

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE



UNIVERSITE HASSIBA BENBOUALI DE CHLEF
FACULTE: SCIENCES & SCIENCE INGENIEUR
DEPARTEMENT D'HYDRAULIQUE

THESE

Présentée par :

HARKAT Samra

En vue de l'obtention du diplôme de Doctorat Sciences d'ingénierie en hydraulique

Intitulée

**APPROCHE SPATIALE ET CARTOGRAPHIE DE
LA VARIABILITE PLUVIOMETRIQUE DANS
LE BASSIN VERSANT DE CHELIFF « ALGERIE »**

Soutenue publiquement en date du 14 / 04 /2016 devant le jury composé de :

- SAAD Hamoudi Abdelamir	Professeur	U.de Chlef	Président
- HARTANI Tarik	Professeur	C.U.de Tipaza	Examineur
- BOUKHETTALA Djamel	Professeur	ENP d'Alger	Examineur
- BESSENESSE Mohamed	Professeur	U.de Blida	Examineur
- REMAOUN Mohamed	MCA	U.de Chlef	Examineur
- DOUAOUI AEK	Professeur	C.U. de Tipaza	Directeur de thèse

■ **REMERCIEMENT**

*« Je tiens à remercier Dieu le tout puissant pour m'avoir préservé,
donné la santé et guidé vers la connaissance et le savoir..... »*

Je ne sais comment ce travail serait fait si ce n'était la présence et le soutien de plusieurs personnes que je dois remercier :

Toute ma profonde et immense reconnaissance va directement à (aux) :

- **Le Professeur M^r Aek. DOUAOUI**, Centre universitaire Tipaza, d'avoir accepté d'être le directeur de la thèse ; je n'oublierai jamais son grand soutien pour terminer ce travail dans les meilleures conditions; toutes mes grands remerciements pour sa disponibilité, ses efforts, ses orientations et sa grande contribution à l'élaboration de ce mémoire, Je le remercie pour m'avoir placé face à un problème aussi passionnant que difficile, et pour la confiance qu'il a su porter en moi
- Un grand honneur, Un grand respect et un très grand merci à **M^{me} Malika .BOUKHAROUBA**, professeur à l'université de Biskra ; qui m'a suivi dès le début pour l'élaboration et l'amélioration de la qualité de ce travail par ses remarques, ses conseils et sa précieuse aide ; J'apprécie sincèrement sa direction technique, sa vision et ces orientations qui ont grandement contribué à ce qui a été finalement réalisé dans cette thèse. Je n'oublierai jamais son grand cœur et surtout sa grande patience avec moi.
- **Dr. Naima Bettahar** : Université de Chlef, d'avoir accepté d'être mon Co-encadreur.
- **Membres de Jury** qui ont accepté de participer à l'évaluation de ce travail, Mon grand respect à vous messieurs ;
- Une pensée particulière est adressée au professeur **Jean LABORDE**, professeur à l'Université de Montpellier en France, qui m'a guidé et m'a aidé à l'interprétation des résultats finale de ce travail.

Je tiens également à remercier les institutions suivantes qui m'ont permis de mener sereinement et jusqu'au bout ce travail:

- Tout le personnel et les enseignants de **l'Université de H.B.B à Chlef** (Faculté: génie civil et architecture/Département d'hydraulique), qui ont bien accueilli et accepté ce thème de recherche ;
- Toute L'Equipe de **Laboratoire**: « Laboratory of Crop Production and Sustainable Development of Natural Resources », Université Khemis-Miliana.
- Je souhaite également remercier la direction de **l'A.N.R.H** d'Alger , Ainsi tout le Personnel des Services d'Hydrologie et d'Hydrogéologie de **l'A.N.R.H** de Soumâa, d'Alger et de Khemis-Miliana pour m'avoir facilité la tâche de la collecte des données.
- Tout le Personnel du Centre d'informatique de **l'E.N.P** d'Alger.
- Touts les étudiants (es) de **l'ENSH** et leur intérêt vif qu'ils ont porté pour cette recherche.

Enfin, je remercie toute personne aime la révolution scientifique, et toutes les personnes qui ont cru en moi en m'accordant leur confiance.de loin

HARKAT. Samra

♥ Á ma très chère **Mère** qui m'a aidé avec son grand cœur...

♥ Á mon **Père**, qui a cru toujours en moi en m'accordant sa confiance, me redonner courage dans les moments difficiles ...

Je dédie cette thèse....

♥ Á ma princesse « **Sadoune Ghada-Aicha** » : qui a vécu souvent mes longues absences elle est toujours dans mon cœur.....

♥ Á mon ange « **Sadoune Malak-Yasmine** »adorable fille...

♥ Á **M^r Sadoune Sami**, mon **mari**, Merci pour ta patience.

♥ Á **M^r Sadoune Djamel**, qui m'a soutenu moralement de loin..... merci pour ta bienveillance.

♥ Á mon amie **Harkat Houda**, ma **Sœur**, à mes cotés toujours

Je dédie ce travail...

HARKAT. Samra

المخلص

هذه الدراسة تندرج في إطار النمذجة العشوائية للظواهر الهيدرولوجية لأغراض التنبؤ. في هذا العمل، يتركز اهتمامنا على نسبة الأمطار وتغيرها حسب التباين الجغرافي والزمني لها في الحوض المائي الشلف (الجزائر). الهدف هو دراسة إمكانية تطبيق تقنية فلتير كالمان في مجال النمذجة و التنبؤ لمواقع متعددة لسواقي الأمطار وكذلك تحسين أخطاء التنبؤ في الفترات الزمنية السنوية والشهرية. تقنية كالمان تملك ميزة كبيرة وهي إعطاء خطأ التنبؤ والذي هو في حد ذاته مؤشر للدقة زيادة على ذلك خوارزمية تعمل في المجال الزمني بطبيعة تكرارية، ويمتلك إمكانية تقديرية مثلى في تقنية الانحدار بطريقة المربعات الصغرى. جانب آخر من جوانب مثاليته هو أنه يدمج جميع المعلومات المتوفرة محطة على النظام، القياسات والأخطاء في حاسب متأقلم والذي يدمج المعلومات الجديدة عندما تتاح ومن أجل ذلك تم فحص معطيات 39 محطة رصد للأمطار بالحوض المائي للشلف وكذلك متابعة تغيرات كمية الأمطار السنوية والشهرية المسجلة على مدار 51 سنة (1959 إلى 2009). النتائج المتحصلة عليها تبين أن التنبؤات في أماكن مختلفة تتبع من قريب الملاحظات المسجلة وأخطاء التنبؤ بسيطة جدا هذا ينطبق على البعد المكاني والزمني للفترة السنوية والشهرية وهذا يبرهن على أن فلتير كالمان طريقة فعالة تأتي بتنبؤات جيدة حتى مع الفصلية. وفي النهاية فعالية فلتير كالمان قد برهنه بالمعدل، بالانزياح، وكذلك على أعلى قيم الأخطاء النسبية للتنبؤ، معدل الخطأ النسبي هو أقل من 10% وهذا مقبول على نطاق واسع.

مفتاح الكلمات: فلتير كالمان، التنبؤ في مواقع متعددة، هطول الأمطار، الحوض المائي الشلف، الجزائر

Résumé

La présente étude s'inscrit dans le cadre de la modélisation stochastique dynamique des phénomènes hydrologiques pour des fins de prédiction. Dans ce travail, nous nous intéressons à la pluviométrie et sa variabilité géographique et temporelle dans le bassin versant de Cheliff (Algérie). L'objectif est l'étude de l'applicabilité de la technique du filtre de Kalman (FK) discret à la modélisation et la prédiction multi-site de la pluviométrie, ainsi que l'amélioration des erreurs de prédiction et ce pour le pas de temps annuel et mensuel. L'outil FK présente l'avantage majeur 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 trente-neuf (39) stations pluviométriques dans le bassin versant de Cheliff ont été investies et les séries chronologiques des précipitations mensuelles et annuelles observées sur une période de 51ans (1959 à 2009) ont été mise en œuvre. 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, autant au pas mensuel qu'annuel, ce qui montre que le FK est un outil assez performant procurant de bonnes prédictions même en présence de saisonnalité. En fin, les performances du FK sont démontrées en termes de moyenne, écart-type et valeurs extrêmes des erreurs de prédiction. 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 Chéloff, Algérie.

Abstract:

The present study is part of the stochastic modeling framework dynamics of hydrological phenomena for prediction purposes. In this work, our attention is focused on rainfall and its geographical and temporal variability in the watershed Cheliff (Algeria).

The objective is to study the applicability of the technique of Kalman filter (KF) discrete modeling and multi-site rainfall prediction, also improvement of prediction errors for Annual and monthly time. KF has the major advantage of providing the prediction error that has to be an accurate indicator. In addition, the algorithm works in the time domain with a recursive nature, and has an optimal estimator in the least squares sense. Another aspect of its optimality is that it incorporates all the information available on the system, measurements and errors in an adaptive operator that is reset each time a new measure becomes available. To this end, data from thirty-nine (39) Cheliff rainfall stations in the watershed Cheliff have been invested; and the time series of annual and monthly precipitation observed over a period of 51 years (1959-2009) have been implemented.

The results obtained show that the multi-site predictions follow more closely recorded observations and prediction errors are minimal. This is true both in the temporal and spatial dimension, both in monthly and yearly step, which shows that the CF is a fairly powerful tool providing good predictions even with seasonality. In the end, the CF's performance is demonstrated in terms of mean, standard deviation and extreme values of the prediction errors. The average relative error of the predictions is less than 10%, which is broadly acceptable.

Keywords: Kalman filter, multi-site prediction, precipitation, Watershed Cheliff, Algeria.

■ *Liste des Figures:*

Fig I.1.1	Lien entre la densité f et la fonction de répartition et les probabilités.
Fig I.2.1	Rudolf Emil Kalman
Fig I.2.2	Les différentes étapes d'évaluation des paramètres du système de FK.
Fig II.1.1	Sous régions du bassin hydrographique Cheliff-Zahrez.
Fig II.1.2	Carte des sous bassins versant
Fig II.1.2	Réseau Hydrographique (Cheliff à l'amont de Boughzoul)
Fig II.1.4	Répartition géographique des précipitations
Fig II.1.5	Les stations pluviométriques retenues
Fig II.1.5	la régression linéaire (la station de Zemalet El Amir AEK (010502))
Fig II.1.6	La corrélation annuelle de 39 stations dans le bassin versant de Cheliff
Fig II.1.7	La corrélation mensuelle de 39 stations dans le bassin versant de Cheliff
Fig II.1.8	La cartographie de la composante F1
Fig II.2.1	Trace de la matrice de covariance d'erreur de prédiction
Fig II.2.2	Gain du filtre de Kalman
Fig II.2.3	Prédictions annuelles à la station pluviométrique Teniet - el-had de 1959 à 2009
Fig II.2.4	Prédictions annuelles des précipitations aux 39 stations pluviométriques (exp : année 1960 et 2009)
Fig II.2.5	Trace de la matrice de covariance d'erreur de prédiction
Fig II.2.6	Gain du filtre de Kalman au pas mensuel
Fig II.2.7	Prédictions mensuelles obtenues à la station de Zmalet El Amir AEK (Sep 1959-Aout 2009)
Fig II.2.8	Observations et prédictions des précipitations aux 39 stations pluviométriques pour le mois d'Octobre
Fig II.2.9	Innovations du FK à Station (St10502 ; St 10701 ; St 10703 ; St 10704 ; St 10804) respectivement (1959-2009)
Fig II.2.10	Fonction d'auto-corrélation des innovations au Station (St10502 ; St 10701 ; St 10703 ; St 10704 ; St 10804) respectivement (1959-2009)
Fig II.2.11	Ajustement Normal aux innovations de (St10502 ; St 10701 ; St 10703 ; St 10704 ; St 10804) respectivement (1959-2009)
Fig II.2.12	Ajustement Normal aux innovations de (St10502 ; St 10704) tronquées des 5 premières valeurs
Fig II.2.13	Ecart-type moyen des observations et des prédictions annuelles aux 39 stations (1959-2009)
Fig II.2.14	Ecart-type moyen des observations et des prédictions mensuelles aux 39 stations Sep.1959 - Aout.2009)

■ *Liste des Tableaux:*

Tableau II.1.1	Caractéristiques morphométriques des bassins versants de la région d'étude.
Tableau II.1.2	Longueur des oueds permanents et temporaires par bassin versant.
Tableau II.2.1	Auto-corrélations des innovations au Station (St10502 ; St 10701 ; St 10703 ; St 10704 ; St 10804) respectivement
Tableau II.2.2	Test de Box-Pierce pour les innovations du filtre.
Tableau II.2.3	Test d'adéquation du Chi-deux de l'ajustement normal à 95% de confiance des innovations du filtre.
Tableau II.2.4	Certaines caractéristiques statistiques moyennes des observations et des prédictions annuelles obtenues aux 39 stations (1959-2009)
Tableau II.2.5	Certaines caractéristiques statistiques moyennes des observations et des prédictions mensuelles obtenues aux 39 stations (Sep.1959 – Aout. 2009)

■ *Liste des Annexes:*

AnnexeII.1.1	Les Cordonnées Lambert des stations pluviométriques de BV Cheliff
AnnexeII.1.2	Ajustement à une loi racine normale
AnnexeII.1.3	Les méthodes de comblement des lacunes
AnnexeII.1.4	La régression linéaire
AnnexeII.1.5	Résultats de l'ACP

AnnexeII.2.1	Prédiction Spatiale pas annuel
AnnexeII.2.2	Prédiction Spatiale pas mensuel
AnnexeII.2.3	Prédiction Temporelle pas annuel
AnnexeII.2.4	Prédiction Temporelle pas mensuel
AnnexeII.2.5	Innovations du FK à St (05) jusqu'à St(39) (1959-2009)
AnnexeII.2.6 (a)	Figures des fonctions d'auto-corrélation des innovations [St (06) à St(39)]
AnnexeII.2.6 (b)	Tableau des fonctions d'auto-corrélation des innovations [St (06) à St(39)]
AnnexeII.2.7	Ajustement aux innovations de 39 stations
AnnexeII.2.8	Ajustement Normal aux innovations des stations [St 01;St 04; St 08; St 11; St 18;St 20; St 22; St 26;St 29;St 30;St 34;St 36 etSt 37] tronquées des 5 premières valeurs
AnnexeII.2.9	Le programme FK

■ TABLE DES MATIERES

Titre : Approche spatiale et cartographie de la variabilité pluviométrique dans le bassin versant de Cheliff » -Algérie-	N° de page :
Remerciement	
Résume, Abstract, ملخص	
LA LISTE DES : FIGURES, TABLEAUX, ANNEXES	
LA LISTE DES SYMBOLES ET DES ABREVIATIONS	
TABLE DES MATIERES	i-ii-iii-iv
INTRODUCTION GENERALE	1-2-3-4
► LA PREMIERE PARTIE : THEORIQUE	
LE CHAPITRE I.1 : NOTIONS DE BASE POUR ABORDER LE FILTRE DE KALMAN	
Introduction :	5
I.1.1. PROBABILITE ET VARIABLES ALEATOIRES	6
I.1.1.1. Loi de probabilité	6
I.1.1.2. Variable aléatoire réelle	7
I.1.1.3. Loi de probabilité à plusieurs variable	7
I.1.1.4. Densité de probabilité	7
I.1.1.5. Densité de probabilité d'une variable aléatoire réelle :	7
I.1.1.6. Espérance, variance et moments d'une variable aléatoire réelle à densité de probabilité	8
I.1.1.7. Densité de probabilité d'un vecteur aléatoire	8
I.1.1.8. Variables aléatoires discrètes	8
I.1.1.9. Fonction de répartition	9
I.1.1.10. Probabilité conditionnelle	9
I.1.1.11. Densité conditionnelle	9
	10
I.1.2. RAPPELLE STATISTIQUE ET METHODES D'ESTIMATION :	10
I.1.2.1. Rappels statistiques	
A- Paramètres de l'échantillon :	
A. 1. <u>Paramètre de tendance centrale:</u>	11
A. 2. <u>Paramètre de dispersion</u>	11
B- Paramètre de forme :	12
I.1.2.2. Test Relatif A L'échantillonnage	12
A- Test de stationnarité et d'homogénéité de Mann et Whitney (1947).	13
B- Test de stationnarité et d'indépendance de Wald et Wolfowitz (1943)	13
C- Test d'Anderson	13
D- Test du cumul des résidus	13
I.1.2.3. Choix Du Modèle Fréquentiel	14
A- Distributions Théoriques	14
A. 1. <u>Loi normale ou loi de Gauss</u>	15
A. 2. <u>Loi log-normale</u>	15
A. 3. <u>Loi de Gumel ou doublement exponentielle</u>	15
B- L'expérience Et La Coutume	15
I.1.2.4. Association Entre Variables	16
A- Modèle linéaire général	
B- Modèle à une variable explicative	16
C- Modèles non linéaires	16
I.1.2.5. Analyse multidimensionnelle (analyse en composantes principales (A.C.P))	16
	17

A-Analyse en composantes principales (A.C.P)	17
B-Principe et mise en œuvre	17
	18
I.1.3. PROCESSUS STOCHASTIQUE ET SERIES CHRONOLOGIQUES	18
	18
I.1.3.1.Processus Aléatoire	19
I.1.3.2.Processus Aléatoire stationnaire	19
I.1.3.3. Processus Bruit Blanc	20
I.1.3.4.Les séries chronologiques:	20
A-Définition	23
B-Types de Séries :	23
	25
I.1.3.5.Estimation stochastique linéaire	25
I.1.3.6.Modèle espace d'états ou modèle état mesure	25
I.1.3.7. Problème du filtrage	26
I.1.3.8.Caractérisation stochastique	26
A-Dans le cas continu	27
B-Dans le cas discret	27
	27
Conclusion	28
<hr/>	
LE CHAPITRE I.2 : METHODE DE FILTRE DE KALMAN	
Introduction	29
I.2.1. GENERALITE SUR LE FILTRE DE KALMAN:	29
	29
I.2.1.1. Historique du Filtre de Kalman	29
I.2.1.2. Domaines d'utilisation du Filtre de Kalman	29
I.2.1.3. Définition du Filtre de Kalman	30
I.2.1.4. Principe du Filtre de Kalman	30
I.2.1.5. Les limites du Filtre de Kalman	30
	30
I.2.2. DE L'ESTIMATEUR OPTIMAL AU FILTRE DE KALMAN	31
	31
I.2.2.1.L'estimation	31
I.2.2.2.Les Estimateurs déterministes	32
A-Le biais	32
B-La variance	32
	32
I.2.2.3.L'estimateur optimal	33
I.2.2.4.Optimalité et Récursivité de l'estimation	33
	34
I.2.3. MECANISME DE LA METHODE DE FILTRE DE KALMAN	34
	34
I.2.3.1.Variables et équations d'état	35
I.2.3.2.Forme matricielle des équations d'état	35
I.2.3.3.Les entrées et sorties du filtre de Kalman	35
	36
I.2.4. PRESENTATION DE L'ALGORITHME RECURSIF DU FK DISCRET	38
I.2.4.1.Les différentes étapes d'évaluation des paramètres du système de FK	38
A-Premier cas :	39
B-Deuxième cas	39
	39
I.2.5. STABILITE DU FILTRE VIS-A-VIS DE LA DIVERGENCE	40
	40
Conclusion	41

► LA DEUXIEME PARTIE : PRATIQUE

LE CHAPITRE II.1 : REGION D'ETUDE ET TRAITEMENT DES DONNEES

Introduction :	42
II.1.1. PRESENTATION DE LA REGION D'ETUDE	42
II.1.1.1. Situation géographique du bassin versant	42
II.1.1.2. Caractéristiques morphométriques de la région	45
II.1.2. LE RESEAU PLUVIOMETRIQUE :	49
II.1.2.1. Acquisition de données et choix de la période de référence	49
II.1.2.2. Caractéristiques du réseau pluviométrique	50
II.1.2.3. Répartition géographique des précipitations	50
II.1.2.4. Acquisition et traitement des données pluviométriques :	51
II.1.2.5. Choix de la période de référence :	52
II.1.2.6. Distribution statistique des pluies annuelles :	52
II.1.3. HOMOGENEISATION DES DONNEES :	53
II.1.3.1. Méthode de la régression linéaire :	55
II.1.3.2. Méthode de l'ACP:	56
Conclusion	58

LE CHAPITRE II.2 : APPLICATION DU FILTRE DE KALMAN

Introduction	59
II.2.1. DONNEES DE LA PREDICTION MULTI-SITE	59
II.2.1.1. HYPOTHESES	60
II.2.1.2. Formulation espace-état du modèle adopté	61
II.2.1.3. Vecteur d'état initial et matrice de covariance d'erreur associée	62
II.2.1.4. Matrice de covariance des bruits du système et celle des bruits de mesure	62
II.2.1.5. Matrice de transition d'état et matrice de mesure	63

II.2.2. LES EQUATIONS RECURRENTEES DU FILTRE DE KALMAN	65
II.2.2.1. Matrice gain du filtre	65
II.2.2.2. Utilisation de la mesure pour actualiser l'estimation	72
II.2.2.3. Matrice de covariance d'erreur- associée à l'estimation corrigée par la mesure	73
II.2.2.4. Prédiction un pas en avant de l'estimation (actualisée) corrigée par la mesure	74
II.2.2.5. Matrice de covariance de l'erreur de prédiction un pas en avant	76
II.2.3. RESULTATS ET INTERPRETATIONS	77
II.2.3.1. Prédiction au pas annuel (temporelle et spatiale)	77
II.2.3.2. Prédiction au pas mensuel (temporelle et spatiale)	79
II.2.3.3. Analyse des innovations du filtre (résidus)	79
II.2.4. MOYENNE ET DEVIATION STANDARD DES PREDICTIONS MULTI-SITES	80
II.2.4.1. Prédictions multi-sites « Pas annuelle »	81
II.2.4.2. Prédictions multi-sites « Pas mensuelle »	92
Conclusion	83
CONCLUSION GENERALE.	84-85
Références Bibliographiques	[iv] [iv] [iv]
Annexes :	16 Annexes
▶ AnnexeII.1.1	▶ AnnexeII.2.1
▶ AnnexeII.1.2	▶ AnnexeII.2.2
▶ AnnexeII.1.3	▶ AnnexeII.2.3
▶ AnnexeII.1.4	▶ AnnexeII.2.4
	▶ AnnexeII.2.5
	▶ AnnexeII.2.6 (a)
	▶ AnnexeII.2.6 (b)
	▶ AnnexeII.2.7
	▶ AnnexeII.2.8
	▶ AnnexeII.2.9

Publications:

- ▶ **Publication international (12p) : Desalination and Water Treatment**
 Publication details, including instructions for authors and subscription information: <http://www.tandfonline.com/loi/tdwt20>
Multi-site modeling and prediction of annual and monthly precipitation in the watershed of Cheliff (Algeria)
 Published online: 20 Oct 2014.
- ▶ **Les attestations de participation aux Colloques de : “2013’2015”**

INTRODUCTION GENERALE :

Dans un pays semi-aride comme l'Algérie, l'eau est un élément de survie qui est directement ou indirectement lié à tout développement économique et social. Malheureusement, les ressources en eau font face à de grands défis émis par le déficit pluviométrique et la rareté, ainsi que leur variabilité géographique. Selon (**Haddad et Rahla, 2004 ; H. Hisdal et L.M. Tallasken, 2003 ; J.M. Garcia-ruiz et al, 2011**), la demande en eau a atteint 5 milliards de mètres cubes par an, avec une alimentation de 170 m³/habitant/jour au lieu d'une norme minimale de 250 m³/habitant /jour, ce qui est déjà considéré comme un déficit. Ce déficit devient de plus en plus problématique d'une année à une autre et requiert, une solution urgente et efficace afin de surmonter le déficit de la demande et d'assurer un approvisionnement régulier en matière d'eau pour tous les usagers.

La nouvelle politique de l'Algérie se concentre sur la gestion intégrée des ressources en eau (**Mate, 2001**) par conséquent, deux types d'actions sont à prendre en considération, la gestion optimale et l'économie des ressources en eau.

Dans sa communication nationale initiale, l'Algérie a déjà prévu un programme d'actions pour faire face aux changements climatiques. Pour ce qui est de l'eau, ce programme intègre la gestion des bassins versants entre autres points. En plus du programme de développement de l'investissement de l'eau, qui a occupé une place importante dans le programme (2010-2014) du gouvernement algérien (**Rouane, 2011**), il serait opportun d'être complété par une recherche scientifique. Ceci va certainement avoir un impact positif sur les ressources en eau en Algérie globalement. Grâce à elle, nous voudrions donner l'un des outils d'aide à la décision aux planificateurs de l'eau ; et la présente étude est une modeste contribution à la réussite de cette nouvelle politique de l'eau.

Au cours de l'histoire des Sciences, les observations ont d'abord été qualitatives puis ce sont raffinées pour devenir quantitatives, alors que parallèlement, les modèles descriptifs se sont renforcés de l'apport des mathématiques, en particulier de l'Analyse Numérique. La motivation et l'enjeu principal de ces deux sources de connaissances (observation-modèle) étaient et sont toujours le besoin immédiat de prévisions météorologiques numériquement calculables. Dans la recherche de méthodes adéquates, s'est imposée l'idée selon laquelle les modèles dynamiques utilisés devaient rendre compte de la propagation des informations dans le temps et dans l'espace. Cette notion fondamentale de propagation compense la disparité et l'imprécision des observations et permet de construire une image consistante, quadridimensionnelle, de l'atmosphère, de climat,...dont la taille rend compliquée voire impossible la description complète. L'un des enjeux de ce type d'étude est donc de confronter une nécessité d'information (météorologie ...), et la complexité des données : bruit, distribution très irrégulière des mesures dans le temps et dans l'espace) (**C. Plumejeaud, 2006**)

La combinaison de modèles numériques rendant compte des lois fondamentales de la physique et des observations disponibles s'appelle assimilation et prédiction de données qui donne lieu à des recherches actuellement en plein essor.

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 (**Hwang and Brown ,1997**).

Il existe à ce jour deux grandes classes de méthodes d'assimilation de données qui correspondent à deux approches bien distinctes. Assimilation séquentielle, qui repose sur des considérations statistiques et qui procède par corrections successives de la prévision du modèle au fur et à mesure que des observations sont disponibles ; et l'assimilation variationnelle, qui consiste à ajuster au mieux une solution du modèle à toutes les observations disponibles tout au long de la période d'assimilation. Les méthodes séquentielles reposent essentiellement soit sur l'interpolation optimale (chercher une combinaison linéaire optimale entre les observations et les états du système aux mêmes instants),, soit sur le filtre de Kalman(utilise les observations dans l'ordre où elles sont disponibles, l'estimation de l'état du système ne dépendra pas des observations futures mais uniquement des observations passées et présentes).(C. **PLUMEJEAUD,2006**)

Dans la prédiction en temps réel, il est nécessaire de travailler dans le mode adaptif, ce qui implique l'utilisation d'un modèle ayant une structure rétroactive qui fait que la sortie du modèle au temps présent est reliée aux sorties observées précédentes, c'est le cas du filtre de Kalman (FK).

Le FK peut être défini d'une manière simple comme étant un algorithme optimal et récursif de traitement de données (**Szollosi, 1978**). 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 à-priori, la seconde, une prédiction basée sur la nouvelle information (nouvelle mesure). C'est un reconstruteur d'état dans un environnement stochastique, c'est un estimateur linéaire minimisant la variance de l'erreur d'estimation. 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 (**Kalman et Bucy,1961**), le filtre de Kalman n'a pas cessé d'intriguer les chercheurs . Plusieurs recherches théoriques ont été stimulées par les problèmes rencontrés lors de l'application de ce filtre à des problèmes pratiques entre autres (**Jazwinski, 1969**), et depuis on a constaté l'explosion de son application à des problèmes divers dans plusieurs domaines d'intérêt.

Le filtre de Kalman est utilisé dans une large gamme de domaines technologiques (radar, vision électronique, communication ...). C'est un thème majeur de l'automatique et du traitement du signal. Mais il est aussi de plus en plus utilisé en dehors de son domaine,

par exemple en météorologie, en océanographie et en hydrologie pour l'assimilation de données dans un modèle numérique (**Jazwinski, 1969**),

Harrison et Stevens (1975a, b) ont montré que toutes les méthodes de prédiction sont des cas spéciaux du FK. Ce Filtre permet des changements de modèle, de paramètres et de variances. La difficulté avec le FK est que plusieurs questions techniques n'ont pas encore été répondues d'une manière satisfaisante. L'approche elle-même ayant évolué au sein des applications dans l'engineering. Par conséquent, plusieurs statisticiens et chercheurs opérationnels connaissent peu sur lui, ou le trouvent difficile à comprendre par ce qu'il est souvent formulé dans l'espace d'état.

Peu d'études ont été réalisées sur le filtre de Kalman en combinaison avec des approches plus avancées de modèle conceptuel pluie débit, Les concepts de l'hydrogramme unitaire et de la précipitation antérieure ont été utilisés par **Husain (1985)** avec un filtre de Kalman standard pour simuler divers événements de crue.

Concernant la modélisation et la prédiction des processus d'écoulements souterrains ou de surface, **Bergman et Delleur(1985)** ont utilisé le filtre de Kalman pour la prédiction des écoulements journaliers, ils ont étudié le filtre de Kalman combiné avec les modèles ARMAX pour les écoulements journaliers, **Jimenez et al. (1989)** ont essayé moyennant le filtre de Kalman d'identifier des modèles PARMA (periodicautogressive-movingaverage) pour des écoulements mensuels

Sen(1991) a combiné le FK avec les séries Walsh orthogonales qui sont proposées comme un modèle effectif au calcul des périodicités dans les séries hydrologiques observées. Cette combinaison a conduit à une procédure de prédiction en temps réel des variables d'état qui sont de variables hydrologiques mensuelles.

Latif (1999), a développé une technique de FK et son application à des précipitations en Turquie. Son software reste applicable à n'importe qu'elle série de mesure de précipitations dans n'importe qu'elle région du monde.

Un filtre de Kalman standard a été intégré avec succès dans le modèle conceptuel pluie-débit IHACRES par **Schreider et al. (2001)**. **Quesney et al. (2000)** ont utilisé un filtre de Kalman étendu pour l'assimilation de mesures d'humidité de surface (par télédétection Radar ERS/SAR) dans un modèle hydrologique conceptuel global

Parmi les publications les plus récentes référant à l'application du filtre de Kalman nous pouvons citer à titre d'exemple **Hartnack&Madsen (2001)** où un filtre de Kalman d'ensemble a été utilisé pour améliorer la performance de la partie hydraulique du modèle MIKE 11 ; **Şen et al. (2004)**, Un modèle dynamique linéaire spatio-temporel a été développé par **Amisigo& van de Giesen (2005)** pour combler les lacunes dans des séries de débit journalier. Ils utilisent à cet effet un filtre de Kalman pour l'estimation des variables d'état ; **Moradkhani et al.(2005)**, qui utilisent une approche d'estimation paramètre-état basée sur le filtre de Kalman d'ensemble (EnKF) pour l'estimation séquentielle à la fois des paramètres et des variables d'états d'un modèle hydrologique conceptuel. **Troch et al. (2003)**, dans un essai sur les méthodes d'assimilation dans la modélisation hydrologique, affirment que l'accent devrait être mis dans les études futures sur l'intégration à l'échelle du bassin versant du filtre de Kalman à travers la modélisation pluie débit ; **Weerts et El Serafy (2006)** et **Ouachani et al. (2007)**.

En Algérie et jusqu'à ce jour-là, 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. Particulièrement dans le domaine de l'eau, 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, 2006**), et (**Boukharouba et Kettab, 2009**) mais qui sont restées inexploitées malgré la qualité des résultats obtenus . L'une des raisons majeures qui n'encouragent pas le recours au FK est justement cette notion d'état et la représentation du système dans son espace. Une autre difficulté réside dans la détermination des conditions initiales entre autres les matrices de covariance des erreurs du système et de la mesure. Cette thèse présente une contribution au comblement de cette lacune dans le domaine de la prédiction hydrologique en Algérie.

Dans ce modeste travail, nous nous intéressons à la quantité d'eau précipitée ainsi que sa distribution qui varie dans le temps et dans l'espace. Cette variabilité est due au concours de plusieurs facteurs climatiques et physiographiques assez complexes. La description et la prédiction de cette variabilité est un élément fondamental dans une large variété d'activités humaines, ainsi que dans l'élaboration et la conception de grands projets techniques et scientifiques. Pour la mise en œuvre de ce filtre, les séries chronologiques des précipitations mensuelles et annuelles enregistrées sur une période de 51ans (1959 à 2009) dans trente-neuf (39) stations pluviométriques dans le bassin versant de Cheliff ont été étudiées.

Les objectifs sont donc l'étude de l'applicabilité de la technique du FK discret à la modélisation et la prédiction multi-site de la pluviométrie annuelle et mensuelle dans le bassin versant de Chéiff, ainsi que l'amélioration des erreurs de prédiction.

Pour se faire, le travail est composé de deux parties :

- **La première partie** : c'est la partie théorique, elle est entièrement dédiée au filtre de Kalman proprement dit, où l'accent est mis sur le modèle mathématique en continu et en discret, sur l'algorithme de calcul et enfin sur les conditions de son application.
 - ✓ **Chapitre I.1** : Notions de base pour aborder le Filtre de Kalman
 - ✓ **Chapitre I.2** : Méthode du Filtre de Kalman

- **La seconde** : est la partie pratique qui se divise aussi en deux chapitres:
 - ✓ **Chapitre II.1** : Traitement des données
 - ✓ **Chapitre II.2** : Application du Filtre de Kalman

Le modèle recherché serait un moyen de prédiction en ligne qui non seulement procure des prédictions multi-sites optimales mais aussi des prédictions qui tiennent en compte de la nature dynamique de la pluviométrie elle-même.

Introduction

Le recours à un outil tel que le filtre de Kalman (FK), en vue d'une prédiction en hydrologie, requiert la connaissance de quelques notions de base sur lesquelles reposent les fondements théoriques de ce dernier. Dans ce sens, le présent chapitre peut être considéré comme une introduction au FK discret, où un rappel au sujet de ces quelques notions élémentaires semble être indispensable.

Ces notions regroupent des concepts mathématiques allant de la simple notion de probabilité et de variables aléatoires ainsi que leur caractérisation mathématique, aux processus stochastiques linéaires et leur modélisation dynamique. D'autres concepts tels que les séries chronologiques, les bruits blancs gaussiens, les modèles espace-état ainsi que le problème du filtrage linéaire seront également introduits.

I.1.1. PROBABILITE ET VARIABLES ALEATOIRES :

I.1.1.1.Loi de probabilité

Une loi de probabilité est une mesure, souvent vue comme la loi décrivant le comportement d'une variable aléatoire, discrète ou continue

Pour (Ω, A) un espace mesurable, P est une loi de probabilité, mesure de probabilité ou plus simplement probabilité si :

P est une application de A dans $[0,1]$; $P(\Omega)=1$;

P est Ω additive, c'est-à-dire pour toute famille finie ou dénombrable d'éléments disjoints $(A_i, i \in I)$ de A ;

$$P(\bigcup_{i \in I} A_i) = \sum_{i \in I} P(A_i) \dots\dots\dots(I.1.1)$$

I.1.1.2.Variable aléatoire réelle

Une variable aléatoire réelle sur l'espace probabilisé (Ω, A, P) , c'est-à-dire une fonction mesurable $X : (\Omega, A) \rightarrow (R, \beta(R))$; La loi de probabilité de la variable aléatoire X est la mesure de probabilité, notée P_X , définie sur l'espace mesurable $(R, \beta(R))$ par :

$$P_X(B) = P(X^{-1}(B)) = P(X \in B) \dots\dots\dots(I.1.2)$$

I.1.1.3.Loi de probabilité à plusieurs variables

La probabilité pour que la variable aléatoire X prenne une valeur numérique inférieure à x alors que Y prend une valeur inférieure à la référence d'équation est insérée à cet emplacement y définit la fonction de répartition :

$$F(x,y) = P(X < x, Y < y) \dots\dots\dots(I.1.3)$$

Celle-ci est non décroissante en x et en y entre la valeur 0 lorsque les deux variables tendent vers $-\infty$ et la valeur 1 lorsqu'elles tendent toutes deux vers $+\infty$.

La densité de probabilité jointe ou loi jointe s'obtient par une double dérivation :

$$f_{XY}(x, y) = \frac{\partial^2 F_{XY}}{\partial_x \partial_y}(x, y) \dots\dots\dots(I.1.4)$$

Une intégration par rapport à y (resp.x) donne la densité de probabilité marginale ou loi marginale de X(resp.Y) :

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y)dy \quad \dots\dots\dots(I.1.5)$$

Le rapport de la densité de probabilité jointe (relative à une valeur x) à la densité marginale de Y (concernant toutes les valeurs x) représente la densité de probabilité conditionnelle de X sous la condition Y=y :

$$f_{X|Y}(x, y) = \frac{f_{XY}(x, y)}{f_Y(y)} \quad \dots\dots\dots(I.1.6)$$

I.1.1.4.Densité de probabilité

Une densité de probabilité est une fonction qui permet de représenter une loi de probabilité sous forme d'intégrales.

$$\int_a^b f(x)dx \quad \dots\dots\dots(I.1.7)$$

Si la variable X a pour densité de probabilité la fonction f, la probabilité que la variable X soit dans l'intervalle [4,3; 7,8] sera

$$P(4,3 \leq X \leq 7,8) = \int_{4,3}^{7,8} f(x)dx \quad \dots\dots\dots(I.1.8)$$

Cela n'implique que l'intégrale de f :

$$\{f(x) \geq 0 \forall x\} \text{ et } \left\{ \int_{-\infty}^{\infty} f(x)dx = 1 \right\} \quad \dots\dots\dots(I.1.9)$$

I.1.1.5.Densité de probabilité d'une variable aléatoire réelle

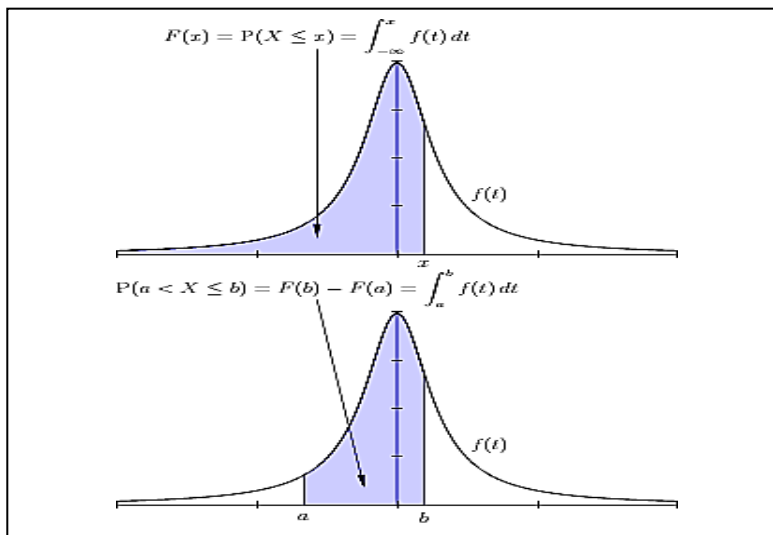


Fig I.1.1.Lien entre la densité f et la fonction de répartition et les probabilités.

On dit qu'une fonction f est une densité de probabilité d'une variable aléatoire réelle X si, pour tout réel x ;

$$f(x) = P(X \leq x) = \int_{-\infty}^x f(t)dt \quad \dots\dots\dots(\text{I.1.10})$$

La probabilité $P(a < X \leq b)$ se calcule alors par la relation suivante :

$$P(a < X \leq b) = \int_a^b f(t)dt \quad \dots\dots\dots(\text{I.1.11})$$

En traçant la représentation graphique de la densité de probabilité, la probabilité $P(a < X \leq b)$ se lit comme l'aire sous la courbe sur l'intervalle $[a \ b]$

En conséquence, la fonction de répartition F_X de X est continue, et $P(X = a) = 0$ pour tout nombre réel a . En cela, le comportement d'une variable à densité est très différent de celui d'une variable discrète. (J.P Laborde, (2003))

I.1.1.6. Espérance, variance et moments d'une variable aléatoire réelle à densité de probabilité

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

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

On a dans ce cas

$$E[X^k] = \int_{-\infty}^{\infty} t^k f(t)dt$$

En particulier, lorsque le moment d'ordre 2 existe :

$$E[X] = \int_{-\infty}^{\infty} t f(t)dt \quad \dots\dots\dots(\text{I.1.12})$$

$$E[X^2] = \int_{-\infty}^{\infty} t^2 f(t)dt \quad \dots\dots\dots(\text{I.1.13})$$

I.1.1.7. Densité de probabilité d'un vecteur aléatoire

On appelle densité de probabilité d'une variable aléatoire X à valeur dans \mathbb{R}^d une fonction f telle que pour toute partie $A \subset \mathbb{R}^d$,

$$P(X \in A) = \int_{\mathbb{R}^d} 1_A(t) f(t)dt = \int_A f(t)dt \quad \dots\dots\dots(\text{I.1.14})$$

Cette définition est en particulier valable pour $d=1$ et est donc équivalente à la première définition, dans le cas particulier $d=1$; Il existe une définition (équivalente) en termes d'espérance mathématique :

Si une fonction f est la densité de probabilité d'une variable aléatoire à valeur dans \mathbb{R}^d , cette fonction vérifie les propriétés suivantes

f est intégrable sur \mathbb{R}^d ; $\int_{\mathbb{R}^d} f(t)dt = 1$; f est presque sûrement positive ou nulle sur \mathbb{R}^d

Réciproquement, si une fonction f vérifie les 3 propriétés ci-dessus, on peut construire une variable aléatoire X à valeur dans \mathbb{R}^d ayant f pour densité de probabilité. Les variables aléatoires qui possèdent une densité de probabilité sont appelées parfois variables à densité, parfois variables continues.

I.1.1.8. Variables aléatoires discrètes

Une variable aléatoire est dite discrète si elle ne prend que des valeurs discontinues dans un intervalle donné (borné ou non borné). L'ensemble des nombres entiers est discret. En règle générale, toutes les variables qui résultent d'un dénombrement ou d'une numération sont de type discret. (J.P Laborde, (2003))

I.1.1.9. 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

$$F_X(x) = P(X \leq x) \dots\dots\dots(I.1.15)$$

Où le membre de droite représente la probabilité que la variable aléatoire réelle X prenne une valeur inférieure ou égale à x La probabilité que X se trouve dans l'intervalle $]a, b]$ est donc, si $a < b$

$$P(a < X \leq b) = F_X(b) - F_X(a) \dots\dots\dots(I.1.16)$$

La fonction de répartition d'une mesure de probabilité P définie sur la tribu borélienne $B(\mathbb{R})$ est la fonction F qui à tout réel x associe

$$F_X = P(]-\infty, x]) \dots\dots\dots(I.1.17)$$

La fonction de répartition d'une variable aléatoire X a les propriétés caractéristiques suivantes :

1. F_X est croissante ;
2. Elle est partout continue à droite ;
3. $\lim_{x \rightarrow -\infty} F_X(x) = 0$
4. $\lim_{x \rightarrow +\infty} F_X(x) = 1$

I.1.1.10. Probabilité conditionnelle

La probabilité conditionnelle d'un événement A , sachant qu'un autre événement B de probabilité non nulle s'est réalisé (ou probabilité de A , sachant B) est le nombre noté $P(A|B)$ défini par :

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \dots\dots\dots(I.1.18)$$

Le réel $P(A|B)$ se lit « probabilité de A, sachant B » ou $P_B(A)$.

Mathématiquement, soient (Ω, β, P) , un espace probabilisé et B un événement de probabilité non nulle. À tout événement A de β , nous associons le nombre noté $P(A|B)$ ou $P_B(A)$ défini par:

$$P_B(A) = \frac{P(A \cap B)}{P(B)} \dots\dots\dots(I.1.19)$$

Nous pourrions vérifier que l'application P_B définie par $A \rightarrow P_B(A)$ est une probabilité.

Si A et B sont indépendants alors:

$$P(A|B) = P(A) \dots\dots\dots (I.1.20)$$

D'autre part, on peut noter que le théorème de Bayes permet d'écrire

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \dots\dots\dots(I.1.21)$$

Ce qui, en cas d'indépendance des événements A et B, se vérifie par: $P(A) = P(B) \frac{P(A)}{P(B)}$ ce qui est vrai.

I.1.1.11. Densité conditionnelle

Soit (Ω, β, P) , et soient X et Y deux variables aléatoires définies sur cet espace. Si l'on suppose que leur loi jointe puisse être définie par une densité bi-variable $f(x, y)$, et si de plus un y_0 vérifie $\int f(x, y_0)dx \neq 0$, alors il existe une loi absolument continue dont la densité est donnée par l'expression

$$g(x|y_0) = \frac{f(x, y_0)}{\int f(x, y_0)dx} \dots\dots\dots(I.1.22)$$

Cette fonction $g(x|y_0)$ est appelée : densité conditionnelle de X sachant $Y = y_0$. Intuitivement, cette expression peut être interprétée comme une formulation continue du théorème de Bayes.(J.P Laborde, (2003))

I.1.2. RAPPELLE STATISTIQUE ET METHODES D'ESTIMATION

En hydrologie, nous ne connaissons jamais la population totale mais nous disposons d'un échantillon non exhaustif tiré de cette population. C'est à partir de cet échantillon que nous choisirons la loi de probabilité adéquate et d'en calculer le mieux que possible les paramètres numériques de cette loi Laborde, (2003).

I.1.2.1. Rappels statistiques

A- Paramètres de l'échantillon

A. 1. Paramètres de tendance centrale

1. La moyenne arithmétique

$$\bar{X} = \frac{1}{N} \cdot \sum_{i=1}^N X_i \dots\dots\dots(\text{I.1.23})$$

N : le nombre d'observation composant l'échantillon

2. Le mode :

Le mode M_0 correspond à la classe de fréquence maximale dans la distribution des fréquences. On peut identifier le mode comme la valeur médiane de la classe de fréquence maximale ou bien effectuer une interpolation linéaire pour obtenir la valeur exacte du mode comme suit :

$$M_0 = x_m + \frac{i\Delta_i}{\Delta_s + \Delta_i} \dots\dots\dots(\text{I.1.24})$$

x_m : Limite inférieure de la classe d'effectif maximal

i : intervalle de classe ($x_{m+1} - x_m$)

Δ_i : Ecart d'effectif entre la classe modale et la classe inférieure la plus proche

Δ_s : Ecart d'effectif entre la classe modale et la classe supérieure la plus proche

3. La médiane :

La médiane M_e est la valeur du caractère pour laquelle la fréquence cumulée est égale à 0,5 ou 50%. Elle correspond donc au centre de la série statistique classée par ordre croissant, ou à la valeur pour laquelle 50% des valeurs observées sont supérieures et 50% sont inférieures.

$$M_e = x_m + (x_{m+1} - x_m) \left[\frac{\frac{n}{2} - N_i}{n_i} \right] \dots\dots\dots(\text{I.1.25})$$

x_m : limite inférieure de la classe dans laquelle se trouve le $ne/2$ individu.

x_{m+1} : limite supérieure de la classe dans laquelle se trouve le $ne/2$ individu.

n_i : effectif de la classe médiane

N_i : effectif cumulé inférieur à x_m

n : taille de l'échantillon

A. 2. Paramètres de dispersion

• L'intervalle total de l'échantillon exprime la différence entre les valeurs extrêmes.

$$R = X_{\max} - X_{\min}$$

• Déviation moyenne : c'est la moyenne arithmétique des valeurs absolues de la différence de chaque X_i de la moyenne \bar{X} .

$$D = \frac{1}{N} \sum_{i=1}^N |X_i - \bar{X}| \dots\dots\dots(\text{I.1.26})$$

4. La variance :

$$S^2 = \frac{1}{N} \sum_{i=1}^N (X_i - \bar{X})^2 \dots\dots\dots(\text{I.1.27})$$

5. L'écart type :

$$S = \sqrt{\frac{1}{N} \sum_{i=1}^N (X_i - \bar{X})^2} \dots\dots\dots(\text{I.1.28})$$

6. Le coefficient de variation :

$$Cv = \frac{S}{\bar{X}} \dots\dots\dots(\text{I.1.29})$$

B- Paramètres de forme

Ces paramètres caractérisent l'aplatissement et l'asymétrie.

B-1 : Mesure de dissymétrie : (selon Pearson)

$$Cs = \frac{\left[\frac{1}{N} \cdot \sum_{i=1}^N (X_i - m^{1/3}) \right]}{\left[\frac{1}{N} \cdot \sum_{i=1}^N (X_i - \bar{X})^2 \right]^{3/2}} \dots\dots\dots(\text{I.1.30})$$

1. Si Cs est positif la distribution est étalée sur la droite, on observe la succession mode - médiane - moyenne la dissymétrie est dite positive.
2. Si Cs est négatif la distribution est étalée sur la gauche, on observe la succession moyenne - médiane - mode la dissymétrie est dite négative.
3. Si Cs est nul, la distribution n'est forcément symétrique.

B-2 : Mesure de d'aplatissement : (selon Pearson)

Pour mesurer l'aplatissement de la courbe, on utilise le **coefficient de Pearson** basé sur le moment centré d'ordre 4:

$$\beta_2 = \frac{\mu_4}{\sigma^4} = \frac{\sum_{i=1}^k f_i (x_i - \bar{x})^4}{\left[\sum_{i=1}^k f_i (x_i - \bar{x})^2 \right]^2} \dots\dots\dots(\text{I.1.31})$$

1. $\beta_2 > 3$ courbe hypernormale

- 2. $\beta_2 = 3$ courbe normale
- 3. $\beta_2 < 3$ courbe hyponormale

I.1.2.2. Tests Relatifs à L'échantillonnage

Des tests statistiques sont généralement indispensables pour vérifier l'homogénéité, l'indépendance, la stationnarité et, parfois, la détermination des points singuliers (Bobée, 1990).

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

Soit une série complète divisée en deux échantillons de taille n_1 et n_2 . La date des séparations étant choisie a priori. Les deux échantillons sont classés par ordre croissant.

Nous considérons la statistique suivante :

$$V = R - n_1(n_2 + 1)/2$$

$$W = n_1 n_2 - V$$

Avec :

R : la somme des rangs des éléments de premier échantillon (taille n_1) et V représentent le nombre de fois des individus du premier échantillon (n_1) qui existe dans le deuxième échantillon (taille n_2).

$n = n_1 + n_2$ avec $n_1 < n_2$ puis on classe l'échantillon global de taille n par ordre croissant.

La statistique U de Mann-Whitney, est définie par le minimum de (V, U), U est supposé suivre une loi normale de moyenne et de variance.

$$\bar{U} = n_1 n_2 / 2, \quad \text{Var}(U) = \left[\frac{n_1 n_2}{n(n-1)} \right] \left[\frac{n^3 - n_2}{12} - \sum T \right] \dots\dots\dots(\text{I.1.32})$$

Avec :

$T = (J^3 - J)/12$, où J est le nombre d'observation lié à un rang donné.

$\sum T$: représente toutes les observations qui sont liées à des rangs pour les deux échantillons, l'hypothèse d'homogénéité et de stationnarité pour un seuil de signification α % stipule que l'indicateur :

$$\left| \frac{U - \bar{U}}{\sqrt{\text{Var}(U)}} \right| \leq U_{\alpha/2} \dots\dots\dots(\text{I.1.33})$$

B- Test d'Anderson

$$R' = \left[\frac{\left(\sum_{i=1}^{n-1} X_i - X_{i+1} + X_n X_1 \right) - \left(\sum_i X_i \right)^2 / n}{\sum_i X_i^2 - \left(\sum_i X_i \right)^2 / n} \right] \dots\dots\dots(\text{I.1.34})$$

On peut écrire : $R' = \frac{R - \frac{S_1^2}{n}}{S_2 - \frac{S_1^2}{n}} \dots\dots\dots(\text{I.1.35})$

De moyenne : $\bar{R}' = -\frac{1}{(n-1)}$ (I.1.36)

Et de variance : $\text{var}(R') = \frac{R' - \bar{R}'}{\sqrt{\text{var}(R')}} \dots\dots\dots(I.1.37)$

$$U = \frac{R' - \bar{R}'}{\sqrt{\text{var}(R')}}$$

- La valeur statistique suit une loi normale centrée réduite.
- Si $U < U_{\alpha/2}$. On dit que l'hypothèse H_0 d'indépendance est vraie.

C- Test du cumul des résidus

C'est une méthode complexe, mais beaucoup plus efficace.

Soient $\bar{X}, \bar{Y}, S_x^2, S_y^2, r$ respectivement les moyennes, la variance, le coefficient de corrélation empirique des séries à vérifier X_t, X et Y_t .

Le résidu ϵ_i est l'écart entre la valeur X_i et la valeur correspondante à Y_i dans la régression linéaire de X et Y :

$$\epsilon_i = X_i - \bar{X} - r \frac{S_x}{S_y} (Y_i - \bar{Y}) \dots\dots\dots(I.1.38)$$

Nous considérons la courbe chronologique cumulée des résidus successifs :

$$Z_i = \sum_{i=1}^t \epsilon_i \dots\dots\dots(I.1.39)$$

Soient les cumuls partiels :

$$T_{t,t+m} = \sum_{i=1}^{t+m} \epsilon_i \dots\dots\dots(I.1.40)$$

Soit l'ellipse de contrôle :

$$Z(m) = +U_{\alpha/2n} \sqrt{m(n-m)(n-1) \frac{S_{y \cdot \sqrt{1-r^2}}}{n}} \dots\dots\dots(I.1.41)$$

Où :

$U_{\alpha/2n}$ Est la valeur de variable.

m : la distance verticale entre deux observations dans la série chronologique normale centrée réduite dont la probabilité de dépassement est

$$\frac{X}{2n} (\alpha = 10\% \text{ ou } 5\%)$$

En déplaçant cette ellipse, dont le grand axe est parallèle à l'axe des abscisses, de façon que l'un de ses sommet coïncide avec les points successifs de la courbe du cumule des résidus, on détecte les écarts $Z_{t,t+m}$ qui sortent des limites de l'ellipse.

D- Test de Buishand

Une rupture peut-être définie par un changement dans la loi de probabilité des variables aléatoires dont les réalisations successives déterminent les séries chronologiques étudiées. Pour déceler d'éventuel changement dans le régime pluviométrique, nous avons utilisé le test

de corrélation sur le rang, les tests statistiques de Pettit, la statistique de Lee et Héghinian (Lee et Héghinian, 1979), le test de Buishand (H. Lubes-Niel, 1998) et la segmentation de Hubert (Hubert et al., 1993). La puissance et la robustesse de ces tests ont fait l'objet d'une revue de méthodes par Lubès-Niel et al. (1998) et dont la dernière a fait l'objet d'une évaluation spécifique avec un recul de 10 ans (Hubert et al, 1998).

La procédure de Buishand est de nature bayésienne et fait référence au modèle simple qui suppose un changement de moyenne de la série. Des valeurs critiques de la statistique U sont données par Buishand (1982) à partir d'une méthode de Monte Carlo. Pour tout changement de moyenne survenant au milieu de la série, la statistique U s'avère performante. La statistique U est robuste et qui reste valide même pour des distributions de la variable étudiée qui s'écartent de la normalité. Les tests sont dits robustes lorsque leurs conditions d'application sont peu strictes. Le test de Buishand se réfère au même modèle et mêmes hypothèses que l'approche de Lee et Héghinian. En cas de rejet de l'hypothèse nulle, aucune estimation de la date de rupture n'est proposée par ce test.

I.1.2.3.Choix Du Modèle Fréquentiel

La validité des résultats d'une analyse fréquentielle dépend du choix du modèle fréquentiel et plus particulièrement de son type. Diverses pistes peuvent contribuer à faciliter ce choix, mais il n'existe malheureusement pas de méthode universelle et infaillible. (Didier Auroux, 2003)

A- Distributions Théoriques

A. 1. Loi normale ou loi de Gauss

La loi normale se justifie, théoriquement par le théorème central-limite, comme la loi d'une variable aléatoire formée de la somme d'un grand nombre de variables aléatoires. En hydrologie fréquentielle des valeurs extrêmes, les distributions ne sont cependant pas symétriques, ce qui constitue un obstacle à son utilisation. Cette loi applique toutefois généralement bien à l'étude des modules annuels des variables hydrométéorologiques en climat tempéré.

A. 2. Loi log-normale

La loi log-normale est préconisée par certains hydrologues dont l'apparition d'un événement hydrologique résulte de l'action combinée d'un grand nombre de facteurs qui se multiplient. Dès lors la variable aléatoire $X=X_1 \cdot X_2 \cdot X_3 \cdot X_4 \dots X_n$. Suit une loi log-normale. En effet, le produit de (x) variables se ramène à la somme de (x) logarithmes de celles-ci et le théorème central-limite permet d'affirmer la log-normalité de la variable aléatoire.

A. 3. Loi de Gumel ou doublement exponentielle

Gumel postule que la loi double exponentielle, ou loi de Gumel, est la forme limite de la distribution de la valeur maximale d'un échantillon de n valeurs. Le maximum annuel d'une

variable étant considéré comme le maximum de 365 valeurs journalières. Cette loi doit ainsi être capable de décrire les séries de maxima annuels.

Il est à remarquer que plus le nombre de paramètres d'une loi n'est grand, plus l'incertitude dans l'estimation n'est importante. Pratiquement il est par conséquent préférable d'éviter l'utilisation de lois à trois paramètres ou plus.

I.1.2.4. Association Entre Variables

L'analyse de régression peut être définie comme la recherche de la relation stochastique qui lie deux ou plusieurs variables entre elles.

La corrélation, pour sa part, définit un coefficient permettant de mesurer le degré de liaison ou d'intensité de la relation entre deux variables.

On peut définir une relation linéaire entre deux variables X et Y par la relation suivante:

$$Y_i = a + bX_i \dots\dots\dots(I.1.42)$$

Avec :

X : variable indépendante (variable explicative).

Y: variable dépendante (variable à expliquer).

a et b sont des constantes, appelées paramétriques. La relation entre les deux variables est représentée par l'ensemble des couples (Xi, Yi), qui constitue toutes les valeurs de X et de Y satisfaisant l'équation. Il s'agit donc d'une relation de correspondance, une équation linéaire décrit une relation linéaire

b. Modèle linéaire général

Si on cherche à expliquer une variable Y en fonction de K-1 variable X₂, X₃,.....X_k, par une équation avec n observations concomitantes de ces K variables, le modèle s'écrira comme suit :

$$\left\{ \begin{array}{l} Y_1 = a_1 + a_2 \cdot x_{21} + \dots\dots\dots + a_k \cdot x_{k1} + \varepsilon_1 \\ Y_2 = a_1 + a_2 \cdot x_{22} + \dots\dots\dots + a_k \cdot x_{k2} + \varepsilon_2 \\ \dots\dots\dots \\ Y_n = a_1 + a_2 \cdot x_{2n} + \dots\dots\dots + a_k \cdot x_{kn} + \varepsilon_n \end{array} \right.$$

Où : ε₁, ε₂.....ε_n terme d'erreur qui prend en considération les influences aléatoires sur Y_i.

Soit le modèle suivant :

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \dots\dots\dots \\ Y_n \end{bmatrix} \quad X = \begin{bmatrix} 1 \dots\dots\dots x_{21} \dots\dots\dots x_{k1} \\ 1 \dots\dots\dots x_{22} \dots\dots\dots x_{k2} \\ \dots\dots\dots \\ 1 \dots\dots\dots x_{2n} \dots\dots\dots x_{kn} \end{bmatrix}$$

$$A = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \quad e = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \quad I = \begin{bmatrix} 10\dots\dots\dots 0 \\ 01\dots\dots\dots 0 \\ \vdots \\ 00\dots\dots\dots 1 \end{bmatrix}$$

Donc on aura le modèle de régression sous forme matricielle suivant :

$$Y = X.a + e \quad \dots\dots\dots(I.1.43)$$

Nous estimons les termes de la matrice A en appliquant la méthode des moindres carrés mais avec certaines hypothèses qui sont les suivantes :

- N > K
- X de rang K
- E(e) = 0
- E (e.e') = σ².I

• **Identification des paramètres du modèle**

En minimisant la somme des carrés des écarts min (e.e') et en développant l'expression matricielle de l'équation précédente, on aura :

L'estimateur Â égal :

$$\hat{A} = (X'.X)^{-1}.X'.Y \quad \dots\dots\dots(I.1.44)$$

Et que la variance de Â est la suivante :

$$Var(\hat{A}) = (X'.X)^{-1}.\sigma^2 \quad \dots\dots\dots(I.1.45)$$

Les valeurs calculées par la méthode linéaire.

$$\hat{Y} = \hat{A}.X$$

Les différences entre les termes de Y et ceux de Ŷ qui sont des réalisations des erreurs « e » sont appelées des résidus et sont notées e.

