

INSTITUT NATIONAL AGRONOMIQUE D'ALGER
*THESE POUR L'OBTENTION DU DIPLOME DE DOCTORAT ES-SCIENCES EN SCIENCES
AGRONOMIQUES*

***MODELISATION PLUIE-DEBIT
MENSUELLE ET JOURNALIERE PAR LES
MODELES CONCEPTUELS ET LES
SYSTEMES NEURO-FLOUS (APPLICATION
AUX BASSINS ALGERIENS)***

PRESENTEE ET SOUTENUE PAR:

BENKACI ALI TARIK

Mr. N. DECHEMI Professeur (Ecole Nationale Polytechnique) Directeur de thèse
Année universitaire : 2005 2006

Devant le jury composé de : Mr. A. AIDAOUI Professeur Président Mme. B. TOUAIBIA Professeur
(Ecole Nationale Supérieure de l'Hydraulique – BLIDA) Examineur Mr. M. CHERRARED M. de
conférences (Université de Bejaia) Examineur Mr. T. HARTANI M. de conférences (Institut National
Agronomique) Examineur

Table des matières

..	1
Remerciements . .	3
Resume .	5
Abstract . .	7
ص غ لم . .	9
Introduction générale . .	11
Chapitre I : La modélisation pluie-débit en hydrologie . .	15
I.1. Introduction : Etude des processus hydrologiques .	15
I.1.1. Les facteurs de génération des débits . .	17
I.2. La modélisation pluie-débit en hydrologie . .	18
I.2.1. Définition d'un modèle mathématique .	19
I.2.2. Pourquoi des modèles pluie-débit ? . .	19
I.2.3. Propriétés requises des modèles .	21
I.2.4. Des modèles en fonction des caractéristiques hydrologiques des bassins .	22
I.3. Classification des modèles pluie-débit . .	23
I.3.1. Les modèles empiriques .	24
I.3.2. Les modèles ' boîte noire' .	25
I.3.3. Les modèles conceptuels .	25
I.3.4. Les modèles physiquement fondés .	25
I.4. Paramétrisation des modèles pluie-débit . .	27
I.4.1. Estimation des paramètres : calage du modèle .	27
I.4.2. Validation des modèles . .	29
I.4.3. Erreurs des modèles .	30
I.5. Synthèse bibliographique .	30
I.5.1 Les modèles ' boîte noire ' .	30

I.5.2. Les modèles conceptuels (à réservoirs) .	32
I.5.3. Les modèles physiquement fondés .	35
I.5.4. Nouveaux outils de la modélisation pluie-débit : Apport de la télédétection . .	36
I.6. Conclusion . .	37
Chapitre II Etude hydrologique et statistique .	39
II.1. Introduction .	39
II.1.1. Utilisation des données hydrométéorologiques . .	40
II.2. Aire d'étude et présentation des bassins .	41
II.2.1. Les données utilisées .	44
II.3. Critique et analyse statistique des données . .	48
II.3.1. Traitements statistiques des données .	49
II.4. Ajustement statistique des données .	53
II.4.1. Estimation des paramètres . .	54
II.4.2. Ajustement des données annuelles à la loi normale .	55
II.4.3. Ajustement des données mensuelles à la loi de Pearson III . .	56
II.5. Critères de validation des modèles pluie-débit . .	58
II.5.1. Les critères numériques . .	59
II.6. Méthodes d'optimisation des modèles pluie-débit . .	62
II.6.1 Les méthodes locales .	62
II.6.2. Les méthodes globales . .	63
II.7. Conclusion .	65
Chapitre III Simulation de la relation pluie-débit par les modèles conceptuels . .	67
III.1. Introduction . .	67
III.1.1. Aspects de la modélisation pluie-débit .	68
III.1.2. Méthode d'optimisation .	69
III.1.3. Mise en œuvre des modèles .	69
III.2. Les modèles conceptuels mensuels . .	70
III.2.1. Présentation des modèles . .	71

III.2.2. Résultats de la modélisation . .	77
III.3. Les modèles conceptuels journaliers . .	83
III.3.1. Présentation des modèles . .	84
III.3.2. Résultats de la modélisation . .	93
III.4. Conclusion . .	101
Chapitre IV : Etude de la relation pluie-débit par les réseaux de neurones artificiels . .	103
IV.1. Introduction . .	103
IV.2. Historique des réseaux de neurones artificiels . .	104
IV.3. Aspects théoriques et généralités sur les réseaux de neurones . .	105
IV.3.1. Le neurone biologique . .	105
IV.3.2. Le neurone formel (artificiel) . .	106
IV.3.3. Fonctionnement mathématique du neurone artificiel . .	107
IV.4. Structure générale des réseaux de neurones . .	108
IV.4.1. Classification des réseaux de neurones artificiels . .	108
IV.5. Apprentissage supervisé des réseaux de neurones artificiels . .	110
IV.5.1. Propriétés des réseaux de neurones à apprentissage supervisé . .	110
IV.5.2. Architecture et apprentissage des réseaux multicouche . .	110
IV.5.3. Réseaux de neurones à base linéaire et à base radiale . .	111
IV.5.4. Algorithmes d'apprentissage des réseaux à base linéaire . .	112
IV.6. Réseaux de neurones multicouche (Perceptron multicouche) . .	113
IV.6.1. Algorithmes d'apprentissage supervisé . .	113
IV.6.2. La Backpropagation . .	114
IV.6.3. Méthodes d'optimisation avec la Backpropagation . .	116
IV.6.4. Propriétés des réseaux de neurones . .	118
IV.7. Résultats de simulation par les réseaux de neurones . .	118
IV.7.1. Données mensuelles . .	120
IV.7.2. Données journalières . .	123
IV.7.3. Résultats des périodes caractéristiques . .	125
IV.8. Conclusion . .	126

Chapitre V : Introduction à la logique floue : généralités et aspects théoriques . .	129
V.1. Introduction .	129
V.2. Notions de base et connaissances préliminaires . .	130
V.2.1. Définition d'une variable linguistique .	130
V.2.2 Sous-ensembles flous : concepts et définition . .	131
V.2.3. Opérations sur les ensembles flous .	132
V.2.4. Les systèmes d'inférence floue (SIF) .	133
V.2.5. Structure interne d'un système d'inférence floue (SIF) . .	134
V.3. Construction de la relation entrée sortie à partir des règles floues . .	137
V.3.1 Proposition floue et inférence floue . .	137
V.3.2. Implication floue .	138
V.3.3. Système d'inférence flou à deux (ou plusieurs) entrées (MISO) .	138
V.3.4. Opérateurs d'implication flous .	139
V.3.5. Approximation par les modèles flous . .	141
V.4. Modélisation par les systèmes flous . .	141
V.4. Différents types de modèles flous . .	142
V.4.1. Les modèles linguistiques (de Mamdani) . .	142
V.4.2. Les modèles de Sugeno (Takagi-Sugeno-Kang) .	143
V.5. Apprentissage des modèles flous .	145
V.5.1. Identification de la structure du modèle . .	146
V.5.2. Identification des paramètres .	146
V.5.3. Optimisation neurofloue : le modèle ANFIS .	148
V.6. Conclusion . .	150
Chapitre VI : Résultats de la simulation par le modèle flou, comparaison entre divers modèles . .	153
VI.1. Introduction . .	153
VI.2. Exemple de simulation par la logique floue : cas d'une relation pluie-débit .	154
VI.2.1. Modèle utilisé .	156
VI.3. Résultats de la simulation .	157

VI.3.1. Données mensuelles . . .	158
VI.3.2. Données journalières . . .	160
VI.3.3. Résultats du modèle flou par périodes caractéristiques . . .	163
VI.4. Modèle flou à deux entrées . . .	164
VI.4.1. Introduction de l'ETP . . .	165
VI.4.2. Introduction de l'humidité du sol . . .	167
VI.5. Analyse comparative des modèles pluie-débit étudiés . . .	173
VI.5.1. Résultats des modèles en période de calage: Robustesse et qualité de simulation des modèles . . .	173
VI.5.2. Simulation des débits caractéristiques . . .	176
VI.5.3. Résultats des modèles en période de validation : Capacité prédictive des modèles . . .	176
VI.5.4. Structure des modèles : Paramétrisation et parcimonie des modèles . . .	177
VI.5.5. Cas du modèle neuro-flou à deux entrées . . .	177
VI.6. Modèle neuro-flou journalier à trois entrées . . .	178
VI.7. Reconstitution des séries hydrométriques tronquées . . .	180
VI.8. Conclusion . . .	180
Conclusion générale . . .	183
Références bibliographiques . . .	187
Annexes . . .	197
Annexe 1 : Présentation des bassins . . .	197
Annexe 2 : Présentation des modèles . . .	201

*Dédicace A ma mère A mon père A mes frères et sœurs A la mémoire de Mr A.
Issolah que Dieu l'accueille dans son vaste paradis*

Remerciements

Au terme de cette thèse je tiens à remercier :

Mr N. Dechemi professeur à l'Ecole Nationale Polytechnique (ENP) de m'avoir intégré dans son équipe de recherche, d'accepté la direction de cette thèse, de m'avoir encadré, il a été pour moi un enseignant privilégié de par sa gentillesse, son expérience, je le remercie pour sa rigueur, ses encouragements ainsi que pour ses critiques et conseils constructifs pour aboutir à cette thèse.

Je tiens à remercier vivement:

Mr A. Aidaoui professeur à l'Institut National Agronomique (INA) d'avoir accepté de présider le jury. Je le remercie pour l'intérêt qu'il a manifesté pour ce travail.

Je remercie Madame B. Touaibia professeur à l'ENSH (Blida), Mr M. Cherrared maître de conférences à l'université de Bejaia, et Mr T. Hartani maître de conférences à l'Institut National Agronomique d'avoir accepté de lire et de juger mes travaux, je les remercie pour leurs critiques et précieux conseils.

Je remercie le personnel du département d'hydrologie de l'Agence Nationale des Ressources Hydraulique, en particulier madame Abda pour sa disponibilité et ses encouragements, et de m'avoir bénéficié des données hydrométriques.

Je tiens à remercier Mr Taibi, d'avoir accepté d'être un invité pour enrichir le débat sur cette recherche.

Je remercie le personnel du service d'exploitation de l'Agence Nationale des Barrages et des Transferts pour leur aide dans l'acquisition des données.

Je remercie également Mr A. Bermad, maître de conférence à l'ENP, Mme Souag chargée de cours à l'USTHB, pour leur aide et encouragements.

Mes remerciements vont à monsieur Pierre-Yves Glorennec, professeur à l'IRISA de Rennes, pour son aide et ses encouragements en particulier dans le domaine de la logique floue.

Je remercie toutes les personnes du département du génie rural, et du service de la post-graduation qui m'ont aidé durant ces cinq années de recherches.

Resume

Ce travail de recherche porte pour l'essentiel sur trois approches distinctes de modèles pluie-débit testés sur les bassins versants du Nord algérien, à deux échelles mensuelle et journalière.

La première approche proposée est une application de quatre modèles conceptuels à réservoirs à deux échelles temporelles. Les modèles testés, nous ont permis d'effectuer une étude analytique sur l'utilisation de ce type de modèles. Nous avons ainsi souligné la simplicité et la parcimonie des modèles conceptuels, ayant remarqué la supériorité des modèles GR, aux deux pas de temps, nous avons dégagé deux principales conclusions : la baisse de niveau de simulation en phase de vérification, en particulier au pas de temps mensuel, ainsi que la difficulté de reconstitution des débits caractéristiques de la série historique.

La seconde approche de modèles utilisée est basée sur les réseaux de neurones artificiels. Ces modèles qui s'inspirent du cerveau humain, très utilisés en hydrologie, permettent d'introduire un nouveau concept de simulation basé sur l'apprentissage de données. Dotés d'algorithmes d'optimisation non-linéaires, ils prennent l'avantage de simuler de façon correcte et parcimonieuse la relation pluie-débit mensuelle. Au pas de temps journalier, leur application s'avère délicate et cela, vu la complexité de la relation pluie-débit à cette échelle.

La troisième méthode de simulation proposée dans cette thèse, se base sur l'utilisation d'une logique appelée floue. En s'appuyant sur la syntaxe de type Si... Alors, ce thème nous a permis de discuter plus en détails des modalités de mise en œuvre de cette nouvelle génération de modèles. Nous présentons ensuite, une analyse quasiment exhaustive de certaines parties du modèle, ses avantages et ses inconvénients.

L'optimisation de la relation pluie-débit a été notre objectif principal. Ainsi, après une analyse comparative, nous avons pu introduire une nouvelle base de simulation à travers une combinaison de modèles. Les sorties simulées d'un modèles conceptuel, telles que l'humidité du sol et les débits, ont été introduites comme entrées aux modèles neuronaux et flous.

Nous avons ainsi amélioré de façon remarquable la qualité de simulation, que ce soit au pas de temps journalier, ou mensuel.

Cette approche offre ainsi des pistes nouvelles et innovantes dans le domaine de la modélisation hydrologique et de son application à des problèmes spécifiques de gestion des eaux.

Abstract

This research task relates basically to three applications of hydrological models in North Algerian watershed on monthly and daily scale.

The first approach suggested is an application of four conceptual models on two time scales.

The models tested allow to execute an analytic study about of conceptual models use. Thus, simplicity and parsimony of the GR category models were evaluated, such as performance and robustness. However, it's should be observed two basic conclusions: the conceptual model inefficiency on verification period, and the incapacity to reproduce characteristics streamflow like peak and base flow.

The second approach of models is based on artificial neural network. These models are inspired by human brain and are recently used in hydrology; allow introducing new concept of simulation, based on data learning process. Using power non-linear algorithms, these models give good results on monthly time scale. However, due daily rainfall-runoff relationship complexity, we have noted the difficulties to simulate runoff on this scale.

The third approach makes use of a new model based on the logic called fuzzy. Concept supported by the syntax If ... Then, this topic allows more detailed discussion of methods to execute of this new generation of models. Exhaustive sensitivity analyses of certain party of models are then presented, his advantages and disadvantages.

Rainfall-runoff process modelling improvement was a principal objective. Then, after comparative analysis between several models, we have introduced a new simulation method, through combination of these models. Thus the simulated outputs like than soil moisture and calculated streamflow are introduced like input to neuro-fuzzy models.

We have improved qualitatively and quantitatively the results, for the two time scales.

That approach offers some news tracks research innovations in the fields hydrological modelling for its application to specific problem in water management.

ص خ لم

.

.

#####. ##### ##### ##### ##### ##### ##### ##### ##### ##### #####

.

#####. ##### ##### ##### #####
.

.

Introduction générale

Depuis la nuit des temps, l'eau a été un facteur prédominant dans le développement socio-économique de l'homme. Ainsi, les grandes civilisations antiques se sont développées sur les berges des grands fleuves tels que le Nil l'Euphrate, où l'homme avait un meilleur accès à la ressource en eau. A travers des siècles, l'homme a essayé de domestiquer le capital eau à travers des barrages et des aménagements hydrauliques, pour mieux se prévenir des effets néfastes de la sécheresse et des risques des inondations.

Avec l'explosion démographique, la révolution industrielle et les diverses formes de vies, les besoins en eau se sont considérablement accrus, et les problèmes de disponibilité de l'eau se sont amplifiés.

Le développement économique et la croissance de la population ont exercé des pressions qui ont eu pour effet de modifier le régime des eaux de surface et le paysage des plaines inondables. Les efforts humains visant à restreindre la zone active des plaines inondables sont variés: construction de barrages, de digues et d'ouvrages de dérivation, dragage et réaligement des lits, et assèchement des marécages. Ces mesures, qui profitent sans nul doute au développement économique, ont par ailleurs entraîné le déclin des habitats des poissons et de la faune, et ont perturbé des écosystèmes entiers.

En ce début du vingt et unième siècle, plusieurs pays du monde seront confrontés à une grave crise de l'eau. Tous les signes suggèrent que cette crise s'intensifiera et que la situation continuera d'empirer si aucune action correctrice ne sera menée.

Au cours des vingt cinq dernières années, de grandes conférences internationales sur l'eau ont été organisées. Les préparatifs qui les ont précédées, ainsi que les discussions qui s'en sont suivies ont modifié notre perception de la crise de l'eau et ont élargi nos connaissances en ce qui concerne les réponses à rechercher.

Depuis quelques décennies, les demandes en eau douce pour l'environnement et les activités humaines tendent à rejoindre globalement, et même dépasser localement, les quantités disponibles. Des études récentes (1998-2000) prévoyaient une pénurie d'eau pour environ 2 milliards d'individus et une augmentation des besoins en eau de l'ordre de 20% d'ici à 2025. L'aspect crucial des solutions à développer réside dans la gestion intégrée tenant compte à la fois des ressources et des demandes, environnementales et sociales à travers les différents flux associés d'eau, d'énergie et d'éléments chimiques.

La demande socio-économique sur la connaissance des ressources en eau, sur l'influence des aménagements et de l'occupation des sols sur le régime et la qualité des eaux, sur l'effet des pratiques agricoles, sur le devenir des rejets ponctuels ainsi que sur la prévention des risques hydrométéorologiques, implique de comprendre les variabilités spatiales et temporelles des apports et des écoulements, et leurs couplages avec les mécanismes météorologiques, climatologiques et géochimiques.

Cependant la gestion de la crise de l'eau ne peut être réalisée qu'avec la maîtrise des différentes formes extrêmes de l'écoulement. C'est ainsi que les chercheurs se sont investis dans la connaissance des études du cycle hydrologique de l'eau. Ce dernier peut être vu comme l'intégrateur spatio-temporel des flux d'eau, d'énergie et de matières (dissoutes, en suspension) dans les différents compartiments constituant les surfaces continentales (sous-sol, sol, végétation, atmosphère, réseaux hydrographiques, océan côtier). La maîtrise de l'évolution du cycle de l'eau a pour but d'anticiper et de se prévenir des catastrophes hydrométéorologiques telles que les crues et les sécheresses.

Ainsi, les ressources en eau en Algérie, limitées vulnérables et inégalement réparties, subissent depuis plus de deux décennies les effets néfastes de la sécheresse et de la pollution.

L'accroissement brutal des besoins, les retards enregistrés dans les programmes de mobilisation et de transfert et l'absence d'outils de planification et de gestion ont généré des situations de déficit chronique, et se sont répercutés de façon négative sur le développement socio-économique du pays.

Face à cette situation, et afin de relever les défis que posent les problèmes de l'eau en Algérie, des mesures d'urgence peuvent être prises en vue d'une meilleure politique de l'eau qui sera fondée sur des principes nouveaux de gestion intégrée de la ressource. Parmi ces mesures nous pouvons citer:

La protection des ressources en eau existantes, par la réhabilitation des réseaux d'alimentation en eau potable, d'assainissement et la lutte contre le gaspillage.

L'utilisation des ressources non conventionnelles, notamment par le recyclage des eaux usées et le dessalement de l'eau de mer.

La gestion rationnelle des ressources en eau.

La mobilisation systématique des ressources disponibles par la conception de grands

barrages et des retenues collinaires, qui peuvent contribuer à l'alimentation en eau au niveau régional et local.

Cependant, la première étape de gestion des ressources en eau consiste en leur estimation, et en la maîtrise de leur évolution.

Pour diverses raisons, plusieurs stations hydrométriques ont été mal gérées, et les séries hydrométéorologiques comportent plusieurs lacunes.

Avec le développement de l'informatique, et grâce au processus croissant de stockage électronique de l'information et du savoir, un nouveau domaine s'ouvre en hydrologie, entre les lieux de production de l'information et du savoir, au moyen de la recherche, et les lieux d'application, au moyen de l'ingénierie et de la gestion pratique, c'est le domaine de l'hydro-informatique. Ainsi, plusieurs axes de recherches se sont développés à travers des modèles qui visent à réduire la complexité des phénomènes hydrologiques.

La modélisation du comportement hydrologique des bassins versants est incontournable dès lors que l'on s'intéresse à des problématiques relatives à l'évaluation et la gestion optimale des ressources en eau. Ceci s'illustre par un aspect quantitatif, dans les pays où l'alimentation en eau est un facteur limitant.

La transformation de la pluie en débit résulte d'un certain nombre de mécanismes complexes qui interviennent simultanément à différentes échelles spatiales et temporelles. De ce fait le modèle pluie-débit trouve sa nécessité dans la mesure où le modèle conçu à partir des chroniques de pluies observées, peut générer des débits qui soient le plus proches possibles des débits observés, c'est à dire à partir des observations de pluie tombée, on peut arriver à reproduire (ou à prédire) la réponse en débit du bassin.

De tels modèles sont particulièrement utiles pour de nombreuses applications, par exemple la prévision des crues, où le fait de remonter à l'origine des débits (les pluies) permet d'allonger le délai de prévision par rapport à des méthodes n'exploitant pas cette information, la prévision des étiages etc.

Décrire la transformation de ces pluies en débits dans les cours d'eau est le sujet de la modélisation pluie-débit. Bien qu'il s'agisse là d'un problème apparemment très ordinaire, aucune solution satisfaisante n'existe à l'heure actuelle pour décrire le mécanisme de cette transformation. Pourtant des décisions d'aménagement, de prévision, de gestion sont à prendre: cela nécessite des modèles.

Cette thèse s'inscrit dans le domaine de la modélisation de la transformation de la pluie en débit et de sa représentation à l'échelle du bassin versant. Notre objectif principal est de tester des modèles de conceptions différentes à différents pas de temps (mensuel et journalier), et de les adapter s'il y a lieu, aux caractéristiques hydroclimatiques des bassins algériens.

La thèse se divise en six chapitres:

Le premier chapitre est dédié à l'étude des processus de génération des débits à l'exutoire d'un bassin, et à la description de la modélisation pluie-débit, des différentes classes de modèles, leurs diverses applications, avec une synthèse bibliographique portant sur les modèles pluie-débit à différentes échelles, leurs avantages et

inconvenients, où nous avons essayé de tracer l'évolution historique des différents modèles.

Le deuxième chapitre est consacré à l'analyse statistique et hydrologique, où nous présentons les bassins versants qui seront étudiés dans notre thèse, leurs caractéristiques climatiques et hydrologiques, ainsi que les différents tests statistiques essentiels à l'analyse des données hydroclimatiques qui seront introduites dans le modèle, et les ajustements aux lois statistiques.

Dans le troisième chapitre, nous détaillons les modèles conceptuels mensuels et journaliers, tant que sur le plan théorique et descriptif que sur le plan pratique. La comparaison entre les résultats des divers modèles, où il sera question de sensibilité des modèles entre les deux phases de calage et de vérification, ce chapitre nous permet d'identifier les différentes problématiques de la modélisation pluie-débit par les modèles conceptuels.

Les réseaux de neurones artificiels ont souvent été l'objet de recherche dans de nombreuses disciplines scientifiques et dans plusieurs optiques telles la classification, la prévision, ou la commande. La relation pluie-débit peut être considérée comme un problème d'extrapolation entre un système de variables d'entrée (pluies) et un système de variables de sortie (débits). C'est dans ce sens, que les réseaux de neurones ont été testés dans un cadre de modèles pluie-débit, à travers une synthèse théorique, mais également pratique dans le quatrième chapitre.

Dans le cinquième chapitre, nous avons abordé la théorie de la logique floue, où nous présentons quelques généralités et les concepts de base des systèmes d'inférence flous (SIF), les variables linguistiques et les significations et descriptions floues. Nos efforts se sont concentrés sur les SIF de type Takagi-Sugeno-Kang (TSK), dont nous avons détaillé leur principe d'apprentissage.

Le sixième chapitre est consacré aux résultats de la modélisation pluie-débit par la logique floue, ainsi qu'aux principales conclusions concernant l'utilisation des modèles neuroflous. Notre objectif dans cette étude, est de montrer comment la logique floue et les systèmes d'inférence floue peuvent contribuer à l'amélioration de la simulation pluie-débit.

Dans ce chapitre, nous avons aussi tenu à explorer une nouvelle optique de modélisation pluie-débit. En effet, après la comparaison de modèles, il s'est avéré que les débits observés peuvent être simulés de façon optimale. Dans ce cas, les sorties calculées par le modèle conceptuel, à savoir l'humidité du sol et le débit simulé constitueront des variables d'entrée aux modèles neuronaux et flous. La combinaison des deux approches conceptuelle et floue nous a permis de calculer des débits très proches de ceux observés, cette nouvelle technique de simulation sera explorée en vue de reconstituer les séries hydrométriques tronquées, nous l'avons intitulée la reconstitution pluie-débit.

Une conclusion générale ressortira les principaux résultats obtenus dans le cadre de cette recherche, ainsi que les perspectives d'amélioration des modèles pluie-débit tant que sur le plan théorique que pratique.

Chapitre I : La modélisation pluie-débit en hydrologie

I.1. Introduction : Etude des processus hydrologiques

L'eau qui intervient dans l'ensemble du fonctionnement du milieu naturel, représente un vecteur principal d'éléments transportés (en solution ou en suspension), et un milieu de vie pour de nombreux végétaux et animaux, est devenue un enjeu écologique et économique considérable.

La demande socio-économique sur la connaissance de la ressource en eau, sur l'influence des aménagements et de l'occupation des sols sur le régime et la qualité des eaux, sur l'effet des pratiques agricoles, sur le devenir des rejets ponctuels ainsi que sur la prévention des risques hydrométéorologiques, implique de comprendre les variabilités spatiales et temporelles des apports et des écoulements, et leurs couplages avec les mécanismes météorologiques, climatologiques et géochimiques.

C'est à ce propos, que l'importance de la tâche des hydrologues dans la résolution des questions environnementales actuelles et futures est considérée de plus en plus reconnue.

Le fonctionnement hydrologique des bassins versants reste encore assez mal connu,

même si les études de terrain menées depuis une cinquantaine d'années sur de petits bassins de recherche ont permis d'apporter des réponses partielles à deux des questions fondamentales qui se posent à l'hydrologue (Ambroise, 1991):

Que devient l'eau de pluie? (Penman, 1963);

D'où provient l'eau des ruisseaux? (Hewlett, 1961).

La représentation la plus connue de la formation des débits, est celle proposée par Horton (Horton, 1933) d'un partage de pluie disponible au niveau de la surface du sol, après interception éventuelle par la végétation:

Dès que l'intensité des pluies dépasse la capacité d'infiltration du sol, il y a saturation de la couche supérieure du sol, l'eau en excès est évacuée rapidement par ruissellement en surface pour former l'écoulement rapide de crue.

L'autre partie s'infiltrate et contribue par un écoulement beaucoup plus lent à travers le sol et les formations superficielles à la recharge des nappes et au débit de base.

Cette théorie nous permet de distinguer quatre cheminements de l'eau des rivières:

Précipitations directes à la surface libre des cours d'eau;

Écoulement de surface, ou ruissellement de surface qui se décompose en ruissellement par dépassement de la capacité d'infiltration (écoulement hortonien) et ruissellement par saturation du sol;

Écoulement de subsurface, c'est un écoulement dû à un cheminement quasi-horizontale de l'eau dans les macropores du sol;

Écoulement souterrain.

Par sa simplicité et son réalisme, cette théorie a connu un énorme succès en ingénierie hydrologique.

Cependant des études de terrain, notamment en milieu tempéré, ont montré que cette représentation correspondait en fait à un cas particulier extrême, puisque par exemple des crues sont fréquemment observées pour des intensités de pluies bien inférieures à l'infiltrabilité des sols (Ambroise, 1991).

D'autre part, ce type de ruissellement est loin d'être généralisé à tout bassin versant, et il n'est actif que sur une partie limitée de ce bassin.

Une part de ruissellement de crue peut provenir essentiellement de l'écoulement hypodermique dans les sols, et aussi par une recharge des nappes peu profondes en aval.

Une des critiques qui a remis en cause cette théorie, concerne une part de ruissellement des crues provenant des écoulements par une recharge rapide des nappes. Cette conception d'une contribution souterraine, évoquée par certains auteurs (Hursch, 1936), a mis beaucoup de temps à s'imposer. Les mécanismes permettant d'expliquer la rapidité des transferts des flux à travers le sol, et l'exfiltration de leur part, ont été difficiles à identifier.

Selon Hursh et Brater (1941), les débits sont générés dans la plupart des cas, par plusieurs processus simultanément ou successivement, dans des combinaisons très

variables dans le temps et dans l'espace, et dépendant dans chaque cas de plusieurs facteurs contrôlant à la fois la présence ou l'absence de chaque processus et son intensité.

De façon objective, les processus de génération des débits sont nombreux et variables selon les bassins et les climats.

I.1.1. Les facteurs de génération des débits

Dans chaque bassin les processus actifs et leurs interactions ne sont pas quelconques, mais dépendent de la combinaison d'un ensemble de facteurs contrôlant leur variabilité spatio-temporelle.

L'apport de précipitations constitue le premier facteur intervenant dans les débits. Durant une averse, le volume et la forme de l'hydrogramme des crues, dépendent à la fois de l'apport total des précipitations, mais aussi de son intensité et de sa durée. La battance peut jouer un rôle important dans la genèse des crues majeures des bassins versants agricoles (Papy et Douyer, 1991).

De même, l'apport d'énergie (radiative ou advective) entre deux événements pluvieux, qui contrôle l'évapotranspiration peut influencer sur la redistribution de l'eau dans les sols.

Les conditions initiales jouent un rôle majeur sur l'intensité des crues. Pour un même apport d'eau et d'énergie, ne correspond pas forcément le même débit. L'état initial hydrique et hydrologique du bassin influence fortement sur la réponse du bassin à l'exutoire.

La transformation de la pluie en débit dépend largement de l'état hydrique et hydrologique initial du bassin, contrôlé par les intensités, durées, fréquences et séquences des périodes pluvieuses et sèches précédentes.

Cette non-linéarité de la réponse en fonction des conditions initiales s'explique par des combinaisons et intensités de processus actifs différentes dans chaque cas.

La variabilité spatio-temporelle pédologique ainsi que la répartition et l'occupation de la végétation constitue un facteur prédominant de la réponse du bassin aux conditions hydrologiques. Ces caractéristiques vont déterminer de façon directe les propriétés des bassins quant aux différents mécanismes intervenants directement dans les processus de génération des débits. La végétation qui par sa densité a une influence sur l'interception, tandis que la structure pédologique du sol, détermine les propriétés hydrodynamiques du sol, et donc sur l'infiltration et le ruissellement de l'eau des pluies vers l'exutoire.

Enfin les caractéristiques géométriques telles que la topographie et la morphométrie vont influencer fortement les conditions d'écoulement (vitesse de propagation des débits), et donc sur l'intensité et la dynamique des crues.

Les mécanismes de génération de l'écoulement n'agissent pas de façon isolée, mais forment un domaine continu de processus. Il s'ensuit que sur un même bassin versant, plusieurs processus concomitants peuvent intervenir durant un épisode pluvieux. De plus, il se peut que le type de processus change selon le type d'événements. Les crues générées durant l'été ou l'hiver ne font pas nécessairement appel aux mêmes processus.

Ceci, reflète la difficulté de proposer une classification de l'ensemble des processus de génération des écoulements.

Cette description sommaire des différents facteurs d'écoulement, reflète le caractère dynamique et complexe du fonctionnement hydrologique d'un bassin versant.

Ainsi, la génération des débits au niveau d'un bassin, résulte d'un grand nombre de processus superficiels et souterrains qui interagissent d'une façon plus générale dans le cycle de l'eau à des échelles spatio-temporelles différentes (Ambroise, 1991).

I.2. La modélisation pluie-débit en hydrologie

Nous avons évoqué la réalité complexe du fonctionnement hydrologique du bassin, et l'hétérogénéité à différentes échelles dans la génération des débits à son exutoire. Dans ce cas, le bassin est considéré comme un système dynamique, dont les divers processus hydrologiques sont assez mal connus.

A partir de cette problématique, plusieurs recherches ont été entamées sur le terrain, en vue d'une meilleure compréhension des processus de génération des débits, et les facteurs qui les contrôlent.

Pour expliquer la formation de l'écoulement à l'exutoire d'un bassin, certains hydrologues ont utilisé des méthodes pratiques, visant à décomposer l'hydrogramme du bassin en ses différentes composantes.

En ingénierie hydrologique, la méthode la plus utilisée est la décomposition graphique de l'hydrogramme de crue en écoulement rapide, écoulement retardé et écoulement de base.

La méthode de décomposition de l'hydrogramme, a été étudiée par différentes approches et n'a pas abouti aux objectifs des hydrologues, et a souvent donné des explications contradictoires.

La mesure directe des différents écoulements de l'hydrogramme qui nécessite un déploiement de matériel expérimental souvent lourd et coûteux n'a fait l'objet que d'essais dans certains bassins expérimentaux, et reste donc spécifiquement dans un cadre de recherche (Linsley *et al.*, 1982).

Le développement des études sur les systèmes complexes et particulièrement les milieux naturels, à l'image des bassins versants, associé au développement de l'informatique, ont provoqué ces dernières années une prolifération de modèles en hydrologie.

Selon Moles (Moles, 1995), la modélisation est la troisième attitude de recherche scientifique après l'observation et l'expérimentation.

Un modèle est donc considéré par son concepteur comme un outil de représentation de l'ensemble des flux dans l'hydrosystème au niveau du bassin.

La mise en œuvre du modèle mathématique, a été motivée, par le degré de

dangerosité du phénomène naturel en question et de ses aléas, et par la volonté de compréhension, et surtout du souhait d'anticipation des problèmes causés par ce phénomène.

I.2.1. Définition d'un modèle mathématique

Un modèle est défini comme étant une représentation mathématique d'un phénomène physique simple ou complexe, qui permet en faisant appel à des entités ou à des processus élémentaires, reconstituer ce phénomène. Le choix des processus considérés comme élémentaires, et la combinaison qui en est faite, dépendent, d'une part des connaissances acquises antérieurement, et d'autre part, par des idées du modélisateur quant à l'origine des phénomènes (Delattre, 1979).

Un modèle mathématique est constitué:

D'un ensemble de variables généralement mesurables choisies pour représenter l'objet étudié;

Et d'un ensemble de relations mathématiques entre ces variables (de forçage), choisies pour représenter son fonctionnement. Ces relations font intervenir des paramètres, qui permettent de calculer à partir des variations spatio-temporelles imposées à certaines de ces variables, les variations correspondantes des autres variables.

Cependant, l'utilité d'un modèle est à juger en fonction de l'utilité de l'objectif poursuivi, le seul critère de validité est son adéquation à la réalité physique.

Le développement de la modélisation est motivé par un grand champ d'application, dont l'hydrologue voit le modèle comme un outil d'amélioration de connaissances, et dont les types d'application est résumé en trois critères:

Le domaine d'application: un modèle peut être exploité dans l'aménagement intégré des bassins versants, des barrages, l'aménagement et l'entretien des cours d'eau, le contrôle et la réduction des rejets, dans l'assainissement etc...

L'échelle de temps: la gestion des ressources en eau, requiert des solutions à des échelles très diverses. Les processus hydrologiques couvrent une large gamme spatiale et temporelle, cette même échelle peut varier selon le type de processus (Higy, 2000). Par exemple, lors d'une transformation d'une pluie en débit, plusieurs processus hydrologiques peuvent intervenir à des échelles temporelles différentes. L'écoulement rapide de crue se déroulera à une échelle horaire, tandis que le ruissellement qui en résulte, va se provoquer à une échelle temporelle plus importante de l'ordre du jour.

L'échelle spatiale: la grande hétérogénéité spatiale des bassins, explique en partie la variabilité spatiale des débits, et l'introduction de modèles ponctuels s'avère utile pour réduire cette complexité, ainsi l'exploitation d'informations dans des endroits non jaugés du bassin reste la préoccupation principale du modélisateur.

I.2.2. Pourquoi des modèles pluie-débit ?

Selon l'utilisateur, l'usage de la modélisation a des fonctions variées. Le modèle est destiné à aider à répondre à des questions, il est conçu en fonction de celles-ci, et ne peut être utilisé que dans ce champ particulier, son objectif principal est la réduction de la complexité naturelle. Les modèles pluie-débit servent en premier lieu à représenter dans le temps le cycle de l'eau, et permettent de transposer l'information climatique (figure I-1) en information sur les écoulements (Colombani et Roche, v. Roche, 1988).

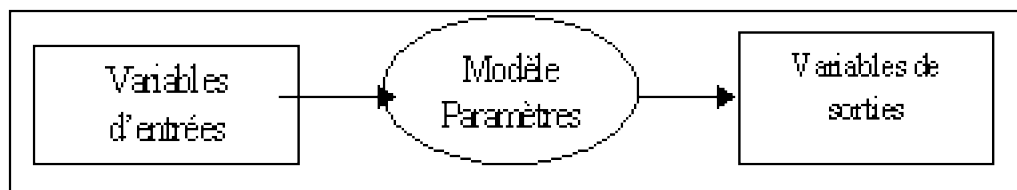


Figure I-1 Représentation simplifiée d'un modèle pluie-débit

Cependant, force en nous est de constater que plusieurs modèles pluie-débit ont été conçus, et dont la majorité n'ont pas trouvé une application pratique en hydrologie.

On peut donc, conclure que la transposition d'un modèle, se mesure par les degrés de robustesse de ses résultats, et de sa représentativité des processus à modéliser.

On peut résumer les fonctions des modèles pluie-débit en:

Extrapolation des débits

Le calcul d'un débit par le modèle est la tâche la plus courante pour ce dernier, dans ce cas, l'objectif reste la simulation et la génération synthétique des séries hydrométriques.

Cette estimation par l'extrapolation de débits peut cependant avoir des objectifs précis tels que :

Le comblement des lacunes des séries hydrométriques tronquées.

L'extension des séries courtes, ou des séries de certaines stations hors service.

Cette simulation quantitative trouve ses intérêts dans le dimensionnement des réseaux d'assainissement et de certains ouvrages hydrauliques.

L'estimation (directe et/ou indirecte) de certains paramètres hydrologiques et hydrogéologiques. Certains modèles trouvent leur utilité dans la déduction des propriétés statistiques des flux de pollution par l'intermédiaire des modèles pluie-débit, et des modèles pluie-débit-MES (Matières en suspension).

En combinant, une approche stochastique (modèle de génération de pluie), et des modèles pluie-débit et de matières en suspension, on peut obtenir une distribution de fréquence de la grandeur dérivée (flux MES) à l'exutoire d'un bassin.

A partir de certaines variables hydrologiques (pluie- ETP), le calage d'un modèle pluie-débit pour une période donnée, permet de calculer des débits, mais aussi de simuler les évolutions des niveaux aquifères d'une nappe en un point du bassin.

Prévision des débits

Connaissant le temps de réponse des débits à l'exutoire, un modèle pluie-débit peut être exploité dans la prévision des débits à court terme (quelques heures), et d'anticiper

l'évolution des débits vers l'exutoire.

Cette prévision 'naïve', permet de relier les débits à un instant t , en fonction des pluies instantanées et des débits observés au pas de temps précédent.

En utilisant des outils de prévision des pluies (tels que les radars), et à partir d'un calage d'un modèle pluie-débit, l'évolution des débits pourrait être étudiée pour une période plus longue que la prévision 'naïve'.

En France, certains modèles pluie-débit sont couramment utilisés dans l'annonce des crues et l'évaluation des risques d'inondations par Météo-France.

Gestion des ressources en eau

La gestion des ressources en eau est un problème délicat, du fait de la variabilité de ces ressources et de l'interaction des opérateurs.

La modélisation pluie-débit peut intervenir dans cette gestion à trois niveaux :

D'abord, par l'évaluation des ressources en eau, et de la valorisation de tous les processus participants aux différents écoulements, c'est le cas par exemple de certains bassins non jaugés, dont les méthodes empiriques aboutissent souvent à des résultats aberrants.

Un modèle peut caractériser l'impact des différents aménagements hydro-agricoles, et de prévoir certains changements des conditions météorologiques, ainsi que de délimiter les zones inondables.

Un modèle pluie-débit qui représente les différents flux du système dynamique au niveau d'un barrage, pourra être exploité dans la gestion des ressources de ce barrage en fonction de la demande des divers opérateurs intervenant sur le bassin. Il ne s'agit pas d'un système d'aide à la décision, mais plutôt d'un outil destiné à tester différents scénarios de développement.

Modèle comme outil de recherche

La connaissance de certains mécanismes intervenant dans le cycle de l'eau, est très limitée, dans ce cas la modélisation est motivée dans un but de compréhension de certains phénomènes, qui peuvent aboutir à une meilleure maîtrise de ces mécanismes, tels que la partition entre infiltration et ruissellement, l'écoulement sub-superficiel.

Les modèles, sont dans ce cas exploités dans la recherche, par l'introduction de nouveaux concepts, et de nouveaux éléments pour étudier et comprendre l'interaction de tous les processus intervenant dans l'écoulement. L'établissement de lois d'érosion et de transport solide lors de la formation d'écoulements peut constituer un objectif du modèle.

I.2.3. Propriétés requises des modèles

La conception des modèles en hydrologie, se base sur des théories scientifiques, ou à partir de certaines expériences réalisées sur le terrain. De ce fait, plusieurs modèles ont été élaborés, dans un contexte théorique ou pratique. On ne peut donc comparer ces divers modèles, sans définir à priori certaines valeurs pertinentes, qui définissent une supériorité remarquable, et une validité suffisante pour un tel modèle.

L'établissement d'une assise de comparaison entre modèles, est défini par certaines qualités qu'on résume en :

Cohérence rationnelle : c'est la première étape de la modélisation, la rationalité scientifique est une exigence de toute recherche scientifique.

Précision du modèle : elle concerne le fonctionnement de ce dernier après ajustement aux données expérimentales. Klemes (1982), stipule pour qu'un modèle puisse être considéré comme un bon modèle, non parce qu'il fonctionne bien, mais surtout parce qu'il fonctionne bien pour de bonnes raisons.

Simplicité : qualité essentielle de tout modèle, la simplicité concerne l'économie de moyens utilisés, c'est à dire un nombre réduit de variables et de paramètres. Elle concerne de toute évidence la réduction de toute complexité de la structure interne du modèle.

Robustesse : c'est la capacité du modèle à garder une certaine précision, même dans les conditions extrêmes, cette qualité reflète le pouvoir de reproduire les différents cycles hydrologiques, et de simuler certaines valeurs caractéristiques.

Falsifiabilité : concerne la possibilité d'introduire des hypothèses dans un modèle, en le modifiant, et en l'adaptant à de nouvelles exigences, sans altérer sa cohérence hydrologique.

Accessibilité et fiabilité : l'acquisition des données calculées en temps réel, est parfois une nécessité pour l'étude d'un phénomène hydrologique. Dans ce cas, l'accessibilité des données mais aussi leur fiabilité, sont jugées comme critères de comparaison entre divers modèles.

Pouvoir de prévision : objectif principal du modélisateur. Le pouvoir de prévision est lié à l'extension du domaine de validité d'un modèle. Le modèle apparaît d'autant mieux justifié, que son champ d'applicabilité s'avère à posteriori plus large.

I.2.4. Des modèles en fonction des caractéristiques hydrologiques des bassins

La transformation de la pluie en débit à l'exutoire par un modèle, est généralement conditionnée par certaines caractéristiques hydro-climatiques, qui gouvernent les différents mécanismes du processus pluie-débit. De ce fait, une des caractéristiques mises en causes par les hydrologues concerne l'universalité d'un modèle.

Il est bien évident, qu'un modèle pluie-débit conçu dans un pays à régime méditerranéen, aura du mal à simuler certains phénomènes climatiques d'un pays à climat nival, et dont les pics d'écoulement sont observés au moins de printemps durant la fonte des neiges. De même qu'un modèle conçu dans les pays nordiques, sera confronté aux difficultés de simulation des formes d'étiage des pays semi-arides.

Cette différenciation, est due non seulement, aux relations fonctionnelles du modèle, mais surtout au type de variables et de paramètres à introduire dans le modèle. Il n'existe donc pas de modèles universels, chaque modèle est adapté à certains types de climats

ou de bassins versants.

I.3. Classification des modèles pluie-débit

Depuis la première conception de modèles pluie-débit, plusieurs modèles sont apparus, et vu les différents objectifs des modélisateurs, les modèles se sont diversifiés, et la classification des modèles peut se baser sur les critères suivants (Higy, 2000):

Système à modéliser: concerne l'unité fonctionnelle fondamentale étudiée, tel que le grand bassin versant, le sous-bassin, le fleuve, les barrages, l'interaction bassin topographique-nappe, etc.

C'est la première étape d'une conception d'un modèle, elle permet de définir les variables et les descripteurs pertinents pour décrire le système à modéliser.

Degrés de causalité : implique la description des relations fonctionnelles des modèles.

Discrétisation spatio-temporelle : aborde l'extrême imbrication des aspects physiques et hydrologiques, et en leur grande variabilité spatiale et temporelle.

Une autre typologie est proposée par (Roche, 1988, Musy et Higy 1998, v. Higy, 2000), elle se fonde sur les différences entre:

Modèle distribué - global

Cette différence est importante sur le plan pratique, car certains modèles abordent la représentation de certains mécanismes dans leur diversité spatiale. Cette modélisation est motivée par le développement et l'exploitation de nouveaux outils tels que la télédétection et les systèmes d'informations géographiques. Dans ce cas, le modèle permet d'étudier l'impact des changements éco-physiologiques (déforestation), et de prévoir certains phénomènes hydrologiques (tels que les débits), dans un endroit bien précis du bassin. Par opposition à ce type de modèles, les modèles (globaux) qui considèrent le bassin versant comme une entité mathématique, ont été plus exploités, de par leur conception simple, et leur parcimonie, en matière de variables et de paramètres à introduire dans le modèle.

Modèle déterministe – stochastique

Les hydrologues ont souvent classé les différents modèles selon leurs degrés de déterminisme. Un modèle est dit déterministe, si à partir de moyens d'observation, on obtient une adéquation correcte entre le modèle et les observations. Cette description permet de caractériser les modèles selon la nature de leur variables et paramètres et/ou de leurs relations.

Modèle empirique - 'boîte noire' – conceptuel – à base physique

La nature des relations mathématiques entre les entrées du système et les sorties, entretient souvent une confusion, des lors que la définition de chaque classe de modèle, n'est toujours pas assez bien définie, et la frontière entre chaque modèle est loin d'être

nette.

Modèle continu – événementiel

Les processus hydrologiques peuvent intervenir à des échelles temporelles différentes. Lors d'un épisode pluvieux, l'estimation de certaines crues nécessaire au dimensionnement d'ouvrages, se déroule durant les averses. Dans ce cas, il est jugé parfois, assez suffisant l'application d'un modèle durant le temps de l'événement. C'est le cas de certains modèles pluie-débit au pas de temps horaire.

Nous pouvons cependant dire que les modèles pluie-débit événementiels, calculent les débits de ruissellement plutôt que les débits tout entiers

Modèle statique ou dynamique

Un modèle est dit statique, si les états du système sont des variables intemporelles, et dit dynamique, si ce sont des processus indexés par le temps.

La littérature hydrologique contient cependant une immense catégorie de modèles, et la classification dépend pour beaucoup de la conception du modélisateur à l'égard de ce modèle, plutôt que de la base ou de la structure du modèle lui même.

En vue d'une description cognitive des modèles et d'une classification qu'on juge meilleure, nous proposons la classification suivante:

I.3.1. Les modèles empiriques

Dans ces modèles, la transformation pluie-débit se fait non par l'application des lois physiques à un milieu et ses conditions aux limites, mais par l'examen des enregistrements des pluies et des débits (Nash et Sutcliffe, 1970).

L'introduction des équations déterministes ou stochastiques dégagées de toute complexité dans ces modèles, est motivée par la conception de modèles aussi simples qu'ils soient.

C'est en 1851, que Thomas Mulvany, présenta ses travaux, dans lequel le débit maximal est estimé par une loi empirique mettant en relation le débit maximal observé avec l'intensité de la pluie (Singh Whooliser, 2002, Kokkonen, 2003). Ces modèles sont basés sur des informations d'ordre général, que l'on possède sur un grand nombre de bassins versants présentant des caractéristiques similaires au bassin étudié.

La méthode empirique de l'estimation des débits vient d'être fondée, et à partir de ces travaux plusieurs équations empiriques ont vu le jour pour estimer certains débits caractéristiques d'un bassin.

Bien plus tard, certains hydrologues, ont essayé de transformer la relation pluie-débit par des modèles régressifs, ou à partir de l'hydrogramme synthétique (Larrieu, 1957), mais cette technique a été rapidement délaissée par son caractère très limité et peu cohérent.

On peut dire que les méthodes empiriques (ou rationnelles) d'estimation pluie-débit sont considérées comme une branche de l'hydrologie descriptive, plutôt que du domaine de l'hydrologie analytique.

I.3.2. Les modèles 'boîte noire'

Vu la complexité du système pluie-débit, certains hydrologues ont pu modéliser la relation pluie-débit par une formulation mathématique plus ou moins complexe, généralement linéaire. Dans ce cas les bassins versants sont considérés comme des 'boîtes noire' qui essaient de représenter théoriquement le prototype hydrologique (Figure I-2).

On peut cependant distinguer certains modèles 'boîte noire', par le fait, que la représentation de certains mécanismes tels que le ruissellement, est modélisée à partir d'un traitement matriciel. Ces modèles ne conviennent pas à une analyse du fonctionnement interne du bassin versant, et sont donc de peu d'utilité de recherche à cette échelle.

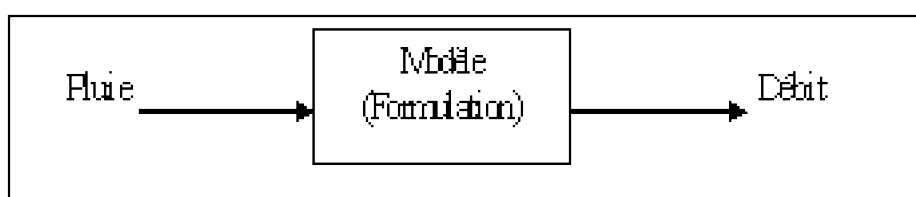


Figure I-2. Structure des modèles 'boîte noire'

I.3.3. Les modèles conceptuels

La transformation de la pluie en débit est une composante très complexe du cycle hydrologique qui engage différents facteurs à une échelle spatiale et temporelle différente.

C'est dans cette approche, que les modèles conceptuels essaient de représenter le cycle hydrologique par une conceptualisation des différents mécanismes, en introduisant des hypothèses jugées capables de traduire la physique de la relation pluie-débit et reliées dans leur ordre conceptuel logique (Sorooshian et Gupta, 1983). Ces modèles considèrent le bassin versant comme un assemblage de réservoirs d'humidité interconnectés, et sont censés représenter plusieurs niveaux de stockage suivant une dimension verticale. La figure I-3, présente l'architecture du premier modèle conceptuel, le modèle de Stanford (Crawford et Linsley, 1966).

En hydrologie, ces modèles sont les plus utilisés, cependant, même s'ils peuvent s'interpréter conceptuellement, les paramètres utilisés sont généralement sans grande signification physique, ni corrélables à des caractéristiques du bassin.

I.3.4. Les modèles physiquement fondés

Sont développés comme une alternative aux modèles conceptuels et globaux, à des fins de recherche parallèlement aux progrès de l'informatique.

D'après leurs concepteurs, ils sont censés rendre compte des problèmes pour lesquels les modèles conceptuels sont a priori jugés inadéquats, tels que la prévision des effets des changements naturels ou anthropiques des caractéristiques du bassin versant;

présentation de catégories de modèles, n'est malheureusement pas exempte d'ambiguïté.

La littérature hydrologique, offre une multitude et une diversité de modèles, dont les fonctions et les paramètres varient sensiblement d'un modèle à l'autre.

Les modèles empiriques se caractérisent par une formulation mathématique dépourvue de toute réalité physique, et la frontière entre ces modèles et les modèles 'boîte noire', n'est pas aussi claire, que certains le laissent supposer.

Un modèle basé sur une régression linéaire entre la variable d'entrée (pluie), et la variable de sortie (débit) peut être considéré comme 'boîte noire' du fait, de son interprétation du bassin comme un système, qui génère des débits à partir d'une série d'impulsions.

Un modèle est décrit comme conceptuel, lorsque le cycle de l'eau est conceptualisé par un agencement de réservoirs, simulant une fonction bien définie du cycle hydrologique (dans certains cas, un modèle conceptuel peut contenir des équations de vidange des réservoirs empiriques).

I.4. Paramétrisation des modèles pluie-débit

Nous avons évoqué le fonctionnement interne des modèles, par des relations entre les variables d'entrées (observées), et celles calculées par des ajustements, par des coefficients ou par des paramètres caractérisant le modèle, de ce fait l'estimation de ces paramètres est une étape essentielle de la modélisation.

Les paramètres sont définis comme des coefficients des modèles hydrologiques, d'une équation ou d'une relation qu'on peut ajuster pour appliquer un modèle à des cas particuliers.

I.4.1. Estimation des paramètres : calage du modèle

Suivant la nature du modèle, les paramètres sont soit, calculés directement sur le terrain, soit estimés par calibration (calage) durant une certaine période plus ou moins longue, en optimisant (manuellement ou automatiquement), l'erreur d'ajustement des variables mesurées à leur variables observées. Le but de ce calage, est de rapprocher le plus possible, le comportement du modèle de celui du bassin modélisé.

En pratique, les modèles à base physique, ne requièrent qu'une estimation directe des paramètres, puisque ces derniers sont mesurables *in Situ*. Ces modèles, qui après découpage du bassin en plusieurs unités spatiales de calcul, présentent le problème de la variabilité des paramètres, et cela même, pour des unités considérées comme homogènes. Ce problème est amplifié d'autant plus que cette variabilité atteint les données de forçage du modèle (pluie). Des travaux récents, ont été menés en vue de résoudre ce problème, par l'apport de la télédétection, et des systèmes d'informations géographiques dans l'estimation directe de certains paramètres (Liu et Todini, 2002).

Pour les modèles globaux, on utilise la paramétrisation inverse, l'estimation des paramètres est obtenue par le biais de données observées, d'où son nom estimation inverse. Dans ce cas, les paramètres sont ajustés après minimisation de l'erreur entre les valeurs observées et celles calculées par le modèle. Cette méthode nécessite une réponse à toute une série de questions aussi pertinentes, telles que la nature des erreurs, la méthode d'estimation, ou encore celle du choix d'une fonction objectif à optimiser.

Certains hydrologues ont mis en question l'existence d'un optimum, c'est à dire l'unicité d'un jeu de paramètres (Edijatno, 1991).

Dans le cas du calage manuel, l'utilisateur cherche à déterminer un jeu de paramètres par tâtonnement qui lui permet d'obtenir un optimum, ce dernier est réalisé lorsque l'erreur est minimisée entre les données simulées et observées, l'inconvénient du calage manuel est qu'il est fastidieux, et nécessite un temps relativement long.

Dans le cas du calage automatique, les données simulées sont ajustées aux données observées, au moyen d'un programme informatique, cet algorithme est plus avantageux que le calage manuel, mais présente l'inconvénient d'être influencé par certaines contraintes, qu'on résume en:

- Interdépendance des paramètres (auto-corrélation) : le changement de la valeur d'un paramètre peut être compensé par la modification d'autres paramètres, ce qui entraîne des solutions équivalentes en terme de valeur du critère d'optimisation;
- La faible sensibilité de la fonction objectif;
- La présence d'optima locaux, ou également la présence de plusieurs zones de convergences.

Certains auteurs (Perrin, 2000, Duan *et al.*, 1992) s'accordent à dire que la qualité de l'estimation des paramètres d'un modèle dépend notamment de la puissance et de la robustesse de l'algorithme utilisé.

Les chercheurs ont relativement bien compris la propagation des erreurs dans le cas d'un modèle à plus de dix paramètres, le nombre élevé de ces paramètres a souvent engendré un calage d'un modèle pour un optimum secondaire. Cette surparamétrisation de certains modèles par rapport à la quantité de données disponibles pour la calibration, peut conduire à une forte indétermination, les rendant non identifiables (Ambroise, 1991).

Les difficultés d'optimiser les paramètres des modèles hydrologiques cachent des problèmes liés à la définition, l'existence et l'identification d'un jeu optimal de paramètres pour le modèle, ceci provenant principalement des défauts internes du modèle et secondairement d'erreurs de données (Perrin, 2000).

Des analyses de sensibilité, ont montré que trois à quatre paramètres étaient en fait suffisants pour calibrer correctement des modèles conceptuels globaux (Sorooshian, 1991).

Un autre problème aussi pertinent reste le nombre de données à caler, à ce propos il n'existe pas de règles ou d'études communes pour tous les bassins. Durant longtemps, la calibration a été effectuée sur les deux premiers tiers de la série historique, le troisième tiers de données est utilisé pour la validation du modèle. Cette méthodologie est perçue

comme une assise de coutume ou d'une expérience, plutôt que d'une analyse scientifique. Néanmoins, les modèles pluie-débit, exigent un nombre d'années d'observations plus ou moins long, pour qu'un jeu de paramètres soit établi.

Le choix de la période de calage est crucial, plus est large la gamme de fonctionnement dont elle est représentative, et plus sont nombreuses les combinaisons de processus et facteurs qu'elle permet d'activer dans le modèle.

De ce fait, la durée de calage, peut différer d'un modèle à l'autre, et des conditions hydrologiques à étudier.

I.4.2. Validation des modèles

La validation (ou vérification) du modèle est une étape essentielle de la modélisation pluie-débit, non seulement pour la validation la structure du modèle, mais aussi pour le choix des paramètres.

Cette période permet de vérifier la robustesse du calage en terme d'indépendance de jeu de paramètres vis à vis des données d'entrées et donc l'aptitude du modèle à fournir des simulations (ou des prévisions) dans diverses situations hydrologiques (Quesney, 1999).

Durant plusieurs années, les hydrologues ne se sont intéressés qu'aux résultats du modèle obtenus lors du calage, et donc la validation de ces modèles a été faite sur des bases erronées. Cependant, il est important de distinguer le calage de la validation, la première opération assurant l'adéquation des paramètres alors que la seconde évalue l'adéquation du modèle tout entier (Fortin 1989, v. Rakem, 1999).

Ainsi, le bon fonctionnement d'un tel modèle, doit être pertinemment dépendant des résultats obtenus durant la phase de validation.

Afin de mener à bien une modélisation hydrologique, plusieurs auteurs se sont posés la question de la stratégie de la validation des modèles. Un des pionniers de cette réflexion est celle de Klemes (1986) qui propose un Schéma hiérarchique de validation en quatre étapes, qui permet de tester la capacité du modèle à prédire les effets des changements climatiques, des changements d'occupation des sols ou d'autres conditions non stationnaires. Ce schéma s'applique essentiellement lorsque la seule variable de validation est le débit à l'exutoire du bassin :

Split sample test : cette approche est la plus utilisée, elle consiste à découper la série modélisée en deux segments, le premier servant à la calibration (calage) du modèle, le second est utilisé pour la validation.

Proxy basin test : aucune calibration directe n'est permise. Son objectif est d'étudier la transposabilité de la modélisation d'un bassin à un autre ayant les mêmes caractéristiques.

Differential split sample test : c'est un test qui permet d'étudier l'adaptabilité d'un modèle, en fonction du changement des conditions climatiques. Le modèle est calibré pour une période sèche, puis validé pour la période humide, ou vice versa.

Proxy basin Differential split sample test : ce test est le plus délicat, la validation du modèle, est confirmée si les caractéristiques hydrologiques et climatiques d'un bassin sont remarquablement transposables par le modèle.

Blind test : diffère des tests précédents par le fait que les données observées sont inaccessibles, c'est une simulation qui concerne généralement les bassins non jaugés

I.4.3. Erreurs des modèles

Comme pour tout modèle hydrologique, l'extrapolation des résultats basée sur la relation pluie-débit est soumise à une certaine incertitude dont les principales sources sont les erreurs dans la structure du modèle, les erreurs de mesure et/ou de calcul des données d'entrée du modèle, les erreurs de calage et les erreurs d'extrapolation. La procédure d'estimation des paramètres (estimation inverse) fait appel à une fonction objectif que l'on nomme aussi fonction d'erreur, qui est fixée en fonction de la taille de l'échantillon, Cependant il est à noter, que l'incertitude de la modélisation augmente avec les valeurs extrêmes (telles que les crues).

I.5. Synthèse bibliographique

Dans cette deuxième partie du chapitre, nous présentons les différents modèles pluie-débit utilisés à différents pas de temps, à travers une chronologie de recherche effectuée par différents hydrologues, ainsi que les récents travaux inspirés de ceux existants dans la littérature hydrologique. Nous avons résumé les différents modèles pluie-débit en trois catégories, et cela en raison de l'immense diversité des modèles pluie-débit.

I.5.1 Les modèles ' boîte noire '

- L'hydrogramme unitaire

La théorie de l'hydrogramme unitaire a été élaborée par L.K. Sherman (1932) (Roche, 1971, Singh, 1995). Elle stipule que le bassin versant est modélisé comme un système soumis à une entrée $E(t)$ (pluie) qui produit l'écoulement et délivrant une sortie $S(t)$ (débit). L'application de la théorie de l'hydrogramme unitaire, nécessite de résoudre le problème de déconvolution, et qui consiste justement à déterminer une fonction de transfert à partir d'un historique assez représentatif des entrées et des sorties.

La théorie de l'hydrogramme unitaire reste très limitée aux bassins de faible surface, pour des grandes étendues, cette théorie est impossible à appliquer pour plusieurs raisons, en particulier les pluies et les caractéristiques pédologiques ne sont pas homogènes.

- Le modèle DPFT (Différence Première de la Fonction de Transfert)

La méthode DPFT a été développée à l'EDF, et présentée pour la première fois par Duband (1978). Cette méthode est une extension de la méthode classique de l'hydrogramme unitaire, et par conséquent décompose le processus global pluie-débit en modèles distincts par leurs structures, un modèle de production qui transforme la pluie brute en pluie efficace ruisselée, et une fonction de transfert permettant le passage de la pluie efficace en débit de ruissellement à l'exutoire du bassin. (Chambaz, 1990, Dechemi et Chambaz, 1994).

Deux principes de base s'imposent à cette méthode:

- Le premier est que les débits sont exprimés en différences premières sur un pas de temps reliant le présent et le futur;
- Le deuxième est que les pluies efficaces sont considérées comme inconnues du problème.
- Les modèles ARMAX (Auto-Regressive Moving Average with exogenous inputs)

Les modèles ARMAX dérivent des modèles ARMA (Auto-Regressive Moving Average) initialement développés par Box et Jenkins (1976), ont été utilisés dès leur première conception pour la simulation des débits.

Les modèles ARMAX, ont été utilisés dans le contexte pluie-débit, à différents pas de temps, journalier (Weeks et Boughton, 1987), et horaire (Rodriguez, 1989) dans le cadre de la prévision des crues.

La forme générale de ces modèles est donnée par l'équation (I-1) (Young et Wallis 1985):

$$y(t) = -\sum_{i=1}^p a_i y(t-i) + \sum_{j=1}^r b_j x(t-j) + \sum_{k=1}^q c_k e(t-k) + e(t) \quad (I-1)$$

Où $y(t)$ représente le débit simulé, p , r , q sont les ordres du modèle, a , b , c sont les paramètres du modèle, $x(t-j)$ est une variable explicative (pluie), $e(t)$ est le bruit blanc, t est l'indice de temps.

Cette approche est capable de modéliser mathématiquement la relation pluie-débit sans prendre en considération les processus physiques (Hsu *et al.*, 1995).

- Les réseaux de neurones artificiels

Relevant du domaine de l'intelligence artificielle, les réseaux de neurones sont une métaphore des structures cérébrales, des assemblages de constituants élémentaires appelés neurones fortement connectés construits par couches, qui réalisent chacun une fonction généralement non-linéaire, pour donnée une sortie unique à partir d'un mécanisme d'apprentissage effectué pendant le calage.

