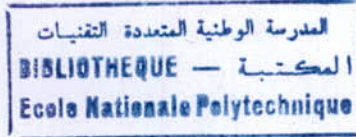


REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA
RECHERCHE SCIENTIFIQUE



P0006/05B

ECOLE NATIONALE POLYTECHNIQUE
DEPARTEMENT HYDRAULIQUE

Projet de fin d'études

**EN VUE DE L'OBTENTION DU DIPLOME D'INGENIEUR D'ETAT
EN HYDRAULIQUE**

**Etude de la relation pluie-débit au pas temps horaire par
différentes approches (DPFT-Réseau de Neurones –
ARMAX)**

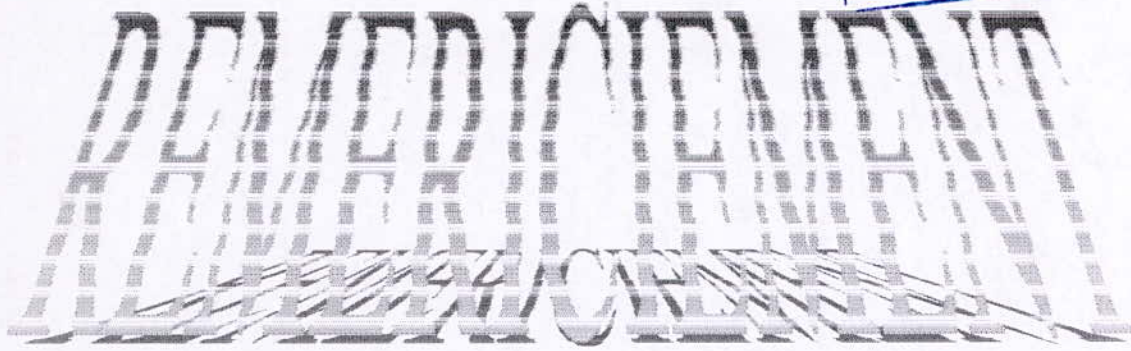
Dirigé par:

Dr: A.BERMAD

Etudié par:

KHELIFI Hocine

Promotion : Juin 2005



« Le difficile c'est ce qui peut être fait tout de suite. L'impossible ce qui prend un peu de temps. »

George Santayana

Je remercie tout d'abord le bon dieu et clément de nous avoir donné la foi et le courage pour terminer ce projet de fin d'études, pour les efforts qui nous a donné, pour pouvoir réaliser ce modeste travail.

Nous tenons à remercier notre promoteur Dr. BERMAD pour sa collaboration, ainsi que M^r le Professeur DECHEMI, qui m'a beaucoup aidé pour ses précieuses remarques et ses orientations scientifiques. Mes remerciements vont également à M^r BENKACI et Mme SOUAG, ainsi, que tous ceux qui ont contribué de près ou de loin à l'élaboration de ce travail.

Comme, nous tenons aussi à remercier tous nos professeurs qui ont contribué à notre formation durant notre cursus formé durant nos cinq années passées, et surtout l'ensemble des enseignants du département d'hydraulique.

DEDICATION

Je dédie ce modeste travail à mes chers parents, qui m'ont encouragés, aidés et soutenus, pour qu'enfin je puisse aboutir ce thème de fin d'études.

Comme je le dédie à :

- Mes frères : Le grand chef " HAMOU", MEHDI, RIAD et BAHIME, et bien sur mes petites sœurs BOUCHRA et HANANE.
- Mes amis : de l'Ecole Nationale Polytechnique du département hydraulique.
-
- Mes amis intimes : LOKMAN, IDRIS, MESBAH, ADEL, SEBTI, LOTFI, AMINE, YACINE et REDA de ma ville natale « ».
- Toute la promotion de l'année 2005 de l'Ecole Nationale Polytechnique, et tous ceux qui connaissent HOUSSAME soit à l'intérieur du pays ou à l'extérieur.

SOMMAIRE

المدرسة الوطنية المتعددة التقنيات
BIBLIOTHEQUE — المكتبة
Ecole Nationale Polytechnique

<i>INTRODUCTION</i>		01
<i>CHAPITRE I</i>	<i>Revue bibliographique des modèles de prévision des débits</i>	
I-1	Introduction	03
I-2	La modélisation hydrologique	03
	I- 2- 1- Variables indépendantes ou variables d'entrée	03
	I- 2- 2- Variables dépendantes ou variables de sortie	03
	I- 2- 3- Variables d'état	03
	I- 2- 4- Paramètres	04
I-3	Les outils et les objectifs de la modélisation	04
	I- 3- 1- Les outils de la modélisation	04
	I- 3- 2- Les objectifs de la modélisation	05
I-4	La typologie des modèles	05
	I- 4- 1- Les modèles déterministes ou stochastiques	06
	I- 4- 2- Les modèles empiriques ou physiques	06
	I- 4- 3- Les modèles globaux ou spatialisés	07
	Conclusion	08
<i>CHAPITRE II</i>	<i>Collecte et critique des données</i>	
II-1	Introduction	09
II-2	Inventaire des données	09
II-3	Critique préliminaire des données	09
	II- 1- Introduction	09
	II- 2- Méthodes de comblement de lacune	10
	II- 3- Détection des valeurs singulières	12
II-4	Traitement statistique des données	13
	II- 4- 1- Test d'homogénéité	13
	II- 4- 2- Test d'indépendance	14
	Conclusion	15
<i>CHAPITRE III</i>	<i>Etude de la méthode de la différence première de la fonction de transfert</i>	
III-1	Présentation	16
III-2	Les principes de base de la méthode	17
	- Principe de base n°1 : La formulation en différences	17
	- Principe de base n°2 : Les pluies efficaces considérées comme inconnues du problème	19
	- Principe de base n°3 : Une solution alternative alternée	20
III-3	Description de la méthode de résolution	20
III-4	Techniques numériques utilisées dans la DPFT	24
	III- 4- 1- Identification de la fonction de transfert	24
	III- 4- 2- Déconvolution des pluies efficaces	25

<i>CHAPITRE IV</i>	<i>Modélisation par les Réseaux de Neurones</i>	
IV-1	Introduction	27
IV-2	Les réseaux de neurones	27
IV-3	Historique	28
IV-4	Définition du neurone formel	29
IV-5	Les types des réseaux de neurones	30
	IV- 1- Réseaux mono-couche	30
	- Perceptron	30
	- Adaline	31
	IV- 2- Réseaux multicouches	32
	- Perceptron	32
	- Propriétés	32
	- Pratique	33
IV-6	Les fonctions d'activation	34
IV-7	Apprentissage d'un réseau de neurones	35
	VII- 1- Apprentissage supervisé	35
	VII- 2- Apprentissage non supervisé	36
	VII- 3- Apprentissage renforcé	37
<i>CHAPITRE V</i>	<i>Modélisation par l'Autorégressif à Moyenne Mobile avec une variable exogène (ARMAX)</i>	
V-1	Introduction	39
V-2	Modèle ARMA à coefficients constants	39
	- Modèle autorégressif (AR)	39
	- Modèle à moyenne mobile (MA)	40
V-3	Modèle ARMA des séries périodiques	40
V-4	Le modèle ARMAX	41
V-5	Le modèle ARMAX sur la variation de débit	43
V-6	Le modèle ARMAX sur le débit	43
<i>CHAPITRE VI</i>	<i>Application sur les données du bassin versant de l'oued EBDA</i>	
VI-1	Introduction	46
	VI- 1- 1- Caractéristiques physiques du bassin de Oued EBDA	46
	VI- 1- 2- Choix des épisodes averse-crue	47
VI-2	Application de la méthode DPFT sur l'Oued EBDA du sous bassin CHELLIF	48
	VI- 2- 1- Procédure de résolution	48
	VI- 2- 2- Présentation des résultats et interprétation	50
	VI- 2- 3- Conclusion	54
VI-3	Application du modèle Réseau de neurones	56
	VI- 3- 1- Introduction	56
	VI- 3- 2- Présentation et interprétation des résultats	56
	Conclusion	60
VI-4	Application du modèle ARMAX	61
	VI- 4- 1- Introduction	61
	VI- 4- 2 - Présentation et interprétation des résultats	61
	Conclusion	64
VI-5	Synthèse des résultats des trois approches	65

VI- 5- 1-	Les résultats de dix crues calage et dix crues validation	67
VI- 5- 2-	Les résultats de cinq crues calage et quinze crues validation	69

CONCLUSION GENERALE

73

Bibliographie

Annexes

INTRODUCTION GENERALE

Le problème des inondations n'est pas un problème mineur. La mise en place des projets d'aménagement des bassins versants tant en milieu rural qu'en milieu urbain consiste à construire des ouvrages et des systèmes de protection contre les crues et les inondations (barrages, digues,...), cette protection passive est beaucoup coûteuse. La protection contre les crues par ces moyens n'est pas vraiment fiable à cause des conséquences désastreuses des inondations survenues récemment en Algérie (Bab El Oued) ou dans le monde qui se traduisent par des pertes humaines et des coûts économiques.

A partir de là, les investissements en ouvrages et systèmes de protection contre les crues peuvent être largement récompensés s'ils permettent de limiter les dégâts matériels et des pertes de vies humaines. Une technique est plus dynamique, basé sur des systèmes d'annonces de crues, qui sont censés alerter la population avec une échéance suffisante pour prendre les dispositions nécessaires à la protection des personnes et des biens.

La performance d'un système d'annonce de crue dépend en partie des modèles utilisés pour représenter le fonctionnement du système hydrologique considéré.

Le travail présenté dans ce mémoire s'inscrit dans le cadre de la modélisation hydrologique et plus particulièrement la modélisation pluie-débit au pas de temps horaire.

Le recours à des modèles à faible nombre de variables, capable de traiter la non-linéarité du phénomène s'avère nécessaire. C'est dans cette optique que nous proposons l'utilisation de trois modèles et leur comparaison, l'une connue sous le nom DPFT (Différence Première de la Fonction de Transfert), c'est une extension de la méthode classique de l'hydrogramme unitaire, la seconde méthode est celle de la modélisation par les réseaux de neurones, issue directement du domaine de l'intelligence artificielle et le dernier c'est le modèle l'autorégressif à moyenne mobile avec une variable exogène (ARMAX).

Ce mémoire comportera en premier lieu, une revue bibliographique sur la modélisation et la typologie des modèles existant dans la littérature, le second chapitre portera sur l'étude détaillée de la méthode DPFT, suivie des théorie de la modélisation par les réseaux de neurones en chapitre trois et l'autorégressif à moyenne mobile en chapitre quatre. L'application et l'interprétation des résultats des trois modèles sur des données réelles est présentée en chapitre cinq, et finalement on termine avec une conclusion générale.

Chapitre I

Revue bibliographique
des modèles de
prévision des débits

I- 1- Introduction :

Le bassin versant (espace géographique collectant les précipitations en amont d'un point de la rivière définissant son exutoire) ne peut plus être considéré comme un tout indifférencié, comme une « boîte noire » qui ne reproduit en sortie que le seul débit observé à l'exutoire.

D'où provient l'eau des ruisseaux ? Quels sont les chemins parcourus depuis l'instant où une goutte de pluie rencontre la surface de la terre ? Telles sont les questions fondamentales qui se posent à l'hydrologie depuis plus d'une quarantaine d'années.

De nombreuses études de terrain, conduites à l'échelle du versant ou des petits bassins versants, ont apporté une bien meilleure compréhension des cheminements de l'eau, qui à son tour a permis, grâce aux progrès concomitants des moyens informatiques, le développement de plusieurs modèles.

Deux étapes sont considérées pour reconstituer le trajet de l'eau de pluie à l'exutoire :

La production : qui assure le passage de la pluie brute en une partie infiltrée et une partie ruisselée ou pluie nette, cette pluie c'est la pluie efficace.

Le transfert : qui représente l'acheminement de la pluie ruisselée à l'exutoire de l'unité hydrologique considérée.

I- 2- La modélisation hydrologique :

La modélisation hydrologique comme la modélisation mathématique d'une manière générale a son vocabulaire propre que nous présentons succinctement ici.

I- 2- 1- Variables indépendantes ou variables d'entrée :

Données d'entrée du modèle. Dans le cas des modèles hydrologiques, il s'agit essentiellement des mesures de pluie. Les modèles hydrologiques sont des modèles dynamiques, les données d'entrée en fonction du temps. Certains modèles utilisent des données d'entrée spatialement distribuées.

I- 2- 2- Variables dépendantes ou variables de sortie :

Il s'agit essentiellement des débits simulés à l'exutoire du bassin versant.

I- 2- 3- Variables d'état:

Variables permettant de caractériser l'état du système modélisé, qui peuvent évoluer en

fonction du temps dans un modèle dynamique. Il s'agit, par exemple, du niveau de remplissage des déferents réservoirs d'eau du bassin versant, du taux de saturation des sols, mais aussi de la profondeur des sols.

1- 2- 4- Paramètres:

La notion de paramètre est intimement liée à celle de modèles conceptuels ou empiriques. Dans de nombreux cas, il n'est pas possible de représenter dans un modèle le processus physique parce que l'échelle de ce processus est trop petite et que les variables d'état contrôlant le processus ne sont pas accessibles à la mesure. Un modèle plus global est alors utilisé pour décrire le processus, mais certaines de ses variables d'état n'ont plus de sens physique, et ne peuvent plus être reliées à des variables mesurables. Ces variables, dont la valeur doit être déterminée par calage, sont appelées paramètres. La loi de Darcy est un exemple de modèle empirique global.

I- 3- Les outils et les objectifs de la modélisation :

I- 3- 1- Les outils de la modélisation :

Les modèles hydrologiques, sont nés bien avant l'avènement de l'informatique. A quel besoin répondaient et répondent encore ces modèles? La littérature scientifique présente une très grande diversité de modèles dont très peu ont trouvé une utilisation opérationnelle. La question de l'utilisation des modèles en hydrologie pour effectuer des prévisions, dimensionner des ouvrages d'art, délimiter des zones inondables, se pose en particulier pour les bassins versants pour lesquels on dispose de séries de débits mesurés. Nous aurons l'occasion de l'aborder dans le cadre du travail consacré à la modélisation hydrologique par les différentes approches (DPFT, réseaux de neurones, ARMAX). Ceci dit, on peut distinguer trois types d'utilisation des modèles mathématiques en hydrologie.

1- La modélisation comme outil de recherche :

La modélisation peut être utilisée pour interpréter des données mesurées. Différents scénarios de fonctionnement hydrologique des bassins versants peuvent être confrontés aux mesures.

2- La modélisation comme outil de prévision :

Anticipation des évolutions futures du débit d'un cours d'eau. Il s'agit de l'utilisation opérationnelle la plus courante des modèles hydrologiques. Dans la plupart des cas cependant,

les modèles développés sont basés sur des régressions linéaires entre les variables indépendantes (pluie, débits amont) et les variables dépendantes (débits aval), et font peu appel aux connaissances sur les processus hydrologiques.

3- La modélisation comme outil d'extrapolation :

Reconstitution de séries de débits plausibles. Dans certains cas, comme par exemple le dimensionnement de déversoirs de sécurité de barrages hydroélectriques ou encore la délimitation de zones inondables dans le cadre d'un plan de prévention des risques, il est nécessaire de proposer des scénarios de crues, ou éventuellement d'étiages, de période de retour nettement supérieure à la durée d'observation des débits sur le site étudié. L'utilisation de longues séries de pluies permet de reconstituer des scénarios rares.

I- 3- 2- Les objectifs de la modélisation :

En hydrologie, les objectifs de la modélisation sont :

1- La prévision :

Les débits futurs à l'exutoire sont évalués par la mesure en temps réel des pluies comme signaux d'entrée.

2- La simulation :

La simulation est l'analyse du comportement du système pour un signal particulier. En hydrologie, la simulation est utilisée :

- Pour compléter des données manquantes ;
- Pour la prédétermination des crues ;
- Pour générer des séries de données utiles à la gestion d'ouvrage de retenue.

3- Le contrôle :

Il consiste à agir sur le signal d'entrée pour obtenir une sortie ayant certaines caractéristiques (gestion de réservoir).

I- 4- La typologie des modèles :

A partir des données observées hydro pluviométriques, la reconstitution théorique de la relation pluie-débit sur un bassin versant peut être résolue selon deux types d'approches.

L'approche mécanique ou dynamique utilise des équations physiques régissant les écoulements en milieu non saturé et les extrapole à l'échelle du bassin versant. L'approche conceptuelle ou systémique assimile le bassin versant à un système. Elle tente de déterminer une opération permettant d'établir la relation entre les données d'entrée (les pluies) et les données de sortie (les débits à l'exutoire).

En fonction du milieu étudié et de la problématique d'étude, de nombreux modèles de transformation pluie débit ont été proposés.

Trois grandes classes de modèles peuvent être établies :

I- 4- 1- Les modèles déterministes ou stochastiques, qui se distinguent par la nature des variables d'état et des paramètres à déterminer. Le modèle déterministe est un modèle paramétrique, basée sur des formulations mathématiques, les différents paramètres sont à déterminer sur la base d'observations.

On distingue deux catégories de modèle déterministe :

- Les modèles BOITE NOIRE : ceux-ci sont construits de façon à transformer l'entrée du système en une réponse du système, mais sans attacher d'importance à la façon de le faire. Il importe seulement que le modèle soit capable de simuler la réponse du système étudié.
- Les modèles conceptuels : contrairement aux modèles Boite Noire, s'attachent à décrire les relations fonctionnelles liant les sollicitations aux réponses du système. On va voir les types des modèles conceptuels.

Pour le modèle stochastique si la probabilité est fonction du temps, on dit que le modèle est stochastique, sinon on parle d'un modèle probabiliste.

I- 4- 2- Les modèles empiriques ou physiques, différenciés par l'utilisation des relations expérimentales ou théoriques (incluant les modèles conceptuels dont le fonctionnement reste décrit par des formules empiriques). Les modèles empiriques caractérisent globalement les relations pluies débits par des traitements de séries chronologiques où n'intervient en général aucune donnée sur la nature physique du bassin : c'est typiquement le cas des modèles régressifs pluie débit et des méthodes basées sur les fonctions de transfert, comme celle de l'hydrogramme unitaire.

On peut distinguer deux types de modèle physique :

- 1- Modèle physique à l'échelle : c'est un modèle réduit de l'état naturel. Ceci nécessite une étude minutieuse de tous les problèmes de distorsion dus à l'échelle. Les difficultés s'accumulent aux plus petites échelles à cause du caractère saisonnier du climat et de la végétation, et de la variabilité spatiale des caractéristiques du bassin. C'est ce que Beven (1995) appelle le problème d'échelle; on ne sait pas vraiment déduire de la description physique et géomorphologique d'un bassin une connaissance des mécanismes qui y prédominent et de la relation pluie débit (Grésillon, 1994).
- 2- Modèle physique analogique : les modèles analogiques exploitent les analogies existantes entre les lois de l'hydrologie et celle de l'électricité et permettent d'approcher d'une manière simple le processus naturel.

I- 4- 3- Les modèles globaux ou spatialisés, dont les échelles spatiales de calcul est l'ensemble du bassin versant où des sous unités discrétisées. Les différents types d'approches des modèles spatialisés sont :

- 1- les modèles conceptuels spatialisés :

Il s'agit des modèles à réservoirs impliquant une discrétisation de l'espace en des unités spatiales que l'on considère homogènes. C'est le cas par exemple du modèle CEQUEAU. Chaque unité est constituée comme un assemblage de réservoirs, et non plus le bassin dans son ensemble. Cela permet en principe de tenir compte de la répartition spatiale des caractéristiques du milieu, aux dépens cependant d'une augmentation du nombre de paramètres à calibrer.

- 2- Les modèles basés sur le concept de similarité hydrologique :

Ces modèles découlent du concept de zones contributives variables et de l'analyse des facteurs qui contrôlent l'extension des zones saturées. L'idée originale revient à postuler que le comportement hydrologique local de tout point du bassin peut être indexé sur la base d'une analyse de la topographie et des caractéristiques des sols. L'aptitude relative des sols d'un bassin à se saturer dépend alors uniquement d'un indice qui fait intervenir la pente locale, la surface amont drainé, et éventuellement la transmissivité du profil si elle est connue. TOPMODEL (Beven et Kirkby, 1979) est le prototype de ce type d'approche, dite semi distribuée, où l'emploi d'un indice de similarité hydrologique régit la spatialisation des

processus du ruissellement et évite ainsi les calculs d'échange entre les mailles des modèles distribués. Ce modèle a été développé, sur la base d'approximations de la physique des écoulements de la zone saturée, dans l'objectif de fournir une description spatiale des processus de façon réaliste tout en ne conservant qu'un minimum de paramètres de calibration.

3- Les modèles physiques distribués :

Historiquement, les approches distribuées ont été développées à des fins de recherche parallèlement aux progrès de l'informatique. Elles utilisent les équations de la physique des écoulements et ont permis d'explorer les conséquences de certaines hypothèses dans tel ou tel cadre théorique particulier, comme par exemple l'influence de la topographie sur la prédominance d'un certain type d'écoulement.

Plus récemment, ces modèles physiques ont cherché à représenter finement le comportement d'un bassin réel. L'intérêt a priori des équations physiques réside dans le fait que leurs paramètres sont en principe mesurables. Beaucoup d'espoirs se sont alors portés sur la capacité prédictive des modèles distribués, en termes de conséquences d'un changement d'occupation du sol notamment. Mais ces modèles se heurtent à la complexité et à l'hétérogénéité du milieu naturel.

Conclusion :

La modélisation pluie débit est donc l'outil obligatoire pour anticiper les conséquences hydrométriques des pluies que l'on observe. Toutefois, un tel modèle ne permet pas de résoudre complètement le problème car tout modèle global pluie débit est par nature beaucoup trop imprécis : des erreurs allant du simple au double sont fréquentes en modélisation pluie débit. Le modèle ne peut donc pas être utilisé au premier degré, mais est indispensable pour exploiter de façon raisonnée les informations pluviométriques et limnimétriques au fur et à mesure de leur arrivée.

La recherche montre que l'utilisation d'un modèle empirique parallèle pour compléter le modèle conceptuel conduit à une bien meilleure prévision du ruissellement que lorsque le modèle conceptuel est utilisé seul.

Finalement, on dira qu'un modèle doit permettre non seulement de calculer mais aussi de comprendre et de raisonner, ce qui permet entre autre de lever l'indétermination dans le cas de défaillance du modèle.

Chapitre II

Collecte et critique des données

II- 1- Introduction :

Les caractéristiques physiques et les données historiques enregistrées concernant la réalisation, la variabilité, les valeurs extrêmes, ...etc des phénomènes hydrologiques constituent les outils essentiels de l'hydrologie.

L'information, apportée par les données historiques, permet à l'hydrologue de connaître certaines caractéristiques physiques d'un phénomène et de prédire son comportement dans le temps ou dans l'espace. La validité théorique d'une étude hydrologique dépend en grande partie de la disponibilité et de la qualité de ces informations, la critique de données permet par le biais de tests statistiques appropriés la recherche, et la détection des observations aberrantes ou erronées, et de vérifier si les séries hydrologiques possèdent certains critères statistiques.

II- 2- Inventaire des données :

Les données recueillies nous ont été aimablement fournies par l'ANRH de BLIDA. Elles sont constituées d'un fichier de précipitations de la station de ARIBS EBDA (Oued EBDA, CHELIFF), s'étalant entre 1972 et 2001, au pas du temps horaire, ainsi que d'un fichier de débits du même sous bassin versant au pas de temps horaire mais s'étalant de 1982-1995.

Les caractéristiques de la station étudiée sont présentées dans le tableau II- 1.

Tableau II- 1: Caractéristiques de la station étudiée.

Station	Oued	Code	Latitude (Nord)	Longitude (EJW)	Altitude (mètre)
ARIBS ABDA	EBDA	18-01	36°19'12"	2° 1'43"E	736

II- 3- Critique préliminaire des données :**II- 3- 1- Introduction :**

Les types de problèmes le plus souvent rencontrés lors du traitement des données en hydrologie sont :

- Le manque des données : station non existante ou fichier incomplet.
- Non- continuité des données : manque d'enregistrement.
- La mauvaise qualité des données, qui peut dépendre de deux causes :

1. Données inconsistantes : lorsqu'elles proviennent d'une erreur systématique due à un instrument de mauvaise qualité ou non calibré ou à un mauvais réglage de l'appareil de mesure.
2. Données non homogènes : lorsqu'il se produit à un moment déterminé, un changement de l'emplacement de l'appareil qui affecte la mesure par exemple.

Afin de garantir la qualité des données, on doit suivre les étapes suivantes :

- Combler les données manquantes dans les séries d'observations.
- Détecter et corriger les erreurs.
- Vérifier l'homogénéité et l'indépendance des données.

II- 3- 2- Méthodes de comblement de lacune :

Dans un enregistrement continu de la pluviométrie d'une station, certaines données sont manquantes pour différentes périodes, afin de ne pas perdre la continuité de l'information, il est nécessaire d'estimer les valeurs manquantes.

Dans ce paragraphe, on va décrire les méthodes de comblement des lacunes, malgré que les deux dernières méthodes ne peuvent pas être utilisées dans ce cas car elles nécessitent plusieurs stations.