Alors : $Y = \hat{Y} + e$

L'estimation de l'écart type peut être donnée par la relation suivante :

Et l'écart type : $S^2 = \frac{e'.e}{n-K}$ et par conséquent l'écart type $S = \sqrt{S^2}$
 e' : l'erreur estimée

$$R^2 = 1 - \frac{e'.e}{Y'.Y - \frac{(\sum Y_i)^2}{n}}$$

Le coefficient de corrélation sera calculé par la formule :

Ou alors :

$$R^2 = \frac{\hat{Y}'.Y - \frac{(\sum Y_i)^2}{n}}{Y'.Y - \frac{(\sum Y_i)^2}{n}} \quad \dots\dots\dots(I.1.46)$$

Si la variable « e » suit une loi de Gauss, on peut même tester l'hypothèse qu'il existe une relation entre Y et les variable X_i.

Pour le coefficient de corrélation multiple, on considère que R est significatif si la variance expliquée est significativement supérieure à la variance résiduelle. L'analyse de ces variances montre que la quantité :

$$F = \frac{n - (K + 1)}{k} \cdot \frac{R^2}{1 - R^2} \dots\dots\dots(I.1.47)$$

Doit être significativement supérieure à 1.

Les valeurs de F ayant une probabilité P d'être dépassée du seul fait du hasard sont données par la table de Snedecor avec des degrés de liberté V₁=K et V₂=n-(k+1)

c. Modèle à une variable explicative

Ce cas est présenté comme un cas particulier du cas général pour un (K =2).

Si l'on a, à notre disposition, « n » observations donc le modèle se présentera comme suit :

$$\begin{cases} Y_1 = a + bx_1 + \varepsilon_1 \\ Y_2 = a + bx_2 + \varepsilon_{21} \\ \dots \\ Y_n = a + bx_n + \varepsilon_n \end{cases}$$

Les hypothèses sont les suivantes :

$E(\varepsilon_i) = 0, E(\varepsilon_i)^2 = \sigma^2, E(\varepsilon_i, \varepsilon_j) = 0,$ tel que $(j \neq i)$

Posons :

$$\begin{aligned} \bar{X} &= \frac{1}{n} \sum_{i=1}^n x_i & ; & & \bar{Y} &= \frac{1}{n} \sum_{i=1}^n y_i \\ S_x^2 &= \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{X}^2 & ; & & s_y^2 &= \frac{1}{n} \sum_{i=1}^n y_i^2 - \bar{Y}^2 \end{aligned}$$

Et le coefficient de corrélation entre x et y s'écrit :

$$r = \frac{\frac{1}{n} \sum x_i \cdot y_i - \bar{X} \cdot \bar{Y}}{S_x \cdot S_y} = \frac{\text{Cov}(x, y)}{S_x \cdot S_y} \dots\dots\dots(I.1.48)$$

A titre indicatif, nous calculons les paramètres de régression par la méthode des moindres carrés, qui consiste à minimiser la somme des carrés des points par rapport à droit de régression.

Et aura :

$$b = \frac{\sum_{i=1}^n \Delta x_i \cdot y_i}{\sum_{i=1}^n \Delta x_i^2} \dots\dots\dots(I.1.49)$$

• **Test d'auto corrélation des résidus**

Ce test est établi par J.Durbin et G.S.Watson. La fonction discriminante du test est la statistique « d » de Durbin-Watson consiste à calculer la statistique d :

$$d = \frac{\sum_{i=1}^n (e_i - e'_i - 1)}{\sum_{i=1}^n e_i^2} \dots\dots\dots(\text{I.1.50})$$

Par ailleurs la statistique « d » est pratiquement égale à : $d \approx 2.(1 - \rho)$
 ρ est le coefficient d'auto corrélation au pas 1 des résidus

$$\rho = \frac{\sum_{i=1}^n (e_i \cdot e_i^2 - 1)}{\sum_{i=1}^n e_i^2} \dots\dots\dots(\text{I.1.51})$$

Les tables de Dubrin-Watson permettent de déterminer deux quantités d1 et d2 en fonction du nombre de variables explicatives et du nombre d'observations.

- $d < d_1$ Auto corrélation positive.
- $d_1 \leq d < d_2$ Doute.
- $d_2 < d < 4 - d_2$ Pas d'auto corrélation.
- $4 - d_2 \leq d \leq 4 - d_1$ Doute.
- $d > 4 - d_1$ Auto corrélation négative.

• **Test de Bruit Blanc (Test de Box-Pierce)**

Lorsque le processus est bien estimé, les résidus entre les valeurs observées et les valeurs estimées par le modèle doivent se comporter comme un bruit blanc. On notera par la suite $\hat{\varepsilon}_t$ le résidu d'estimation du modèle.

Si les résidus $(\varepsilon_t, t \in \mathbb{Z})$; obéissent à un bruit blanc, il ne doit pas exister d'auto-corrélation dans la série.

On note γ_k l'auto-corrélation d'ordre k du processus $(\varepsilon_t, t \in \mathbb{Z})$: Pour un ordre K, le test de Box-Pierce est le test de l'hypothèse $H_0 : \gamma_1 = \dots = \gamma_k = 0$ contre $H_1 : \exists j \in [1; k]$; tel que $\gamma_j \neq 0$, la statistique de ce test est :

$$Q_{BP} = T \sum_{k=1}^K \gamma_k^2 \xrightarrow{\varepsilon} \chi^2 (K - p - q) \dots\dots\dots(\text{I.1.52})$$

L'hypothèse H_0 est rejetée au seuil de 5% si Q_{BP} est supérieur au quantile 0,95 de la loi du χ^2 correspondant.

d. Modèles non linéaires

Ces modèles reposent sur le changement de variables du modèle linéaire dans le but d'obtenir une droite entre les points (x_i, y_i) .

La transformation inverse

Le modèle $Y = a \cdot \frac{1}{U} + b + \varepsilon \dots\dots\dots(\text{I.1.53})$

En posant $\frac{1}{U} = x$

• **La transformation semi- logarithmique**

Le modèle $Y = a + b \cdot \text{Log}U + \varepsilon$ (I.1.54)

En posant $X = \text{Log}U$

Devient $Y = a \cdot b \cdot X + \varepsilon$

• **La transformation log- log**

Le modèle $V = K \cdot U + \varepsilon$ (I.1.55)

En prenant le logarithme des deux membres :

$$\text{Log}V = \text{Log}K + a \cdot \text{Log}U + \text{Log}\varepsilon_i$$

Si on pose ; $y = \text{Log} V$ et $x = \text{Log} U$

On aura : $Y = b + a \cdot x + \eta_i$

• **La transformation racine** $\sqrt{\quad}$

Le modèle s'écrit : $y = a + b \sqrt{Z} + \varepsilon$ (I.1.56)

Si on pose : $x = \sqrt{Z}$

On aura : $y = a + b \cdot x + \varepsilon$

I.1.2.5. Analyse multidimensionnelle

L'analyse à plusieurs variables ou bien l'analyse multi-variée, est utilisée dans le traitement de l'information. Elle est définie comme l'ensemble des méthodes statistiques chargées de l'étude des relations qui existent entre plusieurs variables, dépendantes ou interdépendantes. Les objectifs de cette méthode consistent à :

- Résumer l'information apportée par un grand nombre de variables ;
- Adapter et développer des méthodes d'analyse (A.P.C, Analyse factorielle, classification hiérarchique) ;
- L'aide à une prise de décision ;
- Modéliser les phénomènes présentant une structure temporelle ou spatiale, ou les deux à la fois (Lachance, 1995)

A- Analyse en composantes principales (A.C.P)

Selon **Philippeau (1992)**, l'A.C.P est une méthode qui vise à mettre en évidence graphiquement l'information essentielle contenue dans un tableau des données.

• **Données**

Soit un ensemble de n_v variable indicées 'j', et n_o observation indicées « i », sur ces variables, les données se présentent sous la forme d'une matrice [X] à n_o ligne et n_v colonnes.

$$[X] = \begin{vmatrix} X(1,1) & \dots & X(1,j) & \dots & X(1,nv) \\ \dots & & & & \\ X(i,1) & \dots & X(i,j) & \dots & X(i,nv) \\ \dots & & & & \\ X(n_0,1) & \dots & X(n_0,j) & \dots & X(n_0,nv) \end{vmatrix}$$

Pour chaque variable, on peut évaluer sa moyenne $MX(j)$

et son écart –type $SX(j)$

$$[MX] = |MX(1) \dots MX(j) \dots MX(nv)|$$

$$[SX] = |SX(1) \dots SX(j) \dots SX(nv)|$$

Les calculs peuvent être effectués sur les variables observées et comme elles ne sont pas toujours exprimées dans les mêmes unités donc elles n’ont pas le même poids. Ainsi on procède au centrage de ces variables afin de leurs donner la même importance, en les transformant en une matrice $[x]$ telle que chaque élément est exprimé $x(i,j)$, soit :

$$x(i, j) = \frac{X(i, j) - MX(j)}{SX(j)} \dots \dots \dots (I.1.57)$$

Ainsi toutes les variables ont la même moyenne ($m=0$) et le même écart type (variable réduite ($\sigma=1$)).

Les données transformées sont présentées comme suit :

$$[x] = \begin{vmatrix} x(1,1) & \dots & x(1,j) & \dots & x(1,nv) \\ \dots & & & & \\ x(i,1) & \dots & x(i,j) & \dots & x(i,nv) \\ \dots & & & & \\ x(n_0,1) & \dots & x(n_0,j) & \dots & x(n_0,nv) \end{vmatrix}$$

et :

$$[mx] = |0 \dots 0 \dots 0|$$

$$[sx] = |1 \dots 1 \dots 1|$$

B- Principe et mise en œuvre

L’analyse en composantes principales sur la matrice de coefficient de corrélation $[R]$ entre les variables de la matrice $[x]$ se propose de réduire la matrice $[x]$ en une matrice $[c]$ en construisant un certain nombre nc de variables fictives (variables indépendantes) appelées composantes principales ; soit :

$$[X] = \begin{vmatrix} x(1,1) & \dots & x(1,j) & \dots & x(1,nv) \\ \dots & & & & \\ x(i,1) & \dots & x(i,j) & \dots & x(i,nv) \\ \dots & & & & \\ x(n_0,1) & \dots & x(n_0,j) & \dots & x(n_0,nv) \end{vmatrix} \Rightarrow [C] = \begin{vmatrix} c(1,1) & \dots & c(1,nc) \\ \dots & & \\ c(i,j) & \dots & c(i,nc) \\ \dots & & \\ c(n_0,1) & \dots & c(n_0,nc) \end{vmatrix}$$

Les composantes principales présentent les caractéristiques suivantes :

Chaque composante indiquée K est une combinaison linéaire des nv variables de $[x]$:

$$C_{ik} = \sum_1^{nv} a_{ik} \cdot x_{ij} \dots\dots\dots (I.1.58)$$

C_{ik} : composantes principales.

x_{ij} : observation réduite centrée.

a_{ij} ($j=1$ à nv , $k=1$ à nc) : Cosinus directeur, a_{ij} sont les coordonnées des vecteurs propres de la matrice de coefficient $[R]$ obtenues par diagonalisation de cette matrice.

$$R = \begin{vmatrix} 1 \dots\dots r(1,2) \dots\dots r(1,nv) \\ \dots 1 \dots r(2,3) \dots\dots r(2,nv) \\ \dots\dots\dots 1 \dots\dots\dots \\ \dots\dots\dots 1 \dots\dots\dots \\ \dots\dots\dots 1 \dots\dots\dots \\ \dots\dots\dots 1 \dots\dots\dots \end{vmatrix}$$

Où :

R : matrice de coefficient de corrélation entre observations.

Les valeurs propres λ de cette matrice ne sont autres que les variances des nouvelles variables

C_{ik} :

$$\lambda_k = \frac{1}{n_0} \sum_{i=1}^{n_0} C_{ik}^2 \dots\dots\dots$$

(I.1.56)

Chaque composante est orthogonale aux autres (les composantes sont indépendantes entre elles).

$$C(1,k)C(1,1)+C(2,K)C(2,1)+\dots\dots\dots C(i,k)C(i,1)+C(no,k)C(nv,k)=0$$

Si on impose à chaque composante d'avoir une variance unité

$$C_{ik} = \frac{1}{\sqrt{\lambda_K}} \sum_1^{nv} a_{ik} \cdot x_{ij} \dots\dots\dots (I.1.59)$$

On aura donc une série synthétique de moyenne nulle ($m=0$) et une variance ($\sigma=1$) que qualifiera de vecteur régional soit :

$$[x] = \begin{vmatrix} x(1,1) \dots\dots\dots x(1, j) \dots\dots\dots x(1, nv) \\ \dots\dots\dots \\ x(i,1) \dots\dots\dots x(i, j) \dots\dots\dots x(i, nv) \\ \dots\dots\dots \\ x(no,1) \dots\dots\dots x(no, j) \dots\dots\dots x(no, nv) \end{vmatrix} \Rightarrow [c] = \begin{vmatrix} c(1,1) \dots\dots\dots c(1, nc) \\ \dots\dots\dots \\ c(i, j) \dots\dots\dots c(i, nc) \\ \dots\dots\dots \\ c(no,1) \dots\dots\dots c(no, nc) \end{vmatrix}$$

$$[mx] = |0 \dots\dots\dots 0 \dots\dots\dots 0| \Rightarrow [mc] = |0 \dots\dots\dots 0|$$

$$[sx] = |1 \dots\dots\dots 1 \dots\dots\dots 1| \Rightarrow [sc] = |\sqrt{\lambda_1} \dots\dots\dots \lambda_2|$$

- Les termes $a(k,j)$ donc seront les coefficients de corrélation entre la $i^{\text{ème}}$ variable et la $k^{\text{ème}}$ composante .
- Chaque terme $a(k,j)$ sera appelé : Projection de la variable j sur la composante k .
- Chaque terme $c(i,k)$ sera appelé : Projection de la $i^{\text{ème}}$ observation sur la k composante.

Alors on peut reconstituer chaque terme $x(i,j)$, par son estimation $x_e(i,j)$ de la régression linéaire multiple.

$$x_e(i,j) = a(1,j)C(i,1) + a(2,j)C(i,2) + \dots + a(k,j)C(i,k) + \dots + a(nc,j)C(i,nc).$$

On peut aussi passer aux valeurs de départ $X(i,j)$ en les estimant par $X_e(i,j)$:

$$X_e(i,j) = x_e(i,j) S_x(j) M_x(j).$$

Remarque : les composantes sont non corrélées entres-elles : $E(C_i * C_j) = 0 \quad i \neq j$

I.1.3. PROCESSUS STOCHASTIQUE ET SERIES CHRONOLOGIQUES

I.1.3.1. Processus Aléatoire

Un processus aléatoire est une suite de variables aléatoires indexées dans le temps et définies sur un espace des états de la nature. Ainsi, pour chaque instant du temps, la valeur de la quantité étudiée X_t est appelée variable aléatoire et l'ensemble des valeurs X_t quand t varie est appelé processus.

I.1.3.2. Processus Aléatoire stationnaire

Un processus est un ensemble X_t de fonctions ordinaires x_t , chacune d'elles étant une réalisation du processus. On peut caractériser ce processus en lui associant à chaque instant t_0 une densité de probabilité $P_x(x, t_0)$.

À une réalisation donnée on peut associer les moyennes temporelles :

$$\overline{x(t)^n} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)^n dt \quad \dots \dots \dots (I.1.60)$$

À la densité de probabilité $P_x(x, t_0)$ on peut associer les moments appelés moyennes d'ensemble

$$E[x(t_0)^n] = \int_{-\infty}^{\infty} x^n P_x(x, t_0) dx \quad \dots \dots \dots (I.1.61)$$

Si ces moyennes d'ensemble, et par conséquent la densité de probabilité, ne dépendent pas de l'instant t_0 , on parle de processus stationnaire. Si, de plus, les moyennes temporelles leur mesures d'ensembles sont égales, il s'agit d'un processus non stationnaire.

En fait, il ne s'agit là que des propriétés au premier ordre, les propriétés aux ordres supérieurs faisant intervenir la densité de probabilité jointe à des instants différents. Elles impliquent également des moyennes temporelles et des moyennes d'ensemble. Parmi ces dernières, la plus importante est l'auto-covariance statistique $E[x(t_0)X(t_0 + \tau)]$. Si le processus est

stationnaire et non stationnaire au second ordre, elle est identique à l'auto-covariance temporelle, elle-même équivalente à la densité spectrale

La stationnarité au second ordre est suffisante pour assurer la stationnarité forte lorsque le processus peut être supposé gaussien. (J.P Laborde, (2003)).

- Le processus aléatoire $X(t)$ sera dit stationnaire au sens faible, s'il aura une moyenne et une variance qui ne changeront pas avec le temps, et si la covariance entre les valeurs du processus en deux points dans le temps ne va dépendre que de la distance entre les points

dans le temps et non du temps lui-même. C'est à dire : un processus est stationnaire au second ordre si l'ensemble de ses moments sont indépendants du temps.

I.1.3.3. Processus Bruit Blanc

Parmi la classe des processus stationnaires, il existe des processus particuliers que sont les processus bruit blanc (ou White Noise). Ces processus sont très souvent utilisés en analyse des séries temporelles, car ils constituent en quelque sorte les « briques élémentaires » de l'ensemble des processus temporels. Tout processus stationnaire peut s'écrire comme une somme pondérée de bruits blancs (théorème de Wold)

- Un premier exemple est une collection de variables aléatoires non corrélées, w_t , de moyenne 0 et de variance finie σ_w^2
- Bruit blanc (faible): nom qui vient de l'ingénierie qui utilise ce modèle pour le bruit.
- Blanc: comme la lumière blanche (toutes les oscillations périodiques possibles sont présentes avec une force égale).
- Bruit blanc gaussien: Les variables aléatoires w_t suivent chacune une loi normale.
- Bruit blanc fort: c'est une suite de variables aléatoires indépendantes et identiquement distribuées de moyenne zéro et de variance finie.
 - Bruit blanc faible: $\{w_t\}$ est une suite de variables aléatoires non-corrélées de moyenne zéro et de variance finie.
 - Un bruit blanc Gaussien est un bruit blanc fort.

I.1.3.4. Les séries chronologiques

A- Définition

Une série temporelle, ou série chronologique, est une suite de valeurs numériques représentant l'évolution d'une quantité spécifique au cours du temps. De telles suites de valeurs peuvent être exprimées mathématiquement afin d'en analyser le comportement, généralement pour comprendre son évolution passée et pour en prévoir le comportement futur. Une telle transposition mathématique utilise le plus souvent des concepts de probabilités et de statistique. Parmi ses principaux objectifs figurent la détermination de tendances au sein de ces séries ainsi que la stabilité des valeurs (et de leur variation) au cours du temps.

Un objectif particulièrement important de l'analyse des séries chronologiques consiste à développer des modèles mathématiques qui procurent des descriptions des données.

On peut dire qu'une série chronologique se définit comme une collection de variables aléatoires indexées selon l'ordre qu'elles sont obtenues dans le temps. $\{X_t, t \in Z\}$,

On dit que c'est un processus stochastique.

De manière plus formelle, une série chronologique peut se définir comme une réalisation (finie) d'un processus stochastique.

B- Types de Séries

Une série chronologique ou encore chronique est un ensemble d'observations d'un processus aléatoire $(X_t)_{t \in T}$ se réalisant en un instant spécifié $t \in T$.

- Série continue: Une série chronologique est dite continue si l'ensemble des instants d'observations est continu (non dénombrable).
- Série discrète: Une série chronologique est dite discrète si l'ensemble des instants d'observations est discret (dénombrable).

I.1.3.5. Estimation stochastique linéaire

En pratique, il existe plusieurs approches pour estimer l'état inconnu d'un système linéaire à partir d'un groupe de mesures. Plusieurs de ces méthodes ne prennent pas en compte la nature « bruitée » des mesures, due à des origines diverses.

Sous l'hypothèse de linéarité et d'un bruit gaussien w , l'état x et la sortie y sont également des signaux (vectoriels) gaussiens, qui sont complètement caractérisés par leurs premiers et deuxièmes moments. Avant de procéder au calcul de ces caractéristiques stochastiques proprement dit, ce qui fera l'objet du FK étudié.

I.1.3.6. Modèle espace d'états ou modèle état mesure

C'est un modèle assez général qui englobe plusieurs modèles dynamiques linéaires et multi-variés. A titre d'illustration, un modèle de régression linéaire avec coefficients variables est présenté sous forme d'un modèle espace d'états:

$$y_t = x_t \beta_t + \varepsilon_t \quad \dots\dots\dots(\text{I.1.62})$$

$$\beta_t = T \beta_{t-1} + \mu_t \quad \dots\dots\dots(\text{I.1.63})$$

La première équation est connue comme l'équation de mesure et la deuxième comme équation d'état, où:

- y_t est un élément aléatoire observable qui correspond à la mesure ou output du système à la date t.
- x_t est un vecteur de variables observées déterministes de taille k. Il est connu comme le vecteur de mesure.
- β_t est un vecteur aléatoire de taille k; il représente l'état du système à la date t et il n'est pas observable (dans le cas général il peut être partiellement ou totalement inobservable).
- T est une matrice k*k définie comme la matrice de transition.
- ε_t est une erreur de mesure (ou bruit) à la date t. Elle est inobservable et strictement indépendante avec une moyenne nulle et une variance constante σ^2 .
- μ_t est un vecteur aléatoire de taille k représentant les innovations ou les inputs à la date t. Il est inobservable indépendant avec une moyenne nulle et une matrice de covariances Ω constante.
- $x_t \beta_t$ est appelé signal à la date t.

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

Dans leur version élémentaire, les modèles espace-état reposent sur un certain nombre d'hypothèses principales :

- Les équations de mesure et d'état sont linéaires ;
- Les bruits d'observation et d'innovation sont des bruits blancs
- Un bruit blanc (au sens faible) est un processus aléatoire... les variables cachées suivent à un instant initial donné une loi gaussienne. À ces dernières, sont ajoutées des hypothèses secondaires permettant de déterminer la forme canonique : l'indépendance entre les bruits d'observation et d'innovation. (condition inversible) et l'indépendance entre la variable cachée initiale et ces bruits (condition de causalité).

Toutes ces hypothèses sont destinées à simplifier les procédures d'estimation.

Pour autant, on ne peut pas associer à un processus donné Y_t une unique représentation espace-état. En effet, s'il existe une représentation de vecteur d'état Z_t , on peut formuler facilement une autre représentation Z_t^* et Z_t , Et étant une matrice inversible quelconque. De même, au lieu de modéliser $Z_t + 1$ dans l'équation d'état, on pourrait sans difficulté adapter l'estimation à un modèle d'état de Z_t . Enfin, diverses dimensions du vecteur d'état sont possibles et il convient de rechercher un modèle de dimension minimale (encadré), de manière à ne pas alourdir la procédure d'estimation. (**Gourrierroux C. et A. Montfort, 1990**)

Avec :

- Y_t est appelé observation ou variable de mesure ;
- Z_t est la variable d'état à la date t ;
- ε_t est le vecteur des innovations à la date t ;
- η_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) :

$$\begin{cases} (1): Z_{t+1} = A_t Z_t + B_t X_{1,t} + \varepsilon_t \\ (2): Y_t = C_t Z_t + D_t X_{2,t} + \eta_t \end{cases} \text{ où } \begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix} \square NID \left(0, \begin{pmatrix} Q_t & S_t \\ S_t & R_t \end{pmatrix} \right) \dots\dots\dots (\mathbf{I.1.63})$$

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 de loi $N(m, P)$ indépendant du bruit blanc normal.

Le système (I) peut s'écrire sous la forme développée suivante :

$$\begin{cases} Z_{t+1} = \prod_{j=0}^t A_{t-j} Z_0 + \sum_{j=1}^t \left(A_{t-j} \prod_{k=0}^{j-1} B_{t-k} \right) X_{1,t=j} + \sum_{j=1}^t \left(\prod_{k=0}^{j-1} A_{t-k} \right) \varepsilon_{t=j} + \varepsilon_t \\ Y_t = C_t \left[\prod_{j=0}^t A_{t-j} Z_0 + B_t X_{1,t} + \sum_{j=1}^t \left(A_{t-j} \prod_{k=0}^{j-1} B_{t-k} \right) X_{1,t=j} + \sum_{j=1}^t \left(\prod_{k=0}^{j-1} A_{t-k} \right) \varepsilon_{t=j} + \varepsilon_t \right] + D_t X_{2,t} + \eta_t \end{cases} \dots\dots\dots(\mathbf{I.1.64})$$

Les variables d'état et de mesure s'écrivent donc en fonction de la variable d'état initiale, du passé des erreurs de mesure et des innovations ainsi que des variables exogènes. Cette forme est particulièrement utile lorsqu'on s'intéresse à l'estimateur des moindres carrés généralisés du vecteur d'état ou à l'initialisation du filtre de Kalman. Le système (32) est dit sous forme canonique si et seulement si :

$$E(\varepsilon_t \eta_s) = E(\varepsilon_t Z_0) = E(\eta_t Z_0) = 0 \forall_{t,s} = 1, \dots, T \dots\dots\dots(\mathbf{I.1.65})$$

- Le modèle espace-état est alors dit causal et inversible.
- On appelle dimension minimale d'un système admettant une représentation espace-état, la plus petite dimension possible du vecteur d'état, K^* . En particulier, la représentation est dite minimale si A_t est de taille $(K^* \times K^*)$.

I.1.3.7. Problème du filtrage

D'une manière générale, le problème du filtrage consiste à déterminer des estimateurs de variables du système lorsque l'environnement présente des perturbations aléatoires. Le but est de déterminer un système (filtre) optimal au sens des moindres carrés (**Brown & Hwang, 1997**).

Si l'on dispose d'une série d'observations Z_t qui sont les mesures d'une variable X_t , entachées d'erreurs de mesure e_t d'origines diverses:

$$Z_t = X_t + e_t \dots\dots\dots(\mathbf{I.1.66})$$

Si la variable X_t est décrite par un modèle d'évolution, le problème du filtrage est celui d'obtenir des estimations optimales au sens de la minimisation de la variance d'erreur entre la variable réelle et son estimation notées \hat{X}_t à partir des observations Z_1, Z_2, \dots, Z_t . En d'autres termes, essayer de supprimer en quelque sorte la présence du bruit e_t d'où le nom : filtrage.

Si on considère un système dont on dispose de mesures jusqu'au temps t, On peut chercher à estimer la valeur de l'état X à un instant donné τ sachant les valeurs des mesures précédentes, noté $(\hat{X}_\tau / Z_1, Z_2, \dots, Z_t)$. Suivant la valeur de τ trois sortes de problèmes de l'estimation de l'état du système X_τ au temps τ sont intéressantes:

- Si $\tau < t$ il s'agit d'un problème de lissage ;
- Si $\tau = t$ il s'agit d'un problème de filtrage ;
- Si $\tau > t$ il s'agit d'un problème de prédiction.

Le problème du filtrage optimal, au sens de Kalman, est le même, il s'agit de trouver, pour le système dynamique en question, la meilleure estimation $\hat{X}_\tau / Z_1, Z_2, \dots, Z_t$ de l'état au sens du critère de la variance conditionnelle minimum.

Le filtre de Kalman permet d'appréhender directement le cas d'un système non stationnaire multi variable.

I.1.3.8. Caractérisation stochastique

Sous l'hypothèse d'un système linéaire et d'un bruit ω gaussien, l'état x et la sortie y sont également des processus (vectoriels) gaussiens qui sont donc entièrement caractérisés par leurs moments d'ordre un et deux. Le calcul de ces caractéristiques stochastiques sont données par le théorème suivant pour les deux cas continu et discret :

A- Dans le cas continu

Théorème

Soit le système linéaire continu :

$$\dot{x}(t) = Ax(t) + M\omega(t) \dots\dots\dots(I.1.67)$$

Avec :

$\omega(t)$ est un bruit blanc Gaussien stationnaire centré de densité spectrale de puissance W . On note $m(t_0)$ et $P(t_0)$ la moyenne et la covariance de l'état initial $x(t_0)$ (lui aussi aléatoire).

Alors $x(t)$ est un signal aléatoire gaussien de moyenne $m(t)$ et de covariance $P(t)$ tel que :

$$m(t) = E[x(t)] = \exp(A(t - t_0)) m(t_0)$$

$$P(t) = E[(x(t) - m(t))(x(t) - m(t))^T] \text{ Vérifiant l'équation différentielle}$$

$$\dot{P}(t) = A P(t) + P(t)A^T + MWM^T$$

Si le système est stable (toutes les valeurs propres de A sont à partie réelle négative) on tend vers un régime permanent (stationnaire) :

$\dot{P} = 0$ et $P(t) = P$ Vérifie alors l'équation de LYAPUNOV continue

$$A P + PA^T + MWM^T = 0$$

Si l'on considère l'équation de sortie : $y(t) = Hx(t)$ alors la matrice de covariance $P_y(t)$ de $y(t)$ vérifie : $P_y(t) = HP(t)H^T$

(Si l'on considère un bruit blanc de mesure alors cette covariance est infinie).

B- Dans le cas discret

Théorème

Soit un système linéaire discret :

$$x(k + 1) = Ax(k) + M\omega(k) \dots\dots\dots(I.1.68)$$

$\omega(k)$ est un bruit pseudo-blanc gaussien centré de densité spectrale W ($E[(\omega(k)\omega(k + j))^T] = W\delta(j)$). Soit $m(0)$ et $P(0)$ la moyenne et la covariance de l'état initial $x(k_0) = x(0)$ (lui aussi aléatoire gaussien et indépendant de $\omega(k)$). On montre que $X(k)$ est un signal aléatoire gaussien de moyenne $m(k)$ et de covariance $P(k)$, tel que :

$$m(k) = E[x(k)]$$

$$P(k) = E[(x(k) - m(k))(x(k) - m(k))^T] \text{ Vérifiant l'équation récurrente de LYAPUNOV}$$

$$P(k + 1) = AP(k)A^T + MWM^T$$

Si le système est stable (toutes les valeurs propres de A sont de module inférieur à 1) on tend vers un régime permanent: $P(k+1) = P(k) = P$ vérifie alors l'équation de LYAPUNOV discrète :

$$P = APA^T + MWM^T \quad \dots\dots\dots(\mathbf{I.1.69})$$

Conclusion

Dans ce premier chapitre, nous avons introduit les notions mathématiques élémentaires qui sont fortement requises 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'auto-corrélation, 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.

Introduction

La recherche scientifique repose sur deux sources d'information et d'étude différentes, les observations d'une part, et les modèles de l'autre. Au cours de l'histoire des sciences, les observations ont d'abord été qualitatives puis se sont raffinées pour devenir quantitatives, alors que parallèlement, les modèles descriptifs se sont renforcés de l'apport des mathématiques, en particulier de l'Analyse Numérique.

À ce jour, il existe un grand nombre de méthodes dites d'assimilation de données, qui consistent toutes plus ou moins à fournir une estimation de la condition initiale du système, connaissant des estimations de ce même système à d'autres instants : elles consistent toutes à résoudre un problème dit problème inverse. Il existe actuellement deux grandes classes de méthodes d'assimilation de données, l'une reposant sur des approches statistiques, l'assimilation séquentielle de données, qui repose sur des considérations statistiques et qui procède par corrections successives de la prévision du modèle au fur et à mesure que des observations sont disponibles, Les méthodes séquentielles reposent essentiellement soit sur l'interpolation optimale, soit sur le filtre de Kalman; et l'autre reposant sur la théorie du contrôle optimal, consiste à ajuster au mieux une solution du modèle à toutes les observations disponibles tout au long de la période d'assimilation c'est l'assimilation variationnelles.

Les méthodes séquentielles du type filtre de Kalman et les méthodes variationnelles sont deux approches distinctes pour résoudre le même problème, c'est-à-dire trouver la meilleure estimation linéaire de l'état du système à partir des observations. Ces deux classes de méthodes ne sont bien sûr plus équivalentes lorsque le problème est non linéaire. De plus, pour que le filtre de Kalman soit optimal, les erreurs doivent être décorréliées en temps, alors que les méthodes variationnelles n'ont pas besoin de cette hypothèse. [Didier AUROUX ,2003]

I.2.1. GENERALITES SUR LE FILTRE DE KALMAN

I.2.1.1. Historique du Filtre de Kalman

Le filtre de Kalman doit son nom à Rudolf Kalman bien que Thorvald Nicolai Thiele et Peter Swerling aient développé un algorithme similaire avant lui. La paternité du filtre fait l'objet d'une petite controverse dans la communauté scientifique.

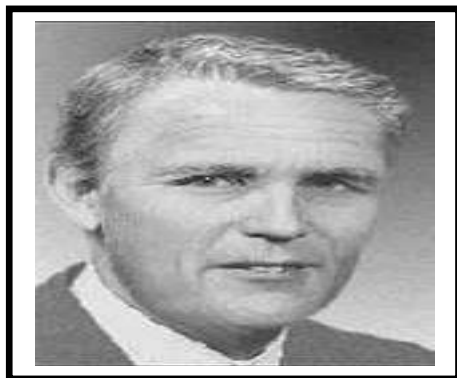


Fig I.2.1. Rudolf Emil Kalman

Le filtre a été décrit dans diverses publications par Swerling (1958), Kalman (1960) et Kalman-Bucy (1961).

Stanley Schmidt est reconnu comme ayant réalisé la première implémentation du filtre. C'était lors d'une visite de Rudolf Kalman au NASA Ames Research Center qu'il vit le potentiel de son filtre pour l'estimation de la trajectoire pour le programme Apollo. Ceci conduisit à l'utilisation du filtre dans l'ordinateur de navigation.

Au début des années **1940**, **Wiener** s'intéressa au problème de filtrage et donna naissance au premier filtre qui porte son nom Filtre de Wiener ; **En 1961**, **Kalman et Bucy** ont introduit un filtre qui enrichit le filtre de Wiener sur deux points essentiels :

- le filtre est récursif.
- le filtre peut être appliqué à des systèmes décrits par des modèles variant avec le temps (Processus non stationnaires).

Le filtre de Kalman a été développé 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.

I.2.1.2. Domaines d'utilisation du Filtre de Kalman

Une grande variété de filtres de Kalman ont été, depuis, développés à partir de la formulation originale dite filtre de Kalman simple. Schmidt développa le filtre de Kalman étendu, Bierman, Thornton et bien d'autres développèrent toute une gamme de filtres racine carrée. Le filtre le plus utilisé est vraisemblablement la phase-locked loop, largement répandue dans les radios, ordinateurs, équipement de communication, etc.

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

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

I.2.1.3. Définition du Filtre de Kalman

Le grand avantage du FK, pour nous en tant qu'ingénieurs, c'est qu'on n'est pas obligé d'être génie en mathématiques pour pouvoir l'assimiler et l'utiliser d'une manière efficace.

Avant de rentrer dans les détails propres du filtre de Kalman, les premières questions qui surgissent sont

- (1) d'abord qu'est-ce qu'un FK ?
- (2) Qu'est ce qui justifie le recours à un tel outil, dans le cadre de cette thèse ?

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 (**Maybeck, 1979**). C'est un reconstruteur d'état dans un environnement stochastique. Lorsque les variances des bruits

sont connues, c'est un estimateur linéaire minimisant la variance de l'erreur d'estimation. Les algorithmes donnant la solution de ce problème ont été déterminés initialement par [Kalman, 1960] dans le cas discret et [Kalman, Bucy, 1961] dans le cas continu. Nous établirons, dans un premier temps, les équations du filtre de Kalman discret puis, celles du filtre de Kalman continu par passage à la limite.

I.2.1.4.Principe du Filtre de Kalman

Cette méthode vise à estimer des paramètres d'un système évoluant dans le temps à partir de mesures bruitées. La force de ce filtre est sa capacité de prédiction des paramètres et de rectification des erreurs, non seulement des capteurs, mais aussi du modèle lui-même. En effet, pour appliquer un filtre de Kalman, il faut avant tout modéliser le système pour lequel on veut estimer les paramètres, de manière linéaire. (Des variantes du filtre de Kalman existent pour la prise en compte de modèles non linéaires) ; Dans une méthode d'estimation classique (par exemple, la méthode des moindres carrés), une simple erreur dans la modélisation du système entraîne inévitablement une erreur au niveau de l'estimation.

- La force du filtre de Kalman est d'intégrer un terme d'imprécision sur le modèle lui-même, ce qui lui permet de donner des estimations correctes malgré les erreurs de modélisation (pour peu que les erreurs restent raisonnables).
- Un autre point fort du filtre de Kalman (mais que l'on retrouve aussi dans la méthode des moindres carrés par exemple) est sa capacité à déterminer l'erreur moyenne de son estimation. En effet, l'outil Kalman fournit un vecteur contenant les paramètres estimés, mais aussi une matrice de covariance de l'erreur ! Cette matrice nous renseigne donc sur la précision de l'estimation, ce qui peut être utile dans de nombreuses applications. Un autre atout du filtre de Kalman est que la convergence de cette erreur est garantie !
- Le fonctionnement du filtre de Kalman peut se diviser en deux étapes :
- Une première étape de prédiction de l'estimation selon le modèle du système. Pour ce faire, le filtre de Kalman reprend l'estimation précédente des paramètres et de l'erreur et prédit les nouveaux paramètres et la nouvelle erreur en fonction de la modélisation du système.
- La seconde étape va faire la mise à jour de cette prédiction grâce aux nouvelles mesures. Ces mesures (par définition bruitées) vont permettre d'obtenir une estimation des paramètres et de l'erreur à partir de la prédiction faite. Si jamais le modèle comporte des erreurs, cette étape de mise à jour permettra de les rectifier.

I.2.1.5.Les limites du Filtre de Kalman

Le filtre de Kalman est donc une méthode d'estimation et de prédiction puissante prenant en compte les modélisations du système. Néanmoins, ce filtre n'est pas forcément l'outil à appliquer dans tous les cas. En effet, comme nous l'avons vu, le développeur a besoin de modéliser le système assez précisément afin de désigner un filtre efficace. Le problème est que certains systèmes sont difficilement modélisables et, encore moins linéairement. 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, elle restera grande.

Pour pallier à ce problème de modélisation linéaire du système, 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 (la précision des estimations) ne converge pas obligatoirement (comme c'était le cas avec une modélisation linéaire).
- Le second défaut est son coût calculatoire plus important. En effet, de nouvelles matrices coûteuses rentrent en jeu (les matrices des dérivés partielles des équations d'états et de mesures

modélisant le système), ce qui peut être une limite à son utilisation, surtout dans des systèmes embarqués très restreints au niveau de la puissance de calcul.

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. Le bruit peut en général être modélisé de façon Gaussienne, mais dans certains cas, un autre type de bruit est requis (notamment en traitement d'images où l'on utilise fréquemment des bruits de Poisson). 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.

I.2.2. DE L'ESTIMATEUR OPTIMAL AU FILTRE DE KALMAN

On commence par faire quelques rappels sur les estimateurs déterministes, présenter l'estimateur des moindres carrés. Ensuite, on complexifiera le modèle progressivement jusqu'à obtenir un filtre de Kalman.

I.2.2.1. L'estimation

L'estimation est une branche des mathématiques statistiques qui permet, à partir de mesures effectuées sur un système, d'estimer la valeur de différents paramètres de ce système. Bien sûr, il est possible de n'estimer qu'un seul paramètre, mais il nous est aussi possible d'estimer plusieurs paramètres en même temps. Dans ce cas, on ne parle plus de paramètre à estimer, mais de vecteur à estimer, ou vecteur d'état.

I.2.2.2. Les Estimateurs déterministes

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

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

A- Le biais

Un estimateur sans biais signifie que l'espérance de l'estimation doit être égale au vecteur à estimer. En gros, ça veut donc dire que l'estimateur, en moyenne, ne commet pas d'erreur. Dit encore autrement, si l'on fait la moyenne des estimations, on trouve la valeur réelle du paramètre. Bien sûr, le biais n'est pas calculable en réalité, car on ignore la valeur réelle du paramètre ou du vecteur à estimer. Par contre, il est très utile pour caractériser nos estimateurs.

B- La variance

La variance peut être assimilée à l'erreur. Un estimateur est dit à variance minimal lorsque l'on a réussi à minimiser les variations entre le vecteur d'état et l'estimation. On a donc minimisé l'erreur.

Il faut bien faire attention, car ces deux notions sont bien distinctes. Un estimateur peut très bien avoir un biais nul, mais si sa variance est forte, cet estimateur ne sera pas performant.

I.2.2.3. L'estimateur optimal

Dans la littérature, nous trouverons qu'un estimateur dont le biais est nul aura une variance toujours supérieure ou égale à ce que l'on appelle la "borne de Cramer-Rao". Finalement, c'est assez logique. On ne peut pas obtenir une variance nulle et donc, on ne peut pas connaître exactement la valeur du vecteur d'état. On aura toujours une incertitude égale à la borne de Cramer-Rao pour l'estimateur optimal. Si la variance est nulle, alors ce n'est plus un problème d'estimation, car ça voudrait dire que l'on connaît la valeur à estimer dès le départ.

Il se trouve que dans certains cas particuliers, il existe un estimateur sans biais à variance minimale calculable. Mais pour cela, il faut que le système respecte quelques contraintes :

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

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

I.2.2.4. Optimalité et Récursivité de l'estimation

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. Cela dit, le FK peut alors être considéré comme tout juste un algorithme de traitement de données dans un processeur central. Comme tel, il incorpore des échantillons de mesures discrètes. Conceptuellement, ce que n'importe quel filtre essaye d'obtenir c'est une estimation « Optimale » des variables en question à partir des données procurées par un environnement bruité. Il s'agit d'extraire une information utile à partir d'une information bruitée. Le terme « optimale » ici veut dire minimum d'erreur. Pour aboutir à cet objectif, il existe plusieurs approches entre autres, l'approche de Bayes (**Harrison & Stevens, 1975b**). Si on adopte le point de vue de Bayes, alors on souhaiterait un filtre qui propage la densité de probabilité conditionnelle des variables en question, conditionnée par la connaissance des données actuelles provenant des systèmes de mesures.

La probabilité de la valeur d'une quantité scalaire x au temps i , $x(i)$, conditionnée par la connaissance des mesures z jusqu'au temps i : ($z(1)=z_1$, $z(2)=z_2$, ..., $z(i) = z_i$) est tracée en fonction des $x(i)$ possibles, pour donner le tracé de la fonction de densité de probabilité conditionnelle notée : $f_{x(i)/z(1),z(2),z(i)}(X/Z_1, Z_2, \dots, Z_i)$. Cette fonction contient toute l'information disponible sur $x(i)$, elle indique pour la valeur donnée de toutes les mesures jusqu'à l'instant i , ce que serait la probabilité de $x(i)$. Elle est appelée « conditionnelle » car sa forme et sa localisation sur l'axe des x dépendent des valeurs des mesures prises. Sa forme révèle sur le taux de certitude que l'on a sur la connaissance de la valeur de x : Si le tracé de la fonction de densité de probabilité conditionnelle est proche d'un « pic » cela signifie que la plupart de la probabilité est concentrée autour d'une plage de valeurs de x , on est alors plus sûr de la valeur de x . Si le tracé est plus disparate et plus plat, cela signifie que l'on est moins sûr de la valeur de x . Une fois une telle densité de probabilité conditionnelle propagée et tracée, l'estimation « optimale » peut alors être définie. Les choix possibles peuvent inclure :

1. La moyenne : la valeur de x correspondant au centre de masse de la probabilité conditionnelle.
2. Le mode : la valeur de x correspondant au maximum de probabilité conditionnelle.

3. La médiane : la valeur de x telle que 50% de probabilité soit à sa gauche et 50% de probabilité soit à sa droite.

Le FK performe cette propagation de la fonction de densité de probabilité conditionnelle, pour les problèmes dans lesquels le système peut être décrit par un modèle linéaire et dans lesquels les bruits du système et de la mesure sont blancs et gaussiens (**Brown & Hwang, 1997**). Dans ces conditions, la moyenne, le mode et la médiane coïncident, de telle sorte qu'il existe en fait, une estimation « meilleure » unique de la valeur de x . En d'autres termes, dans ces trois conditions, le FK est le meilleur filtre parmi toutes les formes concevables, dans le sens où l'estimation qu'il fournit est une estimation à variance minimum (**Welch & Bishop, 2003**). Ceci est la raison principale qui justifie notre recours, dans le cadre de cette thèse, à un tel outil mathématique parmi tant d'autres.

I.2.3. MECANISME DE LA METHODE DE FILTRE DE KALMAN

I.2.3.1. Variables et équations d'état

L'état d'un système est un ensemble de variables, tel que:

- la connaissance exacte de ses variables,
- les fonctions d'entrée du système, et les équations décrivant les dynamiques de ce même système, fournissent l'état futur et la sortie du système.

L'état d'un système dynamique est décrit par un ensemble de variables d'état:
 $[x_1(t), x_2(t), \dots, x_n(t)]$

Les variables d'état (VE) sont celles qui déterminent le comportement futur d'un système quand l'état présent et les signaux d'excitation d'entrée sont connus.

État présent (à $t = t_0$) + Entrées \rightarrow (État futur + Sorties) à l'instant t

I.2.3.2. Forme matricielle des équations d'état

De ces vecteurs d'état (VE), nous obtiendrons un ensemble d'équations différentielles ordinaires de premier ordre, que nous grouperons en utilisant une notation matricielle compacte pour former un modèle temporel de variables d'état.

L'état d'un système est décrit par un ensemble d'équations différentielles de 1er ordre, somme suit:

$$\begin{cases} \bar{x}_1 = a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n + b_{11}u_1 + \dots + b_{1m}u_m \\ \bar{x}_2 = a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n + b_{21}u_1 + \dots + b_{2m}u_m \\ \vdots \\ \bar{x}_n = a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n + b_{n1}u_1 + \dots + b_{nm}u_m \end{cases} \Rightarrow \bar{X}t = \frac{dx}{dt}$$

$$\frac{d}{dt} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1m} \\ b_{21} & b_{22} & \dots & b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nm} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{bmatrix}$$

Avec :

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \begin{cases} \bar{x} = AX + BU \rightarrow \text{Équation d'état} \\ Y = HX + DU \rightarrow \text{Équation de mesure (de sortie)} \end{cases} \dots\dots\dots \text{(I.2.1)}$$

C'est la représentation espace - d'état

Avec :

$X \in R^n$: Vecteur qui représente les n variables d'état

$Y \in R^q$: vecteur qui représente les m mesures

$U \in R^p$: vecteur qui représente les p commandes

$A \in R^{n \times n}$: Matrice de dynamique

$B \in R^{n \times p}$: Matrice de commande

$H \in R^{q \times n}$: Matrice d'observation

$D \in R^{q \times p}$: Matrice d'action directe

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \quad B = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1m} \\ b_{21} & b_{22} & \dots & b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nm} \end{bmatrix} \quad U = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{bmatrix} \quad Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

- Les matrices H et D sont à déduire à partir des dimensions de : Y, X, et U.

- Les matrices A, B, H, D sont souvent invariantes dans le temps, elles deviennent alors des matrices constantes et on parle de représentation d'état continue indépendante du temps.
- Cette représentation d'état se représente sous forme de schéma-bloc

I.2.3.3. Les entrées et sorties du filtre de Kalman

Le but du filtre de Kalman est d'obtenir une estimation du vecteur d'état $X_p = X(t_p)$, compte tenu des informations disponibles à un instant t_q (suivant les cas, on peut avoir $t_q < t_k$, $t_q = t_k$, $t_q > t_k$). Cette estimation est notée de la manière suivante :

$$X(t_p) = X_{p/q} \rightarrow \text{sachant } t_p \text{ les informations disponibles jusqu'à } t_q.$$

Le filtre fournit de plus les matrices de covariance des erreurs d'estimation

$$E(X_p - X_{(p/q)})(X_p - X_{(p/q)})^T = P_{p/q}$$

Les grandeurs sont calculées pas à pas en fonction des données suivantes :

- Matrice de transition $\varphi_{k/k-1}$
- Matrice de mesure H_k
- Estimation initiale du vecteur d'état, notée $\hat{X}_{(0/0)}$
- Covariance de l'erreur d'estimation initiale $X_0 - \hat{X}_{(0/0)}$, notée $P_{0/0}$
- Propriétés statistiques des bruits :

● **Valeurs moyennes nulles**

$$E V_k = 0$$

$$E W_k = 0$$

● **Matrices de covariance**

$$E W_j W_k^T = 0 \text{ si } (j \neq k)$$

$$E W_k W_k^T = Q_k$$

$$E V_j V_k^T = 0 \text{ si } (j \neq k)$$

$$E V_k V_k^T = R_k$$

● **Indépendance des bruits du système et de mesure**

$$\Rightarrow E W_k V_j^T = 0$$

● **Indépendance du bruit de mesure et de l'erreur initiale d'estimation**

$$\Rightarrow E (X_0 - \hat{X}_{0/0}) = 0$$

I.2.4. PRESENTATION DE L'ALGORITHME RECURSIF DU FK DISCRET

L'algorithme est structuré en deux étapes reprises d'itération en itération :

- Les deux premières équations (1 et 2) sont des équations de « mises à jour des mesures » (actualisation)
 - ✓ Les deux suivantes (3 et 4) de « mise à jour du temps » (prédiction).
 - ✓ La première étape concerne les lois de probabilité a posteriori qui tiennent compte de l'information à la date k.
- La seconde étape, à la différence de la première, ne dépend pas des observations à la date k, c'est-à-dire sans utiliser les signaux Y_k .

- La dernière équation (5) actualise la matrice de gain K_k qui intervient dans les équations
- Précédentes.

Chaque itération se résume par les cinq équations suivantes :

1. Mise à jour du vecteur d'état

$$\hat{X}_{k/k} = \hat{X}_{k/k-1} + K_k(Z_k - H_k\hat{X}_{k/k-1}) \dots\dots\dots(\text{I.2.2})$$

2. Correction de la matrice de covariance de l'erreur d'estimation

$$P_{k/k} = (I - K_k H_k)P_{k/k-1} \dots\dots\dots(\text{I.2.3})$$

3. Prédiction du vecteur d'état

$$\hat{X}_{k+1/k} = \phi_{k+1/k}\hat{X}_{k/k} \dots\dots\dots(\text{I.2.4})$$

4. Prédiction de la covariance de l'erreur de prédiction

$$P_{k+1/k} = \phi_{k+1/k}P_{k/k}\phi_{k+1/k}^T + Q_k \dots\dots\dots(\text{I.2.5})$$

5. Estimation du gain

$$K_k = P_{k/k-1}H_k^T(H_kP_{k/k-1}H_k^T + R_k)^{-1} \dots\dots\dots(\text{I.2.6})$$

Où :

- $\hat{X}_{k+1/k}$ la prédiction à l'état k+1
- $P_{k+1/k}$ la covariance de l'erreur à l'instant k+1
- $H_k = H_k + 1$: la matrice d'état non pas pour les k+1 mesures (matrice H), mais uniquement pour la mesure courante. (Dans le cas des moindres carrés ou de man, $H_k = H_k + 1$)
- $R_k = R_k + 1$ la matrice de covariance du bruit de mesure, non plus pour k+1 mesures
- (Matrice Γ), mais pour une seule mesure uniquement. (Pour le filtre de Kalman, on a si $R_k = R_k + 1$)
- I : la matrice identité
- Z_k : la mesure à l'instant k

L'équation (1) calcule l'estimation courante du vecteur d'état $\hat{X}_{k/k}$ comme la somme pondérée de la prévision à la date k-1 du vecteur d'état X_k et de l'erreur de prédiction calculée à partir de la dernière valeur observée Y_k . La pondération K_k , appelée matrice de gain, est actualisée à chaque itération par l'équation (5).

L'équation (3) permet de calculer la prédiction de $\hat{X}_{k/k}$ à la date k+1, $\hat{X}_{k+1/k}$ comme la projection de $\hat{X}_{k+1/k}$ sur son passé (passé synthétisé par $\hat{X}_{k/k}$).

Les équations (2) et (4) sur les matrices de covariance. Ces équations permettent de calculer la suite des gains de Kalman K_k et ce calcul peut être fait « hors-ligne ». La matrice de covariance a posteriori $P_{k/k}$ connaît généralement un gain en précision par rapport à la matrice de covariance a priori $P_{k+1/k}$, grâce au terme $(I - K_k H_k)$

La matrice de covariance a priori en k+1, $P_{k/k+1}$ prend en compte les erreurs liées aux innovations de l'état avec la matrice Q_t , mais est aussi augmentée d'un terme $\phi_{k+1/k}P_{k/k}\phi_{k+1/k}^T$ associé aux erreurs sur l'état à la date t (équation 4).

Lorsque les variables d'état sont stationnaires, la covariance prévue $P_{k+1/k}$ qui part d'une incertitude a priori P, tend vers une constante Σ_∞ [Harvey, 1989]. Après une période transi-

toire, les intervalles de confiance entourant des variables cachées stationnaires ont donc une largeur à peu près constante.

Pour $k+1$ mesures, notre estimateur optimal s'écrit donc comme cela :

$$\begin{aligned}
 K_k &= P_{k/k-1} H_k^T (H_k P_{k/k-1} H_k^T + R_k)^{-1} \\
 \hat{X}_{k/k} &= \hat{X}_{k/k-1} + K_k (Z_k - H_k \hat{X}_{k/k-1}) \\
 P_{k/k} &= (I - K_k H_k) P_{k/k-1} \\
 \hat{X}_{k+1/k} &= \Phi_{k+1/k} \hat{X}_{k/k} \\
 P_{k+1/k} &= \Phi_{k+1/k} P_{k/k} \Phi_{k+1/k}^T + Q_k
 \end{aligned}$$

La phase de Mise à jour (correction)

La phase de prédiction

I.2.4.1. Les différentes étapes d'évaluation des paramètres du système de FK

La figure suivante montre les différentes étapes d'évaluation des paramètres du système. Il y a deux étapes principales pour l'évaluation de ces paramètres; l'évaluation a priori (la prédiction) et l'évaluation a posteriori (la correction). Les paramètres du système sont estimés premièrement à partir de valeurs de l'étape précédente et en suite ils sont corrigés par des mesures dans l'étape de correction.

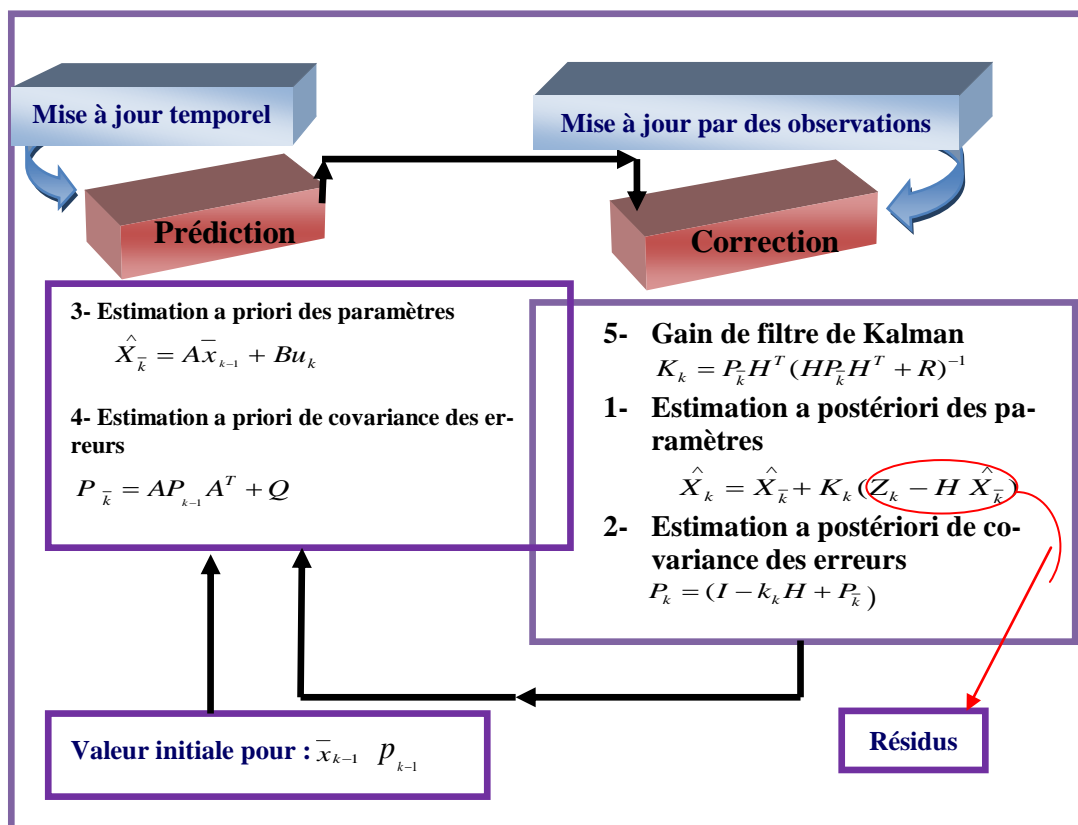


Fig I.2.2. Les différentes étapes d'évaluation des paramètres du système de FK.

Cette théorie fournit à chaque nouvelle observation une nouvelle estimation de la variable d'état. Le filtrage de Kalman utilise les observations dans l'ordre où elles sont disponibles,

d'où l'appellation séquentielle. Par conséquent, l'estimation de l'état du système ne dépendra pas des observations futures mais uniquement des observations passées et présentes.

Considérons une discrétisation en temps de notre problème :

Le problème de filtrage au sens de Kalman est résolu en utilisant les principes de base de l'estimation, simple puis récursive, d'une variable aléatoire, on établit que les équations de fonctionnement d'un filtre de Kalman discret se décomposent en deux étapes :

- Une étape de prédiction
- Une étape de correction

La construction de cet algorithme montre qu'il permet de répondre à deux objectifs différents:

- C'est un filtre linéaire minimisant la variance a priori de l'erreur d'estimation. Dans ces conditions, les bruits peuvent ne pas être gaussiens.
- C'est un filtre maximisant la probabilité a posteriori des grandeurs à estimer. Cela n'est alors applicable que dans l'hypothèse de bruits gaussiens.

Remarque :

Le grand avantage de la méthode du filtre de Kalman est de fournir à chaque itération une estimation des matrices de covariance d'erreur de prédiction. Il faut toutefois savoir initialiser correctement ces matrices à l'instant (t_0) .

I.2.4.2. Critères d'optimisation

Il est temps de préciser les conditions théoriques permettant une démonstration de l'expression de la correction K_k . Il faut dire qu'en général, il existe deux cas de figure, selon les hypothèses dans lesquelles on se met (**Harvey, 1989**):

A- Premier cas

Hypothèses

1. Les bruits obéissent aux relations précédentes ; cependant, leur loi de probabilité n'est pas explicitée.
2. L'opérateur estimation est linéaire par rapport aux mesures et à l'estimation initiale.
3. Les valeurs estimées \hat{X}_k doivent conduire à une minimalisation de la somme de variances $(X - \hat{X})^T (X - \hat{X})$

Conséquences

1. K_k est fixée.
2. Chaque estimation \hat{X}_k est égale à l'espérance mathématique de X , compte tenu des résultats de mesure obtenus :
3. $\hat{X}_{k/k} = E(X_k / Z_1 Z_2 \dots Z_k)$
4. (Loi de probabilité a posteriori, c'est-à-dire compte tenu des résultats de mesure).
5. La minimalisation de $E(X - \hat{X})^T (X - \hat{X})$ implique celle de :
6. $E(X - \hat{X})^T W_k (X - \hat{X})$: W_k étant une matrice réelle, symétrique, définie, positive arbitraire.

B- Deuxième cas**Hypothèses:**

1. Les lois de probabilité des vecteurs d'état $(X_0 X_1 \dots X_k)$ et des mesures $(Z_1 Z_2 \dots Z_k)$ sont gaussiennes.
2. La valeur estimée \hat{X}_k maximalise la probabilité a posteriori $(X_k/Z_1 Z_2 \dots Z_k)$

Conséquences:

1. La formule d'estimation peut être mise sous la forme linéaire
2. La matrice de correction à utiliser est donnée par l'équation :
3. Chaque estimation \hat{X}_k est égale à l'espérance mathématique de X_k , compte tenu des mesures faites (loi de probabilité a posteriori) :
4. $\hat{X}_{k/k} = E(X_k/Z_1 Z_2 \dots Z_k)$
5. Il en résulte une minimalisation de la quantité $E(X - \hat{X})^T W_k (X - \hat{X})$
 W_k étant une matrice réelle, symétrique, définie, positive, arbitraire.