Les réseaux de neurones n'ont été appliqués qu'après des décennies de recherches, et sont pourvus actuellement de fondements mathématiques très sophistiqués ce qui explique leurs utilisations dans de nombreux domaines scientifiques.

L'étude de la relation pluie-débit par les réseaux de neurones a été traitée par plusieurs hydrologues (French *et al.* 1992, Hsu *et al.*, 1995, Shamseldin 1996, Dimopoulos *et al.*, 1996, Minns et Hall 1996, Coulibaly *et al.*, 1999, Anctil *et al.*, 2004).

· La logique floue

La logique floue introduite par Zadeh (1965), est un langage naturel issu de la théorie des ensembles, programmée afin de pallier aux insuffisances de la logique classique. Celle-ci n'utilisant que deux états logiques Vrai et Faux, en langage machine cela donnera les deux valeurs 0 et 1.

A partir des règles floues, le processus à modéliser est divisé en un ensemble de processus plus simples (sous ensembles flous), et en fonction des degrés d'appartenance (degrés de vérité), on engendre des surfaces d'inférences, les valeurs sont obtenues par la projection du centre de gravité de ces surfaces.

La logique floue, par son pouvoir d'extrapolation non-linéaire a été exploitée dans la modélisation pluie-débit (Ozelkan et Duckstein,2001, Dechemi *et al.*, 2003, Benkaci et Dechemi, 2004), et s'est vue une utilisation renommée en hydrologie.

I.5.2. Les modèles conceptuels (à réservoirs)

En 1934, Zoch développa des équations reliant un modèle de transformation de pluie en débit, sous l'hypothèse de la variation du débit en fonction du niveau d'humidité du sol. Les travaux de Zoch, furent probablement, les premières études de représentation de la relation pluie-débit par un assemblage de réservoirs linéaires. Plus tard, Nash (1957), développa un modèle où la pluie est transformée selon une cascade de réservoirs linéaires, en introduisant des paramètres propres au fonctionnement à chaque réservoir. Ses motivations étaient de développer un programme pratique basé sur les paramètres d'hydrogramme unitaire, et les caractéristiques d'un bassin versant, introduisant ainsi le concept de réservoirs linéaires dans la réponse d'un bassin. Cette étude a ouvert la voie à une nouvelle catégorie de modèles, que sont les modèles pluie-débit conceptuels ou modèles à réservoirs.

L'approche des modèles à réservoirs est de modéliser le système hydrologique comme un réservoir, (ou un jeu de réservoirs), qui joue le rôle de tampons emmagasinant la pluie pour la redistribuer par la suite.

C'est au début des années soixante que Linsley et Crawford (université de Stanford) ont utilisé un modèle de simulation pluie-débit, mais cette fois-ci fonctionnant à un pas de temps (continu) journalier. Les auteurs ont publié en 1966 les caractéristiques de ce modèle appelé, modèle de Stanford.

Le modèle comporte six réservoirs à 34 paramètres, et tient compte des stocks neigeux.

Le modèle de Stanford exprime différents mécanismes de tous les phénomènes intervenants dans les processus de vidange, d'évaporation, d'infiltration, et de ruissellement. Les fonctions de transfert sont assez élaborées, il représente ainsi un

système complexe d'équations avec plusieurs paramètres.

A partir de ce modèle, la recherche en hydrologie a vu une émergence de toute une classe de modèles à réservoirs.

C'est le cas par exemple du modèle CREC (Cormary *et al.*, 1967, v. Roche, 1971), à deux réservoirs seulement, ce modèle est caractérisé par le fait que la partie ruisselée dans l'eau se fait par un opérateur

$$(II)$$

, et non par un réservoir. Suivant que la pluie est supérieure ou inférieure à cet opérateur, on aura soit ruissellement, soit infiltration.

La modélisation pluie-débit à réservoirs, se diversifie, en introduisant certains mécanismes qui font l'originalité de certains modèles, c'est le cas par exemple du modèle Siméro, (1972), qui à partir de certains paramètres mesurés in Situ, tels que l'humidité du sol, ou les niveaux acquièrès, permet de simuler les débits, ou le modèle Knapp (1975), qui simule le fonctionnement du bassin et d'une nappe aquifère qui lui est associé (Roche, 1971, Benkaci, 2001).

La variabilité spatiale des paramètres hydrologiques devient une préoccupation des modélisateurs, et la modélisation distribuée devient une propriété de certains modèles. C'est en 1972 que les chercheurs de l'INRS-eau du Québec (Girard *et al.*, 1972), ont développé le modèle Cequeau pour simuler les bassins canadiens.

C'est un modèle distribué, assez élaboré, il permet de représenter toute modification artificielle de l'écoulement dans les rivières et les barrages, et de voir l'évolution des écoulements dans des endroits non jaugées des bassins. Des tests sur des bassins semi-arides (Tunisie) ont été effectués par Ayadi et Bargaoui (1998).

Vers la fin des années soixante dix, Beven et Kirkby (1979) ont élaboré un nouveau modèle dans lequel, sont introduits des paramètres physiques tels que la conductivité hydraulique, la capacité au champ, l'indice de similarité hydrologique du bassin nommé indice topographique, déterminant ainsi l'organisation des états hydriques du sol dans la génération des écoulements (zones contributives variables). Ce modèle appelé TopModel (Beven et Kirkby, 1979), a généré un concept nouveau dans la modélisation à réservoirs, puisque la base physique des relations fonctionnelles du modèle devient une nouvelle assise de comparaison dans la modélisation pluie-débit.

L'approche distribuée a suscité de nombreuses études pour mieux appréhender les processus à l'échelle de la maille choisie (Modèle IHDM, Beven *et al.*, 1987, modèle Swatch, Morel-Seytout, Al hasoun, 1989).

Avec la vulgarisation des ordinateurs dans les années quatre vingt, les modèles à réservoirs se sont diversifiés, la modélisation a séduit de nombreux hydrologues, et la volonté de concevoir des modèles spécifiques était l'ambition de plusieurs Chercheurs.

C'est l'exemple des modèles Amande et Martine (Mazenc *et al.*, 1984), qui exigent parfois des paramètres complexes, tels que la capacité de rétention rapide ou le nombre de jours de pluie par an.

Ces modèles présentent l'avantage d'étudier l'influence des paramètres

physiographiques et morphologiques d'un bassin, sur l'estimation de certains débits caractéristiques.

La particularité de chaque modèle, est déduite du type de climat et de ses caractéristiques. Ainsi le modèle de Tank de Sugawara (1984), a été conçu pour modéliser le régime humide au Japon, ce modèle permet de simuler différents écoulements au niveau d'un bassin. Le dessèchement et la rehumectation du sol se fait selon une succession de dessèchement et de rehumectation des réservoirs superficiels puis souterrains (Singh, 1995). Le modèle Xianjiang (Zhao *et al.*, 1980) a été conçu spécialement pour modéliser les différentes caractéristiques du régime climatique continental de Chine, le modèle Modhydrolog (Chiew et Mc Mahon, 1994) pour affronter les spécificités hydrologiques de l'Australie, ainsi que les apports neigeux et le climat nival ont été représentés par le modèle HBV en Suède (Bergstrom et Forsman, 1973), ou encore le modèle Satt-I en Finlande (Vakkilainen et Karvonen, 1982).

La répartition spatiale des écoulements et sa variabilité, deviennent une préoccupation de certains hydrologues (Baudez *et al.*, 1999) dans l'élaboration des modèles journaliers et même horaires. Le modèle ARNO (Todini, 1996), représente une nouvelle gamme des modèles, ce dernier est de type semi-distribué, dans ce cas, le bassin est divisé en sous-bassins, qui coïncident avec les limites et formes topographiques et hydrographiques du bassin versant.

Connaissant la contribution de la pluie nette dans l'écoulement, certains hydrologues ont pu introduire des fonctions de transfert, permettant de simuler le décalage de cette pluie et la propagation des débits jusqu'à l'exutoire. C'est le cas du modèle GR3j (Edijatno et Michel, 1989, Edijatno, 1991) développé au CEMAGREF et issu des modèles GR (Génie Rural), ce modèle à deux réservoirs et trois paramètres seulement intègre un hydrogramme unitaire capable de simuler le transfert de la pluie nette dans un bassin.

Néanmoins, tous ces modèles, avaient l'inconvénient de ne modéliser que des bassins versants, considérés comme un milieu clos. Certaines conditions hydrologiques tels que les sous-bassins, ou bassins intermittents étaient mal simulées, dès lors la modélisation des échanges entre sous-bassins devient une propriété essentielle au bon fonctionnement des modèles. C'est l'exemple du modèle GR4J (Makhlouf, 1994, Nascimento, 1995,) à quatre paramètres, qui découle du modèle GR3J, ce modèle intègre un opérateur, qui modélise l'échange inter-bassins.

Le modèle GR3J a été spécifiquement étudié dans le cas de la prévision des crues; Yang (1993) a mis au point une méthode de prévision des crues en utilisant le modèle GR3J à un pas de temps horaire, et avec une mise à jour des paramètres.

Dans le bassin de la Somme (France) (Amraoui *et al.*, 2003), ont permis d'étudier la prévision des hautes eaux de l'hiver 2001-2002 par deux modèles pluie-débit .

Les modèles GR, ont démontré leur grande capacité de simulation, avec deux ou trois paramètres seulement, ces modèles étaient parmi les plus robustes, et cela contrairement aux autres modèles, dont le nombre de paramètres à estimer ou à caler était généralement supérieur à dix. Le concept de parcimonie des modèles pluie-débit était la priorité et l'objectif de certains chercheurs. Jakeman *et al.*, (1990) ont pu concevoir un modèle basé sur l'hydrogramme unitaire, ce modèle comporte six paramètres et deux

modules linéaires et non-linéaires pour simuler la pluie nette et déterminer les débits à l'exutoire.

Les modèles à réservoirs présentent un avantage de falsifiabilité important, et peuvent être adaptés aux différentes exigences climatiques et reformulations mathématiques.

C'est le cas de certains modèles, tel que le modèle GR3J (Edijatno *et al.*, 1999) ou le modèle IHACRES (Littlewood *et al.*, 1997, Croke, Jakeman, 2004).

La transformation de la pluie en débit, par les modèles conceptuels, a été exploitée généralement au pas de temps journalier, et dans certaines études au pas de temps horaire. Au pas de temps mensuel, les recherches ont été souvent limitées, à cause de la forte non-linéarité du processus pluie-débit à cette échelle.

L'échelle de temps devient donc une nouvelle assise de conception des modèles, ainsi on a pu voir au début des années quatre vingt certains modèles mensuels tels que le modèle ABCD, (Thomas, 1981), le modèle de Arnell (1992), le modèle de Vandewiele (Vandewiele et Xu, 1992), ou encore le modèle GR2M (Edijatno et Makhlouf, 1994) à deux paramètres. De même qu'un modèle fonctionnant au pas de temps annuel (Bouabdallah, 1997), développé à partir de GR2M, a été testé sur plus quatre vingt bassins français.

I.5.3. Les modèles physiquement fondés

C'est vers la fin des années soixante, que Freeze et Harlan (1969), ont imaginé un nouveau concept de modèles, basé sur des équations issues des théories physiques, et qui prend en considération les variabilités spatiales, des différents processus intervenant au niveau d'un bassin.

Deux ans plus tard, Freeze (1971) met en application ce modèle, qui avait pour originalité de calculer l'écoulement saturé et insaturé dans les trois directions de l'espace. Mais ce modèle avait l'inconvénient d'exiger un nombre de paramètres élevé.

Avec le développement de l'informatique, les modèles fondés physiquement réapparaissent vers le milieu des années quatre vingt. C'est ainsi qu'Abott *et al.*, (1986) développèrent un modèle entièrement basé sur des équations physiques, et qui reproduit finement le comportement d'un bassin réel; le modèle du système hydrologique européen (SHE).

Dans ce modèle, l'interception est calculée à partir du modèle de Rutter (1971), l'évapotranspiration est estimée par la formule de Penman, la zone non saturée est représentée analytiquement par l'équation de Darcy Richard, qui nécessite la connaissance des propriétés hydrodynamiques du sol.

L'écoulement en nappe libre est déduit des équations d'hydraulique telle que l'équation de saint venant. Une version améliorée du modèle (Shetran) a été reformulée par (Ewen *et al.*, 1998), et testée par Higy (2000) pour les bassins versants suisses.

Le modèle MYDHAS (Moussa *et al.*, 2002) représente un autre exemple d'un modèle physiquement fondé. S'appuyant sur des équations d'écoulement en milieu non saturé et

saturé (telles que loi de Philip, ou la loi de Darcy), ce modèle est basé sur une segmentation du bassin en unités hydrologiques.

Les modèles à base physique diffèrent essentiellement par les hypothèses de fonctionnement hydrologique, par la procédure de segmentation du bassin versant, par le mode de représentation des liens entre les versants et le réseau hydrographique, par le mode de représentation des processus hydrologiques, et par les stratégies de paramétrisation, de calage et de validation (Chahinian, 2004).

I.5.4. Nouveaux outils de la modélisation pluie-débit : Apport de la télédétection

A travers des décennies, la modélisation pluie-débit a connu un effort considérable, quant à la conception et à l'utilisation des modèles. Les variables et paramètres hydrologiques ont toujours fait l'objet d'une grande importance dans leur estimation du fait qu'ils conditionnent d'une façon primordiale le comportement hydrologique du système étudié. De ce fait l'erreur d'estimation de ces variables ou paramètres pourraient avoir des conséquences négatives sur le fonctionnement, et donc sur la validation d'un tel modèle.

Les bassins versants montrent une grande variabilité spatiale et temporelle des paramètres tels que les propriétés physiques et hydrodynamiques des sols, la végétation, le taux d'humidité et l'occupation du sol, etc. Ainsi, la non-disponibilité de ces données, constitue le premier obstacle pour le modélisateur.

Pour résoudre le problème d'estimation de ces variables en des lieux non-mesurés, on fait appel à la distribution d'informations spatiales et/ou temporelles, par des procédés analytiques (déterministes) ou probabilistes (stochastiques).

Ces méthodes d'interpolation plus ou moins fiables, présentent l'inconvénient d'être exigeantes en nombre de données, ceci impliquerait la mise en place d'un système de mesure très fastidieux.

Les méthodes basées sur la télédétection et les systèmes d'informations géographiques peuvent constituer une solution à ce type de problème, en ce sens qu'elles permettent de caractériser instantanément des zones spatialement étendues. Dans ce cas, l'objectif est d'étudier l'apport d'une prise en compte des paramètres hydrologiques sur la fiabilité des phénomènes étudiés (telles que les crues) simulés par un modèle. Connaissant le rôle de l'indice hydrique du bassin dans la génération de l'écoulement, des études ont été menées (Chkir, 1994, Loumagne *et al.*, 1996) en vue d'intégrer dans des modèles hydrologiques (tels que les modèles GR) un ou plusieurs indicateurs de l'état hydrique du sol contrôlables par des mesures externes (télédétection), afin d'améliorer le fonctionnement d'un tel modèle hydrologique. Quesney (1999) a testé l'assimilation de l'état hydrique du sol dans le modèle GRhum, sur le bassin de l'Orgeval (France).

I.6. Conclusion

Le bassin versant est un système complexe ouvert, et son fonctionnement en réponse aux forçages atmosphériques dépend des caractéristiques hydrologiques et pédologiques propres à ce bassin. Un grand effort de recherche a été mené sur les différents processus intervenants dans la génération des débits, et des progrès fondamentaux sont accomplis dans la compréhension des cheminements de l'eau au niveau d'un bassin versant.

La modélisation pluie-débit est considérée dans ce cas comme un des moyens les mieux adaptés pour décrire de façon assez sommaire une réalité complexe du fonctionnement d'un bassin versant.

Vu les différences hydroclimatiques à travers le monde, les modèles pluie-débit se sont diversifiés en plusieurs catégories, pour prédire les débits de réponse d'un bassin aux événements pluvieux à partir d'une série d'observations pluies-débits.

Dans cette étude, à travers un exposé des différents modèles, nous avons présenté la diversité des modèles mathématiques utiles à l'hydrologue pour le traitement des problèmes liés à la simulation de la transformation de la pluie en débit.

Les remarques et critiques préliminaires, ainsi justifient à priori le choix des objectifs et méthodes mises en œuvre dans cette recherche, à savoir l'utilisation des modèles conceptuels et des nouvelles techniques d'extrapolation non-linéaires représentées par les réseaux de neurones artificiels et la logique floue.

Chapitre II Etude hydrologique et statistique

II.1. Introduction

L'objectif de notre recherche est l'étude de la modélisation pluie-débit, aux pas de temps mensuel et journalier. Il s'agit donc d'étudier le fonctionnement d'un bassin versant, en réponse à des forçages météorologiques et hydrologiques.

Pour cela, il est nécessaire de disposer d'un échantillon de données hydrométéorologiques à savoir les pluies, les évapotranspirations et les débits de la station à modéliser.

Dans le cadre de cette étude, la première tâche est d'acquérir des séries de données des trois variables météorologiques suffisamment longues, pour mieux représenter les différents processus hydrologiques intervenants dans la modélisation pluie-débit. La difficulté réside dans le fait que ces données doivent non seulement être fiables, mais qu'elles doivent coïncider avec la même période d'observation.

Pour notre part, après avoir dressé une carte thématique des réseaux pluviométriques et hydrométriques des bassins algériens, nous nous sommes aperçus que plusieurs stations pluviométriques comportent des lacunes et parfois sont

relativement très courtes et donc inutilisables. Ce problème est amplifié pour les données hydrométriques, vu que généralement, la densité des stations hydrométriques est lâche, et les données hydrométriques sont parfois entachées d'erreurs, impossibles à corriger, sans avoir recours aux mesures limnimétriques.

Notre travail de modélisation, est confronté au souci d'obtenir de l'information, et donc tributaire de la disponibilité d'un ensemble de données nécessaire à notre recherche.

Les données pluviométriques et hydrométriques disponibles ont été recueillies au niveau de l'Agence Nationale des Ressources Hydrauliques (ANRH) et de l'Agence Nationale des Barrages et des transferts (ANBT), pour les mesures de l'évapotranspiration potentielle (ETP), nous disposons des données établies par l'Office National Météorologique (ONM), et des données de l'ANBT.

II.1.1. Utilisation des données hydrométéorologiques

Les modèles pluie-débit intègrent généralement trois types de variables hydrométéorologiques, à savoir les précipitations, l'évapotranspiration, et le débit observé d'une station hydrométrique.

a) Les précipitations

Représentent les variables de forçage, qui doivent conditionner en grande partie le cycle hydrologique au niveau d'un bassin. Les précipitations sont généralement constituées de quantités de pluie cumulées durant le pas de temps du modèle exprimées en millimètre.

Pour le calage du modèle, il est pratique d'estimer la pluie en prenant la moyenne tombée sur le bassin, dans ce cas on doit prendre en considération toutes les stations se trouvant à l'intérieur de ce bassin. Cependant il est très rare de trouver deux stations pluviométriques ayant fonctionné durant la même période d'observation.

Les diverses stations pluviométriques contiennent plusieurs lacunes dont il est difficile de combler.

Néanmoins, par manque de données, on peut utiliser une seule station dans le cas de bassins de petite taille, on considère donc l'étude spécifique au sous-bassin versant.

Pour les bassins de grande taille (Superficie > 10000 km²), il est impératif de disposer des chroniques des pluies assez représentatives du bassin versant.

b- L'évapotranspiration (ETP)

Représente une donnée climatique importante pour caractériser l'état du sol, elle agit comme la quantité d'eau susceptible d'être perdue dans l'atmosphère sous forme d'évaporation à partir de la surface et des végétaux. Si les précipitations ne peuvent être qu'une donnée représentante du bassin, on peut utiliser les estimations de l'ETP d'une station située en dehors du bassin versant.

Plusieurs méthodes peuvent estimer l'évapotranspiration potentielle, soit par des évaporomètres tels que le bac class A, ou le bac Colorado, soit par des formules

empiriques telles que la méthode de Penman ou celle de Thornthwaite.

A l'échelle mensuelle, l'ETP peut être déduite de la seule variable de température (Dechemi et al., 1997). La méthode de Thornthwaite permet ainsi à partir de la température moyenne du mois, de calculer l'évapotranspiration mensuelle d'une façon relativement précise.

Dans le cas d'un modèle fonctionnant à un pas de temps journalier, les données décennales estimées par la méthode de Penman, peuvent être désagrégées en données journalières (Edijatno, 1991).

En comparant quatre modèles conceptuels sur plus de 300 bassins versants européens et autres, les tests réalisés par Oudin (2004), ont confirmé qu'une donnée précise (datée) d'évapotranspiration n'apporte pas d'amélioration systématique des performances des modèles pluie-débit, et que des données d'ETP à grande échelle fournissent des entrées aussi acceptables pour les modèles.

En effet, l'auteur souligne qu'une information extrêmement simple à obtenir (données interannuelles) peut suffire à estimer des ETP conduisant à des simulations de débits aussi bonnes qu'une information datée difficile à obtenir. De la même manière, une seule station climatique située sur ou à proximité du bassin versant étudié est suffisante.

c- Les débits

Le débit à l'exutoire représente la réponse d'un bassin à un ensemble complexe de plusieurs processus hydrométéorologiques et hydrologiques.

La mesure du débit présente néanmoins une certaine difficulté, puisque ces débits varient sensiblement en fonction du temps, et cela même à l'échelle journalière. Cependant on peut noter que les débits peuvent être estimés par différentes méthodes (moulinet, jaugeage par dilution, par flotteur, par limnigraphe). Chacune de ces méthodes sont relativement fiables mais présentent certains inconvénients.

A titre d'exemple, la courbe de tarage qui permet d'extrapoler les débits à partir d'une relation hauteur-débit, dépend pour beaucoup de la forme du lit et du coefficient de Strickler.

On sous-entend que le tarage est très délicat pour les hautes et basses eaux. En effet, en période d'étiage, le niveau du cours d'eau est tellement bas, qu'on observe une diminution de la précision relative des débits.

En hautes eaux, le problème essentiel concerne l'extrapolation des débits, en effet on observe au limnigraphe des hauteurs supérieures à celles correspondant aux jaugeages des grands débits (Laborde, 2003).

Pour le calage du modèle, les données de pluies et d'évapotranspirations sont exprimées en lame d'eau (mm/jour ou mm/mois), les débits observés sont exprimés généralement en m^3/s .

II.2. Aire d'étude et présentation des bassins

MODELISATION PLUIE-DEBIT MENSUELLE ET JOURNALIERE PAR LES MODELES CONCEPTUELS ET LES SYSTEMES NEURO-FLOUS (APPLICATION AUX BASSINS ALGERIENS)

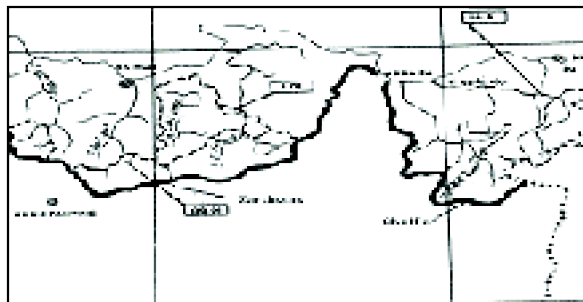
Le réseau hydrologique algérien est divisé en neuf (9) bassins versants tributaires de la Méditerranée tels que le bassin du Cheliff, de la Soumam, ou de la Tafna et huit (8) bassins endoréiques et sahariens.

Les superficies diffèrent d'un bassin à l'autre, le bassin du Cheliff est le plus vaste des bassins versants du Nord, sa superficie atteint 44000 km².

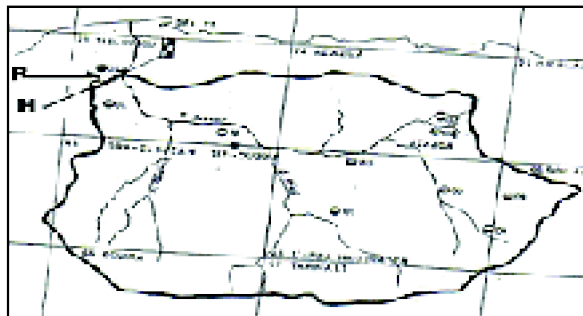
Le régime d'écoulement des grands fleuves algériens est généralement permanent, cependant les affluents secondaires et tertiaires présentent un écoulement faible et temporaire.

Les bassins versants étudiés présentent un climat méditerranéen avec un comportement saisonnier, où on observe une succession de saisons sèches et pluvieuses très contrastées à l'échelle annuelle, et des régimes d'écoulement très variables à l'échelle inter-annuelle.

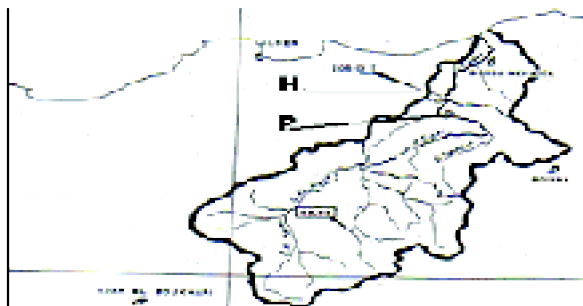
Nous avons les stations des bassins versants de la Cheffia, de l'Isser, du Sebbaou, de la Tafna et de Zardezas (Figure II-1).



(a)

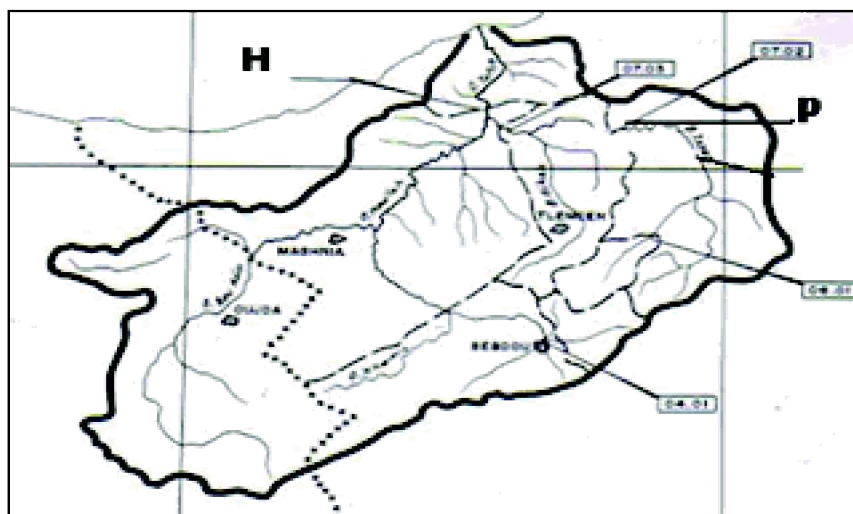


(b)



(c)

Figure II-1 Carte des bassins versants étudiés



(d)

Cheffia et Zardezas (a), Sebbaou (b), Isser (c), Tafna (d)

H (station hydrométrique) et P (station pluviométrique) étudiées

- Le bassin versant de la cheffia est situé dans le Nord-Est algérien et codé dans le côtier constantinois (03), il est alimenté par l'oued Bouna Moussa qui est formé par l'intersection des oueds de Kamouna et de Soudan, et se ramifie en oueds Bou Alalah et El Kebir. Il est limité au Nord par la Méditerranée, à l'Est par la Tunisie, et au Sud-Est par le bassin de Rhumel Kebir. Il est équipé d'une seule station hydrométrique. Le barrage de la Cheffia permet d'alimenter en eau potable et d'irrigation les habitants de cette région.
- Le climat du bassin est méditerranéen, l'hiver est froid où la température atteint parfois 0°C, le bassin est relativement humide, la pluviométrie varie en moyenne de 800 à 900 mm, et dépasse parfois 1000 mm. La période pluvieuse s'étale de Novembre à Mai, l'été est relativement chaud et sec (ANRH (a), 1985).
- Le bassin de l'Isser codé (09) par l'ANRH, est situé dans le Nord de l'Algérie, sa superficie totale est de 4170 km² de Ain Boucif jusqu'à l'embouchure du bassin. Une partie du bassin est drainée par l'oued Mellah et l'oued Hammam, ces deux oueds forment l'Isser, qui se ramifie en plusieurs affluents. Le bassin de l'Isser est limité au Nord par la mer et le bassin du Sebbaou, au Sud par le bassin de Chott Hodna, à l'Est et au Sud est par le grand bassin du Cheliff. Le point culminant est Djebel Dira qui atteint 1810 m. Le climat du bassin est méditerranéen, froid et humide en hiver, les températures descendent jusqu'à 0 degrés, chaud et sec en été. La pluviométrie moyenne est de 800 mm par an.
- Le bassin de la Tafna codé (16), est situé à l'extrême Ouest du pays sa superficie

totale est de 7165 km², dont le tiers est situé au Maroc. Du point de vue relief, on peut distinguer trois zones, les hauts reliefs du mont de Tlemcen, la zone des reliefs moyens et des vallons, et les grandes plaines d'Oujda et de Maghnia. Le climat est de type méditerranéen à semi-aride, où la pluviométrie est inférieure à 500 mm dans les grandes plaines. Le régime hydrographique se présente sous forme de deux branches principales, l'Isser à l'Est et la Tafna à l'Ouest (ANRH (b), 1985).

- Le bassin du Sebbaou (02 20) est un bassin situé dans le Nord de l'Algérie, il couvre une superficie de 2500 km², il est limité à l'Est par le bassin de la Soumam, et à l'Ouest par le bassin de l'Isser. L'oued Sebbaou prend source dans le massif de Durdjura, et se ramifie en plusieurs affluents tels que l'oued Boubhir et oued Aissi. Le massif de Djurdura limite le bassin au Sud-Ouest, au Sud et Sud-Est.
- Le climat du Sebbaou est méditerranéen et relativement tempéré, le module pluviométrique varie de 700 mm à 1000 mm dépassant parfois 1500mm dans les massifs montagneux. L'hiver est froid avec des températures très basses dans les montagnes, on enregistre – 3°C durant les mois de Décembre-Janvier. L'été est chaud et parfois très sec, les mois les plus chauds sont Juillet et Août.
- Le bassin comprend plusieurs postes pluviométriques répartis dans plusieurs régions, mais la densité du réseau hydrométrique est faible, et certaines stations ont cessé de fonctionner telles que la station de Belloua (depuis 1983).
- Le bassin de l'oued Saf-Saf (codé 0309), en amont du barrage de Zardezas résulte de la conjonction de deux cours d'eau : l'oued Bouhadjeb et l'oued Khemkem, et est situé dans le côtier constantinois, d'une superficie de 345 km², il est limité par la Méditerranée au Nord, le bassin de Rhumel Kebir à l'Est et au Sud-Est, le bassin de la Soumam à l'Ouest. Le climat est de type méditerranéen, froid et relativement humide en hiver, et chaud en été. Le bassin présente une forte densité de drainage (3.8 km/km²), son relief est caractérisé par des terrains accidentés. En effet, les massifs montagneux couvrent plus de 50% du bassin, en particulier dans la partie Sud du bassin. Ces fortes pentes favorisent la torrencialité des cours d'eau et rendent les sols du bassin vulnérables à l'érosion (Rouabah, 1999).

II.2.1. Les données utilisées

II.2.1.1. Les stations pluviométriques

Nous avons choisit la station pluviométrique de Cheffia contrôlée par l'Agence Nationale des Barrages (ANB), la station ne présente pas de lacunes pour toute la période étudiée. Cette station fera l'objet de notre étude aux pas de temps mensuel et journalier.

Pour ce qui est de l'Isser, le bassin est équipé d'un grand nombre de postes pluviométriques, ainsi on dénombre trente six (36) postes répartis à travers le grand bassin de l'Isser.

En analysant ces stations, on remarque que la majorité d'entre elles sont incomplètes, présentent des lacunes, et ne peuvent servir d'étude dans un cadre de

modèle pluie-débit.

Pour nos besoins de modélisation au pas de temps mensuel, nous avons choisis deux stations pluviométriques pour la taille et la qualité des données, à savoir la station de Ouled Bouhaddada (090503) et la station de La traïlle (090413). La station de Ouled Bouhaddada comporte deux lacunes (Novembre et Décembre 1982).

Pour combler ces lacunes, il existe plusieurs méthodes (El jabi *et al.*, 1987):

Remplacer la valeur manquante par celle de la station la plus proche.

Remplacer la valeur manquante par la moyenne des stations voisines. Cette méthode est utilisée lorsque les valeurs moyennes annuelles de la station à compléter ne diffèrent pas plus de 10 % des valeurs moyennes annuelles des stations de référence.

Remplacer la valeur manquante par la moyenne inter-annuelle de la série considérée. Cependant dans le cas d'un modèle pluie-débit, cette méthode s'avère risquée, du fait qu'elle ne permet pas de caractériser le cumul pluviométrique au pas de temps considéré, le débit simulé par le modèle peut être fortement biaisé.

Les méthodes basées sur l'analyse de régressions et de corrélations:

Le calcul de régression se fait en déterminant les valeurs des paramètres de la relation entre la variable expliquée et la (ou les) variable(s) explicative(s). Cette régression peut être simple ou multiple.

Nous avons complété deux valeurs manquantes de la station de Ouled Bouhadada (Novembre-Décembre 1982) par une régression linéaire avec la station la plus proche qui est la station de Lakhdaria Gorges 090502, de 09/1976 à octobre 1982.

Le coefficient de régression est très significatif ($R^2 = 0,97$), l'équation de régression est $Y = 0.88 * X + 3.86$

Les valeurs estimées pour ces deux mois sont respectivement de 155 et 119 mm.

Nous avons la station pluviométrique de La Traïlle (090413) qui a fonctionné depuis 1972 à 1996 sans lacunes, qui sera étudiée avec la station hydrométrique du Pont de La Traïlle (090408).

Au pas de temps journalier, la station de Ouled Bouhaddada (090503) sera prise en considération pour une étude pluie-débit, on prendra un échantillon de 1990 à 1995.

Le bassin du Sebbaou fera l'objet d'une étude pluie-débit mensuelle, dans ce cas, nous avons utilisé la station hydrométrique de Baghliia (022002), la station pluviométrique la plus proche est la station de Baghliia village (090220), pour cela nous disposons d'un échantillon de 17 ans (1980/1996). Au pas de temps journalier, l'étude sera consacrée à une période courte : années (1993/1994).

Pour ce qui est du bassin de la Tafna, la station de Ben Sekrane (160702), qui est sans lacunes de 1980 à 1995, sera couplée avec la station hydrométrique de Remchi (160703).

II.2.1.2. Les stations hydrométriques

La Cheffia : la seule station située dans ce bassin est la station de Cheffia, contrôlée par

MODELISATION PLUIE-DEBIT MENSUELLE ET JOURNALIERE PAR LES MODELES CONCEPTUELS ET LES SYSTEMES NEURO-FLOUS (APPLICATION AUX BASSINS ALGERIENS)

l'ANB, les données hydrométriques ne présentent pas de lacunes.

L'Isser : le bassin versant de l'oued Isser présente une superficie totale de 4170 km². Le réseau hydrométrique implanté par l'ANRH, est équipé de sept stations de jaugeage : El Omaria (091205), Dechmia (090309), Beni Slimane (090305), Mezahim (030191), Aomar Gare (090416), Pont de la Traille (090408), et Lakhdaria (090501).

D'une manière générale, la densité du réseau est très lâche, par rapport au vaste étendu de la zone.

La Tafna : nous prendrons la station de Remchi (160703).

Zardezas : station contrôlée par l'ANB.

Pour nos besoins de modélisation nous avons choisis pour leur représentativité et leurs tailles, ainsi que leur correspondance avec les stations pluviométriques six stations réparties de façon homogène sur le territoire national : la Cheffia, Lakhdaria, La Traille, Sebbaou, Tafna et Zardezas.

Stations	Bassins	Coordonnées (Lambert)	Date de mise en service	Superficie du B.V
Cheffia (031501)	Cheffia	X=977,70 - Y=381,32	1978	570 km ²
Lakhdaria (090501)	Isser	X=575,10 - Y=368,75	1954/65	3615 km ²
La Traille (090408)	Isser	X=555,25 - Y=348,25	1968	2570 km ²
Baghnia (022001)	Sebbaou	X=602,70 - Y=390,05	1969	2390 km ²
Remchi (160703)	Tafna	X=122,90 - Y=208,80	1973	1935 km ²
Zardezas (030902)	Zardezas	X=878,75 - Y=370,67	1989	345 km ²

Tableau II-1 Stations hydrométriques étudiées

Les caractéristiques hydrologiques des bassins versants étudiés sont présentées dans le tableau II-2.

Tableau II-2 Caractéristiques hydrologiques des bassins versants

Stations	Cheffia (031501)	Lakhdaria (090501)	La Traille (090408)	Sebbaou (022001)	Tafna (160703)	Zardezas (030902)
Superficie (km ²)	570	3615	2570	2390	1935	345
Coefficient Kc	1.20	1.26	1.24	1.14	1.14	1.26
Altitude moyenne (m)	420	710	600	590	720	628
Indice de pente de Roche	0.11	0.11	0.12	0.20	0.12	0.17
Apport (Hm ³ /an)	134.1	198.7	154.5	643.1	34.7	51.4
Débit spécifique (l/s/km ²)	7.39	1.74	1.90	8.60	0.57	2.83

Au vu du tableau II-2, nous remarquons que les bassins présentent des

caractéristiques hydro-morphométriques assez contrastées, l'indice de pente de Roche le plus élevé est celui du bassin du Sebbaou, et cela pour les terrains accidentés présents dans ce bassin.

L'écoulement est très variable d'un bassin à l'autre, pour le bassin de la Tafna, l'apport est inférieur à $35\text{Hm}^3/\text{an}$, son débit spécifique est inférieur à 0.6 l/s/Km^2 , ce qui prouve que le bassin est sujet à l'aridité, contrairement aux bassins de la Cheffia et du Sebbaou qui paraissent très arrosés, avec un débit spécifique respectivement de 8.60 et de 7.39 l/s/km^2 . L'apport moyen annuel à la station de Lakhdaria est de 643 Hm^3 .

II.2.1.3. Les mesures de l'ETP

Les données de l'évapotranspiration sont indispensables pour mesurer le bilan hydrologique, et donc nécessaires à notre étude. Nous avons les données de la station de Dar El Beida, qui seront utilisées dans le cas du bassin du Sebbaou, Pour les stations de l'Isser, nous disposons d'un échantillon de données décennales de 1990 à 1989 recueillies au niveau de l'Office National de la Météorologie de la ville de Bouira, les valeurs de l'ETP sont estimées par la méthode de Penman.

Pour la période allant de 1972 à 1989, nous utiliserons les données des moyennes de la ville de Bouira (1989-90), vu que nous ne disposons pas d'autres mesures. Pour les modèles journaliers, les données décennales seront désagrégées en données quotidiennes.

Pour le bassin de la Cheffia, nous utiliserons les données de l'ETP journalières de la ville d'Annaba, calculées par la méthode de Penman. Pour le bassin de la Tafna, vu que nous n'avons pas de données d'ETP des stations météorologiques situées dans ce bassin, nous prendrons les mesures de la station de la ville de Cheliff. Les ETP journalières du bassin de Zerdezas calculées par le bac Colorado (coefficient $K=0.7$), seront prises en considération pour les modèles journaliers.

Les caractéristiques des stations étudiées en fonction de la période d'observation sont présentées dans les tableaux II-3 et II-4:

Stations pluviométriques	Période d'observations	Stations hydrométriques	Période d'observations
Cheffia (031501)	1978-1991	Cheffia	1978-1991
Isser (091503)	1990-1997	Isser (090501)	1990-1995
Zerdezas (030502)	1991-1996	Zerdezas (0305)	1991-1996
Sebbaou (022002)	1992	Sebbaou (022002)	1994

Tableau II-3 Périodes d'observations des données mensuelles

Stations pluviométriques	Période d'observations	Stations hydrométriques	Période d'observations
Cheffia (031301)	1978-1991	Cheffia	1978-1991
Isoer (090503)	1990-1997	Isoer (090501)	1990-1995
Zerdezas (030502)	1991-1996	Zerdezas (0305)	1991-1996
Bebbaou (022002)	1992	Bebbaou (022002)	1994

Tableau II-4 Périodes d'observations des données journalières

II.3. Critique et analyse statistique des données

Quelque soit la façon dont les données sont obtenues, il y aura toujours le risque de rencontrer des valeurs erronées. Ces erreurs peuvent être systématiques telles qu'un mauvais appareillage, ou accidentelles telles qu'une mauvaise lecture, ou une erreur lors d'un report.

Dans ce cas, toute étude hydrologique devra commencer par une critique des données disponibles de façon à éliminer si possible les données les plus suspectes.

Ces techniques ne permettent pas de déceler les erreurs; elles permettent seulement de mettre en évidence une valeur, ou des séries de valeur 'anormales', compte tenu des hypothèses que l'on a pu faire sur le comportement de la série des variables utilisées.

Les anomalies dans les séries de données ont été mises en évidence par ces techniques, c'est à l'hydrologue de décider si une donnée est 'bonne ou mauvaise'. Ce choix s'appuie sur des arguments statistiques plus ou moins objectifs, et sur la connaissance plus ou moins subjective que l'on a du phénomène.

Au début de notre travail, il a été question de résoudre le problème de la fiabilité des données, ce problème est d'autant plus pertinent lorsqu'il s'agit de comparer les deux chroniques essentielles à savoir les pluies et les débits.

Le travail d'analyse des données sera divisé en trois parties:

Compléter les données manquantes ou étendre certaines séries d'observation : cette procédure ne s'appliquera pas aux données hydrométriques. Il s'agit de combler certaines lacunes dans les séries de pluie mensuelles, à partir des stations pluviométriques situées dans le bassin étudié.

Vérification des tests d'hypothèse de base : il s'agit de vérifier la crédibilité des estimations en s'assurant qu'aucun phénomène «parasite» ou systématique ne soit intervenu ; dérive temporelle des données du à un aménagement hydro-agricole, interdépendance entre les événements, détection des valeurs suspectes.

Ajustement des données à une loi statistique. Il s'agit d'attribuer aux séries de données une évolution théorique qui nous permettra d'établir la fonction de distribution du processus étudié.

II.3.1. Traitements statistiques des données

C'est la première étape de toute étude, elle consiste à vérifier la présence de certaines valeurs douteuses ou particulières, 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 (Bermad, 2004).

Le test Grubbs et Beck permet de détecter les points singuliers de la série en calculant la statistique suivante :

$$X_n = \text{Exp}(\bar{x} + K_n * s) \quad (\text{II.1})$$

$$X_1 = \text{Exp}(\bar{x} - K_n * s) \quad (\text{II.2})$$

Où \bar{x} et s sont respectivement la moyenne et la variance des logarithmes naturels des éléments constituant l'échantillon.

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

Au niveau de risque de 10%, l'approximation polynomiale suivante a été proposée par Bobée (1991):

$$K_n = 3.62271 + 6.28446 * n^{1/4} - 2.49835 * n^{1/2} + 0.491436 * n^{3/4} - 0.037911 * n \quad (\text{II.3})$$

Où n représente la taille de l'échantillon.

L'application de ce test considère tous les points supérieurs à X_n et inférieurs à X_1 comme des points singuliers ou valeurs aberrantes.

Les résultats du test, confirment qu'il n'y a pas de valeurs singulières ou aberrantes, excepté certaines valeurs nulles des données de pluie, de toutes les stations. Pour notre part, nous n'avons pas rejeté ces valeurs, elles reflètent le caractère sec de certains jours ou mois.

II.3.1.2. Test d'indépendance

Les tests d'indépendance permettent de vérifier si les observations sont stochastiquement indépendantes. C'est à dire qu'elles ne dépendent pas de leur rang et que la moyenne de la répartition étudiée ne doit pas subir de déplacement de nature monotone ou cyclique. L'objectif du test est de ne pas introduire de multiples fois la même information.

Plusieurs tests permettent de vérifier l'indépendance des événements, nous pouvons citer le test d'Anderson, le calcul du coefficient d'auto-corrélation, ou encore le test de Wald-Wolfowitz.

· Test de Wald-Wolfowitz

Il est basé sur l'hypothèse qu'aucune observation ne devrait influencer la valeur suivante.

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

$$\bar{R} = \frac{\sum_{i=1}^{n-1} x_i x_{i+1} + x_1 x_n}{n-1} \quad (\text{II.4})$$

est distribuée sur une loi asymptotiquement normale de moyenne:

$$\bar{R} = \frac{S_1^2 - S_2}{(n-1)} \quad (\text{II.5})$$

Et de variance:

$$\text{Var}(\bar{R}) = \frac{(S_2^2 - S_4)}{(n-1)} + \left(\frac{S_1^4 - 4S_1^2 S_2 + 4S_1 S_3 + S_2^2 - 2S_4}{(n-1)(n-2)} \right) \bar{R}^2 \quad (\text{II.6})$$

Avec:

$$S_r = \sum_i^{i=n} x_i^r = M^r \quad (\text{II.7})$$

M représente le moment d'ordre de l'échantillon.

La valeur statistique

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

suit une loi normale standardisée $N(0,1)$, et peut être utilisée pour tester à un niveau de signification, l'hypothèse d'indépendance en comparant $|T|$ avec la valeur normale standard

$U_{\alpha/2}$ correspondant à la probabilité de dépassement $\alpha/2$.

Si $|T|$ est inférieure à

$$U_{\alpha/2}$$

l'hypothèse d'indépendance est vérifiée, celle-ci est rejetée dans le cas contraire (Bermad, 2004).

Les résultats du test appliqué aux données mensuelles sont présentés dans le tableau II-5.

Stations	Pluies		Débits	
	T	$U_{\alpha/2}$ ($\alpha=5\%$)	T	$U_{\alpha/2}$ ($\alpha=5\%$)
Cheffia	82.8	1.96	100.0	1.96
Lakhdaria (Isser)	97.7	1.96	70.0	1.96
Pont La Traille (Isser)	91.7	1.96	114.0	1.96
Baghlia Village (Sebbaou)	99.1	1.96	86.7	1.96
Ben Sekrane (Tafna)	81.3	1.96	82.1	1.96

Tableau II-5 Résultats du test de Wald-Wolfowitz appliqué aux données mensuelles

Les résultats du Wald-Wolfowitz présentés dans le tableau II.5, montrent que les données mensuelles ne sont pas indépendantes, et cela pour toutes les stations étudiées.

II.3.1.3. Tests d'homogénéité

L'hypothèse d'homogénéité exprime que les échantillons extraits de la même population sont identiques et par conséquent ont les mêmes fonctions de répartition (Aivazian, 1983).

L'objectif des tests d'homogénéité est de détecter la non-homogénéité des données que ce soit entre deux échantillons d'une même station, ou entre deux séries homologues de deux stations différentes.

Une série de données est dite homogène, si les propriétés statistiques de l'erreur de mesure affectant les données sont restées invariables au cours de la période d'observation.

Test de Wilcoxon-Mann-Whitney

Ce test utilise les numéros d'observations dans la série variationnelle obtenue 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 $n=n_1+n_2$ et $n_1 \geq n_2$.

Le test de Wilcoxon-Mann-Whitney considère les quantités suivantes :

$$\begin{cases} V = R - \frac{n_2(n_2 + 1)}{2} \\ W = n_1 n_2 - V \end{cases} \quad (II.8)$$

Avec:

R : Somme des rangs des éléments du premier échantillon de taille n_1 dans la série considérée.

V : Représente le nombre de fois où un élément du premier échantillon soit dans l'ordre un élément du deuxième échantillon.

W : Somme des nombres de dépassements des éléments du premier échantillon par ceux du second.

Lorsque n , n_1 et n_2 son suffisamment grands, V et W suivent une distribution normale :

De moyenne

$$\bar{V} = \bar{W} = \frac{n_1 n_2}{2}$$

Et d'écart type

$$s_V = s_W = \left(\frac{n_1 n_2}{12} (n_1 + n_2 + 1) \right)^{1/2}$$

Nous pouvons alors tester si les deux échantillons proviennent de la même population au niveau de signification α en comparant la grandeur statistique :

$$T = \frac{|V - \bar{V}|}{s_V} \quad (II.9)$$

Avec la variable normale centrée réduite U de probabilité de dépassement

$$\frac{\alpha}{2}$$

Si

$$T \leq U_{\alpha/2}$$

alors les deux échantillons proviennent de la même population, dans le cas contraire ($T \geq U_{\alpha/2}$), l'échantillon étudié est considéré comme non homogène.

Stations	Pluies		Débits	
	U	Seuil (%)	U	Seuil (%)
Cheffia	0.002	5	0.001	5
Isser	0.03	5	0.001	5
Zerdezaz	0.003	5	0.001	5

Tableau II-6 Résultats du test d'homogénéité appliqué aux données journalières

Stations	Pluies		Débits	
	U	Seuil (%)	U	Seuil (%)
Cheffia	0.007	5	0.033	5
Lakhdaria (Isser)	0.012	5	0.073	5
Pont La Traille (Isser)	0.012	5	0.01	5
Baghlia Village (Sebbaou)	0.001	5	0.008	5
Ben Sekrane (Tafna)	0.001	5	0.04	5

Tableau II-7 Résultats du test d'homogénéité appliqué aux données mensuelles

Les résultats du test de Wilcoxon-Mann-Whitney présentés dans les tableaux II-6, II-7, montrent que les observations des pluies et de débits des différentes stations sont homogènes, que ce soit au pas de temps mensuel ou au pas journalier.

II.3.1.4. Test de stationnarité

Il permet de s'assurer que les caractéristiques statistiques soient invariantes avec le temps, ce test est donc indispensable pour la crédibilité de l'estimation des paramètres des lois statistiques.

Le contrôle de cette hypothèse peut être fait en découpant la série de données en deux sous-échantillons 1 et 2, et en testant les moyennes et écart-types du sous-échantillon 2, en vérifiant qu'ils appartiennent aux intervalles de confiances du sous-échantillon 1.

Cependant il est à noter que la stationnarité des variables hydrométéorologiques se mesure sur de longues séries, généralement supérieures à trente (30) ans. En effet, on doit distinguer la stationnarité du changement (ou variation) climatique qui survient d'une période à l'autre. La stationnarité est une propriété qui ne peut être décelée que sur des décennies (Perrin, 2000).

II.4. Ajustement statistique des données

En hydrologie, il est impossible de connaître la population totale, nous disposons à priori des échantillons non exhaustifs, tirés de cette population.

A partir de cet échantillon nous nous proposons de choisir la forme mathématique de la loi de probabilité, et d'en calculer le mieux possible des paramètres numériques de cette loi par des ajustements jugés adéquats.

Cependant, il est indispensable de s'assurer de la bonne adéquation de la loi par :

Un ajustement graphique : permet de voir si la forme de la loi F choisie est bien celle de la distribution des données observées. Dans notre cas la méthode de l'intervalle de confiance sera utilisée.

Un test de dispersion des observations par rapport à la loi F choisie. Le plus souvent on utilise le test χ^2 de Pearson qui consiste à découper l'échantillon de n éléments, en K classes ($K > 4$) de façon à ce que chaque classe contienne théoriquement plus de 5 éléments.

Soit n_i le nombre d'éléments contenus dans la ième classe et $n'_i (>5)$ le nombre théorique qui devrait s'y trouver, on décrit le terme suivant :

$$\chi^2 = \sum_{i=1}^K \frac{(n_i - n'_i)^2}{n'_i} \quad (\text{II.10})$$

Cette variable χ^2 suit une loi de probabilités dite loi du χ^2 , à v degrés de liberté, avec:

$$v = K - p - 1$$

K = nombre de classes ;

p = nombre de paramètres estimés d'après l'échantillon.

II.4.1. Estimation des paramètres

La première tâche d'ajustement à une loi consiste à déterminer les paramètres de cette dernière. Deux méthodes sont couramment utilisées, la méthode du maximum de vraisemblance et la méthode des moments.

Pour notre part, nous nous sommes basés sur la méthode des moments, elle permet de vérifier que les premiers moments de la loi, et les moments empiriques des observations coïncident:

$$m(\bar{x}, \delta) = \int_{-\infty}^{+\infty} Q \frac{\partial F}{\partial Q} \partial Q \quad (\text{II.11})$$

$$m(\bar{x}, \delta)^2 = \int_{-\infty}^{+\infty} [Q - m(\bar{x}, \delta)]^2 \frac{\partial F}{\partial Q} \partial Q \quad (\text{II.12})$$

L'application de cette méthode consiste à estimer d'après l'échantillon autant de moments que la loi a de paramètres, et à évaluer ces estimations aux valeurs théoriques qui ne sont fonction que des paramètres à estimer. On obtient alors un système de n équations à n inconnues.

La loi normale (loi de Gauss)

La loi normale est une loi symétrique et plausible si le phénomène étudié peut être considéré comme une somme d'un grand nombre de petits phénomènes indépendants

(théorème central limite). Cette loi est définie sur $[-\infty, +\infty]$

La fonction de répartition de cette loi s'écrit:

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-\bar{x}}{\sigma}} e^{-\frac{u^2}{2}} du \quad (11.13)$$

Avec:

$$u = \frac{x - \bar{x}}{\sigma}$$

dépend de deux paramètres qui sont la moyenne

$$\bar{x}$$

(paramètre de position) et l'écart-type

$$\sigma$$

(paramètre de forme)

II.4.2. Ajustement des données annuelles à la loi normale

Les pluies annuelles des différentes stations s'ajustent de façon adéquate à la loi normale (figure II-2).

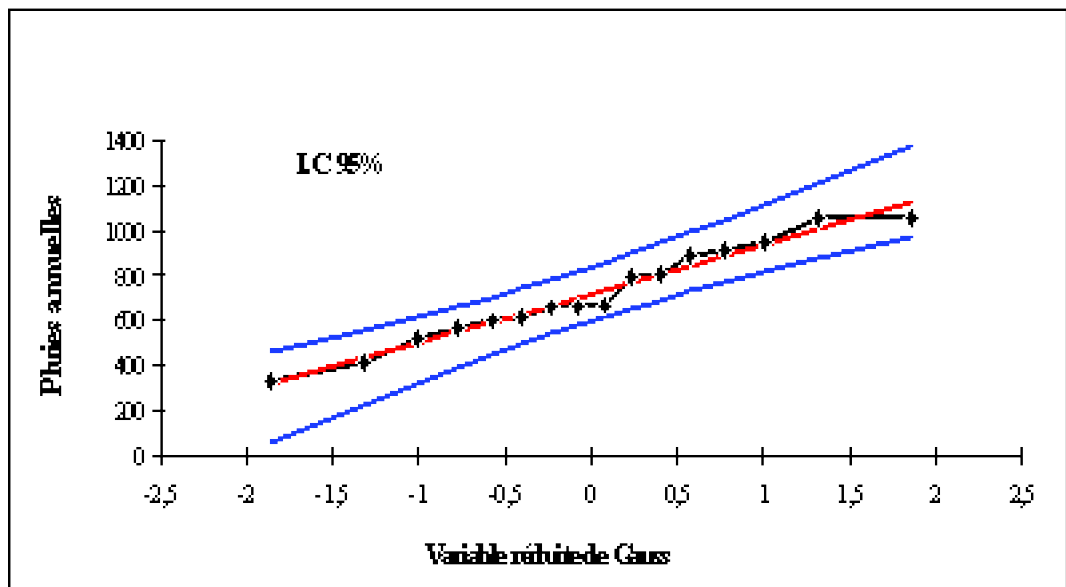


Figure II-2 Ajustement des pluies annuelles à une loi normale

(Station de Lakhdaria)

- Les débits annuels (ou apports annuels) des différentes stations ne s'ajustent pas à la loi normale, cette loi symétrique ne peut pas s'appliquer aux apports liquides annuels,

et cela vu les grands écarts entre les modules hydrométriques annuels. Dans ce cas, Nous pouvons tester la loi de Galton (loi log-normale), on remplace dans ce cas les valeurs des apports annuels (Q) par le logarithme de $Q - Q_0$ (Q_0 moyenne des apports liquides annuels).

La loi Log-normale (loi de Galton)

La fonction de répartition de cette loi est de la forme:

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x}{b}} e^{-\frac{u^2}{2}} du \quad (II 14)$$

$$\text{Avec: } u = \frac{\log(Q - Q_0) - a}{b} \quad (II 15)$$

Dans ce cas nous considérerons que c'est la variable $\log(Q - Q_0)$ qui suit une loi de Gauss à deux paramètres (Laborde et Mouhous, 1998) :

- a : Moyenne des $\log(Q - Q_0)$;
- b : Ecart-type des $\log(Q - Q_0)$.

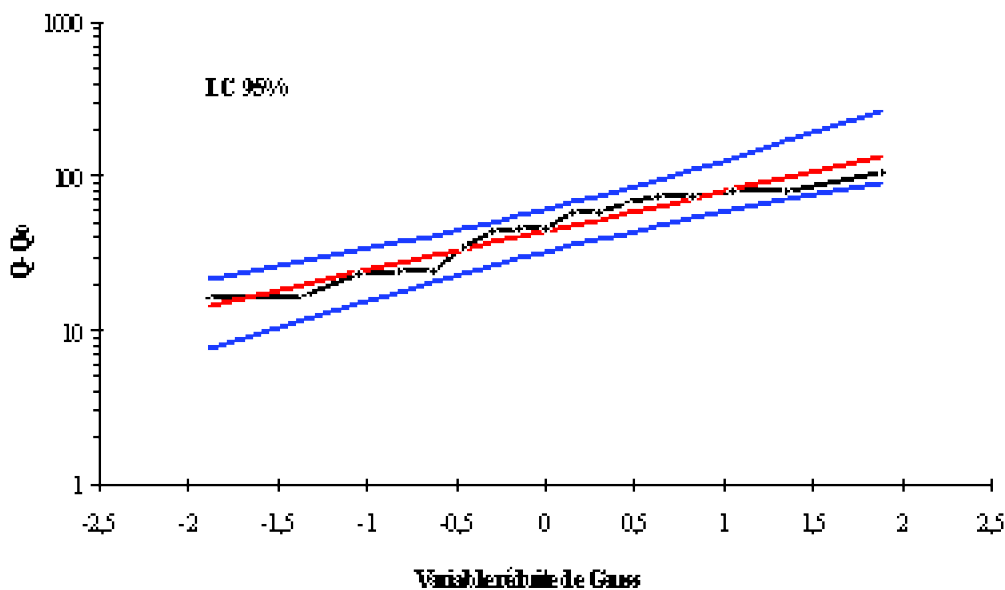


Figure II-3 Ajustement des apports annuels à une loi log-normale,

Q: apports liquides annuels (Station du Sebbaou)

II.4.3. Ajustement des données mensuelles à la loi de Pearson III

L'ajustement des données mensuelles pose problème de fait qu'il y a répétition de plusieurs valeurs (cas des valeurs nulles) qui lors du tri, ces observations peuvent prendre plusieurs fréquences, et donc le contrôle de l'ajustement devient impossible. Nous

proposons l'ajustement par la loi de Pearson.

Parmi les différentes lois de Pearson, la loi de type III ou loi gamma incomplète est la plus utilisée en hydrologie. Sous sa forme générale, la fonction de distribution de cette loi s'écrit :

$$F(x) = \frac{1}{\Gamma(\gamma)} \int_0^u u^{\gamma-1} e^{-u} du \quad \text{avec } u = \frac{x - x_0}{s} \quad (\text{II.16})$$

$\Gamma(\gamma)$: fonction gamma complète :

$$\Gamma(\gamma) = \int_0^{\infty} u^{\gamma-1} e^{-u} du \quad (\text{II.17})$$

L'intervalle de variation de x est $[x_0, +\infty[$, et les trois paramètres d'ajustement sont :

x_0 : paramètre de position (borne inférieure)

s : paramètre d'échelle (de même dimension que x)

γ : paramètre de forme (positif différent de zéro)

Si γ croît indéfiniment, la loi gamma incomplète tend vers une loi de Gauss (Laborde, 2003). La densité de probabilité a pour expression :

$$f(u) = \frac{1}{\Gamma(\gamma)} u^{\gamma-1} e^{-u} \quad (\text{II.18})$$

Selon les valeurs de γ les formes de distribution changent d'allure ; on peut distinguer différents cas:

- $0 < \gamma < 1$: La densité de probabilité décroît continuellement lorsque u augmente, loi dite en J. Cette loi n'a pas de mode réel.
- $\gamma = 1$: Loi exponentielle sans mode réel.
- $1 < \gamma < 2$: Loi modale à $u = \gamma - 1$, densité nulle à $u = 0$ avec tangente verticale.
- $\gamma = 2$: Idem mais : tangente égale à 1 en $u = 0$.
- $\gamma > 2$: Courbe en cloche.

La moyenne et l'écart-type s'expriment simplement à partir de x_0 , γ et s :

$$\bar{x} = x_0 + \gamma s \quad \text{et} \quad \delta = s \sqrt{\gamma} \quad (\text{II.19})$$

Figure II-4 Ajustement des débits mensuels à la loi Pearson III (Station du Sebbaou)

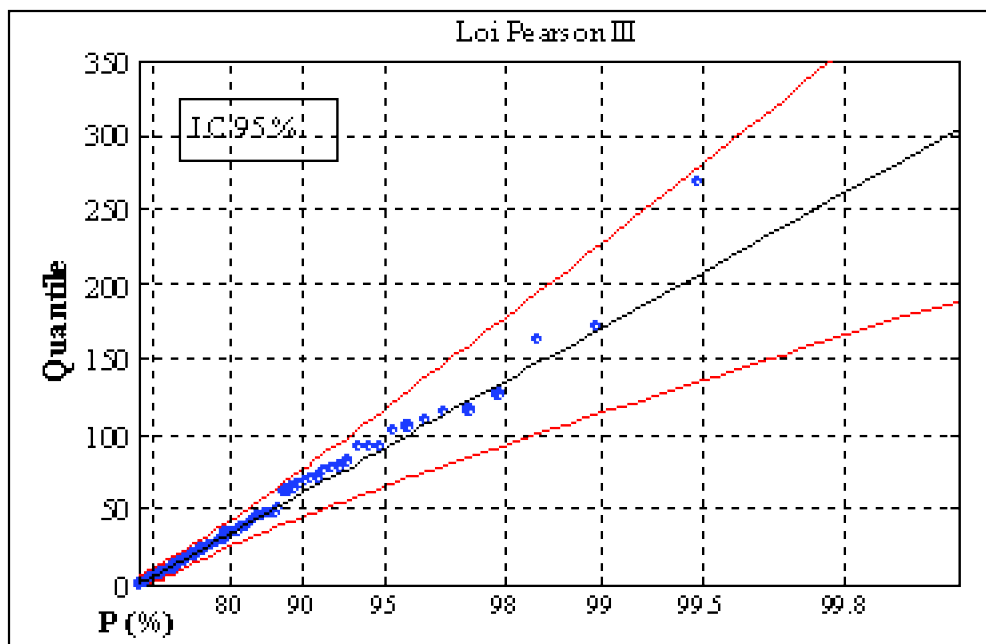


Figure II-4 Ajustement des débits mensuels à la loi Pearson III (Station du Sebbaou)

II.5. Critères de validation des modèles pluie-débit

L'une des tâches les plus délicates dans l'étude d'un modèle pluie-débit est l'estimation de ses performances, et sa robustesse de simuler les débits d'un bassin. Dans ce cas, la qualité de simulation nécessite la comparaison entre les données observées et celles simulées par le modèle.

La technique la plus utilisée est la méthode du double échantillon, qui consiste à caler le modèle sur une période donnée (période de calage ou de calibration), et de le tester sur le reste des observations (période de validation), pour juger son adaptabilité face à des situations dans lesquelles il est censé fonctionner. Les durées de ces deux périodes diffèrent d'un pas de temps à l'autre.

