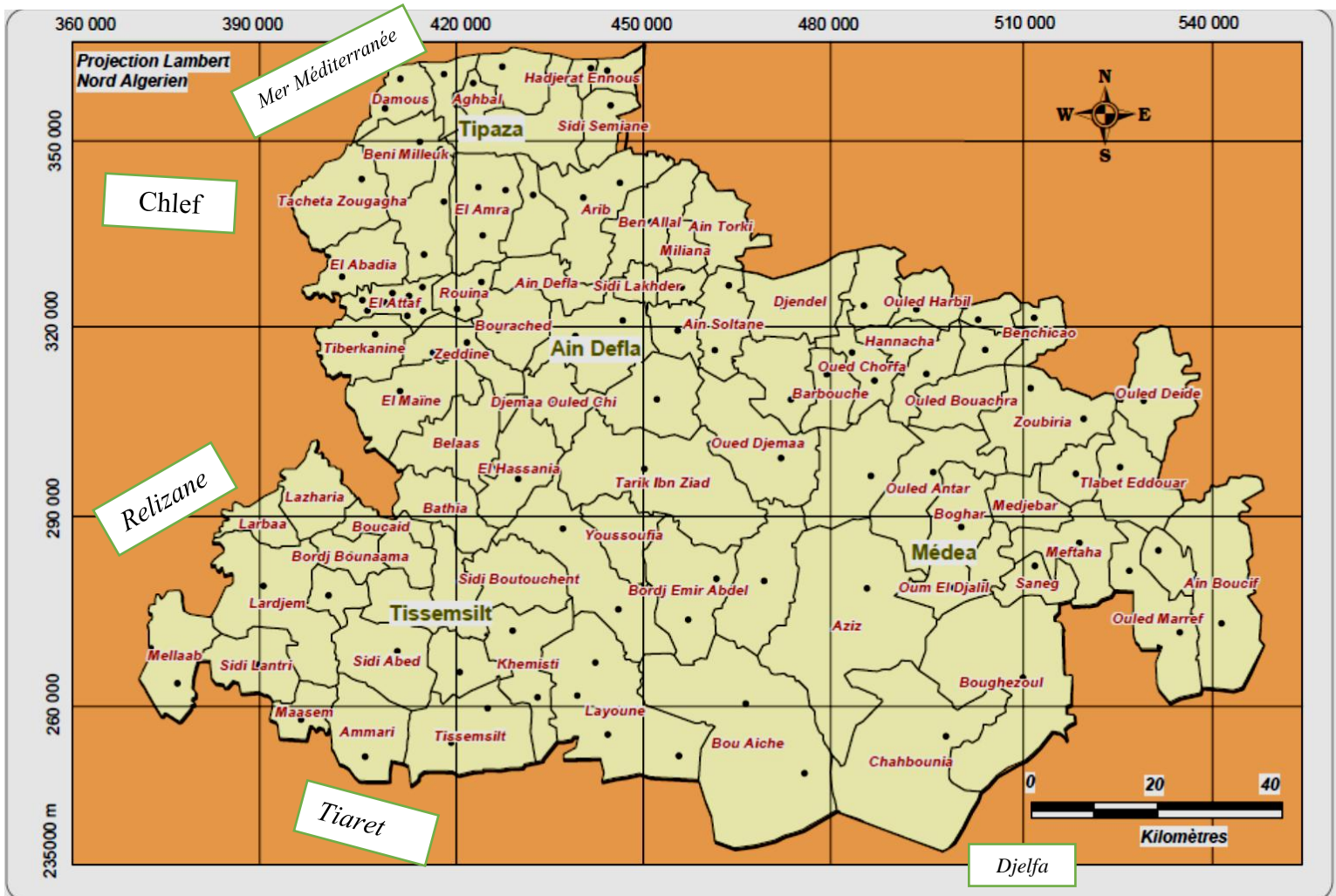


Figure II.3: Découpage administratif de la sous-région n°1.

**II.3.Caractéristiques morpho métriques de la région[13]:**

**II.3.1 Superficie, Altitude et forme :**

Les paramètres morpho-métriques jouent un rôle essentiel dans la répartition des eaux à l'échelle du bassin versant. Ils influent fortement sur l'hydrologie de ce dernier. Plusieurs caractéristiques hydromorphométriques ont été déterminées à savoir :

- Surface du bassin versant " S " en Km<sup>2</sup>
- Périmètre du bassin versant " P " en Km.
- Indice de compacité de Graveluis " Kc '' .

La forme peut être traduite par l'indice de compacité de « Graveluis » Kc,a également une influence certaine sur l'écoulement ; elle détermine, dans une certaine mesure, l'allure de l'hydrogramme de crue. Un bassin très allongé ne réagira pas, toutes choses égales d'ailleurs, de la même manière qu'un bassin de forme ramassée .

Kc est calculé à l'aide de l'expression suivante :

$$Kc = \frac{P}{2\sqrt{\Pi A}} \dots\dots\dots (II-01)$$

Plus que (Kc) approche de l'unité, plus le bassin serait ramassé et le temps de concentration serait plus court.

Les valeurs de l'indice de compacité montrent que tous le bassin est allongé.

**II.3.2. Rectangle équivalent :**

C'est une transformation purement géométrique du bassin versant en un rectangle de dimensions **L** et **l**. Ce dernier permet de comparer les bassins versants entre eux du point de vue de l'écoulement.

Les valeurs de la longueur et de la largeur du bassin sont définies d'après les formules

suivantes : 
$$L = Kc \cdot \frac{\sqrt{S}}{1.128} \left( 1 + \sqrt{1 - \left( \frac{1.128}{Kc} \right)^2} \right) \dots\dots\dots (II-02)$$

$$l = Kc \cdot \frac{\sqrt{S}}{1.128} \left( 1 - \sqrt{1 - \left( \frac{1.128}{Kc} \right)^2} \right) \dots\dots\dots (II-03)$$

**II.3.3. Relief (indice de pente) :**

Le joue un rôle prépondérant dans l'aptitude au ruissellement des terrains. Son appréhension peut être faite à l'aide de l'indice de pente global (Ig).

L'indice de pente globale se calcule d'après la formule suivante :

$$Ig = \frac{D}{L} ; (m/km) \dots\dots\dots (I-04)$$

Ig ; L'indice de pente globale

D : dénivelée simple ou bien la distance verticale qui, exprimée en mètres, sépare les altitudes ayant 5% de la surface du bassin au-dessous d'elles ; elle peut être assimilée à la distance séparant les deux courbes de niveau maîtresses extrêmes qui délimitent le bassin.

L longueur du rectangle équivalent telle que définit par [**M. Roche**] (km)

Pour permettre une comparaison plus poussée des bassins entre eux du point de vue relief, on détermine le dénivelée spécifique (Ds) qui est déduite de l'indice de pente globale.

$$Ds = Ig \cdot \sqrt{A} ; (m) \dots\dots\dots (II-05)$$

La superficie du bassin versant en Km<sup>2</sup>

**Tableau II.1 :** Classes de relief en fonction de dénivelée spécifique.

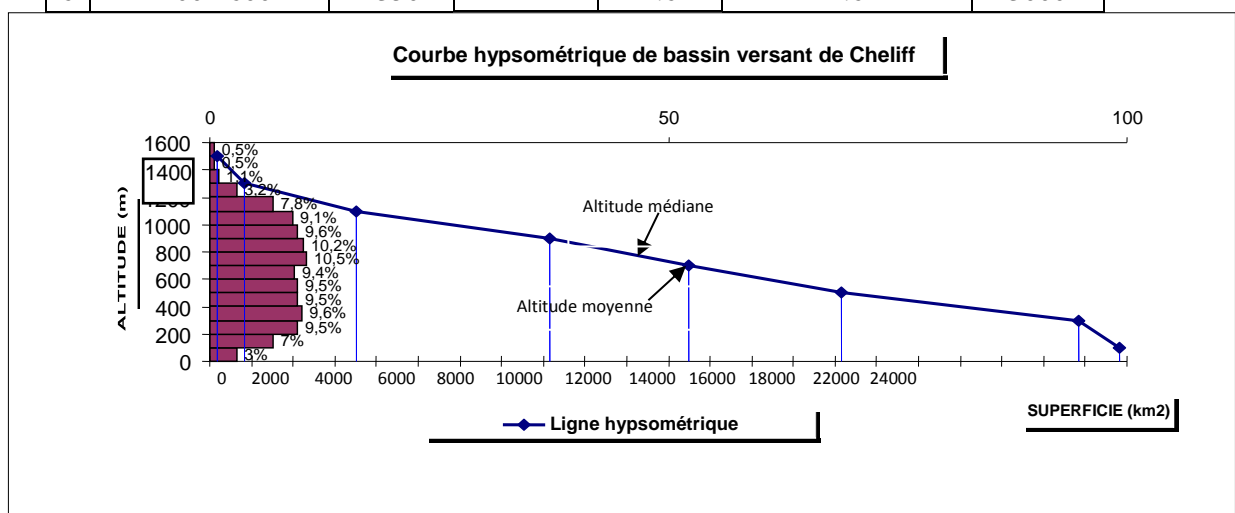
Ds (m)	Classe de relief	Désignation
Ds<10	R1	Relief très faible
10<Ds<25	R2	Relief faible
25<Ds<50	R3	Relief assez faible
50<Ds<100	R4	Relief modéré
100<Ds<250	R5	Relief assez fort
250<Ds<500	R6	Relief fort
Ds>500	R7	Relief très fort

**II.3.4. Répartition et courbe hypsométrique :**

Elle est d'une grande importance dans l'étude du relief d'un bassin versant. Les surfaces partielles par tranche d'altitude sont données dans le Tableau suivant. La courbe hypsométrique est donnée par la figure ci-dessous.

**Tableau II.2** Classes de relief en fonction de dénivelée spécifique.

N°	Altitudes H <sub>i, i+1</sub> (m)	H <sub>i</sub> moy (m)	Surfaces		Surfaces cumulées S <sub>i</sub> (%)	H <sub>i</sub> S <sub>i</sub>
			H <sub>i</sub> (km <sup>2</sup> )	S <sub>i</sub> (%)		
1	200-400	300	80	46.4	100	24000
2	400-600	500	44	27.3	53.5	22000
3	600-800	700	21	14.3	26.3	15000
4	800-1000	900	12	7.7	11.9	10500
5	1000-1200	1100	6	2.6	4.24	6600
6	1200-1600	1350	2	1.6	1.6	3000



**Figure.II.4 :** La courbe hypsométrique du bassin versant de Cheliff « le haut Cheliff »

Sur la courbe hypsométrique, l'altitude médiane ou H<sub>50%</sub> a été déterminée, et l'altitude moyenne calculée par la formule suivante :

$$\bar{H} = \frac{\sum S_i \cdot H_i}{\sum S_i} \quad \text{avec} \quad H_i = \frac{H_{n-1} + H_n}{2} \quad \dots \dots \dots \quad (II-06)$$

H<sub>i</sub> : Altitude moyenne entre deux courbes de niveaux consécutives n-1 et n en (m).

Si = Surface partielle comprise entre deux courbes de niveau consécutives n-1 et n en (Km<sup>2</sup>)

**II.3.4. Indice de pente globale " Ig" :**

Il est défini comme étant le rapport entre l'altitude comprise entre 5% et 95% de la surface du bassin et la longueur du rectangle équivalent.

Il est donné par la formule suivante :  $Ig = \frac{H_{5\%} - H_{95\%}}{L}$  ..... (II-07)

**II.3.5. Densité de drainage :**

Elle est définie comme étant le rapport entre la longueur totale de tous les talwegs "Li"

du bassin versant, à la surface totale " S " :  $Dd = \frac{\sum_{i=1}^n Li}{S}$  ;(Km/Km<sup>2</sup>). ..... (II-I08)

**II.3.6 Temps de concentration du bassin versant :**

C'est le temps que met la particule d'eau la plus éloignée pour arriver à l'exutoire. Il peut être notamment calculé par formule soit :

$tc = \frac{4\sqrt{S} + 1.5Lcp}{0.8\sqrt{Hmoy - H min}}$  ..... (II-09)

S : surface du bassin versant en Km<sup>2</sup>.

L<sub>cp</sub> : longueur du cours d'eau principal en Km. (L<sub>cp</sub>=180km)

Te : temps de concentration en heure.

H<sub>moy</sub> : altitude moyenne en m.

H<sub>min</sub> : altitude minimale en m.