Dans les deux cas précédents, le filtre de Kalman conduit à une minimalisation de la variance $E(X - \hat{X})^T (X - \hat{X})$; et d'une manière générale, l'estimation optimale de AX (A étant une matrice carrée ou rectangulaire) est égale à \widehat{AX} étant l'estimation optimale de X .

I.2.5. STABILITE DU FILTRE VIS-A-VIS DE LA DIVERGENCE

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

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

Les problèmes de stabilité du filtre sont reliés à la connaissance exacte des matrices P_0 , Φ_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.

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

devient leur influence. Ces modifications sont apportées par des ensembles de pondération $(1, e^{-\frac{1}{\tau}}, e^{-\frac{2}{\tau}}, e^{-\frac{3}{\tau}}, \dots, e^{-\frac{s}{\tau}})$ pouvant être appliqués aux observations $(Z_{k-1}, \dots, Z_{k-s})$ où τ est la mémoire du modèle.

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

À 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.

Il est intéressant de préciser que ce filtre nécessite une étape d'initialisation afin de déterminer la première estimation du vecteur d'état ainsi que la matrice de covariance de l'erreur. Ces matrices, si l'on ne connaît pas les conditions initiales du système, peuvent être obtenues soit par une méthode ad-hoc (on fixe le vecteur d'estimation plus ou moins au pif et l'on attribue une erreur importante à cette estimation), soit on utilise un autre estimateur, par exemple, la méthode des moindres carrés classique afin d'initialiser le filtre.

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

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

Introduction

Les régimes hydrologiques varient suivant la répartition des précipitations et suivant les conditions naturelles des bassins versants, En matière de modélisation hydrologique on s'intéresse aux distributions et à l'ampleur des précipitations, au moment et à l'endroit où elles se produisent. L'étude hydrologique des précipitations porte donc essentiellement sur leur quantité et leur dynamique, dans le temps et dans l'espace. A cet effet, la qualité de l'estimation spatiale et temporelle de la pluviométrie suppose la compréhension des modalités de reconstitution de la ressource en eau qui elle-même se base sur une bonne connaissance de la structuration des champs de précipitation dans l'espace étudié et dans le temps.

II.1.1. PRESENTATION DE LA REGION D'ETUDE

II.1.1.1. Situation géographique du bassin versant

Le Bassin Hydrographique Cheliff- Zahrez couvre une superficie d'environ 56 227 km² (soit plus de 22% de la superficie de l'Algérie du Nord). Une région hydrographique limitée naturellement au nord par la mer Méditerranée, à l'Ouest par la région Oranie - Choit Chergui, à l'Est par la région Algéroise - Hoddna- Sommam et au sud par le Sahara.

Cette région hydrographique est découpée en trois grands sous bassins versants (côtier Dahra, Cheliff et le Zahrez), le bassin du Cheliff de son tour se divise en trois grands bassins : le Cheliff à l'amont du barrage de Boughezoul et le Cheliff à l'aval du barrage de Boughezoul et le haut et moyen Cheliff; Il est le plus grand du point de vue superficie en Algérie, il représente plus de 77 % de la superficie totale du bassin Cheliff-Zahrez. Ce même bassin, limité au Nord par les monts du Dahra (1578m d'altitude) et au sud par l'Atlas saharien, comprend au Nord, la vallée du Cheliff, au sud les hauts plateaux de Saida à Tiaret, les plaines de Nahr Ouassel et Ain Oussera et au centre le massif de l'Ouarsenis. (ABH, Chlef 2004, (1)),

Elle est comprise entre :

La longitude 3° 50' Est et 0° 08' Ouest et 33° 53' de latitude Sud et de latitude 36° 34' Nord.

II.1.1.2. Caractéristiques morphométriques de la région

Les différents paramètres morphométriques du bassin (forme, altitude, pente, relief, etc..) interviennent, et souvent de façon combinée, dans les modèles de transformation pluie-débit particulièrement. Devant l'importance de la taille du bassin versant du Cheliff-Zahrez (56 227 km² soit plus de 22% de la superficie de l'Algérie du Nord), ainsi que celle des (03) bassins qu'il englobe; Côtier (3 202 km²), Cheliff (44 035 km²) et le Zahrez (8 990 km²) ainsi que sa complexité géologique, notamment dans sa partie Nord, nous nous limiterons à quelques caractéristiques morphométriques susceptibles d'identifier les bassins de la région d'étude. (ABH, Chlef (2))

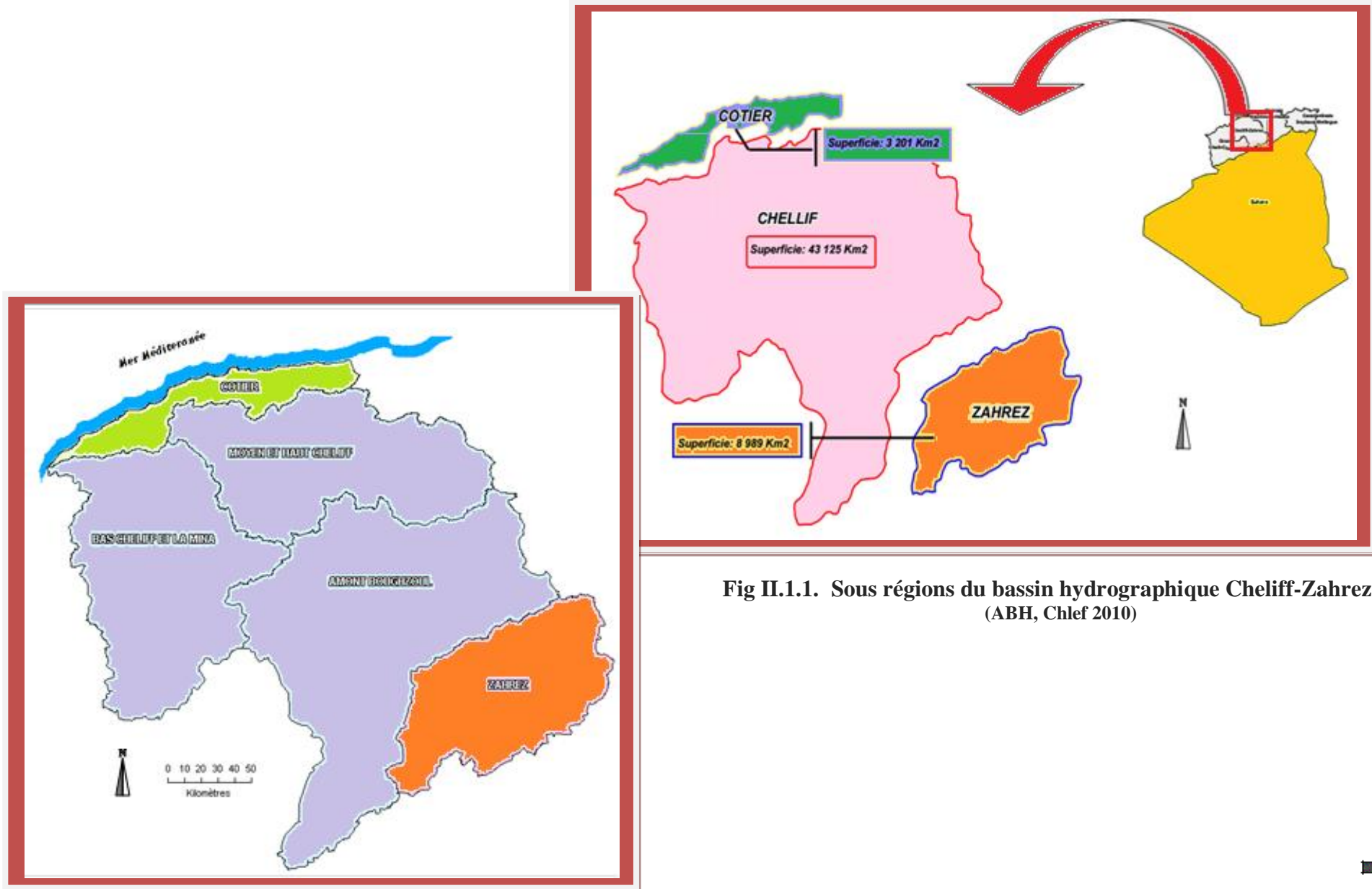


Fig II.1.1. Sous régions du bassin hydrographique Cheliff-Zahrez (ABH, Chlef 2010)

1. Superficie, Altitude et formes des unités hydrologiques du bassin versant

Selon les délimitations hydrographiques, l'aire de l'étude se compose de 03 bassins : côtier Dahra, le Cheliff de son tour se divise en trois grands bassins : le Cheliff à l'aval du barrage Boughzoul, le Cheliff à l'amont du barrage de Boughzoul et le haut et moyen Cheliff, ainsi que le bassin du Zahrez (Fig.II.1.2). Leurs superficies respectives, mentionnées dans le Tableau II.1.1, varient de 8 989 km² (bassin du Zahrez), à 19 979 km² (bassin Cheliff Amont Boughzoul) et à 24 028 km² au bassin Cheliff à l'aval du Boughzoul qui à son tour contrôle 43% de la surface de réception totale de la région hydrographique. Les bassins Amont Boughzoul et Zahrez représentent 51 % de la surface de réception totale du bassin Cheliff-Zahrez, par contre le bassin côtier ne représente que 6%. (ABH, Chlef, 2010)

Ces grands bassins englobent plusieurs sous bassins :

- Côtier : 03 sous bassins;
- l'aval Boughzoul : 24 sous bassins
- l'amont Boughzoul : 12 sous bassins
- Le bassin Zahrez : 06 sous bassins

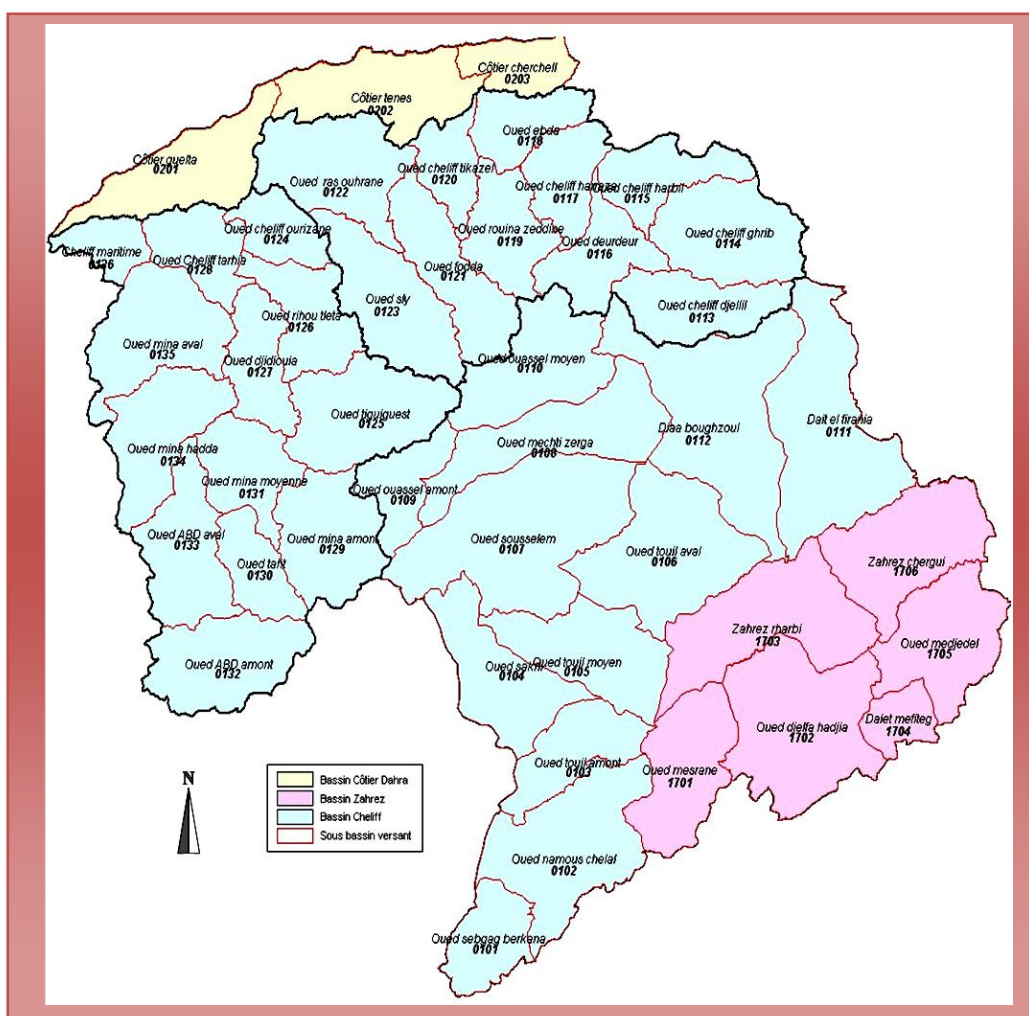


Fig II.1.2. Carte des sous bassins versant (Cheliff à l'amont de Boughzoul /SIG,2010)

2. Orographie

Le bassin versant du Cheliff occupe une portion de l'Algérie dont l'orographie est à la fois simple et typique. Des chaînes montagneuses, parallèles à la côte, séparées par les zones déprimées ou par de hautes plaines, en constituent les principaux traits. Cette disposition correspond à la structure géologique, mais sa réalisation fut assez récente puisqu'elle ne s'est traduite définitivement dans le relief qu'au cours de la période Miocène. La forme du relief, les effets des crêtes ainsi que l'influence du site constituent des facteurs majeurs dans la structuration des champs pluviométriques et leurs orientations dans l'espace.

En fonction de la géologie, de la lithologie et de la topographie la région d'étude s'organise en deux grandes unités structurales : le Système Tellien et les Hautes plaines steppiques.

2. 1. Le Système Tellien

L'atlas Tellien littoral regroupe de l'Est à l'Ouest, le massif de Gontas, les monts de Zaccar le massif de Béni Menasseurs et le Dahra, qui sépare la région d'étude de la mer et qui forme une barrière assez continue de quelque 70 km de longueur s'étageant entre 500 et 900 m d'altitude moyenne. La dépression intertelliene occupe les plaines alluviales du Cheliff, partagée en principales plaines

- La plaine d'El Khemis, reconnue aussi sous le nom du haut Cheliff elle est comprise entre le seuil de Djendel à l'Est et le seuil d'Ain Defla à l'Ouest
- La plaine d'El Abadia et d'Oued Fodda reconnue aussi sous le nom du Moyen Cheliff, elle est comprise entre le seuil de Ain Defla à l'Est et le seuil de Pontéba à l'Ouest
- La basse plaine du Bas Chélif, forme une étroite bande d'une dizaine de kilomètres de largeur qui draine le Cheliff, depuis la ville de Chlef jusqu'au confluent avec l'oued Mina.
- L'atlas Sud tellien borde continuellement au Sud la dépression intertelliene reconnue aussi sous le nom de massif d'Ouersenis, forme l'ossature de la bordure sud de la zone d'étude. C'est une chaîne importante s'allongeant sur plus de 200 km, culminant à 1985 m au djebel Ouerse-nis, son altitude moyenne s'accroît régulièrement d'Ouest en Est à 1300 m environ, à l'Est de ce massif central s'élèvent les monts de Matmata de part et d'autre part les vallées de l'Oued Deurder et du Cheliff. (ABH, Chlef 2004)

2. 2. Les Hautes Plainnes steppiques

Localisées entre l'Atlas Tellien au Nord et l'Atlas Saharien au Sud, à des altitudes plus ou moins importantes de 900 à 1 200 m. Elles sont parsemées de dépressions salées, chotts ou sebknas (Zahrez chergui et Zahrez gharbi) qui sont des lacs continentaux formés au Pléistocène sous l'effet des pluies torrentielles et du ruissellement important qui en découle.

3. Les paramètres morphométriques du bassin versant :

Les paramètres morphométriques jouent un rôle essentiel dans la répartition des eaux à l'échelle du bassin versant. Ils influent fortement sur l'hydrologie de ce dernier.

Tableau II.1. 1. Caractéristiques morphométriques des sous-bassins versants de la région d'étude.

Bassin	Code sous bassin	Superficie (Km2)	Périmètre (km)	Altitude moyenne (m)	Altitude minimale (m)	Altitude maximale (m)	indice de compacité (kc)	Indice de pente globale(Ig) (m/km)	Dénivelée spécifique (Ds) (m)	Relief
	113	992,85	179,96	900	600	1800	1,6	16,91	532,84	Relief très fort
	114	1383,69	176,81	870	500	1600	1,33	17,67	657,29	Relief très fort
	115	767,29	167,95	600	300	1300	1,7	15,01	415,78	Relief assez fort
	116	857,64	185,65	980	300	1800	1,46	19,41	568,42	Relief très fort
	117	743,97	142,68	575	300	1600	1,76	23,87	651,09	Relief très fort
	118	680,55	141,1	610	200	1500	1,51	23,73	619,08	Relief très fort
	119	898,54	173,69	640	200	1700	1,62	21,38	640,87	Relief très fort
	120	570,52	142,58	370	200	1100	1,67	14,49	346,1	Relief assez fort
	121	1161,32	212,75	785	200	1800	1,75	18,08	616,13	Relief très fort
	122	1437,44	224,12	375	100	1100	1,66	11,31	428,8	Relief assez fort
	123	1405,47	225,12	710	100	1700	1,68	17,3	648,58	Relief très fort
Cheliff à l'aval Bouguezoul	124	564,72	122,13	275	75	1000	1,44	20,12	478,12	Relief assez fort
	125	1513,02	202,89	805	300	1300	1,46	13,21	513,83	Relief très fort
	126	928,45	185,18	460	75	1300	1,7	16,02	488,15	Relief fort
	127	799,4	162	550	200	1000	1,6	12,95	366,14	Relief fort
	128	759,63	145	310	50	800	1,47	10,04	276,71	Relief fort
	129	1277,66	176	1095	600	1300	1,38	11,51	411,42	Relief fort
	130	657,57	128	910	600	1300	1,4	15,65	401,32	Relief fort
	131	892,21	162	650	300	1300	1,52	16,21	484,19	Relief fort
	132	1352,8	188	1100	800	1400	1,43	9,22	339,12	Relief fort
	133	1228,61	198	775	300	1300	1,58	13,03	456,72	Relief fort
	134	1250,98	193	490	100	1200	1,53	14,8	523,47	Relief très fort
	135	1462,52	212	345	50	900	1,55	10,03	383,57	Relief fort
	136	431,41	112	215	0	700	1,51	12,85	266,9	Relief fort
Sous total		24028,3		640	0	1800				
	101	791	139,5	1410	1100	1700	2,01	6,45	125,62	Relief assez fort
	102	2019	268,5	1205	900	1700	1,67	5	224,67	Relief assez fort
	103	963,8	177	1050	800	1300	1,6	4,76	147,83	Relief assez fort
	104	1329	156,5	1130	800	1500	1,5	7,69	225,4	Relief assez fort
	105	1403	203,3	970	700	1400	1,52	5,8	217,14	Relief assez fort
Cheliff à l'amont Bouguezoul	106	2033	233,9	870	600	1300	1,45	6,58	296,64	Relief fort
	107	2833	292,7	1000	600	1500	1,54	6,93	368,89	Relief fort
	108	1549	251,3	890	600	1200	1,79	4	157,4	Relief assez fort
	109	699,3	152	1060	800	1300	1,61	5,45	144,24	Relief assez fort
	110	1403	222,1	950	700	1700	1,66	6,1	228,39	Relief assez fort
	111	2339	284,6	840	600	1500	1,65	4,76	230,3	Relief assez fort
	112	2617	338,8	790	600	1600	1,85	4,29	219,24	Relief assez fort
Sous total		19096,8		940	600	1700				
	1701	1330	176,9	1110	800	1600	1,36	11,11	405,21	Relief fort
	1702	2223	252,7	1200	800	1600	1,5	5,88	277,35	Relief fort
Zahrez	1703	1936	257,7	950	800	1500	1,64	5,32	234,04	Relief assez fort
	1704	437,6	95,8	1280	1100	1500	1,28	7,41	154,95	Relief assez fort
	1705	1434	200	1200	800	1600	1,48	4,55	172,13	Relief assez fort
	1706	1628	213	910	700	1600	1,48	8,57	345,84	Relief fort
Sous total		8989		1020	700	1600				

Source : Agence Nationale des Ressources Hydrauliques (ANRH), Alger ,2010)

Tableau II.1. 2. Longueur des oueds permanents et temporaires par bassin versant.

Bassin	Code sous bassin	Nom sous bassin	Longueur permanents Lp (km)	Longueur temporaires Lt (km)	Longueur Total L(km)
Cheliff à l'aval Bougezoul	113	Oued Cheliff Djellil	167	439	606
	114	Oued Cheliff Ghrib	197	598	795
	115	Oued Cheliff Harbil	72	326	398
	116	Oued Deurdeur	176	329	505
	117	Oued Cheliff	107	252	359
	118	Oued Ebda	75	393	467
	119	Oued Rouina	178	575	753
	120	Oued Cheliff Tikazel	55	345	401
	121	Oued Fodda	419	657	1 076
	122	Oued Ras Ouahrane	172	988	1 160
	123	Oued Sly	564	670	1 233
	124	Oued Cheliff	119	298	417
	125	Oued Tiguiguest	546	527	1 073
	126	Oued Rhiou Tleta	277	442	719
	127	Oued Djidiouia	254	359	613
	128	Oued Cheliff Tarhia	316	180	497
	129	Oued Mina amont	263	535	798
	130	Oued Taht	58	393	451
131	Oued Mina moyenne	186	365	551	
132	Oued Abd amont	28	576	604	
133	Oued Abd aval	124	754	878	
134	Oued Mina hadda	287	723	1 011	
135	Oued Mina aval	171	535	707	
136	Oued Cheliff Mari-	87	105	192	
Sous- Total			4 898	11 366	16 264
Cheliff à l'amont Bougezoul	101	Oued Sebgag	114	428	542
	102	Oued Namous	121	1 274	1 395
	103	Oued Touil Amont	38	502	540
	104	Oued Sakni	-	363	363
	105	Oued Touil Moyen	-	641	641
	106	Oued Touil Aval	5	842	847
	107	Oued Sousselem	233	1 340	1 573
	108	Oued Mechti Zerga	260	552	812
	109	Oued Ouassel	113	208	321
	110	Oued Ouassel	319	621	940
	111	Dait El Firania	100	702	802
112	Diaa Boughzoul	165	922	1 087	
Sous Total			1 468	8 395	9 866
Zahrez	1701	Oued Mesrane	-	940	940
	1702	Oued Djelfa Hadjia	116	1 356	1 472
	1703	Zahrez Rharbi	-	700	700
	1704	Daiet Mefiteg	-	190	190
	1705	Oued Medjedel	-	711	711
	1706	Zahrez Chergui	-	545	545
Sous Total			116	4 442	4 558

Source : Agence Nationale des Ressources Hydrauliques (ANRH), Alger

II.1.2. LE RESEAU PLUVIOMETRIQUE

Dans le cadre de l'analyse pluviométrique, nous accentuerons tout d'abord sur la critique et l'homogénéisation des données statiques, Aussi une telle opération est-elle nécessaire pour envisager les possibilités d'amélioration des données d'observation en vue de disposer de séries plus continues et plus fines à l'échelle du terrain d'étude. Après quoi, nous aborderons les principaux caractères de précipitations et l'importance de leur variabilité.

Le bassin Chéelif-Zahrez est équipé d'un réseau pluviométrique de 174 postes, géré par l'Agence Nationale des Ressources Hydrauliques (ANRH), répartis uniformément à travers le bassin, avec une densité d'un poste pour 315km². (Boucefiane AEK, 2010)

II.1.2.1. Caractéristiques du réseau pluviométrique :

- La partie septentrionale du bassin aval-Boughzoul zone des grandes altitudes (hautes plaines), la plus arrosée est la paradoxalement la plus dépourvue en postes pluviométriques.
- 67 postes pluviométriques dans le haut et le moyen Cheliff, soit une densité d'un poste pour 208 km². 61 postes sont en service et 6 sont actuellement à l'arrêt.
- 40 postes pluviométriques dans le bas Cheliff et la Mina avec une densité d'un poste pour 254 km². 31 postes sont opérationnels et 9 postes sont à l'arrêt.
- 28 postes pluviométriques dans le bassin Amont du barrage Boughzoul, soit une densité d'un poste pour 1000 km². 26 postes sont opérationnels et (02) sont actuellement à l'arrêt.
- 09 postes pluviométriques dans le bassin du Zahrez, dont (03) postes sont en service, avec une densité très faible, soit moins d'un poste pour 1000 km².
- 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), (ANRH, Alger)

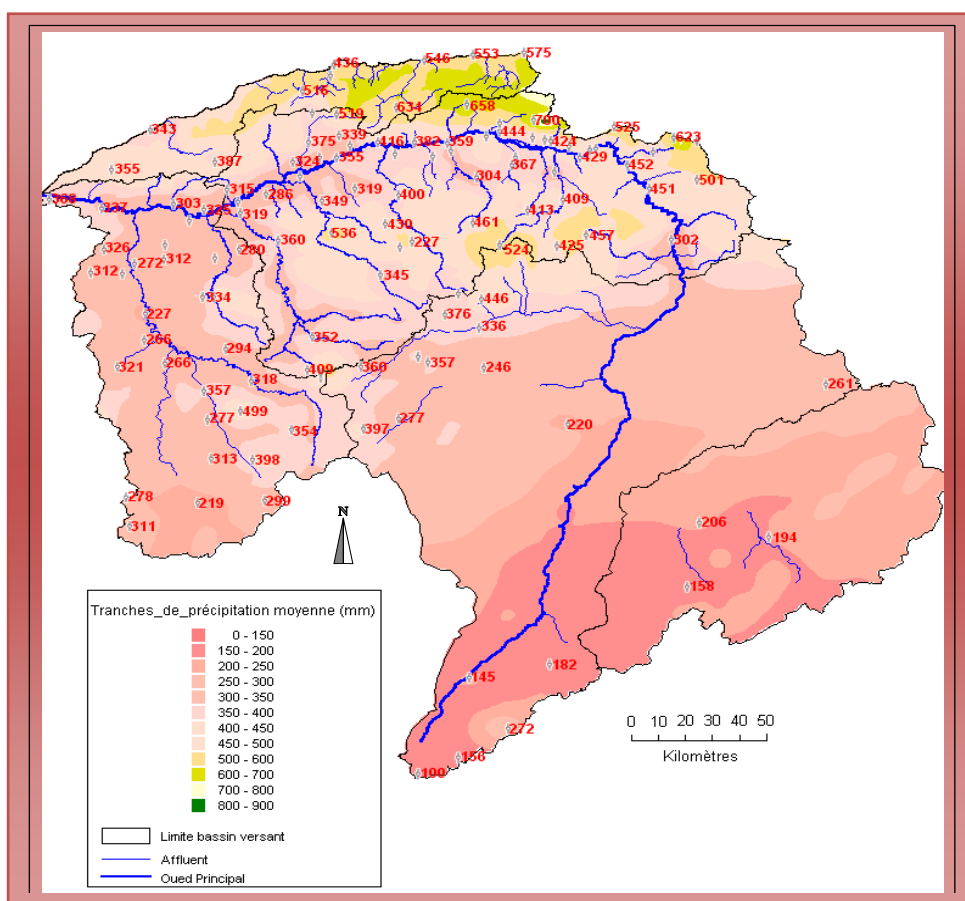
II.1.2.2. Répartition géographique des précipitations

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 seule, pour cerner la disposition globale dans les zones pluviométriques, (ANRH, Alger)

- Une concentration des précipitations sur les versants Est des monts du Dahra et Nord du Zaccar (bassin Côtier), avec une moyenne interannuelle de plus de 600 mm (station F'kirina (Tachta) 020231 ; 634 mm).

- Une disposition longitudinale des isohyètes en bandes plus ou moins parallèles et valeurs décroissantes des précipitations de l'Est vers l'Ouest marquée aussi par une diminution du gradient altimétrique du bassin Cotier (de 634 mm à la station de F'kirina (Tachta) à 355 mm à la station de Sidi Ali, 020107).
- 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.
- La disposition longitudinale des isohyètes en bandes plus ou moins parallèles et les valeurs décroissantes du Nord au Sud, est toutefois, altérée au niveau de la chaîne de l'Ouarsenis où intervient le rôle important de l'altitude ainsi que la dissymétrie pluviométrique entre les versants du Nord-Ouest au vent et les versants Sud-Est sous le vent. Ainsi, on remarque bien une augmentation de l'Ouest vers l'Est marquée aussi par l'augmentation du gradient altimétrique du bassin.
- Pour le bassin versant du Bas Cheliff et de la Mina, les isohyètes illustrent clairement l'augmentation de la concentration des précipitations proportionnellement à l'altitude de 200 mm à 500 mm; les versants Nord-ouest des monts de Tiaret ouverts au vent de la mer sont les plus arrosés, avec une moyenne interannuelle de plus de 400 mm (Sidi Bakht (013103) et Tiaret ANRH (012917)). Les précipitations se réduisent dans la plaine du Bas Cheliff où elles varient entre 200 et 350 mm.
- La région des hautes plaines (bassin Amont Boughzoul et Zahrez) est cernée en gros par la tranche d'eau moyenne précipitée 300-350mm dont le tracé suit étroitement la limite des versants Sud-tellien, sauf au Nord, au niveau de limite de l'Ouarsenis où intervient le rôle important de l'altitude, où la limite remonte jusqu'à 500 mm. Et plus au Sud par l'isohyète 150mm.
- Dans cette région (Amont Boughzoul et Zahrez) le climat semi-aride est particulièrement développé et l'espacement des isohyètes est très important. C'est là où l'effet de la latitude est très sensible : les perturbations cycloniques arrivent appauvries en vapeur d'eau suite à leur passage sur l'écran montagneux du Tell ; la principale remarque qui découle dans la région Amont Boughzoul et Zahrez à trait à la diminution générale des précipitations (145 mm à la station Guelta Sidi Saad, 206 mm à la station de Zaafrane, 246 mm à la station de Recheiga) par rapport aux régions Côtier et Aval Boughzoul

Le phénomène de variabilité spatiale est donc très sensible dans la partie Sud de la région hydrographique, (ABH, Chlef)



Source ABH Cheliff (SIG/ carte pluviométrique 1968/1969-2001/2002)

Fig II.1.4. Répartition géographique des précipitations

II.1.2.3. Acquisition et traitement des données pluviométriques

Les données pluviométriques ont été recueillies auprès de l'Agence Nationale des Ressources Hydrauliques (A.N.R.H) et l'Office Nationale de Météorologie (O.N.M), sous forme de tableaux des cumuls mensuels et annuels. Elles étaient très hétérogènes tant du point de vue fiabilité de mesure que de la durée d'observation, ce qui a conduit à définir une période (la plus longue possible) et d'y apporter le soins du traitement nécessaire, afin qu'on puisse disposer de données suffisamment fiables du point de vue qualité et quantité pour mener à bien notre travail.

II.1.2.4. Choix de la période de référence

Après une analyse plus poussée des données disponibles de la durée des séries d'observations, on a trouvé que les séries qui ont une période de fonctionnement identique et commune et aussi pour répondre aux exigences de l'application de filtre de Kalman multi-site pour la modélisation ; sont situées entre 1959-2009, sur cette base, cette période de 51 ans a été choisi comme période de référence.

Pour cette période, 39 stations pluviométriques ont été sélectionnées, ils disposent dans leurs ensemble 1989 observations annuelles réelles. (Annexe II.1.1)

La figure suivante, représente toutes les stations pluviométriques retenues.

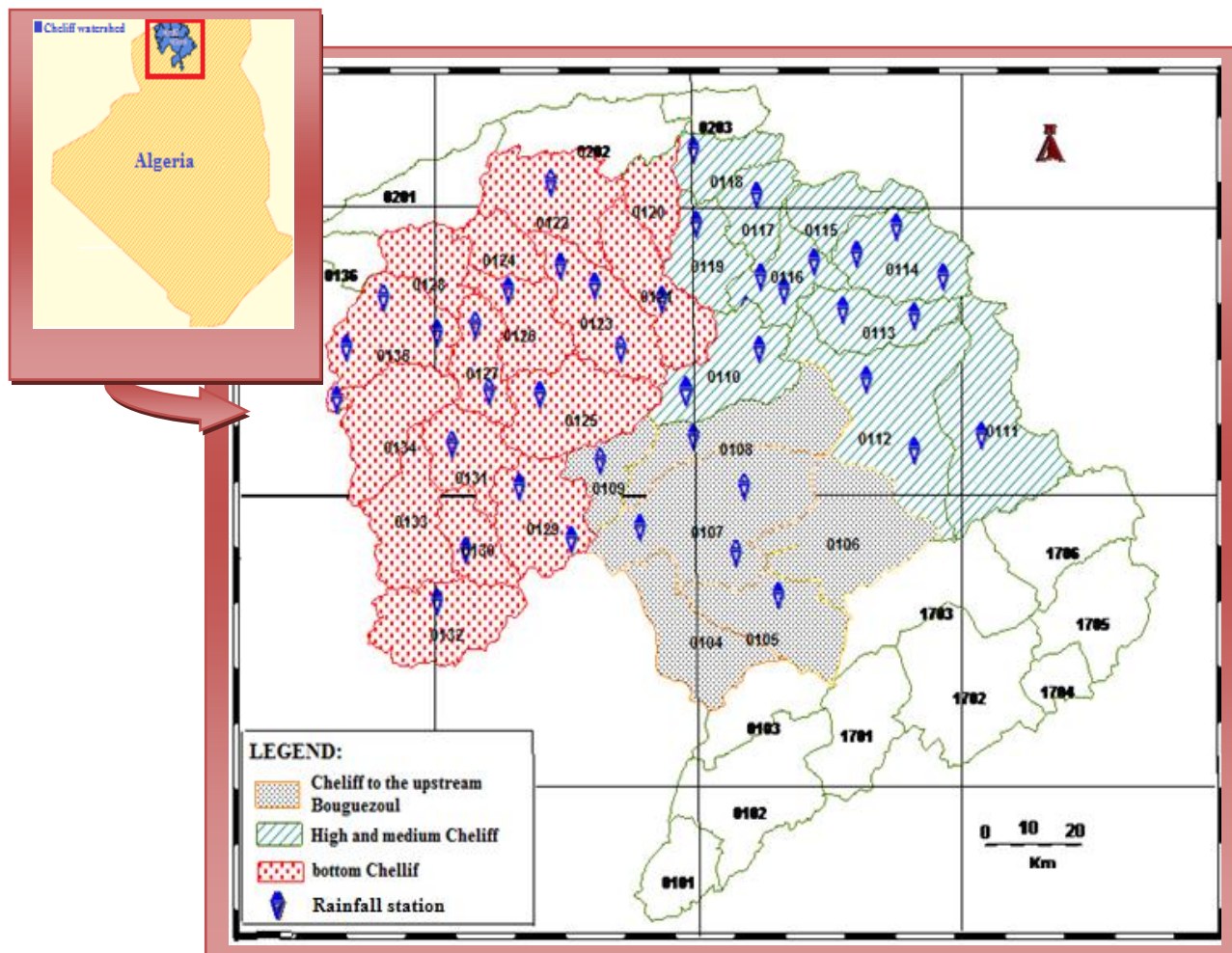


Fig II.1.5. Les stations pluviométriques retenues.

II.1.2.5. Distribution statistique des pluies annuelles :

L'étude de la distribution statistique des précipitations annuelles en Algérie a fait objet de nombreux travaux dont celui de M.CHAUMONT. Ce dernier a constaté que la distribution statistique des précipitations annuelles présentait une nette dissymétrie positive, il préconisait deux modèles de distributions.

Une distribution log-normale en régime « méditerranéen » et une distribution racine-normale en régime « sub-tropical ». Cette constatation a été certifiée par l'agence nationale des ressources hydraulique dans le cadre du projet PNUD1989, et par M.Meddi (1995).

*Annexe (II.1.1 : les stations avec leurs codes et leurs coordonnées Lambert

Nous avons testé ces deux types de lois ainsi que la loi normale sur les 39 stations. Les séries les plus longues et les plus fiables suivent une distribution racine-normale. Ce test a été fait pour les données réellement observées et non encore homogénéisées. (Annexe II.1.2)

II.1.3. HOMOGENEISATION DES DONNEES :

Ayant montré que les données pluviométriques s'ajustent à une loi racine-normale, nous allons passer à l'homogénéisation des données pluviométriques.

L'étude de l'évolution des précipitations ne peut utiliser les données hétérogènes. Plusieurs méthodes statistiques permettent de tester l'homogénéité des séries pluviométriques avant utilisation.

L'analyse statistique des données pluviométriques consiste à déterminer les caractéristiques empiriques d'un échantillon d'une série d'observations de précipitations mensuelles de 51 années.

Deux méthodes ont été adoptées (Annexe II.1.3), pour combler les lacunes dans les données mensuelles et annuelles :

- La régression simple
- L'analyse en composantes principales ACP.

II.1.3.1. Méthode de la régression linéaire :

Le but de cette méthode consiste à estimer les variables pluviométriques inconnues de certaines stations à partir des observations d'autres stations voisines prises pour base, il faut au préalable, que certaines conditions soient satisfaites, notamment la linéarité de la régression et l'ajustement à une loi de normale des variables confrontées.

*Annexe (II.1.2) : Ajustement des séries pluviométriques

*Annexe (II.1.3) : Les différentes méthodes de comblement des lacunes

La méthode des moindres carrés a été utilisée pour combler les données manquantes dans les séries ; Elle s'effectue par le calcul de régression simple de type linéaire entre les séries lacunaires et les séries des stations complètes environnantes aux latitudes semblables et quasi immédiates ; pour cela nous disposons des observations de quatre (04) stations pluviométriques.

Il s'agit des stations pluviométriques de base dans chaque sous bassin versant suivante :

- St 02 : Station Ain baadji sur Oued Sousselem (010701).
- St 17: Station Ghrib sur Oued Cheliff Ghrib (011405).
- St 33 : Station Bordj Bou Naama sur Oued Ras Ouahrane (012304).
- St 39: Station L'Hillil sur Oued Mina (013507).

❖ **Résultats de la régression et la corrélation :**

Le comblement des lacunes par régression simple est basé sur les coefficients de corrélations significatifs. On prend par exemple, la station de Zemalet El Amir Aek et le reste des résultats sont présentées dans : (Annexe II.1.4)

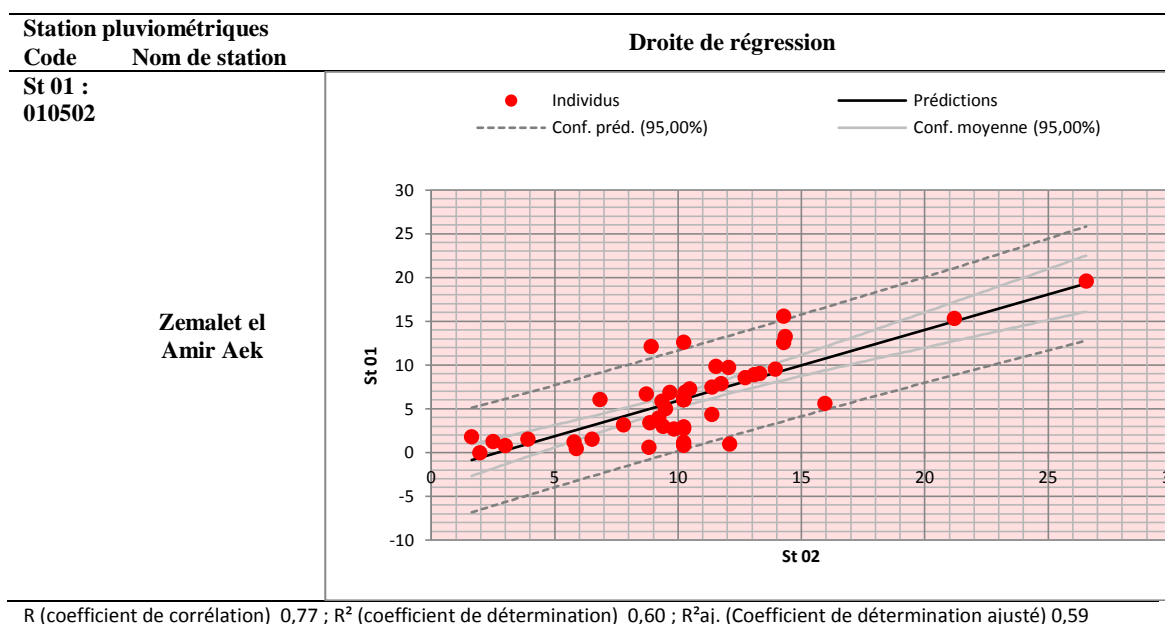


Fig II.1.6. la régression linéaire (la station de Zemalet El Amir AEK (010502))

*Annexe (II.1.4: La régression linéaire

▪ A/ L'échelle Annuel :

Le comblement des lacunes par régression basée sur les coefficients de corrélations les plus significatifs n'est pas toujours fiable aux échelles annuelles car : Les coefficients de corrélations annuelles ne sont pas toujours représentatifs des conditions climatiques réelles du bassin de fait que les régimes climatiques sont très nuancés d'une région à l'autres ; Il a été indispensable de prendre en considération des réalités géographiques de toute la région.

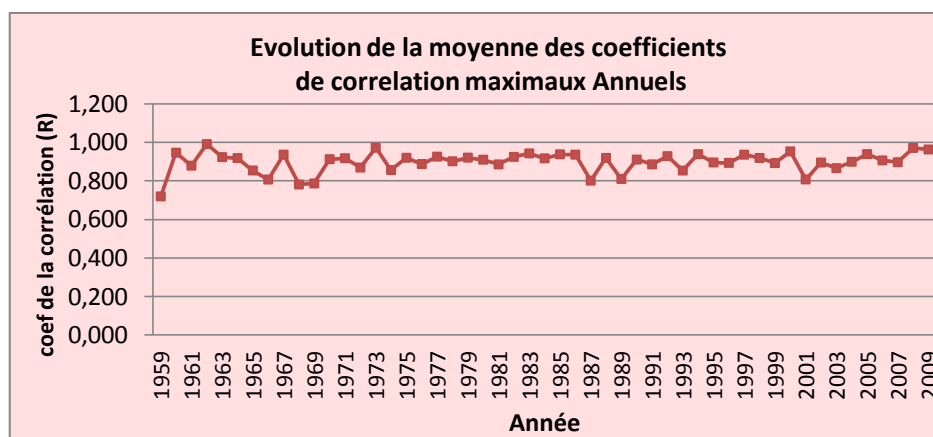


Fig II.1.7. La corrélation annuelle de toutes les stations dans le bassin versant de Cheliff

▪ B/ A L'échelle Mensuel :

Le degré de corrélation est différent d'un mois à un autre ; La meilleure corrélation entre les stations est située en hiver (janviers, décembre, février) et à la fin du printemps (mai) alors que, elle est moins élevée en été et plus particulièrement pour le mois d'août qui est caractérisé par des averses orageuses.

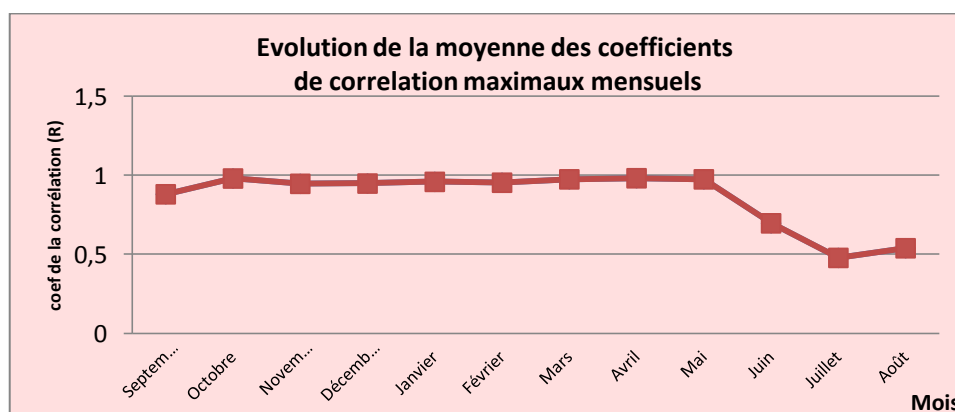


Fig II.1.8. La corrélation mensuelle de toutes les stations dans le bassin versant de Cheliff

Rappelons que cette méthode de comblement des lacunes ne permet jamais de retrouver exactement ce qu'aurait pu être une mesure « vraie » (Laborde, 1999). Au contraire, elle a permis « de remplacer les valeurs manquantes par des estimations présentant des qualités

(vraisemblance) et des indices de confiance (variance d'estimation), mais toujours sous réserve de validité du modèle statistique retenu ».

La présence de ce hasard statistique peut provoquer des erreurs de jugement, car on risque d'établir des régressions entre des stations dont le coefficient de corrélation apparaît très élevé alors qu'elles n'ont en réalité aucun lien climatique. De ce fait, il a été très indispensable de prendre en considération des réalités géographiques de la région.

Le comblement des lacunes par régression avec des stations proches et disponibles, est fastidieuse et peu fiable (à cause des erreurs accidentelles et ou systématiques). Donc, une nouvelle approche plus globale a été définie permettant de lisser les erreurs et anomalies de mesure.

II.1.3.2. Méthode de l'ACP:

L'Analyse en composantes principales est une méthode purement statistique consiste à traiter les séries de données pluviométriques en tenant comptes des réalités géographiques et des meilleures affinités climatiques (voir chapitre I.1). L'avantage de cette méthode est de restituer les données manquantes dans les séries observées à une échelle plus fine (mensuelle non pas annuelle), elle permet d'analyser un nombre important de variables et de faire ressortir la relation entre ces variables dans le cas de son existence.

L'Analyse en Composantes Principales intervient, en premier lieu, dans le but d'analyser le taux de la variance pour chaque axe (chaque composante), l'objectif étant de dégager les axes qui expliquent un taux élevé de la variance. Puis nous analyserons, par la suite, les relations existant entre les axes et les variables. Enfin nous examinerons la matrice de corrélation pour voir la relation qui existe entre les variables.

1. Le travail effectué par la méthode de l'ACP :

■ On a examiné premièrement chaque sous bassin versant à part par la méthode de l'ACP, et on a conclu les résultats suivants :

-Dans l'analyse de la matrice de corrélation. On conclut que la station pluviométrique St02 se corréle bien avec les 03stations pluviométriques St 03, St06 et St07 avec des coefficients de corrélation supérieur ou égal à 0,7.

- Dans l'analyse de la matrice de corrélation. On conclut que la station pluviométrique St17 se corréle bien avec les 02stations pluviométriques St16, St 19 avec des coefficients de corrélation supérieur ou égal à 0,7 et 0,8.

- Dans l'analyse de la matrice de corrélation. On conclut que la station pluviométrique St39 se corréle bien avec les 02stations pluviométriques St36, St 37 avec des coefficients de corrélation supérieur ou égal à 0,7. (Annexe II.1.5)

■ Ensuite on a effectuée sur les racines carrées des données pluviométriques de toutes les 39stations.

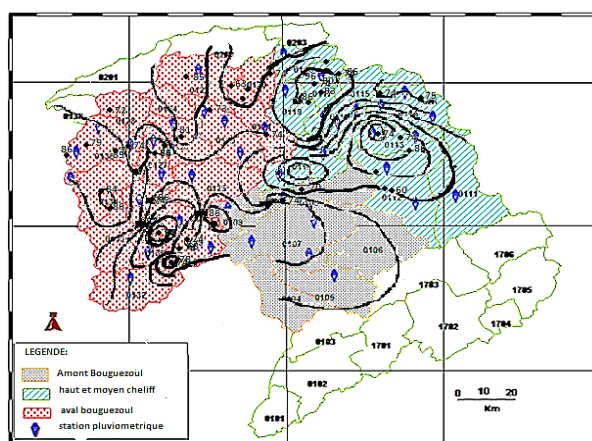
Annexe (II.1.5 : Résultats de l'ACP dans le bassin versant à l'amont Bouguzoul « 010 », le bassin versant Haut et moyens Cheliff « 011 » ; le bassin versant Bas Cheliff « 012 » ; le bassin versant à l'aval Bouguzoul « 013 »)

2. Résultats :

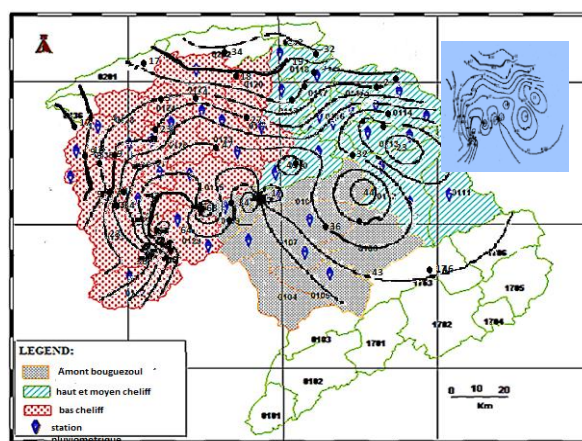
Tableau II.1. 2. Résultats de l'Analyse en composantes principales

Composantes	F1	F2	F3
Valeur propre	38,45	15,09	2,66
Variabilité (%)	53,79	22,18%	5,32%
% cumulé	53,79%	75,97%	81,29%

Les résultats montrent qu'une bonne de part de la variance (76%) est accaparée par les deux seules premières composantes et qu'à partir de la troisième, la part expliquée devient faible, moins de (10%).



(a)



(b)

- a) Première composante principale : représente la cartographie de la répartition temporelle d'ensemble des précipitations annuelles, Elle indique le coefficient de corrélation entre chaque variable et la première composante (F1).
- b) Deuxième composante : représente la cartographie de la répartition spatiale de la variabilité pluviométrique dont le coefficient de corrélation positif de Sud au Nord du bassin, Ce qui explique que la pluviométrie dans le bassin Cheliff diminue du nord au sud (éloignement de la mer).
- c) La troisième composante principale n'as pas de structure spatiale nette, elle explique une part très faible de la variance ($5% < 10%$).

Nous avons considéré donc que seules les deux premières composantes ont une signification climatique avec un pourcentage de variance de 76%. Cela signifie que les séries pluviométriques se ramènent à une combinaison linéaire de ces deux vecteurs avec un Coefficient de corrélation multiple de l'ordre de $(0,7597)^{0,5} = 0,87$

Conclusion

Ce chapitre, a permet d'analyser les précipitations de notre région d'étude (le bassin versant de Cheliff), Nous avons donc utilisé toute information disponible et choisir le nombre définitif des stations pluviométriques retenue de 39, qui ont une série hydrologiques identique, observés sur une même période (51ans) et cela pour répondre aux exigences de la technique de Filtre de Kalman pour la prédiction multi-site des précipitations.

On peut déduire à travers ce chapitre:

- L'étude critique des données d'observations pluviométriques pour choisir la période la plus longue possible possédant peu de lacunes et homogène (méthode de régression simple et méthode de l'ACP)
- Les séries pluviométriques les plus longue et les fiables suivent une distribution racine-normale.
- La détections des erreurs commis lors du traitement des données, nous a permet de choisir le nombre des stations pluviométriques retenue dans notre étude (39)
- Les racines carrées des données pluviométriques de toutes les 39 stations montrent que la variance de 76% est accaparée par les deux seules premières composantes et qu'à partir de la troisième, la part expliquée devient très faible, moins de (10%) ; cela explique que les séries pluviométriques se ramènent à une combinaison linéaire avec un Coefficient de corrélation multiple de l'ordre de 0,87
- les séries chronologiques des précipitations annuelles et mensuelles, couvrant une période d'observation commune de 51 ans (1959 à 2009) ont été investies. Ces dernières sont observées à 39 stations pluviométriques dans le bassin versant de Cheliff.

Introduction

L'objectif de ce chapitre est l'étude de l'applicabilité de la technique du filtre de Kalman (FK) discret à la modélisation et la prédiction multi-site de la pluviométrie annuelle et mensuelle dans le bassin versant de Chélif, ainsi que l'amélioration des erreurs de prédiction. L'avantage majeur du FK est la procuration de l'erreur de prédiction qui présente en soit un indicateur de précision. En outre, 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 résultat escompté serait un modèle optimal de prédiction en ligne qui ne doit être figé, ni dans l'espace ni dans le temps, et qui s'adapterait automatiquement aux changements dans les conditions météorologiques du bassin versant en question.

II.2.1. DONNEES DE LA PREDICTION MULTI-SITE

Les données utilisées sont les précipitations annuelles et mensuelles observées par l'Agence Nationale des Ressources Hydrauliques(ANRH) sur le bassin versant de Chélif.

Ces données sont enregistrées dans trente-neuf stations pluviométriques, quatre d'entre elles sont situées sur des barrages en exploitation. Leur répartition sur le bassin versant du Chélif est tel que montré sur la Figure II.1.6 (voir Chapitre II.1). Ces données forment des séries chronologiques ayant une période d'observation commune de 51 ans allant de 1959 à 2009.

II.2.1.1. HYPOTHESES

- Le système est linéaire. : Cette hypothèse est vérifiée par le filtre qui est linéaire
- Les résidus sont des bruits blancs gaussiens. : Cette hypothèse doit être vérifiée à la fin du travail :(avec les tests de l'autocorrélation pour tester les résidus)

II.2.1.2. Formulation espace-état du modèle adopté

L'une des étapes les plus importantes dans l'application de la technique du Filtre de Kalman est la formulation de l'équation d'état et de la mesure conformément à la structure d'un modèle espace-état tel qu'indiqué dans le Chapitre I.2.

La variable état du système est un vecteur formé selon le cas, des précipitations annuelles et mensuelles observées simultanément aux trente-neuf stations pluviométriques considérées. Ainsi, pour $n=39$ le vecteur d'état du système et de la mesure a pour dimensions (39×1) , alors que les matrices de covariance de transition d'état et celle de mesure sont de dimension (39×39) . Par suite, l'équation d'état et celle de mesure sont respectivement formulées de la manière suivante :

A- Equation d'état

$$\text{Avec : } \begin{bmatrix} X_1 \\ X_2 \\ \cdot \\ \cdot \\ X_{39} \end{bmatrix}_k = \begin{bmatrix} \varphi_{1,1} & \varphi_{1,2} & \cdot & \varphi_{1,39} \\ \varphi_{2,1} & \varphi_{2,2} & \cdot & \varphi_{2,39} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \varphi_{39,1} & \varphi_{39,2} & \cdot & \varphi_{39,39} \end{bmatrix}_{k/k-1} \begin{bmatrix} X_1 \\ X_2 \\ \cdot \\ \cdot \\ X_{39} \end{bmatrix}_{k-1} + \begin{bmatrix} W_1 \\ W_2 \\ \cdot \\ \cdot \\ W_{39} \end{bmatrix}_{k-1} \quad \text{..... (II.2.1)}$$

$\begin{bmatrix} X_1 \\ X_2 \\ \cdot \\ \cdot \\ X_{39} \end{bmatrix}_k$: Vecteur d'état au temps k

$\begin{bmatrix} \varphi_{1,1} & \varphi_{1,2} & \cdot & \varphi_{1,39} \\ \varphi_{2,1} & \varphi_{2,2} & \cdot & \varphi_{2,39} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \varphi_{39,1} & \varphi_{39,2} & \cdot & \varphi_{39,39} \end{bmatrix}_{k/k-1}$: Matrice de transition d'état

$\begin{bmatrix} W_1 \\ W_2 \\ \cdot \\ \cdot \\ W_{39} \end{bmatrix}_{k-1}$: Vecteur de bruit du système

B- Equation de mesure :

$$\begin{bmatrix} Z_1 \\ Z_2 \\ \cdot \\ \cdot \\ Z_{39} \end{bmatrix}_k = \begin{bmatrix} H_{1,1} & H_{1,2} & \cdot & H_{1,39} \\ H_{2,1} & H_{2,2} & \cdot & H_{2,39} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ H_{39,1} & H_{39,2} & \cdot & H_{39,39} \end{bmatrix}_k \begin{bmatrix} X_1 \\ X_2 \\ \cdot \\ \cdot \\ X_{39} \end{bmatrix}_k + \begin{bmatrix} V_1 \\ V_2 \\ \cdot \\ \cdot \\ V_{39} \end{bmatrix}_k \quad \text{..... (II.2.2)}$$

Avec :

$$\begin{bmatrix} Z_1 \\ Z_2 \\ \cdot \\ \cdot \\ Z_{39} \end{bmatrix}_k \quad : \text{Vecteur de mesure au temps } k$$

$$\begin{bmatrix} H_{1,1} & H_{1,2} & \cdot & H_{1,39} \\ H_{2,1} & H_{2,2} & \cdot & H_{2,39} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ H_{39,1} & H_{39,2} & \cdot & H_{39,39} \end{bmatrix}_k \quad : \text{Matrice de mesure au temps } k$$

$$\begin{bmatrix} X_1 \\ X_2 \\ \cdot \\ \cdot \\ X_{39} \end{bmatrix}_k \quad : \text{Vecteur d'état au temps } k$$

$$\begin{bmatrix} V_1 \\ V_2 \\ \cdot \\ \cdot \\ V_{39} \end{bmatrix}_k \quad : \text{Vecteur de bruit de mesure au temps } k$$

II.2.2. CONDITIONS INITIALES

Une fois que la formulation du modèle est accomplie, le filtre de Kalman et l'exécution de son algorithme requiert la spécification des grandeurs suivantes :

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

Ceci va être effectué selon les étapes suivantes :

II.2.2.1. Vecteur d'état initial et matrice de covariance d'erreur associée

En présence d'une information à priori, le vecteur d'état initial $X_{k/k-p}$ pour $k = 1$ est constitué de la moyenne des précipitations respectivement annuelles puis mensuelles observées aux 39 différentes stations pluviométriques en question ($\bar{X}_1 \ \bar{X}_2 \ \bar{X}_3 \ .. \ . \ . \ . \ \bar{X}_{39}$.) comme suit :

$$X_{1/0} = \begin{bmatrix} \bar{X}_1 \\ \bar{X}_2 \\ \cdot \\ \cdot \\ \cdot \\ \bar{X}_{39} \end{bmatrix} \dots\dots\dots \text{(II.2.3)}$$

Cependant, pour spécifier la matrice de covariance d'erreur associée $P_{k/k-1}$ qui n'est pas connue, nous pouvons commencer avec des éléments ayant des valeurs assez élevées dans la diagonale principale. De cette manière, l'algorithme bénéficiera d'une certaine flexibilité afin de s'ajuster aux valeurs sensibles, dans un temps relativement court.

Dans le présent travail, nous avons choisi une telle matrice (39x39) comme suit :

$$P_{1/0} = \begin{bmatrix} 1000 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & 1000 \end{bmatrix} \dots\dots\dots \text{(II.2.4)}$$

Ce choix va conduire à une augmentation de la matrice de covariance $P_{k/k-1}$ et celle du gain du filtre K_k , permettant ainsi au filtre adaptatif de pondérer plus lourdement la nouvelle information (la nouvelle mesure Z_k dans l'équation (1) du (Chapitre I.2 (& 1.2.3.3)). Ainsi, la trace de la matrice de covariance initiale est de 10000. Une telle valeur est supposée décroître d'une manière continue, pour atteindre asymptotiquement une valeur positive et stable proche de zéro. Cette réduction est accomplie au cours de l'exécution de l'algorithme du filtre de Kalman, qui s'adapte automatiquement dès que ce dernier reçoit une nouvelle information objective. Si cette valeur asymptotique de la trace de la matrice de covariance ne change pas significativement durant la procédure de calcul, on peut conclure que le filtre a convergé. Cette caractéristique attribuée à la trace de la matrice de covariance le mérite d'être une mesure fiable de la performance du filtre de Kalman.

II.2.2.2. Matrice de covariance des bruits du système et celle des bruits de mesure

De plus, 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 (39x39) ont été choisies comme suit :

$$Q = \begin{bmatrix} 100 & . & . & . & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ 0 & . & . & . & 100 \end{bmatrix} \dots\dots\dots \text{(II.2.5)}$$

$$R = \begin{bmatrix} 50 & . & . & . & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ 0 & . & . & . & 50 \end{bmatrix} \dots\dots\dots \text{(II.2.6)}$$

II.2.2.3. Matrice de transition d'état et matrice de mesure

L'une des difficultés majeures dans l'application du filtre de Kalman est l'estimation de la matrice de transition d'état $\phi_{k, -1}$. Il faut noter cependant, que la valeur de cette matrice n'affecte pas considérablement les résultats du filtre de Kalman, tel que l'ont montré les études de (Harrison and Stevens, 1975a, b). Par conséquent, pour l'estimation de la matrice de transition d'état, dans le présent travail on considère les inter-corrélations entre les observations des trente-neuf stations pluviométriques.

