

P0008/05B

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



المدرسة الوطنية المتعددة التقنيات
BIBLIOTHEQUE — المكتبة
Ecole Nationale Polytechnique

Ecole Nationale Polytechnique

Département d'Hydraulique

MEMOIRE DE FIN D'ETUDES
En vue de l'obtention du diplôme
d'Ingénieur d'Etat en Génie Hydraulique

Thème :

**CONTRIBUTION À LA MODÉLISATION DE
L'ÉVAPOTRANSPIRATION POTENTIELLE
PAR LES RÉSEAUX DE NEURONES**

Présenté par :
Mr LAHSEN Redouane.

Encadré par :
Dr BERMAD Abdelmalek

Promotion Juin 2005

Ecole Nationale Polytechnique 10, av. Hassen BADI – BP 182- 16200 El Harrach- Alger- ALGERIE
Tél. : 213 21 52 14 94 fax : 213 21 52 79 29

Dédicaces :

A ma mère, Zehira, ... que dieu puisse me la garder.

A la mémoire de Cherif.... Disparu trop tôt....

A mon père, Mohamed, mes sœurs, Dalila et Malika, et ma nièce Melissa.

A mes meilleurs amis, Yacine et Amine en particulier.

Remerciements

Je tiens à remercier en tout premier lieu le Pr N.DECHEMI, qui m'a accueilli les bras ouverts au sein de GESTOP, son équipe de recherche, et auprès duquel je me suis initié aux joies et aux peines du travail de recherche.

Je remercie le Dr Bermad, dont la disponibilité sans failles, la compréhension, et la qualité d'encadrement m'ont permis de mener à terme ce projet de fin d'études.

Merci encore à T.Benkaci, pour l'aide précieuse qu'il m'a fournie.

Je remercie Mlle Benmamar, pour la qualité de sa gestion du département d'hydraulique, elle a grandement contribué à ma formation en assurant tous les moyens pédagogiques et matériels nécessaires.

Merci à tous les enseignants, titulaires et associés, du département d'hydraulique, auprès desquels j'ai reçu une formation d'excellente qualité.

Merci à mes parents, et à DALIL A ma sœur aînée, qui m'ont fourni toutes les commodités matérielles nécessaires à l'aboutissement de ce travail.

Merci à ma petite sœur Malika, pour son soutien moral indéfectible, et pour l'aide apportée lors des vérifications.

Merci à tous mes amis et camarades pour leur soutien moral et leurs encouragements.

يجب التعرف بدقة على معادلة التوازن المائي و هذا لتحسين تسيير موارد مائية نؤول إلى الندرة. إن ظاهرة النتح و التبخر تعد من أهم مكونات معدلة الميزان المائي، ولهذا تكميم هذه الضاهرة ضروري. لهذا الغرض لجأنا إلى تمثيلها بإحدى تقنيات الذكاء الاصطناعي الشبكة العصبونية الطاغية. كلمات مفتاحية: الشبكة العصبونية إصطناعية، النتح و التبخر، تسيير موارد مائية، الميزان المائي،

Abstract :

A precise knowledge of hydrous balance is essential to improve the water resources management, the potential evapotranspiration is one of the most important component of this balance. So, it is necessary to quantifize it corretly.

For this purpose, we have used one of the artificial intelligence ways: The artificial neural network.

Keywords: Artificial neural network, evapotranspiration, water resources management, hydrous balance.

Résumé:

Une connaissance précise du bilan hydrique est indispensable afin de bien gérer les ressources en eau qui se font de plus en plus rare. L'évapotranspiration potentielle est une composante essentielle de ce bilan, il est donc nécessaire de la quantifier au mieux. Le présent travail procédera à cette quantification par l'intermédiaire de la modélisation par une technique d'intelligence artificielle : Les réseaux de neurones artificiels.

Mots clés : Réseaux de neurones artificiels, évapotranspiration, gestion des ressources hydriques, bilan hydrique.

المدرسة الوطنية المتعددة التقنيات
BIBLIOTHEQUE — المكتبة
Ecole Nationale Polytechnique

Sommaire

Introduction	1
Chapitre I : Processus physique de l'évapotranspiration	
I.1. Concept d'évapotranspiration	2
I.1.1. Processus physique de l'évaporation.....	2
I.1.1.1. Facteurs météorologiques	3
I.1.1.2. Facteurs physiques	4
I.1.2. Processus physique de la transpiration des plantes.....	4
I.2. Les différents types d'évapotranspiration	
I.2.1. Evapotranspiration de référence (ET_0).....	5
I.2.1. Evapotranspiration maximale (ET_m).....	5
I.2.3. Evapotranspiration réelle.....	6
Chapitre II : Différentes méthodes de quantification	
II.1. Formules résultant d'ajustements statistiques	7
II.1.1. Formule de BLANEY et CRIDDLE (1945).....	7
II.1.2. Formule de Thornthwaite (1948).....	7
II.1.3. Formule de TURC (1961)	8
II.1.4. Formule de JENSEN et HAISE (1963)	9
II.1.5. Méthode combinée de VAN BAVEL (1968)	9
II.1.6. Formule de JENSEN et WRIGHT (1972)	10
II.1.7. Formule de PRIESTLEY et TAYLOR (1972).....	10
II.2. Formules basées sur le bilan énergétique :	
II.2.1. Formule de BOUCHET (1964)	10
II.2.2. Formule de BROCHET et GERBIER (1974)	11
II.2.3. Formule de Penman (1948)	12.
II.3. Méthodes de mesures de l'ETP	
II.3.1. Les Bacs	12
II.3.2. Les lysimètres	13
Chapitre III : Développement de la formule de Penman	
III.1. Développement mathématique de l'équation de Penman	15
III.2. Equation de Penman-Monteith	20
III.3. Equation de Penman-ANRH	22
Chapitre IV : Notions générales sur les réseaux de neurones	
IV.1. Introduction	25
IV.2. Principe Général	26
IV.3. Les fondements biologiques	26
IV.4. Neurone naturel	27
IV.5. Modélisation du neurone naturel	28
IV.6. Nature des entrées et des sorties	29
IV.7. Différentes Structures de connexions	31
IV.7.1. Les réseaux à couches	31
IV.7.2. Réseau entièrement connecté	32
IV.8. Dynamique des connexions et loi de HEBB	32
IV.9. Les perceptrons	33
IV.9.1. Le perceptron mono-couche	33
IV.9.1.1. Organisation générale	33

IV.9.1.2. Propriétés générales.....	34
IV.9.1.3. Apprentissage du perceptron mono-couche	34
IV.9.1.4. Limites du perceptron mono-couche.....	35
IV.9.2. Le perceptron multicouches	36
IV.9.2.1. Algorithme de back-propagation	37
IV.9.2.2. Apprentissage	38
IV.9.2.3 Avantages	38
IV.9.2.4. Limites de l'algorithme	39
IV.10. Historique	39
IV.10.1. Les prémices	39
IV.10.2. Les pionniers	39
IV.10.3. Le premiers réseaux de neurones artificiels	39
IV.10.4. Les RNA multicouches	40
IV.11. Les différents types de réseaux de neurones artificiels.....	40
IV.11.1. Les réseaux "feed-forward"	40
IV.11.1.1. Le Perceptron mono-couche	40
IV.11.1.2. Le Perceptron multi-couches (PMC)	41
IV.11.2.3. Les réseaux à fonction radiale	41
IV.11.2. Les réseaux "feedback"	41
IV.11.2.1. Les cartes auto-organisatrices de Kohonen	41
IV.11.2.2. les réseaux de Hopfield	42
IV.11.2.3. Les ART	43
 Chapitre V : Modélisation de l'évapotranspiration potentielle	
V.1. Définition de la modélisation	44
V.2. Présentation des données.....	44
V.2.1 Inventaire des données	44
V.3. Méthodes d'évaluation de la modélisation	46
V.3.1 Le critère de Nash-Sutcliffe (E)	46
V.3.2. L'erreur moyenne quadratique (RMSE)	46
V.3.3. Le coefficient de détermination (R^2).....	47
V.3. Description du modèle	47
V.3.1. Architecture de base	47
V.3.2. Architecture réseau	49
V.3.3. Apprentissage du modèle.....	51
V.4. Programmation du modèle.....	52
 Chapitre VI : Présentation des résultats	
VI.1. Modèle au pas de temps décadaire	54
VI.1.1. Modèles à une seule entrée	54
VI.1.2. Modèles à deux entrées	58
VI.1.3. Modèles à trois entrées	62
VI.1.4. Conclusion pour le pas de temps décadaire	66
VI.2. Modèles au pas de temps mensuel	68
VI.3. Modèles au pas de temps journalier	73
 Conclusion Générale	 78

INTRODUCTION



Pour une bonne gestion des ressources en eau, et à plus forte raison lorsque ces ressources sont limitées, il est impératif de quantifier au mieux les apports et les demandes en eau. En d'autres termes, connaître de la manière la plus juste possible les différentes composantes du bilan hydrique, que ce soit à l'échelle réduite de la parcelle ou alors à celle plus globale des bassins versants ou du continent.

Les pertes d'eau par évaporation des sols et transpiration des plantes comptent parmi les plus importants intervenants dans le bilan hydrique, et peuvent représenter jusqu'à 60 ou 70 % des transferts d'eau (Doorenbos, 1977).

Ces deux phénomènes ont été groupés sous un même concept d'évapotranspiration, afin de combiner les effets du rejet de l'eau dans l'atmosphère par les sols (surfaces d'eau comprises) et les plantes et n'en faire qu'une seule entité à étudier. Dans les deux cas, il ne s'agit que d'un seul et même phénomène énergétique, qui est le passage de l'eau de l'état de liquide à celui de vapeur, et du même volume d'eau rejeté dans l'atmosphère.

L'évapotranspiration est un phénomène complexe, multifactoriel, influencé par de nombreux paramètres climatiques, mais aussi par la nature physique et chimique des sols, le couvert végétal ou encore par la géographie. De ce fait il apparaît donc comme étant difficile à modéliser, car sa description nécessite la connaissance de tous ces nombreux facteurs, souvent interdépendants.

Nous nous proposons, dans le cadre de ce projet de fin d'études, de tenter une approche de modélisation de l'évapotranspiration par une technique d'intelligence artificielle : Les réseaux de neurones formels ou artificiels. En effet, cette méthode tente d'imiter le comportement des systèmes nerveux naturels, dont le cerveau humain bien sûr, afin d'en reprendre les qualités. Le cerveau humain est capable d'appréhender des problèmes complexes et inédits grâce à ses capacités de classification, de mémorisation et de généralisation. La connaissance de l'évapotranspiration fait partie de cette famille de problèmes complexes, exigeant ces qualités pour être entièrement appréhendée. C'est pour ces raisons que notre choix s'est porté sur cette méthode.

Le travail proposé se répartit comme suit :

Le chapitre I s'intéressera au phénomène à modéliser, et un rapide descriptif y sera présenté. Le chapitre II sera consacré à un rappel rapide des principales approches de calcul. L'équation de Penman sera entièrement développée au chapitre III, et le chapitre IV donnera des notions générales sur les réseaux de neurones artificiels. Le chapitre V s'intéressera au modèle utilisé et la présentation des résultats se fera au chapitre VI.

CHAPITRE I

« Processus physique de l'évapotranspiration ».

I.1. Concept d'évapotranspiration

L'évapotranspiration combine les deux phénomènes d'évaporation des sols et de transpiration des plantes.

L'évaporation est la conversion de l'eau en vapeur et le transfert de cette eau de la surface évaporante vers l'atmosphère.

La transpiration est l'évaporation de l'eau, captée par le système racinaire de la plante, par les stomates, qui sont des orifices présents à la surface des feuilles (Choisel, 1988).

I.1.1. Processus physique de l'évaporation

C'est par le mouvement désordonné des petites molécules d'eau (H_2O) que commence l'évaporation. Ce mouvement est proportionnel à la température, plus la température est élevée plus il est important. C'est l'énergie associée à ce mouvement qui va permettre à certaines molécules de s'échapper dans l'atmosphère à partir de la surface évaporante. Dalton (1802) a établi une loi qui exprime le taux d'évaporation en fonction de la capacité de l'atmosphère de stocker de l'eau, dont la formulation est la suivante (Calvet, 1988) :

$$E = f(u) \times (es - ea) \quad (1.1)$$

Où :

E : Taux d'évaporation ;

es : Pression de la vapeur d'eau dans l'air à saturation ;

ea : Pression de la vapeur d'eau dans l'air ;

$(es - ea)$: Représente la quantité d'eau que l'air peut stocker ;

$f(u)$: Fonction du vent.

En d'autres termes le processus d'évaporation se poursuit tant que la pression de vapeur actuelle ou effective n'a pas atteint la pression de vapeur saturante.

La pression de vapeur saturante est liée proportionnellement à la température : Plus la température est élevée plus la pression de vapeur de saturation augmente. Elle y est liée par la formulation exponentielle à croissance rapide suivante :

$$es = 611 \times \exp\left(\frac{17.27t}{237.3 + t}\right) \quad (1.2)$$

Où:

t : Température en °C

Il s'ensuit donc que la température est une condition nécessaire mais non suffisante pour que le processus d'évaporation se produise et qu'il faut y ajouter la capacité de l'air à accueillir la quantité d'eau évaporée.

L'évaporation dépend de différents facteurs météorologiques ou physiques :

I.1.1.1. Facteurs météorologiques

Ceux-ci influencent essentiellement la quantité de chaleur disponible et la capacité de l'air à stocker de l'eau, on peut citer (Choisel, 1988) :

1. Rayonnement solaire

Chaleur provenant du soleil sous forme d'onde, et qui entretient le cycle de l'eau par les différents flux d'énergie qu'elle permet de produire. Le rayonnement solaire diffère d'un endroit à un autre et dépend des conditions géographiques (latitudes et altitudes). Il est atténué au cours de sa traversée de l'atmosphère terrestre et son effet est modifié au niveau du sol suivant les caractéristiques de ce dernier. D'où l'introduction de la notion de rayonnement net qui est défini comme la quantité d'énergie radiative disponible à la surface de la terre et, donc, influençant directement l'évapotranspiration et pouvant être transformé en d'autre type d'énergie. Le rayonnement net est le résultat de la combinaison de 3 types de rayonnements dans la formulation suivante :

$$R_n = (R_s + R_a) \times (1 - \alpha) - R_t \quad (1.3)$$

R_n : Rayonnement net ;

R_s : Rayonnement solaire direct atteignant le sol ;

R_a : Rayonnement atmosphérique dirigé vers le sol ;

R_t : Rayonnement terrestre (capté et réfléchi par le sol) ;

α : Albédo, ou pourcentage de rayonnement réfléchi par le sol, il varie selon les caractéristiques du sol et peut aller de 0.03 pour une surface d'eau à 0.95 pour une surface de neige fraîche (Toumi, 2001).

2. Température

Le taux d'évaporation est fonction croissante de la température de l'air qui elle est fortement corrélée au niveau d'ensoleillement et est influencé par de nombreux facteurs, tels que les caractéristiques du sol, du couvert végétal, de l'altitude ou encore de la proximité de la mer.

En outre la température suit un cycle diurne avec un maximum et un minimum observable.

3. Humidité relative de l'air

L'humidité relative est le rapport entre la quantité d'eau présente dans l'air et la quantité d'eau maximale que peut contenir cette masse d'air. C'est une autre manière d'exprimer le déficit de saturation, et s'exprime par :

$$Hr = \frac{ea}{es} \times 100 \quad [\%] \quad (1.4)$$

Où

ea : Pression de vapeur d'eau effective ;

es : Pression de vapeur à saturation ;

Hr : Humidité relative.

Lorsque la température baisse l'air se sature car la pression de vapeur à saturation baisse avec la température et l'évaporation n'est plus possible.

Lorsque l'air est saturé ($Hr=100\%$) et pour une température donnée, la pression de vapeur effective a atteint la pression de vapeur de saturation l'évaporation n'est plus possible et l'eau de l'atmosphère commence à se condenser : cette température est appelée température du point de rosée.

4. Pression atmosphérique

Elle indique la variation de la masse d'air au dessus de la surface évaporante, et influence l'humidité relative et absolue. Mais son effet n'est pas clairement défini dans le processus d'évaporation car les variations de pression atmosphérique sont le plus souvent accompagnées de celle d'autres paramètres tels que la température et l'humidité relative.

5. Le vent

Le vent est un facteur important car il joue le rôle de catalyseur en permettant, par le mélange de l'air ambiant, de remplacer l'air saturé près de la surface évaporante par un air non saturé capable d'absorber l'évaporation.

1.1.1.2. Facteurs physiques

Ils sont divers et variés, et sont le résultat direct de l'environnement, on peut citer en exemple : l'étendue et la salinité, pour les surfaces d'eau, ou alors, la teneur en eau et la couleur pour les sols nus (Allen et al, 1998).

1.1.2. Processus physique de la transpiration des plantes

Lorsqu'un sol est couvert de végétation, il est établi que les échanges par transpiration sont quantitativement plus importants que par évaporation directe à partir du sol.

La transpiration est l'évaporation de l'eau contenue dans les tissus de la plante vivante et son transfert dans l'atmosphère. Ce processus se déroule essentiellement par les stomates, qui sont de petits pores situés sur la surface des feuilles et qui lui permettent de communiquer avec son environnement extérieur par les transferts de chaleurs et de matière. Pratiquement toute l'eau absorbée par la plante est rejetée par transpiration et seulement une infime partie est gardée au sein de la plante. Connaître la quantité d'eau transpirée, c'est connaître les besoins en eau de la plante.

La transpiration est influencée par (Allen et al, 1998):

- **Les caractéristiques de la plante**

L'importance de la transpiration varie d'un type de plante à une autre, et dépend de l'espèce ainsi que de son adaptation à son milieu : Une plante des régions tropicales humides transpirera différemment d'une plante issue d'un milieu aride, et qui aura tendance à limiter au maximum sa transpiration par souci d'« économie ». S'ajoute à cela, la variation de la transpiration selon l'âge et le stade de développement végétatif de la plante.