Les résultats sont récapitulés dans le tableau suivant :

**Tableau II.3 :** Caractéristiques hydrométriques du bassin versant « le haut Cheliff ».

Caractéristiques	Unité	Valeur
Surface	Km <sup>2</sup>	1427
Périmètre	Km	219
Altitude Maximale	M	1600
Altitude moyenne	M	640
Altitude médiane	M	720
Altitude minimale	M	200
Altitude 5 %	%	1300
Altitude 95%	M	240
Indice de compacité	/	1.62
Indice de pente globale	m/Km	11.3
Longueur du rectangle équivalent	Km	94
Largeur du rectangle équivalent	Km	15.4
Longueur du cours d'eau principal	Km	180
Densité de drainage	Km/Km <sup>2</sup>	0.59
Temps de concentration	H	25



**II.3.7. Caractéristiques morphologiques et hydrologiques des sous- bassins versants :**

Nous donnons les caractéristiques morphologiques et hydrologiques de la totalité des sous bassins. [11]

Bassin : « le haut Cheliff »													
Code	sous bassin	Superficie	Périmètre	Indice de compacité	Longueur équivalente	Largeur équivalente	Ig	Ds	Classe de relief	Altitude	Pentes	Sédiments transportés	Ouvrage à l'aval
		(km <sup>2</sup> )	(Km)		(Km)	(Km)		(m/km)		(m)			
113	Oued Cheliff Djellil	992,85	179,96	1,6	77,1	12,88	16,91	532,67	très fort	1150	Forte à moyenne	Fines et paillettes	Non
114	Oued Cheliff Ghrif	1383,69	176,81	1,33	68,08	20,32	17,67	657,11	très fort	460	Forte à moyenne	Fines, sable grossier et galets	retenue
115	Oued Cheliff Harbil	767,29	167,95	1,7	73,54	10,43	15,01	415,86	Fort	320-935	Forte en aval moyen en amont	Fines et sable grossier	Non
116	Oued Deurdeur	857,64	185,65	1,78	82,42	10,41	19,41	568,51	Très fort	447-1160	Forte à moyenne	Fines et paillettes	retenue
117	Oued Cheliff Harraza	743,97	142,68	1,46	58,66	12,6	23,87	650,97	Très fort	230-715	Forte à moyenne	Fines et sable grossier	retenue
118	Oued Ebda	680,55	141,1	1,51	59,02	11,53	23,73	619,02	Très fort	270-850	Forte à moyenne	Fines et sable grossier	Non

**II.4. Caractéristiques du réseau pluviométrique :**

La région hydrographique Cheliff-Zahrez est équipée d'un réseau pluviométrique de 158 postes [12]

La sous-région 01 est équipée de 32 postes pluviométriques soit une densité de 200 Km par poste, bien réparties, celles-ci sont gérées par L'ANRH. Les précipitations sont très irrégulières dans le temps et dans l'espace ; on distingue deux zones extrêmes, une zone pluvieuse avec une moyenne annuelle de 845 mm (sidi medjahed ; côtier) et 564 mm à la station pluviométrique (Thniet El Had, OUARSENIS) ; et une autre zone (La plaine incluse entre les deux zones) moins pluvieuse avec une moyenne annuelle, d'environ, 453mm à la station pluviométrique (Bordj Emir Khaled). (Source : Agence de Bassin Hydrographique Cheliff-Zahrez) [12]

Nous avons choisi quelques stations pluviométriques de la région d'étude. Le tableau représente ces stations avec leurs codes et leurs coordonnées Lambert nord d'Algérie.

**Tableau II.4** Les Stations pluviométriques retenues. [11]

Les sous bassin versants	Code	Stations	X	y	Z
Oued Cheliff Djellil	01-13-02	Derrag	472	290	1150
Oued Cheliff Ghrib	01-14-05	Bge Ghrib	487	318	460
Oued Cheliff Harbil	01-15-01	Tamazguida	497	324	490
	01-15-09	Secteur Médéa	505	330	935
	01-15-10	Domaine Mes bah (Ferroukhi)	467	322	313
	01-15-12	Ain sultane Pépinière	460	320	307
	01-15-14	Djenane ben Ouadhah	476	325	320
Oued Deurdeur	01-16-03	Bordj Emir AEK	461	285	1080
	01-16-05	Teniet El Had	460	295	1160
	01-16-06	Sidi Mokhrifi	462	289	447
	01-16-07	Tarik Ibn Ziad	254	275	650
	01-16-09	Bge Deurder	459	300	575
Oued Cheliff Harraza	01-17-02	Arib Cheliff	440	333	230
	01-17-03	Bordj Emir Khaled	456	314	370
	01-17-06	Khemis-ITGC	459	328	285
	01-17-07	Miliana	451	339	715
	01-17-11	Sidi Lakhdar	452	330	250
	01-17-15	El Ababsa	444	318	320
Oued Ebda	01-17-18	Bge Herraza	445	321	312
	01-18-01	Arib Ebda	440	326	282
	01-18-03	Sidi Medjahed	447	331	850
	01-18-04	Ain Defla	329	319	270

La densité de couverture en informations pluviométriques est très inégale d'une zone de faibles altitudes (plaines du Haut et Moyen Cheliff où le réseau est plus concentré) à une zone de grandes altitudes (monts de l'Ouarsenis et Dahra où la densité du réseau est moins importante),

### II.5. Répartition géographique des précipitations :

L'étude plus ou moins précise des variations spatiales des précipitations annuelles moyennes passe par l'élaboration d'une carte d'isohyètes. En effet, la connaissance de totaux précipités issus de stations plus ou moins dispersées artificiellement dans la région ne suffit pas, à elle seul, pour cerner la disposition globale dans les zones pluviométriques. Il s'agit donc de passer de cette information ponctuelle à une estimation valable de tranche d'eau moyenne précipitée sur l'ensemble du bassin. [1]

Dans le bassin du Haut et Moyen Cheliff; une concentration des précipitations sur les versants Sud des monts du Dahra et du Zaccar, avec une moyenne interannuelle de plus de 600 mm (stations 011803 Sidi Medja et 011806 El Anneb). Les précipitations se réduisent dans la plaine du Haut et Moyen Cheliff où elles varient entre 300 et 400 mm, et dans le sud du bassin où elles atteindraient plus de 500 mm.

**II.6. Acquisition de données :**

Les données concernant les apports liquides la variable de notre étude ont été recueillies auprès de l'Agence Nationale des Ressources Hydrauliques ANRH de BLIDA pour 5 stations hydrométriques bien reparti sur le bassin versant avec les caractéristiques suivant :

**Tableau II.5 :** Caractéristiques des stations .

	Station	Code	Cordonnées			Période D'observation
			X(km)	Y(km)	Z(m)	
1	EL ABABSA	01-17-15	444	318	320	1975-2013
2	ARIB CHELIF	01-17-02	440	333	230	1979-2013
3	BIR OULED TAHAR	01-19-05				1990-2013
4	DJENANE BEN OUADAH	01-15-14	476	325	320	1990-2013
5	MARABOUT BLANC	01-1601				1990-2012

Source : ANRH Blida 2017

**II.7. Paramètre de l'échantillon :**

Dans cette partie on va donner les paramètres de l'échantillon pour chaque station :

**Tableau II.6 :** les paramètres de l'échantillon

	Station I	Station II	Station III	Station IV	Station V
N	33	36	23	22	33
Minimum	0.220	18.8	1.41	17.0	0.220
Maximum	7.33	263	25.7	225	7.33
Moyenne	2.60	97.9	11.5	72.6	2.59
Ecart type	1.80	68.1	7.01	52.2	1.81
Médiane	2.38	73.7	11.2	65.0	2.38
Cv	0.695	0.696	0.608	0.719	0.699
Cs	0.748	0.886	0.173	1.46	0.726
Ca	2.85	2.54	1.75	4.07	2.81

**II.8. Teste d'homogénéité :**

En appliquant le test de Mood on trouve les résultats suivants :

**Tableau II.7 :** test de Mood.

Station	N	Médiane	Ns	Ts	Ns calculé	Ts calculé	RESULTAT
1	33	2,98	14	7	11,29	8,31	Homogène
2	37	2,98	15	8	12,96	8,48	Homogène
3	23	2,98	14	2	7,2	7,79	Homogène
4	22	2,98	14	4	6,8	7,73	Homogène
5	22	2,98	12	4	6,8	7,73	Homogène

**II.9. Conclusion :**

Au terme de ce chapitre, nous concluons que le bassin versant de haut Chélif fait partie major de bassin de chelif-zahrez de Nord d'Algérie et qui comporte un réseau hydrographique très important.

Cette région est équipée de 32 postes pluviométriques soit une densité de 200 Km par poste, bien réparties.

Les séries chronologiques des apports liquides, couvrant une période d'observation commune de 21 ans (1990 à 2011) ont été investies. Ces dernières sont observées à cinq stations hydrométriques dans le bassin versant du haut Cheliff.

Comme on peut encore constater que la couverture en informations est variable selon l'altitude et la densité de réseau hydrographiques

Ces dernières contribueront significativement à la connaissance de la disponibilité des données pour aborder le modèle sujet d'étude.

# *Chapitre III :*

## *Présentation du filtre de KALMAN*

### III.1. Introduction :

Les techniques d'assimilation des données sont couramment employées dans des domaines aussi différents que l'hydrologie, la médecine ou la physique spatiale.

Le Filtre de Kalman est 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.

L'objectif de ce chapitre est d'introduire les notions, les concepts, les définitions et les principes généraux du filtre de Kalman

### III.2 L'estimateur Optimal Au Filtre De Kalman :

#### III.2.1. L'estimation :

L'estimation est un domaine des mathématiques statistiques qui permet, à partir de mesures effectuées sur un système, d'estimer les valeurs de différents paramètres de ce système

##### L'estimateur :

Un estimateur de paramètre inconnu  $\theta$  d'un modèle ou d'une loi de probabilité est une fonction qui fait correspondre à une suite d'observations  $y_1, y_2, \dots, y_n$  issues du modèle ou de la loi de probabilité, la valeur  $\hat{\theta}$  que l'on nomme estimé ou estimation.

$$\hat{\theta}_n = f(y_1, y_2, \dots, y_n) \dots \dots \dots (III - 01)$$

#### III.2.2. les Estimateurs déterministes :

L'estimateurs déterministe 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 (estimateur optimal).

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. Le biais de l'estimateur noté  $B(\hat{\theta})$  est la différence moyenne entre sa valeur et celle du paramètre qu'il estime. Le biais doit être égal à 0 pour avoir un estimateur parfait (dans ce cas, l'estimateur est dit donc sans biais) [14].

$$B(\hat{\theta}) = E(\hat{\theta}) - \theta \dots \dots \dots (III - 02)$$

**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.

$$Var(\hat{\theta}) = E(\hat{\theta} - E(\hat{\theta}))^2 \dots \dots \dots (III - 03)$$

**III.2.3. L'estimateur optimal :**

On trouve souvent qu'un estimateur dont le biais est nul aura une variance toujours supérieur ou égale à ce que l'on appelle la "**borne de Cramer-Rao**" 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".

**III.2.4. Pourquoi utiliser ce filtre de kalman dans une modélisation hydrologique ?**

D'une manière générale, la méthode de Kalman est une version moderne et probabiliste de la méthode des moindres carrés, qui concerne essentiellement les systèmes linéaires.

C'est une méthode d'estimation qui introduit la notion d'état, dans la représentation du système physique étudié et se base sur la notion de filtrage optimal, d'où son appellation : Filtre de Kalman (FK). C'est un outil mathématique assez performant qui est entrain de jouer un rôle de plus en plus important dans toutes les filières des sciences de l'ingénieur. En effet, Le FK permettant de donner une estimation de l'état de système à partir d'une information a priori sur l'évolution de cet état et de mesures réelles, il est utilisé à des fins de prédiction à court terme et des fins d'identification paramétrique

Son grand avantage, pour nous en tant qu'ingénieurs, c'est qu'on n'est pas obligé d'être un génie en mathématiques pour pouvoir comprendre et utiliser ce Filtre d'une manière efficace.

Cependant, avant de rentrer dans les détails propres du filtre de Kalman, les premières questions qui surgissent sont d'abord qu'est-ce qu'un FK ? Qu'est ce qui justifie le recours à un tel outil, dans le cadre de notre travail ?

Pour répondre à la première question, le FK peut être défini d'une manière simple comme étant un algorithme optimal et récursif de traitement de données.

Dans cette définition, le terme « optimal » peut être approché de différentes façons dépendamment des critères choisis pour évaluer la performance du filtre. Un des aspects de cette optimalité est que le FK incorpore toute l'information qui peut lui être procurée. Il traite la totalité des mesures disponibles, indépendamment de leur précision, pour estimer les valeurs courantes des variables d'intérêt. Ceci est fait en utilisant la connaissance des modèles dynamiques du système et des mesures la description statistique des bruits du système, des erreurs de mesure ainsi que les incertitudes dans les modèles dynamiques n'importe qu'elle information disponible sur les conditions initiales des variables en question. Cela étant, l'aspect principal de cette optimalité est que le FK combine toute cette information disponible pour produire une estimation des variables en question, de telle sorte que l'erreur d'estimation est minimale.

Le terme « récursif » dans la description précédente, veut dire que le FK n'exige pas de garder ou 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.

### **III.2.5. Historique du Filtre de Kalman :**

Rudolf Kalman, né en 1930, poursuivit 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 [5].

En 1960, alors que Kalman visitait la NASA, un des responsables, F.Schmidt comprit la possibilité d'appliquer les résultats de Kalman pour le projet "Apollo" pour l'estimation et la commande de la trajectoire [4] [5]. Schmidt travailla sur le sujet et on peut dire que ce fut la première réelle implantation du filtre.



**III.2.6. Domaines d'utilisation :**

Le filtre de Kalman est un outil de traitement utilisé dans une large gamme de domaines technologiques tels que le traitement du signal, l'automatique le radar et les systèmes de communication. Il est également de plus en plus utilisé en dehors du domaine du traitement du signal, par exemple en météorologie et en océanographie, en finance et en navigation.

Le filtre de Kalman peut également être utilisé dans les situations où nous voulons lisser filtrer ou prédire les états d'un système dynamique et linéaire, suivant la quantité d'informations disponibles [8]

**III.2.7. Le filtre de Kalman :****III.2.7.1. Définition :**

Le filtre de Kalman est un filtre récursif, qui utilise une approche temporelle [8]. 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 réelle.

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 [11].

**III.2.7.2. Principe du Filtre :**

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 [7] :

- 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.

**III.2.7.3 Avantage et Inconvénient de FK :****A. Avantage :**

- ✓ La procuration de l'erreur de prédiction qui présente en soit un indicateur de précision.
- ✓ Son algorithme travaille dans le domaine temporel avec une nature récursive et dispose d'un estimateur optimal dans le sens des moindres carrés.
- ✓ Un autre aspect de son optimalité est l'incorporation de toute l'information disponible sur le système, les mesures et les erreurs, dans un opérateur adaptatif qui est recalé à chaque fois qu'une nouvelle mesure devient disponible.
- ✓ Le gros avantage de la méthode est de fournir à chaque itération une estimation des matrices de covariance d'erreur de mesure et d'analyse. Il faut toutefois initialiser correctement ces matrices à

l'instant ( $t_0$ ), et avoir une estimation des matrices de covariance d'erreur modèle et d'erreur d'observation.