- Calcul des moyennes.
- Régression linéaire sur les variables mesurées.
- Simulation.

a) Calcul des moyennes :

Lorsqu'un fichier d'observations comporte moins de trois valeurs manquantes, le comblement de ces lacunes se fait par la méthode de la moyenne :

- On remplace une valeur annuelle manquante par la moyenne interannuelle de la série considérée.
- S'il s'agit d'une valeur mensuelle, on la remplace par la moyenne mensuelle des mois disponibles ou par la moyenne mensuelle interannuelle du mois considéré.
- Une donnée horaire manquante peut être remplacée soit par la moyenne entre la valeur des heures précédentes et suivantes, soit par la moyenne horaire du jour.

b) La régression linéaire :

Soit X et Y deux échantillons, avec n observations (x_i, y_i) ; $i=1..n$, correspondant chacun à des mesures simultanées de précipitations sur deux stations assez rapprochées, les observations sont donc similaires, et la connaissance d'une des variables permet, alors dans une certaine mesure, la prédiction quantitative ou qualitative de l'autre.

La méthode des moindres carrés permet de déterminer l'équation de régression $\hat{y} = a * x + b$ de telle sorte que $\sum (y_i - \hat{y}_i)^2$ soit minimal.

$$\sum (y_i - \hat{y}_i)^2 = \sum (y_i - a - b * x_i)^2 = f(a, b) \text{ Ce minimum est atteint pour } \frac{\partial f}{\partial a} = \frac{\partial f}{\partial b} = 0.$$

c) La simulation :

La notion de la simulation est basée essentiellement sur le concept du hasard. C'est une méthode de tirage d'un échantillon artificiel de (n) valeurs d'une variable aléatoire continue. Elle s'applique lorsque F^{-1} à une forme analytique simple.

Soit F(x) la fonction de répartition de x. la variable $Y = f(x)$ est distribuée sur [0,1] en effet ;

$$\text{Prob}(Y < y) = \text{prob}[F(x) < y] = y$$

Si on tire (n) nombres uniformément répartis sur [0,1] r_1, r_2, \dots, r_n l'échantillon cherché $(x_1, x_2, x_3, \dots, x_n)$ sera déterminé par $x_i = F^{-1}(y_i)$, (fig. II- 1). On dit alors qu'on a simulé la variable x.

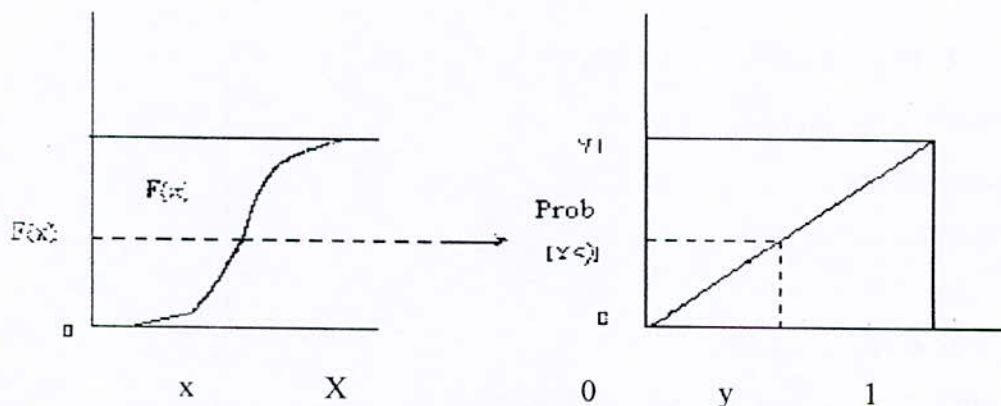


Figure II- 1 : Simulation à partir de la fonction de répartition

L'analyse des fichiers fournis par l'ANRH ne présentent pas de valeurs manquantes.

II- 3- 3- Détection des valeurs singulières :

Les observations qui s'écartent fortement du centre de répartition sont le résultat d'une erreur systématique, d'une lecture incorrecte des indications de l'instrument de mesure ou à cause des valeurs intégrées après le comblement des lacunes.

Il est important avant tout traitement de détecter ces erreurs qui se glissent dans les données de base, afin que l'on puisse effectuer une analyse plus fine pour décider si cette donnée douteuse est aberrante et doit donc être éliminée, ou si c'est une donnée exceptionnelle, qui par conséquent doit être au contraire précieusement conservée.

Test de Grubbs et Beck (valeurs singulières) :

Ce test est très utilisé à cause de sa simplicité. Il peut détecter les points singuliers de la série en calculant les statistiques suivantes :

$$\begin{aligned} X_h &= \text{Exp} (m + K_n * S) \\ X_i &= \text{Exp} (m - K_n * S) \end{aligned} \quad (\text{II. 1})$$

Où :

m et S : sont respectivement, la moyenne et l'écart type des logarithmes naturels des éléments constituant l'échantillon.

K_n : est une valeur statistique du test de Grubbs et Beck, tabulée pour différentes tailles d'échantillons et de niveaux de signification.

X_h, X_i : des paramètres statistiques à déterminer.

Au niveau de risque de 10%, l'approximation polynomiale a été donnée par l'expression de Pilon (1985) :

$$K_n = -3.62201 + 6.28446 * n^{1/4} - 2.49835 * n^{1/2} + 0.491436 * n^{3/4} - 0.037911 * n \quad (\text{II. 2})$$

Avec n : Taille de l'échantillon.

En appliquant ce test, tous les points de l'échantillon supérieurs à X_h ou inférieurs à X_i , sont considérés comme points singuliers.

Analyse des résultats :

Les tableaux (II- 2) et (II- 3) présentent les résultats obtenus après l'application du test de Grubbs et Beck.

Tableau II- 2 : Résultats du test Grubbs et Beck pour les pluies

	n	m (mm)	S (mm)	Kn	Xh (mm)	Xi (mm)
Echelle horaire	1080	0.82	2.56	1.94	331.21	0.01

Tableau II- 3 : Résultats du test Grubbs et Beck pour les débits

	n	m (mm)	S (mm)	Kn	Xh (mm)	Xi (mm)
Echelle horaire	1080	6.11	16.10	1.94	$1.02 \cdot 10^{16}$	$1.16 \cdot 10^{-11}$

L'application de ce test a détecté au pas du temps horaire les valeurs nulles comme singulières, mais on ne peut pas rejeter ces points car ils correspondent à la fin de la précipitation (fin de crue).

II- 4- Traitement statistique des données :

Après la critique des données, il est indispensable de vérifier l'indépendance et l'homogénéité de la série.

II- 4- 1- Test d'homogénéité :

Les échantillons sont dits homogènes, s'ils sont caractérisés par une même loi de probabilité et par les mêmes paramètres, les données recueillies sont alors considérées comme extraites d'une même population.

Il existe plusieurs tests pour vérifier l'homogénéité des séries de données, on a utilisé dans ce cas celui de **WILCOXON-MANN-WHITNEY** :

Ce test utilise les numéros des observations dans la série obtenus en ordonnant l'échantillon global de taille n.

Le numéro d'observation x_i dans l'échantillon ordonné s'appelle ordre ou rang, et sera désigné par $R_{i,n}$.

La série x_i ($i = 1, n$) classée par ordre croissant est subdivisée en deux échantillons de tailles n_1 et n_2 , où $n = n_1 + n_2$ (avec $n_1 \leq n_2$).

On calcule la somme W_x des rangs des éléments du premier échantillon dans la série

commune :

$$W_x = \sum |R_{i,n}| \tag{II. 2}$$

Wilcoxon a montré que dans le cas où les deux échantillons n_1 et n_2 , constituent une série homogène, la quantité W_x est comprise entre deux bornes W_{\max} et W_{\min} , données par les formules suivantes :

$$W_{\min} = \frac{(N_1 + N_2 + 1) * N_1 - 1}{2} - U_{1-\alpha/2} \sqrt{\frac{N_1 N_2 (N_1 + N_2 + 1)}{12}} \tag{II. 3}$$

$$W_{\max} = (N_1 + N_2 + 1)N_1 - W_{\min} \tag{II. 4}$$

$U_{1-\alpha/2}$: représente la valeur de la variable centrée réduite de la loi normale correspondante à $1 - \alpha / 2$.

Si : $W_{\min} < W_x < W_{\max}$, l'hypothèse de l'homogénéité est acceptée, elle est rejetée dans le cas contraire.

Analyse des résultats :

Les résultats du test de WILCOXON sont illustrés dans le tableau (II- 4) appliqué à différents seuils. Ce-ci montre que les observations sont homogènes.

1 : L'hypothèse d'homogénéité acceptée.

0 : L'hypothèse d'homogénéité rejetée.

Tableau II- 4 : résultats du test statistique (pluies et débits).

	Test	$\alpha = 5 \%$	$\alpha = 10 \%$
Pluie	Homogénéité	1	1
Débits	Homogénéité	1	1

II- 4- 2- Test d'indépendance :

Il est basé sur le critère qu'aucune observation appartenant à l'échantillon ne devrait influencer l'observation suivante :

Pour un échantillon de taille n, le test de Wald-Wolfowitz considère la valeur statistique :

$$R = \sum_{i=1}^n x_i x_{i+1} + x_1 x_n \tag{II. 5}$$

Si les éléments de l'échantillon sont indépendants, R suit la loi normale de moyenne :

$$\bar{R} = \frac{S_1 - S_2}{(n-1)} \quad (\text{II. 6})$$

Et de variance :

$$\text{var}(R) = \frac{(S_2^2 - S_4)}{(n-1)} - \bar{R}^2 + \frac{(S_1^4 - 4S_1^2 * S_2 + 4S_1 * S_3 + S_2^2 - 2 * S_4)}{(n-1) * (n-2)} \quad (\text{II. 7})$$

Avec : $S_r * n = \sum x_i^r = M_r'$

M_r' : Moment d'ordre r de l'échantillon.

On calcule la statistique $U = \frac{R - \bar{R}}{\sqrt{\text{var}(R)}}$ qui suit une loi normale N (0,1), et on test l'hypothèse

d'indépendance en comparant $U_{\alpha/2}$ avec la valeur normale standard à la probabilité de dépassement de $\alpha/2$.

- Si $|U|$ est inférieure à $U_{\alpha/2}$ l'hypothèse d'indépendance est vérifiée.
- Si $|U|$ est supérieure à $U_{\alpha/2}$ l'hypothèse d'indépendance est rejetée.

Analyse des résultats :

Les résultats du test de Wald-Wolfowitz illustré dans le tableau (II- 5) :

1 : L'hypothèse d'homogénéité acceptée.

0 : L'hypothèse d'homogénéité rejetée.

Tableau II- 5 : résultats du test statistique (pluies et débits).

	Test	$\alpha = 5 \%$	$\alpha = 10 \%$
Pluie	Indépendance	0	0
Débits	Indépendance	0	0

L'hypothèse d'indépendance pour le pas de temps horaire est rejetée pour les différents seuil utilisés.

Conclusion :

On constate que les fichiers de données (pluie-débit) de l'ANRH sont fiables.

Chapitre III

Etude de la Méthode D.P.F.T

Différences Premières de la Fonction de

Transfert

III- 1- Présentation :

La description de la relation pluie- débit à l'aide d'un modèle capable de reproduire le phénomène naturel, d'une manière rigoureuse, s'avère nécessaire pour effectuer une prévision sur les débits, c'est dans cette optique, que nous proposons une méthode de prévision, qui s'inspire de la méthode classique de l'hydrogramme unitaire, proposé par Sherman en (1932). (JATON, 1982).

Le modèle étudié comme celui de l'hydrogramme unitaire, décompose la relation pluie-débit en deux parties ou sous-modèles :

- Un sous modèle de production, qui transforme la pluie brute en pluie efficace ruisselée, qui se trouve sous forme de débit de ruissellement à l'exutoire. Ce modèle est fortement non linéaire, et doit prendre en compte les conditions antérieures du bassin versant.
- Une fonction supposée linéaire, et temporellement invariante dite 'fonction de transfert', qui étale dans le temps, la pluie efficace en débit de ruissellement, sans oublier un débit de base Q_b , correspondant à la 'réponse lente' du bassin qu'on peut rajouter au débit de ruissellement Q , pour avoir le débit total mesuré Q_T .

La méthode dont l'étude sera faite est connue sous le nom '*Différence Première de la Fonction de Transfert*' (DPFT). Elle s'inspire des travaux effectués par NEWTON et VINYARD (1967) avec la présentation par DUBAND (1978), ou elle n'était qu'une extension améliorée et opérationnellement orientée de l'algorithme des deux chercheurs précédents, ou ils ont proposé un algorithme itératif et alterné permettant l'identification directe d'une fonction de transfert moyenne commune à tous les épisodes, et une série des pluies efficaces.

La DPFT reprend aussi de nombreuses techniques numériques, néanmoins, elle ne se limite pas à identifier un hydrogramme unitaire, comme la plupart des méthodes appartenant à la même famille. Ce procédé nous évite d'imposer à priori des pluies efficaces, ou une fonction de transfert.

Les principes de base posés par NALBANTIS (1987), reflètent l'originalité de cette méthode par rapport aux autres approches classiques de type Hydrogramme Unitaire.

III- 2- Les principes de base de la méthode :

Principe de base n° 1 : La formulation en différences

L'idée de travailler sur les différences premières de débits a deux origines, l'une est opérationnelle : un système d'annonce de crue a souvent besoin de connaître la variation de débits entre le pas présent, ou il est souvent mesurable, et les pas suivants. C'est cette prévision, exprimée sous forme de variation, qui est en réalité diffusée.

L'autre est plus théorique, se rattache aux travaux de BOX et JENKINS (1976) qui nous orientent vers l'utilisation des dérivées d'ordre 1 et 2 des processus réguliers faciles à manipuler, puis revenir ensuite au processus initial. C'est ce qu'on va faire ici et pour le mettre en évidence, écrivons tout d'abord l'équation de la relation pluie-débit donnée par :

$$Q_{dj}^L = \sum_{i=1}^K A_i * PE_{j-i+1} \quad (\text{III } .1)$$

Comme ceci :

$$Q_{dj}^L = A_1 * PE_j + A_2 * PE_{j-1} + A_3 * PE_{j-2} + \dots + A_{K-1} * PE_{j-K+2} \quad (\text{III } .2)$$

Ou :

Q_{dj}^L : Le débit de ruissellement direct sur l'intervalle de temps [j-1, j] de la crue L.

PE_j : La pluie efficace exprimée en termes de volumes d'eau cumulé sur l'intervalle [j-1, j].

$A_1, A_2, A_3, \dots, A_{K-1}$: Les coefficients de l'hydrogramme unitaire ou de la fonction de transfert.

Considérons maintenant les différences premières du débit total, appelées aussi variation de débits entre le pas j-1 et j.

q_j : Différence de débit ($Q_j - Q_{j-1}$).

$a_i = A_i - A_{i-1}$: La différence 1^{ère} de la fonction de transfert, d'où on aura :

$$q_j^L = \sum_{i=1}^K a_i * PE_{(j-i+1)} \quad (\text{III } .3)$$

En hydrologie le débit total peut se décomposer en un débit de ruissellement direct et un débit qui intègre les effet des réponses lentes du bassin appelé débit de **base**. On peut écrire alors :

$$Q_j = Q_{dj} + Q_{bj}$$

$$Q_{j-1} = Q_{dj-1} + Q_{bj-1}$$

D'où :

$$q_j = Q_j - Q_{j-1} = (Q_{dj} - Q_{dj-1}) + (Q_{bj} - Q_{bj-1}) \quad (\text{III .4})$$

On substitue Q_{dj} et Q_{dj-1} par leurs expression données en (III .2) et (III .4) on obtient :

$$q_j = Q_j - Q_{j-1} = A_1 * PE_j + A_2 * PE_{j-1} + \dots + A_{K-1} * PE_{j-K+2} - A_1 * PE_{j-1} - A_2 * PE_{j-2} - \dots - A_{K-2} * PE_{j-2} - A_{K-1} * PE_{j-K+1} + (Q_{bj} - Q_{bj-1})$$

Cette équation peut s'écrire aussi :

$$q_j = (A_1 - 0) * PE_j + (A_2 - A_1) * PE_{j-1} + \dots + (A_{K-1} - A_{K-2}) * PE_{j-K+2} + (0 - A_{K-1}) * PE_{j-K+1} + (Q_{bj} - Q_{bj-1})$$

En définissant la différence première de la fonction de transfert comme le fonction discrète dont les coefficients sont les différences premières des coefficients de la fonction de transfert :

$$\begin{aligned} a_1 &= A_1 \\ a_i &= A_i - A_{i-1} \\ a_K &= -A_{K-1} \end{aligned}$$

$$q_i^L = a_1 * PE_i^L + a_2 * PE_{i-1}^L + \dots + a_K * PE_{i-K+1}^L + (Q_{bi}^L - Q_{bi+1}^L) \quad (\text{III .5})$$

Ou encore de manière condensée :

$$q_{Ti}^L = \sum_{j=1}^K a_j * PE_{i-j+1}^L + (Q_{bi}^L - Q_{bi+1}^L) \quad (\text{III .6})$$

$\sum_{j=1}^K a_j * PE_{i-j+1}^L$: La variation du débit de ruissellement direct.

$Q_{bi}^L - Q_{bi+1}^L$: Variation du débit de base.

Dans la plupart des crues, on peut admettre que les variations de débit de base sont limitées (NALBANTIS (1987)), car il s'agit bien d'une réponse lente du bassin versant.

L'équation (III .6) devient :

$$q_{Ti}^L = q_i^L = \sum_{j=1}^K a_j * PE_{i-j+1}^L \quad (\text{III .7})$$

L'écriture équivalente de (III .7) est de manière analogue à celle de $Q_i = \sum_{j=1}^m PE_j * A_{i-j+1}$:

$$q_i^L = \sum_{j=1}^m PE_j^L * a_{i-j+1} \quad (\text{III .8})$$

Ou : m est le nombre de pluies efficaces mises en jeu.

On remarque que la formulation du phénomène reste la même après l'introduction de la notion de différence première. Les nouvelles écritures des systèmes deviennent :

$$\{q^L\} = [a] * \{PE^L\} \quad (III.9)$$

Avec :

$$\{q^L\} = \begin{Bmatrix} q_1^L \\ q_2^L \\ \dots \\ q_n^L \end{Bmatrix}; \quad \{PE^L\} = \begin{Bmatrix} PE_1^L \\ PE_2^L \\ \dots \\ PE_m^L \end{Bmatrix}; \quad [a] = \begin{bmatrix} a_1 0 \dots \dots \dots 0 \\ a_2 a_1 \dots \dots \dots 0 \\ \dots \dots \dots \dots \dots \dots \\ a_n a_{n-1} \dots \dots \dots a_{n-m+1} \end{bmatrix}$$

Et :

$$\{q^L\} = [PE^L] * \{a\} \quad (III.10)$$

Avec :

$$\{q^L\} = \begin{Bmatrix} q_1^L \\ q_2^L \\ \dots \\ q_n^L \end{Bmatrix}; \quad \{a\} = \begin{Bmatrix} a_1 \\ a_2 \\ \dots \\ a_k \end{Bmatrix}; \quad [PE^L] = \begin{bmatrix} PE_1^L 0 \dots \dots \dots 0 \\ PE_2^L PE_1^L \dots \dots \dots 0 \\ \dots \dots \dots \dots \dots \dots \\ PE_m^L PE_{m-1}^L \dots \dots \dots PE_{m-k}^L \\ \dots \dots \dots \dots \dots \dots \\ 0 \dots \dots \dots PE_m^L \end{bmatrix}$$

Où : n : Nombre total de débit ;

m : Nombre de pluies efficaces ;

- Dans le dernier système d'équation (III.10), les coefficients de la DPFT a_i sont tronqués à k pas de temps qui est la mémoire du système ou mémoire d'identification.

Pour se ramener à la fonction de transfert qui lie la pluie en mm/h et le débit en m^3/s , il suffit

de multiplier les coefficients A_i par : $R_n = \frac{S}{36} * \Delta t$

Avec : S La surface du bassin versant exprimée en Km^2 .

Δt Le pas de temps exprimé en heures.

Principe de base n° 2 : Les pluies efficaces considérées comme inconnues du problème

La méthode DPFT diffère des autres méthodes du type hydrogramme unitaire en considérant les pluies efficaces comme inconnues du problème. Elle va encore plus loin en

recherchant simultanément la fonction de transfert par un traitement multi événement, et les séries des pluies efficaces épisode par épisode.

Principe de base n° 3 : Une solution itérative alternée

On se propose à l'itération (i-1) d'identifier la $DPFT_{(i)}$ moyenne à l'aide des pluies efficaces déconvoluées à l'itération (i-1), en utilisant un jeu de données multi-événements. Avec cette $DPFT_{(i)}$ moyenne ainsi obtenue à l'itération (i), nous allons réestimer un nouveau jeu de pluies efficaces $PE_{(i)}^L$ événement par événement, par déconvolution directe sur le système $q^L = a * PE^L$.

En itérant le processus, on espère converger vers un optimum.

L'initialisation de ce processus alternatif itéré se fait à l'aide des pluies brutes PB^L , en considérant, faute de mieux, que : $PE_{(0)}^L = PB^L$, c'est-à-dire, que les pluies brutes sont une première estimation des pluies efficaces. Les inconnues mises en jeu sont :

➤ Les coefficients a_i de la DPFT : tronqués à k pas de temps.

On résout le système multi-événements (III .10) :

$q = PE * a$ Qui est un système comportant K inconnues et n^{Tot} équations définies par :

$$n^{Tot} = n^1 + n^2 + n^3 + \dots + n^L + \dots + n^N$$

➤ Les $m^{Tot} = m^1 + m^2 + \dots + m^L + \dots + m^N$ pluies efficaces pouvant être considérés comme des paramètres d'entrée.

m^L : Le nombre de pluies efficaces de l'épisode L, et le jeu comporte N crues.

D'où il faut vérifier la condition : $n^L \leq m^L + k - 1$.

Dans la résolution crue par crue du système de déconvolution, nous avons n^L équations et m^L inconnues. La condition de résolution de ce système est $n^L \geq m^L$.

Le nombre maximal de débits n^L doit s'écrire :

$$n^L = m^L + k - 1 \quad \text{(III .11)}$$

III- 3- Description de la méthode de résolution :

La résolution se fait par itération, dans ce qui suit nous donnons une idée générale sur la résolution de l'algorithme comportant la détermination de la fonction de transfert ainsi que la déconvolution des pluies efficaces.

1^{er} Itération :

1- Détermination de la fonction de transfert :

On commence donc par résoudre, dans la phase d'identification de la fonction de transfert, le système multi événements $q = PE * a$ à l'aide des pluies brutes, considérés comme une première approximation des pluies efficaces. On obtient ainsi une première estimation $a_{(1)}^*$ de la DPFT :

$$q = PE_{(0)} * a = PB * a \dots\dots\dots a_{(1)}^* \quad (III .12)$$

La fonction de transfert déduite de cette DPFT estimée, peut ne pas vérifier certaines conditions, pour cela, on applique à posteriori à la fonction de transfert $A_{(1)}^*$ déduite de $a_{(1)}^*$, et dans l'ordre suivant, les contraintes :

1. $A_{(1)i}^* \geq 0$ pour tout $i = 1, 2, \dots, k$ et toute valeur négative ne sera pas acceptée, et sera prise égale à zéro.
2. On lisse la queue de la fonction de transfert en ajustant sur les derniers termes une courbe exponentielle. Celle ci permettra en plus de prolonger la fonction de transfert théoriquement jusqu'à l'infini. (RODRIGUEZ, 1989).

Cette courbe exponentielle est de la forme :

$$A_j = A_L * EXP(-\alpha) \quad j > L$$

Avec :

- A_j : Coefficients de la fonction de transfert au pas j.
- A_L : Premier coefficient de la branche de la fonction de transfert à lisser.
- α : Coefficient du lissage.

3. En fait, on considère au plus n^{\max} ordonnées de la fonction de transfert, avec :

$$n^{\max} = \text{Max}(n^1, n^2, \dots, n^L, \dots, n^N), \text{ et on imposera :}$$

$$\sum_{i=1}^{n^{\max}} A_i = 1$$

Qui traduit la conservation du volume ruisselé. (NALBANTIS, 1987).

Cette contrainte peut être introduite de deux façons :

- Entre les itérations => *contrainte indirecte.*
- Au cour de l'itération => *contrainte directe.*

2- Déconvolution des pluies efficaces :

On repasse aux différences, d'où $a_{(1)}$, puis dans la phase de la déconvolution des pluies efficaces, on résout le système (III.10) crue par crue, et en utilisant ces coefficients pour résoudre le système (III.9) :

$$q = a * PE \dots\dots\dots PE^* \quad (III.13)$$

Globalement, on doit estimer deux jeux d'inconnues, à savoir :

1/ Les coefficients de la différence première de la fonction de transfert moyenne tronquée à k pas de temps, ce qui revient à supposer une fonction de transfert de mémoire k-1.

2/ $m_T = \sum_{L=1}^N m_L$, N étant le nombre de pluies efficaces pouvant être considérées comme des paramètres d'entrée.

Pour l'estimation des pluies efficaces $PE_{(1)}$, et de la différence première de la fonction de transfert, pour chaque épisode séparément du débit, sera donnée par le système (III.9) $\{q_1\} = [a_{(1)}] * \{PE_1\}$.

Ici encore, les pluies efficaces $PE_{(1)}^L$ obtenues numériquement peuvent ne pas convenir parfaitement, on doit vérifier :

- Si elles sont positives ou nulles. la valeur négative ne sera pas acceptée, et sera prise égale zéro.
- A partir de la définition du coefficient de ruissellement et du fait que celui-ci doit être inférieur ou égal à un (1), en sachant que ce coefficient est donné par :

$$C_{ij}^L = PE_j^L / PB_j^L \leq 1$$

C_{ij}^L : Coefficient de ruissellement.

PE_j^L : Pluies efficaces sur l'intervalle [j-1, j] pour l'épisode L.