- **Les caractéristiques environnementales**

Comme le type de sol sur lequel est implanté la culture, et sa capacité à stocker l'eau et à la conduire par capillarité ou par écoulement souterrain. Ainsi que des caractéristiques physico-chimiques de l'eau comme sa salinité ou son PH.

- **Les conditions climatiques**

Une fois la quantité d'eau arrivée au niveau de la surface de la feuille, c'est le processus de l'évaporation qui commence et les paramètres climatiques influents sont donc les mêmes que ceux cités au § (I.1.1.1).

I.2. Les différents types d'évapotranspiration

I.2.1. Evapotranspiration de référence (ET_0)

Ce concept a été introduit pour déterminer la demande évaporative de l'atmosphère, au-delà de toute influence du type de culture, de son stade de développement et des techniques agraires.

C'est la somme des pertes en eau par évaporation et transpiration d'une culture de référence, en général un type de gazon, abondamment arrosée, de hauteur uniforme, en pleine période de croissance, recouvrant entièrement le sol et évoluant dans des conditions sanitaires idéales. Les seuls facteurs influençant cette évapotranspiration sont les facteurs climatiques ; donc, l'évapotranspiration de référence est un paramètre climatique qui indique la capacité d'évaporation de l'atmosphère à un endroit fixé et à une période de l'année bien précise. Une fois l'évapotranspiration de référence établie, des correspondances sont effectuées entre les différents types de culture et la culture de référence.

Elle peut être dite **potentielle** et est notée **ETP** dans ce cas là.

L'évapotranspiration de référence ou potentielle peut-être mesurée ou calculée ; pour ce qui est du calcul, la FAO préconise l'emploi de la formule de Penman-Monteith, afin d'homogénéiser les données et permettre leur comparaison (Allen et al, 1998).

I.2.1. Evapotranspiration maximale (ET_m)

Se définit comme l'ensemble des pertes en eau par évaporation et transpiration d'une culture donnée, autre que le gazon, à différents stades de développement végétatif, lorsque les apports en eau sont suffisants et que les conditions agronomiques sont optimales.

Afin d'éviter de refaire les mesures ou les calculs pour chaque type de culture et dans un souci d'homogénéisation, les chercheurs sont arrivés à déterminer les besoins en eau des cultures, équivalents à l' ET_m , par la correction de l'évapotranspiration de référence ET_0 , par un coefficient appelé "coefficient cultural" (K_c) en utilisant la formule suivante :

$$ET_m = K_c \times ET_0 \quad (1.5)$$

L'échelle de temps sur laquelle les besoins sont calculés peut être l'heure, la journée, la décade, le mois ou la phase de croissance, selon l'objectif poursuivi et la disponibilité de données. La valeur du coefficient K_c est largement affectée par la nature de la culture, sa hauteur, sa durée de cycle, et son taux de croissance, mais aussi par la fréquence des pluies ou de l'irrigation au début du cycle de la culture. K_c est toujours établi expérimentalement au début, pour une région et une culture données, puis ensuite confiné dans des tables pour une utilisation ultérieure dans la même région ou dans une région similaire. (Allen et al, 1998).

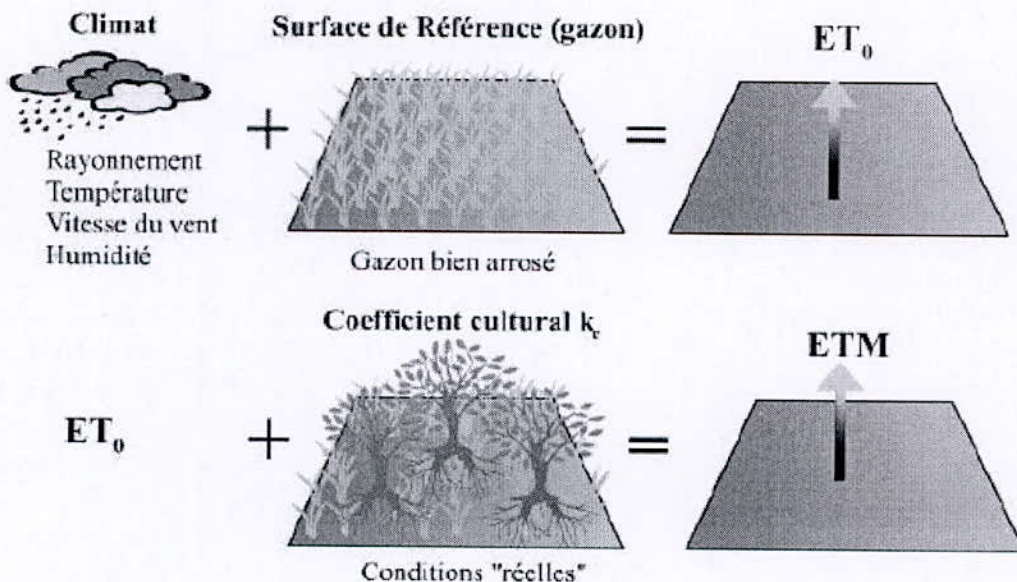


Figure I.1: Relation entre ET_0 et ET_m

I.2.3. Evapotranspiration réelle

Elle se définit comme étant la valeur mesurée par un dispositif approprié permettant de donner la quantité d'eau rejetée par évapotranspiration par une culture dans des conditions climatiques, agraires et sanitaires normales (Allen et al, 1998).

CHAPITRE II

CHAPITRE II

Différentes méthodes
de
Quantification

CHAPITRE II

« Historique des différentes méthodes de quantifications de l'évapotranspiration »

Au sortir de la deuxième guerre mondiale, vers le milieu du XX^e siècles, la population des pays occidentaux a connu une explosion démographique importante. Les besoins en nourriture se sont faits de plus en plus pressants et les techniques agricoles ont du être modernisées afin de satisfaire la demande croissante. L'intensification de l'agriculture devait passer par une utilisation accrue de l'irrigation. C'est pour évaluer les besoins en eau des cultures de manière précise que différentes approches pour évaluer l'évapotranspiration ont été tentées.

II.1. Formules résultant d'ajustements statistiques

II.1.1. Formule de BLANEY et CRIDDLE (1945)

Elaborée à partir d'expérimentations effectuées en zone sub-aride, elle prend la formulation suivante (Toumi, 2001) :

$$Etp = K \times F \quad (II.1)$$

Où :

K : Coefficient peu variable, intégrant les autres facteurs non étudiés il varie entre 0.5 et 0.85 ;

F : Force évaporante , fonction de la température et de la durée d'insolation , elle prend la formulation suivante :

$$F = 0.254 \times (1.8 \times t + 32) \times p \quad (II.2)$$

Où :

t : Température moyenne mensuelle

p : Pourcentage d'heures diurne de la période considérées par rapport au total annuel d'heures diurnes.

Cette méthode permet d'estimer l'ETP mensuelle à partir de données facilement accessibles, que sont la température moyenne mensuelle et l'éclairement de la région. Néanmoins elle n'est efficace réellement que dans les régions sub-arides à aride, c'est-à-dire, là ou la force évaporante dépend essentiellement de la température, et elle marque ses limites dans les régions humides car elle néglige l'humidité de l'air dans la quantification de la force évaporante (Ventura et al, 1999).

II.1.2. Formule de Thornthwaite (1948)

Thornthwaite, botaniste et climatologue américain, avait noté que la croissance des plantes dépendait pour une grande part des échanges d'eau avec l'atmosphère. Il est un des premiers à dégager la notion d'évapotranspiration, et à tenter de la lier à des paramètres climatiques facilement accessibles, dans ce cas la température moyenne de l'air et des données astronomiques (Seyni, 2000).

En utilisant des données expérimentales obtenues sur des cases lysimétriques, et en y appliquant des ajustements statistiques, il obtient la formulation suivante :

$$Etp = 16 \times \left(\frac{10 \times t}{I} \right)^a \times f(\lambda) \quad (II.3)$$

Où

t : Température moyenne en °C ;

a : Fonction de l'indice I ; la formule de a est :

$$a = \frac{1.6}{100} I + 0.5 \quad (II.4)$$

I : Indice thermique annuel, somme des douze indices thermiques mensuels i s'exprimant par :

$$i = \left(\frac{t}{5} \right)^{1.514} \quad (II.5)$$

$f(\lambda)$: Terme correctif, dépendant de la durée théorique d'insolation.

La méthode de Thornthwaite a l'avantage de n'exiger que très peu de données, mais de ce fait reste éloignée de la vérité car négligeant l'humidité relative de l'air, un facteur très important. De plus elle ne permet pas de connaître l'ETP en cours car il faut attendre l'écoulement des douze mois de l'année pour avoir l'indice I . (Ventura et al, 1999)

II.1.3. Formule de TURC (1961)

Turc complète l'étude entreprise par Henin et Ternusien, et remarque que l'évapotranspiration sur cases lysimétriques rendait assez bien compte de l'évapotranspiration des bassins versants et y ajuste une formule quantifiant le phénomène par pas de temps décadaire et plus (Toumi, 2001).

Si l'humidité relative est supérieure à 50%, la formule de Turc s'écrit :

$$Etp = \left(\frac{0.13 \times t}{t + 15} \right) \times (Rg + 50) \quad (II.6)$$

Sinon, elle s'écrit :

$$Etp = \left(\frac{0.13 \times t}{t + 15} \right) \times (Rg + 50) \times \left(1 + \frac{50 - H_r}{70} \right) \quad (II.7)$$

Où :

t : Température moyenne de la période considérée (sous abri)

Rg : Radiation solaire globale évaluée à partir de la formulation suivante :

$$Rg = Iga \times \left(0.18 + 0.62 \times \frac{h}{H} \right) \quad (II.8)$$

Avec :

h : Durée réelle moyenne d'insolation ;

H : Durée astronomique moyenne du jour ;

I_{ga} : Radiation solaire directe en l'absence d'atmosphère en $\text{cal/cm}^2/\text{jour}$;
 Hr : Humidité relative.

La formule de Turc est relativement facile d'emploi car elle n'utilise que des paramètres climatiques facilement mesurables. Contrairement aux deux précédentes, elle prend en compte l'humidité relative de l'air, cela lui permet de donner des résultats assez satisfaisants, mais son efficacité baisse lorsque le climat est moins humide. De plus, elle ne permet pas d'avoir l'ETP à court terme car elle est valable pour des pas de temps décennaux et plus (Seyni, 2000).

II.1.4. Formule de JENSEN et HAISE (1963)

Elaborée pour les régions arides et semi-arides, et dont la formulation est la suivante (Toumi, 2001) :

$$Etp_0 = (0.025 \times t + 0.08) \frac{Rg}{59} \quad (II.9)$$

Etp_0 : Evapotranspiration potentielle d'une culture de référence ;

t : Température moyenne journalière en $^{\circ}\text{C}$;

Rg : Rayonnement de courte longueur d'onde en : $\text{cal/cm}^2/\text{jour}$.

Première formule à ajustement statistique à s'intéresser à l'ETP au pas de temps journalier, elle utilise elle aussi des données facilement accessibles que sont la température moyenne journalière et le rayonnement global. Néanmoins, il a été constaté que cette formule avait tendance à sous estimer l'ETP au printemps et à la surestimer en été. Cela est attribué au fait qu'elle ne prend pas en compte de manière explicite les autres facteurs et qu'elle privilégie la température au détriment du rayonnement. (Seyni, 2000)

II.1.5. Méthode combinée de VAN BAVEL (1968)

Van Bavel propose une solution approximative de l'ETP issue d'une équation combinée appliquée à une herbe courte, et bien irriguée dont la formule est la suivante (Toumi, 2001):

$$Etp = 0.167 \times Rn + 14 \cdot 10^3 \times U \times \frac{(T_{\max} - T_{\min})}{22} - 15 \cdot 10^2 \times (T_{\max} - T_{\min}) \quad (II.10)$$

Avec :

Etp : Evapotranspiration potentielle en (mm/jour) ;

U : Vitesse du vent à 1 m en (km/jour) ;

T_{\max} et T_{\min} : Température journalière maximale et minimale en ($^{\circ}\text{C}$) ;

Rn : Rayonnement net en $\text{cal/cm}^2/\text{jour}$.

Cette formule est plus exigeante en termes de données car elle a besoin de deux mesures de températures journalières, ainsi que de celle de la vitesse du vent à 1 mètre et du rayonnement net. Mais elle marque un progrès car elle est la première méthode à ajustement statistique à introduire la vitesse du vent (Ventura et al, 1999).

II.1.6. Formule de JENSEN et WRIGHT (1972)

Cette formule ne prend en compte que le déficit de saturation de l'air et la vitesse du vent et prend la formulation suivante (Toumi, 2001):

$$E_{tp} = (e_{ws} - e_z) \times (0.347 - 0.186 \times U) \quad (\text{II.11})$$

Avec :

e_{ws} : Tension de vapeur à la surface évaporante en(mb) ;

e_z : Tension de vapeur à l'altitude z en (mb) ;

U : Vitesse du vent (en m/s).

Première formule à ajustement statistique à prendre en compte le déficit de saturation, elle manque pourtant d'efficacité car négligeant les autres facteurs (Seyni, 2000).

II.1.7. Formule de PRIESTLEY et TAYLOR (1972)

Elle s'adapte pour une grande variété de plantes et de climats, et prend la formulation suivante (Toumi, 2001) :

$$E_{tp} = \gamma \times \left[\frac{d}{d+a} \right] \times Rn \quad (\text{II.12})$$

Ou :

γ : Constante psychrométrique en (Hpa/°c) ;

d : Coefficient numérique en fonction de la température de l'air ;

a : Constante ;

Rn : Rayonnement net en cal/cm²/jour.

II.2. Formules basées sur le bilan énergétique :

II.2.1. Formule de BOUCHET (1964)

En appliquant le calcul du bilan énergétique sur un dispositif expérimental constitué d'une pastille de Piche placé sous abri, L'auteur a pu dégager une formule de quantification de L'ETP basée sur l'évaporation à partir de la pastille Piche et qui prend la formulation suivante (Bouchet, 1964) :

$$E_{tp} = \alpha \times (\lambda + 1) \times E_{piche} \quad (\text{II.13})$$

Avec :

E_{piche} : Evaporation mesuré sur la pastille de Piche ;

α : Coefficient de l'abri dans lequel est installée la pastille de Piche ;

λ : coefficient dépendant de θ avec :

$$\theta = \frac{T_{\max} + 3 \times T_{\min}}{4} \quad (\text{II.14})$$

T_{\max} et T_{\min} : Température journalière maximale et minimale.

En effectuant ces mesures sous abri, Bouchet a pu obtenir une formule n'exigeant que très peu de paramètres, et qui sont les températures extrêmes, le point de rosée et l'évaporation, à partir de la pastille de Piche, et de ce fait, négliger l'effet du rayonnement solaire. Le reste des paramètres est inclus dans la mesure de l'évaporation à partir de la pastille de Piche.

Mais cette formule est très dépendante de la qualité de l'appareillage de mesures et de ses différentes caractéristiques, et la quantification de ces paramètres peut être source d'erreur (Seyni, 2000).

II.2.2. Formule de BROCHET et GERBIER (1974)

Ils évaluent séparément la partie radiative et la partie advective. La partie radiative peut être estimée à partir du rayonnement global alors que la partie advective peut être soit estimée à partir de l'évaporation de Piche (Brochet, 1972) :

$$E_{tp} = m \times R_g + n \times E_p \quad (\text{II.15})$$

Avec :

R_g : Rayonnement global en (cal/cm^2) ;

E_p : Evaporation de Piche;

m : Coefficient dépendant du lieu et de la période de l'année ;

n : Coefficient dépendant du type d'abri.

Ou par le déficit de saturation :

$$E_{tp} = m \times R_g + P \times f(v) \times (E - e_a) \quad (\text{II.16})$$

Avec :

$f(v)$: Fonction du vent ;

P : Pression barométrique dépendant de l'altitude de la région étudiée ;

$(E - e_a)$: Déficit de saturation.

Ces deux auteurs combinent la simplicité de la formule de Bouchet, car elle utilise un matériel peu coûteux (mais auquel ils reprochent de ne pas être assez sensible aux autres paramètres) à l'efficacité de la formule de Penman à laquelle ils reprochent le nombre de paramètres trop important. C'est la raison pour laquelle ils ont proposé une deuxième formulation se passant de l'évaporation de Piche.

Cette méthode donne de bons résultats, et même si elle fait appel à un assez grand nombre de paramètres, elle reste assez facile d'emploi (Seyni, 2000).

II.2.3. Formule de Penman (1948)

C'est la formule la plus couramment utilisée, et celle donnant les résultats les plus précis (fao). Elle est l'objet de ce travail et est développée au chapitre III

II.3. Méthodes de mesures de l'ETP

On peut obtenir la mesure de l'évapotranspiration par différents dispositifs (Allen et al, 1998):

II.3.1. Les Bacs

La mesure par bac consiste à mesurer la hauteur d'eau évaporée durant un laps de temps déterminé. Cette hauteur d'eau est ensuite pondérée par un coefficient dépendant du type de bac et des conditions climatiques pour obtenir l'évapotranspiration.

$$E_{tp} = K_b \times E_{bac} \quad (II.17)$$

On peut distinguer différents types de bacs :

- **Bac Colorado**

C'est un bac enterré, de forme carrée de 1m de coté et de 60 cm de profondeur, dont 50 cm sont sous terre. Il est rempli d'eau jusqu'au niveau du sol, c'est-à-dire, à 10 cm du rebord. Il peut contenir une assez grande quantité d'eau, et son inertie thermique est plus tôt importante, son comportement est proche de celui des conditions naturelles.