**B. Inconvénient :**

- Le FK a été développé seulement pour les modèles linéaires gaussiens.
- L'hypothèse des bruits gaussiens n'est pas essentielle pour le fonctionnement du filtre de Kalman, ce dernier approche la densité de l'état sachant l'observation (densité conditionnelle) par une densité gaussienne, déterminée par sa moyenne et sa matrice de covariance. La non linéarité du modèle peut entraîner la multi-modalité de la loi conditionnelle de l'état, et ainsi rend le filtre de Kalman inadapté.
- Lorsque le système est fortement non linéaire le filtre de Kalman étendu peut diverger (Divergeant : 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)

**III.2.7.4. Les limites :**

Le filtrage de Kalman est une méthode d'estimation et de prédiction puissante prenant en compte la modélisation du système. Néanmoins, ce filtre n'est pas forcément l'outil à appliquer dans tous les cas [3].

En effet, le système doit être modélisé assez précisément afin d'obtenir 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.

Pour pallier au problème 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 quelques défauts.

- En premier lieu, la covariance de l'erreur ne converge pas obligatoirement. En effet, si le passage d'une estimation à une autre tend à sortir de la zone de linéarité, alors le modèle linéarisé n'est plus valable et il est possible que le filtre diverge.
- Le second défaut est son coût calculatoire plus important. En effet, de nouvelles matrices coûteuses rentrent en jeu, ce qui peut être une limite à son utilisation.

Une autre limite importante d'une telle méthode est que le filtre de Kalman permet de Prendre en compte uniquement un modèle de bruit gaussien. Cette restriction limite donc l'utilisation du filtre de Kalman.

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



L'équation de mesure :

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}_{k+1} = \begin{bmatrix} H_{1,1} & \dots & H_{1,n} \\ H_{2,1} & \dots & \cdot \\ \vdots & \vdots & \vdots \\ H_{n,1} & \dots & H_{n,n} \end{bmatrix}_k \begin{bmatrix} X_1 \\ \cdot \\ \cdot \\ X_n \end{bmatrix}_k + \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_n \end{bmatrix}_k$$

Vecteur de

Matrice de mesure

Vecteur des bruits de mesure

**III.4. Fonctionnement du filtre de Kalman :**

L'estimation optimale consiste à trouver la meilleure estimation  $\hat{x}_k$  de l'état  $x_k$  en minimisant un critère qui est la variance de l'erreur d'estimation :

$$e_k = x_k - \hat{x}_k$$

Le filtrage de Kalman comporte deux phases distinctes : Prédiction et mise à jour (correction). 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.

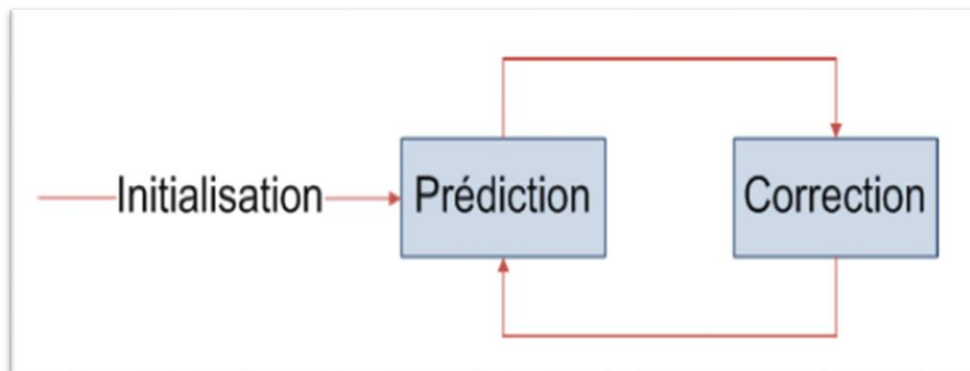


Figure III.1: Les étapes du filtre de Kalman

- **La phase de prédiction :**

On se place à l'instant k. A cet instant, on dispose d'une estimation initiale fondée sur la connaissance du processus et des mesures jusqu'à l'instant précédent, c'est-à-dire k - 1. Cette estimation est appelée estimation à priori.

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

$$e_{k/k-1} = x_k - \hat{x}_{k/k-1}$$

Les équations de la correction sont :

$$\begin{cases} \hat{x}_{k/k-1} = A_{k-1}\hat{x}_{k-1/k-1} + B_{k-1}u_{k-1} & (1) \\ P_{k/k-1} = A_{k-1}P_{k-1/k-1}A_{k-1}^T + Q_{k-1} & (2) \dots\dots\dots(III - 06) \end{cases}$$

• **La phase de mise à jour (correction) :**

Nous allons, maintenant, utiliser la mesure  $y_k$  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 est :

$$e_{k/k} = x_k - \hat{x}_{k/k}$$

Les équations de la phase de correction sont :

$$\begin{cases} \hat{x}_{k/k} = \hat{x}_{k/k-1} + K_k(y_k - H_k\hat{x}_{k/k-1}) & (3) \\ P_{k/k} = (I - K_kH_k) P_{k/k-1} & (4) \dots\dots\dots(III - 07) \end{cases}$$

I : est la matrice identité de même taille que  $P_{k/k-1}$

Gain de Kalman optimal :

$$K_k = P_{k/k-1}H_k^T(H_kP_{k/k-1}H_k^T + R_k)^{-1} \quad (5) \dots\dots\dots(III-08)$$

**III.5. Algorithme du filtre de Kalman :**

On base sur les cinq équations fondamentales de filtre (relations III-6 ,7,8), Les différentes étapes du calcul sont les suivantes :

**1) Initialisation (conditions initiales):**

Lors de l'exécution de l'algorithme il est nécessaire d'initialiser les grandeurs suivantes :

- ✓ Le Vecteur d'état initial,
- ✓ La Matrice de covariance d'erreur associée au vecteur d'état initial,
- ✓ La Matrice de covariance des bruits du système,
- ✓ La Matrice de covariance des bruits de mesures,
- ✓ La Matrice de transition d'état,
- ✓ La Matrice de mesure.

2) Calcul de l'estimé  $\hat{x}_{k/k-1}$  et de la matrice de covariance  $P_{k/k-1}$  à partir des mesures disponibles à l'instant  $k-1$  :

$$\hat{x}_{k/k-1} = A_{k-1}\hat{x}_{k-1/k-1} + B_{k-1}u_{k-1}$$

$$P_{k/k-1} = A_{k-1}P_{k-1}A_{k-1}^T + Q_{k-1}$$

..... (III - 09)

3) Calcul du gain de Kalman :

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

..... (III - 10)

4) Mise à jour de l'estimation :

$$\hat{x}_{k/k} = \hat{x}_{k/k-1} + K_k(y_k - H_k\hat{x}_{k/k-1})$$

..... (III - 11)

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

$$P_{k/k} = (I - K_kH_k) P_{k/k-1}$$

..... (III - 12)

6) Préparation du pas suivant :

$$\hat{x}_{k+1/k} = A_k\hat{x}_{k/k} + B_ku_k$$

$$P_{k+1/k} = A_kP_{k/k}A_k^T + Q_k$$

..... (III - 13)

### III.6 Stabilité du filtre :

Un filtre optimal n'est pas automatiquement stable donc 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. 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. 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. [11]

Les problèmes de stabilité du filtre sont reliés à la connaissance exacte des matrices  $P_0, F_k, H_k, Q$  et  $R$  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. A ce moment,  $P_k$  ne mesure plus exactement la variance de l'erreur de l'estimation, et on assiste à une progression de l'erreur de l'estimation du vecteur  $X_k$  dans le temps.

Ce problème peut être dû 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 peut être accentué, quand les termes de bruit dans l'équation d'état du système, ou celle de mesure deviennent trop petits. Dans ce cas, la covariance de l'erreur devient tellement petite que le gain devienne lui aussi petit et par conséquent, les observations qui viennent au fur et à mesure, ont un effet minime sur l'estimation.

### III.6. Conclusion :

Le filtre de Kalman est un outil essentiel pour les ingénieurs pour résoudre les problèmes de filtrage. C'est une version moderne et probabiliste de la méthode des moindres carrés, le filtre de Kalman concerne essentiellement les systèmes linéaires. On considère que le modèle est linéaire, avec un bruit additif Gaussien et l'observation est une fonction linéaire de l'état entachée d'un bruit additif gaussien,

A partir de l'estimateur optimal linéaire gaussien, nous avons trouvé sa forme récursive et nous l'avons adapté afin qu'il puisse estimer un vecteur d'état variant dans le temps. Nous sommes donc arrivés à écrire un filtre de Kalman. Ce filtre nous assure donc que l'estimateur converge vers l'estimateur optimal.

Le filtre de Kalman est un algorithme d'équations récurrentes optimales qui sont faciles à utiliser en calcul automatique et dont les deux équations de base regroupent toute l'information disponible.

La partie déterministe de la dynamique du système est exprimée par la matrice de transition et celle de la mesure est exprimée par la matrice de mesure, tandis que la partie aléatoire est exprimée par les propriétés statistiques (valeurs moyennes, matrices de covariance) des bruits d'état et de mesures.

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

- Choix de la variable d'état du système étudié et détermination du modèle mathématique conformément aux équations
- Estimation des conditions initiales.

- Calcul conformément aux équations récurrentes du filtre.
- Vérification des hypothèses d'application par l'examen des résidus et critique des résultats.





# *Chapitre IV :*

## *La prédiction*

### *Multi-site des apports annuels*

### IV.1. Introduction :

L'objectif de notre présent chapitre étant de faire des prédictions multi-sites, notre choix est orienté vers le filtre de Kalman discret. L'application est faite sur les données d'apports liquides observés à l'échelle annuelle, sachant que la procédure en gros, reste valable pour des pas de temps inférieurs et valable aussi pour la prédiction multi-site de toute autre variable hydrologique ; pourvu que le choix des variables d'état soit fait d'une manière similaire.

### IV.2. Données De La Prédiction Multi-Site :

Les données regroupent 5 stations hydrométriques, Les observations forment des séries chronologiques ayant une période commune de 21ans

La période commune d'observation entre les 5 stations hydrométriques étant de 21 ans, les observations des premières 15 années au pas de temps annuel ont été utilisées à l'estimation des paramètres du modèle. Les Six dernières années ont servi à la validation. 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 apports liquides annuels pour les stations hydrométriques

Dans notre cas, nous allons travailler avec des séries chronologiques ce qui implique que le terme correspondant à l'entrée du système, dans l'équation d'état, est nul. La variable état du système est un vecteur donnant l'apport liquide annuel observé simultanément aux 5 stations hydrométriques considérées. Donc les équations de système et les différentes matrices citées avant s'appliquent pour  $n=5$ .

### IV.3. Formulation du modèle :

On admet comme hypothèses les points suivants

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

L'état de traitement a priori \*k/k-1\*

On pose :  $K=1$

Le vecteur d'état initial  $\hat{X}_{K/K-1}$  est constitué par la moyenne des apports liquides annuels observés aux 10 différentes stations hydrométriques en question.

$$\hat{X}_{1/0} = \begin{bmatrix} \bar{X}_1 \\ \bar{X}_2 \\ \bar{X}_3 \\ \bar{X}_4 \\ \bar{X}_5 \end{bmatrix}$$

Pour la matrice de l'erreur  $P_{K/K-1}$  on doit donner des valeurs élevées (entre 700 et 1000) à la trace de cette matrice afin de bénéficier de la stabilité de l'algorithme

$$P_{1/0} = \begin{bmatrix} 1000 & 0 & 0 & 0 & 0 \\ 0 & 1000 & 0 & 0 & 0 \\ 0 & 0 & 1000 & 0 & 0 \\ 0 & 0 & 0 & 1000 & 0 \\ 0 & 0 & 0 & 0 & 1000 \end{bmatrix}$$

Pour la matrice de covariance des bruits du système et celle des bruits des mesures du moment que l'on espère que les mesures soient moins bruitées que la dynamique du système, la matrice de covariance du bruit du système  $Q$  et celle du bruit de mesure  $R$ , toutes les deux (5x5) ont été choisies comme telles :

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

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

La matrice de transition d'état est considérée comme le corps de l'algorithme donc pour son estimation on considère les inter corrélations dans nos stations hydrométriques

Pour la matrice de mesure c'est une matrice identité de taille (5x5)

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

Et pour le calcul de gain de filtre qui est un indicateur de l'efficacité de l'algorithme et donc la fiabilité de notre travail, c'est une matrice (5x5) et en injectant les différents paramètres cités pour le pas k=1 on trouve la forme suivante :

$$K_1 = \begin{bmatrix} 1000 & 0 & 0 & 0 & 0 \\ 0 & 1000 & 0 & 0 & 0 \\ 0 & 0 & 1000 & 0 & 0 \\ 0 & 0 & 0 & 1000 & 0 \\ 0 & 0 & 0 & 0 & 1000 \end{bmatrix} \begin{bmatrix} 1000 & 0 & 0 & 0 & 0 \\ 0 & 1000 & 0 & 0 & 0 \\ 0 & 0 & 1000 & 0 & 0 \\ 0 & 0 & 0 & 1000 & 0 \\ 0 & 0 & 0 & 0 & 1000 \end{bmatrix} + \begin{bmatrix} 50 & 0 & 0 & 0 & 0 \\ 0 & 50 & 0 & 0 & 0 \\ 0 & 0 & 50 & 0 & 0 \\ 0 & 0 & 0 & 50 & 0 \\ 0 & 0 & 0 & 0 & 50 \end{bmatrix}^{-1}$$

