

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université des Sciences et de la Technologie Houari Boumediene



Mémoire

Présenté à la :

Faculté de Génie Electrique

Pour l'obtention du grade de :

MAGISTER EN ELECTRONIQUE

Spécialité : Rayonnement Atmosphérique

Par :

ABIB Ghalid Idir

Intitulé :

**MODELISATION DU RAYONNEMENT
SOLAIRE PAR DES RESEAUX DE NEURONES**

Soutenu le 12 Mars 2002 devant le Jury composé de :

Mr A. AISSANI, Professeur (U.S.T.H.B.).....Président
Mr A. MAAFI, Professeur (U.S.T.H.B.).....Directeur de thèse
Mr S. BOUBEKEUR, Chargé de Cours (Université de Blida).....Co-directeur de thèse
Mr A. GUESSOUM, Professeur (Université de Blida).....Examineur
Mr K. BOUKHETALA, Maître de Conférences (U.S.T.H.B.).....Examineur
Mr M. ATTARI, Maître de Conférences (U.S.T.H.B.).....Examineur

A la mémoire de Monsieur le Professeur A. MAAFI.

Dédicaces.

Je dédie ce travail :

A ma très chère et merveilleuse Grand-Mère maternelle "Yemma" qui a toujours été présente à mes côtés et que j'adore.

A mon Père, ma Mère dont les efforts m'ont permis d'arriver à ce jour et que je ne remercierai jamais assez pour leur tendresse, leur compréhension et leurs encouragements.

A mes tantes Zehira et Nour El-Hayatt pour leur dévouement permanent.

A mon oncle El-Bahi et à Manoli pour leurs précieux conseils.

A mes frères Fariz, Nassim et Chahir pour lesquels j'ai énormément d'affection.

A la mémoire de mes Grands Parents paternels.

A la mémoire de mon Grand-Père maternel.

A mes amis.

Remerciements.

Ce travail a été effectué au sein de l'équipe Instrumentation Solaire et Modélisation du Laboratoire d'Instrumentation de la Faculté de Génie Electrique, à l'Université des Sciences et de la Technologie Houari Boumediene (U.S.T.H.B.).

Je tiens à rendre hommage et à témoigner ma profonde reconnaissance à Monsieur A. MAAFI, Professeur à la Faculté de Génie Electrique de l'U.S.T.H.B., pour ses précieux conseils et qui malheureusement n'est plus avec nous aujourd'hui.

Je tiens à exprimer mes plus vifs remerciements à Monsieur S. BOUBEKEUR, Chargé de Cours à l'Université de Blida, qui a été à mes côtés tout au long de ce travail, pour son soutien moral et technique, et pour tout l'intérêt qu'il a témoigné à mon égard.

Ma profonde reconnaissance va également à Monsieur A. AISSANI, Professeur à la Faculté de Génie Electrique, qui a bien voulu nous honorer en acceptant de présider ce jury.

Que Monsieur A. GUESSOUM, Professeur à l'Institut d'Electronique de l'Université de Blida trouve ici mes sincères remerciements, pour avoir bien voulu examiner ce travail.

Je tiens à remercier Monsieur K. BOUKHETALA, Maître de Conférence à la Faculté de Mathématiques de l'U.S.T.H.B., pour avoir bien voulu faire partie de ce jury afin de juger ce travail.

Enfin, je remercie Monsieur M. ATTARI, Maître de Conférence à la Faculté de Génie Electrique de l'U.S.T.H.B., pour avoir accepté d'examiner ce travail et d'être membre du jury.

Sommaire.

	Page
<u>Introduction générale</u>	1
<u>Chapitre I : Généralités sur le rayonnement solaire</u>	4
I-1- Introduction.....	4
I-2- Constitution du rayonnement solaire.....	4
I-2-1- L'irradiation solaire directe.....	4
I-2-2- L'irradiation solaire diffuse.....	5
I-2-3- L'irradiation solaire globale.....	5
I-2-4- L'insolation.....	6
I-2-5- Corrélation entre l'irradiation globale et l'insolation.....	7
I-3- Données expérimentales.....	8
I-4- Modèles développés au sein de l'équipe.....	11
I-4-1- Modèles en chaînes de Markov.....	11
I-4-2- Modèles autorégressifs.....	14
I-4-2-1- Modèles autorégressifs monodimensionnels.....	14
I-4-2-2- Modèles autorégressifs multidimensionnels.....	15
I-4-3- Discussion des modèles développés.....	16
I-5- Conclusion.....	16
<u>Chapitre II : Les réseaux de neurones artificiels</u>	17
II-1- Introduction.....	17
II-2- Les neurones biologiques.....	17
II-3- Le neurone formel.....	18
II-4- Définition de l'apprentissage.....	19
II-5- Classification des réseaux de neurones.....	20
II-6- Les réseaux de neurones multicouches.....	20
II-7- Les réseaux de neurones multicouches d'ordre élevé.....	24

II-8- Les fonctions radiales de base.....	27
II-9- Les réseaux modulaires.....	28
II-10- Les réseaux récurrents.....	30
II-11- Conclusion.....	32
<u>Chapitre III : Description neuronale du rayonnement solaire</u>	33
III-1- Introduction.....	33
III-2- Mise en œuvre des réseaux de neurones dans la modélisation.....	33
III-3- Confection d'une année type.....	35
III-4- Modélisation de HH par les réseaux de neurones multicouches.....	36
III-5- Modélisation de HH par les fonctions radiales de base.....	39
III-6- Modélisation de HH par les réseaux modulaires.....	41
III-7- Modélisation de HH par les réseaux récurrents.....	43
III-8- Modélisation bidimensionnelle de HH et de SS par les réseaux de neurones multicouches.....	45
III-9- Modélisation de HH à partir de SS par les réseaux multicouches.....	47
III-10- Conclusion.....	48
<u>Chapitre IV : Analyse des résultats et prospective</u>	49
IV-1- Introduction	49
IV-2- Résultats de la modélisation par les chaînes de Markov.....	49
IV-3- Résultats de la modélisation autoregressive monodimensionnelle.....	50
IV-4- Résultats de la modélisation autoregressive multidimensionnelle.....	51
IV-5- Comparaison entre les différents résultats.....	52
IV-5-1- Résultats des réseaux de neurones.....	54
IV-5-2- Comparaison avec d'autres modèles.....	55
IV-6- Analogie entre le AR et le réseau multicouches.....	56
IV-7- Analogie entre la régression linéaire et le réseau multicouches.....	58
IV-8- Conclusion.....	59
<u>Conclusion générale</u>	60

<u>Annexe 1 : Conception d'une année type</u>	62
<u>Annexe 2 : Les chaînes de Markov</u>	63
<u>Annexe 3 : Les processus autorégressifs monodimensionnels</u>	67
<u>Annexe 4 : Les processus autorégressifs multidimensionnels</u>	71
<u>Bibliographie</u>	74

Avant-propos.

Le travail présenté dans ce mémoire traite de la modélisation du rayonnement solaire par les réseaux de neurones artificiels. Inspirés du système nerveux humain, ces derniers sont utilisés dans divers domaines tel que le traitement d'image, de la parole, le contrôle et la robotique, en médecine, en météorologie, dans la classification et dans la modélisation de séries temporelles décrivant différents phénomènes.

Dans cette étude, l'application des réseaux de neurones s'est prêtée à la modélisation d'une série temporelle, à savoir l'irradiation solaire. Ceci est motivé par les différentes études récentes publiées dans ce domaine d'une part, et d'autre part par les objectifs de recherche de l'équipe Instrumentation Solaire et Modélisation du Laboratoire d'Instrumentation qui cherche toujours à développer de nouveaux outils d'aide à l'estimation du gisement solaire.

Différentes méthodes furent appliquées avec succès à cette modélisation. Parmi lesquelles, les chaînes de Markov et les processus autorégressifs. Mais les calculs intermédiaires nécessaires lors de l'élaboration de tels modèles, nous poussent à explorer de nouvelles méthodes, dont les réseaux de neurones.

Le but est de tester différents réseaux de neurones, puis de comparer les résultats obtenus à ceux des autres méthodes citées plus haut.

Les résultats obtenus montrent que des réseaux de neurones à structure simplifiée sont suffisants pour reproduire les fluctuations du rayonnement solaire, l'irradiation d'un jour ne dépend que de celle de la veille. Ainsi, le caractère Markovien du rayonnement solaire est bel et bien reproduit par les réseaux de neurones.

Introduction générale.

Le but de ce travail est la modélisation du rayonnement solaire par les réseaux de neurones artificiels, ceci en vue de trouver un modèle décrivant au mieux son évolution temporelle.

Ce rayonnement solaire n'est autre que l'énergie qu'émet le soleil [1-2]. Hors atmosphère, cette évolution est parfaitement décrite par des équations mathématiques faisant intervenir les coordonnées géo-astronomiques [3]. Mais les constituants de l'atmosphère [1-2], font que le rayonnement reçu en un site donné après la traversée de l'atmosphère terrestre, présente des variations aléatoires, dès lors, ces modèles analytiques [3] ne suffisent plus pour décrire son évolution au cours du temps.

Ceci se révèle être un inconvénient dans des applications dépendantes de l'énergie solaire [4-6]. Ainsi, la conversion photovoltaïque [4-5] ou encore les modèles de simulation agronomique ou écologique [6] utilisent le rayonnement solaire comme grandeur d'entrée, leur fonctionnement est directement lié à la bonne description de ce rayonnement.

Pour tenter d'expliquer et de décrire ce phénomène physique, des modèles mathématiques aléatoires [4,7-15] ont été mis à contribution. Ainsi, les chaînes de Markov [7-10] et les processus autorégressifs [4,11-15] ont été appliqués avec succès. Ils permettent de tenir compte du caractère aléatoire du rayonnement solaire.

L'application des chaînes de Markov d'ordre 1 à 2 états [7-8,10] et des processus autorégressifs d'ordre 1 [4,11-15] a conduit à des résultats reproduisant correctement et suffisamment l'évolution temporelle du phénomène physique.

Pour l'application de la première méthode, il est nécessaire de discrétiser les mesures en état, donc de choisir des seuils pour cet effet, de plus la stationnarité de la série temporelle est requise. Dans le cas des processus autorégressifs, une transformation gaussienne et une transformation gaussienne inverse sont nécessaires pour leur application [4,10].

Afin d'éviter ces étapes intermédiaires, les réseaux de neurones artificiels [16-19] sont appliqués à cette modélisation. Inspirés du neurone biologique [18], les premiers travaux datent de 1943 lorsque deux bio-physiciens de l'Université de Chicago McCulloch et Pitts, s'inspirent des récentes découvertes en neurobiologie et conçoivent le premier modèle baptisé Neurone Formel. Puis en 1949, un neurophysiologiste, Donald Hebb, propose une règle de modification des connexions synaptiques. Finalement, en 1958, Rosenblatt conçoit le premier réseau de neurones artificiels, le Perceptron. Il possède une couche de neurones perceptive (entrée) et une couche de neurones décisionnelle (sortie). En 1960, B. Widrow, un automaticien, développe le modèle Adaline (Adaptative Linear Element), ressemblant au perceptron, sa loi d'apprentissage est différente et est à l'origine de l'algorithme de rétropropagation de gradient très utilisé avec les perceptrons multicouches [18]. Mais en 1969, Minsky et Papert publient un livre mettant à jour les limites du perceptron. Au début des années 80, des équipes française et américaine ont découvert simultanément l'algorithme de rétropropagation du gradient permettant la mise en œuvre du réseau multicouches. Le renouveau eut lieu en 1982 par J.J Hopfield où il démontre la relation entre les réseaux complètement connectés et les systèmes physiques (physique statistique).

Les réseaux de neurones artificiels permettent d'effectuer cette modélisation sans nécessiter de traitements annexes [18], ce qui n'est pas le cas pour les chaînes de Markov et les processus autorégressifs.

Depuis, les réseaux de neurones ont connu un essor considérable et sont utilisés dans divers domaines : le traitement d'image [18], de la parole [18], dans le contrôle et la robotique [20], en médecine [21-22], en oncologie [21], dans la classification [21-23], en météorologie [24-29] et dans la modélisation de séries temporelles décrivant différents phénomènes [18] (économiques, météorologiques, industriels, ... etc.).

Parmi les modèles développés, on cite les réseaux de neurones multicouches [16-19], les réseaux multicouches d'ordre élevé [31-32], les réseaux à fonction radiale de base [18,28-30], les réseaux modulaires [18,33] et les réseaux récurrents [18,34].

La structure de ce mémoire est la suivante :

Le premier chapitre est consacré à des généralités sur le rayonnement solaire. Egalement, quelques rappels sur les chaînes de Markov et les processus autorégressifs seront donnés.

Le deuxième chapitre permet la présentation des différents réseaux de neurones utilisés. Ainsi, les réseaux cités plus haut seront définis, ainsi que leurs algorithmes respectifs permettant leur mise en place.

L'application de ces réseaux de neurones s'est faite au chapitre III, où le rayonnement solaire est modélisé par les réseaux décrits dans le chapitre II.

Le quatrième et dernier chapitre, permet d'analyser les résultats trouvés. De là, des structures seront mises en avant en fonction des critères de comparaison retenus et de la complexité du modèle choisi. Les résultats obtenus sont comparés à ceux trouvés par les chaînes de Markov d'ordre un, à deux états [7-8,10] et les processus autorégressifs d'ordre un [4,11-15].

Enfin, une conclusion générale résume tout le travail effectué et les perspectives à observer sont présentées.

Chapitre I : Généralités sur le rayonnement solaire.

I-1- Introduction :

L'énergie qu'émet le soleil est reçue sur terre sous forme de rayonnement solaire [1-2]. Hors atmosphère, ce rayonnement est parfaitement décrit par des équations mathématiques via les paramètres astronomiques [3]. Ainsi, le rayonnement reçu dépend du temps et du site sur lequel on désire estimer ce rayonnement. Mais la présence de l'atmosphère terrestre fait que l'énergie reçue n'obéit plus à ces équations, ceci est dû aux constituants de l'atmosphère tels que les molécules de gaz, les gouttes d'eau et les poussières [1-2]. Ces dernières absorbent une partie de cette énergie avant d'atteindre la terre, la transforment en partie en chaleur et diffusent le reste dans toutes les directions.

Ainsi donc, le rayonnement solaire subit des variations aléatoires en traversant l'atmosphère terrestre. De ce fait, ce caractère aléatoire fait que la connaissance du rayonnement reçu n'est pas précise, et ceci s'avère être un inconvénient pour des applications dépendantes du rayonnement solaire [4-6]. Pour cela, les chaînes de Markov et les processus autorégressifs seront brièvement décrits, ils ont été appliqués avec succès à la modélisation du rayonnement solaire [4,7-15].

I-2- Constitution du rayonnement solaire :

Le rayonnement solaire a deux composantes [1-3] : l'irradiation solaire et la durée d'insolation. De même, l'irradiation solaire se divise en deux parties : l'irradiation directe et l'irradiation diffuse. Elles forment donc l'irradiation solaire globale.

I-2-1- L'irradiation solaire directe :

Notée H_{dir} , l'irradiation directe est le flux énergétique du rayonnement solaire par unité de surface reçu sur un plan posé perpendiculairement aux rayons émis par le soleil [1-2]. Sa mesure se fait grâce à un pyranomètre représenté dans la figure I-1, l'énergie reçue est transformée en chaleur. Les pyranomètres utilisés disposent de piles thermoélectriques à

surface plane, correctement nivelées pour être parfaitement horizontales [7]. L'irradiation est donnée en J/m^2 .

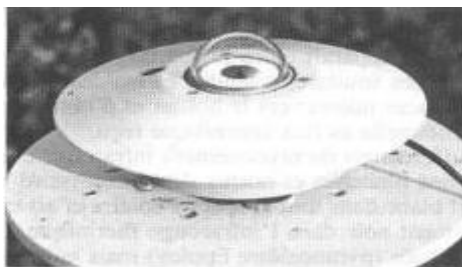


Fig. I-1 : Pyranomètre Cimel.

I-2-2- L'irradiation solaire diffuse :

L'irradiation diffuse notée H_{dif} est le rayonnement diffusé par les composants de l'atmosphère et du sol [1-2]. C'est donc l'énergie qui est reçue au niveau du sol autre que le rayonnement parvenant directement du soleil. Pour sa mesure, un pyranomètre muni d'un cache éliminant le rayonnement direct est utilisé. L'irradiation solaire diffuse est aussi exprimée en J/m^2 .

I-2-3- L'irradiation solaire globale :

L'irradiation globale H est la résultante de l'irradiation directe et diffuse, ainsi [1-2] :

$$H = H_{dir} + H_{dif} \quad (I-1)$$

De même, sa mesure se fait à l'aide de pyranomètres tel que les pyranomètres "Kipp et Zonen", et peut être mesurée sur un plan horizontal ou sur un plan incliné. Vu que peu de stations météorologiques sont équipées d'appareils mesurant l'irradiation directe, elle est déduite de la mesure de l'irradiation globale et de l'irradiation diffuse.

Hors atmosphère, l'irradiation solaire globale ne dépend que des paramètres géo-astronomiques. Elle est alors appelée irradiation potentielle et est notée H_0 . Pour un site de latitude θ et sur plan horizontal, sa valeur pour un jour j de l'année est donnée par l'équation suivante :

$$H_{o_j} = \frac{24 \cdot 3600}{\pi} I_j \cos(\theta) \delta_j \left[\sin(w_j) - \frac{\pi}{180} w_j \cos(w_j) \right] \quad (\text{I-2})$$

Où H_{o_j} est donnée en J/m^2 et peut être convertie en Wh/m^2 .

I_j est la valeur journalière corrigée de la constante solaire I_0 adoptée par l'Organisation Mondiale de la Météorologie (O.M.M.) et est donnée par :

$$I_j = I_0 [1 + 0.033 \cos(0.98 j)] \quad (\text{I-3})$$

avec $I_0 = 1367 \text{ W/m}^2$.

w_j est l'angle horaire au lever du soleil, il est donné comme suit en degré ($^\circ$) :

$$w_j = \text{Ar cos}[-\tan g(\theta) \tan g(\delta_j)] \quad (\text{I-4})$$

δ_j est la déclinaison du soleil, elle est donnée par la relation suivante toujours en degré ($^\circ$) :

$$\delta_j = 23.45 \sin(0.986 j - 80) \quad (\text{I-5})$$

j est le jour de l'année et varie de 1 à 365 ou 366.

En général, l'éclairement (puissance reçue par unité de surface) est mesuré par un pyranomètre à échelle de 10 minutes, la valeur horaire ou journalière est calculée à partir des mesures au pas de 10 minutes par intégration sur la durée désirée (heure ou jour), par la méthode des trapèzes par exemple.

On désigne par fraction d'irradiation journalière notée HH_j , le rapport de l'irradiation globale journalière mesurée H_j à l'irradiation potentielle du même jour H_{o_j} [3].

I-2-4- L'insolation :

L'insolation notée S est la durée pendant laquelle le ciel est dégagé [1-2]. Appelée également durée d'ensoleillement, elle représente pour une journée donnée la somme des intervalles de temps au cours desquels l'intensité du rayonnement solaire (éclairement) a dépassé un seuil de 100 W/m^2 . L'héliographe de la figure I-2 du type "Cambell-Stockes" permet sa mesure. Il est constitué d'une sphère de verre concentrant le rayonnement solaire

sur une bande de carton étalonnée et brûle cette dernière. La longueur de la trace obtenue sur cette bande correspond à la durée d'insolation [7]. Là encore, hors atmosphère, elle est appelée insolation potentielle, notée S_0 . Elle est donnée par l'équation ci-dessous en heures :

$$S_{0j} = \frac{2}{15} \text{Arcos}(w_j) \quad (\text{I-6})$$

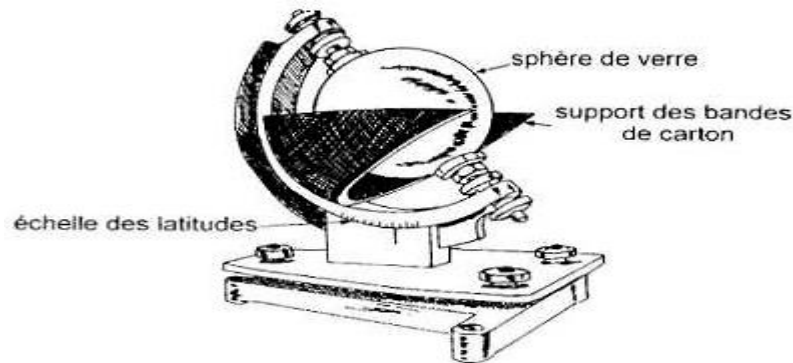


Fig. I-2 : Schéma d'un héliographe.

Nous définissons là aussi, la fraction d'insolation journalière notée SS_j comme étant le rapport de l'insolation journalière mesurée S_j à l'insolation potentielle du même jour S_{0j} [3].

I-2-5- Corrélation entre l'irradiation globale et l'insolation :

En 1924, Angström a démontré que la fraction d'irradiation journalière HH_j et la fraction d'insolation SS_j sont liées par la relation [7] :

$$HH_j = \alpha SS_j + \beta \quad (\text{I-7})$$

α et β sont les coefficients de la droite de régression linéaire [35].

Ainsi, ces deux composantes du rayonnement solaire sont fortement corrélées comme le montre la figure I-3 pour le mois de Septembre cumulé de 1972 à 1982 pour le site d'Alger. Pour cela, nous avons pris toutes les données du mois de Septembre de 1972 à 1982, puis, nous les avons concaténées pour obtenir à la fin une série de données comprenant 30×11 valeurs, où 30 représente le nombre de jours du mois considéré et 11 le nombre d'années disponibles.

Ceci s'avère très utile quand seulement les relevés d'insolation sont disponibles, l'insolation peut être ainsi transposée afin d'estimer l'irradiation.

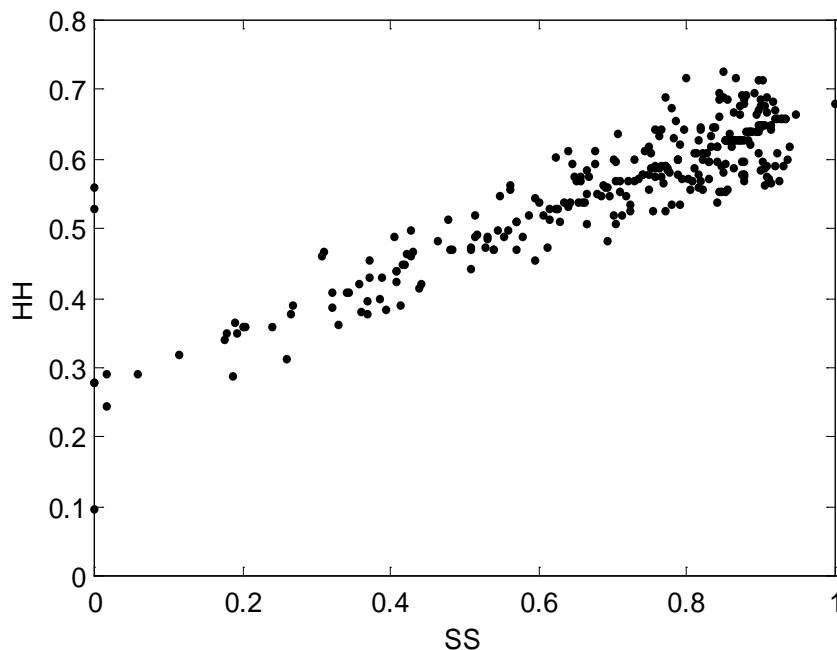


Fig. I-3 : Corrélation entre HH_j et SS_j pour le mois de Septembre cumulé de 1972 à 1982 à Alger.

