

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

2/03

ECOLE NATIONALE POLYTECHNIQUE



المدرسة الوطنية المتعددة التقنيات
Ecole Nationale Polytechnique

المدرسة الوطنية المتعددة التقنيات
BIBLIOTHEQUE — المكتبة
Ecole Nationale Polytechnique

DEPARTEMENT HYDRAULIQUE

**PROJET FIN D'ETUDES
POUR L'OBTENTION DU DIPLOME
D'INGENIEUR D'ETAT EN HYDRAULIQUE**

THEME

**Prévision de la
pluviométrie au pas
de temps mensuel
Par la logique floue**

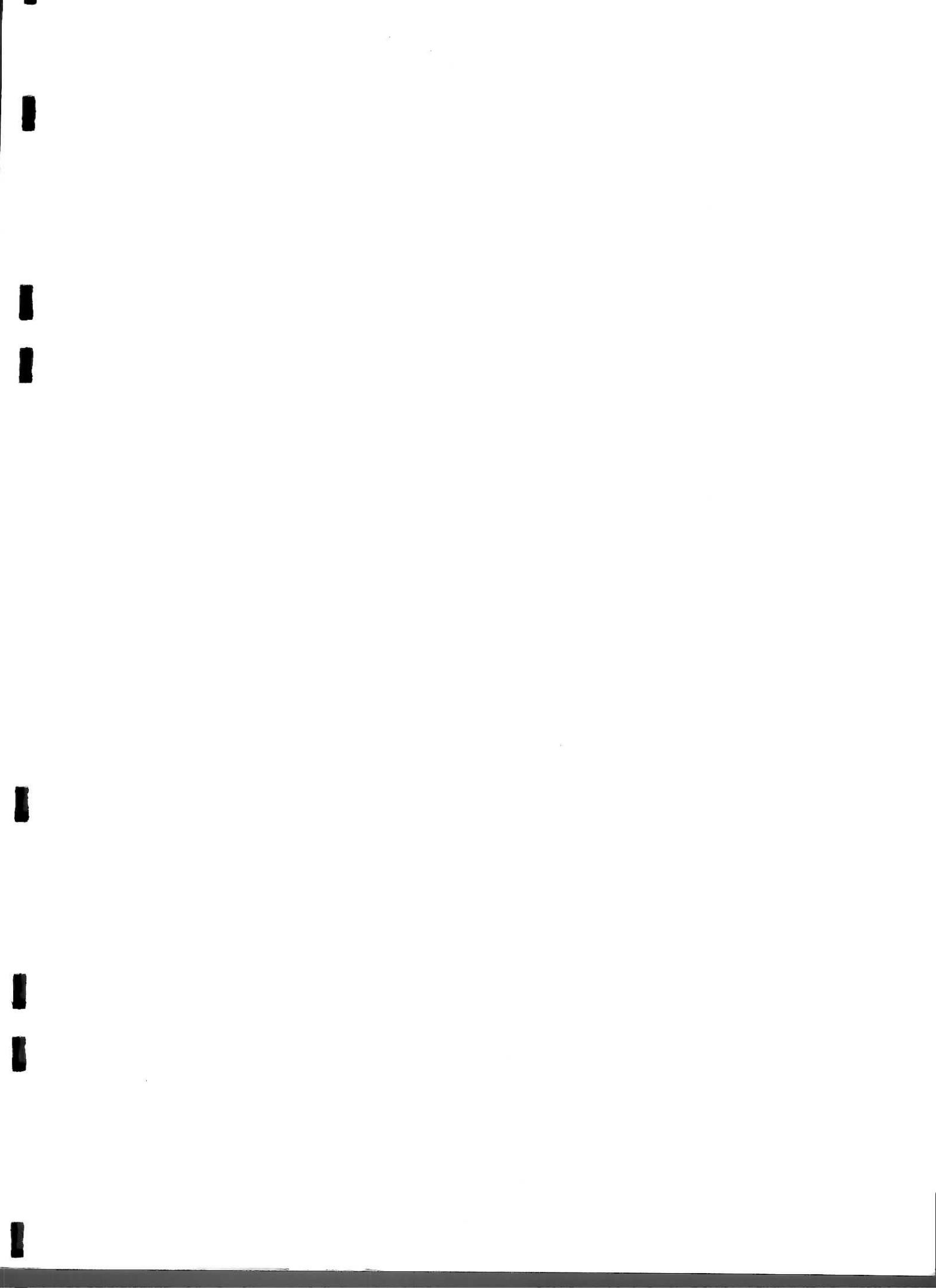
Réalisé par :

**M^r : BILLEL
BENMERZOUG**

Promoteur :

M^r : A. BERMAD

**PROMOTION
2003**



REMERCIEMENTS

*J'adresse mes sincères remerciements à Monsieur BERMAD, mon promoteur ,
qui n'a pas ménagé ses efforts pour m'aider à travers mon travail, à Monsieur
DECHEMI, professeur à l'école Nationale Polytechnique pour son soutien, le
suivi et l'intérêt qu'il a porté à ce travail.*

*J'exprime ma profonde reconnaissance à Monsieur T. BENKACI pour son aide
précieuse et sa disponibilité.*

*Je tiens aussi à exprimer ma profonde gratitude et mes remerciements les plus
sincères à Monsieur MEZGHICHE, Monsieur MAAMRA de l'USTHB pour
leurs coopération et contribution à la réalisation de ce travail.*

*Que tous ceux qui m'ont aidé de près ou de loin, dans ma formation et à la
réalisation de ce présent travail reçoivent ici le témoignage de ma grande
sympathie.*

DEDICACES

Je dédie cette thèse

à la mémoire de mon très cher cousin Fawzi

spécialement à mes parents ABDELKADER et NINO

ainsi qu'à ma sœur AMEL, qui m'ont soutenu tout au long de mes études

à mes grands-parents qui me sont très chers

à mes oncles, en particulier ABDELHAK et KHALIDO, à toutes mes tantes

à mes cousins et cousines AKILA et AMIR

et à tous les miens

Résumé :

Pour une gestion optimale des ressources hydriques, et la **prévision** des situations de crise, tel que, les inondations ou sécheresses. Une technique mathématique nouvelle, introduite par le Professeur ZADEH sous le nom de **la logique floue**, a été appliquée pour la **prévision des précipitations** afin d'apporter d'éventuelle amélioration à la compréhension de ce phénomène.

Mots clés :

Prévision , précipitation, la logique floue....

Abstract :

For an optimal management of the hydrous resources, and the **forecast** of the crises, such as, the floods or drynesses. A new mathematical technique, introduced by Professor ZADEH under the name of **fuzzy logic**, was applied for the **forecast of precipitation** so to bring improving to the comprehension of this phenomenon.

Key words:

Forecast, precipitation, fuzzy logic....

الملخص :

من أجل التسيير الحسن للموارد المائية و أخذ الاحتياطات لحالات الأزمات مثل الفيضانات و الجفاف. أعتمدت في دراستي هاته علي تقنية جديدة تعرف ب (المنطق المبهم) , و التي نستعملها للتحكم مسبقا في هذه الظاهرة الطبيعية التي تعرف بالتساقط حتي نتمكن من تفهمها و تحسين معالجتها.

الكلمات المفتاحية :

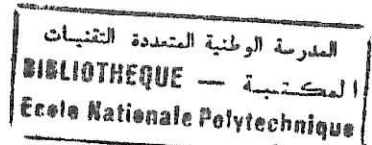
المنطق المبهم , التحكم المسبق , التساقط...

- Introduction	1
- Chapitre I Recherche bibliographique	3
- Chapitre II Collecte et critique des données	9
II .1- Introduction	10
II .2- Inventaire des données disponibles	10
II .3- Critique des données	11
II .3.1- Introduction	11
II .3.2- Méthodes de comblement de lacunes	12
II .3.2.1- Calcul de la moyenne mensuelle interannuelle	12
II .3.2.2- Méthode de régression linéaire simple	13
II .3.2.3- Simulation à partir de la fonction de répartition	14
II .3.3- Comblement de lacune.....	15
II .3.4- Détection des valeurs singulières	15
II .3.4.a- Présentation du test de GRUBBS et BECK	16
II .3.4.b- Analyse des résultats	16
II .4- Traitement statistique des données.....	17
II .4.1- Test d'homogénéité	17
II .4.2- Test d'indépendance	18
II .4.3 - Analyse des résultats	19
- Chapitre III La logique Floue	20
III .1 - Historique de la logique floue	21
III .1.1- Introduction	21
III .1.2- Le concept d'ensemble flou.....	22
III .1.3- Objectifs	22
III .1.4- Autres domaines d'applications	23
III .1.5- Conclusion	24
III .2- La théorie des sous-ensembles flous	25
III .2.1- Introduction	25
III .2.2- Imprécision et incertitude	25
III .2.3- Définition d'un ensemble flou.....	25
III .2.4- Définition d'une Fonction d'appartenance	26
III .2.5- Représentation d'une fonction d'appartenance	27
III .2.6- Méthodes de détermination de fonctions d'appartenances.....	27
III .2.7- Effet du contexte sur les fonctions d'appartenances.....	28
III .2.8- Les concepts de base de la théorie des sous ensembles flous	28
III .2.8.1- Support et noyau d'un sous ensemble flou	28
III .2.8.2- Définitions des T-normes, T-conormes et opérateur de négation	28
III .2.8.2.1- Définition d'une T-norme	28
III .2.8.2. 2- Définition d'une T-conorme	29
III .2.8.2. 3- Définition de la négation	29
III .2.8.3- Normalité d'un ensemble flou.....	29
III .2.8.4- Convexité et concavité d'un ensemble flou	29
III .2.8.5- Décomposition d'un ensemble flou	30

III .2.9- Les opérations sur les ensembles flous	30
III .2.9.1- Le complément d'un ensemble flou	30
III .2.9.2- L'intersection	31
III .2.9.3- L'union	31
III .2.9.4- Egalité.....	32
III .2.9.5- Le produit algébrique	32
III .2.9.6- La somme algébrique	32
III .2.9.7- L'inclusion	32
III .2.10- Indice d'évaluation et de comparaison d'ensembles flous	32
III .2.11- Relations floues	33
III .2.11.1- Définition de relation de similarité et relation de préférence	34
III .2.11.2- Différentes opérations sur les relations floues	34
III .2.11.3- Propriétés des relations floues	34
III .2.12- Conclusion.....	35
- Chapitre IV La théorie des possibilités.....	36
IV . 1- Introduction	37
IV . 2- La modélisation de l'incertitude	38
IV . 2.1- La probabilité	38
IV . 2.2- Mesure de confiance	38
IV . 2.3- La plausibilité et la crédibilité.....	39
IV . 2.4- Mesure de Possibilité	39
IV . 2.5- Mesure de nécessité	41
IV . 2.6- Comparaison entre la mesure de possibilité et la mesure de nécessité	42
IV . 2.7- Mesure d'information.....	42
IV . 3- Sous ensemble flou et le concept de possibilité	42
IV . 4- Distribution de possibilité conditionnelle.....	43
IV . 5- Distribution de possibilité conjointe	44
IV . 6- Principe de minimum de spécificité	44
IV . 7- La Possibilité et la probabilité	44
IV . 8- Conclusion	45
- Chapitre V La modélisation par les systèmes d'inférence flous.....	46
V . 1- Introduction	47
V . 2- Notion de base	48
V . 3- Les différents type de modèles flous	52
V . 3.1- Les modèles linguistiques	52
V . 3.2- Les modèles relationnels.....	54
V . 3.3- Les modèle de Takagi-Sugeno-Kang (TSK)	54
V . 3.3.1- Méthode de calcul	55
V . 3.3.2- Avantages et inconvénients des modèles TSK	56
V . 4- Modélisation des systèmes dynamiques	57
V . 4.1- Application à la prévision des séries temporelles	59
V . 4.2- Méthodes d'apprentissage liées aux modèles de TSK pour la prévision des séries temporelles.....	61
V . 5- Conclusion	65

- Chapitre VI Préviation des pluies par la logique floue	67
VI . 1- Introduction	68
VI . 2- Mécanisme de raisonnement	69
VI . 3- Différentes phases de la modélisation	71
VI . 3.1 - Fuzzification	71
VI . 3.2 - Inférence.....	71
VI . 3.4 - Défuzzification	72
VI . 3.5 - Tests de Performance.....	72
VI . 4 - Organigramme.....	74
VI . 5 - Interprétation des résultats	75
VI . 6 - Conclusion	87
- Conclusion Général.....	90

INTRODUCTION



Les phénomènes météorologiques et climatiques constituent un danger potentiel pour la vie et l'environnement. Ils le sont particulièrement lorsqu'ils revêtent une intensité extrême, inhabituelle ou un caractère persistant (orage, inondation ou sécheresse).

L'étude de l'information météorologique devient donc une nécessité qui permet aux services concernés de mieux contribuer à assurer la sécurité, des biens et des personnes et de mieux maîtriser les apports qui engendrent une meilleure planification du développement.

L'Algérie est un pays semi-aride, il connaît ces dernières années, des variations climatiques distinctes, et les ressources en eau sont faibles, irrégulières, et localisées dans la bande côtière. Cela affecte négativement la ressource hydrique sur qui la demande croît sans cesse..

La connaissance de ces conditions climatiques jouent un rôle important dans les études de planification de la ressource en eau. Elle permet de mieux maîtriser les apports non contrôlés, qui assurent le bon fonctionnement du système de prévision et de gestion des ressources hydriques à court, moyen et long terme.

Afin d'augmenter cette connaissance, la prévision de la pluviométrie doit sans cesse s'améliorer pour essayer d'éviter tout aléas climatique comportant des risques. L'utilisation de nouvelles techniques comme la logique floue semblent apporté une méthodologie flexible, souple et adaptable à des problèmes réels. Pour cela, nous l'avons utilisé pour la prévision des précipitations au pas de temps mensuel.

L'ensembles du travail est regroupe comme suit :

- *Après une introduction, le chapitre I est consacré à la recherche bibliographique sur les travaux réalisés dans le domaine de la prévision et les différentes techniques utilisées.*
- *La collecte et la critique des données sont étudiées dans le chapitre deux.*
- *Par la suite, dans le chapitre III il a été question de notions de base sur la théorie des sous-ensembles flous. Le point le plus important est la définition et la représentation de la fonction d'appartenance.*

-
- *Quatrième chapitre : il sera consacré à la théorie des possibilités, qui est fondée à partir des sous-ensembles flous. Cette théorie nous permet la représentation et le traitement unifié d'informations incertaines et/ ou imprécises.*
 - *Le cinquième chapitre : présentation de la typologie des systèmes l'inférence flous, description des différents types de modèles flous. On s'est basé sur la représentation et l'apprentissage du modèle de Takagi-Sugeno-Kang et son application à la prévision.*
 - *Enfin, dans le dernier chapitre des applications à la prévision des précipitations avec de différents formalismes..*

Chapitre I

Recherche Bibliographique

RECHERCHE BIBLIOGRAPHIQUE

Aujourd'hui, la sécheresse exerce sur les pays non préparés à cette aléa climatique un impact économique et sanitaire qui risque de bouleverser la vie dans certaines régions du monde, et devient l'une des catastrophes climatiques majeures qui affecte aussi l'environnement.

Le rapport de la banque mondiale montre [Conférence, journée mondiale de l'eau, 2002] :

- 40% de la population mondiale vit dans un état de pénurie d'eau chronique.*
- Plus de 80 pays manqueraient d'eau, et plus d'un milliard d'homme n'aurait pas accès à l'eau potable.*
- Au rythme actuel de la croissance démographique, la demande en eau doublerait tous les 20 ans.*
- la situation de la Méditerranée serait grave : la moitié des 20 pays de cette zone aurait une consommation dépassant leurs ressources réelle au cours du siècle prochain.*

L'Algérie aussi a connu, au cours des vingt dernières années, une sécheresse intense et persistante, cette sécheresse, caractérisée par un déficit pluviométrique important, a touché l'ensemble du territoire et a sévi particulièrement dans la région Ouest du pays.

L'augmentation des risques naturels n'est pas seulement dû à la sécheresse , mais aussi à une amplification des Précipitations extrêmes et des crues qui engendrent des dégâts très importants mettant en danger la vie humaine (catastrophe de BAB EL-Oued le 10 Novembre 2001).

Alors si on peut comprendre ces mécanismes physiques, on fera un pas de géant dans la prévision afin de protéger les vies humaines, à optimiser l'utilisation de la ressource et la gestion d'ouvrages hydrauliques.

1- *K-PPV* methode:

le (*k-PPV*) plus proche voisin (*K-nearest-neighbour*) en anglais a ses origine des statistiques non-paramétriques. La définition d'un voisinage joue un rôle important en géométrie algorithmique pour appliquer des techniques d'interpolation en modélisation de données. En statistiques (approches descriptives), le voisinage permet d'associer à un point de l'espace d'entrée, une information de sortie stockée par des représentants (exemple :dans les Systèmes d'Information Géographique , les Réseaux de Neurones et en logique Floue) ou bien de déplacer des représentants (appelés vecteurs «codebook») afin de mieux modéliser la distribution des données à priori inconnue (exemple :en Quantification Vectorielle ou pour la représentation et le regroupement de données («data clustering»). [Michaël Aupetit]

Deux voisinages particuliers, les *k-Plus-Proches-Voisins* (*k-PPV*) et les *Voisins Naturels* (*VN*), nous ont conduits à introduire la notion de voisinage γ -Observable. Ces voisinages sont basés sur une structure fondamentale en géométrie algorithmique connue sous le nom de «diagramme de Voronoï» .

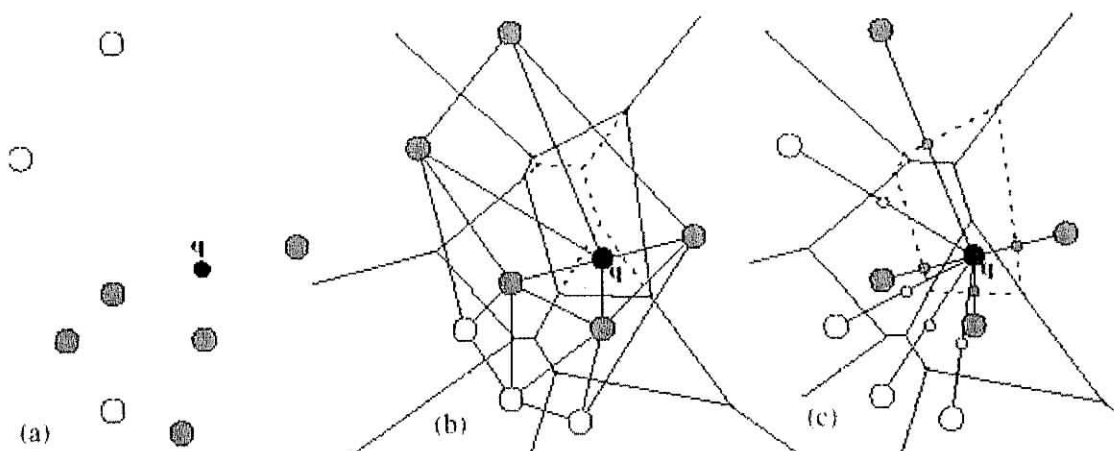


Figure 1.1: La distance Euclidienne est utilisée pour définir (a) les 5-Plus-Proches-Voisins (gris) d'un point q (noir), (b) les Voisins Naturels (gris) de q et (c) les Voisins 0.5-Observable (grands disques gris) de q , « diagramme de Voronoï »

La prévision des séries temporelles est basée sur les approximations locale, utilisant seulement les observations proches. Pour prévoir à un instant (t) la méthode est de constituer un ensemble de données sous forme de vecteur tel que :

$$\bar{X}^d(t) = (x_t, \dots, x_{t-d+1})$$

où $\bar{X}^d(t)$ représente le vecteur de la prévision future. Le vecteur résume l'ensemble des données des observations passées supposé contenir le plus d'informations. La méthode de prévision est faite de façon que des distributions de probabilité de variables présent au hasard conditionnées sous la forme $(x_{t+1} / x_t, \dots, x_{t-d+1})$ ont la même signification que pour des variables présent au hasard formulées pour des observations passées $(x_{t+1} / \bar{X}^d(t))$.

C'est prouvé que si $\bar{X}^d(t)$ ne satisfait pas les propriétés historiques (regroupe le plus d'informations), la prévision de K-NN sera optimale et asymptotique pour toutes les prévisions définies dans le vecteur de valeurs future $\bar{X}^d(t)$.

$$\hat{X}_{t+1} = E \left[x_{t+1} \mid \bar{x}^d(t) \right]$$

alors la prévision est obtenue par une moyenne de l'évolution temporelle du plus proches voisins :

$$\hat{x}_{t+1} = \frac{1}{K} \sum_{j=1}^K x_{tj+1}$$

la généralisation pour toute le série chronologique L est donnée par :

$$\hat{x}_{t+L} = \frac{1}{K} \sum_{j=1}^K x_{tj+L}$$

La prévision des précipitations futures est obtenue par une moyenne des précipitations passées, qui sont les K les plus proches voisins historiquement. Il est important de noter que l'approche K-NN ne requiert pas la sélection des classes du modèle et l'estimation de ses paramètres, alors l'identification de la relation entrées/ sorties ne sera pas son objectif.

2- ARMA :

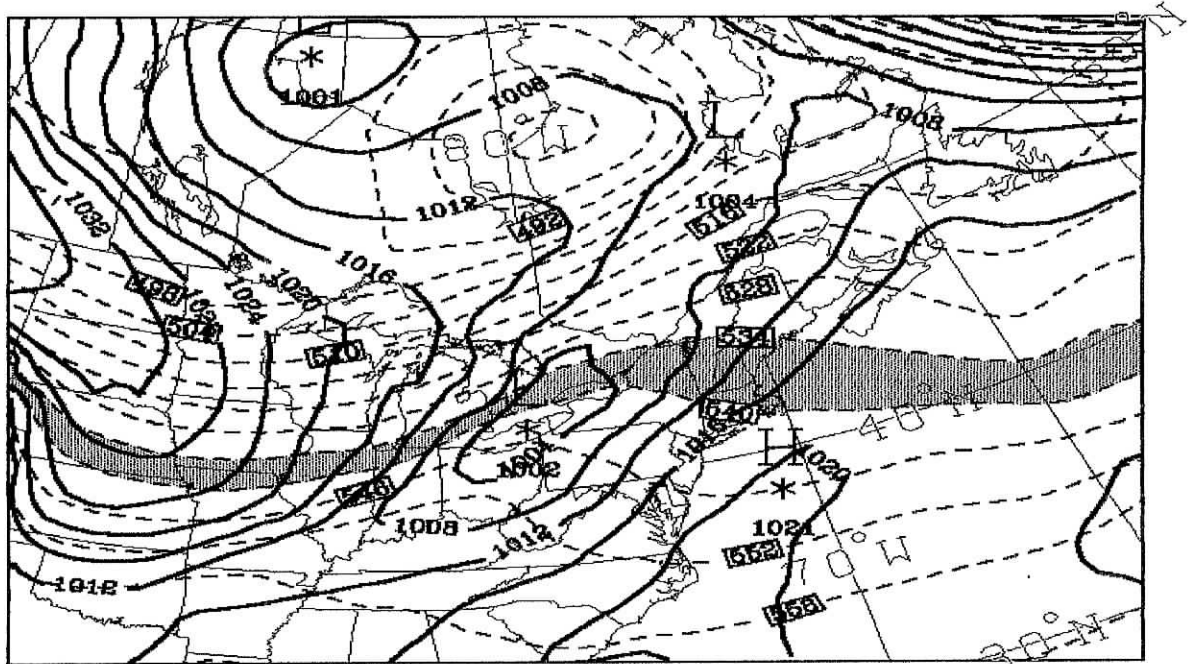
L'application du processus ARMA pour modéliser les précipitations à court terme est considéré et proposé par BRATH (1988) et BURLANDO (1993).

L'utilisation du modèle ARMA pour la prévision demande des données stationnaires ce qui n'est pas toujours le cas pour les observations de précipitations horaires qui ont des propriétés variables selon les saisons. Toutefois, le nombre limité d'observations d'une période prévoit ce risque. Le modèle est paramétré par p , d , q de tel façon que :

$$X_t \sim \text{ARIMA}(p, d, q) :$$

4- Prévisions Météorologiques en Temps Réel :

Des Prévisions en Temps Réel sont faites avec des modèles tel que Le Modèle Communautaire Compressible (MC2)



(V 9 MAR 3 0Z PN OHR D1 \ FCST)

Le programme de prévisions environnementales vise la conception d'outils d'aide aux prévisions de variations météorologiques et de catastrophes environnementales, l'analyse de leurs répercussions sur les infrastructures et les activités économiques, et la prise de décisions concernant la gestion de ces infrastructures et de ces activités. L'utilisation du modèle MC2 (mésoséchelle compressible communautaire) pour la modélisation atmosphérique à haute résolution.

Le CERCA (centre de recherche en calcul appliqué) a amélioré la qualité des prévisions pour les précipitations abondantes (comparaison avec les observations de stations) grâce à l'implantation dans MC2. [Le CERCA, Rapport annuel 1998-1999 - Aspects scientifiques].

Chapitre II

Collecte et critique des données

COLLECTE ET CRITIQUE DES DONNEES

II .1- Introduction :

L'analyse des données disponibles est généralement le premier pas à franchir dans l'étude d'un phénomène hydrométéorologique (précipitation, température, débit...).

Les données hydrologiques constituent la plupart du temps le seul outil dont dispose l'hydrologue pour étudier et comprendre le comportement d'un phénomène météorologique , ou prendre des décisions relatives à un objet particulier.

Il est donc de première ordre, de bien maîtriser les techniques nécessaires à l'analyse et à la critique des données, afin de tirer de l'échantillon le maximum d'informations , de garantir la meilleure qualité des données et de quantifier le risque que comporte une généralisation à partir de données partielles.