**L'état d'actualisation \*k/k\***

Le but est de faire actualiser les vecteurs et les matrices de l'état précédentes pour obtenir une nouvelle mesure et un vecteur d'état amélioré c'est ça l'intérêt de filtre de Kalman c est de faire l'amélioration de l'information dans chaque étape

On utilise un facteur de correction pour améliorer le vecteur d'état donc le vecteur d'état actualisé sera en fonction de ce facteur  $K_1$  et les vecteurs état et mesure de l'état de traitement a priori  $k/k-1$

On obtient :

$$\hat{X}_{1/1} = \begin{bmatrix} \bar{X}_1 \\ \bar{X}_2 \\ \bar{X}_3 \\ \bar{X}_4 \\ \bar{X}_5 \end{bmatrix}_{1/0} + \begin{bmatrix} k_{1,1} & k_{1,2} & k_{1,3} & k_{1,4} & k_{1,5} \\ k_{2,1} & k_{2,2} & k_{2,3} & k_{2,4} & k_{2,5} \\ k_{3,1} & k_{3,2} & k_{3,3} & k_{3,4} & k_{3,5} \\ k_{4,1} & k_{4,2} & k_{4,3} & k_{4,4} & k_{4,5} \\ k_{5,1} & k_{5,2} & k_{5,3} & k_{5,4} & k_{5,5} \end{bmatrix}_1 \left( \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix}_1 - \begin{bmatrix} \bar{X}_1 \\ \bar{X}_2 \\ \bar{X}_3 \\ \bar{X}_4 \\ \bar{X}_5 \end{bmatrix}_{1/0} \right)$$

Matrice de covariance :

$$P_{1/1} = \begin{bmatrix} 1000 & 0 & 0 & 0 & 0 \\ 0 & 1000 & 0 & 0 & 0 \\ 0 & 0 & 1000 & 0 & 0 \\ 0 & 0 & 0 & 1000 & 0 \\ 0 & 0 & 0 & 0 & 1000 \end{bmatrix}_{1/0} - \begin{bmatrix} k_{1,1} & k_{1,2} & k_{1,3} & k_{1,4} & k_{1,5} \\ k_{2,1} & k_{2,2} & k_{2,3} & k_{2,4} & k_{2,5} \\ k_{3,1} & k_{3,2} & k_{3,3} & k_{3,4} & k_{3,5} \\ k_{4,1} & k_{4,2} & k_{4,3} & k_{4,4} & k_{4,5} \\ k_{5,1} & k_{5,2} & k_{5,3} & k_{5,4} & k_{5,5} \end{bmatrix}_1 \begin{bmatrix} 1000 & 0 & 0 & 0 & 0 \\ 0 & 1000 & 0 & 0 & 0 \\ 0 & 0 & 1000 & 0 & 0 \\ 0 & 0 & 0 & 1000 & 0 \\ 0 & 0 & 0 & 0 & 1000 \end{bmatrix}_{1/0}$$

L'état a posteriori \*k+1/k\*

Ceci nous conduit la prédiction un pas en avant de l'estimation corrigée  $\hat{X}_{2/1}$

$$\begin{bmatrix} X_1 \\ X_2 \\ \cdot \\ \cdot \\ X_n \end{bmatrix}_{2/1} = \begin{bmatrix} A_{1,1} & \cdot & \cdot & A_{1,5} \\ A_{2,1} & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ A_{5,1} & \cdot & \cdot & A_{5,5} \end{bmatrix}_{2/1} \begin{bmatrix} X_1 \\ X_2 \\ \cdot \\ \cdot \\ X_n \end{bmatrix}_{1/1}$$

$$\begin{bmatrix} W_1 \\ W_2 \\ \cdot \\ \cdot \\ W_n \end{bmatrix}_k \text{ Ce terme est ignoré généralement pour le cas discret}$$

A la fin la matrice de covariance (voire formule....) :

$$P_{2/1} = \begin{bmatrix} A_{1,1} & \cdot & \cdot & A_{1,5} \\ A_{2,1} & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ A_{5,1} & \cdot & \cdot & A_{5,5} \end{bmatrix}_{2/1} \begin{bmatrix} p_{1,1} & \cdot & \cdot & p_{1,5} \\ p_{2,1} & \cdot & \cdot & p_{2,5} \\ \cdot & \cdot & \cdot & \cdot \\ p_{4,1} & \cdot & \cdot & p_{4,5} \\ p_{5,1} & \cdot & \cdot & p_{5,5} \end{bmatrix}_1 \begin{bmatrix} A_{1,1} & \cdot & \cdot & A_{1,5} \\ A_{2,1} & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ A_{5,1} & \cdot & \cdot & A_{5,5} \end{bmatrix}_{2/1}^T + Q$$

Pour le reste de l’algorithme et pour obtenir l’état suivant 3/2 on prend les résultats de l’itération k=1 comme données initiales

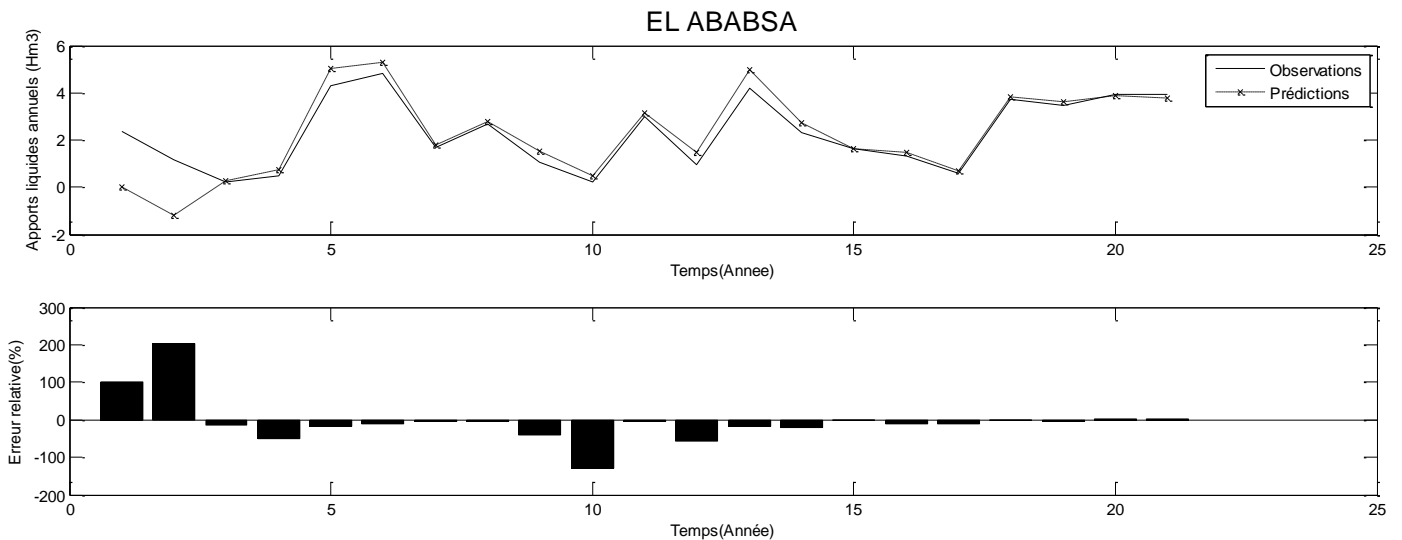
L’algorithme fonctionne de la même manière sur tous les sites et le long de la période d’étude 21 ans

**IV.4. Résultats et interprétations :**

**IV.4.1. Prédications multi-sites :**

Les résultats de ces prédictions sont représentés comme des matrices (21x5) correspondant aux prédictions durant les 21 ans (lignes) pour les 5 stations hydrométriques (colonnes). Ces résultats peuvent donc être appréciés dans les deux dimensions : temporelle si on opte pour une lecture dans le sens des lignes de chaque station, et spatiale si on opte pour une lecture en colonne de chaque année de prédiction.