I-3- Données expérimentales :

Afin de pouvoir comparer les résultats qui seraient obtenus à l'aide des réseaux de neurones à ceux déjà publiés et résultants des autres méthodes, nous avons choisi de travailler avec les séries de mesures de la période 1972 – 1982. En effet, ces séries de mesures ont servi à la mise au point de modèles autorégressifs et en chaînes de Markov qui sont déjà répertoriés dans la littérature [7,10,13-14].

Ainsi, nous disposons des données journalières d'irradiation globale sur plan horizontal et des données d'insolation mesurées à l'Office National de la Météorologie (O.N.M.) à Dar El-Beida à Alger (lat $36^{\circ}43'$ N, long $3^{\circ}15'$ E et alt 25 m).

L'évolution journalière de l'irradiation globale et de l'insolation au cours de la dernière année est représentée dans la figure I-4 ci-dessous, ainsi que H_0 et S_0 . Le caractère aléatoire est superposé à l'évolution saisonnière au cours de l'année. Cette tendance saisonnière est régie respectivement par l'évolution de H_0 pour l'irradiation et S_0 pour l'insolation. En été, les valeurs de H et de S sont les plus élevées de l'année.

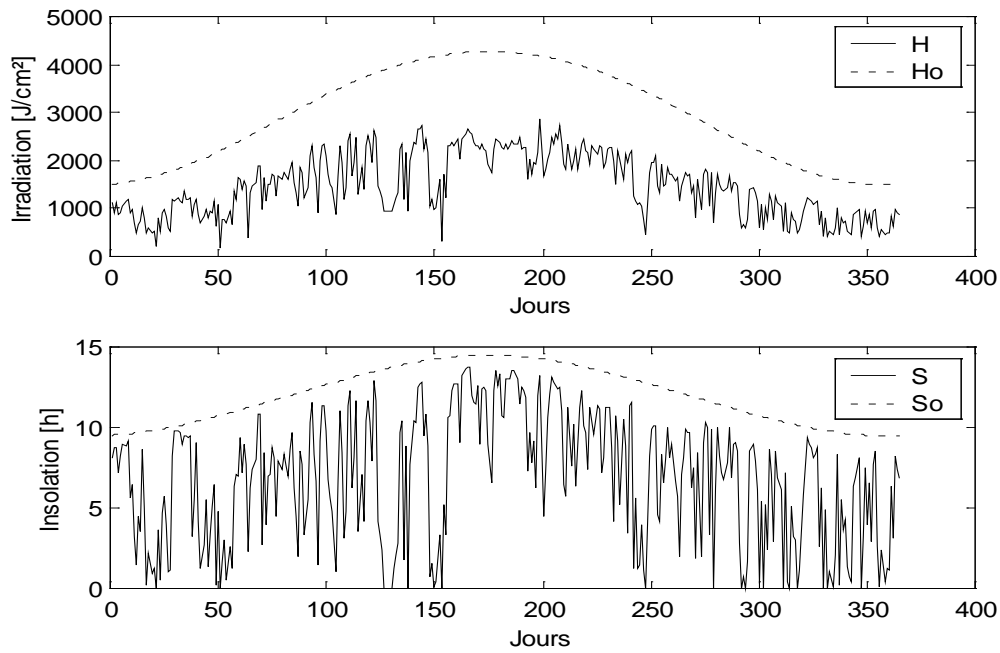


Fig. I-4 : Evolution journalière de H et S de 1982 et de Ho, So.

Afin de mieux apprécier les fluctuations aléatoires de H et de S, la figure I-5 illustre l'évolution journalière des fractions d'irradiation HH_j et d'insolation SS_j de la période 1972-1982.

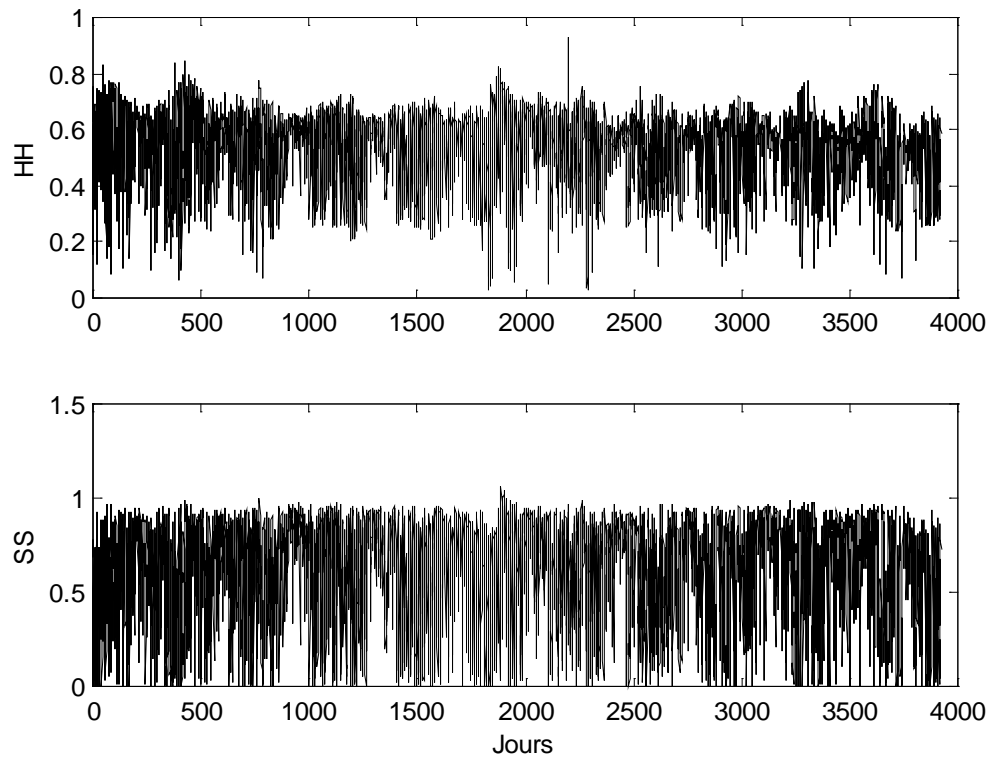


Fig. I-5 : Evolution de HH et de SS de 1972 à 1982 à Alger.

La tendance saisonnière est ainsi écartée. L'évolution au cours du temps est caractérisée par des changements brusques d'une valeur à la suivante. La figure I-6 représente les histogrammes de HH et de SS pour chaque mois cumulé de la période 1972 – 1982 à Alger par classe de 0,1. Nous remarquons que les plus fortes valeurs de HH et de SS sont mesurées durant la saison estivale, caractérisant ainsi le climat méditerranéen régnant à Alger.

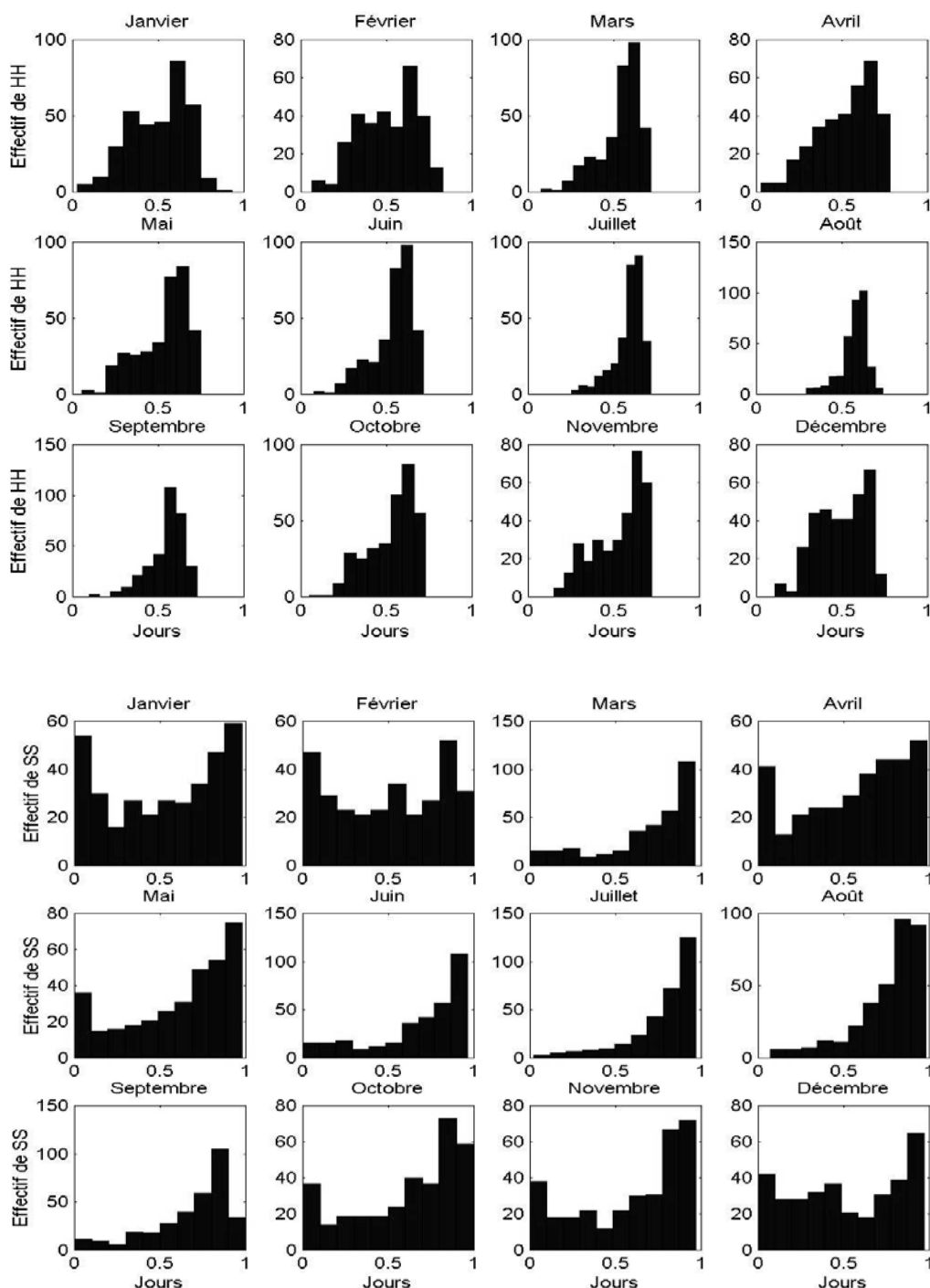


Fig. I-6 : Histogramme de HH et SS de chaque mois cumulé de la période 1972 – 1982 à Alger.

I-4- Modèles développés au sein de l'équipe :

I-4-1- Modèles en chaînes de Markov :

Les chaînes de Markov [7-10,36] sont des processus aléatoires à temps discret et à amplitude discrète et où l'évolution temporelle dépend du passé. Cette évolution est caractérisée par un régime transitoire et un régime permanent (ou stationnaire). Soit un système pouvant se trouver dans m états possibles notés : $\phi_0, \phi_1, \dots, \phi_{m-1}$ et soit X_t (t représente le temps discret) une variable aléatoire qui peut prendre l'une des valeurs discrètes $E_t = 0, 1, \dots, m-1$ associées respectivement aux états $\phi_0, \phi_1, \dots, \phi_{m-1}$. La probabilité (conditionnelle) que le processus effectue une transition de l'état ϕ_i à l'instant (t) ($E_t = i$) vers l'état ϕ_j à l'instant ($t+1$) ($E_{t+1} = j$) est donnée par :

$$P_{ij} = P(X_{t+1} = j / X_t = i) \quad (\text{I-8})$$

Une chaîne de Markov est complètement définie par la connaissance de la matrice stochastique de transition P formée à partir des probabilités de transition P_{ij} et des probabilités à priori, c'est le régime transitoire. La chaîne est dite d'ordre n quand l'état futur du processus dépend des n états précédents, ceci se traduit par [7,10] :

$$P(X_{t+1} = E_{t+1} / X_t = E_t, X_{t-1} = E_{t-1}, \dots, X_{t-(n-1)} = E_{t-(n-1)}, \dots, X_{t-(n'-1)} = E_{t-(n'-1)}) = P(X_{t+1} = E_{t+1} / X_t = E_t, X_{t-1} = E_{t-1}, \dots, X_{t-(n-1)} = E_{t-(n-1)}) \quad (\text{I-9})$$

Ainsi, malgré la connaissance des n' états précédents ($n' > n$), seuls les n plus récents états suffisent à la connaissance de l'état futur. La matrice P pour l'ordre 1 à m états est donnée par [7,10] :

$$P = \begin{pmatrix} P_{00} & P_{01} & \dots & P_{0(m-1)} \\ P_{10} & P_{11} & \dots & P_{1(m-1)} \\ \vdots & \vdots & \ddots & \vdots \\ P_{(m-1)0} & P_{(m-1)1} & \dots & P_{(m-1)(m-1)} \end{pmatrix} \quad (\text{I-10})$$

Dans le cas de l'ordre 2 à 2 états [7] :

$$P = \begin{pmatrix} P_{000} & P_{001} \\ P_{010} & P_{011} \\ P_{100} & P_{101} \\ P_{110} & P_{111} \end{pmatrix} \tag{I-11}$$

Soit l'équation matricielle suivante [7,10] :

$$\lim_{r \rightarrow +\infty} P^r = \lim_{r \rightarrow +\infty} P^{r-1}P \tag{I-12}$$

Sa résolution conduit à une matrice ayant des lignes identiques, le vecteur d'état qui en résulte représente le régime permanent. Ce sont les probabilités à long terme. On peut obtenir ce vecteur en multipliant P par elle même r fois [7,10].

Le problème dans les chaînes de Markov est la détermination de cette matrice P. Dans ce but, on fait appel à des méthodes statistiques telle la méthode du seuil [7] nécessitant le test du χ^2 (Khi-2) [35,37] ou la méthode du maximum d'entropie [10].

Dans ce qui suit, nous illustrons les résultats de la modélisation de la fraction d'irradiation mesurée à Alger de 1972 à 1981 par les chaînes de Markov d'ordre 1 à 2 états [7-10]. La discrétisation en états nécessite le calcul d'un seuil afin de faire la distinction entre ces 2 états, puis vient le calcul des éléments de P [7-10]. Les calculs du seuil, des probabilités de transition, des probabilités à long terme et à priori sont donnés en annexe 2. Ces calculs ont été utilisés dans [7,10].

La figure I-7 représente les 2 états ϕ_0 et ϕ_1 et les probabilités de transition P_{ij} [7-10].

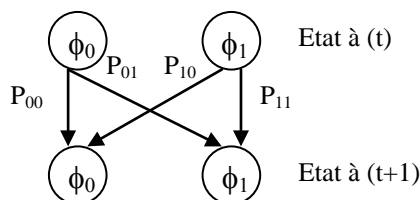


Fig. I-7 : Chaîne de Markov d'ordre 1 à 2 états.

Dans ce cas, X_t est égale à 1 si l'état à (t) est ϕ_1 et égale à 0 si l'état est ϕ_0 . La matrice de transition P est donnée par [7-10] :

$$P = \begin{pmatrix} P_{00} & P_{01} \\ P_{10} & P_{11} \end{pmatrix} \tag{I-13}$$

Avec :

$$P_{00} = P(X_{t+1} = 0 / X_t = 0) \tag{I-14}$$

$$P_{01} = P(X_{t+1} = 1 / X_t = 0) \tag{I-15}$$

$$P_{10} = P(X_{t+1} = 0 / X_t = 1) \tag{I-16}$$

$$P_{11} = P(X_{t+1} = 1 / X_t = 1) \tag{I-17}$$

$$P_{01} = 1 - P_{00} \tag{I-18}$$

$$P_{10} = 1 - P_{11} \tag{I-19}$$

Les 2 probabilités à long terme sont données par le vecteur Pm [7,10] :

$$Pm = (P_0 \quad P_1) \tag{I-20}$$

La figure I-8 illustre l'évolution du seuil S, de P₀₁, P₁₁ et P₁ pour chaque mois cumulé de 1972 à 1981. L'échelle mensuelle [7,10] est adoptée pour satisfaire à la condition de stationnarité de la série temporelle.

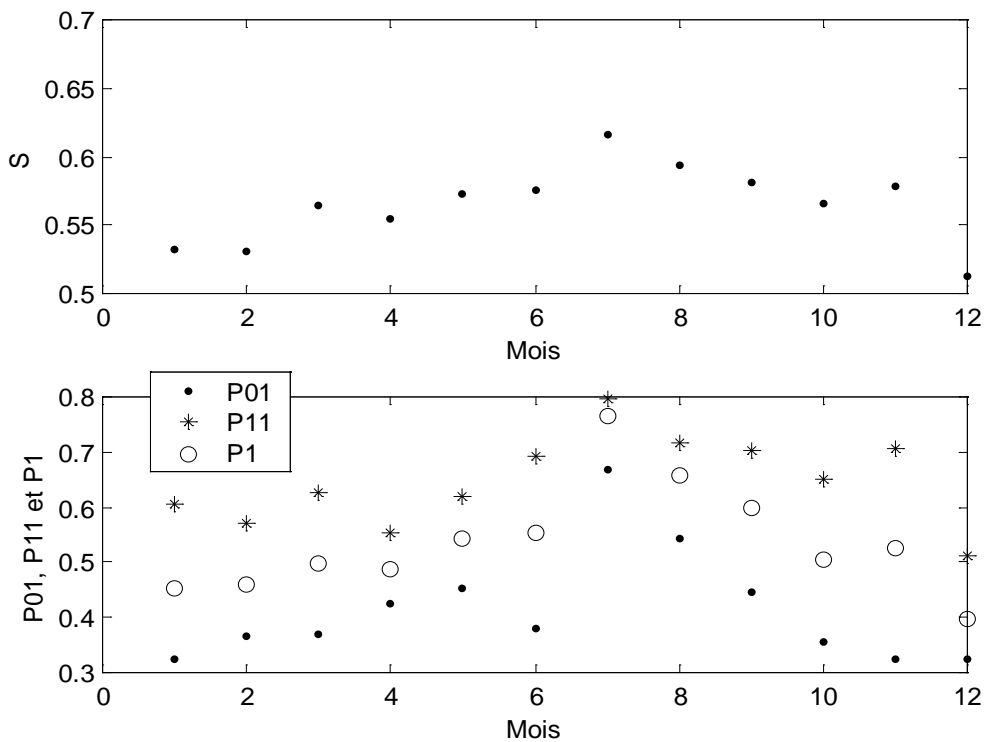


Fig. I-8 : Evolution mensuelle de S, P₀₁, P₁₁ et P₁.

I-4-2- Modèles autorégressifs :**I-4-2-1- Modèles autorégressifs monodimensionnels :**

Soit X_t un processus où le temps est discret, stationnaire, ergodique et gaussien. Un processus AutoRégressif d'ordre p , noté AR(p) est défini par l'équation suivante [4,10-12,15,38-41] :

$$X_t - M = \sum_{i=1}^p a_i (X_{t-i} - M) + e_t \quad (\text{I-21})$$

où M est la moyenne de X_t , e_t un bruit blanc de moyenne nulle et de variance σ^2 et enfin a_i est le $i^{\text{ème}}$ coefficient du modèle.

Ainsi X_t est une combinaison linéaire du processus aux instants $(t-1)$, ..., $(t-p)$ à laquelle on rajoute un bruit blanc gaussien e_t . Les paramètres a_1 , ..., a_p et σ^2 peuvent être déterminés par l'application de l'algorithme de Levinson-Durbin [38-41] décrit dans l'annexe 3. L'ordre optimal est choisi en regard d'un coefficient appelé FPE, lui aussi décrit dans l'annexe 3.

L'application de l'algorithme de Levinson-Durbin jusqu'à l'ordre 3 à une année type HH, créée à partir des données mesurées à Alger de 1972 à 1981, a conduit aux résultats donnés au tableau I-1. Les coefficients du modèle a_i , la variance du bruit σ^2 et le critère FPE y sont affichés. Les données ont dû subir la transformation gaussienne [10,12] avant l'application de l'algorithme pour satisfaire la condition de stationnarité [10]. La confection de l'année type est décrite dans l'annexe 1.

p	a_1, \dots, a_p		σ^2	FPE
1	0.2849		0.8909	0.8958
2	0.3025	-0.0619	0.8875	0.8973
3	0.3048	-0.0730	0.0369	0.8863

Tabl. I-1 : Présentation des paramètres du modèle AR jusqu'à l'ordre 3.

Comme nous le remarquons, l'ordre 1 est suffisant pour la modélisation autorégressive étant donné que le coefficient FPE ne varie pas sensiblement en passant de l'ordre 1 à 3.

I-4-2-2- Modèles autorégressifs multidimensionnels :

Le modèle AutoRégressif Multidimensionnel d'ordre p noté MAR(p) (Multivariate AutoRegressive) se traduit par l'équation suivante [13-14,42] :

$$X_t - M = \sum_{i=1}^p a_i (X_{t-i} - M) + e_t \tag{I-22}$$

où : $X_t = (X_t^1, \dots, X_t^m)^T$, $M = (M_1, \dots, M_m)^T$, et $e_t = (e_t^1, \dots, e_t^m)^T$ tel que $E(e_t) = 0$, et $E(e_i e_j^T) = \sigma^2 \delta_{ij}$. (L'exposant ^T vecteur signifie transposé).

X_t^i est la $i^{\text{ème}}$ série temporelle, M_i est la moyenne de la $i^{\text{ème}}$ série temporelle et e_t^i est le $i^{\text{ème}}$ bruit blanc tel que σ^2 est la matrice de covariance de e_t , de dimension mxm. Enfin, a_i est le $i^{\text{ème}}$ coefficient (matrice) du modèle de dimension mxm.

Là encore, les paramètres a_i et σ^2 peuvent être déterminés par l'algorithme de Levinson-Durbin généralisé [42] décrit dans l'annexe 4.

Pour ce cas de figure, les deux séries temporelles sont la fraction d'irradiation HH et la fraction d'insolation SS mesurées à Alger de 1972 à 1981. La modélisation s'effectue via l'année type de chaque série (voir annexe 1). Le tableau I-2 regroupe les résultats de l'application de l'algorithme de Levinson-Durbin à ces données jusqu'à l'ordre 3. La transformation gaussienne a aussi été appliquée aux 2 séries.

p	a_1, \dots, a_p		σ^2	FPE	
1	$\begin{pmatrix} 0.2855 & -0.0070 \\ -0.0390 & 0.3804 \end{pmatrix}$		$\begin{pmatrix} 0.8909 & 0.0799 \\ 0.0799 & 0.6672 \end{pmatrix}$	0.6010	
2	$\begin{pmatrix} 0.3044 & -0.0192 \\ -0.0308 & 0.3737 \end{pmatrix}$	$\begin{pmatrix} -0.0646 & 0.0255 \\ -0.0278 & 0.0147 \end{pmatrix}$	$\begin{pmatrix} 0.8870 & 0.0782 \\ 0.0782 & 0.6664 \end{pmatrix}$	0.6112	
3	$\begin{pmatrix} 0.3090 & -0.0165 \\ -0.0274 & 0.3753 \end{pmatrix}$	$\begin{pmatrix} -0.0819 & 0.0773 \\ -0.0458 & 0.0115 \end{pmatrix}$	$\begin{pmatrix} 0.0499 & -0.1347 \\ 0.0601 & 0.0132 \end{pmatrix}$	$\begin{pmatrix} 0.8738 & 0.0773 \\ 0.0773 & 0.6629 \end{pmatrix}$	0.6123

Tabl. I-2 : Présentation des paramètres du modèle MAR jusqu'à l'ordre 3.

Il est clair que l'ordre 1 est suffisant pour cette modélisation du fait que le FPE ne varie pas considérablement de l'ordre 1 à 3.