Quant à la matrice de mesure H_k , du moment que toutes les stations pluviométriques fournissent des observations, nous avons opté pour la matrice unité (39x39).

$$H_k = \begin{bmatrix} 1 & . & . & . & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ 0 & . & . & . & 1 \end{bmatrix} \dots\dots\dots \text{(II.2.7)}$$

II.2.3. LES EQUATIONS RECURRENTES DU FILTRE DE KALMAN

Les cinq relations matricielles récursives indiquées au (Chapitre I.2 ;(&) :1.2.4) sont les équations de base dans le calcul du filtre de Kalman. Elles sont utilisées de la manière qui suit ce qui constitue l'algorithme du filtre de Kalman.

- Mise à jour du vecteur d'état
 Equ (1): $\hat{X}_{k/k} = \hat{X}_{k/k-1} + K_k(Z_k - H_k\hat{X}_{k/k-1}) \dots\dots\dots \text{(II.2.8)}$

- Estimation du gain
 Equ (2): $K_k = P_{k/k-1}H_k^T(H_k P_{k/k-1}H_k^T + R_k)^{-1} \dots\dots \text{(II.2.9)}$

- Correction de la matrice de covariance de l'erreur d'estimation

$$\dots\dots\dots Equ (3): \quad P_{k/k} = (I - K_k H_k) P_{k/k-1} \quad \dots (II.2.10)$$

- Prédiction du vecteur d'état

$$\dots\dots\dots Equ (4): \quad \bar{X}_{k+1/k} = \phi_{k+1/k} \bar{X}_{k/k} \quad \dots (II.2.11)$$

- Prédiction de la covariance de l'erreur de prédiction

$$\dots\dots\dots Equ (5): \quad P_{k+1/k} = \phi_{k+1/k} P_{k/k} \phi_{k+1/k}^T + Q_k \quad \dots (II.2.12)$$

II.2.3.1. Matrice gain du filtre

Le calcul de la matrice gain du filtre K_k est la première étape de calcul dans le présent travail. Compte tenu de toute l'information précédente et des hypothèses du filtre, la matrice gain (39x39) peut être calculée pour une prédiction à un pas en utilisant l'équation indiquée au (chapitre I.2 & 1.2.3.3), pour k=1 comme suit :

$$K_1 = \begin{bmatrix} 1000 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & 1000 \end{bmatrix} \left\{ \begin{bmatrix} 1000 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & 1000 \end{bmatrix} \begin{bmatrix} 50 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & 50 \end{bmatrix} \right\}^{-1} \quad \dots\dots\dots (II.2.13)$$

II.2.3.2. Utilisation de la mesure pour actualiser l'estimation

C'est la seconde étape des calculs pour le même pas de temps $k=1$. L'estimation est corrigée moyennant la nouvelle mesure Z_k afin d'améliorer l'estimation initiale $X_{1/0}$. Ceci est fait en faisant intervenir un terme de correction comprenant la différence $(Z_k - H_k X_{k/k-1})$ pondérée par le facteur K_1 tel qu'on obtient :

$$X_{1/1} = \begin{bmatrix} \bar{X}_1 \\ \bar{X}_2 \\ \cdot \\ \cdot \\ \cdot \\ \bar{X}_{39} \end{bmatrix}_{1/0} + \left\{ \begin{bmatrix} k_{1,1} & k_{1,2} & \cdot & k_{1,39} \\ k_{2,1} & k_{2,2} & \cdot & k_{2,39} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ k_{39,1} & k_{39,2} & \cdot & k_{39,39} \end{bmatrix} \begin{bmatrix} Z_1 \\ Z_2 \\ \cdot \\ \cdot \\ Z_{39} \end{bmatrix} - \begin{bmatrix} \bar{X}_1 \\ \bar{X}_2 \\ \cdot \\ \cdot \\ \cdot \\ \bar{X}_{39} \end{bmatrix} \right\} \quad \dots\dots (II.2.14)$$

II.2.3.3. Matrice de covariance d'erreur- associée à l'estimation corrigée par la mesure

Par substitution de la matrice de covariance d'erreur $P_{1/0}$ associée au vecteur d'état initial $X_{1/0}$ de la matrice gain K_1 ainsi que la matrice de mesure H_k indiqué au chapitre I.2, nous pouvons obtenir la matrice de covariance d'erreur associée à l'estimation corrigée par la mesure, comme suit :

$$P_{1/1} = \begin{bmatrix} 1000 & \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & 1000 \end{bmatrix}_{1/0} - \begin{bmatrix} k_{1,1} & k_{1,2} & \cdot & k_{1,39} \\ k_{2,1} & k_{2,2} & \cdot & k_{2,39} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ k_{39,1} & k_{39,2} & \cdot & k_{39,39} \end{bmatrix} \begin{bmatrix} 1000 & \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & 1000 \end{bmatrix}_{1/0} \quad (\text{II.2.15})$$

II.2.3.4. Prédiction un pas en avant de l'estimation (actualisée) corrigée par la mesure

L'estimation corrigée $X_{1/0}$ est projetée un pas en avant par l'intermédiaire de l'équation (Chapitre I.2 (&) 1.2.3.3). Où le terme W_{k-1} est ignoré simplement parce que sa moyenne est nulle et parce qu'il est non corrélé avec les W précédents. Ceci nous conduit à la prédiction un pas en avant de l'estimation corrigée tel que :

$$\begin{bmatrix} X_1 \\ X_2 \\ \cdot \\ \cdot \\ X_{39} \end{bmatrix}_{2/1} = \begin{bmatrix} \phi_{1,1} & \phi_{1,2} & \cdot & \phi_{1,39} \\ \phi_{2,1} & \phi_{2,2} & \cdot & \phi_{2,39} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \phi_{39,1} & \phi_{39,2} & \cdot & \phi_{39,39} \end{bmatrix}_{2/1} \begin{bmatrix} X_1 \\ X_2 \\ \cdot \\ \cdot \\ X_{39} \end{bmatrix}_{1/1} \quad (\text{II.2.16})$$

II.2.3.5. Matrice de covariance de l'erreur de prédiction un pas en avant

De même, en substituant la matrice de transition d'état, la matrice de covariance d'erreur associée au vecteur d'estimation corrigée $P_{1/1}$ et la matrice de covariance du bruit de système Q dans (Chapitre I.2 , (&) 1.2.2.4), Nous pouvons calculer la matrice de covariance d'erreur de prédiction un pas en avant $P_{2/1}$ associée à $X_{2/1}$ comme suit :

$$P_{2/1} = \begin{bmatrix} \phi_{1,1} & \phi_{1,2} & \cdot & \phi_{1,39} \\ \phi_{2,1} & \phi_{2,2} & \cdot & \phi_{2,39} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \phi_{39,1} & \phi_{39,2} & \cdot & \phi_{39,39} \end{bmatrix}_{2/1} \begin{bmatrix} P_{1,1} & P_{1,2} & \cdot & P_{1,39} \\ P_{2,1} & P_{2,2} & \cdot & P_{2,39} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ P_{39,1} & P_{39,2} & \cdot & P_{39,39} \end{bmatrix}_{1/1} \begin{bmatrix} \phi_{1,1} & \phi_{1,2} & \cdot & \phi_{1,39} \\ \phi_{2,1} & \phi_{2,2} & \cdot & \phi_{2,39} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \phi_{39,1} & \phi_{39,2} & \cdot & \phi_{39,39} \end{bmatrix}_{2/1}^T + \begin{bmatrix} 100 & \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & 100 \end{bmatrix} \quad \dots \quad (\text{II.2.17})$$

Pour la prochaine itération dans le calcul, $X_{2/1}$ et $P_{2/1}$ sont considérés comme conditions initiales, et ainsi de suite pour le reste des itérations, jusqu'à la fin des calculs.

Comme tel, avec l'information disponible et les conditions précédentes, les équations multi-sites récursives du filtre de Kalman, ainsi que leurs boucles, ont été calculées durant toute la période d'observation. Ce calcul a été effectué sur deux temps le premier sert à l'estimation des paramètres du modèle, le deuxième sert à la validation de ce dernier.

II.2.4. RESULTATS ET INTERPRETATIONS

Pour développer et appliquer l'approche du modèle de FK à la modélisation et la prédiction multi-site des précipitations dans le bassin versant du Chélif, les observations enregistrées aux 39 stations pluviométriques, observées au pas de temps mensuel et annuel, sur une période de 51 ans (1959 - 2009) ont été étudiées. La variable état du système est un vecteur donnant les précipitations observées simultanément aux trente-neuf stations pluviométriques considérées.

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

II.2.4.1. Prédiction au pas annuel (temporelle et spatiale)

L'estimation des valeurs initiales évoquées plus haut est incontournable pour le démarrage des calculs. A cet effet, le vecteur d'état initial $\mathbf{X}_{k/k-1}$ (pour $k = 1$) est constitué par les moyennes des précipitations annuelles observées aux 39 stations pluviométriques en question ($X_{1, 2, \dots, 39}$) ; la covariance d'erreur associée $\mathbf{P}_{k/k-1}$ est une matrice ayant la valeur de 1000 sur la diagonale principale et zéro ailleurs. Une valeur aussi élevée sur la diagonale principale présente l'avantage de donner une certaine flexibilité à l'algorithme de calcul afin de s'ajuster aux valeurs sensibles dans un temps relativement court. Un tel choix va mener à une augmentation de la matrice de covariance $\mathbf{P}_{k/k-1}$ et celle du gain du filtre, permettant ainsi au filtre de pondérer plus lourdement la nouvelle mesure. Pour la covariance initiale du bruit du système \mathbf{Q} une matrice avec la valeur de 100 sur la diagonale principale et zéro ailleurs a été adoptée, et puisqu'on espère que les mesures soient moins bruitées que la dynamique du système on a pris une matrice de bruit de mesure \mathbf{R} avec la valeur de 50 sur la diagonale principale et zéro ailleurs. En ce qui concerne l'estimation de la matrice de transition d'état $\phi_{k/k-1}$ initiale elle est constituée par les inter-corrélations entre les observations des 39 stations pluviométriques ; **Harrison et Stevens (1975)** a montré que la valeur initiale d'une telle matrice n'affecte pas considérablement les résultats du filtre de Kalman. Quant à la matrice de mesure \mathbf{H}_k , du moment que toutes les stations pluviométriques fournissent des observations, nous avons opté pour la matrice unité (39x39).

La période commune d'observation entre les 39 stations pluviométriques étant de 51 ans, les observations des 30 premières années ont été utilisées à l'estimation des paramètres du modèle alors que les 21 dernières ont servi à la validation de ce dernier. 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 de précipitations annuelles pour les stations pluviométriques mentionnées plus haut.

La minimalisation de la trace de la matrice de covariance de l'erreur de prédiction est un critère de convergence et d'optimalité du filtre. Comme le montre la **Fig.II.2.1** la trace de cette matrice commence par des valeurs élevées, au début des calculs, puis décroît rapidement pour converger vers une valeur positive stable, la plus proche possible de zéro. Cette convergence confirme l'adéquation du modèle ajusté au processus étudié et signifie que l'algorithme de calcul est efficace et que les prédictions qu'il fournit sont optimales.

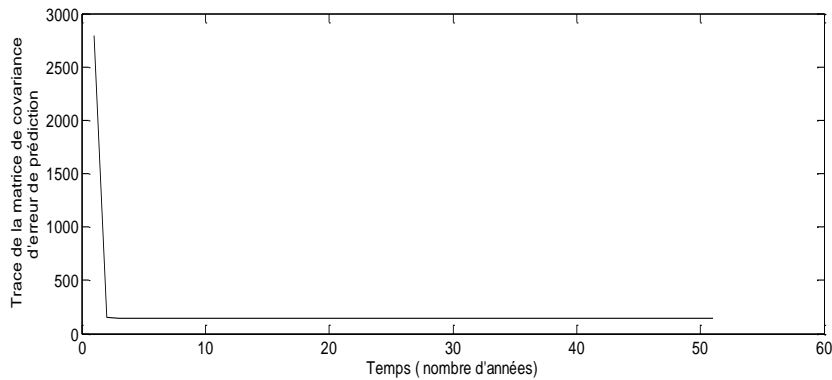


Fig II.2.1. Trace de la matrice de covariance d'erreur de prédiction

En effet, dans cette première phase la mesure est plus crédible que l'estimation fournie par le modèle car les calculs sont influencés par les conditions initiales tout à fait subjectives, mais avec le déroulement des calculs ces dernières sont vite ignorées, et le filtre accorde plus d'importance à la mesure qui sert au fur et à mesure à l'amélioration des estimations des paramètres du modèle.

Donc initialement, quand les paramètres du modèle ne sont que des estimations vagues, la mesure représente toute l'information objective et le rôle de la matrice gain est justement d'assurer que la mesure soit plus lourdement pondérée dans l'estimation des paramètres d'état. Dans ce cas, la matrice gain du FK prend des valeurs importantes, et la conséquence est une estimation automatiquement mauvaise ce qui explique des erreurs relativement importantes durant les premières itérations. Mais, avec le déroulement des calculs quand la confiance assignée à la précision des paramètres du modèle commence à augmenter, les valeurs de la matrice gain commencent à diminuer progressivement pour atteindre asymptotiquement une valeur proche de zéro **Fig.II.2.2** ce qui signifie la diminution de l'influence de la mesure dans la mise à jour de l'estimation des paramètres du modèle et les erreurs associées.

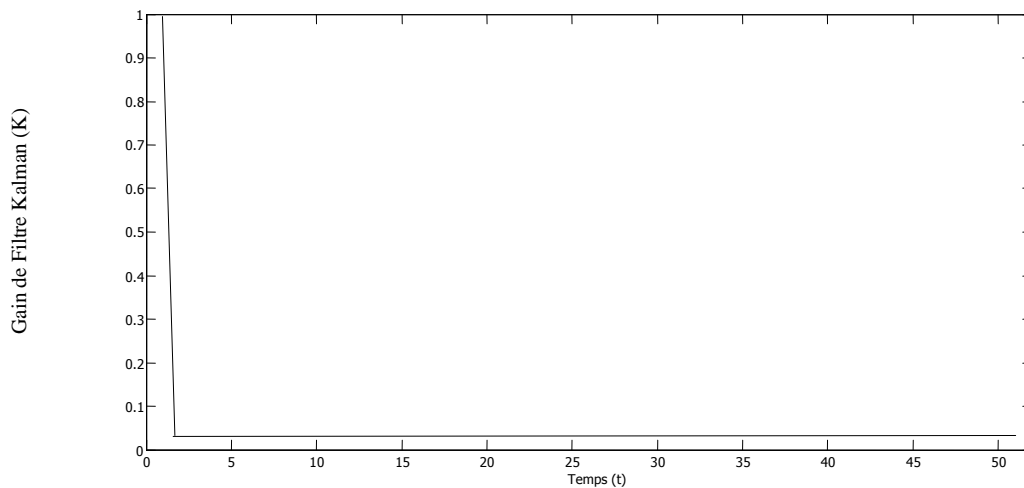


Fig II.2.2. Gain du filtre de Kalman

Les résultats montrent également que les prédictions obtenues pour les 39 stations sur les 51 années suivent de plus près les observations historiques, et les différences entre les deux sont minimales sauf pour les premières itérations. Ces différences sont données en termes de pourcentage d'erreur relative. La **Fig.II.2.3** en est un exemple, elle montre les prédictions obtenues à la station Teniet-el-had avec le pourcentage des erreurs relatives associées. L'examen de cette figure montre la grande concordance dans les variations des observations et des prédictions et montre également le degré de cette concordance en termes d'erreur relative en pourcent.

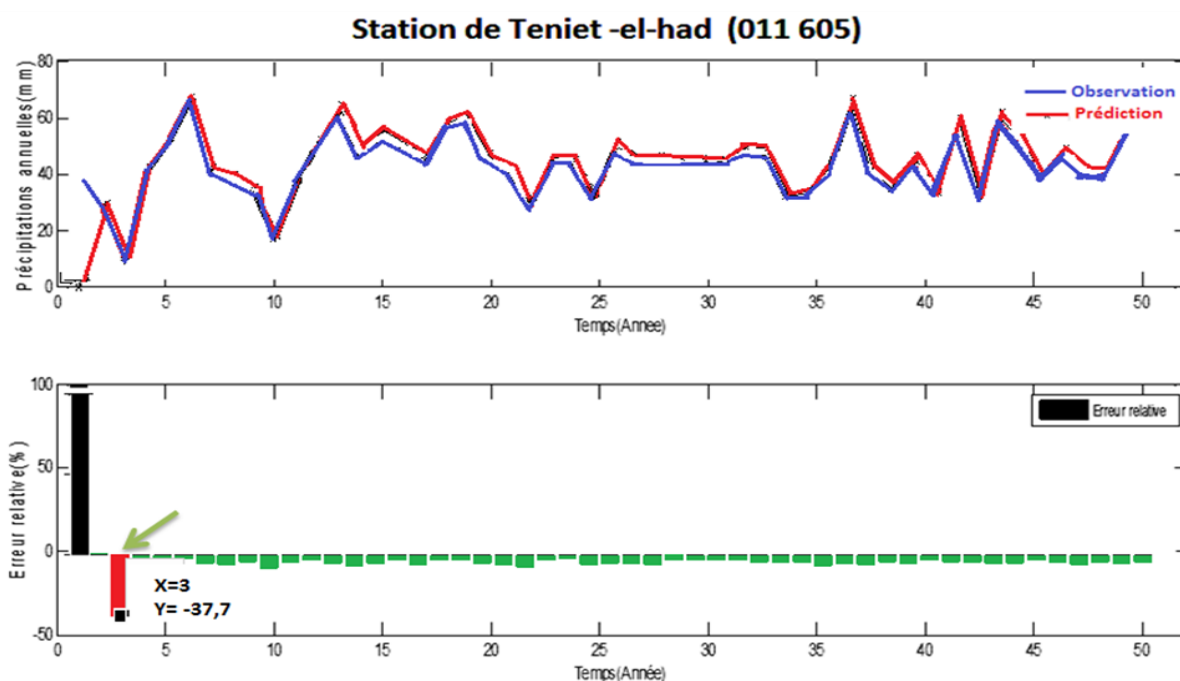


Fig II.2.3. Prédictions annuelles à la station pluviométrique Teniet - el-had de 1959 à 2009

D'un autre point de vue, la **Fig.II.2.3** constituée de deux figures, montre les observations et les prédictions obtenues simultanément dans les 39 stations pluviométriques, avec l'erreur relative de prédiction exprimée en pourcent. Celle de haut, correspond au début des calculs (1960), alors que celle du bas, correspond à la fin de ces derniers (2009). Sur la première, nous pouvons constater l'importante de l'écart entre les observations et leurs estimations (prédictions), ceci est dû au fait qu'au début des calculs, les paramètres du modèle ne sont pas encore bien établis pour donner de bonne estimations. Le modèle qui est initialisé avec des valeurs subjectives, au début des calculs, doit par la suite s'appuyer plus sur la mesure (seule information objective) afin de digérer l'information apportée par cette dernière ; ceci est indispensable afin de pouvoir s'adapter automatiquement par la suite et le résultat est un écart minimum entre l'observation et la prédiction au fil des itérations prochaines, comme le montre la **Fig.II.2.4**. En termes de pourcentage d'erreur relative cette différence qui commence par un maximum de -37,7 % à la station numéro 18 (Teniet-el-had) , à titre d'exemple, se réduit à la valeur de -4,3% vers la fin des calculs (2009).

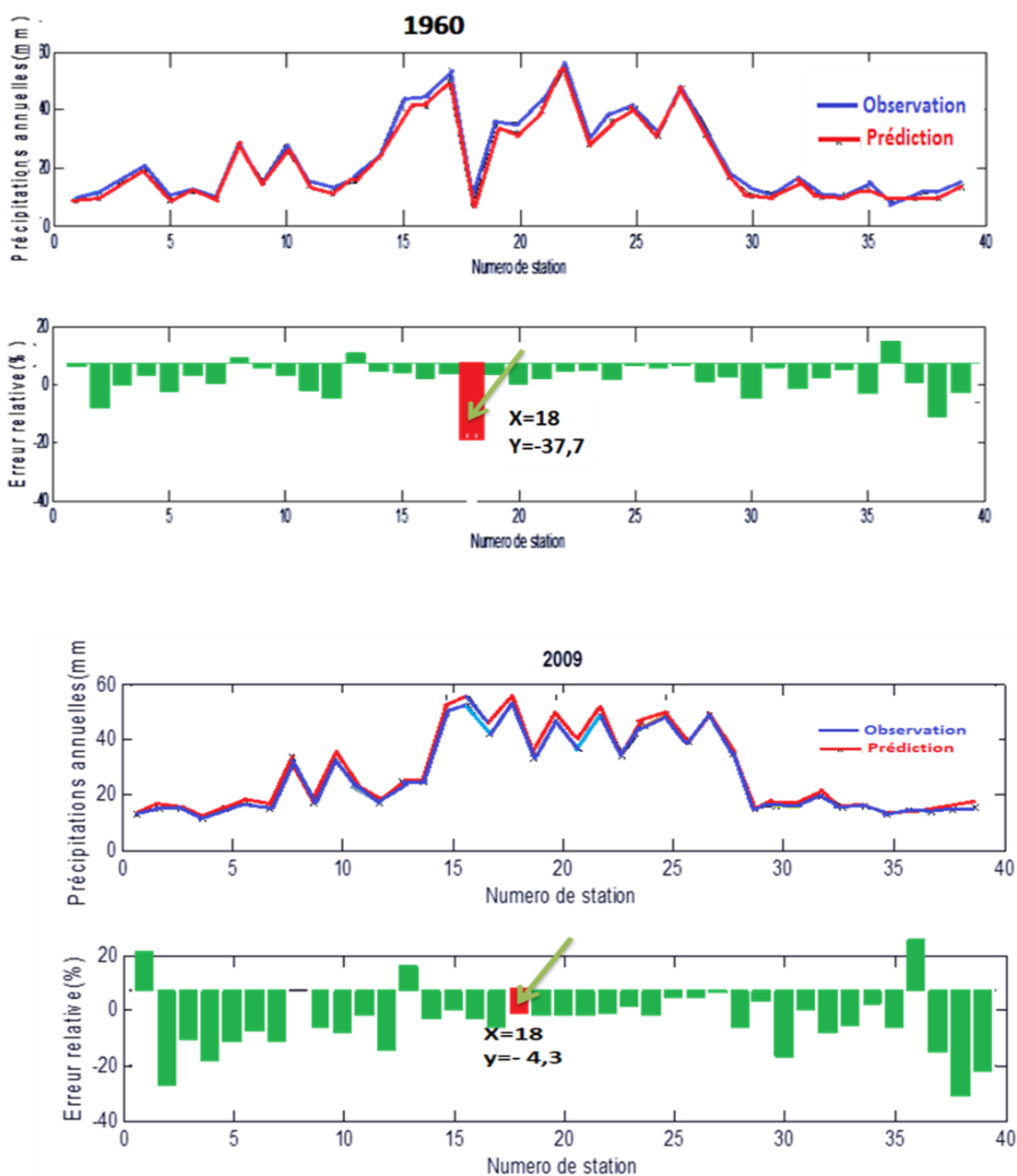


Fig II.2.4. Prédictions annuelles des précipitations aux 39 stations pluviométriques (exp :année 1960 et 2009)

■ Le reste des résultats de la prédiction annuelle dans la dimension spatiale et temporelle est donné en (Annexe II.2.1) et (Annexe II.2.2)

*Annexe (II.2.1 : Prédictions spatiale pas annuelle

*Annexe (II.2.2 : Prédictions temporelle pas annuelle

II.2.4.2. Prédiction au pas mensuel (temporelle et spatiale)

L'application du FK a été également effectuée sur les précipitations mensuelles des 39 stations pluviométriques, celles-ci sont observées du mois de Septembre 1959 au mois d'Aout 2009 présentant ainsi une période d'observation de 612 mois. Comme il a été mentionné plus haut, les 240 premiers mois (20 ans) ont servi à l'estimation du modèle le reste à sa validation.

L'optimalité des résultats ainsi obtenus est exprimée par l'allure de la courbe sur la Fig.II.2.5, celle-ci décrit le comportement décroissant des éléments de la diagonale majeure de la matrice de covariance qui représentent la variance de l'erreur de prédiction. Il est montré comment au début des calculs, cette dernière prend des valeurs importantes et comment après un certain nombre d'itérations, elle décroît rapidement et comment par la suite, elle continue sa décroissance d'une manière régulière pour atteindre des valeurs qui se rapprochent de zéro à la fin des calculs, tout en restant positive.

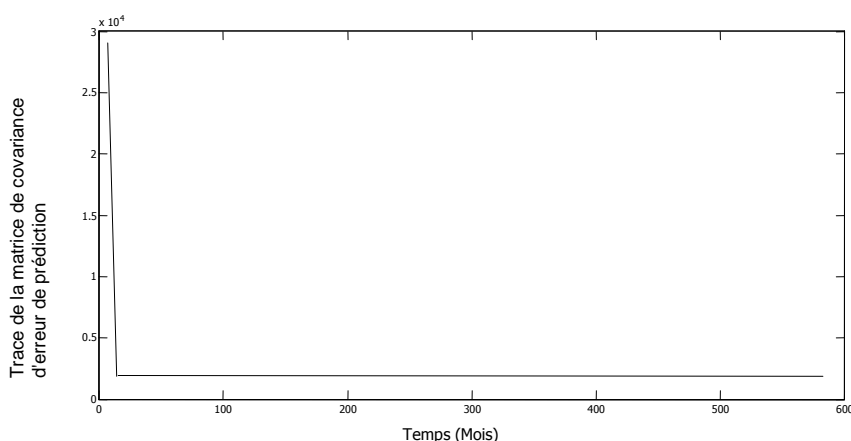


Fig II.2.5. Trace de la matrice de covariance d'erreur de prédiction

La matrice gain Fig.II.2.6 présente une allure similaire à celle de la covariance de l'erreur de prédiction ce qui montre que les paramètres du modèle ont atteint leurs valeurs optimales pendant la période d'estimation (240 premiers mois) et témoigne encore une fois de l'adéquation du modèle.

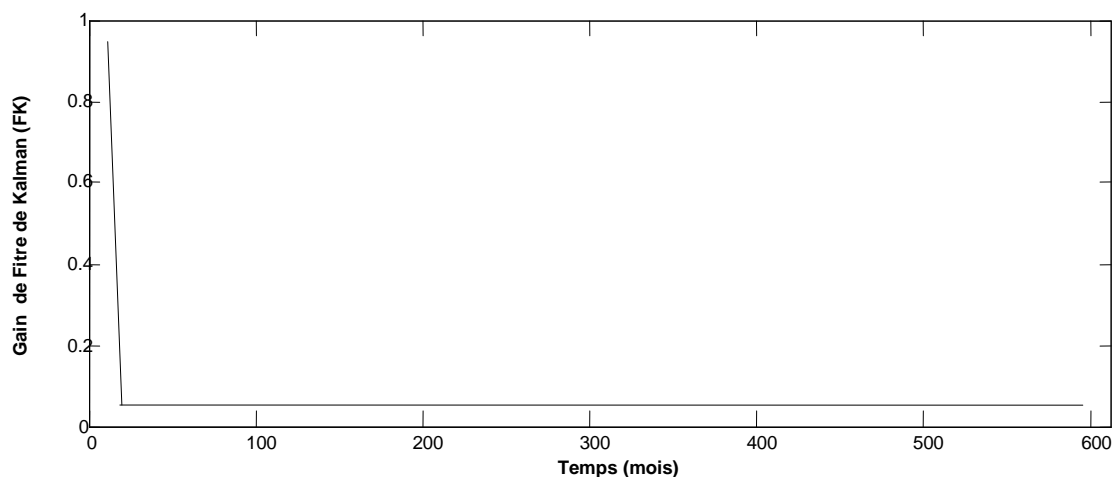


Fig II.2.6. Gain du filtre de Kalman au pas mensuel

La Fig.II.2.7 est un exemple des prédictions multi-sites obtenues au pas mensuel sur les 612 mois allant de septembre 1959 à octobre 2009, elle montre comment les observations et les prédictions à la station Zmalet-el-Amir AEK se suivent si étroitement sur la totalité du parcours que l'erreur relative exprimée en pourcent, varie seulement entre +5 et -5 %.

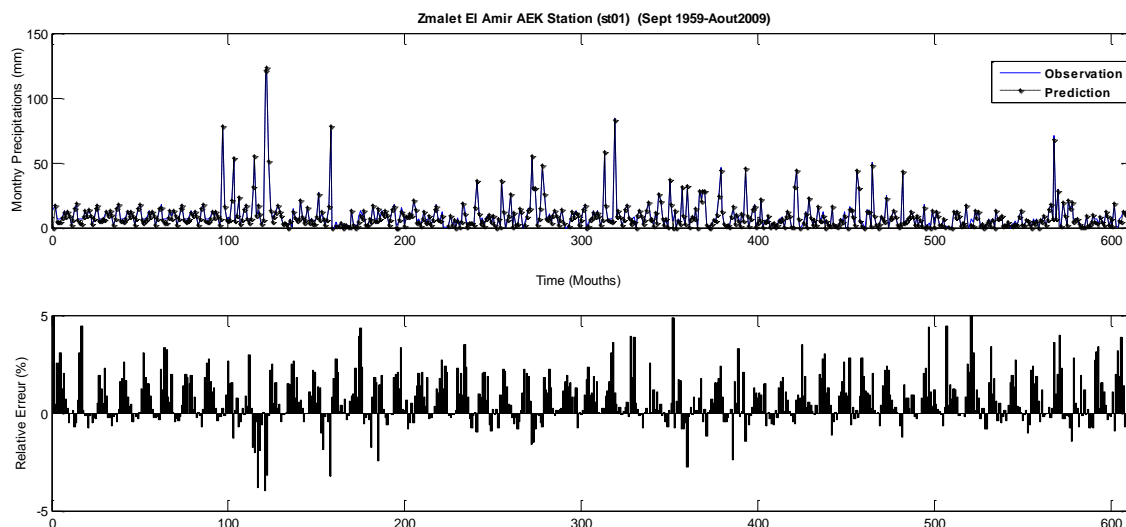


Fig II.2.1. Prédictions mensuelles obtenues à la station de Zmalet El Amir AEK (Sep 1959- Aout 2009)

La qualité des prédictions obtenues peut également être appréciée d'un point de vue spatial. La Fig.II.2.8 en est une preuve. Cette dernière présente les observations et les prédictions mensuelles obtenues aux 39 stations durant le mois d'octobre. Les différences entre les observations et les prédictions correspondantes sont tellement minimales que l'on n'arrive même pas à distinguer entre les deux courbes. Ceci est traduit en termes d'erreur relative en pourcent avec des valeurs qui varient entre +4% et -2% ce qui confirme une fois de plus, la performance du filtre au pas mensuel

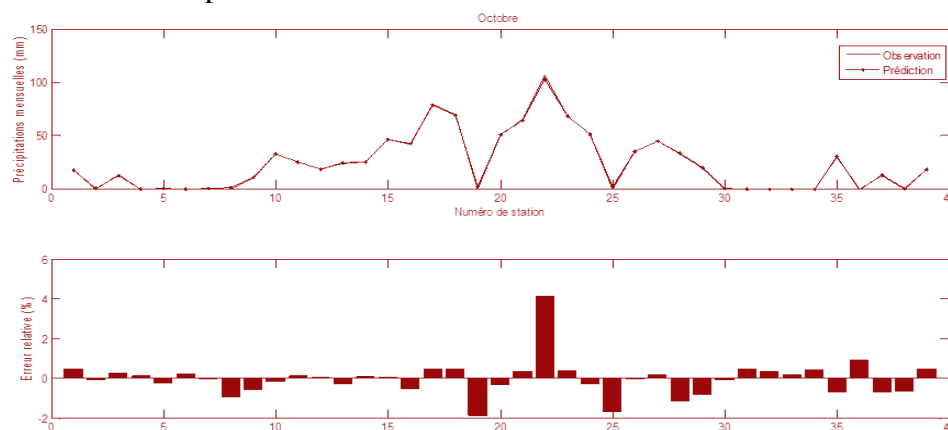


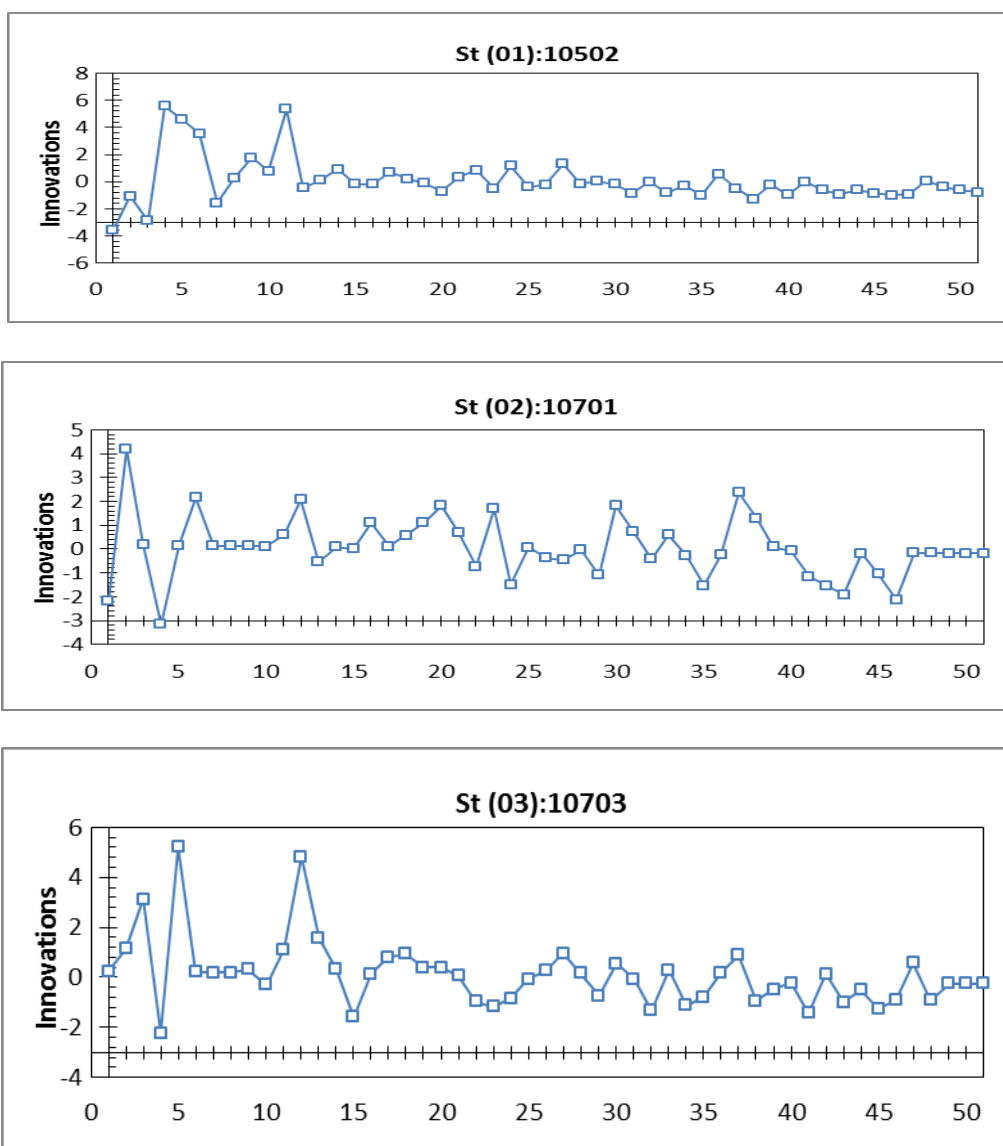
Fig II.2.2. Observations et prédictions des précipitations aux 39 stations pluviométriques pour le mois d'Octobre

■ Le reste des résultats de prédiction mensuelle (temporelle et spatiale) est donné en Annexe :(II.2.3) et (II.2.4)

*Annexe (II.2.3 : Prédictions spatiale pas Mensuel , *Annexe (II.2.4 : Prédictions temporelle pas Mensuel

II.2.4.3. Analyse des innovations du filtre (résidus)

Les erreurs de prédictions ou innovations du filtre pour les des **39** stations étudiées, comme c'est indiqué par la **Fig II.2.9**, présente des valeurs trop élevées durant les cinq premiers pas par rapport au reste des valeurs. Nous rappelons que 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. 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 a figure citée.



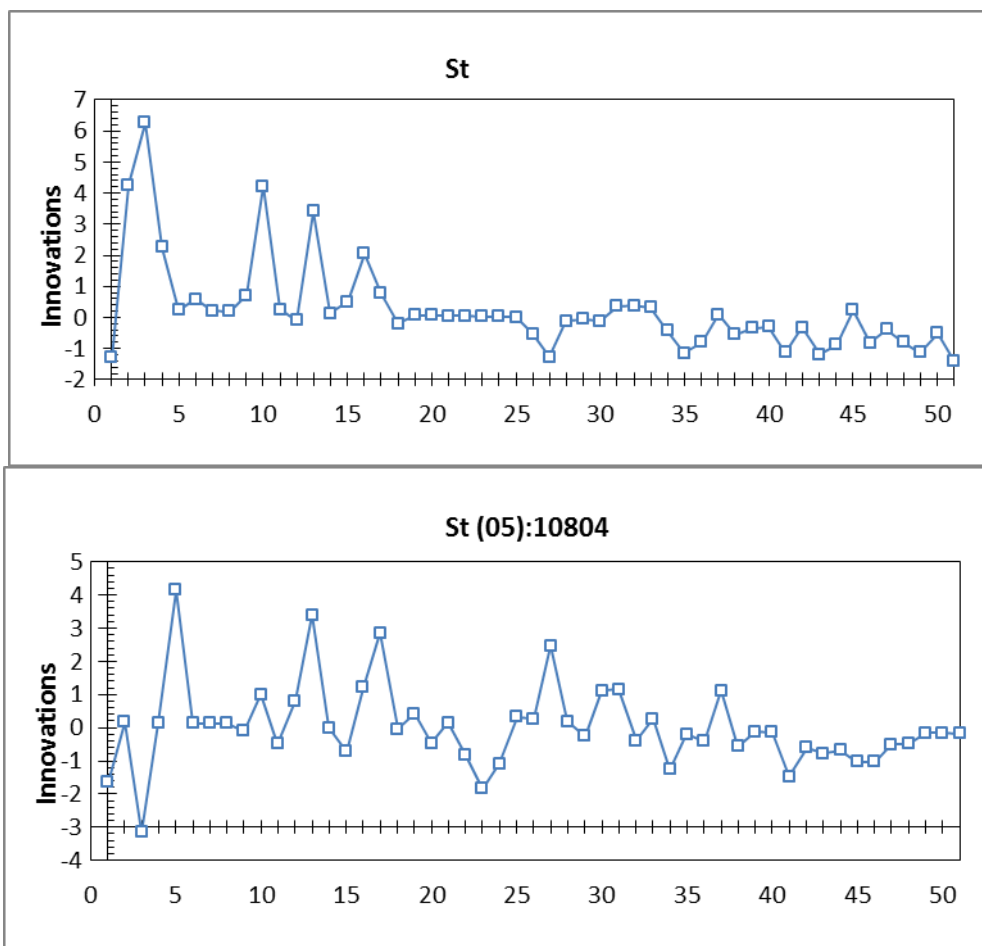


Fig II.2.9. Innovations du FK à Station (St10502 ; St 10701 ; St 10703 ; St 10704 ; St 10804)respectivement (1959-2009)

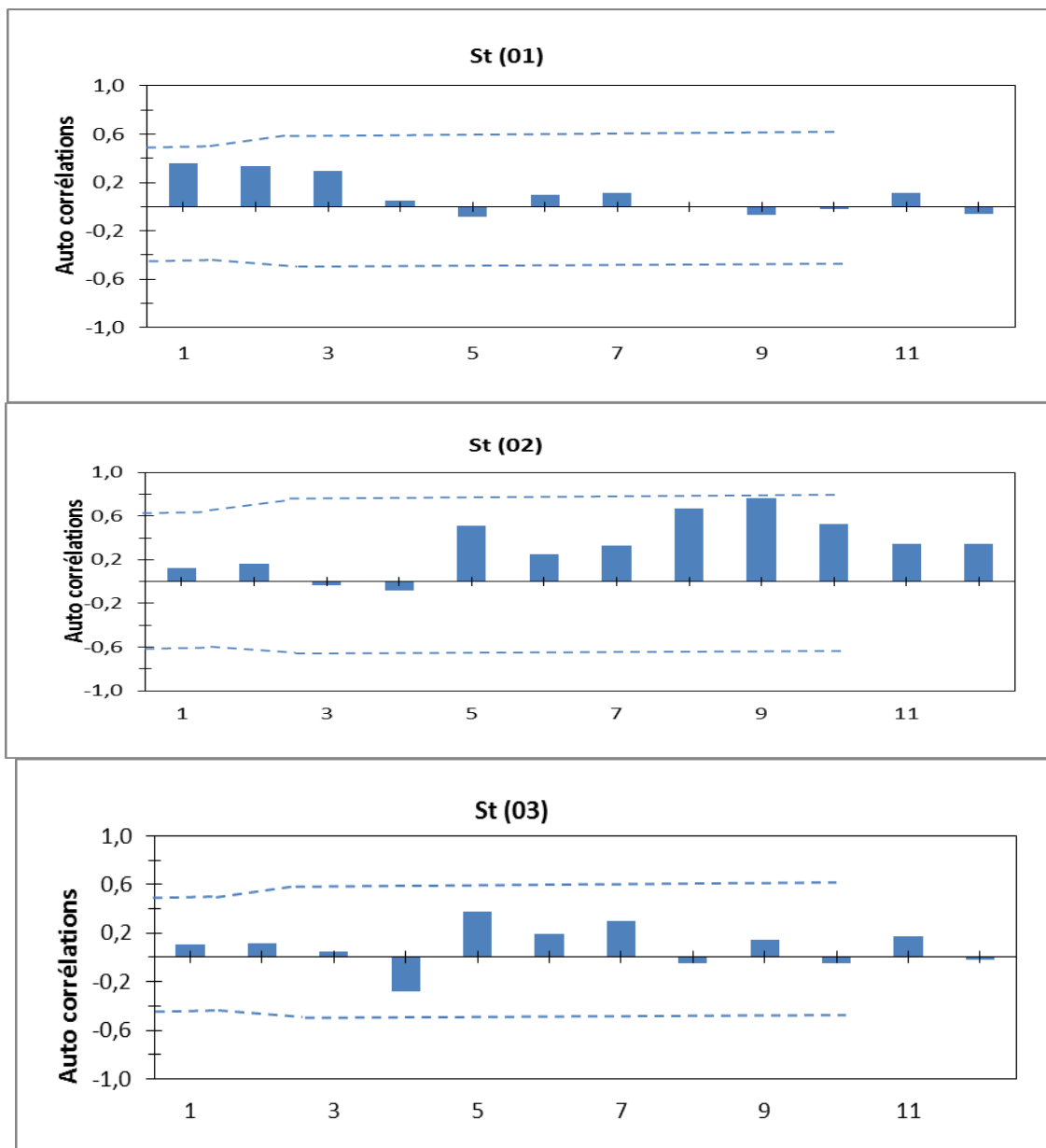
■ Le reste des résultats des innovations du FK est donné en (**Annexe II.2.5**)

L'analyse de toutes ces innovations est effectuée afin de vérifier les hypothèses du filtre de Kalman concernant l'indépendance et la Normalité. Ces analyses sont effectuées uniquement sur la dimension temporelle, car d'un point de vue nombre d'observations, elle est plus intéressante.

*Annexe (II.2.5 Innovations du FK à St (05) jusqu'à St(39) (1959-2009)

- L'indépendance

Pour les prédictions temporelles, les différences entre les Précipitations observés et leurs prédictions pour toutes les stations concernées ont été testées quant à l'indépendance. Ceci est fait moyennant l'examen visuel du tracé de la fonction d'auto-corrélation à 99% de confiance, et le test d'indépendance de Box-Pierce. La Fig II.2.10 montre bien le tracé de cette fonction d'auto- corrélation des innovations pour les 5 stations considérées,et le reste des résultats ,voir **Annexe (II.2.6(a))**.



*Annexe (II.2.6 (a) Figures des fonctions d'auto-corrélation des innovations [St (06) à St(39)]

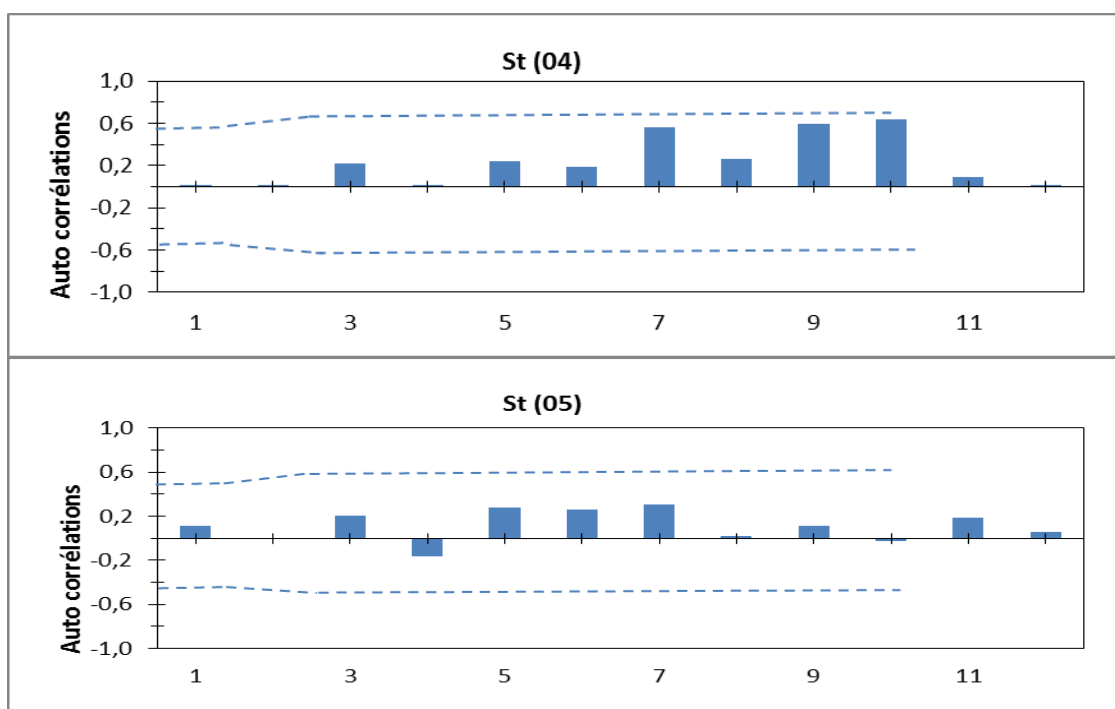


Fig II.2.10. Fonction d’auto-corrélation des innovations au Station (St10502 ; St 10701 ; St 10703 ; St 10704 ; St 10804) respectivement (1959-2009)

L’inspection visuelle de la figure montre que ces innovations ne présentent pas d’auto-corrélations significatives, ce qui affirme l’indépendance des innovations du filtre pour la totalité des stations considérées. Les valeurs des auto-corrélations sont représentées dans l’Annexe **Annexe (II.2.6(b))**.

Ces résultats montrent tous des variations faibles et non significatives des auto-corrélations qui tendent vers zéro, ce qui affirme l’indépendance des innovations du filtre, pour les stations pluviométriques en question. Les résultats du test de Box-Pierce tel que c’est indiqué par le Tableau suivant (Tableau II.2.1) qui est pour confirmer cette indépendance.

Le test est basé sur la somme des carrés des N premières auto-corrélations de la série de données et que si la probabilité P est supérieure ou égale à 0,10 l’on ne peut pas rejeter l’hypothèse que la série en question est aléatoire ou indépendante à 90% de confiance ou plus. C’est le cas de toutes les séries des innovations étudiées, d’où la confirmation de leur indépendance à 90% et plus de confiance.

*Annexe (II.2.6 (b) Tableau des fonction s d’auto-corrélation des innovations [St (06) à St(39)]

Tableau II.2. 1. Test de Box-Pierce pour les innovations mensuel du filtre.

Code station	Nom de la station	N	Statistique du test	P
St 01	10502 Zmalet el amir aek	11	18,66	0,62
St 02	10701 Ain baadj	11	15,89	0,53
St 03	10703 Rechaiga	11	19,55	0,65
St 04	10704 Ksar chellala	11	23,68	0,79
St 05	10804 Mehdiya pepinière	11	20,68	0,69
St 06	10904 Dahmouni trumulet	11	14,53	0,48
St 07	10905 Ain sebain	11	17,38	0,58
St 08	11004 khemisti	11	11,12	0,37
St 09	11006 Tissemsilt	11	15,97	0,53
St 10	11104 Ain boucif	11	18,19	0,61
St 11	11206 Chahbounia	11	9,44	0,31
St 12	11207 Kef lasfar	11	20,2	0,67
St 13	11208 Boughzoul	11	15,6	0,52
St 14	11301 Ksar el boukhari	11	19,73	0,66
St 15	11302 Derrag	11	12,9	0,43
St 16	11404 Zoubiria mongorno	11	14,7	0,49
St 17	11405 Ghrib bge	11	16,12	0,54
St 18	11605 Teniet - el-had	11	15,51	0,52
St 19	11606 Sidi mokrefi	11	19,31	0,64
St 20	11607 Tarik ibn ziad	11	13,05	0,44
St 21	11703 Bordj el amir khaled	11	17,87	0,60
St 22	11803 Sidi medjahed	11	9,45	0,32
St 23	11904 Rouina mairie	11	16,56	0,55
St 24	12108 Fodda bge	11	14,88	0,50
St 25	12205 Benaria	11	12,99	0,43
St 26	12304 Souk-el-had	11	10,05	0,34
St 27	12306 Bordj bounaama	11	16,03	0,53
St 28	12309 Oued- sly	11	16,66	0,56
St 29	12503 Sidi hosni	11	11,56	0,39
St 30	12605 Ammi moussa	11	9,77	0,33
St 31	12706 Mendes	11	16,02	0,53
St 32	12908 Tiaret GRHA	11	13,54	0,45
St 33	12915 Louhou	11	15,44	0,51
St 34	13004 Ain el haddid	11	11,11	0,37
St 35	13202 Rosfa	11	12,35	0,41
St 36	13503 Oued djemaa	11	7,45	0,25
St 37	13504 Saline oued djemaa	11	6,77	0,23
St 38	13506 El matmar DRS	11	12,12	0,40
St 39	13507 L' Hillil	11	15,14	0,50

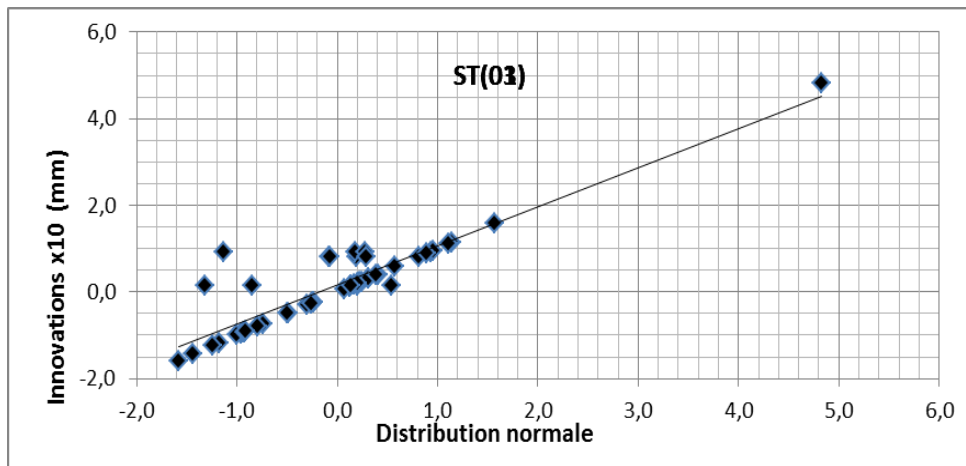
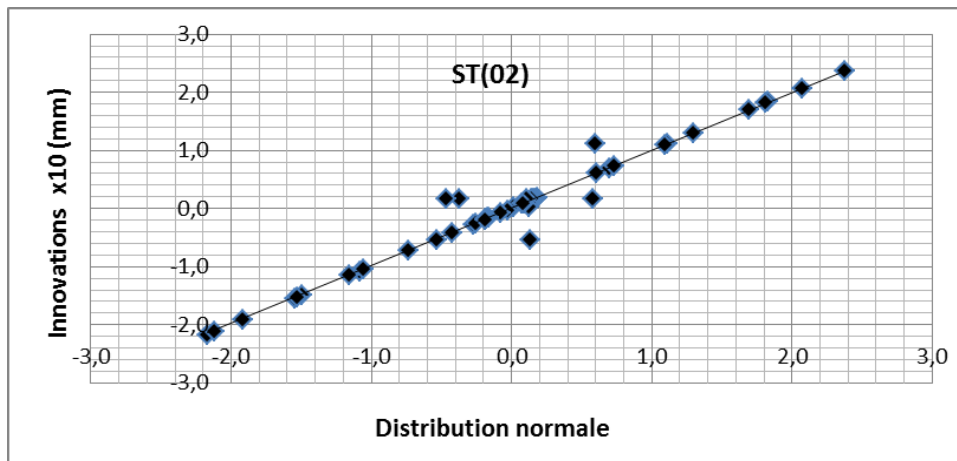
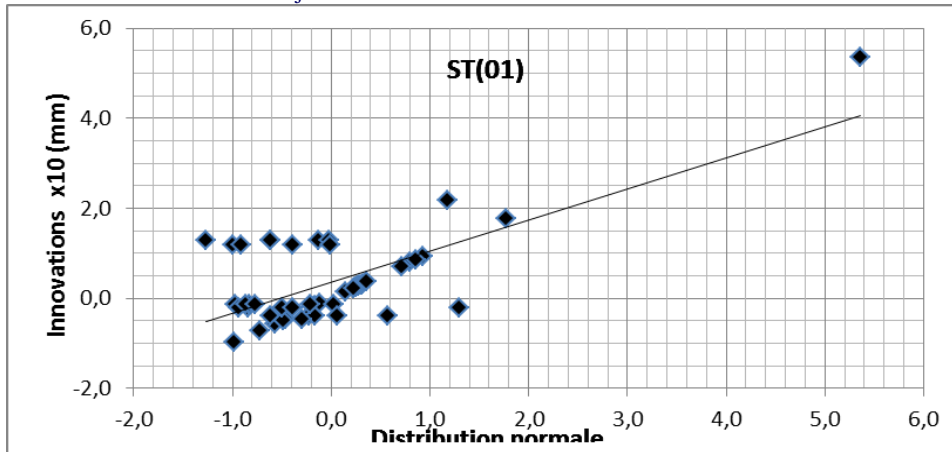
N: nombre des premières auto-corrélations;
P: probabilité.

• La Normalité

La Normalité des innovations (Anderson, 1942) a été testée dans la dimension temporelle, le tracé de ces ajustements Normaux aux innovations des différentes stations considérées de 1959 à 2009 a été effectué à 95% de confiance et la Fig II.2.13 montre la qualité de l'adéquation de ces ajustements pour les différentes stations.

Il ya quelques stations qui s'ajustement à la loi normale ne semble pas être adéquat. L'une des raisons les plus possibles de ce défaut de normalité pourrait être les premières valeurs qui sont dues à l'adaptation du filtre et qui sont inévitablement mauvaises au début des calculs.

Ajustement Normal aux innovations de 39 stations



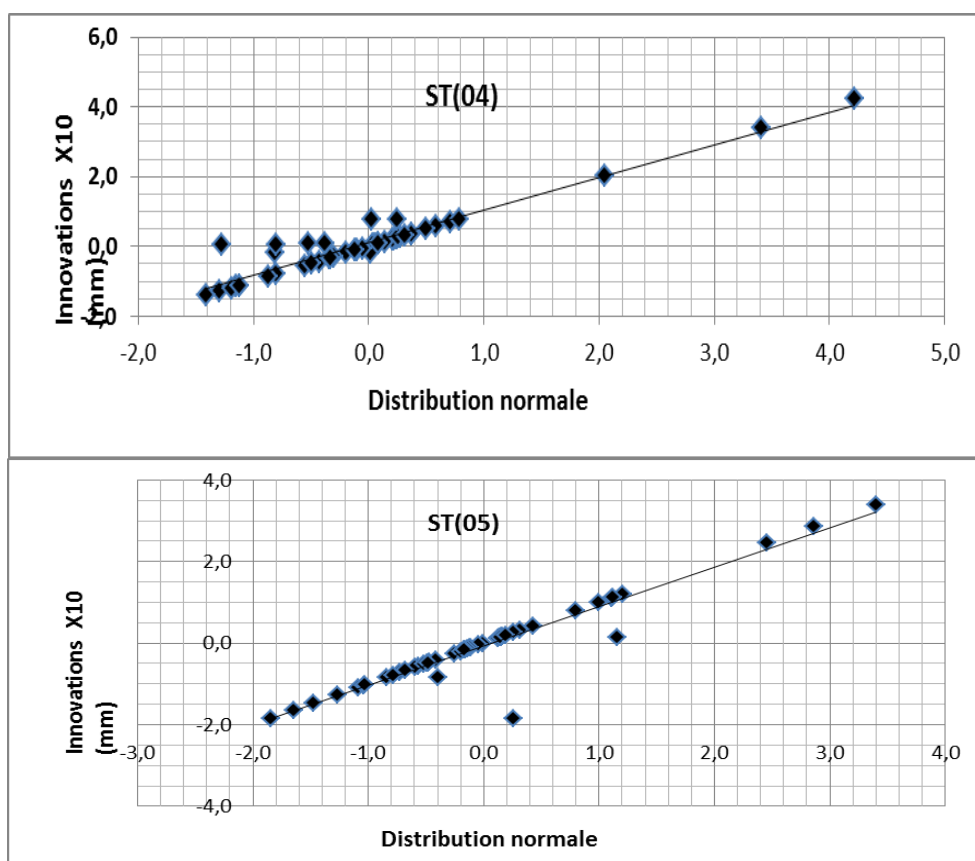


Fig II.2.13. Ajustement Normal aux innovations de (St10502 ; St 10701 ; St 10703 ; St 10704 ; St 10804) respectivement (1959-2009)

■ Pour le reste des resultats : Annexe (II.2.7)

La qualité de l'adéquation de ces ajustements Normaux aux séries des innovations a été vérifiée par le test du Chi-deux à 95 % de confiance dont les résultats sont présentés dans le Tableau II.2.2.

Ces derniers montrent effectivement qu'à l'exception de quelques stations comme :

St 01	10502	Zmalet el amir aek
St 04	10704	Ksar chellala
St 08	11004	Khemisti
St 11	11206	Chahbounia
St 18	11605	Teniet - el-had
St 20	11607	Tarik ibn ziad
St 22	11803	Sidi medjahed
St 26	12304	Souk-el-had
St 29	12503	Sidi hosni
St 30	12605	Ammi moussa
St 34	13004	Ain el haddid
St 36	13503	Oued djemaa
St 37	13504	Saline oued djemaa

affichent une valeur de probabilité P inférieure à 0,05 toutes les autres stations présentent des innovations gaussiennes à 95 % de confiance.

*Annexe (II.2.7 Ajustement Normal aux innovations de 39 stations

Tableau II.2. 2. Test d'adéquation du Chi-deux de l'ajustement normal à 95% de confiance des innovations du filtre.