La validation consiste donc à estimer les performances des modèles, par des critères qui estiment la différence entre les données simulées par le modèle et celles observées.

On distingue deux types de critères:

Les critères graphiques: correspondent à un ajustement graphique des données simulées sur les données observées. Ces critères ne sont pas sans importance, du fait qu'ils permettent un jugement visuel, et donc d'apprécier la qualité du modèle à reproduire les différentes tendances climatiques, tels que la succession des cycles de dessèchement et de rehumectation par le modèle.

Les critères numériques: consistent à déterminer une fonction qui va minimiser l'erreur entre les données simulées par le modèle, et celles observées.

II.5.1. Les critères numériques

II.5.1.1. Le critère quadratique (E)

C'est le critère le plus utilisé, il consiste à minimiser la somme des carrés des erreurs entre les débits observés et les débits calculés, il permet d'évaluer l'erreur entre la variance et le biais :

$$E = \sum_{i=1}^n (Q_{i\text{ obs}} - Q_{i\text{ sim}})^2 \quad (\text{II.20})$$

Où:

Q_{obs} = Débit observé au pas de temps i ;

Q_{sim} = Débit simulé au pas de temps i ;

n = Taille de l'échantillon.

II.5.1.2. Le critère RMSE (Root Mean Squared Error)

Il est défini comme étant la racine carrée de l'erreur quadratique, l'introduction de la racine carrée à pour but d'homogénéiser les écarts entre les débits simulés et observés en particulier dans les cas extrêmes (débits faibles ou élevés), du fait que cette erreur est élevée au carré dans le cas du critère quadratique.

Le critère RMSE exprimé en m^3/s est défini comme suit:

$$RMSE = \sqrt{\sum_{i=1}^n (Q_{i\text{ obs}} - Q_{i\text{ sim}})^2} \quad (\text{II.21})$$

Les modèles sont optimisés si les valeurs des critères sont proches de zéro. Cependant les résultats sont fortement influencés par les valeurs importantes des écarts, correspondant souvent aux crues.

II.5.1.3. Le coefficient de détermination (R^2)

Il permet d'évaluer le degré d'association entre deux variables X et Y et de juger de la qualité de l'ajustement des points par la droite de régression, en expliquant la proportion de la variance de Y par les variations de X.

Il peut être estimé par l'équation suivante:

$$R^2 = \frac{\sum_{i=1}^n (Q_{obs} - Q_{sim})^2}{\sum_{i=1}^n (Q_{obs} - Q_{moy})^2} \quad (L.22)$$

Où Q_{moy} est la moyenne des débits observés. Pour une bonne simulation des débits le coefficient de détermination doit être égal à un.

II.5.1.4. Le critère de Nash

Nash et Sutcliffe (1970) ont proposé un critère, adimensionnel qui est considéré comme la base des critères de performance des modèles pluie-débit.

Ce critère exprime le pourcentage de la variance naturelle que l'on gagne par rapport à un modèle de référence qui donnerait comme débits calculés la moyenne des débits observés.

Il est défini par :

$$Nash = \left[1 - \frac{\sum_{i=1}^n (Q_{obs} - Q_{sim})^2}{\sum_{i=1}^n (Q_{obs} - Q_{moy})^2} \right] \cdot 100 \quad (L.23)$$

Le critère de Nash trouve son intérêt dans le fait qu'il est adimensionnel, ce qui dans la plupart des cas, facilite les comparaisons entre diverses séries hydrométriques. Selon les modèles, le critère de Nash peut prendre des valeurs de $-\infty$ à 100%. Dans ce dernier cas, l'erreur de modélisation est nulle.

En fonction de la valeur du critère de Nash-Sutcliffe, on peut définir des performances des modèles pluie-débit au pas de temps journalier (Rakem, 1999):

Tableau II-8 Qualité des modèles en fonction des valeurs du critère de Nash

Nash (%)	Qualité
Nash <70	Très insuffisant
70 ≤ Nash ≤ 80	Insuffisant
80 ≤ Nash ≤ 85	Médiocre
85 ≤ Nash ≤ 90	Correct
90 ≤ Nash ≤ 95	Bon
Nash >95	Très bon

Cependant, ces appréciations ont été définies pour des petits bassins versants homogènes, dont les débits ne présentent pas de fluctuations importantes.

La validation d'un modèle pluie-débit reste donc problématique, du fait que les appréciations du modèle en fonction d'un critère sont étroitement liées aux caractéristiques statistiques (notamment la variance) des débits modélisés.

En effet, pour la plupart des critères, l'erreur de simulation est plus importante dans le

cas des hautes eaux (débits élevés) que dans le cas des étiages où les débits sont généralement inférieurs à $1.0 \text{ m}^3/\text{s}$.

On note donc que l'appréciation d'un modèle reste tributaire des expériences de chaque hydrologue, où la validation d'un modèle est déterminée en fonction des valeurs du critère choisi.

Au pas de temps mensuel, les écarts entre les données simulées et observées sont tels que la validation d'un modèle à cette échelle est problématique, d'autant plus que la durée de calage est longue.

Dans ce cas, le modélisateur doit se fixer une limite théorique, pour laquelle il doit déterminer l'appréciation du modèle.

II.5.1.5 Autres critères de validation des modèles

Les critères détaillés précédemment permettent une estimation globale de la qualité du modèle à reproduire la transformation pluie-débit. Cette estimation ne permet en aucun cas de mesurer la justesse de reproduction de certains débits caractéristiques, tels que les débits de crues ou d'étiages.

Rapport de crues annuelles

Pour comparer les débits de crue simulés et observés au pas de temps journalier, certains hydrologues utilisent le rapport des crues annuelles, défini comme suit:

$$CR = 100 \frac{Q_{c_{sim}} > 4M}{Q_{c_{obs}} > 4M} \quad (\text{II. 24})$$

Où M est le module interannuel des débits, et $Q_{c_{sim}} > 4M$ est la moyenne des pointes de crues définies chacune comme la valeur maximale d'une succession interrompue de débits simulés par le modèle supérieures à $4M$.

$Q_{c_{obs}}$ est la moyenne des pointes de crues définies chacune comme la valeur maximale d'une succession interrompue de débits observés supérieurs à $4M$.

Connaissant l'importance du débit maximal observé dans la prédétermination des crues, nous proposons de substituer le rapport des crues annuelles, par un indice (Ic) qui va mesurer le degré de reproduction du débit maximal observé durant toute la série historique. Il s'agit d'une comparaison entre le débit maximal simulé (Q_{mxsim}) avec le débit maximal observé (Q_{mxobs}):

$$Ic = \frac{Q_{mxsim}}{Q_{mxobs}} \quad (\text{II. 25})$$

Si l'indice (Ic) est égal à l'unité, la reproduction du débit maximal observé par le modèle est correcte.

Rapport des cumuls d'étiages (CET)

Ce rapport est très important dans le cas des modèles mensuels, il mesure la qualité de reproduction des débits d'étiages par le modèle, son expression est:

$$CET = 100 \frac{\sum_{i=1}^N Q_{sim} < M'}{\sum_{i=1}^N Q_{obs} < M'} \quad (II.26)$$

Où M' représente $\frac{1}{4}$ du module moyen interannuel. Dans le cas d'une bonne reproduction des débits d'étiages le CET égal à 100%.

En Algérie, les mois d'étiages correspondent à la période juin-septembre.

Pour le cas de notre étude, au pas de temps mensuel, le CET sera exprimé en fonction des débits, et non des périodes, vu que dans certains cas, les débits durant ces périodes sont très élevés, cas de la station du Sebbaou, où nous avons relevé un débit de $116 \text{ m}^3/\text{s}$, pour le mois de juin de l'année 1994.

II.6. Méthodes d'optimisation des modèles pluie-débit

L'estimation des paramètres d'un modèle passe par la minimisation de la fonction critère. La calibration manuelle sous forme d'essais-erreurs risque d'être longue et fastidieuse.

Des méthodes automatiques ont été développées, elles présentent l'avantage par rapport aux méthodes manuelles de permettre une estimation plus rapide et une exploration plus large des gammes des valeurs des paramètres.

Si y représente le vecteur des débits observés, les méthodes d'optimisation permettent de minimiser l'erreur de simulation ou fonction critère $E(\theta)$ d'après la relation suivante (Gupta et Sorooshian, 1983):

$$E(\theta) = \hat{y}(\theta) - y = \{e_1(\theta), e_2(\theta), \dots, e_N(\theta)\} \quad (II.27)$$

Où \hat{y} est le vecteur des débits simulés,

$e_1(\theta), e_2(\theta), \dots, e_N(\theta)$ sont les erreurs de simulation, $\theta_1, \theta_2, \dots, \theta_N$ représentent les paramètres du modèle.

Cependant, dans un contexte d'équifinalité, les méthodes automatiques ne permettent pas d'accéder (dans le cas de plusieurs paramètres) à la modélisation optimale, mais à un ensemble de jeux de paramètres conduisant à des simulations équivalentes (Grayson *et al.*, 1992).

On peut classer les méthodes d'optimisation en deux grandes classes:

II.6.1 Les méthodes locales

Ces méthodes adoptent une stratégie itérative. En partant d'un point de l'espace des

paramètres, on se déplace dans une direction qui améliore continuellement la valeur de la fonction critère jusqu'à ne plus pouvoir générer d'amélioration. Le jeu de paramètre ainsi trouvé correspond alors à l'optimum de la fonction (Perrin, 2000). On peut distinguer deux types de méthodes locales (Sorooshian et Gupta, 1995):

- Les méthodes directes : elles se basent seulement sur l'information de la fonction critère. Pour exemple on peut citer la méthode de Rosenbrock (Rosenbrock, 1960), basée sur une méthode de marche en escalier, la méthode pas à pas (Michel (1989), ou encore l'approche GLUE (Generalized Likelihood Uncertainty Estimation) qui est une méthode bayésienne (Beven Binley, 1992).
- Les méthodes du gradient : utilisent comme information à la fois la valeur de la fonction critère et celle du gradient de la fonction pour décider de la stratégie d'évolution dans l'espace des paramètres. La Descente du gradient (Widrow, 1960) utilisée dans les réseaux de neurones artificiels est un exemple de ce type de méthodes d'optimisation.

La méthode pas à pas

Cette méthode consiste à rechercher un minimum s'il existe suivant le chemin de la plus grande pente.

Nous détaillons cette méthode dans le cas d'un modèle à un seul paramètre:

On effectue une transformation du paramètre de façon à rendre à peu près équiprobable tout intervalle $[X, X+\Delta X]$ indépendamment de X , en prenant le logarithme du paramètre ($x = \ln X$), pour certaines valeurs on utilisera sinus X .

Soit F , la fonction objectif. Soit x_0 la valeur initiale d'un paramètre à optimiser, notons $x(1) \dots x(i), \dots x(n)$ les valeurs obtenues successivement et $x(o)$ la valeur optimale du paramètre.

Au pas d'optimisation $i+1$, on fait un essai de chaque côté de $x(i)$ d'une valeur Δx :

$x(i+1)=x(i)-\Delta x$ conduisant au critère $F_1(i+1)$;

$x(i+1)=x(i)+\Delta x$ conduisant au critère $F_2(i+1)$;

$F(i+1)=\max(F_1(i+1), F_2(i+1))$.

Si $F(i+1) > F(i)$, c'est à dire qu'il y a amélioration de la fonction objectif, on prend $x(i+1) = x_j(i+1)$; j correspond à $\max(F_j, j=1,2)$. Le pas d'auscultation peut également varier; Δx est doublé si l'on a obtenu 2n améliorations de la fonction objectif dans une même direction.

Si $F(i+1) < F(i)$, le prochain pas de recherche est divisé par deux, et on réitère à nouveau avec la modification des paramètres.

La condition de convergence porte sur Δx et non sur la fonction objectif, soit $\Delta x \leq \epsilon$ (ϵ valeur minimale pour Δx , soit l'obtention d'un nombre maximal d'essais pour l'optimisation. Cette méthode peut être généralisée aux modèles à plusieurs paramètres.

II.6.2. Les méthodes globales

L'optimisation des modèles peut se heurter à la présence d'un minimum local dans la surface d'erreur, et les méthodes locales sont généralement dépendantes du point de départ de l'espace des paramètres.

Les méthodes globales explorent une partie plus importante de l'espace des paramètres en adoptant diverses stratégies d'exploration, elles peuvent être déterministes ou probabilistes, dans ce cas la convergence de l'algorithme ne se fait plus au sens de distance mathématique mais de distance probabiliste (Duan *et al.*, 1992).

La méthode SCE-UA

La méthode Shuffled Complex Evolution SCE-UA (Duan *et al.*, 1992) est une méthode probabiliste multi-départ, basé sur le même principe les algorithmes génétiques. Ces derniers permettent de retenir d'une itération à la suivante des individus qui n'améliorent pas la qualité de la population totale, ce qui limite la convergence sur des optima locaux.

L'idée est de tirer au hasard un grand nombre de jeux de paramètres, d'évaluer la fonction coût pour chaque jeu et de garder le jeu pour lequel la fonction coût est maximale (pour cet ensemble de tests). En théorie, la probabilité de tomber sur le bon jeu de paramètre est nulle, car la mesure d'un singleton est nulle (dans le cas continu). C'est pourquoi il est fondamental de réaliser un très grand nombre de tels tirages pour espérer être proche de la valeur optimale. Typiquement, des valeurs de l'ordre de 1000 à 100 000 sont envisageables.

La méthode SCE-UA est résumée comme suit:

1- Choix des paramètres de l'algorithme :

- $p \geq 1$, p est le nombre de complexes,
- $m \geq n + 1$, m le nombre de points dans chaque complexe,
- n est le nombre de paramètres à être optimisé,
- s représente la taille de l'échantillon de points ($s = p \cdot m$);

2- Génération d'un échantillon de s points, $x_1, \dots, x_i, \dots, x_s$ dans l'espace des paramètres. Calculer la fonction objectif F à chaque point x_i . En absence d'information a priori, générer l'échantillon de points avec une distribution uniforme.

3- Classement des points : classer les s points dans l'ordre croissant de leur fonction objectif.

4- Préparation des complexes : diviser l'échantillon de s points en p complexes A_1, \dots, A_p , chacun contenant m points.

5- Evolution des complexes : faire évoluer chaque complexe A_1, \dots, A_p en s'appuyant sur un schéma proche de celui du Simplex, la méthode SCE-UA fait évoluer parallèlement des sous groupes de la population qui se mélangent périodiquement (échange d'informations).

6- Mélanger les complexes : remettre les complexes A_1, \dots, A_p dans D , tel que

$D = \{A_k, k = 1, \dots, p\}$. Classer D dans l'ordre croissant de leur fonction objectif.

7- Vérifier la convergence : si le critère de convergence est satisfait, arrêter

l'algorithme, sinon retourner à l'étape (4).

Pour une meilleure exploration du jeu de paramètres, Kuczera (1997) suggère que le nombre de complexes p soit égal au nombre de paramètres du modèle.

II.7. Conclusion

Ce chapitre nous a permis de présenter les bassins étudiés, et leurs principales caractéristiques hydrologiques. Nous avons réuni un échantillon de données mensuelles et journalières de six bassins algériens, dont les superficies varient de 345 à 3615 km², répartis de façon homogène à travers le territoire algérien.

L'étude statistique abordée dans ce chapitre, nous a permis de détailler quelques tests et ajustements nécessaires aux traitements de données hydrométéorologiques. Les tests ont confirmé l'homogénéité des variables, cependant l'indépendance des données n'a pas été confirmée pour les données mensuelles, vu qu'à cette échelle les mesures sont rarement stochastiquement indépendantes.

Concernant la validation des modèles, l'efficacité de Nash, sera utilisée comme étant le principal critère de comparaison des modèles pluie-débit. Cependant, il est indispensable de tester les modèles dans leur aptitude à reproduire certains débits caractéristiques tels que les étiages et les crues.

Chapitre III Simulation de la relation pluie-débit par les modèles conceptuels

III.1. Introduction

Dans le premier chapitre, nous avons évoqué les modèles pluie-débit dans leurs différents mécanismes et structures de base. Nous avons ainsi défini en fonction du découpage spatial deux classes de modèles : les modèles conceptuels globaux, dont le bassin est modélisé dans son entité hydrologique, et les modèles distribués, qui prennent en considération la variabilité spatiale du système considéré. Ces derniers ont l'avantage d'obtenir des informations ponctuelles du bassin, mais nécessitent l'introduction de tout un ensemble de données hydrologiques, et pédologiques, qui généralement, sont difficilement accessibles au niveau d'un seul organisme.

Les modèles globaux permettent une modélisation de la relation pluie-débit en traitant le bassin dans son entité géographique et hydrologique, et le considèrent comme une surface homogène.

Du point de vue pratique, ces deux types de modèles ont un point en commun, ils ne peuvent se passer de l'information hydrologique, c'est à dire des données des réseaux de mesure, et leurs performances sont étroitement liées à la disponibilité des chroniques

d'information. Dans le cas de l'approche distribuée, la problématique des données est amplifiée du fait que la connaissance régionale des paramètres hydrologiques, géomorphologiques et pédologiques (indispensables pour définir la variabilité spatiale des différents processus) est absolument nécessaire pour le fonctionnement de ces modèles.

Dans notre étude, la transformation pluie-débit sera étudiée par l'approche des modèles conceptuels globaux (à réservoirs), qui ont l'avantage de n'exiger que des séries de données météorologiques (pluie, ETP) et hydrométriques (débit).

III.1.1. Aspects de la modélisation pluie-débit

Classiquement, la modélisation pluie-débit par les modèles conceptuels (à réservoirs), est réalisée par deux processus différents, à savoir la fonction de production (fortement non-linéaire), et la fonction de transfert (qu'on considère linéaire) chargée de distribuer les débits vers l'exutoire.

La fonction de production

Représente l'interaction pluie-sol, elle traduit l'aptitude du bassin à produire des écoulements, après un ensemble de pertes, telles que l'évapotranspiration ou la rétention végétale. La pluie brute est donc transformée en pluie nette, après soustraction de l'évapotranspiration. Cette fonction de production est assurée par un réservoir, ou une fonction.

L'évapotranspiration est introduite sous forme d'entrée, elle est destinée dans la majorité des cas à calculer l'évapotranspiration réelle (ETR).

Dans certains modèles, l'évapotranspiration réelle est estimée à partir de l'évapotranspiration potentielle (modèle CREC), qui sera réduite par un paramètre à ajuster. Dans d'autres modèles, l'évapotranspiration réelle est déduite en fonction du niveau de remplissage d'un réservoir (réservoir sol), qui va conditionner l'état hydrique du sol (modèle GR3J).

L'eau qui n'a pas été évapotranspirée va s'infiltrer dans le sol, en fonction de son état hydrique. Si le sol est sec, la pluie va alimenter le réservoir sol, jusqu'à un niveau maximal, puis l'excès d'eau va ruisseler ou percoler vers la zone saturée.

La fonction de transfert

Après le calcul de la pluie nette, une partie de cette pluie va directement s'infiltrer dans le réservoir inférieur, l'autre partie va directement ruisseler vers l'exutoire.

L'infiltration est estimée de façon globale, et dépend des conditions initiales du sol, et peut contribuer dans certains cas, à l'écoulement.

On peut considérer plusieurs composantes de l'écoulement:

- Le ruissellement de surface: il y a partage de la pluie entre l'évapotranspiration, l'eau qui s'infiltré, et le ruissellement de surface dès que l'intensité des pluies dépasse la capacité d'infiltration du sol, ce ruissellement de surface forme l'écoulement rapide;
- L'écoulement de subsurface: une partie des précipitations infiltrées chemine quasi

horizontalement dans les couches supérieures du sol pour réapparaître à l'air libre à la rencontre d'un chenal horizontal.

L'écoulement de base: il représente l'écoulement souterrain en provenance de la nappe phréatique.

La fonction de transfert a pour objectif la distribution de l'écoulement, mais également le décalage dans le temps des débits vers l'exutoire. Elle est assurée par un ou plusieurs hydrogrammes unitaires, ou des réservoirs, ou par des paramètres (coefficients) qui expriment le délai entre la pluie nette et le débit correspondant.

Les débits sont issus des vidanges des différents réservoirs de routage, dont les lois sont des relations empiriques, qui généralement apparaissent sans grande signification physique.

Cette loi de vidange est une fonction de contenu du réservoir de routage, elle est spécifique à chaque type de réservoir utilisé (linéaire ou quadratique) (Rakem, 1999).

III.1.2. Méthode d'optimisation

Le calage d'un modèle pluie-débit consiste à estimer les paramètres du modèle, dans ce cas on doit minimiser l'erreur entre les données simulées et observées. Le calage demande donc le choix à la fois d'un critère de qualité, d'une méthode pour identifier les paramètres, ainsi que des séries de données destinées à fournir l'information nécessaire au calage (Sorooshian et Gupta, 1985). On doit définir une méthode d'optimisation (ou minimisation) par une fonction critère.

La méthode SCE-UA sera utilisée dans notre étude, pour l'estimation des paramètres des modèles conceptuels.

III.1.3. Mise en œuvre des modèles

Notre étude de la modélisation pluie débit concernera des modèles fonctionnant à deux pas de temps; mensuel et journalier.

Afin d'estimer les paramètres des modèles, ces derniers seront testés sur une période d'observation plus ou moins longue, afin de satisfaire au mieux un critère de qualité comparant simulations et observations. De ce fait, la durée de calage doit être déterminée selon le pas de temps du modèle.

Initialisation du système

Les modèles conceptuels permettent de simuler le comportement du bassin versant à travers des réservoirs censés reproduire le cycle hydrologique. Si la simulation par le modèle débute par une période très humide, il est indispensable d'initialiser les états initiaux de ces réservoirs, par des valeurs relativement élevées pour mieux rapprocher la valeur du débit calculé avec celle du premier débit observé. Une autre technique consiste à définir une période de mise en route du modèle, généralement une période n'excédant pas une année, et de tester le modèle après cette durée d'adaptation du modèle.

Dans notre cas, vu que nous ne disposons pas suffisamment de données, nous optons pour la méthode d'initialisation, les niveaux des réservoirs seront initialisés en fonction de chaque réservoir (de chaque modèle) et de chaque station à modéliser.

III.2. Les modèles conceptuels mensuels

La modélisation pluie-débit au pas de temps mensuel n'a été abordée que dans un cadre restreint, ce qui explique le nombre réduit des modèles pluie-débit à cette échelle.

Les hydrologues simulaient les débits mensuels par une simple agrégation des débits calculés au pas de temps journalier. Néanmoins ces modèles sont dépourvus de certains mécanismes déjà présents dans les modèles journaliers, et qui seraient difficilement modélisables à l'échelle mensuelle, tels que les transferts de la pluie par un hydrogramme unitaire, ou encore l'écoulement souterrain.

C'est par exemple le cas du modèle ABC de Thomas (1981) à quatre paramètres, ou encore le modèle Ta de Alley (1985), ou celui de Vandewiele et al. (1991) qui comporte un seul réservoir et trois paramètres, ce modèle a été conçu pour modéliser les bassins pluvieux de la Belgique, ou plus récemment le modèle empirique de Xiong et Guo (1999).

Nous pouvons dire que ces modèles sont généralement des modèles globaux, assez sommaires, et qui ont été conçus pour une application particulière, et d'autres sont le fruit d'une recherche assez élaborée (Mouelhi, 2002).

Selon Schaake et Chunzen (1989), ces modèles sont étudiés de façon à explorer l'impact du changement climatique dans l'aménagement des ressources hydriques, ainsi que dans la prévision des débits à long terme.

La complexité de la modélisation pluie-débit à l'échelle mensuelle peut être résumée par :

- La non-linéarité de la relation pluie-débit, qui s'accroît considérablement à cette échelle;
- La difficulté d'estimation de l'évapotranspiration réelle;
- La succession des jours pluvieux qui diffère d'un mois à l'autre;
- Les erreurs liées aux données hydrométéorologiques.

La modélisation de la relation pluie-débit par les modèles mensuels présente deux fonctions :

- Une fonction de production fortement non-linéaire, équilibre au mieux le bilan en volume du bassin versant, elle traduit l'aptitude du bassin versant à produire des écoulements.
- Une fonction de transfert pour le calcul des débits, issus de la vidange des réservoirs de routage

En Algérie, la modélisation pluie-débit mensuelle a déjà fait l'objet d'études par le département d'hydrologie de l'Agence Nationale des Ressources Hydrauliques, par l'application du modèle SMAP (Saidi, 1994).

Il s'agit d'un modèle à deux réservoirs et six paramètres, deux de ces paramètres à savoir le coefficient de récession de l'écoulement de base, et la capacité de saturation du réservoir sol doivent être estimés en fonction des caractéristiques hydrologiques du bassin étudié.

La conception et l'application de ces modèles restent très tributaires des caractéristiques hydrologiques des bassins étudiés. A cet effet on ne peut démontrer une certaine universalité, déjà difficilement identifiable à un pas de temps inférieur (journalier).

III.2.1. Présentation des modèles

Dans notre thèse nous présentons quatre (4) modèles pluie-débit mensuels, certains de ces modèles ont déjà fait l'objet d'une validation durant des travaux de modélisation pluie-débit (Benkaci, 2001, Mouelhi, 2002, Dechemi et al., 2003). Ces modèles sont constitués de réservoirs avec un nombre réduit de paramètres, et sont considérés comme étant les plus robustes.

III.2.1.1. Le modèle de Thornthwaite et Mather (modifié)

C'est un modèle développé initialement par Thornthwaite et Mather (1955), et repris par Mather (1981), Alley (1985) et Arnell (1992). Nous avons modifié certaines composantes du modèle pour l'adapter aux conditions hydroclimatiques de l'Algérie, en particulier la phase du bilan qui dans notre cas comporte un débit direct avant la neutralisation ($P(i)-ETP(i)$). Le modèle intègre deux réservoirs H et S et trois paramètres à caler, la fonction de vidange est linéaire (Figure III-1).

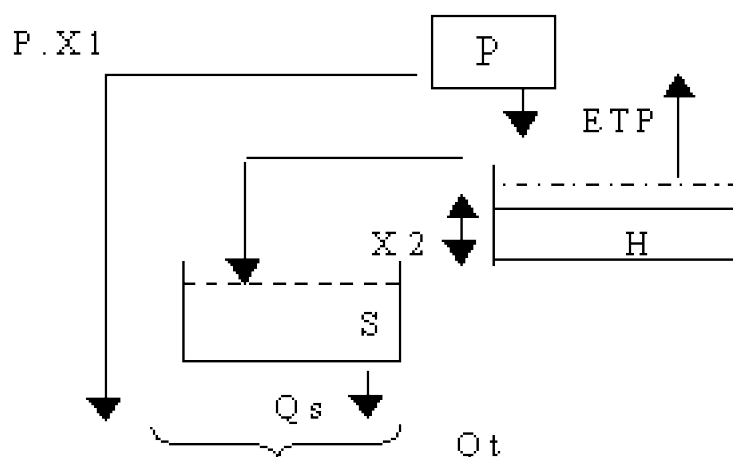


Figure III-1 Architecture du modèle de Thornthwaite et Mather (modifié)

Dans ce modèle, on calcule d'abord la fraction de pluie qui contribue directement à l'écoulement $Q_p(i)$:

$$Q_p(i) = P(i) X1 \quad (III - 1)$$

Où P(i) représente la pluie durant le mois (mm), et X1 est un paramètre à caler (0 < X1 < 1), i est le pas de temps (mensuel).

Le réservoir de stock H reçoit l'autre quantité de pluie (1-X1.P), lorsque la différence entre la pluie est l'évapotranspiration potentielle est positive, cette quantité alimente le réservoir H et devient H'.

Ce réservoir est limité par une capacité X2 (mm), qui lorsque son niveau dépasse ce seuil, l'excès d'eau va alimenter le réservoir de routage S, et va évacuer un débit Qs(i) calculé comme suit:

$$Qs(i) = S(i) X3 \dots \dots \dots (III - 2)$$

X3 est un paramètre du modèle (0 < X1 < 1), i est l'indice du temps (mois).

Dans le cas où le bilan «P(i)-E(i)» est négatif, le réservoir H est soumis à l'évapotranspiration, son niveau diminue par prélèvement selon la loi:

$$H(i) = H(i - 1) \exp(E(i)/X2) \dots \dots \dots (III - 3)$$

Le débit total est donc:

$$Qt(i) = Qs(i) + Qp(i) \dots \dots \dots (III-4)$$

III.2.1.2. Le modèle ABCD

Ce modèle développé par Thomas (1981), est un modèle conceptuel constitué de deux réservoirs, un réservoir sol (S) limité par un seuil maximal (paramètre A), et le deuxième réservoir (R), qui se vidange selon une loi linéaire, et régi par le paramètre D, ce modèle requiert quatre paramètres à caler (figure III-2).

Fonction de production

Pour calculer la pluie nette, la pluie brute (P(i)) est en premier lieu ajoutée au réservoir sol, dont le niveau augmente et devient:

$$S'(i) = S(i - 1) + P(i) \dots \dots \dots (III-5)$$

L'évapotranspiration réelle agit sur le réservoir sol, elle s'exprime comme suit:

$$ETR(i) = S'(i) \left(1 - \exp \left(- \frac{ETP(i)}{A} \right) \right) \dots \dots \dots (III-6)$$

Sous l'effet de l'évapotranspiration réelle, le niveau du sol diminue:

$$S(i) = S(i - 1) - ETR(i) \dots \dots \dots (III-7)$$

La pluie nette pour le mois (i) est calculée de la façon suivante:

$$Pn(i) = \frac{(-B) \frac{S^2(i)}{A}}{1 + (1 - 2B) \frac{S(i)}{A} - \sqrt{1 + \frac{S(i)}{A}} - 4 \frac{B.S(i)}{A}} \dots \dots \dots (III-8)$$

B représente le deuxième paramètre du modèle, il est compris entre 0 et 1.

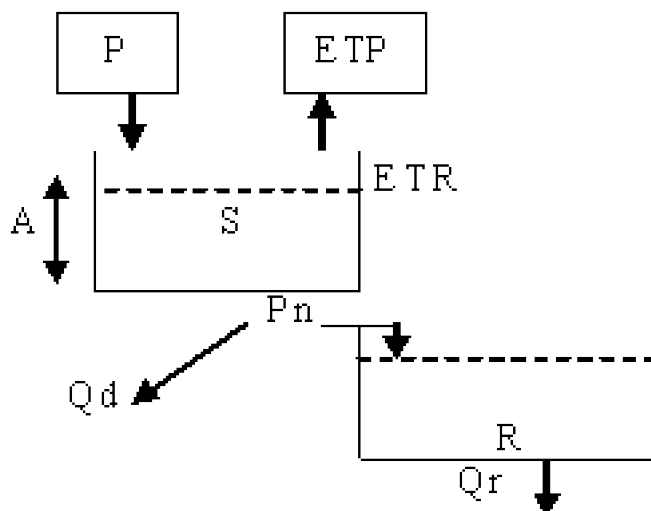


Figure III-2 Architecture du modèle ABCD

Fonction de transfert

Pour calculer le débit à l'exutoire, la pluie nette est divisée en fonction du paramètre C ($0 < C < 1$) en:

- Un écoulement direct $Q_d(i)$ exprimé par l'équation suivante:
- $Q_d(i) = (1 - C) * P_n(i)$(III-9)
- Une pluie nette qui alimente le réservoir de routage $R(i)$:
- $R(i) = R(i) + C.P_n(i)$(III-10)

Ce réservoir évacue un débit linéaire en fonction du stock d'eau:

$$Q_r(i) = D.R(i).....(III-11)$$

Le débit total à l'exutoire calculé par le modèle est:

$$Q_t(i) = Q_d(i) + Q_r(i).....(III-12)$$

III.2.1.3. Le modèle GR2M (1994)

Le modèle GR2M (Makhlouf, 1994, Makhlouf et Michel, 1994) développé au Cemagref, découle des modèles GR journalier (Michel, 1983, Edijatno et Michel, 1989) et a subi plusieurs modifications (Lavabre et al. (1997), Mouelhi (2002)).

Le modèle a été utilisé pour une estimation régionale des débits moyens et d'étiage mensuels des bassins sud-est de la France (Lavabre et al., 1999).

Dans notre étude, nous retenons deux versions, celle de Makhlouf (1994), et la plus récente développée par Mouelhi (2002). Le modèle GR2M (1994) est constitué de deux réservoirs et requiert deux paramètres (Figure III-3).

La pluie $P(i)$ et l'évapotranspiration potentielle $ETP(i)$ sont ajustées par un paramètre U:

$$U = P.(i)ETP(i)/p(i)^\alpha + ETP(i)^\alpha \dots\dots\dots (III- 13)$$

Avec $\alpha = 0.5$

Cette opération permet de quantifier l'interception, et de calculer la pluie nette Pn(i) et l'évapotranspiration nette En(i) :

$$\begin{cases} Pn(i) = P(i) - U(i) \\ En(i) = ETP(i) - U(i) \end{cases} \dots\dots\dots (III- 14)$$

Un paramètre X1 intervient à ce niveau, il sert à calculer Pn'(i) et En'(i):

$$\begin{cases} Pn'(i) = Pn(i)X1 \\ En'(i) = En(i).X1 \end{cases} \dots\dots\dots (III-15)$$

Sous l'effet de Pn'(i) et En'(i) l'état du réservoir H est modifié et donne respectivement H1 et H2.

Sous l'effet de Pn(i)'

$$H1(i) = (H(i) + AV)/(1+HV/A) \dots\dots\dots (III-16)$$

Avec:

H1(i): Humidité du sol à la fin du mois (mm);

H(i) : Humidité du réservoir H au début du mois (mm);

A : Capacité maximale du réservoir (H), fixée à 200mm;

V : Tangente hyperbolique de (Pn'/A).

L'excès d'humidité est donné par l'équation:

$$Pe(i) = Pn'(i) + H(i) - H1(i) \dots\dots\dots (III-17)$$

Un cinquième de l'excès d'humidité (1/5.Pe(i)) ruisselle, et participe au débit total.

Sous l'effet de En'(i)

$$H2(i) = \frac{H1(i)(1-W)}{(1+W(\frac{1-H1(i)}{A}))} \dots\dots\dots (III-18)$$

W = Tangente hyperbolique de (En'(i)/A).

Le deuxième réservoir S(i) avec un niveau initial S(i-1), reçoit la quantité (4/5 Pe(i)) et son niveau devient:

$$S1(i) = S1(i) + 4/5Pe(i) \dots\dots\dots (III-19)$$

Ce réservoir évacue un débit Qs(i):

$$Qs(i) = X2.S1(i) \dots\dots\dots (III-20)$$

Le débit total à l'exutoire calculé par le modèle est donné par:

$$Q_t(i) = Q_s(i) + 1/5P_e(i) \dots \dots \dots (III-21)$$

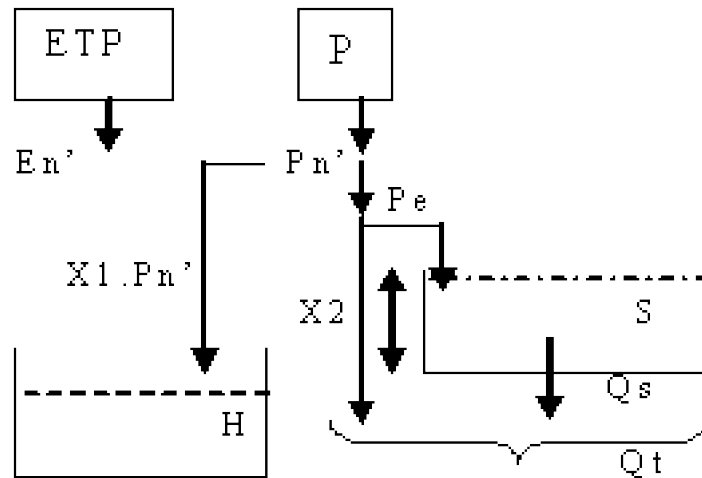


Figure III-3 Architecture du modèle GR2M (1994)

III.2.1.4. Le modèle GR2M (2002)

C'est un modèle issu des travaux de Mouelhi (2002), il a été conçu après une comparaison de plusieurs modèles mensuels, et une adaptation du modèle conceptuel journalier GR4J (Perrin, 2000) à l'échelle mensuelle.

Le modèle GR2M comporte deux réservoirs et deux paramètres X1 et X2, il est caractérisé par deux fonctions (Figure III-4):

Une fonction de production assurée par le réservoir sol S(i) qui est limité par une capacité maximale X2 (paramètre à caler);

Une fonction de transfert représentée par le réservoir R(i), dont le niveau est modifié par l'échange souterrain en fonction du paramètre X1.

En premier lieu, le niveau du réservoir sol S(i) devient S1(i) sous l'action de la pluie:

$$S1(i) = \frac{S(i) + X2.V}{1 + \frac{S(i)}{X2}.V} \quad (III-22)$$

V: Tangente hyperbolique de (P/A)

Une partie de la pluie Pr participe à l'écoulement:

$$Pr(i) = S(i) + P(i) - S1(i) \dots \dots \dots (III-23)$$

Le niveau de ce réservoir est modifié sous l'effet de l'évapotranspiration potentielle, et devient S2(i):

$$S2(i) = \frac{S1(i)(1 - W)}{1 + W(1 - \frac{S(i)}{X2})} \quad (III-24)$$

W = Tangente hyperbolique de ETP(i)/A.

A la fin du mois, le réservoir sol S(i) en fonction du paramètre X2 évacue par percolation une quantité d'eau, le niveau de ce réservoir niveau diminue:

$$S'(i) = \frac{S2(i)}{\left[1 + \left(\frac{S2(i)}{X2}\right)^{2.5}\right]^{0.4}} \quad (\text{III-25})$$

La percolation souterraine Ps est exprimée :

$$Ps(i) = S2(i) - S'(i) \quad (\text{III-26})$$

Le deuxième réservoir R avec un niveau initial R(i) reçoit les quantités de pluie Ps et Pr et devient R' :

$$R'(i) = R(i) + Ps(i) + Pr(i) \quad (\text{III-27})$$

Sous l'effet de la vidange et des échanges souterrains, R'(i) devient R1(i):

$$R1(i) = \frac{R'(i)}{X1 + \frac{R'(i)}{50}} \quad (\text{III-28})$$

L'échange souterrain F(i) agit sur le réservoir de routage (R(i)), il est exprimé par:

$$F(i) = 50.(X1) \ln \left(1 - \frac{R'(i)}{50.X1}\right) \quad (\text{III-29})$$

Le débit total calculé par le modèle, exprimé en lame d'eau est:

$$Q(i) = R'(i) + F(i) - R1(i) \quad (\text{III-30})$$

Figure III-4 Architecture du modèle GR2M (2002)

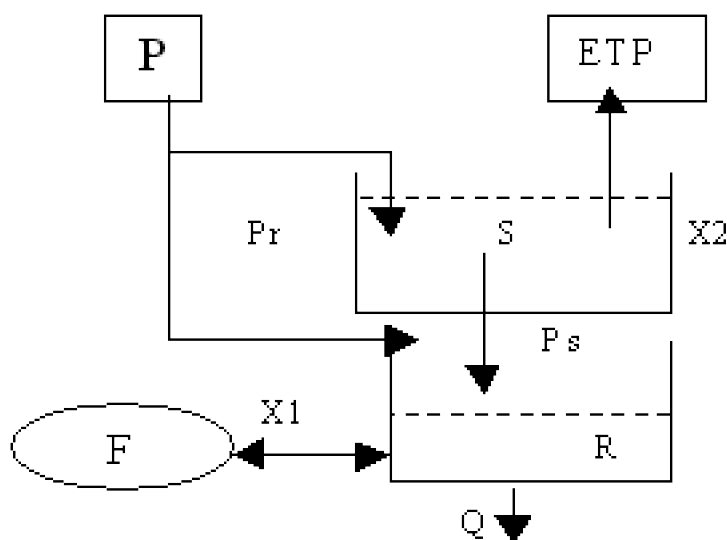


Figure III-4 Architecture du modèle GR2M (2002)

III.2.2. Résultats de la modélisation

Les modèles pluie-débit mensuels testés dans ce chapitre seront appliqués aux cinq stations hydrométriques, réparties sur quatre bassins versants différents, à savoir la Cheffia, l'Isser, le Sebbaou et la Tafna.

Les tailles des séries hydrométriques utilisées dans cette étude, correspondent à une durée d'observations de seize ans. Pour ce qui est des données pluviométriques, les échantillons sont relativement plus longs, de ce fait si le modèle est validé, nous nous proposons de combler les stations hydrométriques tronquées à savoir les stations de l'Isser, de la Tafna et du Sebbaou.

Nous avons été dans l'obligation de choisir les stations hydrométriques qui sont non seulement représentatives des bassins, mais dont la qualité et le nombre de données, soient assez satisfaisants pour appliquer un modèle pluie débit à cette échelle.

Les modèles seront testés sur une période longue de seize ans, dans laquelle la série modélisée est divisée en deux sous-échantillons, de durées homogènes, le premier servant de calage pour l'estimation des paramètres, le deuxième pour la validation du modèle, ce qui faciliterait la comparaison des résultats entre les deux périodes.

Initialisation des modèles

L'initialisation du système est très importante dans l'application des modèles pluie-débit à réservoirs, cette procédure permet le fonctionnement en continu d'un tel modèle.

Etant donné que la période de simulation commence par l'année calendaire (Janvier), les niveaux des réservoirs des modèles seront initialisés par des valeurs élevées. Pour exemple, les niveaux initiaux des réservoirs sol (S) et de routage (R) du modèle GR2M (2002) sont respectivement de 290 mm et 30 mm pour la station de Lakhdaria, pour la station de Tafna, ces valeurs sont respectivement de 70 mm et 30 mm. Pour un même modèle, les différents réservoirs sont initialisés par des valeurs différentes en fonction des

MODELISATION PLUIE-DEBIT MENSUELLE ET JOURNALIERE PAR LES MODELES CONCEPTUELS ET LES SYSTEMES NEURO-FLOUS (APPLICATION AUX BASSINS ALGERIENS)

premiers débits de chaque station.

Pour le bassin de la Cheffia, vu que le premier débit est relativement faible, tous les modèles seront initialisés par des valeurs nulles.

Les modèles ont été optimisés avec la méthode pas à pas (Michel, 1989) détaillée dans le chapitre II.

Les résultats des différents modèles des périodes de calage et de validation de chaque station, sont présentés respectivement dans les tableaux III-1 et III-2.

Stations	Cheffia (1978-84)		Lakhdaria (1980-87)		La Traïlle (1970-77)		Sebbaou (1980-87)		Tafna (1980-1987)	
	Nash (%)	CET (%)	Nash (%)	CET (%)	Nash (%)	CET (%)	Nash (%)	CET (%)	Nash (%)	CET (%)
Thornthwaite	66.2	360	69.2	464	75.9	291	65.8	478	47.9	800
ABC	70.8	112	55.0	508	59.5	336	60.5	593	39.7	800
GR2M (1994)	71.0	102	65.8	256	86.9	77	77.9	197	40.7	235
GR2M (2002)	75.8	73	61.3	34	88.3	64	79.1	527	51.1	43

Tableau III-1 Résultats des modèles mensuels de la période de calage

Stations	Cheffia (1985-91)		Lakhdaria (1988-95)		La Traïlle (1979-85)		Sebbaou (1988-95)		Tafna (1988-1995)	
	Nash (%)	CET (%)	Nash (%)	CET (%)	Nash (%)	CET (%)	Nash (%)	CET (%)	Nash (%)	CET (%)
Thornthwaite	77.5	283	50.0	525	26.9	281	47.0	739	-27.	1000
ABC	70.8	125	34.5	435	-1.8	374	27..	900	-110	1000
GR2M (1994)	79.5	136	66.8	243	43.0	90	51.2	310	16.	386
GR2M (2002)	73.5	93	74.9	20	22.3	65	50.5	760	-57	63

Tableau III-2 Résultats des modèles mensuels de la période de validation

D'après les résultats présentés dans les tableaux III-1 et III-2, nous remarquons que les modèles ont généralement donné des résultats plus performants en période de calage qu'en période de validation. En effet, pour le modèle de Thornthwaite les valeurs du

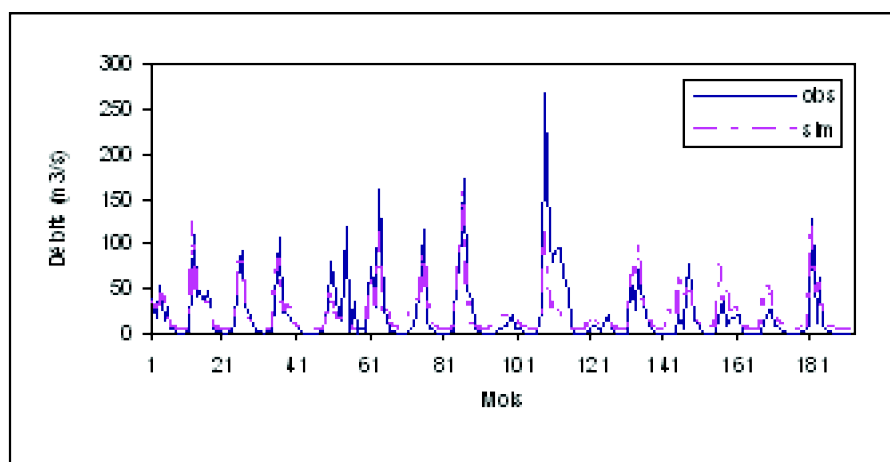
critère de Nash sont plus élevées en période de calage que celles de la validation. Pour exemple, l'efficacité de Nash est égale à 69.2% pour la station de Lakhdaria, en période de calage, elle n'est que de 50% seulement en période de validation, soit une différence de près de 20% entre les deux périodes.

Pour le bassin du Sebbaou, il est à signaler la même remarque, c'est à dire une baisse du niveau de simulation qui atteint en moyenne -19.0%. Pour le cas de la Cheffia, les résultats en période de validation présentent des valeurs relativement élevées par rapport à celle du calage, mais ceci reste limité aux deux modèles de Thornthwaite et de GR2M version 1994.

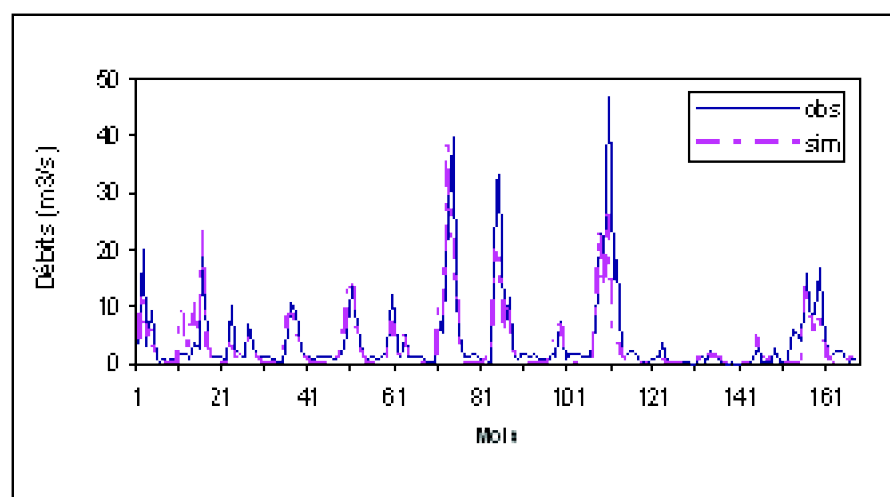
La valeur la moins performante, reste la station de la Tafna avec seulement 39.0% pour le modèle ABCD. Ce modèle enregistre la détérioration la plus remarquable avec un résultat très médiocre égal à -110%, soit une baisse de l'efficacité de Nash atteignant près de -150%.

Les modèles GR2M ont donné des résultats plus performants, en effet, les valeurs du critère de Nash sont très élevées pour la station de la Traille (Nash = 88%), en période de calage. En période de validation, ce modèle présente une baisse de la qualité de modélisation, le critère de Nash atteint seulement 22%.

La version 2002 (Mouelhi, 2002) présente des performances par rapport à l'ancienne version (Makhlouf et Michel, 1994). Le modèle GR2M paraît plus robuste, et permet une meilleure simulation des débits moyens et élevés, en particulier pour les stations du Sebbaou et de Cheffia (figure III-5 (a), (b)).

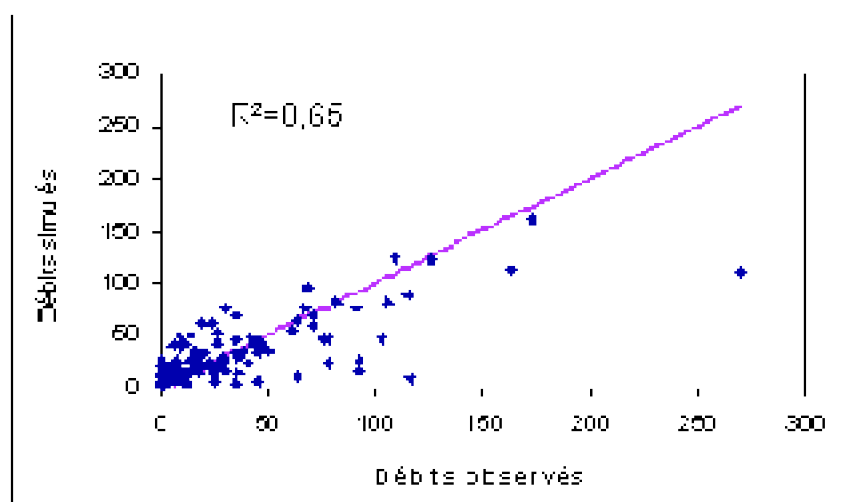


(a)

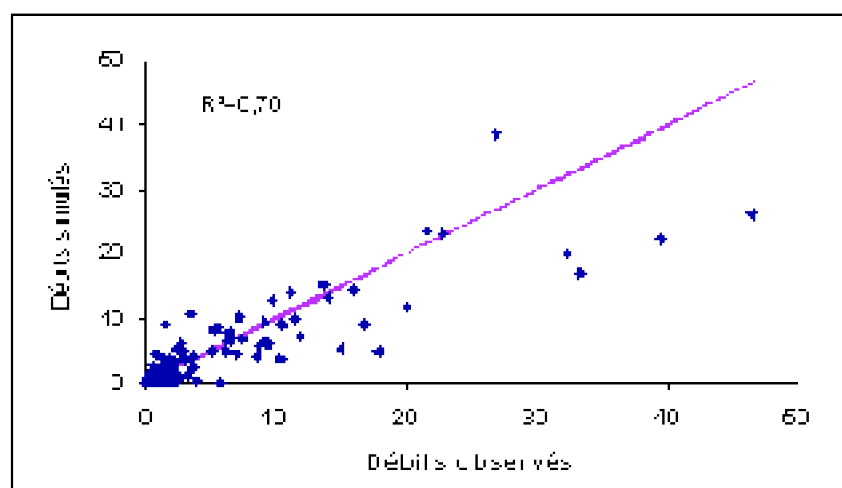


(b)

Figure III-5 Simulation des débits mensuels par le modèle GR2M (2002), (a) station du Sebbaou, (b) station de Cheffia



(a)



(b)

Figure III-6 Comparaison entre débits observés et simulés par le modèle GR2M (2002), (a) station du Sebaou, (b) station de Cheffia

Les résultats des tableaux III-1 et III-2, montrent que les modèles n'arrivent pas à simuler de façon correcte les débits des mois d'étiage. En effet les volumes de ces débits sont nettement sous-estimés pour le modèle GR2M (2002), dont les valeurs du critère d'étiage (CET) sont nettement faibles, elles ne dépassent pas les 30% pour la station de Lakhdaria, ce qui explique une mauvaise reproduction des débits d'étiage.

Ce phénomène est remarqué pour les autres stations, où le rapport CET est excessivement surestimé, le CET dépasse 200% pour le modèle de Thornthwaite et cela pour toutes les stations en phase de calage. Ceci s'explique par les lois de tarissement des modèles. Lorsque la pluie est supérieure à l'évapotranspiration, l'humidité du sol simulée est largement supérieure à son état réel durant les mois d'étiage, le modèle donne des débits dépassant $4 \text{ m}^3/\text{s}$, alors que les débits d'étiages observés sont inférieurs à $1.5 \text{ m}^3/\text{s}$.

Pour les autres modèles, tous les débits des mois d'étiage sont soit surestimés, tels que les modèles de Thornthwaite et ABCD, où on enregistre un rapport du CET très médiocre de 1000% pour la station de la Tafna, soit sont mal reproduits, dans ce cas, les débits sont largement sous-estimés.

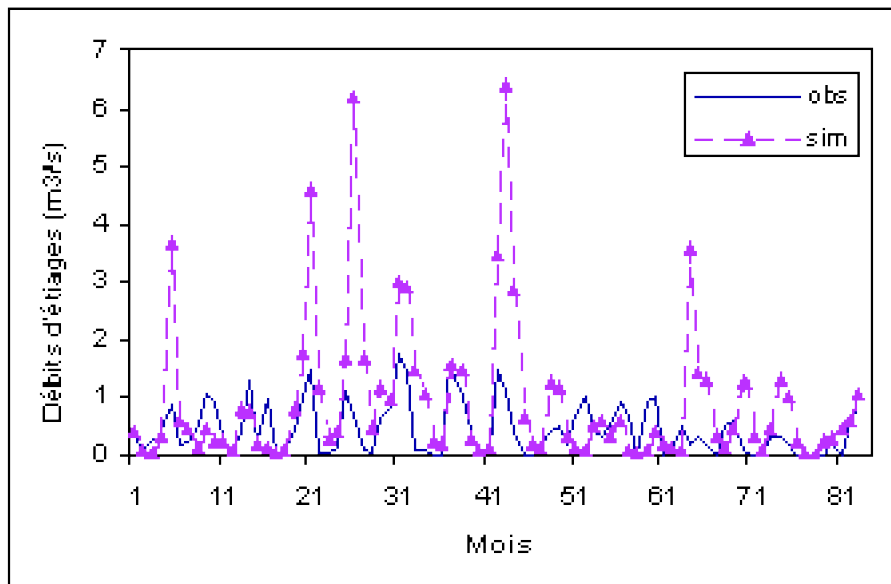


Figure III-7 Surestimation des débits d'étiages par le modèle GR2M (1994), station de Lakhdaria

En comparant les résultats de tous les modèles des différentes stations, l'analyse de variance des valeurs du critère de Nash, confirme que le modèle GR2M (2002) est largement plus performant que les autres modèles (Figure III-8), avec une moyenne de 72%, ce modèle paraît plus robuste en période de calage, suivi du modèle GR2M (1994).

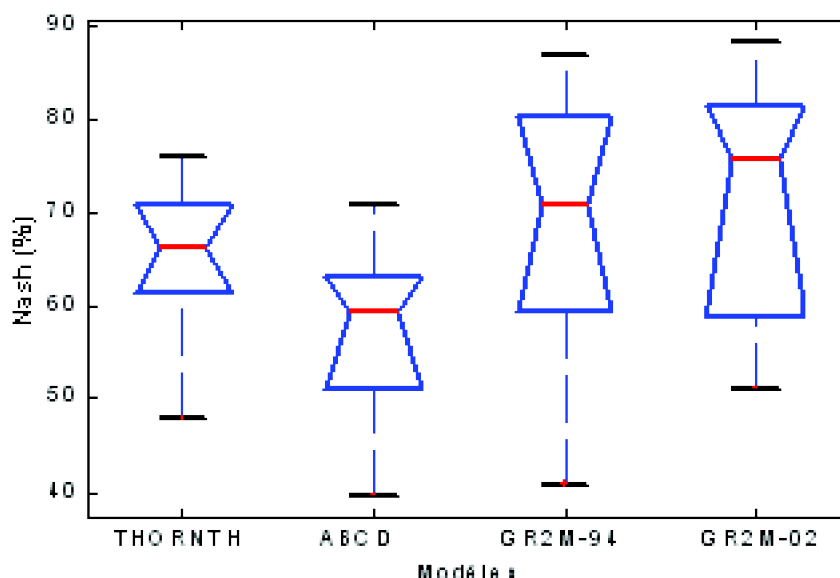


Figure III-8 Performance des modèles en fonction du critère de Nash (période de calage)

En période de validation, le modèle GR2M (1994) est nettement supérieur aux autres modèles (figure III-9), avec une moyenne du critère de Nash de près de 45%, il surclasse

la nouvelle version du modèle GR2M (2002).

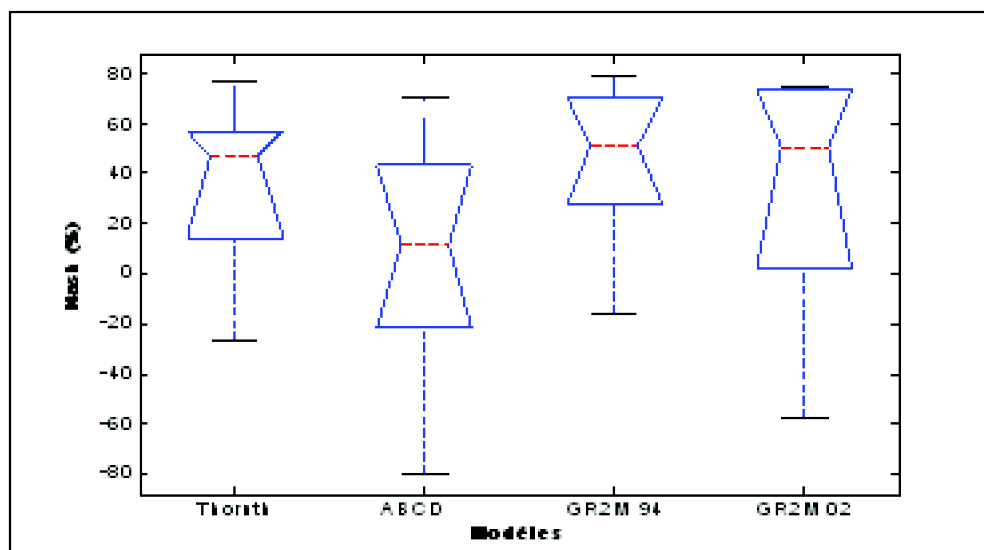


Figure III-9 Performance des modèles en fonction des critères de Nash (période de validation)

Néanmoins ces modèles conceptuels présentent l'inconvénient d'être très sensibles aux changements climatiques qui peuvent survenir entre les différentes périodes. Nous avons souligné une baisse des performances, ou plus exactement, une détérioration de la modélisation pour certains modèles (ABCD), où la diminution du critère de Nash peut atteindre -150 %.

Les valeurs de différents paramètres des modèles sont représentées sous forme de moyenne de toutes les stations dans le tableau III-3.

Modèles	Paramètres			
	Thorntwaite	X1= 0.1	X2= 455	X3 =0.45
ABCD	X1= 82.5	X2= 0.6	X3=0.9	X4= 0.5
GR2M (1994)	X1= 0.4		X2= 0.67	
GR2M (2002)	X1= 4		X2= 620	

Tableau III-3 Moyenne des paramètres optimisés des modèles

III.3. Les modèles conceptuels journaliers

Contrairement à l'échelle mensuelle, les modèles conceptuels journaliers sont nombreux et très utilisés en hydrologie. Cependant, plusieurs modèles présentent des particularités liées aux caractéristiques hydroclimatiques des pays où ils ont été conçus.

Pour éviter toute ambiguïté concernant la transposition du type de ces modèles à cette échelle, nous proposons quatre modèles relativement récents qui ont fait preuve de leur robustesse, et qui ont été sujet à plusieurs recherches.

III.3.1. Présentation des modèles

Les modèles GR ont été initialement développés au Cemagref, au début des années quatre vingt. A partir des structures du modèle CREC (Cormary et Guilbot, 1973), des modifications ont été apportées à ces modèles, Michel (1983) propose un modèle à deux réservoirs limités par un seul paramètre. Edijatno (1987) a réutilisé le modèle en modifiant la fonction de production, et donc un deuxième paramètre a été ajouté au modèle, exprimant la capacité maximale du réservoir R, et a été testé sur plus de cent (100) bassins français.

Une amélioration considérable du modèle a été réalisée (Edijatno et Michel, 1989, Edijatno, 1991) par l'introduction d'un hydrogramme unitaire simulant le transfert de la pluie nette, et par un quatrième paramètre introduit dans le modèle pour simuler les échanges inter-bassins (modèle GR4J).

Nous proposerons deux modèles: la version récente du modèle GR3J développée par Edijatno et al. (1999), et la dernière version du modèle GR4J (Perrin, 2000).

III.3.1.1. Le modèle GR3J

Le modèle GR3J comporte deux réservoirs, et trois paramètres. Les entrées du modèle à savoir la pluie et l'évapotranspiration potentielle sont transformées respectivement en pluie nette (P_n) et évapotranspiration nette (E_n), par l'interception.

$$\begin{cases} \text{Si } P(i) \geq ETP(i) \Rightarrow P_n(i) = P(i) - ETP(i) \text{ et } E_n(i) = 0 \\ \text{Si } P(i) < ETP(i) \Rightarrow E_n(i) = P(i) - ETP(i) \text{ et } P_n(i) = 0 \end{cases} \quad (II-31)$$

La deuxième opération dans ce modèle, se fait par le réservoir S, réservoir de production avec une capacité maximale A, limitée à 330 mm. Ainsi, le niveau du réservoir $S(i)$ augmente si les entrées $P_n(i)$ sont positives, de même que son niveau diminue si les flux $E_n(i)$ sont positifs.

$P_n(i)$ et $E_n(i)$ sont calculées comme suit:

$$P_s(i) = \frac{P_n(i) + \left(1 - \left(\frac{S(i)}{A}\right)^2\right)}{1 + (P_n(i)/A \left(1 + \left(\frac{S(i)}{A}\right)\right))} \quad (\text{III-32})$$

$$E_s(i) = \frac{E_n(i) + S(i)/A \left(1 - \left(\frac{S(i)}{A}\right)\right)}{1 + E_n(i)/A \left(2 - \left(\frac{S(i)}{A}\right)\right)} \quad (\text{III-33})$$

La différence entre (Pn(i)-Ps(i)), constitue la fraction réelle qui va générer des débits, elle se divise en deux:

La première qui représente 90% de (Pn(i)-Ps(i)) sera transformée par le premier hydrogramme unitaire SH1, la seconde (10%) par l'hydrogramme SH2.

Les fonctions de ces deux hydrogrammes sont déterminées comme suit :

$$\left\{ \begin{array}{l} \forall 0 \leq j \leq X3 \Rightarrow SH1(j) = \left(\frac{j}{X3}\right)^3 \\ \forall j > X3 \Rightarrow SH1(j) = 1 \end{array} \right. \quad (\text{III-34})$$

X3 est un paramètre exprimant le temps de montée de l'hydrogramme unitaire SH1.

$$\left\{ \begin{array}{l} \forall 0 \leq j < X3 \Rightarrow SH2(j) = 1/2 \left(\frac{j}{X3}\right)^3 \\ \forall X3 \leq j < 2 X3 \Rightarrow SH2(j) = 1 - 1/2 \left(2 - \frac{j}{X3}\right)^3 \\ \forall j \geq 2 X3 \Rightarrow SH2(j) = 1 \end{array} \right. \quad (\text{III-35})$$

La première quantité translatée par l'hydrogramme SH1, va directement constituer une entrée pour le réservoir de routage R, soit donc Q9.

Ce réservoir est soumis aux échanges de flux F, et dépend du niveau d'humidité de ce réservoir. Lorsque X1 est positif, le flux constitue une sortie du réservoir R, lorsque le paramètre X1 est négatif, F représente une entrée pour le réservoir R.