PB_j^L : Pluie brute sur l'intervalle [j-1, j] pour l'épisode L.

2^{ème} Itération :

Pour réitérer cette procédure alternative, on remplace PB^L par $PE_{(1)}^L$ la pluie efficace estimée à la 1^{ère} itération, on obtient alors $a_{(2)}$ avec laquelle on déterminera PE^L en vérifiant les contraintes précédentes. Ce processus itératif est illustré dans cette figure :

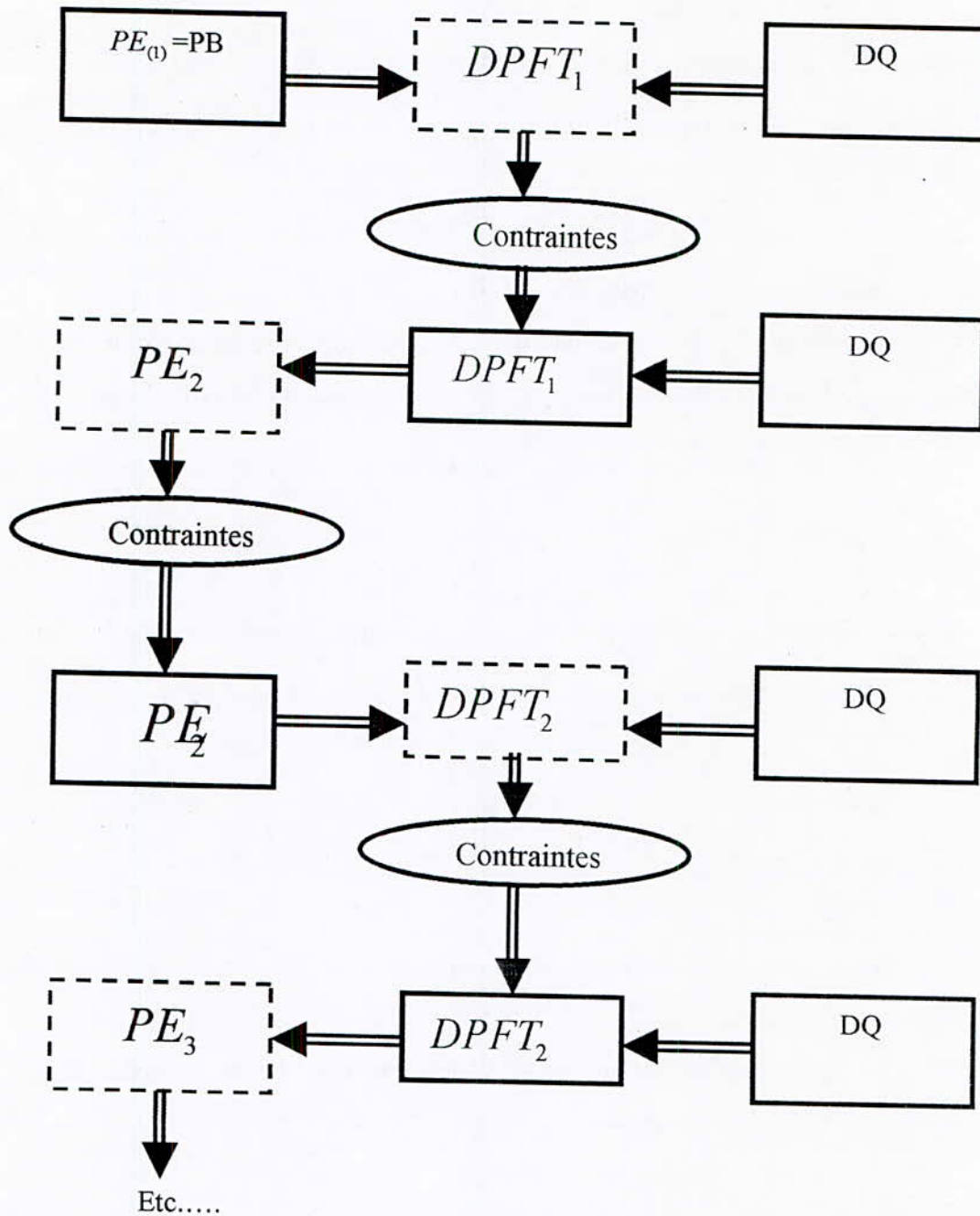


Figure III- 1 : Algorithme alternatif de la DPFT

Cet algorithme alternatif itéré était supposé converger par analogie avec les méthodes de gradients. Nalbantis (1987) a établi la convergence sur les données synthétiques générées 3 à 5 itérations suffisent, selon la qualité des mesures de pluie et de débit dont on dispose, et c'est aussi ce que l'on observe en pratique sur des données réelles.

III- 4- Techniques numériques utilisées dans la DPFT :

Dans cette partie, on présente un résumé des travaux et algorithmes faits à chaque version de la méthode DPFT. On distingue ceux de DUBAND (1978), GUILLOT et DUBAND (1980), VERSIANI (1983), NALBANTIS (1987) et RODRIGUEZ (1989).

III- 4- 1- Identification de la fonction de transfert :

En premier lieu, (Guillot et Duband, 1980) proposèrent un système multi-événements : $\{q\} = [PE] * \{a\}$ par une régression multilinéaire. Le système utilisé est :

$$q_i = \sum_{i=1}^k a_i * PE_{j-i+1}^L + \varepsilon_i^L \quad (\text{III .14})$$

Avec :

ε_i^L : Erreur d'ajustement.

En introduisant le terme de régression, on aura :

$$q_i = \sum_{i=1}^k a_i * PE_{j-i+1}^L + h_a + \varepsilon_i^L \quad (\text{III .15})$$

h_a : Terme constant de régression.

Dans la version de 1983, Versiani a estimé que l'introduction du terme de régression peut engendrer une erreur tout au long de la crue, il a proposé un estimateur par moindres carrés :

$$a^* = [(PE)^T PE]^{-1} (PE)^T q \quad (\text{III .16})$$

A cause du mal conditionnement de la matrice à inverser $[(PE)^T PE]^{-1}$ Versiani en 1983 a utilisé la Ridge Régression et le nouvel estimateur de la fonction de transfert était :

$$a^* = [(PE)^T PE + \tau_k G_L]^{-1} (PE)^T q \quad (\text{III .17})$$

τ_k : Coefficient de la Ridge Régression ;

G_L : Matrice diagonale formée par la diagonale de $(PE)^T PE$.

Nalbantis en (1987), propose une identification par le modèle ARX (modèle du type ARMA), appliqué sur les débits mais il a suggéré le recours aux moindres carrés pour l'estimateur de la DPFT.

III- 4- 2- Déconvolution des pluies efficaces :

Dans le but d'avoir des pluies efficaces robustes et bien stables plusieurs techniques ont été élaborées par différents auteurs.

L'estimateur des pluies efficaces est donné par :

$$PE^L = [(a)^T a]^{-1} (a)^T q^L \quad (III .18)$$

Cet estimateur est influencé par le mauvais conditionnement de la matrice à inverser $[(a)^T a]$. Dans sa version de (1983) et pour faire face à ce problème de mal conditionnement de la matrice inverse, Versiani a proposé un estimateur par la Ridge Régression dont l'expression est la suivante :

$$PE^L = [(a)^T a + \tau_k G_L]^{-1} (a)^T q^L \quad (III .19)$$

G_L : Matrice diagonale formée par la diagonale de $(a)^T a$.

τ_k : Coefficient de la ridge régression.

Nalbantis (1987) a tenu compte de certaines contraintes sur les PE en utilisant la programmation linéaire, il a aussi travaillé par simulation mais il a opté pour la ridge régression proposée par Versiani pour la stabilité de ses résultats.

Rodriguez (1989), proposa une nouvelle technique en utilisant la programmation quadratique (algorithme de moindres carrées sous contraintes), son travail est basé sur l'optimisation du critère de l'erreur quadratique du système de déconvolution noté plus bas et l'application des contraintes sur les PE.

$$q^L = a * PE^L + \varepsilon^L \quad (III .20)$$

ε^L : Erreur d'ajustement.

La contrainte $PE \geq 0$ peut s'écrire matriciellement sous la forme suivante :

$$0_m \leq I_m * PE \quad (III .21)$$

0_m : Vecteur nul de dimension 'm'.

I_m : Matrice identité d'ordre 'm'.

m : Le nombre de pluies efficaces.

Si l'on connaît avec précision la pluie efficace, d'être inférieure ou égale à la pluie brute, pondérée par un coefficient de confiance C_{conf} supérieur ou égale à 1.

On peut aussi écrire : $PE \leq C_{conf} * PB$ sous la forme matricielle :

$$C_{conf} * PB \geq I_m * PE \quad (III .22)$$

Cette technique peut diminuer les oscillations des PE lorsque les données sont entachées d'erreurs.

Rodriguez propose aussi de supprimer les pluies efficaces pour lesquelles les pluies brutes sont nulles, ce qui augmente le degré de liberté du système.

Finalement, le principe de la méthode se résume dans l'identification de la DPFT sur un jeu multi-événements. Pour ce faire, il faut fixer la mémoire du système ou la taille de la fonction de transfert, utiliser la technique la plus simple possible pour la détermination des coefficients de la DPFT, afin d'éviter les très lourds calculs (moindres carrés) et application des différents contraintes citées auparavant.

En ce qui concerne la déconvolution des pluies efficaces les différentes étapes peuvent se résumer comme suit :

- Déconvolution directe de PE par le système (III .9).
- Utilisation de la ridge régression.
- Application à chaque itération des contraintes propres aux pluies efficaces.

Chapitre IV

Modélisation par les Réseaux de Neurones

IV- 1- Introduction :

Avec la souplesse, la puissance et la rapidité des ordinateurs, avec les développements mathématiques, surtout dans les domaines de la statistique et de la recherche opérationnelle, et les efforts réalisés durant les dernières décennies pour augmenter et améliorer les données historiques, le modèle mathématique est devenu l'outil essentiel de l'hydrologue et du planificateur des ressources hydriques.

En hydrologie, l'identification des processus physiques pouvant avoir lieu dans un bassin versant et l'attribution à chacun d'eux une description analytique permettant la prévision nous amène à des expressions complexes et non linéaires utilisant un grand nombre de paramètres rendant ainsi leur exploitation lourde et difficile.

C'est dans ce contexte que nous proposons une méthode de modélisation pluie-débit basée sur l'utilisation des réseaux de neurones dont les performances en hydrologie ont été déjà prouvées par plusieurs chercheurs hydrologues (FRENCH et al. 1992, SHAMSELDIN 1996, HSH et al. 1992, MINN et al. HALL 1996, COULIBALY et al. 1999, SAJIKUMAR et THANDAVESWARA 1999).

IV- 2- Les réseaux de neurones :

Relevant du domaine de l'intelligence artificielle, les réseaux de neurones sont une métaphore des structures cérébrales : des assemblages de constituants élémentaires, qui réalisent chacun un traitement simple, mais dont l'ensemble fait émergé des propriétés globales dignes d'intérêt. Chaque constituant fonctionne indépendamment des autres, de telle sorte que l'ensemble est un système parallèle, fortement interconnecté. L'information détenue par le réseau de neurones est distribuée à travers l'ensemble des constituants. Enfin un réseau de neurone ne se programme pas pour réaliser telle ou telle tâche. Il est entraîné sur des données acquises, grâce à un mécanisme d'*apprentissage* qui agit sur les constituants du réseau afin de réaliser au mieux la tâche souhaitée.

Parmi les tâches particulièrement adaptées au traitement par réseau de neurones, on trouve l'association, la classification, la discrimination et l'estimation. Toutes ces tâches sont très différentes de celles traitées par l'information traditionnelle. Tout l'art de la technique des réseaux de neurones consiste aujourd'hui à modéliser ces divers éléments.

IV- 3- Historique :

C'est en 1943 que le premier neurone artificiel fût inventé par Mc Culloch et Pitts. Ce modèle est inspiré des architectures et des fonctions du système nerveux humain, plus exactement de la cellule nerveuse (neurone).

Cette découverte était comme une étincelle poussant les chercheurs à s'intéresser et s'approfondir dans ce nouveau domaine de recherche et c'est en 1950 que Rosenblatt a présenté son nouveau modèle de réseau de neurones appelé *Perceptron*, ce dernier était la réplique du modèle initiale, présenté par Mc Culloch et Pitts mais avec la possibilité d'apprentissage des différentes couches de neurones (YEDDOU, 1998), ce qui constituait une nouveauté à cette époque surtout lorsqu'il s'agit de résoudre les problèmes non linéaires.

En 1969, Papert et Minsky ont démontré les limites du réseau de Rosenblatt et la nécessité de plusieurs couches identiques pour modéliser un phénomène physique le plus fidèlement possible (DAVALO, 1993). Les deux chercheurs ont reproché au perceptron l'inexistence d'une méthode d'apprentissage de plusieurs couches, ce qui mettra à l'ombre l'idée des réseaux de neurones.

Les recherches dans ce domaine furent réactivées durant les années 70, mais cette fois ci les études portées sur les réseaux interconnectés, mais l'inexistence de moyens capables de réaliser des calculs très poussés obligea les chercheurs à abandonner ce concept.

Au début des années 80, une découverte vient propulser les réseaux de neurones permettant ainsi la résolution de vieux problèmes posés par Papert et Minsky. Cette découverte est due à un groupe de chercheur de l'université de Stanford, leurs études se sont soldées par la mise en œuvre d'un algorithme d'apprentissage qui porta le nom de **Backpropagation** (JODOUIN, 1994). Désormais il est possible d'utiliser les réseaux multicouches sans se soucier de leurs apprentissages. En ce qui concerne le domaine de l'hydrologie, la modélisation avec les réseaux de neurones a été introduite au début des années 90. Les études réalisées portaient généralement sur la prévision des débits d'écoulement soit souterrain au niveau des aquifères (DONNA RIZO, 1994) soit superficiel, ainsi que la prévision spatio-temporelle de la précipitation au niveau des bassins versants (FRENCH et al. 1992) et la gestion d'un réservoir en temps réel (SAKAKIMA et al. 1992).

IV- 4- Définition du neurone formel :

On peut définir un neurone formel (ou artificiel) par les éléments suivants :

1. La nature des entrées ;
2. La fonction d'entrée totale qui définit le prétraitement subit par les entrées ;
3. La fonction de sortie qui calcule la sortie du neurone en fonction de son état d'activation ;
4. La nature de la sortie du neurone.

Le neurone calcule ainsi une somme de ces entrées pondérées par des coefficients, et prend une décision en la comparant à un seuil fixe : si la somme pondérée des entrées dépasse le seuil, la sortie produite vaut +1, sinon la sortie vaut -1. Deux schémas du neurone formel sont proposés sur les deux figures (IV .1) et (IV .2) :

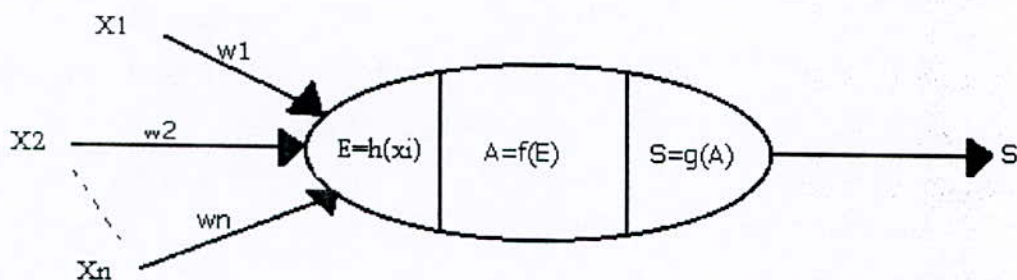


Figure IV- 1 : Schéma relationnel d'un réseau de neurone

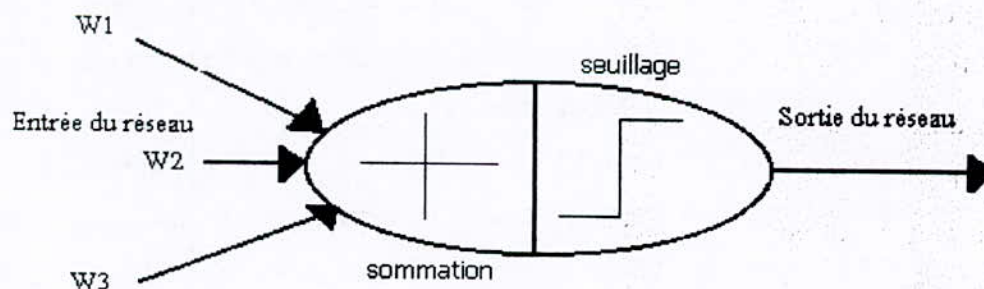


Figure IV- 2 : Représentation schématique d'un réseau de neurone

h : Fonction d'entrée totale ;

f : Fonction d'activation ;

g : Fonction de sortie.

x_i : Différents signaux d'entrées.

w_i : Différents poids des connections.

$E = h(x_i, w_i)$ (E: entrée totale).

$A = f(E)$ (A: état d'activation du neurone).

$S = g(A)$ (S : sortie du neurone).

Le fonctionnement d'un neurone peut-être décrit comme la réaction à une somme pondérée d'excitations.

Le neurone reçoit les entrées x_1, x_2, \dots, x_n et calcule le potentiel ou la somme pondérée des entrées : $x_1W_1 + x_2W_2 + \dots + x_nW_n$. Ensuite, Une décision est prise pour calculer la sortie S en fonction du seuil θ :

Si $W > \theta$, Alors $S = +1$.

Sinon $S = -1$.

Ce qui revient à tester si la différence $(W - \theta) > 0$. Il est donc équivalent de remplacer le seuil par une entrée fixe, de valeur (-1), avec un poids variable. Ceci est représenté par une fonction à seuil, appelée fonction de décision ou d'activation.

IV- 5- Les types des réseaux de neurones :

On distingue deux types de réseaux de neurones :

IV- 5- 1- Réseaux mono-couche :

➤ Perceptron :

A partir des découvertes de Mc Culloch et Pitts qui aboutirent à la découverte du neurone formel en 1943, et des recherches de Hebb qui donnèrent une explication synaptique du mécanisme d'apprentissage en 1949, Rosenblatt a pu établir le premier modèle performant de réseau de neurones formels vers la fin des années 50.

Le but du perceptron mono-couche est d'associer des configurations ou des formes en entrée à des réponses ; c'est un type de réseau utilisé surtout pour le classement. Il se compose de deux couches de neurones formels appelées « rétine » et « couche de sortie ». La couche rétine est une contraction de la couche d'entrée et de la couche de traitement des informations. Cette contraction a été effectuée pour des raisons de simplifications car la couche d'entrée est une couche inerte qui ne participe pas directement au traitement de l'information. Les cellules de la rétine, lorsqu'elles sont excitées, répondent par « oui » ou par « non » (1 ou 0) à leur sortie, car elles sont régies par une fonction binaire à seuil. (MINSKY, PAPERT, 1985).

La rétine est reliée à la couche de sortie grâce à une série de connexions ou des synapses de poids variables. (Figure IV .3).

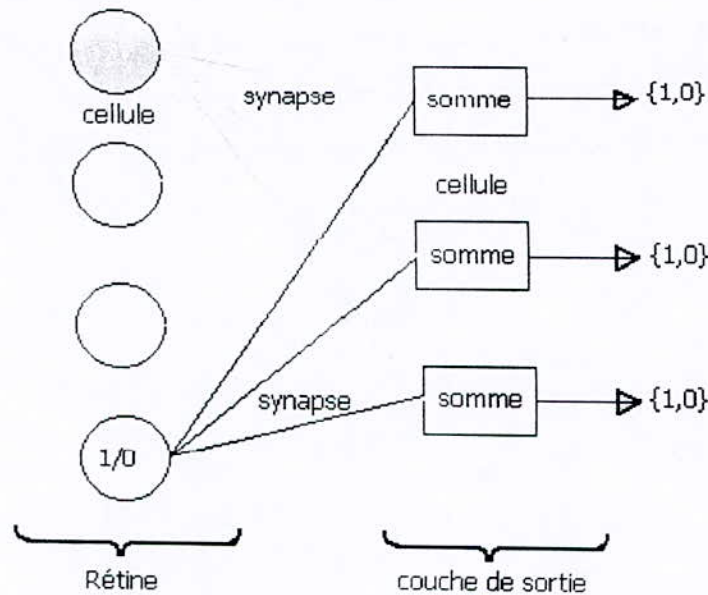


Figure IV- 3 : Schéma perceptron mono-couche

De manière générale, l'apprentissage du perceptron mono-couche se fait en modifiant l'intensité de ces synapses par la variation de leurs poids.

Le signal émis par les neurones formels de la rétine et transmis par l'ensemble des synapses, est évalué par les cellules de sorties.

➤ Adaline :

L'apprentissage d'un réseau de neurones formels se fait en présentant une série de couples entrée-sortie réels et connus, codés sous forme de vecteurs. Il est donc indispensable de disposer d'une base de données fiable et assez vaste et que le codage des informations soit fait correctement pour avoir les résultats les plus significatifs possible. Les poids des connexions sont modifiés selon l'erreur obtenue.

Le modèle mis au point par Widrow et Hoff en 1960 (ADaptive LINar Element) est une amélioration de la méthode d'apprentissage du perceptron. A la différence de la première, celle-ci prend en compte le signal reçu même s'il n'atteint pas le seuil requis pour l'obtention de la réponse.

IV- 5- 2- Réseaux multicouches :

➤ **Perceptron :**

Bien que les règles d'apprentissage soient relativement simples, les réseaux mono-couches demeurent limités au calcul de fonction très simple. A partir de là les réseaux multicouches formés à partir du même neurone formel, mais sont organisés autrement. En effet les cellules sont disposées cette fois ci dans plusieurs couches intermédiaires entre la couche d'entrée et la couche de sortie. Chaque neurone est connecté à l'ensemble des neurones de la couche suivante et précédente par des connexions de poids variables (Figure IV. 4).

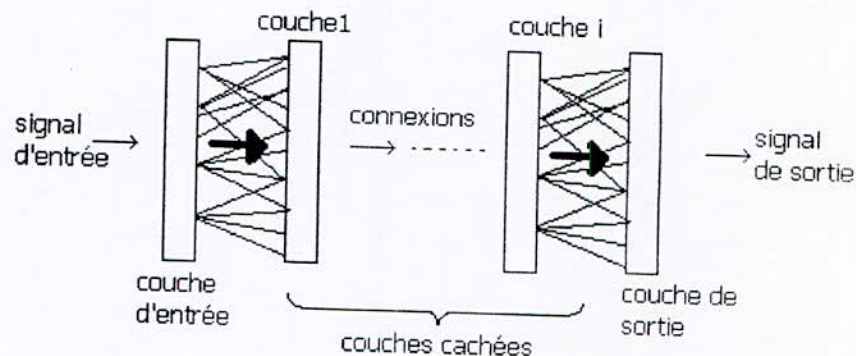


Figure IV- 4 : Réseau multicouche

La différence majeure avec les réseaux mono-couches réside dans les méthodes d'apprentissage qui doivent répondre à une plus grande complexité. En effet, il s'agit maintenant de gérer un plus grand nombre de connexions se trouvant dans plusieurs couches de connexions toutes modifiables. L'algorithme d'apprentissage pour ce genre de réseaux est l'algorithme de rétro-propagation du gradient ou de back propagation du gradient (TOUZET, 1992).

➤ **Propriétés :**

Le perceptron multicouches connaît de nombreuses applications essentiellement grâce à sa propriété d'approximation. Il a été montré qu'un réseau à une couche cachée est capable d'approximer n'importe quelle fonction avec la précision souhaitée, à condition que le nombre de neurones cachés soit suffisant. C'est la propriété essentielle de ce réseau, qui confère la capacité d'un approximateur universel. Néanmoins, cette propriété ne permet pas de choisir, pour un problème donné, le nombre d'unités optimal dans une couche cachée. Cette question fait encore l'objet de recherches théoriques.

➤ **Pratique :**

Le réseau de neurones a été développé à partir de trois couches connues respectivement sous le nom de : couche d'entrée, couche cachée et couche de sortie.

Chaque couche est constituée de plusieurs nœuds appelés neurones, les couches sont interconnectées entre elles par des jeux de poids.

Les nœuds appartenant à la première couche reçoivent l'entrée à partir de l'extérieur du modèle (entrée initiale). Cependant, les neurones des couches cachées et de sortie reçoivent les entrées à partir des interconnexions.

Les nœuds des couches cachées et de sortie transforment les entrées en sortie appelées "firing rate" en utilisant une fonction de transformation associée à chaque nœud connue sous le nom de "fonction sigmoïde".

En présentant un jeu de données aux nœuds de la couche d'entrée " I_i ", $i = 1, \dots, k$, ces valeurs sont multipliées par les poids des interconnexions $(W_{input})_{ij}$, $j = 1, \dots, h$, où $(W_{input})_{ij}$ sont les poids des interconnexions reliant le $i^{\text{ème}}$ nœud d'entrée au $j^{\text{ème}}$ nœud d'entrée caché et " h " étant le nombre de nœud cachés.

Les produits : $I_i * (W_{input})_{ij}$ sont sommés sur l'indice " i " et deviennent les entrées des nœuds cachés, comme le montre l'équation suivante :

$$(H_{input})_j = \sum_{i=1}^k I_i (W_{input})_{ij} \quad (IV.1)$$

Où : $(H_{input})_j$ est l'entrée du " $j^{\text{ième}}$ " nœud caché.

$(W_{input})_{ij}$ est le poids des interconnexions reliant le $i^{\text{ème}}$ nœud d'entrée au $j^{\text{ème}}$ nœud caché.