Code station		Nom de la station	μ	δ	Statistique du test	P
St 01	10502	Zmalet el amir aek	-5,59	8,76	22,1	0,024
St 01	10502	Zmalet el amir aek *	-7,6	3,98	10,35	0,114
St 02	10701	Ain baadj	-0,72	4,19	16,87	0,444
St 03	10703	Rechaiga	-0,33	2,42	15,22	0,401
St 04	10704	Ksar chellala	-16,6	10,08	20,43	0,030
St 04	10704	Ksar chellala *	0,6	4,64	9,515	0,130
St 05	10804	Mehdia pepinière	-0,27	2,55	15,08	0,397
St 06	10904	Dahmouni trumulet	-5,4	4,67	16,73	0,440
St 07	10905	Ain sebain	-1,92	5,82	18,38	0,484
St 08	11004	khemisti	-21,6	0,2	21,77	0,026
St 08	11004	khemisti *	-29,1	0,1	10,185	0,426
St 09	11006	Tissemsilt	-2,5	0,7	14,42	0,379
St 10	11104	Ain boucif	-5,3	2	16,07	0,423
St 11	11206	Chahbounia	-11,7	0,9	20,09	0,002
St 11	11206	Chahbounia *	0,868	0,45	9,345	0,660
St 12	11207	Kef lasfar	-4,5	1,1	12,74	0,335
St 13	11208	Boughzoul	-6,2	0,9	14,39	0,379
St 14	11301	Ksar el boukhari	-11,0	1,4	16,04	0,422
St 15	11302	Derrag	-32,4	1,5	17,69	0,466
St 16	11404	Zoubiria mongorno	-30,9	1,7	19,34	0,509
St 18	11605	Teniet - el-had	-37,0	12,9	24,64	0,040
St 18	11605	Teniet - el-had *	0,6	6,05	11,62	0,160
St 19	11606	Sidi mokrefi	-13,7	1,5	26,52	0,884
St 20	11607	Tarik ibn ziad	-12,7	22,1	23,7	0,013
St 20	11607	Tarik ibn ziad *	-14,7	10,65	11,15	0,213
St 21	11703	Bordj el amir khaled	-8,0	2	16,58	0,553
St 22	11803	Sidi medjahed	-31,6	20,4	20,1	0,020
St 22	11803	Sidi medjahed *	-33,6	9,8	9,35	0,120
St 23	11904	Rouina mairie	-5,6	1,3	12,98	0,433
St 24	12108	Fodda bge	-11,8	1,97	14,86	0,495
St 25	12205	Benaria	-23,6	2,06	16,74	0,558
St 26	12304	Souk-el-had	-31,5	22,4	20,7	0,020
St 26	12304	Souk-el-had *	-33,5	10,8	9,65	0,210
St 27	12306	Bordj bounaama	-27,1	2,58	13,58	0,453
St 29	12503	Sidi hosni	-9,81	0,048	22,21	0,020
St 30	12605	Ammi moussa	-13,8	0,74	20,42	0,020
St 29	12503	Sidi hosni *	0,758	0,024	10,405	0,310
St 30	12605	Ammi moussa *	0,768	0,37	9,51	0,140
St 31	12706	Mendes	-8,62	4,24	17,67	0,589
St 32	12908	Tiaret GRHA	-9,49	3,81	15,19	0,506
St 33	12915	Louhou	-7,57	3,84	17,09	0,570
St 34	13004	Ain el haddid	-10,8	0,04	21,76	0,020
St 34	13004	Ain el haddid *	-12,8	0,02	10,18	0,210
St 35	13202	Rosfa	-2,69	1,44	14	0,467
St 36	13503	Oued djemaa	-9,19	21,11	23,1	0,020
St 37	13504	Saline oued djemaa	-1,32	35,67	25,42	0,020
St 36	13503	Oued djemaa *	0,378	10,155	10,85	0,110
St 37	13504	Saline oued djemaa *	0,248	17,435	12,01	0,320
St 38	13506	El matmar DRS	-3,01	1,73	13,77	0,459
St 39	13507	L' Hillil	-3,53	1,23	16,79	0,560

μ : Moyenne de la loi normale ajustée;
 (*) : tronquée des 5 premières valeurs.
 δ : écart-type de la loi normale
 P: probabilité

Dans le but de déterminer la cause de ce défaut de normalité afin de le corriger, nous avons procédé à l'ajustement normal de la série des innovations de Station (St1) et (St4) en écartant les 5 premières valeurs aberrantes.

Pour cela la Fig.II.2. 14 confirme bien notre thèse en montrant la qualité de cet ajustement aux nouvelles séries d'innovations et le résultat du test du Chi-deux vient à son tour confirmer l'adéquation de ce dernier ce qui nous permet de confirmer la normalité des innovations pour l'ensemble des stations étudiées.

■ Pour le reste des résultats : Annexe (II.2.8)

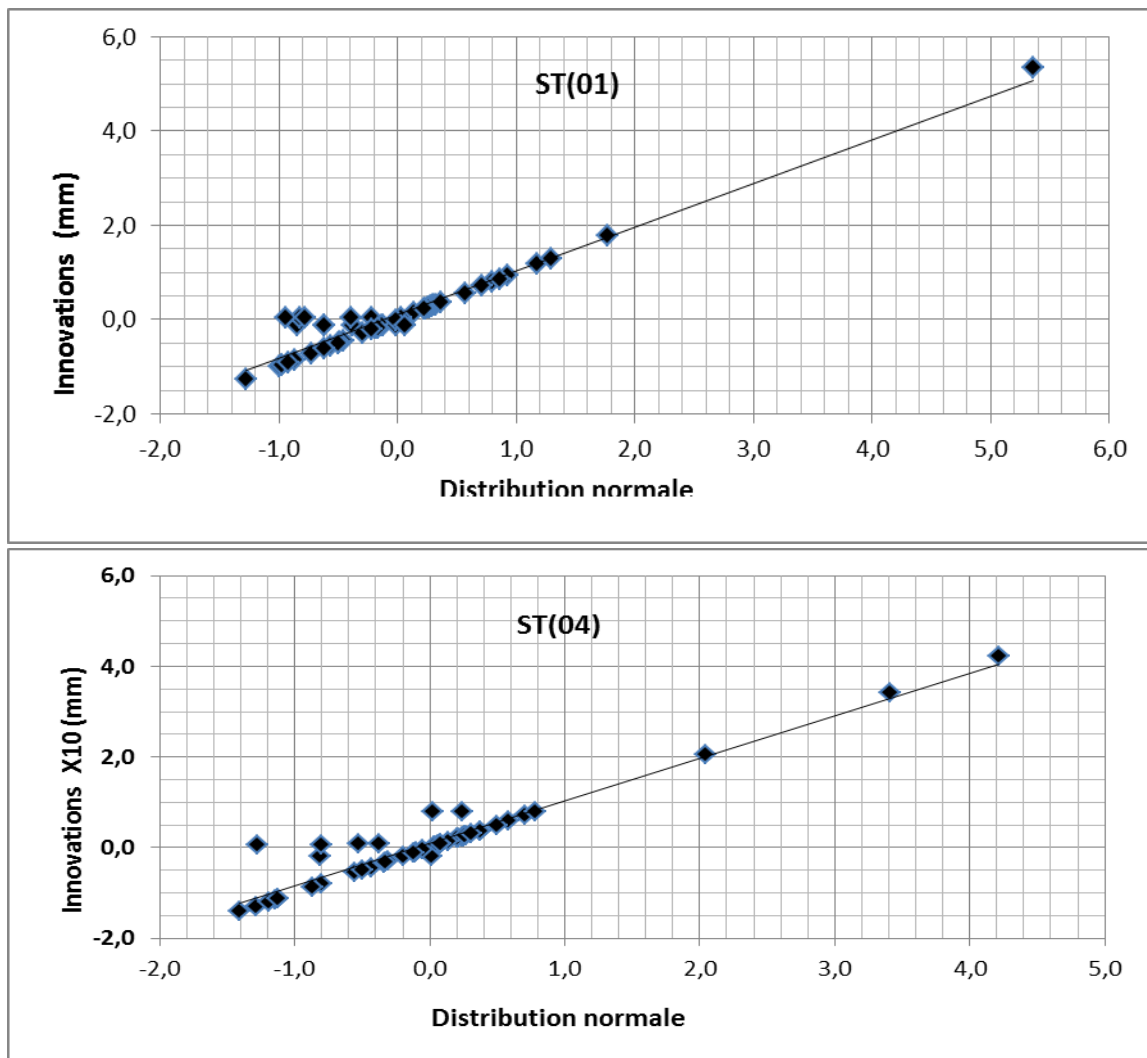


Fig II.2.14. Ajustement Normal aux innovations de (St10502 ; St 10704) tronquées des 5 premières valeurs

*(Annexe II.2.8 Ajustement Normal aux innovations des stations [St 01;St 04; St 08; St 11; St 18;St 20; St 22; St 26;St 29;St 30;St 34;St 36 etSt 37] tronquées des 5 premières valeurs

II.2.5. MOYENNE ET DEVIATION STANDARD DES PREDICTIONS MULTI-SITES

II.2.5.1. Prédications multi-sites « Pas annuelle » :

Le Tableau II.2.4 présente certaines caractéristiques statistiques moyennes des observations et des prédictions obtenues aux 39 postes pluviométriques sur la période d'étude. C'est un résumé d'information qui nous permet de comparer les prédictions obtenues par rapport aux observations enregistrées, en termes de certains paramètres statistiques de localisation centrale (moyenne) et de dispersion autour de cette position centrale (écart-type). Ces paramètres sont des moyennes calculées pour la dimension temporelle et la dimension spatiale. Il est montré que la moyenne des prédictions pour la dimension temporelle est de 22,95 mm contre une valeur de 22,37 mm pour les observations, avec une erreur relative de -4,22% et que l'écart-type moyen est de 18,6 mm contre une valeur de 18,16 mm pour les observations, avec une erreur relative de -4,04%. Le même constat est fait pour la dimension spatiale avec une moyenne des prédictions de 24,24 mm contre une valeur de 22,6 mm pour les observations, avec une erreur relative de -7,28% et un écart-type moyen des prédictions de 22,36 mm contre une valeur de 21,46 mm pour les observations, avec une erreur relative de -11,23%.

Tableau II.2.3. Certaines caractéristiques statistiques moyennes des observations et des prédictions annuelles obtenues aux 39 stations (1959-2009)

	Observation		Prédiction		Erreur relative (%)	
	Moyenne	Ecart-type	Moyenne	Ecart-type	Moyenne	Ecart-type
Prédictions temporelles	22,37	18,16	22,95	18,60	-4,22	-4,04
Prédictions spatiales	22,60	21,46	24,24	22,36	-7,28	-11,23

La Fig.II.2.15 retrace les variations de l'écart-type moyen des 39 stations au cours de la période d'étude (1959-2009) et ce pour les observations et les prédictions annuelles obtenues, elle montre des valeurs qui sont constamment supérieures pour les prédictions. Ce résultat exprime une éventuelle tendance du filtre de Kalman à la surestimation des estimations qu'il fournit.

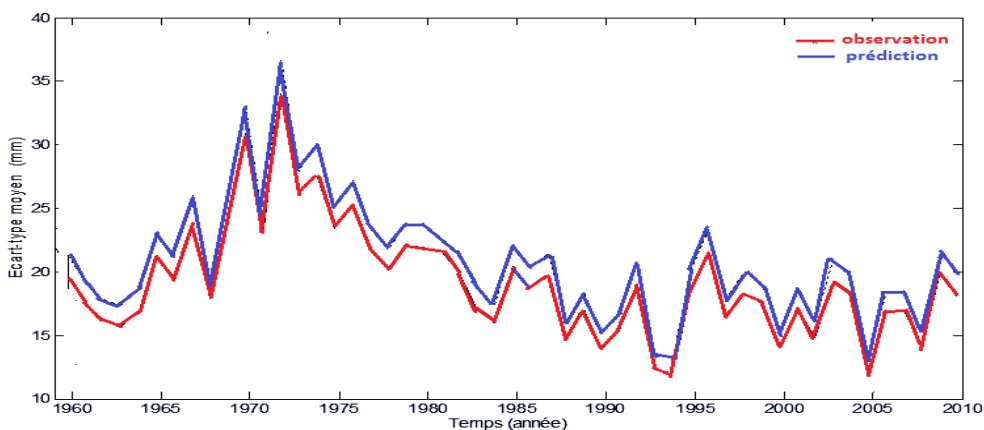


Fig II.2.15. Ecart-type moyen des observations et des prédictions annuelles aux 39 stations (1959-2009)

II.2.5.2. Prédications multi-sites « Pas mensuelle » :

La tendance du modèle développé au cours de ce travail à la surestimation est également confirmée pour les précipitations mensuelles. Le Tableau II.2.2 met en évidence cette tendance en affichant des valeurs moyennes de la moyenne et de l'écart type plus élevées pour les prédictions que pour les observations. Ceci est valable aussi bien du point de vue temporel que spatial. La Fig.II.2.16 le confirme.

Tableau II.2. 4. Certaines caractéristiques statistiques moyennes des observations et des prédictions mensuelles obtenues aux 39 stations (Sep.1959 – Aout. 2009)

	Observation		Prédiction		Erreur relative (%)	
	Moyenne	Ecart-type	Moyenne	Ecart-type	Moyenne	Ecart-type
Temporel	14,62	12,90	15,22	13,03	-3,37	-1,41
Spatial	22,51	10,50	23,62	10,81	-4,95	-2,30

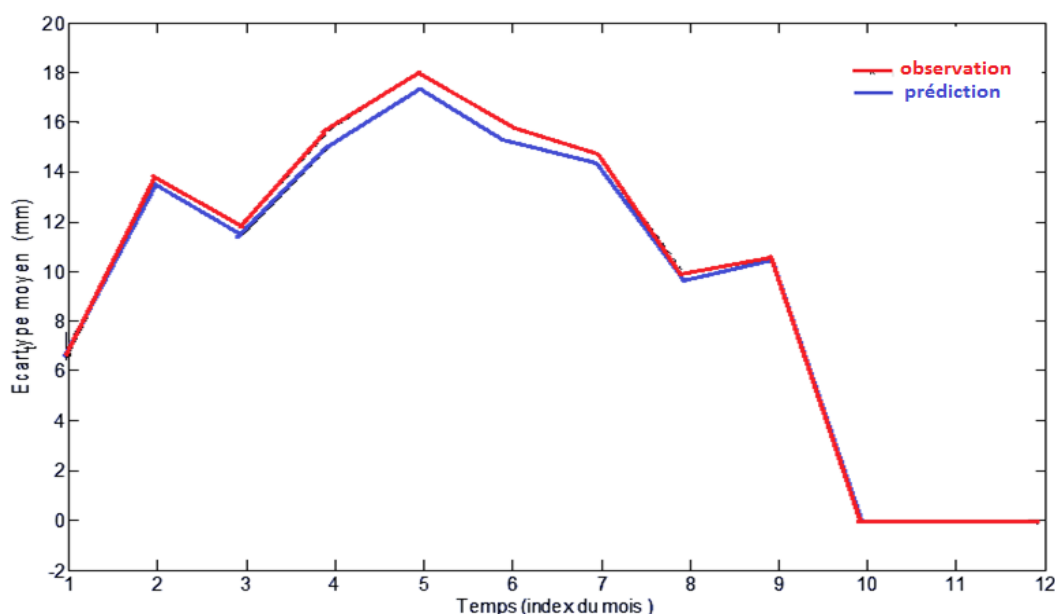


Fig II.2.16. Ecart-type moyen des observations et des prédictions mensuelles aux 39 stations (Sep.1959 - Aout.2009)

Les résultats montrent que les prédictions multi-sites obtenues par le FK suivent de plus près les observations enregistrées avec une erreur relative moyenne en valeur absolue de l'ordre de 7,28 % pour le pas annuel et de 4,95 % pour le pas mensuel. Ces erreurs de prédiction sont minimales, elles sont dans tous les cas largement inférieures à 10% ce qui est très acceptable. Ceci prouve la performance du FK dans la modélisation et la prédiction des précipitations au bassin versant de Chélif, même en présence de saisonnalité et malgré la variabilité géographique et l'indépendance temporelle des séquences pluviométriques.

Ces résultats expriment clairement que les déviations standards des valeurs observées sont plus grandes que celles des valeurs prédites, ce qui pourrait traduire une éventuelle tendance du (FK) multi-site à une sur-estimation.

Conclusion

Au cours de ce chapitre, nous avons développé un modèle de type filtre de Kalman pour la prédiction multi-site des précipitations. A cet effet, les séries chronologiques des précipitations annuelles et mensuelles, couvrant une période d'observation commune de 51 ans (1959 à 2009) ont été investies. Ces dernières sont observées à 39 stations pluviométriques dans le bassin versant de Cheliff, au Nord d'Algérie, et Les prédictions annuelles et mensuelles obtenues dans la dimension temporelle tout comme la dimension spatiale, sont très proches des valeurs observées sur la même période, ce qui indique que le filtre de Kalman multi-site est un outil efficace pour la modélisation et la prédiction des précipitations dans le Chélif.

La performance du modèle développé a été mise en exergue par le calcul du pourcentage de l'erreur relative des prédictions multi-sites ; Ce pourcentage moyen sur toute la période d'observation est nettement inférieure à 10% ce qui est très acceptable.

Une éventuelle tendance du modèle à la surestimation a également été soulignée par des moyennes et des écart-types plus importants pour les prédictions que pour les observations.

Le modèle développé présente l'avantage considérable de prendre en compte non seulement la nature stochastique des précipitations mais aussi leurs variations temporelles et spatiales,

En fin, nous pouvons dire que le filtre de Kalman, qui est une technique basée sur le concept des moindres carrés, possède une propriété très importante à savoir l'optimisation séquentielle, qui veut dire que le modèle est mis à jour d'une manière adaptative au fur et à mesure que les sorties du système hydrologique (mesures) arrivent. L'un des avantages de cette technique est que la stationnarité n'est pas pré requise, comme c'est le cas dans la plupart des modèles en hydrologie, ceci permet des changements dans les paramètres du modèle et dans les variances qui sont une manière de rendre compte de la non linéarité de la réponse du système hydrologique concerné. Un autre avantage de la technique est que l'application est faite dans le domaine temporel, cette caractéristique joue un rôle important dans la prévision en temps réel des séries chronologiques en hydrologie. En plus, l'algorithme du filtre de Kalman peut être amorcé avec un minimum d'information objective disponible, pour s'ajuster automatiquement, au fur et à mesure que les nouvelles données arrivent. Ces caractéristiques font du FK l'un des outils les plus appropriés et donc le plus courtisé, particulièrement dans les situations où tout change et, en plus les résultats qu'il fournit sont optimaux.

CONCLUSION GENERALE

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

La première équation de ce modèle exprime l'état du système hydrologique concerné, dans notre cas, c'est un vecteur donnant les précipitations annuelles (respectivement mensuelles) à plusieurs sites à la fois.

La deuxième 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 de précipitations annuelles (respectivement mensuelles) aux sites précédents.

Le développement de ce modèle a pour objectif, l'application de la technique du filtre de Kalman (FK) à la prédiction multi-site des précipitations annuelles et mensuelles et l'amélioration des erreurs de prédiction.

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 précipitations annuelles (respectivement mensuelles) 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 chronologiques des précipitations annuelles et mensuelles, observées aux 39 stations pluviométriques dans le bassin versant de Cheliff et couvrant une période d'observation commune de 51 ans allant de 1959 à 2009.

Les prédictions multi-sites des précipitations annuelles et mensuelles, obtenues dans la dimension temporelle pour chacune des 39 stations pluviométriques en question, sont très proches des valeurs observées aux 39 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 des précipitations annuelles et mensuelles.

En fin, la justesse du modèle du filtre de Kalman ou son aptitude à représenter le processus naturel des précipitations annuelles et mensuelles, a été vérifiée par l'étude de l'indépendance des innovations du filtre d'une part, par le biais de leur fonction d'auto-corrélation et d'autre, par l'ajustement Normal de ces innovations.

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 est inférieure à 10% ce qui est très acceptable.

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

A la lumière de tous les résultats précédents, il est clair que la technique du FK multi-site est efficace pour la modélisation des précipitations annuelles et mensuelles, 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.

■ *Références bibliographiques :*

- [1] Agence du Bassin Hydrographique Cheliff Zahrez (2004) (1), « Cadastre hydraulique du bassin hydrographique du Cheliff Aval du barrage de Boughzoul, Haut et Moyen Cheliff ».
- [2] Agence du Bassin Hydrographique Cheliff Zahrez (2004) (2). « Cadastre hydraulique du bassin hydrographique du Cheliff aval du barrage de Boughzoul (Bas Cheliff)
- [3] Agence du Bassin Hydrographique Cheliff Zahrez (Mars 2010) « Rapport de synthèse abrégé du plan directeur d'aménagement des ressources en eaux»
- [4] Adlouni, Salaheddine; Ouarda, Taha B.M.J.; Bobée, Bernard, Estimation des paramètres des distributions statistiques par les L-moments Rapport de recherche (R682). INRS-Eau, Terre et Environnement, Québec,2003.
- [5] Anderson, Time Series Analysis and Forecasting: The BOX JENKINS Approach, Butterworth, London. (1976)
- [6] Bertino, G. Evensen, H. Wackernagel, Combining geostatistics and Kalman filtering for data assimilation in an estuarine system, *Inverse Prob.* 18 (2002) 1–23.
- [7] Boucefiane, « Cartographie des précipitations du bassin hydrographique du Cheliff Zahrez », mémoire de magister .Centre Universitaire de Khmis miliana , 2006,p 110
- [8] Boudevillain and Hervé Andrieu : Assessment of Vertically-Integrated Liquid (VIL) Water Content Radar Measurement. *J. Atmos. Oceanic Tech.*, 20, N.6 , 2003 pp 807-819.
- [9] 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.
- [10] Boukharouba, Annual stream flow simulation by ARMA processes and prediction by Kalman filter,
- [11] Box, P.M. Jenkins, ,1970: Times Series Analysis: Forecasting andControl, San Francisco, CA: Holden-Day. Holden
- [12] Chiu, Applications of Kalman filtering theory and technique to hydrology, hydraulics and water resources, in: Proceedings of the American Geophysical Union Chapman Conference on Applications of Kalman Filtering Theory and Technique to Hydrology,Hydraulics and Water Resources, University of Pittsburgh, Pittsburgh, PA, 1978, p. 738.
- [13] Didier AUROUX 2003, Etude de différentes méthodes d'assimilation de données pour l'environnement l'Université de Nice Sophia Antipolis, Spécialité : MATHÉMATIQUES
- [14] Durbin, J. Koopman, Time Series Analysis by State Space Methods, Oxford University Press, Oxford, 2001.
- [15] Jazwinski, Stochastic Processes and Filtering Theory,second ed, Academic Press, New York, NY, 1970.
- [16] Harvey A. C: Forecasting, Structural Time Series Models and theKalman Filter, Cambridge University Press , 1989
- [17] Hisdal, L.M. Tallaksen, Estimation of regional meteorological and hydrological drought characteristics: A case study for Denmark, *J. Hydrol.* 281 (2003) 230–247.

- [18] Haddad, A. Rahla, Investing in Algeria, Ministry for Participation and Investment Promotion, MDPPI, Algiers, 2004.
- [19] Hartnack, H. Madsen, Data assimilation in river flow modelling, in: Proceedings of the Fourth DHI Software Conference, June 6–8, 2001, Scanticon Conference Centre, Helsingør, Denmark.
- [20] Harrison, C. Stevens, Bayesian forecasting (with discussion), *J. R. Stat. Soc.* 38 (1976) 205–247.
- [21] Husain, Kalman filter estimation model in flood forecasting, *Adv. Water Resour.* 8 (1985) 15–21.
- [22] Garcia-ruiz j.m., lopez-moreno j.i., vicente-serrano s.m. lasanta- martinez t. Et begueria -mediterranean water resources in a global change scenario. *Earth-science and reviews*, vol. 105, n° 3-4, p. (2011) 121-139.
- [23] Gourrieroux et A. Montfort: *Séries temporelles et modèles Dynamiques*, 1990
- [24] Gelb, *Applied Optimal Estimation*, MIT Press, Cambridge, MA, 1974.
- [25] Georgakakos, G.F. Smith, On improved hydrologic forecasting—Results from a WMO real-time forecasting experiment, *J. Hydrol.* 114 (1990) 17–45.
- [26] Kalman, R.S. Bucy, New results in linear filtering and prediction theory, *J. Basic Eng.* 83 (1961) 95–108.
- [27] Kalman, A new approach to linear filtering and prediction problems, *J. Basic Eng.* 82 (1960) 35–45.
- [28] Kettab, A. Mitiche, R. and Bennaçar, N. « De l’eau pour un développement durable » enjeux et stratégies.’’, *Revue des Sciences de l’Eau*, N°spécial 20eme anniversaire, (2008) 137-146.
- [29] Laborde, (2003). « Hydrologie de surface».
- [30] Mate, Ministry of Planning and Environment, Development of the National Strategy and Action Plan on Climate Change, Initial National Communication, Algiers, 2001.
- [31] Meddi, (1995): Etablissement d’un modèle expliquant la repartition spatiale des pluies annuelles et mensuelles en fonction de la morphométrie. *Revue DIL INFOEAU*, trimestriel international de liaison Eau environnement N°05, 1995, pp9-21.
- [32] McLaughlin, An integrated approach to hydrologic data assimilation: Interpolation, smoothing, and filtering, *Adv. Water Resour.* 25 (2002) 1275–1286.
- [33] Moradkhani, S. Sorooshian, H.V. Gupta, P.R. Houser, Dual state-parameter estimation of hydrological models using ensemble Kalman filter, *Adv. Water Resour.* 28 (2005) 135–147.
- [34] Ouachani, T. Bargaoui, R. Ouarda, Integration of a Kalman filter in the HBV hydrological model for predicting, *Hydrol. Sci. J.* 52 (2007) 318–337.
- [35] Plumejeaud : *Modèles et méthodes pour l’information spatio-temporelle évolutive*. thèse de Doctorat. Université de Grenoble. France. 2006, p 309.
- [36] Rouane H. (2011) “Contribution à la mise en évidence de l’expérience de l’Algérie dans la gestion intégrée des ressources en eau ”, Euro Mediterranean Scientific Congress on Engineering, Algeciras, Spain.
- [37] Pierre Gagnon, Estimation et interpolation. Note de cours (SCG-65831) département des sciences géomatiques, Université Laval. (1999)

- [38] Schlee, C.J. Standish, N.F. Toda, Divergence in the Kalman filter, AIAA J. 5 (1967) 1114–1120.
- [39] Sorenson, Kalman Filtering Theory and Applications, IEEE Press, New York, NY, 1985, pp. 90–126.
- [40] Schreider, P.C. Young, A.J. Jakeman, An application of the Kalman filtering technique for streamflow forecasting in the upper murray basin, Math. Comput. Modell. 33 (2001) 733–743.
- [41] Şen, Z., Latif, A.M. (2002). "A Kalman filtering approach to spatio-temporal precipitation modeling in turkey." International Conference on the Problems of the Mediterranean Region (EPMR), April 12-15, Near East University, Nicosia, cyprus, TRNC.
- [42] Şen, Z., Altunkaynak, A., and Özger, M. (2004). "Sediment concentration and its prediction by perceptron Kalman filtering procedure." J. Hydraul., Eng., 130 (8), 816-826.
- [43] Sorenson, Kalman Filtering Theory and Applications, IEEE Press, New York, NY, 1985, pp. 90–126.
- [44] Troch, C. Paniconi, D. McLaughlin, Catchmentscale hydrological modeling and data assimilation, Adv. Water Resour. 26 (2003) 131–135.
- [45] Wood, Recent Development in Real Time Forecasting, Control of Water Resources Systems, Oxford Pergamon, Oxford, 1980.
- [46] Wood, A. Szollosi-Nagy, An adaptive algorithm for analyzing short term structural and parameter exchange in hydrologic prediction models, Water Resour. Res. 14(4) (1978) 577–581.
- [47] Weerts, G.Y. El Serafy, Particle filtering and ensemble Kalman filtering for state updating with hydrological conceptual rainfall runoff models, Water Resour. Res. 42 (2006) 140–147.

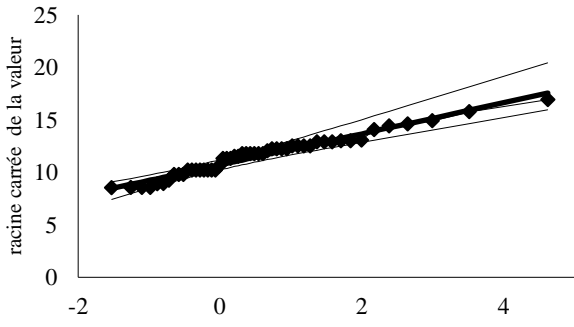
► [Publication international](#) (12p): Desalination and Water Treatment
 Publication details, including instructions for authors and subscription information:
<http://www.tandfonline.com/loi/tdwt20>
 Multi-site modeling and prediction of annual and monthly precipitation in the watershed of Cheliff (Algeria)
[Samra Harkat \(a\)](#), [Khadidja Boukharouba \(b\)](#) & [Abd El Kader Douaoui \(c\)](#)
 Published online: 20 Oct 2014.

Tableau.01: Les coordonnées Lambert des stations pluviométriques

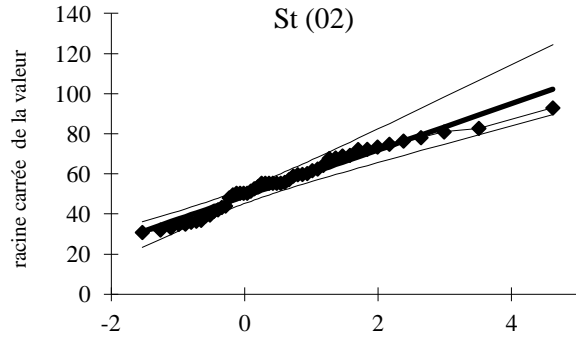
Sous Bassin versant	Code station	Nom de la station		Les coordonnées Lambert (m)		
010 (Cheliff à l'amont Bougezoul)	10502	St 01	Zmalet el amir aek	X : 464.2	Y : 177.75	Z : 850
	10701	St 02	Ain baadj	X : 403,6	Y : 214,45	Z : 1025
	10703	St 03	Rechaiga	X : 434,5	Y : 234,5	Z : 830
	10704	St 04	Ksar chellala	X : 465.4	Y : 212.8	Z : 850
	10804	St 05	Mehdia pepinière	X : 414,8	Y : 240	Z : 900
	10904	St 06	Dahmouni trumulet	X : 414,8	Y : 240	Z : 900
	10905	St 07	Ain sebain	X : 401,3	Y : 238,85	Z : 925
011 (Haut et moyens Cheliff)	11004	St 08	khemisti	X : 433.5	Y : 263.45	Z : 935
	11006	St 09	Tissemsilt	X : 420.2	Y : 257	Z : 858
	11104	St 10	Ain boucif	X : 540.75	Y : 287.8	Z : 1250
	11206	St 11	Chahbounia	X : 491.3	Y : 249.35	Z : 665
	11207	St 12	Kef lasfar	X : 480.8	Y : 252.8	Z : 690
	11208	St 13	Boughzoul	X : 507.1	Y : 272.25	Z : 643
	11301	St 14	Ksar el boukhari	X : 503.8	Y : 287.9	Z : 630
	11302	St 15	Derrag	X :	Y :	Z :
	11404	St 16	Zoubiria mongorno	X : 513.5	Y : 312.6	Z : 932
	11405	St 17	Ghrib bge	X : 487.25	Y : 318.4	Z : 460
	11605	St 18	Teniet - el-had	X :	Y :	Z :
	11606	St 19	Sidi mokrefi	X : 463.2	Y : 304.25	Z : 425
	11607	St 20	Tarik ibn ziad	X : 450.15	Y : 299.5	Z : 660
	11703	St 21	Bordj el amir khaled	X : 455.85	Y : 313.8	Z : 370
	11803	St 22	Sidi medjahed	X : 452.15	Y : 337.05	Z : 850
11904	St 23	Rouina mairie	X : 419.8	Y : 327.3	Z : 190	
012 (Bas Cheliff)	12108	St 24	Fodda bge	X : 402	Y : 305.6	Z : 430
	12205	St 25	Benaria	X : 378.15	Y : 338.4	Z : 320
	12304	St 26	Souk-el-had	X :	Y :	Z :
	12306	St 27	Bordj bounaama	X : 402.6	Y : 284.1	Z : 1050
	12309	St 28	Oued- sly	X :	Y :	Z :
	12503	St 29	Sidi hosni	X : 392,95	Y : 242,05	Z : 790
	12605	St 30	Ammi moussa	X : 256,4	Y : 286,15	Z : 140
	12706	St 31	Mendes	X : 333,7	Y : 262,95	Z : 465
	12908	St 32	Tiaret GRHA	X : 373,4	Y : 232,8	Z : 1100
	12915	St 33	Louhou	X : 363,9	Y : 209,15	Z : 1040
013 (Bas Cheliff)	13004	St 34	Ain el haddid	X : 334,5	Y : 197	Z : 829
	13202	St 35	Rosfa	X : 330,6	Y : 179,5	Z : 960
	13503	St 36	Oued djemaa	X : 315,15	Y : 277,25	Z : 75
	13504	St 37	Saline oued djemaa	X : 315,5	Y : 283,3	Z : 60
	13506	St 38	El matmar DRS	X : 300,1	Y : 271,5	Z : 73
	13507	St 39	L' Hillil	X : 288,65	Y : 272	Z : 125

Source : ANRH Alger(2010)

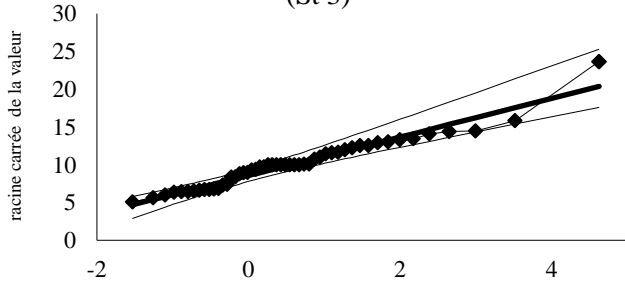
St (01)



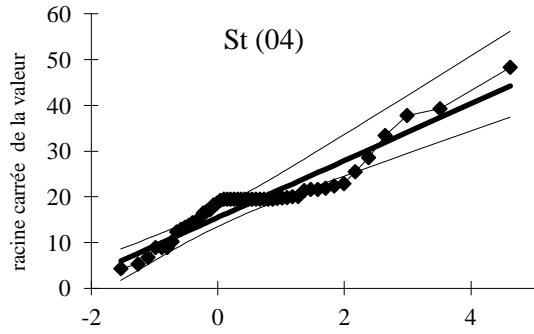
St (02)



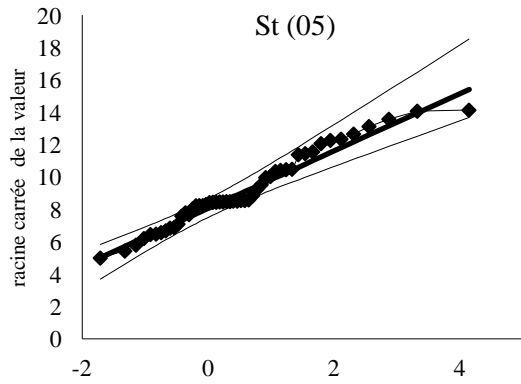
(St 3)



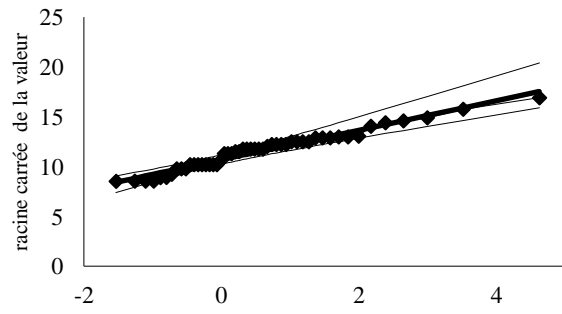
St (04)



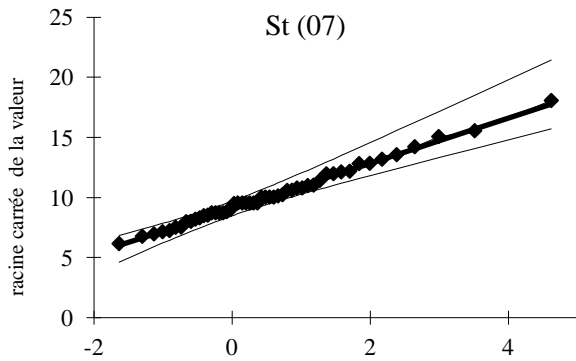
St (05)



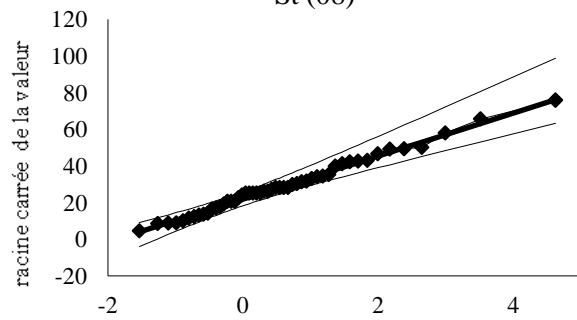
St (06)

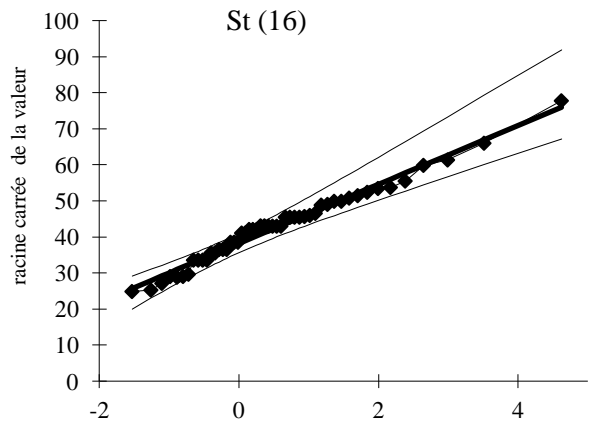
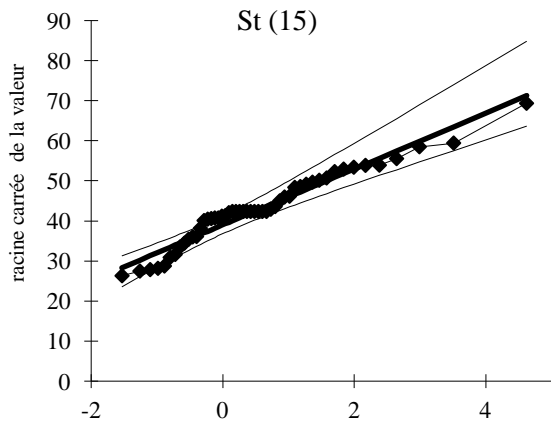
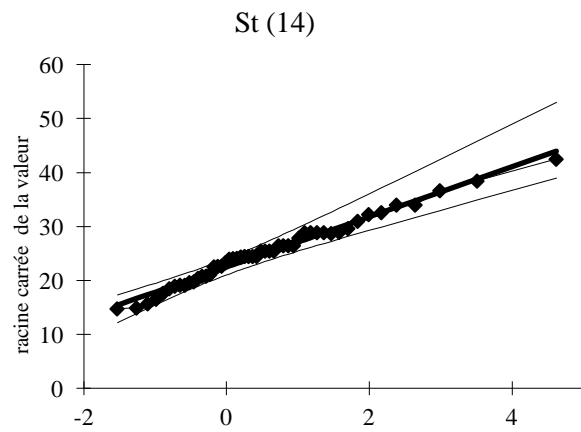
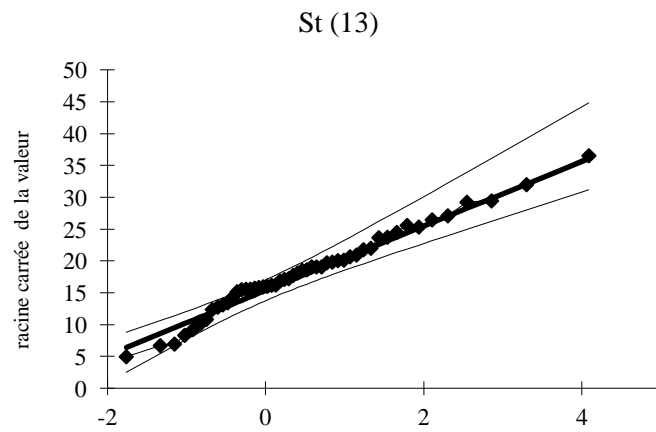
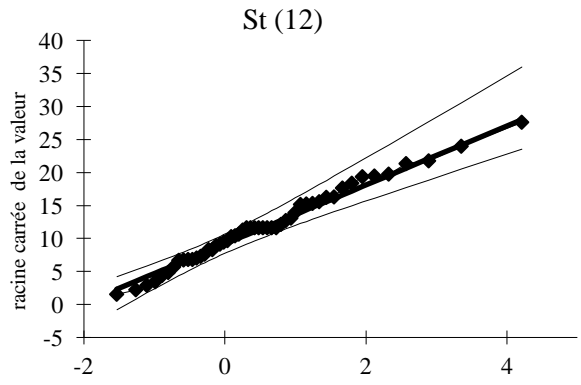
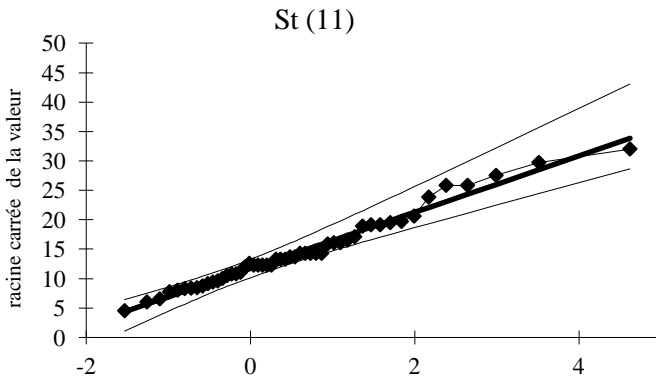
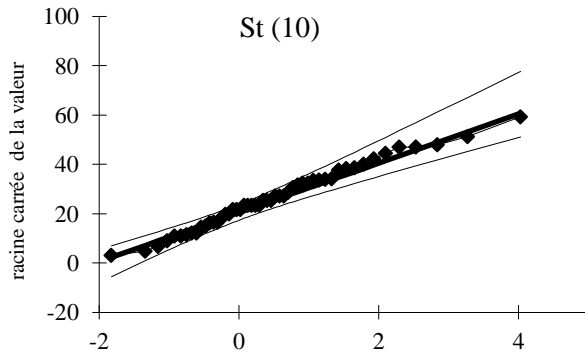
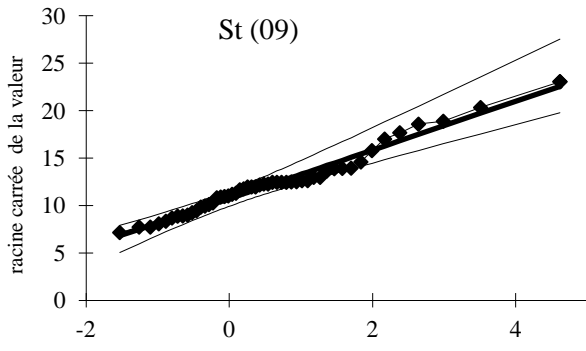


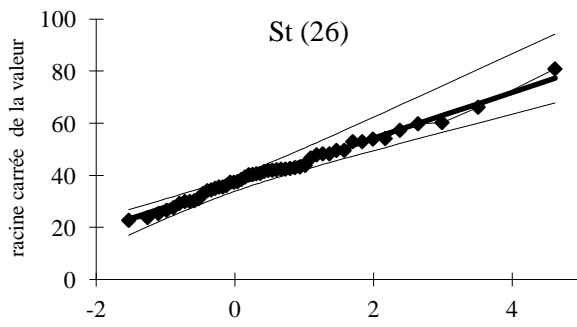
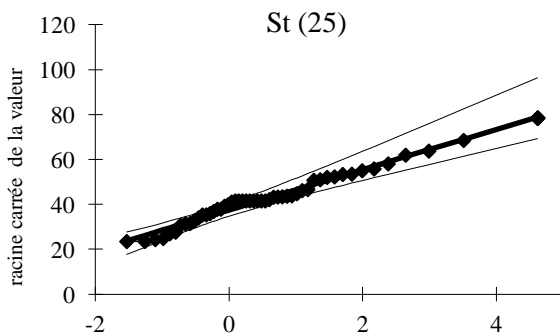
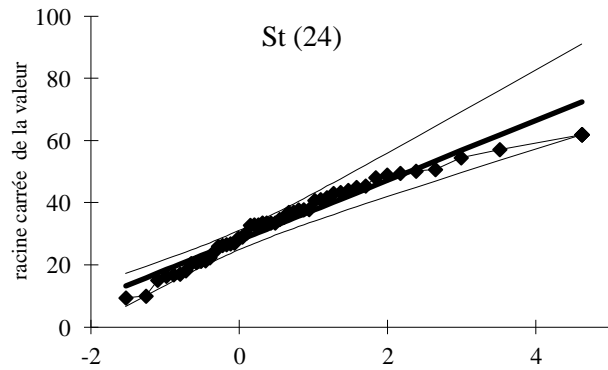
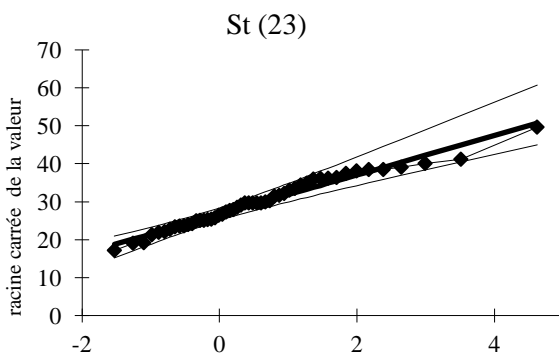
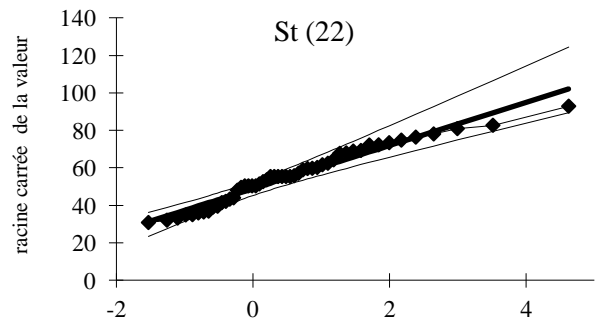
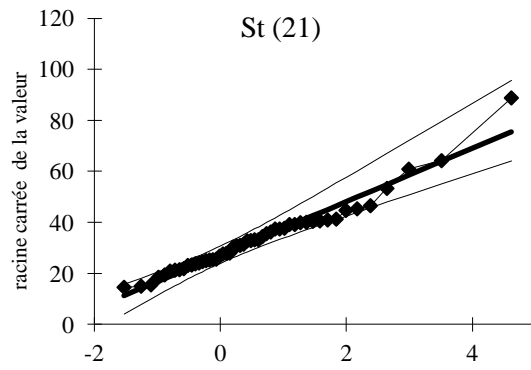
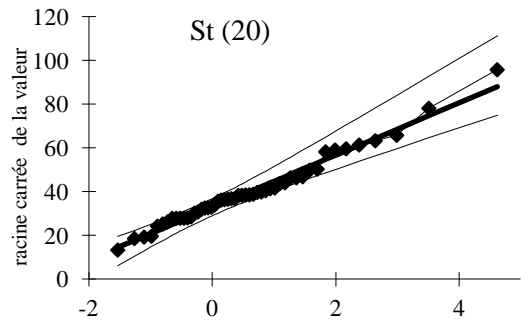
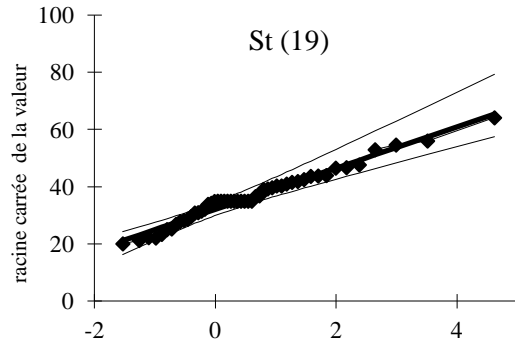
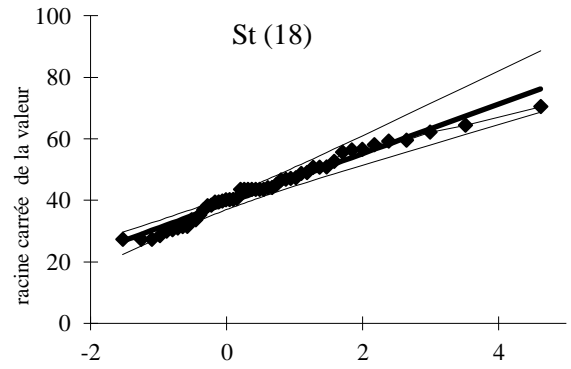
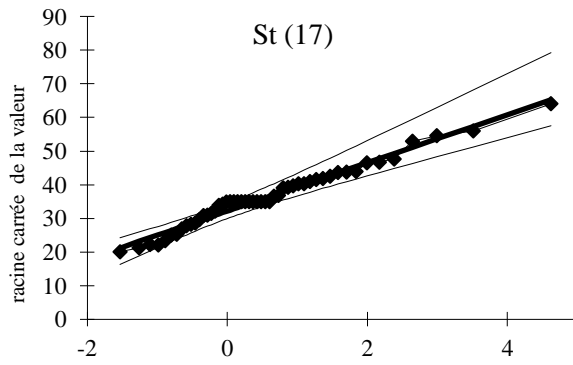
St (07)

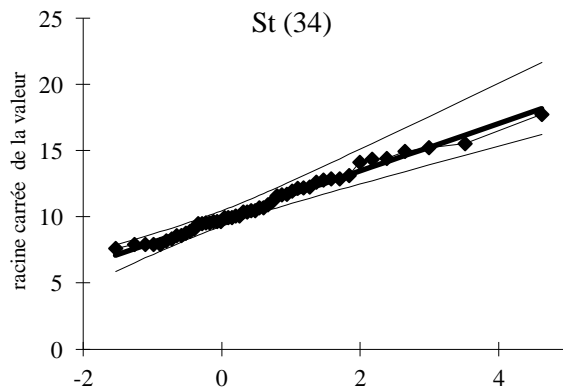
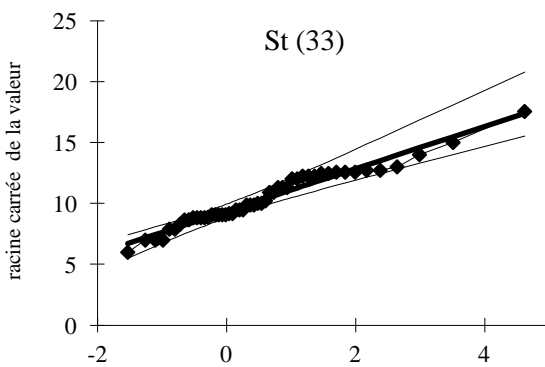
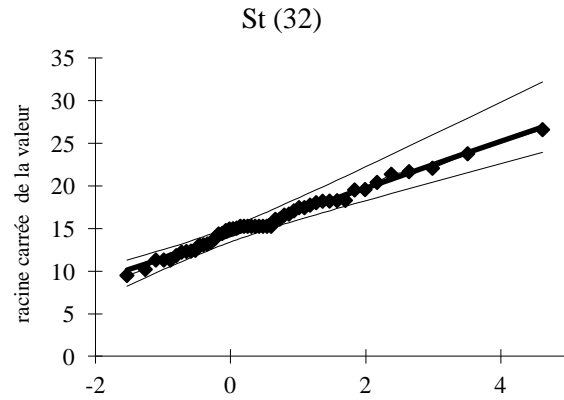
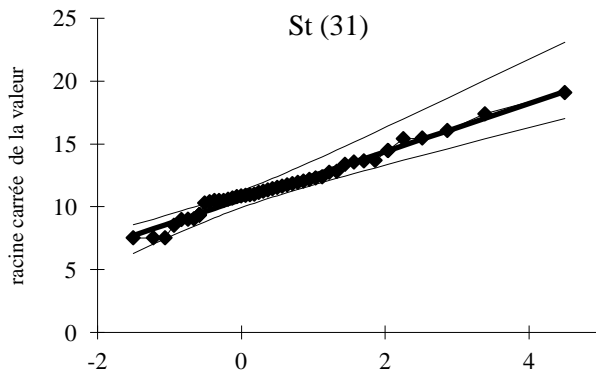
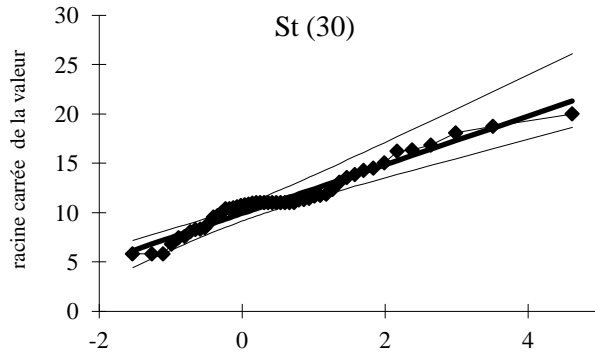
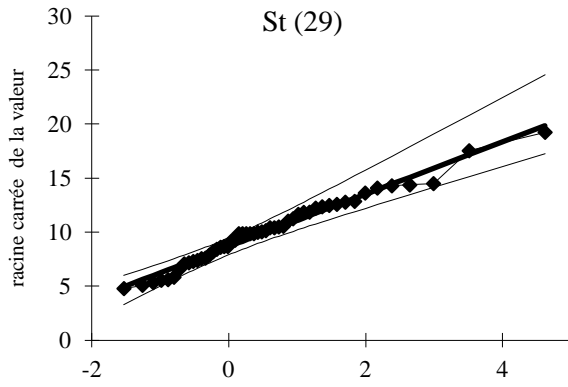
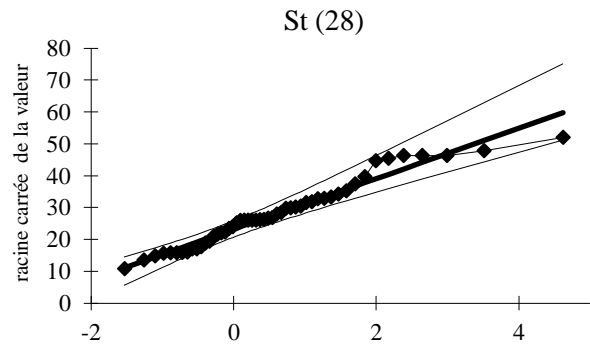
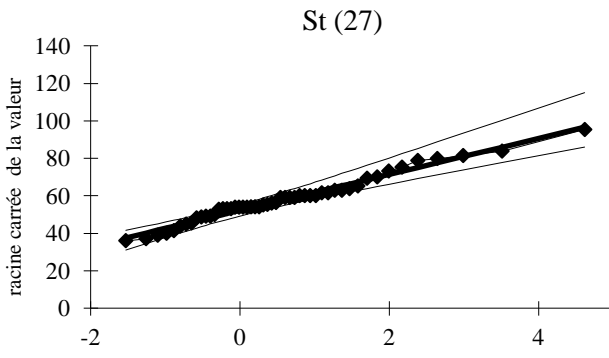


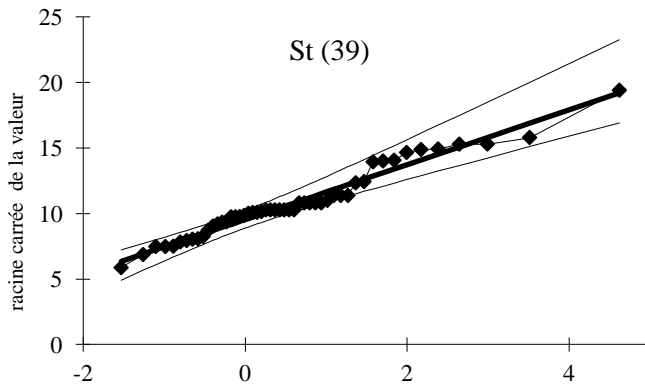
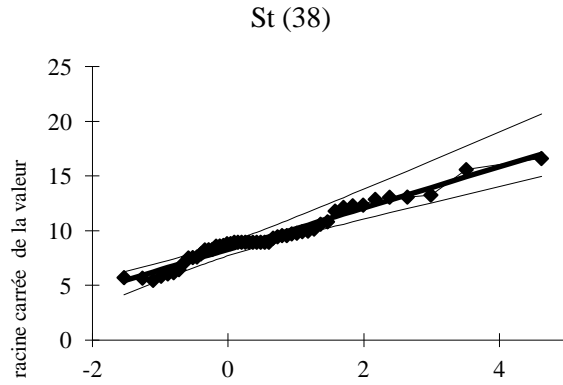
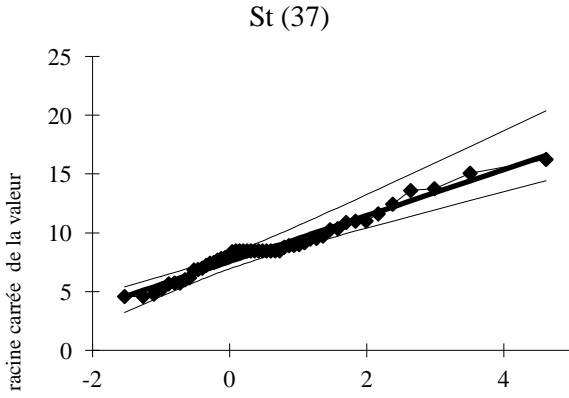
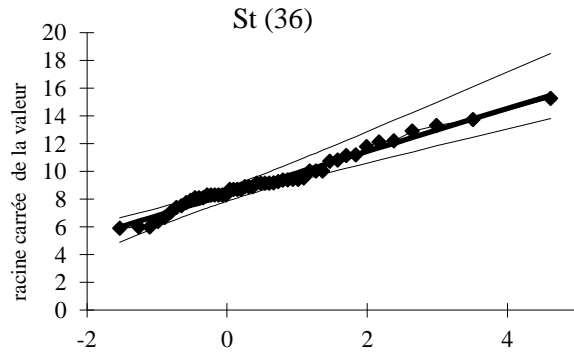
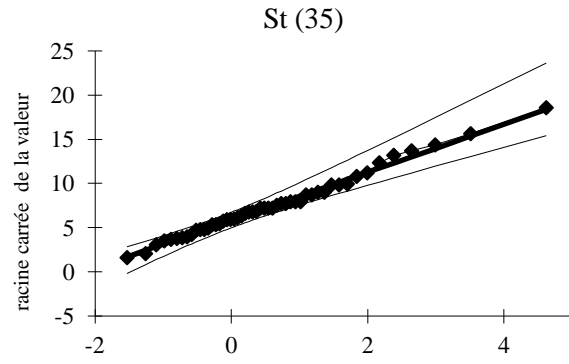
St (08)











a) Méthode des rapports :

La méthode des rapports utilise le rapport entre la somme des précipitations pendant une période homogène de la station à ajuster et la somme des précipitations pour la période correspondante à la station de comparaison.

$$Y = b.X$$

Y : valeur pluviométrique pendant une période homogène de la station à ajuster

X : valeur correspondante à la station de référence

b : constante d'ajustement

$$b = \frac{V}{U}$$

V : Somme des précipitations pendant une période homogène de la station à ajuster.

U : Somme des précipitations pour la période correspondante à la station de comparaison.

Cette méthode a été utilisée pour combler les lacunes à l'échelle mensuelle.

b) Méthode de la régression linéaire :

Cette méthode permet le comblement des lacunes d'observation par rapport à une station de référence, cette dernière doit être proche à une station à corriger.

Le facteur qui lie cette régression est le coefficient de corrélation 'r' estimé à partir de 'n' couples de valeurs observées simultanément, sa formule est :

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \cdot \sum_{i=1}^n (Y_i - \bar{Y})^2}}$$

X_i : variable connue de la station de référence à partir de laquelle est estimée la variable inconnue de la station incomplète.

Y_i : Variable de la station incomplète.

\bar{X} : Moyenne de la série d'observation de référence.

\bar{Y} : Moyenne de la série incomplète.

Cette méthode a été utilisée pour combler les lacunes à l'échelle annuelle.

On se basera sur la valeur du coefficient de corrélation " r ", une erreur de première espèce α de 5 %, un nombre de degré de liberté $\sigma = n - 2$ et la taille de l'échantillon « n » pour valider par la table de Fisher (Dagnellie, 1992) la signification de ce coefficient comparant au coefficient de corrélation critique " r_{cr} " donné par la table.

Les nouvelles séries permet la détermination des différentes caractéristiques (moyenne, écart type et coefficient de variation) d'un échantillon, mensuellement et annuellement .