Il s'agit d'abord de relever les valeurs aberrantes, de détecter et de combler les lacune du fichier recueilli et aussi de vérifier l'indépendance et l'homogénéité de la série étudiée nécessaire à toute étude statistique.

II .2- Inventaire des données disponibles :

Les données utilisées nous ont été aimablement fournies par l'Office National de la Météorologie (O.N.M). Elle comporte une série des précipitations de la station d'Alger (DAR EL BAIDA), s'étalant sur une période de 66 ans (1936-2001), au pas de temps mensuel.

Le tableau II .1 regroupe les caractéristiques de la station considérée :

Tableau II. 1 : Caractéristiques de la station

<i>Station</i>	<i>Numéro de la station</i>	<i>Code</i>	<i>Latitude (Nord)</i>	<i>Longitude (E/W)</i>	<i>Altitude (mètre)</i>
<i>Alger</i>	6	390	36°43	03°15 E	24

II .3- Critique des données :

II .3.1- Introduction :

Les données pluviométriques proviennent d'une mesure qui a été recueillie soit par :

- ✓ **Le pluviomètre:** instrument de base de la mesure des précipitations liquides ou solides. Il indique la quantité d'eau totale précipitée et recueillie à l'intérieur d'une surface calibrée dans un intervalle de temps séparant deux relevés.
- ✓ **Le pluviographe :** instrument captant la précipitation de la même manière que le pluviomètre mais avec un dispositif permettant de connaître, outre la hauteur d'eau totale, leur répartition dans le temps autrement dit les intensités.

Celles ci peuvent comporter deux types d'erreurs ; les erreurs systématiques et les erreurs accidentelles.

a) Les erreurs systématiques : sont dues à un mauvais réglage de l'appareil de mesure, à une procédure de mesure incorrecte, ou à l'installation inadéquate d'appareillage de mesure, donc à des causes produisant des effets constants, ou légèrement variables dans le temps.

b) Les erreurs accidentelles : surviennent inopinément et ne peuvent être ni prévues ni évitées. Elle sont dues à des erreurs de manipulation ou à une défectuosité imprévue de l'appareil, donc à des causes subites qui introduisent des déviations sensibles au sein des séries de données.

En effet, ces erreurs de manipulation peuvent altérer les séries, si certaines précautions ne sont pas prises durant cette phase de transcription de données, ces dernières peuvent être rendues inutilisables à cause de cumul d'irrégularités.

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

- ❖ Détecter et corriger les erreurs ;
- ❖ Vérifier la continuité des données dans le temps et dans l'espace ;
- ❖ Combler les données manquantes dans les séries d'observations.

II. 3.2- Méthodes de comblement de lacunes

Dans la pratique courante, de nombreuses causes contribuent à entraîner l'absence de certaines informations : absence de mesure, défaut de transmission, erreur de stockage, manipulation erronée, etc. Cela fait que dans les fichiers disponibles certaines données soient manquantes.

Ce paragraphe décrit les méthodes de comblement de lacunes pour les valeurs mensuelles. Le comblement des données manquantes peut se faire soit par :

- Le calcul de la moyenne mensuelle interannuelle ;
- La régression linéaire ;
- La simulation.

II .3.2.1- Calcul de la moyenne mensuelle interannuelle

Les données mensuelles se présentent sous la forme d'une matrice à n lignes représentant les années et p colonnes représentant les mois, malheureusement cette matrice peut comporter des vides correspondant aux lacunes d'observations.

Lorsqu'un fichier d'observation comporte moins de trois valeurs manquantes le comblement de lacunes par la méthode de la moyenne mensuelle interannuelle est approprié.

$$[X]= \begin{pmatrix} X(1.1) & \text{inconnu} & & X(1,j) & \dots & X(1.p) \\ X(2.1) & X(2.2) & \dots & X(2.j) & \dots & \text{inconnu} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \text{inconnu} & \dots & \dots & \dots & \dots & \text{inconnu} \\ X(i.1) & X(i.2) & \dots & X(i.j) & \dots & X(i.p) \\ \dots & \text{inconnu} & \dots & \text{inconnu} & \dots & \dots \\ X(n.1) & X(n.2) & \dots & X(n.j) & \dots & X(n.p) \end{pmatrix}$$

le nombre d'observations est de $(n*p)$, le comblement des valeurs manquantes se fait en remplaçant la valeur manquante par la moyenne mensuelle interannuelle, telle que :

$$M_j = 1/q_j \sum_{i=1}^p X(i, j)$$

avec : $q_j = n - (\text{nombre d'inconnues de la colonne } j)$

II .3.2.2- Méthode de régression linéaire simple

soient X et Y deux échantillons, chacun avec n observations $x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_n$, correspondant à deux mesures simultanées sur deux stations différentes. Si ces deux stations sont assez rapprochées, on peut espérer que les observations soient similaires, et la connaissance d'une des variables permet alors, dans une certaine mesure, la prédiction quantitative de l'autre.

Théoriquement, on appelle la régression linéaire sur les données expérimentales, la méthode des moindres carrés reprenant sur l'échantillon la propriété que $E(y/x) = \alpha + \beta x$ est la meilleure approximation de y par x en moyenne quadratique. On cherche donc à ajuster au nuage de points (x_i, y_i) une droite d'équation :

$$y^* = ax + b$$

Dans la plupart des cas, les observations ne sont pas aligner, ce qui fait que la droite de régression n'est pas bien définie ; il faut donc établir certains critères pour obtenir la droite de régression optimale tel que : la somme des carrés des distances de chacun des points à la droite soit la plus petite possible qui peu être formuler comme suit : $\sum c_i^2 = \sum (y_i - y_i^*)^2$ soit minimale (fig.II.1).

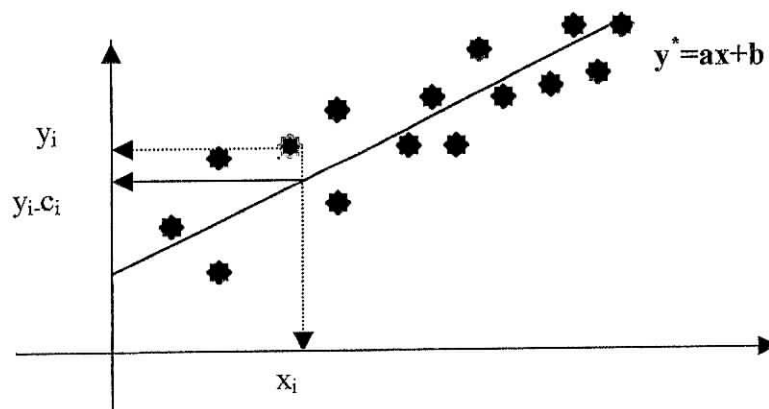


Figure.II.1 : régression linéaire simple

La méthode appliquée pour la détermination des coefficients a , b de la droite est celle des moindres carrées :

$$\sum (y_i - y_i^*)^2 = \sum (y_i - ax_i - b)^2 = F(a, b)$$

le minimum est atteint pour $\frac{\partial F}{\partial a} = \frac{\partial F}{\partial b} = 0$ ceci donne les deux équations :

$$\frac{\partial F}{\partial a} = 0 \Rightarrow -2 \sum (y_i - b - ax_i) = 0 \Rightarrow \bar{y} = a\bar{x} + b \Rightarrow b = \bar{y} - a\bar{x}$$

$$\frac{\partial F}{\partial b} = 0 \Rightarrow \sum (x_i - \bar{x})(y_i - \bar{y}) = a \sum (x_i - \bar{x})(x_i - \bar{x})$$

$$\Rightarrow a = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} = r \frac{\sigma_y}{\sigma_x}$$

d'où la solution sera sous la forme :

$$y^* = \bar{y} + r \frac{\sigma_y}{\sigma_x} (x - \bar{x})$$

Remarque :

Pour la dérivé de F en fonction de a et b , on obtient le minimum et pas le maximum, car F est un trinôme du second degré dont le premier coefficient est positif.

II .3.2.3- Simulation à partir de la fonction de répartition

pour simuler des grandeurs mesurant des phénomènes naturels (tel que la pluie ou le débit), il ne faut pas perdre de vue que cette notion est basée essentiellement sur le concept du hasard.

C'est une méthode de tirage d'un échantillon artificiel de (n) valeurs d'une variable aléatoire continue. Elle s'applique lorsque F^{-1} a une forme analytique simple. Soit $F(x)$ la fonction de répartition de x . la variable $Y=F(x)$ est distribuée sur l'intervalle $[0,1]$, en effet :

$$g(y) = \frac{f[F^{-1}(y)]}{f[F^{-1}(y)]}$$

Si on tire (n) nombre uniformément repartis entre $[0,1]$: $r_1, r_2, r_3, \dots, r_n$, alors l'échantillon cherché ($x_1, x_2, x_3, \dots, x_n$) sera déterminé par $x_i = F^{-1}(r_i)$; cette méthode est dite « de l'anamorphose » (fig.II.2). On dit que l'on a simulé la variable X .

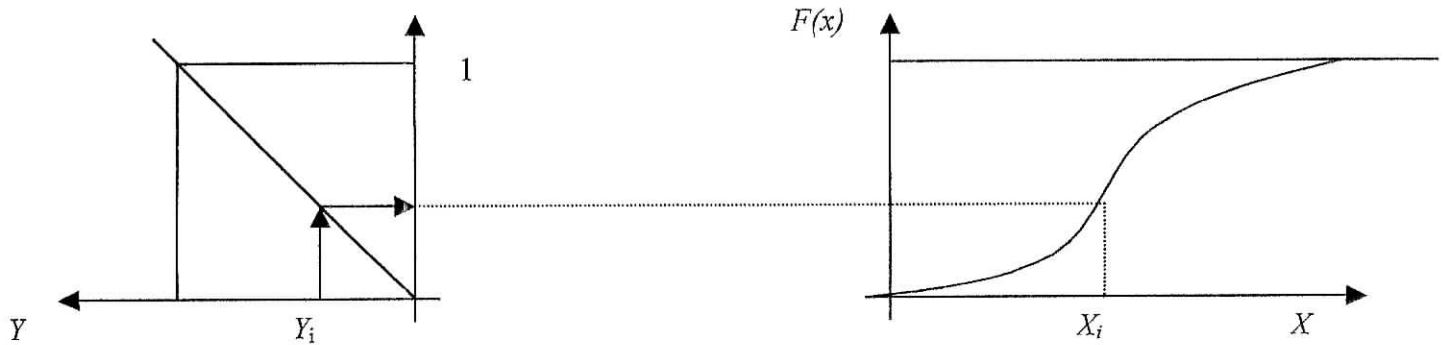


Figure.II.2. Simulation à partir de la fonction de répartition.

II .3.3- Comblement de lacune

Le comblement de lacune n'a pas été nécessaire car le fichiers des données fournis par l'ONM ne comporte pas de données manquantes.

II .3.4- Détection des valeurs singulières

Il arrive fréquemment que des erreurs se glissent dans les données de base et en dénaturent complètement le contenu. Cela peut se produire au cours de la mesure, de la transmission, du codage, du stockage ou du décodage des observations. Il est important avant tout traitement de détecter les erreurs grossières de ce genre 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.

Un test de ce genre a été élaboré par GRUBBS et BECK pour le contrôle initial des erreurs grossiers dans les valeurs observées.

II .3.4.a- Présentation du test de GRUBBS et BECK

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

$$X_h = EXP(\bar{X} + K_n * S)$$

$$X_l = EXP(\bar{X} - K_n * S)$$

Tel que :

\bar{X} et S : sont respectivement la moyenne et l'écart type des logarithmes naturels des éléments constituant l'échantillon.

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

$$K_n = -3.62201 + 6.28446*n^{1/4} - 2.49835*n^{1/2} + 0.491436*n^{3/4} - 0.037911*n.$$

K_n : est la valeur statistique du test de Grubbs et Beck, tabulée pour différente taille d'échantillon de niveaux de signification

Avec n : la taille de l'échantillon.

En appliquant ce test, toutes les valeurs de l'échantillon supérieures à X_h ou inférieures à X_l sont considérées comme valeurs singulières.

II .3.4.b- Analyse des résultats

le tableau II .2 représente les résultats obtenues par l'application du test de Grubbs et Beck.

Tableau II .2 : résultats du test Grubbs et Beck

	n	M (mm)	Σ (mm)	K_h	X_h (mm)	X_l (mm)
Echelle mensuelle	792	1.26	0.93	2.75	6717.7	0.273
Echelle annuelle	66	6.44	0.29	2.87	1439.85	272.51

L'application de ce test a détecté les valeurs nulles comme singulieres, mais nous ne pouvons rejeter ces points car ils correspondent aux mois secs.

II .4- Traitement statistique des données

Après le comblement des lacunes des données pluviométriques disponibles, il est nécessaire de vérifier l'indépendance, l'homogénéité de cette série pluviométrique.

II .4.1- Test d'homogénéité

Avant de passer à l'application des méthodes statistiques sur les données recueillies, il est important de vérifier si elle peuvent être considéré comme extraites d'une même population . c'est ce qu'on appelle le problème d'homogénéité des séries statistiques, ou encore le problème des échantillons.

Il existe plusieurs tests pour vérifier l'homogénéité des séries de données. On utilise le test de Wilcoxon-Mann-Whitney. Ce test est appelé aussi le test de variation de la moyenne. Il utilise les numéros des observations dans la série variationnelle obtenue en ordonnant l'échantillon global de taille N .

Pour ce faire, on divise notre série en deux sous échantillons $\{x_i\}$ et $\{y_j\}$ de taille respectivement n_1 et n_2 , où $N = n_1 + n_2$ (avec $n_1 \leq n_2$), puis on classe l'échantillon global de taille N .

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

Si l'échantillon X est issu de la même population Y , l'échantillon XUY en est également issu. On construit à l'aide des échantillons réunis, la série de variation qui sera composée de n_1+n_2 valeurs, disposées dans l'ordre de croissance.

Par la suite, on s'intéresse au rang de chacun des éléments des deux échantillons dans cette série.

Si la même valeur se répète plusieurs fois, on lui associe le rang moyen correspondant. On calcule la somme W_x des rangs des éléments du premier échantillon dans la série commune :

$$W_x = \sum |\text{Rang}|.$$

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

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

$$W_{\max} = (N_1 + N_2 + 1) N_1 - W_{\min}$$

Avec : $U_{1-\alpha/2}$: représente la valeur de la variable centrée réduite de la loi normale correspondant au seuil de $(1 - \alpha / 2)$.

Si $W_{\min} < W_x < W_{\max}$, l'hypothèse de l'homogénéité est acceptée. (voir le tableau II .3).

II .4.2- Test d'indépendance

L'application des méthodes statistiques repose sur la validité de certaines hypothèses de base relatives à la constitution de l'échantillon observé. En effet le fichier recueilli doit être constitué par une série d'observations indépendantes formant ce que l'on appelle un échantillon extrait au hasard de la population statistique. En d'autres termes l'échantillon $X=(x_1, x_2, \dots, x_j, \dots, x_N)$ est dit « au hasard » cette hypothèse H_0 d'indépendance des observations.

Dans notre étude, nous avons appliqué le test de Wald-Wolfowitz (couramment utilisé) sur les valeurs mensuelles.

Ce test est basé sur le critère qu'aucune observation appartenant à l'échantillon ne devrait influencer l'observation suivante, Bobbée(1991).

La statistique de test est :

$$R = \sum_{i=1}^N X_i X_{i+1}$$

Si l'hypothèse H_0 selon laquelle l'échantillon est extrait au hasard est vraie, la statistique suit une distribution normale :

$$N \left(\frac{m_1^2 - m_2^2}{N-1}, \frac{m_1^2 - m_4}{N-1} + \frac{m_1^4 - 4 m_1^2 m_2 + 4 m_1 m_3 + m_2^2 - 2 m_4}{(N-1) \cdot (N-2)} - \left[\frac{m_1^2 - m_2^2}{N-1} \right]^2 \right)$$

avec : $m_r = \sum_{i=1}^N x_i^r = N * S_r$ et S_r : moment d'ordre r de l'échantillon.

La valeur statistique $U = \frac{R - \bar{R}}{\sqrt{\text{VAR}(R)}}$ suit une distribution normale standardisée $N(0,1)$, et qui

est utilisée pour tester à un niveau de signification, l'hypothèse d'indépendance en comparant la valeur absolue de U avec la valeur normale standard $U_{\alpha/2}$ correspondant à la probabilité de dépassement de $\alpha/2$.

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

II .4.3 - Analyse des résultats du test

Le résultats du test de **Wilcoxon** illustrés dans le tableau II .3 et appliqué aux différents seuils, montre que les observation pluviométriques sont homogènes.

L'hypothèse H_0 d'indépendance pour le pas de temps mensuel est rejetée pour les différents seuil utilisés (tableau II .3).

Tableau II .3 : Résultats des tests statistiques.

Test	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 10\%$
Homogénéité	1	1	1
Independence	0	0	0

1: H_0 acceptée

0: H_0 rejetée

Ces résultats peuvent être expliqués par le fait que l'années comporte des mois humides et secs dont la corrélation entre les mois de la même saison peut être importante.

Chapitre III

La logique Floue

LA LOGIQUE FLOUE

III .1 - Historique de la logique floue

III .1.1- Introduction :

l'intérêt pour la formalisation de l'incertitude remonte essentiellement au XVII^{ème} siècle, avec les travaux, entre autres, de PASCAL et BERNOULLI, même si on trouve des travaux plus anciens abordant ce concept, par exemple chez ARISTOTE. Au départ, l'incertitude est envisagée de deux points de vue : objectif (les jeux de hasard) et subjectif (la fiabilité des témoignages). Même pour BERNOULLI, il ne va pas de soi que les deux types d'incertains, le hasard et la croyance, relèvent de la même approche mathématique [A.Kaufmann].

Les probabilités additives (qu'on appelle alors les chances) semblent naturelles pour capturer les jeux de hasard ; mais BERNOULLI, et quelques épigones plus tardifs, proposeront des modèles non-additifs de la croyance, bien oubliés depuis. En effet, avec l'essor des sciences physiques au XVIII^{ème} siècle, les probabilités, souvent assimilées à des fréquences, deviennent additives et objectives, jusqu'au XX^{ème} siècle.

Dans les années 30, avec les travaux de DE FINETTI et RAMSEY, on assiste à une résurgence de la probabilité subjective, dans le but de formaliser la notion de choix rationnel dans l'incertain. Mais les degrés de croyance ainsi étudiés demeurent additif, en vertu d'une interprétation en termes de mises associées à des paris. Dans le même temps, la logique formelle semble mal adaptée à formaliser le raisonnement humain, notamment à cause de la présence de nombreux termes mal définis dans le langage naturel. C'est au philosophe MAX BLACK que revient l'idée de formalisation du sens des prédicats vagues à l'aide de fonctions numériques d'appartenance.

Parallèlement, et indépendamment, on observe l'éclosion de logique multi-valuées où les degrés de vérité peuvent être différents de « vrai ou faux ». Le statut des valeurs de vérité intermédiaires reste néanmoins controversé car on oscille entre l'idée d' "à moitié vrai" et d' "inconnu".

C'est à LOTFI ZADEH [A.Kaufmann], spécialiste renommé d'automatique et théorie des systèmes, à l'université de BERKELEY, que revient le mérite de faire la jonction en 1965 entre les logiques multivaluées des années 1930 et les travaux de BLACK sur les concepts vagues, grâce à la

*d'ensemble flou et aux logiques associées (extensions de la logique booléenne). C'est aussi ZADEH qui a fait en 1978, le lien entre les ensembles flous et la modélisation non-probabiliste de l'incertitude, grâce à la théorie des possibilités, dans laquelle les fonctions d'appartenances d'ensembles flous servent à représenter les idées de valeur incomplètement spécifiée, mal connue, et de contraintes flexibles. Ensemble flou et théorie des possibilités sont à la base d'une méthodologie de représentation et d'utilisation de connaissances imprécises, vagues et incertaines, que ZADEH appelle **théorie du raisonnement approché** et qui est plus connue sous le vocable de « **logique floue** »*

III .1.2- Le concept d'ensemble flou :

Ce concept permet de considérer des classes d'objets dont les frontières ne sont pas clairement déterminées. Par l'introduction d'une fonction caractéristique (fonction d'appartenance des objets à la classe) prenant des valeurs courantes entre 0 et 1, contrairement aux ensembles "booléens", dont la fonction caractéristique ne prend que deux valeurs possibles 0 et 1.

Ces ensembles flous ont le grand avantage de constituer une représentation mathématique de labels linguistiques largement utilisés dans l'expression des connaissances expertes. qualitatives et manipulées dans le raisonnement approché qui sera fait à partir de cette connaissance.

Ils apparaissent donc comme un moyen de réaliser l'interface entre l'information numérique (quantitative) et l'information symbolique (linguistique, qualitative).

Ces concepts nouveaux, proposés par Zadeh, même s'ils ont induit, assez rapidement, de nombreux travaux de recherche, ont suscité aussi beaucoup d'opposition, de réactions négatives, et soulevé beaucoup de scepticisme sur leur portée, attitude non totalement disparue près de 30 ans après.

En effet, après une "traversée du désert", malgré les nombreux domaines d'application, sous la poussée économique du Japon dès le milieu des années 80 et l'évolution technologique favorisant la mise en oeuvre de ces idées, tous les pays développés s'intéressent maintenant à la logique floue.

III .1.3- Objectifs :

La logique floue interpelle aujourd'hui non seulement de nombreux domaines de la science et de la technologie, mais aussi des disciplines plus spéculatives telles que la philosophie des sciences, même si elle souffre d'un effet de mode qui, bien que passager, risque malgré tout de lui être préjudi-

ciable à terme, en suscitant des attentes auxquelles elle ne peut répondre (comme ce fut le cas avec les systèmes experts).

III .1.4- Autres domaines d'applications :

Il n'y a en fait pas de domaine où l'on ne puisse appliquer les raisonnements propres à la Logique Floue car elle a été conçue pour s'adapter aux techniques de pensée humaine. Tout d'abord il y a les secteurs privilégiés de la commande floue qui sont l'automatisme et la robotique. Dans ces deux secteurs la commande floue a déjà connu un franc succès et est devenu ainsi la source de nombreuses recherches sur la logique floue. Le troisième secteur très important est l'informatique (aussi bien l'intelligence artificielle que les bases de données et la programmation) où l'on doit souvent traiter des informations vagues et imprécises. Ensuite viennent l'engineering, la gestion et la prise de décision. La logique floue est déjà utilisée en médecine pour l'aide au diagnostic. Elle sert beaucoup les mathématiques appliquées. Mais aussi les sciences humaines telles que la sociologie et la psychologie car l'homme est tout ce qu'il a de plus flou et complexe. La logique floue est souvent utilisée en Hydrologie à cause de la complexité des phénomènes et de l'imprécision naturelle des variables observées (pluie, débit, ...)

<i>Domaine d'applications</i>	<i>Exemples</i>
<i>Automatisme</i>	<i>ateliers de production, automates programmables, pilotes d'avions, caméras</i>
<i>Robotique</i>	<i>commande de robots très divers : aide au mouvement, à la détection des formes</i>
<i>Informatique</i>	<i>systèmes experts, base de données relationnelles, langage de programmation flou</i>
<i>Engineering, Gestion, Prise de décision</i>	<i>contrôle de circulation(aérienne ou automobile), télécommunications, théorie des jeux, gestion de personnel d'investissement</i>

<i>Environnement</i>	<i>météorologie(prévisions, évaluation des variations climatiques), évaluation des risques sismiques</i>
<i>Médecine</i>	<i>système expert d'aide aux diagnostics comme par exemple le MYCIN en 1974</i>
<i>Mathématiques appliqués</i>	<i>statistique, recherche opérationnelle</i>
<i>Sciences humaines</i>	<i>sociologie, psychologie</i>

III .1.5- Conclusion :

La logique floue reste, même si elle est déjà beaucoup utilisée, une technique nouvelle dont certains aspects sont encore inexploités. Elle est le sujet de nombreuses recherches et pourrait dans un futur proche devenir une technique incontournable pour l'amélioration de la communication Homme-Machine, par la suite nous verrons l'application de la logique floue pour la prévision hydrologique, qui sera notre sujet à nous, mais préalablement , nous allons donner un aperçu sur la théorie des ensembles flous ainsi que la théorie des possibilités qui sera utilisé par la suite dans la modélisation basée sur les systèmes d'inférences floues .

III .2- LA THEORIE DES SOUS ENSEMBLES FLOUS

III .2.1- INTRODUCTION :

L'importante motivation de la théorie des sous ensembles flous est la conception d'un outil formel qui gère les données vagues et imprécises des connaissances humains exprimées en langage naturel.

En 1971, ZADEH introduit la notion de la sémantique quantitative floue, expliquant comment les termes linguistiques peuvent être exprimés comme des sous ensembles flous de l'univers de discours, et comment des combinaison logique de ces termes linguistique peuvent donner naissance à des termes numériques suite aux opérations sur les ensembles flous.

Ainsi d'après [BON84], la théorie des sous ensembles flous permet la modélisation d'ensembles dont les frontières ne sont pas clairement défini en prenant compte leur imprécision.

III .2.2- Imprécision et incertitude :

Dans ce paragraphe, nous allons faire la distinction entre l'incertitude et l'imprécision , une proposition est incertaine si sa valeur de vérité (vraie ou fausse) ne peut être établi par des informations disponibles. De même une proposition est imprécise si sa valeur n'est pas suffisamment déterminée.

Notons , qu'une proposition précise peut être incertaine et une proposition qui est complètement certaine peut être imprécise.