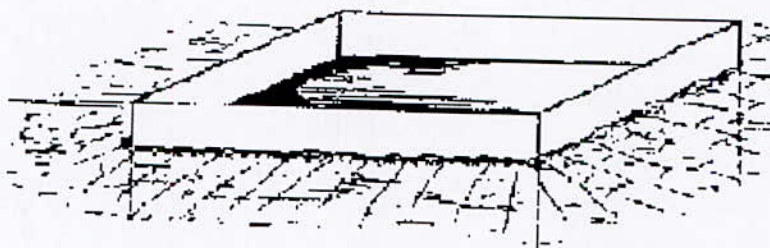


Figure II.1 : Bac Colorado

- **Bac classe A**

Ce bac est constitué d'un récipient cylindrique, posé sur un support en bois, de 15 cm de hauteur, l'isolant du sol et permettant son aération. Il a un diamètre de 121.9 cm pour 25.4 cm de hauteur. On y maintient la hauteur d'eau entre 17.5 et

20 cm de profondeur. L'inertie thermique de la quantité d'eau contenue dans ce bac est assez faible, et il est très sensible aux variations de température.

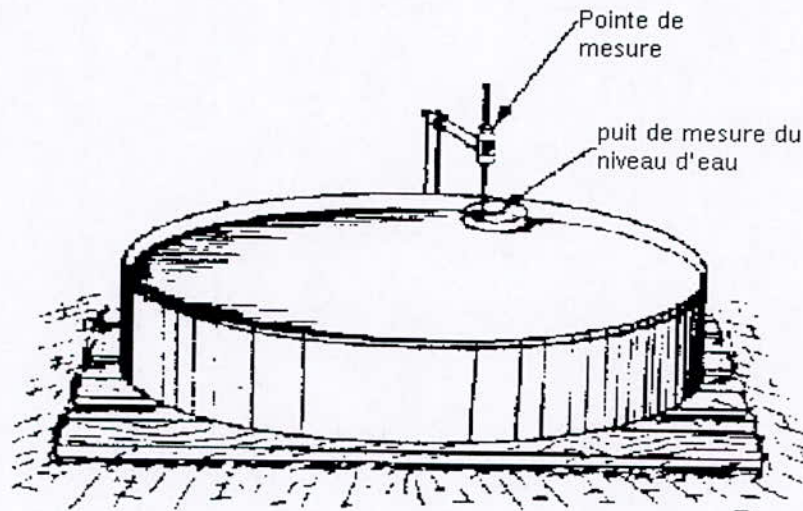


Figure II.2 : Bac Classe A

- **Bac flottant**

De formes diverses, ils servent à déterminer l'évaporation de grandes étendues d'eau tel que les lacs et les retenues de barrages.

II.3.2. Les lysimètres

Il s'agit d'une cuve enterrée au milieu d'une parcelle irriguée, servant de zone tampon et l'isolant de l'influence de son environnement direct. Elle est plantée de la même culture et est drainée de manière à pouvoir y calculer le bilan hydrique. La différence entre les volumes d'eau apportés et les volumes drainés représente la quantité d'eau perdue par évapotranspiration. C'est à l'aide d'un lysimètre (ou cuve lysimétrique) que l'on mesure l'évapotranspiration de référence.

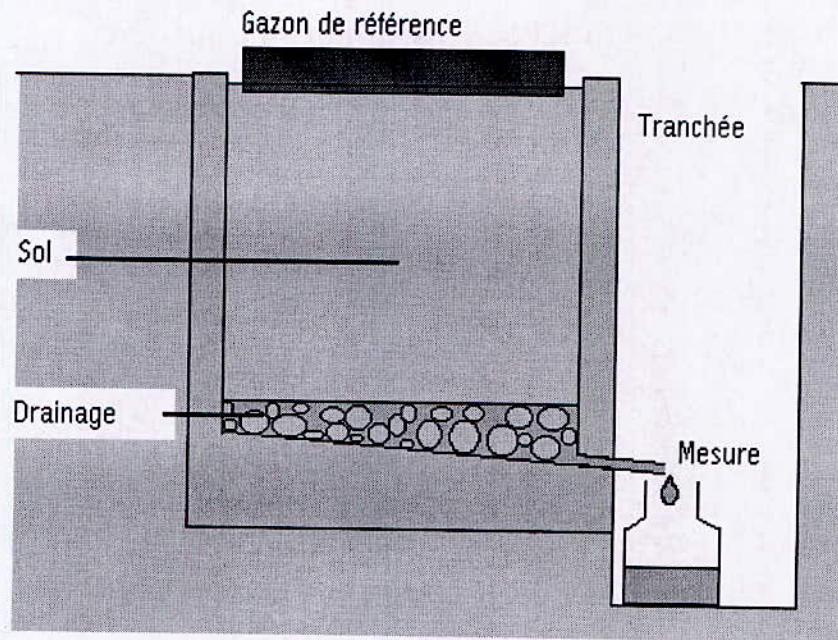


Figure II.3 : Lysimètre

CHAPITRE III

CHAPITRE III

Développement de
l'équation
de Penman

CHAPITRE III

« Développement de la formule de Penman »

La formule de Penman résulte de la résolution de l'équation du bilan d'énergie, dont voici les principales étapes de développement (Lehre, 2001).

III.1. Développement mathématique de l'équation de Penman

Penman s'est d'abord intéressé à l'évaporation à partir d'un plan d'eau et a développé une équation en partant du bilan énergétique à la surface de celle-ci. Il s'est intéressé à l'évaporation à partir d'un bac, en négligeant les échanges énergétiques à travers les parois du bac ce qui donne une équation du bilan énergétique réduite à l'expression (III.1)

$$Q_n = Q_h + Q_e \quad (\text{III.1})$$

Où :

Q_n : Radiation nette du soleil en (cal/cm²) ;

Q_h : Conduction atmosphérique en (cal/cm²);

Q_e : Energie utilisée pour l'évaporation en (cal/cm²).

En divisant l'équation (III.1) par ρL :

Où

L : Chaleur latente de vaporisation ;

ρ : Masse volumique de l'eau.

Ce qui donne l'équation (III.2) :

$$\frac{Q_n}{\rho L} = \frac{Q_h}{\rho L} + \frac{Q_e}{\rho L} \quad (\text{III.2})$$

En posant que :

$H = \frac{Q_n}{\rho L}$: Rayonnement net du soleil, exprimé en hauteur d'eau que ce rayonnement peut évaporer (m) ;

$K = \frac{Q_h}{\rho L}$: Flux de chaleur atmosphérique, exprimé en hauteur d'eau que ce flux pourrait évaporer (m) ;

$E = \frac{Q_e}{\rho L}$: Hauteur d'eau évaporée (m).

$$H = K + E \quad (\text{III.3})$$

En introduisant R , la constante de Bowen, comme étant le rapport de Q_h sur Q_e

$$R = \frac{Q_h}{Q_e} = \gamma \times \frac{(T_s - T_a)}{(e_s - e_a)} \quad (\text{III.4})$$

Avec :

γ : Constante Psychrométrique = 0.66 mb/°C

T_s : Température de la surface d'eau ;

T_a : Température de l'atmosphère ;

e_s : Pression de vaporisation de la surface d'eau ;

e_a : Pression de vaporisation de l'atmosphère.

En posant que $\frac{Q_n}{Q_e} = \frac{K}{E}$, de (III.4) on trouve :

$$K = E \times R \quad (\text{III.5})$$

En remplaçant dans (III.3) on trouve :

$$H = E \times R + E = E \times (R + 1) \quad (\text{III.6})$$

Et sachant que :

$$R = \gamma \times \frac{(T_s - T_a)}{(e_s - e_a)}$$

On peut alors écrire (III.6) sous la forme suivante :

$$\frac{H}{E} = R + 1 = 1 + \gamma \times \frac{(T_s - T_a)}{(e_s - e_a)} \quad (\text{III.7})$$

Dalton a donné l'expression (III.8) comme formule de quantification du taux du transfert de masse par évaporation :

$$E = f(u) \times (e_s - e_a) \quad (\text{III.8})$$

Où :

$f(u)$: Fonction du vent ;

e_s : Pression de vaporisation de la surface d'eau ;

e_a : Pression de vaporisation de l'atmosphère.

Penman a quand à lui proposé une expression légèrement modifiée, en gardant la même fonction du vent :

$$E_a = f(u) \times (e'_a - e_d) \quad (\text{III.9})$$

Où :

E_a : Transfert de masse par évaporation en (cm) ;

e'_a : Pression de vapeur à saturation de la surface évaporante à température égale à celle de l'air en (mb) ;

e_d : Pression de vapeur à saturation de l'atmosphère en (mb).

En divisant l'expression de Penman (E_a) par celle de Dalton (E), La fonction du vent se simplifie et l'on obtient l'expression suivante :

$$\frac{E_a}{E} = \frac{(e'_a - e_d)}{(e_s - e_d)} = 1 - \frac{(e_s - e'_a)}{(e_s - e_d)} \quad (\text{III.10})$$

Penman a ensuite introduit le paramètre Δ , car en sachant que la pression de vapeur à saturation est proportionnelle à la température, et que donc la représentation de la courbe $e=f(T)$ nous donne une courbe rapidement croissante, On peut admettre que Δ est l'approximation de la pente de cette courbe et est donnée par l'expression suivante :

$$\Delta = \frac{(e_s - e'_a)}{(T_s - T_a)} \quad (\text{III.11})$$

Que l'on peut écrire encore sous la forme suivante :

$$T_s - T_a = \frac{(e_s - e'_a)}{\Delta} \quad (\text{III.12})$$

En remplaçant (III.12) dans (III.7) on obtient l'expression suivante :

$$\frac{H}{E} = 1 + \gamma \times \frac{\left(\frac{e_s - e'_a}{\Delta} \right)}{e_s - e_a} \quad (\text{III.13})$$

Ou encore :

$$\frac{H}{E} = 1 + \left(\frac{\gamma}{\Delta} \right) \times \left(\frac{e_s - e'_a}{e_s - e_a} \right) \quad (\text{III.14})$$

De l'expression (III.10), on a que :

$$\frac{e_s - e'_a}{e_s - e_a} = 1 - \frac{E_a}{E} \quad (\text{III.15})$$

En remplaçant (III.15) dans (III.14), on trouve :

$$\frac{H}{E} = 1 + \frac{\gamma}{\Delta} \times \left[1 - \frac{E_a}{E} \right] \quad (\text{III.16})$$

En multipliant l'expression (III.16) par Δ , on obtient :

$$\frac{H}{E} \times \Delta = \Delta + \gamma - \frac{\gamma \times E_a}{E} \quad (\text{III.17})$$

Et en dégagant l'expression de E de (III.17), on a :

$$E = \frac{H \times \Delta + \gamma \times E_a}{\Delta + \gamma} \quad (\text{III.18})$$

Que l'on peut encore mettre sous la forme suivante :

$$E = \frac{\left(\frac{\Delta}{\gamma} \times H + E_a \right)}{\left(\frac{\Delta}{\gamma} + 1 \right)} \quad (\text{III.19})$$

L'équation (III.19) est la formule de Penman pour quantifier l'évaporation d'un plan d'eau. Penman a donné une relation empirique pour l'estimation de E_a :

$$E_a = 0.47 \times (0.5 + 0.01 \times U_2) \times (e'_a - e_2) \quad (\text{III.20})$$

Avec :

U_2 : Vitesse du vent à 2 m au dessus de la surface évaporante ;

e_2 : Pression de vapeur à 2 m de la surface d'eau évaporante ;

$\frac{\Delta}{\gamma}$: Est une fonction tabulée de la température, ce qui nous permet de calculer E

sans avoir à mesurer T_s ;

$H = \frac{Q_n}{\rho L}$: Est obtenue par mesure directe ou par différentes formules empiriques.

Penman a ensuite modifié cette équation pour l'adapter à l'évapotranspiration d'un couvert végétal, en modifiant l'équation de départ du bilan énergétique et en posant les hypothèses suivantes :

- Pour une période d'un jour au moins, les échanges de chaleur entre la plante et le sol peuvent être négligés.
- Tous les types de transferts par advection peuvent être négligés.

En partant de ces simplifications l'équation de départ devient :

$$Q_n = Q_h + Q_{et} \quad (\text{III.21})$$

Avec :

Q_n : Rayonnement solaire net ;

Q_h : Energie transférée du couvert végétal à l'atmosphère par conduction ;

Q_{et} : Energie utilisée pour l'évapotranspiration.

En réintroduisant, comme précédemment, la constante de Bowen R :

$$R = \frac{Q_h}{Q_{et}} = \gamma \times \left(\frac{T_s - T_a}{e_s - e_a} \right) \quad (\text{III.22})$$

Où dans ce cas là :

T_s : Température de la surface évaporante (sol + couvert végétal) ;

T_a : Température de l'air ;

e_s : Pression de vapeur de la surface évaporante (sol + couvert végétal) ;

e_a : Pression de vapeur de l'air.

En donnant la valeur de Q_h à partir de (III.22) et en l'injectant dans (III.21) :

$$Q_n = Q_{et} + R \times Q_{et} \quad (\text{III.23})$$

$$Q_n = Q_{et} \times (1 + R) \quad (\text{III.24})$$

D'ou:

$$Q_{et} = \frac{Q_n}{1 + R} \quad (\text{III.25})$$

En remplaçant R par l'expreson (III.22) on obtient :

$$Q_{et} = \frac{Q_n}{\left(1 + \gamma \times \frac{T_s - T_a}{e_s - e_a} \right)} \quad (\text{III.26})$$

En divisant Q_{et} par ρL

$$Et = \frac{Q_{et}}{\rho L} \quad (III.27)$$

Qui devient :

$$Et = \frac{1}{\rho L} \times \frac{Q_n}{\left(1 + \gamma \times \frac{T_s - T_a}{e_s - e_a}\right)} \quad (III.28)$$

Où :

T_s et e_s étant pratiquement impossibles à déterminer à cause de la diversité de la surface évaporante constituée du sol et du couvert végétal, ils sont remplacés par T_2 et e_2 qui sont des mesures prises à un autre niveau de l'atmosphère.

Finalement :

$$Et = \frac{1}{\rho L} \times \frac{Q_n}{\left(1 + \gamma \times \frac{T_2 - T_1}{e_2 - e_1}\right)} \quad (III.29)$$

Penman a encore émis l'hypothèse que si le couvert végétal était homogène, on pouvait négliger Q_h et que donc $Q_n = Q_{et}$.

III.2. Equation de Penman-Monteith

Monteith (1981) a modifié la formule de Penman pour l'évapotranspiration, en y introduisant la résistance du couvert végétal à l'action des variations atmosphériques (Allen et al, 1998) :

$$Et_0 = \frac{R_n \times \Delta + \frac{\rho \times C_p \times \delta_e}{r_a}}{\lambda \times \left[\Delta + \gamma \times \left(1 + \frac{r_s}{r_a}\right) \right]} \quad (III.30)$$

Où :

Et_0 : Evapotranspiration de référence calculée par la relation de Penman [mm/s] ;

R_n : Rayonnement net [W/m^2] ;

Δ : Pente de la courbe de pression de vapeur à la température moyenne de l'air en [$kPa/^\circ C$] ;

ρ : Densité de l'air à pression constante [Kg/m^3] ;

C_p : Capacité thermique de l'air humide ;

δ_e : Différence entre la pression de vapeur saturante e_s et la pression de vapeur

e_a effective dans l'air [kPa], avec :

$$\delta_e = e_s - e_a \quad (\text{III.31})$$

r_a : Résistance aérodynamique [s/m] (descripteur météorologique traduisant le rôle des turbulences atmosphériques dans le processus d'évaporation) ;

r_s : Résistance du couvert végétal [s/m] ;

λ : Chaleur latente de vaporisation de l'eau [j/Kg] ;

γ : Constante psychrométrique [kPa/°C].

Pour les calculs pratiques, certaines grandeurs sont considérées comme constantes et les autres sont calculées à base de données météorologiques.

La pression de vapeur à saturation se détermine par :

$$e_s = 0.611 \times \exp\left(\frac{17.27 \times T}{T + 273.3}\right) \quad (\text{III.32})$$

La relation liant la pression de vapeur à l'humidité relative s'écrit

$$e_a = e_s \times \frac{H_r}{100} \quad (\text{III.33})$$

H_r : Humidité relative en [%] ;

La pente quand à elle s'obtient par :

$$\Delta = \frac{4098 \times e_s}{(T + 273.3)^2} \quad (\text{III.34})$$

Δ : Pente de la courbe tension de vapeur fonction de la température en [kPa/°C²]

La résistance aérodynamique s'obtient par la formule suivante :

$$r_a = \frac{\ln\left[\frac{Z_m - d}{Z_{om}}\right] \times \ln\left[\frac{Z_h - d}{Z_{oh}}\right]}{K^2 \times U_z} \quad (\text{III.35})$$

Où :

r_a : Résistance aérodynamique [s m⁻¹] ;

Z_m : Hauteur à laquelle est mesurée la vitesse du vent [m] ;

Z_h : Hauteur à laquelle est mesurée l'humidité de l'air [m] ;

d : niveau de référence des mesures [m] ;

Z_{om} : Hauteur ou rugosité influençant la vitesse [m] ;

Z_{oh} : Hauteur ou rugosité influençant le transfert de chaleur et de vapeur [m] ;

k : Constante de Von Karman (K=0.41) ;

u_z : Vitesse du vent à la hauteur z [m s⁻¹]

r_s : Exprime la résistance opposée au flux de vapeur à travers le couvert végétal

$$r_s = \frac{r_l}{\pi} \quad (\text{III.36})$$

Où :

r_s : Résistance de la surface évaporante [$s \text{ m}^{-1}$],

r_l : Résistance individuelle des feuilles [$s \text{ m}^{-1}$],

π Index du couvert végétal ; sans dimensions

r_l exprime la résistance des feuilles des végétaux pris individuellement , et dépend du type et du stade de développement de la plante ; il augmente en général avec l'âge de la plante jusqu'à un seuil limite, et est déterminé par des études de physiologie.

π représente la densité du couvert végétal, il est le résultat du rapport de la surface occupée par les feuilles de la canopée en [m^2] à la surface du sol considéré en [m^2]. Il dépend du type de végétation, du stade de développement végétatif ainsi que des saisons.