A.1. Le bassin versant Cheliff à l'amont Bougezoul « 010 »

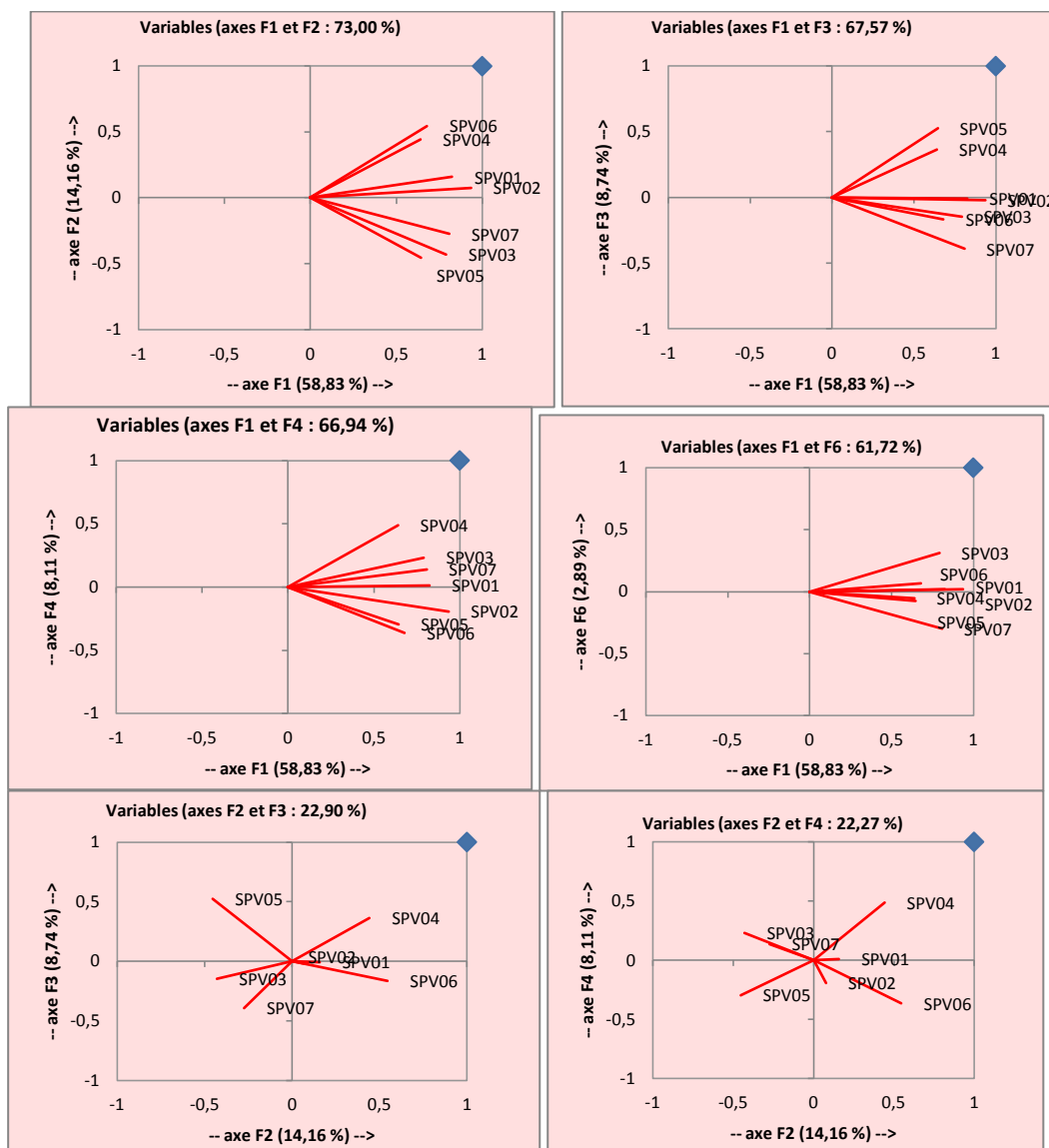


Fig II.1.1. Cercle de corrélation (Analyse graphique)

A.2. Le bassin versant Haut et Moyen Cheliff « 011 » :

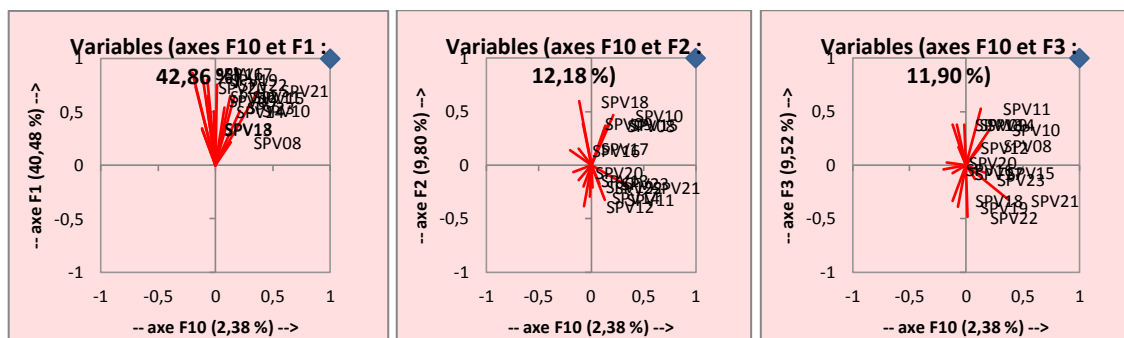


Fig II.1.2. Cercle de corrélation (Analyse graphique)

A.3. Le bassin versant du Bas Cheliff « 012 »:

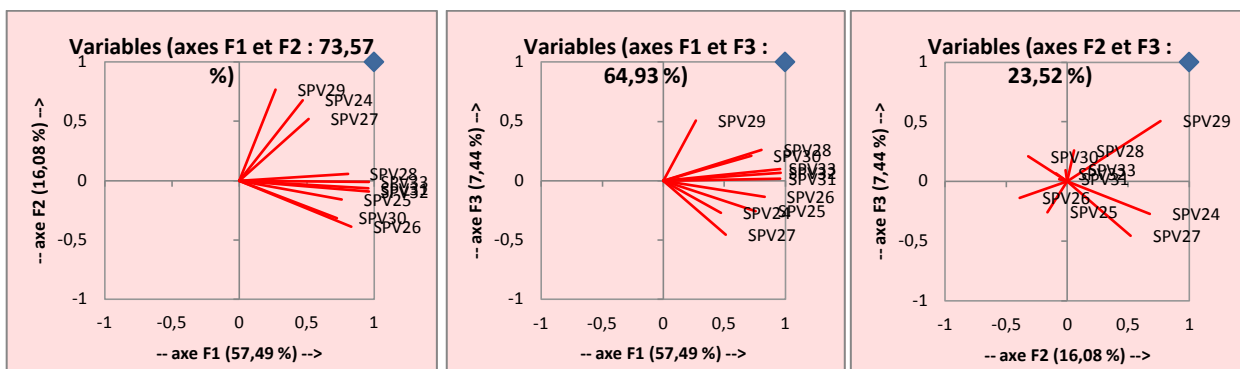


Fig II.1.3. Cercle de corrélation (Analyse graphique)

A.4. Le bassin versant du Bas Cheliff « 013 »:

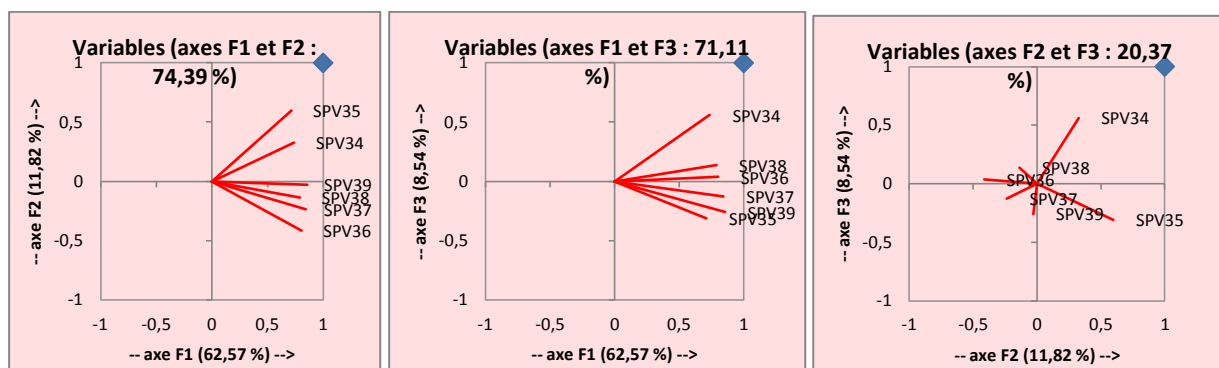
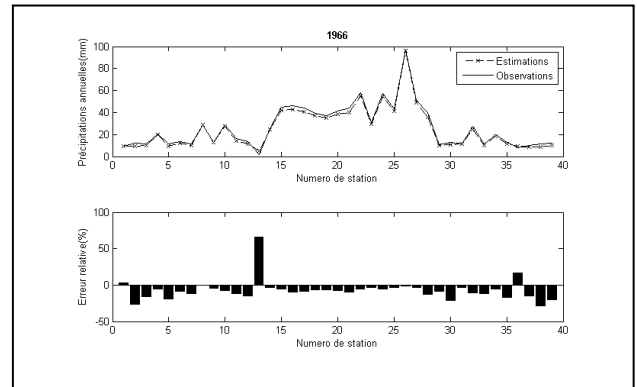
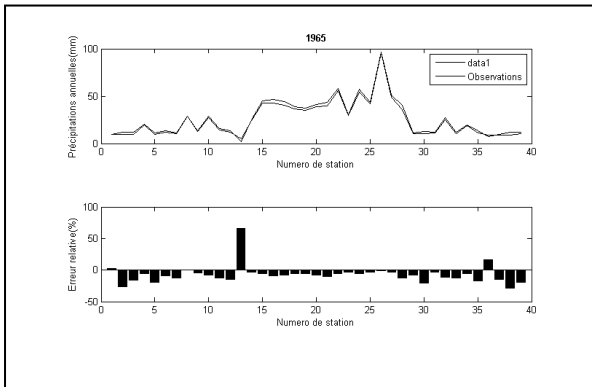
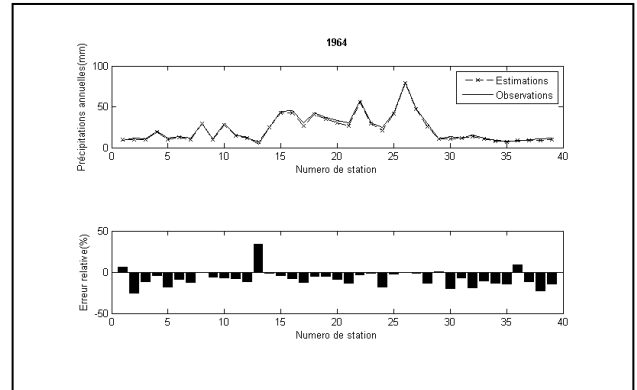
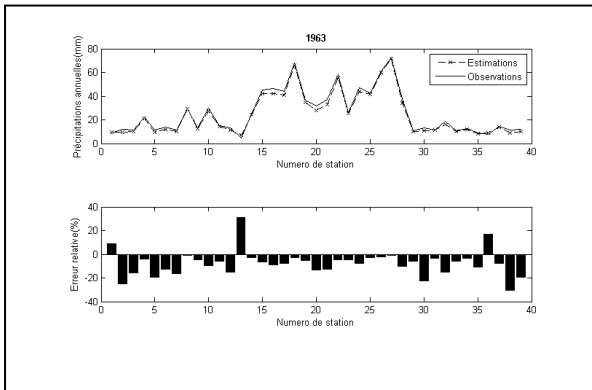
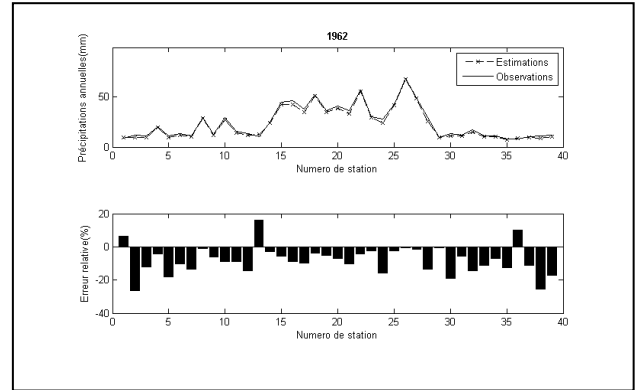
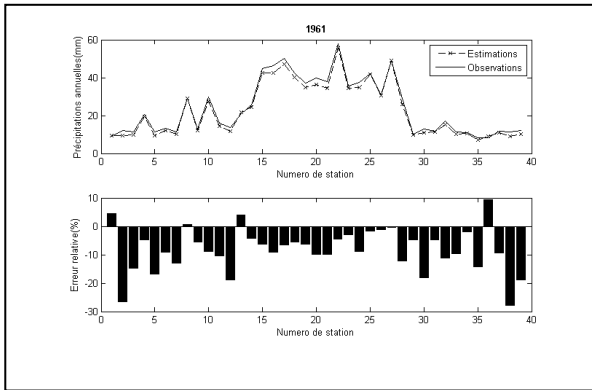
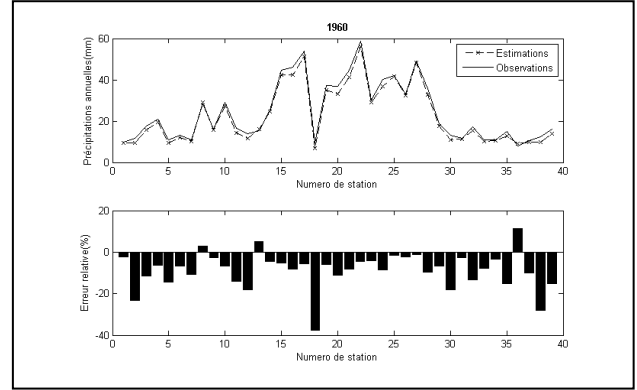
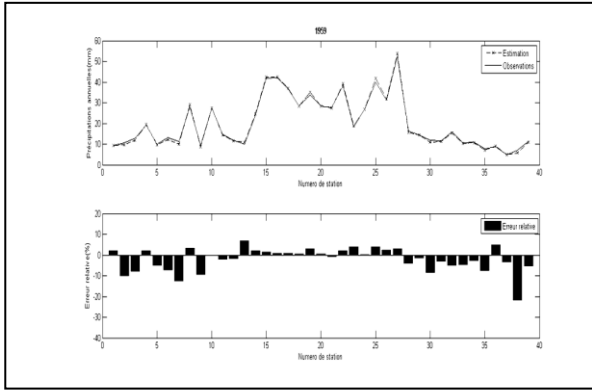
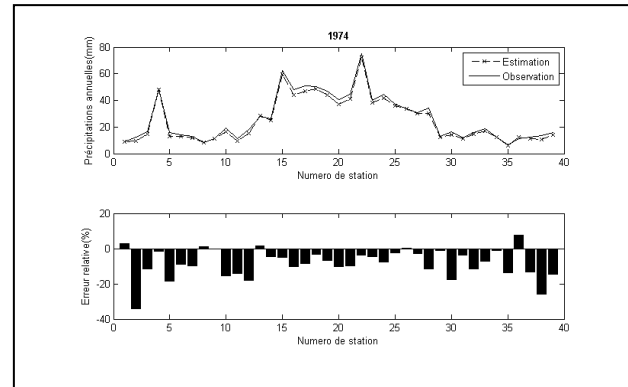
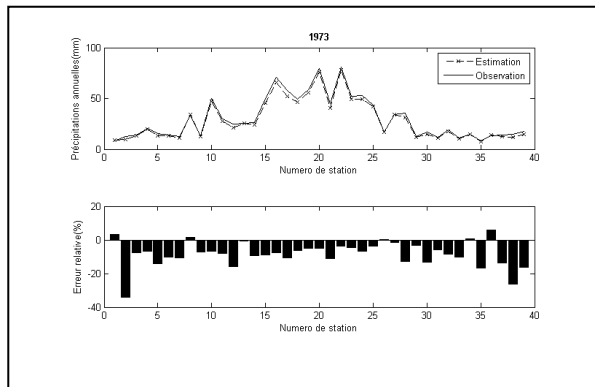
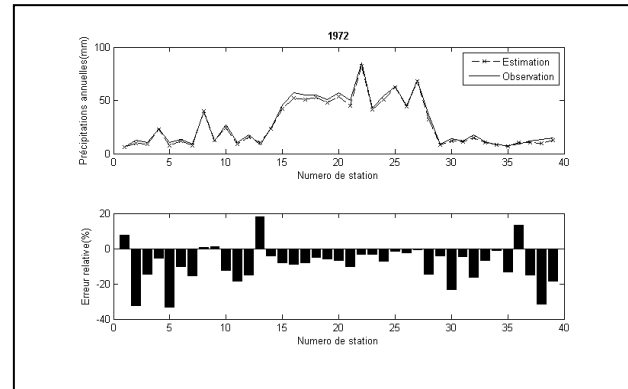
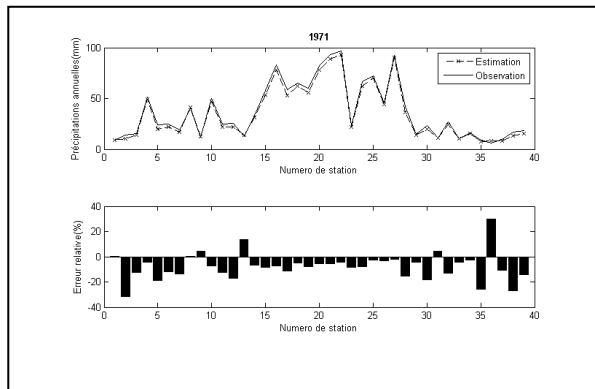
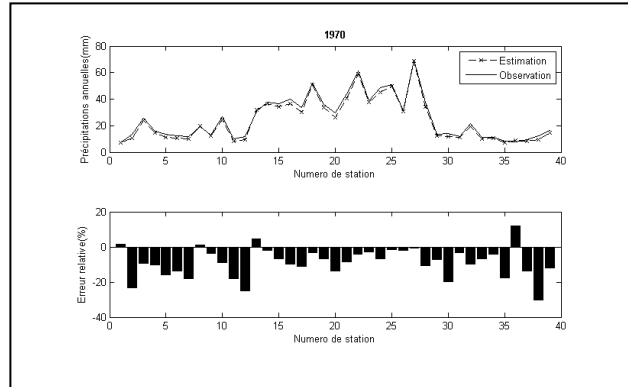
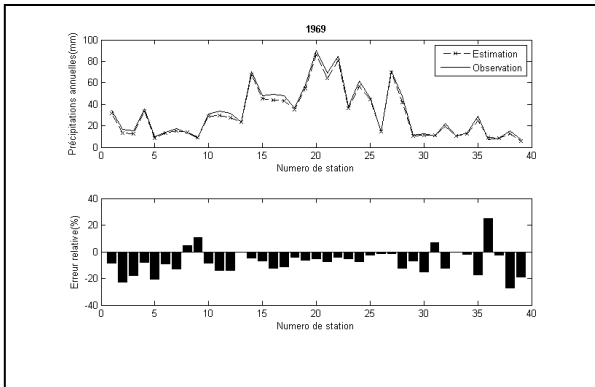
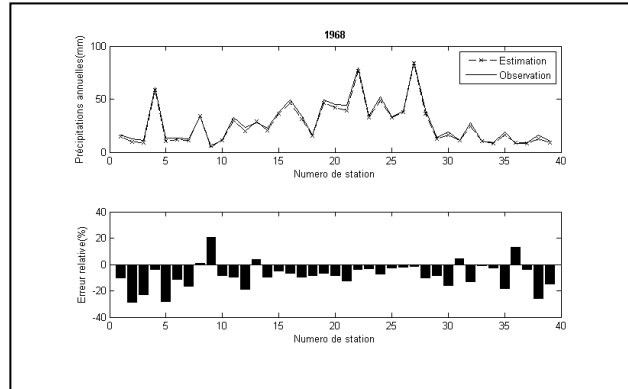
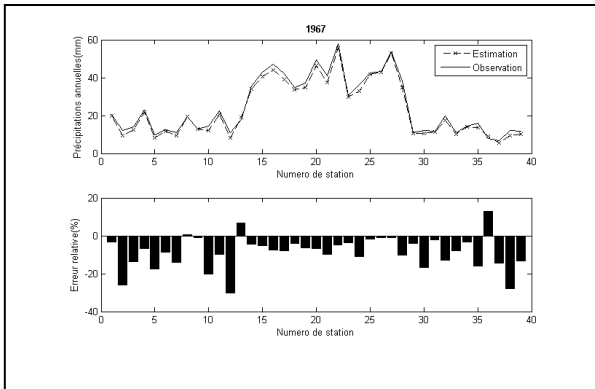
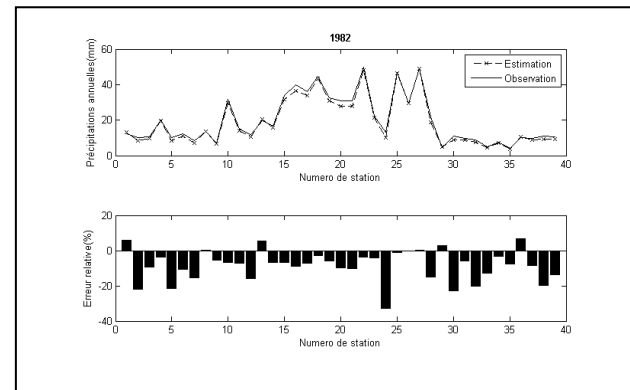
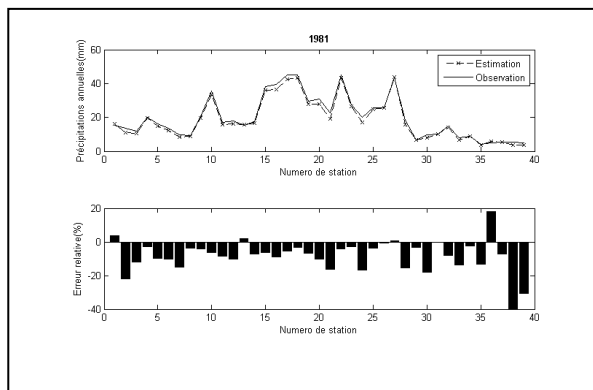
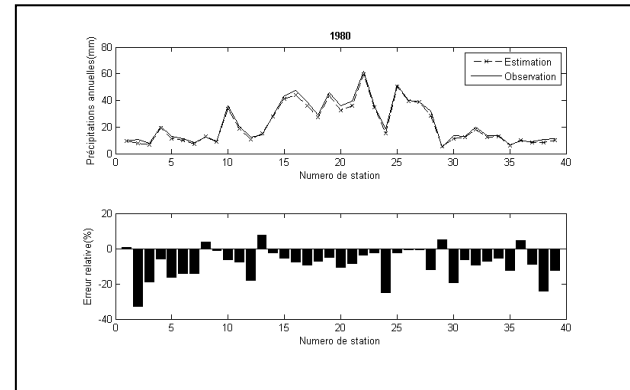
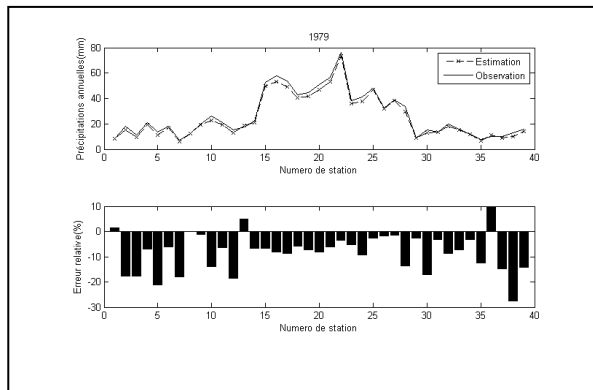
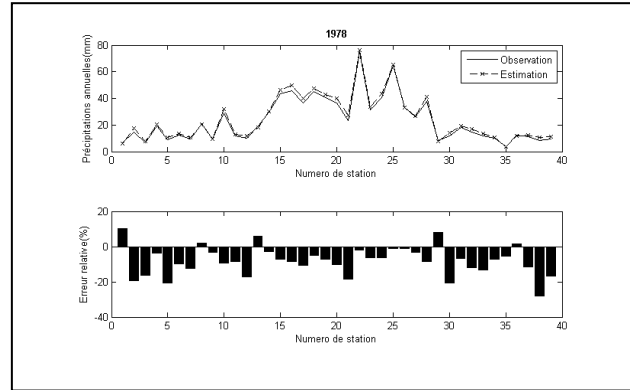
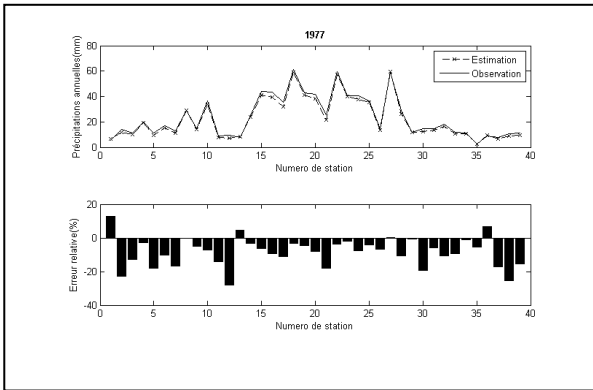
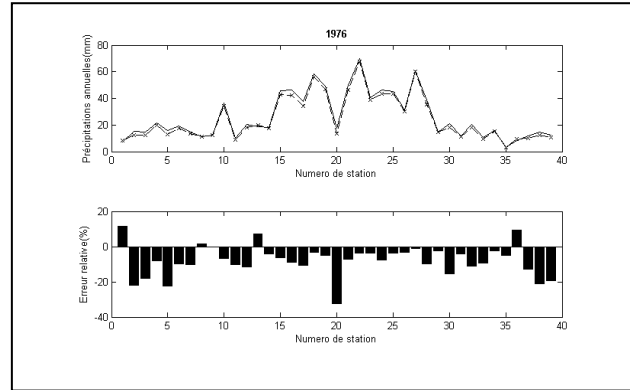
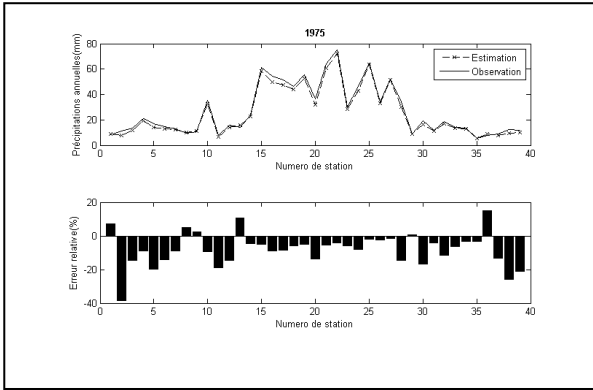
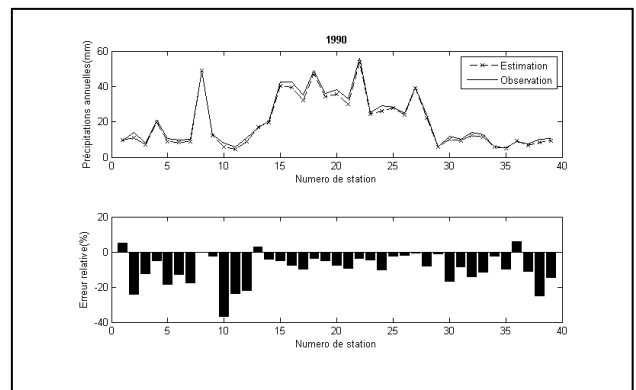
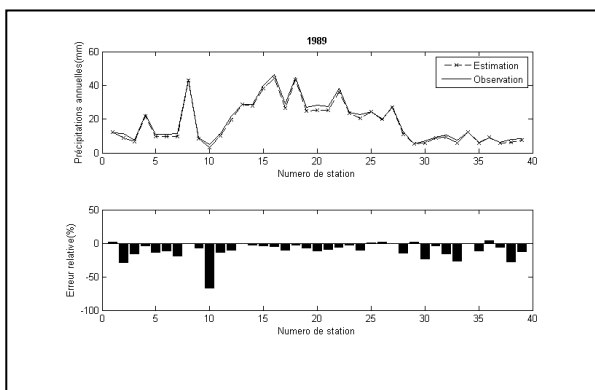
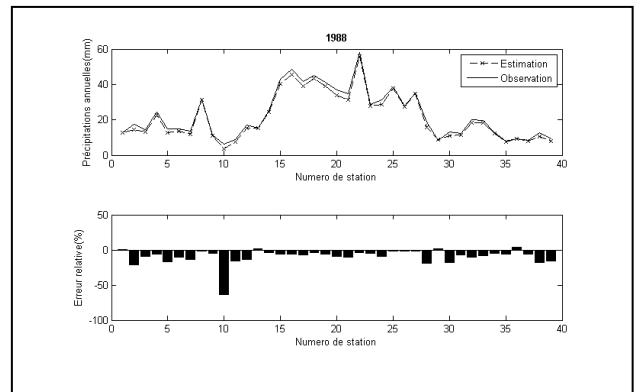
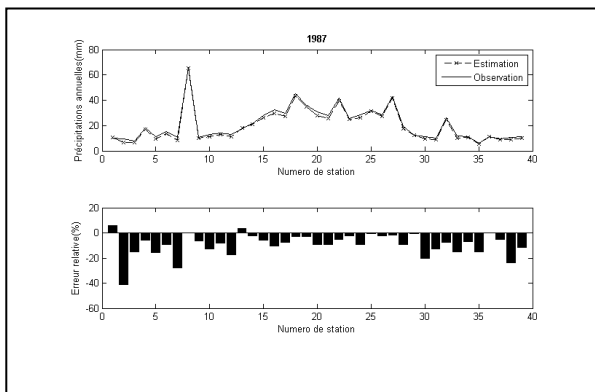
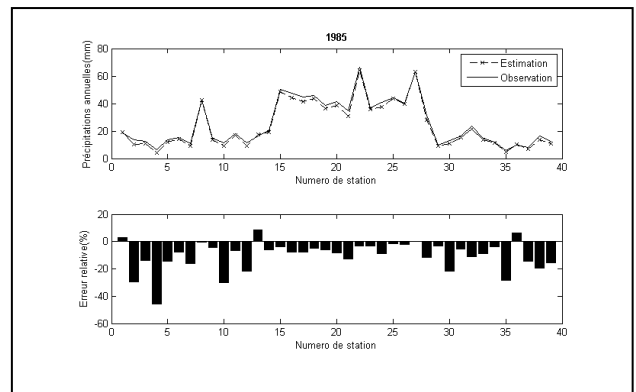
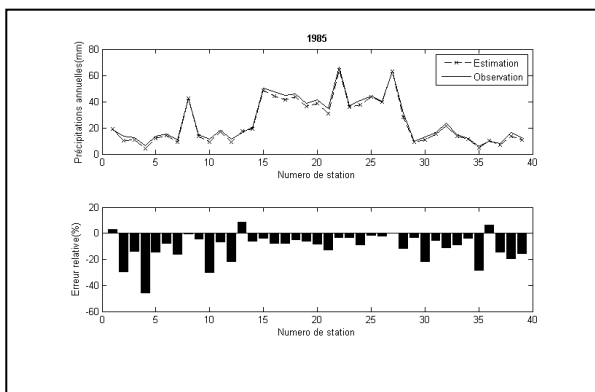
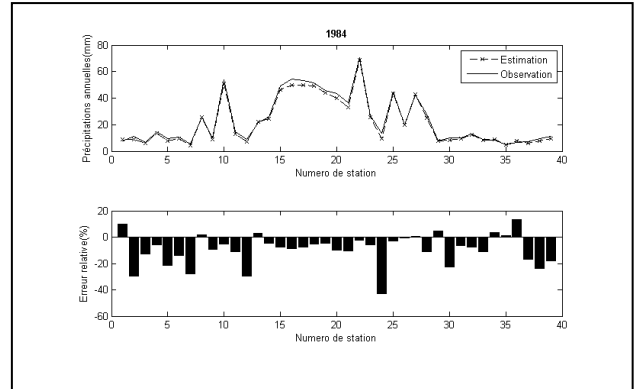
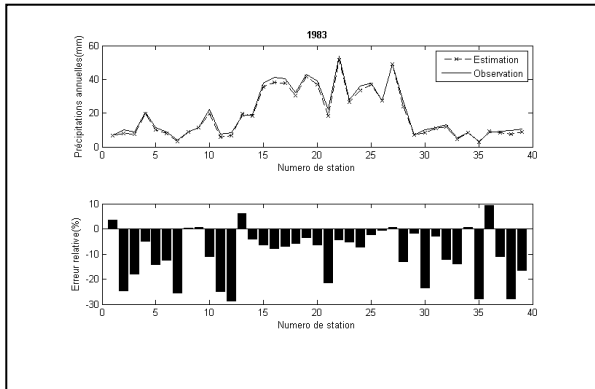


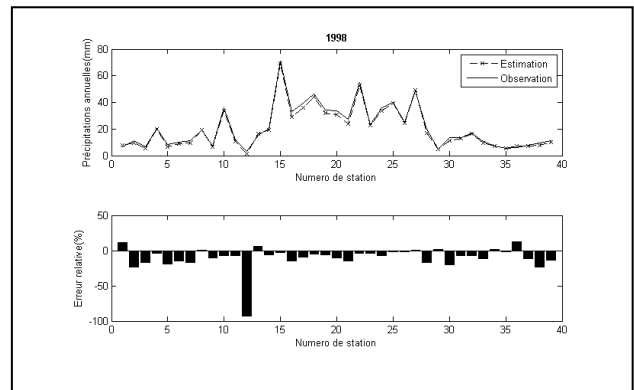
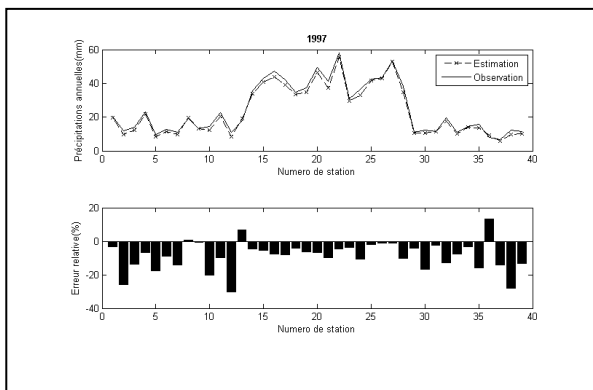
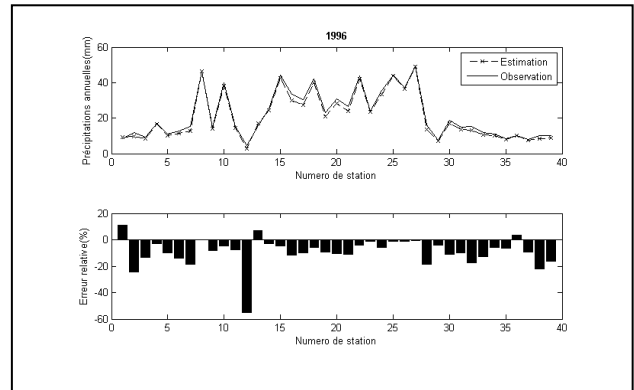
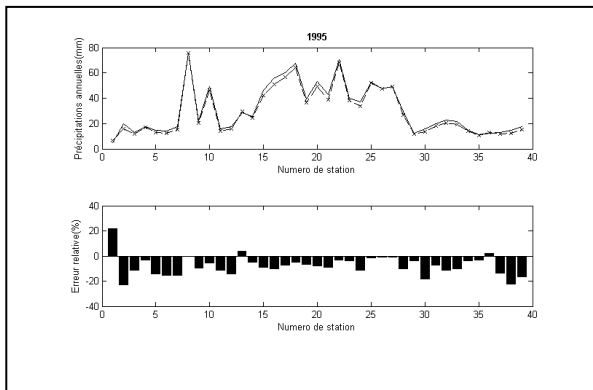
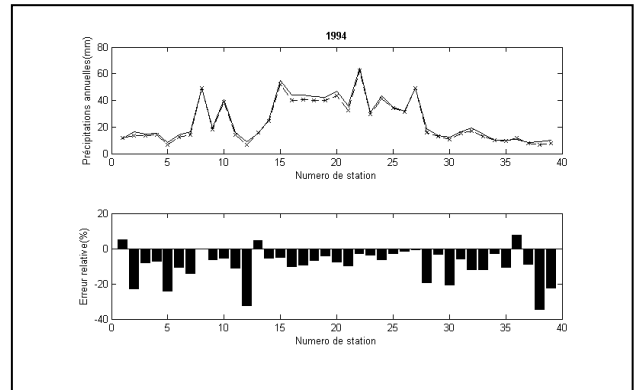
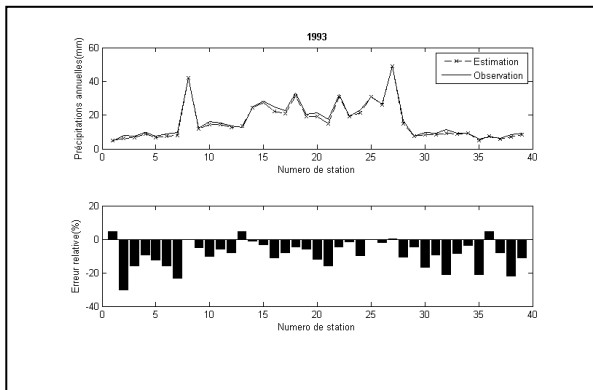
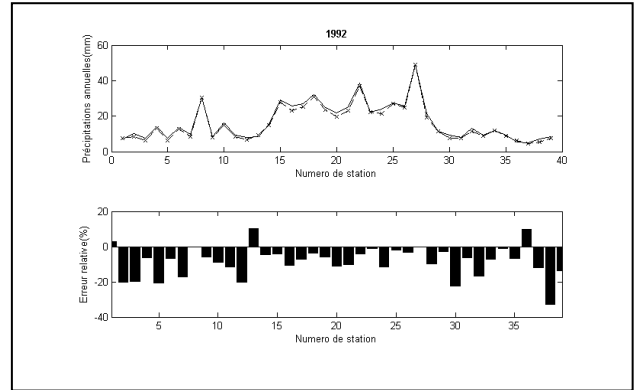
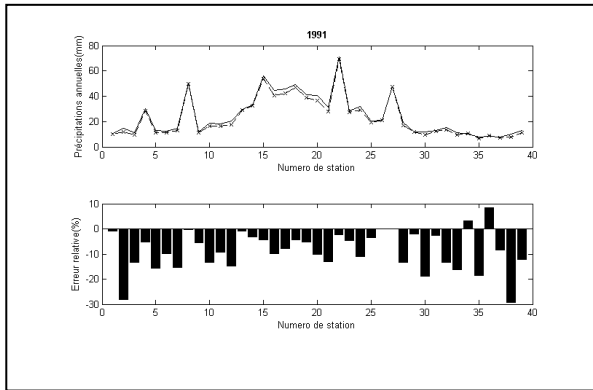
Fig II.1.4. Cercle de corrélation (Analyse graphique)

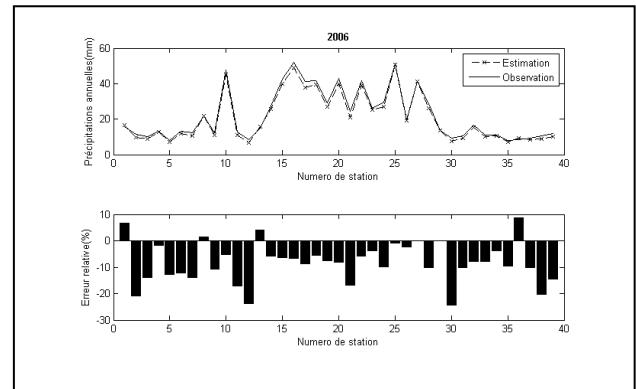
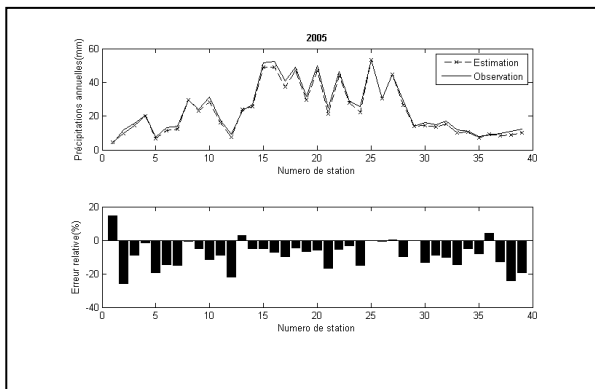
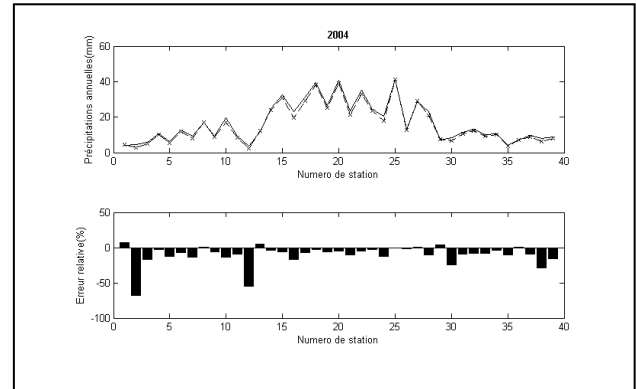
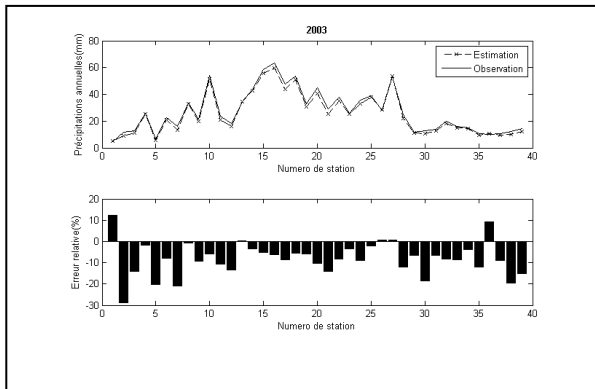
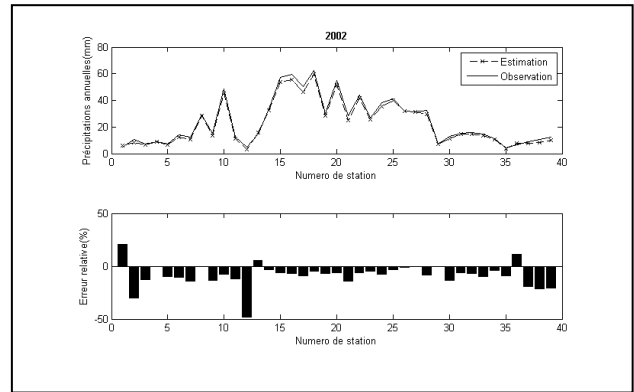
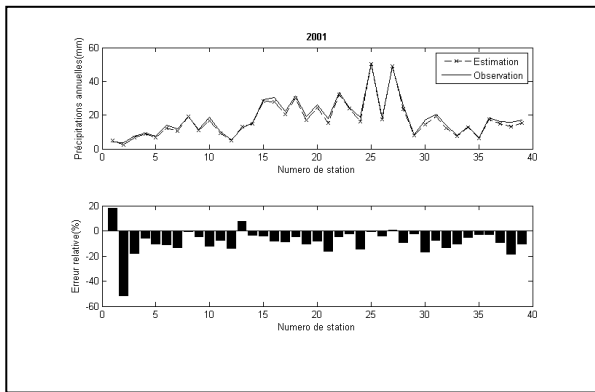
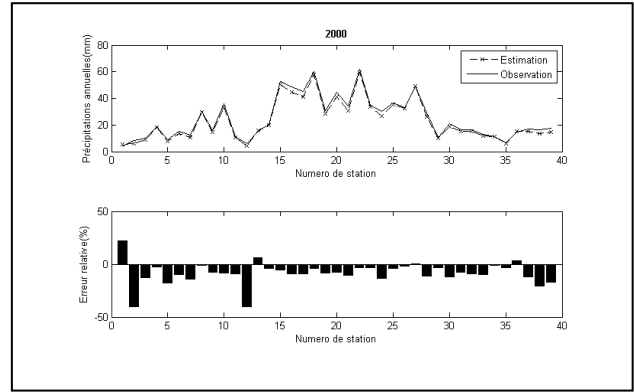
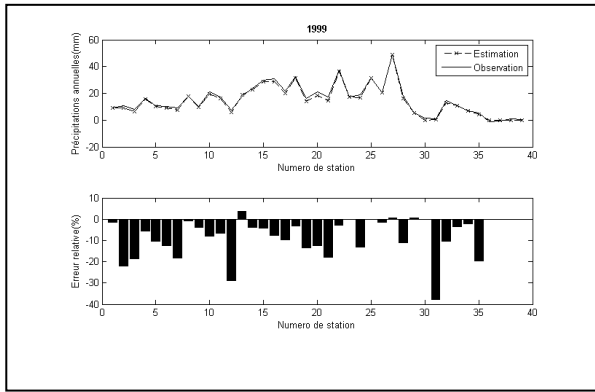


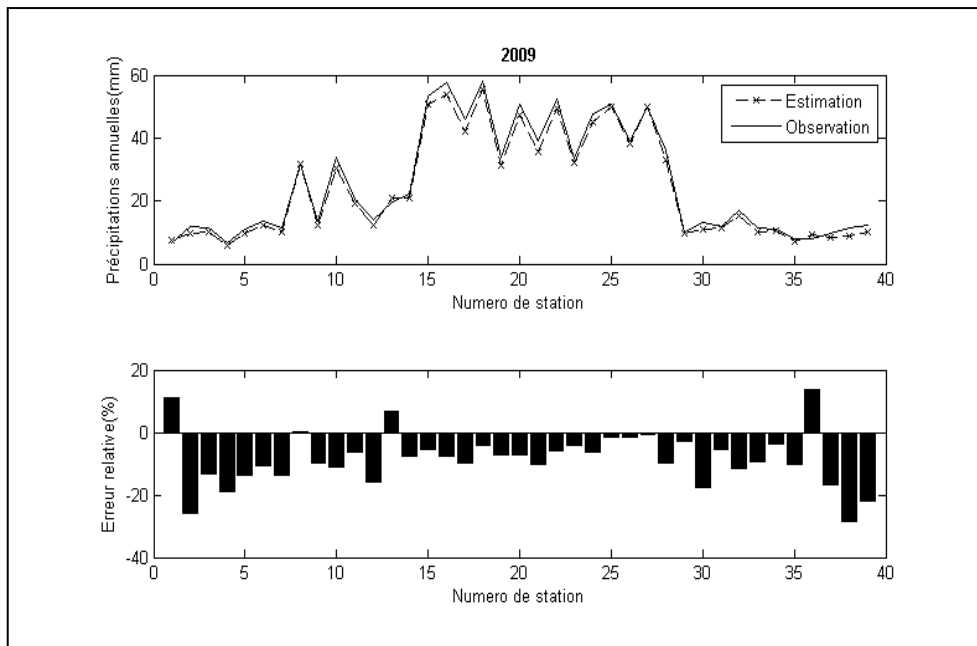
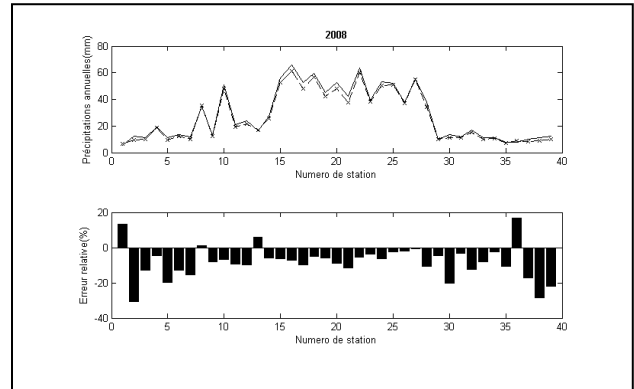
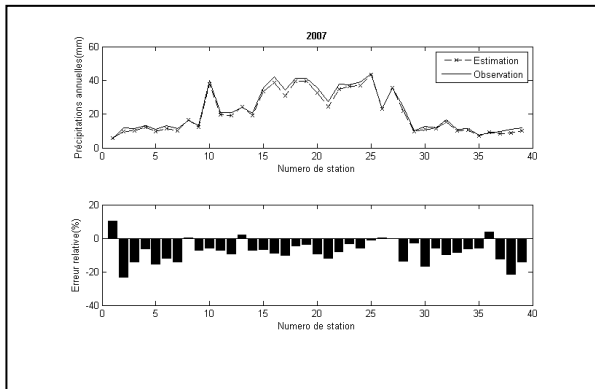




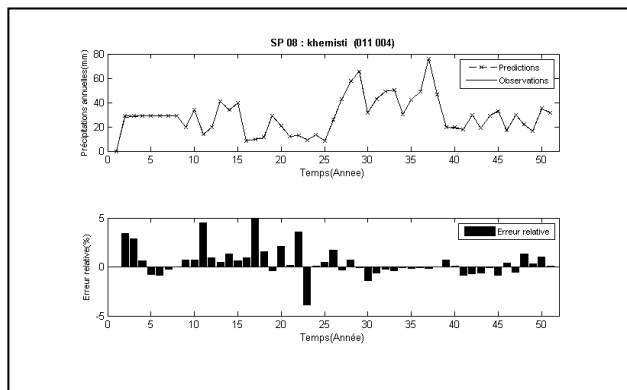
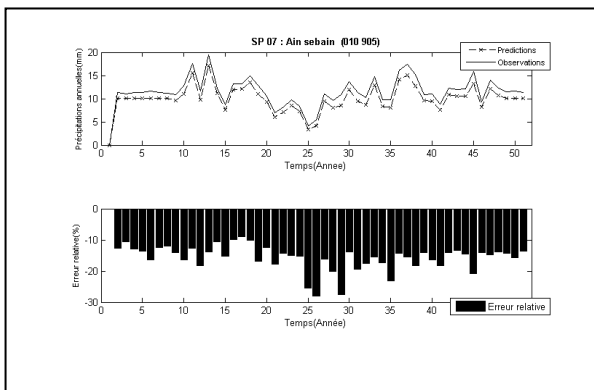
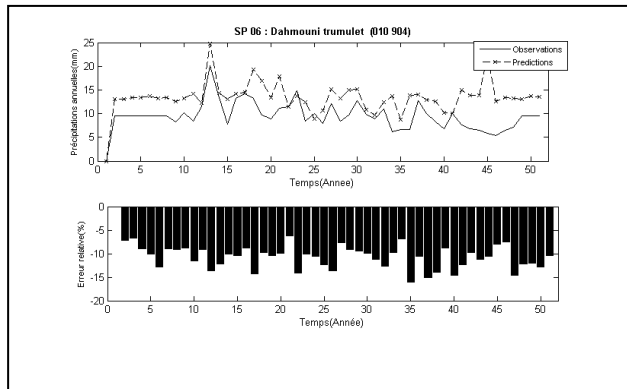
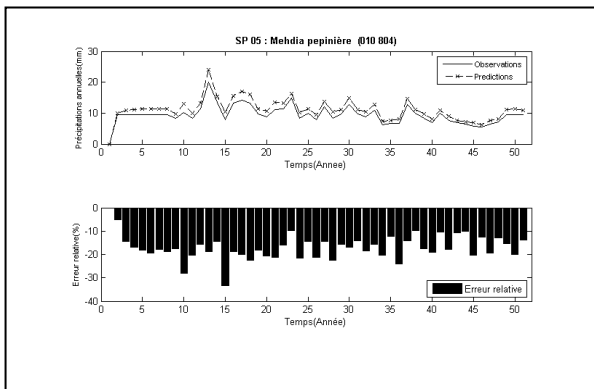
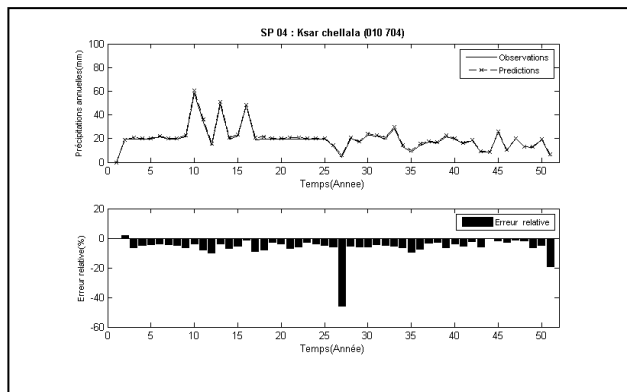
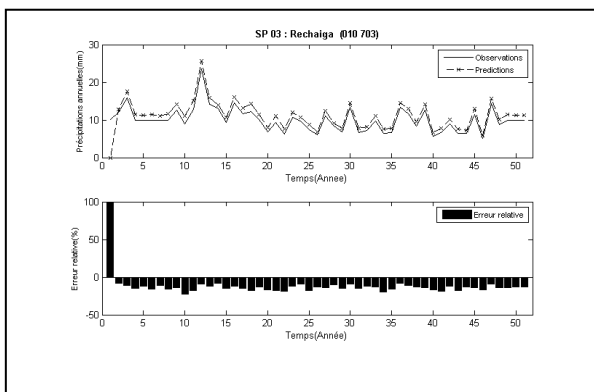
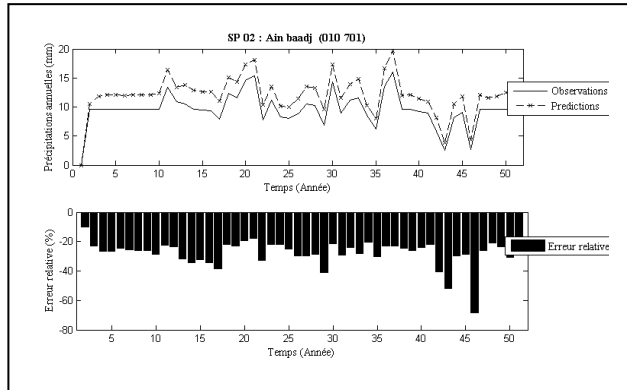
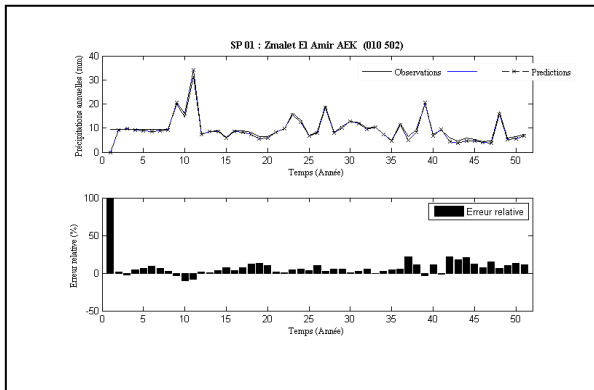




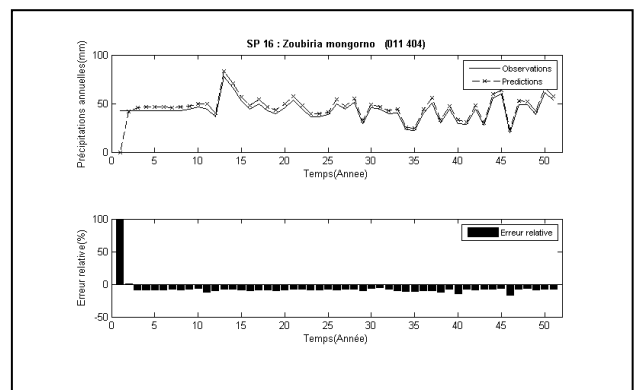
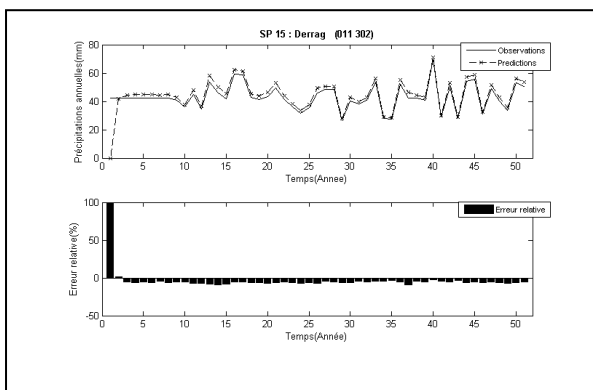
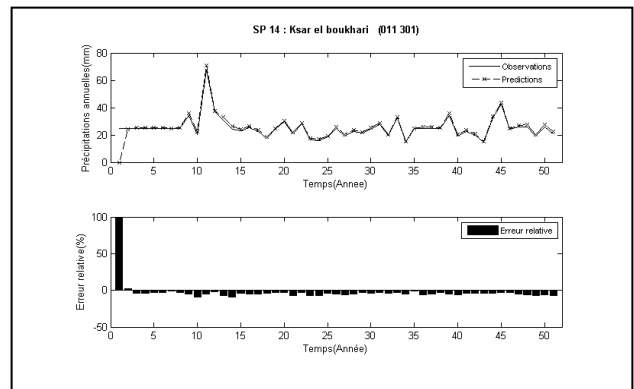
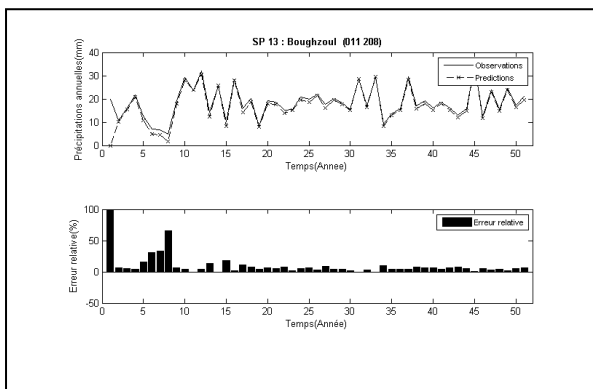
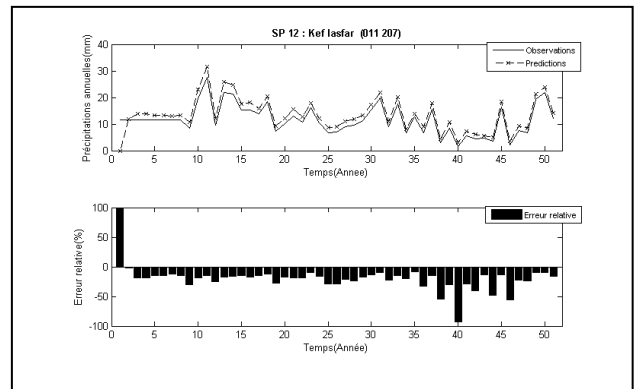
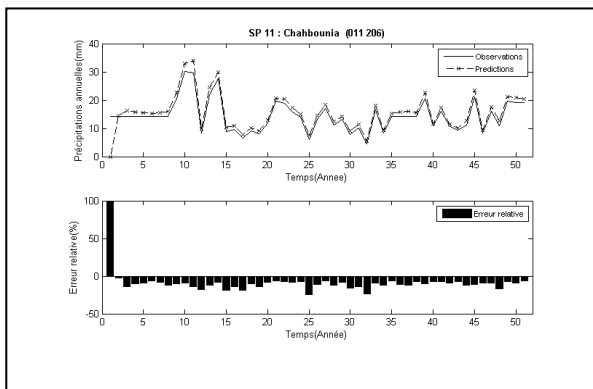
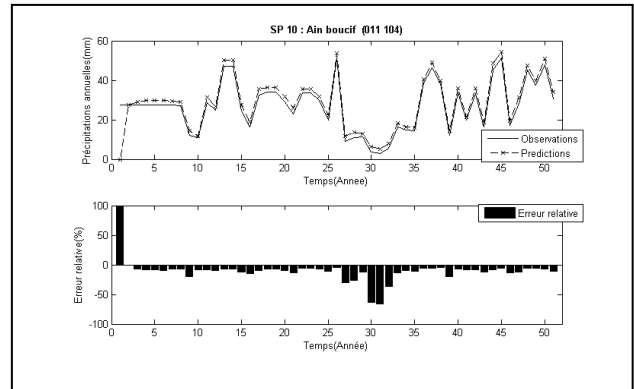
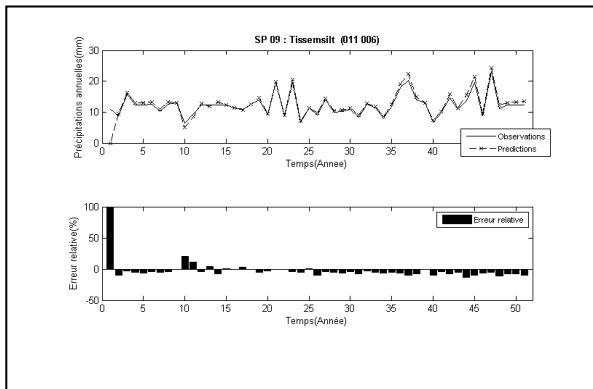