III .2.3- Définition d'un ensemble flou :

Un ensemble flou est un ensemble dont la frontière n'est pas nette: tel que l'on passe graduellement mais non brusquement de l'appartenance à la non appartenance. Soit $U=\{u\}$ une collection d'objets (points) notés généralement par u , alors un ensemble flou F dans U est un ensemble de couple tel que:

$$F = \{u, \mu_F(u)\}, \forall u \in U.$$

où $\mu_F(u)$ détermine le degré d'appartenance de l'élément u dans l'ensemble F .

III .2.4- Définition d'une Fonction d'appartenance :

Un sous ensemble F d'un référentiel U est caractérisé par une fonction notée μ_F de U dans l'intervalle unité $[0,1]$ tel que:

$$\begin{aligned} \mu_F : U &\rightarrow [0,1] \\ u &\mapsto \mu_F(u) \end{aligned}$$

où $\mu_F(u)$ est le degré d'appartenance de u dans le sous ensemble flou F .

Dans le cas où F est un ensemble ordinaire (précis, imprécis-non-flou), les valeurs de la fonction d'appartenance μ_F sont données par:

$$\mu_F(u) = \begin{cases} 1 & \text{si } u \in F \\ 0 & \text{sinon} \end{cases}$$

avec $\mu_F(u)=1$ correspond à l'appartenance de u dans F .

$\mu_F(u)=0$ correspond à la non appartenance de u dans F .

Exemple

Soit : F : concept de température normal (TN).

U : référentiel de température.

On peut représenter le concept TN par des degrés d'appartenance :

$$TN = 0/10 + 0.3/15 + 0.7/20 + 1/25 + 1/30 + 0/40.$$

où: "+" note l'union, et un élément de cette union EXEMPLE : « 0.3/15 » signifie qu'une température qui est de 15 °c appartient au sous ensemble flou "TN" avec le degré 0.3.

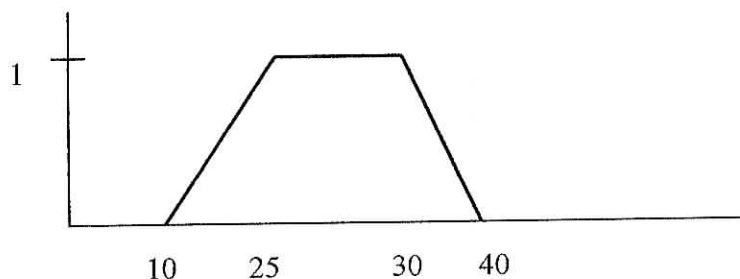


Figure III. 1 : Fonction d'appartenance du concept température normal

III .2.5- Représentation d'une fonction d'appartenance :

La fonction d'appartenance μ_F associée au sous ensemble flou F est défini à l'aide de cinq paramètres (a, α, b, β, q) , la valeur de q dans l'intervalle $[0,1]$ décrit l'incertitude entachant le fait.

On a les conditions suivantes à vérifier:

$$\mu_F(a) = \mu_F(b) = 1 \quad , \quad \mu_F(a - \alpha) = \mu_F(b + \beta) = q$$

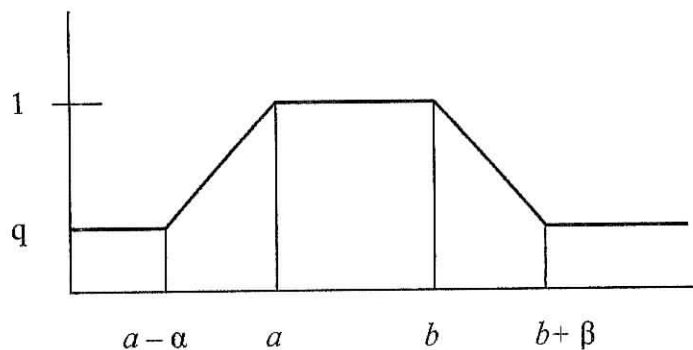


Figure III. 2 : Représentation d'une fonction d'appartenances

les fonctions d'appartenances ont des formes trapézoïdales dans ce cas, et peuvent être gaussiennes ou triangulaire comme on le verra par la suite.

III .2.6- Méthodes de détermination de fonctions d'appartenances :

D'après [Dub84] pour trouver les fonctions d'appartenances, il faut faire la distinction entre les catégories simples qui sont délivrées sur un référentiel linéaire, exemple: "grand", et les autres catégories complexes qui font intervenir plusieurs référentiels simultanément, exemple : "trapu" qui signifie "petit et large d'épaule" et parfois les échelles de référence sont difficiles à déterminer exemple : "beau".

- Catégorie simple :

L'évaluation d'une fonction d'appartenance est un exercice de psychométrie, qui est un ensemble de méthodes de mesures des phénomènes psychologiques (tests, etc).

- Catégorie complexe :

Le référentiel est défini par un produit cartésien d'échelles linéaires, la fonction d'appartenance peut être obtenue par un processus d'agrégation.

III .2.7- Effet du contexte sur les fonctions d'appartenances :

Affirmer que "la mer est calme" n'est pas représentée de la même façon dans le domaine de la prévision météorologique ou celui de la navigation.

Dans le cadre de la médecine "le choix d'un traitement par un médecin est non seulement conditionné par les connaissances, la technologie, médicaments disponibles, mais aussi par la culture du pays en question , c'est à dire par son individualité et sa philosophie nationale", D'où la conclusion que les fonctions d'appartenances dépendent du contexte dans lesquelles elles sont mesurées.

III .2.8- Les concepts de base de la théorie des sous ensembles flous :

III .2.8.1- Support et noyau d'un sous ensemble flou :

*Soit F un sous ensemble flou, défini par sa fonction d'appartenance $\mu_F(u)$, l'ensemble $F_\alpha = \{ u \in U / \mu_F(u) \geq \alpha, \forall \alpha \in [0,1] \}$ est appelé: **coupe de niveau α** de l'ensemble flou F . Un ensemble flou, peut être défini à partir de ces coupes via:*

$$\mu_F(u) = \text{SUP} \{ \alpha \in [0, 1] / u \in F_\alpha \}$$

Le support $\text{supp}(F)$ est une coupe tel que: $\text{supp}(f) = \{ u \in U / \mu_F(u) > 0 \}$ et le noyau d'un ensemble flou est donné par: $\text{noy}(f) = \{ u \in U / \mu_F(u) = 1 \}$. Le noyau correspond à tous les éléments qui appartiennent totalement à l'ensemble, on remarque que: $\text{noy}(F) \subseteq \text{supp}(F)$.

III .2.8.2- Définitions des T-normes, T-conormes et opérateur de négation :

Ces opérations ont été suggéré par [DUB84] , pour calculer les valeurs d'appartenance de l'intersection, l'union et le complément d'un sous ensemble flou.

III .2.8.2.1- Définition d'une T-norme :

Une T-norme $$ est une application de $[0,1] \times [0,1]$ dans $[0,1]$ vérifiant:
 $\forall u, \forall v, \forall w \in [0,1]$: la commutativité, la monotonie, l'associativité, 1 est l'élément neutre, et 0 est l'élément idempotent.*

Exemple

$$u * v = \text{Min}(u, v)$$

$$u * v = u \cdot v$$

$$u * v = \text{Max}(0, u + v - 1)$$

III .2.8.2. 2- Définition d'une T-conorme :

La commutativité, la monotonie, l'associativité, 0 est l'élément neutre, et 1 est l'élément idempotent $\forall u, \forall v, \forall w \in [0, 1]$.

Exemple

$$u * v = \text{Max}(u, v)$$

$$u * v = \text{Min}(1, u + v)$$

III .2.8.2. 3- Définition de la négation :

Un opérateur de négation c est une application de $[0, 1]$ dans $[0, 1]$ vérifiant:

$$c(0) = 1 \text{ et } c(1) = 0.$$

$c(u) \leq c(v)$ si $u \geq v$ monotonie strictement décroissante.

$$C(c(u)) = u \quad c \text{ est involutive}$$

Exemple $c(u) = 1 - u$

III .2.8.3- Normalité d'un ensemble flou :

Un ensemble flou F est dit normal si et seulement si $\text{Sup}_X \mu^F(u) = 1$, et un ensemble est non-normal s'il n'est pas normal. Un ensemble flou non-vide et non-normal, peut être normal en divisant ces éléments par le Sup de cet ensemble.

III .2.8.4- Convexité et concavité d'un ensemble flou :

Soit F un ensemble flou dans $U = \mathbb{R}^n$, alors F est convexe si et seulement si pour tout couple de point u, v dans U , la fonction d'appartenance de F satisfait l'inégalité:

$$\mu_F(\lambda u + (1-\lambda)v) \geq \text{Min}(\mu_F(u), \mu_F(v)) \quad \text{tel que} \quad 0 \leq \lambda \leq 1.$$

Par dualité F est concave s'il n'est pas convexe.

III .2.8.5- Décomposition d'un ensemble flou :

Soit F et G deux sous ensembles flous dans U et V (respectivement), et H un sous ensemble flou dans $W=U \times V$ défini par la fonction d'appartenance $\mu_H(u, v)$, Alors H est décomposable suivant U et V si et seulement si H admet la représentation $H=F \cap G$ ou bien d'une façon équivalente: $\mu_H(u, v) = \mu_F(u) * \mu_G(v)$, avec $\mu_F(u)$ et $\mu_G(v)$ sont les fonctions d'appartenance de F et G (respectivement).

III .2.9- Les opérations sur les ensembles flous :

On fait avec les ensembles flous, ce que l'on fait avec les ensembles classique, ZADEH a ainsi généralisé les opérations d'inclusion, d'égalité, de complément, d'intersection et d'union, sur les ensembles flous. Pour toutes les définitions, considérons qu'on a deux sous ensembles flous F et G , avec leur fonctions d'appartenances respectives μ_F et μ_G .

III .2.9.1- Le complément d'un ensemble flou :

Le complément d'un sous ensemble flou F de fonction d'appartenance μ_F , est un sous ensemble flou F' de fonction d'appartenance $\mu_{F'}$ tel que: $\forall u \in U : \mu_{F'}(u) = 1 - \mu_F(u)$ [Zad65].

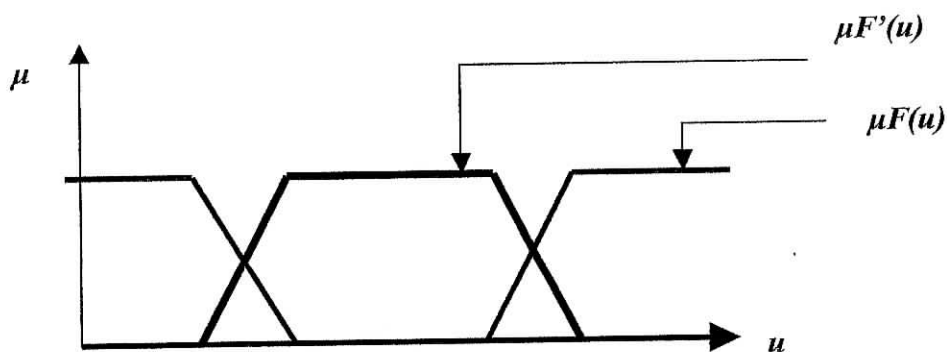


Figure III.4 : Le complément d'un ensemble flou

III .2.9.2- L'intersection :

L'intersection de deux sous ensemble flous F et G est défini par la fonction d'appartenance: $\mu_{F \cap G}(u) = \text{Min}(\mu_F(u), \mu_G(u))$ [Zad65]. D'une manière générale, [DUB84] a défini l'intersection à partir d'une T-norme notée : (*)

$$\forall u \in U : \mu_{F \cap G}(u) = \mu_F(u) * \mu_G(u)$$

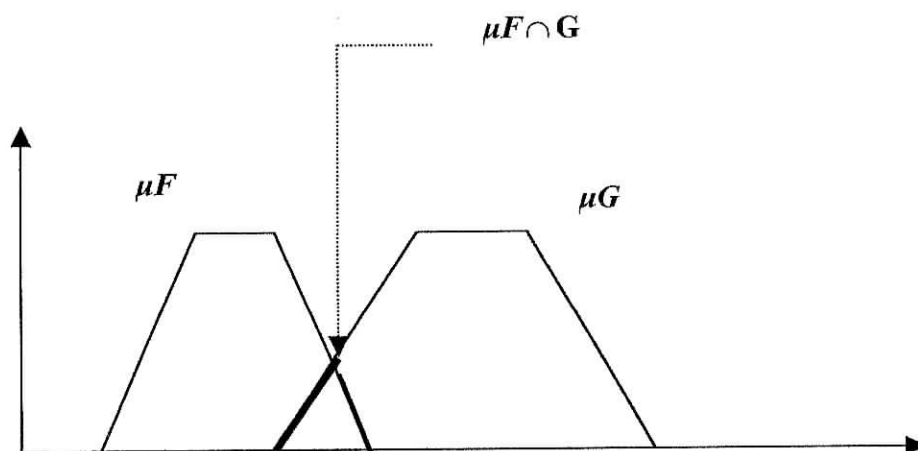


Figure III. 5 : Intersection de deux sous ensembles flous

III .2.9.3- L'union :

L'union de deux sous ensembles flous F et G est défini par la fonction d'appartenance $\mu_{F \cup G}(u) = \text{Max}(\mu_F(u), \mu_G(u))$ [ZAD65]. Toujours, d'une manière générale, Dubois a défini l'union à partir d'une T-conorme \perp .

$$\forall u \in U : \mu_{F \cup G}(u) = \mu_F(u) \perp \mu_G(u)$$

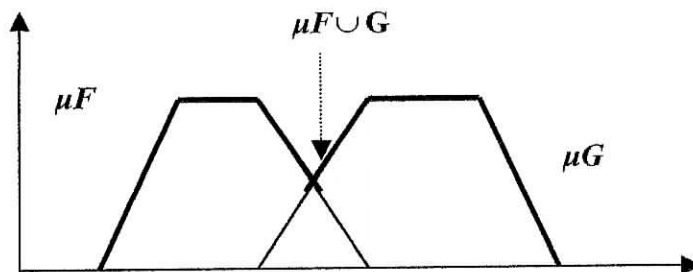


Figure III. 6 : Union de deux sous ensembles flous

III .2.9.4- Egalité :

Deux sous ensembles sont égaux, on écrit $F=G$, si seulement si:

$$\forall u \in U : \mu F(u) = \mu G(u).$$

III .2.9.5- Le produit algébrique :

Le produit algébrique de deux sous ensembles flous F et G est un produit noté $F.G$ est défini par:

$$\mu F.G(u) = \mu F(u) \cdot \mu G(u) \quad , \quad \forall u \in U.$$

III .2.9.6- La somme algébrique :

La somme algébrique de deux sous ensembles flous F et G notée $F \otimes G$ est défini par:

$$\mu F \otimes G(u) = \mu F(u) + \mu G(u) - \mu F(u) \cdot \mu G(u) \quad , \quad \forall u \in U$$

On peut vérifier que $F \otimes G = (F'.G')$.

III .2.9.7- L'inclusion :

L'ensemble flou F est dit inclus dans l'ensemble flou G si et seulement si :

$$\mu F(u) \leq \mu G(u) \quad , \quad \forall u \in U.$$

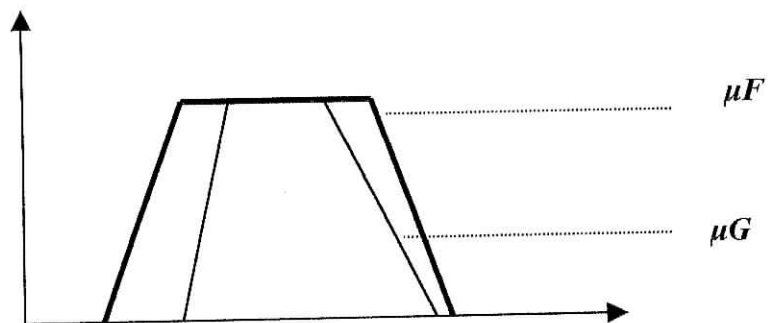


Figure III. 7 : Inclusion d'ensemble flou

III .2.10- Indice d'évaluation et de comparaison d'ensembles flous :

On peut évaluer un ensemble flou par rapport au nombre d'éléments qu'il contient et à quel point il se distingue d'un ensemble classique ; à quel point un ensemble flou est flou et à quel point il est non-vide , cela se fait par les moyens suivants :

❖ Notion de cardinalité :

$$|F| = \mu^F(u) .$$

❖ Indice de flou :

$$f(F) = 1 - 1/|U| * |\mu^F(u) - \mu^G(u)| .$$

❖ Hauteur d'un ensemble flou :

$$Ht(F) = \sup_u \mu^F(u) .$$

❖ Indice de chevauchement :

$$I_1[F, G] = ht(F \cap G) .$$

III .2.11- Relations floues :

Un ensemble flou R défini sur un produit cartésien $X*Y$ donne naissance à une relation floue. Soit μ_R sa fonction d'appartenance, $\mu_R(x, y)$ estime à quel point x est en relation avec y ; plus $\mu_R(x, y)$ est proche de 0, moins x est en relation avec y ; quand $\mu_R(x, y)$ est égale à 1, x et y sont complètement et indiscutablement en relation.

Les relations floues permettent de représenter des égalités approchées, des relations d'ordre tel que "être beaucoup plus grand que" des relations de préférence graduées, des relations d'implication pondérées.

Plus généralement, une relation floue dans un produit d'univers $X = X_1 \times X_2 \times \dots \times X_n$, est un ensemble flou dans X caractérisé par la fonction d'appartenance $\mu_R(x_1, x_2, \dots, x_n)$, $x_i \in X_i$ pour $i=1..n$.

Exemple

Un sous ensemble flou de couple d'éléments (u, v) avec leur degré d'appartenance $\mu_R(x, y)$ qui évalue le degré de liaison entre u et v ; soit $U=V=R^1$ où R^1 est la ligne réelle $]-\infty, +\infty[$, alors $(u > v)$ est une relation floue dans R^2 , une expression subjective de μ_R dans ce cas doit être:

$$\mu_R(u, v) = 0 \text{ si } u \leq v$$

$$\mu_R(u, v) = (1 + (u - v)^{-2})^{-1} \text{ si } u > v .$$

A l'aide des relations floues, on peut modéliser des prédicats graduels complexes, en combinant à l'aide des opérations ensemblistes floues, des prédicats graduels plus simples.

III .2.11.1- Définition de relation de similarité et relation de préférence :

Une relation de similarité est une extension de la relation d'équivalence usuelle, en généralisant facilement les propriétés de réflexivité, de symétrie et de transitivité. De même, une relation de préférence est une extension de la relation d'ordre, elle est réflexive, anti-symétrique et transitive.

III .2.11. 2- Différentes opérations sur les relations floues :

Soient deux relations floues R_1 et R_2 caractérisées par les fonctions d'appartenance respectives μ_{R_1} et μ_{R_2} , alors les opérations usuelles sur les relations floues sont données par:

- ❖ L'image de u par R_1 $U \times V : R_1(u) : \mu_{R_1(u)}(v) = \mu_{R_1}(u, v), \forall v.$
- ❖ L'image inverse de w par R_2 $V \times W : R_2^{-1}(w) : \mu_{R_2^{-1}(w)}(v) = \mu_{R_2}(v, w), \forall v.$
- ❖ Projection d'une relation floue R $U \times V$ dans $U : \mu_{R \text{ sur } U}(u) = \sup_v \mu_R(u, v).$
- ❖ Composition de relations floues $\mu_{R_1 \circ R_2}(u, v) = \sup_w \min[\mu_{R_1}(u, w), \mu_{R_2}(w, v)]$

III .2.11. 3- Propriétés des relations floues :

Les propriétés les plus importantes des relations floues sont:

- ❖ Réflexible : $\mu_R(u, u) = 1, \forall u \in U.$
- ❖ Symétrique : $\mu_R(u, v) = \mu_R(v, u), \forall u, \forall v.$
- ❖ Mi-transitive : $\mu_R(u, v) \geq \min(\mu_R(u, w), \mu_R(w, v)), \forall u, \forall v, \forall w.$

III .2.12- CONCLUSION :

L'introduction de l'idée d'ensemble flou a été motivée par le besoin de représenter des spécifications non strictes telles que des contraintes souples, des critères flexibles , des prédicats vagues etc.

L'utilisation de la théorie de sous ensembles flous pour modéliser de vagues systèmes avec l'aide d'une représentation linguistique apparaît très prometteuse, contrairement à l'utilisation des modèles numériques, car bon nombre de connaissances exprimées en langage naturel ont un sens vague (Flou). Afin de traiter avec ce type de connaissance, il est important de disposer d'une représentation de leur sens , donc représenter des connaissances dans le cas du raisonnement approximatif revient à les traduire en des sous ensembles flous.

On peut constater que la théorie des sous ensembles flous a attiré sur elle l'attention de nombreux scientifiques et a été utilisée dans de nombreux domaines tel que l'automatisme, la météorologie, l'électricité. Elle possède désormais un corpus de résultats et d'outils formels établis qui semble être parvenu à une certaine maturité.

La théorie des sous ensembles flous peut être appliquée dans trois (03) sortes de problèmes:

- *Les problèmes de satisfaction et de propagation de contraintes : problèmes de conception , d'aide à la décision et recherche documentaire dans les bases de données.*
- *Les problèmes d'étiquetages symboliques et de classification : reconnaissance des formes, vision par ordinateur, l'analyse des données , les capteurs intelligents.*
- *Les problèmes de représentation de l'imprécis et de l'incertain et leurs propagations : intelligence artificielle, analyse du risque basé sur des opinions d'experts.*

Chapitre VI

*Prévision des pluies par la logique
floue*

LA THEORIE DES POSSIBILITES

IV . 1- INTRODUCTION :

La théorie des possibilités a été proposée dans son principe en 1978 par ZADEH [ZAD78] à partir de la notion d'ensemble flou qu'il avait lui-même introduite une dizaine d'années auparavant [ZAD65], un nouveau cadre était ainsi offert pour la représentation et le traitement unifié:

- i) d'informations imprécises et/ou incertaines (grâce à la théorie des possibilités) et*
- ii) de contraintes souples, de critères flexibles et de classes aux frontières vagues (grâce à la notion d'ensemble flou).*

Nous insistons dans la suite sur la distinction qu'il y'a lieu de faire entre ensemble flou et mesure de possibilité, les deux pouvant s'envisager de façons complètement séparée mais peuvent aussi être utilisées conjointement de façon harmonieuse quand l'application considérée le nécessite.

Une information est dite imprécise, si elle est incomplète, insuffisante pour l'usage à la réponse à une question claire. L'imprécision se traduit par une hésitation face à plusieurs choix, elle est toujours représentée par une disjonction ; un ensemble d'alternatives qui exprime l'imprécis est appelé l'ensemble disjonctif.

Exemple

« Le mardi prochain, il pleuvra entre 10, 15 ou 20 mm. »

Cet ensemble disjonctif qui décrit l'imprécis, indique quelles sont les valeurs possibles et quelles sont les valeurs impossibles pour une variable et comme la proposition est considérée aujourd'hui, la proposition est aussi incertaine. Ainsi comme la notion de possibilité est continue et graduelle, de l'impossibilité totale, en passant par des degrés intermédiaires ordonnés , on arrive à la possibilité absolue. La fonction caractéristique d'un ensemble disjonctif peut être vu comme une distribution de possibilité.

IV . 2- La modélisation de l'incertitude :

Du point de vue vocabulaire, tout événement peut être: probable, plausible, crédible, possible ou nécessaire (certain), plusieurs théories mathématiques ont été proposés pour modéliser ces différentes formes d'incertitudes.

IV . 2.1- La probabilité :

Le calcul de la probabilité est la façon la plus ancienne et la mieux connue pour quantifier l'incertain, une mesure de probabilité est défini par une application de U (un référentiel), pouvant être un ensemble de propositions dans l'intervalle unité $[0,1]$ vérifiant les axiomes suivant:

- I. $\text{Prob}(\emptyset) = 0.$**
- II. $\text{Prob}(U) = 1.$**
- III. Axiome de l'additivité : $\forall A \in U, \forall B \in U$
Si $A \cap B = \emptyset$ alors $\text{Prob}(A \cup B) = \text{Prob}(A) + \text{Prob}(B).$**

(A et B sont dits mutuellement exclusif ou incompatible puisque si l'une des propositions est vrai, l'autre est fausse : ceci équivaut à $A \Rightarrow \text{non}B = 1.$

On déduit: $\forall A \in U : \text{Prob}(A) + \text{Prob}(\text{non}(A)) = 1;$ qui signifie que la connaissance de la probabilité d'un événement détermine la probabilité de l'événement contraire. Cette mesure de probabilité s'avère insuffisante pour la représentation des données [DUB88a], puisqu'elle nécessite d'énumérer un ensemble exhaustif d'alternatives mutuellement exclusive, or rien ne prouve que l'ensemble de faits (observations et données) soit exhaustif et mutuellement exclusif. Cette insuffisance a conduit à généraliser cette mesure de probabilité, ceci a donné, entre autre la mesure de confiance.

IV . 2.2- Mesure de confiance :

Elle désigne les fonctions qui évaluent en quelque sorte la confiance que nous avons en la vérité d'une proposition P. Prenons la proposition suivante : " x est A ", elle doit être vraie ou fausse , selon que la valeur prise par la variable x appartient ou non à l'ensemble A. La valeur de vérité sera le caractère vrai ou faux de la proposition; si de plus , l'information est incomplète, cette valeur de vérité peut ne pas être établie avec certitude.

La mesure de confiance g est une application de U dans $[0,1]$ qui vérifie les conditions suivantes :

- I. $g(\emptyset) = 0$.
- II. $g(U) = 1$.
- III. Si $A \Rightarrow B=1$ alors $g(B) \geq g(A)$, (*).
(B sera vrai au moins chaque fois que A le Sera).