I-4-3- Discussion des modèles développés :

Que ce soit les chaînes de Markov ou les processus autorégressifs, la principale conclusion est que le caractère Markovien de la fraction d'irradiation solaire d'Alger se réduit à l'ordre 1 [4,7-15]. Ainsi, 2 mesures journalières consécutives sont fortement corrélées. La différence entre ces deux méthodes est que pour les chaînes de Markov l'amplitude est discrète alors que pour les processus autorégressifs, elle est continue.

I-5- Conclusion :

La modélisation par les chaînes de Markov et les processus autorégressifs, nécessite la stationnarité de la série temporelle. Pour satisfaire cette condition, cette modélisation se fait à échelle mensuelle dans le cas des chaînes de Markov, et les données doivent subir la transformation gaussienne dans le cas de la modélisation à échelle annuelle par les processus AR. Tout ceci nous incite à appliquer de nouvelles techniques afin de pallier à ces inconvénients. Parmi ces méthodes, les réseaux de neurones permettent cette modélisation et aucune condition n'est requise sur les données. Le chapitre suivant présente les réseaux de neurones et leurs algorithmes associés.

Chapitre II : Les réseaux de neurones artificiels.

II-1- Introduction :

Les réseaux de neurones artificiels [16-19] sont inspirés des neurones biologiques du système nerveux humain. Ils ont été mis à contribution dans divers domaines tel que la vision, le traitement de la parole, la reconnaissance de formes, la modélisation de séries temporelles, ... etc.

Le neurone biologique est une cellule vivante traitant des signaux électriques. Les neurones sont reliés entre eux par des liaisons appelées axones. Ces axones conduisent les signaux électriques de la sortie d'un neurone vers l'entrée (synapse) d'un autre neurone. Ils font une sommation des signaux reçus en entrée et en fonction du résultat obtenu, vont fournir un courant en sortie [16-19].

Les réseaux de neurones présentés dans ce travail sont : les réseaux de neurones multicouches [16-19], les réseaux multicouches d'ordre élevé [31,32], les réseaux à fonction radiale de base [18,28-30], les réseaux modulaires [18,33] et les réseaux récurrents [18,34]. Le choix de ces réseaux vient du fait qu'ils aient été appliqués dans la modélisation de série temporelle de différentes natures [18,34,43-44]. Il s'adaptent donc à la modélisation du rayonnement solaire. Les résultats de cette modélisation seront présentés dans le prochain chapitre.

II-2- Les neurones biologiques :

Des cellules interconnectées appelées neurones sont les constituants du cerveau humain. Le neurone biologique se compose d'un corps cellulaire (soma) et de deux prolongements : les dendrites qui sont les entrées ou récepteurs du neurones, et l'axone qui est la sortie ou émetteur de la cellule tel que le montre la figure II-1. Il reçoit des signaux (impulsions électriques) émanant d'autres neurones au niveau des dendrites et transmet l'information, générée dans son corps cellulaire, via l'axone [18].

Chaque branche de cette arborisation se termine en un bouton synaptique autour duquel se trouvent les synapses. La synapse est une structure essentielle entre deux neurones permettant la transmission des signaux entre un axone et une dendrite. A son niveau, il existe un espace vide à travers lequel le signal électrique ne peut pas se propager. La transmission se fait alors par l'intermédiaire de substances chimiques (acides aminés, protéines), appelées neurotransmetteurs. Quand un signal arrive au niveau de la synapse, ces substances sont libérées, elles le traversent et se fixent sur les récepteurs du neurone récepteur [18].

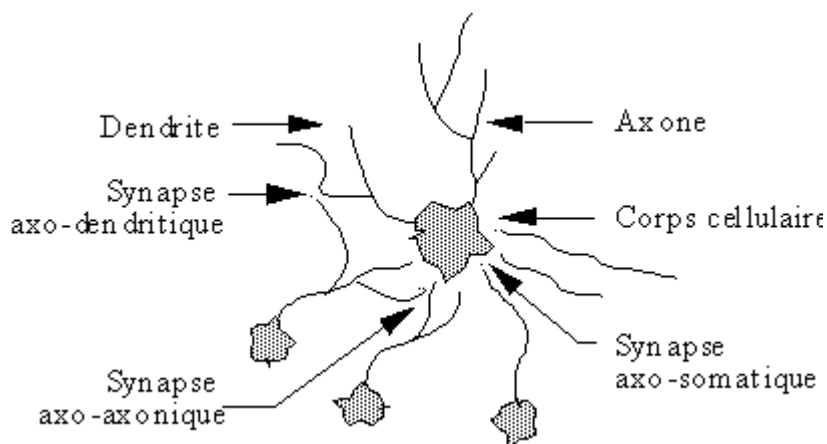


Fig. II-1 : Le neurone biologique.

II-3- Le neurone formel :

Il a été présenté par McCulloch et Pitts en 1943 [18,45], c'est une cellule ayant plusieurs entrées et une sortie. Pour chaque entrée, un poids synaptique lui est associé représentant ainsi la force de connexion entre cette entrée et le neurone. Ces entrées proviennent des neurones en amont et sa sortie alimente les neurones en aval [18]. La figure II-2 illustre un neurone formel ayant p entrées.

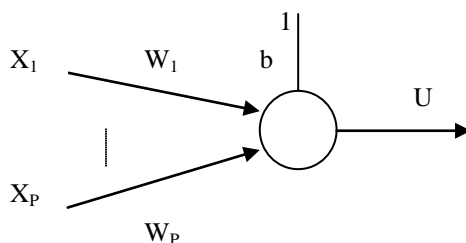


Fig. II-2 : Le neurone formel.

avec : X_i : la $i^{\text{ème}}$ entrée.
 p : le nombre d'entrées.

W_i : le poids synaptique de la $i^{\text{ème}}$ entrée.

b : le seuil ou biais, il sera remplacé par un poids synaptique comme on le verra en dessous.

U : la sortie du neurone.

Le neurone est un opérateur qui calcule sa sortie U à partir de p entrées X_i , d'autant de poids synaptiques W_i et d'un seuil b comme suit [18]:

$$U = F\left(\sum_{i=1}^P W_i X_i + b\right) \quad (\text{II-1})$$

F est la fonction d'activation, elle peut être :

Sigmoïde :

$$F(x) = \frac{e^{kx}}{1 + e^{kx}} \quad (k \in \mathfrak{R}). \quad (\text{II-2})$$

Tangente hyperbolique :

$$F(x) = \tanh(x) \quad (\text{II-3})$$

Linéaire :

$$F(x) = x \quad (\text{II-4})$$

Pour simplifier la notation de l'équation (II-1), on peut remplacer le seuil b par un poids synaptique W_0 associé à une entrée X_0 égale à 1, ainsi [18] :

$$U = F\left(\sum_{i=0}^P W_i X_i\right) \quad (\text{II-5})$$

II-4- Définition de l'apprentissage :

L'apprentissage est la caractéristique la plus importante des réseaux de neurones [16-19]. Durant cette phase, le comportement du réseau est modifié et ce jusqu'à l'obtention du comportement désiré, ainsi les variables modifiées sont les poids en faisant appel à des exemples de comportement. A la fin de l'apprentissage, les poids sont fixés, vient alors la phase d'utilisation. L'apprentissage est dit supervisé lorsque les exemples sont constitués de couples de valeurs du type : (entrée, sortie désirée). Il est dit non supervisé quand seules les valeurs d'entrée sont disponibles et en sortie, un signal d'erreur est disponible.

II-5- Classification des réseaux de neurones :

Il existe une variété importante de réseaux de neurones dans la littérature [16-19], leurs domaines d'applications étant la robotique [18,20], l'optimisation [18], le traitement d'images et de la parole [18], la médecine [21-22], la modélisation de séries temporelles [18]... etc. Ainsi le réseau de Hopfield [17-18] et une de ses généralisations, la machine de Boltzmann [18,23], reposent sur le concept de la thermodynamique statistique. Ils ont été utilisés avec succès dans le traitement d'images, d'optimisation, et également en tant que mémoire à contenu adressable [17-18]. Egalement les cartes auto-organisatrices (ou carte de Kohonen) [17-18] ont été utilisées pour générer des cartes phonétiques, diagnostic de pannes, compression d'images, robotique, reconnaissance de formes ... etc.

Mais l'application qui nous intéresse, est la modélisation de séries temporelles en l'occurrence le rayonnement solaire. Les réseaux cités ci-dessus n'ont pas été utilisés dans ce but. Les réseaux et les algorithmes présentés dans ce chapitre permettent la réalisation de cette tâche en tenant compte du caractère aléatoire du rayonnement solaire.

II-6- Les réseaux de neurones multicouches :

Les réseaux multicouches dont la structure est illustrée à travers la figure II-3, utilisent un apprentissage supervisé en 2 étapes [16-19]. Dans un premier temps, une propagation en avant permet de calculer la sortie, puis dans un second temps, une propagation en arrière ajuste les poids dont le but est de minimiser l'erreur entre la sortie calculée par le réseau et la sortie désirée. Les neurones sont arrangés par couche, il n'y a pas de connexion entre neurones d'une même couche, et chaque neurone d'une couche est connecté à tous les neurones de la couche suivante et celle-ci seulement.

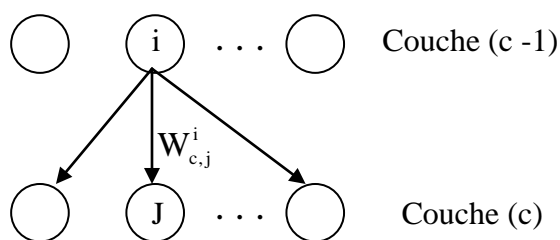


Fig. II-3 : Un réseau multicouches.

On considère un réseau comprenant $L+1$ couches (de 0 à L), on désigne par :

N_c : le nombre de neurones de la couche c .

$W_{c,j}^i$: le poids entre le $j^{\text{ème}}$ neurone de la couche c et le $i^{\text{ème}}$ neurone de la couche $c-1$.

$U_{c,j}$: la sortie du neurone j de la couche c .

X^g : le $g^{\text{ème}}$ exemple d'entrée.

d_j^g : la sortie désirée du $j^{\text{ème}}$ neurone de la couche de sortie associée à l'exemple g .

$U_{0,j}$: la $j^{\text{ème}}$ composante du vecteur d'entrée X^g .

F : la fonction d'activation (sigmoïde, tangente hyperbolique,...etc).

N : le nombre d'exemples.

E : l'erreur du réseau.

μ : le taux d'apprentissage.

K_{\max} : le nombre d'itérations maximales.

Pour un neurone j de la couche c , la sortie est donnée par :

$$U_{c,j} = F\left(\sum_{a=0}^{N_{c-1}} W_{c,j}^a U_{c-1,a}\right) \quad (\text{II-6})$$

avec $U_{c-1,0} = 1$.

L'algorithme d'apprentissage est basé sur la méthode du gradient [18,46], où il s'agit de trouver les $W_{c,j}^i$ qui minimisent l'erreur. Pour un exemple g , cette dernière est donnée par :

$$E_g = \frac{1}{2} \sum_{q=1}^{N_L} (U_{L,q}(X^g) - d_q^g)^2 \quad (\text{II-7})$$

$U_{L,q}(X^g)$ est la sortie $U_{L,q}$ obtenue pour l'entrée X^g .

L'erreur sur tous les exemples (ou erreur totale) est :

$$E = \frac{1}{N} \sum_{g=1}^N E_{\sigma g} \quad (\text{II-8})$$

La modification des poids d'une itération k à la suivante, se fait comme suit [18] :

$$W_{c,j}^a(k+1) = W_{c,j}^a(k) + \mu \sum_{g=1}^N \frac{dE_g}{dW_{c,j}^a} \quad (\text{II-9})$$

La dérivée de E_g est donnée par :

$$\frac{dE_g}{dW_{c,j}^a} = \frac{dE_g}{dU_{c,j}} \frac{dU_{c,j}}{dW_{c,j}^a} \quad (\text{II-10})$$

Où :

$$\frac{dU_{c,j}}{dW_{c,j}^a} = \frac{d}{dW_{c,j}^a} F\left(\sum_{m=0}^{N_{c-1}} W_{c,j}^m U_{c-1,m}\right) \quad (\text{II-11})$$

$$\frac{dU_{c,j}}{dW_{c,j}^a} = F'\left(\sum_{m=0}^{N_{c-1}} W_{c,j}^m U_{c-1,m}\right) \frac{d}{dW_{c,j}^a} \left(\sum_{m=0}^{N_{c-1}} W_{c,j}^m U_{c-1,m}\right) \quad (\text{II-12})$$

$$= F'(U_{c,j}) U_{c-1,a} \quad (\text{II-13})$$

$\frac{dE_g}{dU_{c,j}}$ est la sensibilité de E_g par rapport à $U_{c,j}$, vu l'influence du neurone dont la sortie est

$U_{c,j}$ sur les neurones de la couche suivante, on peut exprimer $\frac{dE_g}{dU_{c,j}}$ par rapport aux sorties de

ces derniers, d'où :

$$\frac{dE_g}{dU_{c,j}} = \sum_{m=1}^{N_{c+1}} \frac{dE_g}{dU_{c+1,m}} \frac{dU_{c+1,m}}{dU_{c,j}} \quad (\text{II-14})$$

Ainsi,

$$\frac{dE_g}{dU_{c,j}} = \sum_{m=1}^{N_{c+1}} \frac{dE_g}{dU_{c+1,m}} F'(U_{c+1,m}) W_{c+1,m}^j \quad (\text{II-15})$$

$$F'(x) = \begin{cases} F(x)(1-F(x)) & (\text{pour une fonction sigmoïde}). \\ 1-F^2(x) & (\text{pour la fonction tanh}). \\ 1 & (\text{pour une fonction linéaire}) \end{cases} \quad (\text{II-16})$$

Pour la couche de sortie, on a :

$$\frac{dE_g(W)}{dU_{L,j}} = U_{L,j}(X_g) - d_j^g \quad (\text{II-17})$$

L'algorithme de rétropropagation du gradient comprend les étapes suivantes [18]:

i- Initialisation des poids à des valeurs aléatoires.

ii- Présentation d'une entrée.

iii- Propagation de cette entrée et calcul de la sortie correspondante du réseau.

iv- Calcul du gradient pour la couche de sortie puis pour les autres couches pour cet exemple.

v- Présentation de l'entrée suivante et retour à l'étape 3.

Une fois que tous les exemples sont présentés, modifier les poids. Répéter l'algorithme jusqu'à la convergence (i.e. jusqu'à ce que l'erreur soit inférieure à une valeur fixée à l'avance) où jusqu'à ce que le nombre d'itérations maximales soit atteint.

Afin d'accélérer l'apprentissage, un terme supplémentaire est ajouté à l'équation (II-9) comme ceci [18] :

$$W_{c,j}^a(k+1) = W_{c,j}^a(k) + \mu \frac{dE_g}{dW_{c,j}^a} + \alpha (W_{c,j}^a(k) - W_{c,j}^a(k-1)) \quad (\text{II-18})$$

α est appelé momentum.

Soit ϕ une fonction croissante et bornée, telle que la fonction sigmoïde ou la tangente hyperbolique, soit I_p l'hypercube de dimension p $[0, 1]^p$, et soit $C(I_p)$ l'espace des fonctions continues sur I_p . Alors, pour toute fonction $f \in C(I_p)$ et pour tout $\varepsilon > 0$, existe l'entier M et des constantes réelles w_i , w_{ij} et b_i ($i = 1$ à M et $j = 1$ à p) tel qu'on définit F comme suit [18] :

$$F(X_1, \dots, X_p) = \sum_{i=1}^M w_i \phi \left(\sum_{j=1}^p w_{ij} X_j + b_i \right) \quad (\text{II-19})$$

F est une approximation de f tel que :

$$|F(X_1, \dots, X_p) - f(X_1, \dots, X_p)| < \varepsilon \quad (\text{II-20})$$

pour tous les $\{X_1, \dots, X_p\} \in I_p$.

L'analogie avec le réseau multicouches apparaît en considérant un réseau ayant, P entrées, une couche cachée ayant M neurones de fonction de transfert ϕ , et une sortie linéaire donnée par l'équation (II-19) [18]. Ce modèle est donc un approximateur universel dont l'erreur d'ajustement dépend du nombre de neurones cachés M [18,34].

II-7- Les réseaux de neurones multicouches d'ordre élevé :

Les travaux de recherche sur les réseaux de neurones ont pour but l'amélioration des algorithmes d'apprentissage, du temps de calcul et de la taille du réseau. Pour cela, les réseaux de neurones d'ordre élevé sont apparus. Ainsi, un traitement consiste à multiplier entre eux tous les éléments du vecteur d'entrée et des sorties des couches cachées par groupe de 2 (ordre 2), ..., par groupe de n (ordre n) [31].

Par exemple, dans le cas d'une entrée à 3 composantes (X_1, X_2, X_3) et pour l'ordre 2, nous aurons comme vecteur appliqué à l'entrée du réseau, le vecteur suivant :

$$(X_1, X_2, X_3, X_1^2, X_1 \cdot X_2, X_1 \cdot X_3, X_2^2, X_2 \cdot X_3, X_3^2)$$

On applique le même traitement à tous les vecteurs d'entrée d'une couche c, d'où la représentation de la figure II-4 pour un ordre n. L'ordre 1 correspond au réseau multicouches classique décrit dans le paragraphe précédent, car les sorties des neurones de chaque couche sont appliquées directement à la couche suivante sans aucun traitement.

Pour l'étude de ce réseau, on pose :

n : l'ordre du réseau.

$W_{c,j}^{a_1, \dots, a_n}$: le poids entre le neurone j de la couche c et le produit des sorties des neurones

a_1, \dots, a_n de la couche c-1.

Pour le reste, la même notation est conservée.

Pour un neurone j d'une couche c :

$$U_{c,j} = F \left(\sum_{a_1=0}^{N_{c-1}} W_{c,j}^{a_1} U_{c-1,a_1} + \sum_{a_1=1}^{N_{c-1}} \sum_{a_2=a_1}^{N_{c-1}} W_{c,j}^{a_1, a_2} U_{c-1,a_1} U_{c-1,a_2} + \dots + \sum_{a_1=1}^{N_{c-1}} \dots \sum_{a_n=a_{n-1}}^{N_{c-1}} W_{c,j}^{a_1, \dots, a_n} U_{c-1,a_1} \dots U_{c-1,a_n} \right) \quad (\text{II-21})$$

Dans l'équation ci-dessus et pour le premier terme (première somme), a_1 débute de 0 à cause du seuil et donc :

$$U_{c-1,0} = 1 \quad (\text{II-22})$$

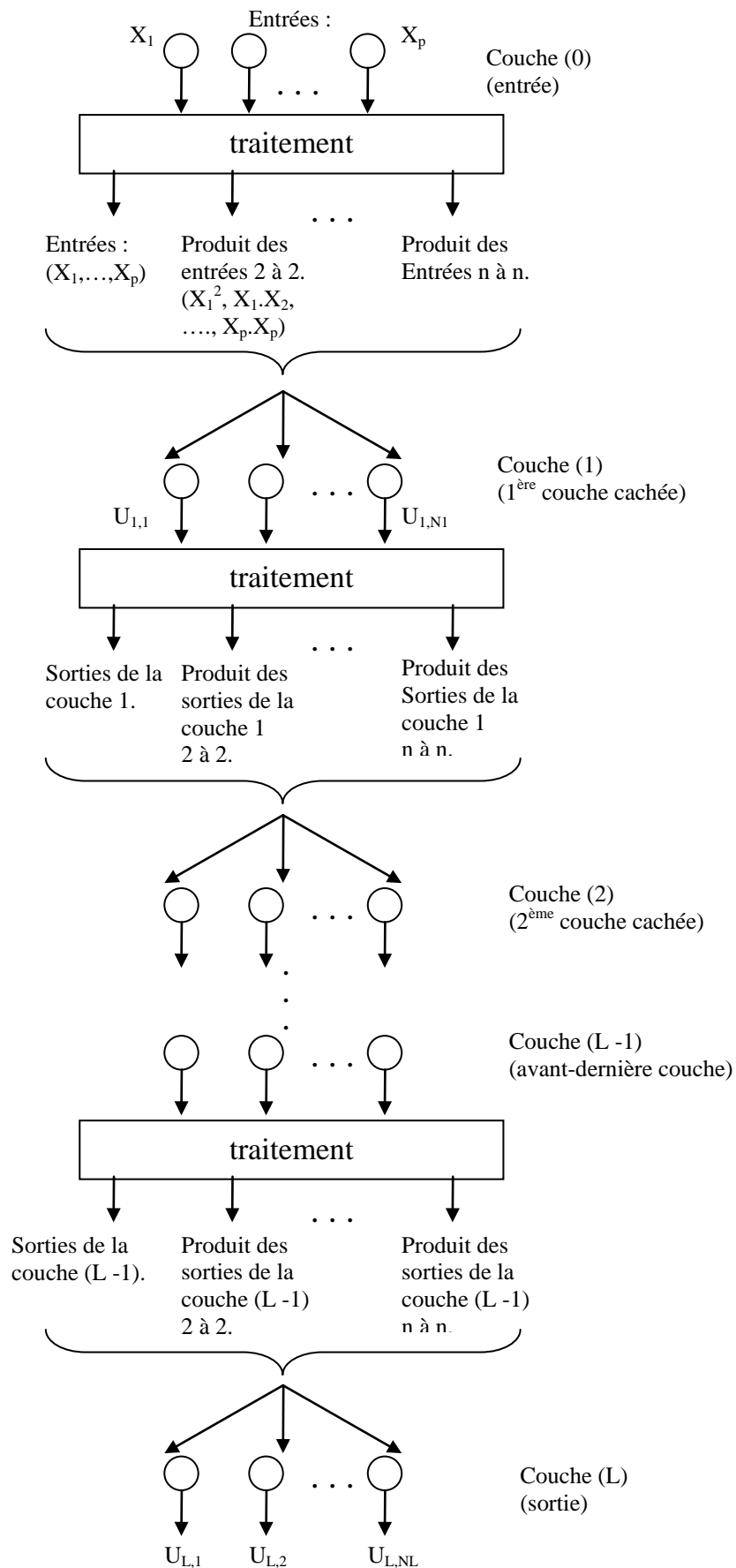


Fig. II-4 : Réseau de neurones multicouches d'ordre n.