La quantité F est donnée par la relation:

$$F(i) = X1 \left(\frac{R(i)}{X2} \right)^1 \tag{III-36}$$

Où X1, X2 représentent les paramètres du modèle à déterminer.

Le niveau du réservoir de routage R sera la somme du flux F, et du débit translaté par l'hydrogramme SH1:

$$R^*(i) = \max(R(i) + F(i) + Q9(i)) \dots \dots \dots \tag{III-37}$$

Où R(i) est le niveau initial du réservoir R. Le terme max, est utilisé pour éviter d'obtenir des niveaux négatifs.

Si F est négatif, il sera additionné au débit routé par l'hydrogramme SH2, et donnera ainsi le débit Qd. De même que si ce résultat est négatif ce débit est remis à zéro (Figure III-8).

Le réservoir de routage évacue un débit qui est égal à:

Le débit total (Qt(i)) sera la somme des débits QR(i) et Qd(i):

$$Qt(i) = Qr(i) + Qd(i) \dots \dots \dots \tag{III-39}$$

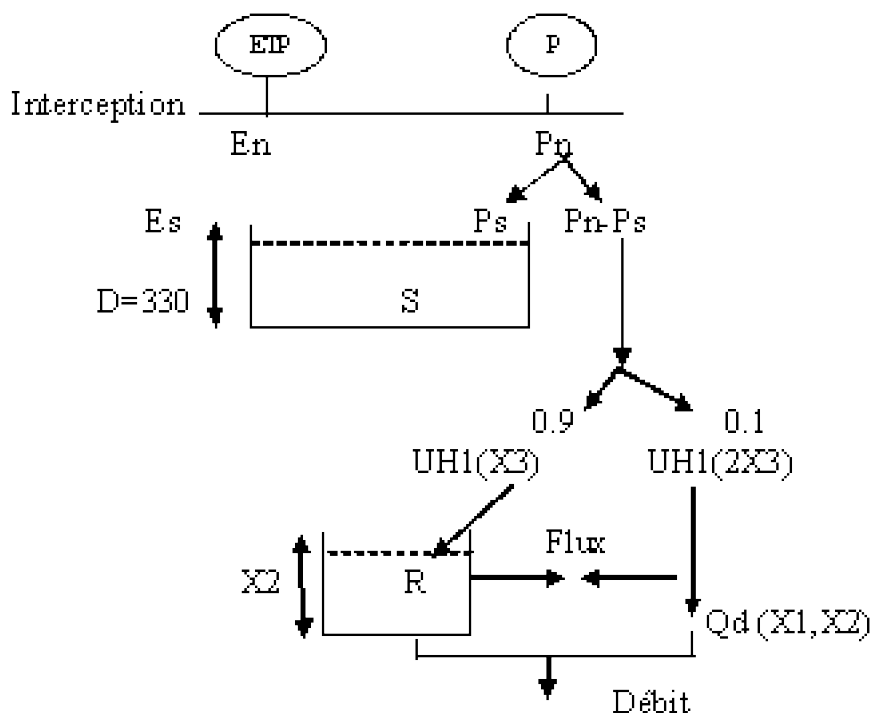


Figure III-10 Architecture du modèle GR3J

III.3.1.2. Le modèle GR4J (2002)

Le modèle GR4J a été développé par Perrin (2000), (Perrin *et al.*, 2001, 2002), à la suite d'une étude comparative de 38 modèles testés sur 429 bassins européens. Le modèle comporte deux réservoirs et nécessite quatre paramètres à caler.

Le premier réservoir est le réservoir Sol, destiné à la production de l'écoulement, le

deuxième est le réservoir R, qui est un réservoir de Routage. Le modèle intègre deux hydrogrammes unitaires, avec un paramètre commun (C).

Fonction de production

Les deux variables d'entrée sont la pluie, et l'évapotranspiration potentielle. La fonction de production se divise en deux phases:

La phase d'interception: dans cette partie, on calcule la pluie nette (Pn) et l'évapotranspiration nette (En), suivant deux cas:

$$\begin{cases} S: P(t) \geq ETP(t) \Rightarrow P_n(t) = P(t) - ETP(t) \text{ et } E_n(t) = 0 \\ S: P(t) < ETP(t) \Rightarrow E_n(t) = ETP(t) - P(t) \text{ et } P_n(t) = 0 \end{cases} \quad (\text{III-40})$$

La phase de Production (Réservoir sol) : ce réservoir est limité par une capacité maximale A (mm), qui est le premier paramètre du modèle, paramètre à caler.

Le réservoir Sol, sert à conditionner l'humidité du bassin. Si la pluie nette (Pn) est nulle, l'évapotranspiration fait diminuer le niveau du réservoir sous l'effet l'évapotranspiration réelle (ES).

Si la pluie nette est supérieure à zéro, le niveau du réservoir augmente, et cette quantité de pluie (PS) est stockée dans ce réservoir.

PS et ES sont données par:

$$PS(t) = \frac{A \left(1 - \left(\frac{S(t)}{A} \right)^2 \right) W}{1 + \frac{S(t)}{A} W} \quad \text{si } P_n(t) > 0 \quad (\text{III-41})$$

$$ES(t) = \frac{S(t) \left(2 - \frac{S(t)}{A} \right) V}{1 + \left(2 - \frac{S(t)}{A} \right) V} \quad \text{si } E_n(t) > 0 \quad (\text{III-42})$$

A la sortie du réservoir sol, les percolations (Per) contribuent à l'écoulement:

$$Per(t) = S(t) \left[1 - \left(1 + \left(\frac{S(t)}{2.25.A} \right)^{-4} \right)^{-1/4} \right] \quad (\text{III-43})$$

Fonction de transfert

A cette étape il y a séparation des écoulements, la partie complémentaire de la pluie efficace (Pn-Ps), à laquelle on additionne les percolations, est séparée en deux

composantes:

- Un écoulement (Q1) pseudo-direct (écoulement superficiel), qui atteint l'exutoire après un décalage par l'hydrogramme unitaire SH2, il représente 10% de la pluie efficace.
- L'autre partie soit 90% de la pluie efficace, est routée par l'hydrogramme (lent) SH1, et permet d'alimenter le réservoir de routage R (Figure III-9)

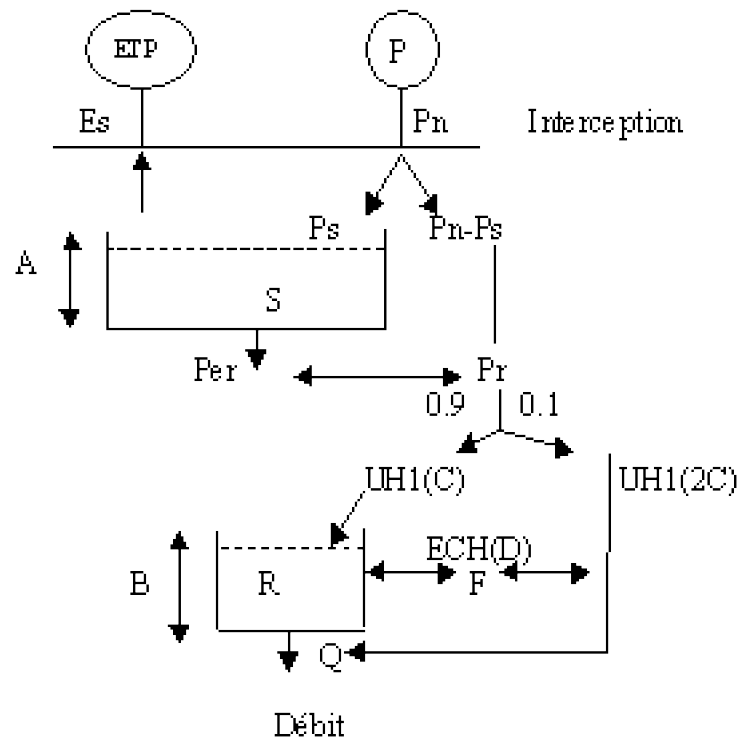


Figure III-11 Architecture du modèle GR4J

L'hydrogramme unitaire

Les expressions des fonctions des deux hydrogrammes sont décrites comme suit:

$$\begin{cases} \text{Si } 0 \leq j \leq C \Rightarrow SH1 = \left(\frac{j}{C}\right)^{5/2} \\ \text{Si } j > C \Rightarrow SH1 = 1 \end{cases} \quad (III-44)$$

$$\begin{cases} \text{Si } 0 < j < C \Rightarrow SH2(j) = 1/2 \left(\frac{j}{C}\right)^{3/2} \\ \text{Si } C \leq j < 2C \Rightarrow SH2(j) = 1 - 1/2 \left(2 - \frac{j}{C}\right)^{3/2} \\ \text{Si } j > 2C \Rightarrow SH2(j) = 1 \end{cases} \quad (III-45)$$

Les deux hydrogrammes utilisés (SH1, SH2) dépendant d'une même durée de référence (C) exprimée en jour (paramètre à optimiser).

Pour la nouvelle formulation du modèle GR4J, la variable j peut prendre des valeurs non entières.

Le réservoir de routage

La principale composante de l'écoulement transite par un réservoir de routage, qui dans le cas des modèles GR est non linéaire, limitée par une capacité (paramètre B). La loi de vidange instantanée est:

$$Qr(t) = K.[R(t)]^{5/2} \dots\dots\dots (III-46)$$

Où K est une constante et R, le niveau du réservoir de routage.

Après intégration, le niveau de ce réservoir est modifié comme suit:

$$\frac{1}{R_{i+1}^4} = \frac{1}{R_i^4} + \frac{1}{B^4} \dots\dots\dots (III-47)$$

B représente la capacité maximale (mm) du réservoir. Dans cette équation on a posé

$$\frac{1}{B^4} = 4K\Delta t$$

, et Δt est le pas de temps (égal à un jour).

Le débit évacué par le réservoir à la fin du pas de temps est:

$$Qr(i) = R_i - R_{i+1} \dots\dots\dots (III-48)$$

L'échange inter-bassins (T)

Le modèle GR4J permet de prendre en considération l'échange inter-bassins, et cela non pas par l'addition d'un nouveau réservoir, mais plutôt par une fonction qui correspond à un flux net (superficiel et souterrain) qui est ajouté au débit routé par le réservoir R.

L'échange inter-bassins est caractérisé par un quatrième paramètre (D), et exprimé en millimètre:

$$T(i) = D \cdot \left(\frac{R(i)}{B} \right)^{7/2} \dots\dots\dots (III-49)$$

Le paramètre D, peut prendre des valeurs négatives ou positives. Si D est négatif, le terme d'échange réduit l'écoulement pseudo-direct et le contenu du réservoir R. Dans ce cas, l'échange représente comme une perte en eau pour le bassin. Dans le cas contraire, l'échange contribue à l'écoulement, et le débit total à l'exutoire représente la somme du débit Qr, du débit Q1 et de l'échange T:

$$Qt(i) = Qr(i) + Q1(i) + T(i) \dots\dots\dots (III-50)$$

III.3.1.3. Le modèle CREC (à huit paramètres)

Le modèle CREC est un modèle à réservoirs développé initialement au laboratoire d'hydrologie de l'université de Montpellier par Cormary et Guilbot (1973). Ce modèle a été ensuite repris et modifié par plusieurs hydrologues, Drogue et al. (1982), dont la version contenait dix (10) paramètres, Servat (1986), Servat et Dezetter (1988), Servat et Dezetter (1993).

Le modèle schématise le bassin par trois réservoirs, un réservoir superficiel, un réservoir sol H, et le dernier permettant de simuler l'écoulement souterrain, il nécessite huit paramètres à caler.

Fonction de production

Elle est assurée par le réservoir superficiel S caractérisé par un seuil d'absorption (I), qui suivant que la pluie soit inférieure ou supérieure à ce seuil, on aura une infiltration ou un ruissellement direct.

$$I(i) = X8 - \sqrt{X6.S(i)} \quad (\text{III-51})$$

Avec:

S(i): Niveau d'humidité (mm) du jour précédent.

X8 et X6: Paramètres du modèle déterminants la capacité d'infiltration.

Le niveau du réservoir S est diminué sous l'effet de l'évapotranspiration réelle dont l'expression est:

$$ETR(i) = ETP(i) \cdot (1 - \exp(-\frac{S(i)}{X7})) \quad (\text{III-52})$$

X7: paramètre d'évaporation.

Si la pluie est supérieure au seuil (I), un débit de ruissellement est évacué vers l'exutoire:

$$Qs(i) = P(i) - I(i) \dots \dots \dots (\text{III-53})$$

La variation de stock d'eau de ce réservoir est définie par:

$$\Delta S(i) = PS(i) / (1 + \exp(\frac{S(i)}{X3 - X4})) \quad (\text{III-54})$$

X3 et X4 représentent les paramètres du modèle.

Fonction de transfert

Elle est assurée par les deux réservoirs de routage (H) et souterrain (G) qui sont caractérisés par des paramètres à optimiser.

Le réservoir de routage (H) est alimenté par la pluie infiltrée dans le réservoir superficiel (interception):

$$\Delta H(i) = P(i) - \Delta S(i) \quad (\text{III-55})$$

A cet instant le niveau d'humidité du réservoir augmente, il s'ensuit:

- Un débit qui est évacué directement à l'exutoire, ce débit est représenté par une relation de vidange quadratique, issue des lois de l'hydrodynamique:

$$Qh(i) = 1/X1.H(i)^2 \dots\dots\dots(\text{III-56})$$

Dans ce cas, X1 représente la constante de vidange du réservoir H, le terme $(1/X^2)$ peut être assimilé, à la conductivité hydraulique du sol.

- Une quantité d'eau (ΔG) qui est percolée (par une loi linéaire) vers le réservoir inférieur (G), ce débit est provoqué lorsque le niveau d'humidité du réservoir sol (H), est supérieur à une lame de 0.03 mm, il est en fonction du paramètre X5:

$$\Delta G(i) = \frac{H(i) - 0.03}{X5} \quad (\text{III-57})$$

Le réservoir G, simule la réserve souterraine (Figure III-10), et participe à l'écoulement, il se vidange directement dans la rivière par le débit QG:

$$Qg(i) = X2.Qg_0(i) + (1 - X2) \cdot \Delta G(i) \dots\dots\dots(\text{III-58})$$

$Qg_0(i)$ représente le débit évacué par le réservoir G au jour précédent, X2 représente un paramètre à caler ($0 \leq X2 < 1$).

Le débit total (exprimé en mm) sera donc la somme du débit de ruissellement, du débit routé par le réservoir sol, et du débit souterrain:

$$Q(i) = Qs(i) + Qh(i) + Qg(i) \dots\dots\dots(\text{III-59})$$

Deux critiques peuvent être apportées à ce modèle:

D'abord le nombre élevé de paramètres à optimiser, en effet avec huit paramètres fortement corrélés, on ne peut en aucun cas obtenir un optimum, même en utilisant un calage automatique.

Les débits des réservoirs H et S sont immédiatement évacués à l'exutoire sans effet de retard.

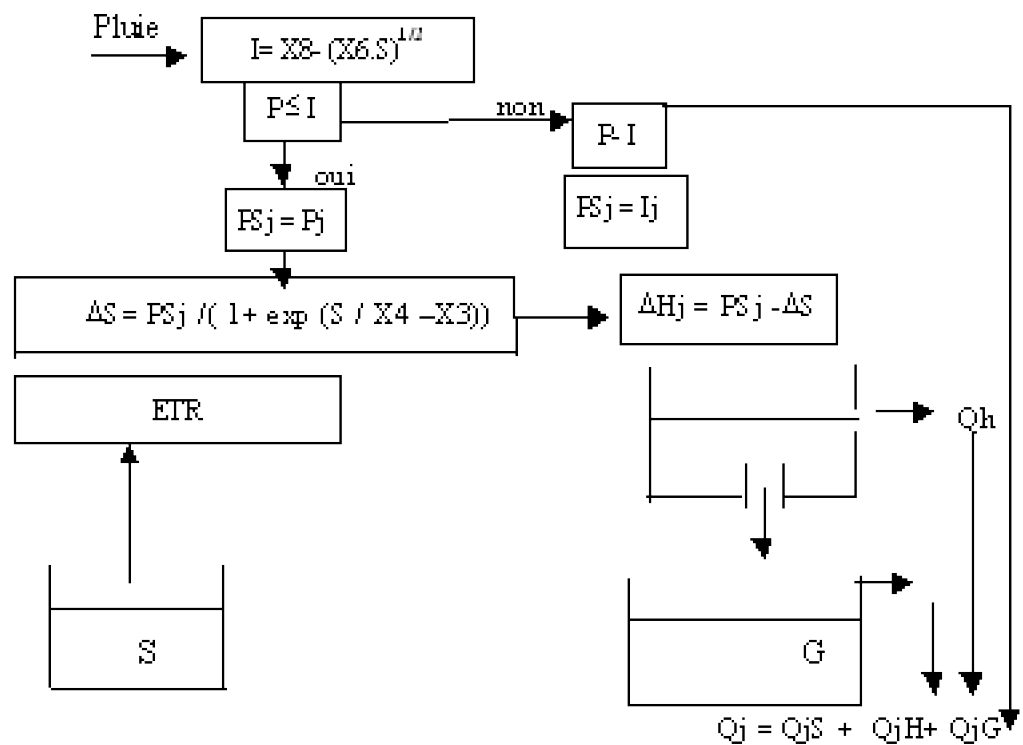


Figure III-12 Architecture du modèle CREC

III.3.1.4. Le modèle IHACRES (modifié)

Le modèle IHACRES (Identification Hydrographs And Components from Rainfall, Evaporation and Streamflow) a été développé par Jakeman et al. (1990), et a été l'objet de plusieurs recherches (Littlewood et al. 1997, Schreider et al. 1997, Perrin et al., 2001, Kokkonen, 2002).

Le modèle comporte deux modules, le premier non-linéaire servant à déterminer la pluie efficace après neutralisation avec l'évapotranspiration. Le deuxième module de type linéaire sert à calculer le débit à l'exutoire du bassin.

La pluie nette $P_n(i)$ dépend d'un indice de saturation des sols qui est calculée à chaque pas de temps à partir de la pluie brute et de l'évapotranspiration en fonction des paramètres X_2 , X_6 :

$$I_n(i) = \left[I_n(i-1) + \left(\frac{P(i)}{X_2} \right) \right] \cdot \exp\left(\frac{-ETP(i)}{X_2 \cdot X_6} \right) \quad (\text{III-60})$$

Avec $I_n(i)$ l'indice de pluie nette, $P(i)$ et $ETP(i)$ sont respectivement la pluie et l'évapotranspiration au jour (i) , X_2 et X_6 représentent les paramètres du modèle.

La pluie nette est estimée comme suit:

$$P_n(i) = P(i) \left(\frac{I_n(i) + I_n(i-1)}{2} \right) \quad (\text{III-61})$$

Cette pluie nette est ensuite répartie (en fonction du paramètre X_5) dans deux

réservoirs, entre un écoulement lent et un écoulement rapide:

$$\begin{cases} R(i) = R(i-1) + X5 \cdot Pn(i) \\ T(i) = T(i-1) + (1 - X5) \cdot Pn(i) \end{cases} \quad (\text{III-62})$$

R(i) et T(i) sont les niveaux des deux réservoirs R et T, X5 est un paramètre à optimiser ($0 < X5 < 1$).

Le débit de routage de ces réservoirs est de type linéaire, il est fonction de leur niveau d'humidité, et régit par les paramètres X1, X3. Ainsi le réservoir T libère un débit (QT) qui constitue un écoulement superficiel (Figure III-11) :

$$QT(i) = T(i) / X3 \dots \dots \dots (\text{III-63})$$

Tandis qu'un écoulement de base est provoqué par le réservoir R:

$$QR(i) = R(i) / X3 \cdot X1 \dots \dots \dots (\text{III-64})$$

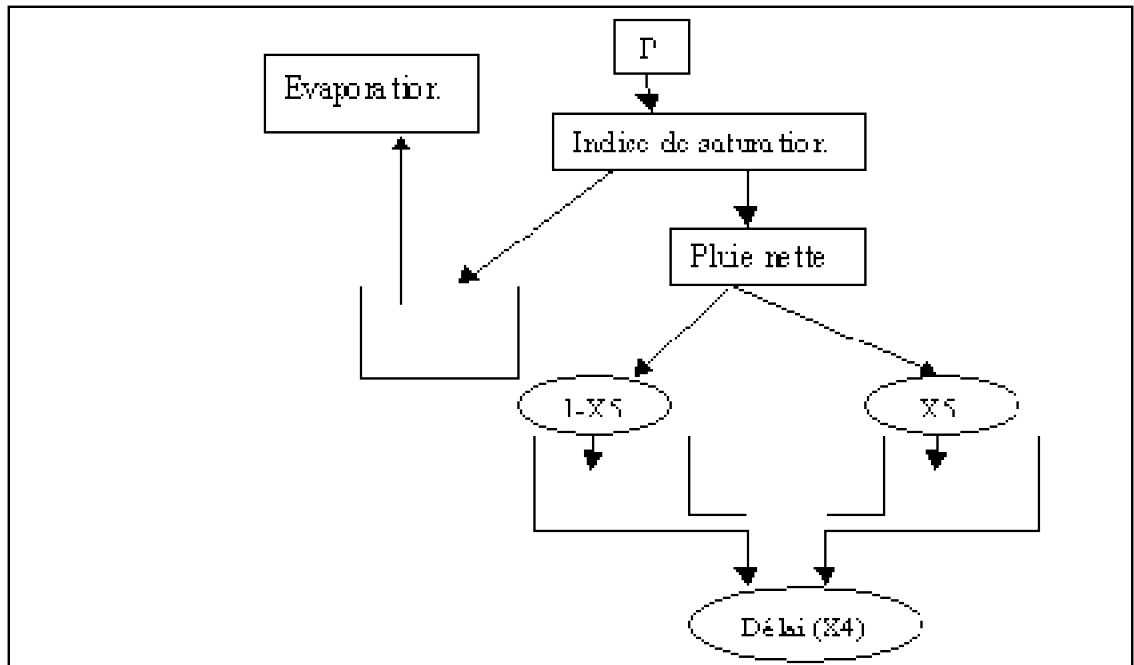


Figure III-13 Architecture du modèle IHACRES (modifié)

Le débit total calculé par le modèle, est donc la somme des deux débits routés par les deux réservoirs R et T:

$$Q(i)_{total} = QR(i) + QT(i) \dots \dots \dots (\text{III-65})$$

Les hydrogrammes calculés, sont décalés dans le temps, cet effet retard est assuré par le paramètre X4 :

$$Q(i) = Q_{total} \cdot X4 \dots \dots \dots (\text{III-66})$$

III.3.2. Résultats de la modélisation

La problématique dans l'utilisation d'un modèle pluie-débit journalier est de trouver une

durée de calage qui sera suffisamment longue pour pouvoir représenter le processus pluie-débit du bassin modélisé.

Au pas de temps journalier, la durée généralement souhaitée de ces périodes de calage et de validation est de trois ans, ce qui permet dans la majorité des cas, de couvrir des conditions hydrologiques variées.

Néanmoins, les modèles journaliers seront testés sur une période particulièrement courte, afin d'étudier ou de démontrer la robustesse de ces modèles face à des phénomènes hydrologiques tels que la succession des périodes sèches et humides (differential split sample test)

Les modèles journaliers seront testés sur:

Une période d'observation de six ans, dans ce cas les modèles seront testés sur leur capacité à maintenir un éventail d'erreur lorsque la période de validation est assez longue (trois ans et plus).

Enfin, nous ne pouvons valider ces modèles sans que ces derniers soient aptes à reproduire les débits dans des conditions extrêmes tels que les effets de sécheresse et d'humidité. Nous prendrons deux périodes distinctes, l'une sèche et l'autre humide, les modèles seront calés durant une de ces périodes particulières, et validés sur l'autre période.

Il s'agit de démontrer s'il y a lieu la capacité de robustesse de ces modèles.

Initialisation des modèles

Comme cité précédemment, l'initialisation du système est indispensable dans le cas de la modélisation conceptuelle, cette procédure permet le fonctionnement en continu d'un tel modèle, en particulier si le calage débute par une période pluvieuse où le premier débit observé est assez élevé.

Dans notre cas, pratiquement toutes les séries hydrométriques modélisées présentent des débits relativement faibles, ces derniers sont nettement inférieurs à $1.5\text{m}^3/\text{s}$. De ce fait, les niveaux des réservoirs superficiels, des réservoirs sols et des réservoirs de routage ont été initialisés par des valeurs nulles.

Cependant le modèle CREC englobe dans sa structure une simulation de la zone saturée (nappe), le niveau de ce réservoir est rarement en tarissement en particulier pour les mois pluvieux.

Dans ce cas, le niveau du réservoir souterrain a été initialisé après des tests de simulations sur les trois bassins à une valeur de 200mm.

III.3.2.1. Simulation en période longue

Les modèles sont testés sur les phases de calage et de validation de durées égales à trois (3) ans, soit une période d'observations de six ans.

Les performances de ces modèles, sont démontrées par trois critères importants à savoir, le critère de Nash, l'indice de reproduction du débit maximal (I), et le rapport des cumuls d'étiage (CET).

Les résultats sont présentés dans le tableau III-4 pour la phase de calage, et dans le tableau III-5 pour celle de la validation.

Stations	Cheffia (1986-88)			Isser (1990-92)			Zardezas (1991-93)		
	Nash (%)	Ic	CET (%)	Nash (%)	Ic	CET (%)	Nash (%)	Ic	CET (%)
GR3J	74.7	0.75	573	76.2	0.87	270	83.0	0.91	185
GR4J	74.0	0.58	517	67.2	0.65	472	85.6	1.02	327
CREC	70.3	0.60	323	51.3	0.50	163	82.4	0.98	363
IHACRES	71.0	0.48	713	44.8	0.36	429	85.4	0.93	275

Tableau III-4 Résultats des modèles journaliers de la période de calage

Stations	Cheffia (1989-91)			Isser (1993-95)			Zardezas (1994-96)		
	Nash (%)	Ic	CET (%)	Nash (%)	Ic	CET (%)	Nash (%)	Ic	CET (%)
GR3J	81.7	0.91	751	73.7	0.86	188	59.9	0.59	136
GR4J	69.0	0.52	877	50.1	0.50	445	71.6	0.59	174
CREC	66.1	0.50	400	68.4	1.10	167	71.5	0.51	214
IHACRES	64.1	0.60	856	54.1	0.46	400	76.6	0.54	168

Tableau III-5 Résultats des modèles journaliers de la période de validation

D'après les tableaux III-5 et III-6, nous constatons que les résultats des simulations varient d'un modèle à l'autre et d'une station à l'autre, la valeur maximale est réalisée par le modèle IHACRES avec 85.6% pour la station de Zardezas en période de calage.

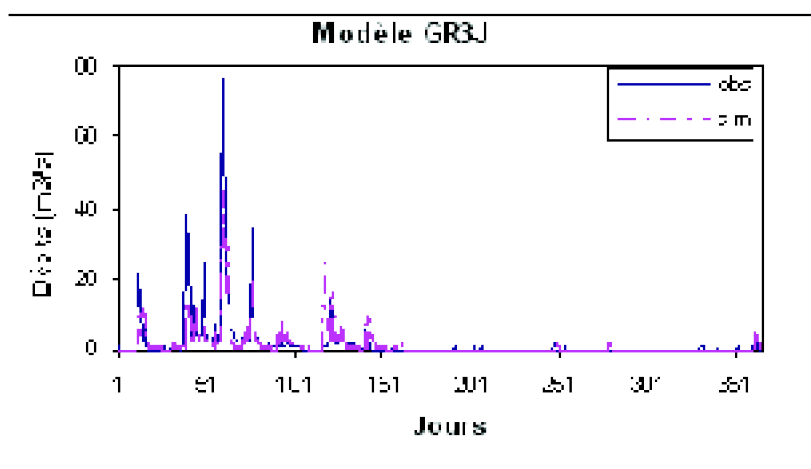
Nous remarquons que pour l'ensemble des modèles, les résultats de la période de calage des différents bassins sont nettement meilleurs que ceux de la période de validation. En effet pour la station de Zardezas, les résultats des modèles GR3J et GR4J de la période de calage qui sont respectivement de 83.0 et 85.6%, se détériorent lorsqu'on passe à la phase de validation soient respectivement 59.9 et 71.6 %, ce qui

représente un écart moyen de -18.5% .

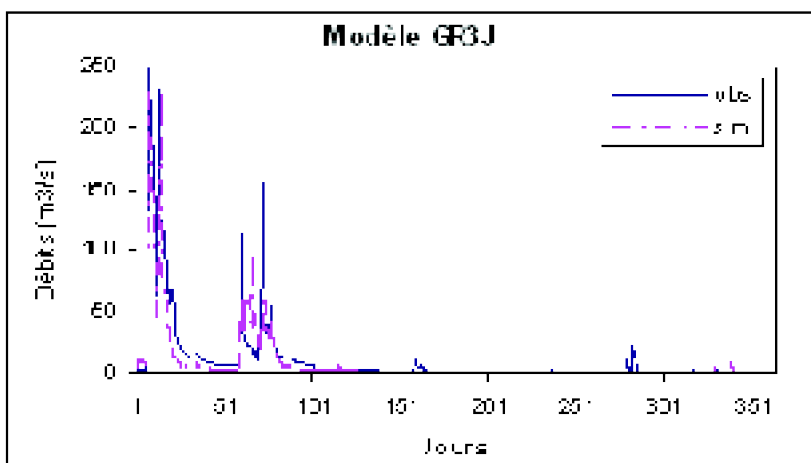
La baisse des résultats est plus importante pour la station de l'Isser, où le modèle enregistre une diminution des valeurs de l'efficience de Nash, et cela même pour la période de calage, avec des résultats nettement moins performants que ceux des autres stations, que ce soient en période de calage ou de validation, le modèle IHACRES donne une valeur très médiocre de moins de 45.0% en phase de calage, et moins de 50.0% pour le modèle GR4J.

Ceci confirme la difficulté de simuler ce bassin, qui est caractérisé par la présence d'averses locales.

Cette baisse significative de performance dans la phase de validation, est remarquée pour tous les modèles.



(a)

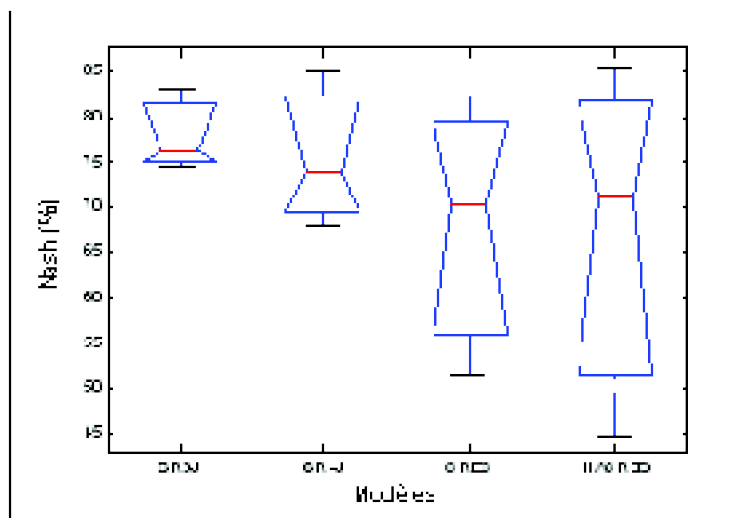


(b)

Figure III-14 Simulations par le modèle GR3J

(a) station de Zardezas (année 1996), (b) station de l'Isser (année 1995)

Les résultats des modèles (analyse de variance) sont présentés sous forme graphique Figure III-13, en période de calage (a) et de validation (b). Le modèle GR3J (1999) paraît le plus robuste des modèles conceptuels testés.



(a)

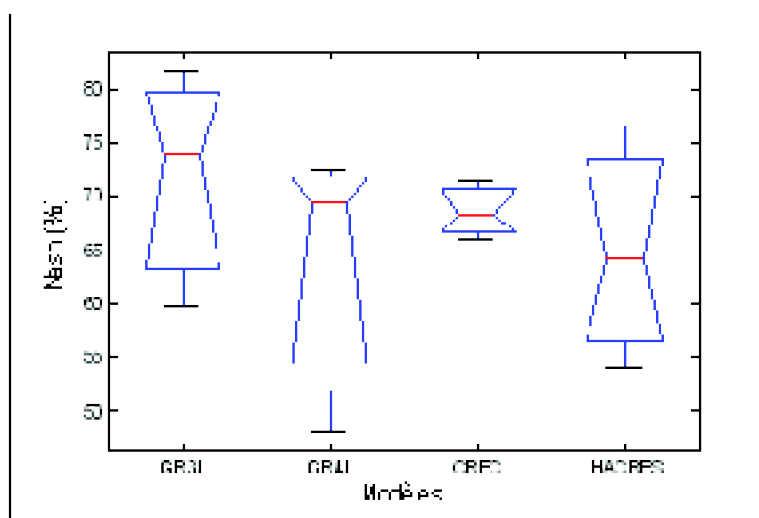


Figure III-15 Performances des modèles

(a) phase de calage, (b) phase de validation

Le dimensionnement de certains ouvrages hydrauliques nécessite la connaissance du volume maximal que doit évacuer cet ouvrage. De ce fait la reproduction du débit maximal par le modèle est indispensable, et représente un critère d'appréciation important.

L'indice I_c , qui mesure la justesse de reproduction du débit maximal pour les deux périodes, indique que pratiquement tous les modèles conceptuels testés n'arrivent pas à simuler les valeurs maximales de la série hydrométrique observée.

En effet, les valeurs de cet indice sont inférieures à l'unité, ce qui explique une

mauvaise reproduction du débit maximal, en particulier dans la phase de validation.

Il est à noter une forte corrélation entre la simulation des valeurs élevées et les performances d'un modèle, en effet on remarque généralement une efficacité de Nash relativement élevée lorsque le débit maximal observé est fidèlement reproduit par le modèle, et vice versa.

Ceci, est de même pour les débits d'étiage, qui sont largement surestimés par le modèle. En effet, nous remarquons généralement des rapports des cumuls d'étiage largement supérieurs à cent.

III.3.2.2. Simulation en période courte (Differential split sample test)

Cette simulation comporte deux procédures, la première consiste à simuler le processus pluie-débit durant une période d'observation humide dont la durée est d'une année, et de tester le modèle sur une autre période sèche ou moins arrosée, la seconde procédure, permet de caler le modèle sur une période sèche, et de le vérifier sur la période humide, sachant que dans cette simulation les périodes sont de durées homogènes et ne dépassant pas deux ans.

C'est le cas du bassin du Sebbaou, où l'année 1993 présente un débit moyen très élevé de $9.0 \text{ m}^3/\text{s}$, cette période humide est suivie de l'année 1994 moins humide, le débit moyen journalier ne dépasse pas $6.0 \text{ m}^3/\text{s}$.

Les modèles seront donc calibrés et testés sur une période d'une année.

Les durées et les caractéristiques hydrométriques de cette période sont présentées comme suit:

Tableau III-6 Caractéristiques de la série du bassin du Sebbaou

Caractéristiques	Période de calage (1993)	Période de validation (1994)
Moyenne (m^3/s)	9.00	5.97
Ecart-type (m^3/s)	16.96	17.55
Débit maximal journalier (m^3/s)	157.0	232.0

Les résultats de la modélisation des différents modèles sont présentés dans le tableau III-7

Périodes	Calage (1993)			Validation (1994)		
	Nash (%)	Ic	CET (%)	Nash (%)	Ic	CET (%)
GR3J	81.0	0.60	1500	61.6	0.35	760
GR4J	75.0	0.65	1280	54.7	0.30	700
CREC	74.3	0.53	1374	49.5	0.26	898
IHCARES	70.7	0.44	1100	45.9	0.22	710

Tableau III-7 Résultats de la modélisation (Station Sebbaou)

Les résultats présentés dans le tableau III-7, montrent que les modèles donnent des résultats différents, le modèle GR3J donne une efficacité de Nash la plus importante en période de calage (81.0%), les autres modèles présentent des performances moins élevées mais relativement similaires. En période de validation, tous les modèles présentent une baisse de la qualité de simulation, avec des critères de Nash de moins de 50.0% pour les modèles CREC et IHCARES.

Les débits caractéristiques sont mal reproduits par les modèles. Pour les deux périodes on note une sous-estimation du débit maximal, l'indice *Ic* est très médiocre pour le modèle IHCARES, qui calcule les débits à partir d'une loi de vidange linéaire.

Les débits d'étiages sont largement supérieurs à ceux observés, en particulier pour la phase de calage.

La deuxième tentative de simulation d'une période caractéristique, est celle du bassin du Zardezas, dont nous remarquons durant la période d'observation, une succession d'une phase sèche durant les années 1993-1994, suivie d'une période relativement pluvieuse, représentée par les années 1995-1996.

De ce fait, la simulation de cette période se fera en calibrant le modèle pour les années où le débit journalier moyen est relativement faible (années 1993/94), et de tester le modèle (phase de validation) durant les années (1995/96) où le débit moyen est plus élevé, afin d'en tirer les conclusions nécessaires concernant la robustesse des modèles conceptuels durant ces phases.

Les caractéristiques hydrométriques de cette série sont présentées comme suit :

Tableau III-8 Caractéristiques de la série du bassin du Zardezas

MODELISATION PLUIE-DEBIT MENSUELLE ET JOURNALIERE PAR LES MODELES CONCEPTUELS ET LES SYSTEMES NEURO-FLOUS (APPLICATION AUX BASSINS ALGERIENS)

Caractéristiques	Période de calage (1993-94)	Période de validation (1995-96)
Moyenne (m ³ /s)	1.17	2.20
Ecart-type (m ³ /s)	6.2	8.45
Débit maximal journalier (m ³ /s)	124.3	108.5

Les résultats des différents modèles appliqués à la série courte de la station de Zardezas sont résumés dans le tableau III-9.

Périodes	Calage (1993-94)			Validation (1995-1996)		
	Nash (%)	Ic	CET (%)	Nash (%)	Ic	CET (%)
GR3J	86.1	0.75	361	57.0	0.57	218
GR4J	80.1	0.75	276	75.6	0.64	196
CREC	85.3	1.06	245	77.7	1.06	284
IHACRES	81.5	0.67	293	78.9	0.56	210

Tableau III-9 Résultats de la modélisation (Station Zardezas)

Les résultats du tableau III-9, montrent des performances des modèles en période de calage relativement similaires, pour les modèles GR3J et CREC où l'efficacité de Nash est supérieure à 85%. En phase de validation, on note des valeurs du critère de Nash relativement faibles, on enregistre un écart important pour le modèle GR3J entre les deux périodes, ces résultats confirment la baisse des performances des modèles lorsque les caractéristiques climatiques varient entre les deux phases.

Néanmoins, la baisse du niveau de simulation en phase de validation pour l'ensemble des modèles est moins importante par rapport à la série modélisée du Sebaou. Cela est dû essentiellement aux débits assez fluctuants de ce bassin. En effet, en passant à une période plus humide caractérisée par de fortes crues, l'effet seuil des modèles (capacité maximale du réservoir de routage) limite considérablement la reproduction des débits de pointe.

Cependant, cette conclusion devrait être confirmée et étudiée sur un large éventail de bassins versants.

Nous pouvons affirmer, la robustesse des modèles GR par rapport aux autres modèles étudiés, qui avec un nombre de paramètres réduit, ont donné des résultats assez performants en période de calage.

III.4. Conclusion

Dans ce chapitre, nous avons exploité la modélisation pluie-débit à deux échelles mensuelle et journalière.

En comparant quatre (4) modèles mensuels les plus utilisés, avec un nombre de paramètres qui varie de deux à quatre, à savoir le modèle de Thornthwaite, le modèle ABCD, et deux versions des modèles GR2M (1994, 2002), nous avons démontré la robustesse de ces deux derniers modèles, et leurs performances à simuler les débits mensuels à partir des variables hydrométéorologiques (Pluie, ETP), et cela en dépit de la complexité de la relation pluie-débit à cette échelle.

Ces modèles ont donné des résultats relativement satisfaisants en phase de calage, néanmoins, en période de validation, les résultats sont généralement moins performants en particulier pour certaines stations (station de Tafna), où les écarts entre les débits observés et simulés deviennent importants pour utiliser un modèle conceptuel dans l'estimation des apports, ou dans la reconstitution des séries hydrométriques tronquées.

La modélisation pluie-débit au pas de temps journalier, a été étudiée par le biais de quatre modèles, connus pour leur robustesse, à savoir le modèle GR3J, le modèle GR4J, le modèle CREC à huit paramètres, et le modèle IHACRES.

Les modèles testés dans trois bassins versants algériens pour une période longue de six ans ont donné des résultats satisfaisants, en particulier en phase de calage où l'efficacité de Nash est généralement assez correcte.

Cependant, vu que la validation d'un modèle pluie-débit est tributaire de ses performances en phase de validation, on note une baisse significative de la qualité de simulation de ces modèles dans cette phase. Certains modèles (modèle CREC) se détériorent lorsqu'on passe à la phase de validation, et sont très sensibles aux variations climatiques.

Nous avons noté la difficulté de ces modèles à reproduire les débits extrêmes (débits maxima et étiages) observés dans toute la période de simulation, ce qui limite leur utilisation dans la prédétermination des débits caractéristiques.

Malgré un effort considérable apporté dans l'amélioration des modèles pluie-débit conceptuels, nous ne pouvons dire que les tâches d'optimisation de ces modèles soient accomplies.

Chapitre IV : Etude de la relation pluie-débit par les réseaux de neurones artificiels

IV.1. Introduction

Le développement de l'informatique au milieu du vingtième siècle a provoqué une révolution dans les procédés de calculs numériques, et a ouvert la voie à une nouvelle branche scientifique dédiée au traitement automatique de l'information.

Néanmoins avec le développement des ordinateurs et l'augmentation de leurs puissances, et des approches mathématiques théoriques, certaines opérations ne trouvent pas de solutions, et se heurtent à des difficultés, tels que le grand nombre de données, leur nature et leur variabilité, ainsi que la nécessité des calculs en temps réel.

Les programmes informatiques conçus avaient pour objectif de résoudre la complexité des calculs numériques, et du temps de leur exécution. Ainsi on voit naître la première approche du traitement informatique, l'approche algorithmique.

Fondée sur le développement récent de la recherche opérationnelle, cette approche

qui nécessite l'écriture et la transcription dans un langage de programmation du processus à suivre pour résoudre le problème. Cependant, la nature très complexe de certains phénomènes (physiques ou autres) ne permet pas à cette technique d'aboutir devant une situation imprévue ou inadaptée.

Face à ces problèmes, les chercheurs se sont consacrés à de multiples approches, théoriques et opérationnelles dans un but essentiel de donner à l'ordinateur des compétences de raisonnement proches des compétences humaines. Cette faculté de mémoriser et de réutiliser des expériences permet à l'ordinateur d'atteindre un stade complet d'autonomie. C'est dans ce contexte que les chercheurs se sont investis dans l'intelligence artificielle, en vue d'améliorer les capacités de résolution d'un système informatique, est de lui inclure des connaissances sur le domaine qu'il doit gérer.

L'approche de l'intelligence artificielle, (ou plus exactement basée sur des expériences), paraît donc indispensable avec pour applications les plus connues les systèmes experts. L'hypothèse principale, à la base de l'essor des réseaux de neurones artificiels, est que le comportement intelligent est sous-tendu par un ensemble de mécanismes mentaux. Ces mécanismes étant basés sur des processus neurophysiologiques, nous supposons donc que la structure du système nerveux central est à la base du développement d'un comportement intelligent.

IV.2. Historique des réseaux de neurones artificiels

Le principe des réseaux de neurones artificiels est né dans les années 1940 à partir d'une analogie avec le système nerveux humain. En 1943 deux chercheurs Mc Culloch et Pitts ont inventé le premier "neurone artificiel", qui correspondait à un produit d'un vecteur d'entrées, et un vecteur poids suivi d'un élément de seuillage. Ces deux chercheurs ont montré que ces réseaux de neurones formels peuvent réaliser des fonctions logiques, arithmétiques et symboliques complexes.

Cette modélisation du comportement du neurone humain a ouvert la voie à une nouvelle recherche, puisqu'en 1949, le physiologiste américain D. Hebb expliqua le conditionnement chez l'animal par les propriétés des neurones humains et proposa une théorie fondamentale de l'apprentissage.

Les travaux de recherches avaient pour objectif de s'approcher le plus possible du fonctionnement du cerveau humain, afin de bénéficier de plusieurs de ses fascinantes caractéristiques.

Le concept d'apprentissage par adaptation des poids synaptiques des réseaux a vu le jour, puisque Minsky (1951) construit ce qui fut le premier modèle réel de réseaux de neurones baptisé Snarl.

Plusieurs recherches ont été entamées dans un contexte de modèles d'apprentissage, en 1957 Rosenblatt fut le premier à développer un modèle neuronal de classifications appelé Perceptron. Le Perceptron avait comme originalité de modifier ses poids en fonction des exemples que l'on veut lui faire apprendre. Trois ans plus tard,

Widrow et Hoff (1960) développèrent un nouveau modèle neuronal, l'Adaline (Adaptative Linear Neurone). Ce modèle avait la même structure que le Perceptron, mais intègre un algorithme destiné à l'approximation de fonctions. En 1969, Minsky et Papert publièrent un ouvrage critique qui met en exergue les limitations théoriques du Perceptron, concernant l'impossibilité de traiter des problèmes non linéaires.

A partir des années quatre-vingt, avec le développement des mathématiques et de la physique, les recherches sur les réseaux de neurones s'intensifient, ainsi le physicien Hopfield (1982), puis Grossberg (1987), ont pu résoudre pour la première fois le problème de stabilité des réseaux dynamiques.

En 1986, une équipe de chercheurs dirigée par Rumelhart, a pu concevoir un nouvel algorithme d'apprentissage, constitué d'un ensemble de couches (appelé Perceptron multicouches). Sa principale caractéristique est l'approximation de fonction non-linéaire, par la rétropropagation de gradient (ou Backpropagation). Cette découverte, a permis de réaliser une fonction non linéaire d'entrée/sortie, et a ouvert une nouvelle ère pour les réseaux de neurones.

Actuellement, les réseaux de neurones disposent de fondements théoriques solides et constituent donc un ensemble d'opérateurs linéaires et non-linéaires, qui permettent grâce à leur adaptabilité une large utilisation dans différentes disciplines scientifiques.

En hydrologie connaissant la non-linéarité de la relation pluie-débit, certains auteurs ont exploité les réseaux de neurones artificiels dans la modélisation pluie-débit à différentes échelles (French *et al.*, 1994, Shamseldin, 1996).

IV.3. Aspects théoriques et généralités sur les réseaux de neurones

Le développement des connaissances de la structure interne du cerveau, et la compréhension d'une partie de son fonctionnement a contribué de façon remarquable à l'émergence des réseaux de neurones artificiels.

IV.3.1. Le neurone biologique

Le neurone est l'unité fonctionnelle de base du système nerveux, et peut présenter certaines différences fonctionnelles selon les espèces.

Du point de vue anatomique, le neurone est une cellule composée d'un corps cellulaire et d'un noyau. Le corps cellulaire se ramifie pour former les dendrites. C'est par les dendrites que les informations sont acheminées de l'extérieur vers le soma, qui représente le corps du neurone (Figure IV-1).

Chaque neurone est une unité autonome au sein du cerveau. Le neurone reçoit en continu des entrées (informations). La réponse, unique, à ces signaux est envoyée au travers de l'axone. L'information réside, principalement dans l'architecture des

connexions, et dans la force de ces connexions (Touzet, 1992)

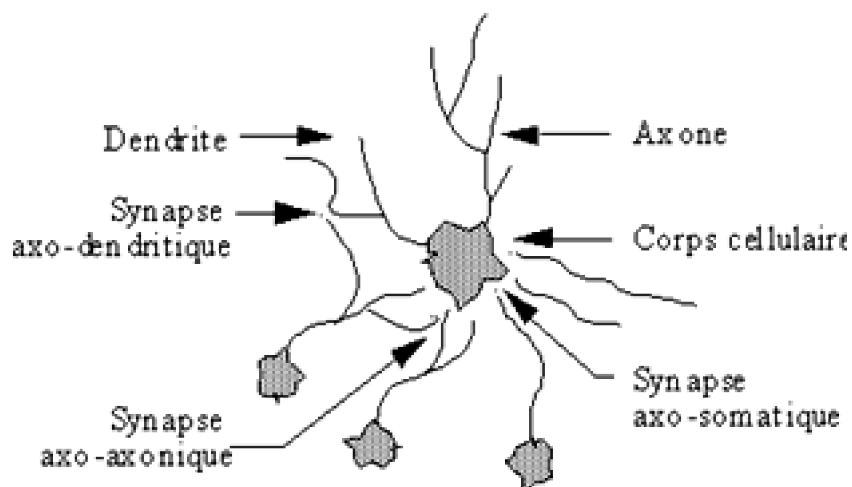


Figure IV-1 Schéma simplifié du neurone

IV.3.2. Le neurone formel (artificiel)

Développé par Mc Culloch et Pitts (1943), le neurone formel est essentiellement constitué d'un intégrateur qui effectue la somme pondérée de ses entrées.

Chaque neurone artificiel reçoit un nombre variable d'entrées en provenance de neurone amont. A chacune de ces entrées est associé un poids (w) représentatif de la force de la connexion (Figure IV-2)

Un neurone est donc une fonction algébrique, non-linéaire paramétrée, à valeurs bornées (Dreyfus et al., 2004).

Les variables sur lesquelles opèrent les réseaux sont habituellement désignées sous le terme du neurone, et la valeur de la fonction sous celui de sortie.

Un poids d'un neurone artificiel représente donc l'efficacité d'une connexion synaptique. Un poids négatif vient inhiber une entrée, alors qu'un poids positif vient l'accentuer.

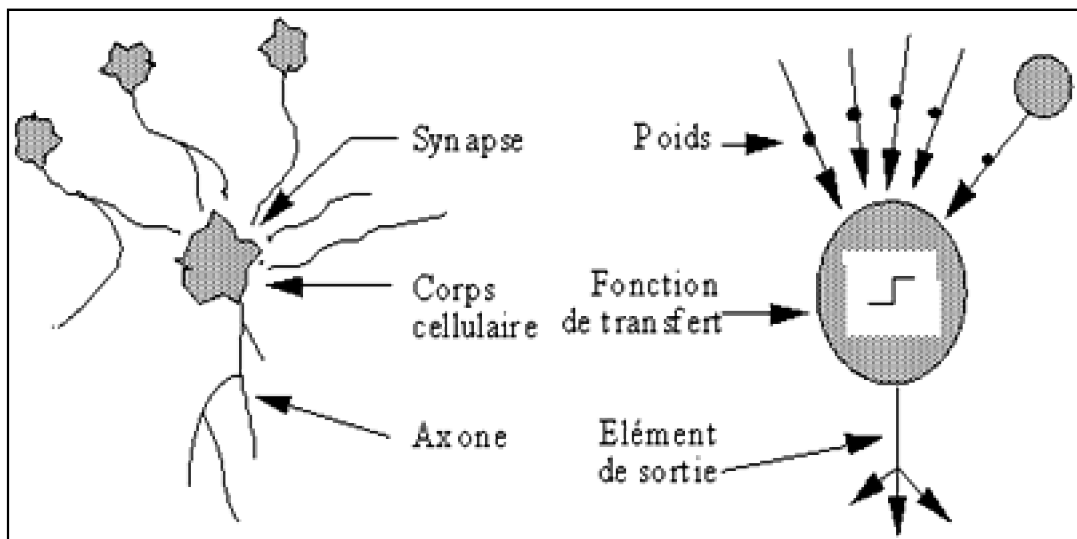


Figure IV-2 Mise en correspondance neurone biologique / neurone artificiel

IV.3.3. Fonctionnement mathématique du neurone artificiel

Il est régi par trois fonctions:

- *La fonction discriminant* (fonction de base), elle définit l'activité du neurone, la plus utilisée est la fonction linéaire, elle est définie par:

$$\varphi(Z,W) = ZW^T - b \dots\dots\dots(IV.1)$$

Où W représente la matrice des poids, Z sont les entrées qui viennent de l'extérieur ou des autres neurones vers le neurone en question, b représente le biais.

La sortie des neurones se fait par le calcul de la somme pondérée des entrées avec des coefficients appelés poids auxquels on retranche le biais.

- *La fonction dynamique linéaire*, elle détermine la dynamique du réseau.
- *La fonction dynamique*, c'est la fonction dite d'activation qui a pour objectif de rendre l'activité du neurone bornée. Cette fonction doit être non-linéaire, continue dérivable et monotone (Yeddou, 1998). La fonction d'activation (transfert) calcule la valeur de l'état du neurone. C'est cette valeur qui sera transmise aux neurones avals.

Parmi les fonctions les plus utilisées, on retrouve la fonction «seuil» (Hard limit) la fonction linéaire et la fonction sigmoïde.

La fonction seuil applique un seuil sur son entrée. Une entrée négative ne passe pas le seuil, alors retourne dans ce cas la valeur zéro (qui peut être interprétée comme signifiant faux) alors qu'une entrée positive ou nulle dépasse le seuil et retourne la valeur 1 (vrai). La fonction linéaire affecte directement son entrée à sa sortie. Dans certains cas, il est recommandé d'utiliser pour *f* une fonction "sigmoïde" (c'est-à-dire une fonction en forme de "s" symétrique par rapport à l'origine.

IV.4. Structure générale des réseaux de neurones

Un réseau de neurones est un maillage de plusieurs neurones, on distingue deux types de réseaux (Figure IV-3):

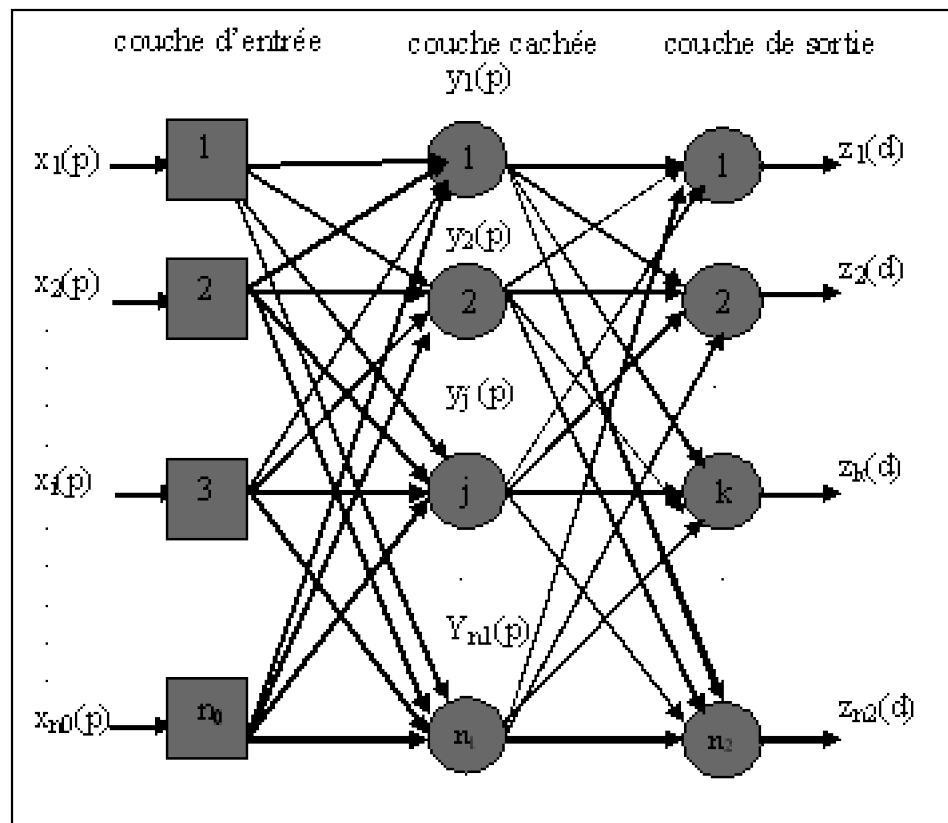


Figure IV-3 Structure d'un réseau multicouche

Les réseaux de neurones monocouche: tels que le Perceptron et l'Adaline, ces réseaux sont exploités essentiellement dans la classification des données.

Les réseaux de neurones multicouche: développés par Minsky et Papert (1969) (Rumelhart et al. 1986) sont plus puissants que les réseaux monocouche. Les réseaux multicouche peuvent constituer une solution aux problèmes non-linéaires, et cela par l'introduction de fonctions non-linéaires et de couches cachées.

IV.4.1. Classification des réseaux de neurones artificiels

Vu la multitude des modèles des réseaux neuronaux, nous pouvons classer ces réseaux en fonction de leur architecture et de leur apprentissage.

IV.4.1.1. Architecture des réseaux de neurones

Elle définit le fonctionnement du réseau, on distingue:

- Les réseaux non bouclés (statiques) qui sont organisés en couches de neurones interconnectés entre eux (Figure IV.3), et dont l'information à l'intérieur du réseau est sans retour en arrière (Feed-back).
- Les réseaux bouclés ou récurrents (dynamiques), dont les connexions récurrentes ramènent l'information en arrière par rapport au sens de propagation défini dans un réseau multicouche (figure IV-4).

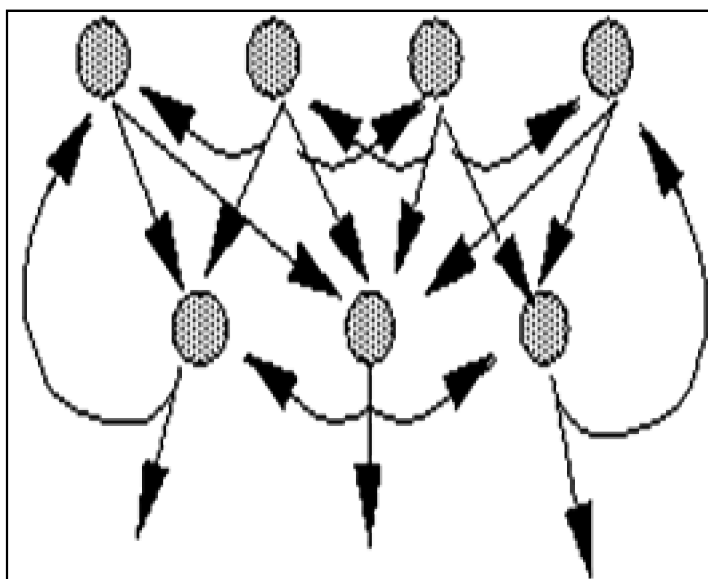


Figure IV-4 Réseau de neurones à connexions récurrentes

IV.4.1.2. Apprentissage des réseaux de neurones

L'apprentissage est une procédure adaptative pour laquelle les connexions (synapses) des neurones sont ajustées face à une source d'informations (Rumelhart *et al.*, 1986).

L'apprentissage est défini comme étant n'importe quel changement opéré dans la mémoire du réseau. Ainsi cette modification affecte les paramètres (poids synaptiques) qui relient les neurones entre eux (Yeddou, 1998).

Il existe plusieurs règles d'apprentissage qui peuvent être regroupées en trois grandes catégories (Coulibaly *et al.*, 1999, Dreyfus *et al.*, 2004) :

L'apprentissage supervisé: dans ce cas on fournit au réseau la donnée à traiter mais aussi la réponse attendue. Le réseau effectue une évaluation de la donnée, puis compare la valeur obtenue avec la valeur désirée. La procédure s'effectue progressivement, jusqu'à ce que l'erreur entre les sorties du réseau et les résultats désirés (données observées) soit minimisée.

L'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. L'objectif de cet apprentissage est de partager l'espace d'entrées en

plusieurs entrées (classes) représentant des régions de décisions différentes.

L'apprentissage semi-supervisé ou par renforcement (apprentissage par Récompense et Pénalité) : représente une technique nouvelle, dans ce cas on fournit des exemples au réseau et des indications sur le résultat par un superviseur qui est régi par un de degrés de performance (vrai ou faux).

IV.5. Apprentissage supervisé des réseaux de neurones artificiels

Au cours des dernières années, le développement des réseaux de neurones à apprentissage supervisé a été le plus marquant, vu que cette classe d'apprentissage présente des propriétés fondamentales telle que l'approximation universelle, et le développement des algorithmes plus performants leur a ouvert de nouvelles perspectives d'utilisation.

IV.5.1. Propriétés des réseaux de neurones à apprentissage supervisé

L'apprentissage est vraisemblablement la propriété la plus intéressante des réseaux neuronaux.

Définition

L'apprentissage est un processus dynamique et itératif permettant de modifier les paramètres du réseau en fonction des entrées qu'il reçoit, jusqu'à l'obtention du comportement désiré.

Dans le cas des réseaux de neurones artificiels, on ajoute souvent à la description du modèle l'algorithme d'apprentissage. Dans la majorité des algorithmes actuels, les variables modifiées pendant l'apprentissage sont les poids des connexions. Il est souvent impossible de décider a priori des valeurs des poids des connexions d'un réseau pour une application donnée. A l'issue de l'apprentissage, les poids sont fixés : c'est alors la phase d'utilisation.

L'apprentissage supervisé consiste à calculer les coefficients de telle manière que les sorties du réseau de neurones soient, pour les exemples utilisés lors de l'apprentissage aussi proches que possibles des sorties désirées, qui peuvent être la classe d'appartenance de la forme que l'on veut classer, ou encore la sortie souhaitée du processus à commander.

IV.5.2. Architecture et apprentissage des réseaux multicouche

Un réseau de neurones multicouche est formé d'un assemblage d'éléments (neurones), chacun d'eux est lié à d'autres neurones. Ceux-ci sont disposés dans une série de couches, la première couche (couche d'entrée), la (ou les) seconde(s) couche(s)

intermédiaire(s) (appelées couches cachées), et la dernière appelée couche de sortie.

L'entrée du neurone est transformée en sortie (dans notre cas débit simulé), chaque neurone reçoit des entrées sous forme vectorielle (X), effectue une somme pondérée (α), et génère à l'aide d'une fonction de transfert (f) un résultat réel (Y), cet argument est appelé potentiel du neurone:

$$Y = f(Y_{net}-b) \dots \dots \dots (IV.2)$$

$$Y_{net} = \sum (W_{i_n} \cdot X_{i_n}) \dots \dots \dots (IV.3)$$

Y = représente la sortie du neurone;

$W = (w_{i1}, w_{i2} \dots, w_{in})$ représente les poids du neurone;

$X = (x_{i1}, x_{i2} \dots, x_{in})$ sont les entrées du neurone i (ou vecteur d'entrée);

b = est le biais du neurone ou valeur seuil.

Il est à noter, que la fonction tangente hyperbolique est une version symétrique de la sigmoïde, elle est donnée par:

$$f = \frac{1}{1 + e^{-Y_{net}}} \quad (IV.4)$$

IV.5.3. Réseaux de neurones à base linéaire et à base radiale

La fonction f définie dans (IV-2) peut être paramétrée de manière quelconque. Deux types de paramétrages sont fréquemment utilisés :

- Les paramètres sont attachés aux entrées du neurone : la sortie du neurone est une fonction non linéaire d'une combinaison des entrées $\{x_i\}$ pondérées par les paramètres $\{w\}$, qui sont alors souvent désignés sous le nom de "poids" ou, en raison de l'inspiration biologique des réseaux de neurones, «poids synaptiques». Conformément à l'usage (également inspiré par la biologie), cette combinaison linéaire sera appelée "potentiel". Le potentiel Y le plus fréquemment utilisé est la somme pondérée, à laquelle s'ajoute le "biais".

Ces modèles forment des réseaux de neurones appelés à base linéaire, et sont les plus répandus.