Remarque :

L'axiome (*) caractérise une large famille de fonctions :

- la fonction de croyance introduite par Shafer [SHA 76] , appelée généralement "fonction de crédibilité".

- fonction de plausibilité, obtenue par dualité de la fonction de crédibilité.

IV . 2.3- La plausibilité et la crédibilité :

Une mesure de crédibilité est une mesure floue qui est exprimée à partir d'une fonction m appelée fonction de croyance, m est une fonction de U dans $[0,1]$ tel que:

- I. $m(\emptyset) = 0$.
- II. $\sum_{p \in U} m(p) = 1$.

ou $m(p)$ représente "la part de croyance" associée à p et à p seulement. La fonctions de crédibilité Cr , basée sur m , est définie par :

$$\forall q \in P(U), \quad Cr(q) = \sum_{p \text{ entraîne } q} m(p) \quad \text{ou « } p \text{ entraîne } q \text{ » signifie } (A \Rightarrow B=1).$$

Par dualité, la fonction de plausibilité PL est définie à partir de Cr :

$$\forall p \in P(U), \quad PL(p) = Cr(\text{non } p) \quad \text{i.e. : } \forall q \in P(U) : PL(q) = \sum_{q \text{ entraîne } p} m(p)$$

$$\forall p \in P(U), \quad \text{on a : } Cr(p) + Cr(\text{non } p) \leq 1.$$

$$PL(p) + PL(\text{non } p) \geq 1$$

Deux cas particuliers et importants de fonctions de crédibilité et de plausibilité sont données par la "mesure de nécessité et possibilité".

IV . 2.4- Mesure de Possibilité :

Étant donné un référentiel U , une fonction d'ensemble, appelée "mesure de possibilité Π " définie sur l'ensemble $P(U)$ des parties de u à valeurs dans $[0,1]$ peut se définir par :

$$1 \text{ si } A \cap E = \emptyset$$

$$\forall A \in U, \quad \Pi(A) = \begin{cases} 1 & \text{si } A \cap E = \emptyset \\ 0 & \text{sinon} \end{cases}$$

Soulignons que la mesure de possibilité Π est induite à partir de l'ensemble E , nous ne faisons cependant pas apparaître E dans la notation afin de l'alléger.

L'appellation "mesure de possibilité" est en accord avec l'interprétation suivante: sachant qu'une variable prend sa valeur dans E , il est possible que cette variable prenne sa valeur dans (A) si et seulement si (A) a une intersection non vide avec E .

L'idée de mesure de possibilité se généralise naturellement si l'on souhaite exprimer que les valeurs dans E sont inégalement possible pour la variable ; on convient alors de donner le degré de possibilité maximale, soit 1, à la valeur ou aux valeurs considérée(s) comme totalement possibles, et un degré de possibilité d'autant plus petit que l'élément de E considéré correspond à une valeur moins possible pour la variable; une valeur jugée totalement impossible recevra un degré de possibilité égal à 0.

Cette extension autorise $\Pi \{(x)\}$, qu'on notera dans la suite $\pi(x)$, à prendre n'importe quelle valeur entre 0 et 1, soit donc $\pi(x)$ le degré de possibilité que x soit la valeur de la variable, π est appelée distribution de possibilité. La mesure de possibilité [ZAD78a] est alors définie par : $\Pi(A) = \text{SUP}_{x \in A} \pi(x)$.

où π est une application de U dans $[0,1]$, peut être vue comme la fonction d'appartenance d'un ensemble flou [ZAD 65].

La fonction d'ensemble Π définie ci-dessus satisfait la propriété (caractéristique dans le cas fini) : $\forall A, \forall B, \quad \Pi(A \cup B) = \text{Max}(\Pi(A), \Pi(B))$ (*), qui exprime que réaliser « A ou B » est exactement aussi possible que de réaliser le plus possible de A et de B . Ainsi les axiomes régissant les mesures de possibilité sont :

I. $\Pi(\emptyset) = 0$.

II. $\Pi(U) = 1$.

III. $\Pi(A \cup B) = \text{Max}(\Pi(A), \Pi(B)) : \forall A, B \in U$.

- $\Pi(A)$ évalue à quel point A est possible compte tenu de l'information disponible; ainsi on a:
- $\Pi(A) = 0$ évalue à quel point A n'est pas du tout possible, ce qui entraîne que A est certainement fausse.
- $\Pi(A) = 1$ signifie que A est totalement possible, ce qui est beaucoup moins fort que " A est certainement vrai ", en effet en l'absence d'informations, A et $\text{non}(A)$ sont toutes deux possibles, mais l'une des deux seulement est vraie, on a $\text{Max}(\Pi(A), \Pi(\text{non } A)) = 1$.

IV . 2.5- Mesure de nécessité :

La mesure duale associée à la mesure de possibilité est appelée "mesure de nécessité" elle est définie par :

$$\forall A \in U, N(A) = 1 - \Pi(\square A'), \quad (**)$$

où A' est le complément de A dans U , ce qui donne :

$$\forall A \in U \quad N(A) = \begin{cases} 1 & \text{si } A \subseteq E \\ 0 & \text{sinon} \end{cases}$$

L'appellation "mesure de nécessité" traduit que la nécessité d'un événement correspond à l'impossibilité de l'événement contraire, dualité usuelle entre le possible, le nécessaire depuis Aristote . On pourrait exprimer à quel point on est certain que chaque valeur u de U n'est pas la valeur de la variable considérée, puisque :

$$N(\{u\}) = 1 - \pi(x), \text{ d'où d'après } (**) \text{ on a :}$$

$$N(A) = \text{INF}_{u \in A} (1 - \pi(x)).$$

Les axiomes régissant la mesure de nécessité sont :

- I.** $N(U) = 1$.
- IV.** $N(\emptyset) = 0$.
- II.** $\forall A, \forall B \in U, N(A \cup B) = \text{Min}(N(A), N(B))$.

Remarque :

- $N(A) = 1$ signifie que A est totalement certaine.
- $N(A) = 0$ signifie seulement que A n'est pas du tout certaine, qui est beaucoup moins forte que " A est fausse ".

IV . 2.6- Comparaison entre la mesure de possibilité et la mesure de nécessité :

Nous avons les propriétés suivantes :

$$-\forall A, N(A) \leq \Pi(\Box A)$$

$$-\forall A, N(A) > 0 \Rightarrow \Pi(\Box A) = 1 \Leftrightarrow \forall A, \Pi(\Box A) < 1 \Rightarrow N(A) = 1.$$

Ce qui exprime qu'un énoncé doit être totalement possible avant d'être quelque peu certain ,par conséquent , le couple (N(A) , $\Pi(\Box A)$) a nécessairement l'un des termes suivants:

$N(A)$	$\Pi(\Box A)$	L'état du fait A
1	1	A est certainement vrai
$A \in]0,1[$	1	A est plus certaine que -A
0	1	Ignorance totale quand à la vérité ou la fausseté de A
0	$\alpha \in]0,1[$	-A est plus certain que A
0	0	A est certainement fausse

IV . 2.7- Mesure d'information :

Il est possible de construire des mesures scalaires relatant l'idée d'imprécision au sens introduit ci-dessus ; le plus simple est sans doute la cardinalité. La cardinalité scalaire |F| d'un ensemble flou a été initialement défini par:

$$|F| = \sum_u \mu_F(u) .$$

où l'ensemble support de F est $\{ u, \mu_F(u) > 0 \}$ est supposé fini, |F| est encore égal à :

$$|F| = \sum_i (\alpha_i - \alpha_{i+1}) \cdot |F\alpha_i| \text{ ou } F\alpha_i \text{ est le } \alpha_i\text{-coupe de F défini par :}$$

$$F\alpha_i = \{u, \mu_F(u) \geq \alpha_i\}$$

IV . 3- Sous ensemble flou et le concept de possibilité :

La distribution de possibilité $\mu: U \rightarrow [0,1]$ est une fonction qui associe à chaque élément u de U la possibilité que x peut prendre u comme valeur ainsi, on a: $Poss\{x = u\} = \mu(u)$.

Pour comprendre la relation entre un sous ensemble flou et la distribution de possibilité, il convient de considérer cet exemple: "aujourd'hui, la température est entre 18 et 20°C" cette proposition exprime :

- Il est possible qu'une valeur comprise entre 18 et 20 soit la température exacte.
- Il est impossible qu'une valeur en dehors de [18,20] soit la température d'aujourd'hui.

Cette proposition induit une distribution de possibilité où x est la température, qui associe à chaque valeur u de [18,20] la possibilité μ de [0,1] que x puisse être la température d'aujourd'hui. on a :

$$\begin{aligned} \text{Poss}\{x=u\} &= 1 & \text{si} & & 18 \leq u \leq 20. \\ \text{Poss}\{x=u\} &= 0 & \text{si} & & u > 20, u < 18. \end{aligned}$$

D'une manière générale, soit la proposition floue suivante: "la température est normal", à partir de sous ensemble défini dans la théorie des sous ensembles flous auparavant .

$TN = 0/10+0.3/15+0.7/20+1/25+1/30+0/40$. cette proposition induit une distribution de possibilité tel que :

- $\text{Poss}\{x=10\} = \text{Poss}\{x=40\} = 0$.
- $\text{Poss}\{x=15\} = 0.3$.
- $\text{Poss}\{x=20\} = 0.7$.
- $\text{Poss}\{x=25\} = \text{Poss}\{x=30\} = 1$.

Nous remarquons donc, que le degré d'appartenance $\mu_A(u)$, est interprété comme le degré de possibilité que u soit la valeur de x , conformément à la proposition "x est A", où A est un sous ensemble flou de l'univers de discours U , x est une variable de A , et μ_A c'est la fonction d'appartenance de A , ainsi :

$$\forall u \in U : \pi(x) = \mu_A(u),$$

IV . 4- Distribution de possibilité conditionnelle:

Une distribution de possibilité conditionnelle est une fonction :

$$\begin{aligned} \pi_{y/x} : \quad & \forall x \in U \rightarrow [0,1] \\ & (v, u) \rightarrow \pi_{y/x}(v, u) \end{aligned}$$

tel que $\pi_{y/x}$ est la possibilité que y prenne la valeur v sachant que x prend la valeur u . $\pi_{y/x}$ limite donc les valeurs possibles de y pour une valeur donnée de x . Cette distribution peut être représentée par des règles de la forme: "Si x est A Alors y est B ".

IV . 5- Distribution de possibilité conjointe :

Soit la distribution de possibilité $\pi(x)$ qui restreint les valeurs possibles de la variable x , les valeurs de distribution de possibilité conjointes réduisant les valeurs possibles du couple (x, y) , sont obtenues en combinant les distributions de possibilité $\pi(x)$ et $\pi_{y/x}$:

$$\pi(u, v) = (\pi(u) * \pi_{y/x}(v, u)) \quad \forall u, \forall v \quad \text{où } (*) \text{ est une opération, généralement T-norme .}$$

IV . 6- Principe de minimum de spécificité :

Le principe de minimum de spécificité consiste à choisir parmi un ensemble de distributions de possibilités envisageables, celle qui est la plus grande : maximisé les degrés de possibilité des valeurs envisageables, tel que $\pi(x) < \pi'(x)$ alors $\pi(x)$ est dit « **plus spécifique** » que $\pi'(x)$.

IV . 7- La Possibilité et la probabilité :

La théorie de possibilité peut représenter l'ignorance tandis que la théorie de probabilité ne peut exprimer l'ignorance par une distribution unique et elle ne peut distinguer entre possible et certitude, en prolongeant la théorie des possibilités dans un cadre probabiliste, alors la mesure de possibilité et la mesure de nécessité engendrent la valeur de probabilité.

IV . 8- Conclusion :

les mérites de la théorie de possibilité pour la formulation du raisonnement résident dans ses capacités à représenter de manière naturelle et simple les informations disponibles quand elles sont imprécises ou incertaines, dans le caractère quantitatif des représentations à utiliser en pratique (les opérations de max, min et le complément à un) , et dans le fait qu'elle conduit à des calculs qui restent en général assez simples à mettre en œuvre.

La théorie de possibilité et la théorie de probabilité sont des outils complémentaires pour la représentation de facettes différentes de l'incertitude, les probabilités correspondant à des états de connaissance précis mais contradictoire (un dé tombe tantôt sur une face, tantôt sur une autre), tandis que la théorie des possibilités permet de représenter des états de connaissance imprécis et cohérents (dans la mesure où un ensemble flou peut être vu comme une combinaison convexe).

Chapitre V

*La modélisation par les systèmes
d'inférence flous*

LA MODELISATION PAR LES SYSTEMES D'INFERENCE FLOUS

V. 1- Introduction :

La complexité d'un système peut être caractérisée par plusieurs facteurs : le nombre d'éléments qui le composent, les relations entre ces éléments, leurs caractéristiques, la dynamique du système ... etc.

*Lorsque la complexité augmente, les méthodes de modélisation traditionnelles s'avèrent souvent incapables de représenter le comportement global du système. L'impuissance des techniques quantitatives traditionnelles pour décrire correctement les systèmes complexes a été exprimée par Zadeh sous la forme du principe d'incompatibilité [Zadeh 71] : « **Lorsque la complexité d'un système augmente, notre capacité à faire une description précise et significative de son comportement diminue jusqu'à un seuil au delà duquel précision et signification (pertinence) deviennent des caractéristiques mutuellement exclusives** ».*

*Le traitement des systèmes complexes nécessite souvent la manipulation d'informations vagues, imprécises et incertaines. Pour cela la théorie des sous-ensembles flous constitue en fait une interface entre les mondes linguistique et numérique. L'introduction des concepts flous dans la théorie des systèmes a impliqué la naissance d'une nouvelle classe de systèmes appelés **systèmes flous**, qui sont capables de décrire des systèmes complexes en introduisant des informations vagues et en imitant le mécanisme de raisonnement approximatif utilisé chez l'homme.*

De nos jours, les applications des systèmes flous se sont multipliées. Entre autres, on peut citer, les contrôleurs flous, les modèles flous, les systèmes experts flous, les systèmes de diagnostic, de classification et de reconnaissance des formes flous etc.

*Malgré les différences dues aux particularités de chaque application, ces systèmes ont la même structure interne, et ils utilisent tous un mécanisme d'inférence. Ils sont baptisées **SYSTEMES D'INFERENCE FLOUS (SIF)**.*

V. 2- Notion de base :

une variable linguistique [Zad 71] est une variable dont les valeurs sont des mots ou des phrases exprimés dans une langue naturelle ou un langage artificiel. Une variable linguistique est définie par le quadruplet suivant [Driankov 93] :

$$\langle x_nom, L(x), X, Mx \rangle \quad \text{où}$$

- x_nom : Est le nom de la variable linguistique, comme température, précipitation, débit... etc.
- $L(x) = \{L_1, L_2, \dots, L_n\}$ l'ensemble des valeurs linguistiques que peut prendre la variable x_nom , exemple : $L(x) = \{\text{froid, frais, tiède, chaud}\}$. L'ensemble $L(x)$ est aussi appelé ensemble de termes (terme set), ensemble de symboles ou de référence associé à la variable (x_nom).
- X : Est le domaine physique associé à la variable x_nom , il est aussi appelé univers de discours. C'est en fait l'ensemble de toutes les valeurs numériques que peut prendre la variable numérique associée à la variable linguistique x_nom . Par exemple, X peut être définie comme étant l'intervalle $[0 \dots 1000]$.
- Mx : Est une fonction sémantique qui associe à tout symbole de $L(x)$ une signification floue.

La définition d'une variable linguistique regroupe donc à la fois des informations de nature numérique et linguistique. Le lien entre ces deux types d'information nécessite la spécification d'interfaces. Si la signification floue de Mx introduite dans la définition d'une variable linguistique représente l'interface (linguistique - numérique), la notion duale d'interface (numérique - linguistique) n'apparaît pas explicitement dans le concept de variable linguistique.

Elle a toutefois été introduite par Zadeh par le biais de la description floue d'une valeur numérique. Dans l'optique d'introduire un raisonnement linguistique sur des variables de nature numérique, les interfaces de conversion sont essentielles. C'est pourquoi, celles-ci sont utilisées.

Soit Z un ensemble quelconque. $F(Z)$ représente l'ensemble des sous ensembles flous de Z . Les conventions classiques seront utilisées pour la notion des sous-ensembles flous [Zadeh 73].

- Si Z est discret, tout élément F de $F(Z)$ sera noté :

$$F = \sum_{z \in Z} \mu_F(z) / Z$$

- dans le cas où Z est continu, le symbole Σ est remplacé par le symbole d'intégrale \int :

$$F = \int_{z \in Z} \mu_F(z) / Z$$

Dans le cas d'une variable linguistique $\langle x_nom, L(x), X, Mx \rangle$, $F(x)$ représente l'ensemble des sous-ensembles flous de l'univers de discours X . comme X correspond à un domaine physique, tout élément de $F(x)$ est un sous-ensemble flou numérique, c'est-à-dire un sous ensemble flou défini sur un ensemble numérique. De la même manière, $F(L(x))$ est l'ensemble des sous-ensembles flous de l'ensemble de termes de $L(x)$. tout élément de $F(L(x))$ est un sous ensemble flou linguistique, autrement dit un sous ensemble-flou défini sur un ensemble linguistique, discret par nature. Ces sous ensembles flous linguistiques sont aussi appelés des parties floues de niveau 2 [Zadeh 71].

a) Signification floue d'un terme linguistique :

La signification floue est une injection de $L(x)$ vers $F(x)$ associant à tout terme linguistique un sous ensemble-flou numérique le représentant.

$$M_X: L(x) \rightarrow F(x)$$

$$L \rightarrow M_X(L)$$

La signification floue (fuzzy meaning [Zadeh 73]) du terme linguistique L est donc le sous-ensemble flou numérique $M_X(L)$ de fonction d'appartenance $\mu_{M_X}(L)$. Elle définit le sens attribué au terme L sur l'univers de discours X . L'ensemble des significations floues de tous les termes de $L(x)$ définit une partition floue de l'univers X .

Dans la littérature, les symboles et leur signification floue sont souvent confondus. L'écriture $\mu_L(x)$ fait alors référence au degré d'appartenance de la valeur numérique $x \in X$ à la signification floue du symbole L .

b) Description floue d'une valeur numérique :

La description floue d'une valeur numérique, également définie dans [Zadeh 71], est une application du domaine X dans l'ensemble des sous-ensembles flous symbolique $F(L(x))$.

$$D_{L(X)} : X \rightarrow F(L(X))$$

$$x \rightarrow D_{L(X)}(x)$$

$D_{L(X)}(x)$ est donc un sous ensemble flou symbolique qui donne la description de la mesure x par l'ensemble des symboles de $L(x)$. En effet, la description floue d'une valeur x est une partie floue linguistique qui représente pour chaque terme linguistique L , le degré avec lequel ce terme décrit la valeur x [Zadeh 71]. Le degré d'appartenance d'un terme linguistique L à la description floue d'une mesure x peut être assimilé au degré de possibilité que la mesure soit x si celle-ci qualifiée par L ([Dubois 88],[Foulloy 95],[Yager 84]).

Le degré d'appartenance de chaque symbole $L \in L(x)$ à $D_{L(X)}(x)$ est égale au degré d'appartenance de x à $M_X(L)$. Ceci implique la relation suivante qui relie la signification floue et la description floue :

$$\forall L \in L(X), \forall x \in X, \mu_{M_X(L)}(x) = \mu_{D_{L(X)}(x)}(L) \quad (1)$$

Signification et description sont finalement deux manières de spécifier la relation existant entre les nombres et les symboles, autrement dit :

$$\forall L \in L(X), \forall x \in X, \mu_R(L, x) = \mu_{M_X(L)}(x) = \mu_{D_{L(X)}(x)}(L) \quad (2)$$

où R représente la relation liant le symbole L au nombre x .

Exemple :

Soit la variable température. Supposons que l'ensemble des symboles relatifs à cette dernière soit défini par $L(T) = \{B, N, H\}$, où $B = Basse$, $N = Normale$ et $H = Haute$, et que l'univers de discours associé à T soit défini par $T = [0, 50]$. La signification floue des différents symboles est donnée par la fonction M_T . La figure (V.1.a) donne un exemple de la partition obtenue sur T par définition de la signification floue de chacun des termes linguistiques. A partir de la connaissance complète de la variable linguistique $\langle \text{température}, L(T), T, M_T \rangle$, on peut obtenir la description floue de toute mesure numérique de la température en appliquant la relation (1). La figure(V.1.b) montre par exemple la description floue de la mesure $T=30^\circ\text{C}$:

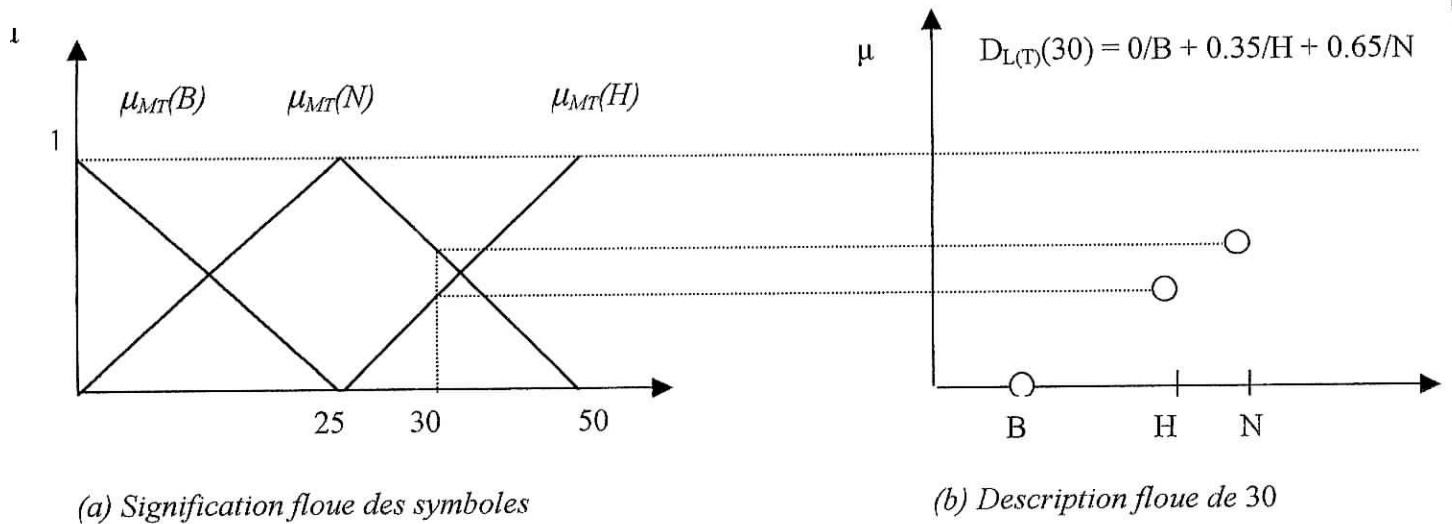


Figure V.1 : Lien entre signification floue et description floue

V . 3- Les différents type de modèles flous :

La modélisation d'un système est une représentation de son comportement par les concepts de la théorie des sous-ensembles flous. Cette représentation caractérise la relation entre les variables d'entrées et de sorties du système. Elle remplace le système d'équations ou la fonction analytique utilisée dans la modélisation classique.

Dans la théorie classique des systèmes , les modèles sont classifiés en modèles linéaires, non-linéaires, dynamiques, statiques ... etc. Cette classification est faite selon la nature des équations définissant la relation entre les variables d'entrée et de sortie du système. En modélisation floue cette classification disparaît. D'un point de vue structurel, les modèles flous peuvent être classés en trois classes essentielles : les modèles linguistiques, ou modèles de Mamdani, les modèles flous relationnels et les modèles à conséquence, ou modèles de Tagagi-Sugeno-Kang (TSK).

V . 3.1- Les modèles linguistiques :

Ces modèles reposent sur l'approche proposée par Zadeh [Zadeh 73]. Ils permettent une description linguistique du système par une base de règles floues. En fait, ils correspondent tout simplement à l'utilisation d'un SIF pour modéliser les relations entrée/sortie d'un système. Des les années 70, Mamdani et ses collaborateurs ont adopté l'approche linguistique pour transformer les règles heuristiques de contrôle, données par l'opérateur humain en une stratégie de contrôle ([Mamdani 74], [Mamdani75], [Mamdani76]). Dans ce sens, un contrôleur flou constitue un modèle de l'opérateur humain et non pas du système lui même.

Du point de vue structure interne, un modèle linguistique peut être implanté avec une vision numérique (SIF_1) ou une vision symbolique (SIF_4).

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 naturelle.

Ces modèles ont les avantages suivants :

- *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. C'est un avantage indéniable quand les données ou l'expertise, ou les deux sont incomplètes.*
- *L'interpolation dépend de la forme des fonction d'appartenance, des opérateurs d'inférence et de la défuzzification. 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*
- *Ces modèles sont bien adaptés à l'utilisation des données floues . Dans ce cas, la partie fuzzification n'est gardée que pour un modèle de type SIF₄. En revanche, pour un modèle de type SIF₁ les entrées sont déjà fuzzifiées. Une méthode d'utilisation des entrées floues à été présentée dans [Palm 94].*

En revanche, les modèles linguistique ont les limitations suivantes :

- *Leur nature non-structurée se traduit par une capacité de représentation limitée. En d'autres termes, un grand nombre de règles est souvent nécessaire pour approcher le comportement global d'un système. Par conséquent, le nombre de paramètres à identifier est élevé [Babuska 95a].*
- *Le grand nombre de paramètres exige une base d'information importante et complète pour déterminer tous les paramètres. Ceci veut dire, un nombre de données d'identification important, mais aussi des données répartie sur la totalité des espaces d'entrée.*
- *Dans le cas des systèmes dynamiques, l'obtention d'un tel ensemble de données constitue une limitation, car certains états ne seront pas atteint lors de l'acquisition des données d'apprentissage. Ces états ne seront alors représentées par aucune règle du modèle résultant. Un problème se pose donc lorsque le modèle est utilisé, et que des situations absentes de base de règles se présentent.*

V . 3.2- Les modèles relationnels :

Un modèle relationnel flou est tout simplement constitué d'une relation floue R décrivant le lien existant entre les différents éléments du système. L'utilisation d'un tel modèle nécessite alors l'application de la règle compositionnelle d'inférence.