Chaque entrée des nœuds cachés est transformée à l'aide de la fonction sigmoïde définie plus haut pour donner les sorties de nœuds cachés $(H_{output})_j$, de là chaque sortie est multipliée par le poids interconnecté $(W_{output})_{jn}$, $n = 1, \dots, m$, entre le $j^{\text{ème}}$ nœud caché et chaque " m " nœuds de sortie.

L'entrée des " m " nœuds de sortie $(O_{input})_n$ sera la somme des produits exprimée comme suit :

$$(O_{input})_n = \sum_{j=1}^h (H_{input})_j * (W_{output})_{jn} \quad (IV.2)$$

Ces entrées $(O_{input})_n$ sont transformées à travers la fonction sigmoïde pour donner les sorties finales des " m " nœuds existants $(O_{output})_n$.

IV- 6- Les fonctions d'activation :

Différentes fonctions de transfert pouvant être utilisées comme fonction d'activation du neurone. Les trois les plus utilisées sont les fonctions à «seuil» (en anglais «hard limit»), «linéaire» et «sigmoïde», qui prennent les formes suivantes :

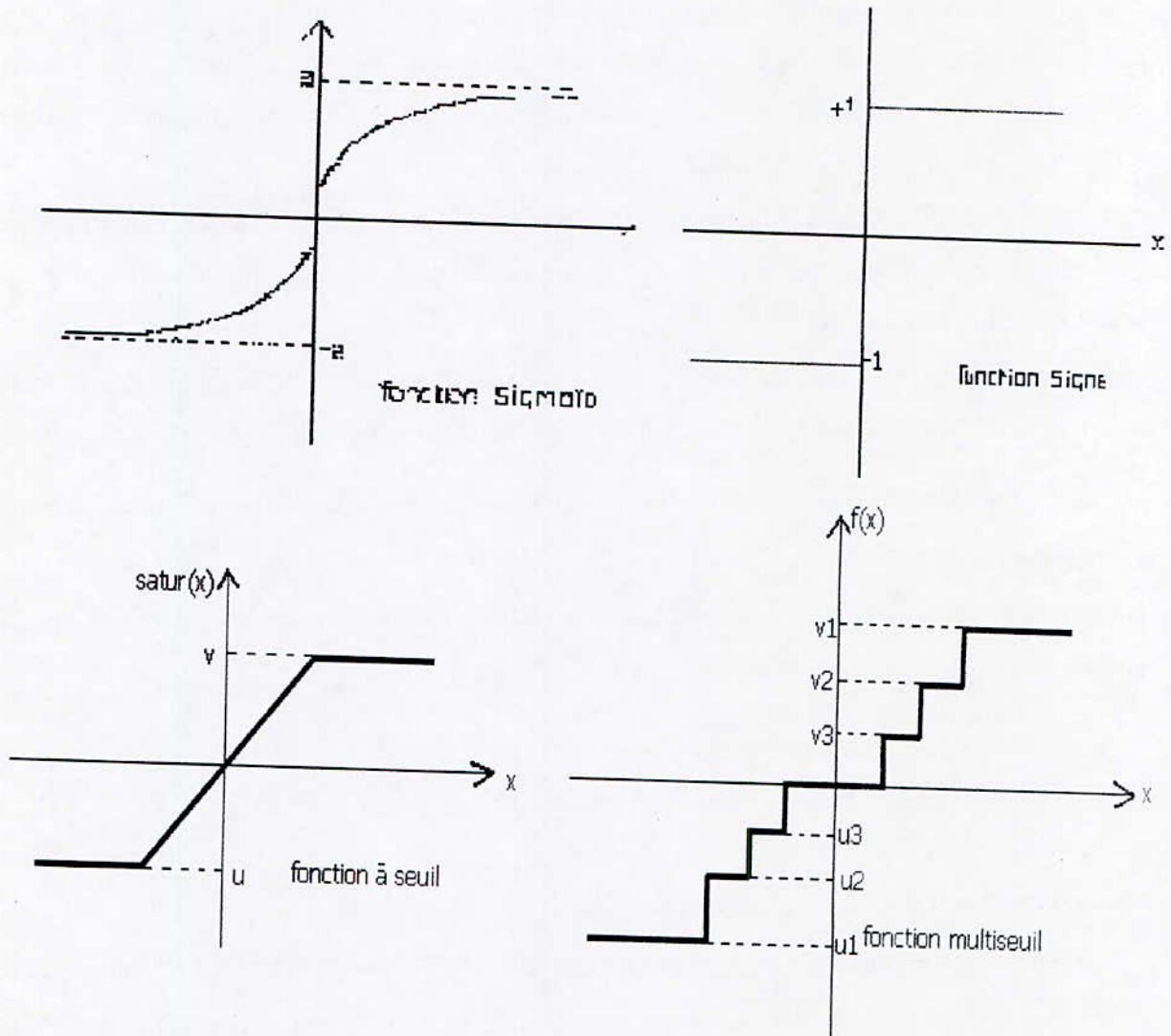


Figure (IV- 5) : Différentes formes de fonctions d'état

Enfin La fonction d'état peut-être tout autre fonction choisie croissante et impaire.

IV- 7- Apprentissage d'un réseau de neurones :

L'apprentissage est une phase du développement d'un réseau de neurones durant lequel le comportement du réseau est modifié jusqu'à l'obtention du comportement désiré.

L'apprentissage est défini comme étant n'importe quel changement opéré dans la mémoire du réseau. Ainsi cette modification affecte les poids synaptiques qui relient les neurones entre eux.

L'apprentissage a comme objectif l'amélioration des performances futures du réseau, sur la base d'une connaissance acquise au fur et à mesure des expériences passées.

Il existe plusieurs règles d'apprentissage qui peuvent être regroupées en trois catégories :

IV- 7- 1- Apprentissage supervisé : associe une réponse spécifique à chaque signal d'entrée. La procédure s'effectue progressivement, jusqu'à ce que l'erreur entre la sortie du réseau et le résultat désiré soit minimisée.

La technique d'apprentissage supervisé généralement utilisée est la rétropropagation de l'erreur appelée "*Backpropagation*".

➤ **Backpropagation :**

Cette technique a été élaborée en 1972 par Werbos dont le nom était "*The Dynamic Algorithm feed-back*".

Cet algorithme est une généralisation aux réseaux de neurones formels multicouches du modèle de Windrow Hoff. Il a été mis au point par deux équipes de recherche simultanément : une équipe française (Fogelman-soulié, Gallinari et Lecun) et une équipe américaine (Rumelhart, Hinton et Williams (DAVALO, 1993)). L'idée consiste à l'utilisation d'une fonction sigmoïde dérivable en remplacement de la fonction à seuil couramment utilisée.

➤ **Algorithme de Rétropropagation de l'erreur :**

Rétro propager l'erreur de la couche de sortie vers les couches internes, en suivant le chemin inverse du signal provenant de la couche d'entrée et allant vers la couche de sortie.

Cet algorithme permet de dépasser le problème du « Credit Assignment Problem ».

On considère un réseau de "*n*" neurones de sorties, avec un nombre quelconque de couches cachées.

Rappelons les notations utilisées :

- $X = (X_1, X_2, \dots, X_n)$ est le vecteur des entrées.
- $Y = (Y_1, Y_2, \dots, Y_m)$ est le vecteur des sorties désirées.
- $S = (S_1, S_2, \dots, S_m)$ est le vecteur des sorties effectivement obtenues.
- f est la fonction sigmoïde d'un neurone, f' est sa dérivée.
- O_j est la sortie du neurone "j".
- I_i est l'entrée du neurone "i".
- $E(k)$ est le pas du gradient à l'étape "k".

L'erreur quadratique entre la sortie effective et la sortie désirée est considérée comme une fonction des poids variables des connexions :

$$E^k(w) = \sum_i (Y_i^k - S_i^k)^2 \quad i = 1, m \quad (\text{IV.4})$$

Où $E^k(w)$ est l'erreur commise sur un exemple k.

L'algorithme de rétropropagation consiste à effectuer une descente de gradient sur "E".

En général on utilise une version qui est une approximation, où on effectue une modification de chaque poids après chaque passage d'un exemple "k", d'où la règle de modification des poids :

$$W_{ij}(k) = W_{ij}(k-1) - e(k) * d_i * O_j \quad (\text{IV.5})$$

Tel que :

k : Indice de l'étape de l'algorithme.

$e(k)$: Pas du gradient à l'étape k.

d_i : Coefficient calculé de proche en proche.

Avec :

$$1. \quad d_i = 2 * (S_i - Y_i) * f'(I_i) \quad \text{pour la couche de sortie.} \quad (\text{IV.6})$$

$$2. \quad d_i = \sum_h d_h * W_{hi} * f'(I_i) \quad \text{pour les couches cachées.} \quad (\text{IV.7})$$

Où h : Indice des neurones vers lesquels le neurone i envoie des connexions.

IV- 7- 2- Apprentissage non supervisé : consiste à ajuster les poids à partir d'un seul ensemble d'apprentissage formé uniquement de données, aucun résultat désiré n'est fourni au réseau. C'est le réseau qui assure la correction des poids.

IV- 7- 3- Apprentissage renforcé : représente une technique nouvelle, qui a en commun avec l'apprentissage supervisé, la présence d'un superviseur régi par une mesure de degrés de performance qui juge l'évolution de l'apprentissage du réseau après quelques itérations, et une similitude avec l'apprentissage non supervisé, puisqu'il nécessite que des variables d'entrées, sans définir les variables à la sortie du réseau.

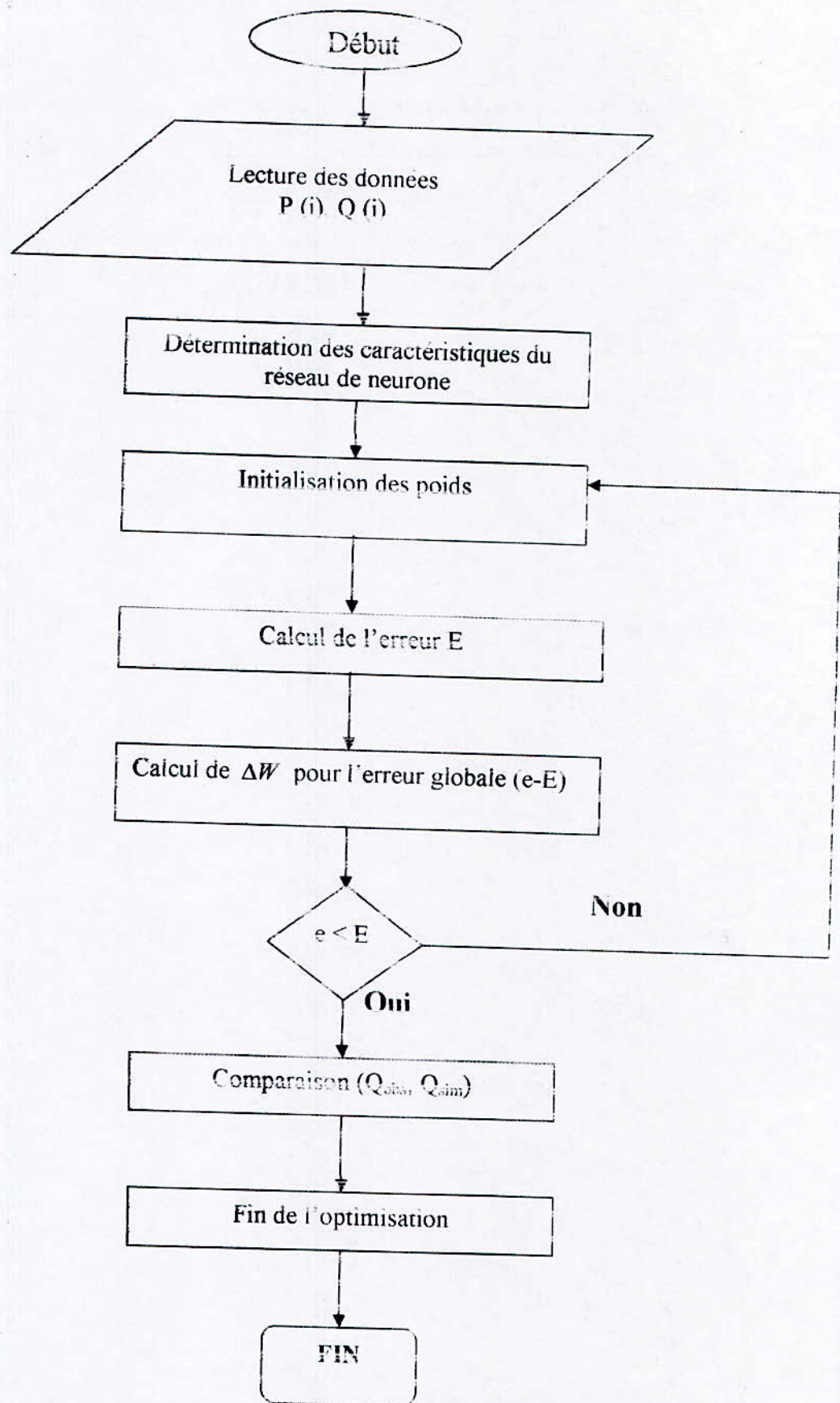


Figure (IV- 6) : Organigramme des réseaux de neurones

Chapitre V

Modélisation par l'Autorégressif à
Moyenne Mobile
avec une variable exogène

ARMAX

V- 1- Introduction:

La plupart des modèles utilisés de nos jours sont des modèles stochastiques. Selon FELLER (1968) : il s'agit de tous modèle utilisant la théorie des probabilités pour représenter l'incertitude.

Dans ce chapitre, un aperçu sur les modèles ARMA pour le cas de variables à coefficients constants et aussi pour le cas de variables périodiques est présenté. Nous donnons également une description sur une famille du modèle ARMA qui sera utilisée dans ce travail et que l'on appelle le modèle ARMAX.

V- 2- Modèle ARMA à coefficients constants : (Box et Jenkins, 1970)

La famille des modèles de Box-Jenkins, dans sa formulation la plus générale, permet de présenter un vaste ensemble de séries temporelles. Il s'agit d'un modèle linéaire. Une hypothèse souvent plus contraignante que la linéarité, et la stationnarité : Les modèles Box-Jenkins ne permette que de modéliser de séries dont les propriétés statistiques n'évoluent pas dans le temps. Cependant, il est en général possible d'effectuer une transformation sur les séries de données pour les rendre stationnaires. Par exemple la transformation $Z_t = Q_t - Q_{t-1}$, qui correspond à la première différence du débit est souvent satisfaisante. Lorsque cette première différence n'est pas stationnaire, on peut considérer la deuxième différence, notée :

$$z_t = (Q_t - Q_{t-1}) - (Q_{t-1} - Q_{t-2}) = Q_t - 2Q_{t-1} + Q_{t-2} \quad (V.1)$$

Un modèle Box-Jenkins comprend deux parties : une composante autorégressif (AR) et une composante à moyenne mobile (MA).

➤ **Modèle autorégressif (AR) :**

Un modèle autorégressif linéaire d'une transformation z, du débit Q, relie la valeur de z, au temps t à ses valeurs précédentes de la façon suivante :

$$Z_t = \gamma + \varphi_1 Z_{t-1} + \varphi_2 Z_{t-2} + \dots + \varphi_p Z_{t-p} + \varepsilon_t \quad (V.2)$$

Où $\gamma, \varphi_i (i = 1, 2, \dots, p), p$ étant des coefficients d'ajustement, ε_t est un bruit, $N(0, \sigma)$. Un modèle à p termes (en plus de ε_t et γ), est dénoté $AR(p)$. La valeur de p, que l'on nomme l'ordre du modèle, doit être fixée à priori, alors que les autres paramètres sont estimés par calibration. En général il est nécessaire d'essayer plusieurs valeurs de p et observer la qualité des ajustements ainsi obtenus avant de faire un choix final. Il existe plusieurs méthodes

permettant de choisir une valeur de p donnant un ajustement adéquat, pour cela il est recommandé d'utiliser les tests statistiques (Box and Jenkins 1970).

➤ **Modèle à moyenne mobile (MA) :**

Un modèle à moyenne mobile d'une transformation z_t du débit Q_t relie la valeur z_t au pas de temps t aux erreurs ' ε ', précédentes de la façon suivante :

$$z_t = \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t = \sum_{i=1}^q \theta_i \varepsilon_{t-i} + \varepsilon_t \quad (V.3)$$

Où : θ_i ($i = 1, 2, \dots, q$) et q sont des coefficients d'ajustement, et ε_t le bruit normal $N(0, \sigma)$. Un modèle comportant q termes, est dénoté $MA(q)$. La valeur de q , que l'on nomme l'ordre du modèle, doit être fixée à priori, alors que les autres paramètres sont estimés par calibration. Comme dans le cas d'un modèle AR, différentes valeurs de q peuvent être essayées avant d'effectuer un choix final.

La combinaison d'un modèle autorégressif AR et d'un modèle de moyenne mobile MA conduit à un modèle ARMA qui peut modéliser les débits des différents pas de temps des cours d'eau.

L'avantage de ce modèle est qu'il conduit à un nombre réduit de paramètres par rapport à un modèle autorégressif d'un grand ordre.

V- 3- Modèle ARMA des séries périodiques :

Soit une série originale $X_{v,\tau}$ où v désigne l'année et $\tau = 1, 2, \dots, w$, avec w le nombre d'intervalles de l'année (mois, jours,...). Si $X_{v,\tau}$ est asymétrique, une transformation doit être utilisée pour la rendre symétrique, soit $Y_{v,\tau}$ la nouvelle série.

Le modèle ARMA périodique d'une variable $Y_{v,\tau}$ peut être écrit sous la forme .

$$Y_{v,\tau} = \mu_\tau + \sigma_\tau \cdot z_{v,\tau} \quad (V.4)$$

Où : μ_τ : est la moyenne périodique de $Y_{v,\tau}$;

σ_τ : est l'écart type périodique de $Y_{v,\tau}$.

$z_{v,\tau}$ peut être modélisée par un modèle ARMA dont les coefficients sont périodiques.

Le modèle ARMA périodique s'écrit comme suit :

$$z_{V,t} = \sum_{j=1}^p \phi_{j,t} \cdot z_{V,t-j} - \sum_{i=1}^q \theta_{i,t} \cdot \varepsilon_{V,t-i} + \varepsilon_{V,t} \quad (\text{V. 5})$$

Où : $\phi_{j,t}$ et $\theta_{i,t}$: Sont les coefficients autorégressifs et de moyennes mobiles ;

$\varepsilon_{V,t}$: Variables périodiques indépendantes et distribuées selon la loi normale.

Tao et Delleur (1976) , Delleur et al. (1976) avaient utilisé un modèle ARMA (p, q) avec des paramètres périodiques pour la généralisation des débits des rivières.

Il faut signaler que l'utilisation de ce modèle ARMA périodique n'est pas fréquente.

V- 4- Le modèle ARMAX :

Le modèle ARMAX (AutoRegressive Moving Average with eXogenous input), issue de la théorie des systèmes linéaires en automatique (BOX and JENKINS, 1970) constitue un outil mathématique dont les possibilités ont été très tôt reconnues par les hydrologues.

La structure du modèle ARMAX (p, q) simplifié à une entrée et une sortie, sans délai entre l'entrée et la sortie (DELLEUR et BASTIN, 1984), s'écrit :

$$\widehat{Y}_j + c_1 * Y_{j-1} + c_2 * Y_{j-2} + \dots + c_p * Y_{j-p} = b_0 * x_j + b_1 * x_{j-1} + b_2 * x_{j-2} + b_q * x_{j-q} + e_j \quad (\text{V. 6})$$

Avec :

\widehat{Y}_j : Sortie modélisée au temps j.

Y_j : Sortie observée au temps j.

x_j : Entrée observée au temps j.

c_1, c_2, \dots, c_p : Coefficients de la partie autorégressive du modèle.

b_0, b_1, \dots, b_q : Coefficients de la partie moyenne mobile sur la variable exogène (x_j).

e_j : L'erreur de prévision ou de reconstitution du modèle données par :

$$e_j = Y_j - \widehat{Y}_j \quad j = 1, 2, 3, \dots \quad (\text{V. 7})$$

En substituant dans l'équation (V.7), on obtient :

$$\begin{aligned} & \hat{Y}_j + c_1 * \hat{Y}_{j-1} + c_2 * \hat{Y}_{j-2} + \dots + c_p * \hat{Y}_{j-p} \\ & = b_0 * x_j + b_1 * x_{j-1} + \dots + b_q * x_{j-q} + c_1 * e_{j-1} + c_2 * e_{j-2} + \dots + c_p * e_{j-p} \end{aligned} \quad (V.8)$$

Et introduisant l'opérateur délai Z^{-s} tel que :

$$Z^{-s} * Y_j = Y_{j-s}$$

$$\text{D'où : } Z^{-s} * \hat{Y}_j = \hat{Y}_{j-s}$$

$$Z^{-s} x_j = x_{j-s} \quad (V.9)$$

On obtient :

$$\begin{aligned} & (1 + c_1 * Z^{-1} + c_2 * Z^{-2} + \dots + c_p * Z^{-p}) * \hat{Y}_j \\ & = (b_0 + b_1 * Z^{-1} + b_2 * Z^{-2} + \dots + b_q * Z^{-q}) * x_j - (c_1 * Z^{-1} + c_2 * Z^{-2} + \dots + c_p * Z^{-p}) * e_j \end{aligned} \quad (V.10)$$

Et en posant :

$$C(Z^{-1}) = c_1 * Z^{-1} + c_2 * Z^{-2} + \dots + c_p * Z^{-p}$$

$$B(Z^{-1}) = b_0 + b_1 * Z^{-1} + \dots + b_q * Z^{-q} \quad (V.11)$$

On peut écrire :

$$(1 + C(Z^{-1})) * \hat{Y}_j = B(Z^{-1}) * x_j - C(Z^{-1}) * e_j \quad (V.12)$$

Ou encore :

$$\hat{Y}_j = \frac{B(Z^{-1})}{1 + C(Z^{-1})} * x_j - \frac{C(Z^{-1})}{1 + C(Z^{-1})} * e_j \quad (V.13)$$

On voit que la sortie du modèle \hat{Y}_j dépend non seulement des entrées mais aussi des erreurs de prévision (ou de reconstitution) des pas de temps précédents (terme $\frac{C(Z^{-1})}{1 + C(Z^{-1})} * e_j$).

V- 5- Le modèle ARMAX sur la variation de débit :

Si l'on considère la variation de débit comme variable endogène et la pluie efficace comme variable exogène, on peut réécrire l'équation (VI.6) du modèle ARMAX(p, q) simplifié comme ceci :

$$q_{j,t} + c_1 * q_{j-1,t} + c_2 * q_{j-2,t} + \dots + c_p * q_{j-p,t} = b_0 * PE_{j,t} + b_1 * PE_{j-1,t} + \dots + b_q * PE_{j-q,t} + e_{j,t} \quad (V. 14)$$

Avec :

$q_{j,t}, PE_{j,t}, e_{j,t}$: Respectivement la variation de débit, la pluie efficace et l'erreur pour le $j^{ième}$ pas de $t^{ième}$ épisode.

Ou alternativement :

$$(1+C(Z^{-1})) * q_{j,t} = B(Z^{-1}) * PE_{j,t} + e_{j,t} \quad (V. 15)$$

Avec :

$C(Z^{-1}), B(Z^{-1})$: Polynômes de Z^{-1} .

V- 6- Le modèle ARMAX sur le débit :

Dans ce cas, le modèle ARMAX (p, q) analogue à celui de l'équation (V.14) peut s'écrire comme ceci :

$$Q_{j,t} + c'_1 * Q_{j-1,t} + c'_2 * Q_{j-2,t} + \dots + c'_p * Q_{j-p,t} = b'_0 * PE_{j,t} + b'_1 * PE_{j-1,t} + \dots + b'_q * PE_{j-q,t} + e_{j,t} \quad (V. 16)$$

Avec :

$Q_{j,t}$: Le debit total au $j^{ième}$ pas de temps du $t^{ième}$ épisode.

Ou autrement :

$$(1+C'(Z^{-1})) * Q_{j,t} = B'(Z^{-1}) * PE_{j,t} + e_{j,t} \quad (V. 17)$$

Avec :

$$C'(Z^{-1}) = c'_1 * Z^{-1} + c'_2 * Z^{-2} + \dots + c'_p * Z^{-p}$$

$$B'(Z^{-1}) = b'_0 + b'_1 * Z^{-1} + \dots + b'_q * Z^{-q} \quad (V. 18)$$

$e_{j,t}$: L'erreur de prévision (ou reconstitution) du modèle.

Le système de transformation de la pluie efficace en débit de ruissellement direct présenté sur la figure (V- 1), peut être considéré comme linéaire.

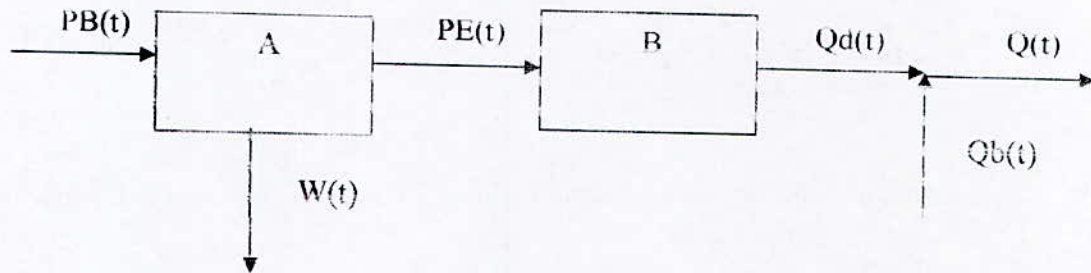


Figure (V- 1) : Décomposition du modèle globale pluie-débit en sous-modèles

Tel que :

PB : Pluie brute ;

PE : Pluie efficace ;

W : Pertes ;

Qd : Ruissellement direct ;

Qb : Débit de base ;

Q : Débit total ;

A : Sous-modèle de production de la pluie efficace ;

B : Sous-modèle de transformation (PE engendre le débit de ruissellement Qd).