L'algorithme d'apprentissage se base sur celui du gradient, d'où la mise à jour des poids :

$$W_{c,j}^{a_1, \dots, a_i}(k+1) = W_{c,j}^{a_1, \dots, a_i}(k) + \mu \sum_{g=1}^N \frac{dE_g}{dW_{c,j}^{a_1, \dots, a_i}} \quad (\text{II-23})$$

où $i=1, \dots, n$

L'erreur est donnée par :

$$E_g = \frac{1}{2} \sum_{q=1}^{N_L} (U_{L,q}(X^g) - d_q^g)^2 \quad (\text{II-24})$$

Le calcul du gradient est identique à celui suivi dans le chapitre précédent, on a donc :

$$\frac{dE_g}{dW_{c,j}^{a_1, \dots, a_i}} = \frac{dE_g}{dU_{c,j}} \frac{dU_{c,j}}{dW_{c,j}^{a_1, \dots, a_i}} \quad (\text{II-25})$$

avec :

$$\frac{dU_{c,j}}{dW_{c,j}^{a_1, \dots, a_i}} = F'(U_{c,j}) U_{c-1,a_1} \dots U_{c-1,a_i} \quad (\text{II-26})$$

et :

$$\frac{dE_g}{dU_{c,j}} = \sum_{m=1}^{N_{c+1}} \frac{dE_g}{dU_{c+1,m}} \frac{dU_{c+1,m}}{dU_{c,j}} \quad (\text{II-27})$$

selon la relation (II-18),

$$\begin{aligned} U_{c+1,m} &= F \left(\sum_{a_1=0}^{N_c} W_{c+1,m}^{a_1} U_{c,a_1} + \sum_{a_1=1}^{N_c} \sum_{a_2=a_1}^{N_c} W_{c+1,m}^{a_1, a_2} U_{c,a_1} U_{c,a_2} + \dots + \sum_{a_1=1}^{N_c} \dots \sum_{a_n=a_{n-1}}^{N_c} W_{c+1,m}^{a_1, \dots, a_n} U_{c,a_1} \dots U_{c,a_n} \right) \\ &= F(V) \end{aligned} \quad (\text{II-28})$$

d'où :

$$\frac{dE_g}{dU_{c,j}} = \sum_{m=1}^{N_{c+1}} \frac{dE_g}{dU_{c+1,m}} F'(U_{c+1,m}) \frac{dV}{dU_{c,j}} \quad (\text{II-29})$$

où :

$$\frac{dV}{dU_{c,j}} = \sum_{n'=1}^n \left[\sum_{h=1}^{n'} h (U_{c,j})^{h-1} \left(\sum_{a_1=1}^{N_c} \dots \sum_{a_{n'-h}=a_{n'-h-1}}^{N_c} \overbrace{W_{c+1,m}^{j, \dots, j, a_1, \dots, a_{n'-h}}}^{h \text{ fois}} U_{c,a_1} \dots U_{c,a_{n'-h}} \right) \right] \quad (\text{II-30})$$

Dans les sommes, les indices $a_1, \dots, a_{n'-h}$ ne peuvent être égaux à j , ainsi :

$a_1, \dots, a_{n'-h} = 1, \dots, j-1, j+1, \dots, N_c$. (j est exclu).

$$\text{Si } h = n', \left(\sum_{a_1=1}^{N_c} \dots \sum_{\substack{a_{n'-h}=a_{n'-h-1} \\ a_{n'-h} \neq j}}^{N_c} W_{c+1,m}^{j,a_1,\dots,a_{n'-h}} U_{c,a_1} \dots U_{c,a_{n'-h}} \right) = W_{c+1,m}^{\overbrace{j,j,\dots,j}^{\text{h fois}}}.$$

L'algorithme pour les réseaux multicouches d'ordre élevé est identique à celui décrit dans le paragraphe précédent. On note par M_c ($p \times N_c \times S \times n$) un réseau multicouches d'ordre n ayant p entrées, une couche cachée de N_c neurones et une couche de sortie de S neurones.

II-8- Les fonctions radiales de base :

Les fonctions radiales de base (radial basis functions) [18,28-30] sont connues comme étant un puissant outil dans la résolution de problèmes tels que l'interpolation et l'approximation de fonctions.

Ce type de réseau a la même structure que le réseau multicouches ayant une seule couche cachée, la fonction radiale de base est appliquée aux neurones cachés [18,28-32], elle est choisie comme étant une gaussienne définie par sa moyenne T et sa variance C^2 , la couche de sortie peut être par exemple linéaire. La détermination des paramètres du réseau suit la même procédure que celle du réseau multicouches [18]. C'est aussi un approximateur universel [18,28-29].

Soit un vecteur X ayant p composantes X_j formant la couche d'entrée du réseau RBF, et soit une couche cachée comprenant N_c neurones et une couche de sortie, la sortie Y est donnée par [18]:

$$Y = \sum_{i=1}^{N_c} W_i F(\|X - T_i\|^2) = \sum_{i=1}^{N_c} W_i \exp\left(-\frac{\sum_{j=1}^p (X_j - T_{ij})^2}{2C_i^2}\right) \quad (\text{II-31})$$

où $\|\cdot\|$ est la norme euclidienne. $T_i = (T_{i1}, \dots, T_{ip})$ est le vecteur moyenne du neurone caché i , et dont l'élément T_{ij} est le poids (la moyenne) entre l'entrée j et le neurone caché i . C_i^2 est la variance par rapport au neurone caché i . W_i est le poids liant le neurone caché i à la sortie.

La détermination des paramètres T_{ij} , C_i et W_i se fait par un apprentissage en utilisant la méthode de la décroissance du gradient. Le but étant la minimisation de l'erreur E définie par [18]:

$$E = \frac{1}{2} \sum_{g=1}^N e_{g_g} = \frac{1}{2} \sum_{g=1}^N (d_{g_g} - Y_{g_g})^2 \quad (\text{II-32})$$

où N est le nombre d'exemples, Y_g est la sortie calculée pour le $g^{\text{ème}}$ exemple d'entrée et d_g est la sortie désirée du $g^{\text{ème}}$ exemple d'entrée. Soit X^g le $g^{\text{ème}}$ exemple d'entrée (g allant de 1 à N). La modification de T_i , C_i et W_i d'une itération k à la suivante se fait donc selon les équations suivantes [18] :

$$W_i(k+1) = W_i(k) + \mu_1 \sum_{g=1}^N e_{g_g}(k) F\left(\|X^g - T_i(k)\|^2\right) \quad (\text{II-33})$$

$$T_i(k+1) = T_i(k) + 2\mu_2 W_i(k) \sum_{g=1}^N e_{g_g}(k) F'\left(\|X^g - T_i(k)\|^2\right) C_i(k) (X^g - T_i(k)) \quad (\text{II-34})$$

$$C_i(k+1) = C_i(k) + \mu_3 W_i(k) \sum_{g=1}^N e_{g_g}(k) F'\left(\|X^g - T_i(k)\|^2\right) (X^g - T_i(k)) (X^g - T_i(k))^t \quad (\text{II-35})$$

μ_1, μ_2 et μ_3 sont les taux d'apprentissage par rapport à W_i , T_i et C_i respectivement. Dans l'équation (II-35), t signifie vecteur transposé. Ce réseau sera noté par (RBF $p \times N_c$).

II-9- Les réseaux modulaires :

Les réseaux modulaires sont des modèles d'approximation de fonction non-linéaire [18], en particulier, ils peuvent être utilisés pour une régression non-linéaire dans la modélisation de séries temporelles. Leur but est de diviser les unités de calcul par module pour approcher la solution finale [18].

Chaque module calcule sa sortie comme étant une combinaison linéaire de l'entrée, puis la sortie finale du réseau est une somme pondérée des sorties des différents modules. La figure II-5 illustre un réseau modulaire comprenant M modules experts et un module superviseur, tous les modules ont une constitution identique. L'entrée X est appliquée à tous les modules.

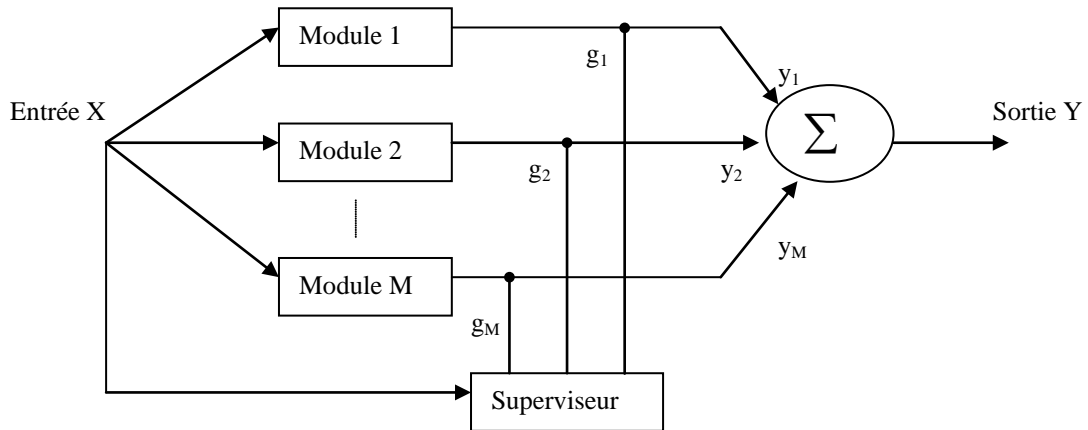


Fig. II-5 : Le réseau modulaire.

Le but de l'apprentissage est la modélisation de la distribution de probabilité des exemples d'entrée-sortie désirée (X, d) . Chaque module expert comporte p neurones d'entrée, dont l'entrée est représentée par le vecteur X (ayant p composantes) et S neurones de sortie, la sortie y_m^i de chaque neurone de sortie i pour un module m est donnée par [18]:

$$y_m^i = X^t W_m^i \quad (\text{II-36})$$

$W_m^i = (W_{m,1}^i, \dots, W_{m,p}^i)^t$ étant un vecteur de p composantes relatif au neurone i du module m .

La $i^{\text{ème}}$ sortie du réseau est :

$$Y^i = \sum_{m=1}^M g_m y_m^i \quad (\text{II-37})$$

où g_m est la probabilité à priori que le module m génère la sortie désirée d associée à l'entrée X . Elle est donnée par [18]:

$$g_m = \frac{\exp(X^t a_m)}{\sum_{m=1}^M \exp(X^t a_m)} \quad (\text{II-38})$$

a_m étant le vecteur poids ayant p composantes du module superviseur relatif au module expert m . La probabilité de distribution de la sortie désirée ayant l'entrée X et le $m^{\text{ème}}$ module expert est choisie comme étant [18]:

$$P(d|X, m) = \frac{1}{(2\pi)^{S/2}} e^{-\frac{1}{2} \|d - y_m\|^2} \quad (\text{II-39})$$

L'application de la méthode du gradient et la méthode du maximum de vraisemblance conduit aux deux équations représentant la modification des paramètres W_m^i et a_m d'une itération k à la suivante [18]:

$$W_m^i(k+1) = W_m^i(k) + \mu h_m(k) e_m^i(k) X \quad (\text{II-40})$$

$$a_m(k+1) = a_m(k) + \mu (h_m(k) - g_m(k)) X \quad (\text{II-41})$$

Avec : μ étant le taux d'apprentissage. $e_m^i(k)$ et $h_m(k)$ sont donnés par les expressions (II-42) et (II-43) :

$$e_m^i(k) = d^i - y_m^i(k) \quad (\text{II-42})$$

et

$$h_m(k) = \frac{g_m(k) \exp\left(-\frac{1}{2} \|d - y_m(k)\|^2\right)}{\sum_{j=1}^M g_j(k) \exp\left(-\frac{1}{2} \|d - y_j(k)\|^2\right)} \quad (\text{II-43})$$

d^i est la $i^{\text{ème}}$ composante du vecteur d .

Ce réseau sera noté (MOD $p \times M \times S$).

II-10- Les réseaux récurrents :

Soit X_t une série temporelle, la modélisation linéaire autorégressive à moyenne mobile (ARMA(p,q): AutoRegressive Moving Average) [34,38-39] se traduit par l'équation suivante :

$$X_t = C + \sum_{i=1}^p A_i X_{t-i} + \sum_{j=1}^q B_j e_{t-j} + e_t \quad (\text{II-44})$$

e_t est un bruit blanc de moyenne nulle. C , A_i et B_j sont les paramètres du modèle. Ainsi, X_t est une combinaison linéaire de ses valeurs et erreurs précédentes avec en plus un bruit e_t .

Si une modélisation ARMA non-linéaire (Nonlinear ARMA) s'avère nécessaire [18,34], une fonction F doit être trouvée comme le montre l'équation (II-45) :

$$X_t = F(X_{t-1}, \dots, X_{t-p}, e_{t-1}, \dots, e_{t-q}) + e_t \tag{II-45}$$

Cette fonction F peut être représentée par le biais des réseaux récurrents. Pour cela, on considère un réseau multicouche bouclé ayant 3 couches : une couche d'entrée, une cachée ayant N_c neurones cachés et une de sortie, comme le montre la figure II-6 [34]. La couche de sortie est bouclée sur la couche d'entrée, des termes d'erreur notés e_i sont donc présents en entrée comme le montre la figure II-6 et l'équation (II-46).

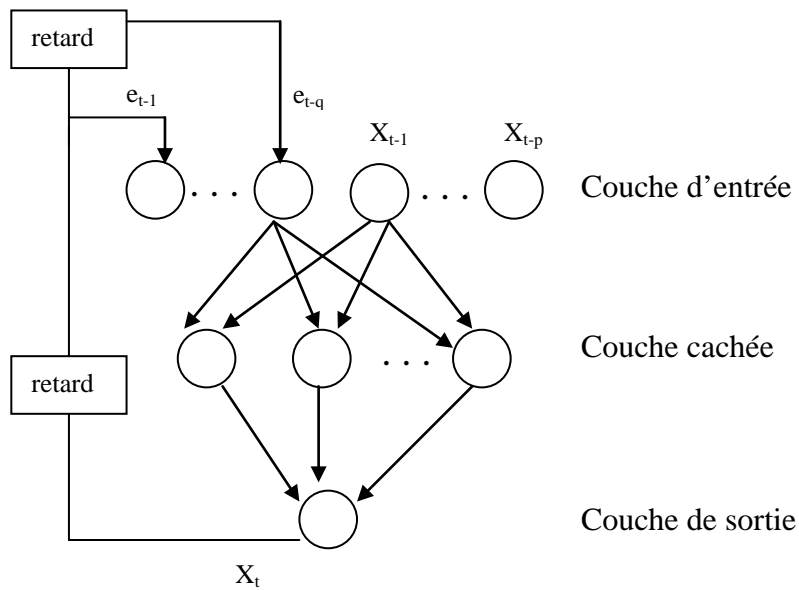


Fig. II-6 : Le réseau récurrent.

Cette structure comporte $(p + q)$ neurones d'entrée, qui sont [34] :

$$\text{Entrée} = \begin{cases} X_{t-1}, \dots, X_{t-p} \\ e_{t-1}, \dots, e_{t-q} \end{cases} \tag{II-46}$$

Avec :

$$e_{t-q} = X_{t-q} - \hat{X}_{t-q} \tag{II-47}$$

et \hat{X}_i la sortie simulée du $i^{\text{ème}}$ exemple.

Les poids du réseau sont déterminés par la méthode du gradient classique comme dans le cas du réseau multicouche, minimisant l'erreur quadratique moyenne

$$E = \frac{1}{2} \sum_{i=1}^N (X_i - \hat{X}_i)^2 \tag{II-48}$$

[34], tout en utilisant comme conditions initiales :

$\hat{X}_0 = \dots = \hat{X}_{-p+1} = e_0 = \dots = e_{-q+1} = 0$ [34]. Dans la suite, on représente ce réseau par NARMA ($p \times q \times N_c$).

II-11- Conclusion :

Les réseaux présentés dans ce chapitre sont applicables à la modélisation de séries temporelles, contribuant ainsi à la modélisation du rayonnement solaire. Le chapitre suivant comporte les résultats obtenus pour les différents réseaux et paramètres.

Chapitre III : Description neuronale du rayonnement solaire.

III-1- Introduction :

Le but de ce travail est la modélisation du rayonnement solaire grâce à la technique des réseaux de neurones. Nous avons appliqué cette méthode uniquement aux données d'irradiation dans le cas d'une modélisation monodimensionnelle, vu que l'insolation et l'irradiation sont reliées par la régression linéaire. Dans le cas multidimensionnel, les deux séries doivent être utilisées simultanément, pour cela, le réseau multicouches s'y est prêté. Les réseaux utilisés sont ceux présentés dans le chapitre II, à savoir : les réseaux de neurones multicouches, les réseaux multicouches d'ordre élevé, les réseaux à fonction radiale de base, les réseaux modulaires et les réseaux récurrents.

Le but de cette modélisation est la génération de séquences synthétiques, offrant les mêmes caractéristiques statistiques que la série réelle [10]. Tout ceci est décrit dans les paragraphes qui suivent.

III-2- Mise en œuvre des réseaux de neurones dans la modélisation :

S'agissant de trouver un modèle mathématique liant l'irradiation du jour « J » à celle des jours précédents, la modélisation d'une série temporelle X_t (en l'occurrence la fraction d'irradiation journalière HH_j) partant de ses p valeurs précédentes, peut se traduire par l'équation suivante [34,40-41,43] :

$$X_t = F(X_{t-1}, \dots, X_{t-p}) + e_t \quad (\text{III-1})$$

F est une fonction inconnue à déterminer, e_t est un bruit blanc de moyenne m nulle, de variance σ^2 et satisfaisant $E(e_t | X_{t-1}, \dots, X_{t-p}) = 0$ (e_t est indépendant des valeurs passées de X_t). Dans le cas des processus autorégressifs (prédiction linéaire), F est une fonction linéaire. Autrement, l'équation (III-1) est une modélisation autoregressive non-linéaire [18,34].

Ainsi, la modélisation de X_t à partir de ses p valeurs précédentes au sens des moindres carrés se traduit par [34,40-41] :

$$\hat{X}_t = E(X_t | X_{t-1}, \dots, X_{t-p}) = F(X_{t-1}, \dots, X_{t-p}) \quad t \geq p+1 \quad (\text{III-2})$$

E étant l'opérateur espérance mathématique. Ce prédicteur a pour erreur quadratique moyenne σ^2 .

Les réseaux de neurones permettent la réalisation de cet objectif, car en déterminant les poids synaptiques, ils trouvent une fonction liant une sortie de dimension S à une entrée de dimension p à partir d'un échantillon N. L'objectif étant la minimisation de la variance σ^2 de l'erreur e_t . Nous pouvons écrire :

$$\hat{X}_t = \hat{F}(X_{t-1}, \dots, X_{t-p}) \quad (\text{III-3})$$

où \hat{F} est une estimation de F, elle dépend du réseau (multicouches, rbf, récurrents), et est définie par la structure du réseau en question. Selon le nombre de séries, la modélisation peut être monodimensionnelle ou multidimensionnelle.

En effet, deux types de modélisations ont été effectués : monodimensionnelle et multidimensionnelle [42,47-48]. Dans le cas monodimensionnel, une seule série est prise en compte, sa modélisation ne se fait qu'à partir de ses propres valeurs antécédentes. Cette modélisation n'a été faite que pour les fractions d'irradiation étant donné que la fraction d'insolation est liée à la fraction d'irradiation par une régression linéaire (relation (I-7) du chapitre I).

Pour ce qui est de la modélisation multidimensionnelle, toutes les séries sont prises en compte simultanément : les fractions d'irradiation journalière et d'insolation sont estimées en même temps par le réseau multicouches, à partir de leurs valeurs précédentes.

De ce fait, la prédiction de la fraction d'irradiation du jour « J » s'est faite à partir de ses valeurs antécédentes (selon la valeur de p) et du bruit e_t , en essayant plusieurs réseaux et en changeant les paramètres de chaque réseau.

III-3- Confection d'une année type :

A partir des données dont on dispose, nous avons créé une année type pour HH et une autre pour SS servant à l'apprentissage, à l'aide des années 1972 à 1981. La dernière année quant à elle, servira comme année de test. La figure III-1 représente l'évolution de HH et de SS pour les années type et test.

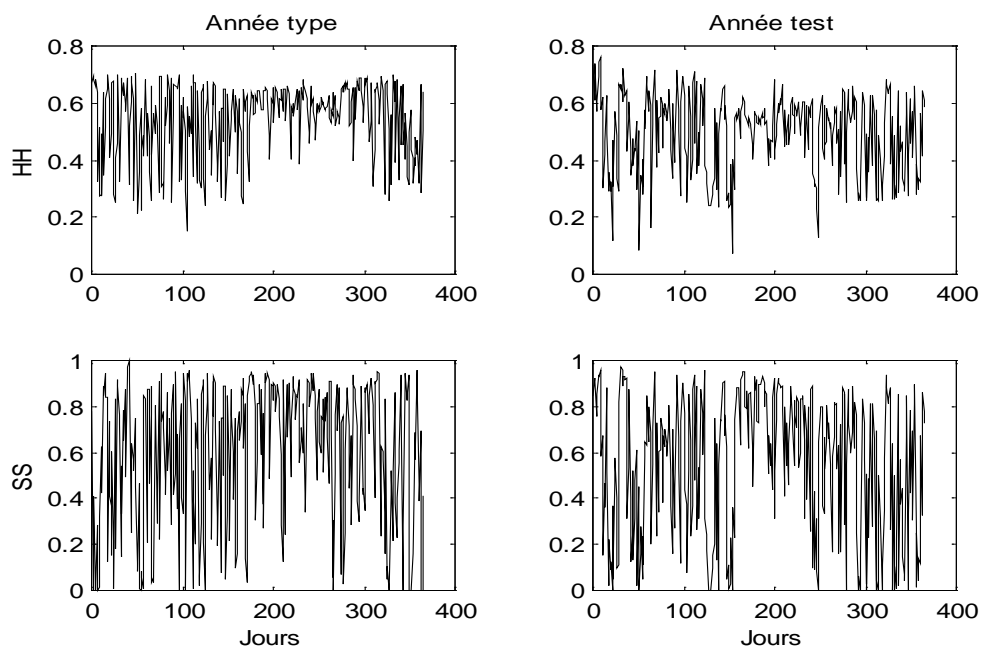


Fig. III-1 : Evolution annuelle des années type et test pour les séries HH et SS.

Le tableau III-1 regroupe les paramètres statistiques (moyenne, variance et coefficient d'autocorrélation de décalage égal à 1) des fractions d'irradiation et d'insolation, pour les années type et test, ainsi que pour la série de données HH cumulées de 1972 à 1981.

série	année	moyenne	variance	auto-corrélation
HH	72-81	0.53	0.019	0.32
	Type	0.54	0.016	0.27
	Test	0.49	0.018	0.43
SS	Type	0.60	0.091	0.37
	Test	0.55	0.090	0.49

Tabl. III-1 : Caractéristiques statistiques des séries HH et SS mesurées.

Nous avons pu obtenir des données d'irradiation plus récentes s'étalant de 1990 à 2000. En confectionnant l'année type de cette période de la même manière que pour l'autre période, nous avons trouvé que les paramètres statistiques (moyenne, variance et autocorrélation de décalage égal à 1) sont identiques pour les 2 années types.

III-4- Modélisation de HH par les réseaux de neurones multicouches :

Les réseaux de neurones multicouches sont utilisés de la manière suivante : tout d'abord, on choisit une architecture : nombre d'entrées p ($HH_{j-1}, \dots, HH_{j-p}$), nombre de neurones cachés N_c et enfin la fonction de transfert. On définit aussi le nombre d'itérations maximales, le taux d'apprentissage et le momentum. Puis l'apprentissage de l'année type débute. L'étape finale consiste à estimer les performances du réseau pour l'année de test, on rappelle que celle-ci n'a pas été utilisée pour la détermination des poids.

Egalement, les réseaux multicouches d'ordre élevé seront traités dans ce paragraphe. Ceci vient du fait que l'ordre 1 correspond au multicouches classique. Ainsi, l'ordre n fait partie des paramètres spécifiant l'architecture.

Pour toutes les simulations, le nombre d'itérations est de 1000, la fonction de transfert de la couche cachée est la fonction sigmoïde, la sortie quant à elle est linéaire et compte un seul neurone. Dans le cas de l'absence de neurones cachés, la fonction de transfert de la sortie est la sigmoïde.

Le tableau III-2 regroupe la variance du bruit σ^2 obtenue pour les différentes structures, de même que le nombre de poids définissant le réseau (nombre de paramètres).

Puis, la génération des séquences de fractions d'irradiation se fait en ajoutant le bruit e_t , après avoir sélectionné le nombre d'entrée p et le nombre de neurones cachés N_c . Ce choix est conditionné par la valeur de σ^2 obtenue et de la taille du réseau. En effet, un nombre élevé de neurones cachés pourrait conduire à une mauvaise généralisation des données de test [21]. Alors qu'un faible nombre de neurones cachés ne reproduirait pas suffisamment les variations de la série d'apprentissage.