Dans le cas d'un système à plusieurs entrées, les équations sont souvent réécrites de manière compacte.

$$F = E_1 \circ E_2 \circ \dots \circ E_r \circ R \quad (3)$$

Dans le cadre des modèles relationnels, la relation R est discrète, c'est-à-dire construite sur des univers discrétisés, qui peuvent être numériques ou symboliques. La relation est alors stockée en mémoire sous la forme d'un tableau multidimensionnel. La différence essentielle avec les modèles linguistique réside dans le fait que la construction de la relation floue n'est pas faite à partir d'un ensemble de règles. De même, un modèle relationnel n'intègre généralement pas les interfaces de fuzzification et défuzzification. Les données disponibles en entrée du modèle sont considérés comme étant floues par nature. De même, aucune défuzzification de la sortie produite n'est envisagée. En résumé, un modèle relationnel peut être vu comme un modèle linguistique discrétisé auquel on a enlevé toutes les interfaces (base de règles, fuzzification et défuzzification). Il n'est alors composé que d'une relation floue discrète.

A l'inverse, lorsque des interface de fuzzification/défuzzification symbolique son associées à un modèle relationnel ([Pedrecz 89]), celui-ci devient équivalent à un modèle linguistique implanté par un SIF₄ ([Rondeau 95]).

En fait, plus qu'à la forme proprement dite du modèle, l'appellation modèle relationnel fait référence à sa méthode de construction, basée sur la théorie des équations floues initiée par Sanchez [Sanchez 76].

V . 3.3- Les modèle de Takagi-Sugeno-Kang (TSK) :

Les modèles linguistiques et relationnels ne permettent pas d'intégrer directement des connaissances objectives sur le système qui ne sont pas exprimées par des sous ensembles flous. Ainsi des connaissance exprimées sous forme analytique décrivant la structure physique du système, par exemple, les équations de conservation d'énergie ou du bilan de masses ne peuvent pas être prise en

compte directement. Sugeno et ses collaborateurs ([Takagi 85],[Sugeno 86]) ont proposé une autre classe de modèles qui permet de représenter ce type d'information. Les modèles de cette classe sont appelés « Modèles de Takagi-Sugeno-Kang (TSK) ». Pour un système MISO (multi input single output) à r entrées et une sortie, un modèle TSK est constitué d'une base de la forme :

$$\text{Si } x_1 \text{ est } A_1^{i1} \text{ et } x_2 \text{ est } A_2^{i2} \text{ et } \dots x_r \text{ est } A_r^{ir} \text{ alors } y = f_i(x_1, x_2, \dots, x_r) \quad (4)$$

où x_1, x_2, \dots et x_r sont les variables numériques d'entrée du modèle flou et f_i 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 flou 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.

V . 3.3.1. Méthode de calcul :

On considère une collection de règles de la forme (4). Soit une entrée (x_1, x_2, \dots, x_r) , la méthode de calcul de la sortie y est composée de 3 étapes :

a) **Calcul des sorties individuelles** : la sortie de chaque règle i est donnée par conséquence, c'est-à-dire :

$$y_i = f_i(x_1, x_2, \dots, x_r) \quad (5)$$

b) **Calcul des degrés d'activation de chaque règle** : le degré d'activation w_i de chaque règle i est calculé par l'équation :

$$w_i = \mu_{M X_1}(A_1^{i1})(x_1) \cdot \dots \cdot \mu_{M X_r}(A_r^{ir})(x_r) \quad (6)$$

c) **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, soit :

$$y^* = \frac{\sum_i w_i \cdot y_i}{\sum_i w_i} \quad (7)$$

Le dénominateur de cette équation étant une constante, on peut reformuler l'expression de y^* sous la forme :

$$y^* = \sum_i v_i y_i \quad (8)$$

où $v_i = \frac{w_i}{\sum_i w_i}$. On peut noter que $\sum_i v_i = 1$

Dans la littérature, les f_i sont souvent des fonction linéaires, soit :

$$f_i(x_1, x_2, \dots, x_r) = b_{i0} + b_{i1}x_1 + b_{i2}x_2 + \dots + b_{ir}x_r \quad (9)$$

où $b_{i0}, b_{i1}, \dots, b_{ir}$ sont des coefficients constants. Chaque règle peut alors être considéré comme un modèle local linéaire selon les variables d'entrée x_1, x_2, \dots, x_r .

remarque :

Dans le cas d'un système dynamique, les différentes instances des variables d'entrée u et de sortie y du système, aux instants antérieurs sont comprises dans les variables du modèle flou x_1, x_2, \dots, x_r .

V . 3.3.2- Avantages et inconvénients des modèles TSK :

Ces modèles ont les avantages suivants :

- 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, généralement linéaire.
- Les modèles TSK ont une capacité de représentation importante. Le nombre de règle nécessaire pour approcher un système à un certain degré de précision est généralement plus faible que dans le cas des modèles linguistiques [Babuska 95a]. L'identification de modèle à

partir de données en est facilitée et nécessite moins de données que pour des modèles linguistiques.

- *La relation (7) relie directement la variable de sortie aux variables d'entrées. Ceci permet une application relativement facile des techniques d'optimisation pour identifier les différents paramètres du modèle.*

Toutefois, les modèles TSK ont les inconvénient suivants :

- *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.*
- *L'interpolation entre les modèles locaux peut engendrer une erreur importante qui réduit la précision globale du modèle [Babuska 95a].*
- *Ces modèle ne permettent pas d'intégrer l'expérience humaine 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. Une solution à ce problème a été proposée dans [Palm 94]. Elle consiste à utiliser les données floues pour calculer les degrés d'activation des règles, et à défuzzifier ces données sur leurs espaces respectifs pour obtenir leur valeurs numériques, qui sont ensuite utilisées dans le calcul de la sortie.*

V . 4- Modélisation des systèmes dynamiques :

Dans la suite du chapitre, nous allons étudier comment un modèle flou peut représenter la dynamique d'un système. Un système dynamique SISO (single input single output) d'ordre n à une seule entrée $u(k)$ et une seule sortie $y(k)$ peut être décrit par :

$$y(k+1) = f[u(k), u(k-1), \dots, u(k-n), y(k), y(k-1), \dots, y(k-n)] \quad (10)$$

où f est une fonction non linéaire. En tant qu'approximateur universel ([Kosko 92b], [Nguyen 93], [Sudkamp 94], [Wang 92b]), un système flou à $2(n+1)$ entrées peut approximer toute fonction non linéaire de $2(n+1)$ variables, en particulier la fonction f de la relation (10).

Ainsi un système dynamique SISO peut être modélisé par un système flou MISO dont la forme des règles est dérivée de (10). Par conséquent, chaque règle est exprimée par :

$$\begin{aligned} \text{Si } u(k) \text{ est } A_0 \text{ et } u(k-1) \text{ est } A_1 \text{ et ... et } u(k-n) \text{ est } A_n \text{ et} \\ y(k) \text{ est } B_0 \text{ et } y(k-1) \text{ est } B_1 \text{ et ... et } y(k-n) \text{ est } B_n \\ \text{alors } y(k+1) \text{ est } C \quad \text{avec le poids } \omega, \end{aligned} \quad (11)$$

où $A_l, l = 0..n$, est un symbole utilisé pour décrire la variable d'entrée u à l'instant $k-l$; $B_l, l = 0..n$, est aussi un symbole, mais cette fois, relié à la variable de sortie y à l'instant $k-l$; C est le symbole caractérisant la variable de sortie y à l'instant $k+1$.

Un modèle flou basé sur un ensemble de règles de la forme (11) représente l'équation de régression non linéaire (10). Il peut donc être vu comme un modèle NARX (Nonlinear AutoRegressive with exogenous input). Dans le cas des systèmes dynamiques MISO, la même discrétisation temporelle peut être appliquée sur chaque variable d'entrée, u_1, u_2, \dots, u_m , par exemple.

La représentation d'un système d'ordre n et à m variables d'entrée conduit à utiliser un modèle flou MISO à $(m+1)(n+1)$ entrées. Bien que la mise en œuvre d'un tel modèle pour de grandes valeurs de m et n ne pose pas de problème d'un point de vue pratique, il est évident que la lisibilité de la base de règles est perdue et que la complexité du modèle est élevée.

Le choix d'une structure NARX pour représenter un système dynamique n'est pas le seul choix possible. Un modèle d'espace d'état peut aussi être considéré. Dans ce cas, le modèle flou représente un système d'équations différentielles utilisant un vecteur d'état auxiliaire [Nguyen 93]. Une collection de systèmes flous MISO doit alors être considérée, où chaque système flou donne l'approximation d'une composante particulière du vecteur d'état.

Nous supposons que le type et la taille du modèle, c'est-à-dire le choix des variables d'entrée/sortie et la détermination des données d'apprentissage, de l'ordre, et des délais ainsi que le but de la modélisation (la prévision dans notre cas), ont fait l'objet d'un travail préalable à la modélisation floue de ce type de système. Dans la littérature de la modélisation floue, le choix de variables et des délais a été souvent effectué par des études empiriques qui consistent à essayer plusieurs choix possibles et à adopter celui qui donne la meilleure performance ([Lee 94], [Yoshinari 93]). Dans les exemples de ce paragraphe, nous gardons le type et la taille de modèle choisi par d'autres chercheurs en modélisation. A chaque fois nous faisons référence à l'origine de notre choix.

V . 4.1- Application à la prévision des séries temporelles :

Le problème de prévision d'une série temporelle $x(t)$ est un problème pratique très important [Box 70]. Le problème de prévision consiste en l'utilisation des valeurs passées de la série jusqu'au point $x(t)$ afin de prévoir la valeur à un instant futur $x(t+P)$. La méthode standard pour ce type de prévision est de constituer un ensemble d'apprentissage à D entrées, échantillonnées à la période S , c'est-à-dire, $(x(t-(D-1)S), \dots, x(t-S), x(t))$, et une sortie (la valeur $x(t+P)$ à prévoir).

Les applications de la prévision peuvent être trouvées dans les domaines de l'économie, prévision météorologique, traitement de signal, contrôle et de nombreux autres domaines.

Pour la suite de notre étude, nous avons choisi les valeurs $D=4$ et $S=P=6$. A partir des données, nous avons aussi extrait N points d'entrée/sortie de la forme :

$$[x(t-18), x(t-12), x(t-6), x(t) ; x(t+6)], \text{ où } t=18 \text{ à } N+17.$$

Les premiers $N/2$ points sont utilisés pour l'apprentissage du système, alors que les autres $N/2$ point sont utilisés pour valider le modèle identifié.

1- Dans le premier cas, on a essayé d'utiliser une nouvelle méthode de modélisation linguistique. Il s'agit d'un algorithme itératif d'apprentissage supervisé qui construit à la fois les fonctions d'appartenances et la base de règle à partir des données numériques sous la forme suivante :

$$\text{Si } x_1 \text{ est } A_1^{i1} \text{ et } x_2 \text{ est } A_2^{i2} \text{ et } x_3 \text{ est } A_3^{i3} \text{ et } x_4 \text{ est } A_4^{i4} \text{ alors } y \text{ est } \omega_i / B_i \quad (12)$$

où x_1, x_2, x_3, x_4 et y représentent respectivement les entrées $x(t-18), x(t-12), x(t-6), x(t)$ et la sortie $x(t+6)$ du système.

L'algorithme part d'un modèle initial avec une paramétrisation minimale, celui-ci évolue à chaque itération jusqu'à obtention d'un modèle acceptable. L'apprentissage peut être arrêté selon un critère d'erreur d'apprentissage ou un nombre de symbole.

L'architecture interne du modèle le place dans la famille SIF4 avec des tâches de fuzzification, inférence et défuzzification purement linguistique. La relation associée à la base de règles est une relation symbolique de petit taille. Cette relation est directement construite à partir de la base de règles.

Dans [Jang 95], la performance du modèle a été mesurée par l'indice d'erreur non-dimensionnel (Non-Dimensional Error Index NDEI). Cet indice est défini par :

$$NDEI = \frac{MSE}{STD} \quad (13)$$

où MSE est l'erreur quadratique moyenne, et STD :

$$STD = \sqrt{\frac{\sum_{i=1}^N (y_i - y_m)^2}{N}}$$

est l'écart type de la série cible. Les y_i sont les valeurs de la sortie et y_m est leur valeur moyenne.

Bien que cette méthode est nouvelle et a fait ces preuves sur des modèles à une ou deux entrées, le modèle reste très difficile à mettre en œuvre pour des systèmes à plusieurs entrées car il se constitue avec un très grand nombre de règles dans le cas multidimensionnelle. Parfois même, comme c'est le cas dans notre application, il peut atteindre 625 règles avec 5 symboles sur chaque entrées, de ce fait il sera illisible et impossible de le représenter.

2- Comme nous l'avons illustré précédemment et vu les inconvénients de la modélisation linguistique et relationnel, nous avons opté pour le choix du modèle de Takagi-Sugeno-Kang.

V . 4.2- Méthodes d'apprentissage liées aux modèles de TSK pour la prévision des séries temporelles

Depuis l'apparition des modèles TSK [Takagi 83], plusieurs chercheurs ont développé des méthodes d'identification de ce modèles. Celles ci identifient le modèle entier ou une partie. Dans ce paragraphe, nous allons présenter quelques unes de ces méthodes.

Les modèles TSK permettent d'établir une relation analytique qui relie les entrées à la sortie. Ceci permet d'appliquer les méthodes d'optimisation ou d'estimation de paramètres ; en effet, les méthodes d'apprentissage des modèles TSK suivent la stratégie suivante :

"Etant donné un modèle initial TSK (4), et un ensemble de données d'entrée/sortie $T=\{(x_1^k, \dots, x_r^k, y_k) ; k= 1..K\}$ du système, trouver les paramètres du modèle qui minimisent un certain critère ?".

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 fonction choisie. Dans le cas de fonction Gaussiennes 2 paramètres (moyenne et écart type) suffisent, soit $p= 2$. dans le cas de fonction triangulaires, on a $p=3$ (valeur modale et les deux extrémités du support). Les paramètres de sortie sont ceux des fonctions apparaissant dans les conclusions de règle. Dans le cas linéaire, $r+1$ paramètres sont nécessaire pour chaque règle.

Si l'on considère un modèle TSK à m règles, avec la définition de I_1, I_2, \dots, I_r termes linguistiques respectivement pour chaque variable x_1, x_2, \dots, x_r , le modèle est alors entièrement paramétré par le vecteur :

$$P = \left[a_{11}, \dots, a_{1I_1}, a_{21}, \dots, a_{2I_2}, \dots, a_{r1}, \dots, a_{rI_r}, b_1, \dots, b_m \right]^T \quad (14)$$

avec a_{ij} vecteur des paramètres relatif au $j^{\text{ième}}$ symbole défini pour la $i^{\text{ième}}$ variable, $i=1..r, j=1, \dots, I_i$, soit $a_{ij} = [a_{ij}^1, a_{ij}^2, \dots, a_{ij}^p]^T$

et b_i vecteur des $(r+1)$ coefficients relatifs à la $i^{\text{ième}}$ règle, $i=1..m$, soit $b_i = [b_{i0}, b_{i1}, \dots, b_{ir}]^T$.

Pour une entrée (x_1^k, \dots, x_r^k) , la sortie du modèle est $\hat{y}_k = \sum_{i=1}^m v_i y_i^k$ (voir équation 8).

La différence entre cette sortie et celle du système pour la même entrée (x_1^k, \dots, x_r^k) est donnée par :

$$e_k = \hat{y}_k - y_k$$

On définit l'erreur quadratique en ce point k par :

$$E_k = \frac{1}{2} \cdot e_k^2 = \frac{1}{2} \cdot (\hat{y}_k - y_k)^2 \quad (15)$$

C'est une erreur instantanée sur un 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 ([Jang 92], [Nomura 92], [Yager 94]).

L'erreur quadratique moyenne est définie par :

$$E = \frac{1}{K} \cdot \sum_{k=1}^K E_k \quad (16)$$

C'est une fonction objective qui constitue un deuxième critère d'optimisation. Ce critère est utilisé pour un apprentissage hors ligne en utilisant tout l'ensemble des données d'apprentissage ([Jang 92], [Yager 93], [Takagi 85], [Ragot95]).

Ces deux critères sont les plus couramment utilisés dans l'identification de modèles TSK. D'autres critères ont été définis pour juger de la pertinence de modèle par rapport au système [Walter 94] et [Tong 80].

Dans la littérature, les méthodes d'optimisation de ces critères sont nombreuses. Nous présentons ici deux méthodes, qui sont l'optimisation par descente de gradient et optimisation par l'utilisation d'un filtre de Kalman.

a) Optimisation par une méthode de descente de gradient :

Notons J l'un des critères (15) ou (16). Une méthode de descente de gradient [Curry 44] est un algorithme itératif de recherche du vecteur de paramètres P qui minimise la fonction objective $J(P)$. A chaque itération, le vecteur P est ajusté de façon, à décrémente la valeur de $J(P)$. Les règles d'apprentissage sont exprimées par les formules suivantes :

$$a_{ij}^l(t+1) = a_{ij}^l(t) - \alpha_{a_{ij}^l} \frac{\partial}{\partial a_{ij}^l} J(P), \text{ avec } i=1..r, j=1..I_i \text{ et } l=1..p \quad (17)$$

$$b_{ij}(t+1) = b_{ij}(t) - \alpha_{b_{ij}} \frac{\partial}{\partial b_{ij}} J(P), \text{ avec } i=1, \dots, m, j=0, \dots, r \quad (18)$$

où t est le nombre d'itérations d'apprentissage, $\alpha_{a_{ij}^l}$ et $\alpha_{b_{ij}}$ sont des constantes positives appelées taux d'apprentissage. Elles jouent un rôle important dans la vitesse de convergence de l'algorithme. Cette méthode peut être appliquée aux deux critères (15) et (16).

b) Utilisation d'un filtre de Kalman :

Ce filtre est une méthode de minimisation de l'erreur quadratique moyenne sur un ensemble de points. La minimisation est effectuée en considérant que les fonctions d'appartenance des entrées sont fixées, c'est-à-dire que la fonction objectif ne dépend que des paramètres des conséquences. Cet algorithme considère le critère (16). Si nous supposons que les f_i sont des fonctions linéaires de la forme $f_i(x_1, x_2, \dots, x_r) = b_{i0} + b_{i1}x_1 + b_{i2}x_2 + \dots + b_{ir}x_r$, à partir de l'équation (8) on déduit :

$$\hat{y}_k = V_k^T B, k = (1..K) \quad (19)$$

$$\text{où, } V_k^T = \left[v_{1k}, \dots, v_{mk}, (v_{1k}, \dots, v_{mk}) \cdot x_1^k, \dots, (v_{1k}, \dots, v_{mk}) \cdot x_r^k \right] \quad (20)$$

$$B^T = \left[b_{10}, \dots, b_{m0}, b_{11}, \dots, b_{m1}, \dots, b_{1r}, \dots, b_{mr} \right] \quad (21)$$

sont deux vecteurs de dimension $m.(r+1)$. L'objectif E est alors défini par :

$$E = \frac{1}{2K} \cdot \|VB - y\| \quad (22)$$

où $V = \left[V_1^T, \dots, V_K^T \right]^T$ et $Y = \left[y_1, \dots, y_K \right]$ est le vecteur composé des sorties des points d'apprentissage.

Il est évident que l'erreur quadratique moyenne atteint son minimum lorsque l'égalité :

$$VB = Y \quad (23)$$

est satisfaite. Malheureusement, le nombre K de points d'apprentissage est généralement très supérieur au nombre de paramètres $m.(r+1)$, et l'équation (23) n'a pas de solution exacte.

La solution au sens des moindres carrés est alors obtenue par pseudo-inversion de la matrice V ce qui donne :

$$B = \left(V^T V \right)^{-1} V^T Y \quad (24)$$

Le problème de cette méthode réside en la nécessité d'inverser la matrice $V^T V$. Takagi et Sugeno [Takagi 85] ont proposé une solution consistant à utiliser un filtre de Kalman qui permet de trouver une estimation de B par un calcul récursif :

$$B_{k+1} = B_k + W_{k+1} V_{k+1} \left(y_{k+1} - V_{k+1}^T B_k \right) \quad (25)$$

$$\text{avec } W_{k+1} = W_k \frac{W_k V_{k+1} V_k^T W_k}{1 + V_k^T W_k V_{k+1}}, k = 0..K-1 \quad (26)$$

et les valeurs initiales suivantes : $B_0 = 0$ et $W_0 = \alpha \cdot I$ ($\alpha = \text{grand nombre}$, et I est la matrice identité).

c) Optimisation par une approche neuro-floue :

Les réseaux de neurones (RN) multi-couches constituent des approximateurs universels. L'atout principal de ces réseaux réside dans leur capacité d'apprentissage. Par contre, leur structure et leurs paramètres n'ont pas toujours des justifications physiques. De plus, la connaissance humaine ne peut pas être exploitée pour les construire. Les systèmes d'inférences flous sont également des approximateurs universels [Wang92b]. ces systèmes possèdent deux point forts par rapport aux RN. D'une part, ils sont généralement construits à partir de la connaissance humaine, d'autre part, ils ont une capacité descriptive élevée due à l'utilisation de variables linguistiques. Il est donc apparu naturel de construire des systèmes hybrides qui combinent les concepts des systèmes d'inférence flous et des RN. Ainsi l'approche Neuro-floue est apparue.

Plusieurs auteurs ont utilisés cette approche ([Takagi91], [Jang93a], [Lin95]). Jang a proposé un type spécial de réseaux neuro-flous. Il s'agit de l'approche ANFIS (système d'inférence flou organisé en réseaux adaptatif), qui est un réseau multi-couches dont les connexions ne sont pas pondérées [Jang93a].

V . 5- Conclusion :

Nous avons consacré une partie de ce chapitre à l'étude des systèmes d'inférence flous. Dans un premier temps, nous avons abordé la notion de variable linguistique qui permet de représenter par des mots des informations numériques imprécises. Malgré la diversité des application possibles, la structure globale d'un SIF est toujours la même. Un SIF se compose de trois parties essentielles : la fuzzification, l'inférence et la défuzzification. La nature de la défuzzification est déterminée par celle du résultat de l'inférence.

La seconde partie a été consacrée à la présentation des différents types de modèles flous et leurs apprentissages, qui sont, les modèles de Mamdani ou(Linguistiques), les modèles relationnels et les modèles de TSK.

Dans le contexte général de la modélisation floue, les modèles TSK sont également introduits et utilisés par la suite . Ils correspondent , en fait à une approche multi-modèle dans laquelle le flou est utilisé pour agréger les sorties produites par chacun des modèles locaux.

Finalemment la modélisation floue (linguistique et TSK) a été utilisé pour la modélisation des systèmes dynamiques, qui représentent les systèmes sur lesquels nous nous baserons pour la prévision des séries temporelles .

En conclusion les modèles linguistiques ont une représentation moins efficace que les modèles TSK et demandent par conséquent plus de règles pour décrire le comportement d'un système, ces deux type de modèles flous nous paraissent complémentaires. Le choix de l'un ou de l'autre dépend en effet des données du problème et du but de la modélisation.

Chapitre IV

La théorie des possibilités

PREVISION DES PLUIES PAR LA LOGIQUE FLOUE

VI. 1- INTRODUCTION

Dans le but et la nécessité d'assurer une prévision des précipitations futures adéquates pour une gestion efficace des ressources hydriques et pour permettre une prise de décision idoine durant les situations de crise telles que celles créées par des inondations ou sécheresse, la prévision des précipitations confirme et semble de plus en plus indispensable. Les prévisions des apports hydriques futurs permettent la répartition des volumes d'eau stockés au niveau des réservoirs afin de satisfaire les besoins industriels, agricole, alimentation en eau potable ... etc.

Certains problèmes reliés à la prévision ont été abordés dans des projets . Pour plusieurs raisons, il sera intéressant d'élargir, le cadre de la recherche portant sur les modèles de prévision en profitant de l'expérience acquise par le chercheurs. Il s'agit d'un problème complexe et très important puisque des prévisions adéquates à court, moyen et long termes contribuent à prévenir les risques inhérents à la gestion d'ouvrages hydrauliques et à optimiser l'utilisation de la ressource. Une éventuelle amélioration des prévisions pourra alors avoir des répercussions économiques importantes. De ce fait le domaine de la prévision a connu de grands développements au cours de ces dernières années.

Compte tenue de la limitation de la prévision météorologique classique qui ne pourra atteindre des échéances supérieures à deux ou trois semaines, les possibilités de prévision à plus long terme reposent principalement sur le développement d'une méthodologie nouvelle.

*À notre avis pour résoudre ce problème il est nécessaire de refuser des méthodes conventionnelles d'analyse statistique (des fonctions de corrélation, de l'analyse de régression etc.), qui assurent l'exposition seulement des relations linéaires. En même temps la plupart de processus et de rapports météorologiques devraient être référés (par leur nature) à des relations non linéaires, ainsi notre objectif est d'utiliser **la modélisation basée sur la logique floue** pour la prévision de ses paramètres météorologiques.*

VI . 2- Mécanisme de raisonnement :

1- tout d'abord, afin de prévoir la valeur à un instant futur $x(t+P)$. La méthode conforme pour ce type de prévision est de former un ensemble d'apprentissage à D entrées, échantillonnées à la période S , c'est-à-dire, $(x(t-(D-1)S), \dots, x(t-S), x(t))$, et une sortie (la valeur $x(t+P)$ à prévoir).