Alors on est capable de modéliser ce système par un modèle ARMAX général qui peut s'écrire comme suit :

$$Q_{dj} + \sum_{i=1}^m c_i * Q_{dj-i} = \sum_{i=0}^n b_i * PE_{j-i-1} + e_j \quad (V. 19)$$

Où :

Qd, PE : Sont respectivement le débit de ruissellement direct et la pluie efficace.

$c_i, i = 1, \dots, m$: Coefficients de la partie autorégressive sur Qdj.

$b_i, i = 0, \dots, n$: Coefficients de la partie moyenne mobile sur la variable exogène PE_j .

e_j : Erreur de reconstitution ou de prévision du modèle selon qu'on s'intéresse à la phase de calage ou à la prévision en dehors de la période de calage. Cette erreur est égale à la différence entre la sortie réelle et la sortie modélisée.

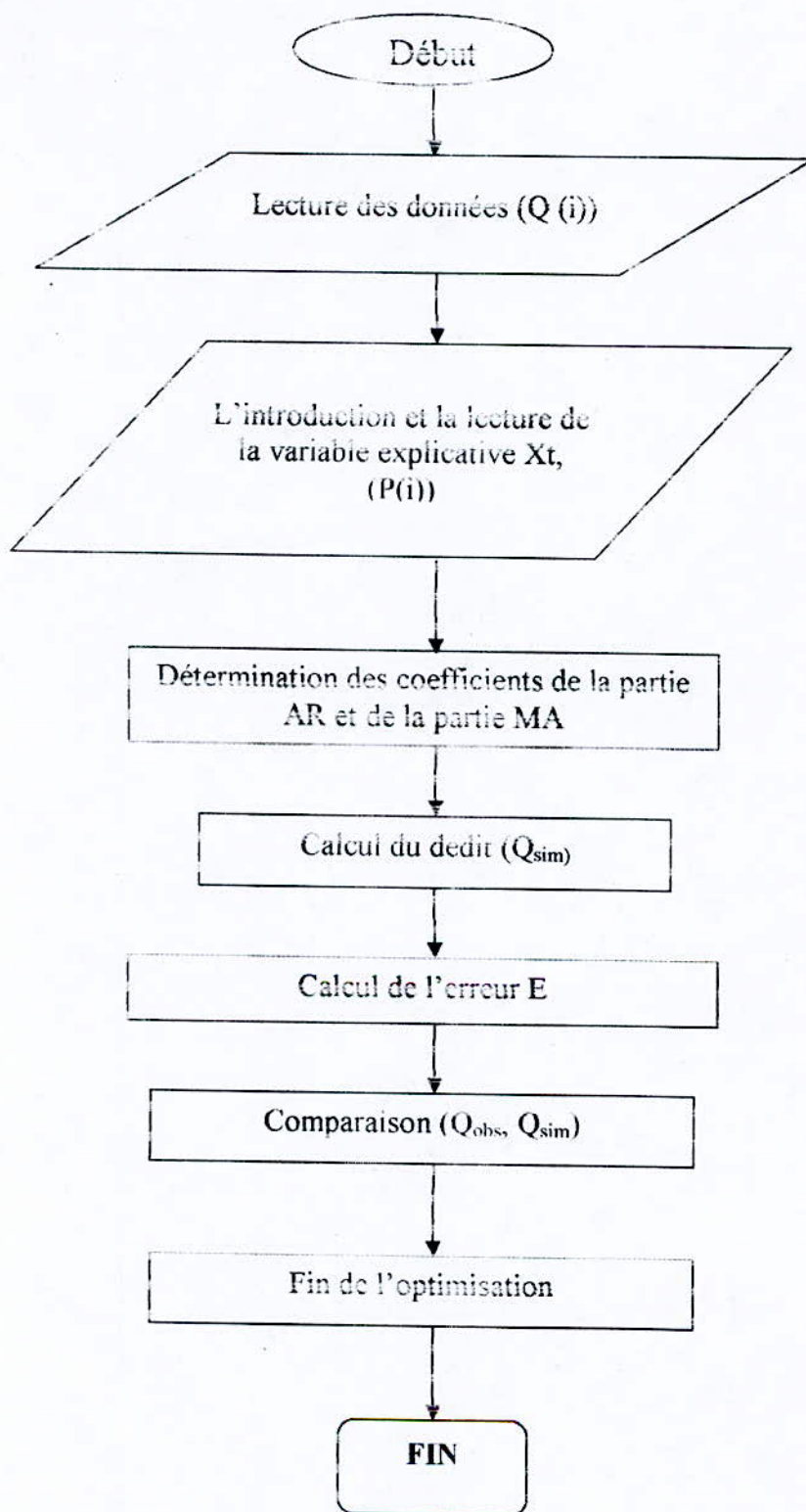


Figure (V- 2) : Organigramme du modèle ARMAX

Chapitre VI

Application sur les données du bassin
versant de l'oued EBDA

VI- 1- Introduction :

Après une étude théorique des différentes approches, nous procédons à leur application sur des données réelles, qui nous permettra de juger les performances et la fiabilité de ces méthodes ainsi que la qualité des résultats qu'elles génèrent.

Ce chapitre comportera trois parties : la première sera consacrée à la présentation du bassin versant et les événements averse-crue, la deuxième portera sur l'application des méthodes DPFT, Réseaux de Neurones et ARMAX sur les données du bassin de l'oued EBDA ainsi que l'interprétation des résultats obtenues, et la dernière sera consacrée à la modélisation par le modèle qui donne des meilleurs résultats avec une variation dans le nombre de crues retenus pour le calage et la validation pour avoir le seuil suffisant de la modélisation des débits dans notre cas.

VI- 1- 1- Caractéristiques physiques du sous bassin de Oued EBDA :

Tableau (VI- 1) : Les caractéristiques du sous bassin de l'Oued EBDA

Surface	270 km ²
Périmètre	70 km
Coefficient de compacité de Gravélius	$K = (0,28 P / \sqrt{S}) = 1,19$
Longueur du cours d'eau principale	37
Longueur du rectangle équivalent	23.50
Indice de pente de Roche (I. P)	0.20
Densité de drainage (Dd)	5.40 km/km ²
Coefficient de torrentialité	128
Altitude Max	1417 m
Altitude Min	275 m
Altitude Moy	736 m

Le cours d'eau étudié est celui de l'oued EBDA, sous bassin versant du CHELIFF, de la station hydrométrique ARIBS EBDA de code 01-18-01. Celle-ci a été mise en service en 1971 et permet le contrôle des hauteurs d'eau. Les données de débits recueillies sont au pas de temps horaire et s'étalent entre 1982 et 1995.

En ce qui concerne les données des pluies, au pas de temps horaire au niveau de la station pluviométrique de ARIBS EBDA, elles s'étalent de 1972 jusqu'à 2001.

Ces fichiers de données nous ont été aimablement fournis par l'ANRH de BLIDA.

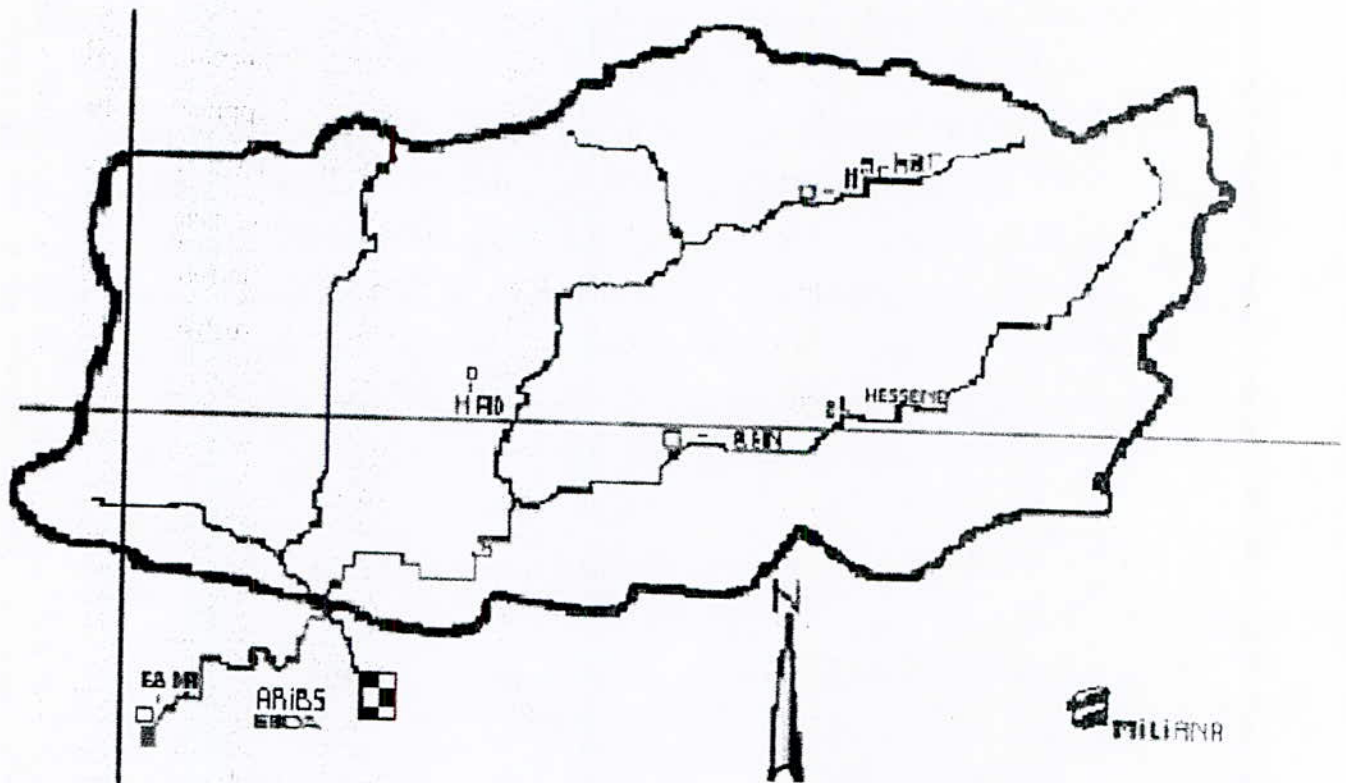


Figure (VI-1) : Le sous bassin versant du Cheliff Oued EBDA

VI- 1- 2- Choix des épisodes averse-crue :

Bien que les données de pluies et de débits soient disponibles et s'étalent sur des périodes considérables, nous n'avons pas pu faire sortir un nombre important d'événements averse-crue ; car dans la plupart des cas rencontrés nous avons eu affaire à des crues complexes ou multiples. Ces dernières nécessitent une décomposition très minutieuse pour pouvoir séparer les crues successives, et déterminer leur début et leur fin, en plus, il fallait mettre en évidence l'averse qui a généré chaque crue. Quelques épisodes ont été éliminés à cause du grand décalage entre les pluies et les débits, c'est-à-dire que les valeurs des débits n'étaient pas significatives et ne reflétaient pas celles enregistrées pour les précipitations.

Néanmoins et à défaut des crues simples (crue à un seul pic) nous étions contraints de prendre quelques épisodes contenant plusieurs crues composées et complexes pour la constitution de l'échantillon d'étude. Enfin, la taille de l'échantillon retenu pour notre étude est de vingt (20) épisodes averse-crue.

VI- 2- Application de la méthode DPFT sur l'Oued EBDA du bassin Cheliff :**VI- 2- 1- Procédure de résolution :**

La méthode est appliquée selon la démarche citée au chapitre III.

On utilisera :

- Nombre d'itération : 5.
- Identification des coefficients de la DPFT par la méthode des moindres carrés.
- Application de la contrainte de positivité sur les coefficients de la DPFT identifiés.
- Lissage exponentiel de l'extrémité de la fonction de transfert (l'allure).
- Normalisation de la fonction de transfert.
- Déconvolution des pluies efficaces.
- Application de la contrainte : pluies efficaces inférieures aux pluies brutes.

Les paramètres que nous avons fait varier sont :

- La mémoire d'identification de la fonction de transfert.
- Le coefficient de la ridge régression.
- Point du début de lissage de la fonction de transfert.

Pour une comparaison plus rigoureuse entre les débits calculés par la méthode DPFT et les débits observés, une série de critères de validation numérique seront établis pour les période calage et test ainsi qu'un critère graphique qui permet de juger l'évolution de l'échantillon calculé par rapport à l'observé.

* critères de validation numériques :

1- Le coefficient de détermination est donné par :

$$R^2 = \left[\frac{Cov(x, y)}{\sigma_x \cdot \sigma_y} \right]^2 \quad (VI. 1)$$

Où :

R^2 : Coefficient de détermination entre les débits calculés et observés :

$$Cov(x, y) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y}) \quad (VI. 2)$$

σ_x et σ_y : Ecart type des séries de débits observés et calculés ;

x et y : Débits observé et calculé.

2- Le critère de comparaison de Nash est défini par la formule suivante :

$$Nash (\%) = 100 \cdot \left(1 - \frac{\sum_{i=1}^n (Q_{i,c} - Q_{i,o})^2}{\sum_{i=1}^n (Q_{i,o} - \bar{Q}_o)^2} \right) \quad (VI. 3)$$

Avec :

$Q_{i,o}$: Débit observé en (m^3/s).

$Q_{i,c}$: Débit calculé en (m^3/s).

\bar{Q}_o : Moyenne des débits observés en (m^3/s).

3- La RMSE c'est l'erreur quadratique moyenne standard (Root Mean Standard Error), tel que :

$$RMSE = \left(\frac{1}{N} \left(\sum_{i=1}^N (Q_c - Q_o)^2 \right) \right)^{1/2} \quad (VI. 4)$$

a- Identification de la fonction de transfert :

Le calage du modèle DPFT consiste à déterminer la fonction de transfert. Le nombre d'épisodes pluie-débit utilisé pour le calage est égale à quinze (15) crues c'est-à-dire 75% de la taille de l'échantillon, ce qui constitue un nombre relativement faible par rapport aux recommandations de NALBANTIS, qui préconise l'utilisation de vingt (20) à cinquante (50) épisodes (NALBANTIS, 1987), et le nombre de crues de validation est de cinq (5) soit 25% du nombre totale de l'échantillon.

Par ailleurs, les caractéristiques de la fonction de transfert retenues après avoir varié les paramètres cités précédemment sont les suivants :

- La mémoire d'identification : $k = 13$ pas de temps ;
- Un lissage effectué à partir du 6^{ième} point avec un coefficient de lissage $\alpha = 0.40$.
- Le coefficient de ridge régression.

b- Déconvolution des pluies efficaces :

Les pluies efficaces sont obtenues à partir du système de déconvolution :

$$\{q\} = [a] * \{PE\} \quad (VI. 5)$$

$$\{q\} = \begin{Bmatrix} q_1 \\ q_2 \\ \dots \\ q_n \end{Bmatrix}; \quad \{PE\} = \begin{Bmatrix} PE_1 \\ PE_2 \\ \dots \\ PE_m \end{Bmatrix}; \quad [a] = \begin{bmatrix} a_1 0 \dots \dots \dots 0 \\ a_2 a_1 \dots \dots \dots 0 \\ \dots \dots \dots \dots \dots \dots \\ a_n a_{n-1} \dots \dots \dots a_{n-m+1} \end{bmatrix}$$

La déconvolution consiste à déterminer les pluies efficaces à partir de la fonction de transfert calculée à l'itération (i-1) et du débit observé.

Le coefficient de ridge régression a été pris égale à 10^{-1} . Il a été remarqué que pour des valeurs supérieures à 10^{-1} , le système de déconvolution était perturbé, et le coefficient de détermination entre le débit observé et calculé est très faible, par contre les valeurs inférieures à 10^{-1} étaient sans effet sur le système de déconvolution.

VI- 2- 2- Présentation des résultats et interprétation :

Pendant la période de calage, quinze épisodes averse-cruve ont été utilisés afin de tester le modèle par cinq crues, ce qui représente les (3/4) de l'échantillon d'origine. Les résultats obtenus pour les quinze premières crues sont résumés dans le tableau VI-2.

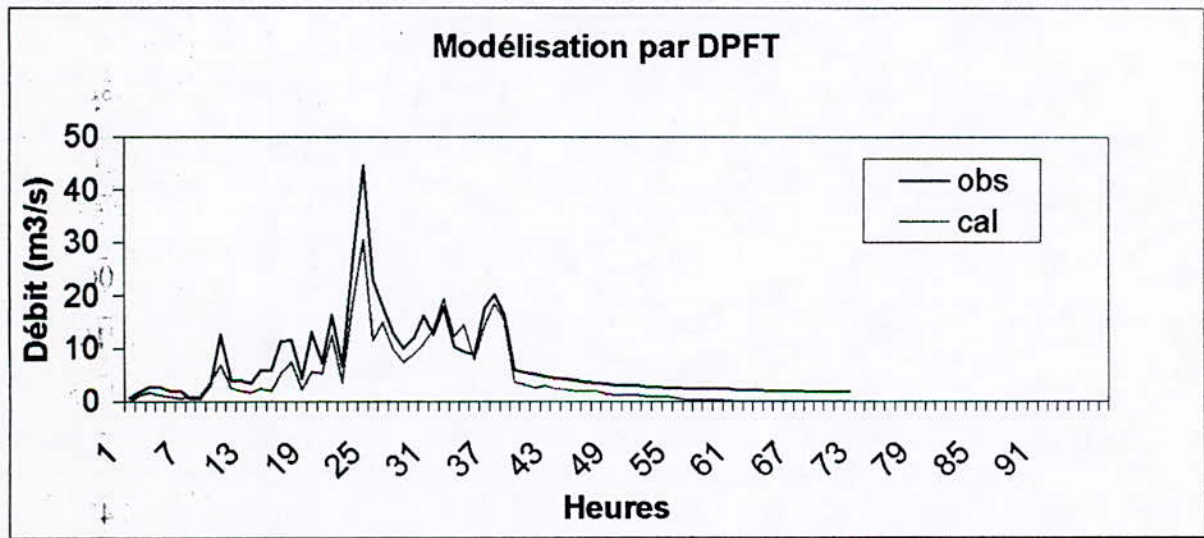


Figure (VI- 2) : Reconstitution de la crue n°1 (calage)

D'après cette figure, on remarque que le modèle n'arrive pas à reproduire les pics des valeurs observées, alors que dans la figure VI-3, l'allure est presque parfaite.

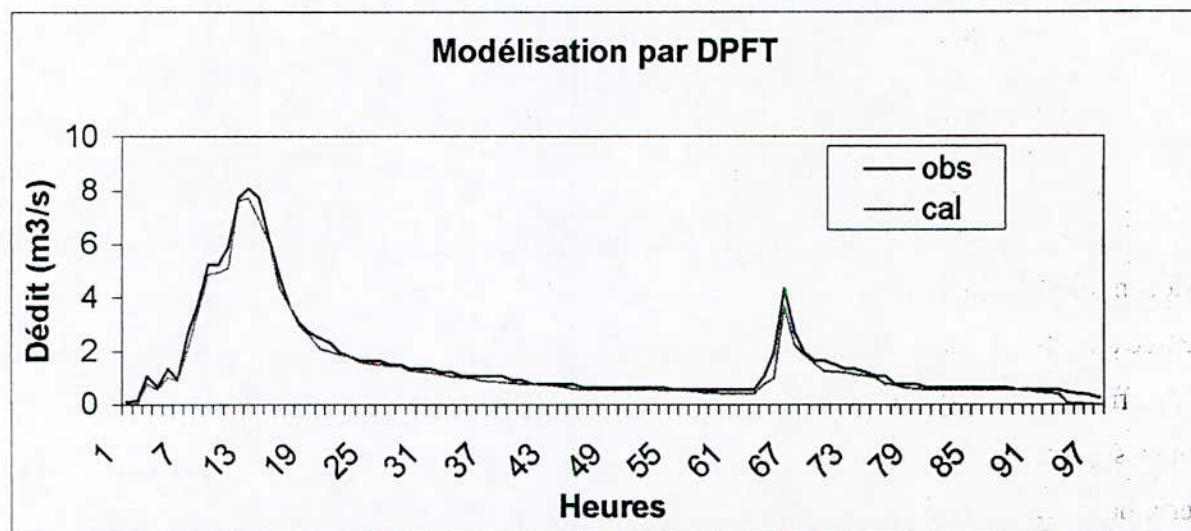


Figure (VI- 3) : Reconstitution de la crue n°8 (calage)

Globalement les critères de validation numériques pour la période de calage sont très satisfaisants sauf pour quelques crues des valeurs faibles des deux coefficients (Nash et R^2) ont été relevé, cela est dû essentiellement à l'existence des crues complexes dans l'échantillon de calage.

Tableau (VI- 2) : Critères de validation de la période de calage de la DPFT.

Crues	n°1	n°2	n°3	n°4	n°5
Nash (%)	78.47	94.04	63.50	91.66	19.11
R^2 (%)	74.65	88.41	58.48	97.05	33.55
RMSE	0.63	0.26	26.32	1.28	1.06
Crues	n°6	n°7	n°8	n°9	n°10
Nash (%)	68.51	84.41	97.63	92.26	94.65
R^2 (%)	67.42	94.77	95.07	93.06	96.79
RMSE	0.26	0.39	0.16	0.10	0.47
Crues	n°11	n°12	n°13	n°14	n°15
Nash (%)	96.28	66.72	87.16	94.12	80.55
R^2 (%)	89.54	62.87	84.47	92.66	91.61
RMSE	0.52	4.29	15.87	1.47	0.45

Après le calage du modèle, il nous a été possible de tracer la fonction de transfert et de voir son allure. La figure (VI-4) montre l'évolution de la fonction de transfert à chaque itération, où il est possible de constater sa stabilité à partir de la 3^{ième} itération et que son allure est presque confondue avec celle identifiée à la 4^{ième} et la 5^{ième} itération.

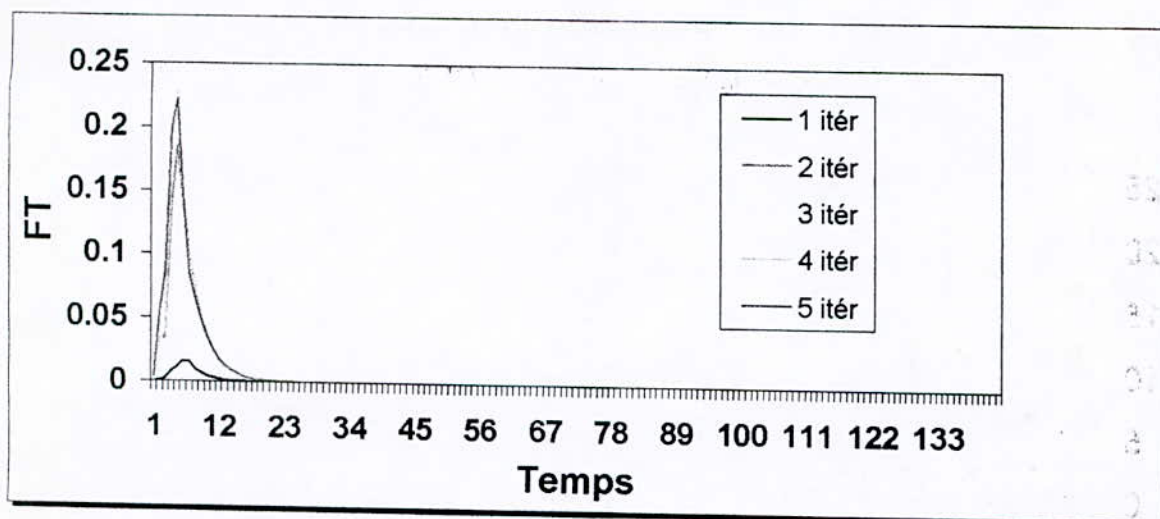


Figure (VI- 4) : Evolution du calage de la fonction de transfert.

Pour vérifier cette constatation, nous avons utilisé la contrainte de forme relative à la fonction de transfert qui exige que $\sum FT_{lissé} = 1$.

Le tableau suivant donne à chaque itération la somme des coefficients de la fonction de transfert :

Tableau (VI- 3) : Valeurs de $\sum FT$ lissée à chaque itération.

Itération	1	2	3	4	5
$\sum FT$ lissée	0.08043	0.8681	0.9421	0.9978	0.9988

On remarque bien que le processus de la méthode DPFT converge au bout de la 3^{ième} itération. Après avoir identifié la fonction de transfert moyenne pour tous les épisodes ainsi que les pluies efficaces, il est possible de déterminer les débits issus de la résolution du système : $\{q\} = [a] * \{PE\}$.