III.3. Equation de Penman-ANRH

L'approche de l'équation de Penman pour le calcul de l'évapotranspiration pour le nord Algérien a été effectuée par l'Agence Nationale des Ressources Hydrauliques dont voici un résumé des principales étapes (ANRH, 2004):

Le type de formulation choisi est le suivant :

$$Etp = (H + a) \times (b \times T + c) \quad (\text{III.38})$$

Où :

Etp : Evapotranspiration potentielle mensuelle ;

H : Durée théorique d'insolation ;

T : Température moyenne mensuelle ;

a, b, c : Coefficients à déterminer par ajustements statistiques.

Les ajustements statistiques ont été effectués sur une base de données de mesures de températures et d'évapotranspirations prises dans 19 stations réparties à travers tout le Nord Algérien.

Les calculs ont abouti à la formule suivante :

$$Etp = (H - 187) \times (0.032 \times T + 0.077) \quad (\text{III.39})$$

Formule qui a été successivement corrigée par deux coefficients K_r et K_m :

$$Etp = K_r \times (H - 187) \times (0.032 \times T + 0.077) \quad (\text{III.40})$$

$$Etp = K_r \times K_m \times (H - 187) \times (0.032 \times T + 0.077) \quad (\text{III.41})$$

Où :

K_r : Coefficient régionale, spécifique à chaque région ;

K_m : Coefficient mensuel, différent pour chaque mois.

Les valeurs des différents coefficients sont connues et tabulées.

CHAPITRE IV

CHAPITRE IV

Notions générales
sur les
réseaux de neurones

Introduction à l'intelligence artificielle

Les réseaux

de neurones

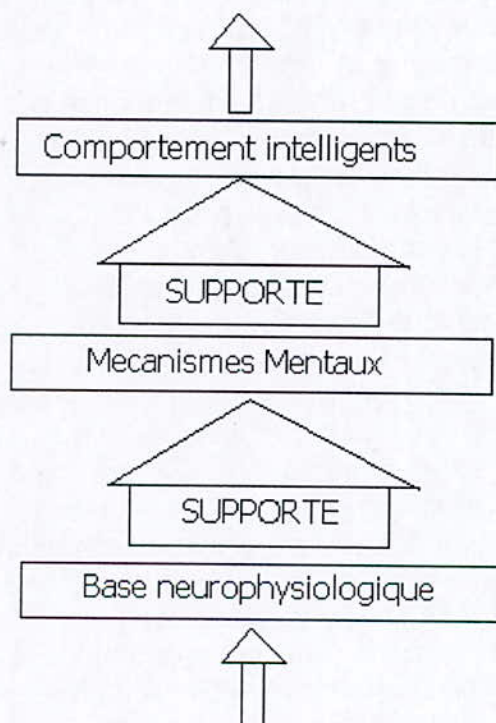
artificiels

CHAPITRE IV

« Notions générales sur le réseaux de neurones »

IV.1. Introduction

La démarche qui mène aux réseaux de neurones est semblable à celle qui mène à l'intelligence artificielle. Avec l'intelligence artificielle, les chercheurs ont voulu modéliser le raisonnement et la connaissance pour recréer de manière explicite les différentes étapes de la réflexion d'un expert du domaine considéré et obtenir à partir d'une question la réponse appropriée. L'hypothèse sous-jacente est que le comportement intelligent est obtenu grâce à ces mécanismes mentaux. La modélisation des mécanismes mentaux par des moteurs d'inférence et de la connaissance par des bases de règles permet d'obtenir un "comportement intelligent" (Touzet, 1992).



Cette approche s'applique relativement bien aux phénomènes qui peuvent être représentés sous forme symbolique, par exemple la démonstration de théorèmes. Par contre, elle est mal adaptée à des problèmes où les données se présentent sous forme numérique (Touzet, 1992).

Une autre école de pensée consiste à construire un modèle des mécanismes physico-chimiques du cerveau. Ces mécanismes cellulaires sont ceux qui font évoluer le cerveau et calculent la réponse de l'organisme à des stimuli extérieurs. Ils créent le raisonnement et donc le comportement intelligent. Si le modèle construit est suffisamment fidèle, un tel système devrait également avoir un comportement

intelligent. Les réseaux de neurones sont issus de ce courant de pensée et des études menées depuis le début du siècle sur l'anatomie du cerveau.

De tels modèles sont bien adaptés aux traitements de données numériques, puisqu'ils utilisent pour acquérir leurs entrées, des mécanismes analogues aux capteurs humains qui acquièrent un signal sous la forme de différents potentiels électriques. Ces signaux sont représentés par des valeurs numériques réelles ou discrètes. Il ne faut cependant pas en déduire que le modèle de réseaux de neurones est une imitation fidèle du fonctionnement du cerveau humain. La comparaison fonctionnelle, entre réseau de neurones et cerveau humain est de la même nature que celle entre avions et oiseaux. Un avion ne bat pas des ailes mais il vole tout de même. Un parallèle analogue conduit à dire que des fonctions utiles peuvent être réalisées avec des réseaux de neurones même si ceux-ci ne sont pas des modèles du cerveau aussi complexes que le cerveau lui-même (Davallo, 1991).

IV.2. Principe Général

La nature a su armer les animaux d'un système nerveux doté de capacité de raisonnement, certes plus ou moins complexe, mais toujours redoutablement efficace. Le système nerveux permet aux animaux (ainsi qu'à l'homme bien sur) d'appréhender un environnement très changeant et de réagir en temps réel à des situations inédites, tout en se référant aux expériences passées (Dowla, 1995).

Tous ces mécanismes de raisonnements ont besoin de très grandes capacités de « calculs », nécessitant la mise en place de systèmes robustes pour leur prise en charge. La nature a choisie la simplicité pour le constituant de base du système nerveux afin d'assurer cette robustesse (Dowla, 1995).

C'est ainsi que le système nerveux est constitué d'une multitude de petites cellules appelées neurones, possédant chacune une petite capacité de calcul limitée. Ces capacités individuelles, sont assemblées entre elles selon des dispositions variables afin d'obtenir des potentialités de calcul importantes. Toutes ces cellules sont à la fois autonomes et liées entre elles car elles assurent leur propre subsistance du milieu extérieur tout en communiquant, et en s'influençant. Cette propriété d'indépendance permet la sauvegarde de l'ensemble du système en cas de perte d'un nombre limité de neurones, alors que la communication intercellulaire permet au système d'apprendre (Encarta, 2004).

L'homme, en observant la nature, a essayé de dégager ces principes fondamentaux, afin de les appliquer et de créer un comportement intelligent artificiellement. En reprenant ces principes, l'homme a essayé de modéliser ce comportement en créant de petites unités de calcul, limitées et indépendantes, et en les assemblant afin de mettre en place des réseaux de neurones formels.

IV.3. Les fondements biologiques

Les premières études faites par les philosophes grecs pour localiser le siège de l'intelligence dans le corps humain attribuaient la capacité de raisonner au cœur et laissaient au cerveau un rôle très mineur. Cette croyance restera en vigueur jusque dans les écrits de Descartes, qui réaffirme le rôle prépondérant du cœur dans la capacité du raisonnement. Cependant, à partir du dix-huitième siècle, La Mettrie et Cabanis affirment que le cerveau est l'organe de contrôle central du corps humain, reprenant en cela des études oubliées de Platon, Hippocrate et Galien (Encarta 2004).

Les expérimentations sur le cerveau commencent au dix-neuvième siècle, avec les travaux de Broca. Ces expériences ont pour but l'étude neuropsychologique du cerveau, c'est à dire la mise en relation de comportements extérieurs avec les zones du cerveau qui reçoivent ou qui génèrent les stimuli électriques et chimiques associés. Par exemple, le déplacement d'un membre a une relation avec une activité électrique et chimique localisée dans un endroit précis du cerveau. Les études de Broca mettent fin aux théories globalistes du fonctionnement du cerveau en montrant que les fonctions cérébrales sont localisées dans des zones cellulaires clairement délimitées. Bien que réductrice car ne tenant pas compte des interactions entre les différentes zones, cette approche a porté ses fruits et nous en avons hérité les notions de **zones corticales** et de **couches** utilisées dans la théorie des réseaux de neurones. Elle a également permis de mettre en évidence différents niveaux d'abstraction pour la représentation de l'information. Plus un signal traverse de couches corticales, et plus sa représentation en sortie est synthétique et abstraite (Encarta, 2004).

Les études plus récentes, depuis le début du siècle et surtout après la deuxième guerre mondiale, s'intéressent au cerveau d'un point de vue microscopique et cellulaire. Il s'agit cette fois de relier les faits du comportement et les réactions électriques et chimiques qui se produisent dans notre cerveau. Ceci a permis d'isoler le composant cellulaire de base du cerveau, **le neurone**, qui est l'unité de traitement de l'information et qui constitue environ quarante pour cent de la masse totale du cerveau, le reste étant constitué de cellules de support pour protéger, nourrir et soutenir physiquement l'ensemble.

IV.4. Neurone naturel

Les neurones sont les cellules de base du système nerveux, elles ressemblent beaucoup aux autres cellules du corps humain, mais elles ont quelques caractéristiques spécifiques qui sont (Davalò, 1991):

1. Recevoir des signaux des neurones voisins
2. Intégrer ces signaux et les interpréter
3. Engendrer un influx nerveux
4. Le conduire
5. Le transmettre à un autre neurone capable de le recevoir

Schéma descriptif d'un neurone naturel

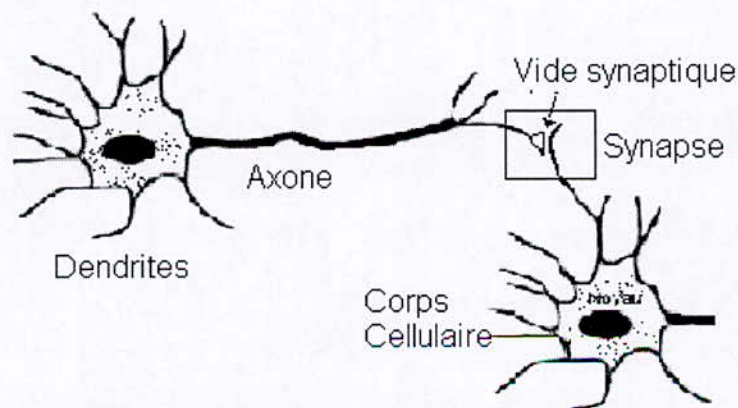


Figure IV.1: Structure du neurone naturel

Mode de fonctionnement (Davalò, 1991) :

- ✓ Le neurone capte les signaux qui lui sont transmis grâce à sa **chevelure de dendrites**.
- ✓ Le signal est traité au niveau du **corps cellulaire**.
- ✓ Le neurone réagit au signal reçu en déclenchant lui-même un **signal qui lui est spécifique** et qu'il transmet grâce à son **axone** aux autres neurones auxquels il est connecté.
- ✓ La connexion se fait par les **synapses** qui peuvent se trouver entre deux axones ou entre un axone et des dendrites.
- ✓ La transmission du signal se fait par la libération de substances chimiques dans **l'espace synaptique**, ces substances sont appelées neurotransmetteurs.

Le corps cellulaire fait la somme de toutes les entrées, si la somme dépasse un certain seuil le neurone réagit, il agit donc comme une fonction seuil. C'est cette particularité qui va permettre la modélisation du neurone naturel en un neurone formel.

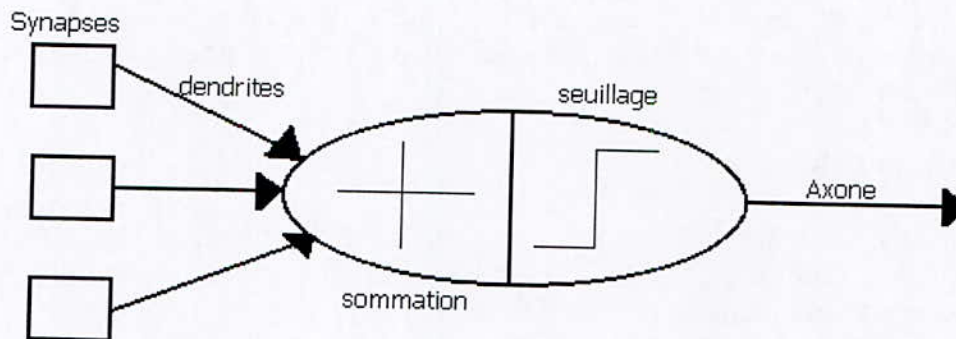


Figure IV.2 : Schématisation du neurone naturel

IV.5. Modélisation du neurone naturel

On peut définir un neurone formel (ou artificiel) par les éléments suivants (Touzè, 1992) :

- La nature des entrées.
- La fonction d'entrée totale qui définit le pré traitement subit par les entrées .
- La fonction de sortie qui calcule la sortie du neurone en fonction de son état d'activation.
- La nature de la sortie du neurone (Figure IV.3)

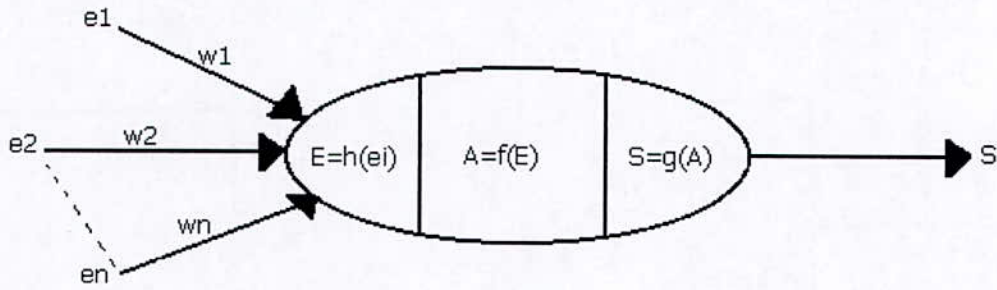


Figure IV.3 : Schéma relationnel

Définition des différentes fonctions

h = Fonction d'entrée totale
 f = Fonction d'activation
 g = Fonction de sortie

Définitions des différents états

\$e_i\$: Différents signaux d'entrées .
 \$W_i\$: Différents poids des connections.

\$E=h(e_i, w_i)\$ (E : entrée totale)
 \$A=f(E)\$ (A : état d'activation du neurone)
 \$S=g(A)\$ (S : sortie du neurone)

Finalement

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

IV.6. Nature des entrées et des sorties :

Les **entrées** peuvent être soit binaires, soit réelles. La fonction h d'entrée totale peut prendre une des formes suivantes :

- ✓ Booléenne
- ✓ Linéaire : $h(e_1...e_n)=\sum_i W_i * e_i$ avec $i=1,n$ (IV.1)
- ✓ Affine
- ✓ Polynomiale de degrés supérieur à 2

La fonction d'état ou fonction d'activation peut prendre les différentes formes suivantes :

CHAPITRE IV

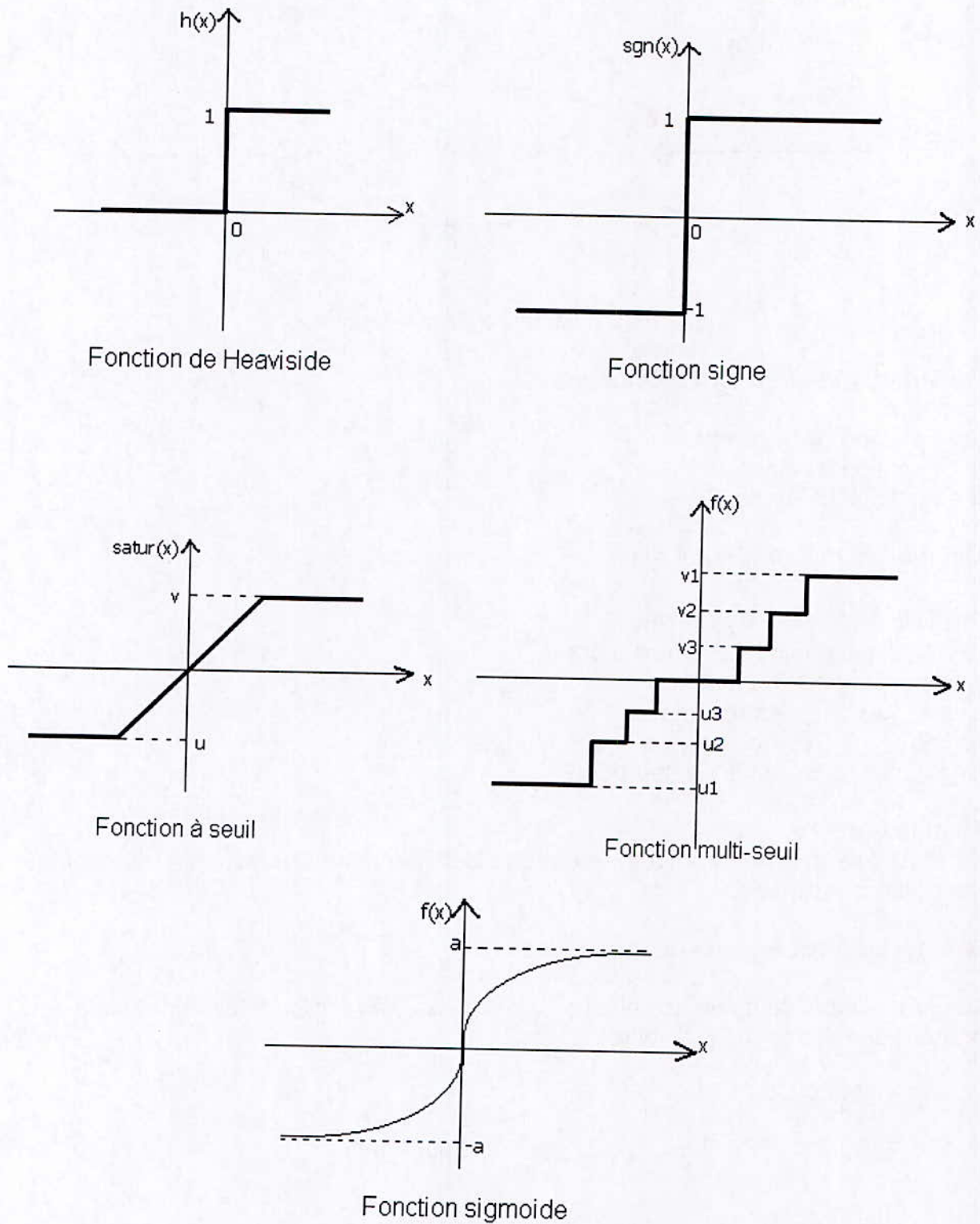


Figure IV.4 : Différentes formes de fonctions de transferts

Remarque

Toute autre fonction choisie croissante et impaire peut jouer le rôle d'une fonction de transfert.