Pour notre modèle le choix de P va être modifié et changé au cours de cette étude pour pouvoir constituer l'ensemble d'apprentissage le mieux adapté pour le modèle le plus performant.

Pour des raisons de capacité d'exécution du programme informatique et de la croissance exponentielle du nombre de règles, qui pour deux descriptions sur chaque fonction d'appartenance (les symboles d'entrées), pour chaque entrée, donne $2^4 = 16$ règles pour quatre entrées. Alors on a constitué un ensemble d'apprentissage avec quatre entrées et une sortie de la forme :

$$[x(t-3P), x(t-2P), x(t-P), x(t) ; x(t+P)] \quad t= 3P \dots (N-P).$$

tel que $x(t-3P) = X_1$, $x(t-2P) = X_2$, $x(t-P) = X_3$, $x(t) = X_4$ et la sortie $y = x(t+P)$. les entrées et la sortie sont reconstituées d'après la série chronologique des données recueillies $P(t) \quad t=0 \dots N$.

2- Comme le montre (la figure VI .1) , le traitement réalisé par un Système d'Inférence Flou est constitué de trois étapes : **la fuzzification, l'inférence et la défuzzification .**

- La fuzzification transforme les valeurs des quatre entrées numériques X_1, X_2, X_3, X_4 en partie floue. Pour notre application, on a opté pour l'utilisation de fonction triangulaire.
- L'inférence produit l'image de la partie floue issue de la fuzzification par la relation floue R généralement construite à partir de règles, pour notre cas , il s'agit de déterminer les sorties y_i^* de chaque règle ainsi que son degré d'activation w_i .
- La défuzzification transforme la partie floue issue de l'inférence en une valeur numérique y , de sortie.

Cette structure est utilisée dans tous les SIF quelle que soit l'application considérée (modélisation pour la prévision des précipitations). La mise en œuvre de ces trois étapes de traitement dépendent de la nature des sous-ensembles flous véhiculés à l'intérieur du système.

De même, la relation floue R , peut être une relation floue numérique (liant entre eux des nombres), comme il c'est le cas pour notre modèle :

$$f_i(x_1, x_2, \dots, x_r) = b_{i0} + b_{i1} x_1 + b_{i2} x_2 + \dots + b_{ir} x_r$$

une relation floue symbolique (liant entre eux des symboles), ou une relation mixte (liant entre eux des nombre et des symboles).

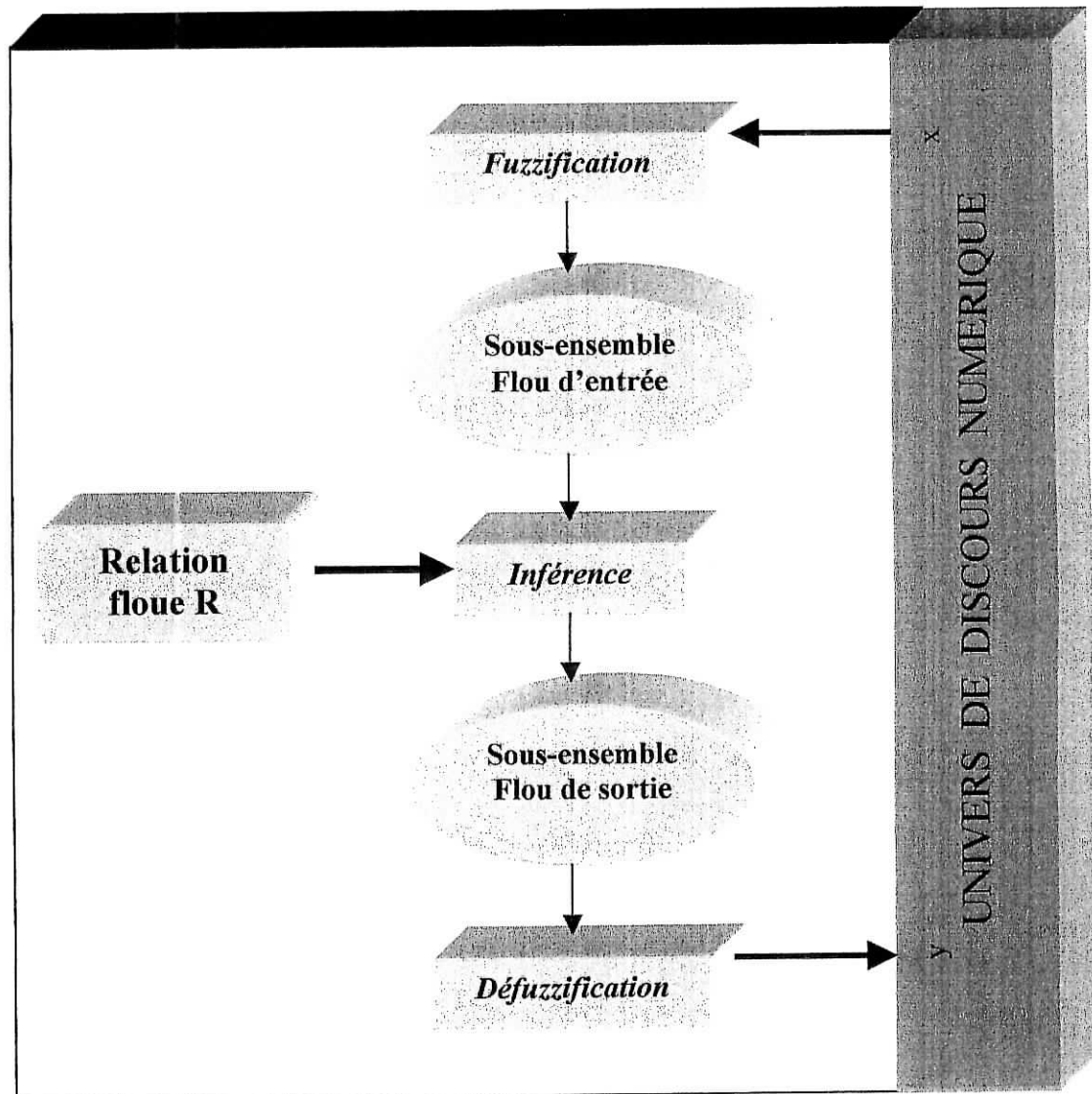


Figure VI . 1: Structure d'un Système d'Inférence Flou

VI . 3- différentes phases de la modélisation

Supposons qu'un quadruplet de données $(x_1^0, x_2^0, x_3^0, x_4^0)$ soit disponible. La sortie numérique y est obtenue par application successive d'une fuzzification, inférence et défuzzification.

VI . 3.1- fuzzification :

Les quatre entrées numériques x_1^0, x_2^0, x_3^0 et x_4^0 sont fuzzifiées indépendamment en utilisant les fonctions de description floue de forme triangulaire $D_L(X_1), D_L(X_2), D_L(X_3), D_L(X_4)$. On obtient alors les quatre sous-ensembles flous d'entrées :

$$E_1 = D_L(X_1)(x_1^0), E_2 = D_L(X_2)(x_2^0), E_3 = D_L(X_3)(x_3^0) \text{ et } E_4 = D_L(X_4)(x_4^0).$$

Par l'application des règles des normes, nous obtenons l'entrée floue composée sur l'espace $L(X_1) \times L(X_2) \times L(X_3) \times L(X_4)$:

$$\forall A_i \in L(X_1), \forall B_j \in L(X_2), \forall C_k \in L(X_3), \forall D_l \in L(X_4) :$$

$$\mu_E(A_i, B_j, C_k, D_l) = T(\mu_{D_L(X_1)}(x_1^0)^{(A_i)}, \mu_{D_L(X_2)}(x_2^0)^{(B_j)}, \mu_{D_L(X_3)}(x_3^0)^{(C_k)}, \mu_{D_L(X_4)}(x_4^0)^{(D_l)})$$

VI . 3.2 - Inférence :

Après avoir "fuzzifié" les variables d'entrée et de sortie, il faut établir les règles liant les entrées aux sorties. En effet, il ne faut pas perdre de vue le but final que nous poursuivons, et qui consiste à analyser la valeur des entrées du système pour déterminer la valeur de la sortie.

Soit la relation R floue sur $L(X_1) \times L(X_2) \times L(X_3) \times L(X_4) \times L(Y)$ induite par la base de règle (4), et qui représente une fonction linéaire reliant les quatre entrées à la sortie y , de tel sorte que le quadruplet de données d'entrée sera lié à la sortie y représentée par $X(t+p)$ de la forme (9) :

$$y_i = f_i(x_1, x_2, x_3, x_4) = b_{i0} + b_{i1} x_1 + b_{i2} x_2 + b_{i3} x_3 + b_{i4} x_4$$

Chaque sortie (y_i) pour chaque donnée d'entrée représente une règle de la forme (4) avec un degré d'activation w_i calculé par l'équation (6).

VI. 3.4 - Défuzzification

Comme nous l'avons vu précédemment, cette opération peut être effectuée par la méthode du centre de gravité, qui donne une agrégation des sorties de l'ensemble des règles par la moyenne des sorties individuelles :

$$y^* = \frac{\sum_i w_i \cdot y_i}{\sum_i w_i}, \text{ cette méthode est illustrée par la figure VI.2 :}$$

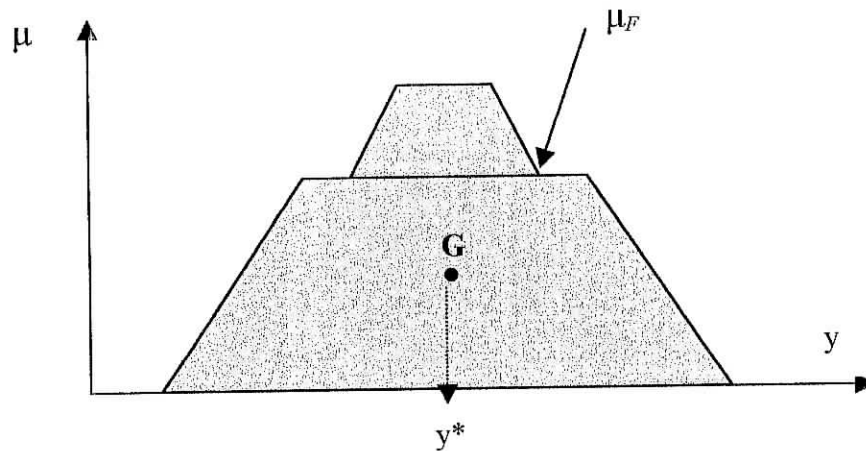


Figure VI.2 : Méthode du centre de gravité

VI. 3.5 - Tests de Performance :

Il existe plusieurs critères statistiques qui permettent d'apprécier quantitativement et qualitativement les résultats et la performance des modèles, et peuvent être utilisés comme des critères d'optimisation :

a) Les critères graphiques :

l'analyse graphique de n'importe quel modèle est indispensable et essentiel, cela est obtenu en portant sur un graphique les valeurs observées et les valeurs prévues sur une période de test.

b) les critères numériques :

- **Le critère de NASH** : introduit par NASH et Sutcliffe (1970), est défini par [A. Rahai]:

$$Ns = \left(1 - \frac{\sum_{i=1}^N (P_i - \hat{P}_i)^2}{\sum_{i=1}^N (P_i - \bar{P}_i)^2} \right) * 100.$$

Le critère de NASH peut être interprété comme étant la proportion de la valeur calculée par le modèle. Si $Ns = 100\%$, la prévision est parfaite, par contre si $Ns < 0$, cette dernière est mauvaise.

- **L'erreur quadratique moyenne** : $RMSE = \frac{1}{N} \sum_{i=1}^N \frac{1}{2} (P_i - \hat{P}_i)^2$.

Ce critère est souvent utilisé aussi bien pour un test de performance, que pour l'optimisation. Plus le RMSE est proche de zéro, plus l'écart entre la valeur observée et calculée diminue.

- **Moyenne observée** : $m_l = \frac{1}{N} \sum_{i=1}^N P_i$.

- **Moyenne prévue** : $m_l = \frac{1}{N} \sum_{i=1}^N \hat{P}_i$.

- **Coefficient adimensionnelle** $D = \frac{\hat{P}_{\max}}{P_{\max}}$. Si $D=1$ la prévision des pics est parfaite.

Tel que :

- \hat{P}_i : précipitation prévue.

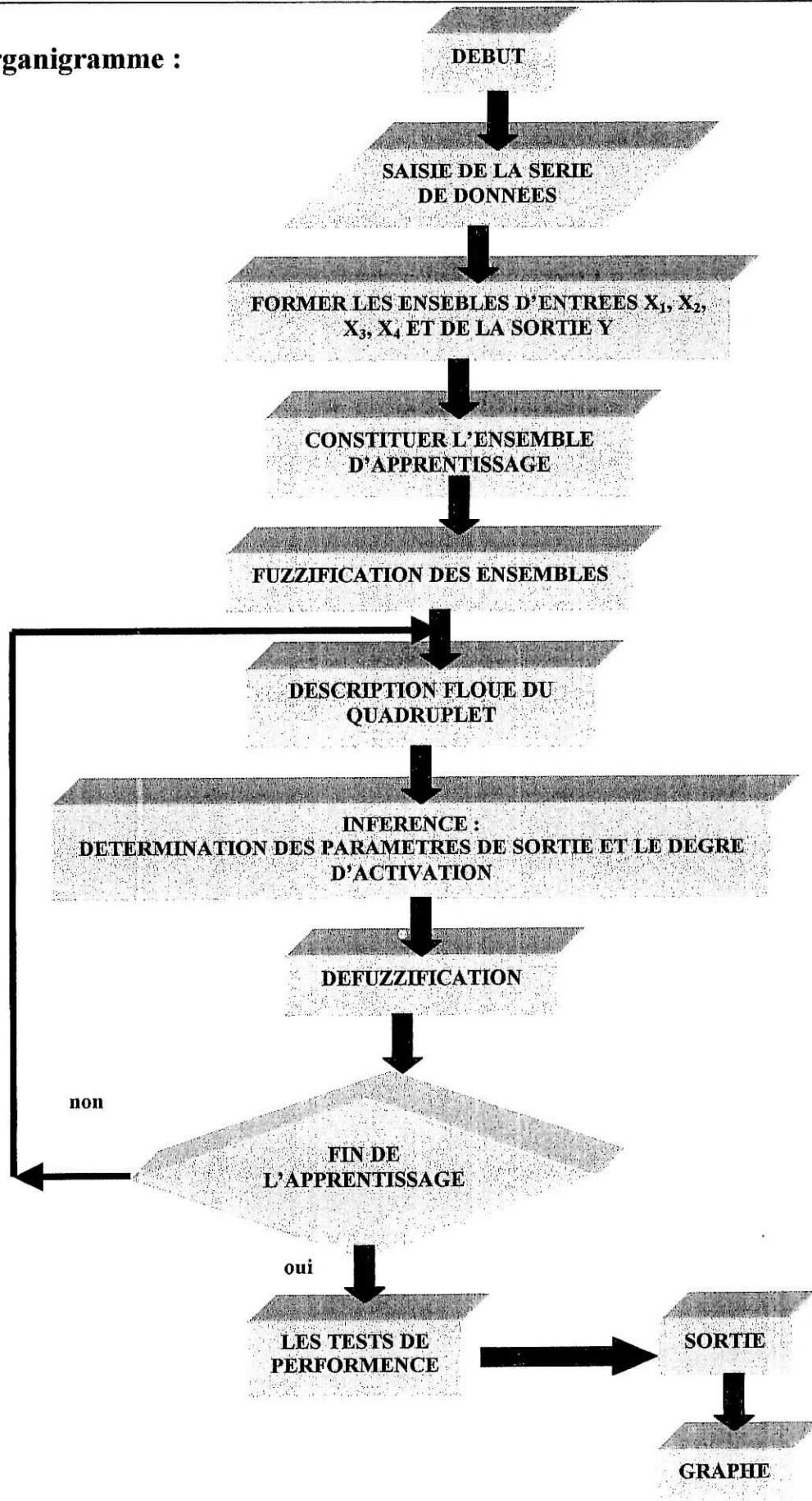
- P_i : précipitation observée.

- N : nombre de donnée.

- \hat{P}_{\max} : précipitation prévue maximal.

- P_{\max} : précipitation observée maximal.

VI. 4- Organigramme :



VI. 5 - Interprétation des résultats :

Avant de détailler l'interprétation des résultats obtenus, il est important de noter qu'un modèle flou généré par la méthode proposée par **Takagi-Sugeno-Kang** ne permet qu'un traitement élémentaire de l'extrapolation par simple prolongation du comportement du modèle.

1- 1^{ère} approche :

Dans le premier cas, le modèle est utilisé avec quatre entrées et une sortie, pour :

- $P=6$.
- Et une fuzzification de 4, 2, 2, 2 respectivement pour chaque entrée.

Ainsi le modèle s'écrit de la manière suivante : $[x(t-18), x(t-12), x(t-6), x(t) ; x(t+6)]$.

Avec l'idée, que pour prévoir le mois de janvier de l'année prochaine, on utilise comme valeurs d'entrées les mois de janvier et juin de l'année en cours et de l'année précédente.

:

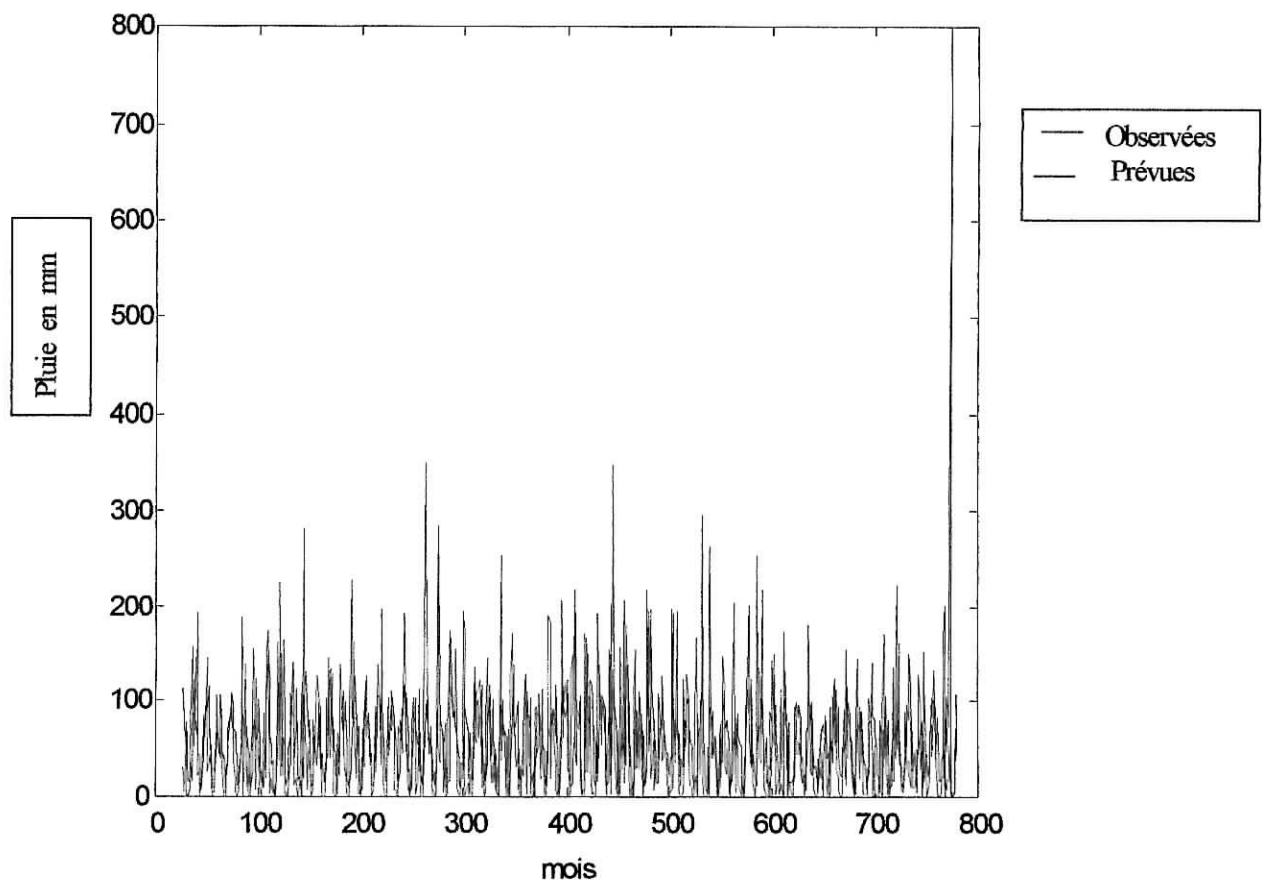


Figure VI.3 : Résultats de la prévision du modèle $P=6$

Les prévisions ont tendances à suivre la même évolution pendant le calage, tandis que pour la validation, les pics prévues dépassent de loin ceux observés.

Tableau VI.1 : Tests de performances pour $P=6$

	les tests de performance	modèle 6
calage	m1	57.53
	m2	57.89
	Ns	-2.59
	RMSE	60.80
	D	1.25
validation	m1	58.60
	m2	57.89
	Ns	-112.95
	RMSE	74.41
	D	0.32

D'après le tableau VI.1, le modèle semble bien représenter les moyennes pendant le calage et la validation avec une grande précision. Par contre, pour les tests de performances avec le critère de Nash, la modélisation dans les conditions d'apprentissage donne des résultats aberrants en validation et n'arrive pas à reproduire les pics, cela confirme que ce formalisme n'a aucune signification physique en hydrologie, car comme nous l'avons conçu, l'ensemble d'apprentissage ne donne par une réelle description du problème de la prévision (prévoir le mois de janvier de l'année prochaine avec les mois de janvier et juin de l'année en court et de la précédente).

Pour tenter de contourner le problème et améliorer les résultats de la prévision, nous avons reconstitué l'ensemble d'apprentissage pour le rendre en phase avec les exigences d'hydrologie (prendre les mois qui ont une forte corrélation entre eux).

2- 2^{ème} approche :

Le modèle est reformulé avec :

- $P=12$.
- 4, 2, 2, 2 symboles sur chaque entrée, tel que l'ensemble d'apprentissage s'écrit sous la forme : $[x(t-36), x(t-24), x(t-12), x(t) ; x(t+12)]$.

C'est à dire, la prévision du mois (**M**) prochain, nécessite l'introduction du mois (**M**) de cette année et celui des trois années précédentes figure VI.4.

Remarque : (**M**) peut représenter un des douze mois .

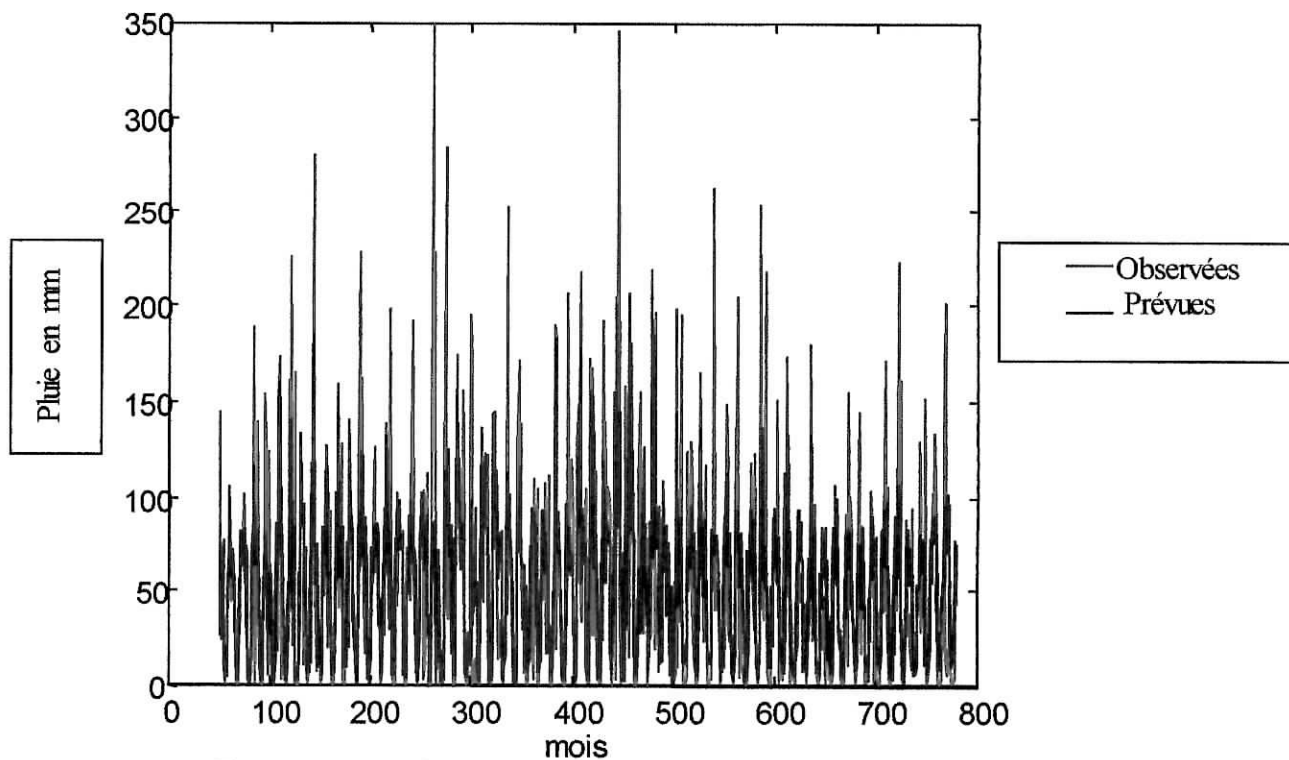


Figure VI . 4 : Résultats de la prévision du modèle P=12

Cette application nous a permis d'améliorer les résultats de la prévision, avec une évolution de la tendance plus au moins bien reproduite.