- Les paramètres sont attachés à la non-linéarité du neurone: ils interviennent directement dans la fonction f ; cette dernière peut être une *fonction radiale* ou RBF (*Radial Basis Function*), ou encore une ondelette ; les premières tirent leur origine de la théorie de l'approximation (Powell, 1987), les secondes de la théorie du signal (Mallat, 1989).

Pour exemple, la sortie d'un neurone RBF à non linéarité gaussienne a pour équation:

$$Y = \text{Exp} \left[\frac{\sum_{i=1}^n (x_i - w_i)^2}{2 w_{n+1}^2} \right] \quad (\text{IV.5})$$

Où les paramètres w_i sont les coordonnées du centre de la gaussienne, et w_{n+1} son écart-type.

La différence pratique essentielle entre les deux types de neurones que nous venons d'introduire est la suivante : les neurones tels que les RBF ou les ondelettes ont des non-linéarités *locales*, qui tendent vers zéro dans toutes les directions de l'espace des entrées; leur zone d'influence est donc limitée dans l'espace, ce qui n'est pas le cas des neurones à potentiel à fonction d'activation sigmoïde (à base linéaire) (Dreyfus *et al.*, 2004).

IV.5.4. Algorithmes d'apprentissage des réseaux à base linéaire

La plupart des algorithmes d'apprentissage supervisé des réseaux de neurones artificiels sont des algorithmes d'optimisation, ils cherchent à minimiser de manière itérative, une fonction de coût qui constitue une mesure de l'écart entre les réponses réelles du réseau et ses réponses désirées.

On peut cependant classer les réseaux à apprentissage supervisé en deux classes:

Les réseaux basés sur la décision, et les réseaux basés sur l'optimisation.

IV.5.4.1. Les réseaux de neurones basés sur la décision

Le Perceptron et l'Adaline (Perceptron simple)

Ce type de réseau effectue la séparation entre deux zones de l'espace. La position de l'hyperplan séparant ces deux régions, constitue les contours de décision. L'Adaline (ADAPtative LInear NEurone) est une extension du Perceptron (simple), l'idée est de remplacer la fonction seuil sur les neurones de sortie par la fonction identité ($f(x)=x$). L'avantage de cette approche est de ne plus être restreinte à une réponse binaire, et d'offrir tout l'éventail de valeurs possibles.

Widrow (1960) a proposé une solution qui évite l'estimation de l'espérance. Pour cela l'algorithme agit à la présentation de chaque exemple en entrée. Ainsi, l'algorithme estime la trajectoire à prendre vers le point minimum à chaque présentation d'un vecteur d'entrée, c'est la méthode de descente de gradient.

Méthode de Descente de Gradient:

Cette méthode consiste à déterminer les adaptations qu'il faut apporter aux vecteurs des poids synaptiques à chaque présentation d'une entrée, afin de faire décroître l'erreur de sortie jusqu'à obtention de la solution optimale.

L'algorithme proposé par Widrow, permet d'estimer la trajectoire à prendre vers le point minimum à chaque présentation d'entrée. Ainsi à chaque vecteur présenté, on enregistre une surface d'erreur, dont on estime le gradient en fonction des poids synaptiques, et on modifie ces poids pour faire décroître l'erreur: c'est la méthode du Delta. De cette manière, le gradient ne pointera que le minimum instantané à chaque étape.

L'erreur quadratique instantanée à chaque étape:

$$E^k = \frac{1}{2} (e^k)^2 = \frac{1}{2} (d^k - y^k)^2 \quad (IV.6)$$

Avec E^k est la fonction quadratique, e^k est la fonction objectif, d et y sont respectivement les valeurs observées et simulées par le réseau de neurones.

Afin de progresser dans le sens opposé du gradient, on a:

$$W_{ij}^{k-1} = W_{ij}^k + \eta_k (d_i^k - y_i^k) x_j^k \quad (IV.7)$$

Où W_{ij} sont les poids ou coefficients, η_k est un paramètre appelé taux d'apprentissage, dont le rôle est de réguler la vitesse de convergence et contrôler la stabilité de ce processus.

Cependant, les réseaux de neurones basés sur la décision représentent un approximateur linéaire, et sont donc limités aux fonctions linéaires.

IV.6. Réseaux de neurones multicouche (Perceptron multicouche)

Vers le début des années quatre vingt, avec l'essor qu'a connu les réseaux de neurones artificiels notamment par la vulgarisation de l'algorithme de rétrogradation de l'erreur (Backpropagation) (Rumelhart *et al.*, 1986), ce modèle constitué de plusieurs couches (appelé Perceptron multicouches) a permis l'émergence de toute une catégorie de classe de réseaux de neurones, dont la propriété reste l'approximation non-linéaire.

IV.6.1. Algorithmes d'apprentissage supervisé

Pour les réseaux de neurones, les algorithmes les plus utilisés, sont les modèles d'optimisation non linéaires par rapport aux paramètres, les vecteurs des entrées et de la sortie sont centrés:

$$Y = f(X, \theta) + \omega \quad (IV.8)$$

Où f est une fonction de régression (d'activation).

Ces modèles utilisent des méthodes itératives pour obtenir une estimation des paramètres. Leur principe est de se placer en un point, de trouver une direction de descente du coût dans l'espace des paramètres \square , et de se déplacer d'un pas suivant cette direction, on poursuit cette démarche jusqu'à satisfaction de ce critère d'arrêt (Dreyfus *et al.*, 2004)

IV.6.2. La Backpropagation

C'est Werbos (1974), qui a été le premier à mettre en œuvre, cet algorithme avec notamment, son fondement mathématique. Il avait appelé cette méthode *The Dynamic Algorithm Feed-back* (Werbos, 1990). Son initiative resta cependant sans succès.

En 1986 un groupe de chercheurs de l'université de Standford (Etats Unis) Rumelhart, Hinton et Williams, l'ont développé et l'ont complètement adapté aux réseaux multicouches. Cet algorithme a été publié dans leur ouvrage sur les processus distribués parallèles (Parallel Distributed Processing) (Rumelhart *et al.*, 1986). Ils ont appelé la méthode Backpropagation en référence à l'erreur qui se rétro-propage, à travers les couches du réseau (Yeddou, 1998).

Il s'agit d'une méthode d'apprentissage supervisé, fondée sur la modification des poids du réseau dans le sens contraire à celui du gradient de l'erreur par rapport à ces poids.

IV.6.2.1. Algorithme de Backpropagation

L'algorithme de Backpropagation est décrit comme suit:

1) Initialiser les poids synaptiques à de petites valeurs aléatoires comprises en général entre -1 et 1

2) Présenter un nouvel exemple X_p à l'entrée du réseau.

Calculer la sortie du réseau et l'erreur en sortie:

$$E_p = \sum_{i=1}^{NL} (d_i^p - y_i^L)^2 \quad (IV.9)$$

Où i parcourt les indices des neurones de sortie, d_i et y_i représentent respectivement l'activation mesurée et l'activation désirée pour les neurones.

3) les poids du réseau sont modifiés suivant la règle:

$$\left[\Delta w_{ij}^l \right]_p = \eta \delta_j^l (z_i^l) \delta_i^{l-1} \quad (IV.10)$$

Où

η

est une constante positive appelée pas du gradient (Taux d'apprentissage), et contrôle la vitesse de l'apprentissage.

$$\text{Avec : } \delta_i^l = \begin{cases} (d_i^l - y_i^l) & \text{pour les neurones de la couche de sortie} \\ \sum_{k=1}^{n^{l+1}} v_k^{l+1} w_{ki}^l & \text{pour les neurones des couches cachées} \end{cases} \quad (\text{IV.11})$$

S'il s'agit d'un Data Learning (FFNpattern), réadapter les poids:

$$1) \text{ Forcer : } \Delta w_{ij}^l(t) = \left[\Delta w_{ij}^l(t) \right]_p \quad (\text{IV.12})$$

S'il s'agit d'un Block Learning (FFN Batch):

Aller à 5.

Si $p \neq M$ (M étant le nombre d'exemple d'entraînement), retourner à 1.

$$\text{Si non : } \Delta w_{ij}^l(t) = \sum_{p=1}^{p=M} \left[\Delta w_{ij}^l(t) \right]_p \quad (\text{IV.13})$$

5) Réadapter les poids.

$$w_{ij}^l(t+1) = w_{ij}^l(t) + \Delta w_{ij}^l(t) \quad (\text{IV.14})$$

6) Répéter de 1. A 5. Autant de fois que nécessaire jusqu'à convergence vers le seuil d'erreur fixé.

Dans cet algorithme nous avons noté que l'apprentissage (réadaptation des poids) peut se faire de deux manières différentes, à savoir Block Adaptive Learning, et l'apprentissage par données Data Adaptive Learning.

Block Adaptive Learning (Batch learning)

Dans cette technique, on ne réadapte les poids qu'après passage de tous les exemples d'entraînement. La réadaptation est donc plus prudente. Le réajustement se fait suivant la moyenne de tous les exemples et la méthode est donc beaucoup moins sensible au bruit que peuvent contenir des exemples singuliers pouvant se présenter, la fonction de coût qui calcule l'erreur est:

$$E_p = \sum_{i=1}^{nL} (d_i^p)^2 - \sum_{i=1}^{nL} (v_i^p)^2 \quad (\text{IV.15})$$

Cette méthode est jugée plus robuste, cependant les données simulées correspondent à une moyenne de données observées.

Data Adaptive Learning (Incremental Learning)

En utilisant cette méthode d'apprentissage, les poids sont réadaptés au passage de chaque exemple :

$$E_P = \sum_{i=1}^{nL} (d_i^P - y_i^L)^2 \quad (IV.16)$$

De cette manière le processus devient sensible à chaque exemple individuellement, ce qui le rend facilement influençable par les bruits que peuvent contenir ces entrées durant l'entraînement. Cette technique n'est donc utilisée que pour un apprentissage en temps réel.

La simulation pluie-débit au pas de temps mensuel par les réseaux de neurones doté d'un apprentissage en block a déjà été testée au bassin versant de la Cheffia (Dechemi *et al.*, 2003).

IV.6.2.2. Etude du taux d'apprentissage

La version théorique de la Backpropagation réclame une variation infinitésimale des valeurs des poids à chaque itération. Cette variation est contrôlée par le taux d'apprentissage η . En pratique, ce facteur doit être petit assurant la convergence mais d'une manière très lente.

Si la valeur du paramètre η est grande, l'apprentissage est effectué de façon rapide, et la surface d'erreur est parcourue d'une manière 'téméraire', ce qui se répercute sur la recherche du minimum.

Précédemment nous avons mentionné, qu'une valeur importante du taux d'apprentissage était nécessaire pour accélérer l'apprentissage. Le problème d'oscillation nous empêche de le faire augmenter.

Rumelhart *et al.*, (1986) ont proposé une solution qui consiste à utiliser les changements précédents des poids pour la réadaptation des poids actuels.

L'équation d'adaptation devient donc:

$$w_{ij}^l(c+1) = w_{ij}^l(c) + \eta \Delta w_{ij}^l(c) + \alpha w_{ij}^l(c) - 1 \quad (IV.17)$$

Le terme ajouté est appelé Momentum (élan, quantité de mouvement), il représente un paramètre qui est utilisé pour pondérer l'effet de ce terme. Sa valeur est généralement comprise entre 0.7 et 0.9.

L'utilisation de cette méthode permet de faire sortir les poids des minima locaux, afin de chercher d'autres optimums, ce qui donne beaucoup de possibilités d'aboutir à un minimum global.

IV.6.3. Méthodes d'optimisation avec la Backprogataion

Nous présentons dans notre étude deux méthodes, celle de Newton et celle de Levenberg-Marquardt, qui est une méthode d'optimisation du second ordre. Cette dernière est basée sur l'utilisation de la dérivée seconde de la fonction coût (fonction

objectif) par rapport aux poids. L'introduction de la matrice Hessian H nous informe sur la forme de la surface d'erreur dans l'espace des poids.

Méthode de Newton

La méthode de Newton utilise la courbure (dérivée seconde) de la fonction de coût pour atteindre le minimum plus rapidement. Elle consiste à minimiser la fonction d'énergie:

$$E(k+1) = E(k) - E'(k) \Delta W(k) + \frac{1}{2} \Delta W(k)^T E''(k) \Delta W(k) \tag{IV.18}$$

La modification des paramètres est donnée par:

$$\begin{aligned} \theta_k &= \theta_{k-1} - H_{k-1}^{-1} \nabla J \\ \text{et } d_{k-1} &= - H_{k-1}^{-1} \nabla J \\ \nu_{k-1} &= 1 \text{ (constant)} \end{aligned} \tag{IV.19}$$

Méthode de Levenberg-Marquardt

Cette méthode consiste à considérer la surface d'erreur quadratique, en se basant sur la fonction d'énergie (V.18).

En posant pour chaque neurone *i* d'une couche *l* le vecteur de poids lui parvenant des neurones de la couche qui le précède, et qui inclue en plus le biais:

$$W_i^l = [w_{i1}^l, w_{i2}^l, \dots, w_{i_{n_{l-1}}}^l, \theta_i^l] \tag{IV.20}$$

Soient F_i^l : les vecteurs des dérivées de l'erreur à la sortie du réseau par rapport à ces vecteurs poids w_i^l , $l=1 \dots L$, $i=1 \dots n_l$

$$F_i^L = \sum_{p=1}^M \sum_{n_i=1}^{n_i} (d_i^p - y_i^p) \left[- \frac{\partial y_i^p}{\partial W_i^L} \right] \tag{IV.21}$$

M représente le nombre d'entraînement n_l , le nombre de neurones dans la couche *l*, Y_i^L la *i*ème sortie du réseau, et d_i^p sa sortie désirée pour la *p*ème entrée.

Soit F : Vecteur rassemblant des vecteurs F_i^L . Pour trouver l'optimum on a:

$$H \Delta W = -F \tag{IV.22}$$

H représente la matrice Hessian de la fonction erreur en sortie. La méthode de Levenberg-Marquardt remplace le calcul du Hessian par une approximation numérique donnée par :

$$H = FF^T + \lambda \Omega \tag{IV.23}$$

H est la valeur du calcul Hessien, FF^T est une matrice positive qui sert à conditionner la matrice H . Ω représente la matrice diagonale définie positive, λ est un coefficient positif.

IV.6.4. Propriétés des réseaux de neurones

Il a été démontré que moyennant le choix d'une architecture appropriée (i.e. nombre de neurones cachés), les réseaux multi-couche sont capables d'approcher n'importe quelle fonction (Cybenko, 1989). On parle d'approximateurs universels de fonctions.

Une propriété fondamentale de l'apprentissage réalisé concerne les capacités de généralisation de ces réseaux. Dans le cas où l'architecture initiale est correctement choisie, on constate généralement que les exemples ne sont pas appris, mais que le réseau est capable d'étendre les connaissances acquises à des exemples proches ou intermédiaires.

Cependant un certain nombre de problèmes sont liés à l'utilisation de ce type de réseaux (et de manière plus générale de la majeure partie des RNA). On peut remarquer notamment la lenteur de l'apprentissage et surtout l'absence de résultats théoriques garantissant sa convergence. On constate souvent des problèmes liés au blocage de l'apprentissage dans des minima locaux de la fonction d'erreur (Gauthier, 1999).

IV.7. Résultats de simulation par les réseaux de neurones

La simulation d'un processus par les réseaux de neurones requiert une certaine connaissance quant aux différents modèles et algorithmes que caractérisent ces réseaux.

Cette différenciation peut être résumée en:

a- Choix d'un réseau de neurones

Il existe un grand nombre de types de réseaux de neurones, avec pour chacun des avantages et des inconvénients (Lippman, 1987). Le choix d'un réseau peut dépendre:

- De la tâche à effectuer (classification, association, contrôle de processus, séparation aveugle de sources...);
- De la nature des données (dans notre cas, des données présentant des variations au cours du temps);
- D'éventuelles contraintes d'utilisation en temps réel: certains types de réseaux de neurones, tels que la 'machine de Boltzmann' nécessitant des tirages aléatoires et un nombre de cycles de calculs important, et présentent plus de contraintes que d'autres réseaux pour une utilisation en temps réel.

b- Choix du nombre d'unités cachées

Le nombre d'unités cachées (neurones dans les couches cachées) joue un rôle

crucial dans le contrôle de la capacité du réseau de neurones. Si la valeur de ce nombre est trop petite, alors le réseau possède trop peu de paramètres et ne peut capter toutes les dépendances qui servent à modéliser et prédire les valeurs du processus observé. À l'inverse, si l'on choisit une valeur trop grande, alors le nombre de paramètres du modèle augmente et il devient possible, pendant la phase d'optimisation des paramètres, de modéliser certaines relations qui ne sont que le fruit de fluctuations statistiques propres à l'ensemble d'entraînement utilisé plutôt que des relations fondamentales de dépendance entre les variables.

c- Paramètres d'apprentissage

Le choix des paramètres d'un algorithme d'apprentissage influe beaucoup sur la rapidité de calculs. Dans le cas de l'algorithme de rétropropagation, le calcul du gradient consiste à définir, dans un espace contenant autant de dimensions qu'il y a de poids, la direction dans laquelle doit s'effectuer la modification des poids.

Le principe de descente de gradient consiste alors à effectuer de manière itérative (pas par pas) une modification des poids suivant cette direction jusqu'à arriver à un minimum sur la fonction de coût représentant l'écart entre les sorties obtenues et les sorties de référence.

Autrement dit, un réseau de neurones peut apprendre par cœur un ensemble d'entraînement. Afin de s'assurer que le réseau de neurones s'en tient aux relations fondamentales de dépendance, on utilise, en plus de l'ensemble d'entraînement, un second ensemble appelé ensemble de validation : à la fin de chaque époque d'entraînement, on mesure non seulement l'erreur d'entraînement mais aussi l'erreur de validation, c'est à dire l'erreur totale commise sur tous les exemples de l'ensemble de validation. Cette erreur de validation est calculée une fois que la phase d'optimisation des paramètres est terminée.

Après avoir entraîné quelques modèles, chacun avec un nombre différent d'unités cachées, on peut comparer les erreurs d'entraînement et de validation.

L'erreur d'entraînement diminue au fur et à mesure que le nombre d'unités cachées augmente. L'erreur de validation, quant à elle, est élevée lorsque le nombre d'unités cachées est faible, décroît avec l'augmentation du nombre d'unités cachées, atteint un minimum pour un certain nombre optimal d'unités cachées, puis croît lorsque le nombre d'unités devient trop grand.

C'est donc l'utilisation d'un ensemble de validation, distinct de l'ensemble d'entraînement (calage), qui nous permet de choisir le nombre optimal d'unités cachées ou neurones.

d. Caractéristiques du modèle neuronal utilisé

Le modèle testé s'appuie sur un apprentissage de type supervisé, l'algorithme de backpropagation est celui de Levenberg-marquardt.

Les paramètres d'apprentissage sont:

- La fonction coût est l'erreur quadratique, inférieure ou égale à 0.01

- Le taux d'apprentissage η est fixé à 0.7.
- La fréquence d'apprentissage des données est de 10.

IV.7.1. Données mensuelles

Nous avons cinq stations hydrométriques à modéliser à savoir, les stations de Cheffia, Lakhdaria, La Traille, Sebbaou et la station de la Tafna.

Vu le nombre de couches cachées et le nombre de neurones par couches possibles, deux modèles seront appliqués:

Un modèle à une couche cachée à une couche cachée avec un neurone soit le modèle (2,1).

Un modèle à deux couches cachées avec respectivement deux (02) et quatre (04) neurones dans la première et la deuxième couche, modèle (4,2,1).

La couche cachée des réseaux sera modélisée avec un neurone, et une fonction de transfert linéaire. Pour les couches cachées, la fonction sigmoïde sera utilisée.

Les résultats des différents modèles des deux périodes en fonction de chaque station sont présentés dans le tableau IV-1 et IV-2.

Stations	Cheffia (1978-84)		Lakhdaria (1983-87)		La Traille (1970-77)		Sebbaou (1980-87)		Tafna (1980-1987)	
	Nash (%)	CET (%)	Nash (%)	CET (%)	Nash (%)	CET (%)	Nash (%)	CET (%)	Nash (%)	CET (%)
(4,1)	63.9	253	45.6	593	61.4	394	41.3	1000	-1.0	700
(2,4,1)	57.8	326	55.0	556	66.8	405	46.3	1000	-1.8	956

Tableau IV-1 Résultats des modèles neuronaux au pas de temps mensuel (calage)

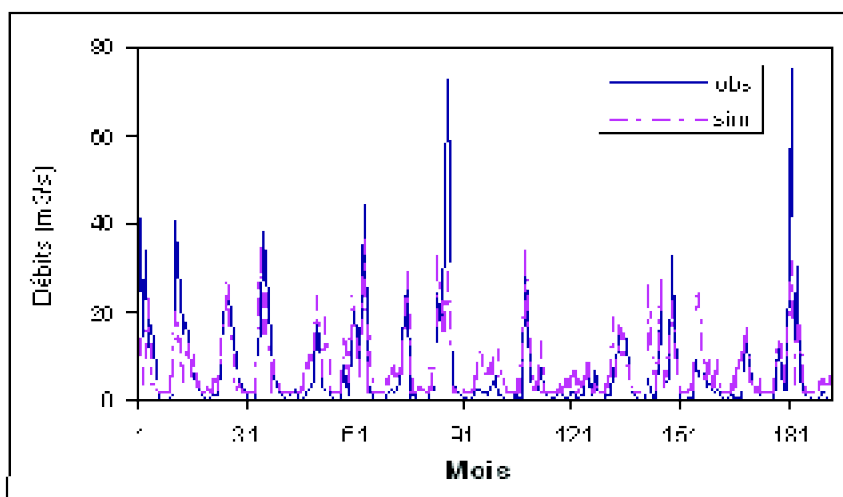
Stations	Cheffia (1985-91)		Lakhdara (1988-95)		La Traille (1979-85)		Sebbaou (1938-55)		Tafna (1988-1995)	
	Nash (%)	CET (%)	Nash (%)	CET (%)	Nash (%)	CET (%)	Nash (%)	CET (%)	Nash (%)	CET (%)
(4,1)	41.9	113	45.5	195	4.15	192	41.9	1000	5.0	47
(4,2,1)	45.2	129	52.0	197	6.0	190	35.4	410	10.0	52

Tableau IV-2 Résultats des modèles neuronaux au pas de temps mensuel (validation)

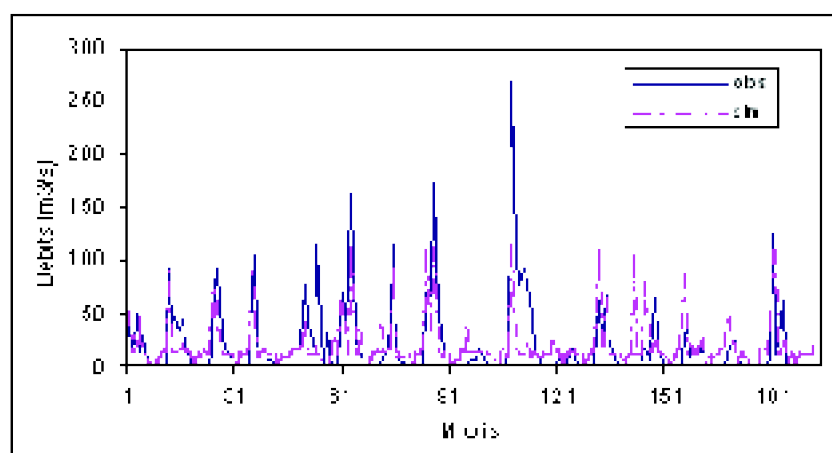
D'après les résultats des tableaux VI-1 et VI-2, nous remarquons que les modèles donnent des résultats sensiblement différents, en fonction du nombre de chaque station et du nombre de couches cachées. En effet, pour le modèle à une couche cachée l'efficacité de Nash est relativement importante pour les stations de Cheffia et de Sebbaou avec respectivement des valeurs de 63% et 61%.

Pour la station de Tafna on note une très mauvaise simulation, le critère de Nash est en effet négatif, il est de -1% en période de calage, et ne dépasse pas 5% en phase de validation. Les résultats des autres stations sont sensiblement moins performants.

Excepté la station de Sebbaou, toutes les séries des bassins versants présentent une baisse de la qualité de simulation en phase de validation, et cela quelque soit le modèle.



(a)



(b)

Figure IV-5 Résultats de simulations mensuelles par les réseaux de neurones

(a) station de La Traille Lakhdaria, (b) station du Sebbaou

Pour ce qui est des débits d'étiages, le rapport CET qui mesure la justesse de reproduction des faibles, les résultats concernant cet indice, montrent une mauvaise simulation de ces débits, à titre exemple, le CET est de 1000% pour la station du Sebbaou, et de 700% pour la station de Tafna pour la période de calage.

Le modèle à deux couches cachées (4,2,1) qui utilise sept paramètres (neurones) dans l'optimisation, ne donne généralement pas de meilleurs résultats, pour la station de la Chaffia, une baisse de simulation est à noter entre les deux modèles, qui est de l'ordre de 6%.

En phase de validation, ce modèle enregistre une mauvaise qualité de simulation en comparant avec ceux obtenus en période de calage.

Il est à signaler, que le mode d'apprentissage (réadaptation des poids) est en block (batch training), dans ce cas la réadaptation des poids ne se fait qu'après passage de

tous les exemples, dans ce cas, le modèle se trouve à simuler la moyenne de tous les débits observés, plutôt que de calculer chaque donnée en correspondant avec la donnée observée.

IV.7.2. Données journalières

Au pas de temps journalier, les modèles seront testés sur trois stations hydrométriques, la station de Cheffia, la station de l'Isser et celle de Zarzedas.

La procédure de validation des modèles est identique à celle des modèles conceptuels. Les données seront divisées en deux périodes, la période de calage et une autre pour la validation du modèle. Dans le cas des données journalières, nous disposons d'un échantillon de six ans, et cela pour toutes les stations à modéliser.

Deux modèles seront testés pour chaque station, un modèle avec un nombre de couches et de neurones différents. Deux modèles seront testés en fonction du nombre de couches cachées.

Un modèle à une couche cachée avec quatre neurones soit le modèle (4,1).

Un modèle à deux couches cachées avec respectivement deux (02) et quatre (04) neurones dans la première et la deuxième couche, modèle (4,2,1).

Les critères d'appréciation des modèles sont ceux appliqués aux modèles conceptuels, à savoir le critère de Nash, l'indice de reproduction du débit maximal, et le rapport des débits d'étiages.

Stations	Cheffia (1986-88)			Isser (1990-92)			Zarzedas (1991-93)		
	Nash (%)	Ic	CET (%)	Nash (%)	Ic	CET (%)	Nash (%)	Ic	CET (%)
(4,1)	38.6	0.52	1600	35.0	0.99	1400	26.8	0.74	2800
(4,2,1)	36.8	0.42	1669	6.05	0.1	1530	24.8	1.18	2400

Tableau IV-3 Résultats des modèles journaliers en période de calage

Stations	Cheffia (1989-91)			Isser (1993-95)			Zarzedas (1994-96)		
	Nash (%)	Ic	CET (%)	Nash (%)	Ic	CET (%)	Nash (%)	Ic	CET (%)
(4,1)	1.0	0.72	87.0	-63.0	0.99	172	20.0	0.2	139
(4,2,1)	12.3	0.62	86.0	6.35	0.1	179	16.5	0.12	135

Tableau IV-4 Résultats des modèles journaliers en période de validation

En examinant les résultats des tableaux Tableau VI-3 et VI-4, nous remarquons que les performances des modèles neuronaux varient sensiblement d'un modèle à un autre.

Pour la période de calage, les résultats sont mauvais, et cela quelque soit la station, en effet le critère de Nash ne dépasse pas 40% pour la station de Cheffia, alors que la station de Zardezas enregistre la plus mauvaise efficacité avec 26 % seulement.

En période de validation, il est à signaler une détérioration de la simulation pour toutes les stations, avec une efficacité négative de -63% pour l'Isser, c'est-à-dire des débits aussi mauvais que la moyenne du débit observé.

Cependant, les résultats obtenus pour l'ensemble des modèles, sont nettement moins performants que ceux des modèles conceptuels, que ce soit en période de calage ou en période de validation. Pour ce qui est des débits d'étiage, excepté la station de Zardezas en période de calage, le rapport CET n'est toujours pas optimisé, et ses valeurs sont largement aberrantes.

Pour ce qui est des étiages, les résultats montrent une nette différence entre la phase de calage où les débits sont nettement surestimés, et celle de la validation, dans cette période, le rapport CET est relativement correct, CET est égal à 135% pour la station de Zardezas.

Il nous a paru donc, que les modèles neuronaux présentent des problèmes de reconstitution de certaines valeurs qui se répètent successivement en particulier les valeurs nulles (pluie nulle).

En effet, les modèles se trouvent face à des cas difficilement modélisables (ex: pluie nulle débit faible, pluie nulle débit élevé).

Ainsi, les diverses tentatives que nous avons menées en vue d'optimiser les débits journaliers par le modèle flou en augmentant le nombre de couches cachées.

Le résultat de cette modélisation est donc une mauvaise simulation des débits de décrue, (décrues brusques), et une surestimation des débits d'étiage.

Ainsi, on conclut que la modélisation pluie-débit à l'échelle journalière par le modèle neuronal à une seule entrée (pluie) est jugée délicate pour caractériser les différents cas qui peuvent se présenter au modèle.

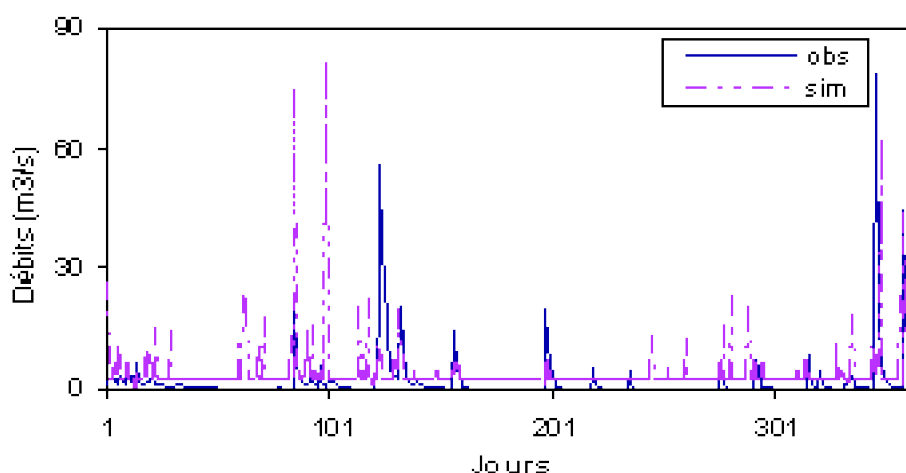


Figure IV-6 Résultats de simulations journalières par les réseaux de neurones, station de l'Isser (année 1991)

IV.7.3. Résultats des périodes caractéristiques

Les variations climatiques, représentent l'épreuve majeure que doit affronter un modèle pluie-débit, et cela pour juger de sa capacité d'adaptation.

Cette procédure de simulation, nous permet de tester le modèle sur un éventail de données relativement court, et de déduire la sensibilité des réseaux de neurones aux différents cycles hydrologiques. Le modèle neuronal sera testé sur les séries courtes des stations du Sebbaou et de Zardezas. Les résultats sont détaillés dans les tableaux VI-5 et VI-6.

Sebbaou	Calage (1993)			Validation (1994)		
	Nash (%)	Ic	CET (%)	Nash (%)	Ic	CET (%)
(4,1)	4.4	0.12	4600	-2.7	0.10	336
(4,2,1)	3.1	0.10	4300	-2.1	0.10	338

Tableau IV-5 Résultats des modèles neuronaux par périodes caractéristiques (Station du Sebbaou)

Zardezaz	Calage (1994-95)			Validation (1996-97)		
	Nash (%)	Ic	CET (%)	Nash (%)	Ic	CET (%)
(4,1)	42.0	0.33	818	28.4	0.38	161
(4,2,1)	49.5	0.43	742	21.8	0.52	161

Tableau IV-6 Résultats des modèles neuronaux par périodes caractéristiques (Station de Zardezaz)

Il apparaît clairement dans les tableaux VI-5 et VI-6 que les résultats obtenus pour la station du Sebbaou sont très médiocres quelque soit le nombre de neurones introduits dans le modèle, et cela pour les deux périodes de calage et de validation. Le critère de Nash ne dépasse par 5 % en phase de calage pour le modèle à une couche cachée.

En revanche pour la station de Zardezaz, les résultats sont nettement meilleurs en particulier pour le modèle à 2 couches, où on obtient une efficacité de Nash de près de 50% ce qui est très moyen. Cependant, en période de validation, une baisse de performance est à signaler, elle atteint en moyenne 18%.

Pour ce qui est des débits caractéristiques, l'indice *Ic* est médiocre, quelque soit le modèle, en effet la reproduction du débit maximal est pratiquement mauvaise pour toutes les stations.

IV.8. Conclusion

Dans ce chapitre, nous avons procédé à une présentation dans l'univers des réseaux de neurones. Nous avons abordé quelques définitions et notions essentielles pour une étude des réseaux de neurones. Nous avons insisté sur le caractère très varié de ces réseaux, ainsi que leur évolution, et leur application dans de nombreux domaines.

Ayant détaillé les diverses techniques d'apprentissage effectuées par les réseaux de neurones, nous avons défini les propriétés de chaque apprentissage, et présenté quelques algorithmes exploités dans de nombreuses disciplines.

Dans ce contexte, nous avons souligné qu'une des fonctions renommées effectuées par les réseaux est l'approximation non-linéaire des données lorsque la linéarité des méthodes classiques limite l'extrapolation de ces données. Cette tâche est généralement effectuée par les réseaux basés sur l'apprentissage supervisé.

Connaissant le caractère complexe de la relation pluie-débit, nous tentons à travers cette thèse d'exploiter au mieux la technique des réseaux de neurones, et d'en juger ses

performances en matière de la modélisation pluie-débit.

Cependant, le type d'apprentissage utilisé dans cette thèse correspond à un apprentissage en block (batch training), ainsi les résultats obtenus ont été relativement très limités voir médiocres au pas de temps journalier. L'approche «batch» représente une simulation des données par moyenne, ainsi les critères d'appréciation sensibles à la différence des erreurs par données, tels que le critère de Nash sont souvent très faibles.

Nous concluons que l'utilisation des réseaux de neurones en block, que ce soient au pas de temps mensuel ou journalier, n'est pas adaptée à la modélisation pluie-débit qui est un processus se calculant en temps réel.

Chapitre V : Introduction à la logique floue : généralités et aspects théoriques

V.1. Introduction

Dans les chapitres III et IV, nous avons exploité la modélisation pluie-débit par deux approches différentes à savoir les modèles conceptuels déterministes et les réseaux de neurones artificiels, basés sur un apprentissage de données. Néanmoins ces deux modèles, ont en commun la propriété d'être des méthodes analytiques, qui permettent une simulation des débits à partir des variables numériques.

A la différence de ces deux modèles, la logique floue permet une représentation du processus étudié par un langage naturel sous la forme des règles en introduisant des nuances ' faible' 'élevé' 'très élevé', et modélise en partie l'approche humaine, ce qui permet d'introduire une interprétabilité au système modélisé.

L'approche floue a été introduite pour la première fois par Zadeh (1965) comme un outil permettant de prendre en considération les incertitudes sur les grandeurs mesurées.

Conscients des limites des méthodes classiques dans la modélisation de systèmes complexes, Zadeh cherchait à définir un nouveau formalisme susceptible de nuancer les deux concepts d'incertitude et d'imprécision. Il eut l'idée de généraliser la notion de

fonction caractéristique d'un ensemble net régissant l'appartenance, ou la non appartenance, d'un élément quelconque de l'univers à cet ensemble de façon partielle et continûment graduelle. Cette approche constitue une rupture avec le paradigme de la théorie classique des ensembles qui consiste à discerner deux états exclusifs (oui (vrai), ou non (faux)) de l'appartenance d'un élément à un ensemble net.

Cependant, les premiers travaux pratiques sur la logique floue, ne furent réalisés que dans les années soixante dix (Mamdani, 1974), l'idée de base consistait à exploiter l'expérience des opérateurs humains pour construire une loi de commande.

De nos jours, les applications des systèmes flous se sont multipliées. Entre autres, nous pouvons citer les contrôleurs flous, les systèmes de classification et de reconnaissance des formes flous. Malgré les différences dues aux particularités de chaque application, ces systèmes ont pratiquement la même structure interne, et se basent sur un mécanisme d'inférence, ils sont baptisés systèmes d'inférence flous (SIF).

V.2. Notions de base et connaissances préliminaires

La logique classique caractérisée par ses deux valeurs de vérité (vrai et faux) et ses règles d'inférence telles que le modus ponens et le modus tollens, est inappropriée à une description fine et souple des connaissances. En particulier, elle est inadéquate à la représentation et aux traitements des connaissances vagues, imprécises ou incertaines et ne peut de ce fait être utilisée comme modèle de base du raisonnement humain. En effet, la logique floue ne se limite pas à deux valeurs de vérité 'classiques' du vrai et du faux absolus, mais à une infinité de valeurs dans l'intervalle [0,1].

V.2.1. Définition d'une variable linguistique

La notion de variable linguistique a été introduite par Zadeh (1971, 1975), elle correspond à une variable dont les valeurs sont des mots ou des phrases exprimées dans une langue naturelle ou un langage artificiel (Zadeh, 1971). Elle est définie par le quadruplet suivant (Driankov et al., 1993):

$$\{x(\text{nom}), L(x), (U), 3 \text{ tre}(x)\} \dots \dots \dots (V.1)$$

Où:

$x(\text{nom})$ est le nom de la variable linguistique, comme taille, température, pluie, âge, etc.

$L(x) = \{L_1, \dots, L_n\}$ est l'ensemble des valeurs linguistiques que peut prendre la variable $x(\text{nom})$. Par exemple $L(X) \square \text{grand, petit, moyen}$. $L(x)$ est aussi appelé ensemble des étiquettes associé à la variable $x(\text{nom})$ ou ensemble des termes.

U est dénommé univers du discours, il représente le domaine associé à la variable $x(\text{nom})$. Il correspond à l'ensemble de toutes les valeurs (ou appréciations) que peut prendre la variable numérique associée à la variable linguistique $x(\text{nom})$. Par exemple

pour la variable linguistique taille, U peut être défini dans l'intervalle [1,60m –1,90m]

\mathfrak{S}

est une fonction sémantique qui associe à tout symbole de U une signification floue.

Une variable linguistique peut être plus informative et plus généralisable qu'une variable numérique (Nakoula, 1997).

La figure V-1 présente un exemple de variable linguistique taille avec l'ensemble d'étiquettes correspondant:

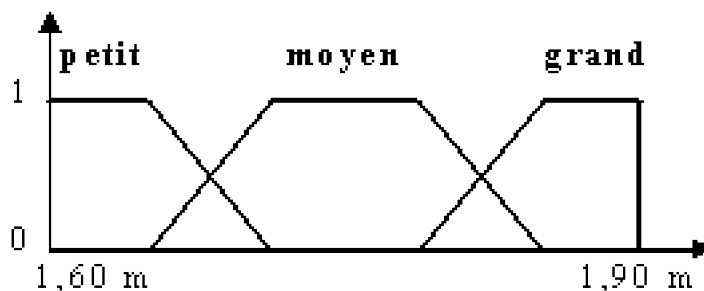


Figure V-1 Etiquette pour la variable linguistique taille

V.2.2 Sous-ensembles flous : concepts et définition

Un sous ensemble flou $F(U)$ est défini sur un univers de discours U , et caractérisé par une fonction d'appartenance \mathfrak{S}_F à valeurs dans un intervalle $[0,1]$, soit donc $\mathfrak{S}_F \rightarrow [0,1]$ (Zadeh, 1975).

Le sous ensemble flou F est défini par :

$$F = \{(u, \mathfrak{S}_F(u)) / u \in U\} \quad (V.2)$$

Si Z est discret, tout élément de $F(Z)$ sera noté:

$$F = \sum (u, \mathfrak{S}_F(u)) / u \in U \quad (V.3)$$

Dans le cas où z est continu:

$$F = \int \sum (u, \mathfrak{S}_F(u)) / u \in U \quad (V.4)$$

La figure V-2 présente un exemple de fonction d'appartenance d'un ensemble flou défini sur un univers de discours U , ainsi que la fonction d'appartenance d'un ensemble classique.

Dans le cas d'un sous-ensemble flou F , on distingue trois cas pour la fonction d'appartenance :

- $\mathfrak{S}(x) = 0$, si x n'appartient pas au sous-ensemble flou $F(U)$;
- $\mathfrak{S}(x) = 1$, si x appartient au sous-ensemble flou $F(U)$;

- $\exists (x) \neq 0$, et $\exists (x) \neq 1$, si x appartient partiellement au sous-ensemble flou $F(U)$.

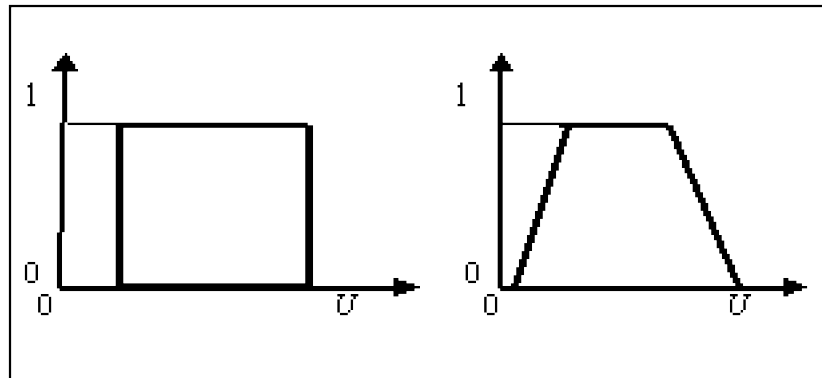


Figure V-2 Fonction d'appartenance d'un ensemble précis (classique) et d'un ensemble flou

Le support est l'ensemble des éléments qui appartiennent à l'univers E tels que la fonction d'appartenance soit non nulle sur cet intervalle.

$$\text{Support}(F) = x \in U / \exists (x) \neq 0 \quad (V.5)$$

Le noyau est l'ensemble des éléments qui appartiennent à l'univers U tels que la fonction d'appartenance soit égale à 1 sur cet intervalle.

$$\text{Noyau}(F) = x \in U / \exists (x) = 1 \quad (V.6)$$

Une α coupe correspond à l'intervalle :

$$N_{\alpha}(F) = \{x \in R / \mu_F(x) \geq \alpha\}$$

(V.7)

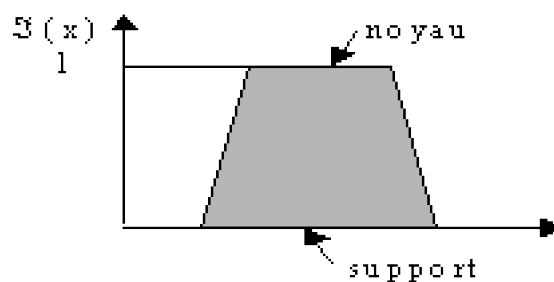


Figure V-3 Exemple d'un support et d'un noyau

V.2.3. Opérations sur les ensembles flous

Pour pouvoir manipuler les sous-ensembles flous, il a fallu généraliser les opérations ensemblistes classiques. Considérons que A et B sont deux sous-ensembles flous définis dans l'univers de discours U .

Définition (1) : A et B sont égaux si:

$$\mathfrak{F}_A(x) = \mathfrak{F}_B(x) \quad \forall x \in U$$

(V.8)

Définition (2): Le complément de A ou (A^c) de A par rapport à U est défini par la fonction d'appartenance suivante:

$$\mathfrak{F}_{A^c}(x) = 1 - \mathfrak{F}_A(x) \quad \forall x \in U$$

(V.9)

Définition (3): A est inclus dans B si et si seulement

$$\mathfrak{F}_A(x) \leq \mathfrak{F}_B(x) \quad \forall x \in U$$

(V.10)

Définition (4): L'union de A et B est l'ensemble d'appartenance ayant la fonction d'appartenance suivante:

$$\mathfrak{F}_{A \cup B}(x) = \max \{ \mathfrak{F}_A(x), \mathfrak{F}_B(x) \} \quad \forall x \in U$$

(V.11)

Définition (5): L'intersection de A et B est l'ensemble d'appartenance ayant la fonction d'appartenance suivante:

$$\mathfrak{F}_{A \cap B}(x) = \min \{ \mathfrak{F}_A(x), \mathfrak{F}_B(x) \} \quad \forall x \in U$$

(V.12)

De ce fait nous pouvons conclure que:

Deux sous-ensembles flous sont égaux si leurs fonctions d'appartenance prennent la même valeur en tout point de U.

Chaque élément qui appartient à A, (A étant inclus dans B), appartient à B au moins de la même façon qu'il appartient à A. Tout d'abord, un élément $x \in U$ appartient à $A \cup B$ s'il appartient à A ou à B. Pour que cet élément x appartienne à $A \cap B$ il faut qu'il soit dans A et B.

En définissant l'union comme le 'max', on détecte le plus petit ensemble flou qui contient A et B.

En définissant l'intersection comme le 'min', on détecte le plus grand ensemble flou qui est contenu dans A et B. De cette façon, un ensemble flou qui contient A et B contient $A \cup B$, qui à son tour contient $A \cap B$.

V.2.4. Les systèmes d'inférence floue (SIF)

Les systèmes d'inférence floue (SIF) permettent de modéliser la plupart des applications continues d'un espace à n dimensions dans \mathbb{R} . A la différence des modèles

mathématiques ou des modèles de type "boite noire", la représentation est faite en langage naturel, sous la forme de règles *Si ... Alors*. Ceci permet une interprétabilité immédiate, qui peut être exploitée de deux façons :

A priori, en donnant une valeur, même approximative, aux différents paramètres du SIF,

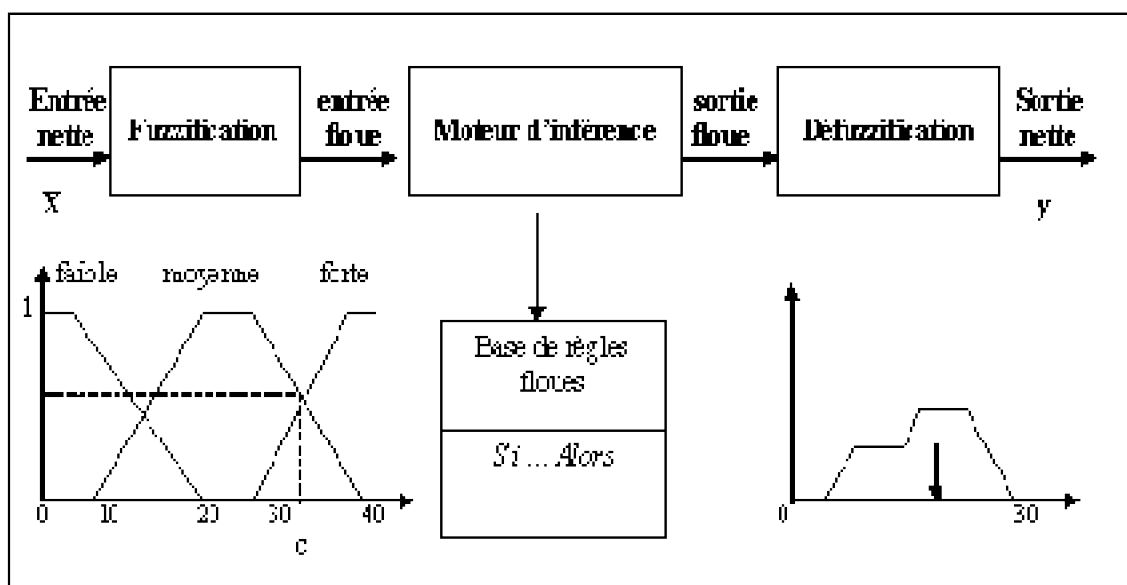
A posteriori, en permettant l'examen de la connaissance extraite lors d'une optimisation (on parle aussi d'apprentissage). Cet examen peut prendre plusieurs formes :

La validation, une par une, des règles extraites;

La détection de solutions aberrantes, pouvant être dues à des erreurs de mesure, à du bruit ou à l'influence de variables non modélisées, à la détection de zones de l'espace d'entrées non ou insuffisamment couvertes et qui correspondent à des configurations physiquement impossibles, à de mauvaises conditions d'expérimentation ou à des cas particuliers qui demandent un traitement spécifique (Glorennec & Orseau, 1992).

V.2.5. Structure interne d'un système d'inférence floue (SIF)

Elle est déterminée par trois étapes essentielles: la fuzzification, l'inférence floue, et la défuzzification (Figure V-4).



FigureV-4. Structure d'un SIF (Guillaume, 2001)

V.2.5.1. La fuzzification

La fuzzification ou quantification floue permet de modéliser les entrées d'un système entrée/sortie principalement sous forme de courbes (fonction d'appartenance), et permettent de déterminer le degré d'appartenance d'une valeur à un état donné. La fonction d'appartenance associée à un terme linguistique est appelée signification floue. L'ensemble des parties floues associées aux termes flous représente une partition floue (Bouchon-Meunier et al., 2003).

Les fonctions d'appartenance peuvent prendre différentes formes: singleton, triangulaire ou gaussienne, etc.

Les fonctions d'appartenance doivent être correctement construites, car elles sont à la base du système. Une courbe possède en abscisse les valeurs de la variable d'entrée qui peut être une distance, une vitesse, une mesure de pluie ou de débit, et en ordonnée le degré d'appartenance à l'état, compris entre 0 et 1. Dans ce cas, la transition d'appartenance à non-appartenance est graduelle (Kauffman, 1973).

On peut distinguer deux types de fuzzification, la fuzzification numérique et la fuzzification symbolique.

fuzzification numérique:

Elle transforme toute valeur numérique x de l'entrée en une partie floue numérique, c'est à dire un élément de $F(x)$. Cette partie floue caractérise l'imprécision qu'on peut avoir sur la mesure x .

Pour effectuer cette fuzzification deux méthodes sont possibles:

- Fuzzification singleton, est utilisée lorsque la valeur fournie par le capteur est supposée connue (Figure V-5). La fonction d'appartenance est définie par:

$$F(x) = \begin{cases} 1 & \text{Si } x = x_0 \\ 0 & \text{Si } x \neq x_0 \end{cases}$$

(V.13)

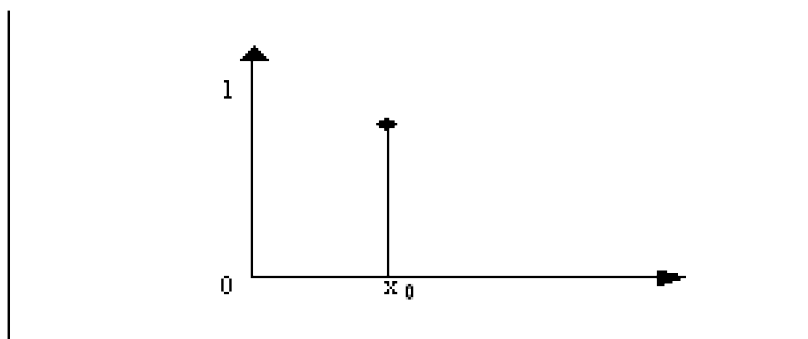


Figure V-5 Fuzzification singleton

Fuzzification non singleton, on considère dans ce cas, que le capteur peut effectuer une erreur de mesure.

b) Fuzzification symbolique

Elle transforme toute valeur numérique x en une partie floue linguistique, c'est à dire un élément de $F(L(x))$.

V.2.5.2. L'inférence floue

L'inférence floue peut être définie comme étant l'opération qui étend au cas flou le calcul de l'image d'une partie dont on connaît le graphe (Josserand, 1996).

Elle est constituée de règles floues du type *Si .Alors*. Pour chaque règle, elle permet de calculer à partir des degrés d'appartenance, le degré de vérité de x , et en déduit la conclusion de la règle pour le vecteur d'entrée présenté, et agrège ensuite les conclusions de toutes les règles activées pour ce vecteur.

Zadeh (1994) a défini L'inférence comme un problème d'interpolation tel qu'exprimé:

Soit U l'ensemble d'entrée, et V l'ensemble de sortie, la relation floue R est un sous-ensemble flou défini sur le produit cartésien $U \times V$, soit $R \in F(U \times V)$. La connaissance E sur les valeurs admises en entrée est exprimée par un sous-ensemble défini sur U , soit $E \in F(U)$. L'interpolation de la connaissance sur la sortie consiste à déterminer l'ensemble F des valeurs que peut prendre la sortie, F étant un sous-ensemble défini sur V , soit $F \in F(V)$.

V.2.5.3. La défuzzification

La défuzzification représente la conclusion d'un calcul par le modèle flou. Elle permet de transformer la partie floue issue de l'inférence en une valeur physique ou numérique (Josserand, 1996). On peut distinguer deux types de défuzzification: numérique et symbolique. Dans le cadre de notre travail nous insistons sur la défuzzification numérique qui ne nécessite aucune connaissance sur les symboles associés à la sortie.

Défuzzification numérique

Diverses méthodes de défuzzification numériques ont été développées (Driankov *et al.*, 1993), telles que la méthode du maximum, de la moyenne pondérée ou encore celle du centre de gravité.

La technique du centre de gravité

Cette technique est la plus utilisée malgré le nombre de calcul important, elle fera l'objet de notre travail de recherche.

La valeur de sortie est égale au centre de gravité de l'enveloppe fournie par l'union de zones du vecteur de sortie (Figure V-6). Dans ce cas, tous les termes calculés ont une influence et la valeur numérique de sortie peut prendre une infinité de valeur.

Soit $F \in F(Y)$, la partie floue numérique issue de l'inférence, la valeur numérique y^* produite par l'application de la méthode du centre de gravité est donnée par:

$$y^* = \tilde{O}_1(F) = \frac{\int_{y \in Y} \mu_F(y) \cdot y \cdot dy}{\int_{y \in Y} \mu_F(y) dy} \quad \text{si } Y \text{ est continu} \quad (V.14)$$

$$y^* = \tilde{O}_1(F) = \frac{\sum_{y \in Y} \mu_F(y) \cdot y}{\sum_{y \in Y} \mu_F(y)} \quad \text{si } Y \text{ est discret} \quad (V.15)$$

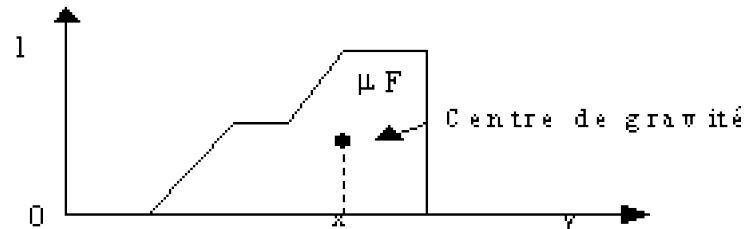


Figure V-6 Défuzzification par la méthode du centre de gravité

V.3. Construction de la relation entrée sortie à partir des règles floues

V.3.1 Proposition floue et inférence floue

Une proposition floue élémentaire est une expression de la forme: "Si x est A",

Où x est une variable linguistique et A une caractérisation floue de V.

Une règle floue est symboliquement exprimée comme une règle conditionnelle:

Si < proposition floue 1 > Alors < proposition floue 2 >

En termes flous, on décrit ces règles par:

si (prémisse i) Alors (conséquence i)

Si on considère:

$L(X) = \{A_i, i=1,2, \dots, I\}$ l'ensemble des symboles associés à l'entrée,

$L(Y) = \{B_j, j=1,2, \dots, J\}$ l'ensemble des symboles associés à la sortie,

$I = \{1,2, \dots, I\}$ l'ensemble indices possibles pour les symboles d'entrée,

$J = \{1,2, \dots, J\}$ l'ensemble indices possibles pour les symboles de sortie.

Toute règle floue a la forme générale:

Si entrée est A_i , Alors sortie est B_j(V.16)

Avec $(i,j) \in B \subset I \times J$.

Dans le cas présent, la prémisse est une proposition élémentaire floue. Celle-ci peut être plus complexe et combiner plusieurs propositions avec les opérateurs de conjonction, disjonction et de négation (Hartani, 1995).

Les règles du système définissent les relations logiques entre les ensembles des variables d'entrées et les ensembles des variables de sortie à l'aide d'opérateur et des étiquettes linguistiques.

Un exemple d'une règle floue est donné par:

Si la pluie est forte, Alors le débit est élevé.

Cette règle nécessite la définition des termes linguistiques forte et élevé.

V.3.2. Implication floue

Une implication floue associée à une règle floue est définie par la fonction d'appartenance de la relation floue R entre les ensembles de référence U et V .

Elle s'exprime en fonction d'appartenance intervenant dans la prémisse et conséquent d'une règle floue.

B définit la base de règles constituées de l'ensemble des règles élémentaires indexées par le couple (i,j) .

Selon le type d'architecture de SIF choisie, la construction fait appel à différentes interprétations des règles.

Les t-normes les plus utilisées dans ce type de systèmes flous sont celle de Mamdani ($T=\min$) et celle de Larsen ($T=$ produit).

Approche implicative

Dans cette approche le " *Si... Alors*" est interprété par:

$$\exists! x \in M_x(A_i) \text{ Alors } y \in M_y(B_j) \Rightarrow x \in M_x(A_i) \Rightarrow y \in M_y(B_j) \quad (V.17)$$

Où \Rightarrow représente un opérateur d'implication floue.

On obtient ainsi une nouvelle construction de chaque relation élémentaire R_{ij}

$$\mu_{R_{ij}}(x, y) \Rightarrow (\mu_{M_T}(A_i)(x), \mu_{M_T}(B_j)(y))$$

(V.18)

V.3.3. Système d'inférence flou à deux (ou plusieurs) entrées (MISO)

Une relation floue binaire R entre les ensembles de référence U et V est un sous-ensemble flou défini sur le produit cartésien U x V.

Considérons un système d'inférence à r variables d'entrées (MISO) :

$\langle x_1, L(X_1), X_1, M_{x1} \rangle \langle x_r, L(X_r), X_r, M_{xr} \rangle \dots$ et une variable de sortie:

$\langle y, L(Y), Y, M \rangle$

Nous prenons le cas d'un système d'inférence appelé SIF1, définie par la règle compositionnelle:

$$(U = X, V = Y) \forall x \in X, \forall y \in Y$$

$$\mu_F(y) = \bigwedge_{x \in X} (\mu_E(x) T \mu_R(x, y))$$

(V.19)

L'extension de la règle compositionnelle d'inférence (V.19) est donnée par l'équation:

$$\forall y \in Y, \mu_F(y) = \bigwedge_{(x_1, \dots, x_r) \in X_1 \times \dots \times X_r} (\mu_E(x_1, \dots, x_r) T \mu_R(x_1, \dots, x_r, y)) \quad (V.20)$$

$$\forall y \in Y, \mu_F(y) = \bigwedge_{(x_1, \dots, x_r) \in X_1 \times \dots \times X_r} (\mu_E(x_1, \dots, x_r) T \mu_R(x_1, \dots, x_r, y)) \quad (V.21)$$

ϵ représente une partie floue multidimensionnelle généralement obtenue par le produit cartésien des parties floues issues de la fuzzification de chacune des variables d'entrée.

Les définitions précédentes peuvent être appliquées dans le cas d'un SIF à plusieurs sorties (MIMO). Cependant, dans notre thèse, le cas multi-entrées/multi-sorties ne fera pas l'objet d'un travail de recherche, vu que son utilisation requiert plusieurs séries de données.

V.3.4. Opérateurs d'implication flous

Il existe une logique binaire pour les ensembles classiques et des logiques multivalentes (avec degré de vérité) pour les sous-ensembles flous. Les opérateurs sur les sous-ensembles flous sont créés en essayant de respecter un certain nombre de propriétés et de retrouver les opérations de base existant en logique classique telles que l'union ou l'intersection.

Soit X et Y un couple d'univers de discours, une relation floue R entre X et Y est définie par:

$$R = \{ ((x, y), \mu_R(x, y)) \mid (x, y) \in X \times Y \} \dots \dots \dots (V.22)$$

$$\mu_R: X \times Y \rightarrow [0, 1]$$

La définition des opérateurs *et*, *ou* se fait à partir des fonctions d'appartenance aux ensembles flous correspondants:

Le climat est chaud et humide $\Leftrightarrow x \in A_{chaud} \cap B_{humide} \Leftrightarrow \mu_{A_{chaud} \cap B_{humide}}(x)$

Le climat est chaud ou humide $\Leftrightarrow x \in A_{chaud} \cup B_{humide} \Leftrightarrow \mu_{A_{chaud} \cup B_{humide}}(x)$

Les définitions les plus courantes sont alors :

- Et (intersection): $\mu_{A \cap B}(x,y) = \mu_{A \cap B}(x,y) = \min(\mu_A, \mu_B)$ (V.23)

- Ou (union): $\mu_{A \cup B}(x,y) = \mu_{A \cup B}(x,y) = \max(\mu_A, \mu_B)$ (V.24)

Exemple :

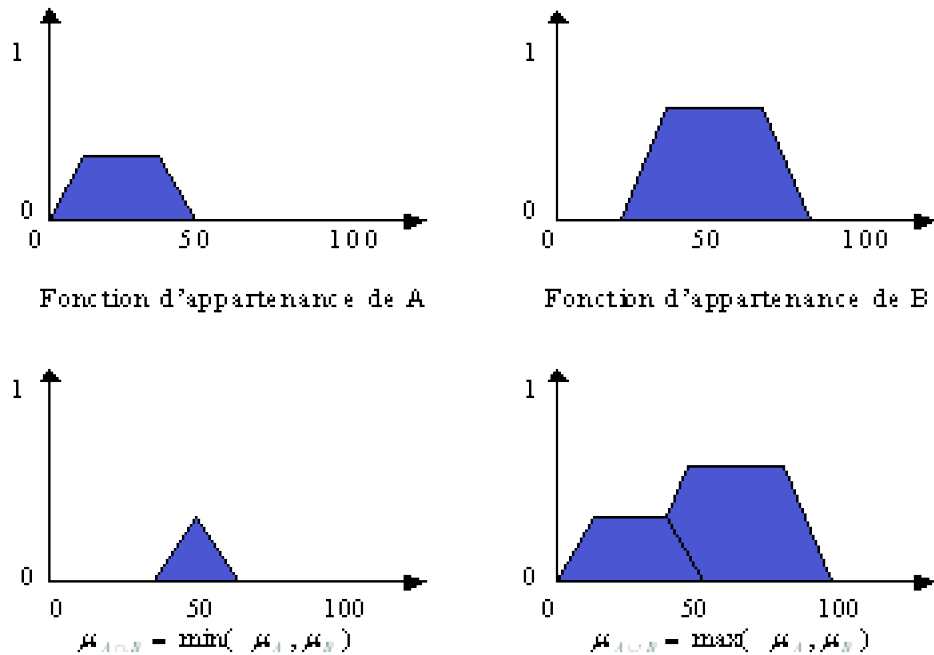


Figure V-7 Exemples d'intersection et d'union des sous ensembles flous

D'une manière plus générale, l'opérateur *et* peut être défini à l'aide d'une norme triangulaire (T_norme) :

$T : [0,1] \times [0,1] \rightarrow [0,1]$(V.25)

$x, y \rightarrow z = x T y$(V.26)

Avec les propriétés suivantes:

- Commutativité $x T y = y T x$
- Associativité $(x.T.y) T z = x.T.(y.T.z)$
- Non-décroissance par rapport aux arguments si $x \leq y, w \leq z$, alors $x T w \leq y T z$

De même, l'opérateur *Ou* peut être défini par la donnée d'une S_norme (T_conorme) :

$S : [0,1] \times [0,1] \rightarrow [0,1]$(V.27)

Avec les propriétés suivantes:

- commutativité $x S y = y S x$

- Associativité $(x \text{ S } y) \text{ S } z = x \text{ S } (y \text{ S } z)$
- non-décroissance par rapport aux arguments si $x \leq y, w \leq z$, alors $x \text{ S } w \leq y \text{ S } z$.