Pour la validation, on compare la moyenne, la variance et le coefficient d'autocorrélation de décalage égal à 1 (lag 1) des séries mesurées et simulées. On calcule également l'erreur quadratique moyenne Mse entre ces deux séries et on leur applique le test de Kolmogorov-Smirnov [36] dont K_s est l'écart maximum entre les distributions des séries mesurées et simulées, pour l'année type et l'année de test.

simulation n°	n	p	Nc	σ^2	nombre de paramètres
1	1	1	0	0.0150	2
2	1	1	2	0.0148	7
3	1	1	3	0.0148	10
4	1	1	5	0.0149	16
5	1	1	8	0.0147	25
6	1	2	3	0.0149	13
7	1	2	8	0.0146	33
8	1	2	15	0.0146	61
9	1	3	5	0.0148	26
10	1	3	8	0.0147	41
11	2	1	0	0.0148	3
12	2	1	2	0.0146	12
13	5	1	0	0.0148	6
14	5	2	0	0.0146	21

Tabl. III-2 : Evolution de σ^2 pour différentes structures du réseau multicouches.

Le tableau III-3 présente les propriétés statistiques des séries réelle et simulée des années type et test dans le cas de la simulation n° 4.

année	n	p	Nc	moyenne	variance	auto- corrélacion	K_s	Mse
Type	1	1	5	0.54	0.016	0.18	0.17	0.014
Test				0.54	0.016	0.24	0.14	0.017

Tabl. III-3 : Paramètres statistiques de la série HH
pour la simulation n° 4 du réseau multicouches.

La figure III-2 illustre l'évolution des séries HH réelle et simulée, ainsi que leur fonction cumulative, pour l'année type et l'année de test pour la 4^{ème} simulation. La figure III-3 représente l'évolution de la variance σ^2 en fonction des itérations, toujours pour la 4^{ème} simulation.

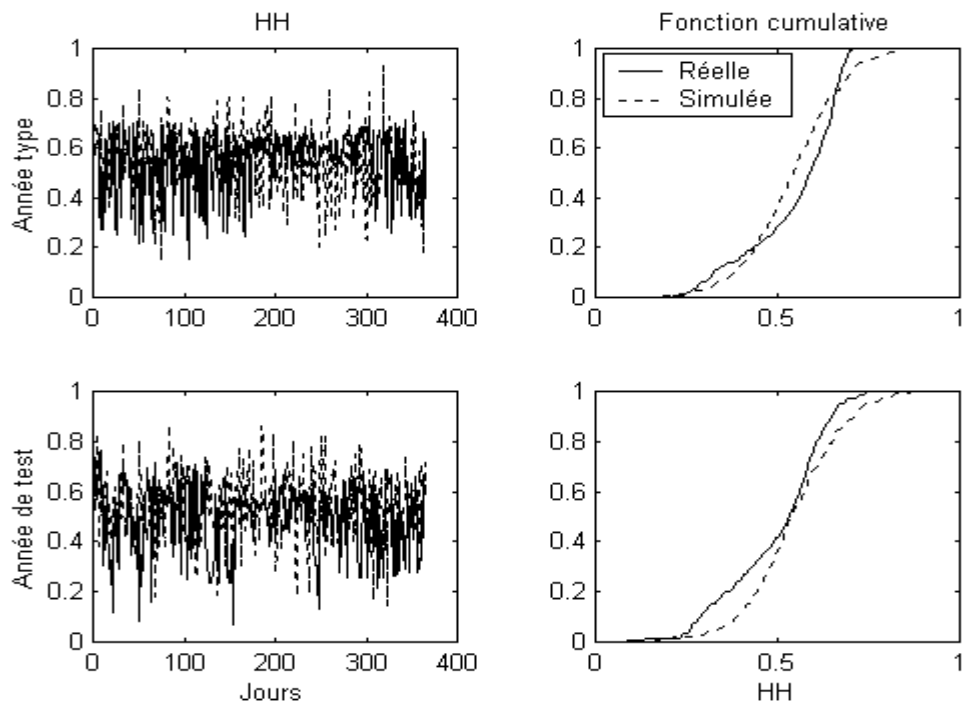


Fig. III-2 : Evolution des séries réelles et mesurées ainsi que leur fonction cumulative pour la simulation n° 4 du réseau multicouches.

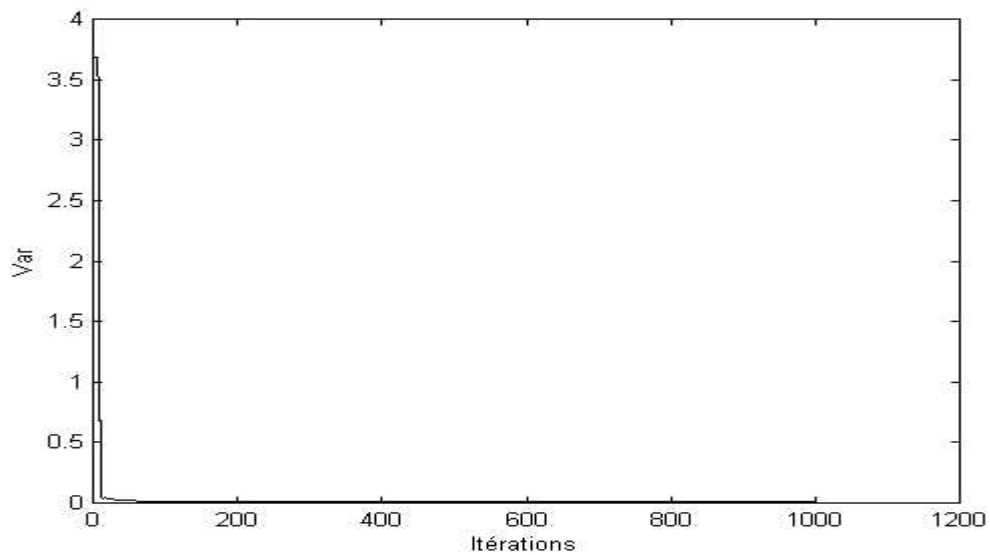


Fig. III-3 : Evolution de σ^2 en fonction du nombre d'itérations pour la simulation n° 4 du réseau multicouches.

III-5- Modélisation de HH par les fonctions radiales de base :

Les seuls paramètres à présenter sont les taux d'apprentissage des poids W , des moyennes T et des variances C^2 . La modélisation suit les mêmes étapes que pour le multicouches, p étant le nombre d'entrées et N_c le nombre de neurones cachés. Le nombre d'itérations est de 1000. La variance σ^2 est donnée dans le tableau III-4 pour différentes valeurs de p et N_c .

simulation n°	p	Nc	σ^2	nombre de paramètres
1	1	2	0.0152	7
2	1	3	0.0150	10
3	1	5	0.0152	16
4	1	8	0.0150	25
5	2	3	0.0152	13
6	2	5	0.0150	21
7	3	5	0.0150	26

Tabl. III-4 : Evolution de σ^2 pour différentes structures du réseau RBF.

Le tableau III-5 ci-dessous représente les paramètres statistiques de la série HH pour les années type et test dans le cas de la simulation n°2 du réseau RBF.

année	p	Nc	moyenne	variance	auto-corrélation	Ks	Mse
Type	1	3	0.55	0.015	0.15	0.16	0.015
Test			0.55	0.015	0.16	0.14	0.016

Tabl. III-5 : Paramètres statistiques de la série HH pour la simulation n° 2 du réseau RBF.

Les figures III-4 et III-5 représentent respectivement les évolutions des séries HH (réelle et simulée) pour les données d'apprentissage et de test ainsi que leurs fonctions cumulatives et l'évolution de σ^2 en fonction des itérations pour la 2^{ème} simulation du réseau RBF.

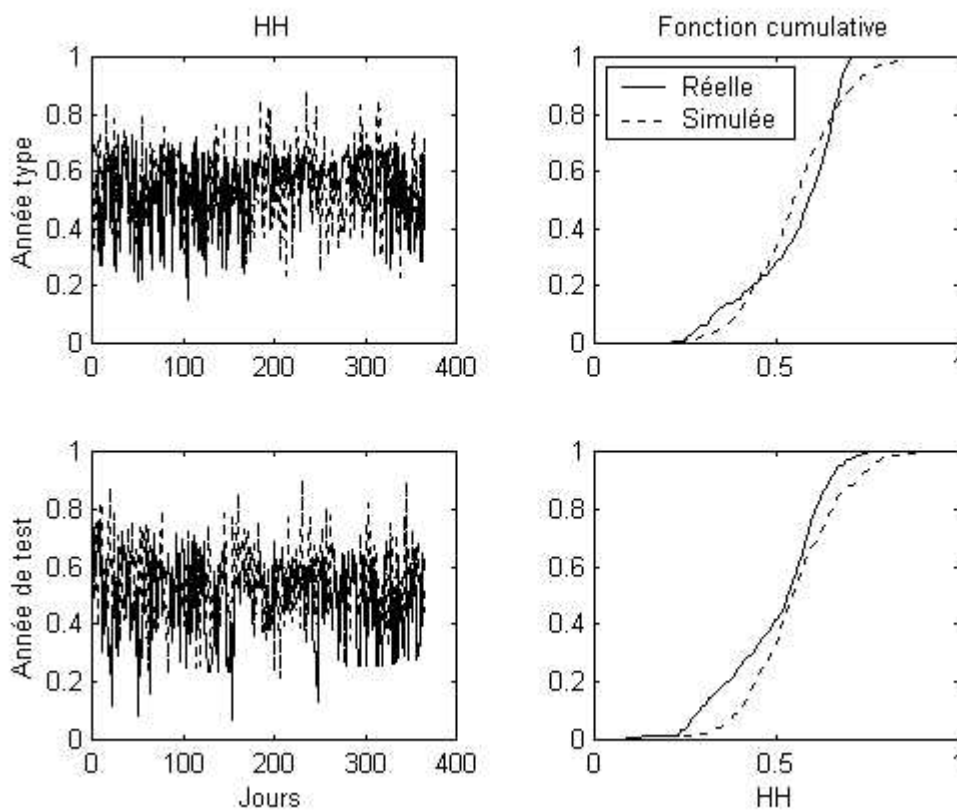


Fig. III-4 : Evolution des séries réelles et mesurées ainsi que leur fonction cumulative pour la simulation n° 2 du réseau RBF.

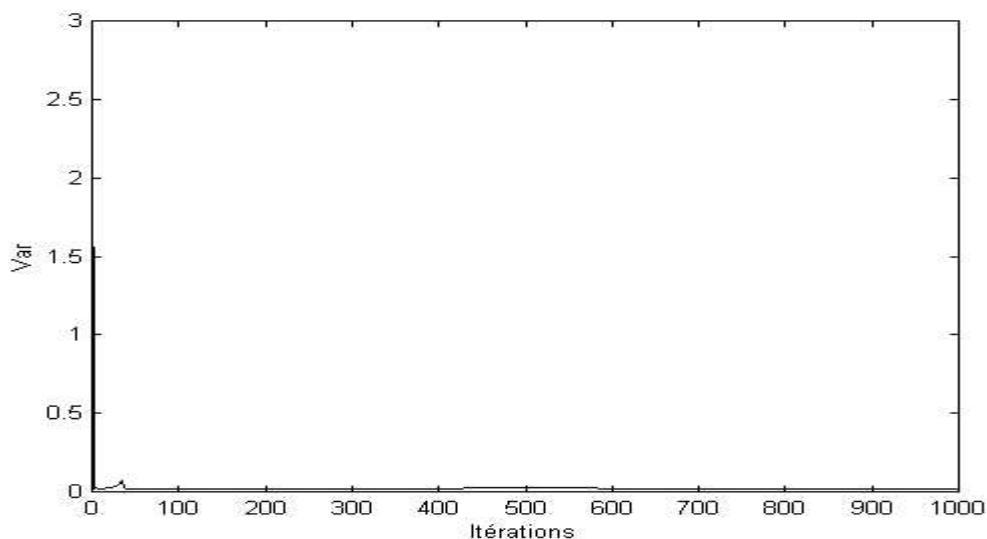


Fig. III-5 : Evolution de σ^2 en fonction du nombre d'itérations pour la simulation n° 2 du réseau RBF.

III-6- Modélisation de HH par les réseaux modulaires :

Le paramètre propre à ce type de réseau est le nombre de modules M . Le nombre d'itérations est 1000 et le nombre de neurones de sortie S est égal à 1. Les résultats obtenus sont résumés dans le tableau III-6.

simulation n°	p	M	σ^2	nombre de paramètres
1	1	3	0.0162	7
2	1	5	0.0152	11
3	1	10	0.0152	21
4	2	3	0.0157	14
5	3	5	0.0166	33

Tabl. III-6 : Evolution de σ^2 pour différentes structures du réseau modulaire.

Les paramètres statistiques de la série HH résultant de la simulation n° 2 du réseau modulaire, sont présentés dans le tableau III-7.

année	p	M	moyenne	variance	auto- corrélacion	Ks	Mse
Type	1	5	0.53	0.014	0.28	0.17	0.011
Test			0.48	0.017	0.43	0.10	0.010

Tabl. III-7 : Paramètres statistiques de la série HH
pour la simulation n° 2 du réseau modulaire.

Dans la figure III-6, l'évolution de la série HH pour les années type et test dans le cas de la simulation n° 2 du réseau modulaire est affichée, de même que leur fonction cumulative. Puis la figure III-7 donne l'évolution de l'erreur σ^2 en fonction du nombre d'itérations, toujours pour la 2^{ème} simulation.

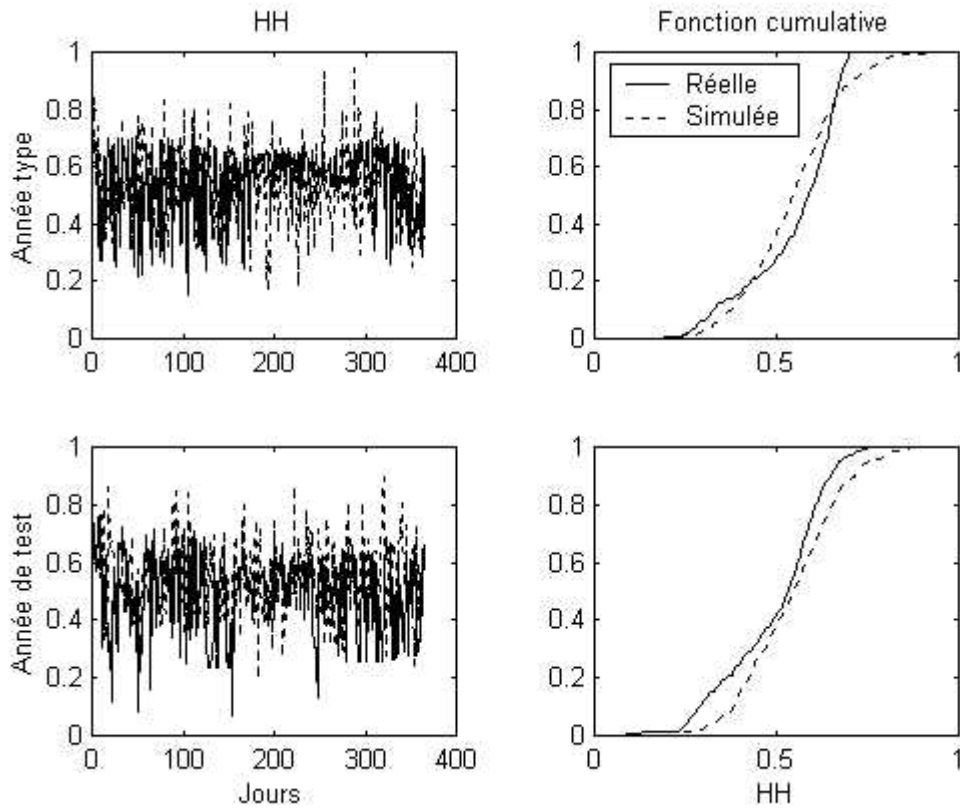


Fig. III-6 : Evolution des séries réelles et mesurées ainsi que leur fonction cumulative pour la simulation n° 2 du réseau modulaire.

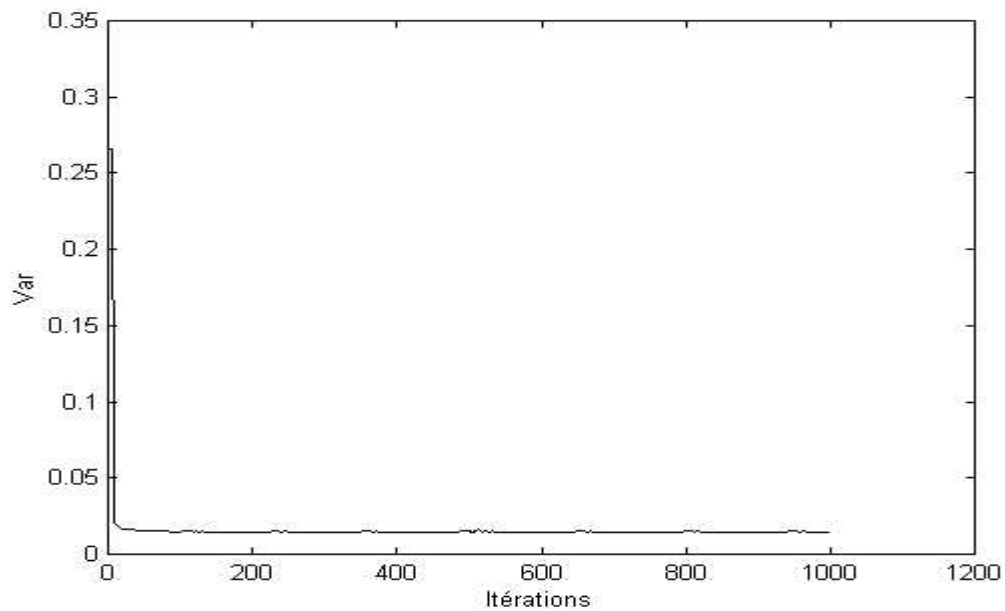


Fig. III-7 : Evolution de σ^2 en fonction du nombre d'itérations pour la simulation n° 2 du réseau modulaire.

III-7- Modélisation de HH par les réseaux récurrents :

En plus des paramètres du réseau multicouches, les réseaux récurrents nécessitent l'ajout d'une nouvelle variable notée q qui représente le nombre d'erreurs passées à introduire. Là encore, la couche de sortie comporte une seule unité, et le nombre d'itérations est de 500. Le tableau III-8 présente l'évolution de l'erreur σ^2 pour différentes valeurs de p , q et N_c .

simulation n°	p	q	N_c	σ^2	nombre de paramètres
1	1	1	0	0.0172	3
2	0	1	0	0.0146	2
3	1	1	2	0.0136	9
4	0	1	2	0.0134	7

Tabl. III-8 : Evolution de σ^2 pour différentes structures du réseau récurrent.

La génération des séquences de la série HH à l'aide de la simulation n° 3 du réseau récurrent a conduit aux caractéristiques statistiques données au tableau III-9.

année	p	q	N_c	moyenne	variance	auto- corrélacion	Ks	Mse
Type	1	1	2	0.55	0.013	0.07	0.15	0.015
Test				0.55	0.014	0.15	0.16	0.018

Tabl. III-9 : Paramètres statistiques de la série HH
pour la simulation n° 3 du réseau récurrent.

L'évolution de la fraction HH des 2 années (type et test) ainsi que les fonctions cumulatives sont données dans la figure III-8. Quant à la figure III-9, elle représente la variation de σ^2 en fonction des itérations. Toutes deux dans le cas de la simulation n° 3.

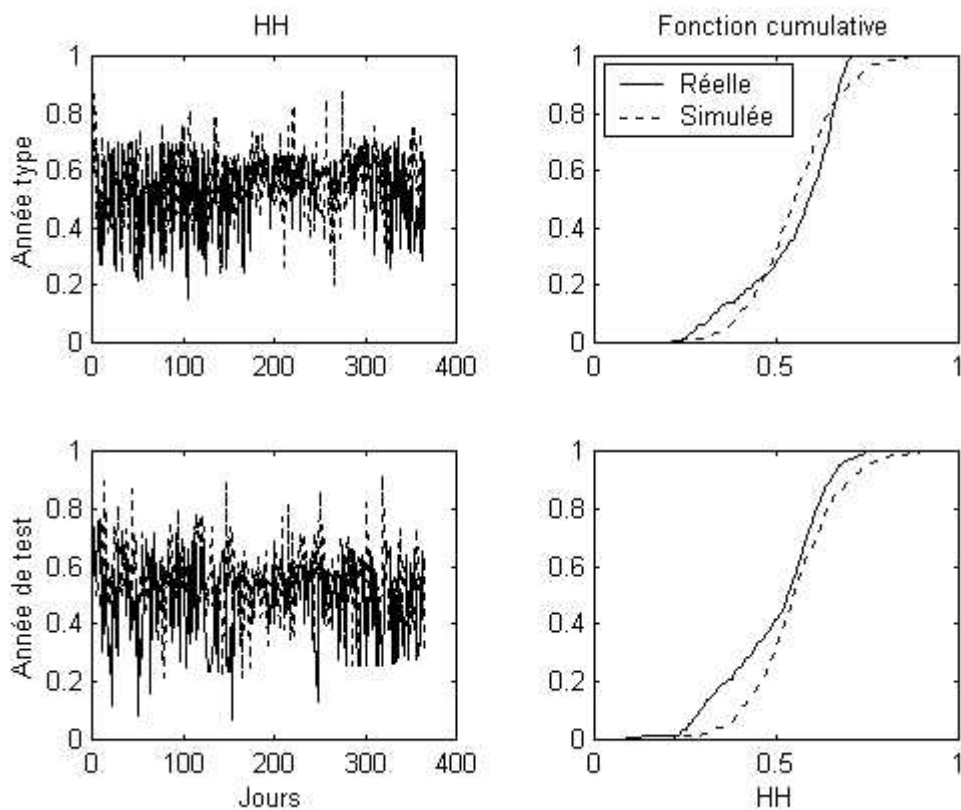


Fig. III-8 : Evolution des séries réelles et mesurées ainsi que leur fonction cumulative pour la simulation n° 3 du réseau récurrent.

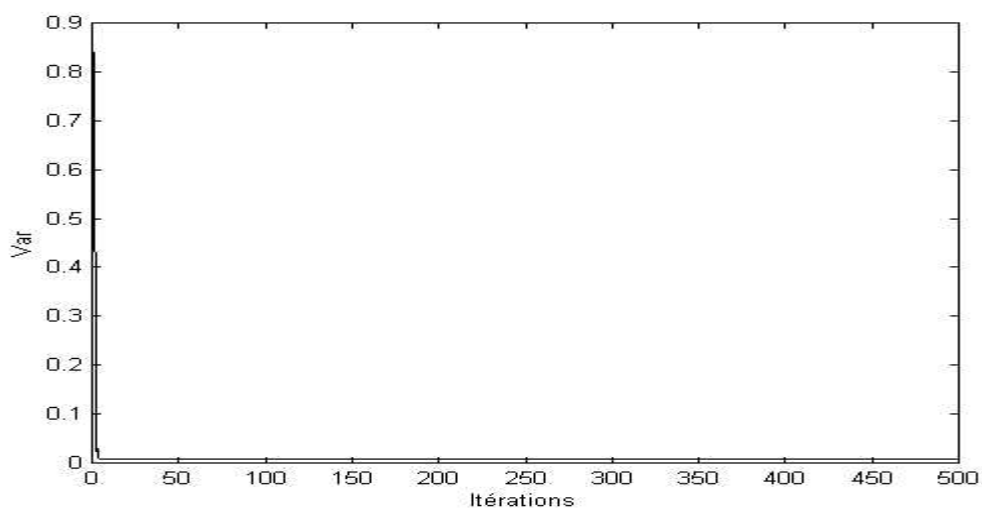


Fig. III-9 : Evolution de σ^2 en fonction du nombre d'itérations pour la simulation n° 3 du réseau récurrent.