IV.7. Différentes Structures de connexions :

Les réseaux de neurones naturels présentent des connexions (ou liaisons) dans toutes les directions de l'espace, et avec un très grand nombre d'autres neurones.

On peut aussi remarquer un certain nombre de groupements de neurones, ayant entre eux des liens ou connexions privilégiées, formant ainsi des couches.

Cette configuration nous conduit à des réseaux d'une complexité « gigantesque », et qui sont donc très difficile à modéliser, et demandent des capacités de calcul considérables.

Pour simplifier la procédure de modélisation, on peut passer d'un état de connectivité totale (un réseau où tous les neurones sont reliés entre eux) à un état de connectivité partielle ou locale (où un certain nombre de liaisons sont supprimées).

L'utilisation de structures régulières pour la disposition des neurones est couramment adoptée (Patterson, 1996).

IV.7.1. Les réseaux à couches :

Le système naturel peut être simplifié en disposant les neurones de la manière suivante (Figure IV.5) :

- ✓ Les neurones sont groupés en plusieurs couches successives ;
- ✓ Les neurones ne sont pas connectés à l'intérieur d'une même couche ;
- ✓ Chaque neurone est entièrement connecté aux neurones de la couche suivante (Figure IV.6) ;
- ✓ Chaque couche reçoit l'information de la couche précédente, la traite et l'envoie à la couche suivante.

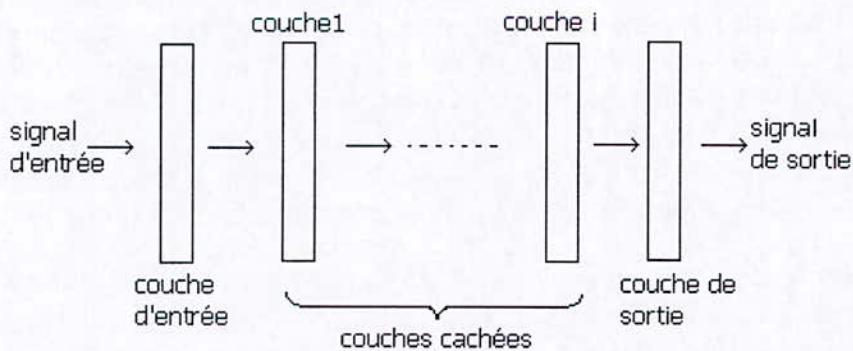


Figure IV.5 : Structure d'un réseau multicouche

Les couches situées entre la couche d'entrée et la couche de sortie sont appelées **couches cachées**. Leur nombre varie d'un modèle de réseau de neurones à un autre. Les réseaux à une seule couche cachée sont appelés réseaux **mono-couche**, alors que les réseaux en possédant plusieurs sont appelés réseaux **multicouches**.

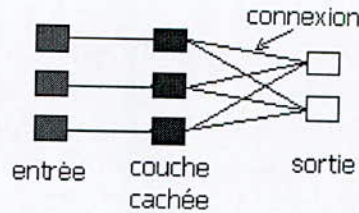


Figure IV.6 : Réseau mono-couche, état de connectivité de chaque neurone

IV.7.2. Réseau entièrement connecté :

C'est la disposition qui se rapproche le plus de la réalité biologique, ou chaque neurone est relié à tous les autres, et possède même un retour sur lui-même.

IV.8. Dynamique des connexions et loi de HEBB

Dans les réseaux de neurones formels, les différents neurones sont reliés entre eux par des connexions. Celles-ci représentent les synapses du modèle naturel et elles peuvent jouer différents rôles comme l'inhibition ou l'excitation du neurone. Ces connexions peuvent être plus ou moins importantes les unes par rapport aux autres, et leur importance peut varier dans le temps. Pour modéliser cette variabilité, on affecte à chaque connexion un **poids** qui pondère le signal transmis par cette connexion. Comme son titre l'indique, cette partie traite de l'évolution dans le temps du poids des connexions.

L'évolution des connexions ne se fait pas de manière aléatoire mais est régie par des mécanismes et des lois qui lui sont propres. Hebb fut le premier à proposer dès 1949 un mécanisme d'évolution pour les synapses dans les réseaux de neurones naturels. Ce mécanisme permet d'expliquer l'importance prise par une connexion par rapport aux autres connexions que possède la cellule. Ce changement d'état est dû à l'insistance que met une cellule A, à exciter de manière répétée et prolongée dans le temps, une autre cellule B. Il se crée alors une « affinité » entre les deux cellules qui se manifeste par l'accroissement de la connexion les reliant. L'interprétation de la loi de Hebb pour les réseaux de neurones formels pourrait s'exprimer comme suit :

On considère que si deux neurones formels interconnectés sont activés en même temps, la connexion les reliant doit être renforcée pour qu'elle puisse jouer entièrement son rôle. Dans le cas contraire, elle ne doit pas être modifiée. Le renforcement d'une connexion se traduit par l'augmentation de son poids (Parizeau, 2004)

Formulation mathématique

$$W_{ij}(t+\Delta t) = W_{ij}(t) + \mu \times A_i \times A_j \quad (IV.2)$$

- W_{ij} : Poids de la connexion reliant le neurone i au neurone j ;
- t : Temps ;
- A_i et A_j : Activation des neurones i et j ;
- μ : Paramètre de l'intensité de l'apprentissage.

IV.9. Les perceptrons

Les perceptrons sont les premières approches faites dans le domaine des réseaux de neurones artificiels, Ils tentent d'imiter de manière directe les systèmes neuronaux naturels et en tout premier lieu le cerveau humain.

IV.9.1. Le perceptron mono-couche

A partir des découvertes de Mc Culloch et Pitts qui aboutirent à la découverte du neurone formel en 1943, et des recherches de Hebb qui donnèrent une explication synaptique du mécanisme d'apprentissage en 1949, Rosenblatt a pu établir le premier modèle performant de réseau de neurones formels vers la fin des années 1950.

IV.9.1.1. Organisation générale :

Le but du perceptron monocouche est d'associer des configurations ou des formes en entrée à des réponses. C'est un type de réseau utilisé surtout pour le classement. Il se compose de deux couches de neurones formels appelées « rétine » et « couche de sortie ». La couche rétine est une contraction de la couche d'entrée et de la couche de traitement des informations. Cette contraction a été effectuée pour des raisons de simplification, car la couche d'entrée est une couche inerte qui ne participe pas directement au traitement de l'information. Les cellules de la rétine, lorsqu'elles sont excitées, répondent par « oui » ou par « non » (1 ou 0) à leur sortie, car elles sont régies par une fonction binaire à seuil (Davalò, 1992).

La rétine est reliée à la couche de sortie grâce à une série de connexions ou de synapses de poids variables. (Figure IV.7)

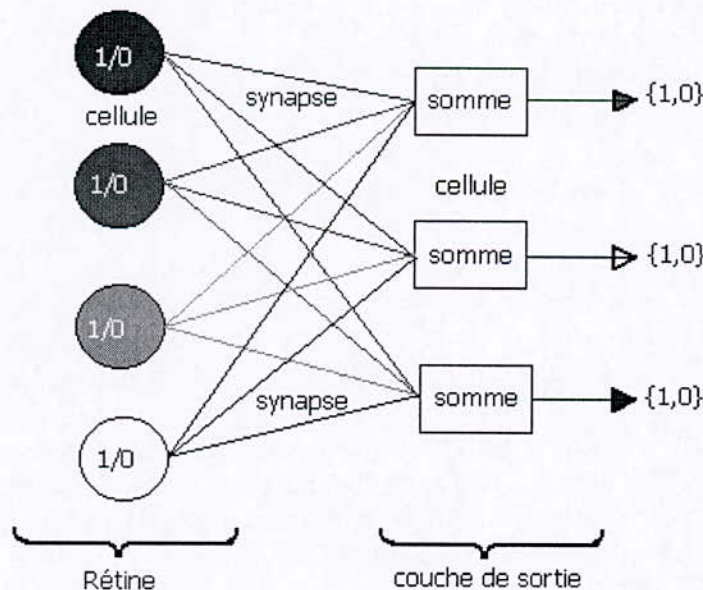


Figure IV.7 : Schéma perceptron mono-couche

De manière générale, l'apprentissage du perceptron mono-couche se fait en modifiant l'intensité de ces synapses par la variation de leurs poids. Le signal émis par les neurones formels de la rétine et transmis par l'ensemble des synapses, est

CHAPITRE IV

évalué par les cellules de sortie. Ce mécanisme peut être résumé par l'équation IV.3 et est représenté sur la figure IV.8 :

$$A_j = \sum_i (X_i \times W_{ij}) \quad i=1,n \quad (IV.3)$$

Avec :

A_j : Activation (ou excitation) de la j^{eme} cellule de sortie ;

X_i : Valeur (0 ou 1) de la i^{eme} cellule de la rétine ;

W_{ij} : poids de la connexion allant de la i^{eme} cellule de la rétine à la j^{eme} cellule de sortie.

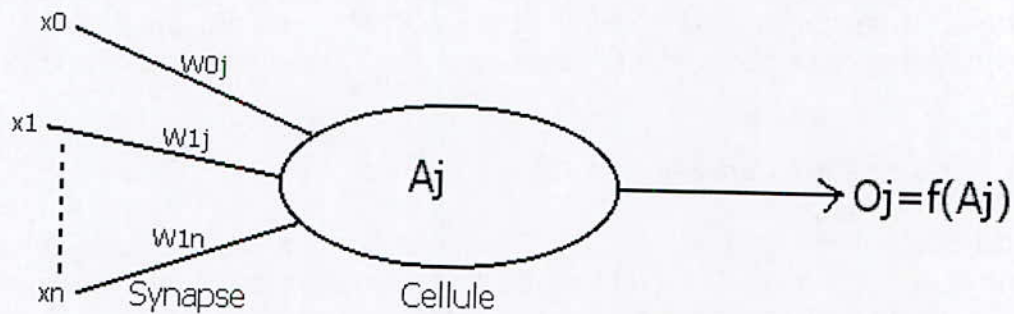


Figure IV.8: Schématisation du neurone artificiel

Où :

O_j : Sortie calculée par le neurone.

IV.9.1.2. Propriétés générales

- ✓ Précis ;
- ✓ Assez Complexe pour donner des résultats intéressants ;
- ✓ Assez simple pour être prédictible ;
- ✓ En accord avec les faits biologiques (Touzet, 1992).

IV.9.1.3. Apprentissage du perceptron mono-couche

Le perceptron mono-couche présente une seule couche de connexions modifiables, c'est la couche reliant les cellules d'association ou rétine, dans laquelle se fait le traitement de l'information, à la couche de sortie qui donne les résultats obtenus. L'apprentissage d'un réseau de neurones formels se fait en présentant une série de couples entrée-sortie réels et connus, codés sous forme de vecteurs. Il est donc indispensable de disposer d'une base de données fiable et assez vaste et que le codage des informations soit fait correctement, pour avoir les résultats les plus significatifs possibles. Les poids des connexions sont modifiés selon l'erreur obtenue, par la règle de Windrow-Hoff (1960) (Davaló, 1991).

L'apprentissage se fait en trois étapes :

1. Présentation du vecteur entrée E
2. Calcul de l'erreur Δ en sortie par la formule suivante :

$$\Delta = T - \sum_j W_j \times E_j \quad (IV.4)$$

Où

Δ : Erreur ;

$\sum_j W_j \times E_j$: Entrée pondérée du neurone ;
 T : Vecteur de la sortie désirée.

3. Modification des poids des connexions :

$$\Delta W_{jj} = K \times \Delta \times E_j \quad (IV.5)$$

Où

ΔW_{jj} : Modification à apporter au poids;

Δ : Erreur calculée en sortie;

E_j : Activation du neurone j;

K : Constante numérique aléatoire.

IV.9.1.4. Limites du perceptron mono couche

L'algorithme d'apprentissage du perceptron ne converge que si le problème est séparable linéairement. C'est à dire que le neurone à seuil doit être capable de séparer les vecteurs présentés en entrée en deux classes de la manière suivante (Davallo, 1991) :

W_0 seuil du neurone

$$\text{si } \sum_j W_j \times E_j > W_0 \rightarrow \text{réponse} = 1 \quad (IV.6)$$

$$\text{si } \sum_j W_j \times E_j \leq W_0 \rightarrow \text{réponse} = 0 \quad (IV.7)$$

La schématisation graphique de cette séparation nous donnerait deux classes séparée par un hyperplan :

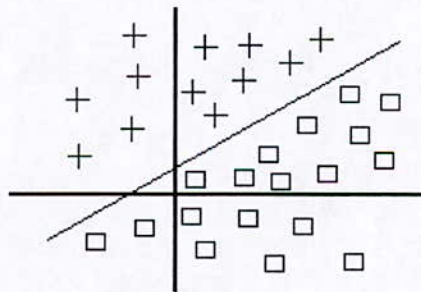


Figure IV.9 : Schématisation du principe de séparation linéaire en deux dimensions

Pour cette raison, le perceptron mono-couche n'est pas capable de traiter les problèmes de type XOR (ou exclusif).

Tableau IV.1 : Table de la loi XOR :

X	Y	XOR(X,Y)
1	1	0
1	0	1
0	1	1
0	0	0

Un réseau de neurones à seuil égale à 1 devrait répondre par 0 à l'entrée {0,0} (et c'est ce qu'il fait) et par 1 à l'entrée {1,1}. Mais cela ne correspond pas à la réalité

La séparation linéaire obtenue par le réseau serait :
 Les entrées {1,0} et {0,1} dans la classe 1 → ayant une réponse 1;
 Les entrées {0,0} et {1,1} dans la classe 2 → ayant une réponse 0.

La représentation graphique de cette séparation n'est pas possible (Figure IV.11). Il n'est pas possible de séparer les points blancs des points noirs par une droite.

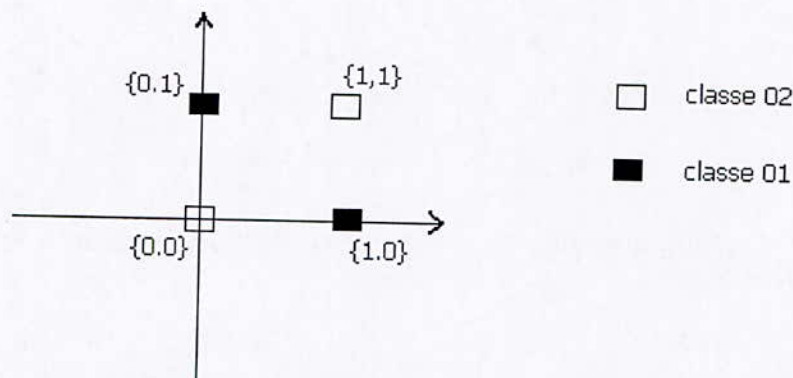


Figure IV.11 : Schématisation de l'impossibilité de la séparation linéaire de la loi XOR

Cette découverte faite dans les années 1960 a donné un sérieux coup de frein aux développements des réseaux de neurones et a poussé un grand nombre de chercheurs à abandonner cette voie de recherche dans le domaine de l'intelligence artificielle, pour se diriger vers les « systèmes experts » entre autres.

Ce n'est qu'avec le développement des algorithmes d'apprentissage avec descente du gradient dans les années 1980 que les réseaux de neurones ont de nouveau intéressé les chercheurs.

IV.9.2. Le perceptron multicouches

Il est formé à partir même neurone formel, mais est organisé autrement. En effet, les cellules sont disposées cette fois dans plusieurs couches intermédiaires entre la couche d'entrée et la couche de sortie. Chaque neurone est connecté à l'ensemble des neurones de la couche suivante et précédente par des connexions de poids variables (Figure IV.11) (Touzet, 1992).

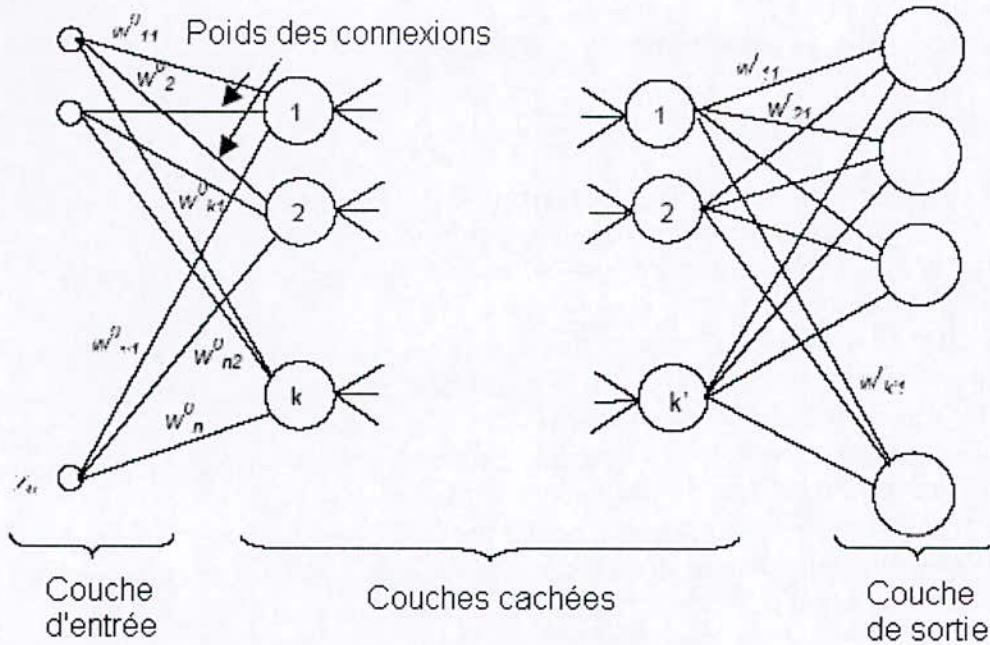


Figure IV.12 : Perceptron multicouche

La *différence majeure* avec les réseaux mono-couches, réside dans *les méthodes d'apprentissage* qui doivent répondre à une plus grande complexité. En effet, il s'agit maintenant de gérer un plus grand nombre de connexions se trouvant dans plusieurs couches de connexions toutes modifiables. L'algorithme d'apprentissage pour ce genre de réseaux est l'algorithme de rétro-propagation du gradient ou de back-propagation, son principe de fonctionnement est revu avec plus de détails au paragraphe (§V.3) .