**IV.4.2. Prédications temporelles (selon les années) :**



*Figure. IV.1.1 Prédications obtenues à la station hydrométrique EL ABABSA*

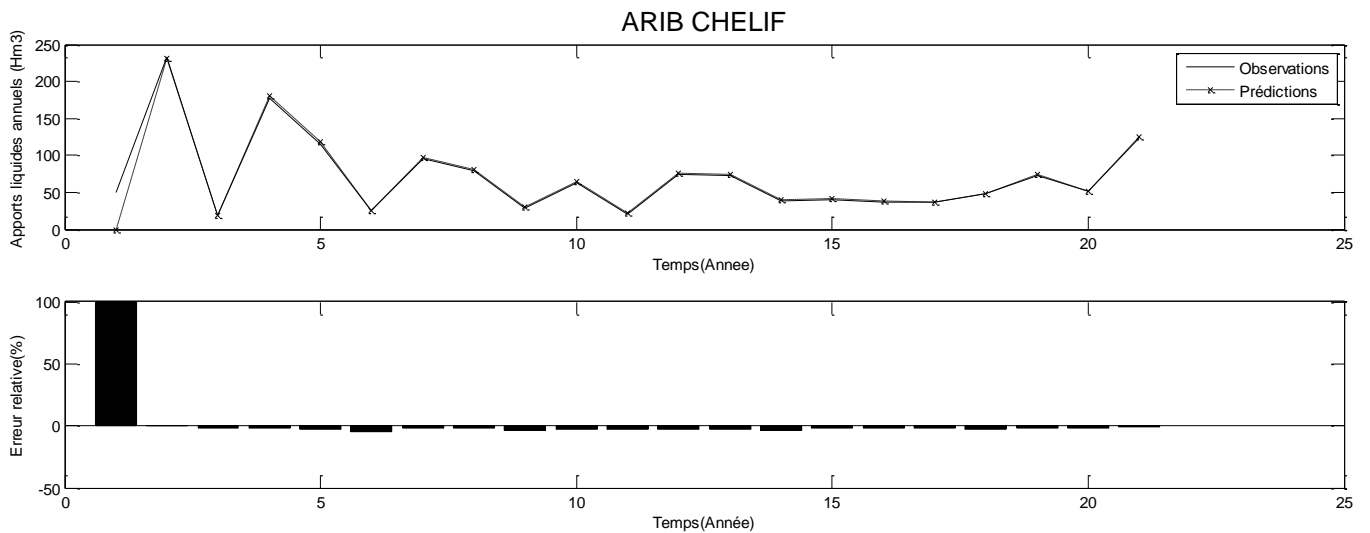


Figure. IV.1.2 Prédications obtenues à la station hydrométrique ARIB CHELIF

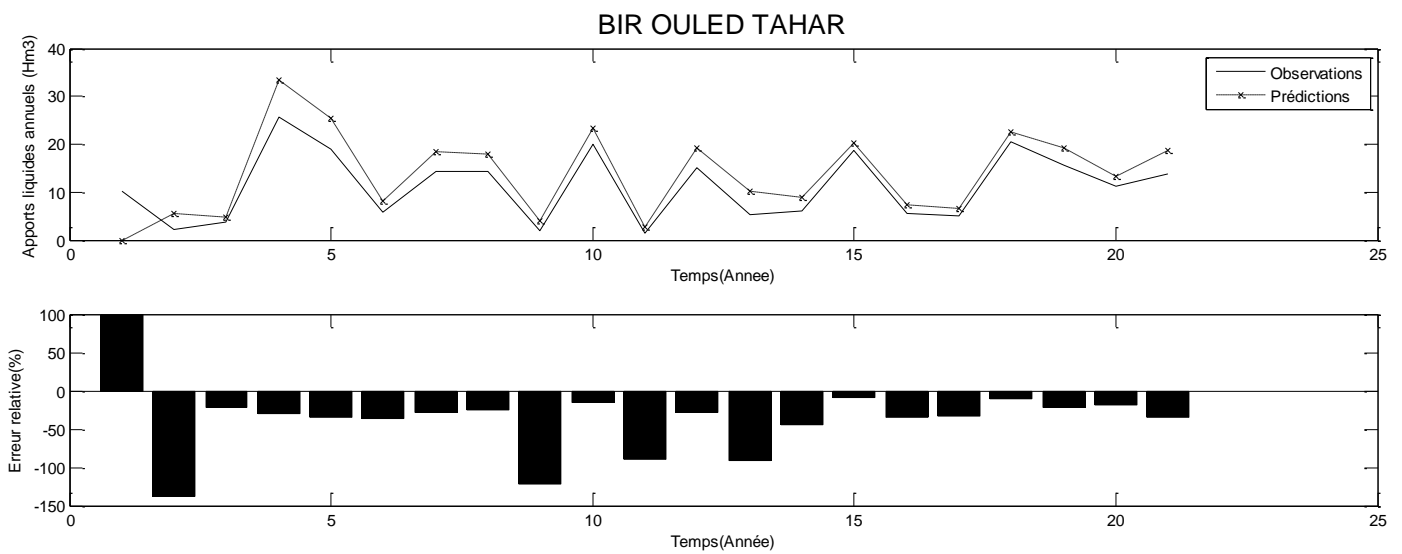


Figure. IV.1.3 Prédications obtenues à la station hydrométrique BIR OULED TAHAR

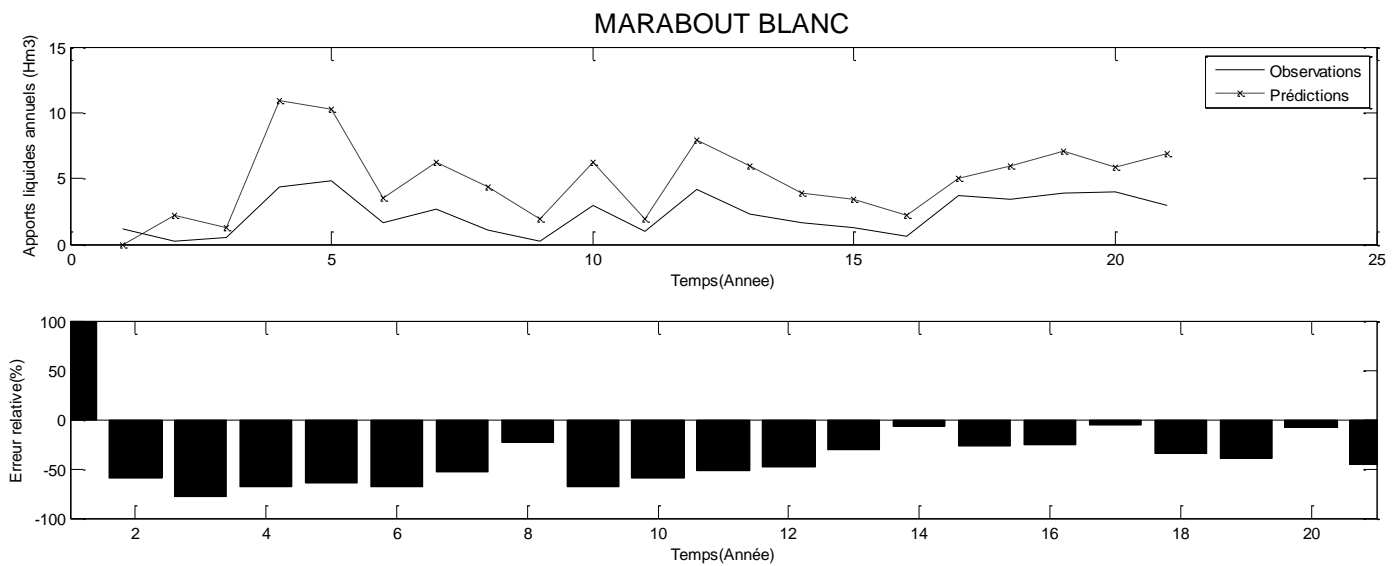


Figure. IV.1.4 Prédiction obtenue à la station hydrométrique MARABOUT BLANC

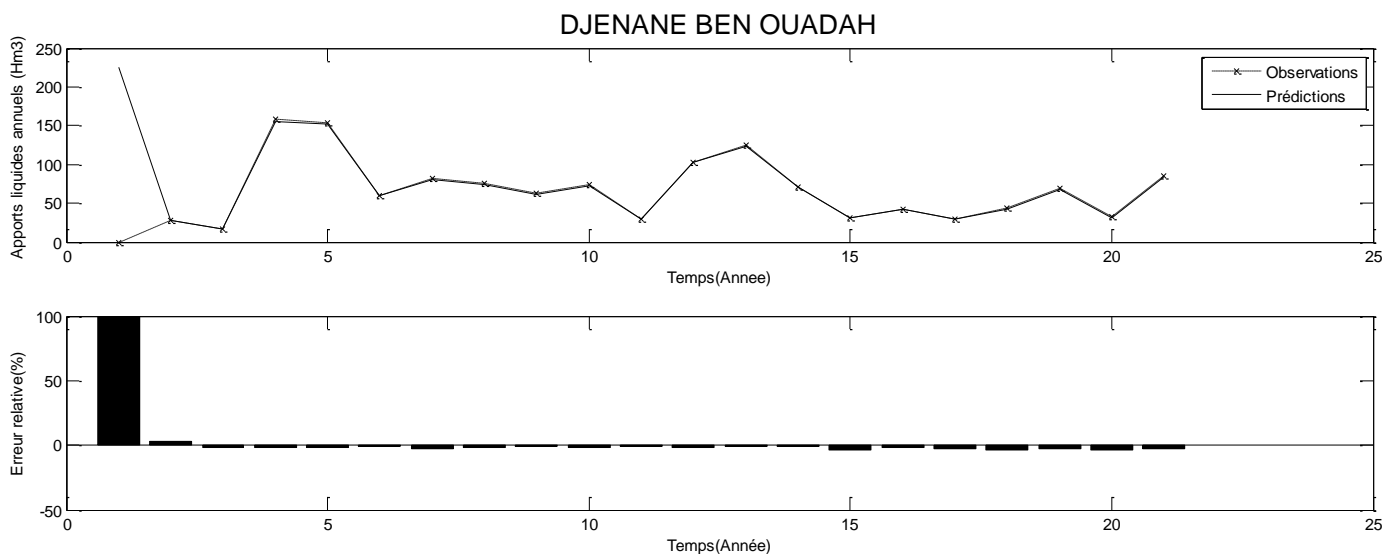


Figure. IV.1.5 Prédiction obtenue à la station hydrométrique DJENANE BEN OUADAH

Pour la prédiction multi-site dans la dimension temporelle, les Figures. IV.1.1-.IV.1.5 montrent les prédictions obtenues simultanément dans les 21 ans sur chaque station hydrométrique comparées aux observations enregistrées et donnent en même temps les erreurs relatives de prédiction exprimées en pourcentage.

On remarque pour chaque station que l'erreur au début est trop élevée (les 15 ans premiers) car le filtre est dans la phase d'estimation des paramètres par la suite le modèle se converge rapidement pour atteindre des valeurs plus réduites (la validation)

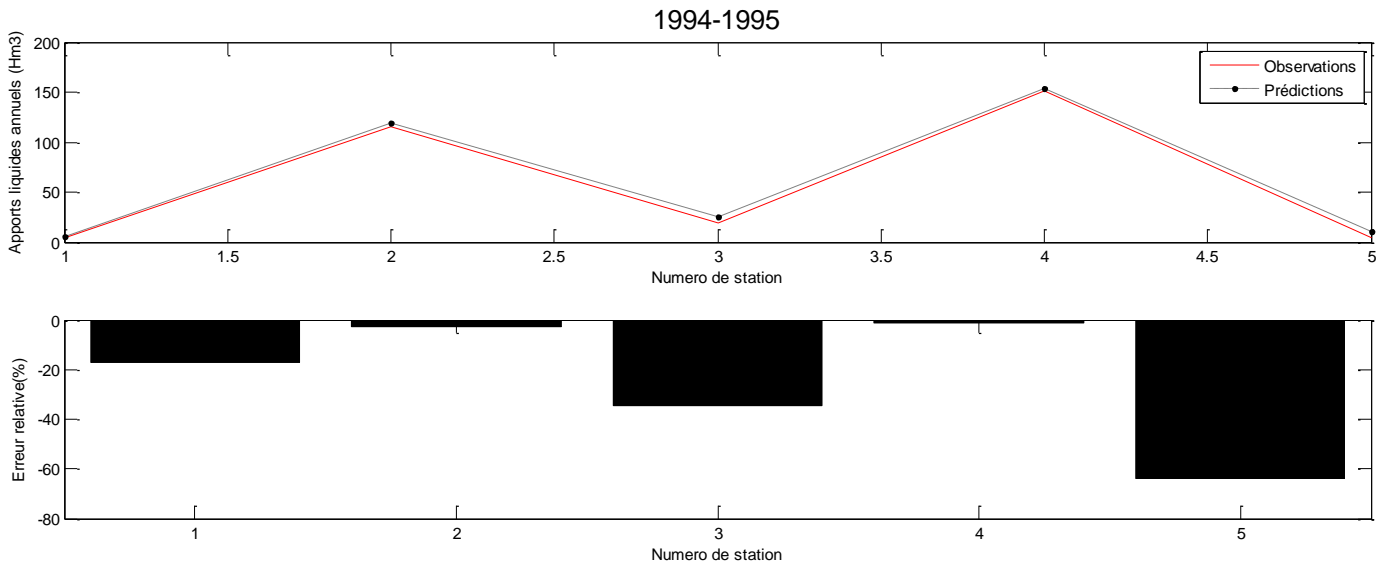
- 1% dans la station ABABSA pour l'année hydrologique 2010-2011,
- Entre 1.2% et 5% pour la station ARIB CHELIF.



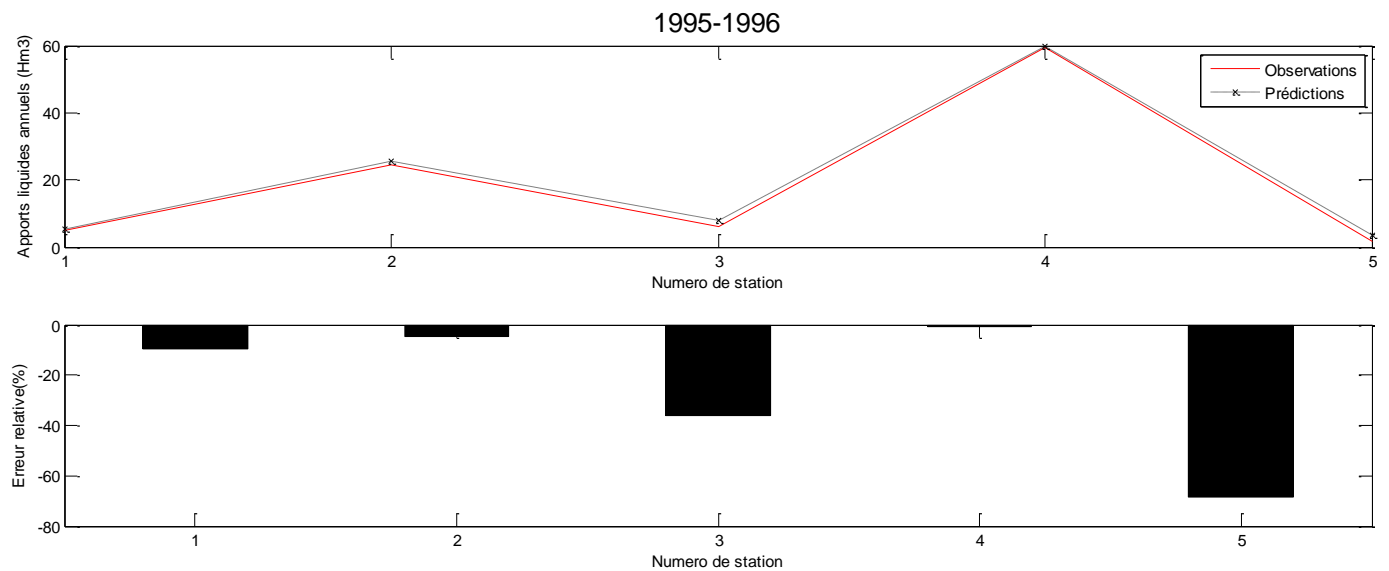
- Des erreurs de -8% pour l'année 2005-2006 dans la station BIR OULED TAHAR.
- La station DJENANE BEN OUADAH donne des résultats remarquables avec erreur entre -1% et -3%
- Pour la dernière station MARABOUT BLANC on trouve que l'erreur diminue avec le temps mais pas rapide comme le reste des stations,

**IV.4.3. Prédictions spatiales (selon les station) :**

Vu le nombre élevé des graphes résultants on prend quelque année dans la période d'étude .



*Figure.IV.2.1 Prédictions des apports annuels (1994-1995)*



*Figure.IV.2.2 Prédictions des apports annuels (1995-1996)*

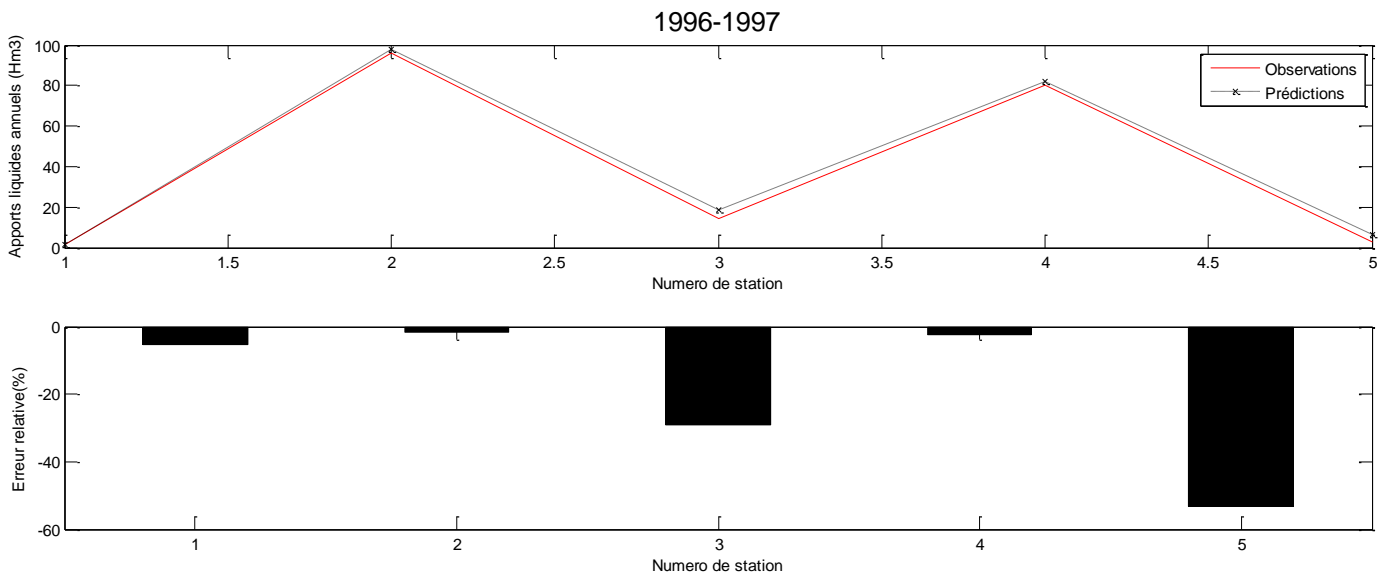


Figure.IV.2.3 Prédictions des apports annuels (1996-1997)