Les cinq crues restantes des épisodes averse-crues ont été utilisées pour la reconstitution du modèle. L'allure des hydrogrammes issus des débits calculés et observés est bien

Les cinq crues restantes des épisodes averses-crues ont été utilisées pour la reconstitution du modèle. L'allure des hydrogrammes issus des débits calculés et observés est bien reproduite, avec une légère sous-estimation des valeurs max (pics), figure (VI-5) pour la crue n°16 et la figure (VI-6) pour la crue n°17 :

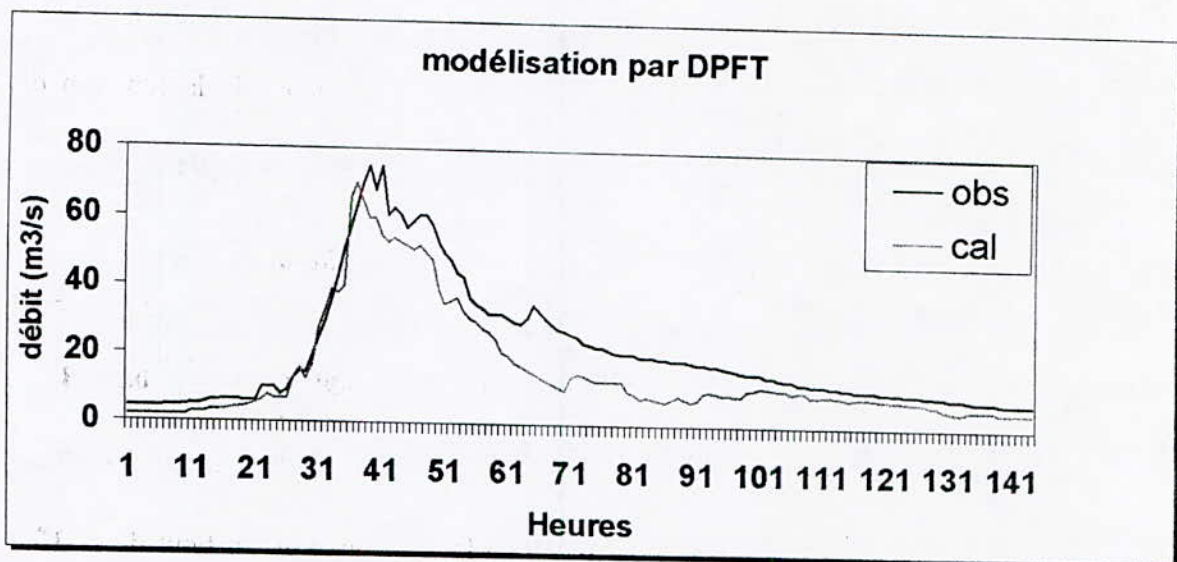
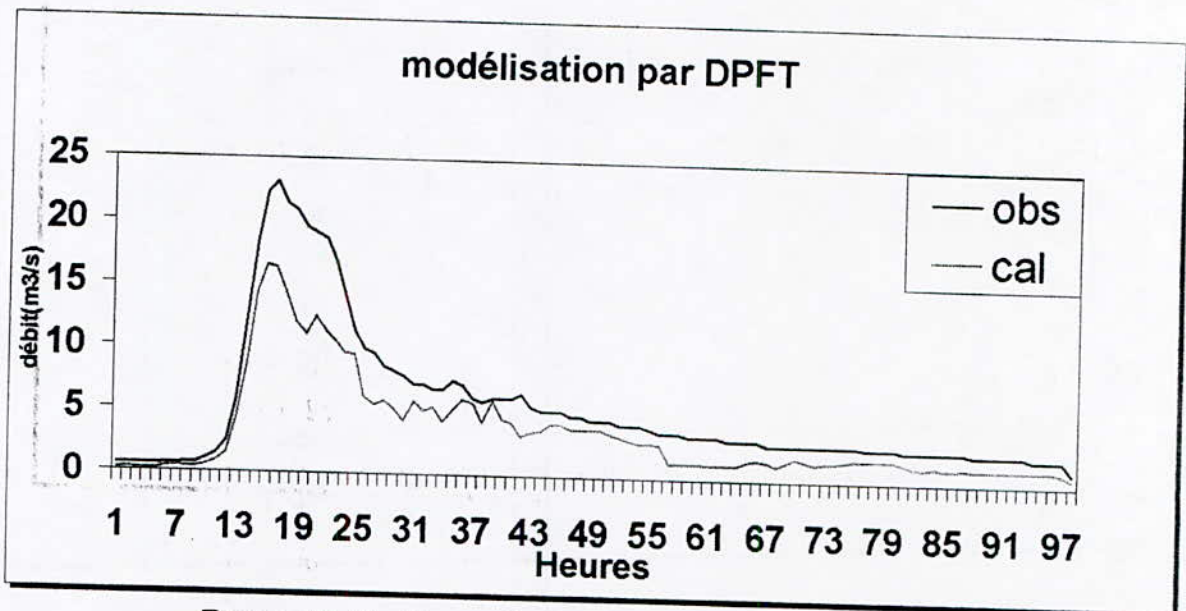


Tableau (VI- 4) : Critères de validation pour la période test

Crues	n°16	n°17	n°18	n°19	n°20
Nash (%)	86.53	77.52	72.52	65.34	38.35
R^2 (%)	96.25	91.61	94.91	88.16	67.72
RMSE (%)	0.51	0.80	0.67	0.71	0.41

En outre, les critères de comparaison numériques (Nash, R^2 et RMSE) sont très satisfaisants, surtout pour les quatre premières crues reconstituées qui sont des crues simples possédant un seul pic, ceci nous amène à constater la capacité et la fidélité de la méthode DPFT à reconstituer ce type de crue.

Pour la reconstitution de la crue n°20 dont les résultats numériques sont moins bons ainsi que l'allure des valeurs simulées qui n'est pas correctement reproduites, cela est dû essentiellement au fait que cette crue est complexe (2 pics).

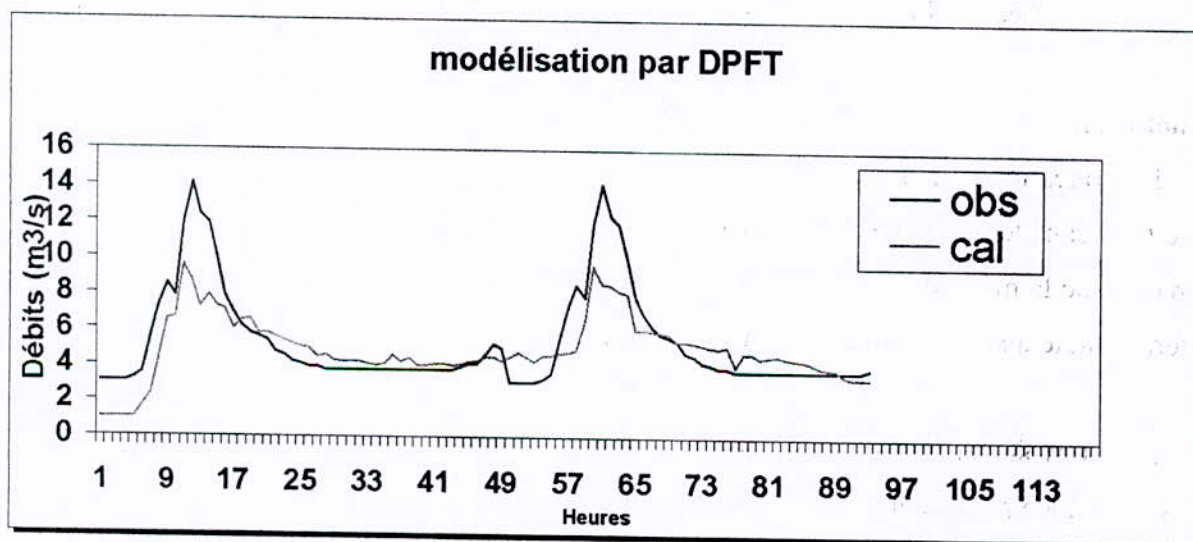


Figure (VI- 7) : Reconstitution de la crue n°20 (période de test)

L'histogramme suivant présente une comparaison entre les critères de validation de Nash et R^2 obtenus par le modèle DPFT pour la période test.

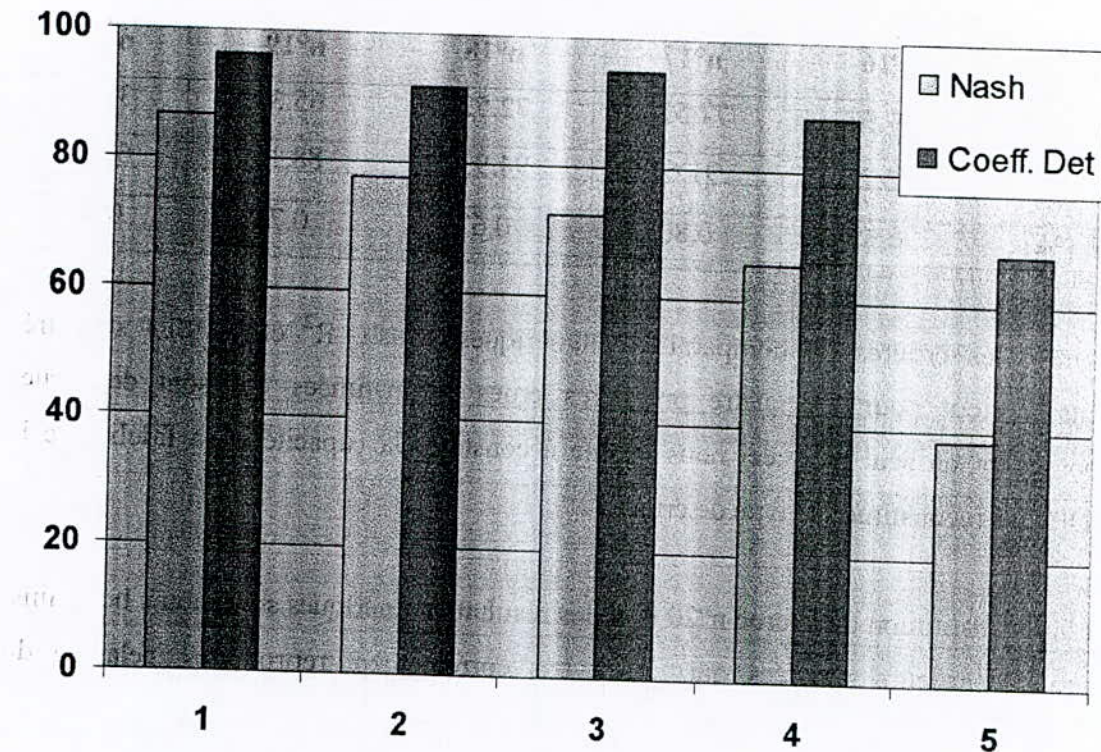


Figure (VI- 8) : Les valeurs des critères de validation (Nash et R^2)

Conclusion :

L'application de la méthode DPFT au sous bassin versant de l'Oued EBDA (CHELIFF), a été faite dans le but de tester les performances de cette méthode sur des données réelles. La démarche de la méthode a été suivie en respectant toutes les consignes recommandées par les différents auteurs de la partie théorique de la DPFT.

Après avoir déterminé la fonction de transfert à travers sa différence première (DPFT), qui s'est soldée par la détermination d'une fonction vérifiant la contrainte de conservation de masse avec une $\sum FT = 0.9988$, et une série des pluies efficaces.

La combinaison de ces deux parties nous a conduit à l'étape la plus importante de cette étude, qui réside dans la détermination des débits (test du modèle), en reconstituant les crues.

Les résultats obtenus sont très satisfaisant, en effet, la fonction de transfert identifiée et les pluies efficaces déconvoluées sont satisfaisantes et acceptables, pour les critères de comparaison utilisés, surtout pour les crues simples avec des coefficients de détermination et de Nash très acceptables, et moins appréciable pour les crues complexes où les mêmes coefficients étaient plutôt moyens.

Enfin, nous pouvons conclure que malgré les difficultés rencontrées lors de la constitution de l'échantillon qui était de petite taille, et la complexité de 50% des crues, que la méthode d'une manière générale, donne des résultats très appréciables.

VI- 3- Application du modèle Réseau de neurones :

VI- 3- 1- Introduction :

L'application du modèle des réseaux de neurones et sa validation se feront sur les mêmes événements averse crue déjà utilisés dans l'application de la DPFT.

La validation et la performance du modèle seront jugées par les mêmes critères de validation numériques et graphique utilisés pour la DPFT.

VI- 3- 2- Présentation et interprétation des résultats :

Dans cette partie, un perceptron à trois couches a été utilisé pour modéliser la relation pluie-débit. La technique d'apprentissage supervisé utilisée et la rétropropagation de l'erreur. Le modèle en question est constitué d'une couche d'entrée, comportant un neurone d'entrée, une couche cachée, comportant neuf neurones cachés et une couche de sortie comportant un neurone de sortie. Le choix du nombre de couches et du nombre de neurones a été décidé après plusieurs essais de différentes combinaisons. Nous avons effectivement pu constater, à travers ces essais, que l'augmentation du nombre de couches cachées ou celui des neurones ne conduisait plus à une amélioration des résultats, au contraire, il rendait le réseau plus difficile à caler et son temps d'apprentissage plus long.

Interprétation :

L'apprentissage du modèle réseaux de neurones consiste en la détermination des poids interconnectés, entre les couches composantes le réseau, pour cela, un échantillon de quinze crues a été utilisé. Le calage du réseau nous a permis d'avoir une reconstitution des crues avec des coefficients de validation très appréciables et remarquables en général (tableau : VI- 5) :

Tableau (VI- 5) : Les critères de validation de la période de calage

Crues	n°1	n°2	n°3	n°4	n°5
Nash (%)	66.05	78.77	45.65	83.30	52.00
R^2 (%)	66.77	70.65	41.23	79.61	63.20
RMSE	2.71	1.37	21.19	0.9794	0.32
Crues	n°6	n°7	n°8	n°9	n°10
Nash (%)	95.58	93.71	96.23	80.86	91.28
R^2 (%)	95.50	93.91	91.98	80.96	91.33
RMSE	0.17	0.31	0.47	0.16	0.32
Crues	n°11	n°12	n°13	n°14	n°15
Nash (%)	96.91	90.20	89.68	97.82	91.15
R^2 (%)	96.66	90.26	89.89	97.83	91.70
RMSE	0.10	1.08	8.74	0.22	2.81

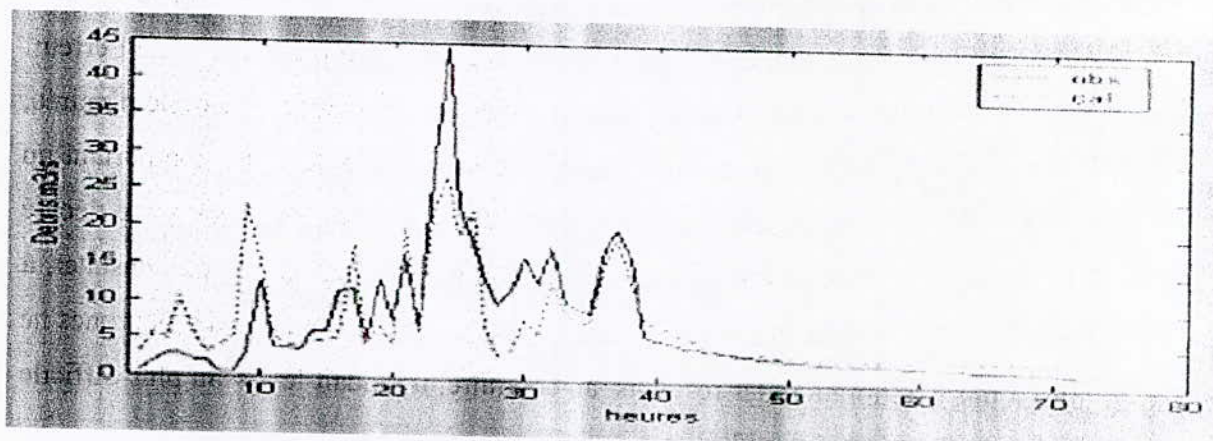


Figure (VI- 9) : Reconstitution de la crue n°1

Pour la crue n°1, la figure (VI-9) montre que l'allure n'est pas fidèlement reproduite et que les valeurs calculées par le réseau de neurones ne suivent pas parfaitement celles observées, ce résultat étant confirmé par les valeurs obtenues par les critères numériques (Nash, RMSE et R^2). Il en est de même pour les crues n°3 et 5.

La seule explication c'est la complexité de ces crues (au minimum quatre pics pour les crues n°1 ,3).

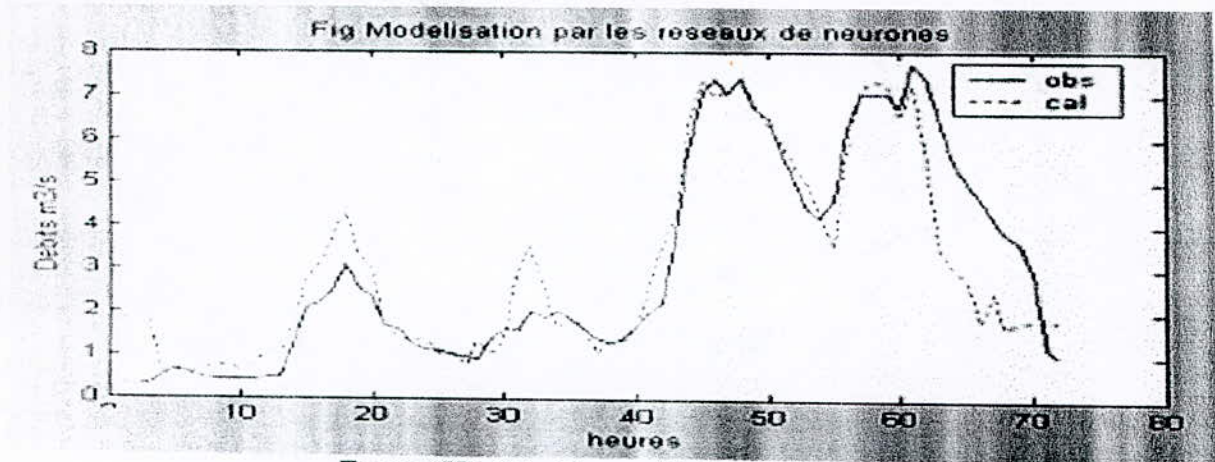


Figure (VI- 10) : Reconstitution de la crue n°4

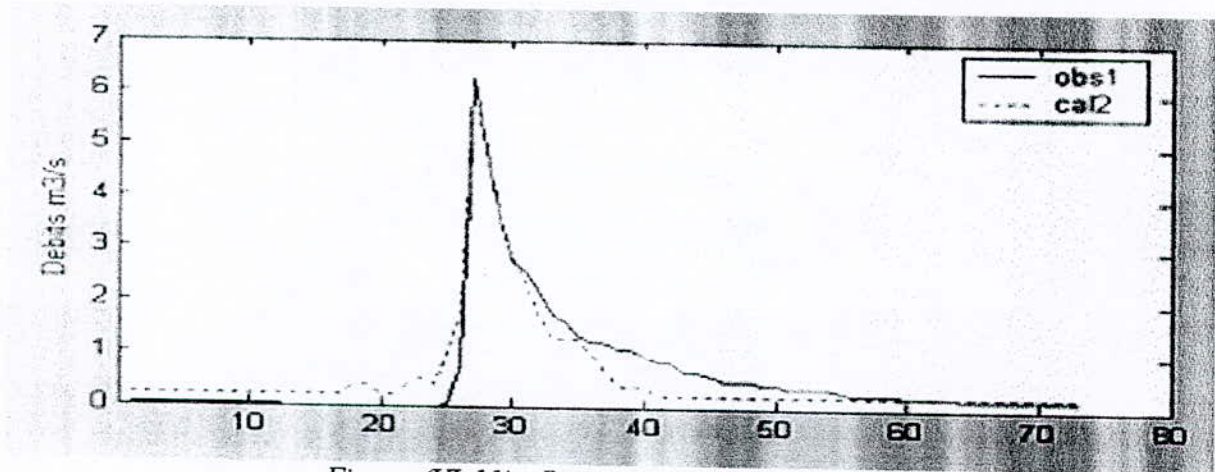


Figure (VI-11) : Reconstitution de la crue n°10

Pour le reste des crues la tendance des courbes simulées suit bien l'allure des courbes observées (Fig VI- 10 et Fig VI- 11), et donc d'une manière générale les crues de la période de calage sont fidèlement reproduites et les résultats des différents critères de validation sont très appréciables.

Le test du modèle réseau de neurones a été effectué sur les cinq crues restantes. Nous avons constaté la capacité et le potentiel du modèle à reproduire les débits, en effet, les résultats obtenus montre que l'allure des hydrogrammes est très bien reproduite par le modèle comme le montre les figures (VI- 12 et VI- 13) :

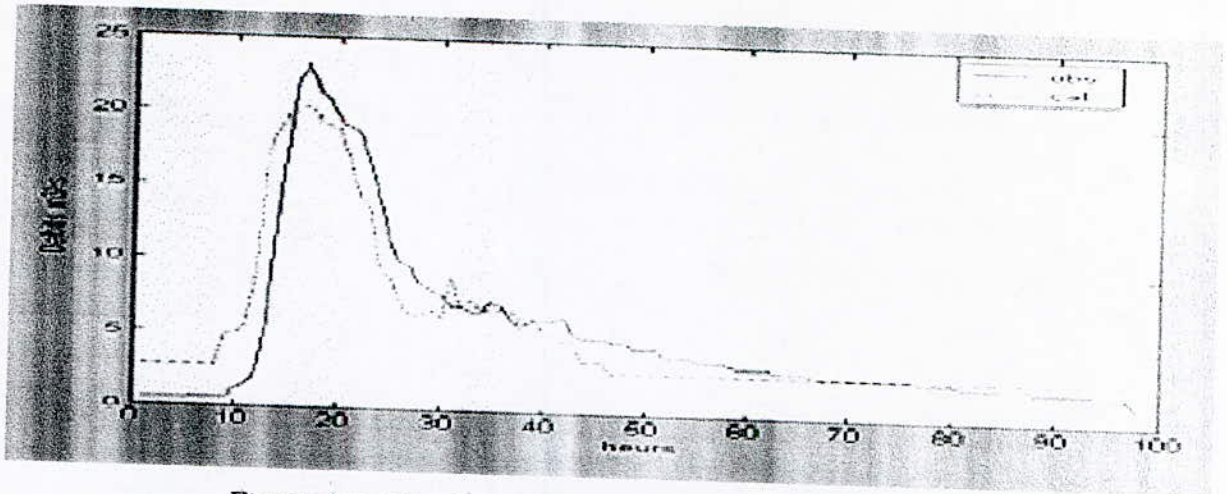


Figure (VI- 12) : Reconstitution des crues test (Crues n°16)

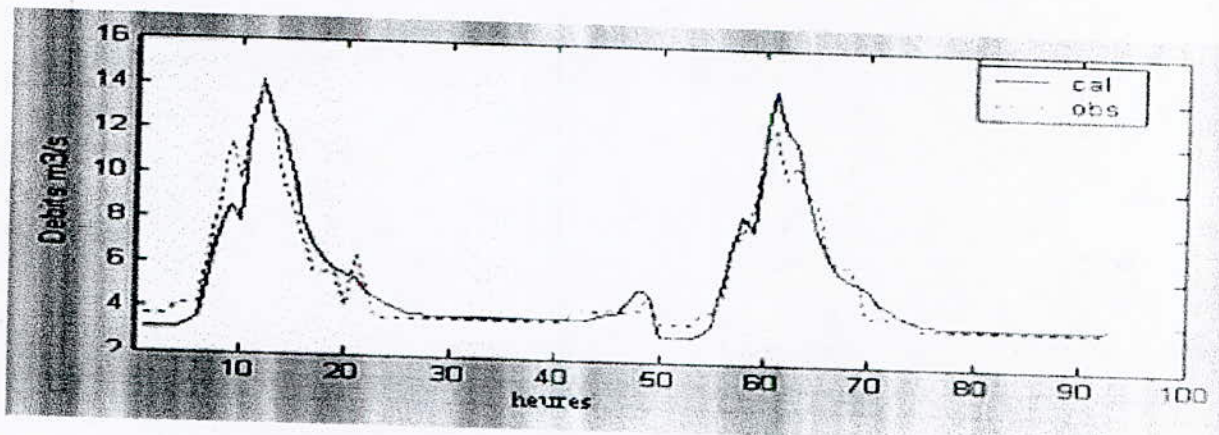


Figure (VI- 13) : Reconstitution des crues test (Crue n°20)

Et que par conséquence les crues test sont très bien reconstituées.

Tableau (VI- 6) : Les critères de validation des crues test

Crues	n°16	n°17	n°18	n°19	n°20
Nash (%)	83.53	86.60	88.79	90.24	92.55
R^2 (%)	83.59	87.86	88.61	90.54	93.95
RMSE	1.74	1.97	1.24	2.66	0.57

D'après ce tableau, on remarque que les résultats numériques sont très satisfaisants.

Conclusion :

Dans cette partie, un modèle de réseau de neurones à trois couches, avec un neurone à la couche d'entrée, neuf neurones à la couche cachée et un neurone à la couche de sortie a été utilisé afin de simuler un phénomène hydrologique complexe pluie-débit.