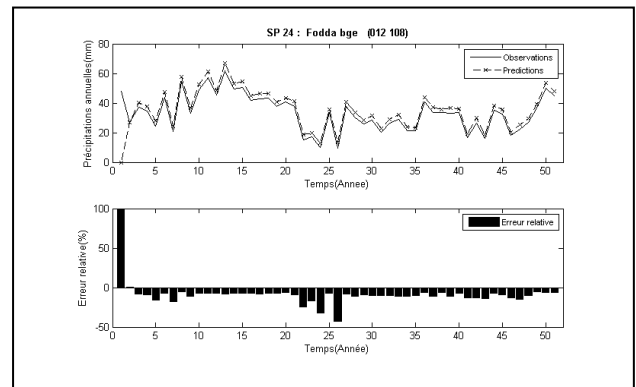
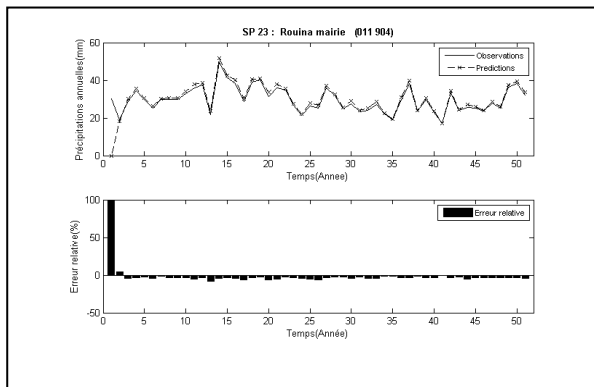
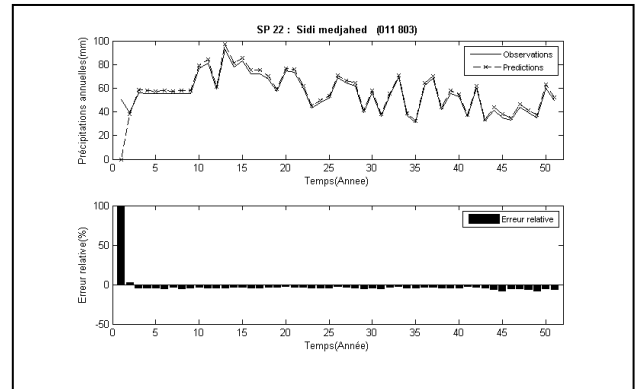
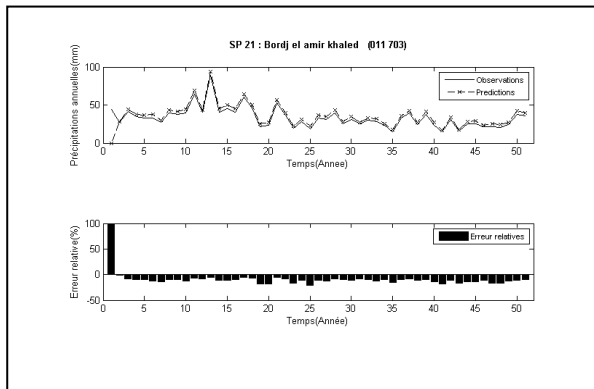
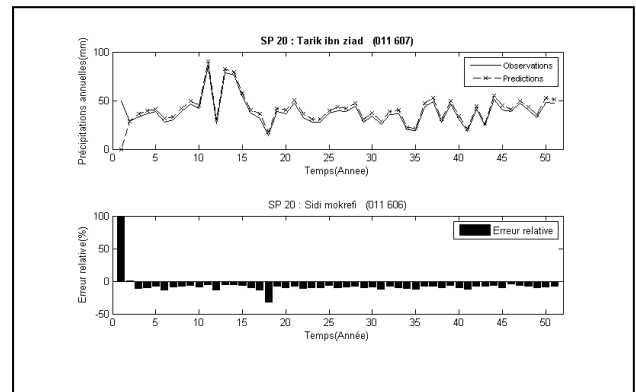
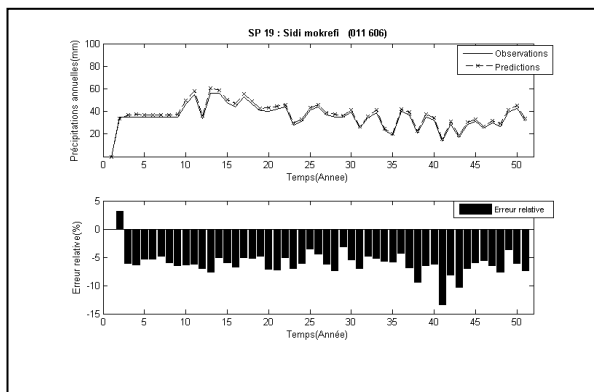
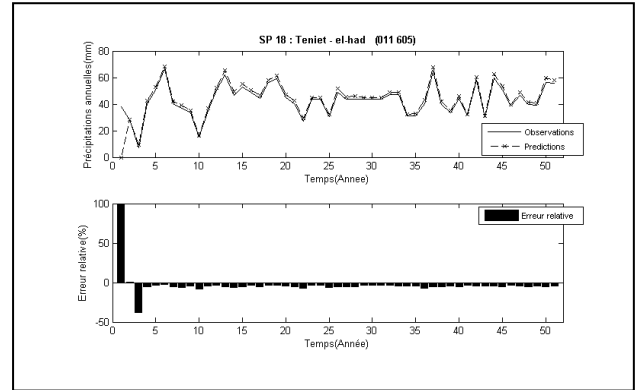
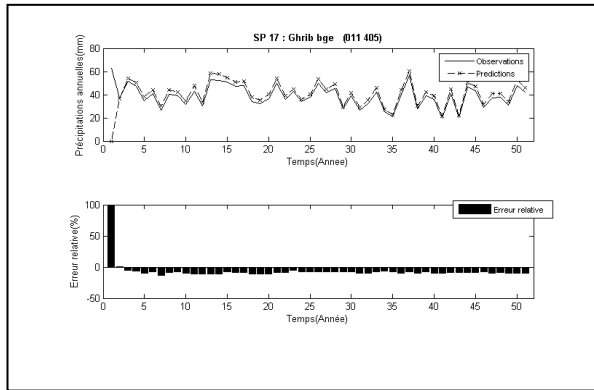
Prédiction Temporelle : Pas Annuel



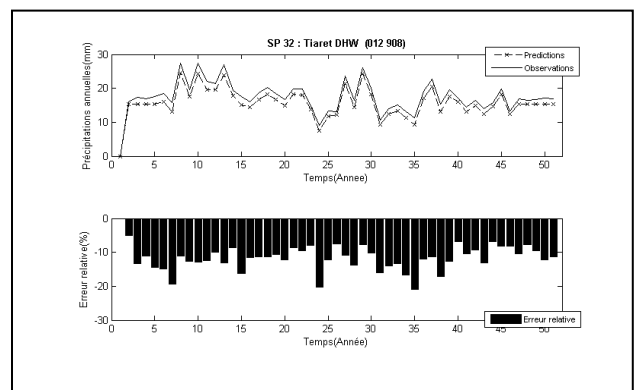
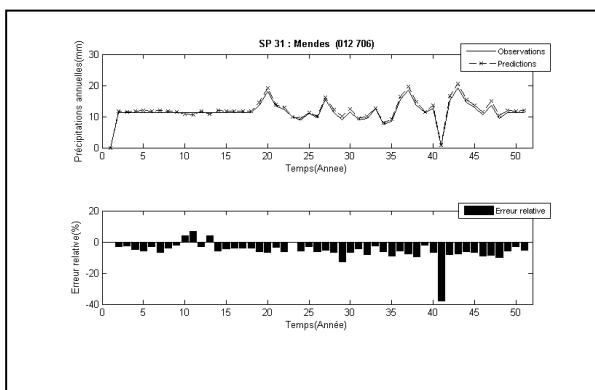
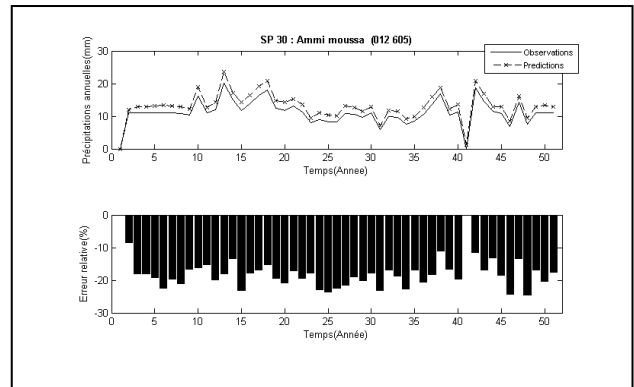
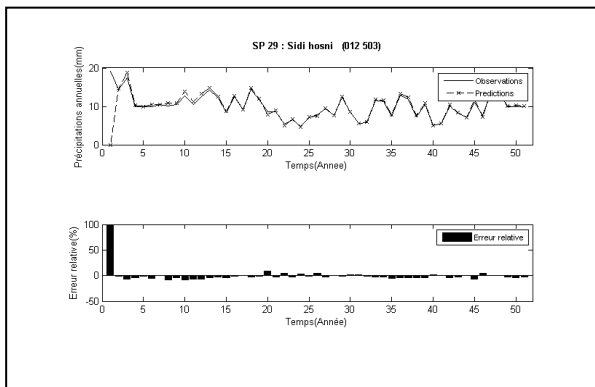
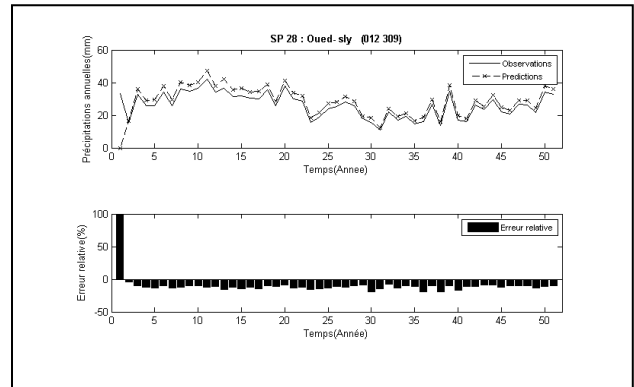
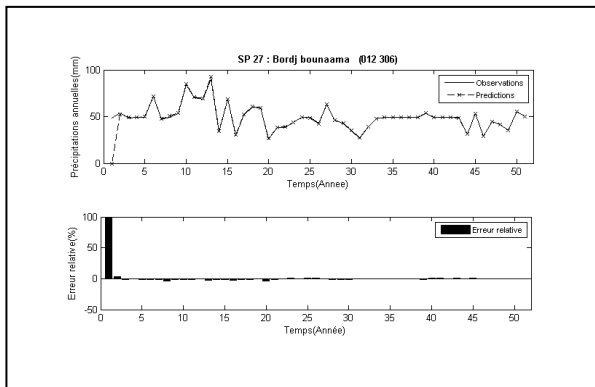
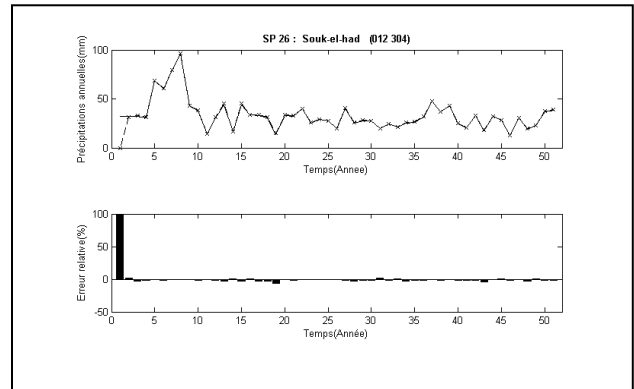
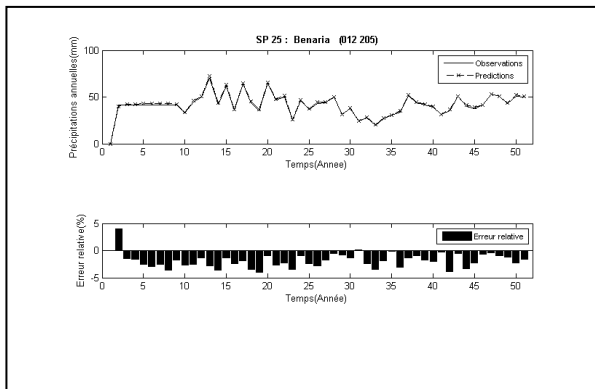
Prédiction Temporelle : Pas Annuel



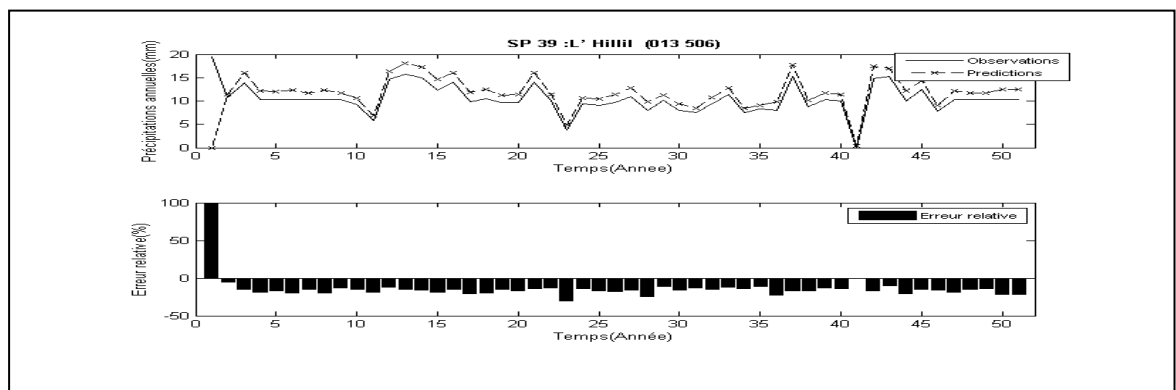
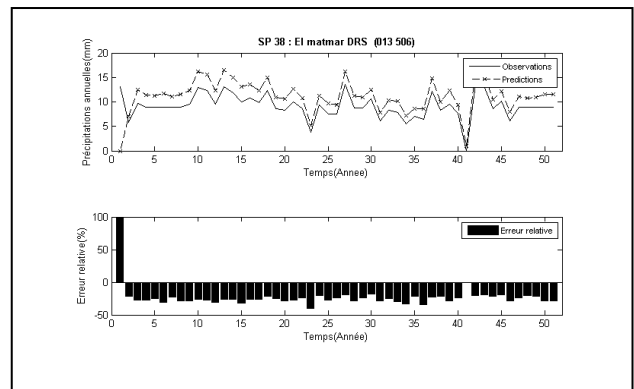
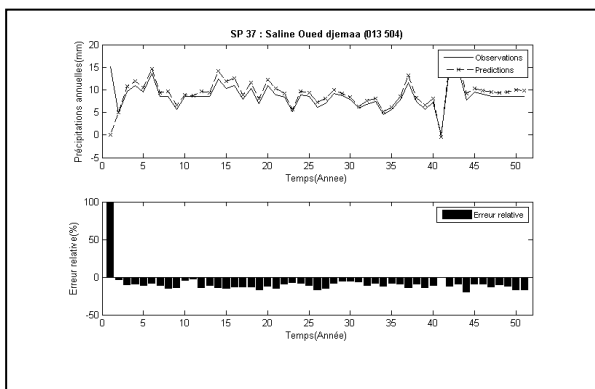
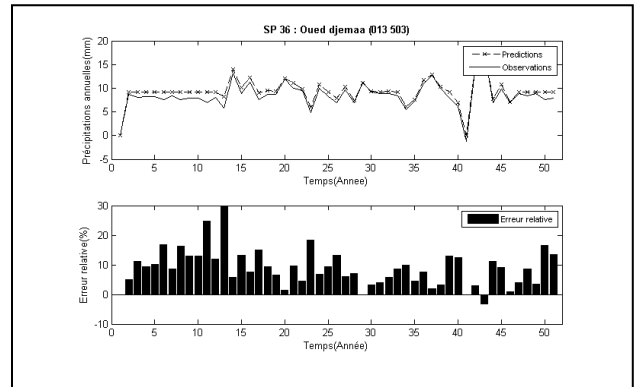
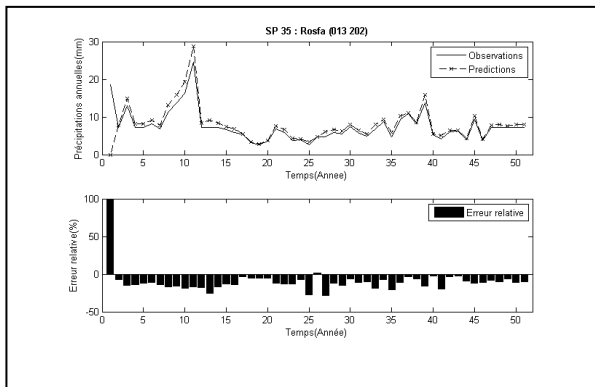
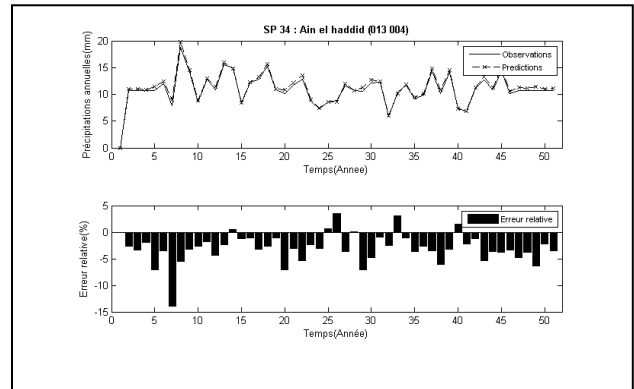
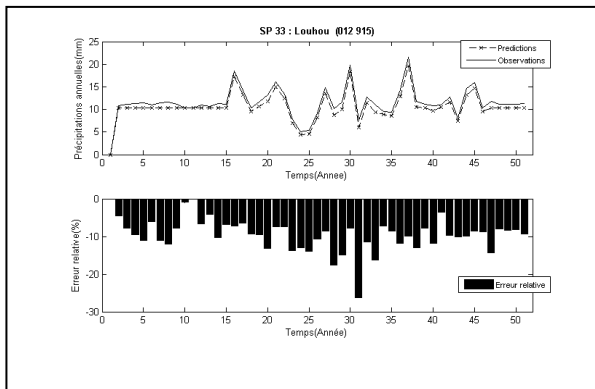
Prédiction Temporelle : Pas Annuel



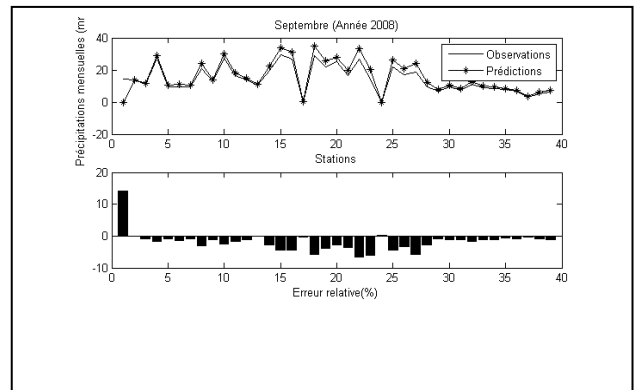
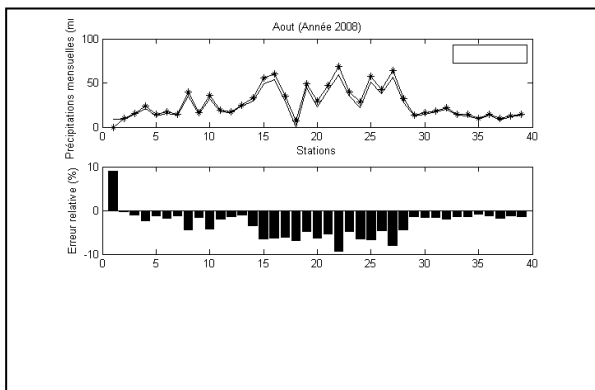
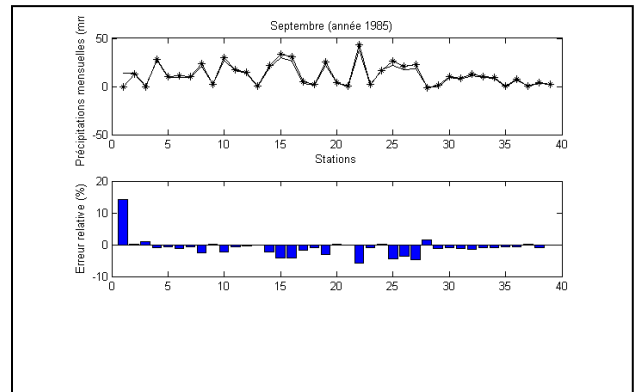
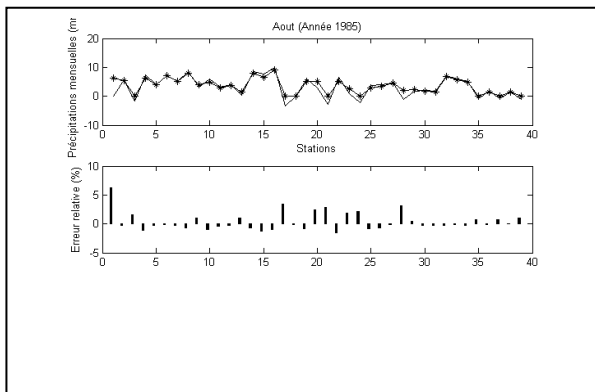
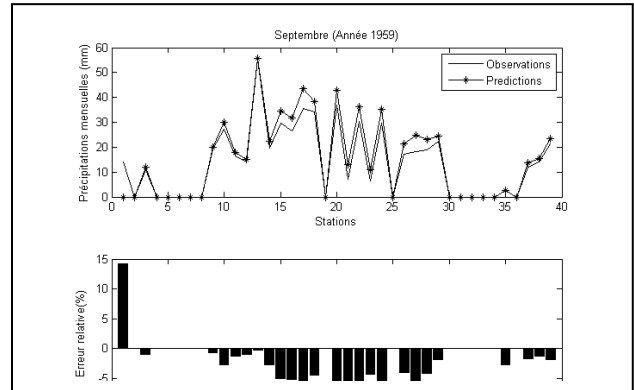
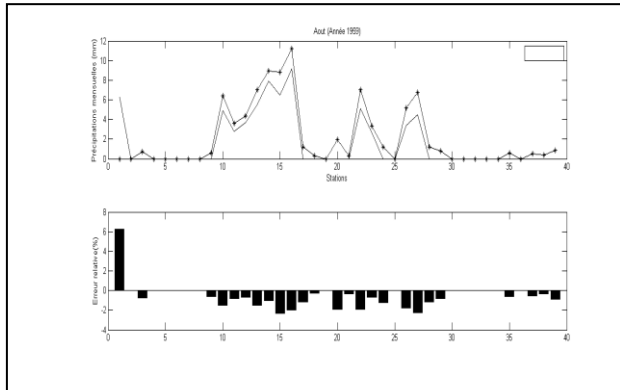
Prédiction Temporelle : Pas Annuel

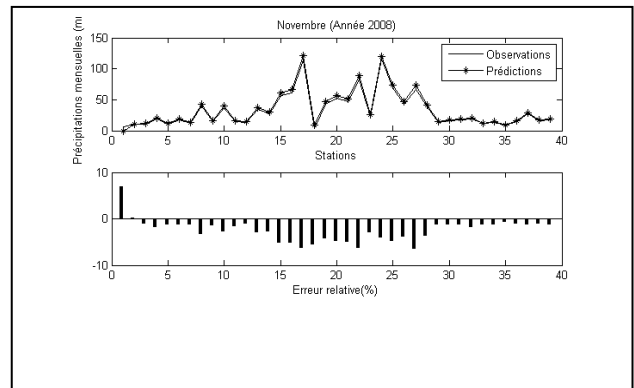
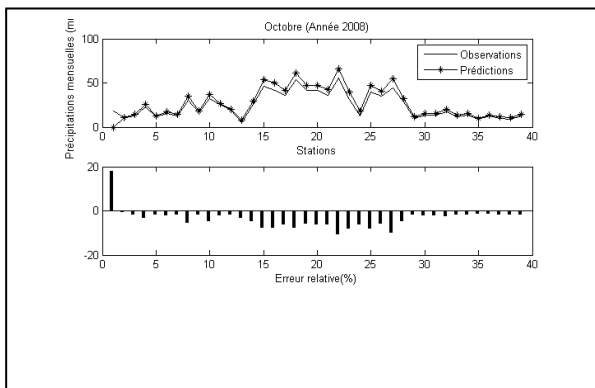
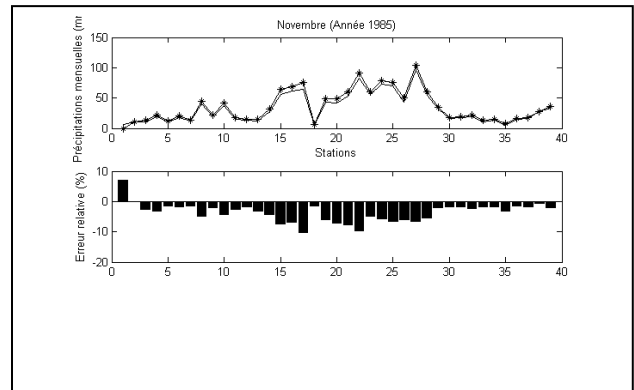
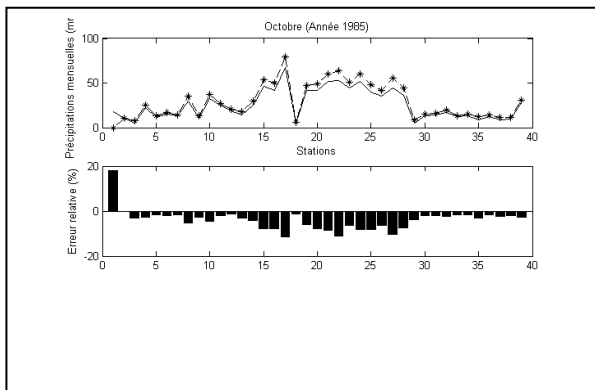
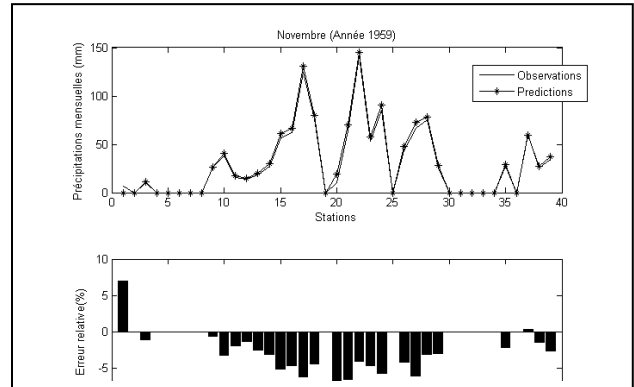
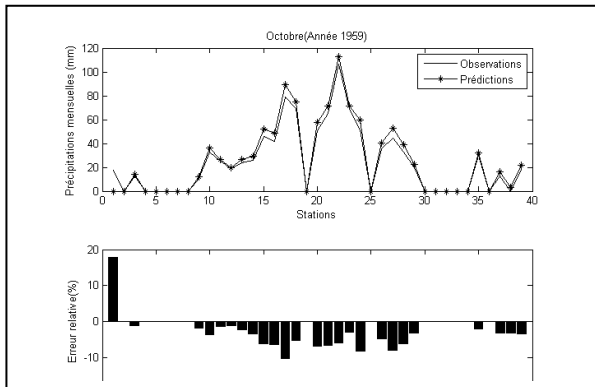


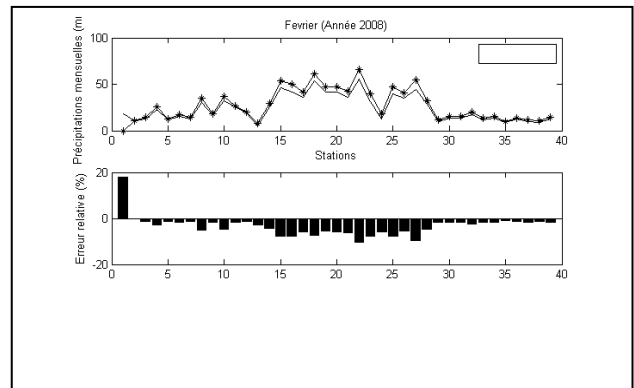
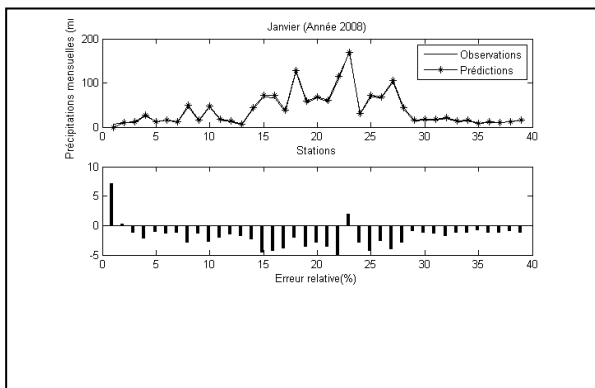
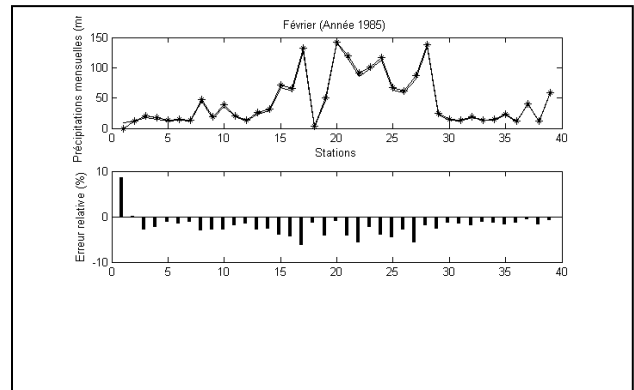
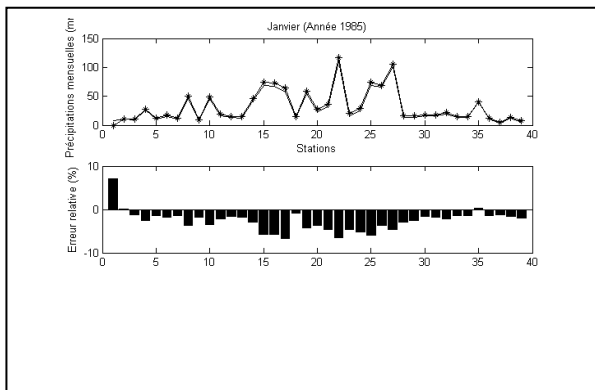
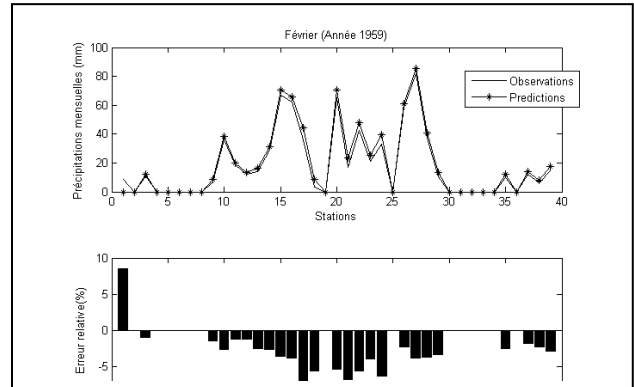
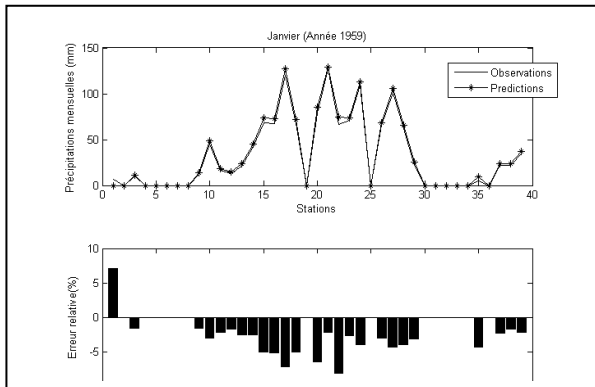
Prédiction Temporelle : Pas Annuel

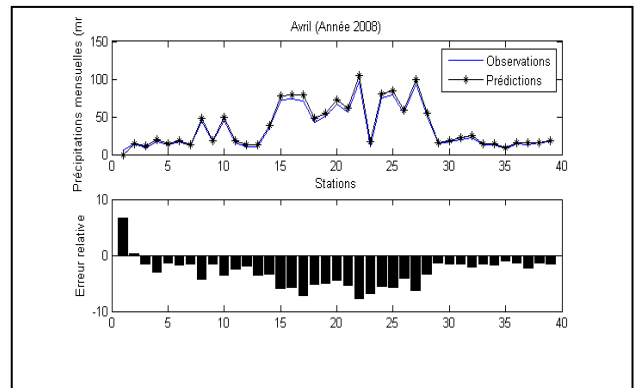
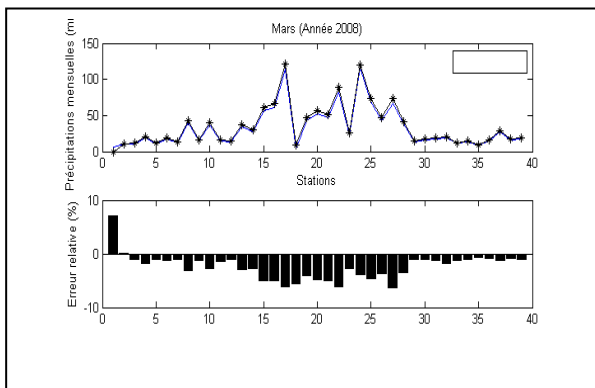
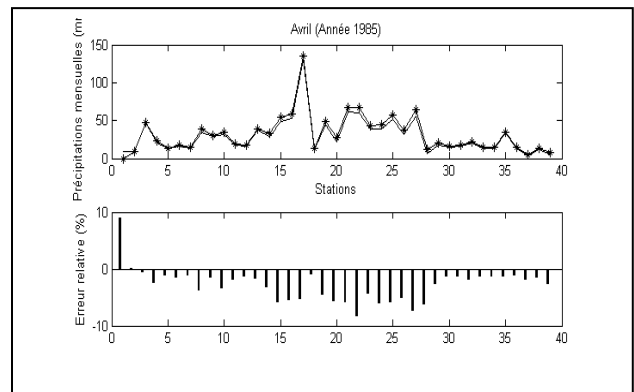
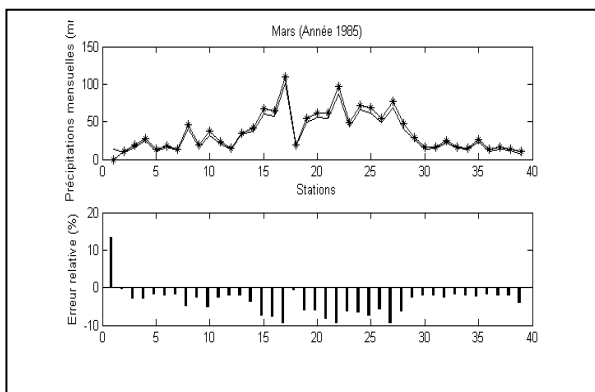
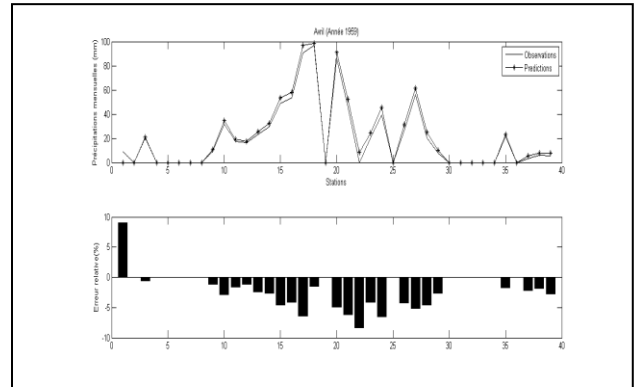
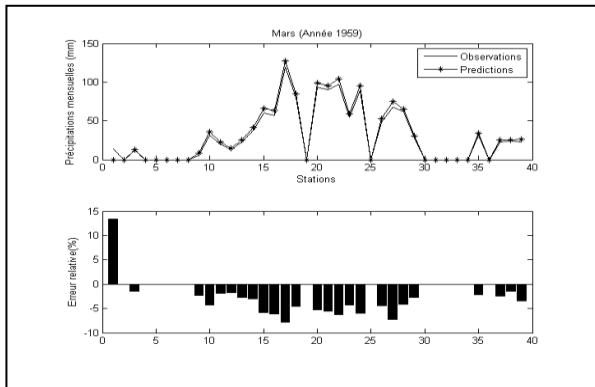


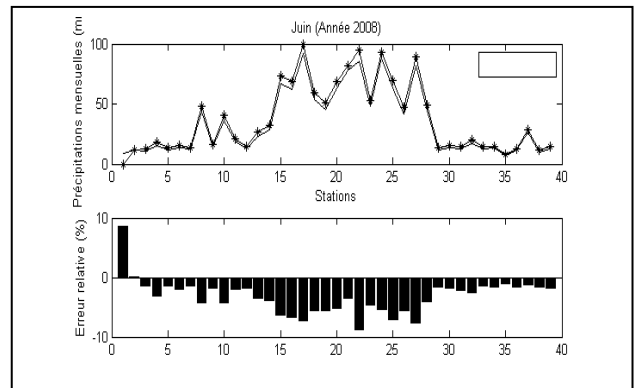
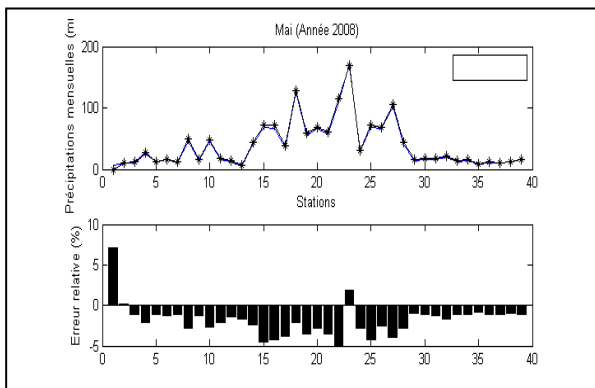
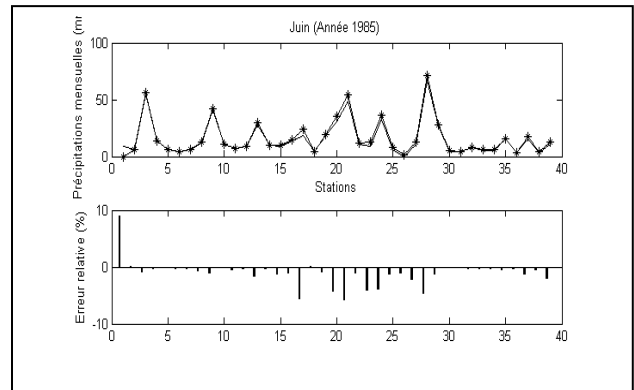
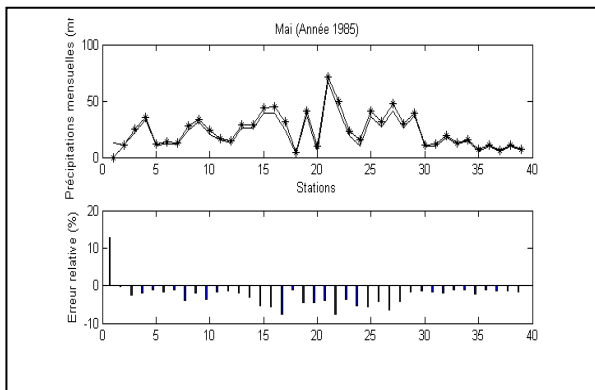
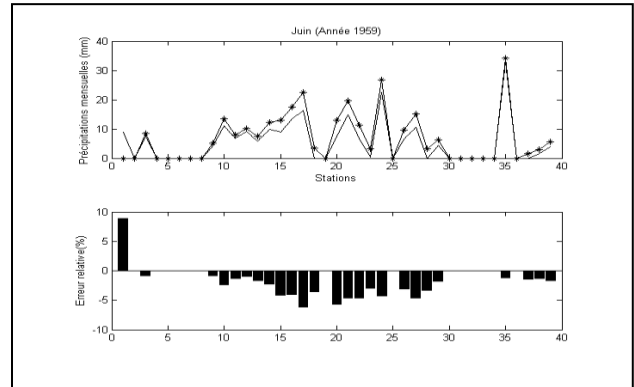
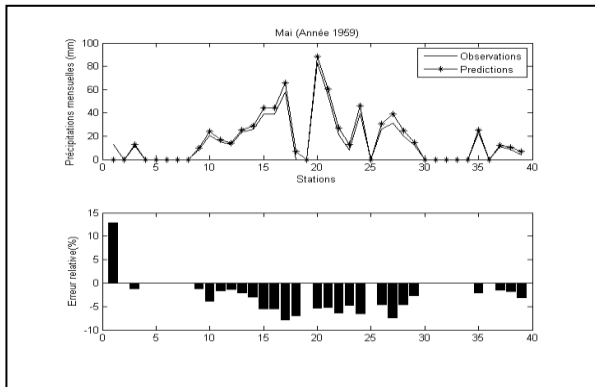
Année / mois	Aout	Sept	oct	nov	dec	jan	fevr	mar	avril	Mai	juin	juillet
1959				
1985		
2008		

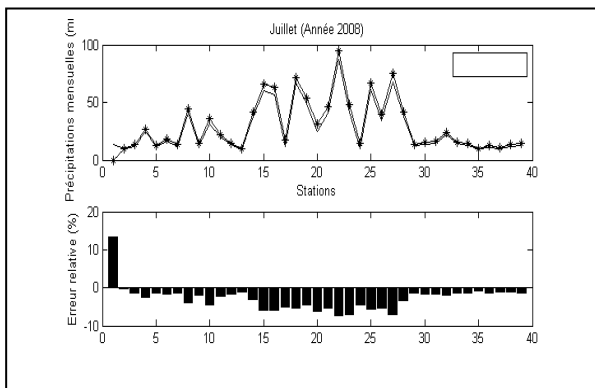
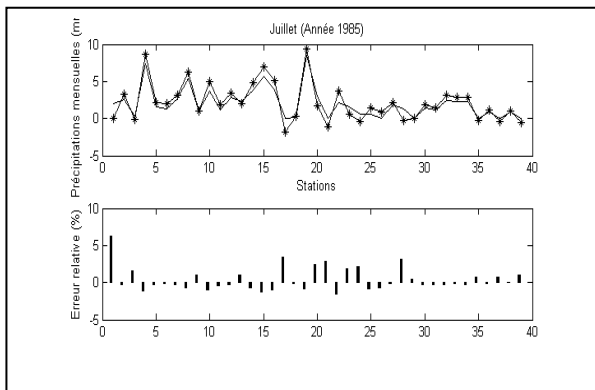
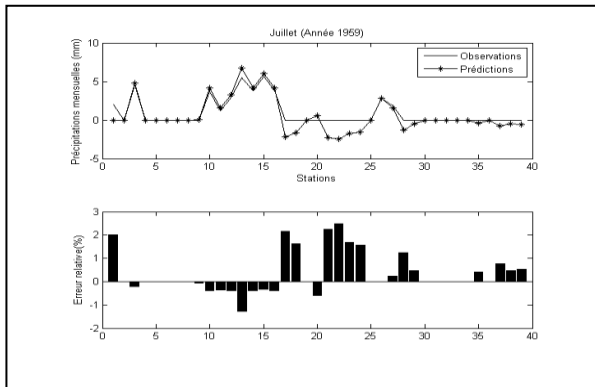


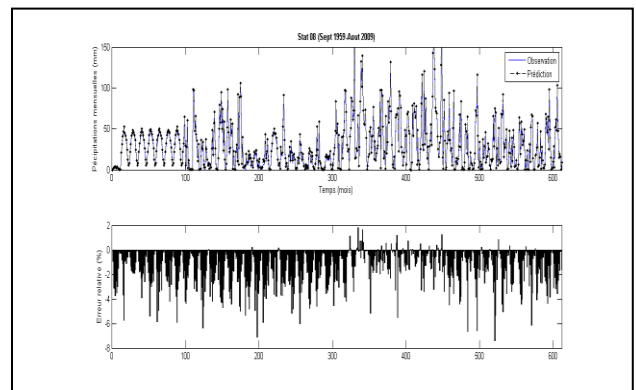
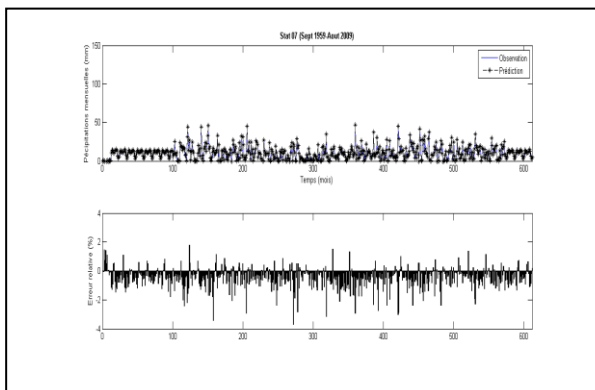
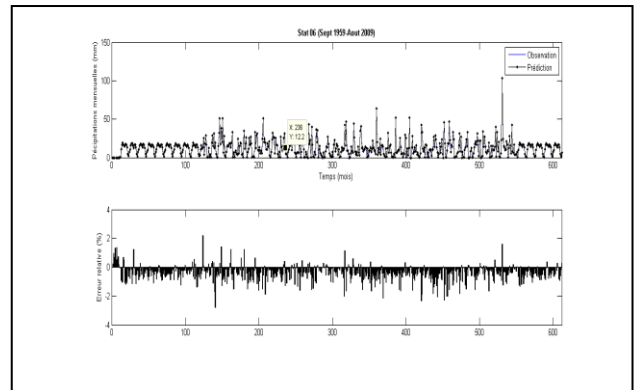
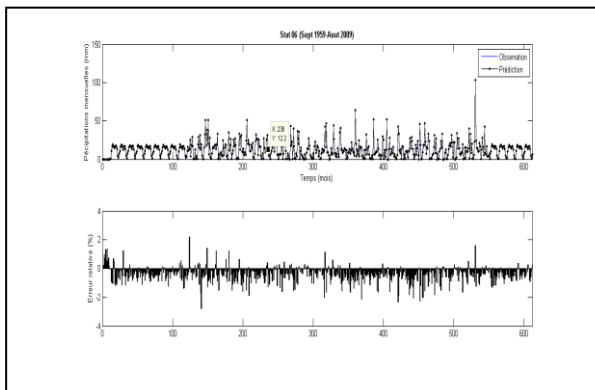
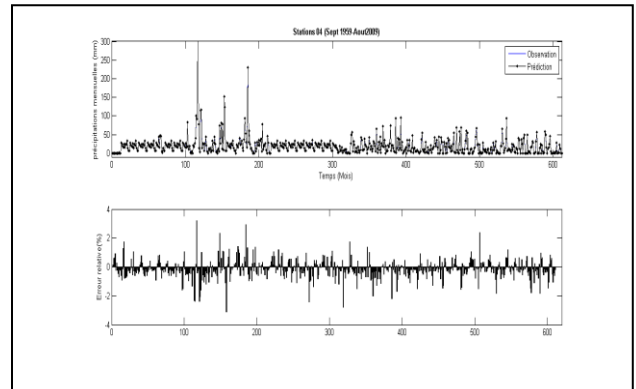
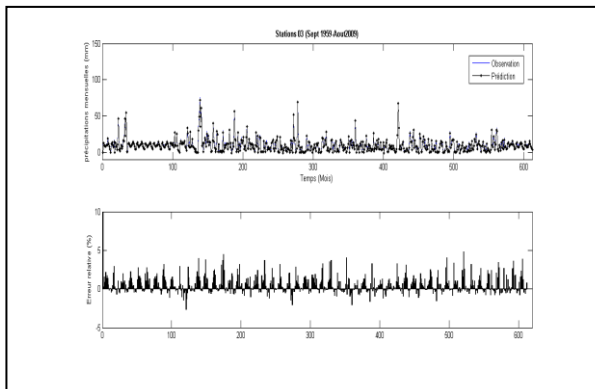
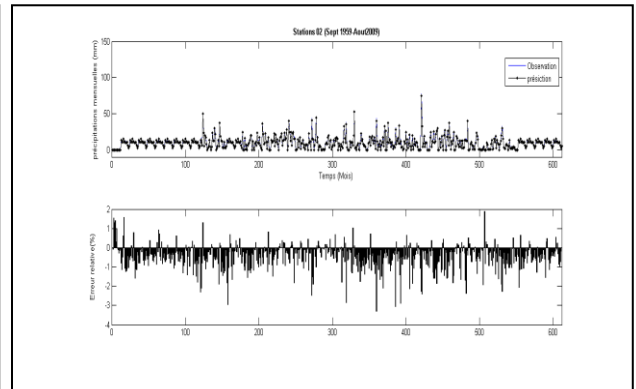
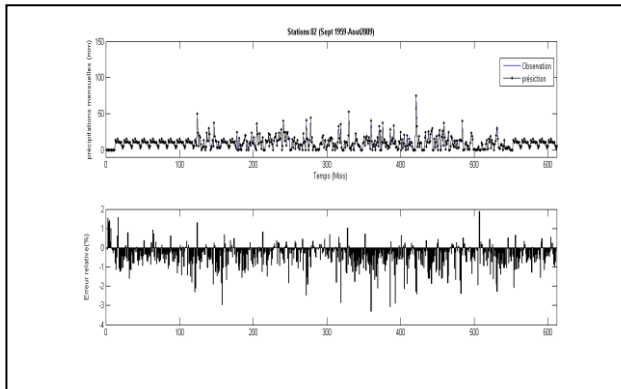


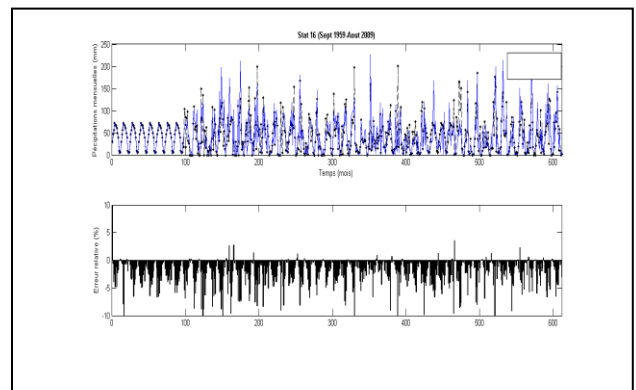
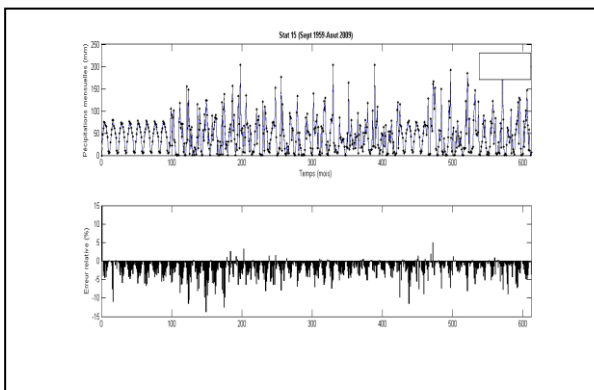
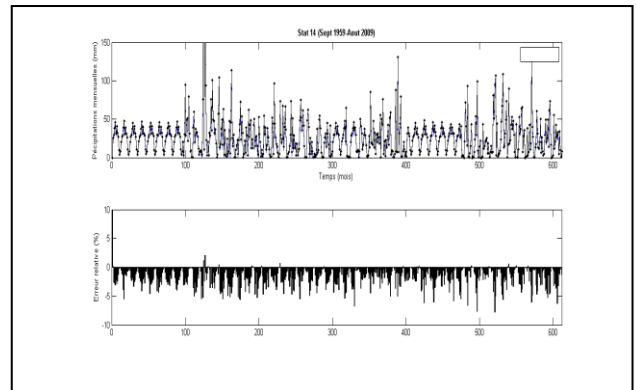
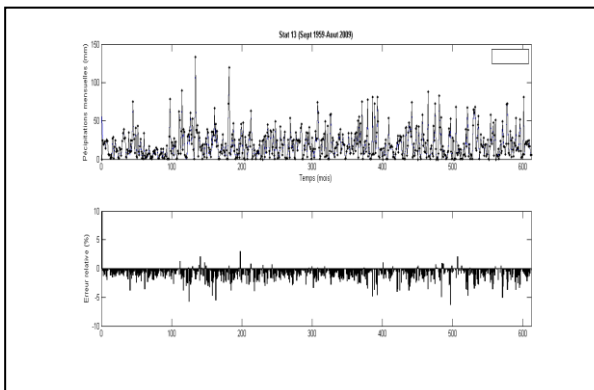
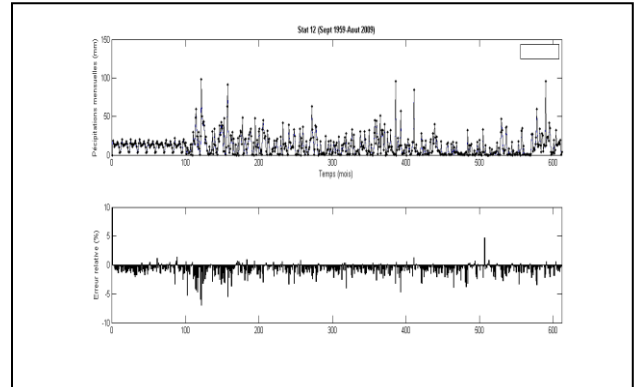
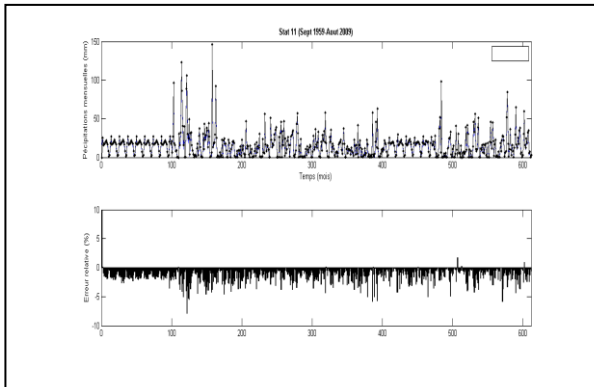
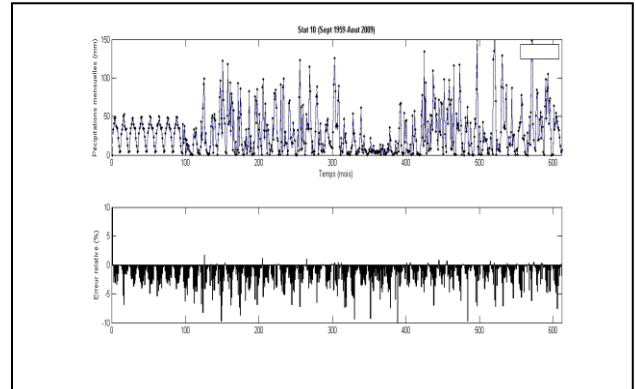
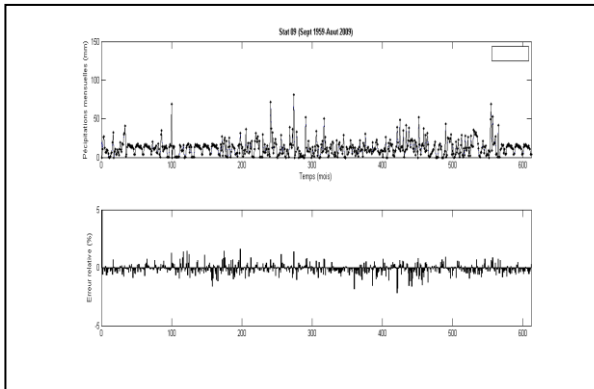