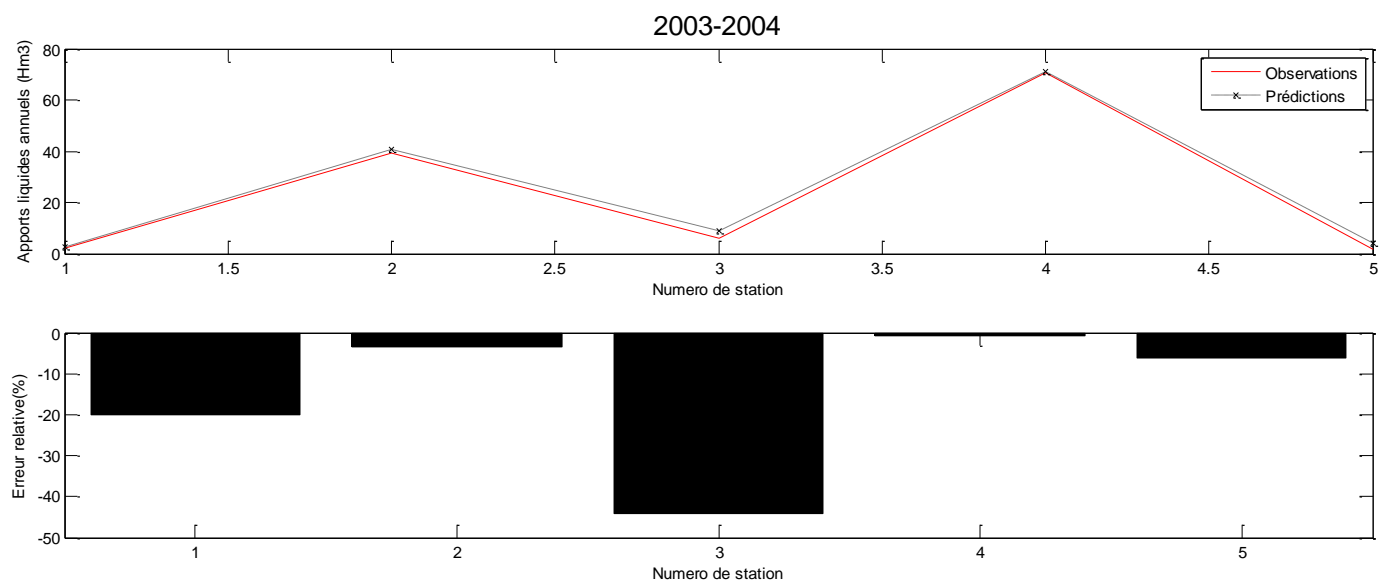


Figure.IV.2.4 Prédictions des apports annuels (2003-2004)

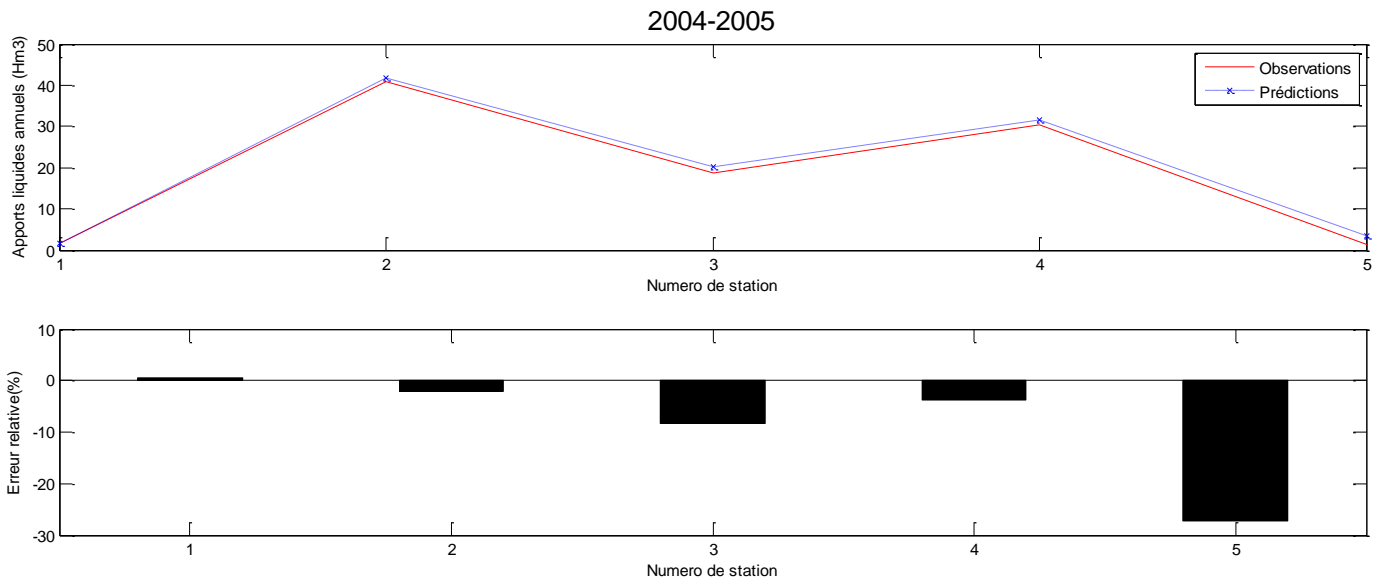


Figure.IV.2.5 Prédications des apports annuels (2004-2005)

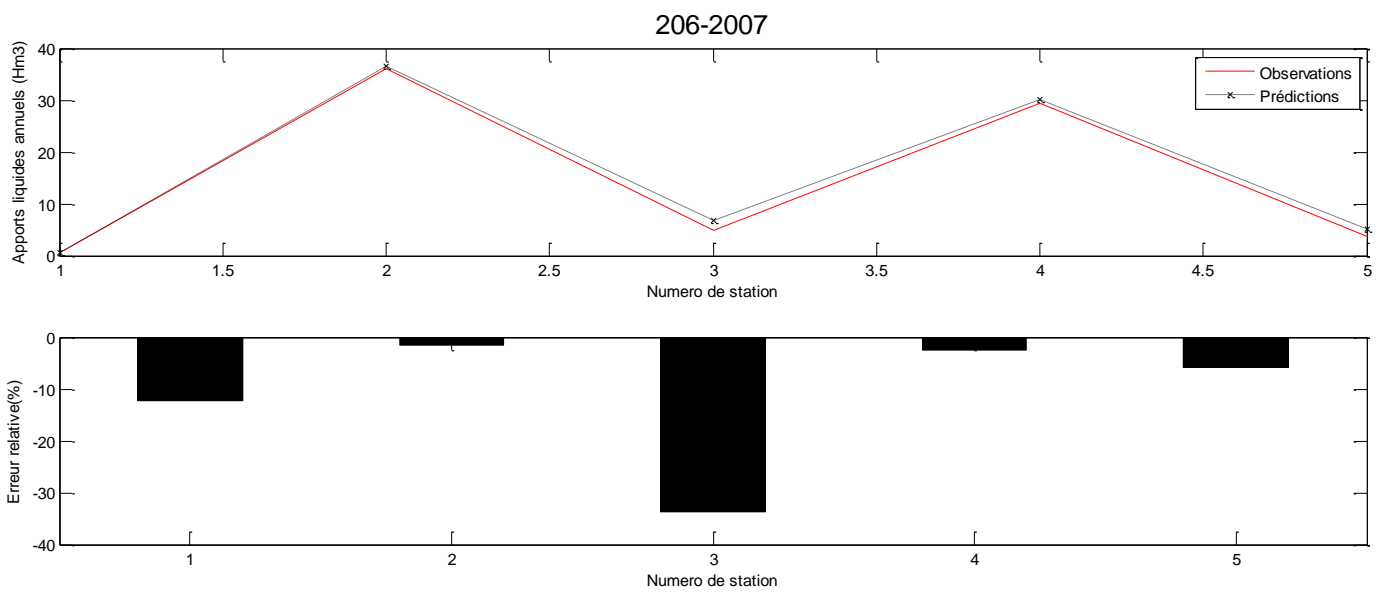


Figure.IV.2.6 Prédications des apports annuels (2006-2007)

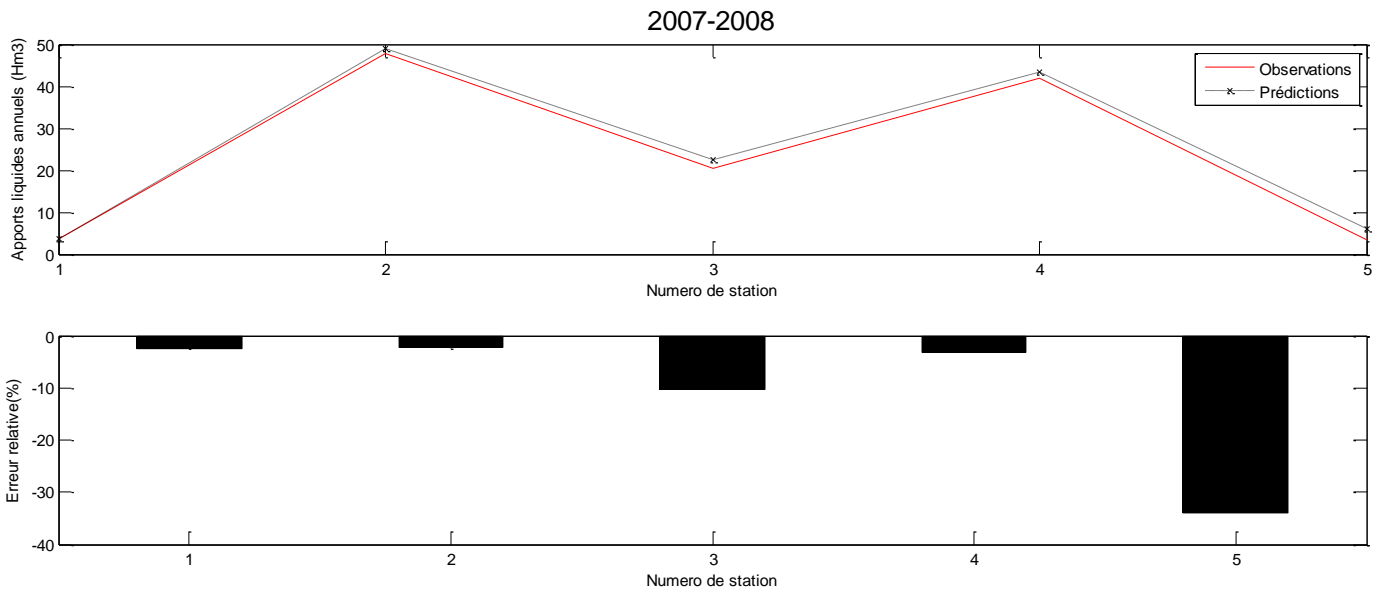


Figure.IV.2.7 Prédiction des apports annuels (2007-2008)