Les résultats numériques et graphiques obtenues dans cette étude confirme le potentiel des réseaux de neurones pour la simulation en hydrologie, et montrent clairement qu'il peuvent apporter des réponses très satisfaisantes pour la modélisation pluie-débit et la prédiction des débits au pas de temps horaire avec un très haut degré de fidélité, d'où la capacité des réseaux de neurones à bien modéliser les phénomènes physiques non-linéaires au vu des résultats obtenus. Ceux-ci sont résumés par l'histogramme suivant :

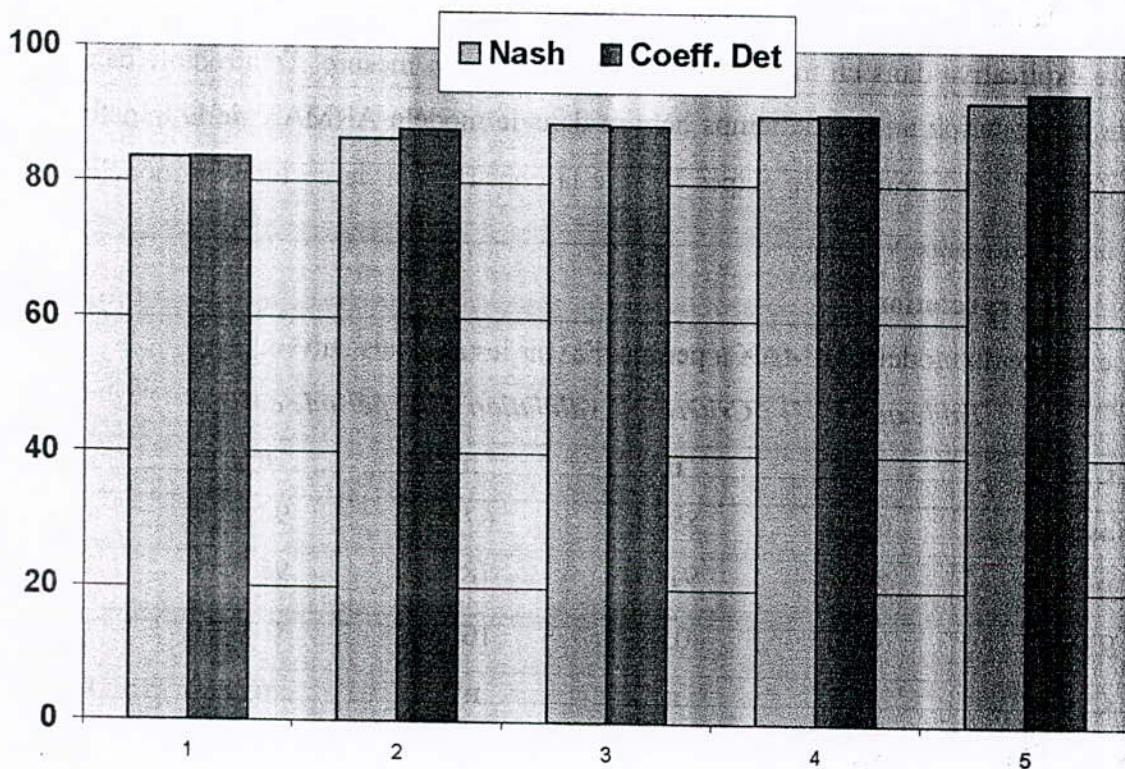


Figure (VI-14) : Les valeurs des critères de validation (test)

VI- 4- Application du modèle ARMAX :**VI- 4- 1- Introduction :**

L'application du modèle ARMAX et sa validation se feront aussi sur les mêmes événements averse crue déjà utilisés dans l'application de la DPFT et des réseaux de neurones. L'algorithme général du modèle ARMAX utilisé est :

$$Q_{dj} + \sum_{i=1}^m c_i * Q_{dj-i} = \sum_{i=0}^n b_i * PE_{j-i-1} + e_j \quad (\text{VI. 6})$$

Les mêmes critères de validation numériques et graphiques, sont utilisés pour juger la performance du modèle ARMAX.

VI- 4- 2- Présentation et interprétation des résultats :

Pour la modélisation de la relation pluie-débit par le modèle ARMAX on a introduit une variable explicative dans un modèle de Box et Jenkins. La mesure X_t introduite dans ce cas c'est la précipitation au pas de temps horaire dans le modèle ARMAX de la transformation z_T du débit Q_T . Les principales étapes lors de la construction de ce type de modèle ont été citées dans le chapitre V.

Interprétation :

Le calage du modèle ARMAX a permis d'avoir les résultats suivants :

Tableau (IV- 7) : Critères de validation de la période calage.

Crues	n°1	n°2	n°3	n°4	n°5
Nash (%)	60.49	89.82	67.01	94.89	17.64
R² (%)	59.98	91.68	78.69	95.06	30.61
RMSE	4.88	0.89	16.69	0.54	0.47
Crues	n°6	n°7	n°8	n°9	n°10
Nash (%)	83.10	89.40	96.38	95.00	89.61
R² (%)	94.90	92.65	96.41	94.34	95.41
RMSE	0.36	0.39	0.31	0.08	0.35
Crues	n°11	n°12	n°13	n°14	n°15
Nash (%)	97.53	76.31	77.11	93.83	76.53
R² (%)	97.93	82.79	84.76	96.05	80.92
RMSE	0.46	0.67	16.22	1.17	1.97

Ceux-ci montrent des valeurs très satisfaisantes pour les différents critères numériques, et un ajustement très appréciable des valeurs observées par rapport à celle calculées pour les critères graphiques figure (VI- 15), sauf pour la 1^{ère}, 3^{ème} et 5^{ème} crue (crues complexes à plusieurs pics) ou l'allure des courbes simulées n'est pas bien reproduite figure (VI.16).

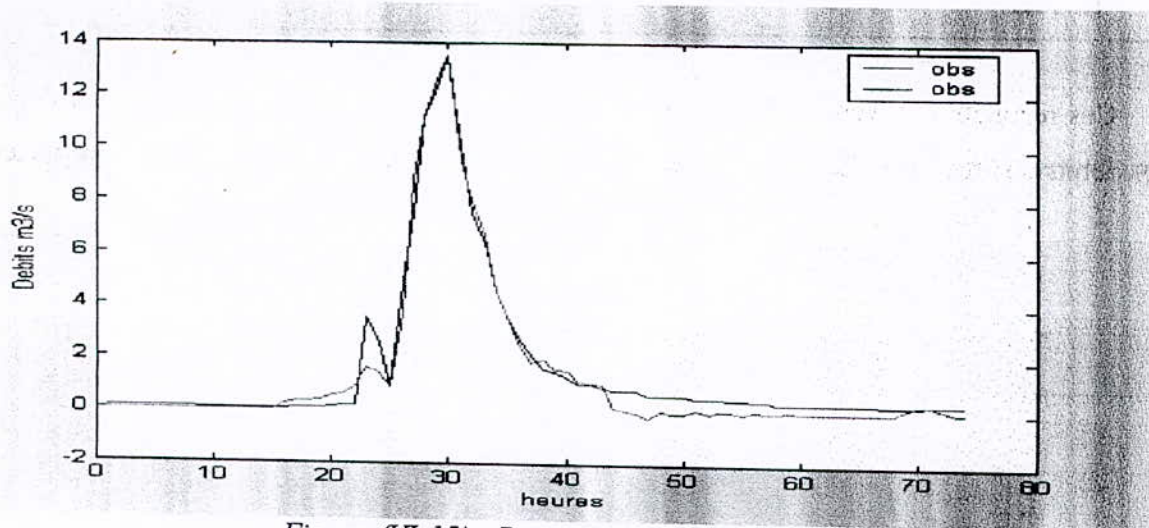


Figure (VI-15) : Reconstitution de la crue n°1 (calage)

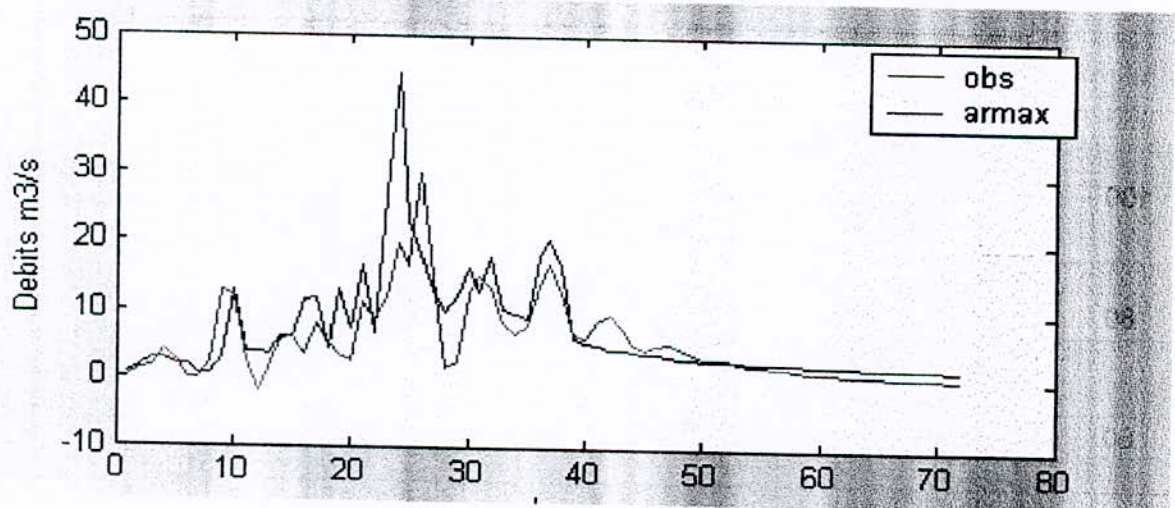


Figure (VI-16) : Reconstitution de la crue n°11 (calage)

En général, on peut dire que la reconstitution des crues de calage par le modèle ARMAX est satisfaisante et appréciable.

L'application du modèle ARMAX aux crues tests a montré une fidélité moins bonne pour quelques crues au niveau de l'allure du débit modélisé Figure (VI-17).

Tableau (VI- 8) : Critères de validation des crues de test

Crues	n°16	n°17	n°18	N°19	n°20
Nash (%)	90.13	57.02	19.14	81.41	26.70
R^2 (%)	96.26	65.26	34.44	91.06	83.66
RMSE	0.60	4.42	4.88	0.16	3.07

Ces résultats prouvent si besoin est la puissance et la capacité de l'approche à modéliser les débits de crues simples alors que le modèle montre ces limites pour les crues complexes.

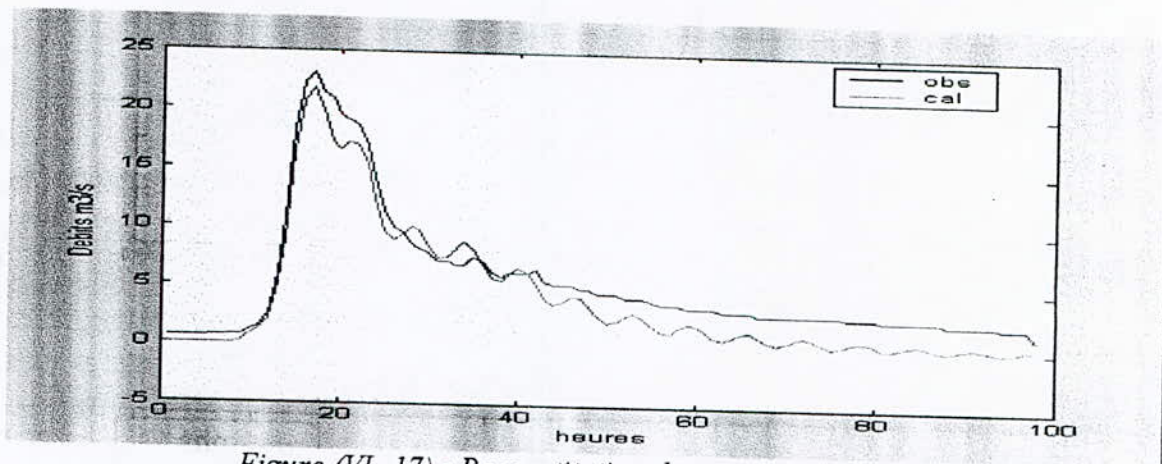


Figure (VI-17) : Reconstitution des crues test (Crue n°16)

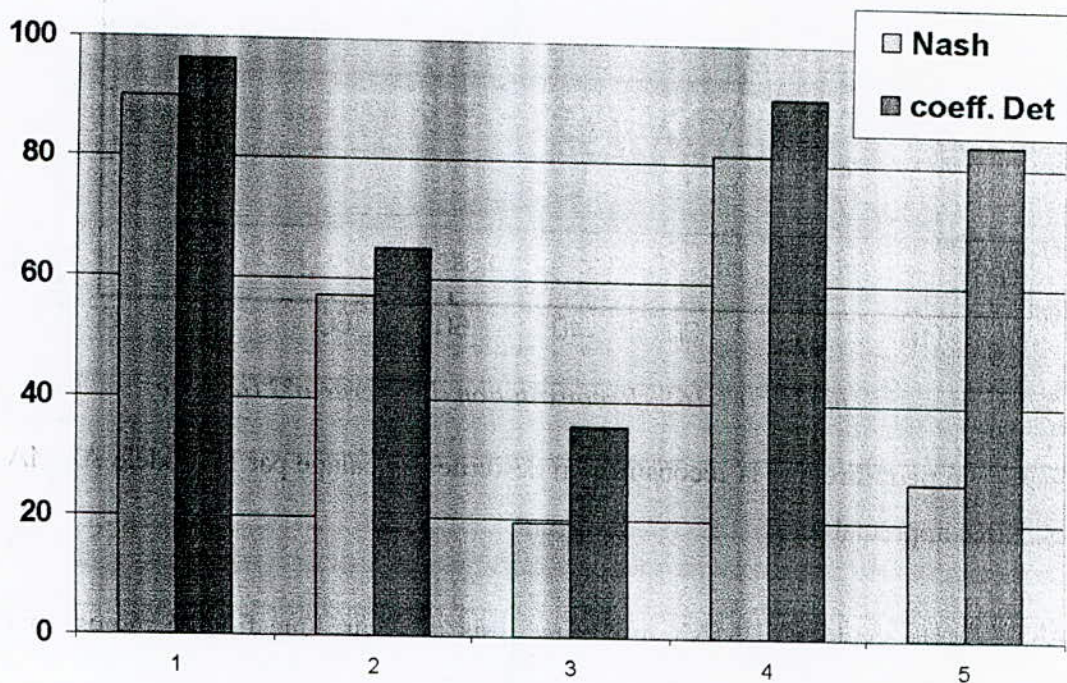


Figure (VI-18) : Les valeurs des critères de validation (test)

Conclusion :

Sur la base des résultats obtenus par les critères de validation, nous constatons d'une manière générale, que les crues de calage sont bien reconstituées (petites erreurs relatives pour certaines crues) est que les crues tests sont bien reproduites, surtout pour les crues simples.

Nous notons également, que le potentiel de la reconstitution des crues n°18 et n°20 donne un résultat plutôt faible.

Enfin, on remarque qu'il y pas une amélioration au niveau des résultats des critères de validation, surtout par rapport au modèle de réseau de neurones.

VI- 5- Synthèse des résultats des trois approches :

D'après les résultats trouvés dans la partie précédente, on peut faire une comparaison entre les trois modèles utilisés, tel que on a calculé la moyenne des différents critères de validation de la période de test comme suit :

Tableau (VI- 9) : Les valeurs moyenne des critères de validation pour chaque modèle

	Réseau de neurone	DPFT	ARMAX
Nash (%)	88.34	68.05	54.88
R ² (%)	88.91	87.73	74.11

Une autre façon plus représentative par des histogrammes donne :

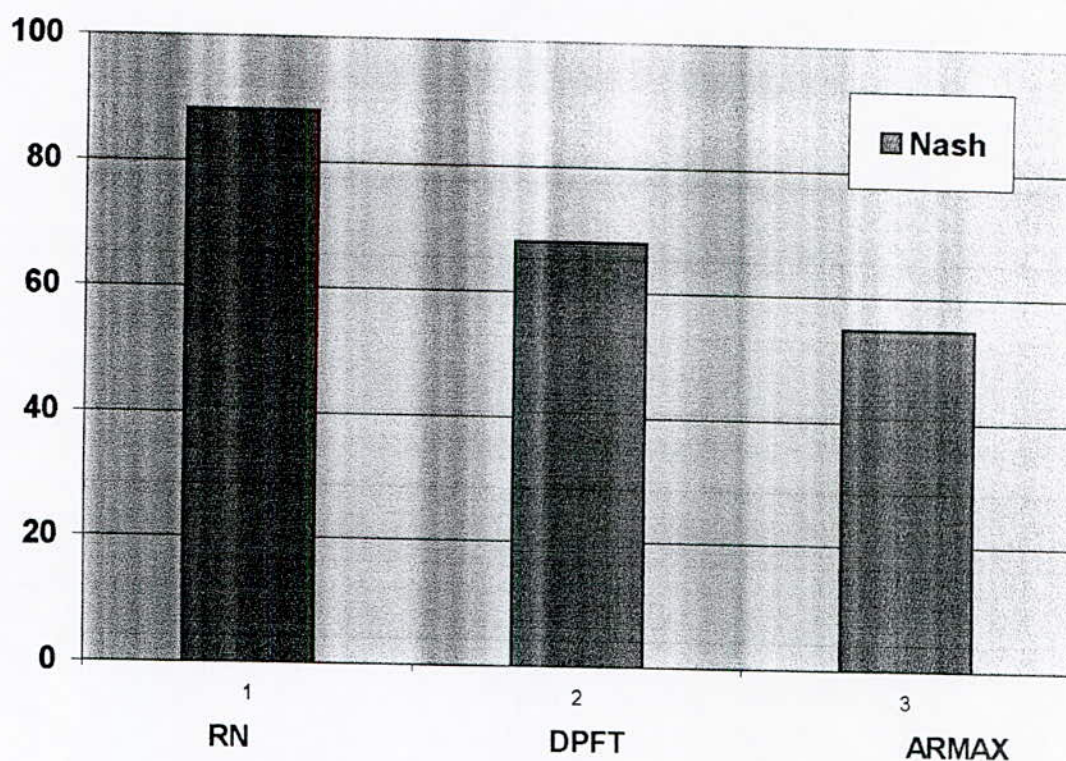


Figure (VI- 19) : Les valeurs moyennes de Nash des différents modèles

D'après le tableau précédent et l'histogramme de la figure (VI-19) on remarque clairement la différence de la valeur de Nash entre les trois modèles utilisés, une différence de 20.29% entre le modèle réseau de neurone et la DPFT, et une différence de 33.46% entre le réseau de neurone et ARMAX.

Pour le coefficient de détermination :

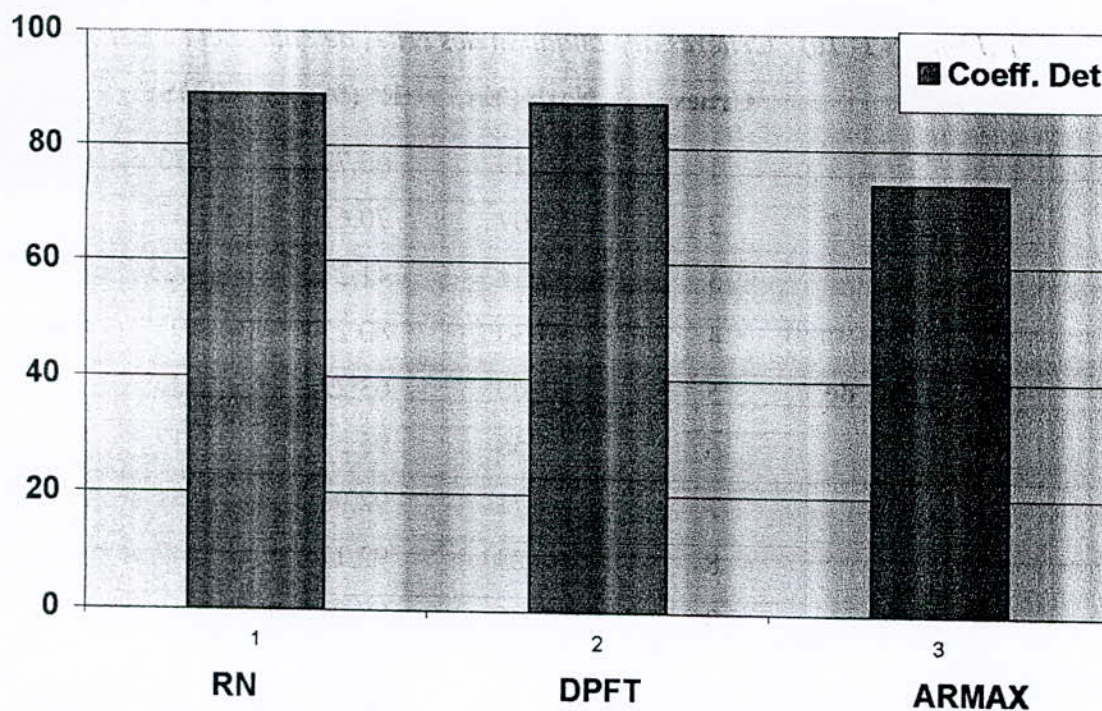


Figure (VI- 20) : Les valeurs moyennes de coefficient de détermination des différents modèles

Les moyennes des coefficients de détermination du réseau de neurones et la DPFT sont presque des même ordre, mais c'est le réseau de neurone qui donne le meilleur coefficient de détermination (88.91 %), alors que R^2 de ARMAX c'est le plus faible.

Au terme de cette étude nous pouvant affirmer que le modèle RN est le mieux adapté à la modélisation complexe pluie-débit pour le site étudié, mais cela ne diminue en rien le potentiel et la capacité de la DPFT et ARMAX à reproduire la non-linéarité des événements averse-crue étudiés.

A partir de ces résultats, on a essayé de trouver un seuil du nombre de crues nécessaire et suffisant au calage pour pouvoir modéliser les débits, à partir des vingt (20) événements averse crue disponible, pour le modèle retenu (réseau de neurone).

Pour cela on a pris pour le premier cas dix (10) crues pour le calage et dix crues pour la validation, ensuite on a diminué la période calage à cinq crues et donc quinze crues pour le test dans le deuxième cas.

- Résultats et interprétation :

VI-5-1- Les résultats de dix crues calage et dix crues validation :

Tableau (VI- 10) : Critères de validation des crues de calage et de test

	Crues	Nash (%)	R ² (%)	RMSE
La période de calage	1	65.29	66.77	2.73
	2	78.07	70.62	1.37
	3	45.34	41.21	21.25
	4	83.30	80.23	0.97
	5	52.00	63.22	0.32
	6	95.58	95.51	0.17
	7	93.71	92.99	0.31
	8	96.23	92.12	0.47
	9	80.86	80.96	0.16
	10	91.28	91.33	0.32
La période de test	11	91.10	97.84	1.33
	12	70.24	88.52	1.08
	13	73.64	91.53	18.59
	14	95.63	95.35	1.019
	15	88.56	92.36	2.81
	16	82.36	85.12	2.02
	17	91.21	86.13	0.44
	18	87.00	81.15	1.10
	19	68.17	86.33	4.98
	20	77.48	91.68	6.10
	Moyenne	82.53	89.60	

Les résultats obtenus pour le 1^{ier} cas sont très appréciables et satisfaisants, avec une moyenne de Nash de 82.53% et une moyenne de coefficient de détermination de 89.60%, le réseau de neurone a bien reproduit l'allure de l'hydrogramme des débits figure (VI-21 et VI-22).

Nous remarquons ainsi une diminution de la valeur moyenne de Nash par rapport au test précédent (cinq crues de validation), par contre le coefficient de détermination n'a pas beaucoup changé.