Tableau VI.2 : Tests de performances pour P=12

	les tests de performance	modèle 12
calage	m1	55.20
	m2	55.24
	Ns	42.45
	RMSE	43.55
	D	0.59
validation	m1	56.22
	m2	55.24
	Ns	5.41
	RMSE	53.88
	D	0.48

Le tableau VI.2 illustre une excellente reproduction des moyennes pour le calage et la validation.

Le critère de Nash est amélioré de façon considérable, ainsi nous avons pu atteindre une bonne performance du modèle mais sans toutefois arriver à reproduire les pics observés.

3- 3^{ème} approche :

Pour plus d'ajustement dans la prévision, il faut donc trouver le bon P "le bon ensemble d'apprentissage" qui explique mieux le phénomène. Alors, on a opté au partage de l'ensemble des données en deux sous-ensemble représentant la période sèche et la période humide.

La méthode consiste à calculer la moyenne de chaque mois ($M_i, i=1..12$) pendant les 66 ans observés, calculer la moyenne mensuelle interannuelle (A), représenter et diviser l'ensemble en deux parties

- l'ensemble des mois M_i supérieur à (A) représente la période humide.
- l'ensemble des mois M_i inférieurs à (A) représente la période sèche.

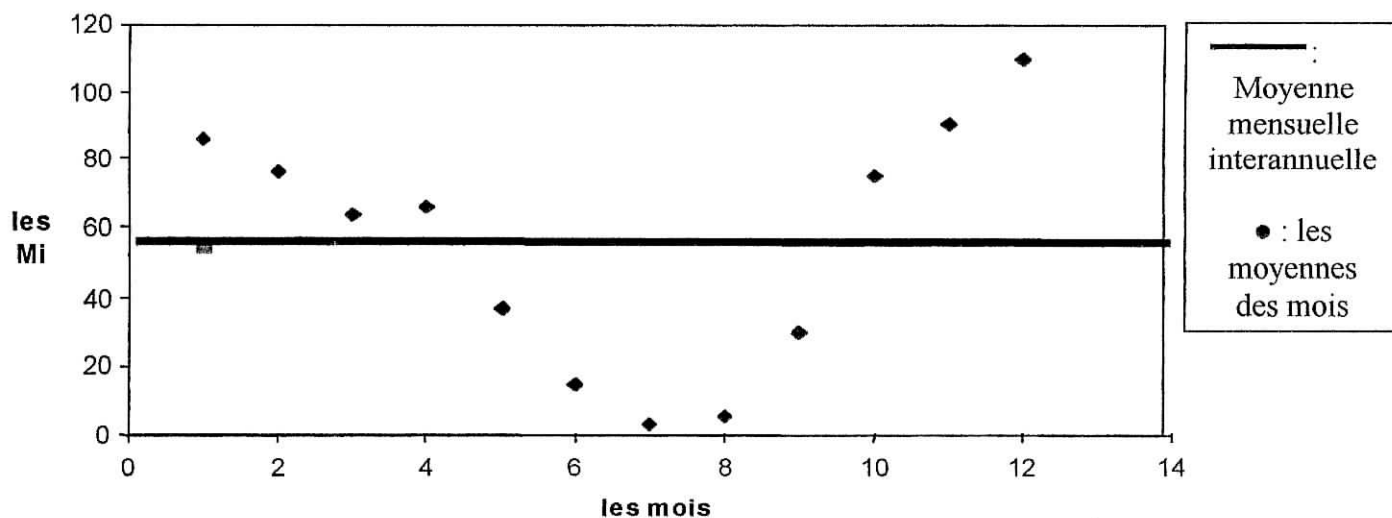


Figure VI.5.a : Représentation des périodes humide et sèche

Après la séparation des points par la droite représentant la précipitation moyenne mensuelle interannuelle, on perçoit sur la (figure VI.5.a) la formation de deux nuages de point, l'un au dessus et l'autre au dessous :

- Les points qui sont au dessus de la droite représentent les mois humides qui sont {octobre, novembre, décembre, janvier, février, mars, avril}.
- Alors que les points qui sont au dessous représentent la période sèche {mai, juin, juillet, août, septembre}.

Le modèle est configuré avec :

- $P=1$.
- 2, 2, 2, 2 symboles sur chaque entrée pour la période humide (figure VI. 5.b), et la période sèche (figure VI. 5.c) .

L'ensemble d'apprentissage est repris pour les mêmes quadruplets mais avec une autre formulation : $[x(t-3), x(t-2), x(t-1), x(t) ; x(t+1)]$ qui permet de respecter la logique de la succession des entrées.

Remarque : même si les points d'apprentissages sont analogues aux modèles précédents, le mécanisme de l'inférence flou, lui change de façon radicale, car le changement s'opère pendant l'étape de fuzzification des données d'entrées.

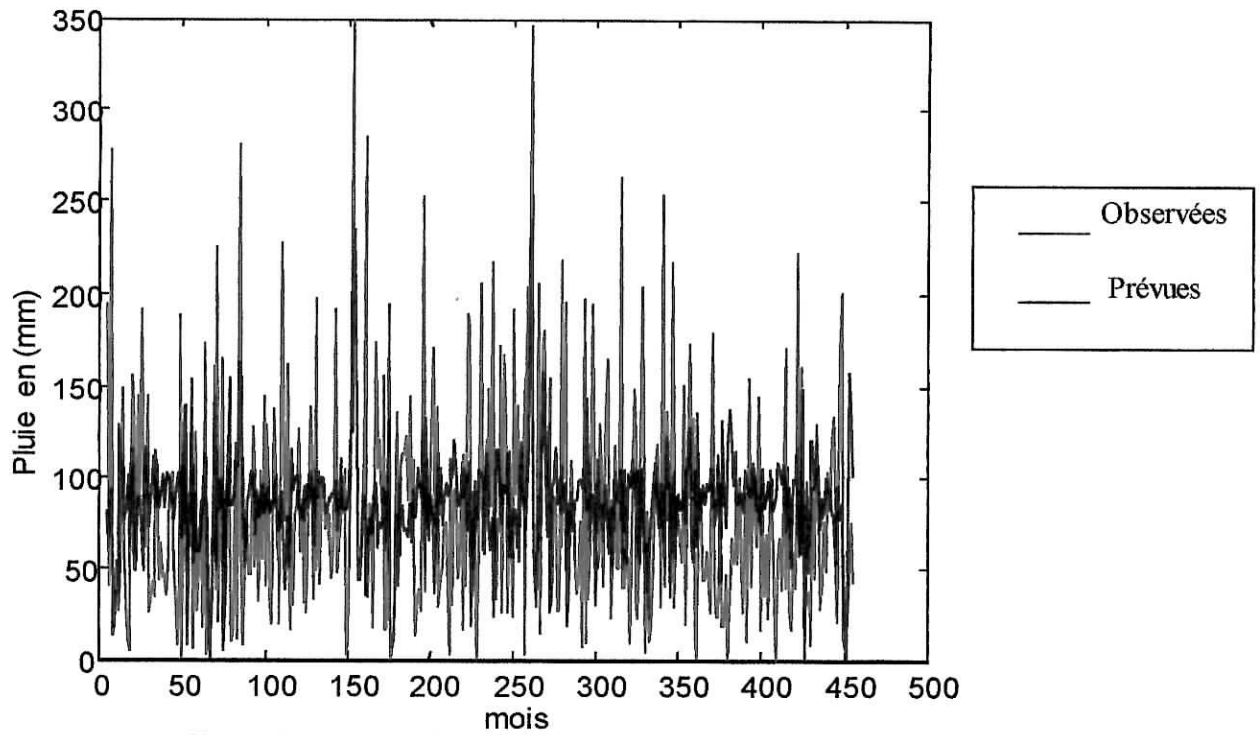


Figure VI.5. b : Prévion en période humide

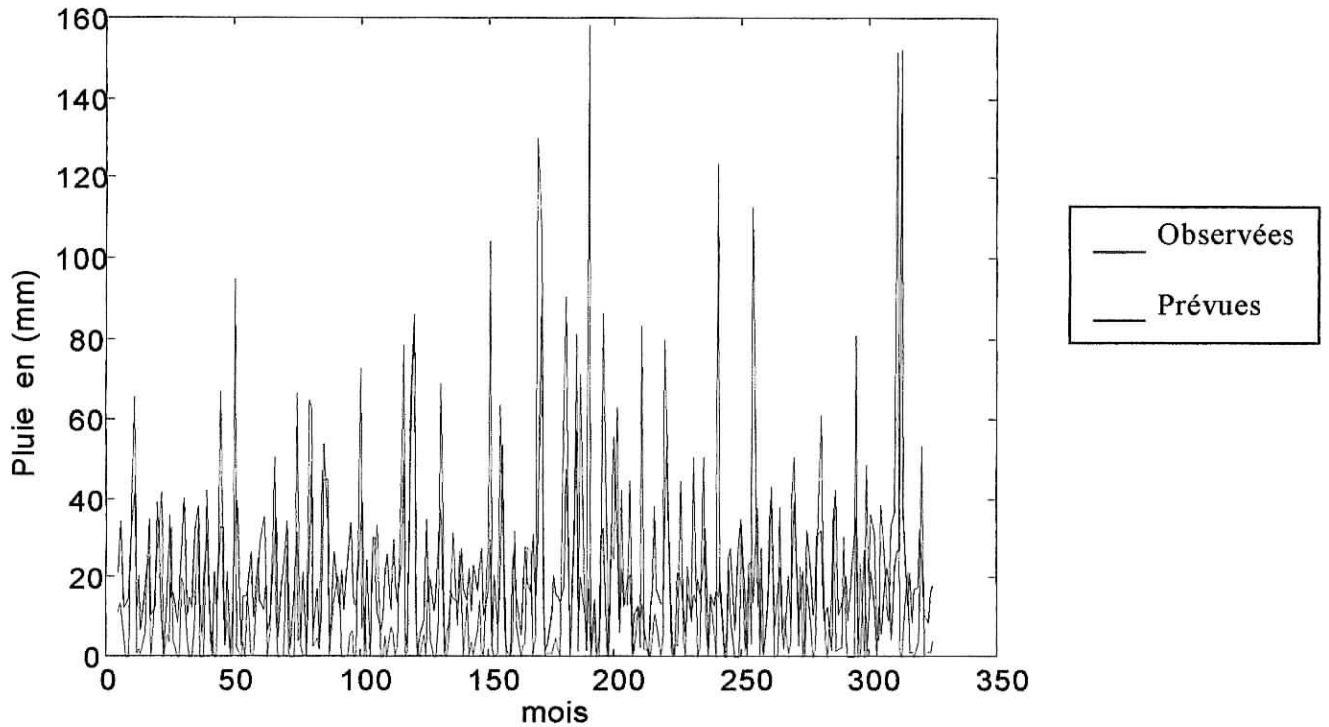


Figure VI.5. c : Prévion en période sèche

Il faut être prudent quand à l'interprétation des résultats des figures. En effet pour :

L'utilisation de cette méthode a apporté des résultats plus au moins meilleurs, et la tendance du phénomène a été généralement reproduit.

Tableau VI.3 : Tests de performances pour $P=12$

	les tests de performance	modèle humi	modèle sèch
calage	m1	85.10	18.44
	m2	84.45	18.49
	Ns	34.98	31.08
	RMSE	49.11	21.83
	D	0.70	0.56
validation	m1	82.29	18.62
	m2	84.45	18.49
	Ns	-61.41	-35.40
	RMSE	59.71	31.67
	D	0.86	1.00

- **La période humide :** Nous remarquons que le modèle arrive à reproduire des moyennes identiques à celles des observations pendant l'apprentissage et la validation, alors que le critère de Nash pour la validation est très loin de 100% ce qui démontre que la nature de la variable à étudiée est très difficile à reproduire surtout si celle-ci a une évolution ne permettant de suivre aucune loi (purement aléatoire).
- **La période sèche :** Les résultats confortent les remarques faites pour la période humide et montre que le modèle a perdu de sa performance comparé au second cas, cela est dû à l'écart très important entre les mois (voir la figure VI.5.a).

Le comportement du modèle permet pendant le calage de reproduire les observations, mais arrivé à la validation, les valeurs prévues sont parfois contradictoires avec celles observées. Les mois qui ont des écarts importants les uns par rapport aux autres font que le modèle perd de sa performance, ainsi donne des résultats erronés. Ce problème est propre à la nature aléatoire du phénomène météorologique (précipitations).

4- 4^{ème} approche :

Finalemment, après les différents types de changements opérés dans les modèles et les différents formalismes utilisés, la voie qui pouvait engendrer de meilleurs prévisions est peut être de reconstituer des ensembles d'apprentissage avec deux mois successives. Toutefois reconstituer cet ensemble, engendre une base de données de $2*66=132$ valeurs, ce qui nous oblige à utiliser que 2 symboles sur chaque entrée seulement.

Le modèle est appliqué avec :

- $P=1$.
- 2 symboles sur chaque entrée, $x(t-3)$, $x(t-2)$, $x(t-1)$, $x(t)$.

Remarque :

Nous avons considéré deux mois successives pour assurer une plus forte corrélation entre les données d'apprentissage .