Il existe différents opérateurs d'implication flous : T_norme, S_norme, Minimum de Mandani et de Zadeh, Maximum de Lukasiewicz, ou encore le produit de Goguen.

Les 2 opérateurs classiquement utilisés dans les applications de la théorie des ensembles flous sont les fonctions Max. et Min. Cependant, les T_normes peuvent être employées comme opérateur de conjonction en lieu et place du Min., tandis que les S_normes peuvent remplacer le Max. dans les disjonctions.

Les T_normes et S_normes les plus utilisées sont données dans le tableau V-1.

Tableau V-1 T-normes et S-normes utilisées dans les modèles flous

Nom	T-norme	S-norme
Zadeh	$\min(x,y)$	$\max(x,y)$
Probabiliste	$x.y$	$x + y - xy$
Lukasiewicz	$\max(x + y - 1, 0)$	$\min(x + y, 1)$

V.3.5. Approximation par les modèles flous

Wang (1992) a montré que les systèmes flous sont des approximateurs universels s'ils respectent certains critères. D'un point de vue mathématique, la construction d'un système flou consiste à trouver une relation de correspondance entre l'espace d'entrée et l'espace de sortie pour approximer avec un degré de précision donné, la fonction ou la dynamique désirée.

Selon Buckley (1993), les modèles flous de Takagi-Sugeno sont des approximateurs et des contrôleurs universels.

Théorème

Soit f une fonction non linéaire définie sur un ensemble compact E . Pour toute constante positive ε il existe un système flou

$$\sum_f \text{ Tel que : } \sup_{e \in E} \left| f(e) - \sum_f(e) \right| < \varepsilon \quad (\text{V.28})$$

V.4. Modélisation par les systèmes flous

Hartani (1995) a proposé de diviser un travail de modélisation par un système flou en trois étapes : l'identification de sa structure externe, l'identification de sa structure interne, et l'approximation linguistique.

Identification de la structure externe:

Elle représente la première étape de la modélisation, et nécessite la détermination des variables d'entrée et de sortie à utiliser lors de la modélisation, cela correspond évidemment à la période de calage.

Le système flou utilise un concept d'apprentissage (similaire aux réseaux de neurones artificiels) qui à partir d'une série de données d'entrée, permet de calculer des valeurs et de les comparer aux données mesurées (données de sortie).

Identification de la structure interne:

Représente la base de règles floues, cette étape permet la détermination des différents éléments constituant le modèle flou, tels que le nombre de règles et sous-ensemble flou, le type de fonctions d'appartenance, ainsi que l'identification de leurs paramètres.

l'approximation linguistique:

Elle se base sur l'utilisation de variables linguistiques (Zadeh, 1975). Elle consiste à associer à tout ensemble flou une étiquette linguistique.

Les différents éléments intervenant dans la conception d'un système flou relèvent généralement de l'expérience du modélisateur, et les différentes questions relatives à la modélisation trouvent leur réponse d'une connaissance à priori établi par le concepteur.

V.4. Différents types de modèles flous

Nous pouvons classer les modèles flous en deux grandes catégories; les modèles linguistiques ou modèles de Mamdani, et les modèles à conséquences fonctionnelles ou modèles de Takagi-Sugeno-Kang.

V.4.1. Les modèles linguistiques (de Mamdani)

Basés sur l'approche de Zadeh, ces modèles permettent une description linguistique par une base de règles floues. Ils correspondent à l'utilisation d'un SIF pour modéliser les relations entrées/sorties d'un système.

Dans les modèles linguistiques, les quantités floues sont décrites par des termes linguistiques, et le modèle flou constitue une description du système exprimée dans une langue naturelle. Cette expression décrit le comportement du système en utilisant la langue sémantique claire au modèle. Leurs avantages sont:

Les prémisses et les conclusions des règles sont de nature entièrement linguistique. Ceci les rend facilement compréhensibles, et donne une interprétation sémantique claire au modèle.

Il est tout à fait possible de combiner des règles de ce type identifiées à partir de données par une méthode d'apprentissage, avec des règles formulées par un expert.

L'interpolation dépend de la forme des fonctions d'appartenance, des opérateurs d'inférence et de la fuzzification. Elle peut être améliorée par un choix adéquat de ces éléments. De ce point de vue, l'interpolation dans les modèles linguistiques peut être qualifiée de contrôlable.

En revanche les inconvénients des modèles linguistiques sont:

Leur nature non-structurée se traduit par une capacité de représentation limitée. En revanche d'autres termes, un grand nombre de règles est souvent nécessaire pour approcher le comportement global d'un système.

Le grand nombre de paramètres exige une base d'information importante et complète pour déterminer tous les paramètres.

V.4.2. Les modèles de Sugeno (Takagi-Sugeno-Kang)

Sugeno, Takagi et Kang (1985) ont proposé une autre classe de modèles où la partie conclusion utilise un formalisme d'état (Nakoula, 1997).

Les modèles de cette classe sont appelés ' Modèles de Sugeno ou de Takagi-Sugeno-Kang (TSK).

La particularité d'un modèle de Sugeno est que la logique floue est seulement utilisée dans la prémisse des règles. La partie conclusion est décrite par des valeurs numériques.

Exemple : Si x_1 est faible et x_2 est élevé Alors $y=z$(V.29)

Pour un système MISO à r entrées et une sortie, un modèle TSK est constitué d'une base de règles de la forme:

$$\text{Si } x_1 \text{ est } A_1^i \text{ et } x_2 \text{ est } A_2^i \text{ et } \dots x_r \text{ est } A_r^i \text{ alors } y = f_i(x_1, x_2, \dots, x_r)$$

(V.30)

Où $x_1, x_2 \dots x_r$ sont les variables numériques d'entrée du modèle flou et f est une fonction numérique de $X_1 \times X_2 \times \dots \times X_r$ dans Y .

Chacune des règles représente un modèle local sur une région floue d'entrée, ou sur un sous-espace d'entrée. Dans chaque région, le modèle est défini par la fonction f_i qui relie les entrées à la sortie numérique. Le modèle global est constitué par interpolation des modèles locaux. Dans ce cas, l'interpolation dépend des données d'entrées, et cela contrairement au modèle de Mamdani.

V.4.2.1. Méthode de calcul

Soit une entrée $(x_1, x_2 \dots x_r)$, le de calcul de la sortie y est composé de 3 étapes:

Calcul des sorties individuelles : la sortie de chaque règle i est donnée par sa conséquence, c'est à dire :

$$y_i = f_i(x_1, x_2 \dots x_r) \dots \dots \dots (V.31)$$

Calcul des degrés d'activation de chaque règle : le degré d'activation w_i de chaque règle i est calculé par l'équation :

$$w_i = \mu_{M_{X_1}}(A_1^i(x_1)) \dots \mu_{M_{X_r}}(A_r^i(x_r))$$

(V.32)

Agrégation des sorties individuelles : la valeur finale de la sortie résultante de l'ensemble des règles est donnée par la moyenne des sorties individuelles pondérées par le degré d'activation des règles, c'est à dire :

$$y^* = \frac{\sum_i w_i \cdot y_i}{\sum_i w_i}$$

(V.33)

Le dénominateur de cette équation étant une constante, on peut reformuler l'expression de l'ensemble y^* sous la forme :

$$y^* = \sum_i v_i \cdot y_i$$

(V.34)

$$\text{Où } v = \frac{w_i}{\sum_i w_i} . \text{ On peut noter que } \sum_i v_i = 1 .$$

Dans l'équation (V.31), deux cas peuvent se présenter :

- Les f_i sont des constants, dans ce cas le modèle est dit d'ordre zéro.
- Les f_i sont des fonctions linéaires, on aura donc un modèle d'ordre un, soit :

$$f_i(x_1, x_2, \dots, x_r) = b_{i0} + b_{i1}x_1 + b_{i2}x_2 + \dots + b_{ir}x_r$$

(V.35)

Où $b_{i0}, b_{i1}, \dots, b_{ir}$ sont des coefficients constants. Chaque règle peut alors être vue comme un modèle local, linéaire selon les variables d'entrée x_1, x_2, \dots, x_r .

Les opérateurs logiques dans le cas de Sugeno pour l'union et l'intersection sont respectivement :

$$\mu_{R \cup S}(u, v) = \mu_R(u, v) + \mu_S(u, v) - \mu_R(u, v) \cdot \mu_S(u, v)$$

(V.36)

Et :

$$\mu_{R \cap S}(u, v) = \mu_R(u, v) \cdot \mu_S(u, v)$$

(V.37)

Les modèles de Sugeno présentent certains avantages qu'on résume en :

Ils permettent de décrire des situations où la structure physique du système est bien connue, par exemple, une fonction linéaire exprimant la relation entrée-sortie.

Ils permettent d'approcher le comportement d'un système complexe par une collection de modèles locaux.

Les modèles TSK ont une capacité de représentation importante. Le nombre de règles nécessaires pour approcher un système à un certain degré de précision est généralement plus faible que dans le cas des modèles linguistiques.

La relation (V-33) relie directement la variable de sortie aux variables d'entrées, ce qui permet une application relativement facile des techniques d'optimisation pour identifier les différents paramètres du modèle.

Les modèles TSK ont certains inconvénients, qu'on résume en :

De part la nature fonctionnelle des conclusions, les règles n'ont pas de signification sémantique claire comme dans le cas des modèles linguistiques. Si un objectif de la modélisation est de fournir une meilleure compréhension de la relation entre les variables du système, ce type de modèle est donc inadapté.

Ces modèles ne permettent pas d'intégrer l'expérience formulée par des règles linguistiques.

La méthode de calcul ne permet pas de manipuler des entrées floues, car le calcul de la sortie exige les valeurs précises des variables d'entrées. Néanmoins une solution à ce problème a été proposée, elle consiste à utiliser les données floues pour calculer les degrés d'activation des règles, et de défuzzifier ces données sur leurs espaces respectifs pour obtenir leurs valeurs numériques, qui sont ensuite utilisées dans le calcul de la sortie.

V.5. Apprentissage des modèles flous

Dans le cadre de l'approximation de fonction entre l'entrée et la sortie du système, l'apprentissage (supervisé) consiste à définir la base de règles flous, c'est-à-dire leurs nombres, leurs prémisses et leurs conclusion, qui minimisent l'écart entre les sorties désirées et celles inférées par le système formé par de l'ensemble flou (Guillaume, 2001).

Le problème de modélisation floue, ou de construction de modèles flous peut être vu comme un processus d'identification de système flou. D'un point de vue conceptuel,

l'identification d'un système d'inférence flou peut être effectuée selon deux phases : l'identification de la structure du modèle, et l'estimation des paramètres du modèle à partir d'un ensemble de données.

V.5.1. Identification de la structure du modèle

Les systèmes d'inférence flous se basent sur l'induction des règles floues, qui à partir de ces règles permettent une formalisation des relations qui existent entre les entrées et les sorties du système modélisé. Dans ce cas, il est primordial de respecter la sémantique des règles, qui permet d'interpréter les sous-ensembles flous, en terme linguistique.

Si l'interprétabilité est plus ou moins importante (dans le cas de l'approximation), la cohérence reste une propriété essentielle de la base des règles, quelque soit l'utilisation du modèle. Elle signifie que les conclusions des règles simultanément activées, ne soient pas contradictoires (Guillaume, 2001).

L'identification de la structure du modèle est liée au partitionnement de l'univers de discours.

Une des techniques utilisées pour générer des règles floues est le partitionnement de l'espace multidimensionnel, ce partitionnement définit un certain nombre d'ensembles flous pour chacune des règles. L'apprentissage consiste à optimiser le partitionnement et les conséquences des règles en fonction des données.

Dans cette approche dite subjective, le modélisateur définit un nombre de termes linguistiques sur les espaces d'entrée et de sortie. Ces termes constituent le langage qu'il utilise pour décrire le système, et forment également les partitions des différents espaces. Les fonctions d'appartenance sont alors connues.

La deuxième approche est celle utilisant une coalescence floue (fuzzy clustering), dans ce cas l'identification de la structure sera faite à partir des données, la coalescence floue permet de partitionner l'espace des données en plusieurs classes. Chacune de ces classes ou régions flous est caractérisée par un vecteur appelé centre de classe.

L'utilisation d'un algorithme de coalescence floue a l'avantage essentiel de permettre la génération automatique de fonctions d'appartenance.

La technique de coalescence floue, est une méthode d'apprentissage non supervisé qui ne demande aucune connaissance à priori sur la structure des données. La méthode des C-moyennes floues (Dunn, 1973) représente un exemple de partitionnement par une coalescence floue.

V.5.2. Identification des paramètres

Dans notre travail, nous nous intéresserons qu'aux modèles de Sugeno (Takagi-Sugeno-Kang). A ce titre nous pouvons noter que plusieurs méthodes ont été élaborées pour l'identification des paramètres de ces modèles.

Les modèles de Sugeno permettent d'établir une relation analytique qui relie les

entrées à la sortie du système modélisé. L'estimation des paramètres du modèle se fait par l'introduction par des méthodes d'optimisation qui reviennent à minimiser un certain critère (Nakoula, 1997).

Les paramètres à identifier sont de deux types : les paramètres d'entrée et les paramètres de sortie. Les paramètres d'entrée sont ceux des fonctions d'appartenance dans les partitions des espaces d'entrée. On suppose que chacune de ces fonctions d'appartenance peut être décrite par p paramètres où p dépend de la forme de la fonction choisie.

Dans le cas de fonctions Gaussiennes, p est égal à 2 paramètres (moyenne et écart-type), dans le cas d'une fonction triangulaire, on a 3 paramètres (valeur modale et les deux extrémités du support).

Les paramètres de sortie sont ceux apparaissant dans les conclusions des règles.

Si l'on considère un modèle de Sugeno (d'ordre 1) à m règles, avec la définition de l_1, l_2, \dots, l_r termes linguistiques respectivement pour chaque variable x_1, x_2, \dots, x_r , le modèle est alors entièrement paramétré par le vecteur :

$$P = [a_{11}, \dots, a_{l_1 l_1}, a_{21}, \dots, a_{l_2 l_2}, \dots, a_{r1}, a_{r l_r}, b_1, b_2, \dots, b_m]^T \quad (V.38)$$

(V.38)

Avec a_{ij} vecteur des paramètres relatifs au $j^{\text{ième}}$ symbole défini pour la $i^{\text{ième}}$ variable,

$i=1..r, j=1, \dots, l_i$, soit :

$$a_{ij} = [a_{ij}^1, a_{ij}^2, \dots, a_{ij}^P]^T$$

(V.39)

Et b_i vecteur des $(r+1)$ coefficients relatifs à la $i^{\text{ième}}$ règle, $i=1 \dots m$, soit :

$$B_i = [b_{i0}, b_{i1}, \dots, b_{ir}]^T \quad (V.40)$$

Pour une entrée donnée, (x_1^k, \dots, x_r^k) , la sortie du modèle est

$$\hat{y}_k = \sum_{i=1}^m v_i \cdot y_i^k$$

(équation (V-40)). La différence entre cette sortie et celle du système pour la même entrée

$$(x_1^k, \dots, x_r^k) \quad e_k = \hat{y}_k - y_k$$

, l'erreur quadratique instantanée en ce point k par :

$$E_k = \frac{1}{2} \cdot e_k^2 = \frac{1}{2} (\hat{y}_k - y_k)^2$$

(V.41)

C'est une erreur instantanée en tout point. Elle constitue un premier critère d'optimisation pour les modèles TSK. Ce critère est en général utilisé pour un apprentissage en ligne à chaque fois qu'un nouveau point arrive.

Il existe plusieurs méthodes d'apprentissage, nous pouvons citer les plus utilisées telles que la méthode de descente de gradient (v. chapitre IV), ou le filtre de Kalman (1960) qui consiste à minimiser l'erreur quadratique moyenne. Cette minimisation est effectuée en considérant les fonctions d'appartenance des entrées fixées.

V.5.3. Optimisation neurofloue : le modèle ANFIS

Les réseaux de neurones artificiels présentent l'avantage d'être des approximateurs universels qui procèdent par apprentissage pour minimiser l'erreur entre un système d'entrée/sortie. Les techniques d'apprentissage neuronal utilisant différentes architectures neurofloues permettent d'allier le symbolique au numérique et d'en réduire les inconvénients mutuels.

C'est dans cette optique que plusieurs chercheurs (Glennec *et al.*, 1992, Bersini, 1993, Jang, 1993) ont exploité les réseaux de neurones artificiels et combiné les deux approches floue et neuronale d'où le modèle neuro-flou. Cette propriété permet alors d'exploiter des mécanismes d'optimisation numérique pour la synthèse de contrôleurs flous. Ainsi, Sugeno voit la solution dans les algorithmes des moindres carrés, alors que Bersini utilise une méthode de descente de gradient pour minimiser un critère quadratique (Boukezzoula, 2000).

Considérons un modèle TSK, composé des deux règles suivantes:

Si x_1 est A_1 et x_2 est B_1 alors $y_1 = f_1(x_1, x_2) = a_1 x_1 + b_1 x_2 + c_1$

Si x_1 est A_2 et x_2 est B_2 alors $y_2 = f_2(x_1, x_2) = a_2 x_1 + b_2 x_2 + c_2$

Le modèle neuroflou est représenté par le réseau adaptatif proposé par Jang (1993). Il s'agit d'un réseau adaptatif multicouche dont les connexions ne sont pas pondérées, ou ont toutes un poids de 1. Les nœuds sont de deux types différents selon leur fonctionnalité: les nœuds carrés (adaptatifs) contiennent des paramètres, et les nœuds circulaires (fixes) n'ont pas de paramètres. Toutefois, chaque nœud (carré ou circulaire) applique une fonction sur ces signaux d'entrées. La sortie O_i^k du nœud i de la couche k (appelé nœud (i,k)) dépend des signaux provenant de la couche $k-1$ et des paramètres du nœud (i,k) , c'est à dire:

$$O_i^k = f(O_1^{k-1}, \dots, O_{n_{k-1}}^{k-1}, a, b, c \dots)$$

(V.42)

Où n_{k-1} est le nombre de nœuds dans la couche $k-1$, et a, b, c, \dots sont les

paramètres du nœud (i, k) . Pour un nœud circulaire ces paramètres n'existent pas.

Ainsi les nœuds de chaque couche réalisent une fonction:

Couche 1: Chaque nœud est donc un nœud carré avec une fonction:

$$O_i^1 = \mu_{A_i}(x)$$

(V.43)

Où x est l'entrée du nœud i , et A_i le terme linguistique associé à chaque fonction. En d'autres termes, O_i^1 est le degré d'appartenance de x à A_i . Les paramètres d'un nœud de cette couche sont ceux de la fonction d'appartenance correspondante.

Couche 2 : Chaque nœud i de cette couche est un nœud circulaire appelé \square qui engendre en sortie le produit de ses entrées. Ce produit représente le degré d'activation d'une règle:

$$w_i = \mu_{A_i}(x_1) \cdot \mu_{B_i}(x_2), i = 1, 2$$

(V.44)

Couche 3: Chaque nœud de cette couche est un nœud appelé N. La sortie du nœud i est le degré d'activation normalisé de la règle i :

$$v_i = \frac{w_i}{w_1 + w_2} \quad (V.45)$$

Couche 4: Chaque nœud de cette couche est un nœud carré avec une fonction réalisant le calcul:

$$O_i^4 = v_i \cdot f_i = v_i (a_i x_1 + b_i x_2 + c_i), i = 1, 2$$

(V.46)

Où v_i est la sortie de la couche 3, et $\{a_i, b_i, c_i\}$ est l'ensemble des paramètres de sortie de la règle i .

Couche 5: Le seul nœud de cette couche est un nœud circulaire qui effectue la somme des signaux provenant de la couche 4, c'est à dire:

$$O_i^5 = y = \sum_i v_i \cdot f_i$$

(V.47)

La généralisation du réseau à un système à r entrées ne pose aucun problème particulier. Le nombre de nœuds de la couche l est toujours égal au nombre total de termes linguistiques définis.

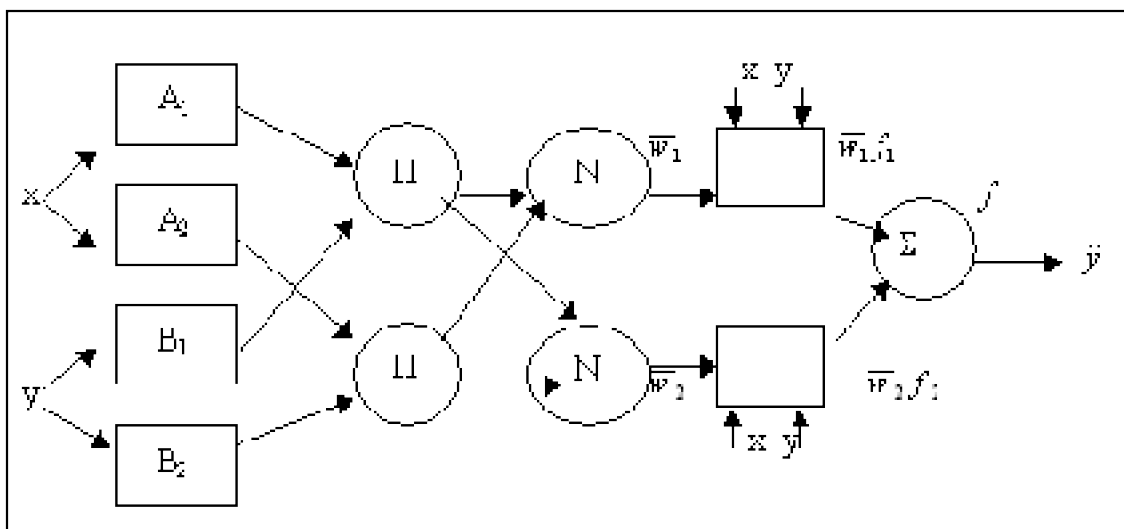


Figure V-8 Réseau adaptatif du modèle ANFIS

L'apprentissage à partir d'un ensemble de données concerne l'identification des paramètres des prémisses et des conséquences, la structure du réseau étant fixée. L'algorithme d'apprentissage commence par construire un réseau initial, ensuite on applique une méthode d'apprentissage par rétro-propagation de l'erreur. Jang (1993) a proposé d'utiliser une règle hybride qui combine un algorithme de descente de gradient avec une estimation par moindres carrées.

V.6. Conclusion

Dans ce chapitre, nous avons abordé la théorie floue, ainsi que les définitions nécessaires au fonctionnement d'un système d'inférence floue.

Nous avons ainsi détaillé les différentes classes des systèmes d'inférence floue (SIF), et les détails de leur conception.

Notre démarche a porté sur le choix de représentation des modèles flous, nous avons également consacré une grande partie à la description des modèles entrée-sortie au lieu des modèles d'état.

Les modèles flous de type Sugeno présentent un grand intérêt dans l'extrapolation des données à travers un système entrée/sortie. En effet ils présentent une structure beaucoup plus simple que les modèles de Mamdani.

Cependant, la procédure d'apprentissage que nous avons étudiée, indispensable à l'identification des paramètres, est relativement complexe, en effet l'architecture ANFIS qui associe deux modèles d'apprentissage (neuronal et moindres carrés) présente une certaine difficulté dans sa conception, mais reste nécessaire au bon fonctionnement du modèle.

L'étude de la relation pluie-débit aux pas de temps mensuel et journalier par un modèle flou sera détaillée dans le chapitre VI.

Chapitre VI : Résultats de la simulation par le modèle flou, comparaison entre divers modèles

VI.1. Introduction

La complexité de la relation pluie-débit peut être caractérisée par plusieurs éléments: le nombre de facteurs qui la composent, l'interaction entre ces facteurs, leurs caractéristiques, la dynamique de cette relation ... etc.

Ainsi, de nombreux modèles ont été conçus en vue d'une meilleure compréhension de la relation pluie-débit et une reconstitution acceptable des séries hydrométriques à différents pas de temps.

La logique floue représente une nouvelle technique de simulation, qui manipule des concepts linguistiques comme 'faible', 'moyen', 'élevé, et fournit un cadre conceptuel et opérationnel utile pour chacune des problématiques de la simulation pluie-débit. Elle peut constituer une interface commode pour la mise en forme d'appréciations du processus pluie-débit, en proposant un cadre de fusion d'informations et d'agrégations multicritères paramétrables, robuste et éprouvé dans un large champ d'applications.

Dans ce chapitre, il sera question de l'application d'un modèle flou au processus pluie-débit. L'objectif principal est de tester cette nouvelle technique, en vue d'une amélioration de la qualité de la simulation, soit à travers une comparaison des différents modèles testés (conceptuel, neuronal et flou), soit par une combinaison des ces trois approches à deux échelles mensuelle et journalière.

VI.2. Exemple de simulation par la logique floue : cas d'une relation pluie-débit

La modélisation pluie-débit par la logique floue peut être considérée comme une opération d'extrapolation à travers un système d'entrée/sortie. La première étape d'une simulation par les systèmes d'inférence floue consiste à déterminer les variables d'entrée et de sortie.

Les pluies et les débits à l'échelle mensuelle seront utilisés pour cet exemple de simulation.

Dans ce cas, la règle floue sera de la forme:

Si la pluie est A Alors le débit est B.

Un exemple est donné par la règle suivante:

Si la pluie est forte Alors le débit est élevé.

Une simulation par un modèle flou est traditionnellement décomposée en trois parties essentielles: la fuzzification, l'inférence floue, et la défuzzification.

a- Fuzzification

C'est la première étape de simulation floue, elle permet de modéliser les entrées d'un système principalement sous forme de courbes, elle est effectuée par le partitionnement flou.

L'extrapolation à partir d'un jeu de règles nécessite l'utilisation d'un ensemble de termes linguistiques (définis dans le chapitre V), tels que '*faible*', '*moyen*', '*élevé*'.

Dans notre exemple nous avons considéré:

- L'univers du discours pour les variables d'entrée (Pluie): [0, 50], et l'univers du discours pour les débits [0, 30];
- Le nombre de règle est de trois (3) règles, chaque variable sera divisée en trois zones;
- Les fonctions d'appartenance sont trapézoïdales pour les variables d'entrée (pluie), et triangulaires pour les variables de sortie (débit).

On définit donc des zones de pluie et des zones de débits (Figure VI-1).

Dans le cas du sous-ensemble flou '*moyenne*' pour la variable entrée (pluie), veut

dire que c'est une zone de pluie, qui commence à 10mm, qu'elle est pleinement réalisée entre 20mm et 30mm, et qu'elle cesse à partir de 40mm.

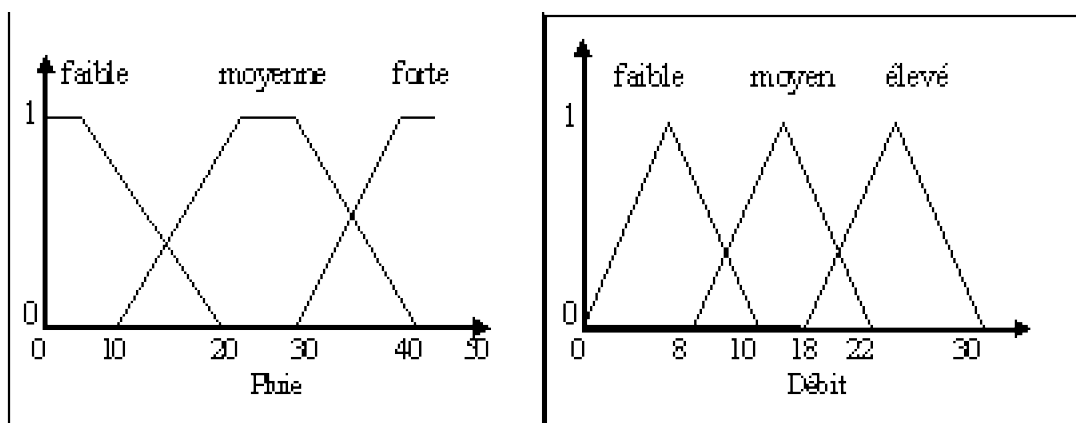


Figure VI-1 Partitionnement flou à 3 sous-ensembles flous

b- Règles d'inférence floue et calcul des degrés de vérités

A la présentation de chaque exemple (valeur de pluie), en fonction des règles d'inférence floue, on détermine son degré d'appartenance à un sous-ensemble flou. Dans notre cas, nous prenons en considération la valeur de pluie 18mm.

Ainsi, en fonction de l'appartenance de la valeur $x=18\text{mm}$ aux différents sous-ensembles flous, nous aurons les règles suivantes:

- Il est vrai à 33 % (degrés = 0.33) que la valeur 18 mm appartient au sous-ensemble flou 'faible'.
- Il est vrai à 67 % (degrés = 0.67) que la valeur 18 mm appartient au sous-ensemble flou 'moyenne'.
- Il est vrai à 0 % (degrés = 0) que la valeur 18 mm appartient au sous-ensemble flou 'forte'.

Pour chaque degré d'appartenance calculé, on engendre des surfaces d'inférence (Figure VI-2), ce qui permet de délimiter des \square -coupes pour chaque sous-ensemble flou.

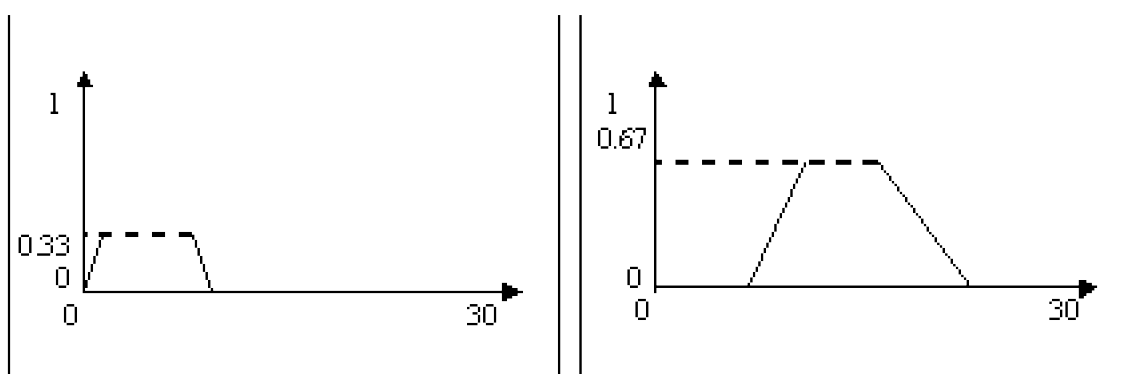


Figure VI-2 Calcul des \square -coupes sur la variable de sortie (Débit)

c- Défuzzification

C'est la dernière étape de simulation par les systèmes d'inférence flous de type Mamdani, elle permet d'obtenir une valeur physique par le modèle. La méthode de centre de gravité permet de calculer la valeur simulée (Figure VI-3).

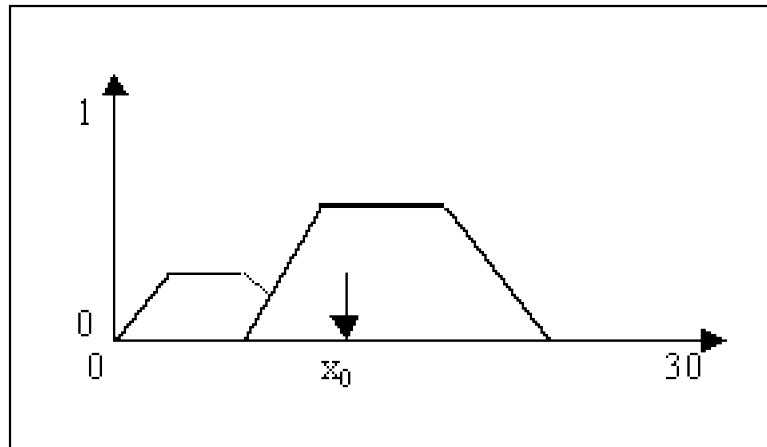


Figure VI-3 Défuzzification sur l'axe des débits par la méthode du centre de gravité

VI.2.1. Modèle utilisé

La description précédente caractérise le modèle flou de Mamdani, très utilisé dans le cas des contrôleurs et des capteurs flous.

Dans le cadre de notre travail, il sera question du modèle de Sugeno (modèle TSK) qui ne nécessite pas à priori aucune connaissance sur la relation entrée-sortie (Gloennec et Orseau, 1992).

Soient la variable d'entrée x (pluie), la variable de sortie y (débit), soit l'appréciation A , la règle floue i s'exprime comme suit:

$$\text{Si } x_i \text{ est } A_i \text{ Alors } y_i = ax_i + b$$

Avec $i=1..k$ règles floues.

Ce modèle se traduit donc par une interpolation de plusieurs modèles locaux ou «mutli-linéaires», ce qui permet d'accroître la précision du système lorsque le nombre de règles i augmente (Figure VI-4).

Le modèle neuro-flou ANFIS par sa robustesse sera utilisé dans notre étude. Ce modèle représente un système d'inférence flou organisé en réseau adaptatif.

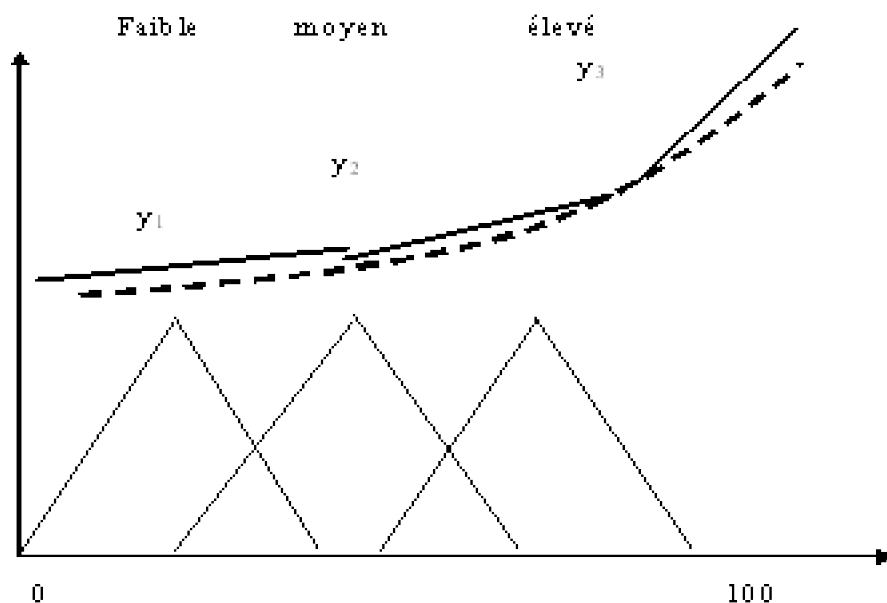


Figure VI-4 Exemple d'un modèle de Sugeno d'ordre 1 à trois règles

Ce modèle combine la logique floue et une structure neuronale à apprentissage supervisé (JANG, 1993) qui permet l'estimation des paramètres non-linéaires (de prémisse). Les paramètres linéaires (de conclusion) sont ajustés par la méthode des moindres carrés.

Pour ce qui des fonctions d'appartenance, l'utilisation des fonctions gaussiennes a un intérêt du fait du nombre de paramètres à optimiser,

soit μ la moyenne et δ l'écart-type de cette fonction

.Dans le cas des fonctions triangulaires ou trapézoïdales le nombre de paramètres augmente (3 ou 4 paramètres).

VI.3. Résultats de la simulation

Le modèle neuro-flou (ANFIS) sera testé aux deux pas de temps : mensuel et journalier.

Pour l'estimation des paramètres, le procédé est le même que celui des modèles conceptuels, c'est à dire que les modèles seront testés sur une période longue de seize ans, dans laquelle la série modélisée est divisée en deux sous-échantillons de durées homogènes, le premier servant de calage, le deuxième pour la validation du modèle, ce qui faciliterait la comparaison des résultats entre les deux périodes.

Les critères d'appréciation des modèles, sont les mêmes que ceux des modèles conceptuels, à savoir l'efficacité de Nash (%), et le rapport de reproduction des débits d'étiage, pour juger de la qualité de simulation des débits d'étiage.

VI.3.1. Données mensuelles

Nous avons cinq stations hydrométriques à modéliser à savoir, les stations de Cheffia, Lakhdaria, La Traille, Sebbaou et la station de la Tafna.

Trois modèles seront testés pour chaque station, un modèle avec un nombre de cinq (5) règles, ce qui représente un nombre de règles relativement faible, un nombre de douze règles (12) et le troisième nombre sera en fonction de chaque station à modéliser.

Après calage des modèles et après augmentation du nombre de règles pour chaque station, nous avons remarqué qu'à partir de trente (30) règles floues, l'optimisation des modèles est réalisée, et que toutes les séries présentent des débits modélisés assez satisfaisants.

Stations	Cheffia (1978-84)		Lakhdaria (1980-87)		La Traille (1970-77)		Sebbaou (1980-87)		Tafna (1983-1987)	
	Nash (%)	CET (%)	Nash (%)	CET (%)	Nash (%)	CET (%)	Nash (%)	CET (%)	Nash (%)	CET (%)
Flou (5)	71.1	262	60.5	462	55.4	366	52.0	960	49.7	756
Flou (12)	86.0	268	74.5	365	78.2	361	66.8	880	83.8	833
Flou (30)	92.6	225	87.4	337	84.4	311	75.3	760	92.5	742

Tableau VI-1 Résultats du modèle flou au pas de temps mensuel (calage)

Les résultats des différents modèles de la période de validation en fonction de chaque station sont présentés dans le tableau VI-2.

Stationne	Cheffia (1985-91)		Lakhdaria (1988-95)		La Traille (1979-85)		Sebbaou (1988-95)		Tafna (1988-1995)	
	Nash (%)	CET (%)	Nash (%)	CET (%)	Nash (%)	CET (%)	Nash (%)	CET (%)	Nash (%)	CET (%)
Flou (5)	47.7	132	79.2	141	28.1	159	63.8	414	35.6	82
Flou (12)	58.0	128	87.7	136	46.6	128	70.0	323	55.3	82
Flou (30)	80.5	110	91.2	134	50.1	130	83.0	270	68.8	90

Tableau VI-2 Résultats du modèle flou au pas de temps mensuel (validation)

D'après les résultats des tableaux VI-1 et VI-2, nous remarquons que les modèles donnent des résultats sensiblement différents, en fonction du nombre de règles floues et pour diverses stations. En effet, pour le modèle flou à 5 règles, les résultats sont peu performants en période de calage, l'efficacité de Nash est en moyenne de 60%, alors qu'en période de validation les résultats se détériorent. Nous observons donc une baisse de la qualité de la modélisation, excepté la station de Lakhdaria dont les résultats s'améliorent et atteignent près de 80%.

En augmentant le nombre de règles à 25, les résultats s'améliorent considérablement, le critère de Nash est en moyenne de 85%, la station de Cheffia enregistre la valeur la plus performante avec un critère de Nash de 92.6%.

Ces résultats confirment la robustesse du modèle en période de validation, en moyenne les résultats s'améliorent de 44%.

Le modèle flou à 30 règles représente le modèle optimal pour l'ensemble des stations, ceci est confirmé par les bons résultats obtenus pour les cinq stations modélisées.

Ainsi, les stations de Cheffia, Lakhdaria (Figure VI-4) et Tafna présentent des performances assez élevées, les valeurs du critère de Nash dépassent la barre des 90%, les débits sont ainsi reproduits de façon correcte, ainsi que pour la station du Sebbaou. Nous pouvons considérer les résultats comme satisfaisants, et cela, vu la grande amplitude que présentent les débits de cette station.

En période de validation, il est à remarquer une baisse des résultats moins importante que pour les modèles à 5 et 25 règles floues, excepté la station de La Traille, où l'efficacité de Nash ne dépasse pas 51% en phase de validation.

Pour ce qui est de la reproduction des débits faibles, le rapport des cumuls d'étiages varie considérablement d'une station à l'autre. Néanmoins, nous constatons qu'en période de calage, le rapport CET est largement supérieur à 100, ce qui démontre une mauvaise reproduction des débits. Ceci est observé en particulier pour la station du Sebbaou, où le rapport CET dépasse 1500.

En période de validation, une meilleure simulation des débits d'étiage est remarquée pour l'ensemble des stations, où les valeurs du CET se rapprochent de 100% essentiellement pour les modèles flous à 12 et 30 règles, on enregistre un rapport de CET égal à 110% pour la station de la Cheffia.

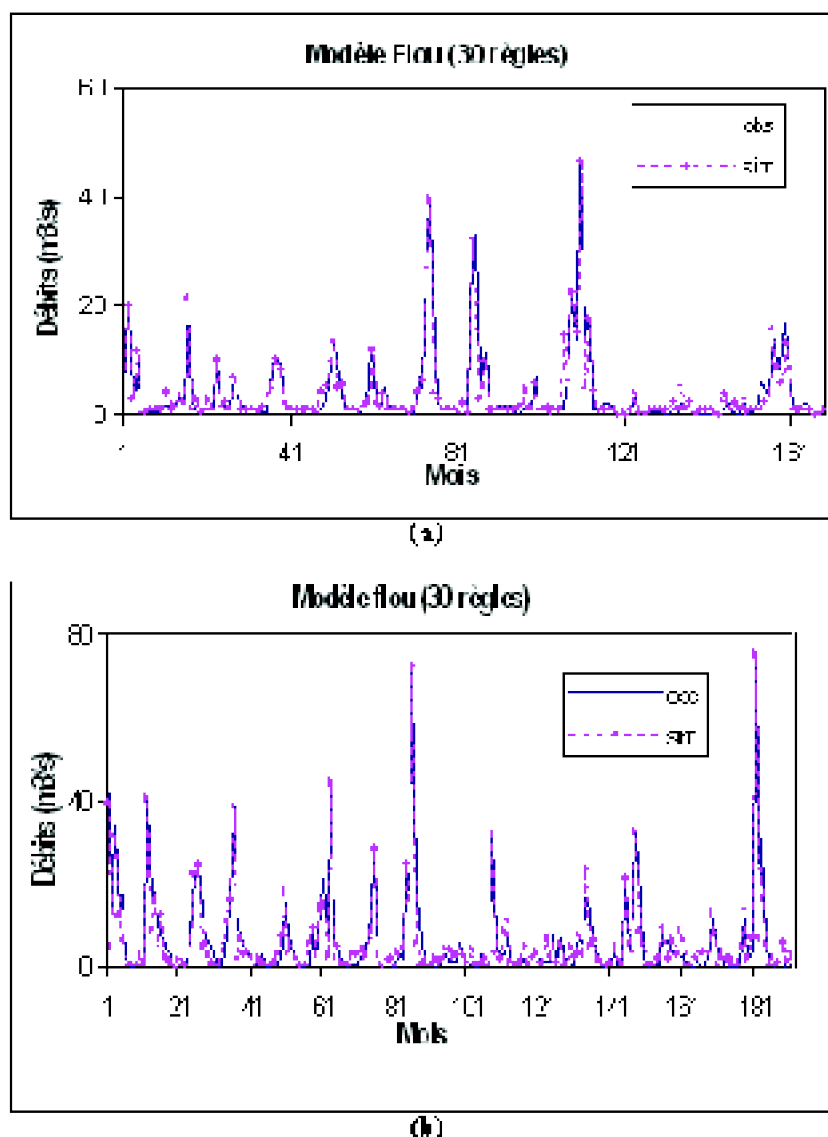


Figure VI-5 Résultats de simulations mensuelles du modèle flou (30 règles),
(a) station de la Cheffia, (b) station du Sebbaou

VI.3.2. Données journalières

Au pas de temps journalier, les modèles seront testés sur trois stations hydrométriques, la station de Cheffia, la station de l'Isser et celle de Zardezas.

La procédure de validation des modèles est identique à celle des modèles conceptuels. Les données seront divisées en deux périodes, la période de calage et une autre pour la validation du modèle. Dans le cas des données journalières, nous disposons d'un échantillon de six ans, et cela pour toutes les stations à modéliser.

Trois modèles seront testés pour chaque station, un modèle avec un nombre de cinq (5) règles, un nombre de vingt cinq règles (25), ainsi qu'un nombre de quatre vingt (80) règles floues. Ce dernier nombre paraît assez important, mais vu la complexité de la

relation pluie-débit au pas de temps journalier et le nombre de données à traiter ainsi que l'univers de discours qui est largement supérieur à celui des données mensuelles. Dans ce cas, l'optimisation nécessite un nombre de règles relativement plus élevé.

Les critères d'appréciation des modèles sont ceux appliqués aux modèles conceptuels, à savoir le critère de Nash, l'indice de reproduction du débit maximal, et le rapport des débits d'étiages.

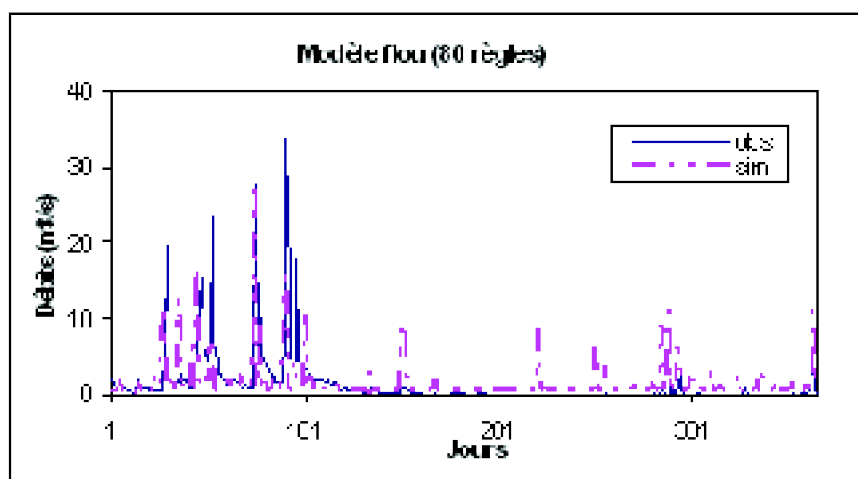
Stations	Cheffia (1986-88)			Isser (1990-92)			Zarzedas (1991-93)		
	Nash (%)	Ic	CET (%)	Nash (%)	Ic	CET (%)	Nash (%)	Ic	CET (%)
Flou (5)	38.3	1.29	1258	4.37	0.22	1600	53.6	0.99	1200
Flou (25)	55.7	1.18	1248	13.4	0.19	1564	64.5	1.0	1200
Flou (80)	61.2	1.0	1200	29.9	0.29	1500	80.5	0.99	1100

Tableau VI-3 Résultats des modèles journaliers en période de calage

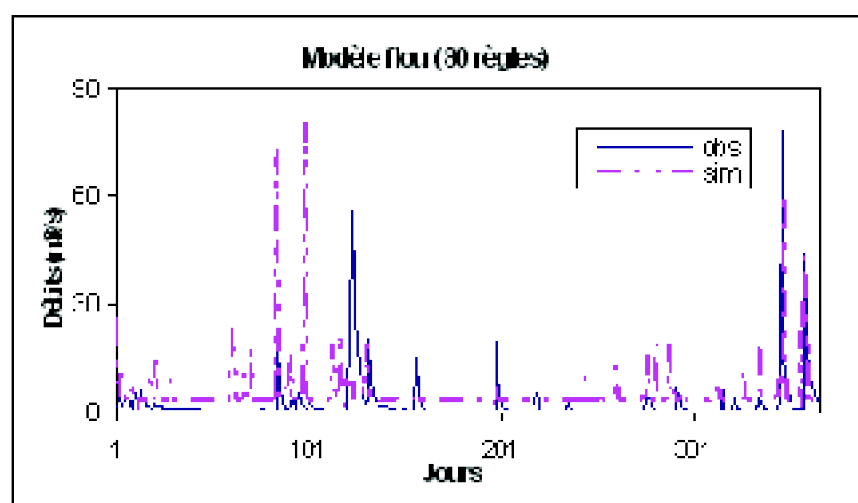
Stations	Cheffia (1989-91)			Isser (1993-95)			Zarzedas (1994-96)		
	Nash (%)	Ic	CET (%)	Nash (%)	Ic	CET (%)	Nash (%)	Ic	CET (%)
Flou (5)	42.4	1.0	70	14.9	0.59	205	49.6	1.0	167
Flou (25)	52.8	1.18	73	31.4	0.93	195	59.0	1.0	164
Flou (80)	64.9	1.0	62	41.8	0.93	203	65.5	1.0	145

Tableau VI-4 Résultats des modèles journaliers en période de validation

En examinant les résultats des tableaux Tableau VI-3 et VI-4, nous remarquons que les performances des modèles neuro-flous augmentent sensiblement avec le nombre de règles floues. Ainsi, pour la station de Cheffia, les résultats du modèle flou à règles sont nettement supérieurs que ceux du modèle flou à 5 règles, le critère de Nash passe de 38% à 61%. Ce phénomène est remarqué pour toutes les stations, la valeur maximale est celle obtenue par la simulation des débits de la station de Zarzedas (Figure VI-5 (a)), où l'efficacité de Nash dépasse 80% en période de calage.



(a)



(b)

Figure VI-6 Résultats de simulation par le modèle flou, a) station de Zardezas (année 1991),

(b) station de l'Isser (année 1991)

En période de validation les résultats sont parfois meilleurs que ceux obtenus dans la phase de calage.

Cependant, les résultats obtenus pour l'ensemble des modèles, sont nettement moins performants que ceux des modèles conceptuels, que ce soit en période de calage ou en période de validation.

Pour ce qui est des débits d'étiage, excepté la station de Zardezas en période de validation, le rapport CET n'est toujours pas optimisé, et ses valeurs sont largement aberrantes.

Néanmoins, les modèles présentent un grand avantage de simulation du débit maximal, pratiquement, pour tous les modèles testés, l'indice I_c est généralement égal à un (1), ce qui suppose une bonne aptitude du modèle flou à reconstituer le débit maximal

observé.

Pour ce qui est des étiages, les résultats montrent une nette différence entre la phase de calage où les débits sont nettement surestimés, et celle de la validation, dans cette période, le rapport CET est relativement correct, CET est égal à 145% pour la station de Zardezas.

Il nous a paru donc, que les modèles neuro-flous (et neuronaux) présentent des problèmes de reconstitution de certaines valeurs qui se répètent successivement en particulier les valeurs nulles (pluie nulle).

En effet, les modèles se trouvent face à des cas difficilement modélisables (ex: pluie nulle débit faible, pluie nulle débit élevé).

Ainsi, les diverses tentatives que nous avons menées en vue d'optimiser les débits journaliers par le modèle flou en augmentant le nombre de règles restent vaines. Le modèle donne une mauvaise simulation des débits de décrue (décrues brusques), et une surestimation des débits d'étiages.

Ainsi, on conclut que la modélisation pluie-débit à l'échelle journalière par le modèle neuro-flou (et neuronal) à une seule entrée (pluie) reste insuffisante pour caractériser les différents cas qui peuvent se présenter au modèle.

VI.3.3. Résultats du modèle flou par périodes caractéristiques

Les variations climatiques, représentent l'épreuve majeure que doit affronter un modèle pluie-débit, et cela pour juger de sa capacité d'adaptation.

Cette procédure de simulation, nous permet de tester le modèle sur un éventail de données relativement court, et d'enduire la sensibilité du modèle flou aux différents cycles hydrologiques.

Le modèle flou sera testé sur les séries courtes des stations du Sebbaou et de Zardezas. Les résultats sont détaillés dans les tableaux VI-5 et VI-6.

Sebbaou	Calage (1993)			Validation (1994)		
	Nash (%)	Ic	CET (%)	Nash (%)	Ic	CET (%)
Flou (5)	4.6	0.14	2990	0.10	0.1	286
Flou (25)	11.8	0.5	2900	4.60	0.2	280
Flou (80)	23.4	0.6	2900	10.0	0.3	267

Tableau VI-5 Résultats des modèles flous par périodes caractéristiques (Sebbaou)

Zardezas	Calage (1994-95)			Validation (1996-97)		
	Nash (%)	Ic	CET (%)	Nash (%)	Ic	CET (%)
Flou (5)	54.7	0.99	953	50.3	1.0	186
Flou (25)	81.5	1.0	929	66.6	1.0	175
Flou (80)	86.1	1.0	900	72.0	1.0	180

Tableau VI-6 Résultats des modèles flous par périodes caractéristiques (Zardezas)

Il apparaît clairement dans les tableaux VI-5 et VI-6 que les résultats obtenus pour la station du Sebbaou sont très médiocres quelque soit le nombre de règles introduites dans le modèle, et cela pour les deux périodes de calage et de validation. Le critère de Nash ne dépasse par 24 % en phase de calage pour le modèle à 80 règles.

En revanche pour la station de Zardezas, les résultats sont nettement meilleurs en particulier pour le modèle à 80 règles, où on obtient une efficacité de Nash de 86%, ce qui est très correct. Cependant, en période de validation, une baisse de performance est à signaler, elle atteint en moyenne 18%.

Pour ce qui est des débits caractéristiques, l'indice *Ic* est correct, quelque soit le modèle, en effet la reproduction du débit maximal est pratiquement systématique.

Néanmoins, les débits d'étiage sont surestimés, le rapport CET est supérieur à 100%, ce qui indique une mauvaise simulation des débits faibles.

En conclusion, la simulation pluie-débit journalière par un modèle flou à partir des pluies observées seulement est relativement délicate, ceci est dû à la multiplication des cas telles que les valeurs nulles (jours secs).

VI.4. Modèle flou à deux entrées

Les modèles pluie-débit, simulent le processus pluie-débit à partir d'une chronique d'observations de pluie plus ou moins longue. Les modèles doivent ainsi reproduire les divers cycles de sécheresse et d'humidité que caractérisent la série à modéliser.

Pratiquement pour tous les modèles conceptuels, le débit calculé est fonction de deux états, de la pluie à l'instant (*t*), et du niveau du réservoir observé à l'instant (*t-1*), soit:

$$Q(t) = F(P(t), R(t-1)) \dots \dots \dots \text{VI-1}$$

Où $R(t-1)$, est le niveau du réservoir de routage (cas des modèles GR).

Pour un modèle flou (ou neuronal), la simulation se fait en mode statistique, c'est à dire sans connaissance de l'état du jour précédent, de ce fait le débit calculé n'est fonction que de la pluie à l'instant t, soit:

$$Q(t) = F(P(t)) \dots \dots \dots \text{VI-2}$$

Cependant, les jours les plus fréquents en hydrologie, représentent le jour sec, (Pluie nulle). La problématique se pose donc dans la répétition de ces valeurs que le modèle ne peut modéliser.

Ainsi, avec la succession des jours humides et secs, les niveaux d'humidité des sols, et donc l'intensité des débits diffèrent d'une pluie à l'autre. Après l'arrêt des précipitations, plusieurs cas peuvent se présenter pour le modèle flou (neuronal), ces cas sont résumés comme suit :

Après une pluie intense, Pluie = 0 → Débit élevé

Après une pluie moyenne, Pluie = 0 → Débit moyen

En période d'étiage, pluie = 0 → Débit nul

Donc, le résultat simulé pour ces valeurs nulles, est la moyenne de tous les débits observés pour le cas des valeurs nulles (pluie = 0). De ce fait une baisse de la qualité de simulation est remarquée, ce qui engendre une grande perte du critère de Nash, pour les deux phases de calage et de validation.

Ainsi l'introduction d'une deuxième entrée qui va orienter le modèle est d'une nécessité absolue, elle permettra ainsi de caractériser les différents cas qui peuvent se présenter au modèle.

Pour cela, plusieurs variables hydrométéorologiques peuvent être prises en considération à savoir la température de l'air, l'humidité relative, l'évapotranspiration potentielle, etc.

Il nous est paru après réflexion, que l'évapotranspiration (ETP) et l'humidité du sol, qui jouent un rôle majeur dans le bilan hydrologique peuvent être introduites comme des variables principales (explicatives) au modèle neuro-flou.

VI.4.1. Introduction de l'ETP

L'évapotranspiration est définie comme étant la quantité d'eau évaporée par le sol et transpirée par les plantes. Elle agit directement dans le bilan énergétique du sol, elle permet entre autres, d'extraire l'eau contenue dans le sol et de la rejeter dans l'atmosphère sous forme gazeuse.

Dans cette tentative, nous essayerons d'étudier l'influence de l'évapotranspiration (valeurs mesurées) comme entrée décisive permettant de simplifier la complexité pluie-débit journalière. L'objectif est d'améliorer la qualité de simulation pluie-débit par une prise en compte de l'évapotranspiration potentielle (ETP).

En effet, la règle floue ainsi obtenue aura la forme suivante :

Si la pluie est nulle Ou l'évapotranspiration est faible Alors le débit est élevé.

Si la pluie est nulle Ou l'évapotranspiration est moyenne Alors le débit est moyen.

Si la pluie est nulle Ou l'évapotranspiration est élevée Alors le débit est faible

Il est à préciser que le nombre total de règles NR est égal à :

$$NR = N^P$$

Où N représente le nombre de règles effectives (pour chaque entrée) et P est le nombre d'entrées.

Dans notre étude nous testons trois modèles: un modèle à 25 règles (5^2), un modèle à 64 règles (8^2) et un troisième à 121 règles (11^2). Ce qui permet d'étudier au mieux l'influence du nombre de règles sur l'optimisation des résultats.

Les résultats du modèle flou* à deux entrées (Pluie, ETP) des trois stations journalières, respectivement pour la période de calage et de validation sont présentés dans les tableaux VI-7 et VI-8.

Stations	Cheffia (1986-88)			Isser (1990-92)			Zarzedas (1991-93)		
	Nash (%)	Ic	CET (%)	Nash (%)	Ic	CET (%)	Nash (%)	Ic	CET (%)
Flou (5^2)	50.2	1.13	1100	8.17	0.16	1300	55.5	1.0	1000
Flou (8^2)	57.1	1.0	1140	7.18	0.15	1300	65.4	0.99	892
Flou (11^2)	61.5	0.99	900	12.0	0.14	1213	76.2	1.0	900

Tableau VI-7 Modèle flou* appliqué aux données journalières en période de calage

Stations	Cheffia (1989-91)			Isser (1993-95)			Zarzedas (1994-96)		
	Nash (%)	Ic	CET (%)	Nash (%)	Ic	CET (%)	Nash (%)	Ic	CET (%)
Flou (5^2)	48.4	1.0	70	24.6	0.88	154	61.9	1.0	100
Flou (8^2)	66.2	1.0	70	31.7	0.93	152	66.7	1.0	102
Flou (11^2)	71.3	1.0	71	43.0	0.93	153	66.7	1.0	103

Tableau VI-8 Modèle flou* appliqué aux données journalières en période de validation

D'après les résultats des tableaux VI-7 et VI-8, nous remarquons qu'en effet le nombre de règles permet d'améliorer les résultats, le modèle à 121 règles est largement optimal par rapport à celui programmé à 25 règles (Nash =57%) en période de calage. Le critère de Nash s'améliore légèrement pour le cas de l'Isser en période de validation, mais

n'excède pas la barre des 50%.

Pour la station de Cheffia (Figure VI-6), on enregistre une efficacité de Nash la plus performante pour le modèle à 121, avec un résultat maximal pour l'année 1987 (Figure VI-10), où la valeur du critère de Nash est de 74.2%.

Cependant, nous n'avons pas obtenu une amélioration significative de la modélisation, les résultats obtenus n'ont pas été à la hauteur de ceux attendus, et cela quelque soit le nombre de règles utilisé.

En effet, l'introduction de l'évapotranspiration ne permet en aucun d'orienter le modèle aux différents cas étudiés. Cette donnée, par son caractère cyclique présente une faible amplitude, et les pertes par évaporation ont un rôle mineur en période de crue et de hautes eaux.

Le modèle est donc peu sensible à l'ETP et ne permet pas de calculer les débits de façon précise, ceci est du notamment aux règles contradictoires.

Un exemple est donné dans le cas où la pluie est nulle et l'évapotranspiration est faible, le débit peut être dans ce cas là très faible, ce qui constitue une contradiction avec la règle générale floue.

VI .4.2. Introduction de l'humidité du sol

L'humidité du sol joue un rôle majeur à l'interface entre l'atmosphère et les surfaces continentales. En caractérisant les flux d'eau et d'énergie dans le continuum, elle représente une variable importante en hydrologie qui détermine la réponse d'un bassin.

L'humidité contenue dans le sol participe à l'alimentation du cours d'eau, et cela par un écoulement quasi-horizontal et souterrain. Entre autres, elle peut caractériser le niveau du fleuve, et son niveau évolue proportionnellement à celui du sol.

Dans les modèles pluie-débit, les précipitations se répartissent en fonction de l'état hydrique du sol entre évapotranspiration, ruissellement de surface, infiltration, stockage de l'eau dans le sol.

De nombreuses méthodes expérimentales existent pour estimer l'humidité du sol. Cependant, le caractère ponctuel de ces mesures associé à la forte variabilité spatiale de l'état hydrique de surface au sein d'une région en réponse à des différences de contraintes atmosphériques, de types d'occupation du sol, de travail du sol ou même de nature de sol impliquerait la mise en place d'expérimentations fastidieuses pour obtenir des indices moyens représentatifs à l'échelle du bassin versant (Quesney, 1999).

Vu la difficulté d'accéder aux valeurs d'humidité réelles (mesurées) du sol, nous avons pris en considération des contenus en eau du sol obtenus par simulation d'un modèle pluie-débit conceptuel.

Dans cette opération nous avons introduit au modèle flou (en plus de la pluie observée) le niveau du sol simulé par le modèle GR3J. Ce choix est justifié par le fait que ce modèle a donné de bons résultats, et que le réservoir Sol simule de façon assez correcte le niveau d'humidité évoluant dans le sol.