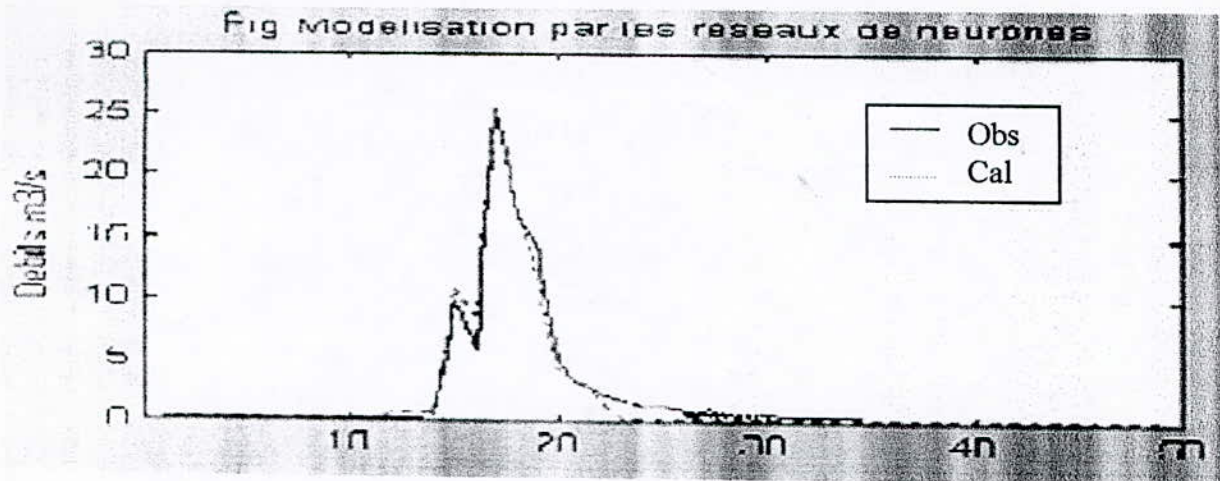


Figure (VI- 21) : Reconstitution des crues test (Crue n°14)

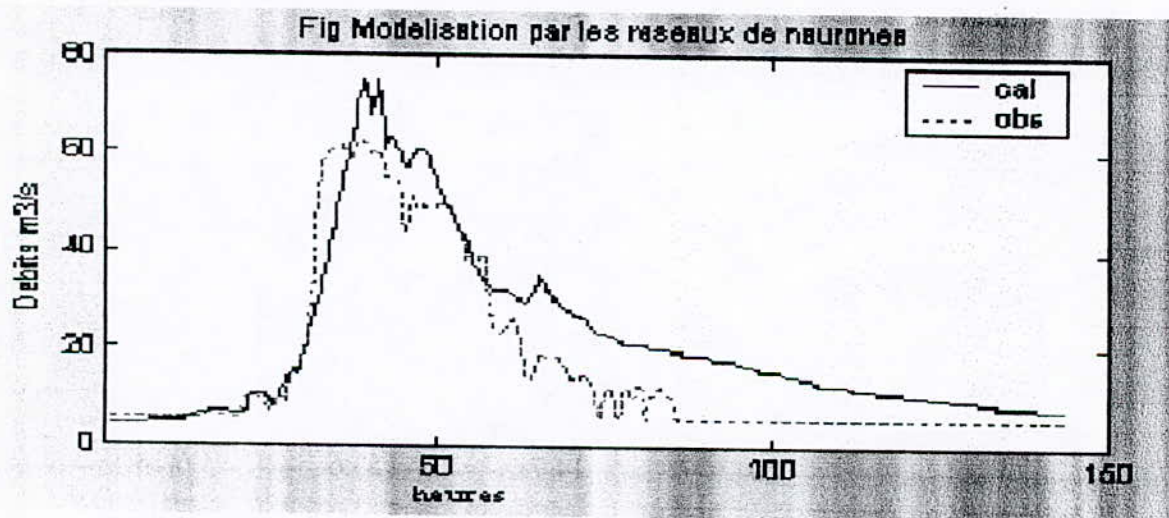


Figure (VI- 22) : Reconstitution des crues test (Crue n°16)

VI-5-2- Les résultats de cinq (5) crues calage et quinze (15) crues validation :

Tableau (VI- 11) : Critères de validation des crues de calage et de test

	Crues	Nash (%)	R ² (%)	RMSE
Période de calage	1	68.02	68.15	2.71
	2	74.66	74.67	0.52
	3	55.60	57.65	13.20
	4	81.07	81.65	0.32
	5	51.71	51.72	0.14
Période de validation	6	92.36	92.85	0.04
	7	93.72	91.69	0.15
	8	78.27	86.39	1.51
	9	77.69	77.45	1.62
	10	44.28	69.00	2.01
	11	18.61	26.53	4.02
	12	6.48	34.71	3.12
	13	7.63	32.52	31.24
	14	-10.79	23.52	8.67
	15	22.32	17.09	10.56
	16	34.26	0.93	7.27
	17	-80.09	49.46	12.56
	18	9.63	34.70	6.42
	19	7.70	21.35	8.00
	20	14.71	53.6	6.18

D'après le tableau (VI-11), on remarque que les résultats des crues qui viennent juste après la période de calage sont très satisfaisants, mais qu'à partir de la huitième crue le modèle commence à diverger et a donné même des résultats négatifs de Nash. D'une manière générale les résultats sont très faibles et le modèle réseau de neurone ne peut donner de bons résultats à partir de cinq (5) crues de calage.

Exemples des crues :

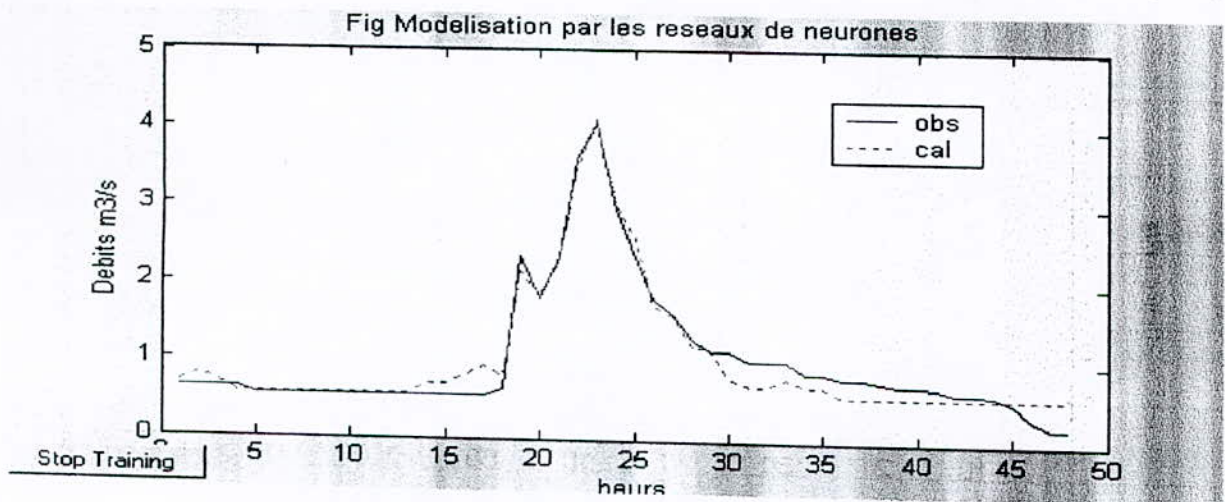


Figure (VI- 23) : Reconstitution des crues test (Crue n°6)

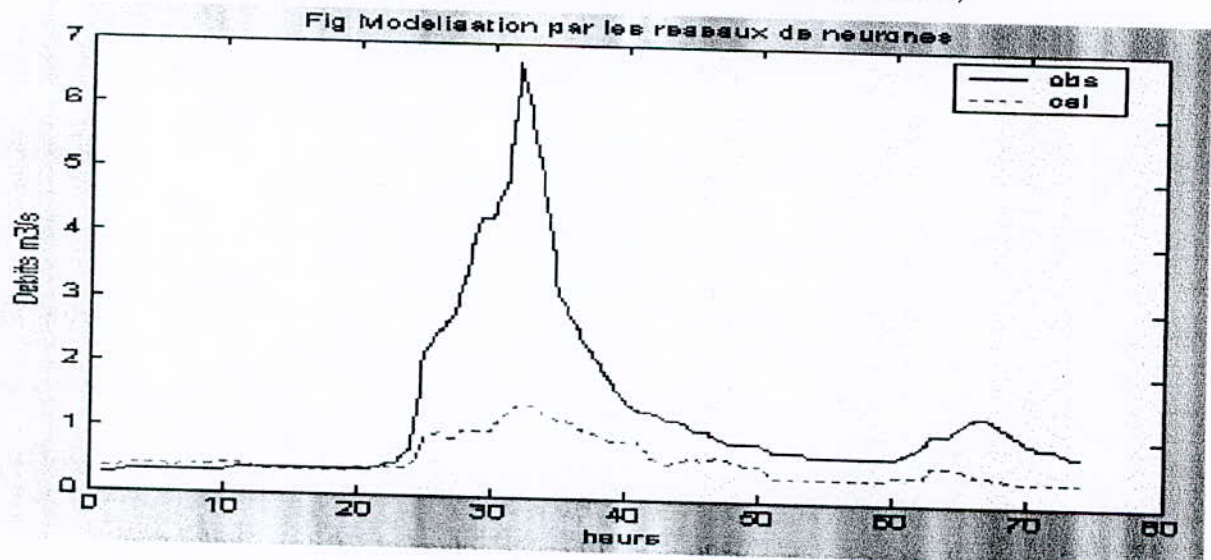


Figure (VI- 24) : Reconstitution des crues test (Crue n°12)

D'après cette étude nous pouvons dire que le seuil du nombre de crues de calage dans notre cas (20 crues au totale), ne doit pas être inférieure à huit (8) pour pouvoir modéliser un phénomène aussi complexe que la relation pluie-débit.

Pour confirmer le fait que le modèle diverge à la fin de son exécution, un ultime test a été utilisé, qui consiste à faire varier l'ordre de la vingtième crue pour le cas cinq crues de calage, et on la met juste après la période de calage.

Le résultat obtenu montre effectivement une nette amélioration des différents critères de comparaison numériques (Nash, R^2 et RMSE) par rapport surtout au troisième cas.

Les résultats de la validation de la crue n°20 pour les différentes périodes de calage sont donnés avec leurs figures correspondantes ci-après :

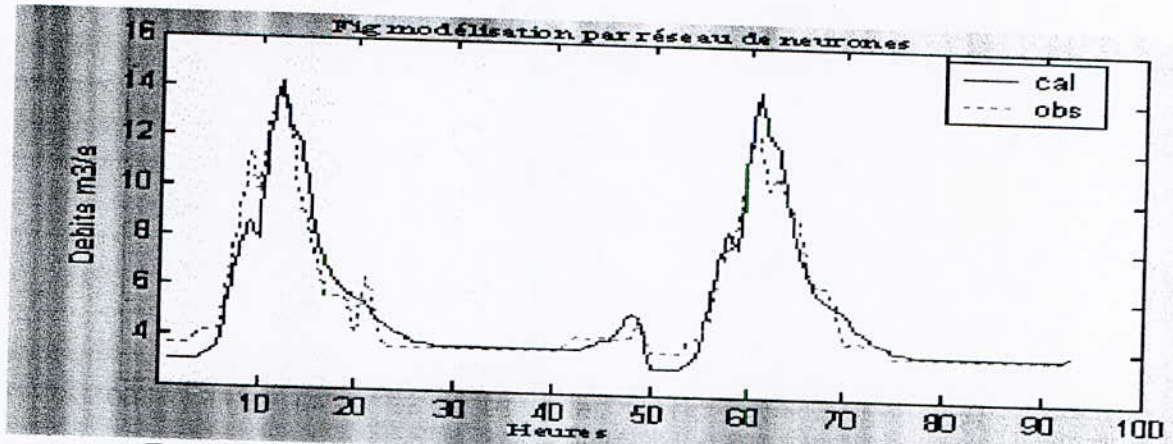


Figure (VI- 25) : La crue n°20 pour 15 crues calage et 5 crues validation
 Nash = 92.55% $R^2 = 92.55\%$ RMSE = 0.51

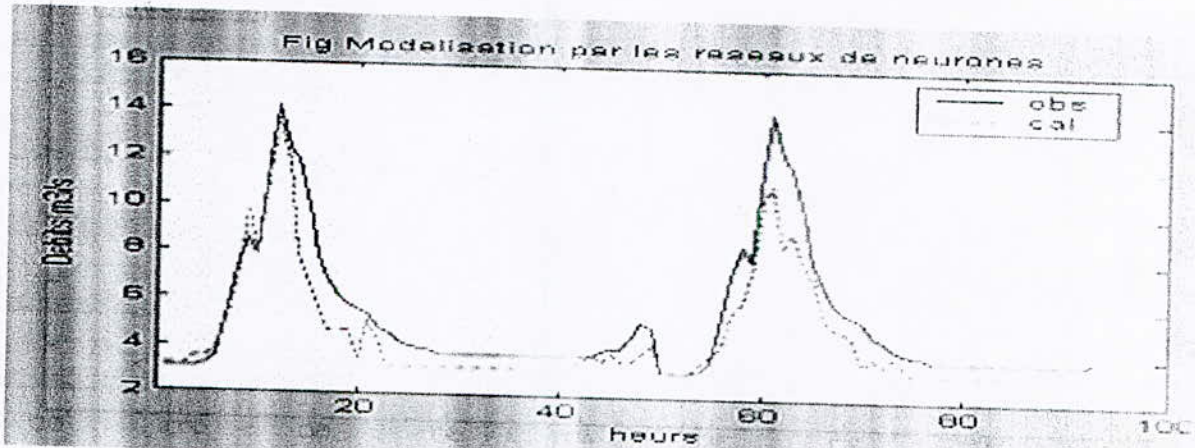


Figure (VI- 26) : La crue n°20 pour 10 crues calage et 10 crues validation
 Nash = 77.48% $R^2 = 91.68\%$ RMSE = 2.10

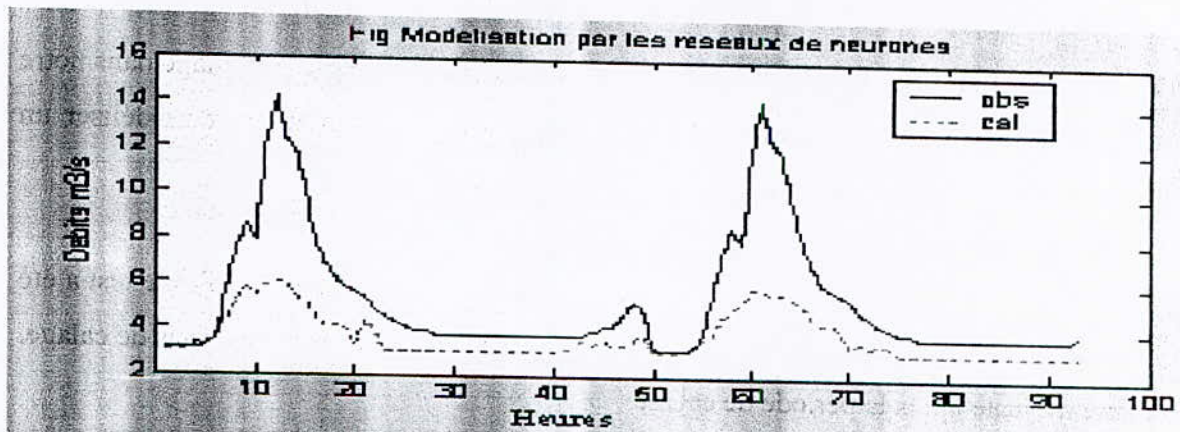


Figure (VI- 27) : La crue n°20 pour 5 crues calage 15 crues validation
 Nash = 14.71% $R^2 = 53.62\%$ RMSE = 6.18

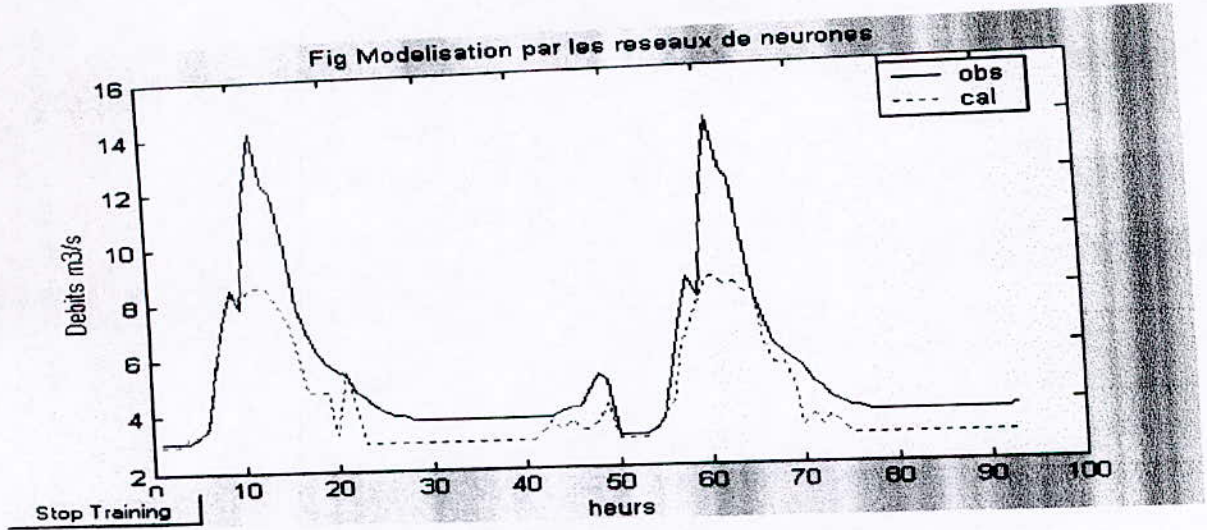


Figure (VI- 28) : La crue n°20 à la sixième position

Nash = 65.28%

$R^2 = 88.64\%$

RMSE = 2.81

CONCLUSION GENERALE

L'étude de la relation pluie débit a fait l'objet de plusieurs recherches qui ont essayé de modéliser cette relation à différents pas de temps afin d'appréhender le problème des inondations et de mettre au point un système d'annonce des crues par un modèle au pas de temps horaire.

L'étude présentée dans ce mémoire consiste à étudier la relation pluie-débit au pas de temps horaire par trois modèles différents, issus de différents domaines, pour trouver une approche de modélisation robuste afin de choisir le modèle le plus fiable, qui permettra de cerner les problèmes liés aux crues, et de minimiser les dégâts qu'ils peuvent engendrer.

Les trois techniques de modélisation hydrologique utilisées sont :

Le premier modèle est la DPFT (Différence Première de la Fonction de Transfert), qui repose sur trois principes de base, en faisant intervenir les différences des débits et des coefficients de transfert pour la transformation des pluies en débits, avec l'introduction des contraintes pour filtrer les données des pluies. L'application du modèle sur des épisodes pluie-débit a reproduit le plus fidèlement les crues lorsque celles-ci sont de type simple (un seul pic), par contre les crues complexes résultants de la succession de plusieurs crues sont moins satisfaisantes, et ce sont celles qui engendrent les forts débits.

Le second modèle, issu du domaine de l'intelligence artificielle, s'inspire du mode de fonctionnement des neurones du cerveau humain. Après plusieurs essais, un perceptron à trois couches a été utilisé : un neurone dans la couche d'entrée, neuf neurones pour la couche caché et un neurone pour la couche de sortie.

L'application du modèle réseau de neurones sur les crues tests a montré une puissance et une capacité appréciable quelque soit le type de crues (complexe ou simple) à reproduire.

Une autre approche a été utilisée dans cette étude, c'est le modèle ARMAX, qui s'inspire des travaux de Box et Jenkins. Le modèle a donné de bons résultats, mais on a remarqué une certaine limite de ce modèle, surtout au niveau des crues complexes (plusieurs pics).

Une étude comparative entre les trois modèles a été effectuée ensuite, pour aboutir au choix du meilleur modèle à retenir à savoir les réseaux de neurones.

A partir de là, nous avons essayé de trouver un seuil au nombre de crues nécessaires et suffisantes à partir duquel la capacité et le potentiel du modèle restent satisfaisants et appréciables.

On a remarqué qu'à chaque fois qu'on diminuait le nombre de crues du calage, les résultats devenaient moins appréciables, jusqu'à atteindre un seuil limite (8 crues de calage).

Enfin, au terme de cette étude et pour le site étudié, on peut dire que le modèle réseau de neurones est le mieux adapté à la modélisation de la relation pluie-débit au pas de temps horaire.

Bibliographie

AFFANE A. et **HOUGLAOUANE M.**, (1997): " Contribution à l'étude et à la prévision de la pluviométrie sur le littoral Algérien". Mémoire de fin d'étude, ENP, Alger.

BASTIN G. (1984): "Identification de relation pluie-débit, prévision des débits en temps réel ", dans session de formation continue "Crues et précipitations intenses", I.N.P.G., Fascicule III bis, pp. 1-30.

BENKACI T. (2001): "Contribution à la modélisation de la relation pluie-débit au pas de temps mensuel et journalier par les modèles conceptuels et les réseaux de neurones". Th. Magister. Institut national agronomique d'Alger.

BOBEE B., (1999): "The gamma family and derived distribution applied in hydrology". Water resource application, Colorado, U.S.A, pp 50-69, 1999.

BOUZIANE A. (1989): " Contribution a la modélisation stochastique de débits mensuels et a la recherche de la meilleure règle de gestion d'un réservoir", Thèse présentée de l'obtention du grade de docteur en sciences appliquées, Université Catholique de LOUVAIN (Belgique), unité Génie Civil.

BOX G.E.P et **JENKINS G.M.**, (1970): "Time series analysis, Forecasting and control, Holden Day", SAN FRANCISCO, p 578.

CEMEGREF, (1989): "Recherche sur la mise au point d'un modèle Journalier Pluie Débit".

CHAMBAZ H., (1990): "Contribution à l'étude de la Méthode DPFT", Thèse de Magister, ENP, Alger.

DAVALO E. et **NAIM P.**, (1993): "Des réseaux de Neurones", Edition Eyrolles, Paris.

DECHEMI N., **BENKACI T.** and **ISSOLAH A.**, (2003): "Modélisation des débits mensuels par les modèles conceptuels et les systèmes Neuro-Flous", p. 407-425.

DELLEUR J.W. (1984): "Contribution to the recursive estimation of the rainfall-runoff relationship for flood forecasting-Application to the Gardon d'Anduze". Rapport CE-HSE-84-4. Purdue Univ. USA, 124 p.

DONNA RIZZO M., **DAVID DOUGHERTY E.**, (1994): "Characterization of Aquifer properties using Artificial Neural Network", Water Resources Research, Vol. 30(2), p. 483-497.

DUBAND D., (1978): "Fonction de transfert pluie-débit sur des bassins versants de l'ordre de 1000 Km²". Note Interne, EDF-DTG, Grenoble.

FRENCH M.N., **KRAJEWSKI W.F.**, **CUYKENDALL. R.R.**, (1992): "Rainfall Forecasting in Space and Time using a Neural Network", Journal of Hydrology, Vol.137, p. 1-31.

HARTANI R. (1995): "Modélisation des systèmes Neuro-Flous, contribution théorique et applications". Th. doct. Université Paris VI.

JATON J. F., (1982): "Contribution à l'étude des Relation Pluie Débit dans les petits Bassins Versant Ruraux", Thèse de Doctorat, Ecole Polytechnique Fédérale de Lausanne.

JODOUIN J. F., (1994): "Les réseaux de neurones", Edition Hermès, Paris.

LIAMAS J., (1985): "Hydrologie générale: principes et applications. Gaetan Morin, Québec, pp. 38-60.

MINSKY M., POPERT S. (1985): "Perceptron. The MIT Press".

NALBANTIS I. (1987): " Identification des modèles pluie-débit du type hydrogramme unitaire : développements de la méthode DPFT et validation sur données générées avec et sans erreur ", Thèse présentée de l'obtention du titre de docteur en mécanique, Institut National Polytechnique de Grenoble.

NEWTON D. W. et VINYARD J. W. (1967): "Computer Determinate Unit Hydrographs From Floods". Journal of Hydraulic Division, ASCE, VOL. 93, N°HY5. PROC. Paper 5449. pp. 219-234.

RAHAI A. (2002) : "Prévision de la pluviométrie par combinaison de l'ACP et des chaînes de MARKOV", Mémoire de P.F.E, ENP, Alger.

RODRIGUEZ J. Y., (1989): "Modélisation Pluie-Débit par la méthode DPFT", p. 42-65, Thèse de Doctorat, INPG, Grenoble.

SAIDOUNI Dj. (2003) : "Modélisation de la relation pluie débit au pas de temps horaire par différentes approches", Thèse de Magister, ENP, Alger.

SAJIKUMAR N., THANDAVESWARA B.S., (1999): "A Non-Linear Rainfall-Runoff model using an Artificial Neural Network", Journal of Hydrology, Vol. 216, p. 32-55.

SAKAKIMA S., KOJIRI T., ITOH K., (1993): "Real-Time reservoir operation with neural network concept", Computational Mechanics Publications, Southampton, p 502-514.

SHAMSELDIN A. Y., (1996): "Application of a Neural Network Technique to Rainfall-Runoff Modeling", Journal of Hydrology, Vol. 199, p. 272-249.

TOUZET C. (1992) : "Les réseaux de neurones artificiels". Cours. EERIE.

YEDDOU Y., (1998): "Etude de Synthèse sur les Réseaux de Neurones", Thèse de Magister, ENP, Alger.

ملخص:

لقد حاولنا في إطار هذا البحث إيجاد وسيلة فعالة لدراسة توقعات منسوب الماء خاصة الذي ينجر عن الفيضانات و الذي يعد اليوم مشكلا عويصا لازال العلماء يولون عناية فائقة دون الوصول إلى نتيجة مرضية لتفادي هذا المشكل.

ففي هذا الموضوع وفي نفس الاتجاه نقترح دراسة و تطبيق على معطيات حقيقية لثلاثة طرق فنية للتوقعات و هي: الاختلافات الأولية لدالة التحويل ، شبكات الخلايا العصبية الاصطناعية و التناقص الذاتي معدل الحركة مع إقحام متغير صريح.

تعتبر هذه الدراسة وسيلة لإبراز قدرة و نجاعة كل طريقة ليتسنى اختيار الطريقة المناسبة التي يكون لها الأثر الحسن في التحكم في الخطر المتوقع.

الكلمات المهمة: ش خ ا ، ا ا د ت ، ت د م ح ، توقعات ، التحكم في الخطر.

Résumé :

Dans le cadre de ce travail, nous avons essayé de trouver un moyen efficace pour la prévision de débits en particulier ceux engendrés par les crues, un problème sur lequel les scientifiques se penchent actuellement, sans pour autant avoir trouver un palliatif.

C'est dans cette optique que nous proposons l'étude et l'application sur des données réelles de trois techniques de prévision à savoir la différence première de la fonction de transfert (DPFT), les réseaux de neurones artificiels (RNA) et l'autorégressif moyenne mobile avec l'introduction d'une variable exogène (ARMAX).

Cette étude sera un moyen de mettre en évidence la capacité et la fiabilité de chacune des méthodes et d'en sortir celle qui aura un meilleur impact sur la gestion du risque.

Mots clés : RNA – DPFT – ARMAX – prévision – gestion du risque.

Abstract:

In this work we tried to find effective for rain off forecasting in particular those generated by streams, problem whose scientist are currently trying to find a solution.

The study that we propose is based on the application of three methods of forecasting, the First Difference of the Transfer Function (FDTF), the Artificial Neural Network (ANN) and the AutoRegressive Moving Average with eXogenous input (ARMAX).

This study will be a mean of highlighting the capacity and the reliability of each method and which of them will have a better impact on risk management.

Keywords : ANN – FDTF – ARMAX – forecasting – impact on risk.