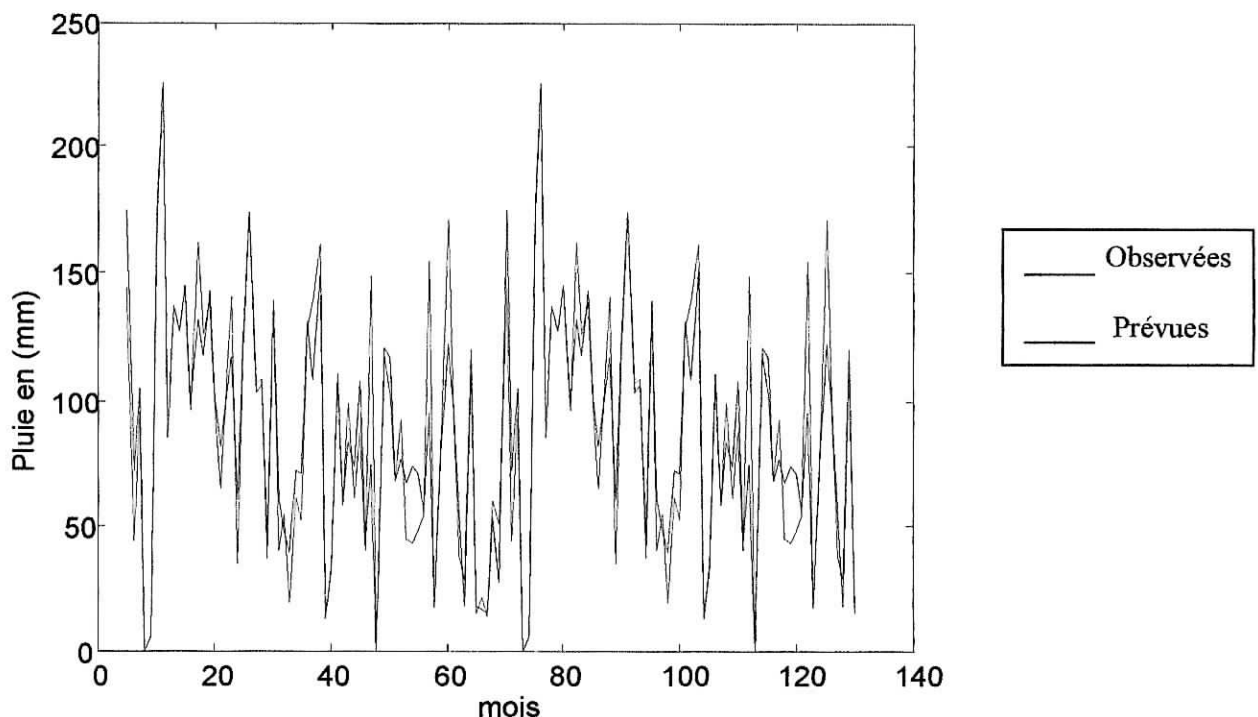


Figure VI.6.a : Prévision pour les mois janvier- février

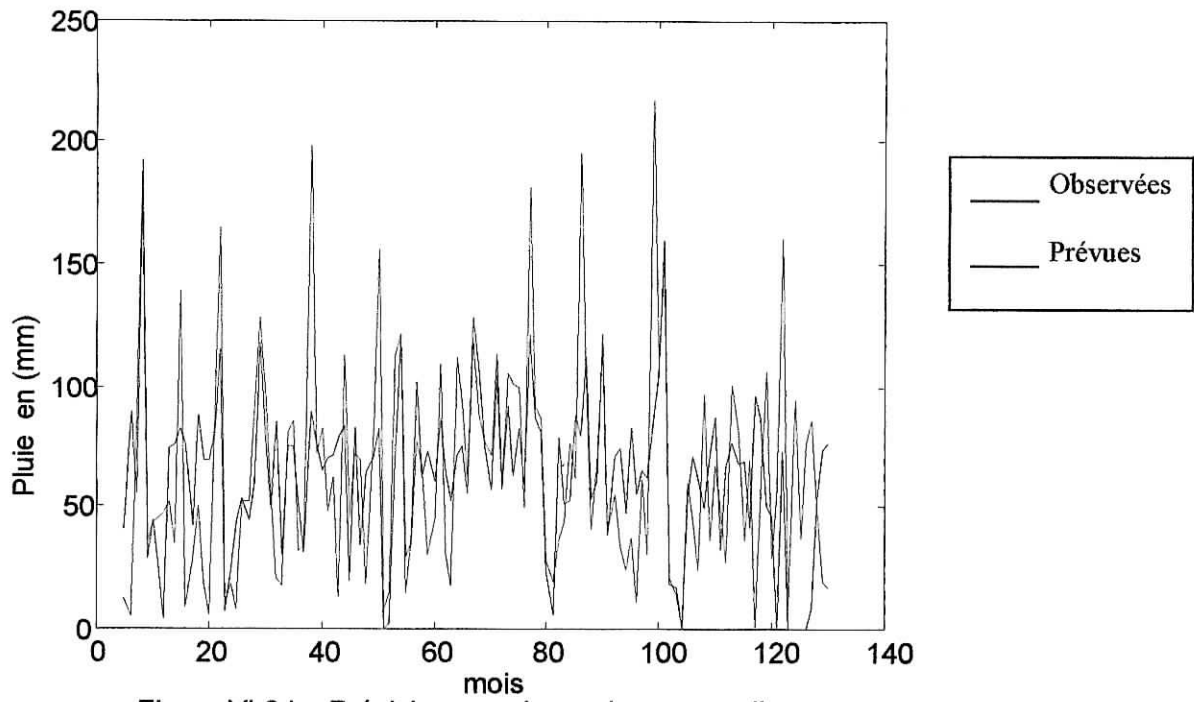


Figure VI.6.b : Prévion pour les mois mars-avril

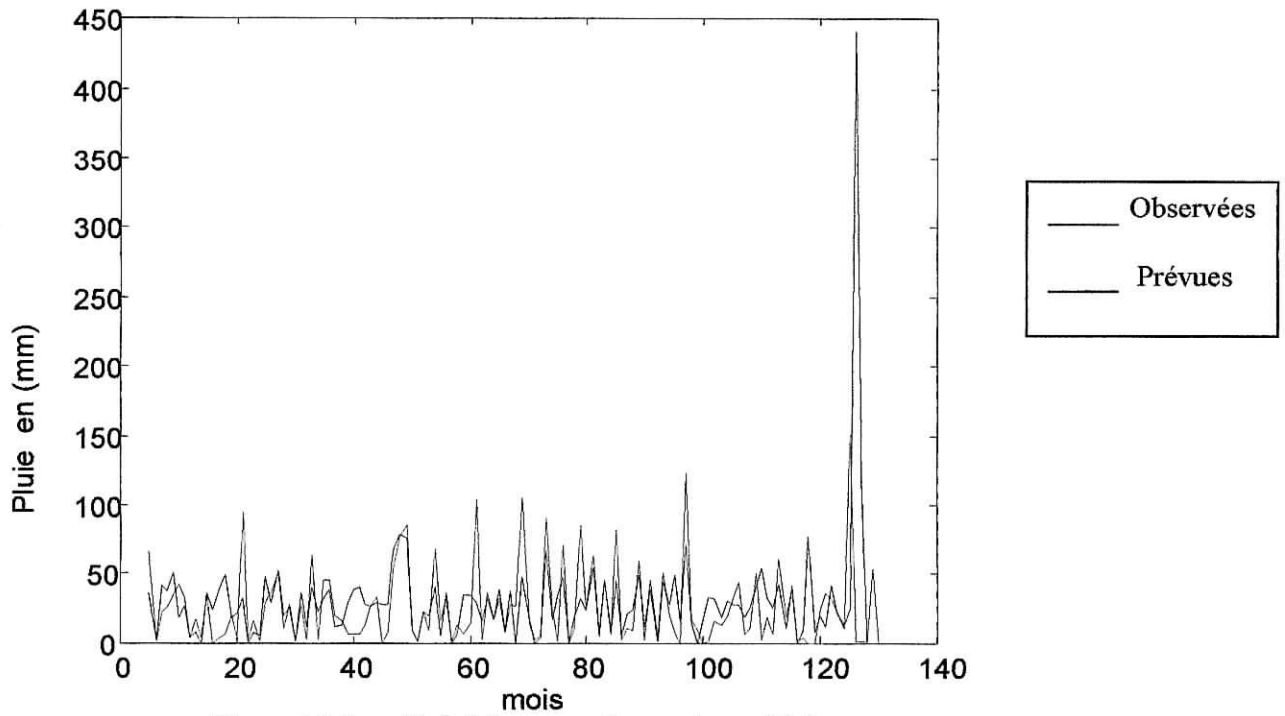


Figure VI.6.c : Prévion pour les mois mai-juin

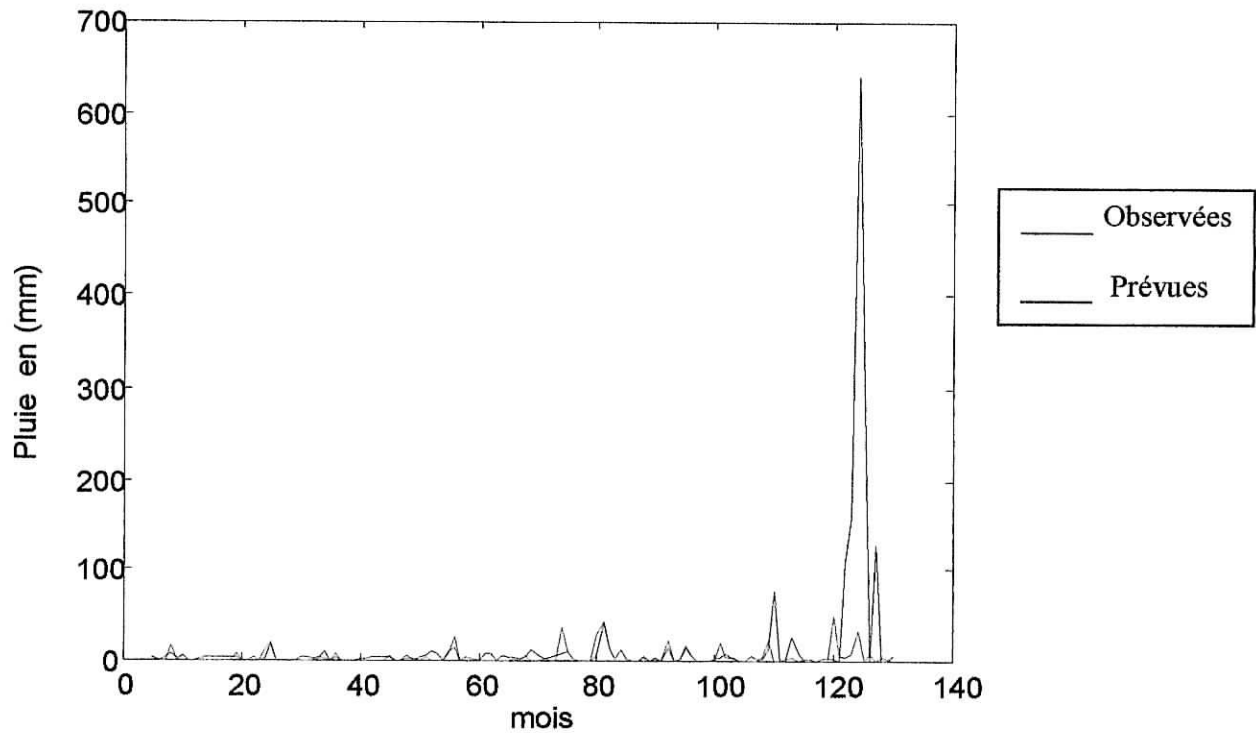


Figure VI.6.d : Prévion pour les mois juillet-août

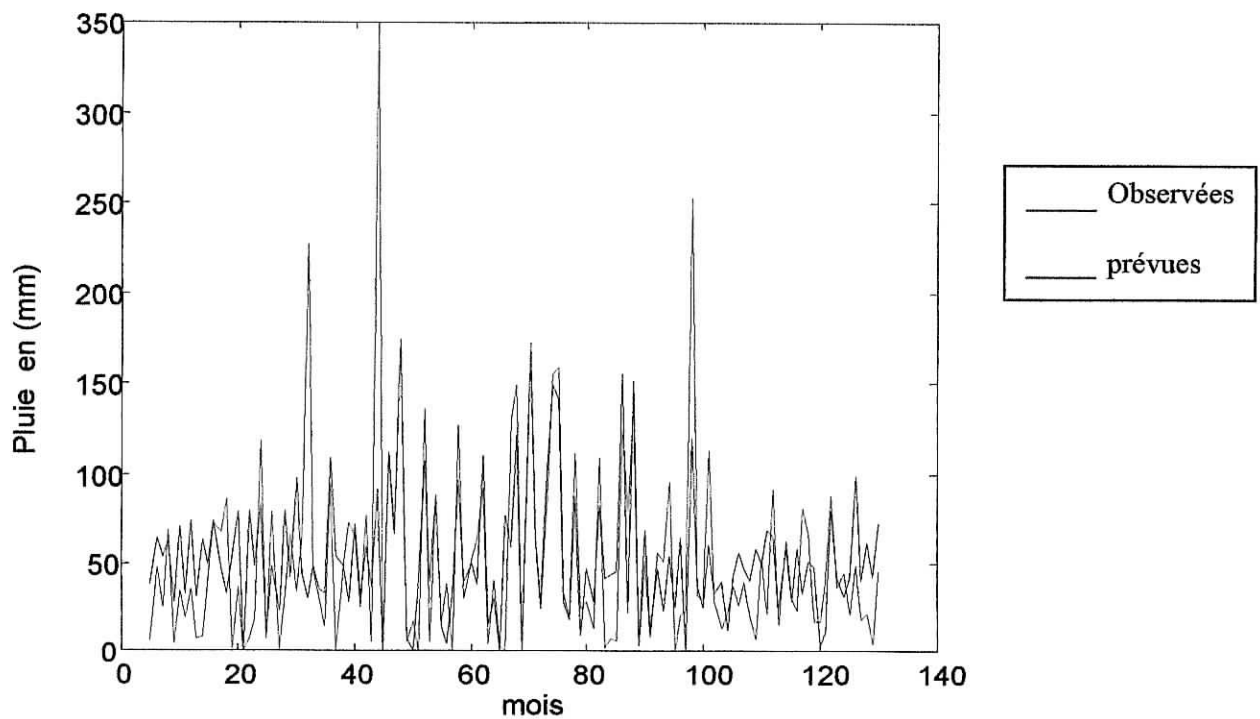


Figure VI.6.e : Prévion pour les mois septembre-octobre

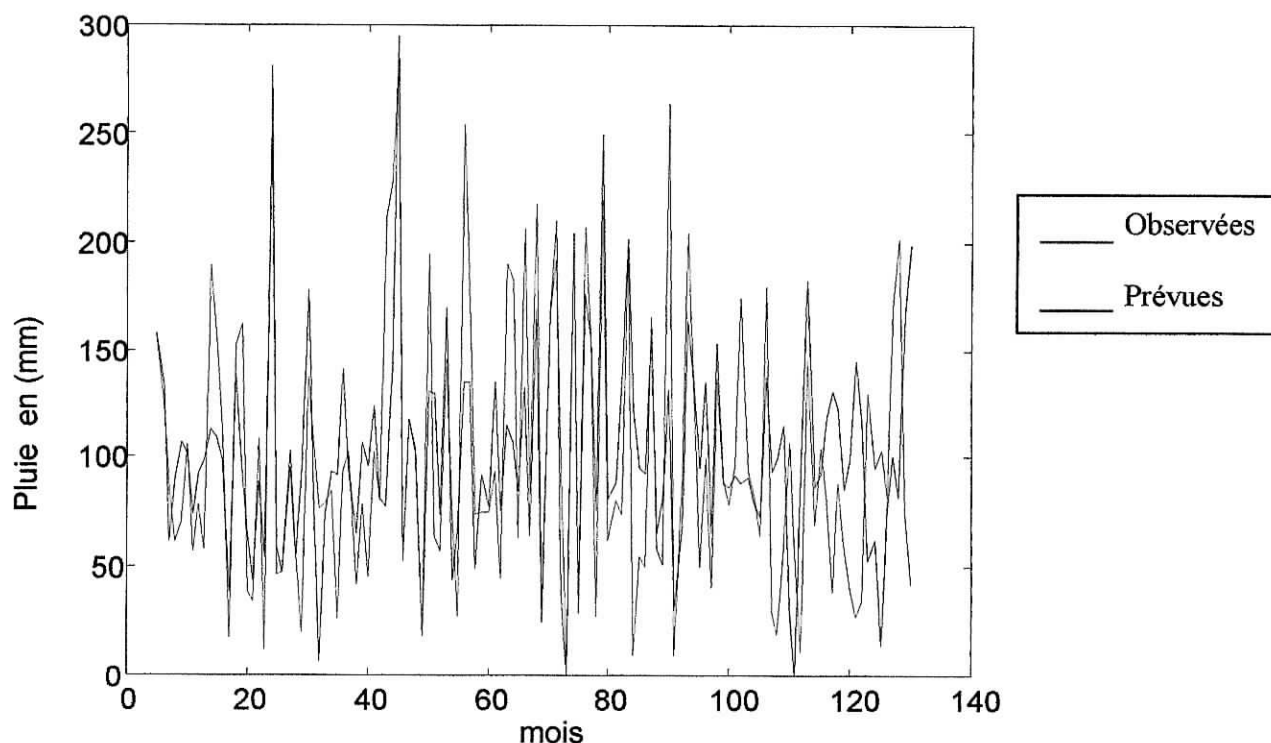


Figure VI.6.f : Prédiction pour les mois novembre- décembre

Tableau VI.4 : Les tests de performance pour la 4^{ème} approche

	les tests de performance	Modèle j-f	modèle ma-a	modèle m-j	modèle j-j-a	modèle s-o	modèle n-d
calage	m1	90.3808	65.87	26.76	4.15	54.59	103.21
	m2	87.1298	245.29	30.18	14.84	53.57	102.37
	Ns	89.7972	45.25	45.83	52.05	42.88	58.08
	RMSE	16.7559	35.30	20.37	5.48	44.71	43.29
	D	0.9999	0.83	0.64	0.95	0.49	1.03
validation	m1	70.8747	47.68	47.28	68.26	48.47	98.16
	m2	87.1298	245.29	30.18	14.84	53.57	102.37
	Ns	68.2841	-94.83	-885.52	-1.67e+004	-22.66	-154.54
	RMSE	23.7898	47.19	95.41	139.0925	24.69	67.92
	D	0.7145	0.59	2.91	13.1410	1.07	0.98

On remarque que les résultats de la prévision dans les six cas se divisent en deux groupes, qui sont clairement visible par leur performance .

Les prévisions des pics et les mouvements tendanciels sont très bien reproduits (Figure VI.6.a) , des tests de performance assez bons (modèle j-f) avec une moyenne parfaitement reproduite et un critère de Nash aussi bon pour le calage que pour la validation.

Ceci implique une performance de prévision supérieure à toutes les autres. L'explication, et que les mois de janvier et février ont des valeurs de précipitation qui suivent une certaine logique dans leur évolution chronologique et sont rarement nulles (cas exceptionnelle : deux fois en 66 ans) .

Un second groupe, avec des calages assez bons, vérifiés aussi bien pour la moyenne que pour le critère de Nash (à examiner dans le tableau VI.4) , néanmoins les prévisions des validations sont parfois insensées comme c'est le cas pour le modèle mai-juin et modèle juillet-août qui donnent des pics pendant la validation erronés.

L'apparition de l'anomalie pendant la validation est une conséquence directe de la difficulté du problème lié surtout au modèle de prévision de précipitation . Les données étant purement aléatoire et ne suivent aucune loi. Le modèle conçu pour un apprentissage de données chaotiques (composition de période) lui permettant de faire des extrapolations , de ce fait , il reproduit des valeurs pendant la validation qui ne peuvent correspondre à celles observées.

les tests de performance :

Le tableau regroupe les résultats de performance obtenus pour chaque modèle :

- *Modèle 6 : prévision avec $P=6$.*
- *Modèle 12 : prévision avec $P=12$.*
- *Modèle humi : prévision pour la période humide.*
- *Modèle sèch : prévision pour la période sèche.*
- *Modèle j-f : prévision pour les mois janvier-février.*
- *Modèle ma-a : prévision pour les mois mars-avril.*
- *Modèle m-j : prévision pour les mois mai-juin.*
- *Modèle j-a : prévision pour les mois juillet-août.*
- *Modèle s-o : prévision pour les mois septembre-octobre.*
- *Modèle n-d : prévision pour les mois novembre-décembre.*

Tableau VI.4 : Les tests de performance

	les tests de performance	modèle 6	modèle 12	modèle humi	modèle sèch	Modèle j-f	modèle ma-a
calage	m1	57.5367	55.2041	85.1095	18.4497	90.3808	65.8727
	m2	57.8983	55.2487	84.4546	18.4964	87.1298	245.2901
	Ns	-2.5972	42.4505	34.9860	31.0848	89.7972	45.2517
	RMSE	60.8034	43.5586	49.1128	21.8335	16.7559	35.3006
	D	1.2560	0.5926	0.7030	0.5676	0.9999	0.8309
validation	m1	58.6074	56.2246	82.2968	18.6259	70.8747	47.6838
	m2	57.8983	55.2487	84.4546	18.4964	87.1298	245.2901
	Ns	-112.9575	5.4162	-61.4154	-35.4073	68.2841	-94.8345
	RMSE	74.4149	53.8812	59.7155	31.6726	23.7898	47.1950
	D	0.3293	0.4862	0.8630	1.0020	0.7145	0.5979

	les tests de performance	modèle m-j	modèle j-a	modèle s-o	modèle n-d
calage	m1	26.7634	4.1579	54.5993	103.2143
	m2	30.1835	14.8425	53.5784	102.3728
	Ns	45.8338	52.0595	42.8825	58.0895
	RMSE	20.3729	5.4848	44.7101	43.2988
	D	0.6444	0.9595	0.4946	1.0376
validation	m1	47.2839	68.2656	48.4735	98.1652
	m2	30.1835	14.8425	53.5784	102.3728
	Ns	-885.5236	-1.6789e+004	-22.6694	-154.5442
	RMSE	95.4196	139.0925	24.6920	67.9259
	D	2.9108	13.1410	1.0773	0.9844

VI.6 - Conclusion :

Nous avons utilisé les modèle de *Takagi-Sugeno-Kang* avec une typologie SIF "Système d'Inférence Flou" pour la prévision d'une série de précipitation au pas de temps mensuel.

La méthode proposé s'est montré capable de modéliser un système dynamique, une fois le type et la taille du modèle choisis. Du point de vue résultats numériques, cette méthode a une performance satisfaisante.

Il est tout à fait possible d'améliorer cette performance, en modifiant la stratégie d'apprentissage, le modèle pourra alors prendre en considération cette variabilité des données et faire en sorte qu'il soit adaptatif pour chaque période d'apprentissage distincte.

Cependant l'approche est basée sur la connaissance et elle n'est pas suffisante, car le modèle est entièrement construit par un expert qui peut se tromper, mal exprimer ses connaissances, ou encore ne pas connaître tous les états du système étudié.

Conclusion générale

CONCLUSION GENERALE

En introduisant la théorie des sous-ensembles flous, Zadeh a offert un outil formel pour la modélisation des systèmes complexes, pour lesquels on ne dispose que d'une spécification approximative ou imprécise. Le but d'un modèle est de capturer la relation entre les entrées et les sorties d'un système. A l'encontre d'un modèle conventionnel qui décrit cette relation par une loi mathématique, le modèle flou la décrit linguistiquement.

Les modèles flous ont deux propriétés essentielles :

- Le traitement se fait au niveau linguistique. En fait, ces modèles sont conçus pour manipuler des valeurs linguistiques comme c'est le cas chez l'homme. Ceci est devenu possible grâce aux variables linguistiques et à la représentation des significations des valeurs linguistiques par des sous-ensembles flous.*
- Ils sont capables de représenter l'imprécision et l'incertitude d'un expert humain. Cette propriété est particulièrement intéressante, car les modèles flous sont souvent inspirés de la connaissance humaine. Il est donc nécessaire d'inclure l'imprécision et l'incertitude que contient cette connaissance dans le modèle flou.*

Dans ce travail, nous nous sommes intéressés à l'étude des différentes approches de modélisation floue. Nous avons tout d'abord exposé des notions de base sur les variables linguistiques, puis nous avons introduit la typologie des Systèmes d'Inférence Flous. Par la suite, nous avons identifié trois classes essentielles de modèles flous dans la littérature : les modèles linguistiques ou de Mamdani, les modèles relationnels et les modèles de Takagi-Sugeno-Kang.

Le développement des modèles de Takagi-Sugeno-Kang à partir de la connaissance d'un expert est difficilement envisageable. Ceci est dû à la forme fonctionnelle des conclusions des règles. De plus les modèles flous conçus par un expert ne sont généralement pas optimaux d'un point de vue performance numérique pour la prévision des séries aléatoires. Au cours des vingt dernières années, des méthodes d'identification de modèles flous à partir de données numériques se sont alors développées, et l'approche de la modélisation floue a gagné de plus en plus d'espace dans les revues

spécialisées. Malgré cela, l'identification d'un modèle flou reste un problème mal défini, et les limites entre identification de structure de paramètres ne sont pas clairement spécifiées. Par exemple, le choix des variables d'entrée du modèle fait partie de l'identification de structure pour certains auteurs.

L'apport principal de cette thèse consiste en l'utilisation de la logique floue et les systèmes d'inférence floue pour la prévision des précipitations au pas de temps mensuel:

- La prévision des précipitations se confond avec la modélisation des systèmes dynamiques représentés par un modèle flou de Takagi-Sugeno-Kang .*
- La nature linéaire de la fonction représentant la relation entrée-sortie n'est pas toujours représentative dans notre application. On peut le constater dans les résultats des modèles mai-juin et juillet août (période sèche, où les valeurs nulles sont importantes et faussent les résultats pendant la validation) .*
- Le changement de la taille et la forme des ensembles d'apprentissages influent sur les résultats de prédiction de façon très importante.*
- Le travail effectué avec des variables aléatoires pour la prévision est unique , car la conception des modèles par les experts esté fondé sur des données différentes de celles étudiées. En effet, les variables utilisées dans leur modélisation s'avèrent de nature chaotique, qui signifie que c'est une composition cyclique ce qui n'est pas le cas chez nous.*

En conclusion, la modélisation floue nécessite un expert ayant des connaissances très développés dans ce domaine. La prévision des précipitations avec la logique floue se résume à un système dynamique SISO qui peut être représenté par un modèle d'un processus dynamique MISO. Les résultats comme première expérience dans ce domaine sont satisfaisantes, seule la nature de la variable étudiée pendant la période sèche donnent des prévisions erronés, néanmoins , si on considère que pendant la période sèche, les valeurs sont quasi-nulles et par conséquent, il n'y a pas lieu de faire de la prévision, le modèle sera parfait dans l'ensemble.

Perspectives :

Ce travail ouvre une voie à l'application de la logique floue par les chercheurs hydrologues qui pourraient d'une part le compléter et d'autre part, en tirer profit.

Comme nous l'avons vu dans les applications , le problème essentiel de la méthode proposée est la difficulté à reproduire une série purement aléatoire. Nous proposons des modifications qui peuvent améliorer la performance .

- *Le type de fonction linéaire utilisée dans les modèles de Takagi-Sugeno-Kang pour lier les entrées à la sortie peut être changé de façon qu'elle soit plus représentative de l'évolution temporelle des précipitation, la structure NARX(NonLinear AutoRegressive inputs) peut représenter ce type de relation .*
- *L' apprentissage du modèle flou demande un échantillon très grand. Il faut donc une base de données plus importante.*
- *Apporter des corrections à la structure du modèle. Le comportement sera alors adaptatif à tout changement d'évolutions temporelle des variables étudiées.*

Référence bibliographique

[Babuska 95a]

R. Babuska, H.B Verbruggen, "A new identification method for linguistic fuzzy models", *Proc. of FUZZ IEEE '95, Yokohama, Japan, March 20-24, 1995*, pp. 905-912.

[Bon 84]

A. Bonnet, "L'intelligence artificielle promesses et réalités", 1984, inter édition.

[Box 70]

G.E. Box, G.M. Jenkins, *Time Series Analysis, Forecasting and control* (Holden Day, San Fransisco, CA, 1970).

[Curry 44]

H.B. Curry, "The method of Steepest Descent for nonlinear Minimization Problems", *Quart. J.Appl. Math*, 2,1944,pp.258-261

[Driankov 93]

D. Driankov, H. Hellendoorn & M. Reinfrank, *An introduction to Fuzzy Control*, Springer-Verlag,1993.

[Dubois 84]

D. Dubois and H. Prade, "A theorem on implication function defined from triangular norms", *Busfal N:18, printemps 1984*, p 33-41.

[Dubois 88]

D. Dubois and H. Prade, *Théorie des possibilités: application à la représentation des Connaissances en informatique*, Masson, 1988, Paris.

[Dubois 88a]

D. Dubois et H. Prade, "Théorie des possibilités", 1988.

[Foulloy 95]

L. Foulloy, S. Galichet, "Typology of fuzzy controllers", in *Theoretical Aspect of Fuzzy*

[Jang 95]

J.-S. Jang, C.-T. Sun, "Neuro-Fuzzy Modelling and Control", *Proceeding of the IEEE*, vol. 83, No.3, pp. 378-406, March 1995.

[Jang 93a]

J.S.R. Jang, "ANFIS: Adaptive-Network-Based Fuzzy Inference System", *IEEE Trans. On System, Man, and Cybernetics*, Vol .23, No .3, pp. 665-685, May 93.

[Jang 92]

J.-S. Jang, *Neuro-Fuzzy Modeling: Architecture, Analyses and Applications*, PHD thesis, Dep. Of Electrical Engineering and Computer Sciences, Univ. of California, Berkeley, July 1992.

[Kosko 92b]

B. Kosko, "Fuzzy systems as universal approximators, " *Proc. of Fuzz-IEEE'92, San Diego, USA, Mar. 1992, pp. 1153-1162.*

[Lee 94]

Y.C. Lee, C. Hwang, Y.P. Shih, " A combined Approach to Fuzzy Model Identification" , *IEEE Trans. On systems, Man and Cybernetics, Vol SMC 24 , No.5, May 94, pp 736-744.*

[Lin 95]

Y. Lin and G. A. Cunningham , " A New Approach to Fuzzy-Neural Systems Modeling" , *IEEE Trans. on Fuzzy Systems, Vol. 3, No. 2, May 1995 , pp. 190-198.*

[Mamdani 76]

E.H. Mamdani , v *Advances in the linguistic synthesis of fuzzy controllers"*, *Int. Journal Of Man-Machines Studies, 8, pp. 669-678, 1975.*

[Mamdani 75]

E.H. Mamdani , S. Assilian , " An experiment in linguistic synthesis with a fuzzy logic Controller " , *Int. journal of Man-Machines Studies, 7, pp.1-13, 1975.*

[Mamdani 74]

E.H. Mamdani , " Application of fuzzy algorithms for control of simple dynamic plan", *Proc. of the IEE. Control & Science, Vol. 121, No.12, Decembre 1974, pp.158-1588*

[Nguyen 93]

H.T. Nguyen, V. Kreinovich, "On approximation of controls by fuzzy systems" , in *Proc. 5th IFSA World Congress, Seoul, Korea, July 93, vol.II, pp. 1414-1417*

[Nomura 92]

H. Nomura, I. Hayashi, N. Wakami, "A Learning Method of Fuzzy Inference Rules by Descent Method", *Proc. of Fuzz-IEEE'92, San Diego, USA, Mar. 1992. pp. 203-210.*

[Palm 94]

R. Palm, "Processing of signal distributions in control loops", *Proc. EUFIT'95, Sep 1994, Aachen, Germany, pp.310-321.*

[Pedrycz 89]

W. Pedrycz, *Fuzzy Control and Fuzzy Systems, Research Studies Press LTD, Taunton, Somerset, England. 1989.*

[Ragot 95]

J. Ragot, K. Adjallah, "Identification of ARX fuzzy models", *Int. Workshop on Fuzzy Technologies and in Automation and Intelligent Systems Fuzzy Duisburg '95, Duisberg, Germany, July 17 1995.*

[Rondeau 95]

L. Rondeau, R. Ruelas, E. Levrat, M. Lamotte, "Fuzzy Rules Learning Method", *Porc. EUFIT'95,28-31 August 1995, Aachen, Germany, pp. 515-519.*

[Shafer 76]

G. Shafer , "A mathematical theory of evidence in medicine", *Mathematical Biosciences* Vol 23, 1976 , p 351-379.

[Sanchez 76]

E. Sanchez, "Resolution of composite relational equations", *Information & control* 30, 1976, pp 38-48.

[Sudkamp 94]

T. Sudkamp, R.J. Hammell II, " Interpolation, Completion, and Learning Fuzzy RulesV , *IEEE Trans. On Systems, Man, and Cybernetics*, vol. 24, No.2, Feb.94.

[Sugeno 86]

M. Sugeno and G.T. Kang, "Fuzzy modelling and control of multiplayer incinerator" , *Fuzzy Sets and Systems*, vol. 18, pp. 329, 1986.

[Takagi 83]

T. Takagi, M. Sugeno, "Fuzzy Identification of Systems and Its applications to Modeling and Control", *IEEE Trans. On Systems, Man, and Cybernetics*, vol. 15, NO. 1, Feb.85.

[Takagi 83]

T. Takagi, M. Sugeno, "Derivation of fuzzy control rules from human operator's control actions" , *Proc of the IFAC on Fuzzy information*, pp 55-60, Marseille, France, 1983.

[Takagi 81]

H ; Takagi and I. Hayashi, "NN-driven fuzzy reasoning " , *Int . J. of approximate reasoning*, 5(3), 1991, pp.; 191-212.

[Tong 80]

R. M. Tong, "The evaluation of fuzzy models derived from experimental data", *Fuzzy Set and Systems*, 4(1980), pp.1-12.

[Walter 94]

E. Walter, L. Pronzato, *Identification de modèles paramétriques à partir de données expérimentaux*, Masson, Paris 1994.

[Wang 92b]

L.-X. Wang, "Fuzzy systems are universal approximators" , in *Proc. 1st IEEE Conference on Fuzzy Systems*, 1992, San Diego, pp.1163-1169.

[Yager 93]

R.R. Yager, D.P. Filev, "Unified Structure and Parameter Identification of Fuzzy Models" , *IEEE Trans. On Systems, Man, and Cybernetics*, vol. 23, No.4, July 93.

[Yager 84]

R.R. Yager, "Approximate reasoning as a basis for Rules-Based Expert Systems" , *IEEE Trans. On Systems, Man, and Cybernetics*, vol. SMC-14, No.4, July/August 1984, pp. 636-643.

[Yoshinari 93]

Y. Yoshinari, W. Pedrycz, K. Hirota, "Construction of fuzzy models through clustering techniques", *Fuzzy Sets and Systems*, Vol 54, 1993, pp. 157-165.

[Zadeh 65]

Zadeh, L.A "Fuzzy sets". *Information control* , vol 8, 1965, pp 338-353.

[Zadeh 65]

Zadeh, L .A ., "Quantitative fuzzy semantics", *Information Sciences*, vol.3, 1971, pp. 159-176

[Zadeh 73]

Zadeh, L .A ., "Outline of a New Approach to Analysis of Complex Sytems and Decision Process ", *IEEE Trans. On Systems, Man, and Cybernetics*, 3(1)(1973), pp. 28-44.

[Zadeh 73]

Zadeh, L .A ., "Quantitative fuzzy semantics Zadeh, L .A ", *Information Sciences*, vol. 3, 1971, pp. 159-176.

[Zadeh 78]

Zadeh, "Fuzzy sets as a basis for a theory possibility ", *fuzzy sets and systems*, 1, 1978, p 3-28.

[Zadeh 78a]

Zadeh L. A , "pruf-a meaning representation languages" , *international .J. Man-machine studies* vol 10, 1978, p 395-460.

B. Bouchon-Meunier , R.R. Yager and L.A. Zadeh , "Uncertainty in Knowledge Bses", IPMU '90 Paris, France, july 2-6 1990.

D. Dubois et H.Prade, "THEORIE DES POSSIBILITES", application à la représentation des connaissances en informatique, 1985.

D. Dubois , H.Prade, Y. Kuvisutto et J.P. Ceccato, pour le Laboratoire associe CNRS, "Bulletin pour les Sous Ensembles Flous et leurs Applications", Printemps 1985.

A.Kaufmann , Prefave du Pr. L.A. Zadeh, "INTRODUCTION A LA THEORIE DES SOUS-ENSEMBLES FLOUS ", tome 1, éléments théoriques de base à l'usage des ingenieurs, 1973.

Remenieras, G. "L'hydrologie de L'ingénieur", Edition Eyrolles, 1972. pp 55-69.

Mr. Tounsi Mohamed et Melle Animaer Djamila "L'UTILISATION DE LA LOGIQUE FLOUE POUR L'EVALUATION ET LA CLASSIFICATION DES PROSPECTS D'HYDROCARBURES", USTHB, 1995.

Haouchine Zahida et Saker Farah, " ETUDE DE LA THEORIE DES SOUS ENSEMBLES FLOUS CONCEPTION ET REALISATION D'UN MONITEUR FLOU, USTHB, 1994.

Mr. K. Belgaid et Mr. N. Tebbal " Etude et Prévision du Régime Pluviométrique sur le Littoral Algérien", ENP, 2000.

A. Rahai " Prévission de la pluviométrie par combinaison de l'ACP et des CHAINES DE MARKOV", ENP, 2002.

Jean .A. Vergnes : " Conférence, journée mondiale de l'eau, 2002 ".

OcCC : "Organe consultatif en matière de recherche sur le climat et les changements climatiques", Berne, décembre 1998.

Centre de recherche en calcul appliqué (Le CERCA) , "Rapport annuel 1998-1999 - Aspects scientifiques ".

[Michaël Aupetit] Réseaux de neurones et traitement de données 2002.

[Bobee],

THE GAMMA FAMILY AND DERIVED DISTRIBUTION APPLIED IN HYDROLOGY. Water resource application, Colorado, USA, pp 50-69, 1999.