L'organigramme du modèle est présenté dans la figure VI-6:

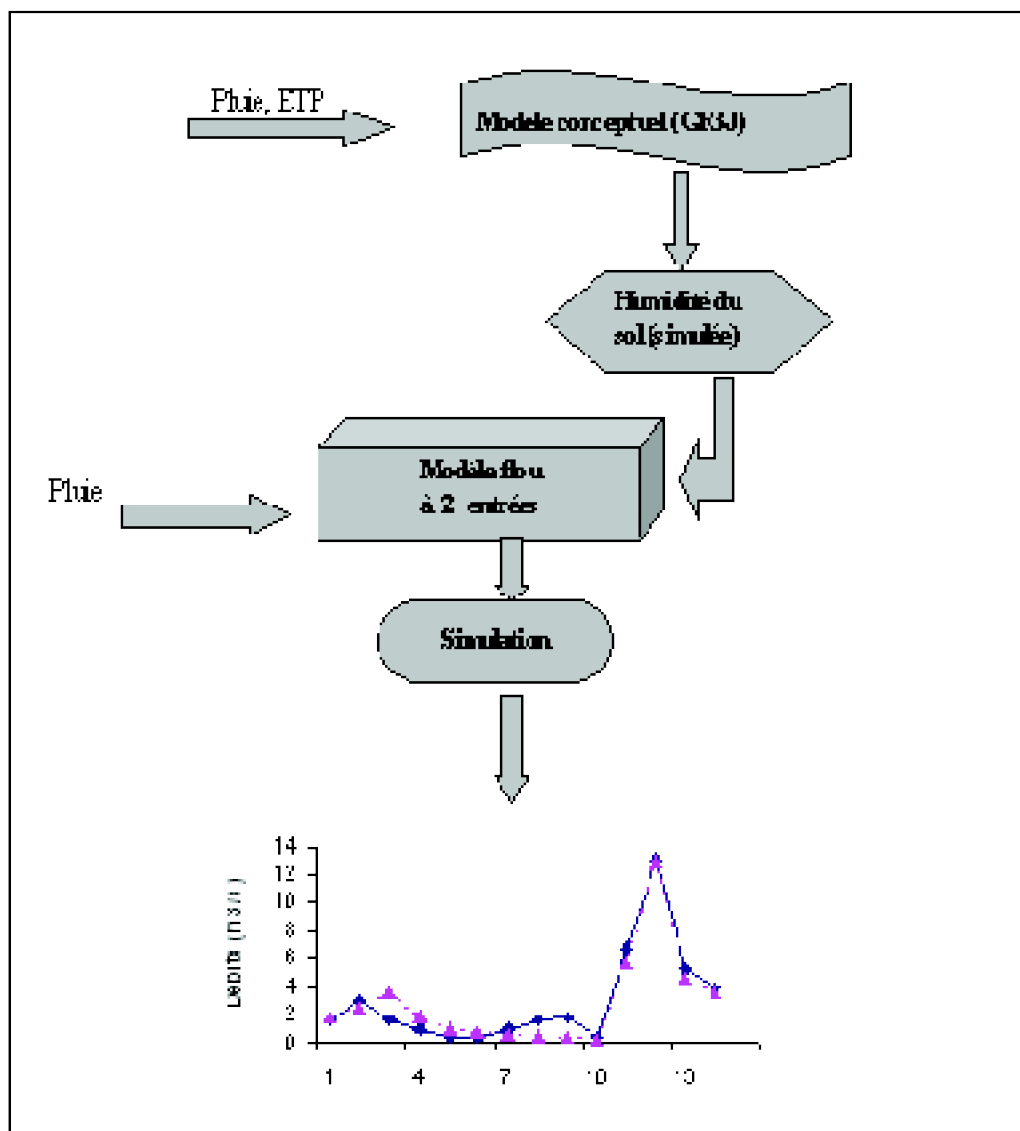


Figure VI-7 Organigramme du modèle Flou(2) à deux entrées

En intégrant cette donnée, le modèle flou (Flou(2)) à deux entrées (pluie et humidité du sol) aura pour règle:

- Si la pluie est nulle Ou le niveau du sol est faible Alors le débit est faible
- Si la pluie est nulle Ou le niveau du sol est moyen Alors le débit est moyen.
- Si la pluie est nulle Ou le niveau du sol est élevé Alors le débit est élevé.

VI.4.2.1. Résultats du modèle Flou (2) appliqué aux données journalières

Les résultats du modèle flou(2) appliqué aux stations journalières sont présentés dans le tableau VI-9 et V-10 pour les deux phases de calage et de validation. Le nombre de règles est identique au modèle flou* (Pluie, ETP).

Stations	Cheffia (1986-88)			Isser (1990-92)			Zarzedas (1991-93)		
	Nash (%)	Ic	CET (%)	Nash (%)	Ic	CET (%)	Nash (%)	Ic	CET (%)
Flou (5 ²)	81.2	1.0	473	29.7	0.35	572	71.8	1.0	496
Flou (8 ²)	84.5	1.0	429	37.1	0.40	525	85.1	1.0	483
Flou (11 ²)	88.9	1.0	421	61.1	0.64	520	91.8	1.0	451

Tableau VI-9 Résultats du modèle flou(2) (période de calage)

Stations	Cheffia (1986-88)			Isser (1990-92)			Zarzedas (1994-96)		
	Nash (%)	Ic	CET (%)	Nash (%)	Ic	CET (%)	Nash (%)	Ic	CET (%)
Flou (5 ²)	72.1	1.0	87	52.3	0.93	98	68.9	0.99	111
Flou (8 ²)	87.1	1.0	86	67.5	0.93	103	76.0	1.0	111
Flou (11 ²)	86.0	1.0	87	72.4	1.01	95	81.5	1.0	110

Tableau VI-10 Résultats du modèle flou(2) (période de validation)

Au vu des résultats exposés dans les tableaux VI-9 et VI-10, on remarque que les résultats sont assez performants par rapport aux précédents, en particulier pour les stations de Cheffia et de Zarzedas.

Pratiquement pour tous les modèles, l'efficacité de Nash est supérieure à 80.0%, en période de calage excepté le bassin de l'Isser, avec un résultat maximal pour le modèle à 121 règles de la station de Zarzedas (Figure VI-8), où le critère de Nash avoisine les 92.0%, ce qui représente une performance assez élevée.

En période de validation, les résultats sont autant appréciés, puisque les valeurs du critère de Nash dépassent 80.0% pour les modèles à 64 et 121 règles, on assiste même à une amélioration de la simulation pour de la station Cheffia, les valeurs du critère augmentent de 3% pour le modèle à 64 règles (Figure VI-9).

Pour ce qui est du bassin de l'Isser, si on compare le modèle flou*(1) (Introduction de l'ETP) à celui du modèle flou(2) (Introduction de l'humidité du sol), on voit nettement que ce dernier est plus performant, et donne des résultats nettement meilleurs, et qui s'améliorent en augmentant le nombre de règles.

Il est à signaler, des performances remarquables obtenues par années, pour exemple, l'année 1992 de la station de Zarzedas simulée avec une efficacité de Nash de plus de 93.0% (Figure VI-8), et l'année 1987 du bassin de la Cheffia simulée avec une efficacité de Nash dépassant 92.0% (Figure VI-9).

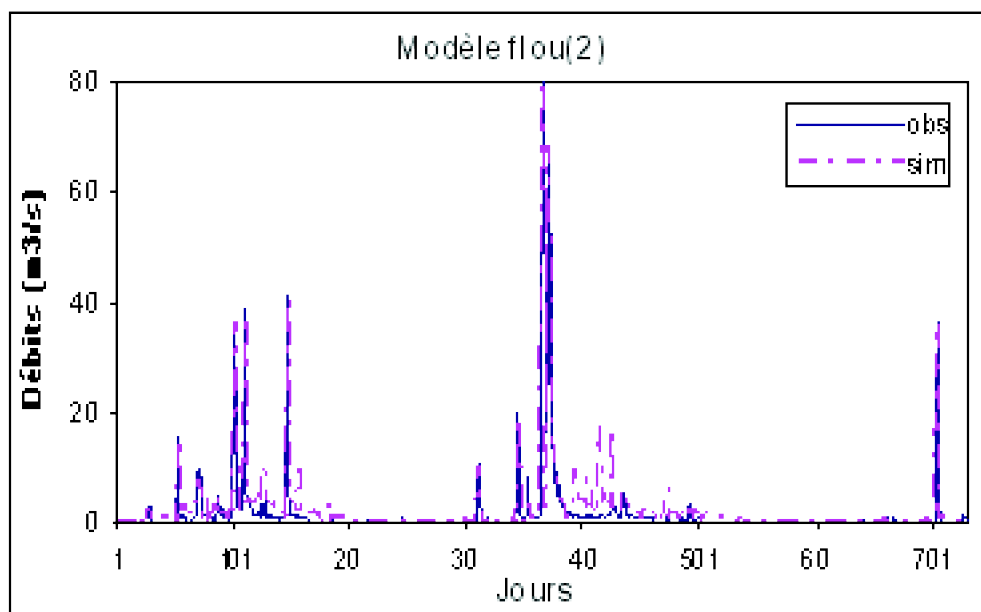


Figure VI-8 Simulation de l'année 1992 par le modèle Flou (2), (Nash=93.7) pour la station de Zardezas

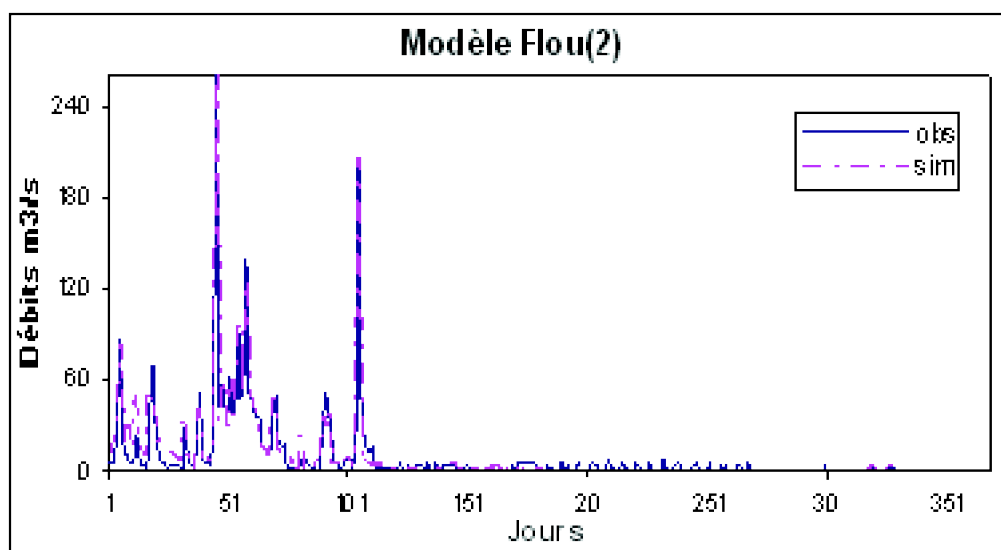


Figure VI-9 Simulation de l'année 1987 par le modèle Flou(2) (Nash=92.3) pour la station de Cheffia

Cependant, même si le modèle neuro-flou appartient aux modèles "boîte-noire", nous pouvons conclure que l'humidité du sol joue un rôle majeur dans la simulation des débits journaliers par le modèle flou, cette donnée permet non seulement de déterminer l'écoulement pour les valeurs de pluie nulle mais aussi d'approcher avec une précision correcte les débits d'étiages.

VI.4.2.2 Résultats du modèle flou (2) appliqué aux données caractéristiques

Il est indispensable de tester le modèle flou(2) aux données caractéristiques, vues les améliorations obtenues en période de calage longue, et cela après les résultats

relativement moyens donnés par le modèle flou à une seule entrée.

Isser	Calage (1993)			Validation (1994)		
	Nash (%)	Ic	CET (%)	Nash (%)	Ic	CET (%)
Flou (5 ²)	55.1	0.6	411	48.1	0.20	90.0
Flou (11 ²)	67.3	1.0	210	53.3	0.34	93.7

Tableau VI-11 Résultats du modèle flou(2) du bassin Sebbaou

Zardezas	Calage (1994-95)			Validation (1996-97)		
	Nash (%)	Ic	CET (%)	Nash (%)	Ic	CET (%)
Flou (5 ²)	80.1	0.99	434	81.0	1.3	103
Flou (11 ²)	92.5	1.0	364	94.5	1.0	114

Tableau VI-12 Résultats du modèle flou(2) du bassin Zardezas

Les résultats exposés dans les tableaux VI-11 et VI-12 montrent une amélioration significative de la simulation pour le bassin du Sebbaou, en particulier pour le modèle à 11 règles, cependant ces résultats restent relativement limités, puisque l'efficacité de Nash ne dépasse pas 68.0%.

Pour la station de Zardezas, les résultats sont très satisfaisants, on obtient une efficacité de Nash de plus de 80.0% pour le modèle à cinq règles (5²) par entrée, et cela pour les deux périodes de calage et de validation.

La simulation à 121 (11²) paraît plus robuste et optimale, le modèle donne des valeurs du critère de Nash très élevées de 92.5% et 94.5% respectivement pour les phases de calage et de validation. La reproduction des débits caractéristiques est relativement correcte, en effet le débit maximal est correctement simulé, et les étiages sont bien estimés en particulier en phase de validation.

VI.4.2.3. Résultats du modèle Flou (2) appliqué aux données mensuelles

Les résultats obtenus par le modèle flou à une entrée (pluie) au pas de temps mensuel (modèle à 12 et 30 règles) sont assez corrects, en effet les valeurs du critère de Nash sont supérieures à 84% pour le modèle à 25 règles, et dépasse 92% pour le modèle utilisant 60 règles floues en période de calage.

Dans ce qui suit, les stations hydrométriques étudiées seront modélisées par un modèle flou introduisant l'humidité du sol simulée en plus de la pluie mensuelle, ce qui permet l'analyse de sensibilité des modèles neuro-flous au pas de temps mensuel à l'humidité du sol. Cette initiative est motivée par la recherche d'une meilleure simulation

des débits, et cela par une minimisation des erreurs des modèles.

Dans le chapitre III, le modèle GR2M (version 2004) s'est révélé le plus robuste. En effet, en le comparant avec d'autres modèles conceptuels tels que les modèles ABCD et de Thornthwaite, ce modèle donne une efficacité la plus importante

Le modèle flou(2) ainsi appliqué aura la règle floue suivante:

Si la pluie est élevée ou l'humidité du sol est faible Alors le débit est élevé;

Si la pluie est élevée ou l'humidité du sol est élevée Alors le débit est très élevé.

Ce qui permet donc, d'optimiser la relation pluie-débit en apportant une information sur l'état hydrique du sol durant le mois.

Les modèles seront testés avec un nombre de quatre et de six règles par entrée, soit respectivement un nombre total de 4^2 et 6^2 règles.

Les résultats de ce modèle sont présentés dans les tableaux VI-13 pour la période de calage et dans le tableau VI-14 pour celle de la validation.

Stations	Cheffia (1978-84)		Lakhdaria (1980-87)		La Traïlle (1970-77)		Sebbaou (1980-87)		Tafna (1980-1987)	
	Nash (%)	CET (%)	Nash (%)	CET (%)	Nash (%)	CET (%)	Nash (%)	CET (%)	Nash (%)	CET (%)
Flou (4^2)	94.0	189	91.0	226	95.3	211	84.2	353	95.1	401
Flou (6^2)	98.1	149	96.1	181	93.0	183	90.2	360	98.0	344

Tableau VI-13 Résultats des modèles flou(2) en période du calage

Stations	Cheffia (1985-91)		Lakhdaria (1988-95)		La Traïlle (1979-85)		Sebbaou (1988-95)		Tafna (1988-1995)	
	Nash (%)	CET (%)	Nash (%)	CET (%)	Nash (%)	CET (%)	Nash (%)	CET (%)	Nash (%)	CET (%)
Flou (4^2)	93.5	105	90.1	100	69.7	102	82.1	148	75.1	83
Flou (6^2)	96.1	100	94.1	95	84.0	103	92.1	143	85.0	90

Tableau VI-14 Résultats des modèles flou(2) en période de validation

En analysant les résultats détaillés dans les tableaux VI-13 et VI-14, on remarque en effet, les performances élevées des modèles, pour le modèle à 6 règles, l'efficacité de Nash dépasse 90.0% en phase de calage, et cela pour toutes les stations, en phase de validation la moyenne du critère de Nash est de 90.0%. Une amélioration de la simulation appréciée est donnée pour le cas de la station de La traïlle, dont on a signalé une baisse du niveau de simulation avec le modèle à une seule entrée en période de validation (50.1%), le modèle à 2 entrées donne une efficacité de 84.0%.

VI.5. Analyse comparative des modèles pluie-débit étudiés

Trois types de modèles pluie-débit ont été testés dans cette recherche; les modèles conceptuels, les modèles basés sur les réseaux de neurones artificiels, et un modèle flou.

Ces modèles présentent des différences assez nettes, que ce soient dans leurs structures mathématiques, ou dans leurs méthodes de simulation des débits.

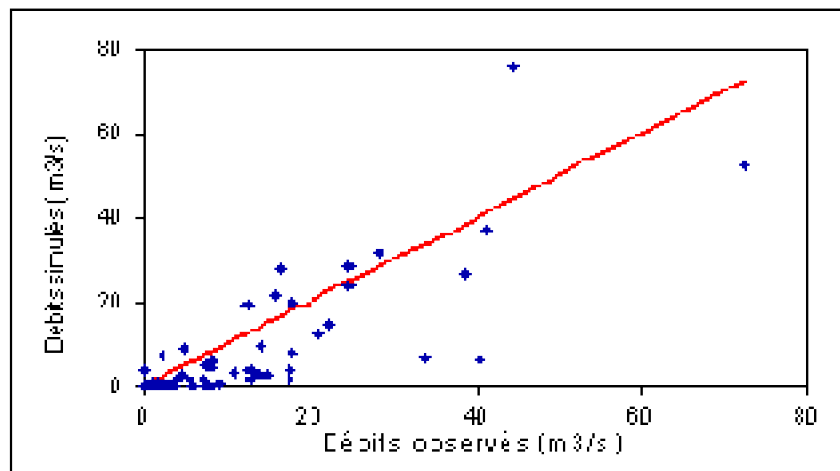
En vue d'une meilleure analyse de ces modèles, notre étude sera basée sur plusieurs assises de comparaison:

VI.5.1. Résultats des modèles en période de calage: Robustesse et qualité de simulation des modèles

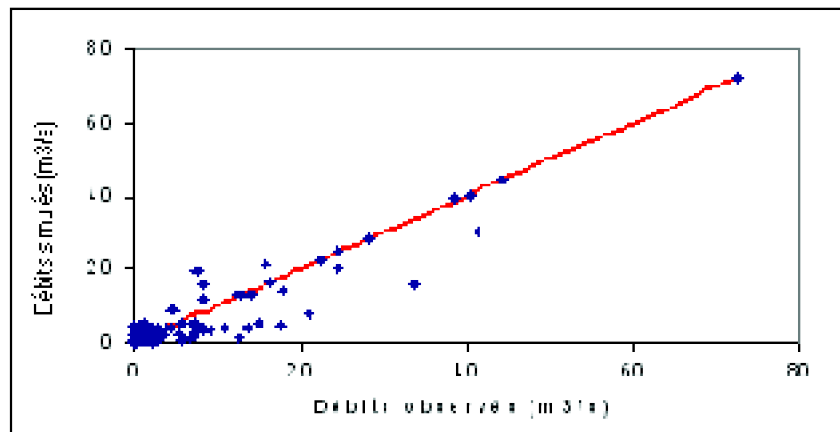
La qualité de simulation d'un modèle pluie-débit représente un important paramètre de sa validité, cependant vu le différent contexte du processus pluie-débit entre le pas de temps mensuel (non linéarité) et journalier (complexité et non linéarité), cette analyse sera étudiée en tenant compte des deux pas de temps de simulation.

Au pas de temps mensuel, nous avons démontré les performances des modèles neuro-flous. En effet à cette échelle, vu la non linéarité croissante de la relation pluie-débit, et malgré les performances relatives des modèles conceptuels obtenues pour certains bassins, ces modèles donnent généralement des résultats moyens voir faibles (Nash < 50%).

Ceci, contrairement aux modèles 'boîte noire' où on a enregistré des performances très élevées pour l'ensemble des stations, les modèles neuro-flous optimaux donnent de bonnes appréciations avec une efficacité dépassant 80% (Figure VI-10).



(a)



(b)

Figure VI-10 Comparaison des résultats des modèles: (a) GR2M (2002) et (b) neuro-flo de la période de calage de la station de Lakhdaria

On peut dire que les modèles neuro-flous déclassent les modèles conceptuels au pas de temps mensuel.

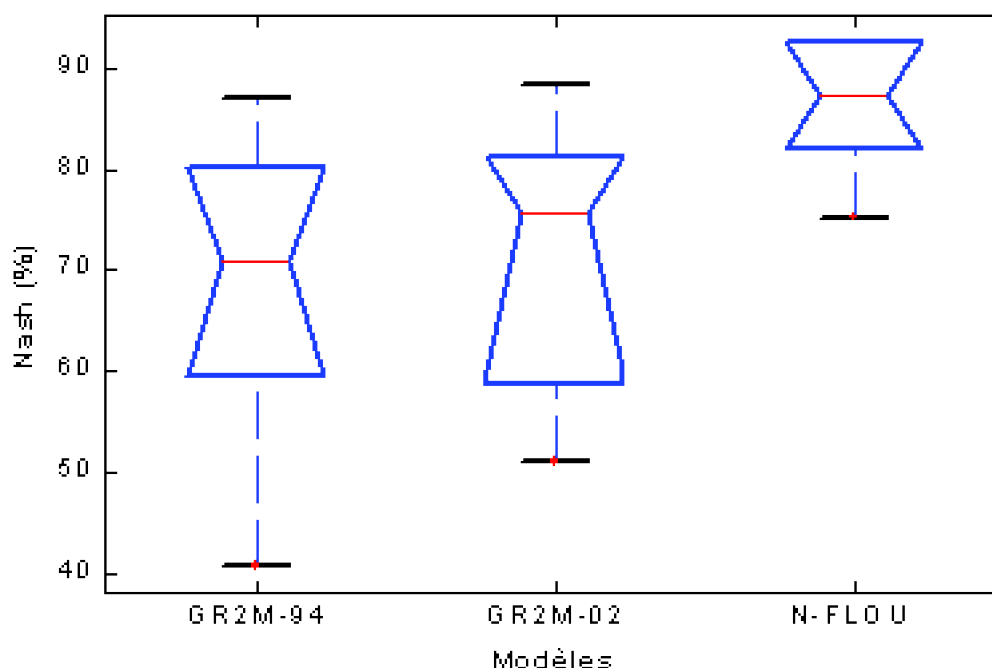


Figure VI-11 Comparaison des résultats des modèles GR et neuro-flou au pas de temps mensuel (période de calage)

De plus, le modèle neuro-flou (à 30 règles) mensuel présente une faible marge de simulation, l'écart-type sur les valeurs du critère de Nash est de 7.0 seulement (Figure VI-11).

Au pas de temps journalier, si on compare les modèles conceptuels avec les modèles neuro-flous (à une seule entrée: pluie), on voit nettement que ces derniers n'arrivent pas à reproduire les débits observés, la relation complexe du processus pluie-débit fait que les modèles (neuronaux et flous) donnent des débits mal simulés, en effet les décrues brusques et la surestimation des débits faibles représentent les inconvénients de ces modèles.

L'efficacité de Nash pour les modèles conceptuels (tels que les modèles GR3/4J ou IHACRES) est généralement meilleure, pour toutes les stations (Figure VI-12).

Le modèle GR3J a donné une moyenne de Nash de 78% en période de calage, contre 57.4 % seulement pour le modèle neuro-flou optimal (à 80 règles).

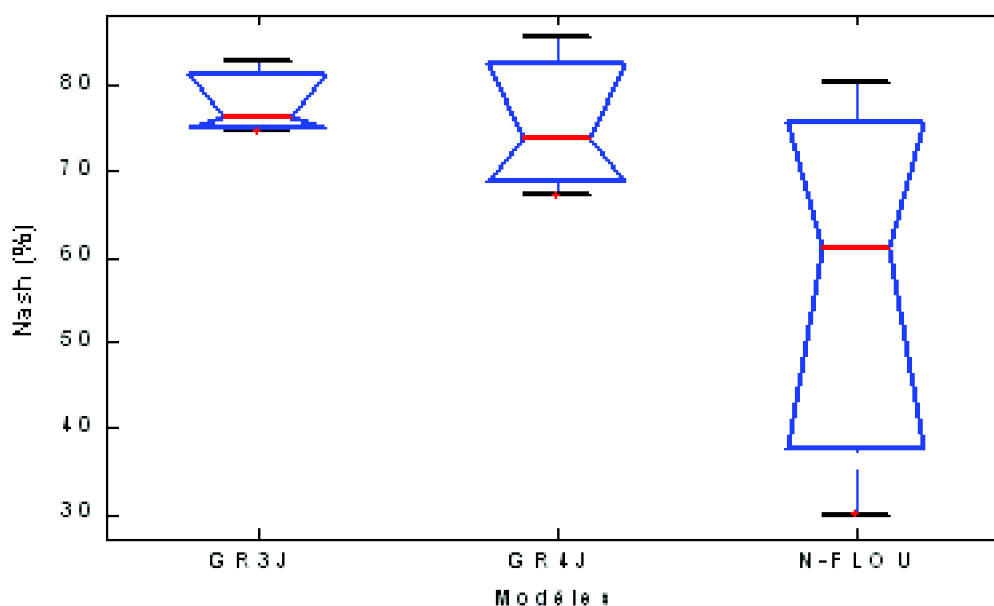


Figure VI-12 Comparaison des résultats des modèles GR et neuro-flou au pas de temps journalier (période de calage)

VI.5.2. Simulation des débits caractéristiques

Pour ce qui est des débits d'étiages, les modèles basés sur les réseaux de neurones et la logique floue, permettent de mieux reproduire les débits de crue et maximaux observés. Pour les modèles conceptuels, vu l'effet seuil (Capacité maximale d'un réservoir), la non-reconstitution du débit de pointe représente leur inconvénient majeur, en particulier en phase de validation où l'indice I_c est généralement inférieur à 0.5.

Pour l'estimation des débits d'étiage, nous nous sommes basés sur le rapport des cumuls d'étiages 'CET'. Pour l'ensemble des modèles, nous avons noté la difficulté de simulation de ces débits en particulier en phase de calage, pour les deux pas de temps. Une amélioration de cette simulation a été enregistrée en phase de validation pour les modèles mensuels neuronaux et flous optimaux, le CET est généralement satisfaisant.

VI.5.3. Résultats des modèles en période de validation : Capacité prédictive des modèles

La prévision des débits reste l'objectif principal d'un modèle pluie-débit. Dans ce cas, son appréciation reste tributaire des résultats obtenus en période de validation.

Pour tous les modèles testés, nous avons remarqué une baisse de qualité de simulation lorsqu'on passe en phase de validation.

Cependant les modèles conceptuels mensuels (ABCD, et Thornthwaite) présentent une baisse des performances, qui a été relativement importante, pour certaines stations telle que la station de Tafna où le critère de Nash est négatif.

Pour les modèles 'boîte noire' en particulier neuro-flous, nous avons souligné leur

pouvoir prédictif, à travers une bonne simulation des débits calculés en période de validation.

Pour exemple, au pas de temps mensuel, l'efficacité de Nash reste au dessus de 80% pour les stations de Cheffia, Lakhdaria et Sebbaou, obtenu par le modèle neuro-flou à 30 règles.

VI.5.4. Structure des modèles : Paramétrisation et parcimonie des modèles

Les réseaux de neurones et le modèle ANFIS utilisent des algorithmes non-linéaires robustes mais très complexes, avec un nombre de paramètres pouvant être augmenté pour différentes stations, ceci représente donc un avantage pour ces modèles (Lefkir *et al.*, 2006).

Pour ce qui est du modèle neuro-flou (ANFIS), les résultats exposés dans les tableaux précédents démontrent que ce modèle avec un faible nombre de règles donnent des résultats médiocres et moins performants que le modèles conceptuels GR2M ou GR3J, respectivement aux pas de temps mensuel et journalier.

En effet, avec un nombre de cinq (5) règles, le modèle donne des résultats moyens voir très faibles (<60%).

Ainsi, les performances de l'approches ANFIS ne se manifestent que si le nombre de règles floues (et/ou le nombre de paramètres) est important, en moyenne, supérieur à vingt (20) règles.

Cependant, les modèles conceptuels présentent une simplicité dans leur structure. Pour exemple, au pas de temps journalier, les résultats obtenus par le modèle GR3J sont optimisés avec trois (3) paramètres seulement, et cela contrairement au modèle flou qui a nécessité 80 règles.

Les modèles conceptuels GR présentent dans ce cas des performances très remarquables avec un nombre réduit de paramètres (2 ou 3), ils permettent d'approcher les débits observés avec une précision acceptable en particulier en phase de calage.

Ces paramètres sont généralement liés à une signification physique ou hydrologique; tels que la capacité maximale du réservoir sol (paramètre A) ou encore le temps de montée de l'hydrogramme unitaire (paramètre C) dans le cas du modèle GR4J.

VI.5.5. Cas du modèle neuro-flou à deux entrées

Dans le chapitre III, après essai des modèles conceptuels, le modèle GR3J a été jugé comme étant le modèle le plus robuste des modèles à réservoirs. Cependant, nous avons noté certains inconvénients de ce modèle, en particulier la reproduction des débits caractéristiques (crues et étiages).

En comparant les résultats des modèles à réservoirs avec ceux du modèle neuro-flou à deux entrées flou(2) à 11² règles (121 règles), nous avons remarqué que ce dernier a

donné des résultats meilleurs et cela pour les deux périodes, le modèle enregistre une efficacité de Nash de plus de 86% pour la station de Cheffia, et atteint 92% pour la station de Zardezas en phase de calage.

Ce modèle est largement plus robuste en phase de validation, le critère de Nash dépasse 81% pour la station de Zardezas, et cela contrairement au modèle GR3J où la baisse de performance entre les deux périodes reste très importante.

Pour ce qui est des débits caractéristiques, il est à noter, que le modèle Flou(2) reconstitue correctement le débit maximal, pour les deux phases.

VI.6. Modèle neuro-flou journalier à trois entrées

Les modèles neuronaux et neuro-flous s'appuient sur des algorithmes qui utilisent une méthode d'optimisation à partir des couples de pluie/débit. L'introduction d'une deuxième entrée nous a permis d'affronter les difficultés de simulation pluie-débit journalière et donc une meilleure reproduction des débits.

Cependant, il reste une marge d'erreurs de simulation dans les deux périodes de calage et de validation. La recherche d'un modèle assez robuste qui nous permettra de minimiser ces erreurs reste l'objectif de cette thèse.

Dans ce qui suit, les stations journalières seront modélisées à partir de trois entrées:

(a) Les pluies (variables de forçage), (b) l'humidité du sol simulée par le modèle conceptuel, (c) et les débits simulés par le modèle GR3J (Figure IV-9).

Cette procédure sera appliquée aux modèles neuro-flous.

La règle floue générale sera de la forme:

*Si la pluie est nulle Ou le niveau du sol est faible Ou les débits simulés sont faibles
Alors le débit est faible.*

Les résultats du modèle flou*(3) appliqué aux stations journalières sont présentés dans le tableau VI-15 et VI-16.

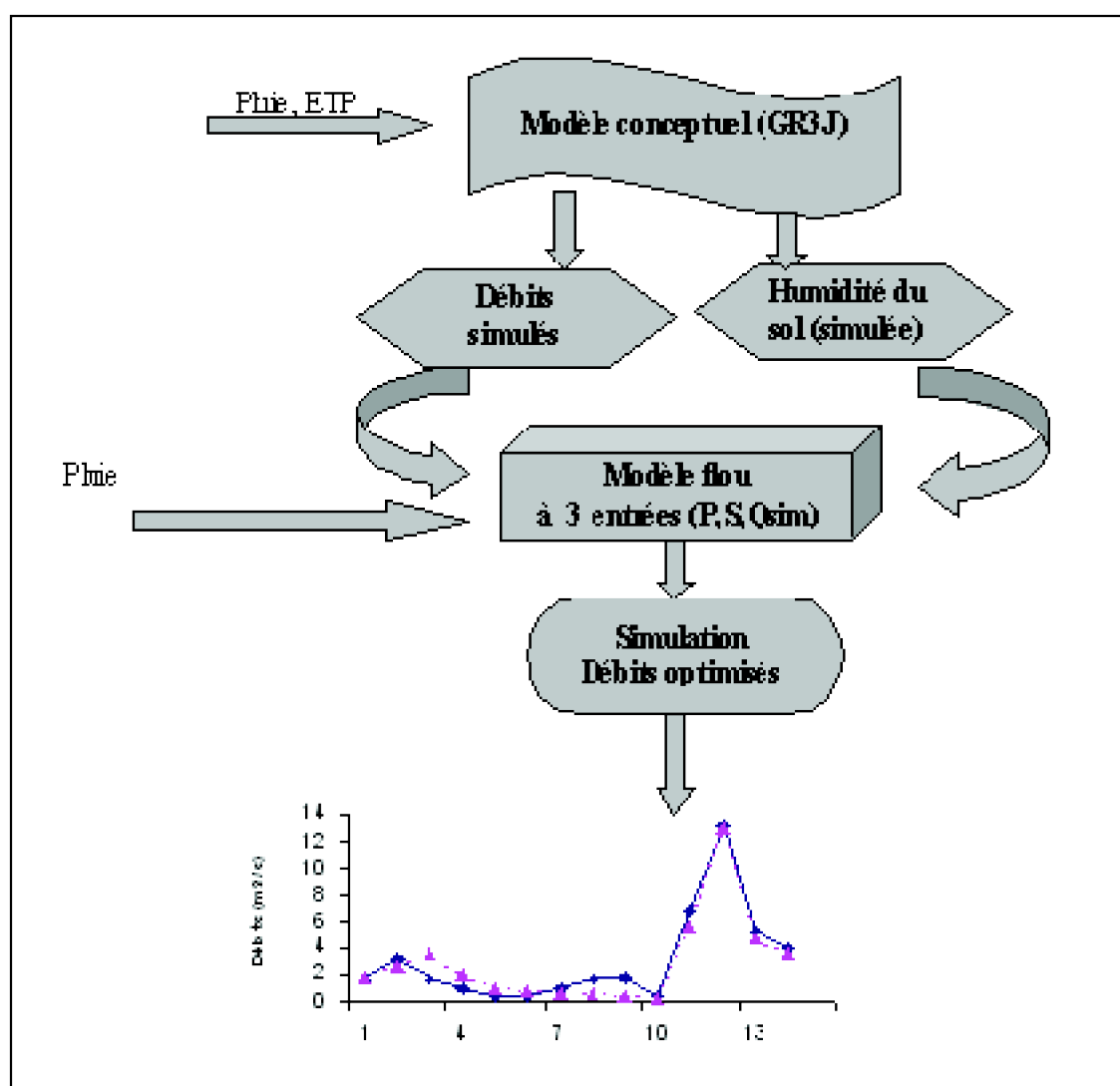


Figure VI-13 Organigramme du modèle flou à 3 entrées

Stations	Cheffia (1986-88)			Isser (1990-92)			Zarzadas (1991-93)		
	Nash (%)	Ic	CET (%)	Nash (%)	Ic	CET (%)	Nash (%)	Ic	CET (%)
Flou (2 ³)	90.7	1.0	410	87.9	1.0	475	91.7	1.0	342
Flou (3 ³)	94.4	1.0	423	91.3	0.99	379	94.5	1.0	326

Tableau VI-15 Modèle flou*(3) (période de calage)

Stations	Cheffia (1986-88)			Isser (1990-92)			Zarzedas (1994-96)		
	Nash (%)	Ic	CET (%)	Nash (%)	Ic	CET (%)	Nash (%)	Ic	CET (%)
Flou (2 ^s)	88.4	1.0	86	80.6	1.0	121	80.0	1.0	114
Flou (3 ^s)	91.2	1.0	88	85.3	1.0	114	91.0	1.0	113

Tableau VI-16. Résultats du modèle flou(3) en période de validation*

Au vu des tableaux précédents, les résultats montrent une amélioration considérable de la simulation quand le modèle intègre trois entrées. Ainsi le modèle neuro-flou à neuf (3³) règles permet une meilleure approche des débits (cas de l'Isser), et une simulation correcte des débits maxima.

Des performances très élevées sont à remarquer pour toutes les stations, avec une simulation plus acceptable des débits d'étiages en période de validation.

VI.7. Reconstitution des séries hydrométriques tronquées

Un des objectifs d'un modèle pluie-débit est la simulation des débits à l'exutoire d'un bassin versant. L'extrapolation de chroniques de débits pour des séries hydrométriques courtes est possible, si le modèle reste assez performant dans la période de validation.

Certaines séries des stations du réseau hydrométrique algérien présentent des lacunes importantes, qu'il est difficile de combler. Les modèles neuro-floous à 2 et 3 entrées dont les performances ont été démontrées précédemment, peuvent être explorés dans la reconstitution des séries hydrométriques tronquées, que ce soient aux pas de temps mensuel ou au pas de temps journalier, telles que les stations de Lakhdaria (090501) ou de Béni Slimane (090305) dans le bassin de l'Isser qui présente d'énormes lacunes entre 1996-1998, les stations du bassin de la Tafna ou encore certaines stations hydrométriques qui ont cessé de fonctionner, telle que la station de Belloua, depuis 1983.

Cet apport d'informations permet une meilleure prédétermination des phénomènes extrêmes tels que les sécheresses et les inondations et une gestion plus rationnelle des ressources en eau.

VI.8. Conclusion

La logique floue représente une nouvelle technique de simulation, qui à partir de valeurs

linguistiques, et s'appuyant sur la syntaxe de simulation *Si-Alors* permet de calculer de nouvelles valeurs.

En explorant ce modèle dans la modélisation du processus pluie-débit nous avons pu établir quelques avantages et inconvénients, quant à l'utilisation du modèle neuro-flou.

Au pas temps mensuel, nous avons obtenu des résultats très performants, en effet, les débits simulés ont été correctement simulés pour des modèles à plusieurs règles (15 ou 30 règles). Nous avons ainsi noté des performances plus importantes que ceux des modèles conceptuels, et cela pour la qualité de simulation, ainsi que pour la bonne reproduction des débits d'étiages en particulier en phase de validation.

Au pas de temps journalier, quelque soit le nombre de règles utilisé, nous n'avons pas obtenu une bonne qualité de simulation.

En introduisant une deuxième entrée au modèle neuro-flou journalier, nous avons enregistré une amélioration considérable des débits simulés, l'efficacité de Nash est généralement supérieure à 80% pour les deux périodes, et les débits caractéristiques sont généralement correctement reproduits.

Notre initiative nous a permis d'explorer la simulation par les modèles pluie-débit sous un nouveau volet, en combinant les deux approches conceptuelle et neuro-floue.

L'essai de cette technique au pas de temps mensuel nous a permis d'améliorer la qualité de simulation pour certaines stations (telle que la station de Tafna), et d'affiner les débits modélisés. La convergence du modèle a été telle que les valeurs du critère de Nash dépassent largement 90% pour l'ensemble des stations.

Notre analyse comparative des deux approches de modèles : conceptuelle et 'boîte noire', nous a permis de dégager les principales conclusions de ces modèles dans le cadre de simulation pluie-débit.

Ainsi, si au pas de temps mensuel, un modèle neuro-flou est considérablement meilleur, au pas de temps journalier, l'amélioration de la qualité de simulation n'est possible qu'avec la combinaison de ces deux approches.

En effet, cette procédure nous a permis non seulement d'approcher les débits avec une précision acceptable pour les deux périodes de calage et de validation, mais de simuler correctement les débits de crues et d'étiages.

Conclusion générale

En réponse à un événement pluvieux l'hydrologie d'un bassin versant se trouve modifiée. Ceci peut notamment s'observer à travers les composantes suivantes : l'écoulement de surface, l'écoulement dans la zone saturée et non saturée du sol et l'évapotranspiration. Cependant, la réponse du bassin (c'est à dire la variation du débit à cet événement pluvieux) varie en fonction de plusieurs facteurs d'abord propres aux variables climatiques telles que les pluies et leur intensité, puis des caractéristiques du bassin versant telles que sa superficie, les conditions pédologiques et morphométriques.

Cette complexité de la réponse du bassin versant aux variables de forçage telles que les pluies, a fait l'objet de plusieurs recherches durant plusieurs décennies.

Les modèles hydrologiques sont des outils numériques qui servent à représenter le fonctionnement d'un bassin versant. L'objectif de cette modélisation est de simuler l'une ou plusieurs de ces composantes pour une échelle de temps et d'espace.

La relation pluie-débit représente un axe de recherche important en hydrologie du fait qu'elle permet de comprendre les mécanismes de génération de débits et d'anticiper les phénomènes hydrologiques extrêmes tels que les sécheresses et les inondations.

L'objectif essentiel de cette thèse est d'évaluer les qualités des modèles existants, de les tester en milieu semi-aride algérien, caractérisé par une forte variabilité annuelle et inter-annuelle, et des cycles de sécheresses qui sévissent d'une période à l'autre pour explorer les voies d'amélioration possibles de ces modèles. La recherche de modèles plus satisfaisants a en effet, des enjeux importants. Elle permet tout d'abord une meilleure

compréhension du comportement hydrologique du bassin versant. Elle garantit ensuite une meilleure fiabilité aux applications qui utilisent les simulations du modèle dans un contexte opérationnel.

Trois approches de modèles ont été exploitées dans cette thèse: les modèles conceptuels globaux à réservoirs, les modèles basés sur les réseaux de neurones artificiels utilisant des algorithmes non-linéaires et les modèles flous à apprentissage neuronal dits modèles neuro-flous.

Ces modèles ont été appliqués aux bassins versants du Nord algérien répartis à travers le territoire national, à deux pas de temps: mensuel caractérisé par une forte non linéarité du processus pluie-débit, et journalier, par la complexité de ce processus.

Etude de la relation pluie-débit par les modèles conceptuels

La modélisation pluie-débit par les modèles conceptuels (à réservoirs) est généralement réalisée par deux processus différents à savoir la fonction de production non linéaire, et la fonction de transfert qu'on considère linéaire

En comparant quatre (4) modèles mensuels robustes, avec un nombre de paramètres qui varie de deux à quatre, à savoir le modèle de Thornthwaite, le modèle ABCD, et deux versions des modèles GR2M (1994, 2002), les résultats ont démontré la robustesse de ces deux derniers modèles, et leurs performances à simuler les débits mensuels à partir des variables hydrométéorologiques (Pluie, ETP), et cela en dépit de la complexité de la relation pluie-débit à cette échelle.

Au pas de temps journalier, quatre modèles ont été pris en considération à savoir le modèle GR3J, le modèle GR4J, le modèle CREC à huit paramètres et le modèle IHACRES.

Les modèles testés dans trois bassins versants algériens ont donné des résultats satisfaisants, en particulier en période de calage où l'efficacité de Nash est généralement assez correcte.

Cependant, vu que la validation d'un modèle pluie-débit est tributaire de ses performances en phase de vérification, on note une baisse significative de la qualité de simulation de ces modèles dans cette phase. Certains modèles (modèles ABC, CREC) se détériorent lorsqu'on passe à la phase de vérification, et sont très sensibles aux variations climatiques.

On a noté la difficulté de ces modèles à reproduire le débit maximal observé dans toute la période de simulation, ce qui limite leur utilisation dans la prédétermination des débits des crues.

Etude de la relation pluie-débit par les réseaux de neurones artificiels

Les réseaux de neurones artificiels sont une nouvelle technique d'extrapolation, exploitée dans diverses disciplines scientifiques.

C'est à ce propos qu'ils ont été utilisés dans le domaine de la modélisation pluie-débit. Vu la richesse et l'immensité des modèles basés sur les réseaux de neurones artificiels que ce soit sur le plan théorique ou pratique, on a donc débuté l'étude par une

synthèse sur les différents types de réseaux de neurones.

Ces réseaux sont dotés d'un pouvoir d'extrapolation non-linéaire et agissent par apprentissage de données et permettent ainsi de simuler le processus pluie-débit à travers des paramètres appelés poids, constitués dans chaque neurone, et dans chaque couche cachée.

Trois remarques essentielles peuvent être notées concernant ces modèles:

- Les réseaux de neurones de neurones en block (batch training) ne donnent pas une simulation correcte du processus pluie-débit par rapport aux modèles conceptuels, les résultats sont nettement inférieurs à ceux obtenus par les modèles à réservoirs.
- Au pas de temps journalier, quelque soit le nombre de neurones et de couches cachés, les débits sont mal simulés, cela est dû à la complexité de la relation pluie-débit à cette échelle.
- Il est pratiquement impossible de définir au préalable le modèle optimal pour une relation pluie-débit, et cela vu le nombre important de paramètres intervenants dans l'exécution d'un modèle basé sur les réseaux de neurones artificiels.
- Etude de la relation pluie-débit par la logique floue

La logique floue représente une nouvelle méthodologie de simulation, basée sur un langage naturel, elle permet à partir des termes de grandeurs flous d'engendrer des valeurs simulées dans un espace entrée/sortie. C'est donc une théorie mathématique formelle pour la prise en compte de l'imprécision et des incertitudes.

Les résultats obtenus par les modèles flous au pas de temps mensuel montrent qu'avec cette approche les débits sont nettement meilleurs que ceux obtenus par les modèles conceptuels, et présentent un net avantage par rapport aux réseaux de neurones artificiels sur le plan pratique. En effet avec une augmentation du nombre de règles floues, le modèle converge vers un optimum difficilement réalisable par les modèles de réseaux de neurones artificiels.

Au pas de temps journalier, il est à signaler la difficulté de la modélisation pluie-débit à cette échelle, et cela vu la complexité de cette relation à ce pas de temps. Cependant la reconstitution des débits de pointe par le modèle flou représente un avantage majeur par rapport aux modèles conceptuels.

- Modèles conceptuels ou modèles «boîte-noire»: combinaison des deux approches conceptuelle et floue:

Deux approches de modèles pluie-débit ont été explorées dans cette thèse: l'approche conceptuelle qui met en relation la pluie en débit par un jeu de réservoirs sensés reproduire le cycle hydrologique, et l'approche dite boîte-noire, qui permet une extrapolation des débits à partir d'une chronique de pluies observées, et cela aux pas de temps mensuels et journalier.

L'amélioration de la simulation de cette relation pluie-débit déjà étudiée par plusieurs hydrologues, était l'objectif principal de cette thèse.

Après la comparaison des deux modèles conceptuels et neuro-flous, la simulation pluie-débit a été exploitée sous une nouvelle optique: combiner les deux approches.

Une deuxième entrée a été introduite au modèle neuro-flou. Les résultats ont montré que toutes les stations étudiées au pas de temps journalier sont modélisées de façon très correcte, et les débits ainsi calculés sont très proches de ceux observés.

Cette nouvelle technique a été aussi exploitée au pas de temps mensuel, ce qui a permis d'obtenir des résultats très satisfaisants que ce soient en phase de calage ou en phase de validation.

Cette recherche nous a conduits à une meilleure connaissance de la relation pluie-débit. En effet, les débits simulés optimisés par les modèles conceptuels, ont été introduits comme troisième entrée dans le modèle neuro-flou en plus des pluies et de l'humidité du sol simulée.

Ainsi, toutes les séries hydrométriques journalières simulées, sont très proches de celles observées et le critère de Nash est souvent supérieur à 90%, ce qui confirme la robustesse de cette technique.

Cette recherche nous a ouvert donc la voie à une meilleure exploration de la modélisation de la pluie en débit, et a permis une meilleure simulation du processus pluie-débit aux deux pas de temps mensuel et journalier.

Références bibliographiques

- Abbott M.B, Bathurst J.C, Cunge J.A, O'Connell P.E Rasmusen J.L. (1986) An introduction to the European Hydrological System 'SHE'. Structure of a physically-based distributed modelling system. *Journal of hydrology*, 87; 61-67.
- Aivazian S., Enikov F., Mochalkine L. (1983) Elément de modélisation et traitement primaire des données. Ed Mir, Moscou.
- Alley M.W. (1985) Water balance models in one-month-ahead stream flow forecasting. *Water resources research*,21(4); 597- 606.
- Ambroise B. (1991) Hydrologie des petits bassins versants ruraux en milieu tempéré, Processus et Modèles. Centre d'Etudes et de Recherches Eco-Géographiques. Université Louis Pasteur, Strasbourg. 33 p.
- Ambroise B. (1998) La dynamique du cycle de l'eau dans un bassin versant. Processus, facteurs, modèles. HGA, Bucarest, 200 p.
- Amraoui N., Golaz C., Mardhel V. Pinault J. (2003) Evaluation du risque d'inondation dans le bassin de la Somme: apport de l'approche globale et de l'approche distribuée. SIRNAT (Orléans); 1-9.
- Anctil F., Michel C., Perrin C. Andréassian V. (2004) A soil moisture index as an auxiliary ANN input for stream flow forecasting. *Journal of hydrology*, 286; 155-167.
- ANRH (a). (1985) Ressources en eaux de surface et possibilités de leur aménagement. Algérie orientale. Hydrotechnic Corporation.

- ANRH (b). (1985) Hydrologie de la Tafna. (Rapport préliminaire). Chaumont et collaborateurs. Etude SES 1080/HF.
- Arnell N. (1992) Factors controlling the effects of climate change on river Flow regimes in humid temperatures environments. *Journal of hydrology*, 132; 321-342.
- Ayadi M. Bargaoui Z. (1998). Modélisation des écoulements de l'oued Miliane par le modèle CEQUEAU. *Hydrological Sciences Journal*, 43(5); 741-758.
- Bardossy A. Duckstein L. (1995) Fuzzy rule based modelling with applications to geophysical, biological and engineering systems. *CRC press*, p 232.
- Baudez J.C., Loumagne C., Michel C., Palalgos B., Gomendy V. #Bartoli F. (1999) Modélisation hydrologique et hétérogénéité des bassins. *Etude et gestion des sols*, 6 (4); 165-184.
- Benkaci Ali T. (2001) Contribution à la modélisation de la relation pluie-débit aux pas de temps mensuel et journalier par les modèles conceptuels et les réseaux de neurones artificiels. Thèse de Magister, Institut National Agronomique d'Alger, 150 p.
- Benkaci Ali T. Dechemi N. (2004) Modélisation pluie-débit journalière par les modèles conceptuels et 'boite noire', test d'un modèle neuroflou. *Journal des sciences hydrologiques* 4(21).
- Berenji R., Malkani A., Copeland C. (1995) Tether control using fuzzy reinforcement learning, *IEEE/IFES*, Yokohama, Japan. 1315-1319.
- Bergstrom S. Forsman A. (1973) Developed of a conceptual deterministic rainfall-runoff model. *Nordic hydrology*, 4; 47-70.
- Bermad A. (2004) Elaboration d'une méthodologie de simulation et de prévision des variables aléatoires et cycliques par l'analyse en composantes principales. Thèse de doctorat, Ecole Nationale Polytechnique d'Alger, 140 p.
- Bersini H., Gorinni V. (1993) FUNNY, fuzzy or neural net methods for adaptive process control. *Proc of EUFIT*, Germany, 55-61.
- Beven K., Kirkby M. (1979) A physically based model, variable contributing area model of basin hydrology. *Hydrological Sciences Bulletin* 24, 43-69.
- Beven K., Calvert A., Morris E.M. (1987) The Institute of Hydrology Distributed Model. Institut d'hydrologie, rapport (8), Wallingford, Royaume uni, 30 p.
- Bobée (1991). The Gamma family and derived distribution applied in hydrology. Water resources application, Colorado, USA.
- Boukezzoula R. (2000) Commande floue d'une classe de systèmes non linéaires : application au problème de suivi de trajectoire. Thèse de doctorat Université de Savoie, 245 p.
- Box G. Jenkins G.M. (1976) Time series analysis: Forecasting and control. *Holden Day Inc. Oakland, CA*. 575 p.
- Bouabdallah F. (1997) Mise au point d'un modèle de transformation pluie-débit au pas de temps annuel. Rapport D.E.A, CEMAGREF, 45 p.
- Bouchon-Meunier B. (1995) La logique floue et ses applications Paris, Addison-Wesley France.
- Bouchon-Meunier B., Dubois D., Godo L. & Prade H. (2003). Fuzzy sets and systems,

-
- International Fuzzy Systems Association (IFSA-03), 2003.
- Buckley J. (1993) Sugeno type Controllers are Universal Controllers, *Fuzzy sets and Systems* 53, North-Holland Publishing Company, pp. 299-303, 1993.
- Carte pluviométrique de l'Algérie du Nord (1/500000) (1992) Projet ANRH-PNUD : Etude des ressources en Algérie. (ALG/88/021).
- Chambaz H. (1990) Contribution à l'étude de la méthode DPFT- différence première de la fonction de transfert- Application au bassin versant de l'oued Dis. Thèse de magister, Ecole Nationale Polytechnique d'Alger (ENP), 200 p.
- Chahinian N. (2004) Paramétrisation multi-critère et multi-échelle d'un modèle hydrologique spatialisé de crue en milieu agricole. Thèse de doctorat Université Montpellier II, 260 p.
- Chiew F. Mc Mahon T. (1994) Application of the daily rainfall-runoff model MODHYDROLOG to 28 Australian catchments. *Journal of hydrology*, 153; 383-416.
- Chkir, N., (1994) Mise au point d'un modèle hydrologique conceptuel intégrant l'état hydrique du sol dans la modélisation pluie-débit, Thèse de l'Ecole Nationale des Ponts et Chaussées de Paris 327p.
- Cormary Y. Guilbot A. (1973) Etude des relations pluie-débit sur trois bassins versants d'investigation. IAHS Madrid Symposium, *IAHS Publication*, (108); 265-279.
- Coulibaly P., Anctil F., Bobee B. (1999) Prévision hydrologique par réseaux de neurones artificiels : Etat de l'art. *Journal of civil engineering*, 26; 293-301.
- Crawford N.H Linsley, R.K. (1966) Digital simulation in hydrology. Stanford Watershed Model IV. *Technical report 39*, université de Standford , Californie.
- Croke B.W Jakeman A.J. (2004) A catchment moisture deficit module for the IHACRES rainfall-runoff model. *Environmental Modelling & Software*, 19; 1-5.
- Cybenko G. (1989) Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signal and Systems*, 2(4):303-314, 1989.
- Dechemi N. Chambaz H. (1994) La prévision des crues du bassin versant de l'oued Dis (Sebbaou) par la méthode DPFT. *Revue Sciences de l'eau*, 7; 331-342.
- Dechemi N., Bermad A., Tachet K. (1997) Etude et simulation de l'évapotranspiration (ETP) par l'analyse en composantes principales (ACP). *ICID Journal*, 46; 45-60.
- Dechemi N., Benkaci Ali T., Issolah A. (2003) Modélisation des débits mensuels par les modèles conceptuels et les systèmes neuro-flous. *Revue Sciences de l'eau*, 3; 407-425.
- Delattre, P. (1979) Le problème de la justification des modèles dans le cadre du formalisme des systèmes de transformation. Actes du colloque 'Elaboration et Justification des Modèles', Tome I, Ed Maloine; 97-128.
- Dimopoulos I., Lek S., Lauga J. (1996). Modélisation de la relation pluie-débit par les réseaux connexionnistes et le filtre de Kalman. *Journal des sciences hydrologiques*, 41; 177-193.
- Dreyfus G., Samuelides M., Martinez J., Gordon M., Badran F., Thiria S., Hérault L. (2004) Réseaux de neurones - Méthodologies et applications. Editions Eyrolles 200 p.
-

- Driankov D., Hellendoorn C., Reinfrank M. (1993) An introduction to fuzzy Control, Springer-Verlag.
- Duan Q., Sorooshian S., Gupta V.K. (1992) Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resources Research* 28 (4); 1015-1031.
- Duband D. (1978) Fonctions de transfert pluie-débit sur des bassins versants de l'ordre de 1000 km². Note interne, EDF-DTG, Grenoble.
- Dunn J.C (1973) A fuzzy relative of the isodata process and its use in detecting compacting well-separated clusters. *J Cybernetics* 3 (3); 32-57.
- Drogue C., Gdalia L., Razack M. Guilbot, A. (1982) Application of a rainfall-runoff mathematical model to compute karstic aquifers outflows. In: Applied Modelling in catchment Hydrology, V. P. Singh (Ed.), *Water Resources Publications*, 323-336.
- Edijatno (1987). Amélioration des modèles simples de transformation pluie - débit au pas de temps journalier sur des petits bassins versants. Mémoire de D.E.A, Cemagref, 45 p.
- Edijatno. Michel C. (1989) Un modèle pluie-débit journalier à trois paramètres. *La Houille Blanche*, 2; 113-121.
- Edijatno. (1991) Mise au point d'un modèle élémentaire pluie-débit au pas de temps journalier. Thèse de Doctorat, Université Louis Pasteur, Strasbourg. 220 p.
- Edijatno, Nascimento O., Yang X., Makhlouf Z., & Michel C. (1999) GR3J: A daily watershed model with three parameters. *Hydrological Sciences journal*, 2; 263-278.
- El Jabi N., Rousselle J. (1987) Hydrologie fondamentale. Ecole Polytechnique de Montréal.
- Ewen J., Parkin G., O'Connell P.E (1998) SHETRAN : a coupled surface/subsurface modelling system for 3D water flow and sediments and solute transport in river basins. *Water resources research*.
- Foulloy L. Galichet S. (1995) Typology of fuzzy controllers, in Theoretical aspects of fuzzy Control. Willey Publishers, 65-90.
- Freissinet C. (1997) Estimation des imprécisions dans la modélisation du devenir des produits phytosanitaires dans les sols; une méthode basée sur la logique floue. Thèse de Doctorat, Université Joseph Fourier, Grenoble. 200 p.
- Freissinet C. (1998) Estimation des imprécisions dans la modélisation du devenir des produits phytosanitaires dans les sols; une méthode basée sur la logique floue. *La Houille Blanche*, 8; 14-16.
- Freeze, R.A. Harlan, R.L., 1969. Blueprint for a physically-based digitally-simulated hydrologic response model. *Journal of Hydrology*, 9; 237-258.
- Freeze A.R. (1971) Three-dimensional, transient, saturated-unsaturated flow in a groundwater basin. *Water Resources Research*, 32; 2161-2173.
- French M., Krajewski W., Cuykendall R. (1992) Rainfall forecasting in space and time using neural network. *Journal of hydrology*, 137; 1-31.
- Gauthier E. (1999) Utilisation des réseaux de neurones artificiels pour la commande d'un véhicule autonome. Thèse de doctorat, Institut Polytechnique de Grenoble. 216 p.