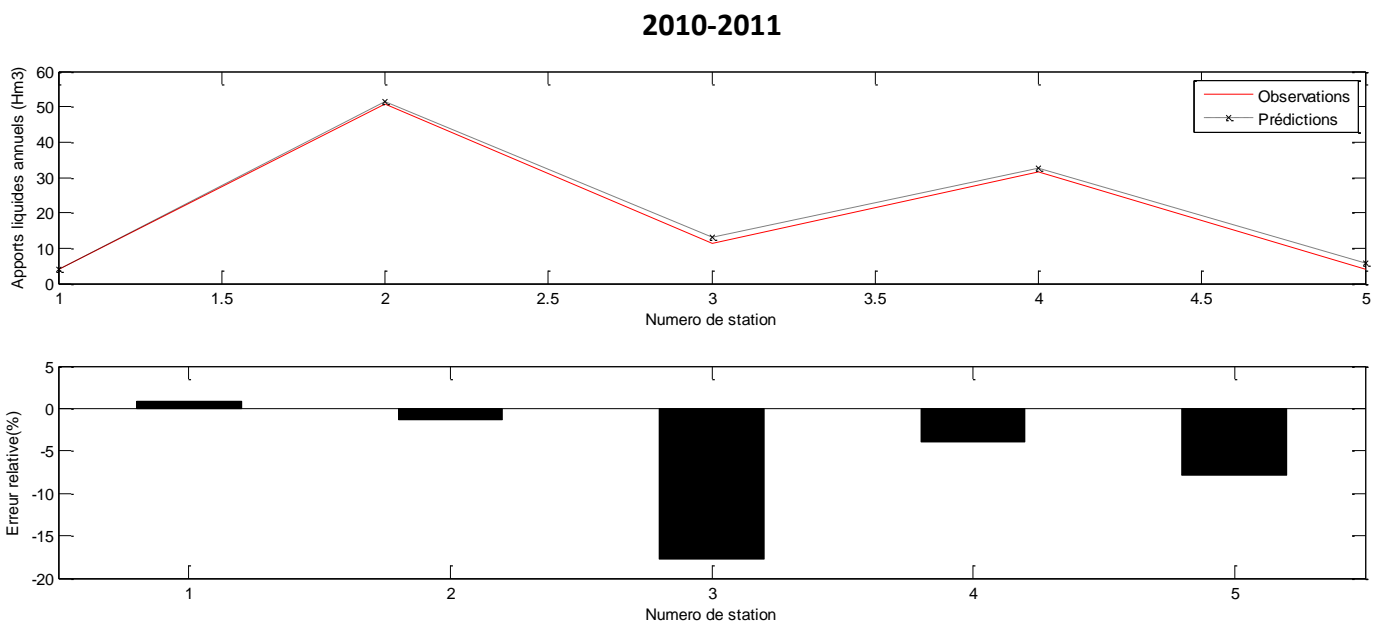


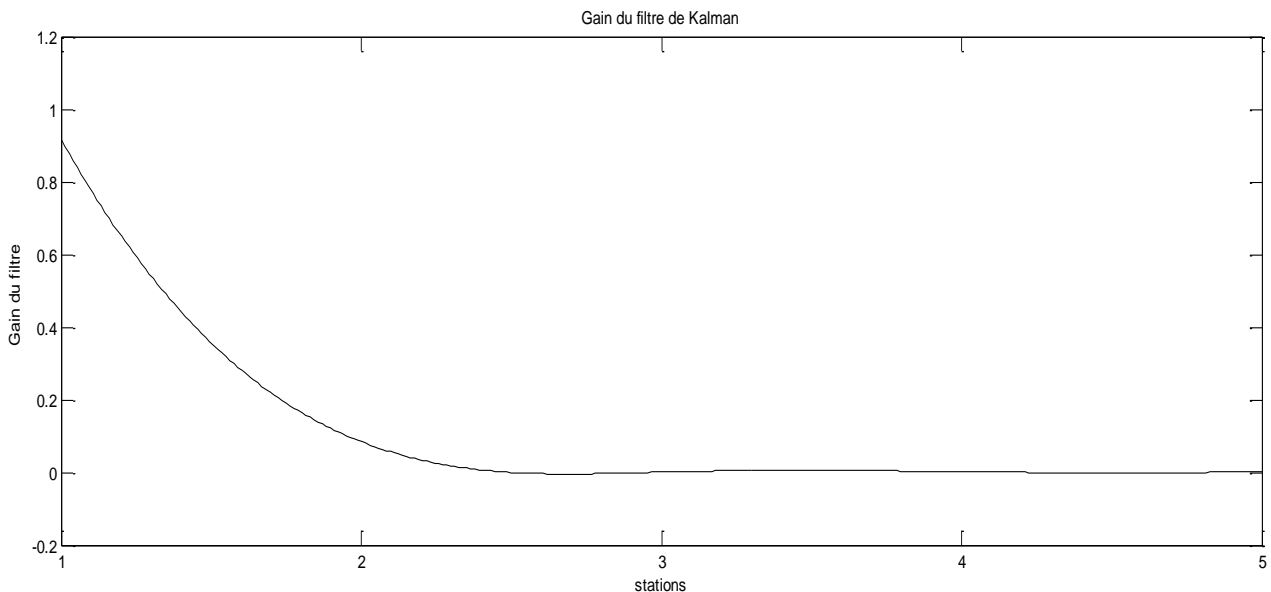
Figure.IV.2.8 Prédiction des apports annuels (2010-2011)

Pour la prédiction multi-site dans la dimension spatiale, la Figure IV.2.1 et Figure.IV.2. sont également des exemples, elles montrent les prédictions obtenues simultanément dans les cinq stations hydrométriques comparées aux observations enregistrées et donnent en même temps les erreurs relatives de prédiction exprimées en pourcentage.

La première figure est l'une des figures correspond au début des calculs (année 1994-1995), nous pouvons y constater le décalage important qui existe entre les observations et leurs estimations (prédictions). Ce décalage est bien exprimé par l'erreur relative qui atteint un maximum de -60 % à la station MARABOUT BLANC. La seconde figure correspond quant à elle à la fin des calculs (2009-2010), il est évident que le décalage entre les observations et leurs prédictions est nettement plus réduit et l'erreur relative qui était maximale à MARABOUT BLANC en 1994-1995 a atteint la valeur de -7 % en 1992.

#### IV.4.4. Optimalité des résultats

##### A. Gain du filtre :



**Figure.IV.3 : Gain de filtre FK**

Cette figure montre la variation continuellement décroissante du gain du filtre, du début à la fin des calculs.

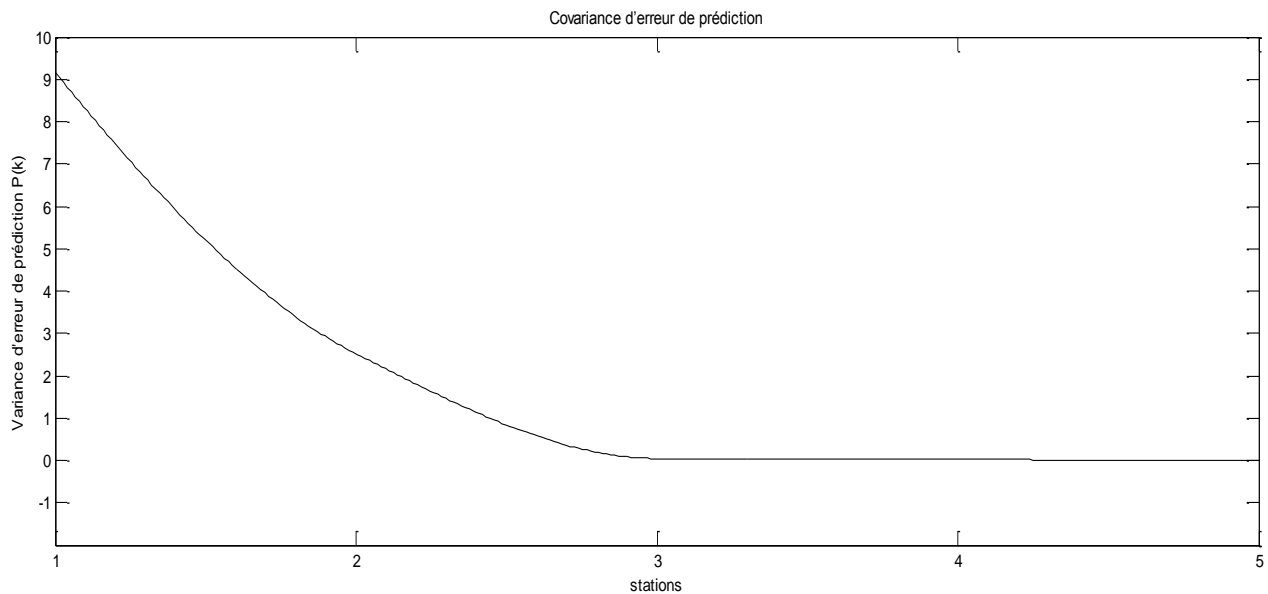
Le gain de Kalman varie selon la confiance (précision des mesures) que l'on peut accorder aux observations, c'est-à-dire le gain de Kalman augmente si les observations deviennent plus précises et inversement.

- Le gain diminue si l'estimation par le modèle devient plus précise .
- Le gain augmente si les observations deviennent plus précises .

Durant les premières itérations de l'algorithme de Kalman, la matrice gain  $K_k$  prend des valeurs importantes dont la conséquence est une estimation automatiquement mauvaise. Dans cette première phase, la mesure  $Z_k$  étant une information objective est plus crédible que l'estimation fournie par le modèle. En d'autres termes, on fait plus confiance à la mesure qu'au modèle. Après quelques itérations, la confiance assignée au modèle, comme mécanisme de prédiction, prend de plus en plus de l'espace, cette confiance croît et s'améliore d'une manière continue.

Dans notre cas, la valeur du gain du filtre  $K_k$  devient plus petite, par conséquent, la prédiction fournie par le modèle devient très proche de la valeur observée tel que montré par les figures de la prédiction spatiale et temporelle.

### A. La variance d'erreur



**Figure.IV.3 Covariance d'erreur de filtre FK**

L'optimalité des résultats est prouvée par la convergence de la matrice de covariance de l'erreur de prédiction et plus exactement par les éléments de sa diagonale majeure.

Cette convergence vers une valeur stable devient évidente juste après les premières itérations de l'algorithme tout en restant positive. Ce résultat étant conforme à la performance espérée du (FK), vient confirmer l'adéquation du modèle ajusté au processus étudié.

### IV.5. Analyse des innovations du filtre :

A la fin des calculs, il s'avère que certaines des valeurs prédites sont surestimées. La différence entre les mesures et leurs prédictions conduit à une série de résidus (innovations) pour chaque station.

Les erreurs de prédictions ou innovations du filtre pour les des 5 stations étudiées, présentent des valeurs trop élevées durant les premiers pas par rapport au reste des valeurs.

Cette différence est tout à fait normale et inévitable. Elle est due 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.[11]

Le résultat est un écart important entre l'observation et son estimation par le filtre, mais une fois que ce dernier converge tout rentre dans l'ordre : le modèle ajusté reprend beaucoup plus de confiance et les innovations qu'il produit sont par conséquent beaucoup plus petites à la fin des calculs, c'est le cas de toutes les stations étudiées comme c'est indiqué par les figures citées.