III-8- Modélisation bidimensionnelle de HH et de SS par les réseaux de neurones multicouches :

La modélisation multidimensionnelle a pour but d'améliorer la précision des résultats de la modélisation monodimensionnelle, car l'introduction de séries supplémentaires pourrait reproduire de façon plus précise les fluctuations aléatoires de cette même série. Pour notre part, la deuxième série est la fraction d'insolation. Il s'agit donc de modéliser les fractions d'irradiation et d'insolation HH_j et SS_j à partir de leurs valeurs précédentes HH_{j-1} , $SS_{j-1}, \dots, HH_{j-p}$, SS_{j-p} , par le réseau multicouches. Disposant en sortie des fractions HH_j et SS_j , et en entrée des p fractions précédentes, il est clair que le nombre de neurones d'entrée est de $2p$ et le nombre de neurones de sortie est de 2. Le nombre d'itérations est de 1000. Le tableau III-10 regroupe les valeurs de σ^2 obtenues en changeant la structure du réseau.

simulation n°	p	Nc	σ^2	nombre de paramètres
1	1	0	0.0931	6
2	1	3	0.0950	17
3	1	5	0.0928	27
4	1	8	0.0918	42
5	2	3	0.0937	23
6	3	3	0.0944	29

Tabl. III-10 : Evolution de σ^2 pour différentes structures du réseau multicouches dans le cas bidimensionnel.

Le tableau III-11 présente les paramètres statistiques des séries HH et SS obtenus à l'aide de la simulation n°3.

série	année	p	Nc	moyenne	variance	auto-corrélation	Ks	Mse
HH	Type	1	5	0.54	0.016	0.29	0.15	0.016
	Test			0.54	0.016	0.29	0.14	0.019
SS	Type			0.60	0.093	0.40	0.15	0.090
	Test			0.60	0.093	0.40	0.12	0.086

Tabl. III-11 : Paramètres statistiques des séries HH et SS pour la simulation bidimensionnelle n° 3 .

La figure III-10 représente l'évolution de HH et SS pour l'année type et l'année de test dans le cas de la simulation n°3 pour la modélisation bidimensionnelle, ainsi que les fonctions cumulatives de chaque série.

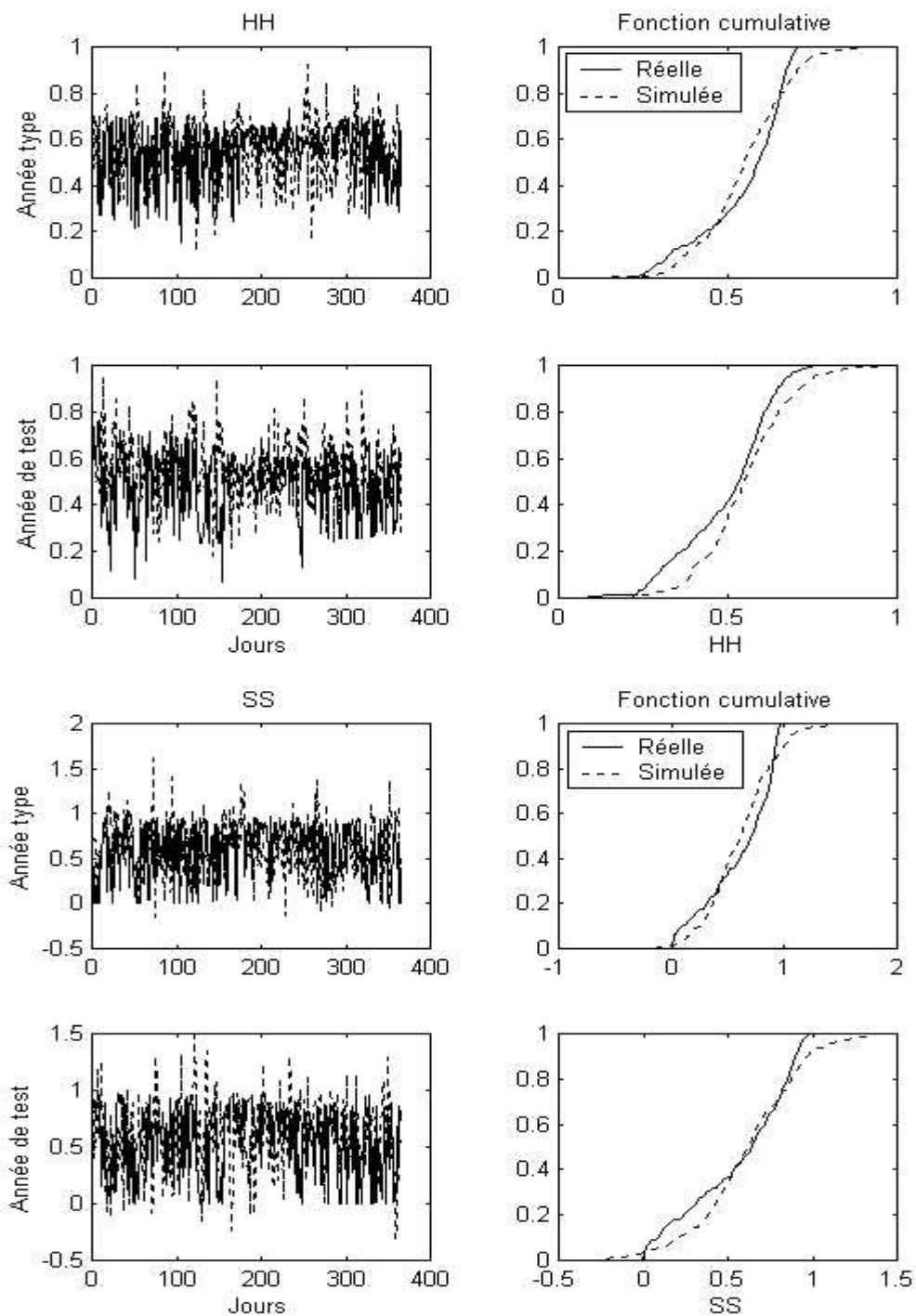


Fig. III-10 : Evolution des séries réelles et mesurées ainsi que leur fonction cumulative pour la simulation n° 3 du cas bidimensionnelle.

L'évolution de σ^2 en fonction du nombre d'itérations est donnée dans la figure III-11.

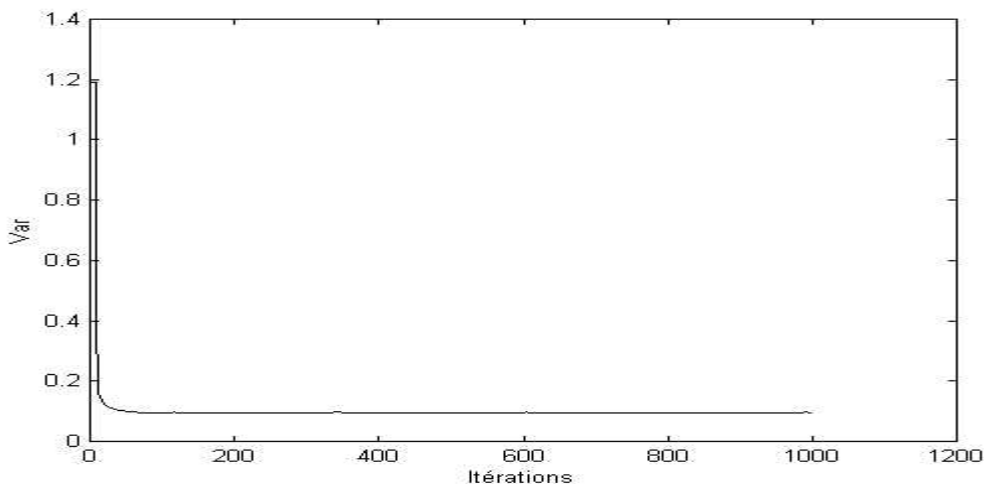


Fig. III-11 : Evolution de σ^2 en fonction du nombre d'itérations pour la simulation n° 3 du cas bidimensionnel.

III-9- Modélisation de HH à partir de SS par les réseaux multicouches :

Dans ce cas, la prédiction de HH_j s'est faite à partir de SS_j , d'où le nombre de neurones d'entrée égal à 1 ainsi que celui de la sortie. Afin de confronter les résultats à ceux de la régression linéaire, les données d'apprentissage sont les données cumulées mois par mois de 1972 à 1981. Ainsi, pour chaque mois, un réseau est calculé puis sera utilisé pour le test à l'aide de l'année 1982. Les résultats de chaque mois sont concaténés afin de former l'année simulée 1982. Le tableau III-12 regroupe les paramètres statistiques de la série HH de l'année 1982 obtenus pour différentes structures.

Nc	moyenne	variance	auto-corrélation	Mse	Ks
0 ^(L)	0.51	0.018	0.49	0.0011	0.09
0 ^(S)	0.51	0.018	0.48	0.0012	0.11
1	0.51	0.019	0.47	0.0011	0.12
3	0.51	0.019	0.48	0.0011	0.11
5	0.51	0.019	0.48	0.0011	0.11

Tabl. III-12 : Paramètres statistiques de HH simulée pour l'année de test.

^(L) : sortie linéaire, ^(S) : sortie sigmoïde.

La figure III-12 illustre l'évolution de la série HH de l'année 1982 obtenue pour la première simulation, ainsi que les fonctions cumulatives.

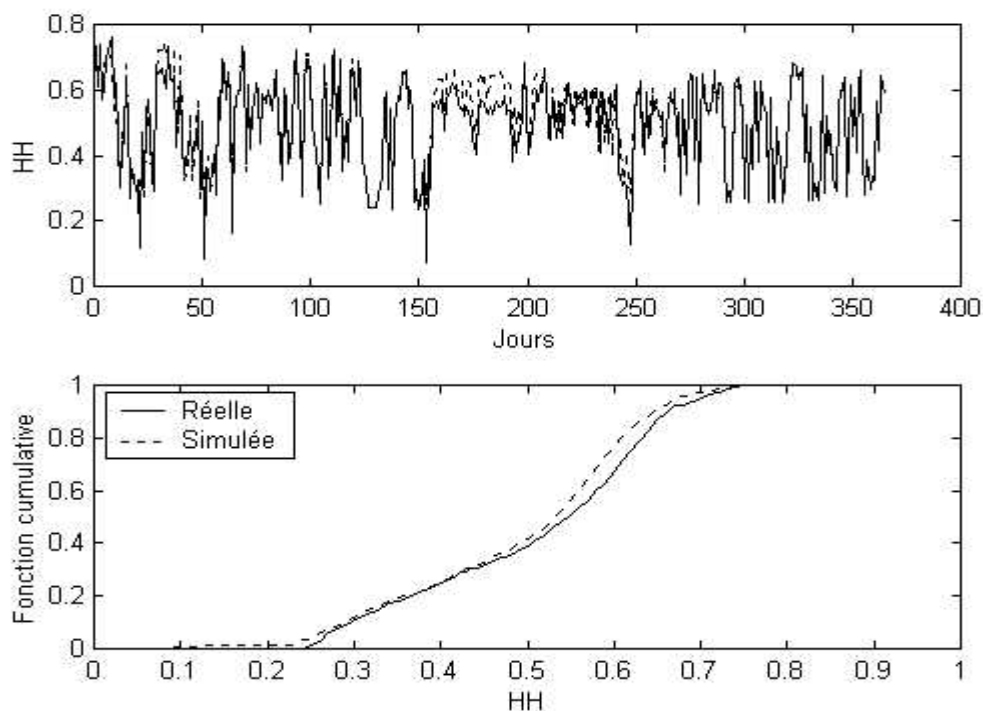


Fig. III-12 : Evolution de HH réelle et simulée ainsi que leur fonctions cumulatives pour l'année de test.

Remarque :

Les programmes informatiques ont été écrits sous MATLAB 5.3, sur un micro-ordinateur Pentium II 300 Mhz avec 32 Mo de Ram.

III-10- Conclusion :

Dans ce chapitre, l'utilisation des réseaux de neurones a été faite au profit de la modélisation de la fraction d'irradiation solaire dans les deux cas : mono et multidimensionnel. En changeant de type de réseaux et leurs paramètres spécifiques, différents résultats furent trouvés. Selon les critères de comparaison, l'analyse de ces résultats sera faite dans le prochain chapitre, la comparaison avec les chaînes de Markov et les processus autorégressifs sera également effectuée.

Chapitre IV : Analyse des résultats et prospective.

IV-1- Introduction :

Ce chapitre est consacré à l'étude comparative des résultats obtenus pour les différents réseaux, un ou plusieurs réseaux s'adaptant à la modélisation du rayonnement solaire seront sélectionnés et seront classés selon différents critères de comparaison tel que l'erreur, l'implémentation ...etc. A la fin, ces mêmes résultats seront comparés à ceux des chaînes de Markov d'ordre 1 à 2 états et des processus autorégressifs monodimensionnels et multidimensionnels d'ordre 1. A cet effet, les principes de la modélisation par les chaînes de Markov ainsi que les processus autorégressifs sont décrits dans les annexes 2, 3 et 4.

Egalement , le passage des processus autorégressifs aux réseaux de neurones et vice versa sera abordé, ainsi que l'analogie avec la régression linéaire.

IV-2- Résultats de la modélisation par les chaînes de Markov :

Les caractéristiques statistiques de la fraction d'irradiation HH simulée par les chaînes de Markov d'ordre 1 à 2 états, suivant l'algorithme [10] décrit dans l'annexe 2, sont affichées dans le tableau IV-1. Nous rappelons que le calcul des probabilités nécessite une longue série de mesures, pour cela, et à cause de la nécessité de la stationnarité, la matrice P est calculée pour chaque mois cumulé de la période 1972-1981. Puis, lors de la simulation, les mois sont cumulés et concaténés sans interruption [10]. Enfin, la fraction d'irradiation de l'année 1982 est simulée à son tour grâce à la matrice P.

année	moyenne	variance	auto- corrélation	Ks	Mse
72-81	0.50	0.039	0.20	0.11	0.029
Test	0.44	0.018	0.18	0.46	0.020

Tabl. IV-1 : Caractéristiques statistiques de la série
HH simulée par les chaînes de Markov.

La figure IV-1 illustre l'évolution de la série HH durant les mois cumulés concaténés de la période 1972-1981 et de l'année de test, ainsi que des fonctions cumulatives, obtenues à l'aide des chaînes de Markov d'ordre 1 à 2 états.

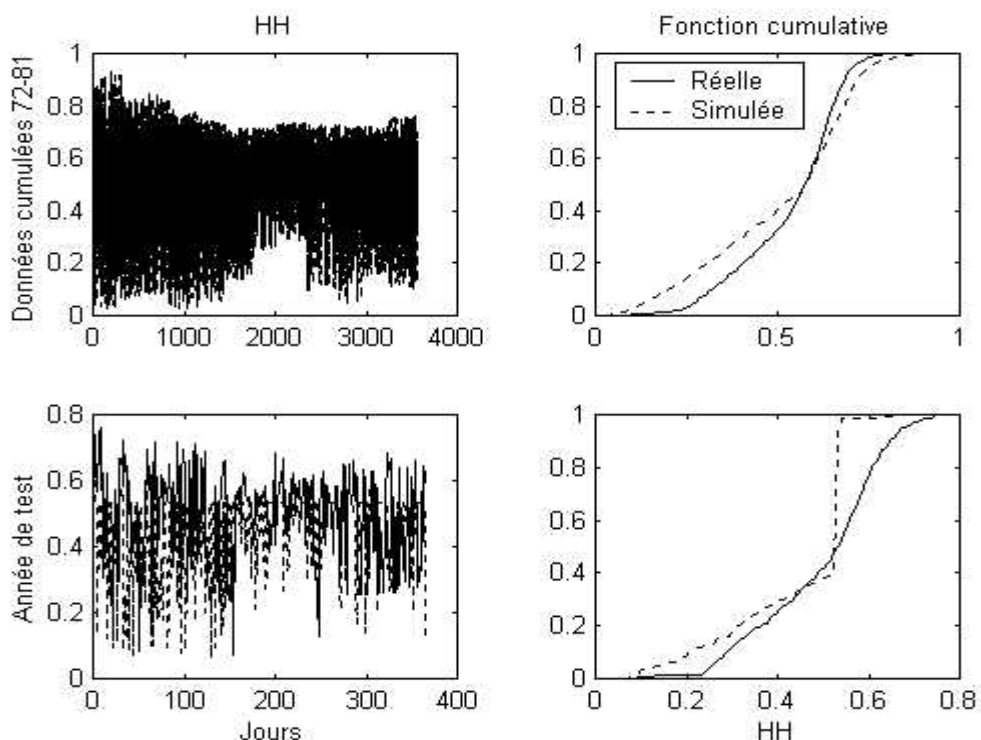


Fig. IV-1 : Evolution de la série HH des mois cumulés concaténés 1972-1981, de l'année test, ainsi que des fonctions cumulatives obtenus par les chaînes de Markov.

IV-3- Résultats de la modélisation autorégressive monodimensionnelle :

Les étapes de la simulation par les processus autorégressifs d'ordre 1 sont décrites dans l'annexe 3. Pour cela, la transformation gaussienne et son inverse sont utilisées. Elles nous ont conduits aux résultats du tableau IV-2, où les caractéristiques statistiques (moyenne, variance et autocorrélation de décalage égal à 1) de HH mesurée (réelle) et simulée pour l'année type et l'année test sont affichées.

année	moyenne	variance	auto-corrélation	Ks	Mse
Type	0.54	0.015	0.22	0.04	0.014
Test	0.54	0.015	0.26	0.21	0.017

Tabl. IV-2 : Propriétés statistiques de la série HH obtenues par le AR(1).

La figure IV-2 représente l'évolution des séries réelles et simulées par les processus autorégressifs d'ordre 1 ainsi que des fonctions cumulatives, des années type et test, obtenues par les processus AR(1).

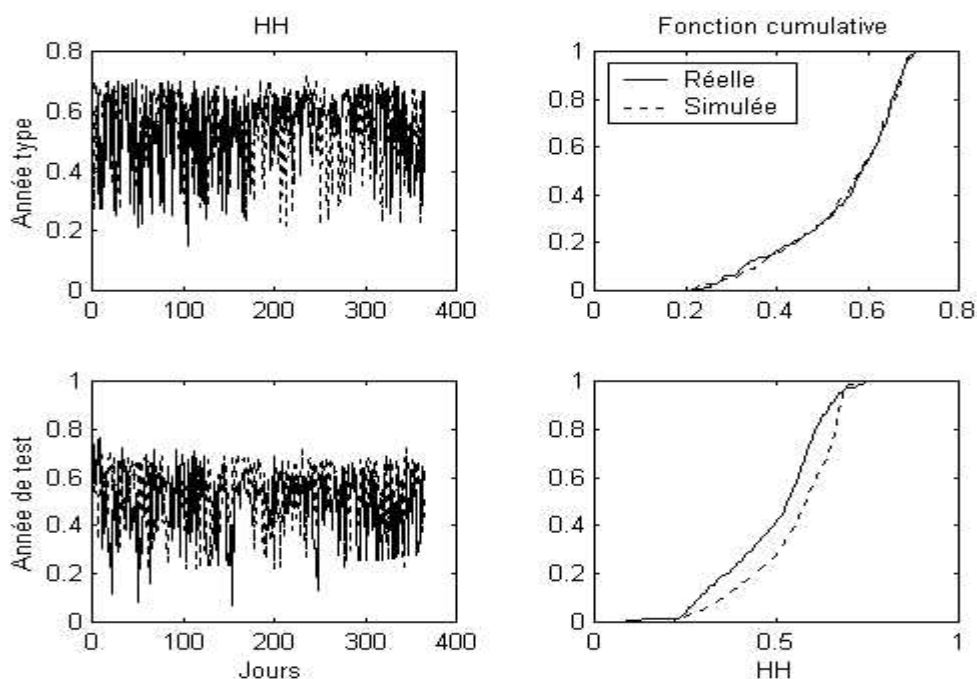


Fig. IV-2 : Evolution de la série HH et des fonctions cumulatives réelle et simulée pour les années type et test par le modèle AR(1).

IV-4- Résultats de la modélisation autoregressive multidimensionnelle :

Les caractéristiques statistiques des fractions d'irradiation HH et d'insolation SS des années type et test simulées par les processus autorégressifs multidimensionnels d'ordre 1 sont regroupées dans le tableau IV-3. Là encore, la transformée gaussienne et son inverse sont utilisées (voir annexe 3 et 4).

série	année	moyenne	variance	auto-corrélation	Ks	Mse
HH	Type	0.54	0.015	0.22	0.04	0.014
	Test	0.54	0.015	0.18	0.25	0.017
SS	Type	0.58	0.086	0.38	0.12	0.084
	Test	0.57	0.100	0.49	0.07	0.098

Tabl. IV-3 : Propriétés statistiques des séries HH et SS obtenues par le MAR(1).

La figure IV-3 représente l'évolution des séries HH et SS des années type et test simulées par les processus MAR(1), ainsi que les fonctions cumulatives.

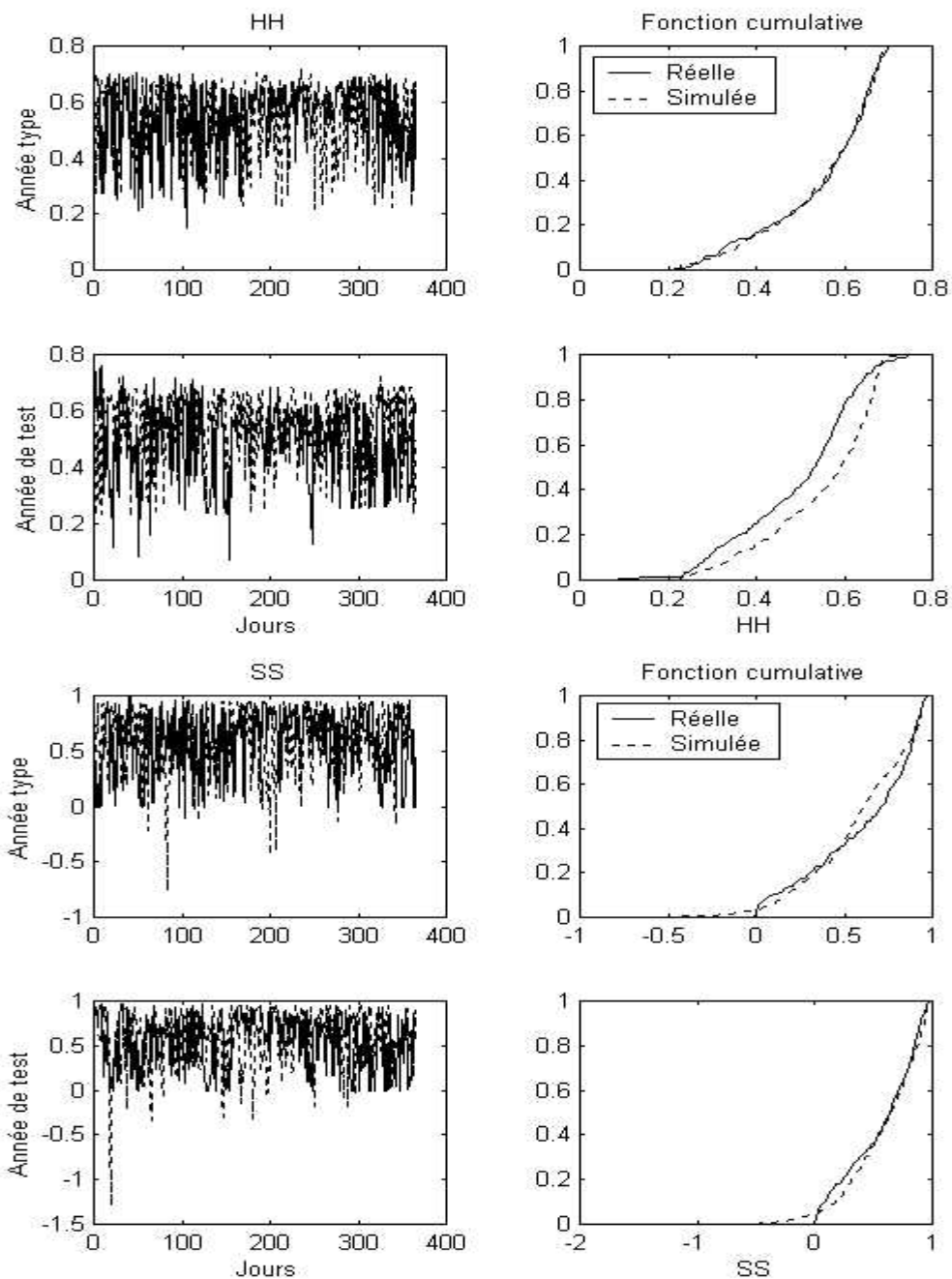


Fig. IV-3 : Evolution des séries HH et SS et des fonctions cumulatives des années type et test simulées par le MAR(1).