IV.9.2.1. Algorithme de back-propagation

Cet algorithme est une généralisation aux réseaux de neurones formels multicouches du modèle de Windrow-Hoff. Il a été mis au point par deux équipes simultanément : une équipe française (Fogelman-soulié, Gallinari, Lecun) et une équipe américaine (Rumelhart, Hinton, Williams). Son idée consiste en l'utilisation d'une fonction sigmoïde dérivable en remplacement de la fonction à seuil couramment usitée. La fonction sigmoïde prise en général est (Davallo, 1991) :

$$f(x) = \frac{1}{1 + e^{-x}} \quad (IV.8)$$

L'algorithme permet de *Réinjecter* l'erreur, calculée en sortie, vers les couches internes, en suivant le chemin *inverse* du signal provenant de la couche d'entrée et allant vers la couche de sortie.

Cet Algorithme permet de dépasser le problème du « Credit assignment problem ».

- *Qu'est-ce que le « Crédit assignment problem »*

Dans le cas des réseaux de neurones multicouches, le plus grand problème qui se pose durant l'apprentissage est de comment répercuter sur chacune des connexions le signal d'erreur qui n'a été mesuré que sur la seule couche de sortie .

IV.9.2.2. Apprentissage

Se fait en plusieurs étapes (Debbbar, 1993) :

1. On dispose d'un ensemble d'exemples de couples {entrée – sortie désirée} que l'on présente un par un au réseau. Le calcul se fait de proche en proche, et à chaque présentation, une sortie réelle est calculée : c'est la phase « propagation avant ».
2. On calcule l'erreur qui est égale à la somme des erreurs quadratiques sur chaque cellule de sortie. Cette erreur est représentée par la différence de la sortie réelle avec la sortie désirée.
3. Cette erreur est rétro-propagée ou réinjecter dans le réseau donnant lieu à la modification des poids de chaque connexion.
4. Le processus est répété, en présentant successivement chaque exemple, jusqu'à ce que le réseau converge, c'est à dire que l'erreur devienne inférieure à un certain seuil que l'on s'est fixé pour chaque exemple.

IV.9.2.3 Avantages

- **Les représentations internes**

Les couches cachées du réseau de neurones permettent d'avoir un grand nombre de variables intermédiaires pour résoudre les problèmes de classification ou d'association; le réseau n'est plus obligé de trouver un lien direct entre la sortie et l'entrée. Les unités cachées peuvent représenter intérieurement le problème posé, alors que ces représentations n'étaient pas explicites, lors de la formation du vecteur entrée. Ces représentations internes s'expriment par la spécification de certains neurones à certaines propriétés. L'étude de ces représentations internes peut donner lieu à des solutions astucieuses grâce à l'apparition de liens privilégiés qui n'étaient pas explicites au départ (Debbbar, 1993).

- **La généralisation**

La capacité de généralisation d'un réseau de neurones est son aptitude à donner une réponse précise et satisfaisante à une entrée qui ne fait pas partie de l'ensemble des entrées avec lesquelles a été effectué l'apprentissage. Les réponses doivent être proches si les entrées le sont ; en outre , le réseau agit par construction ;Il n'a aucune idée préconçue sur une nouvelle information, mais il peut la relier à des informations qu'il connaît déjà et modifiera son appréciation en fonction des nouvelles informations qu'il va recevoir. De manière générale, la capacité d'approximation dépend des frontières de classification établies par le réseau. Ces dernières sont elles mêmes dépendantes des exemples d'apprentissages qui ne doivent pas être trop proches ou n'être qu'une série d'un seul cas particulier (Debbbar, 1993).

IV.9.2.4. Limites de l'algorithme

Cet algorithme donne de bons résultats, mais son efficacité n'est pas systématique. Cela est dû au fait que son formalisme mathématique n'a pas été prouvé. En plus cet algorithme ne permet pas de préciser la pédagogie de l'apprentissage et se révèle être un gros consommateur de temps de calculs (Davalò, 1991).

IV.10. Historique :

Un rapide historique est présenté ci-dessous, dans le but de connaître l'évolution de cette technique dans le temps et de comprendre ainsi son cheminement, du niveau de science pluridisciplinaire au stade de technique de pointe (Yeddou, 1998).

IV.10.1. Les prémices

Dans la période de 1940 à 1956, les sciences et techniques de la cognition s'élaborent peu à peu et de l'ancienne cybernétique naissent les sciences cognitives. La volonté de fonder une science "naturelle" de l'esprit émerge quasiment au même moment à cette époque, en Europe et aux États-Unis. Par les efforts conjugués de chercheurs d'origines très diverses (Neurosciences, Intelligence artificielle, Linguistique, Psychologie et Epistémologie principalement), un nouveau champ d'investigations pluridisciplinaires, que sont les sciences cognitives, vit donc le jour. L'apparition des réseaux de neurones artificiels s'inscrit complètement dans cette genèse des sciences cognitives, et leur origine tient dans la volonté de modéliser, de façon mathématique, les neurones biologiques.

IV.10.2. Les pionniers

Les premiers à proposer un modèle, sont deux bio-physiciens de Chicago, **McCulloch et Pitts**, qui en **1943** inventent le premier neurone formel qui portera leurs noms (neurone de McCulloch-Pitts ou automate à seuil). Quelques années plus tard, en **1949**, c'est **Hebb** qui proposera une formulation du mécanisme d'apprentissage, sous la forme d'une règle de modification des connexions synaptiques (règle de Hebb). Cette règle, basée sur des données biologiques, modélise le fait que si des neurones, de part et d'autre d'une synapse, sont activés de façon synchrone et répétée, la force de la connexion synaptique va aller croissant.

IV.10.3. Le premiers réseaux de neurones artificiels

C'est en **1958** qu'apparaît, à proprement dit, le premier réseau de neurones artificiels, grâce aux travaux de **Rosenblatt** qui conçoit le fameux **Perceptron**. Le Perceptron est inspiré du système visuel et possède une couche de neurones d'entrée ("perceptive") ainsi qu'une couche de neurones de sortie ("décisionnelle"). Ce réseau parvient à apprendre à identifier des formes simples et à calculer certaines fonctions logiques. Il constitue donc le premier

système artificiel présentant une faculté jusque là réservée aux êtres vivants : la capacité d'apprendre par l'expérience. Malgré tout l'enthousiasme que soulève le travail de Rosenblatt dans le début des années 1960, la fin de cette décennie sera marquée en **1969**, par une critique violente du Perceptron par **Minsky et Papert**. Ils montrent dans un livre (« *Perceptrons* ») toutes les limites de ce modèle et soulèvent particulièrement l'incapacité du Perceptron, à résoudre les problèmes non linéairement séparables, tels que le problème du XOR. Il s'en suivra alors, face à la déception, une période noire d'une quinzaine d'années dans le domaine des réseaux de neurones artificiels.

IV.10.4. Les RNA multicouches

Il faudra attendre le début des années 80 et le génie de **Hopfield** pour que l'intérêt pour ce domaine soit de nouveau présent. Hopfield démontre, en **1982**, tout l'intérêt d'utiliser des réseaux récurrents (dit "feed-back") pour la compréhension et la modélisation des processus mnésiques (ou faisant appel aux propriétés de la mémoire). Les réseaux récurrents constituent alors la deuxième grande classe de réseaux de neurones, avec les réseaux type perceptron (dit "feed-forward").

En parallèle des travaux de Hopfield, **Werbos** conçoit son algorithme de rétropropagation de l'erreur, qui offre un mécanisme d'apprentissage, pour les réseaux multi-couches de type Perceptron (appelés MLP pour Multi-layer Perceptron), fournissant ainsi un moyen simple d'entraîner les neurones des couches cachées. Cet algorithme de "back-propagation" ne sera pourtant popularisé qu'en **1986**, via un article publié dans *Nature* et un livre ("*Parallel Distributed Processing*") de **Rumelhart et al.**

Depuis la fin des années 80, l'intérêt pour les réseaux de neurones artificiels ne fait que croître et ceci dans de plus en plus de domaines...

IV.11. Les différents types de réseaux de neurones artificiels

On peut grossièrement classer les RNA en deux grandes catégories que sont, les réseaux dit "feed-forward" (type Perceptron) ou "feedback" (encore appelés réseaux récurrents). Comme leurs noms l'indiquent, dans les réseaux "feed-forward", l'information se propage de couche en couche sans retour en arrière possible, contrairement aux réseaux "feedback" (Patterson, 1996).

IV.11.1. Les réseaux "feed-forward"

Ce sont les réseaux dans lesquels l'information circule à sens unique, Ils peuvent se répartir comme suit :

IV.11.1.1. Le Perceptron mono-couche

C'est historiquement le premier RNA, c'est le Perceptron de Rosenblatt. C'est un réseau simple, puisque il ne se compose que d'une couche d'entrée et d'une couche de sortie. Il est calqué, à la base, sur le système visuel et de ce fait a été conçu dans un but premier de reconnaissance des formes. Cependant, il peut aussi être utilisé pour faire de la classification et pour

résoudre des opérations logiques simples (telle "ET" ou "OU"). Sa principale limite est qu'il ne peut résoudre que des problèmes linéairement séparables. Il suit généralement un apprentissage supervisé selon la loi de Hebb.

IV.11.1.2. Le Perceptron multi-couches (PMC)

C'est une extension du précédent, avec une ou plusieurs couches cachées entre l'entrée et la sortie. Chaque neurone dans une couche est connecté à tous les neurones de la couche précédente et de la couche suivante (excepté pour les couches d'entrée et de sortie), et il n'y a pas de connexions entre les cellules d'une même couche. Les fonctions d'activation utilisées dans ce type de réseaux sont principalement les fonctions à seuil ou sigmoïdes. Il peut résoudre des problèmes non linéairement séparables et des problèmes logiques plus compliqués, et notamment le fameux problème du XOR. Il suit aussi un apprentissage supervisé selon la règle de correction de l'erreur.

IV.11.2.3. Les réseaux à fonction radiale

Ce sont les réseaux que l'on nomme aussi RBF (pour "Radial Basic Functions"). L'architecture est la même que pour les PMC cependant, les fonctions de base utilisées ici sont des fonctions Gaussiennes. Les RBF seront donc employés dans les mêmes types de problèmes que les PMC à savoir, en classification et en approximation de fonctions, particulièrement. L'apprentissage le plus utilisé pour les RBF est le mode hybride et les règles sont : soit, la règle de correction de l'erreur, soit la règle d'apprentissage par compétition.

IV.11.2. Les réseaux "feedback" :

Ce sont les réseaux dans lesquelles l'information n'a pas de sens particulier de circulation, fixé à l'avance et peut donc faire des allers-retours. Ils peuvent se répartir comme suit :

IV.11.2.1. Les cartes auto-organisatrices de Kohonen :

Ce type de RNA vient de l'analogie faite avec le cerveau ,qui démontre qu' il existe des zones dans le cerveau qui ont la même topologie que les capteurs auxquels elles sont reliées , par exemple deux zones proches dans le cortex visuel correspondront à deux zones proches dans la rétine.

Ce type de réseau, aussi appelé SOM (pour : "Self Organised Maps"), est un réseau à apprentissage non supervisé qui établit une carte discrète, ordonnée topologiquement, en fonction de patterns(ou règles) d'entrée. Le réseau forme ainsi une sorte de treillis où, chaque noeud du treillis est un neurone associé à un vecteur de poids. La correspondance entre chaque vecteur de poids est calculée pour chaque entrée. Par la suite, le vecteur de poids ayant la meilleure corrélation, ainsi que certains de ses voisins, vont être modifiés afin

d'augmenter cette corrélation. Le schéma de l'apprentissage peut être résumé par les étapes suivantes :

- Présentation du vecteur de données en entrée ;
- Compétition entre les neurones (calcul des distances entre tous les neurones et les vecteurs entrée) ;
- Désignation du neurone vainqueur (ayant la meilleure des corrélations) ;
- Evolution des poids synaptiques pour le neurone vainqueur ainsi que pour certains de ses voisins

Les LVQ (pour "Learning Vector Quantization") sont un cas particulier des SOM, où seul le noeud ayant la meilleure corrélation est adapté. On parle, pour ce type de réseau où seul le "vainqueur" est sélectionné, de réseaux à compétition.

IV.11.2.2. les réseaux de Hopfield

Les réseaux de Hopfield sont des réseaux récurrents et entièrement connectés. Dans ce type de réseau, chaque neurone est connecté à chaque autre neurone et il n'y a aucune différenciation entre les neurones d'entrée et de sortie. Ils fonctionnent comme une mémoire associative non-linéaire et sont capables de trouver un objet stocké en fonction de représentations partielles ou bruitées. L'application principale des réseaux de Hopfield est l'entrepôt de connaissances, mais aussi la résolution de problèmes d'optimisation. Le mode d'apprentissage utilisé ici est le mode non-supervisé .

IV.11.2.3. Les ART

Les réseaux ART (pour "Adaptative Resonance Theorie") sont des réseaux à apprentissage par compétition. Le problème majeur qui se pose dans ce type de réseaux est le dilemme « stabilité/plasticité ». En effet, dans un apprentissage par compétition, rien ne garantit que les catégories formées vont rester stables. La seule possibilité, pour assurer la stabilité, serait que le coefficient d'apprentissage tende vers zéro, mais le réseau perdrait alors sa plasticité. Les ART ont été connus spécifiquement pour contourner ce problème. Dans ce genre de réseau, les vecteurs de poids ne seront adaptés que si l'entrée fournie est suffisamment proche, d'un prototype déjà connu par le réseau. On parlera alors de résonance. A l'inverse, si l'entrée s'éloigne trop des prototypes existants, une nouvelle catégorie va alors se créer, avec pour prototype, l'entrée qui a engendrée sa création. Il est à noter qu'il existe deux principaux types de réseaux ART : les ART-1 pour des entrées binaires et les ART-2 pour des entrées continues. Le mode d'apprentissage des ART peut être supervisé ou non.

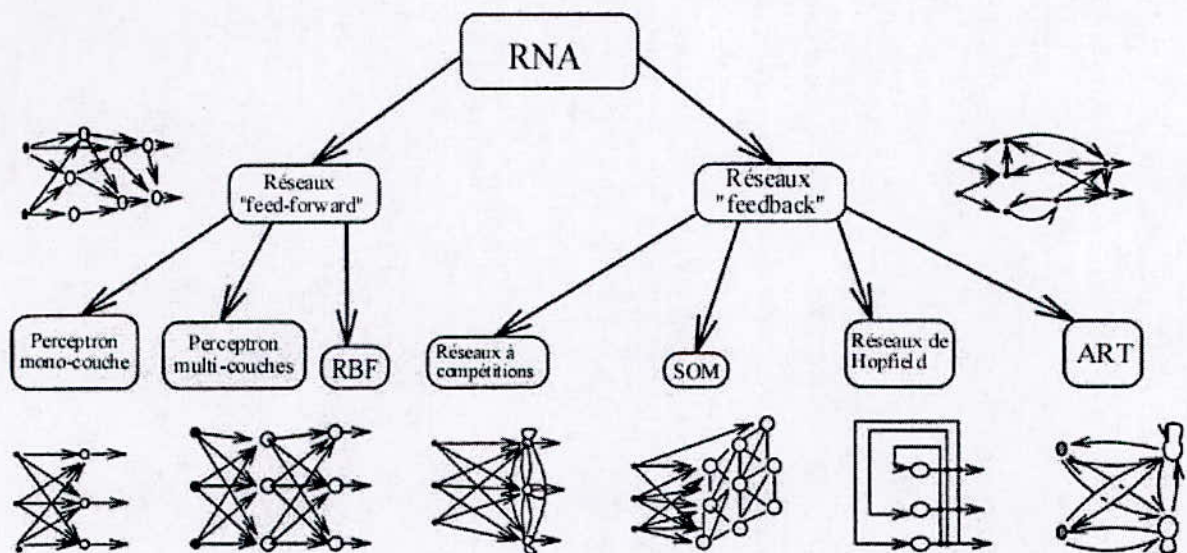


Figure IV.14 : Différents types de RNA

CHAPITRE V

CHAPITRE V

Modélisation de
l'évapotranspiration
potentielle.