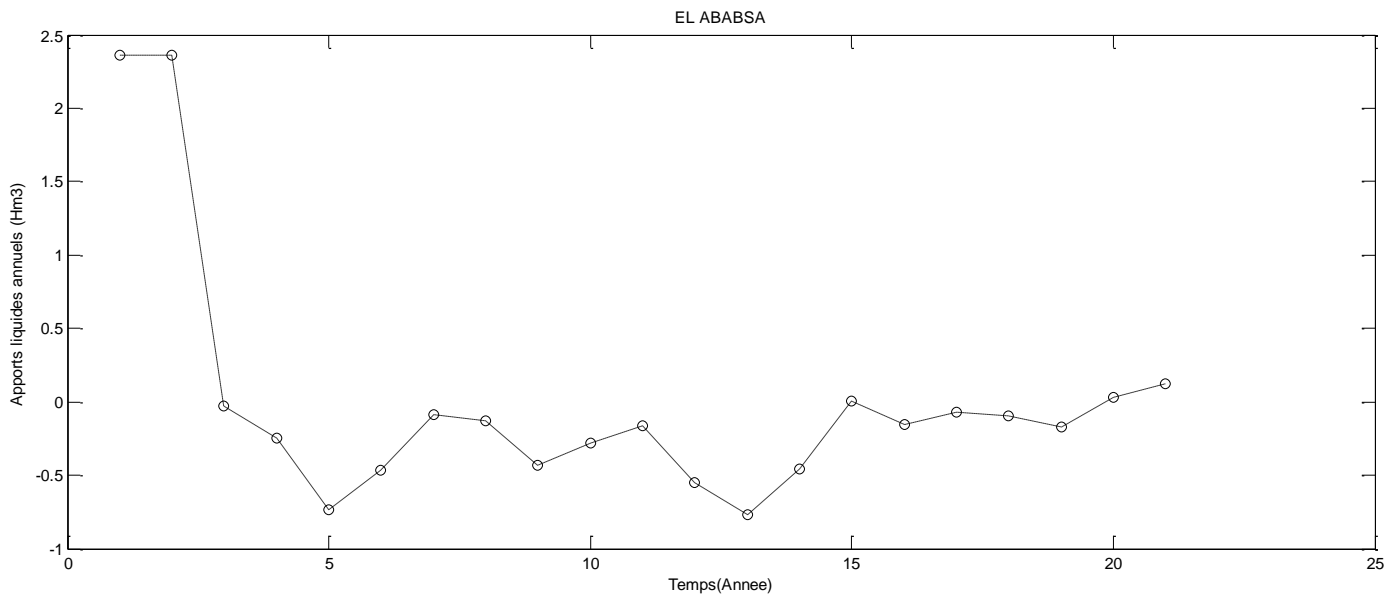


Fig.IV.3.1 Innovations du FK à ABABSA

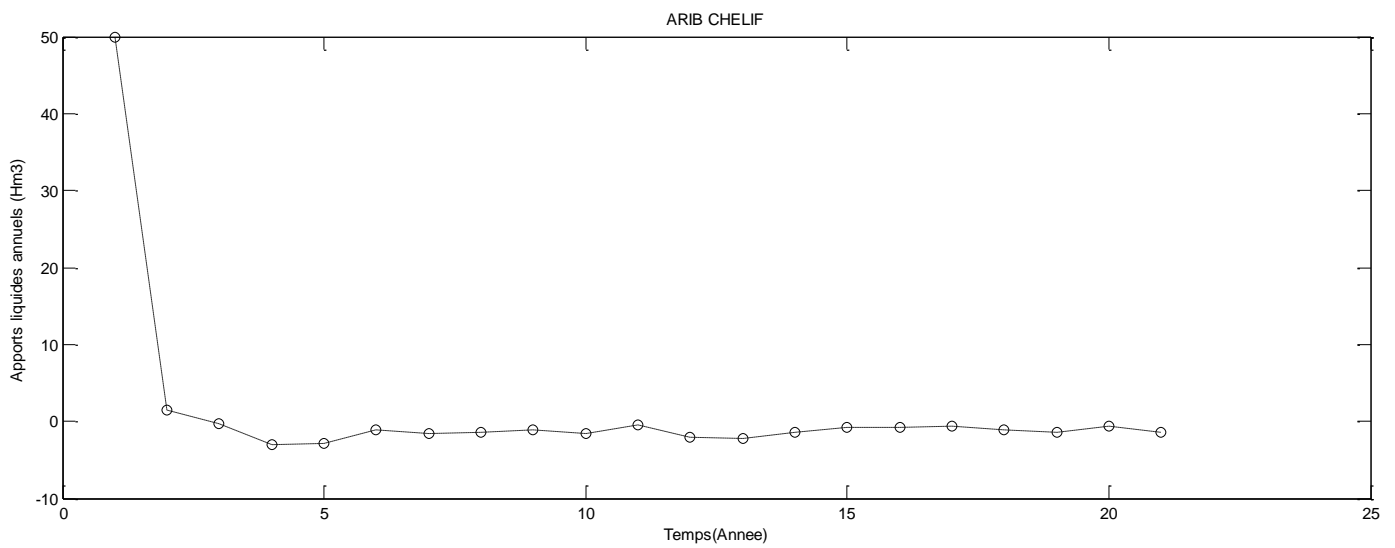


Fig.IV.3.2 Innovations du FK à ARIB CHELIF

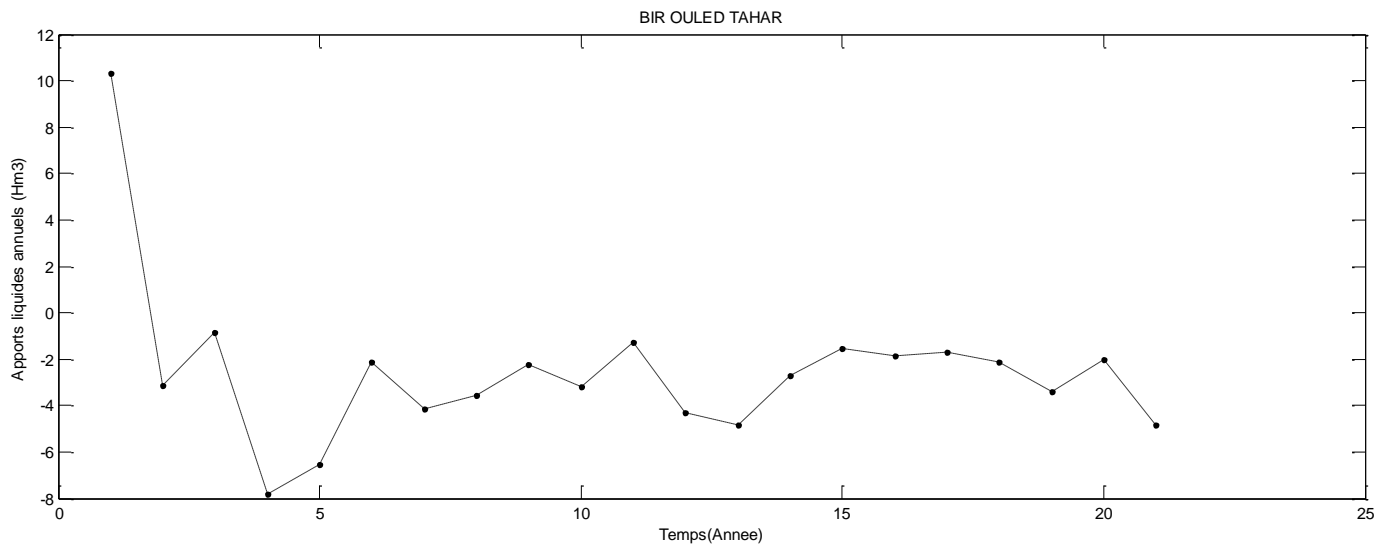


Fig.IV.3.3 Innovations du FK à BIR OULED TAHAR

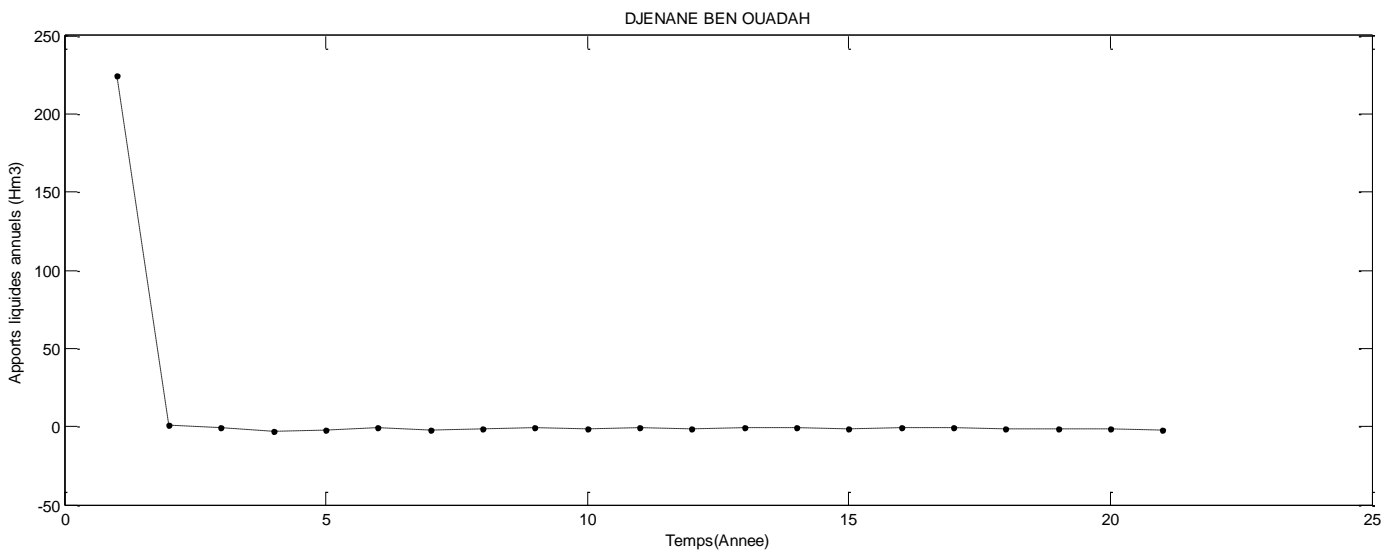
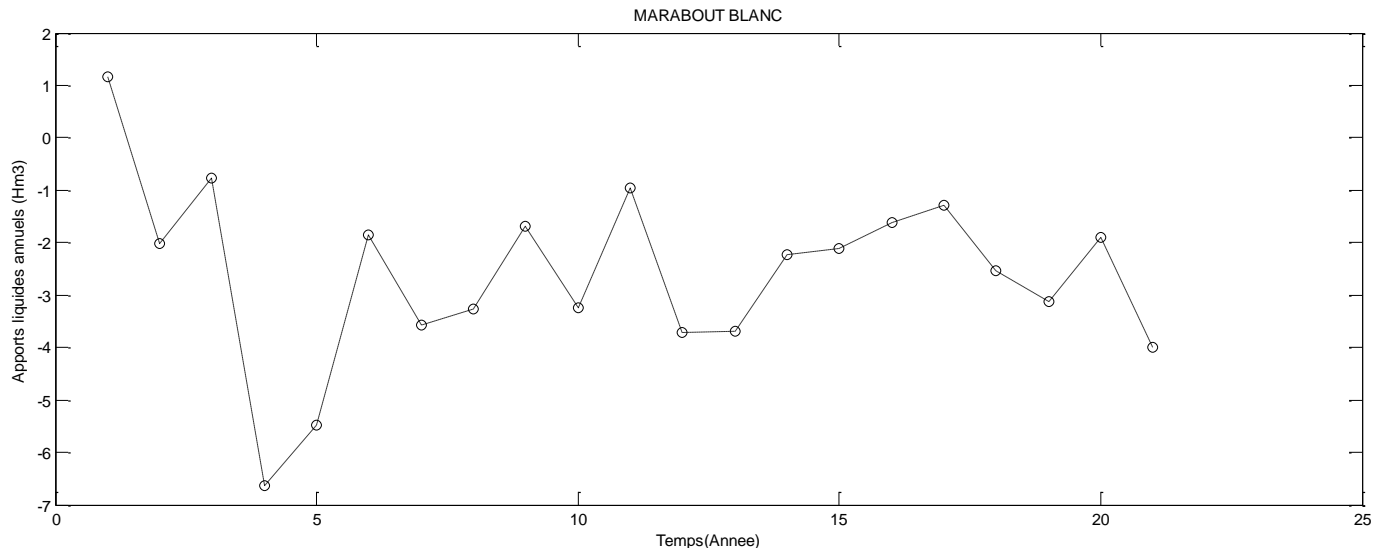


Fig.IV.3.4 Innovations du FK à DJENANE BEN OUADAH





**Fig.IV.3.5 Innovations du FK à MARABOUT BLANC**

#### IV.5. Conclusion :

Dans ce chapitre nous avons présenté Le développement de ce modèle a pour objectif, l'application de la technique du filtre de Kalman(FK) à la prédiction multi-site des apports liquides annuels et l'amélioration des erreurs de prédiction.

Dans la présente étude, nous nous intéressons aux séries chronologiques des apports liquides annuels, observés à 5 stations hydrométriques au Nord d'Algérie et couvant une période d'observation commune de 21 ans allant de 1990 à 2011.

Le travail a été fait sur deux périodes de temps. La première a été utilisée dans la procédure d'estimation des paramètres du modèle, tandis que la deuxième a servi à la validation du modèle.

L'apport liquide moyen annuel observé à chacune des 5 stations hydrométriques a été adopté comme vecteur d'état initial. Pour la matrice de covariance d'erreur initiale, nous avons opté pour de grandes valeurs dans les éléments de la diagonale principale afin de donner plus de poids à la nouvelle information (mesure) qui vient d'être disponible.

Les prédictions multi-sites des apports liquides, obtenues dans la dimension temporelle pour chacune des 5 stations hydrométriques en question, sont très proches des valeurs observées aux 5 stations sur la même période. De même pour les prédictions multi-sites dans la dimension spatiale, les valeurs observées et les prédictions correspondantes se suivent d'une manière assez proche l'une de l'autre pour toute la période d'observation, ceci indique que le filtre de Kalman multi-site est un outil efficace pour la modélisation et la prédiction des apports liquides annuels en temps et en espace.

En fin, la justesse du modèle du filtre de Kalman ou son aptitude à représenter le processus naturel des apports liquides a été vérifiée par l'étude de l'indépendance des innovations du filtre, le gain et la covariance de l'erreur, Avec un pourcentage moyen de l'erreur relative de prédiction sur toute la période d'observation environ 10%.



## *Conclusion et perspectives*

Dans ce travail nous avons étudié le filtre de Kalman dans le but de l'appliquer dans le domaine d'hydrologie et plus exacte l'étude des apports liquides qui ils sont une partie major dans l'étude de transport solide

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. La première équation de ce modèle exprime l'état du système hydrologique concerné, dans notre cas, c'est un vecteur donnant les apports annuels à plusieurs sites à la fois. La deuxième équation exprime la mesure de cette grandeur d'état qui est dans ce cas un vecteur de même dimension que l'état obtenu par observation aux sites précédents.

L'estimateur développé au cours de ce travail a la particularité de pouvoir s'adapter automatiquement chaque fois qu'une nouvelle information (mesure) devient disponible. Par conséquent, des prédictions optimales des apports annuelles sont obtenues aussi bien dans le temps que dans l'espace.

En effet, dans le cas où l'on dispose d'une information historique au préalable, une fois que le modèle a été sélectionné, d'abord les conditions initiales concernant le vecteur d'état, les covariances des bruits du système et de la mesure doivent être déterminées, puis la matrice de transition d'état et la matrice de mesure doivent être spécifiées, pour qu'en fin les calculs puissent être amorcés.

Dans la présente étude, nous nous intéressons aux séries des apports liquides annuels, observées aux 5 stations hydrométriques dans le bassin versant de haut Cheliff et couvrant une période d'observation commune de 21 ans allant de 1990 à 2011.

Les prédictions, obtenues dans la dimension temporelle, sont très proches des valeurs observées aux 5 stations sur la même période. De même pour les prédictions multi-sites dans la dimension spatiale, les valeurs observées et les prédictions correspondantes se suivent d'une manière assez proche l'une l'autre pour toute la période d'observation, ceci indique que le filtre de Kalman multi-site est un outil efficace pour la modélisation et la prédiction temporelle et spatiale de notre variable.

Le pourcentage de l'erreur relative des prédictions multi-sites a également été un moyen d'apprécier cette justesse. Le pourcentage moyen de l'erreur relative de prédiction sur toute la période d'observation environ 10% ce qui est très acceptable.

A la lumière de tous les résultats précédents, il est clair que la technique du FK multi-site est efficace pour la modélisation des apports liquides, aussi bien dans la dimension temporelle que spatiale. Le résultat obtenu est un modèle dynamique qui présente l'avantage considérable de prendre en compte les variations dans le temps des paramètres du modèle, de la variance des bruits d'état et celle de la mesure et dont les erreurs de prédictions sont minimales.

À l'aide de ce filtre on peut créer la carte des erreurs sur l'apports liquide pour le bassin de haut cheliff en se basant sur un multiple des sites et sur une longue durée d'observation.

## *Références Bibliographiques*

- [1].ATLAS du Cheliff Agence de Bassin Hydrographique Cheliff-Zahrez , 2005
- [2].Belmahdi Fatiha, Magister en Electronique Option télédétection Application du filtre de Kalman pour le débruitage des signaux université mouloud Mammeri, Tizi ousou 2015, p 16-29
- [3].Dennis Wackerly, William Mendenhall et Richard L. Schaeffer, Mathematical Statistics with applications, Brooks Cole, 2008, 7e éd., 922 p
- [4].Hilairret M.. Application des outils du traitement du signal à la commande des Machines tournantes. Thèse de Doctorat. Université de Nantes. France 2001
- [5].Jutten C.. Filtrage linéaire optimal. Cours de cinquième année du département 3i Options Image et Signaux et Automatique. Université Joseph Fourier-Polytech' Grenoble.France 2010
- [6].K. 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.
- [7].Kalman R. E. (1960). A New Approach to linear Filtering and Prediction Problems.
- [8].Transaction of the ASME- Journal of Basic Engineering, 82 (series D): 35-45.
- [9].Patrick Bogaert, Probabilités pour scientifiques et ingénieurs, De Boeck Université, 2006, Bruxelles
- [10].ROTELLA D. Filtrage Optimal. Ecole Nationale d'Ingénieurs de Tarbes.
- [11].S.Harkat, Application du Filtre de Kalman sur la variabilité pluviométrique dans le bassin versant de Chellif » Algerie,Université de Chlef,2016,188pp
- [12].Touhari, Fadhila Etude de la qualité des eaux de la vallée du Haut Cheliff Doctorat : Aménagement hydraulique : ENSH : Blida : 2015 p 20-28
- [13]. Yves DuceL, Introduction à la théorie mathématique des probabilités, Ellipses ,1998, (ISBN 2-7298-9820-4)