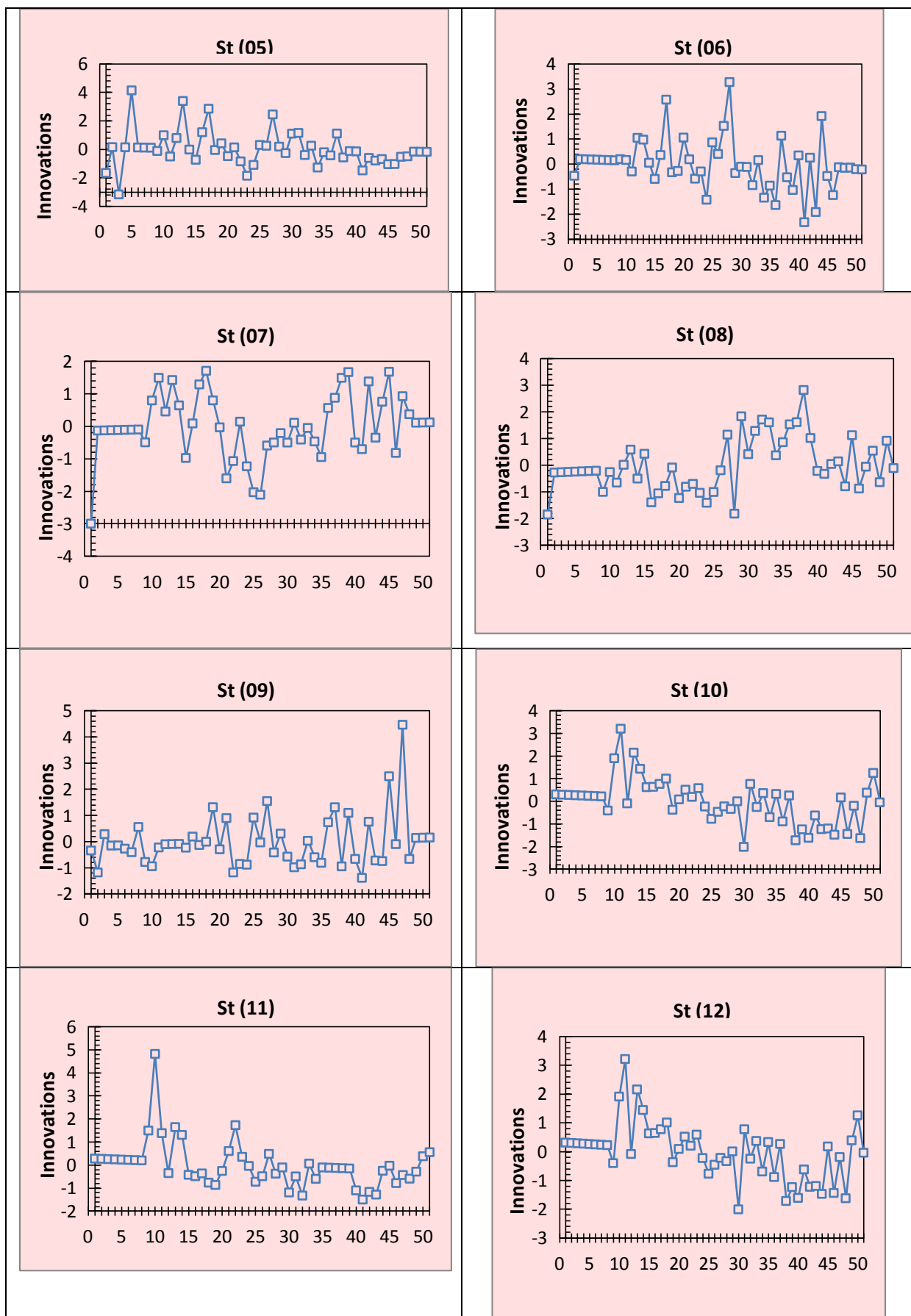


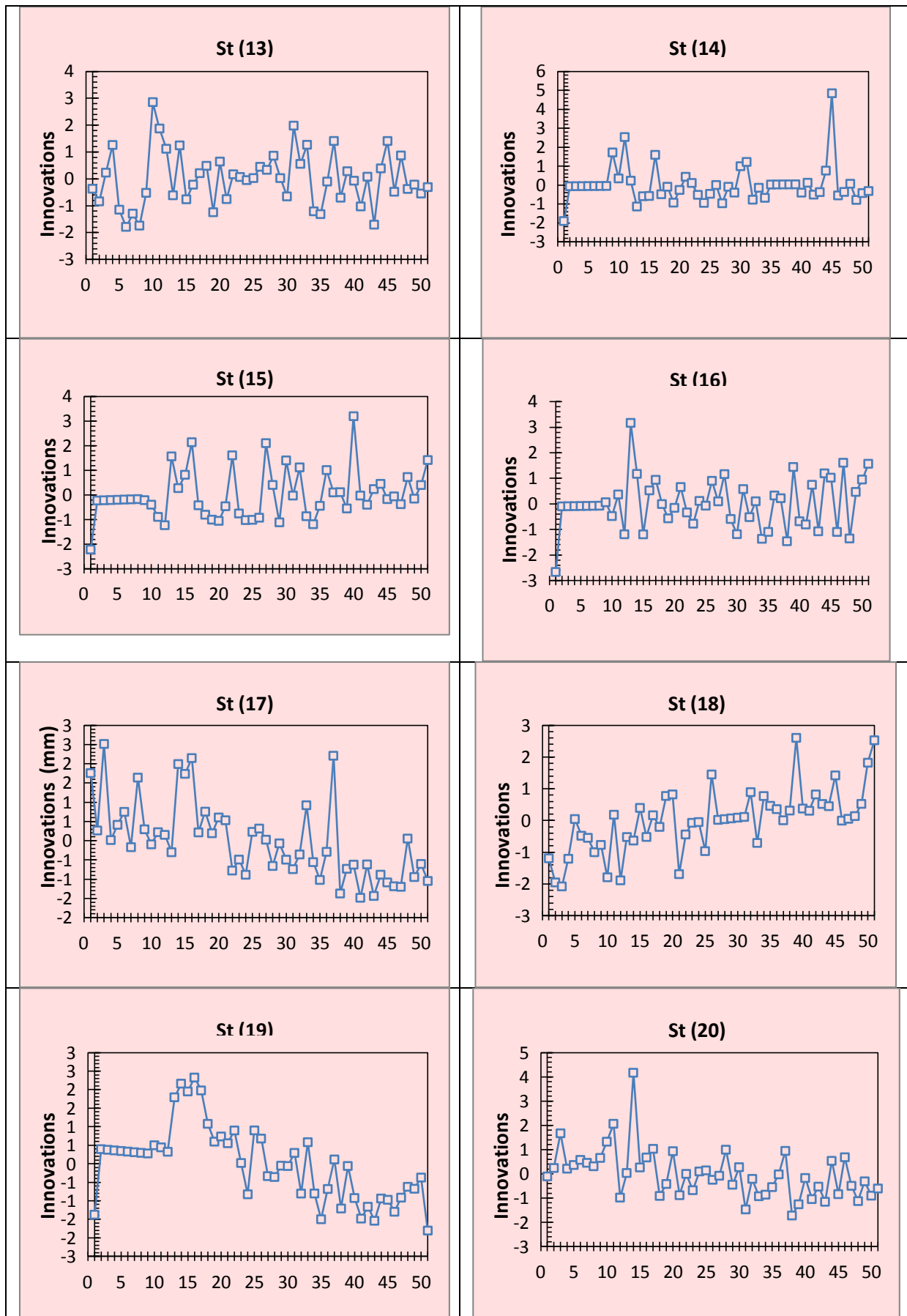


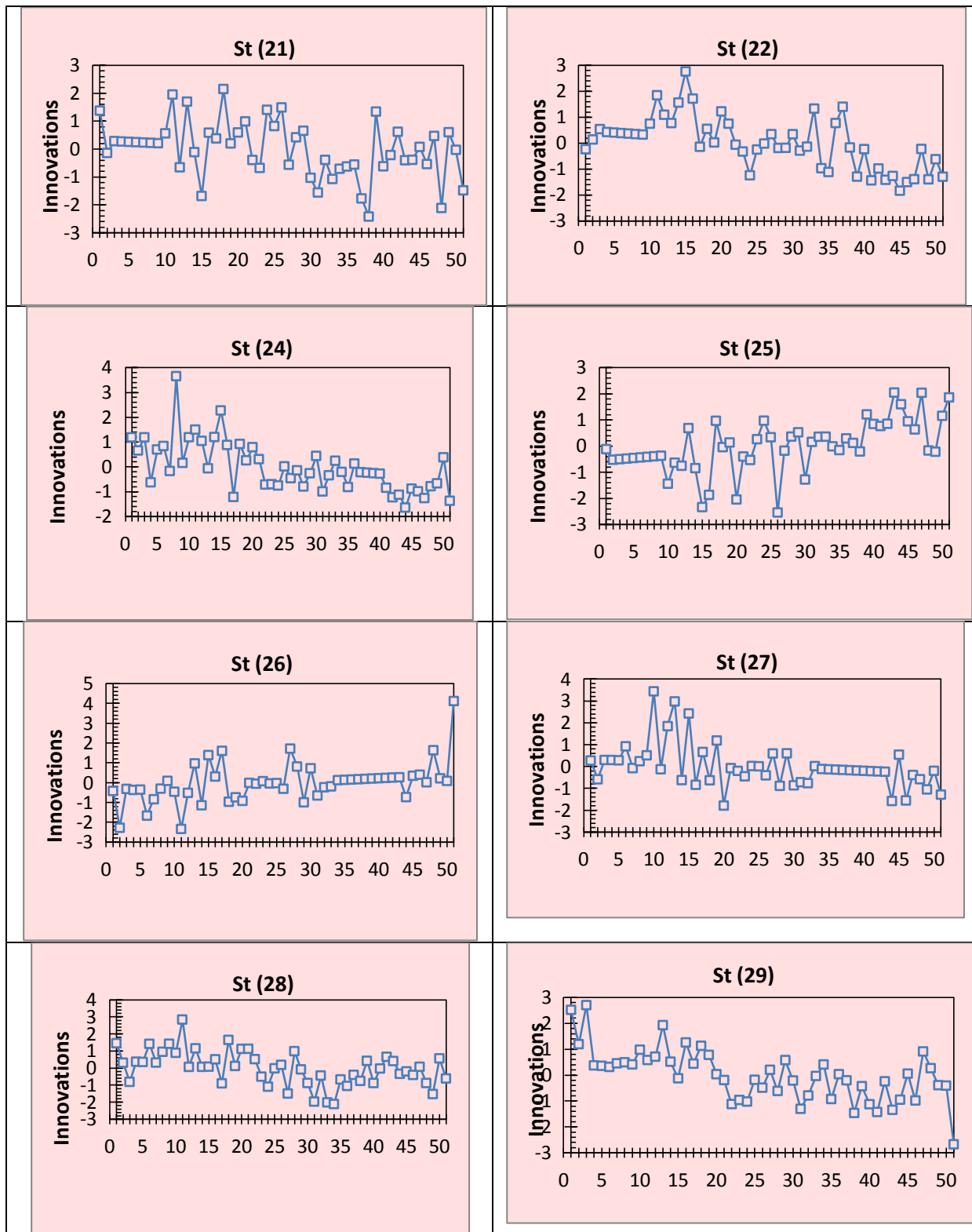


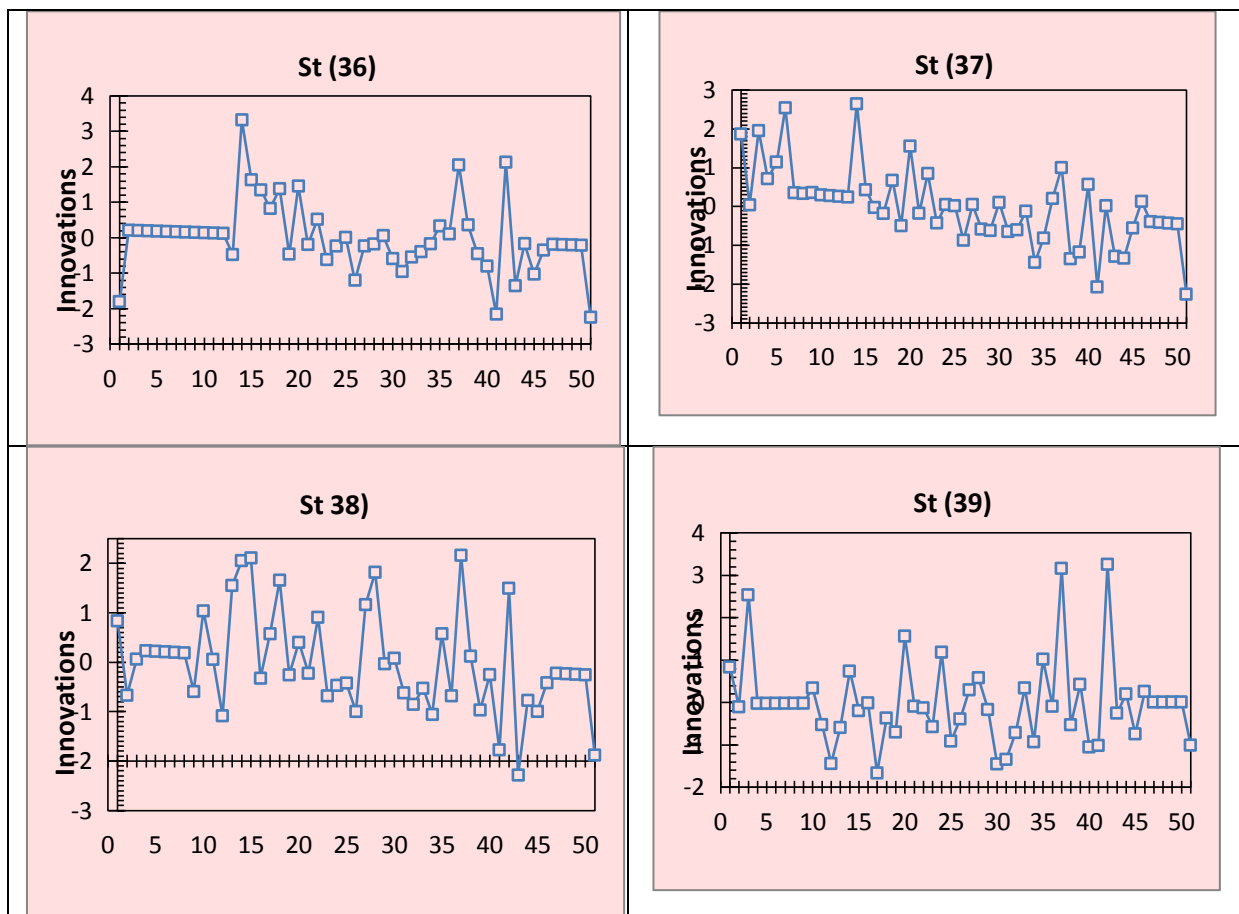


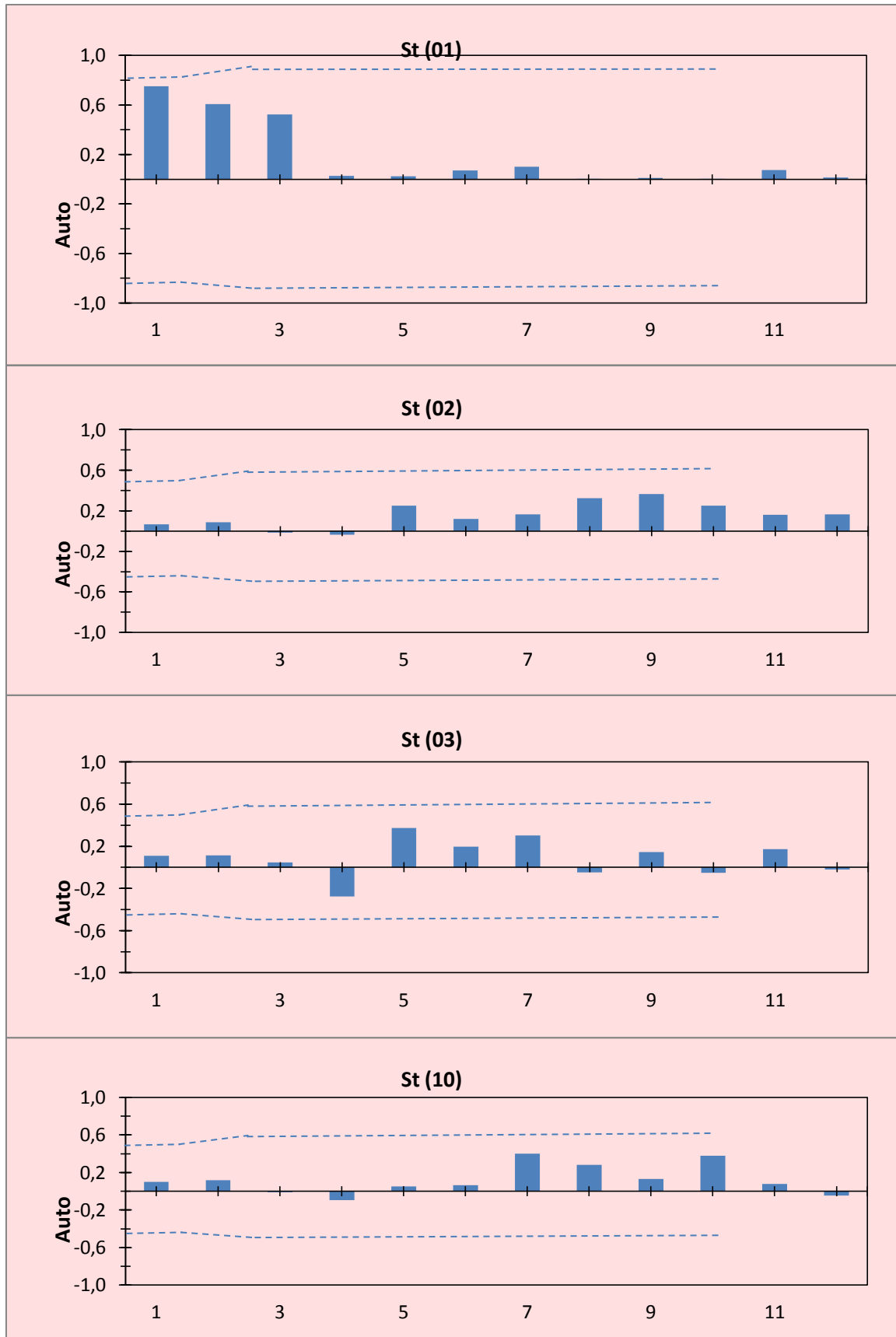


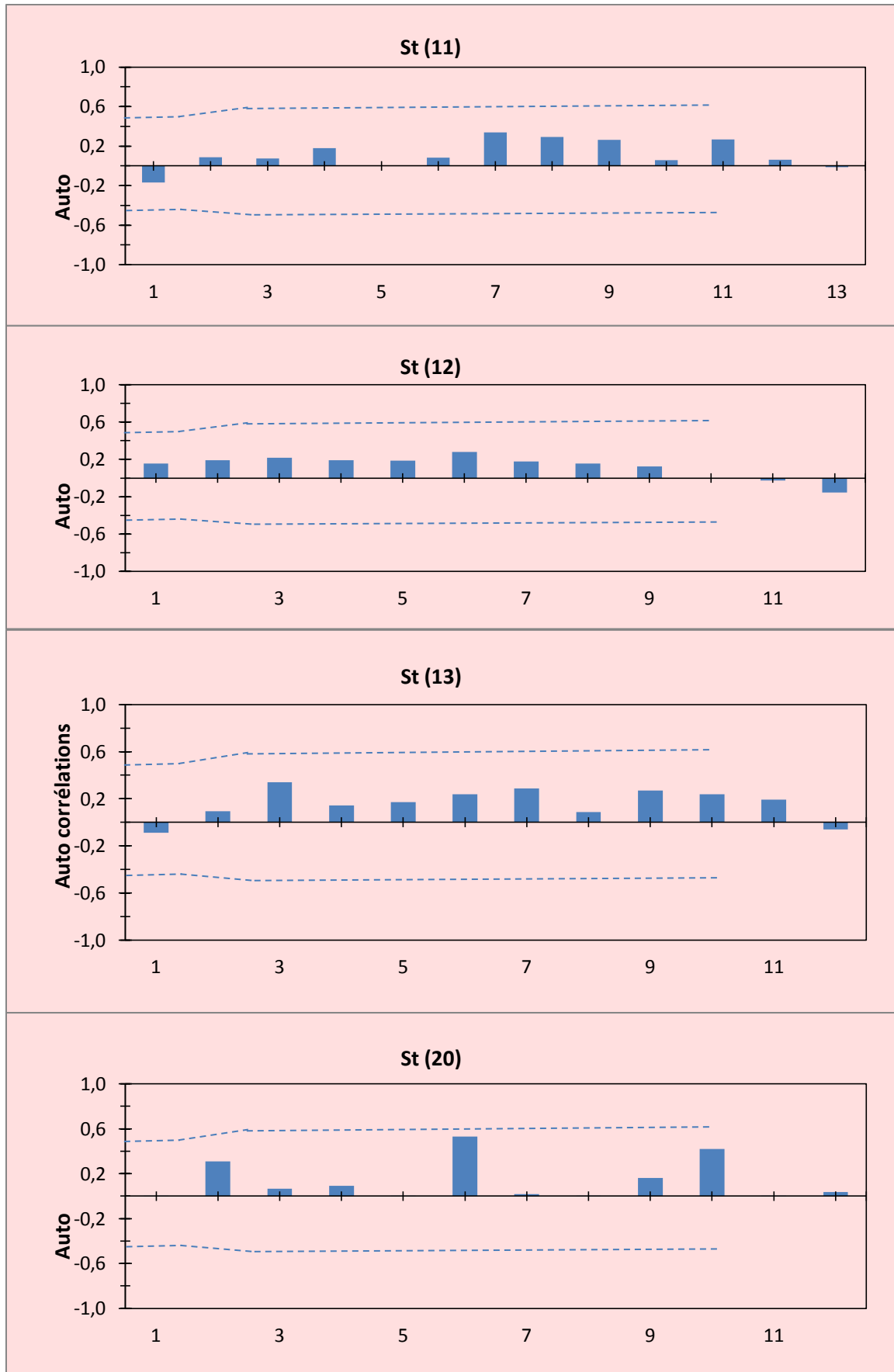


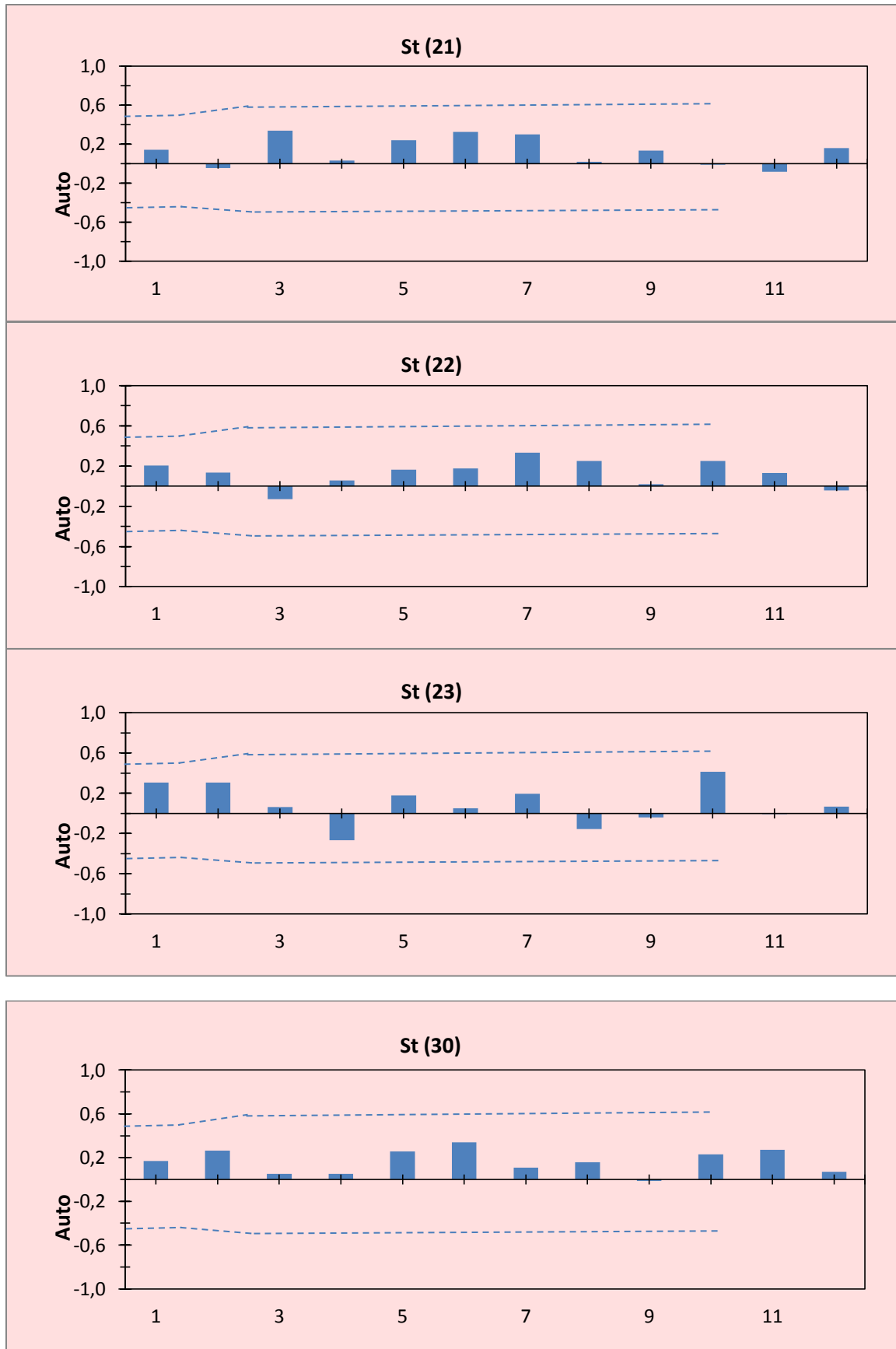


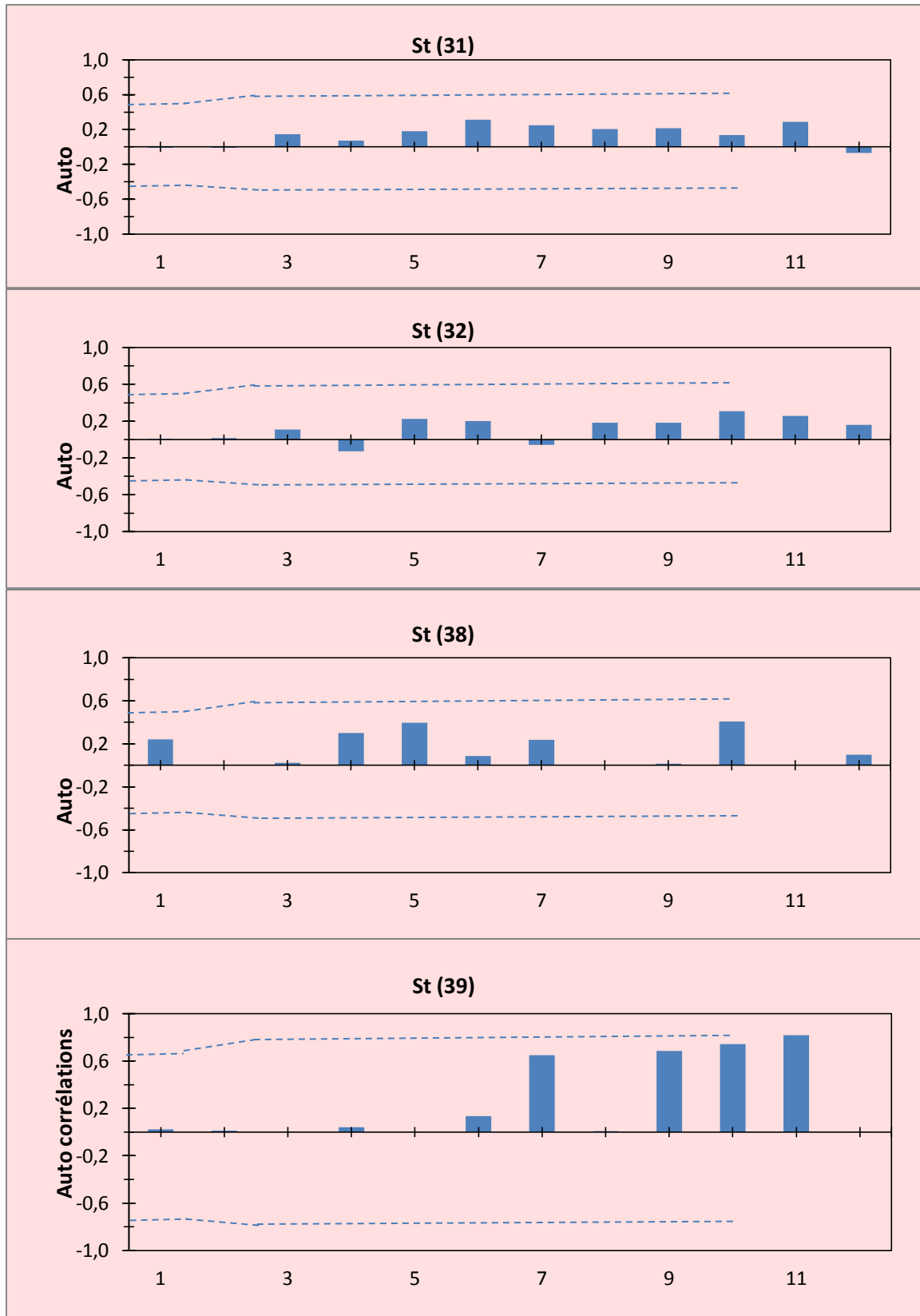


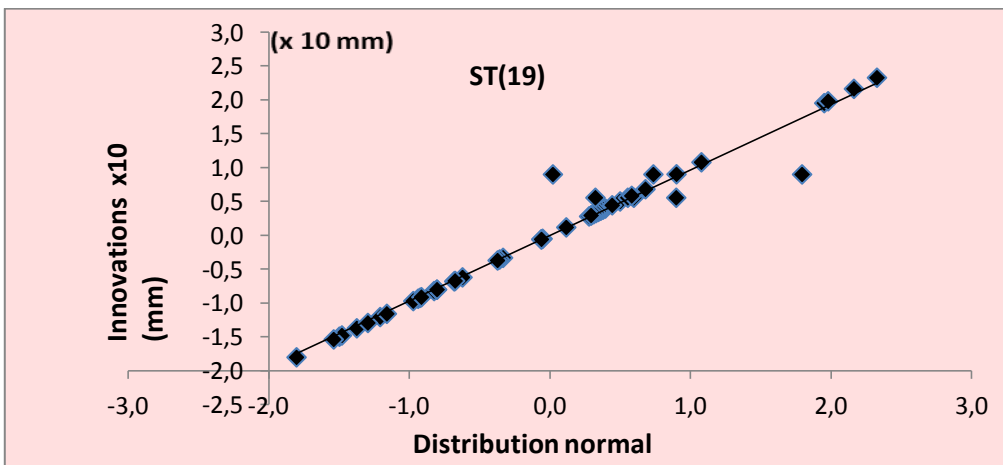
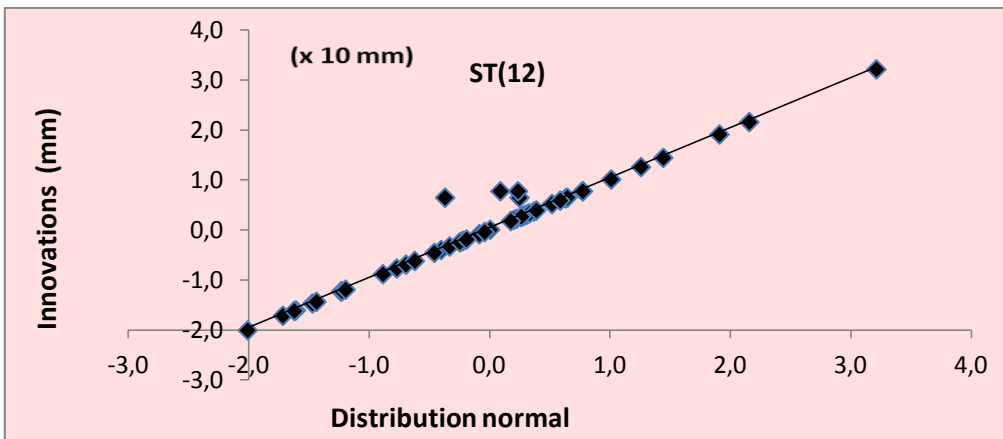
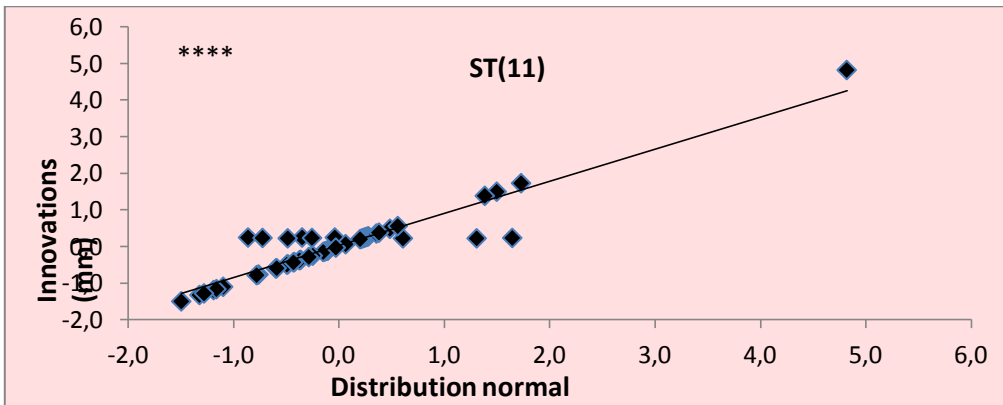
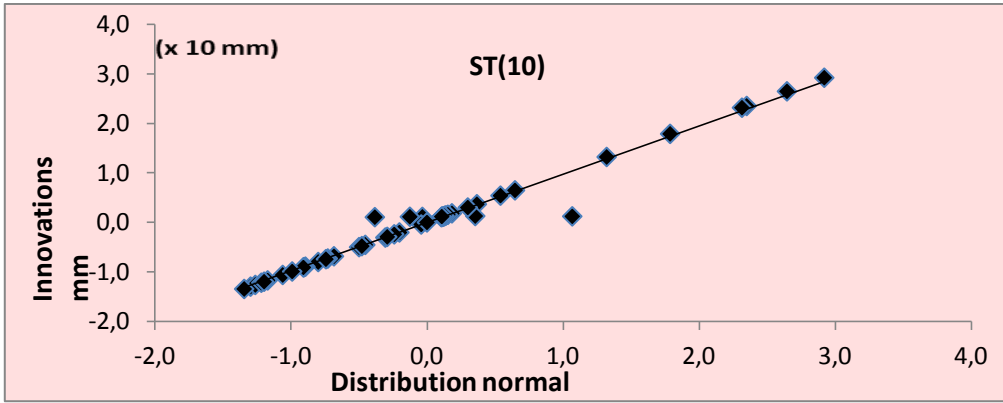


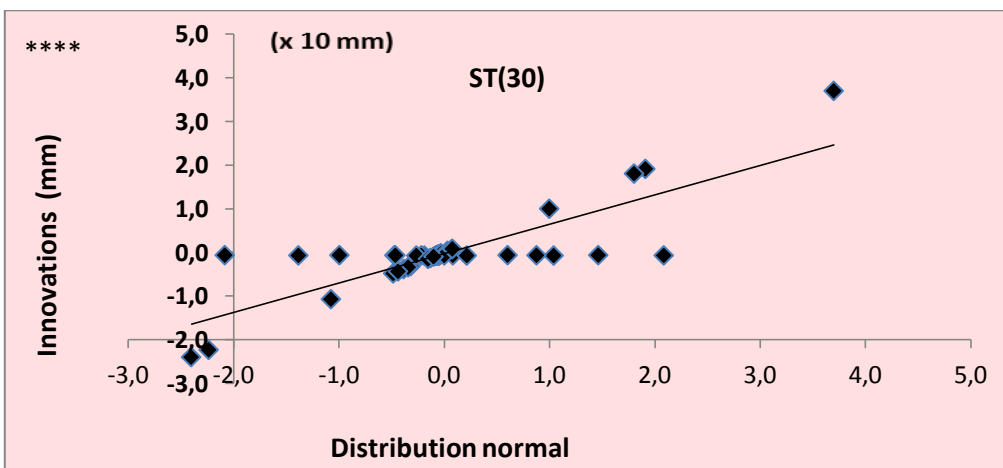
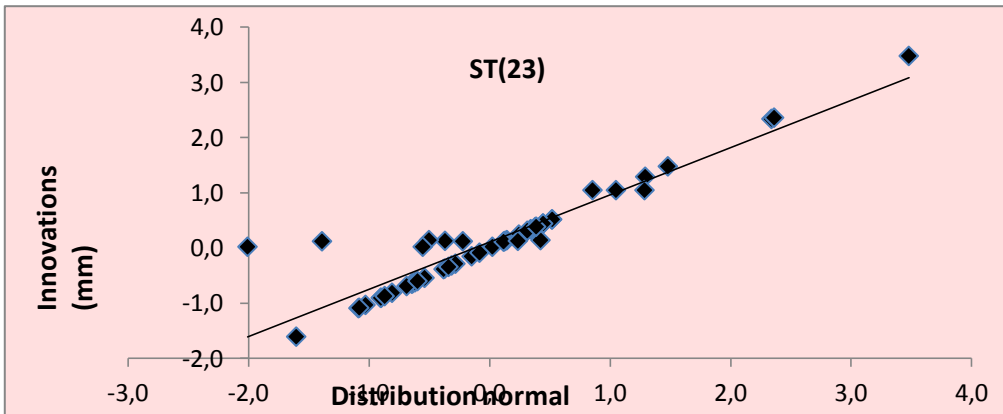
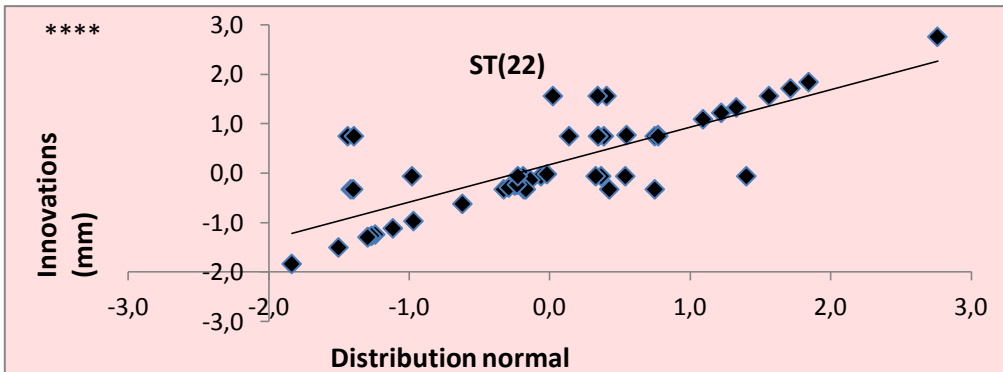
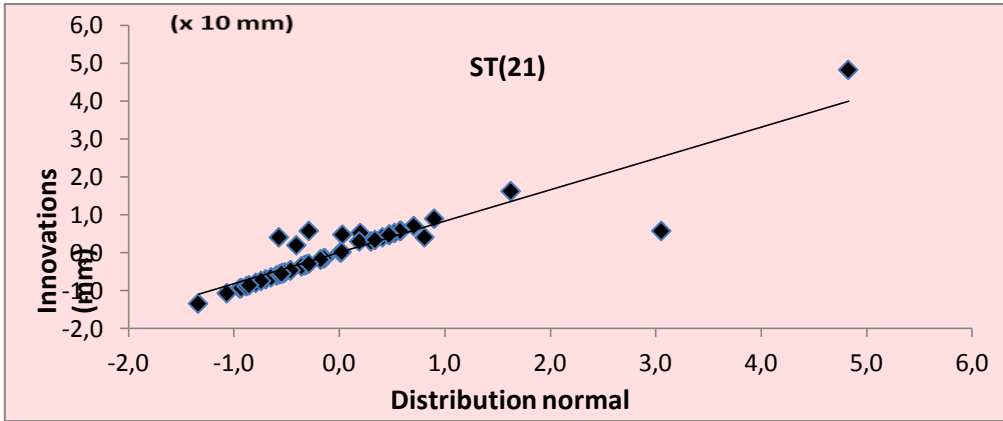


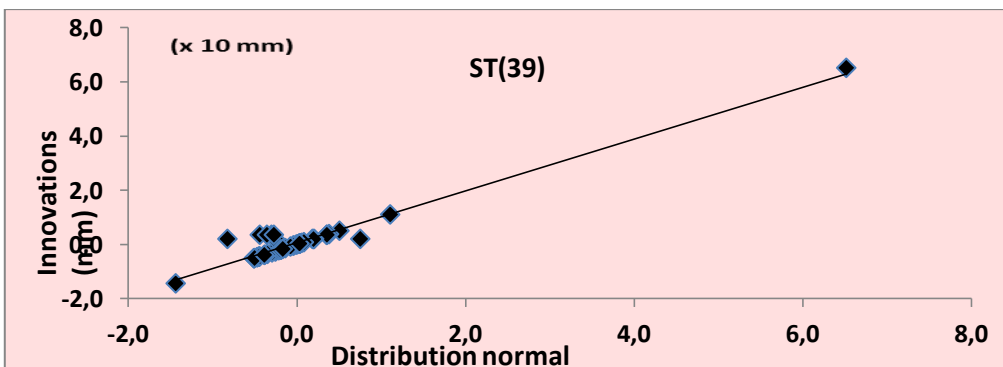
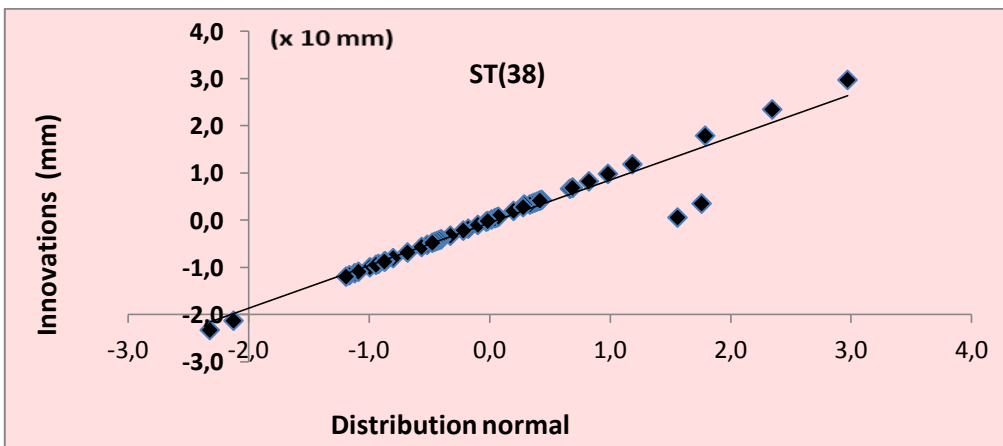
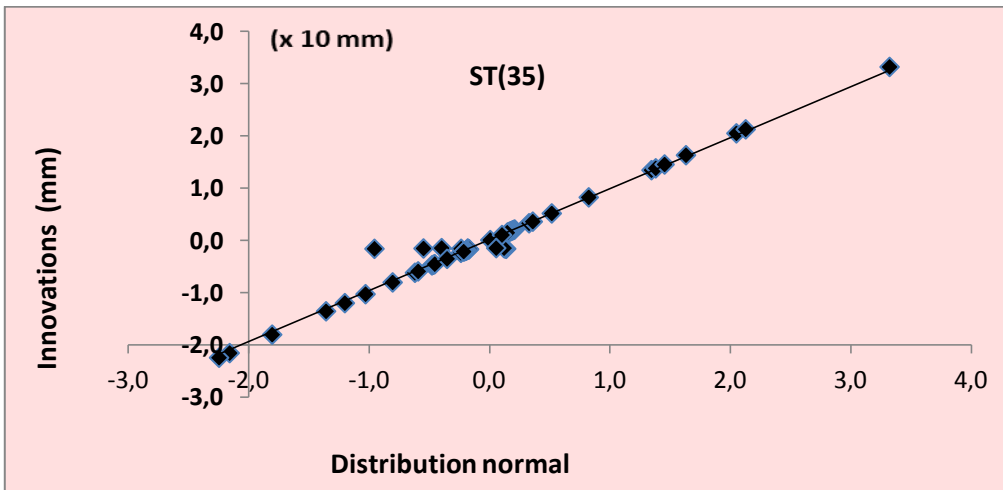
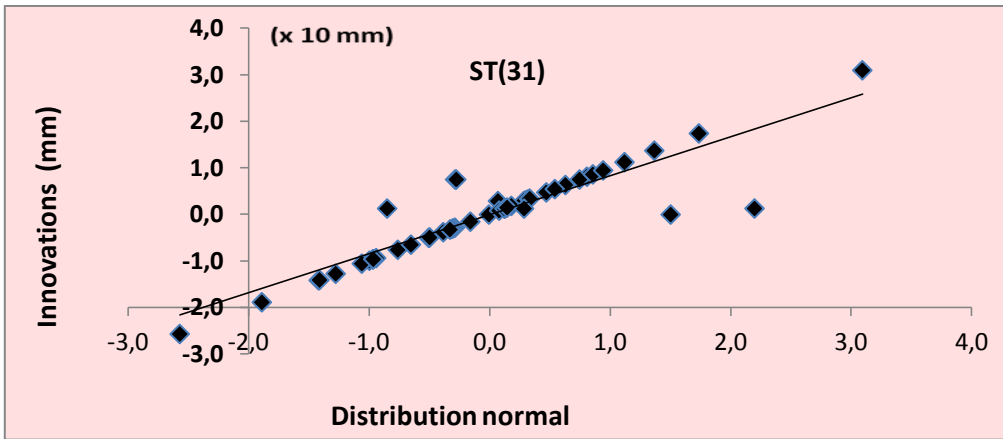


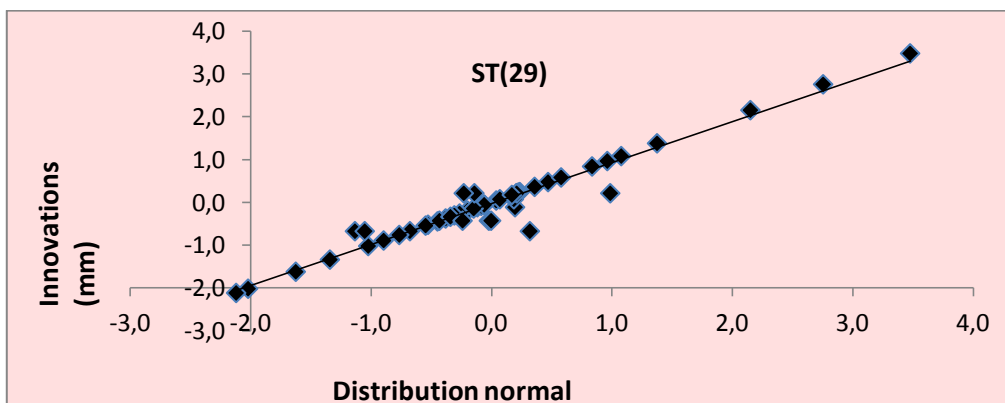
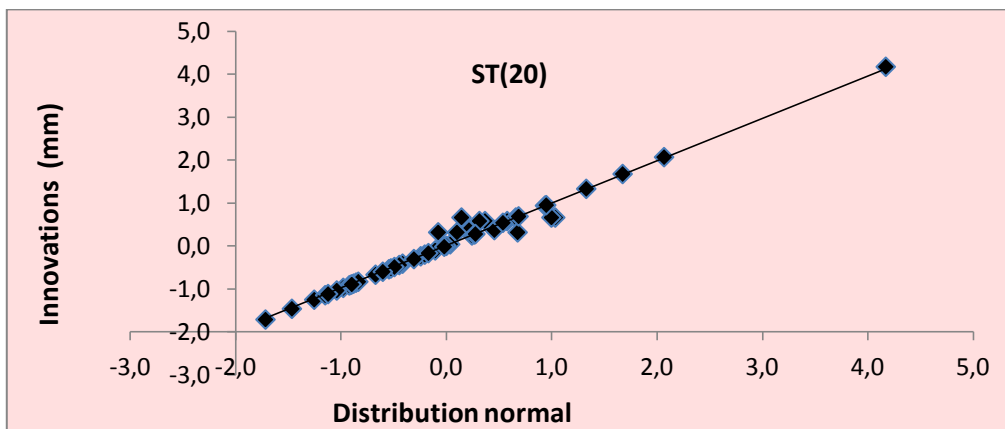
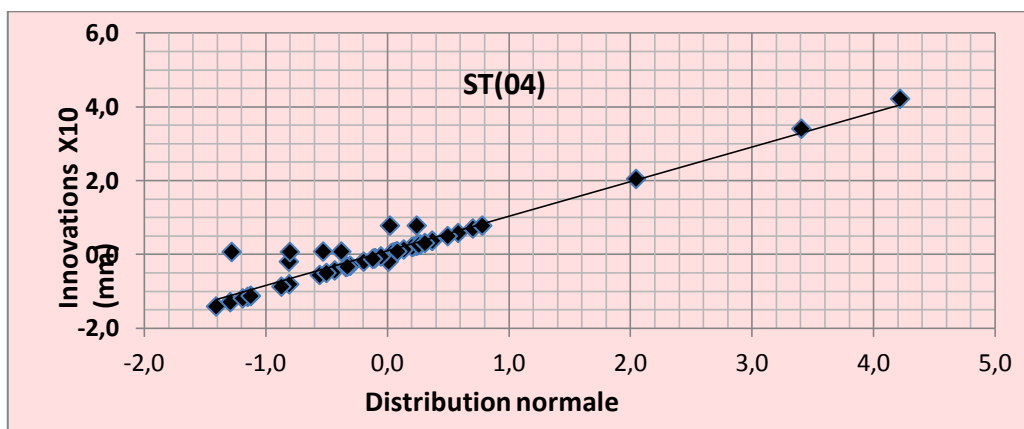
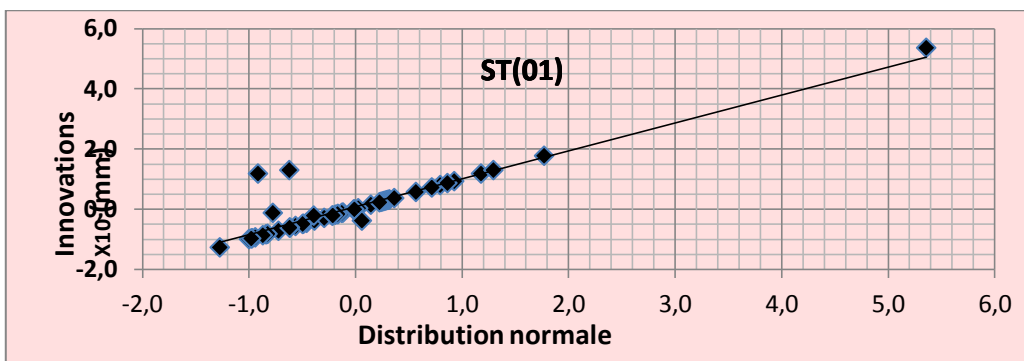


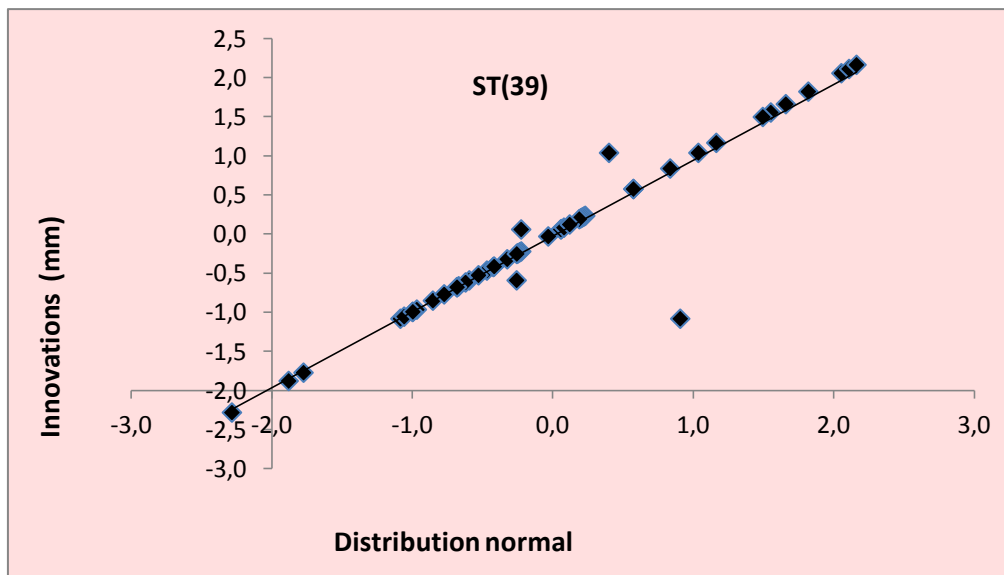
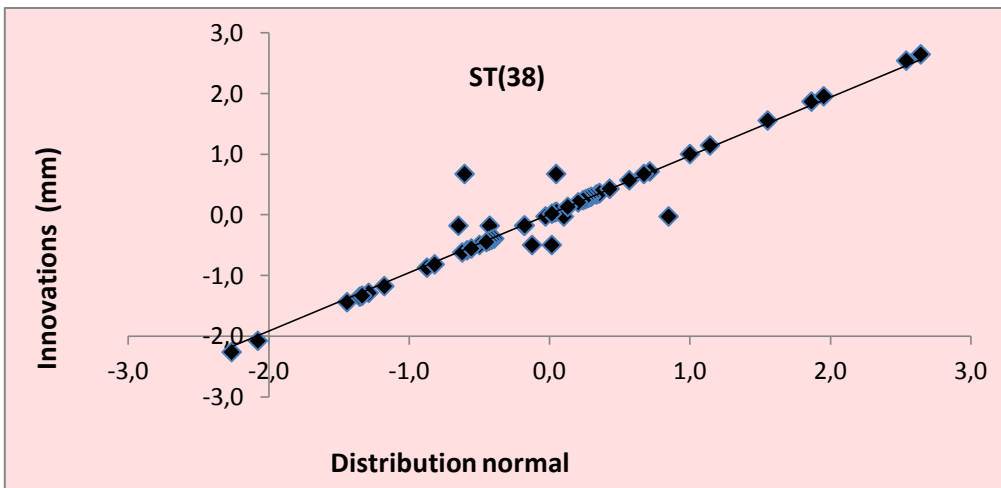
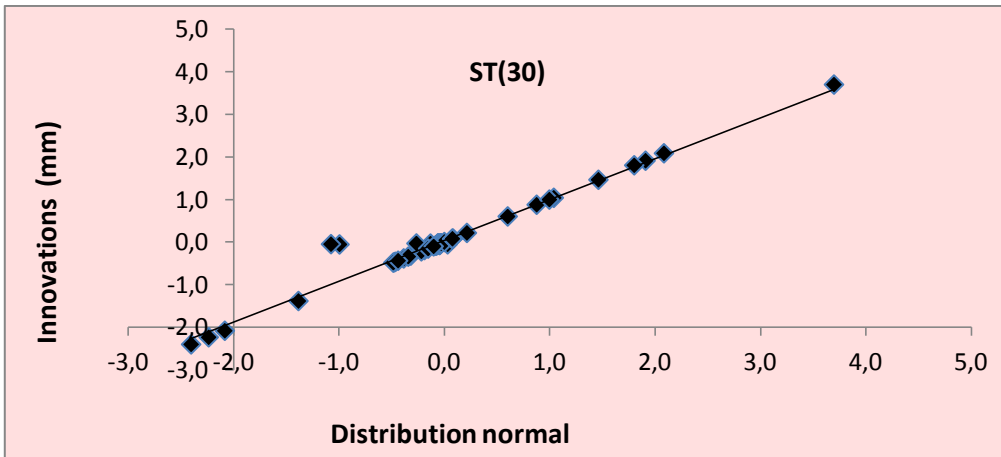












% PROGRAMME FK: This program performs discrete Kalman filter (Variations temporelles et Spatiales)
By Samra Harkat and Malika Boukharouba (2013)



```
PHI=corrcoeff(D);
P=cov(D);
[XV,KG,PT] = KalmanShort(0.39,51,PHI,P);
Pr=XV;
% Pas Temporelle: pour i =1:39 (stations)
39figure.subplot(2,1,1),plot(0:(t:51:4)),hold on,plot(Pr(t:51:4));xlabel('Temps(Année)');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(t:51:4)),xlabel('Temps(Année)');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(t:51:5)),hold on,plot(Pr(t:51:5));xlabel('Temps(Année)');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(t:51:5)),xlabel('Temps(Année)');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(t:51:6)),hold on,plot(Pr(t:51:6));xlabel('Temps(Année)');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(t:51:6)),xlabel('Temps(Année)');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(t:51:7)),hold on,plot(Pr(t:51:7));xlabel('Temps(Année)');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(t:51:7)),xlabel('Temps(Année)');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(t:51:8)),hold on,plot(Pr(t:51:8));xlabel('Temps(Année)');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(t:51:8)),xlabel('Temps(Année)');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(t:51:9)),hold on,plot(Pr(t:51:9));xlabel('Temps(Année)');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(t:51:9)),xlabel('Temps(Année)');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(t:51:10)),hold on,plot(Pr(t:51:10));xlabel('Temps(Année)');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(t:51:10)),xlabel('Temps(Année)');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(t:51:11)),hold on,plot(Pr(t:51:11));xlabel('Temps(Année)');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(t:51:11)),xlabel('Temps(Année)');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(t:51:12)),hold on,plot(Pr(t:51:12));xlabel('Temps(Année)');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(t:51:12)),xlabel('Temps(Année)');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(t:51:13)),hold on,plot(Pr(t:51:13));xlabel('Temps(Année)');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(t:51:13)),xlabel('Temps(Année)');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(t:51:14)),hold on,plot(Pr(t:51:14));xlabel('Temps(Année)');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(t:51:14)),xlabel('Temps(Année)');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(t:51:15)),hold on,plot(Pr(t:51:15));xlabel('Temps(Année)');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(t:51:15)),xlabel('Temps(Année)');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(t:51:16)),hold on,plot(Pr(t:51:16));xlabel('Temps(Année)');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(t:51:16)),xlabel('Temps(Année)');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(t:51:17)),hold on,plot(Pr(t:51:17));xlabel('Temps(Année)');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(t:51:17)),xlabel('Temps(Année)');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(t:51:18)),hold on,plot(Pr(t:51:18));xlabel('Temps(Année)');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(t:51:18)),xlabel('Temps(Année)');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(t:51:19)),hold on,plot(Pr(t:51:19));xlabel('Temps(Année)');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(t:51:19)),xlabel('Temps(Année)');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(t:51:20)),hold on,plot(Pr(t:51:20));xlabel('Temps(Année)');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(t:51:20)),xlabel('Temps(Année)');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(t:51:21)),hold on,plot(Pr(t:51:21));xlabel('Temps(Année)');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(t:51:21)),xlabel('Temps(Année)');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(t:51:22)),hold on,plot(Pr(t:51:22));xlabel('Temps(Année)');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(t:51:22)),xlabel('Temps(Année)');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(t:51:23)),hold on,plot(Pr(t:51:23));xlabel('Temps(Année)');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(t:51:23)),xlabel('Temps(Année)');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(t:51:24)),hold on,plot(Pr(t:51:24));xlabel('Temps(Année)');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(t:51:24)),xlabel('Temps(Année)');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(t:51:25)),hold on,plot(Pr(t:51:25));xlabel('Temps(Année)');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(t:51:25)),xlabel('Temps(Année)');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(t:51:26)),hold on,plot(Pr(t:51:26));xlabel('Temps(Année)');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(t:51:26)),xlabel('Temps(Année)');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(t:51:27)),hold on,plot(Pr(t:51:27));xlabel('Temps(Année)');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(t:51:27)),xlabel('Temps(Année)');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(t:51:28)),hold on,plot(Pr(t:51:28));xlabel('Temps(Année)');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(t:51:28)),xlabel('Temps(Année)');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(t:51:29)),hold on,plot(Pr(t:51:29));xlabel('Temps(Année)');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(t:51:29)),xlabel('Temps(Année)');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(t:51:30)),hold on,plot(Pr(t:51:30));xlabel('Temps(Année)');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(t:51:30)),xlabel('Temps(Année)');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(t:51:31)),hold on,plot(Pr(t:51:31));xlabel('Temps(Année)');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(t:51:31)),xlabel('Temps(Année)');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(t:51:32)),hold on,plot(Pr(t:51:32));xlabel('Temps(Année)');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(t:51:32)),xlabel('Temps(Année)');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(t:51:33)),hold on,plot(Pr(t:51:33));xlabel('Temps(Année)');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(t:51:33)),xlabel('Temps(Année)');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(t:51:34)),hold on,plot(Pr(t:51:34));xlabel('Temps(Année)');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(t:51:34)),xlabel('Temps(Année)');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(t:51:35)),hold on,plot(Pr(t:51:35));xlabel('Temps(Année)');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(t:51:35)),xlabel('Temps(Année)');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(t:51:36)),hold on,plot(Pr(t:51:36));xlabel('Temps(Année)');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(t:51:36)),xlabel('Temps(Année)');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(t:51:37)),hold on,plot(Pr(t:51:37));xlabel('Temps(Année)');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(t:51:37)),xlabel('Temps(Année)');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(t:51:38)),hold on,plot(Pr(t:51:38));xlabel('Temps(Année)');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(t:51:38)),xlabel('Temps(Année)');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(t:51:39)),hold on,plot(Pr(t:51:39));xlabel('Temps(Année)');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(t:51:39)),xlabel('Temps(Année)');ylabel('Erreur relative(%)');
```

```
% Pas Spatiales : Pour n=1:51 (nombres d'années)
figure.subplot(2,1,1),plot(0:(1:1:39)),hold on,plot(Pr(1:1:39));xlabel('Numero de station');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(1:1:39)),xlabel('Numero de station');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(2:1:39)),hold on,plot(Pr(2:1:39));xlabel('Numero de station');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(2:1:39)),xlabel('Numero de station');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(3:1:39)),hold on,plot(Pr(3:1:39));xlabel('Numero de station');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(3:1:39)),xlabel('Numero de station');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(4:1:39)),hold on,plot(Pr(4:1:39));xlabel('Numero de station');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(4:1:39)),xlabel('Numero de station');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(5:1:39)),hold on,plot(Pr(5:1:39));xlabel('Numero de station');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(5:1:39)),xlabel('Numero de station');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(6:1:39)),hold on,plot(Pr(6:1:39));xlabel('Numero de station');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(6:1:39)),xlabel('Numero de station');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(7:1:39)),hold on,plot(Pr(7:1:39));xlabel('Numero de station');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(7:1:39)),xlabel('Numero de station');ylabel('Erreur relative(%)');
figure.subplot(2,1,1),plot(0:(8:1:39)),hold on,plot(Pr(8:1:39));xlabel('Numero de station');ylabel('Précipitations annuelles(mm)');subplot(2,1,2),plot(Er(8:1:39)),xlabel('Numero de station');ylabel('Erreur relative(%)');
```

```
function [XV,KG,PT] = KalmanShort(D,NC,NR,PHI,P)
% This program performs discrete Kalman filter
% D = This is the data matrix with NC columns (stations) and NR rows (data)
% PHI = Transition matrix (NCxNC) IT CAN
% BE CALCULATED AS corrcoeff(D)
% P = Covariance matrix (NCxNC) IT CAN
% BE CALCULATED AS cov(D)
% X = The Kalman prediction (NCxNR)
% KG= The Kalman gain (NCxNC)
% PT= The covariance trace (NR)
H(1:NC,1:NC)=0;
Q(1:NC,1:NC)=0;
R(1:NC,1:NC)=0;
for i=1:NC
    H(i,i)=1;
    Q(i,i)=100;
    R(i,i)=10;
end
% Kalman operations
PT(1)=trace(P);
Z(1,1:NC)=D(1,1:NC);
for kk=2:NR
    Z(kk,1:NC)=D(kk,1:NC);
    % State update X(k/k)=PHI(k/k-1)*X(k-1/k-1)
    X(1:NC,1)=PHI*D(kk,1:NC);
    % State estimate update, P(k/k-1)=PHI(k/k-1)P(k-1/k-1)PHIT(k/k-1)+Q(kk)
    P=PHI*P*PHI'+Q;
    % Kalman gain calculation, K(k)=P(k/k-1)H(k)T[H(k)P(k/k-1)H(k)T + R(k)]
    KG=P*H'/(H'*P*H+R);
    % State variable update calculation,
    % X(k/k)=X(k/k-1)+K(k)*[Z(k)-K(k)*X(k/k-1)]
    X(1:NC,1)=X(1:NC,1)+KG*(Z(kk,1:NC)-H'*X(1:NC,1));
    XV(1:NC,1:NC)=X(1:NC,1);
    % Estimation error covariance matrix calculation,
    % P(k/k)=P(k/k-1)-K(k)*H(k)*P(k/k-1)
    P=P-KG*H'*P;
    % Trace of the covariance matrix
    PT(kk)=trace(P);
end
```