IV-5- Comparaison entre les différents résultats :

Il s'agit d'analyser les performances des réseaux de neurones appliqués à la modélisation monodimensionnelle et bidimensionnelle du rayonnement solaire dans le

chapitre précédent. Ceci se fait selon différents critères : erreur, coefficient K_s , moyenne, variance, coefficient d'autocorrélation et enfin le nombre de paramètres (poids) décrivant le réseau. Ensuite, la comparaison avec les chaînes de Markov d'ordre 1 à 2 états et les processus AR et MAR d'ordre 1 sera effectuée.

Tout d'abord, le tableau IV-4 rappelle les caractéristiques statistiques (moyenne, variance et autocorrélation de décalage égal à 1) de HH et SS mesurées (réelles). Puis les tableaux IV-5 et IV-6 récapitulent les résultats des différentes simulations, respectivement dans le cas monodimensionnel et multidimensionnel.

série	année	moyenne	variance	auto-corrélation
HH	72-81	0.53	0.019	0.32
	Type	0.54	0.016	0.27
	Test	0.49	0.018	0.43
SS	Type	0.60	0.091	0.37
	Test	0.55	0.090	0.49

Tabl. IV-4 : Caractéristiques statistiques des séries HH et SS mesurées.

année	structure	moyenne	variance	auto-corrélation	K_s	Mse	nombre de paramètres
Type	MC	0.54	0.016	0.18	0.17	0.014	16
Test	1x5x1x1	0.54	0.016	0.24	0.14	0.017	
Type	RBF	0.55	0.015	0.15	0.16	0.015	10
Test	1x3	0.55	0.015	0.16	0.14	0.016	
Type	MOD	0.53	0.014	0.28	0.17	0.011	11
Test	1x5x1	0.48	0.017	0.43	0.10	0.010	
Type	NARMA	0.55	0.013	0.07	0.15	0.015	9
Test	1x1x2	0.55	0.014	0.15	0.16	0.018	
Type	AR	0.54	0.015	0.22	0.04	0.014	7
Test	(1)	0.54	0.015	0.26	0.21	0.017	
72-81	Markov	0.50	0.039	0.20	0.11	0.029	5
Test	ordre 1, 2 états	0.44	0.018	0.18	0.46	0.020	par mois

Tabl. IV-5 : Synthèse des résultats des différents réseaux,
des processus AR(1) et des chaînes de Markov.

série	année	structure	moyenne	variance	auto-corrélation	Ks	Mse	nombre de paramètres
HH	Type	MC	0.54	0.016	0.29	0.15	0.016	21
	Test		0.54	0.016	0.29	0.14	0.019	
SS	Type	2x5x2x1	0.60	0.093	0.40	0.15	0.090	par série
	Test		0.60	0.093	0.40	0.12	0.086	
HH	Type	MAR	0.54	0.015	0.22	0.04	0.014	18
	Test		0.54	0.015	0.18	0.25	0.017	
SS	Type	(1)	0.58	0.086	0.38	0.12	0.084	par série
	Test		0.57	0.100	0.49	0.07	0.098	

Tabl. IV-6 : Synthèse des résultats de la modélisation multidimensionnelle.

IV-5-1- Résultats des réseaux de neurones :

En observant le tableau IV-5, nous constatons que dans tous les cas, la moyenne de l'année type est fidèlement reproduite. Mais dans le cas de l'année de test, la moyenne estimée reste proche de celle trouvée pour l'année type. Concernant la variance, les valeurs pour l'année type et de test sont proches, la valeur trouvée par le multicouches est plus proche de la valeur réelle en comparaison avec les autres réseaux.

Le coefficient Ks nous renseigne sur la distribution des données mesurées et simulées. Ainsi, pour un échantillon d'effectif N ($N > 35$), et en considérant un risque d'erreur égal à 5%, on compare Ks à la valeur $\frac{1.36}{\sqrt{N}}$. Si $Ks > \frac{1.36}{\sqrt{N}}$, on rejette l'hypothèse d'ajustement au seuil de signification de 0,05. De là et pour l'année type (N égal à 365), cette hypothèse est rejetée dans le cas des réseaux de neurones.

Considérons maintenant l'erreur quadratique moyenne comme critère de comparaison. Nous remarquons que c'est le réseau modulaire qui présente la plus petite erreur pour les séries d'apprentissage et de test.

Si l'analyse doit se faire selon le nombre de paramètres du modèle (principe de parcimonie), le réseau multicouches offre le meilleur compromis entre encombrement, précision et généralisation, sa formulation (simple) et les temps de calcul sont en son avantage par rapport aux autres réseaux.

Nous avons fait varier le nombre d'entrées p de 1 à 3, la précision des résultats ne s'est pas améliorée, pour tous les réseaux. Ainsi, une seule entrée ($p = 1$) suffit à cette modélisation. De même, dans le cas des réseaux récurrents (NARMA), l'introduction de termes d'erreur (e_{t-i}) n'a pas donné de meilleurs résultats. Pour les réseaux multicouches d'ordre élevé, l'ordre 1 suffit, ni le nombre d'itérations ni le temps de calcul ne furent améliorés en augmentant l'ordre.

IV-5-2- Comparaison avec d'autres modèles :

Tout d'abord, concernant les chaînes de Markov, les résultats obtenus en utilisant une série de 1972 à 1981 sont presque identiques avec ceux obtenus à l'aide d'une série de 11 années (1972-1982) [10]. Ce qui confirme que la série que nous avons utilisée est suffisante pour rendre compte des caractéristiques statistiques du rayonnement solaire. Dans le cas des processus autorégressifs, la transformation gaussienne et sa transformation inverse ont été utilisées.

De la comparaison de ces méthodes avec les réseaux de neurones, on note que le coefficient d'autocorrélation dans les cas des réseaux de neurones est légèrement plus faible que celui trouvé à l'aide des processus autorégressifs, ceci est dû à l'utilisation de la transformation gaussienne. C'est une transformation artificielle des données.

Concernant le test de Kolmogorov-Smirnov, les processus AR donnent les meilleurs résultats, ceci est dû aussi à l'utilisation de la transformation gaussienne, l'hypothèse d'ajustement est donc acceptée. Toutefois, pour l'année de test, le modèle AR(1) présente le coefficient K_s le plus élevé, l'hypothèse d'ajustement est rejetée car K_s est supérieur au seuil donné dans les tables. La généralisation à l'année de test est meilleure à partir des réseaux de neurones. Dans ce cas, les réseaux de neurones sont préférables aux processus AR, vu que les données ne subissent pas de transformations artificielles bien que la précision soit moindre. Les mêmes remarques restent valables pour les chaînes de Markov.

En confrontant le réseau multicouches au modèle AR(1), nous notons une légère amélioration de l'erreur pour l'année type mais pour l'année de test, les valeurs sont semblables. Les chaînes de Markov présentent la plus forte erreur pour l'année de test. Ceci

est dû à la discrétisation en 2 états, l'amplitude de la série n'est plus continue de 0 à 1, elle ne peut être égale qu'à 0 ou 1.

Nous pouvons dire que le réseau multicouches ayant une entrée et 5 neurones cachés suffit à la modélisation monodimensionnelle de la fraction d'irradiation journalière. Ainsi, seule l'irradiation du jour « j » influe sur l'irradiation du jour suivant, ce qui est en accord avec les résultats des processus AR(1) et les chaînes de Markov d'ordre 1. Ce modèle est une autorégression non-linéaire, c'est une généralisation des processus autorégressifs linéaires.

Pour le cas multidimensionnel, en comparant la moyenne, la variance et le coefficient d'autocorrélation dans le tableau IV-6, le réseau multicouches est plus performant par rapport aux processus MAR. Pour les données d'irradiation, le Ks des processus MAR est meilleur que celui du multicouches pour l'année type. Mais pour l'année de test, les résultats sont en faveur du multicouches. Par contre, pour les données d'insolation, le Ks est identique pour les 2 modèles pour l'année type, alors que pour l'année de test, le MAR est meilleur. Pour l'erreur MSE, le MAR est légèrement plus performant, le nombre de paramètres est quasi identique à celui du multicouches. Dans le cas de la série HH de l'année de test, l'erreur obtenue lors de la modélisation monodimensionnelle est proche de celle obtenue grâce à la modélisation multidimensionnelle, ceci est dû à la régression linéaire qui existe entre les séries HH et SS. Il n'y a donc pas de gain d'information avec l'ajout d'une série dont la corrélation est linéaire. Il est à signaler que pour les 2 modèles (MAR et multicouches), quelques valeurs de la fraction d'insolation pour l'année type et de test se situent hors de l'intervalle [0,1]. Pour y remédier, nous pouvons soit tronquer les valeurs qui se situent hors de cet intervalle [6], ou alors générer un nouveau nombre aléatoire gaussien au niveau de l'équation (III-1) pour le jour où la fraction d'irradiation ou d'insolation est négative, et répéter cette étape jusqu'à ce que la valeur obtenue soit positive [12].

IV-6- Analogie entre le AR et le réseau multicouches :

Nous rappelons que le but de la modélisation de la fraction d'irradiation journalière HH_j à partir de ses p valeurs précédentes, est de trouver une fonction F tel que :

$$HH_j = F(HH_{j-1}, \dots, HH_{j-p}) + e_j \quad (IV-1)$$

Considérons le cas d'un réseau multicouches, ayant p neurones d'entrée, sans neurones cachés, et ayant une sortie linéaire. L'équation (IV-1) se traduit par :

$$HH_j = \sum_{i=1}^p W_i HH_{j-p} + b + e_j \quad (IV-2)$$

où W_i et b sont les poids et seuil du réseau.

Par ailleurs, à échelle mensuelle, nous pouvons supposer que HH_j est stationnaire [7,10]. L'application de la modélisation AR a été faite dans [7,10] à échelle mensuelle, nous obtenons ainsi les coefficients a_i pour chaque mois. Pour un mois quelconque, la modélisation autorégressive se traduit par l'équation (IV-3) après avoir centré HH_j :

$$HH_j = \sum_{i=1}^p a_i HH_{j-i} + e_j \quad (IV-3)$$

L'analogie entre ce réseau et le AR apparaît en superposant l'équation (IV-2) et (IV-3), ainsi :

$$W_i = a_i \text{ et } b = 0 \quad (IV-4)$$

Pour vérifier ceci, nous avons calculé les coefficients a_i du modèle AR(1) et les poids (et seuil) du réseau multicouches défini plus haut, pour chaque mois de l'année type centrée. Nous obtenons donc les résultats affichés dans le tableau IV-7. Le nombre d'itérations du réseau est 500.

mois	AR(1)	réseau multicouches	
	A	W1	b
Janvier	0.33	0.35	0.0074
Février	0.28	0.29	0.0020
Mars	0.04	0.04	0.0050
Avril	0.13	0.14	0.0060
Mai	0.12	0.12	0.0042
Juin	0.35	0.36	0.0121
Juillet	0.33	0.34	0.0000
Août	-0.01	-0.01	-0.0017
Septembre	0.12	0.13	-0.0089
Octobre	0.27	0.27	0.0006
Novembre	0.18	0.19	0.0031
Décembre	-0.04	-0.04	0.0024

Tabl. IV-7 : Comparaison du AR(1) avec le réseau multicouches.

Nous remarquons que $W_1 \approx a_1$ et $b \approx 0$, ceci confirme bien cette analogie. Elle a été également vérifiée pour l'ordre 2 et 3.

IV-7- Analogie entre la régression linéaire et le réseau multicouches :

L'analogie avec la régression linéaire apparaît en considérant un réseau multicouches sans couche cachée et de sortie linéaire. Dès lors, l'équation de ce réseau en considérant la grandeur SS_j comme unique entrée et HH_j comme sortie, devient :

$$HH_j = W_1 SS_j + b \quad (\text{IV-5})$$

où W_1 et b sont respectivement le poids synaptique et seuil du réseau.

Nous rappelons que la relation de régression entre HH_j et SS_j est (formule d'Angström [7]) :

$$HH_j = \alpha SS_j + \beta \quad (\text{IV-6})$$

Etant donné que les équations (IV-5) et (IV-6) sont analogues, les valeurs de W_1 et b trouvées par l'algorithme de rétropropagation gradient doivent être égales respectivement à α et β . Le tableau IV-8 regroupe les valeurs trouvées de ces différents paramètres :

mois	régression linéaire		réseau multicouches	
	α	β	W_1	b
Janvier	0.48	0.25	0.48	0.25
Février	0.50	0.25	0.50	0.25
Mars	0.54	0.22	0.54	0.22
Avril	0.54	0.21	0.54	0.21
Mai	0.47	0.24	0.47	0.24
Juin	0.43	0.25	0.43	0.25
Juillet	0.41	0.27	0.42	0.26
Août	0.35	0.30	0.36	0.29
Septembre	0.39	0.28	0.39	0.28
Octobre	0.43	0.26	0.43	0.26
Novembre	0.44	0.26	0.44	0.26
Décembre	0.42	0.27	0.42	0.27

Tabl. IV-8 : Comparaison de la régression linéaire
avec le réseau multicouches.

Ainsi, sous ces conditions, le réseau multicouches n'est autre qu'une simple régression. Dans le cas de plusieurs entrées, nous aurons une régression linéaire multiple.

IV-8- Conclusion :

A partir des résultats obtenus, le réseau multicouches à 5 neurones cachés suffit à la description de l'évolution temporelle de l'irradiation solaire. Ce modèle est une autorégression non-linéaire. C'est une généralisation des processus autorégressifs linéaires.

Afin de montrer l'aptitude des réseaux de neurones à la modélisation multidimensionnelle, l'irradiation et l'insolation solaire s'y sont prêtées, les résultats sont comparatifs à ceux des processus autorégressifs multidimensionnels.

Aussi, le passage des réseaux de neurones aux processus autorégressifs fut abordé et démontré, de même que pour la régression linéaire.

Conclusion générale.

Les réseaux de neurones artificiels ont été mis au profit de la modélisation de l'évolution temporelle du rayonnement solaire à Alger. Il a été trouvé que le caractère Markovien (dépendance des valeurs successives entre elles) est fidèlement reproduit.

La modélisation par les processus autorégressifs d'ordre 1 est suffisante pour cette description. Cependant les étapes nécessaires à l'élaboration du modèle sont plus compliquées dans le cas des processus autorégressifs, qui nécessitent une variable stationnaire et gaussienne. Dans le cas des réseaux de neurones, aucun traitement intermédiaire n'est nécessaire.

Il a été démontré que les chaînes de Markov du premier ordre à deux états étaient suffisantes à leur tour pour la modélisation du rayonnement solaire [6-9]. Mais l'erreur obtenue est supérieure à celle obtenue à l'aide des réseaux de neurones.

Il existe plusieurs types de réseaux, et pour chaque réseau, quelques paramètres sont modifiables afin d'améliorer la précision des résultats. Les résultats obtenus montrent que le réseau multicouches ayant une seule couche cachée convient à cette modélisation. De plus, la généralisation à une année de test, qui elle n'a pas été introduite dans la phase d'apprentissage des poids du réseau, est meilleure que dans le cas des processus autorégressifs et des chaînes de Markov. Toujours dans le cas du réseau multicouches et en changeant le nombre de neurones d'entrée et le nombre de neurones cachés, l'erreur n'a pas diminué, ainsi, une entrée et 5 neurones cachés suffisent à la modélisation du rayonnement solaire.

Le passage des réseaux de neurones aux processus autorégressifs et à la régression linéaire a été réalisé, les résultats obtenus de part et d'autre sont en accord.

L'aptitude des réseaux de neurones à la modélisation multidimensionnelle, en l'occurrence l'irradiation et l'insolation solaire, fut comparée aux processus autorégressifs multidimensionnels, ce qui nous a conduit à des résultats en accord les uns avec les autres.

Les perspectives à envisager concernent la nature des données. Ainsi, l'application pourrait être élargie aux autres composantes du rayonnement solaire, à savoir l'irradiation solaire directe et diffuse. De même, le pas de temps (10 minutes, demi-heure, heure) pourrait être changé. Aussi, la diversité des données pourrait contribuer à l'affinement de la modélisation. Par exemple, la direction et vitesse du vent, la pression atmosphérique, la température, l'humidité, la pluviométrie pourraient mieux décrire l'évolution du rayonnement solaire.

Une autre extension concernerait la modélisation spatiale (en fonction des coordonnées géographiques) du rayonnement solaire. Il s'agit d'estimer l'irradiation en un lieu à partir de celle mesurée dans un autre lieu. Les réseaux de neurones pourraient contribuer à cet objectif, ceci s'avère intéressant dans le cas où des mesures ne sont pas disponibles en ce lieu [29].

Il est à signaler que diverses méthodes peuvent être associées aux réseaux de neurones : la logique floue [49], les chaînes de Markov cachées, les algorithmes génétiques et les systèmes chaotiques pourraient éventuellement améliorer la précision des modèles actuels.

Bibliographie.

- [1] : J. R. Vaillant, "Utilisations et promesses de l'énergie solaire", 1978, Eyrolles, Paris.
- [2] : R. Bernard, G. Menguy, M. Schwartz, "Le rayonnement solaire, conversion thermique et applications", 1980, Paris : techniques et documentation.
- [3] : M. Capderou, "Atlas solaire de l'Algérie", 1983, Vol. 1, OPU, Alger.
- [4] : M. Egido and E. Lorenzo, "The sizing of stand alone PV-systems : a review and a proposed new method", 1992, Solar Energy Materials and Solar Cells 26, pp. 51–69.
- [5] : S. A. Klein and W. A. Beckman, "Loss-of-load probabilities for stand-alone photovoltaic systems", 1987, Solar Energy, Vol. 39, N° 6, pp. 499-512.
- [6] : J. W. Hansen, "Stochastic daily solar irradiance for biological modeling applications", 1999, Agricultural and Forest Meteorology, 94, pp. 53–63.
- [7] : A. Maafi, "Mise en évidence d'aspects physiques du modèle Markovien du premier ordre à deux états en météorologie solaire. Application à la conversion photovoltaïque", 1991, thèse de Doctorat d'Etat en Electronique, USTHB, Alger.
- [8] : A. Maafi, "Markov models in discrete time for solar radiation", April 1998, IMACS-IEEE, Tunisia, Vol. 2, pp. 319-322.
- [9] : R. J. Aguiar, M. Collares-Pereira and J. P. Conde, "Simple procedure for generating sequences of daily radiation values using a library of Markov transition matrices", 1988, Solar Energy, Vol. 40, N° 3, pp. 269-279.
- [10] : F. Hammadou, "Analyse des performances de la modélisation Markovienne du rayonnement solaire", 1993, thèse de Magister en Electronique, USTHB, Alger.
- [11] : B. J. Brinkworth, "Autocorrelation and stochastic modelling of insolation sequences", 1977, Solar Energy, Vol. 19, p. 343.
- [12] : R. Aguiar and M. Collares-Pereira, "TAG : time-dependent, autoregressive, gaussian model for generating synthetic hourly radiation", 1992, Solar Energy, Vol. 49, N° 3, pp. 167–174.
- [13] : S. Boubekour and A. Maafi, "Assessment of solar radiation data using multivariate autoregressive modelling", September 1994, WREC III, UK, part III, pp. 2342-2344.
- [14] : S. Boubekour and A. Maafi, "Univariate and bivariate autoregressive modelling applied to solar data", Actes de la 2^{ème} Conférence Internationale sur les Signaux, Systèmes et Automatique, Blida, Algérie 10-12 Mai 1999, Vol. 1, pp. 141-143.

- [15] : M. Muselli, P. Poggi, G. Notton and A. Louche, "Improved procedure for stand-alone photovoltaic systems sizing using Meteosat satellite images", 1998, *Solar Energy*, Vol. 62, N° 6, pp. 429–444.
- [16] : E. Davalo, P. Naïm, "Les réseaux de neurones", 1989, Eyrolles.
- [17] : J. E. Dayhoff, "Neural network architectures. An introduction", 1990, Van Nostrand Reinhold, New York.
- [18] : S. Haykin, "Neural networks. A comprehensive foundation", 1994, Macmillan, New York.
- [19] : J. F. Jordain, "Les réseaux de neurones, principes et définitions", 1994, Hermes.
- [20] : F.-C. Chen, "Back-propagation neural networks for nonlinear self-tuning adaptive control", 1990, *IEEE Control System Magazine*, pp. 44–48.
- [21] : G. Schwarzer, W. Vach and M. Schumacher, "On the misuses of artificial neural networks for prognostic and diagnostic classification in oncology", 2000, *Statistics in Medicine*, 19, pp. 541–561.
- [22] : N. Stamatis, D. Parthimos and T. M. Griffith, "Prediction of arterial burst activity and transitions between chaotic attractors with a multilayer perceptron optimized by a new stopping criterion", September 1996, *International Journal of Neural Systems*, Vol. 7, N° 4, pp. 417-428. Special Issue on the Role and Control of Random Events in Biological Systems.
- [23] : E. Yair and A. Gersho, "Maximum a posteriori decision and evaluation of class probabilities by Boltzmann perceptron classifiers", October 1990, *Proceedings of the IEEE*, Vol. 78, N° 10, pp. 1620–1628.
- [24] : E. Zorita and H. Von Storch, "The analog method as a simple statistical downscaling technique : Comparison with more complicated methods", 1999, *Journal of Climate*, Vol. 12, pp. 2474-2488.
- [25] : M. Yousif and M. K. Darwish, "Prediction of clearness index using artificial neural network technique", 1998, *WREC*, pp. 2595–2598.
- [26] : J. L. Duomarco, "Calculation of hourly values of direct radiation using artificial neural networks", 1998, *WREC*, pp. 2567–2570.
- [27] : Y. Kemmoku, S. Orita, S. Nakagawa and T. Sakakibara, "Daily insolation forecasting using a multi-stage neural network", 1999, *Solar Energy*, Vol. 66, N° 3, pp. 193–199.
- [28] : A. Sfetsos and A. H. Coonick, "Univariate and multivariate forecasting of hourly solar radiation with artificial intelligence techniques", 2000, *Solar Energy*, Vol. 68, N° 2, pp. 169-178.