CHAPITRE V

« Modélisation de l'évapotranspiration potentielle »

V.1. Définition de la modélisation

Initialement le modèle fut "la référence" à imiter ou à reproduire, l'exemple ; puis il devint progressivement le résultat de cette imitation : du "modèle du peintre" on passe au modèle représenté par l'image, ou à la représentation que le peintre a établie. La recherche scientifique s'est progressivement approprié le concept de modèle, au cours du XXe siècle, par l'intermédiaire sans doute de son usage dans les "arts et métiers" (modèle réduit, ou maquette), et elle l'utilise couramment aujourd'hui pour désigner les **"représentations" des phénomènes qu'elle cherche à comprendre** ou à **expliquer**. Le modèle est alors un **"système de symboles"**, indifféremment et parfois simultanément : graphiques, discursif, mathématique, iconique, dont l'extrême souplesse potentielle permet de rendre compte de la plupart des perceptions dont on dispose lorsqu'on souhaite décrire un phénomène, afin de **l'interpréter intelligiblement**. La modélisation ouvre des espaces immenses à l'investigation de type scientifique, sans contraindre la représentation par un langage trop "fermé". L'observateur peut aujourd'hui "commencer par modéliser", et ceci de façon intelligible, reproductible et communicable. Les développements des méthodes de programmation informatique ouvre à la méthode des modèles le champ des simulations les plus diverses, permettant de **concevoir** et **d'évaluer** les comportements **spatio-temporels** des **phénomènes modélisés**. Les phénomènes hydrologiques, par leur complexité, sont des cibles de choix pour la modélisation ; l'évapotranspiration n'échappe pas, bien-sûr, à cette règle (Bermad, 2004). C'est dans le but d'améliorer la compréhension de ce phénomène et de simplifier sa méthode de calcul que nous nous proposons de modéliser l'évapotranspiration potentielle.

V.2. Présentation des données :

Les données utilisées, nous ont été fournies à titre gracieux par l'Office National de la Météorologie dans le cadre d'une convention les liant à l'université afin de promouvoir le travail scientifique. Elles ont été recueillies au niveau de la station de Dar El Beida (Wilaya d'Alger), dont voici les caractéristiques :

Tableau V.1 Caractéristiques de la station de Dar El Beida

Station	Numéro	Code	Lat.(Nord)	Long.(E/W)	Alt.(mètre)
D.E.B	6	390	36° 43	03°15E	24

V.2.1 Inventaire des données

- **L'évapotranspiration potentielle :**

Cette valeur est calculée par la formule de Penman classique suivante :

$$E_{tp} = \left[\frac{Iga}{59} \times (1 - \alpha) \times \left(0.18 + 0.62 \times \frac{h}{H} \right) - \frac{\sigma}{59} T^4 \times (0.56 - 0.08\sqrt{e}) \times \left(0.1 + 0.9 \frac{h}{H} \right) \right] \times \frac{\frac{\Delta}{\gamma}}{1 + \frac{\Delta}{\gamma}} + \frac{0.26}{1 + \frac{\Delta}{\gamma}} \times (e_s - e) \times (1 + 0.54V) \quad (V.1)$$

Où :

E_{tp} : Evapotranspiration potentielle en [mm] ;

Iga : Radiation solaire directe en l'absence d'atmosphère exprimée en [cal/cm²/j] ;

α : Albédo de la surface évaporante, prise dans ce cas égale à 0.2 ;

h : Durée réelle d'insolation en heure de la période considérée ;

H : Durée d'insolation maximale possible de la période considérée ;

σ : Constante de STEFANO-BOLTZMAN, et $\sigma = 1.18 \cdot 10^{-7}$ [cal/cm²/j/°K]

T : Température moyenne de la période considérée de l'air sous abri, en°K
 $T = 273.3 + t$ si la température mesurée en °C ;

e : Pression de vapeur moyenne de la période considérée, mesurée sous abri exprimée en [mb] ;

Δ : Pente de la courbe de tension de vapeur saturante en fonction de la température ;

γ : Constante psychrométrique ;

e_s : Pression de vapeur à saturation pour la température T ;

$$e_s = 0.611 \times \exp\left(\frac{17.27 \times T}{T + 273.3}\right) \quad (V.2)$$

V : Vitesse du vent mesurée à 2m de hauteur de la surface évaporante en [m/s] ;

• Rayonnement global

Le rayonnement global est calculé à partir du taux d'insolation par la formulation générale suivante (Yahiaoui, 1982)

$$Rg = Iga \times \left(a + b \times \frac{h}{H} \right) \quad (V.3)$$

Où a et b sont deux coefficients déterminés par régression linéaire, spécifique à chaque station de mesure et dépendant de la situation géographique de celle-ci. Pour la station de Dar El Beida , $a=0.21$ et $b=0.51$.

Ainsi que les données suivantes : Température moyenne, Vitesse du vent (à 2m de hauteur), Insolation, Humidité relative et tension de vapeur.

V.3. Méthodes d'évaluation de la modélisation

Une simple analyse visuelle des résultats, bien qu'étant un premier élément de comparaison et d'appréciation, ne saurait être suffisante pour évaluer de manière objective les résultats d'un modèle. Il est en effet difficile de voir à l'œil nu, toutes les différences qui peuvent exister entre deux courbes ayant la même allure générale. Il est donc nécessaire d'utiliser des critères mathématiques objectifs pour estimer et quantifier la distance entre les observations et les simulations (Dagnelie, 1992). Il en existe plusieurs, et les critères utilisés dans le cadre de ce travail sont les suivants :

V.3.1 Le critère de Nash-Sutcliffe (E)

Il s'exprime par :

$$E = 1 - \frac{\sum_{i=1}^{n_{obs}} (Y_{i\ mod} - Y_{i\ obs})^2}{\sum_{i=1}^{n_{obs}} (Y_{i\ mod} - \bar{Y}_{obs})^2} \quad (V.4)$$

Où :

$Y_{i\ mod}$: Variable simulée ;

$Y_{i\ obs}$: Variable observée ;

\bar{Y}_{obs} : Moyenne des observations ;

n_{obs} : Nombre d'observations.

Ce critère quantifie la fraction de la variance observée qui est prise en compte par le modèle. Il prend des valeurs allant de $-\infty$ à 1 ; plus l'efficacité est proche de 1, meilleure est l'accord entre la modélisation et l'observation.

Lorsque les valeurs de E sont négatives, cela signifie que les valeurs simulées conduisent à de moins bons résultats que si la valeur moyenne des observations était utilisée comme prédicteur.

Mais le critère de Nash-Sutcliffe a l'inconvénient de donner beaucoup de poids aux valeurs fortes. Dans ce cas là, on peut utiliser la même formulation sur le logarithme de la variable, (si elles ne sont pas nulles), ou sur leur racine carrée afin d'augmenter le poids des valeurs faibles.

V.3.2. L'erreur moyenne quadratique (RMSE) :

Est un critère utilisée pour calculer l'écart entre les valeurs simulées et observées ; il s'exprime par :

$$RMSE = \sqrt{\left(\frac{1}{n_{obs}} \sum_{i=1}^{n_{obs}} (Y_{i\text{mod}} - Y_{i\text{obs}})^2 \right)} \quad (\text{V.5})$$

Il donne une idée sur la dispersion entre les valeurs observées et simulées, plus le critère RMSE sera faible plus l'écart entre les valeurs sera limité.

De même que pour le critère de Nash-Sutcliffe, le critère RMSE tend à donner plus de poids aux valeurs fortes et peut être appliqué sur le logarithme ou la racine carrée des valeurs afin d'améliorer la représentation des valeurs faibles.

Comme dans le cadre de ce travail, il peut y avoir un rapport de 1 à 50 entre les différentes valeurs simulées, c'est la RMSE appliquée aux racines carrées qui sera adoptée, et dont la formulation est la suivante :

$$RMSE = \sqrt{\left(\frac{1}{n_{obs}} \sum_{i=1}^{n_{obs}} (\sqrt{Y_{i\text{mod}}} - \sqrt{Y_{i\text{obs}}})^2 \right)} \quad (\text{V.6})$$

V.3.3. Le coefficient de détermination (R^2)

C'est le carré du coefficient de corrélation, Il s'exprime par :

$$R^2 = \frac{\sum_{i=1}^n ((Y_{i\text{obs}} - \bar{Y}_{\text{obs}}) \times (Y_{i\text{mod}} - \bar{Y}_{\text{mod}}))^2}{\sum_{i=1}^n (Y_{i\text{obs}} - \bar{Y}_{\text{obs}})^2 \times \sum_{i=1}^n (Y_{i\text{mod}} - \bar{Y}_{\text{mod}})^2} \quad (\text{V.7})$$

Il représente le taux de corrélation entre les séries simulées et observées, il faut donc chercher à le maximiser, pour qu'il prenne des valeurs proches de 1 (ou 100%), afin de vérifier l'accord entre les deux séries simulées et observées.

V.4. Description du modèle

La technique de modélisation par les réseaux de neurones adoptée dans le cadre de ce travail, est un réseau de neurone multicouches de type Perceptron, dont les principes de fonctionnement généraux sont explicités au (§ IV.9). Néanmoins, il convient d'aller plus en détails dans la description du fonctionnement de ce type de réseaux de neurones, afin de mieux comprendre son comportement :

V.4.1. Architecture de base

Le perceptron est constitué de plusieurs unités de calcul représentant chacune un neurone artificiel. Tel qu'illustré par la figure V.1, un neurone est essentiellement constitué d'un intégrateur qui effectue la somme pondérée de ses entrées (Parizeau, 2004).

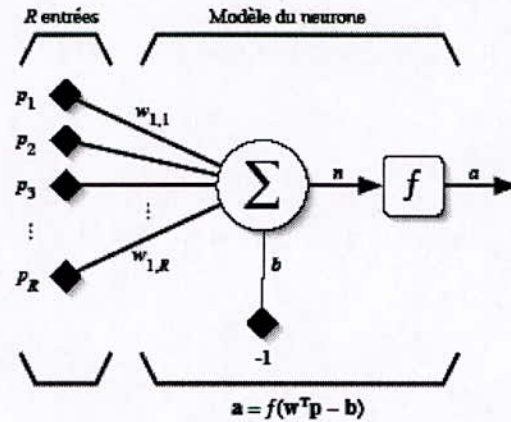


Figure V.1 : Modèle de neurone artificiel

Où :

- p : Vecteur entré de taille R et de composante p_i ;
- w_{ij} : Poids de la synapse ;
- f : Fonction de transfert ;
- b : Biais ;
- n : Sommes des entrées pondérées et biaisées ;
- a : Sortie du neurone.

Ce modèle peut être résumé sous forme de calculs matriciels ; tel que représenté par la figure V.2 :

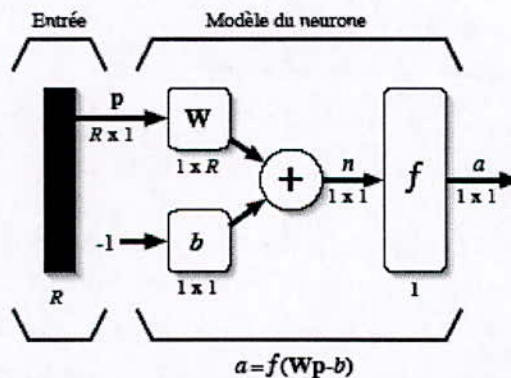


Figure V.2 : Représentation matricielle du modèle d'un neurone artificiel

Où :

- p : Matrice d'entrée ou stimuli de taille $1 \times R$;
- W : Matrice des poids des connexions de taille $R \times 1$;
- n, a et b scalaires de taille 1×1 dans le cas d'un perceptron simple.

Le travail d'un neurone artificiel peut donc se résumer à sa sortie qui elle est le résultat d'une série d'étapes de calculs matriciels, tel qu'exprimé par la formule V.8 :

$$a = f(W \times p - b) \tag{V.8}$$

Cette manière d'appréhender le comportement d'un neurone artificiel en tant que série de calculs matriciels, facilitera le passage aux réseaux multicouches.

V.4.2. Architecture réseau :

Pour constituer un réseau de neurone de type perceptron multicouche, il suffit d'assembler plusieurs unités simples de calcul, et de les disposer en couche comme sur la figure V.3 (Parizeau,2004) :

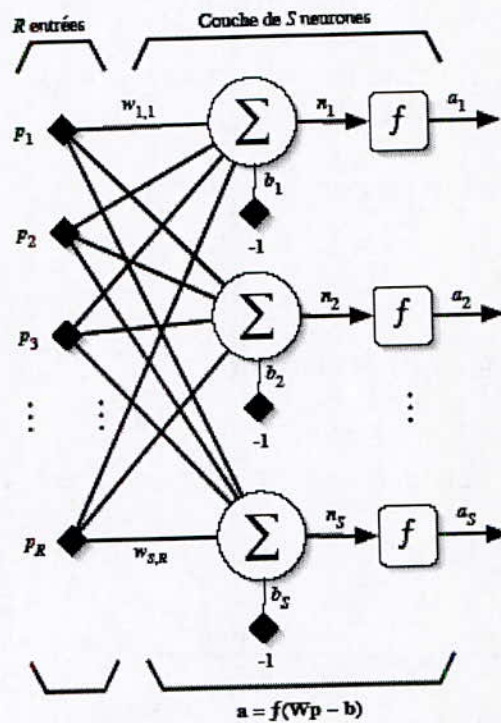


Figure V.3 : Couche de s neurones

Le travail de chaque neurone reste le même et peut se faire indépendamment du reste des neurones de la couche. Chaque élément p_i du vecteur entré p est envoyé à chacun des s neurones n_j , de la couche k , à travers une connexion, qui elle, est représentée par le poids qui lui a été attribué w_{ij} . Chaque neurone doit donc faire la somme des entrées pondérées à son niveau, somme à laquelle il ajoute un biais b_j qui lui est propre, avant de calculer sa sortie, notée a_j , par sa propre fonction de transfert f .

La sortie de chaque neurone j peut être résumée par l'expression (V.12) :

$$a_j = f\left(\sum_{i=1}^r w_{ij} \times p_i + b(j)\right) \tag{V.9}$$

Les sorties de chaque élément a_j de la couche k , contenant s neurones, peuvent être groupées et constituer un vecteur de même taille que le nombre de neurones présents dans cette couche. Ce vecteur est noté a^k et il est de taille s . Il en va de même pour les biais de chaque neurone, ils peuvent être groupés en un vecteur noté b^k de taille s .

Les poids w_{ij} , peuvent quand à eux, constituer une matrice des poids des connexions W de taille $s \times r$, dont les éléments w_{ij} constituent les poids des connexions allant de l'entrée j au neurone i .

On entend par vecteur entré dans ce cas, aussi bien le vecteur stimuli si l'on parle de la première couche du réseau, ou alors le vecteur résultant des calculs de la couche précédente s'il s'agit d'une couche intermédiaire ; la matrice W est spécifique aux connexions présentes entre deux couches.

La figure ci-dessous exprime la représentation matricielle du modèle d'une couche de s neurones :

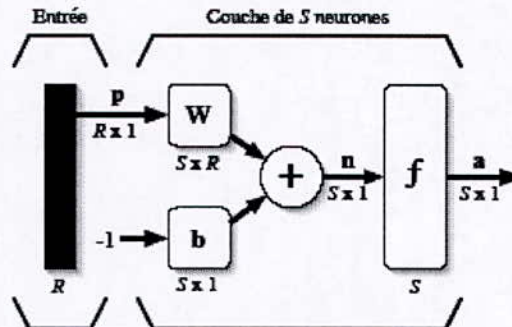


Figure V.4 : Représentation matricielle d'une couche de s neurones.

La sortie de la couche peut alors prendre l'expression suivante s'il s'agit de la première :

$$a^1 = f(W_{11} \times p - b^1) \quad (\text{V.10})$$

Ou alors prendre cette deuxième expression s'il s'agit d'une couche intermédiaire :

$$a^k = f(W_{k-1,k} \times a^{k-1} + b^k) \quad (\text{V.11})$$

Où :

a^{k-1} : Sortie calculée par la couche $k-1$;

Par suite, le signal traverse, de manière successive, les différentes couches du réseau, que l'on dit *cachées*, jusqu'à atteindre la dernière couche ou *couche de sortie*.

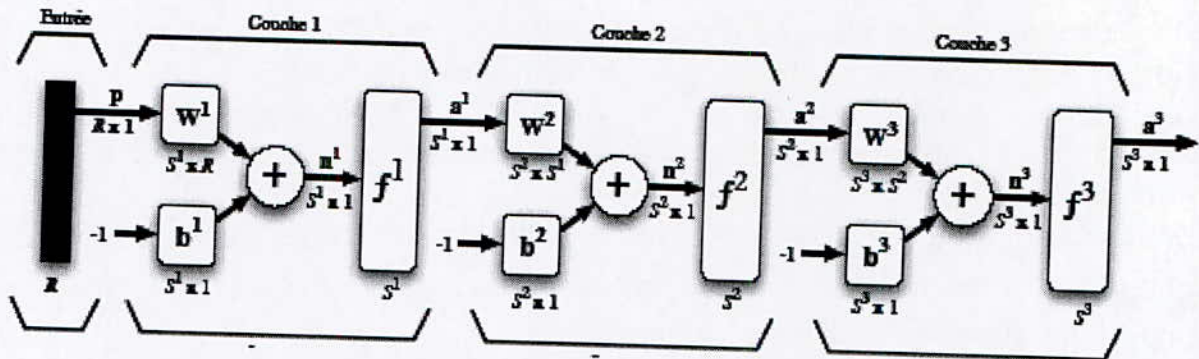


Figure V.5 : Représentation matricielle d'un réseau à trois couches cachées

V.4.3. Apprentissage du modèle

L'apprentissage du perceptron multicouche est un apprentissage supervisé par l'Algorithme de rétro propagation de l'erreur ou de « backpropagation ».

Il consiste en la modification des poids des connexions entre les différentes couches, par rapport à une erreur calculée en sortie de réseau : on dit alors que l'erreur est rétro-propagée (Davalo, 1991).

Cette erreur notée $e(t)$ au temps t est déterminée par la présentation d'une série $\{(p_1, d_1), (p_2, d_2), \dots, (p_Q, d_Q)\}$ de couples d'entrée p et de sortie désirée d .

On confronte alors à chaque fois la sortie effective $a(t)$ à $d(t)$, la sortie que l'on désire obtenir, l'erreur est alors calculée par l'expression qui suit (Parizeau, 2004) :

$$e(t) = d(t) - a(t) \tag{V.12}$$

Où :

$a(t)$: Désigne la sortie calculée par le réseau au temps t ;

Un indice de performance est alors calculé : il s'agit de l'espérance de l'erreur instantanée

$$F(x) = E[e^T(t) \times e(t)] \tag{V.13}$$

Où :

$[e^T(t) \times e(t)]$: Désigne le produit de $e(t)$ et de $e^T(t)$, deux erreurs calculées à un temps T d'intervalle.