-
- Girard G., Morin G., Charbonneau R. (1972) Modèle précipitations-débits à discrétisation spatiale. *Cahiers ORSTOM, Série Hydrologie*, IX(4), 35-52.
- Glorennec P.Y, Barret C. & Brunnet M., (1992) Application of neuro-fuzzy network to identification and control of nonlinear dynamical systems, Proc of IPMU, Mallorca, Spain, 507-510.
- Glorennec P.Y Orseau L. (1992) Optimisation en-ligne de modèles flous, application à la prévision de pics d'ozone. INSA Rennes, 12 p.
- Gupta, V.K. Sorooshian, S. (1983) Uniqueness and observability of conceptual rainfall-runoff model parameters: the percolation process examined. *Water Resources Research*, 19(1), 269-276.
- Grosseberg (1987) *The Adaptive Brain*. New York: North-Holland.
- Hartani R. (1995) Modélisation de systèmes neuro-flous, contributions théoriques et applications. Thèse de Doctorat, Université Paris VI, 200 p.
- Hebb D. (1949) *The organization of behaviour*, New york : Wiley 1949.
- Hewlett J.D. (1961) Watershed Management. In : Annual Report 1961, USDA Forest Service, Asheville (INC, USA), 61-66.
- Higy C. (2000) Modélisation conceptuelle et à base physique des processus hydrologiques : application du bassin versant de la haute Mentue. Thèse de Doctorat, Ecole Polytechnique Fédérale de Lausanne, 300 p.
- Hopfield J.J. (1982) Neural networks and physical systems with emergent collective computational abilities," *Proceedings of the National Academy of Sciences*, 79; 2554-2558.
- Horton R.E. (1933) The role of infiltration in the hydrological cycle. National Research Council, *Trans of American geophysical Union*.
- Hsu K., Gupta H., Sorooshian S. (1995) Artificial neural network modelling of the rainfall- runoff process. *Water resources research*, 31; 2517- 2530.
- Hursch C.R. (1936) Storm water and absorption. *Trans of American geophysical Union*, 17; 301- 302.
- Hursch C.R. Brater E.F. (1941) Separating storms hydrographs from small drainage areas into surface and subsurface flow. *Trans of American geophysical Union*, 22; 863-870.
- Jakeman A.J., Littlewood I.G., Whitehead P.G. (1990) Computation of the instantaneous unit hydrograph and identifiable components flows with application to two small upland catchments. *Journal of hydrology*, 117; 275-300.
- Jang J.-S. R. (1993) ANFIS: Adaptive-Network-based Fuzzy Inference Systems. *IEEE, Man, and Cybernetics*, 23; 665-685.
- Josserand J.F (1996) Cellules floues : contribution à la commande floue distribuée. Thèse de doctorat Université de Savoie, 165 p.
- Kalman R.E. (1960) A new approach to linear filtering and prediction problem theory. *Journal of basic engineering*, 82; 35-45.
- Kauffman (1973) Introduction à la théorie des sous-ensembles flous. Volume 1, Masson, 1973.
-

- Khodja L. (2000) Contribution à la classification floue non supervisée. Thèse de doctorat Université de Savoie, 165 p.
- Klemes V. (1982) Empirical and causal models in hydrology. In: Scientific basis of water-resource management, *National Academy Press, Washington, D.C.*, 95-104.
- Kokkonen T. (2003) Rainfall-runoff modelling, comparison of modelling strategies with a focus on ungauged predictions and model integration. Thèse de Doctorat, Université de Helsinki, 66 p.
- Kuczera G. (1997) Efficient subspace probabilistic parameter optimization for catchment models, *Water Resour. Res.*, 33(1), 177-185.
- Laborde J.P. Mouhous N. (1998) Logiciel Hydrolab, version 98.2, équipe Gestion et Valorisation de l'environnement de l'UMR, «Espace» du CNRS.
- Laborde J.P. (2003) Eléments d'hydrologie de surface. ANRH Editions 200 p.
- Larrieu J. (1957) Evaluation des crues catastrophiques par la méthode des hydrogrammes synthétiques. AIHS, Tome III, Assemblée de Toronto, 160-173.
- Lavabre J., Cambon J.P., Breil D. (1997) Plaquette de présentation du logiciel LOIEAU, journée d'études des débits de référence, 26 juin, Paris 4.
- Lavabre J. Folton Michel, C. (1999) L'estimation régionale de la ressource en eau par modélisation pluie-débit. Application à la détermination des débits de référence des débits de la région méditerranéenne française. *Hydropôle*, 15; 443-440.
- Linsley R.K., Kohler A., Paulhus J. (1982) Hydrology for Engineers. Mc Graw Hill, 508 p.
- Lippman R. (1987) An introducing to computing with neural nets. IEEE ASSP magazines, 4-22.
- Littlewood I.G., Down K., Parker J.R. Post D.A. (1997) The PC version of IHACRES for catchments-scale rainfall-streamflow modelling. Version 1.0. User Guide. *Institute of Hydrology Ed.*, 89 p.
- Liu Z. Todini E. (2002) Towards a comprehensive physical-based rainfall-runoff model *Hydrology and Earth System Sciences*; 6(5), 859-881.
- Loumagne, C., Chkir, N., Normand, M., Ottlé, C., et Vidal-Madjar, D. (1996) Introduction of the soil-vegetation-atmosphere continuum in a rainfall-runoff model for remote sensing data assimilation. *Hydrological Sciences Journal*, (41), 889-902.
- Makhlouf Z. (1994) Compléments sur le modèle pluie-débit GR4 et essai d'estimation de ses paramètres. Thèse de Doctorat, Université Paris-Sud, CEMAGREF, 426 p
- Makhlouf, Z. et C. Michel (1994). A two-parameter monthly water balance model for French watersheds. *Journal of Hydrology* 162; 199-318.
- Mallat, S. (1989) A theory for multiresolution signal decomposition : the wavelet transform, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11; 674-693.
- Mamdani E.H. (1974) Application of Fuzzy Algorithms for Control of a simple Dynamic Plant, Proc. of the IEE Control and Science, 121; (12), 1585-1588.
- Mamdani E.H. Assilian S. (1975) An experiment in Linguistic Synthesis with a fuzzy logic Controller, Int. Journal of Man-Machines Studies, 7; 1-13.
- Marsi S.F., Bkey G.A., Safford F. (1976) An adaptive random search method for

- identification of large scale nonlinear systems. Proceeding of the IFAC Symposium, 1645-1654.
- Mazenc B., Sanchez M., Thiery D. (1984) Analyse de l'influence de la physiographie d'un bassin versant sur les paramètres d'un modèle hydrologique global et sur les débits caractéristiques à l'exutoire. *Journal of Hydrology*, 69; 97-118.
- McCulloch W. Pitts W. (1943) A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5; 115-113.
- Michel C. (1983) Que peut-on faire en hydrologie avec un modèle conceptuel à un seul paramètre? *La Houille Blanche*, 1; 39-44.
- Michel C. (1989) Hydrologie appliquée aux petits bassins versants ruraux. Rapport technique CEMAGREF, Anthony.
- Minns A. Hall J. (1996) Artificial neural network as rainfall-runoff models. *Hydrological Sciences journal*, 41; 399-417.
- M. Minsky S. Papert (1969). *Perceptrons*, Expanded Edition. MIT Press.
- Moles A.A. (1995) Les sciences de l'imprécis. Seuil, Points sciences, S105, 360 p.
- Morel-Seytout H.J. Al Hasoun S. (1989) The unsaturated components of Swatch. Theory and practice. NATO ASI series, *Kluwer Academy publication*, 275; 413,433.
- Moussa R, Voltz M, Andrieux P. (2002) Effects of spatial organization of agricultural management on the hydrological behaviour of farmed catchment during flood events. *Hydrological Processes* 16(2): 393-412.
- Mouelhi S. (2002) Vers une chaîne cohérente de modèles pluie-débit conceptuels globaux aux pas de temps pluriannuel, annuel, mensuel et journalier. Thèse. Doctorat, Université Paris 6 ENGREF, France. 300 p
- Nakoula Y., (1997) Apprentissage des modèles linguistiques flous par jeu de règles pondérées. Thèse de doctorat université de Savoie.
- Nascimento N.O. (1995) Appréciation d'un modèle empirique des effets d'actions anthropiques sur la relation pluie-débit à l'échelle du bassin versant. Thèse de Doctorat, CERGRENE/ENPC, 550 p.
- Nash J.E. (1957). The form of the instantaneous unit hydrograph. *International Association of Scientific Hydrology*, Wallingford; 114-121.
- Nash J.E. # Stoulliffe J.V. (1970) River flow forecasting through conceptual model. Part 1- a discussion of principles. *Journal of hydrology*, 10; 282-290.
- Oudin L. (2004) Recherche d'un modèle d'évapotranspiration potentielle pertinent comme entrée d'un modèle pluie-débit global. Thèse de Doctorat Cemagref/ENGREF, 495p.
- Özelkan E. Duckstein L. (2001) Fuzzy conceptual rainfall-runoff models. *Journal of hydrology*, 253, 41-68
- Papy F. Douyer C. (1991) Influence des états de surface du territoire agricole sur le déclenchement des inondations catastrophiques. *Agronomie*, p 200 –215.
- Penman H.L. (1963) Vegetation and Hydrology. Commonwealth bureau of soil Technical communication (53), Farnham Royal (U.K) 124 p.
- Perrin C. (2000) Vers une amélioration d'un modèle global pluie-débit au travers d'une

- approche comparative. Thèse de Doctorat, Cemagref, Anthony, Institut National Polytechnique de Grenoble, 530 p.
- Perrin, C., C. Michel et V. Andréassian (2001) Does a large number of parameters enhance model performance? *Journal of Hydrology*, 242 (3-4), 275-301.
- Perrin C., Michel C., Andréassian V. (2002) Improvement model of a parsimonious for streamflow simulation. *Journal of hydrology*.
- Powell M., (1987) Radial basis functions for multivariable interpolation : a review, *Algorithms for approximation*, 143-167.
- Quesney A. (1999) Assimilation de mesures d'humidité de surface dans un modèle hydrologique conceptuel global. Apport de la télédétection radar ERS/SAR. Thèse de Doctorat, Université Paris 7, 350 p.
- Rakem Y. (1999) Analyse critique et reformulation mathématique d'un modèle pluie-débit empirique (GR4J). Thèse de Doctorat, Ecole Nationale des Ponts et Chaussées, 267 p.
- Roche P. A. (1971) Les divers types de modèles déterministes. *La Houille Blanche* 2; 111-129.
- Roche P. A. (1988) Les modèles. In : Gestion des eaux : automatisation, Information et télégestion, Tome 3, Presses de l'Ecole Nationale des Ponts et Chaussées. 133-166.
- Rodriguez. J.Y. (1989) Modélisation pluie-débit par la méthode DPFT. Thèse de Doctorat. INPG, Grenoble
- Rosenblatt F. (1958) The Perceptron: a probabilistic model for information storage and organization in the brain, *Psychological Review* (65) pp 386-408.
- Rosenbrock H. (1960) An automatic method for finding the greatest or least value of a function. *Computer journal*, 3; 175-184.
- Rouabah, T. (1999) Etude des transports solides en suspension de l'oued SFA-SAF dans la wilaya de Skikda. Institut National Agronomique d'Alger, 100 p.
- Rumelhart D.E., Hinton G.E., Williams J. (1986) Learning internal representation by error propagation, in *Parallel Distributed Processing*, MIT Press, Cambridge, 1; 318-362.
- Rutter A.J. (1971) A predictive model of rainfall interception in forests. Derivation of the model form observations in a plantation of corsicane pine. *Agricultural Meteorology*, 9; 367-384.
- Saidi A. (1994) Evaluation des ressources en eau superficielles, modèle de simulation pluie - débit "SMAP". ANRH.
- Sajikumar N. Thandaveswara S. (1999) A non-linear rainfall-runoff model using an artificial neural network. *Journal of hydrology*, 21; 32-55.
- Schaake J. C. Chunzen L. (1989) Development and application of simple water balance models to understand the relationship between climate and water resources. *IAHS*, 181; 343-352.
- Schreider S.Y., Jakeman A.J., Dyer B.G. #Francis R.I. (1997). A combined deterministic and self adaptive stochastic algorithm for streamflow forecasting with application to catchments of the Upper Murray basin, Australia. *Environmental Modelling and*

- Software*, 12 (1); 93-104.
- Servat E. (1986). Présentation de trois modèles globaux conceptuels déterministes: CREC 5, MODGLO, MODIBI, ORSTOM, Département F, Unité de Recherche 604.
- Servat E. Dezetter A. (1988). Modélisation globale de la relation pluie-débit: des outils au service de l'évaluation des ressources en eau. *Hydrologie Continentale, ORSTOM*, 3(2), 117-129.
- Servat E. Dezetter A. (1993) Rainfall-runoff modelling and water resources assessment in north-western Ivory coast. Tentative extension to ungauged catchments. *Journal of hydrology*, 148; 231-248.
- Singh V.P. (1995) Computer models of watershed hydrology. 1130 p.
- Singh V.P. Whooliser D. (2002) Mathematical modelling of watershed hydrology. *Journal of Hydrological engineering*, 7; 270- 92.
- Shamseldin Y. (1996) Application of neural network technique to rainfall-runoff Modeling. *Journal of hydrology*, 272-294.
- Sherman L.K. (1932) Streamflow from rainfall by the unit-hydrograph method engineering news record, 108; 501-505.
- Sorooshian, S. Gupta, V.K. (1983). Automatic calibration of conceptual rainfall-runoff models: the question of parameter observability and uniqueness. *Water Resources Research*, 19(1), 260-268.
- Sorooshian S. Gupta H. V. (1985). The analysis of structural identifiability: theory and applications of conceptual rainfall-runoff models. *Water Resources Research*, 21(4); 487-495.
- Sorooshian S. (1991) Parameter estimation, model identification and model validation: conceptual-type models. NATO ASI series, Kluwer academic publ., 443-467.
- Sugeno M. (1985) Industrial applications of fuzzy control. *Elsevier sciences publications*.
- Thomas H.A. (1981) Improved methods for rational water assessments. US Water Resources Council. Report Number.
- Takagi H. Sugeno M. (1985) "Fuzzy Identification of Systems and its Application to Modeling and Control", IEEE Trans. on Syst. Man and Cybernetics, 15; 116-132.
- Thornthwaite C. W. # Mather J. R. (1955). The water balance. Publ. Climatol. Lab. Drexel. Inst. Techol, 8(1); 1-104.
- Todini E. (1996) The ARNO rainfall-runoff model. *Journal of hydrology*, 175; 339- 382.
- Touzet C. (1992) Les réseaux de neurones artificiels: introduction au connexionnisme.
- Vakkilainen P. Karvonen T. (1982) Adaptive rainfall-runoff models SATT-I. Civil engineering and building construction series. Université de Helsinki, 54 p.
- Vandewiele G. L., Xu C.-Y., W. Huybrecht (1991) Regionalisation of physically - based water balance models in Belgium. Application to ungauged catchments. *Water Resources Management*; 199 - 208.
- Vandewiele G. Xu C. (1992) Methodology and comparative study monthly water balance model in Belgium. *Journal of hydrology*, 134; 315-347.
- Wald A. Wolfowitz J. (1943) An exact test for randomness in the non parametric case

- based on serial correlation. *Ann Math. Statistics* 15; 378-388.
- Wang L.X. (1992) Fuzzy systems an universal approximators. IEEE, conference on fuzzy systems, San Diego 1163-1169.
- Weeks D. Boughton W. C. (1987) Tests of ARMA model forms for rainfall-runoff modelling. *Journal of hydrology*, 91; 29-47.
- Werbos (1974) Beyond regression: new tools for prediction and analysis in the behavioral sciences. PhD thesis, Harvard University.
- Werbos P. J. (1990) Backpropagation through time: what it does and how to do it. *Proceedings of the IEEE*, 78(10), 1990.
- Widrow B. Hoff M. (1960) Adaptive switching circuits, IRE WESCON Convention Record, New York: IRE, 96-104.
- Xiong L. Guo. S. (1999) Short term streamflow forecasting using artificial neural network. *Journal of hydrology*, 216; 111-123.
- Yang X. (1993) Mise au point d'une méthode d'utilisation d'un modèle pluie-débit conceptuel pour la prévision des crues en temps réel. Thèse de Doctorat, Ecole Nationale des Ponts et Chaussées, Paris, 350 p.
- Yeddou H. (1998) Etude de synthèse sur les réseaux de neurones et leurs applications. Thèse de Magister, ENP, Alger, 200 p.
- Young P. Wallis S. (1985) Recursive estimation : a unified approach to the identification, estimation and forecasting of hydrological sytem. *Appl. Math. Computer*, 17; 299-334.
- Zadeh L. (1965) Fuzzy sets. *Information and control*, 8; 338-353.
- Zadeh L.A. (1971) Quantitative fuzzy semantics, *Information Sciences*, 3; 159- 176.
- Zadeh L.A. (1975) "The concept of a linguistic variable and its application to approximate theory", *Information Sciences*, 8-9; Part 1: (8), 199-249.
- Zadeh L. A. (1994) Soft computing and fuzzy logic, *IEEE Software*, November issue, 48-56.
- Zhao R.J., Zhang Y.L., Fang L.R., Xiu X.R. # Zhang Q.S. (1991) The Xinanjiang model, Hydrological Forecasting. *IAHS Publication*, 129; 351-356.
- Zoch R.T. (1934) On the relation between rainfall and streamflow. *Monthly whether review*, 62; 315-322.

Annexes

Annexe 1 : Présentation des bassins

Introduction

Les données pluviométriques et hydrométriques disponibles ont été recueillies au niveau de l'Agence Nationale des Ressources Hydrauliques (ANRH) et de l'Agence Nationale des Barrages (ANB), pour les mesures de l'évapotranspiration potentielle (ETP), nous disposons des données établies par l'Office National Météorologique (ONM), et des données de l'ANB.

Les modèles pluie-débit intègrent trois types de variables hydrométéorologiques, à savoir les précipitations, l'évapotranspiration, et le débit observé d'une station hydrométrique.

Nous avons les stations des bassins versants de la Cheffia, de l'Isser, du Sebbaou, de la Tafna et de Zardezas

Les tableaux A-I.1 et A-I.2 présentent les caractéristiques géographiques et hydrologiques des bassins étudiés.

MODELISATION PLUIE-DEBIT MENSUELLE ET JOURNALIERE PAR LES MODELES CONCEPTUELS ET LES SYSTEMES NEURO-FLOUS (APPLICATION AUX BASSINS ALGERIENS)

Stations	Bassins	Coordonnées	Date de mise en service
Cheffia (031501)	Cheffia	X=977.70 - Y=381.32	1978
Lakhdaria (090501)	Isser	X=579.10 - Y=368.75	1954/65
La Traïlle (090408)	Isser	X=559.25 - Y=348.25	1968
Baghdia (022001)	Sebbaou	X=603.70- Y=390.05	1969
Remchi (160703)	Tafna	X=122.90- Y=208.80	1973
Zardezaz (030902)	Zardezaz	X=878.75 - Y=370.67	1989

Tableau A-I.1 caractéristiques géographiques des bassins versants

Le bassin de l'Isser jusqu'à la station de Lakhdaria présente une grande superficie, cette station est exploitée sans arrêt depuis 1965.

Par contre la station de Zardezaz sa superficie est relativement faible, et la série est récemment exploitée.

Tableau A-I.2 Caractéristiques hydrologiques des bassins versants

Stations	Cheffia	Lakhdaria (090501)	La Traïlle (090408)	Sebbaou (022001)	Tafna (160703)	Zerdezaz
Superficie (km ²)	575	3615	2570	2390	1935	345
Coefficient Kc	1.20	1.26	1.24	1.14	1.14	1.26
Altitude moyenne (m)	420	710	600	590	720	628
Indice de pente de Roche	0.11	0.11	0.12	0.20	0.12	0.17
Apport (Hm ³ /an)	134.1	198.7	154.5	643.1	34.7	51.4
Débit spécifique (l/s/km ²)	7.39	1.74	1.90	8.60	0.57	2.83

Le bassin du Sebbaou présente un débit spécifique élevé, il est considéré comme étant le plus arrosé, et cela Contrairement au bassin de la Tafna, ou son débit spécifique est inférieur à 0.6 l/s/ km², il est caractérisé donc par un climat semi-aride.

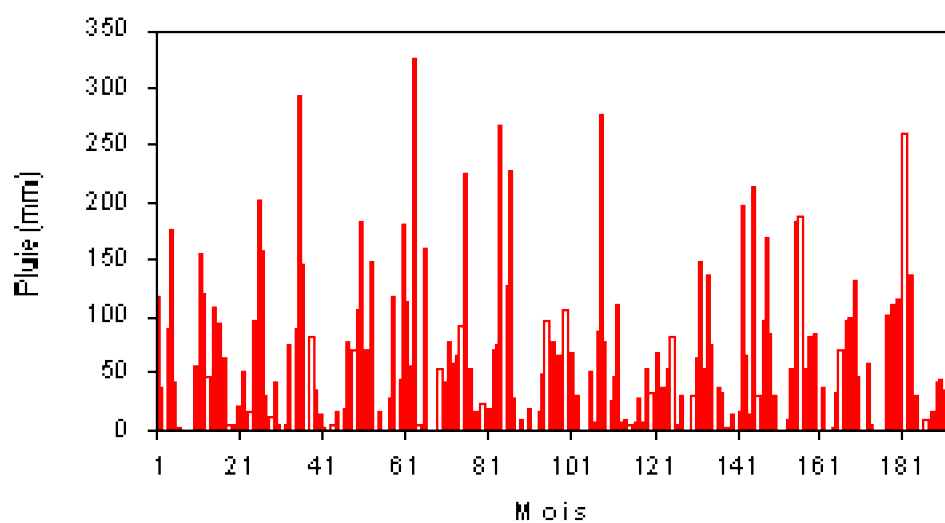


Figure A-1.1 Evolution de la pluie mensuelle de la station de la Cheffia

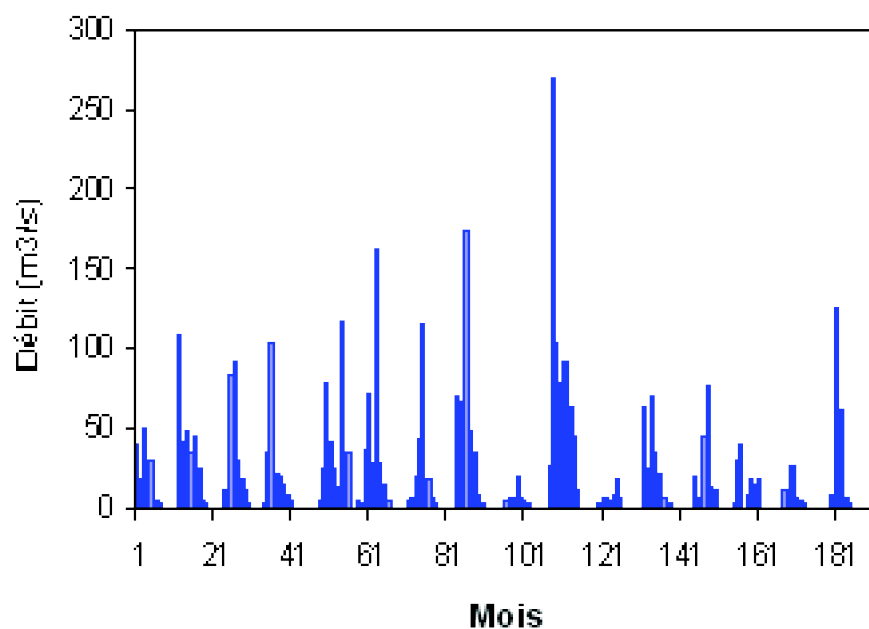


Figure A-1.2 Evolution des débits mensuels du bassin du Sebaou

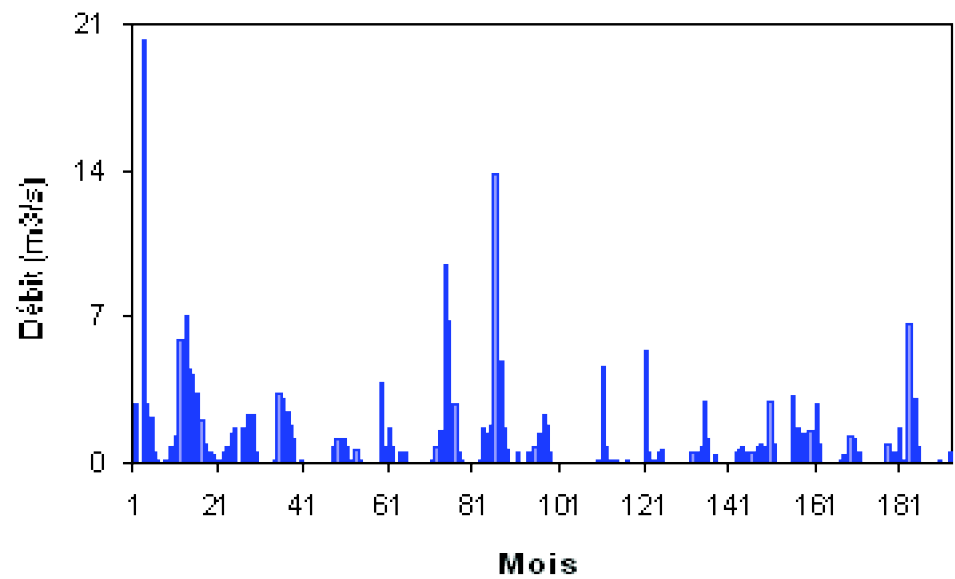


Figure A-1.3 Evolution des débits mensuels du bassin de la Tafna

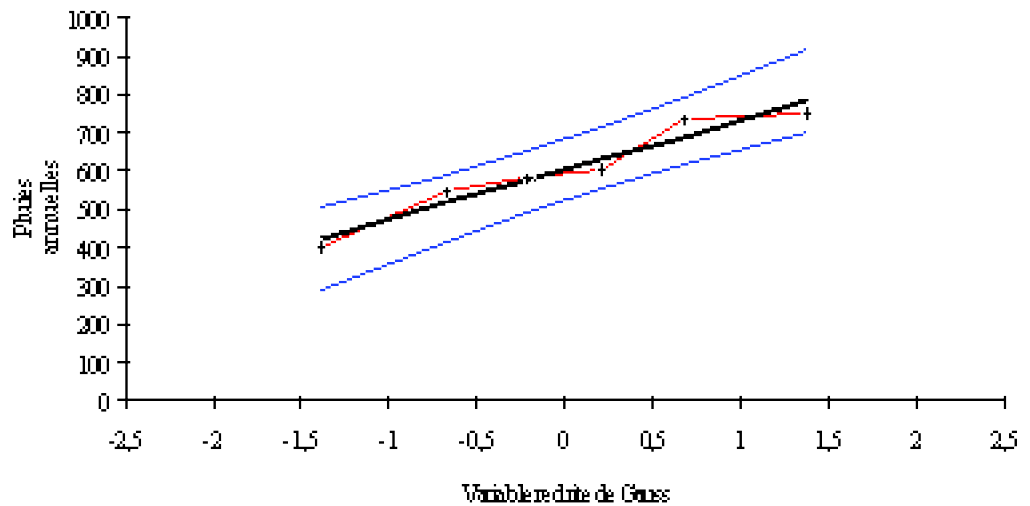


Figure A-1.4 Ajustement des pluies annuelles (1991-1996) à la loi normale (Station de Zardezas)

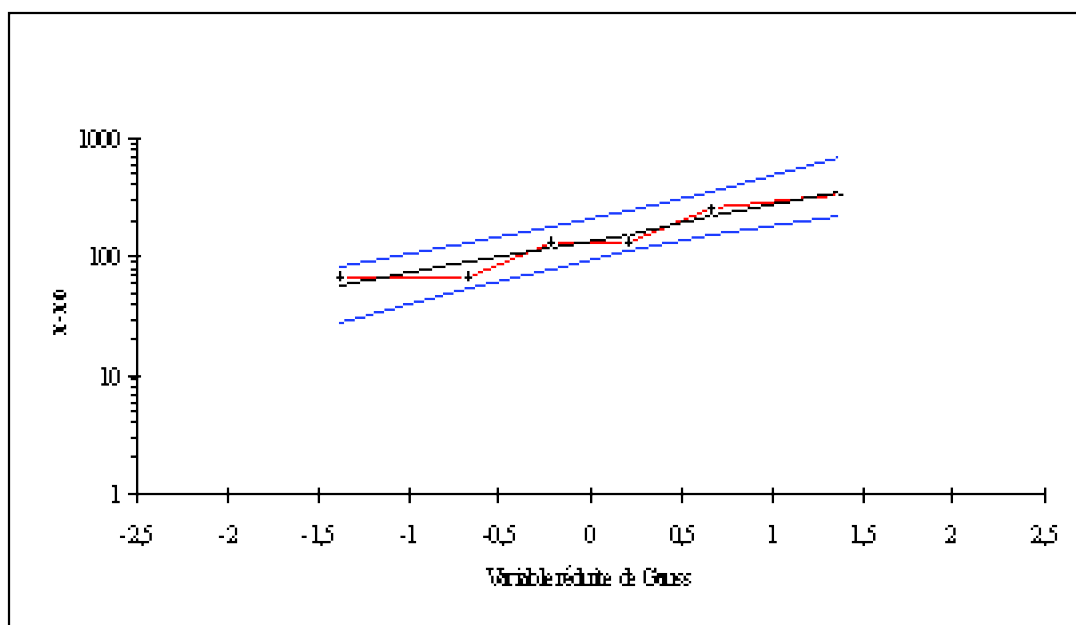


Figure A-1.5 Ajustement des apports annuels (1990-1995) à la loi log-normale (Station de l'Isser)

Annexe 2 : Présentation des modèles

Introduction

Les modèles pluie-débit permettent de simuler la réponse d'un bassin versant à partir de certaines variables hydrométéorologiques telles que les pluies. Plusieurs modèles ont été conçus dans différents pays, selon leurs caractéristiques climatiques.

Dans cette thèse, nous avons travaillé avec deux grandes catégories des modèles pluie-débit globaux à savoir les modèles conceptuels, et les modèles «boîte noire», aux pas de temps mensuel et journalier appliqués aux bassins versants du Nord algérien.

1. Les modèles conceptuels

Ces modèles simulent le processus pluie-débit par un agencement de réservoirs censés reproduire les différentes phases du cycle hydrologique.

Les modèles conceptuels, par leur simplicité représentent la catégorie des modèles pluie-débit la plus utilisée en hydrologie. Certains de ces modèles présentent une particularité qui est déduite du type de climat et de ses caractéristiques tels que le modèle HBV en Suède (Bergstrom, 1976), le modèle TOPMODEL (Beven, 1979) au Royaume uni, le modèle Xianjiang (Zhao et al., 1980) en chine, ou encore le modèle GR3j (Edijatno Michel, 1989, Edijatno et al., 1999) pour les climats méditerranéens.

Au pas de temps mensuel, les modèles GR2M issus de la famille des modèles GR ont été les plus exploités et appréciés par leur parcimonie et leur qualité de simulation. Ces modèles ont été comparés avec deux autres modèles que sont le modèle de

Thornthwaite et le modèle ABCD.

Le modèle GR2M (2002) issu des travaux de Mouelhi (2002), a été conçu après une comparaison de plusieurs modèles mensuels, et une adaptation du modèle conceptuel journalier GR4J (Perrin, 2000) à l'échelle mensuelle.

Ce modèle comporte deux réservoirs et deux paramètres $X1$ et $X2$, il est caractérisé par deux fonctions (Figure III-4):

Une fonction de production assurée par le réservoir sol $S(i)$ qui est limité par une capacité maximale $X2$ (paramètre à caler);

Une fonction de transfert représentée par le réservoir $R(i)$, dont le niveau est modifié par l'échange souterrain en fonction du paramètre $X1$.

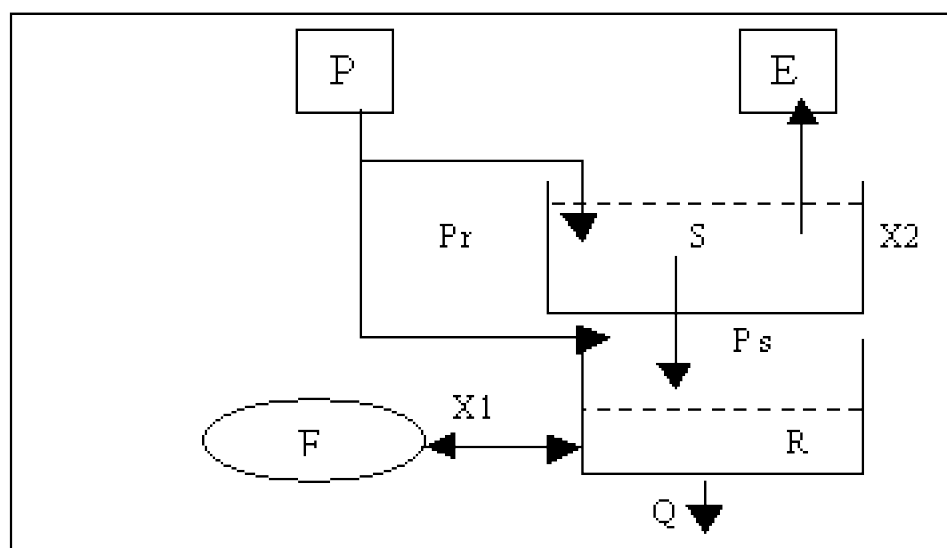


Figure A-II.1 Architecture du modèle GR2M (2002)

Au pas de temps journalier, une étude comparative de quatre modèles les plus robustes à savoir le modèle GR3J (1999), le modèle GR4J (2000), le modèle CREC à 8 paramètres (1988) ainsi que le modèle IHACRES démontre la robustesse du modèle GR3J (1999), en particulier en phase de calage du modèle.

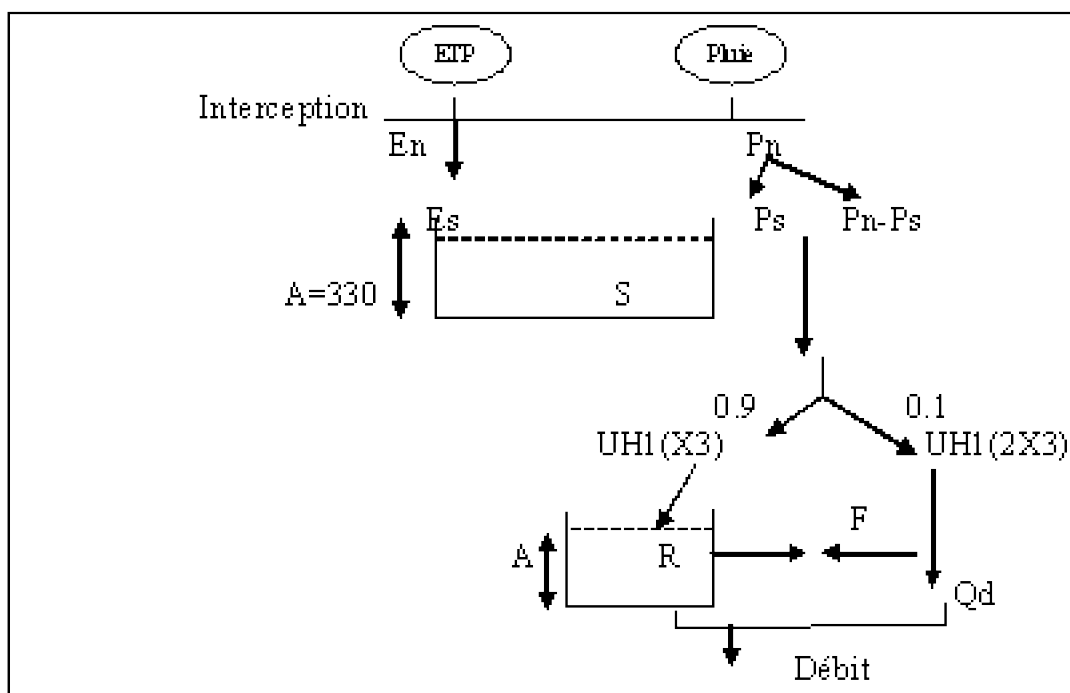


Figure A-II.2 Architecture du modèle GR3J (1999)

Cependant, la baisse significative du niveau de simulation des modèles en phase de validation, et la mauvaise reconstitution des débits caractéristiques, tels que les étiages et les crues, nous a conduit à tester d'autres approches.

Les programmes informatiques des deux modèles GR2M (1994) et GR3J (1999) sont détaillés comme suit:

```

C----- SUBROUTINES DU MODELE      GR2M VERSION 1994

Subroutine GR2MP(E,P1,P2,H,S,SUP,Q)
A=200
EN=E*(P**0.5)+(E**0.5)**2)
EN=E*(P**0.5)+(E**0.5)**2)
DNI=20*EN
EN1=20*EN
V=TANH(PN/A)
W=TANH(EN1/A)
C   CALCUL DU NIVEAU D'HUMIDITE SOUS LA PLUIE P
H1=(H+(A**V)/(1+(E**V/A)))
C   CALCUL DE L'EXCES D'EAU P
DE=DNI+H-H1
C   CALCUL DU NIVEAU D'HUMIDITE SOUS ETP
E2=H1*(1-W)/(1+(W*(1-(H1/A))))
E-E2
C   NIVEAU D'HUMIDITE DU RESERVOIR SOL.
S=S+4*PE/S
C S=20*E
S-S Q S
C   DEBIT TOTAL -----
Q=(QS+(PE/S))

C   TRANSFORMATION (mm-->R/c)
IR= SUP*1000/(24*3600*24)
C SUP= SUPERFICIE
Q=Q*IR
Write(4,20)Q
2) format(' 2f10.3)
RETURN
END

```

Figure A-II.3 Ecriture informatique (Fortran) du modèle GR2M (1994)

multicouches qui seront utilisés dans notre thèse.

Le réseau multicouche : Le **réseau multicouche** est caractérisé par une distribution en couches des neurones. On distingue trois types de couches : **La couche d'entrée** qui transmet uniquement les valeurs d'entrée pour les distribuer aux neurones de la première couche cachée. **La ou les couches cachées** qui extrai(en)t des caractéristiques du vecteur d'entrée : ce qui correspond à traiter l'information. **La couche de sortie** qui donne la réponse du réseau. Chaque neurone est alors connecté : Soit vers une couche suivante immédiate (on parle alors de **réseaux à couches simples**), Soit vers une couche quelconque située en aval (ces réseaux sont alors appelés **réseaux à couches non simples**). Les connexions entre les neurones peuvent aussi se faire d'une couche sur elle-même mais on se rapproche alors fortement de la connexion des réseaux bouclés.

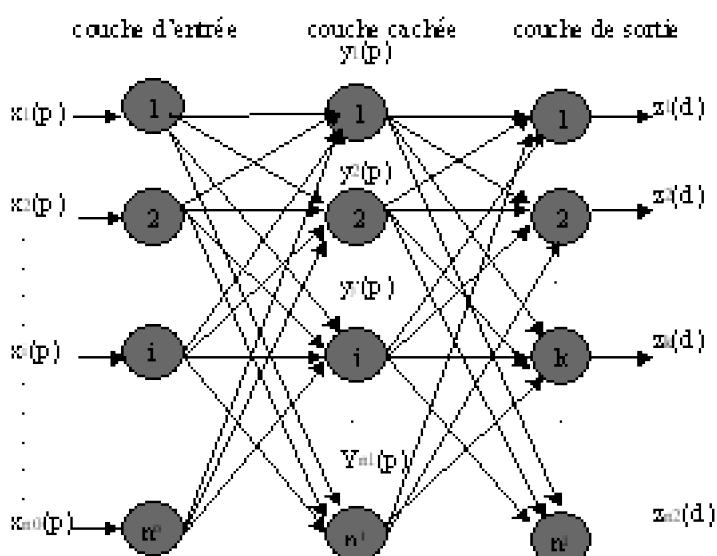


Figure A-II.5 Présentation d'un réseau neuronal multicouche

L'apprentissage : l'apprentissage ou modification des poids w_i est la phase importante car toute la connaissance est mémorisée dans les coefficients synaptiques qui sont fixés lors de la phase d'apprentissage. On distingue trois types d'apprentissage : L'apprentissage supervisé où des associations entrée-sorties sont présentées au réseau par un "professeur", l'apprentissage non supervisé où on ne connaît pas les sorties par rapport aux entrées, et l'apprentissage par renforcement qui associe les deux techniques. Il existe différentes méthodes d'apprentissage (La méthode des moindres carrés, la rétropropagation du gradient, La méthode de Windrow-Hoff,...) dont le but est de minimiser un critère souvent appelé "fonction de coût" par ajustement des poids.

La méthode par rétropropagation du gradient (Backpropagation) présente des avantages de non-linéarités et de calculs pratiques: L'intérêt de cette règle est de prendre en compte les modifications antérieures (individuelles) des poids dans le but d'éviter des oscillations perpétuelles.

```

SUBROUTINE GRADIENT
  COMMON /GRADIENT/ W(100,100), V(100), U(100), Z(100), X(100), Y(100),
  & W1(100,100), W2(100,100), W3(100,100), W4(100,100), W5(100,100),
  & W6(100,100), W7(100,100), W8(100,100), W9(100,100), W10(100,100),
  & W11(100,100), W12(100,100), W13(100,100), W14(100,100), W15(100,100),
  & W16(100,100), W17(100,100), W18(100,100), W19(100,100), W20(100,100),
  & W21(100,100), W22(100,100), W23(100,100), W24(100,100), W25(100,100),
  & W26(100,100), W27(100,100), W28(100,100), W29(100,100), W30(100,100),
  & W31(100,100), W32(100,100), W33(100,100), W34(100,100), W35(100,100),
  & W36(100,100), W37(100,100), W38(100,100), W39(100,100), W40(100,100),
  & W41(100,100), W42(100,100), W43(100,100), W44(100,100), W45(100,100),
  & W46(100,100), W47(100,100), W48(100,100), W49(100,100), W50(100,100),
  & W51(100,100), W52(100,100), W53(100,100), W54(100,100), W55(100,100),
  & W56(100,100), W57(100,100), W58(100,100), W59(100,100), W60(100,100),
  & W61(100,100), W62(100,100), W63(100,100), W64(100,100), W65(100,100),
  & W66(100,100), W67(100,100), W68(100,100), W69(100,100), W70(100,100),
  & W71(100,100), W72(100,100), W73(100,100), W74(100,100), W75(100,100),
  & W76(100,100), W77(100,100), W78(100,100), W79(100,100), W80(100,100),
  & W81(100,100), W82(100,100), W83(100,100), W84(100,100), W85(100,100),
  & W86(100,100), W87(100,100), W88(100,100), W89(100,100), W90(100,100),
  & W91(100,100), W92(100,100), W93(100,100), W94(100,100), W95(100,100),
  & W96(100,100), W97(100,100), W98(100,100), W99(100,100), W100(100,100)
  DIMENSION W(100,100), V(100), U(100), Z(100), X(100), Y(100),
  & W1(100,100), W2(100,100), W3(100,100), W4(100,100), W5(100,100),
  & W6(100,100), W7(100,100), W8(100,100), W9(100,100), W10(100,100),
  & W11(100,100), W12(100,100), W13(100,100), W14(100,100), W15(100,100),
  & W16(100,100), W17(100,100), W18(100,100), W19(100,100), W20(100,100),
  & W21(100,100), W22(100,100), W23(100,100), W24(100,100), W25(100,100),
  & W26(100,100), W27(100,100), W28(100,100), W29(100,100), W30(100,100),
  & W31(100,100), W32(100,100), W33(100,100), W34(100,100), W35(100,100),
  & W36(100,100), W37(100,100), W38(100,100), W39(100,100), W40(100,100),
  & W41(100,100), W42(100,100), W43(100,100), W44(100,100), W45(100,100),
  & W46(100,100), W47(100,100), W48(100,100), W49(100,100), W50(100,100),
  & W51(100,100), W52(100,100), W53(100,100), W54(100,100), W55(100,100),
  & W56(100,100), W57(100,100), W58(100,100), W59(100,100), W60(100,100),
  & W61(100,100), W62(100,100), W63(100,100), W64(100,100), W65(100,100),
  & W66(100,100), W67(100,100), W68(100,100), W69(100,100), W70(100,100),
  & W71(100,100), W72(100,100), W73(100,100), W74(100,100), W75(100,100),
  & W76(100,100), W77(100,100), W78(100,100), W79(100,100), W80(100,100),
  & W81(100,100), W82(100,100), W83(100,100), W84(100,100), W85(100,100),
  & W86(100,100), W87(100,100), W88(100,100), W89(100,100), W90(100,100),
  & W91(100,100), W92(100,100), W93(100,100), W94(100,100), W95(100,100),
  & W96(100,100), W97(100,100), W98(100,100), W99(100,100), W100(100,100)
  DO 100 I=1,100
    DO 100 J=1,100
      W(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W1(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W2(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W3(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W4(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W5(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W6(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W7(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W8(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W9(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W10(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W11(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W12(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W13(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W14(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W15(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W16(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W17(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W18(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W19(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W20(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W21(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W22(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W23(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W24(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W25(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W26(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W27(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W28(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W29(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W30(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W31(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W32(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W33(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W34(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W35(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W36(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W37(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W38(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W39(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W40(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W41(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W42(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W43(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W44(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W45(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W46(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W47(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W48(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W49(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W50(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W51(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W52(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W53(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W54(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W55(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W56(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W57(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W58(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W59(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W60(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W61(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W62(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W63(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W64(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W65(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W66(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W67(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W68(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W69(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W70(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W71(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W72(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W73(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W74(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W75(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W76(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W77(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W78(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W79(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W80(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W81(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W82(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W83(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W84(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W85(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W86(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W87(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W88(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W89(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W90(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W91(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W92(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W93(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W94(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W95(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W96(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W97(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W98(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W99(I,J)=0.0
    END DO
  END DO
  DO 100 I=1,100
    DO 100 J=1,100
      W100(I,J)=0.0
    END DO
  END DO
  RETURN
END

```

Figure A-II.6 Subroutine du Gradient conjugué (fortran), réseaux de neurones artificiels

2.2 La logique floue

Introduction : Le raisonnement humain se base généralement sur des connaissances imparfaites, soit parce qu'on a un doute sur leur validité, elles sont alors incertaines, soit parce que nous éprouvons une difficulté à les exprimer clairement, elles sont alors imprécises. Les méthodes probabilistes s'attaquent aux événements incertains alors que la logique floue est un concept qui permet de traiter les raisonnements imprécis de l'être humain. La théorie des ensembles flous permet alors de généraliser des événements naturels tout en faisant appel à des outils mathématiques. Elle présente au moins 2 intérêts : Elle permet de modéliser des connaissances imprécises ou vagues; elle est le seul moyen de traiter dans un même cadre des connaissances fournies numériquement et des connaissances exprimées symboliquement.

La logique floue s'appuie sur la syntaxe *Si ... alors*. Un exemple est donné en hydrologie : *Si* la pluie est forte *alors* le débit est élevé.

Le schéma représentant la méthodologie générale de l'utilisation de la logique floue :



Figure A-II.7 Etape de modélisation par la logique floue

La fuzzification C'est l'étape de passage du domaine numérique au domaine symbolique. Cette étape est nécessaire dès que l'on veut manipuler à l'aide de la théorie des ensembles flous des grandeurs physiques mesurables précises ou non. Elle permet entre autre de déterminer la fonction d'appartenance d'une variable à un ensemble flou.

Ensemble flou la définition d'un ensemble flou d'après Zadeh (1965, 1971) :

Soit un ensemble X dénombrable ou non et x un élément de X alors un ensemble flou A de X est un ensemble de couples tels que :

$$\forall x \in X \{ (x, \mu_A(x)) \}$$

Avec $\mu_A(x)$ la fonction d'appartenance qui prend des valeurs sur l'intervalle $[0;1]$.

Si cet intervalle se réduit aux valeurs $\{0;1\}$ alors l'ensemble A est un ensemble classique (ordinaire).

Fonction d'appartenance Elle détermine 3 zones :

Une zone où les éléments du référentiel appartiennent totalement à l'ensemble flou ;
==> $\mu_A(x)=1$

Une zone où les éléments sont exclus; ==> $\mu_A(x) = 0$

Une zone floue (appelée aussi zone de transition) où l'appartenance s'estime à des degrés divers. ==> $0 < \mu(x) < 1$

La fonction d'appartenance est donc le degré d'appartenance de x à l'ensemble flou A . Son rôle est de préciser numériquement la signification d'un ensemble flou énoncé en principe sous forme linguistique. C'est donc le passage du domaine numérique au domaine symbolique. Il existe différentes fonctions d'appartenance dont : Fonction en forme de trapèze; Fonction en forme de pointe; Fonction Gaussienne (continuellement dérivable). La forme et les paramètres des différentes fonctions d'appartenance sont à adapter au cas particulier à traiter.

L'inférence Cela consiste à traiter l'information par des règles floues.

Les règles floues Elles expriment un lien entre des propositions floues élémentaires : "La pluie est faible ou l'humidité du sol est élevée, alors le débit est élevé". Elles forment le noyau des systèmes d'inférence flous.

Une règle floue se décompose en 2 parties :

La prémisse : endroit où une ou plusieurs propositions floues sont énoncées.

La conclusion ou la conséquence : c'est-à-dire la réponse.

Une règle décrit un lien causal entre les valeurs x_i de la prémisse et celles de la conclusion à l'aide d'un opérateur d'implication flou.

Opérateurs d'implication flous Il existe différents opérateurs d'implication flous : T_norme, T_conorme, Minimum de Mandani et de Zadeh, Maximum de Lukasiewicz et de Giles, produit de Goguen, etc.

Les 2 opérateurs classiquement utilisés dans les applications de la théorie des ensembles flous sont les fonctions Max. et Min. Cependant, les T_normes peuvent être employées comme opérateur de conjonction en lieu et place du Min., tandis que les T_conormes peuvent remplacer le Max. dans les disjonctions.

La défuzzification est l'opération qui permet de passer d'une représentation sous forme de variable linguistique à une variable numérique applicable physiquement: c'est la phase inverse de la fuzzification. Il existe plusieurs méthodes possibles dont Le Max. et le centre de gravité. Lorsque la conclusion de la règle est précise, spécifique et certaine, c'est-à-dire lorsque toutes les règles floues en présence ont une conclusion non floue alors la défuzzification est implicite.

La fuzzification va donc nous donner le degré d'appartenance qui va nous permettre d'obtenir la conséquence grâce aux règles floues. La défuzzification nous donnera le résultat sous forme numérique.

Association d'un modèle flou neuronal: l'approche neurofloue

On peut représenter cette analyse sous la forme suivante :

Les neurones de la première couche cachée vont représenter les différentes zones de la logique floue. Les différentes sorties de cette couche donneront donc les degrés d'appartenance correspondant aux différentes zones. On introduit alors dans une deuxième couche cachée les règles d'inférence de la logique floue.

ce qui nous permet d'obtenir la règle : **si x ET si y ALORS Conséquence**

La couche de sortie va nous donner le résultat qui peut être obtenu par défuzzification. Le schéma ci-dessus met en évidence le parallélisme massif entre le raisonnement flou et la représentation connexionniste. On retrouve les réseaux de neurones lors de l'incorporation des poids par apprentissage.



On peut ajouter des poids plus ou moins prédéfinis ce qui nous donnera un basculement entre la **connaissance explicite** (Logique Floue : Les poids sont tous prédéfinis) et le **raisonnement implicite** (Réseaux de Neurones : Poids établis par apprentissage).

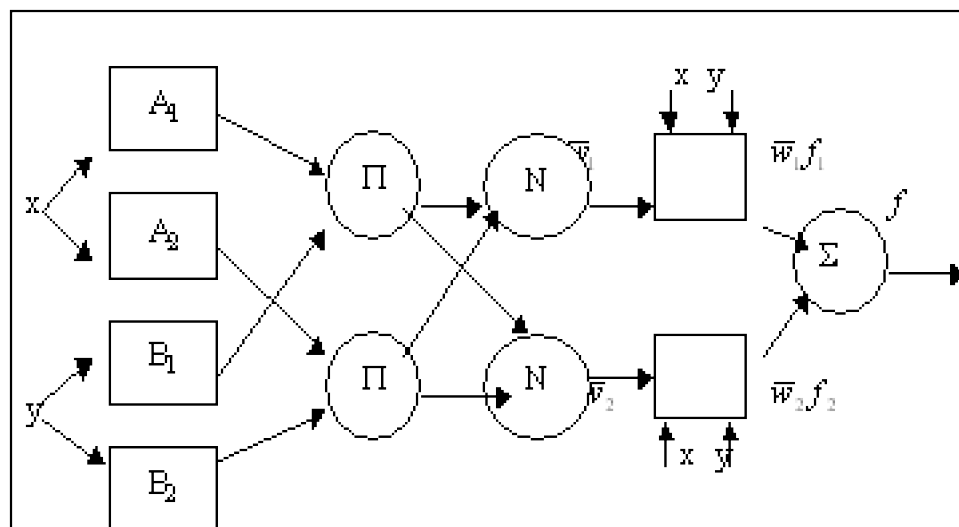
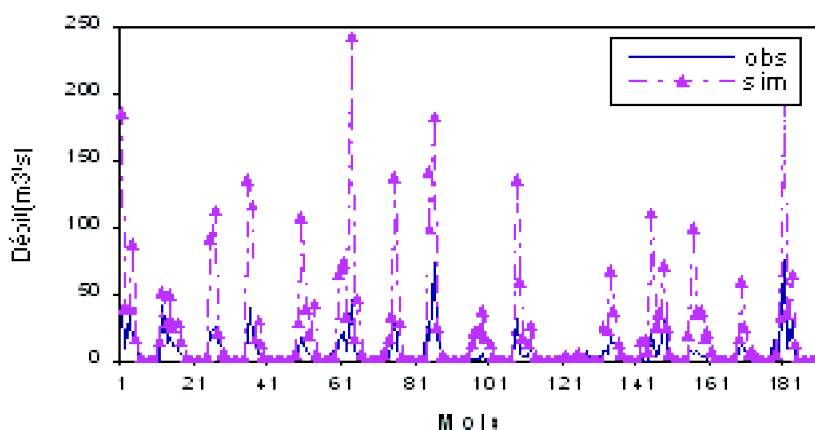


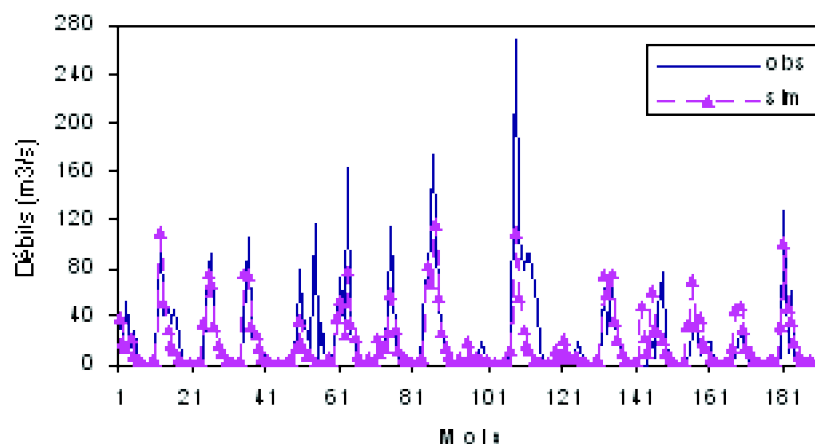
Figure A-II.8 Modèle adaptatif ANFIS

Conclusion : Cette association semble apporter au niveau de l'apprentissage qui est ainsi moins complexe (certains poids étant préétablis) et au niveau de la compréhension du cheminement de la connaissance dans les réseaux de neurones.

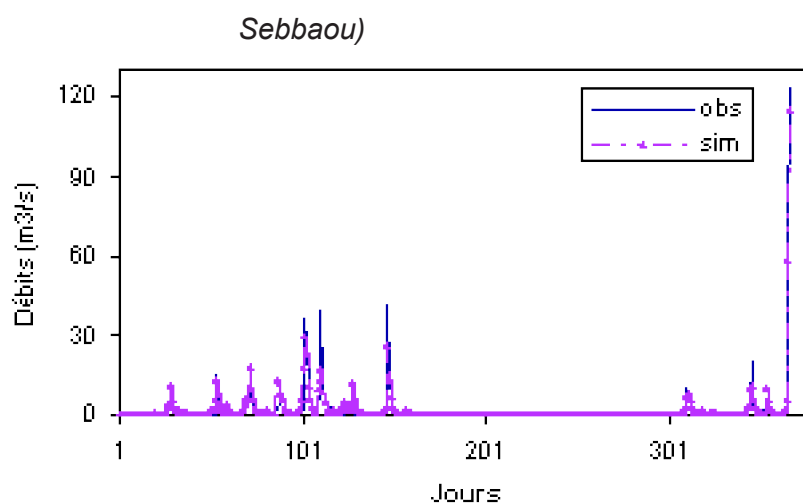
Résultats des modèles conceptuels



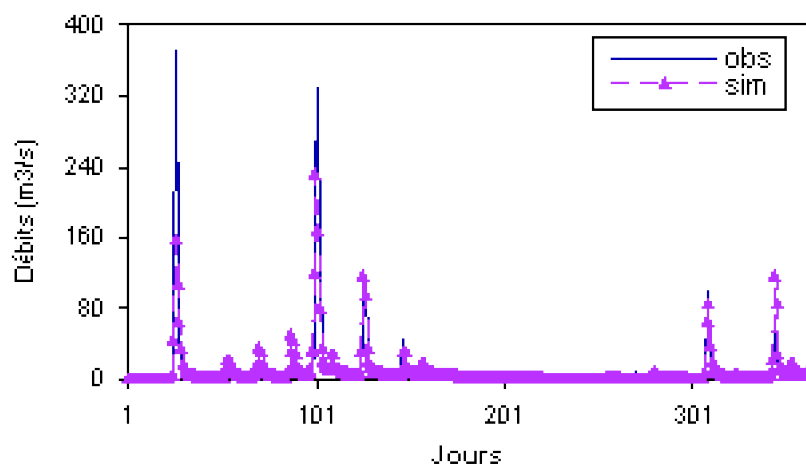
A-III.1 Simulation des débits mensuels par le modèle GR2M (2002) (Station de Lakhdaria)



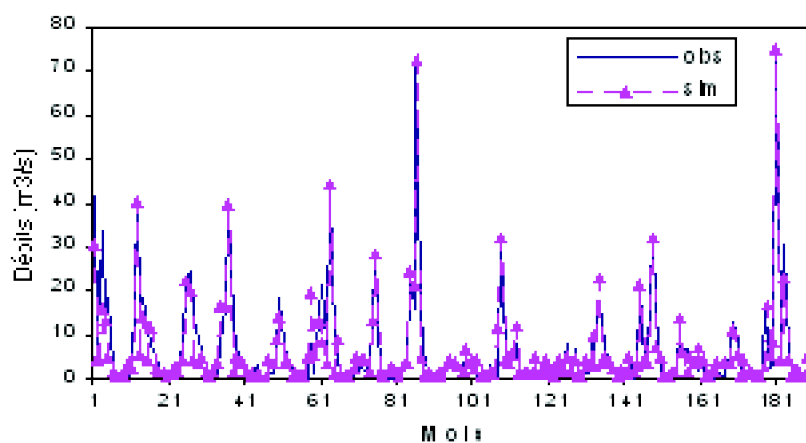
A-III.2 Simulation des débits mensuels par le modèle Thornthwaite (modifié) (Station du)



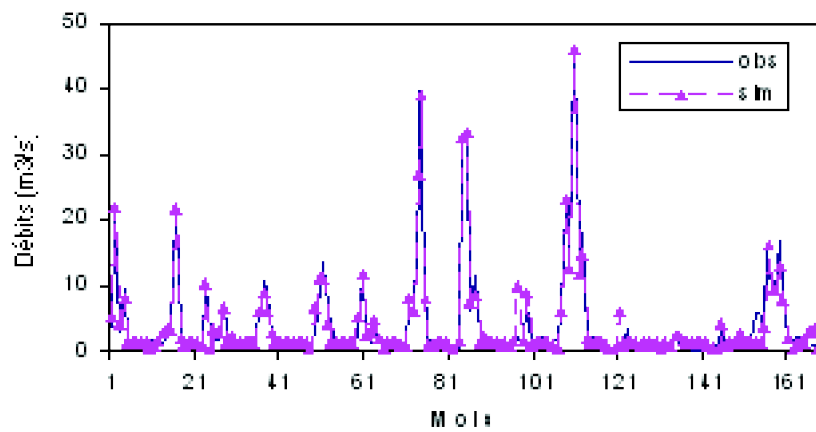
A-III.3 Simulation des débits mensuels par le modèle GR2M (1994) (Station de Latriille)



A-III.4 Simulation des débits journaliers par le modèle GR3J (Station de Cheffia – Année 1991)

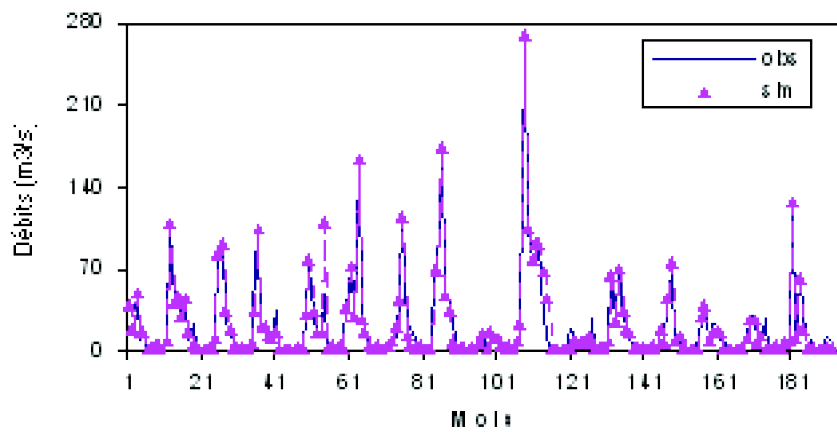


A-III.5 Simulation des débits journaliers par le modèle IHACRES (Station de Zardezas – Année 1992)

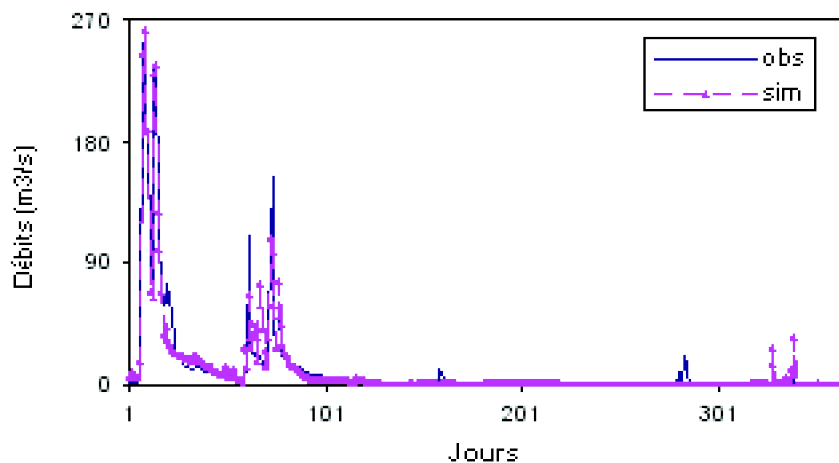


A-III.6 Simulation des débits journaliers par le modèle GR4J (Station de l'Isser – Année 1992)

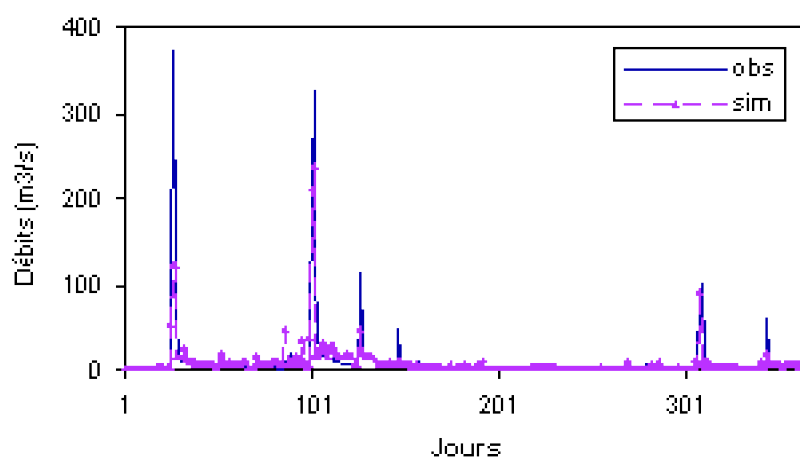
Résultats des modèles neuroflous



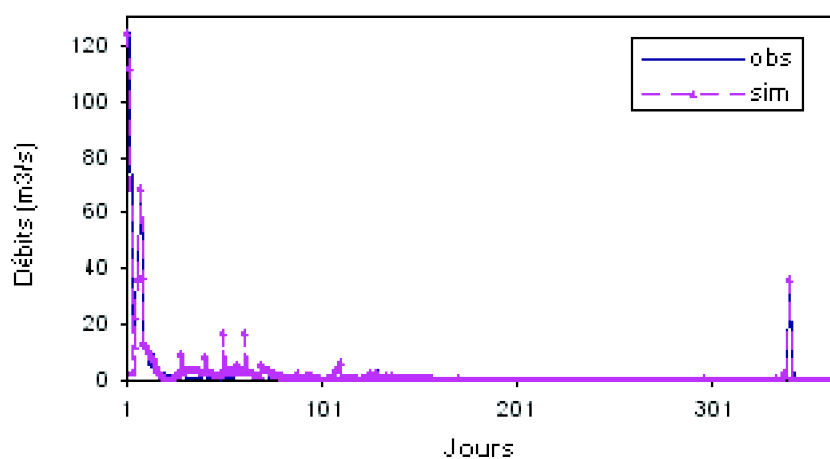
A-III.7 Simulation des débits mensuels par le modèle neuroflou à 30 règles (Station de Lakhdaria)



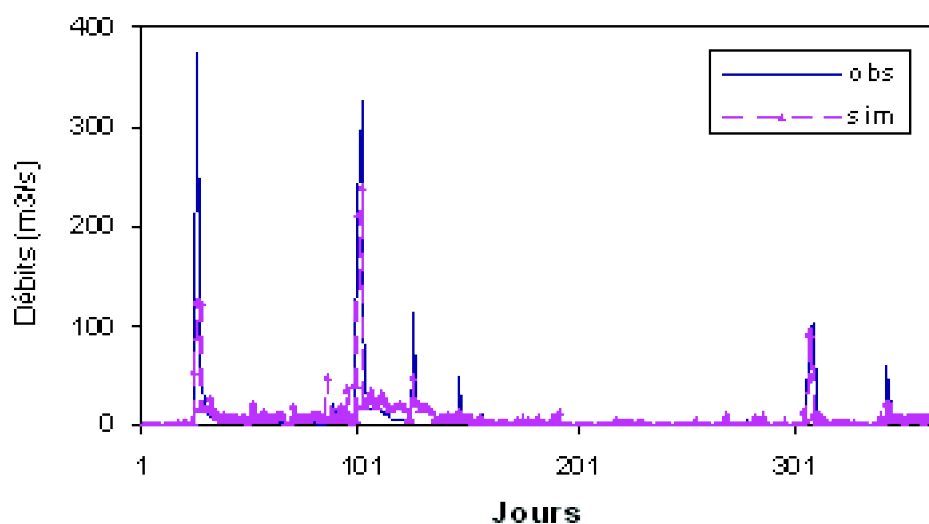
A-III.8 Simulation des débits mensuels par le modèle neuroflou à 2 entrées (5 règles) (Station de Cheffia)



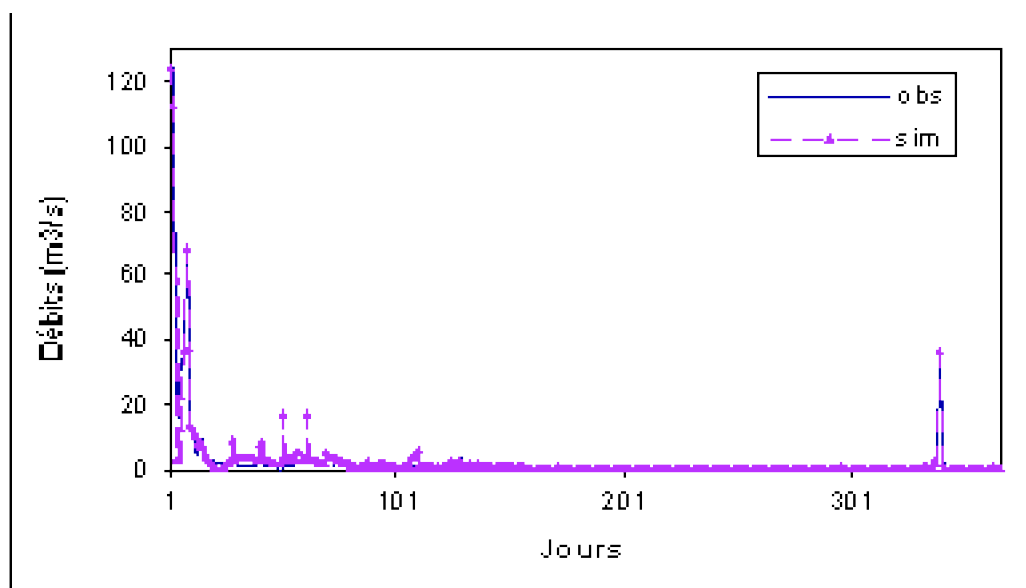
A-III.9 Simulation des débits mensuels par le modèle neuroflou à 2 entrées (7 règles)
(Station du Sebbaou)



A-III.10 Simulation des débits journaliers par le modèle neuroflou à 3 entrées (Station de
l'Isser – Année 1995)



A-III.11 Simulation des débits journaliers par le modèle neuroflou à 2 entrées (11 règles)
(Station de l'Isser -1992)



A-III.12 Simulation des débits journaliers par le modèle neuroflou à 2 entrées (11 règles)
(Station de Zardezas – Année 1993)