- [29] : M. Mohandes, A. Balghonaim, M. Kassas, S. Rehman and T. O. Halawani, "Use of radial basis functions for estimating monthly mean daily solar radiation", 2000, Solar Energy, Vol. 68, N° 2, pp. 161-168.
- [30] : A. Guessoum, S. Boubekeur and A. Maafi, "A global irradiation model using radial basis function neural networks", September 20-25 1998, WREC V, Italy, part IV, pp. 2533-2536.
- [31] : G. Thimem and E. Fiesher, "High order and multilayer perceptron initialization", 1997, IEEE Trans Neural Networks, Vol. 8, pp. 349-359.
- [32] : B. Widrow and M. A. Lehr, "30 years of adaptive neural networks : perceptron, madaline, and backpropagation", September 1990, Proceedings of the IEEE, Vol. 78, N° 9, pp. 1415–1442.
- [33] : M. I. Jordan and R. A. Jacobs, "Hierarchical mixtures of experts and the EM algorithm", 1994, Neural Computation, Vol. 6, pp. 181–214.
- [34] : J. T. Connor, R. Douglas Martin and L. E. Atlas, "Recurrent neural networks and robust time series prediction", March 1994, IEEE Transactions on Neural Networks, Vol. 5, N° 2, pp. 240-254.
- [35] : O. Admane, H.– KY, N. Ouakli, "Statistique (cours & exercices). Pour les étudiants du tronc commun biomédical", 1992, OPU, INESSM Alger.
- [36] : M. Babes, "Statistiques, files d'attente et simulation", 1992, OPU, Alger.
- [37] : Ph. Lazar et D. Schwartz, "Eléments de probabilités et statistique", 1967, Flammarion, Paris.
- [38] : A. Bensaber, B. Blense-Trillon, "Pratique des chroniques et de la prévision à court terme", 1989, Paris : Masson.
- [39] : G. E. P. Box and G. M. Jenkins, "Time series analysis, forecasting and control", 1970, Holden-Day, San Fransisco.
- [40] : B. Picinbono, "Signaux aléatoires et modèles avec problèmes résolus", tome 2, 1994, Dunod.
- [41] : S. M. Kay and S. L. Marple JR., "Spectrum analysis – A modern perspective", November 1981, Proceedings of the IEEE, Vol. 69, N° 11, pp. 1380–1418.
- [42] : B. R. Upadhyaya, M. Kitamura and T. W. Kerlin, "Multivariate signal analysis algorithms for process monitoring and parameter estimation in nuclear reactors", 1980, Journal of Nuclear Energy, Vol. 7, pp. 1-11.
- [43] : V. Tresp and R. Hofmann, "Nonlinear time-series prediction with missing and noisy data", 1998, Neural Computation, 10, pp. 731–747.

- [44] : H. A. Ceccatto, H. D. Navone and H. Waelbroeck, "Learning persistent dynamics with neural networks", 1998, Neural Networks, Vol. 11, N° 1, pp. 145–151.
- [45] : O. Kinouchi, M. H. R. Tragtenberg, "Modeling neurons by simple maps", 1996, International Journal of Bifurcation and Chaos, Vol. 6, N° 12 A, pp. 2343–2360.
- [46] : P. J. Werbos, "Backpropagation through time : what it does and how to do it", October 1990, Proceedings of the IEEE, Vol. 78, N° 10, pp. 1550–1560.
- [47] : S. Boubekour, G. I. Abib et A. Maafi, "Modélisation bidimensionnelle des données d'énergie solaire par les réseaux de neurones multicouches", Communication acceptée pour être présentée au 3^{ème} Séminaire National en Génie Electrique, 29-31 Octobre 2001, Biskra.
- [48] : L. Ramirez Santigosa, LI. Mora Lopez, F. Martin Morillas and M. Sidrach-de-Cardona, "Calculation of daily global solar radiation using multivariate linear models", 1998, WREC, pp. 2612–2615.
- [49] : Y.-Q. Zhang and A. Kandel, "Compensatory neurofuzzy systems with fast learning algorithms", January 1998, IEEE Transactions on Neural Networks, Vol. 9, N° 1, pp.83–105.

Annexe 1 : Confection d'une année type.

Supposons que l'on dispose de A années de données. Pour chaque mois, le mois de l'année type est celui ayant une distribution de probabilité proche de celle du mois cumulé. Soit x la série temporelle (fraction d'irradiation ou d'insolation). Pour chaque mois, on procède de la manière suivante :

- On calcule la fonction de distribution de probabilité F(x) pour le mois cumulé.
- On calcule la fonction de distribution de probabilité F_i(x) de ce mois pour l'année i (i variant de 1 à A), $x \in [x_{\min}, x_{\max}]$.
- On calcule la moyenne des différences absolues entre les 2 fonctions F(x) et F_i(x) pour l'année i :

$$D_i = \frac{1}{b} \sum_{j=1}^b |F(jh) - F_i(jh)| \quad (\text{A1-1})$$

et $b = \frac{x_{\max} - x_{\min}}{h}$, h étant le pas de quantification de x dans l'intervalle $[x_{\min}, x_{\max}]$, il est choisi comme étant égal à 0,01.

Parmi les A mois disponibles, on choisit le mois ayant le plus petit D_i.

Annexe 2 : Les chaînes de Markov.

A2-1- Discrétisation en états :

La modélisation d'une suite de données Y_t par les chaînes de Markov d'ordre 1 à m états nécessite la discrimination entre ces m états. Ainsi, à partir des données réelles et continues Y_t , nous obtenons les données discrètes X_t . Pour cela, les données réelles continues doivent être discrétisées en états grâce à l'utilisation de $(m-1)$ seuils notés (par ordre croissant) : S_1, \dots, S_{m-1} , ainsi [7-8,10] :

- Si $Y_t < S_1$ alors : $X_t = 0$.
- Si $S_i \leq Y_t < S_{i+1}$ alors : $X_t = i$.
- Si $Y_t \geq S_{m-1}$ alors : $X_t = m-1$.

A2-2- Calcul des seuils :

Les seuils optimaux correspondent à ceux donnant la convergence rapide de la matrice P vers le régime permanent. Pour déterminer ces seuils, on fait appel à des méthodes statistiques telle la méthode du seuil [7] nécessitant le test du χ^2 (Khi-2) ou la méthode de maximum d'entropie [10] pour calculer les P_{ij} et ainsi trouver P . C'est cette dernière qui a été utilisée.

L'application de cette méthode nécessite d'abord le calcul de la fonction cumulative F_c des données réelles Y_t [10]. Puis, dans le cas de 2 états, le seuil S_1 est choisi parmi les données réelles donnant F_c égale à $\frac{1}{2}$. Pour 3 états, S_1 et S_2 sont choisis parmi les données réelles donnant F_c égale respectivement à $\frac{1}{3}$ et $\frac{2}{3}$. Dans le cas général de m états, le seuil S_i (i allant de 1 à $(m-1)$) est choisi parmi les données réelles donnant F_c égale à $\frac{i}{m}$.

A2-3- Calcul des probabilités de transition :

Une fois les seuils trouvés, on discrétise les données réelles en états. Maintenant, il est possible de calculer les probabilités P_{ij} , elles sont calculées par la méthode du maximum de vraisemblance et sont données par [7-8,10] :

$$P_{ij} = \frac{U_{ij}}{\sum_{j=0}^{m-1} U_{ij}} \quad (\text{A2-1})$$

où U_{ij} est le nombre de transitions de l'état i vers l'état j .

A2-4- Calcul des probabilités à long terme :

Appelées aussi probabilités Markovienne ou vecteur d'état stationnaire, les probabilités à long terme sont obtenues par multiplication de la matrice P par elle-même plusieurs fois. Nous obtenons donc une matrice ayant des lignes identiques, dont le vecteur d'état noté P_m est de la forme [7-8,10] :

$$P_m = (P_0 \ P_1 \ \dots \ P_{m-1}) \quad (\text{A2-2})$$

où P_i est la probabilité à long terme de l'état i .

Pour valider la modélisation, chaque probabilité P_i peut être comparée à la probabilité à priori d'avoir l'état (i) notée \hat{P}_i . Elle est calculée directement à partir des données, tel que [7-8,10]:

$$\hat{P}_i = \frac{N_i}{\sum_{j=0}^{m-1} N_j} \quad (\text{A2-3})$$

N_i est le nombre de données dont l'état est (i).

A2-5- Simulation :

Elle a pour but de générer les données réelles à partir de la matrice de transition et des probabilités à long terme [10]. La simulation par le modèle en chaînes de Markov d'ordre 1 à m états se fait en considérant les fonctions de répartition linéaire par parties : la première valeur simulée est obtenue à partir d'une valeur z d'un générateur aléatoire uniformément répartie entre 0 et 1. On teste $z < \sum_{i=0} P_i$. La première valeur de i qui satisfait l'inégalité sera l'état de la donnée simulée dont la valeur sera obtenue en utilisant la droite D_i de la figure A2-1.

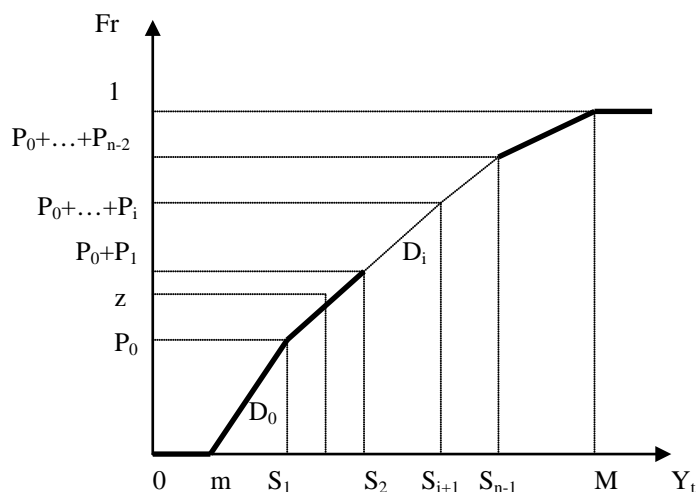


Fig. A2-1 : Simulation de la première valeur.

Pour simuler les valeurs suivantes, supposons que e est l'état de la donnée simulée précédente. Un nombre z est généré à partir d'un générateur aléatoire uniformément réparti. On teste $z < \sum_{i=0} P_{ei}$. La première valeur de i satisfaisant l'inégalité sera l'état de la donnée simulée dont la valeur sera obtenue en utilisant la droite D_i ' de la figure A2-2.

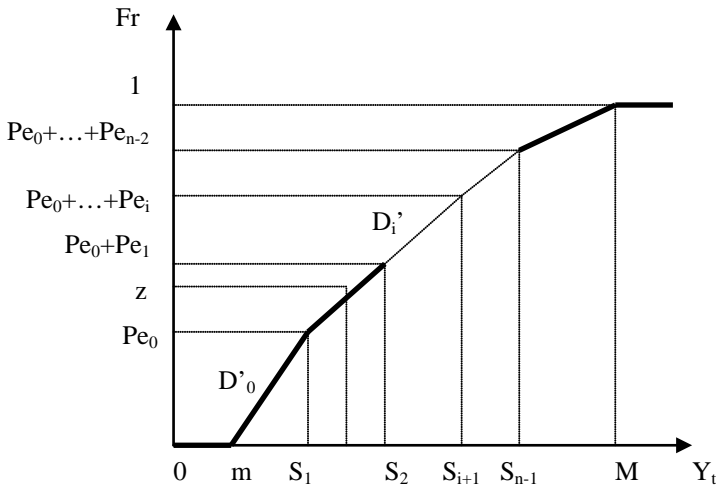


Fig. A2-2 : Simulation des valeurs suivantes.

Annexe 3 : Les processus autorégressifs monodimensionnels.

A3-1- Algorithme de Levinson-Durbin :

Soit X_1, X_2, \dots, X_N , N observations d'une série temporelle notée X_t , stationnaire ou faiblement stationnaire et de moyenne M . La modélisation par les processus autorégressifs se traduit par l'équation suivante [4,10-12,15,38-41] :

$$X_t - M = \sum_{i=1}^p a_i (X_{t-i} - M) + e_t \quad (\text{A3-1})$$

où p est l'ordre du modèle, a_i est le $i^{\text{ème}}$ coefficient du modèle d'ordre p , et e_t est un bruit

blanc vérifiant : $E(e_t) = 0$ et $E(e_t e_{t-i}) = \begin{cases} \sigma^2 & \text{pour } i = 0 \\ 0 & \text{pour } i \neq 0 \end{cases}$

Les coefficients a_i sont déterminés par la méthode des moindres carrés qui permet de minimiser la covariance de la série temporelle, soit :

$$\sum_{t=p+1}^N [(X_t - M) - \dots - a_p (X_{t-p} - M)]^2 = \sum_{t=p+1}^N e_t^2 \quad (\text{A3-2})$$

La minimisation de l'équation (A3-2) par rapport à a_1, \dots, a_p conduit à :

$$\sum_{t=p+1}^N (X_{t-i} - M)(X_{t-j} - M) = NR_{j-i} \quad (\text{A3-3})$$

avec R_r la fonction d'autocovariance de X_t définie par :

$$R_r = \frac{1}{N} \sum_{t=1}^{N-|r|} X_t X_{t+|r|} \quad (\text{A3-4})$$

l'équation (A3-3) peut se mettre sous la forme :

$$R_j - a_1 R_{j-1} - \dots - a_p R_{j-p} = 0, \quad j = 1, \dots, p \quad (\text{A3-5})$$

sous forme matricielle, nous obtenons :

$$\begin{pmatrix} R_0 & R_1 & \dots & R_p \\ R_1 & R_0 & \dots & R_{p-1} \\ \vdots & \vdots & \ddots & \vdots \\ R_p & R_{p-1} & \dots & R_0 \end{pmatrix} \begin{pmatrix} 1 \\ a_1 \\ \vdots \\ a_p \end{pmatrix} = \begin{pmatrix} \sigma^2 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad (\text{A3-6})$$

La résolution de l'équation (A3-6) peut se faire par la méthode de Gauss par exemple, mais Levinson et Durbin ont développé une méthode itérative afin de minimiser le nombre d'opérations à effectuer, le faisant passer de P^3 à P^2 opérations. Pour cela, on note par a_i^p le $i^{\text{ème}}$ coefficient du modèle d'ordre p et par S_p l'estimation de σ^2 . L'algorithme de Levinson-Durbin [4,10-12,15,38-41] débute par les initialisations suivantes :

$$S_0 = R_0 \quad (\text{A3-7})$$

$$a_1^1 = \frac{R_1}{R_0} \quad (\text{A3-8})$$

$$S_1 = \left(1 - (a_1^1)^2\right) S_0 \quad (\text{A3-9})$$

Puis vient le calcul des coefficients, dont l'étape générale est : (pour $i = 2, \dots, p$)

$$a_i^i = \frac{R_i - \sum_{m=1}^{i-1} a_m^{i-1} R_{i-m}}{S_{i-1}} \quad (\text{A3-10})$$

et

$$S_i = \left(1 - (a_i^i)^2\right) S_{i-1} \quad (\text{A3-11})$$

$$a_j^i = a_{j-1}^i - a_i^i a_{i-j}^{i-1} \quad (j = 1, \dots, i-1) \quad (\text{A3-12})$$

L'estimation de l'ordre p du modèle AR peut se faire par le critère d'Akaike. Il s'agit de calculer pour chaque ordre un coefficient noté FPE (Final Prediction Error) donné par :

$$\text{FPE}(p) = \frac{N + p}{N - p} \sigma^2 \quad (\text{A3-13})$$

L'ordre p retenu correspond à la valeur minimale du FPE.

A3-2- Modélisation et simulation par les processus autorégressifs :

Nous rappelons que la modélisation par les processus autorégressifs nécessite d'une part la stationnarité de la série temporelle [10], et d'autre part que sa distribution soit normale [10]. Le but étant d'exprimer l'évolution de cette série à partir de ses valeurs antérieures, et donc de trouver l'ordre p du modèle. Dans ce qui suit, nous décrirons les étapes suivies pour la simulation d'une série (données de l'année type par exemple) à partir du modèle établie.

Tout d'abord, les données de l'année type doivent subir la transformation gaussienne. Ainsi, partant de la fonction de répartition $\text{Fr}(X_t)$ des données non gaussiennes centrées et réduites X_t , nous obtenons les nouvelles données de distribution normale X_{gauss_t} , définies par [10,12] :

$$X_{\text{gauss}_t} = \sqrt{2} \text{erf}^{-1}(2\text{Fr}(X_t)) \quad (\text{A3-14})$$

tel que :

$$\text{erf}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt \quad (\text{A3-15})$$

De là on applique l'algorithme de Levinson-Durbin et le critère d'Akaike sur les données gaussiennes X_{gauss_t} pour la détermination de l'ordre p et des coefficients du modèle. L'étape suivante consiste à simuler de nouvelles séquences grâce à un générateur de bruit blanc gaussien e_t de moyenne nulle et de variance calculée en même temps que les coefficients. Partant des p valeurs précédentes, la sortie simulée \hat{X}_{gauss_t} se fait de la manière suivante [10] :

$$\hat{X}_{\text{gauss}_t} = \sum_{i=1}^p a_i \hat{X}_{\text{gauss}_{t-i}} + e_t \quad (\text{A3-16})$$

tout en initialisant les p premières valeurs de \hat{X}_{gauss_t} aux valeurs réelles X_{gauss_t} comme suit :

$$\hat{X}_{\text{gauss}_t} = X_{\text{gauss}_t} \text{ pour } t = 1 \text{ à } p \quad (\text{A3-17})$$

\hat{X}_{gauss_t} a le même nombre d'échantillons que X_{gauss_t} .

Nous obtenons ainsi une série simulée par les processus autorégressifs notée \hat{X}_{gauss_t} . Cependant, cette nouvelle série est de distribution normale, ce qui n'est pas le cas pour les données d'énergie solaire, nous devons transformer cette série en une série ayant une distribution de même forme que les données réelles. Pour cela, un polynôme P_5 d'ordre 5 suffit pour revenir aux données réelles, ses coefficients sont calculés en ajustant les données réelles gaussiennes X_{gauss_t} aux données réelles X_t , tel que [10]:

$$X_t = \sum_{i=0}^5 C_i (X_{\text{gauss}_t})^i \quad (\text{A3-18})$$

La série réelle notée \hat{X}_t simulée par le AR est donnée en fonction de \hat{X}_{gauss_t} par :

$$\hat{X}_t = \sum_{i=0}^5 C_i (\hat{X}_{\text{gauss}_t})^i \quad (\text{A3-19})$$

Dans le cas de la série de test et en partant de ses p valeurs initiales, nous la simulons à la manière de l'équation (A3-1), puis nous la transposons via le polynôme P_5 selon l'équation (A3-19).

Ainsi, la modélisation par le AR revient à trouver les coefficients a_i et la variance du bruit. L'ordre du modèle se détermine par l'observation d'un coefficient noté FPE. Nous choisissons l'ordre quand la variation du FPE n'est plus importante.

Tout ceci concerne les processus autorégressifs monodimensionnels, le même principe est suivi dans le cas multidimensionnel (annexe 4).

Annexe 4 : Les processus autorégressifs multidimensionnels.

Dans le cas où l'on dispose de plusieurs séries chronologiques, les processus AR monodimensionnels ne font pas apparaître les relations entre les différentes séries, ils traitent chaque série séparément des autres. D'où l'intérêt des processus AR multidimensionnels (MAR). La modélisation par les processus autorégressifs multidimensionnels se traduit par l'équation suivante [13-14,42] :

$$X_t - M = \sum_{i=1}^p a_i (X_{t-i} - M) + e_t \quad (\text{A4-1})$$

où $X_t = (X_t^1, \dots, X_t^m)^T$ (l'exposant T signifie vecteur transposé) représente m séries temporelles ayant chacune N observations, $M = (M_1, \dots, M_m)^T$ est le vecteur moyenne, tel que M_i est la moyenne de la $i^{\text{ème}}$ série temporelle X_t^i et a_i est le $i^{\text{ème}}$ coefficient du modèle d'ordre p. Il est sous forme de matrice de dimension (m x m). p est l'ordre du modèle. Enfin, $e_t = (e_t^1, \dots, e_t^m)^T$ est un bruit blanc vérifiant : $E(e_t) = 0$ et $E(e_t e_{t-i}) = \begin{cases} \sigma^2 & \text{pour } i = 0 \\ 0 & \text{pour } i \neq 0 \end{cases}$

Pour la détermination des coefficients a_i , les étapes sont les mêmes que dans le cas monodimensionnel, nous obtenons l'équation suivante en multipliant l'équation (A4-1) par X_{t-L}^T :

$$E(X_t X_{t-L}^T) = \sum_{i=1}^p a_i E(X_{t-i} X_{t-L}^T) \quad (\text{A4-2})$$

La matrice de corrélation est donnée par :

$$R_L = E(X_t X_{t-L}^T) \quad (\text{A4-3})$$

L'équation (A4-2) se réécrit :

$$R_L = \sum_{i=1}^p a_i R_{L-i} \quad L = 1, \dots, p \quad (\text{A4-4})$$

C'est l'équation de YULE-WALKER, on peut également écrire :

$$E(\mathbf{X}_t \mathbf{X}_t^T) = \sum_{i=1}^p a_i E(\mathbf{X}_{t-i} \mathbf{X}_t^T) + E(\mathbf{e}_t \mathbf{X}_t^T) \quad (\text{A4-5})$$

Ceci donne :

$$\mathbf{R}_0 = \sum_{i=1}^p a_i \mathbf{R}_{-i} + \sigma^2 \quad (\text{A4-6})$$

Sachant que :

$$\mathbf{R}_i = \mathbf{R}_{-i} = \frac{1}{N} \sum_{t=1}^{N-i} \mathbf{X}_t \mathbf{X}_{t-i}^T \quad (\text{A4-7})$$

Un des algorithmes permettant la résolution de l'équation de YULE-WALKER est l'algorithme de LEVINSON-DURBIN généralisé. Soit \mathbf{a}_i^p et \mathbf{b}_i^p deux matrices de dimension $(m \times m)$ représentant les $i^{\text{èmes}}$ coefficients du modèle MAR d'ordre p , en considérant respectivement la prédiction progressive (Forward) et rétrograde (Backward), l'initialisation commence en posant [42] :

$$\mathbf{S}_0^f = \mathbf{S}_0^b = \mathbf{R}_0 \quad (\text{A4-8})$$

$$\mathbf{a}_1^1 = \frac{\mathbf{R}_1}{\mathbf{R}_0} \quad (\text{A4-9})$$

$$\mathbf{b}_1^1 = \frac{\mathbf{R}_1^T}{\mathbf{R}_0} \quad (\text{A4-10})$$

L'étape générale est :

$$\mathbf{a}_p^p = \frac{\mathbf{R}_p - \sum_{i=1}^{p-1} a_i^{p-1} \mathbf{R}_{p-i}}{\mathbf{S}_{p-1}^b} \quad (\text{A4-11})$$

$$\mathbf{b}_p^p = \frac{\mathbf{R}_p^T - \sum_{i=1}^{p-1} b_i^{p-1} \mathbf{R}_{p-i}^T}{\mathbf{S}_{p-1}^f} \quad (\text{A4-12})$$

$$\mathbf{a}_k^p = \mathbf{a}_k^{p-1} - \mathbf{a}_p^p \mathbf{b}_{p-k}^{p-1} \quad (\text{A4-13})$$

$$\mathbf{b}_k^p = \mathbf{b}_k^{p-1} - \mathbf{b}_p^p \mathbf{a}_{p-k}^{p-1} \quad (\text{A4-14})$$

$$\mathbf{S}_p^f = (\mathbf{I} - \mathbf{a}_p^p \mathbf{b}_p^p) \mathbf{S}_{p-1}^f \quad (\text{A4-15})$$

$$\mathbf{S}_p^b = (\mathbf{I} - \mathbf{b}_p^p \mathbf{a}_p^p) \mathbf{S}_{p-1}^b \quad (\text{A4-16})$$

Le critère de sélection de l'ordre p est le critère d'Akaike :

$$\text{FPE}(p) = |\sigma^2| \left[\frac{N+1+mp}{N-1-mp} \right]^m \quad (\text{A4-17})$$

$|\sigma^2|$ est le déterminant de la matrice de covariance du bruit. Là aussi, l'ordre p retenu correspond à la valeur minimale du FPE.