L'indice de performance précédemment défini peut être approximé par l'erreur instantanée $\hat{F}(x)$:

$$\hat{F}(x) = e^T(x) \times e(t) \tag{V.14}$$

C'est finalement cette erreur instantanée qui devra être optimisée, et qui permettra de modifier les poids des différentes connexions ainsi que les biais des différents neurones, par les expressions suivantes :

Modification du poids :

$$\Delta W^k(t) = -\eta \times S^k(t) \times (a^{k-1})^T(t) \quad (\text{V.15})$$

$$\Delta b^k(t) = -\eta \times S^k(t) \quad (\text{V.16})$$

Où :

η : Taux d'apprentissage (une constante aléatoire comprise entre 0 et 1, et qui permet, selon les cas, d'accélérer la convergence de l'algorithme d'apprentissage.

S^k : Définit comme *la sensibilité* aux changements dans le niveau d'activation des neurones elle s'exprime par :

$$S^k = \frac{\partial \hat{F}}{\partial n^k} \quad (\text{V.17})$$

Où :

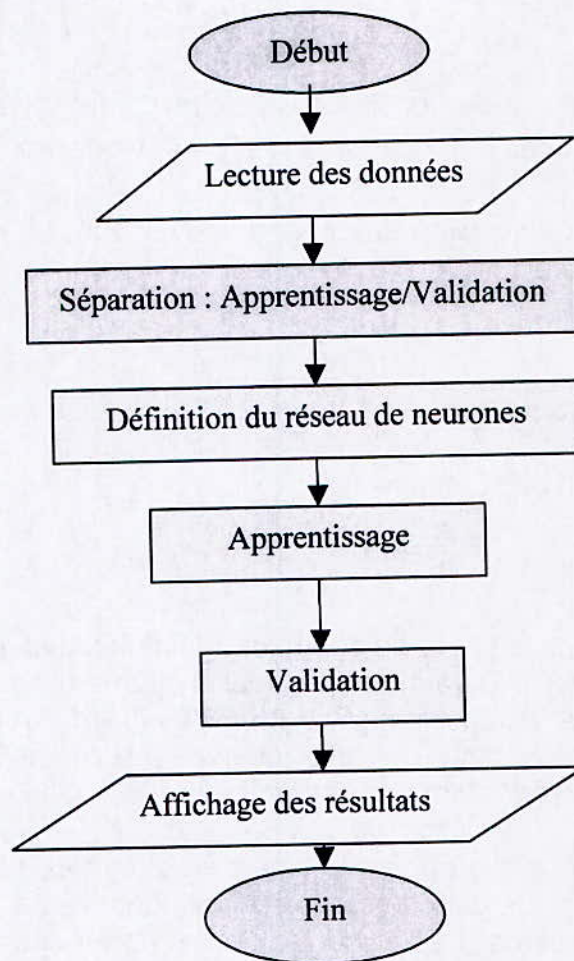
$n^k(t)$: Désigne l'activation de la couche k .

V.5. Programmation du modèle

Le programme élaboré dans le cadre de ce projet de fin d'études, utilise le logiciel de calculs mathématiques MATLAB version 6p5, sous système d'exploitation Windows, mis au point par la firme américaine MATHWORKS LABORATORY. Ce logiciel dispose de sous-programmes, dans sa partie réservée aux réseaux de neurones, qui permettent la création et l'entraînement de perceptrons multicouches de manière personnalisée.

Le perceptron multicouche mis au point à l'aide de ce logiciel, pour les besoins du présent travail de modélisation, est un réseau à deux couches cachées. La première couche contient dix (10) neurones alors que la deuxième en contient treize (13). La fonction de transfert des neurones des 2 couches cachées est la fonction sigmoïde (voir § IV.9.2). La couche de sortie n'est constituée que d'un seul neurone dont la fonction de transfert est de type linéaire. La taille de la couche d'entrée est variable et dépend de la taille du vecteur stimuli. On rappelle que la couche d'entrée ne sert qu'à transmettre l'information, et n'a pas de rôle actif dans calcul.

• Organigramme du programme



CHAPITRE VI

CHAPITRE VI

Résultats de la modélisation

CHAPITRE VI

« Résultats de la modélisation »

Dans le cadre de ce travail, plusieurs modèles ont été élaborés et testés et leurs différents résultats confrontés et comparés. Ils peuvent être séparés principalement par le pas de temps auquel ils s'intéressent comme suit :

VI.1. Modèle au pas de temps décadaire

Ce sont les modèles utilisés afin de quantifier l'évapotranspiration potentielle de Penman au pas de temps décadaire. Eux même peuvent être aussi répartis selon le nombre de paramètres en entrée comme suit :

VI.1.1. Modèles à une seule entrée

Dans ce cas là il s'agit du perceptron multicouches à deux couches décrit aux §IV.9. et §V.3. La démarche entreprise a consisté en la recherche du meilleur modèle à une entrée, en faisant varier 2 paramètres : la longueur de la série d'apprentissage et le type de paramètres en entrée. Les différents essais ainsi que leurs résultats sont résumés et classés sur le tableau VI.1 :

Tableau VI.1 : Résultats du modèle à entrée unique

Paramètres	2 années de calage			5 années de calage			8 années de calage		
	E(%)	R ²	RMSE	E(%)	R ²	RMSE	E(%)	R ²	RMSE
Rg	82,56	0,83	0,69	83,06	0,85	0,62	85,80	0,87	0,60
Tmoy	60,98	0,61	1,24	63,43	0,65	1,15	67,23	0,70	0,98
Ew	55,21	0,57	1,31	57,70	0,60	1,27	59,71	0,64	1,10
Ins	35,01	0,43	1,27	36,25	0,44	1,26	37,91	0,46	1,24
Hr	24,26	0,21	1,51	27,33	0,25	1,48	30,55	0,32	1,38
Vvent	21,58	0,22	1,49	19,42	0,21	1,52	6,53	0,08	1,62

Ces résultats sont reportés sur les graphes suivants :

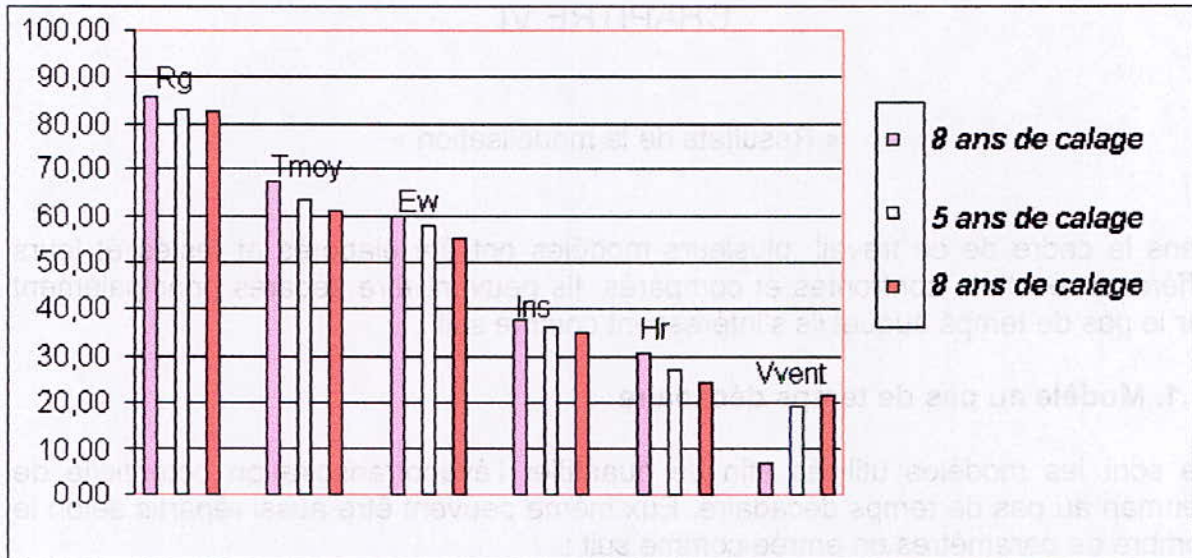


Figure VI.1 : Evolution du critère de Nash par Type de paramètre en entrée et par années de calage.

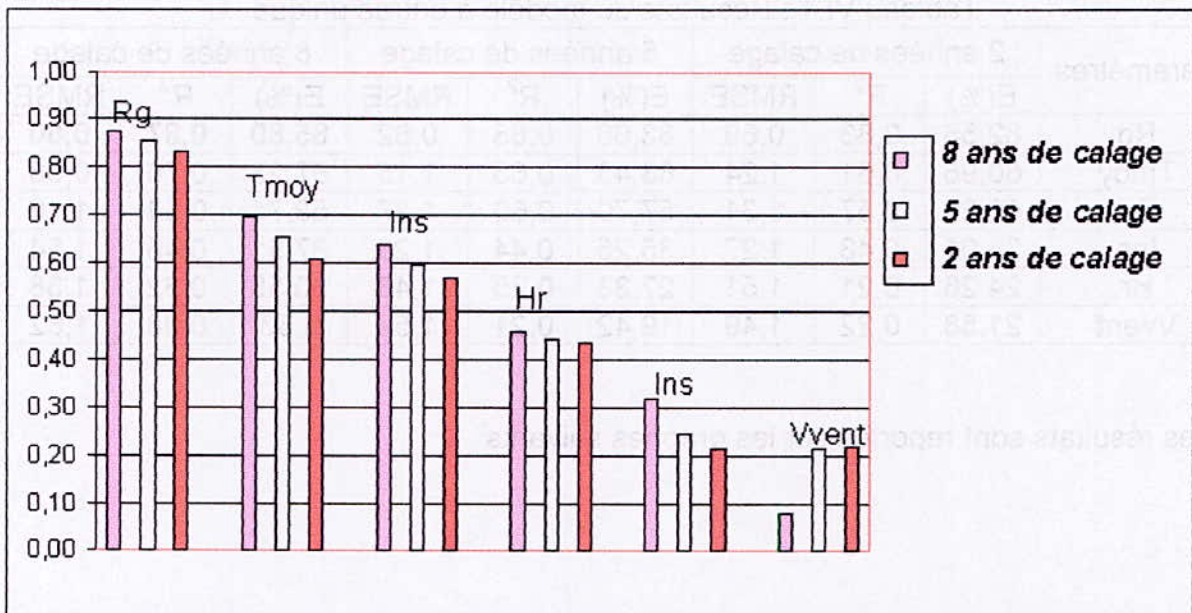


Figure VI.2 : Evolution du coefficient de détermination par type de paramètre en entrée et par années de calage.

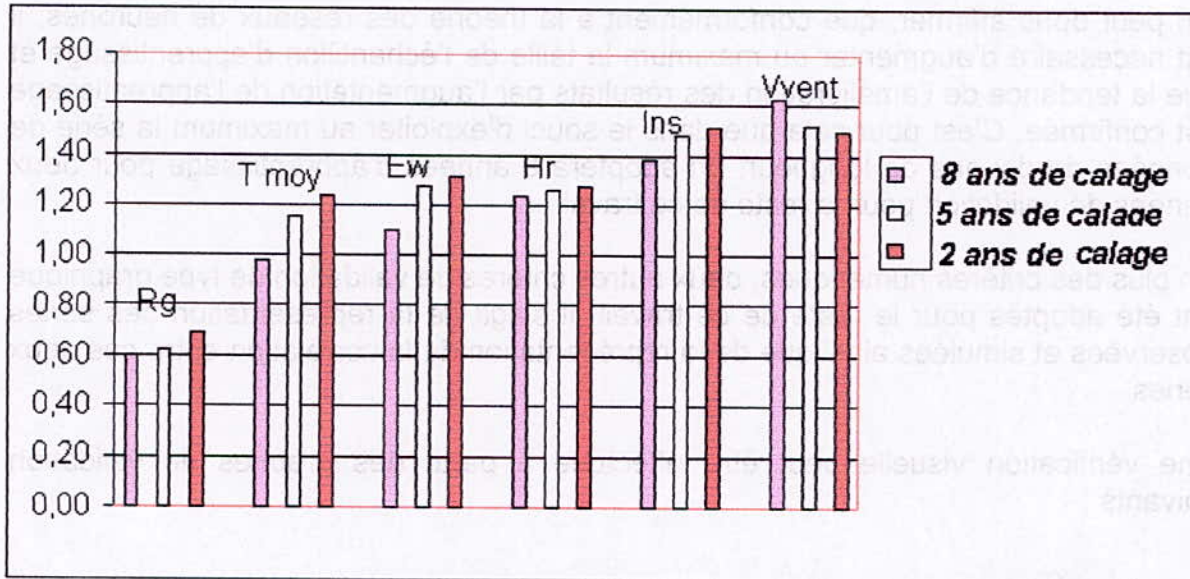


Figure VI.3 : Evolution de la RMSE par type de paramètre en entrée et par années de calage.

Interprétation

Les essais ayant été réalisés en faisant varier deux facteurs, Il convient donc de séparer dans un premier temps l'interprétation des résultats comme suit :

- **Facteur temps de calage**

La tendance générale qui se dégage au vu des figures VI.1,VI.2 et VI.3, est que plus le nombre d'années d'apprentissage augmente plus les résultats des différents critères sont meilleurs : cela cadre avec la théorie des réseaux de neurones. Néanmoins, le gain le plus remarquable est pour la RMSE qui apparaît comme étant la plus sensible à la variation du facteur temps de calage.

- **Facteur type de paramètre en entrée**

Le tableau V.1 et les figures VI.1,VI.2 et VI.3, montrent que les résultats varient de manière significative d'un paramètre à un autre et ce, pour les trois critères retenus à titre de comparaison. Il apparaît aussi que les meilleurs résultats sont obtenus pour le paramètre **Rayonnement global** en entrée, suivi à moindre mesure par la température moyenne et la tension de vapeur.

De manière plus globale, seul le modèle ayant comme entrée la vitesse du vent ne suit pas la tendance générale de l'amélioration des résultats avec l'augmentation des années de calage ; il se trouve aussi que ce modèle obtient les plus mauvais résultats en modélisation et les séries simulées à partir de ce modèle ne sont que très peu corrélées aux séries observées, cela nous permet de déduire que les valeurs obtenues par les trois critères à partir de ce modèle ne sont que des particularités numériques sans signification physique.

On peut donc affirmer, que conformément à la théorie des réseaux de neurones, il est nécessaire d'augmenter au maximum la taille de l'échantillon d'apprentissage et que la tendance de l'amélioration des résultats par l'augmentation de l'apprentissage est confirmée. C'est pour cela que dans le souci d'exploiter au maximum la série de données de dix ans de longueur, on adoptera 8 années d'apprentissage pour deux années de validation pour le reste de ce travail.

En plus des critères numériques, deux autres critères de validation de type graphique ont été adoptés pour le reste de ce travail. Il s'agit de la représentation des séries observées et simulées ainsi que de la représentation de la corrélation entre ces deux séries.

Une vérification visuelle peut être effectuée à partir des graphes de validation suivants :

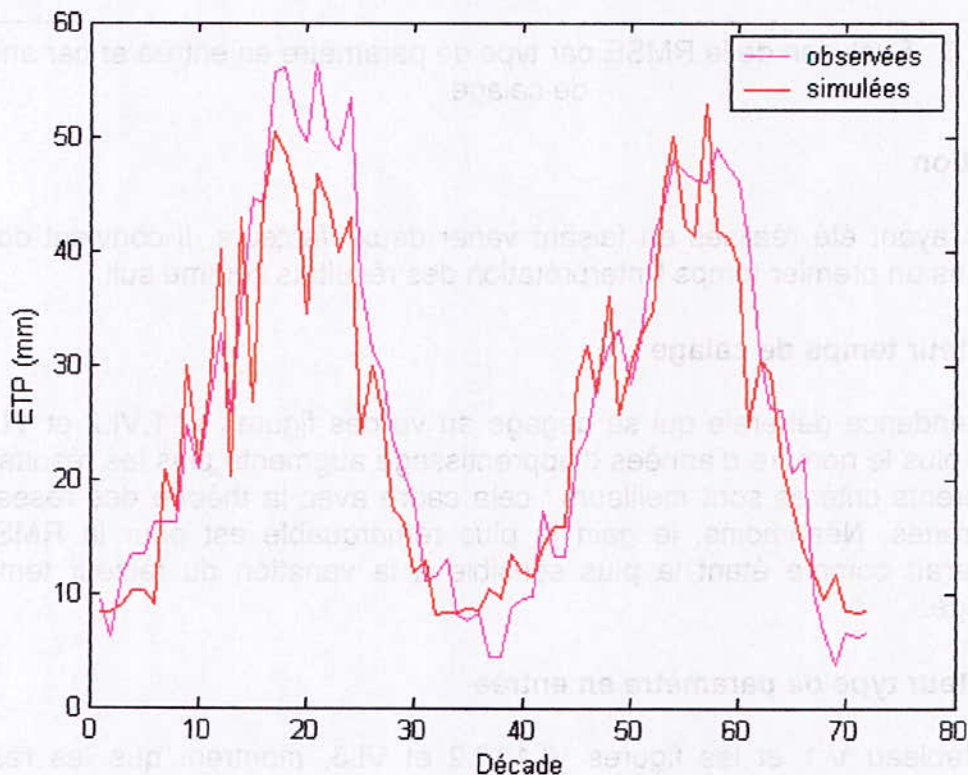


Figure VI.4 : Comparaison graphique des séries simulées et observées pour les deux années de validation.

L'observation de ce graphe nous permet de confirmer les bons résultats obtenus par les critères statistiques de comparaison. Cette figure montre que la série simulée suit bien la tendance même si elle rate quelques pics et qu'elle surestime de manière automatique les petites ETP.

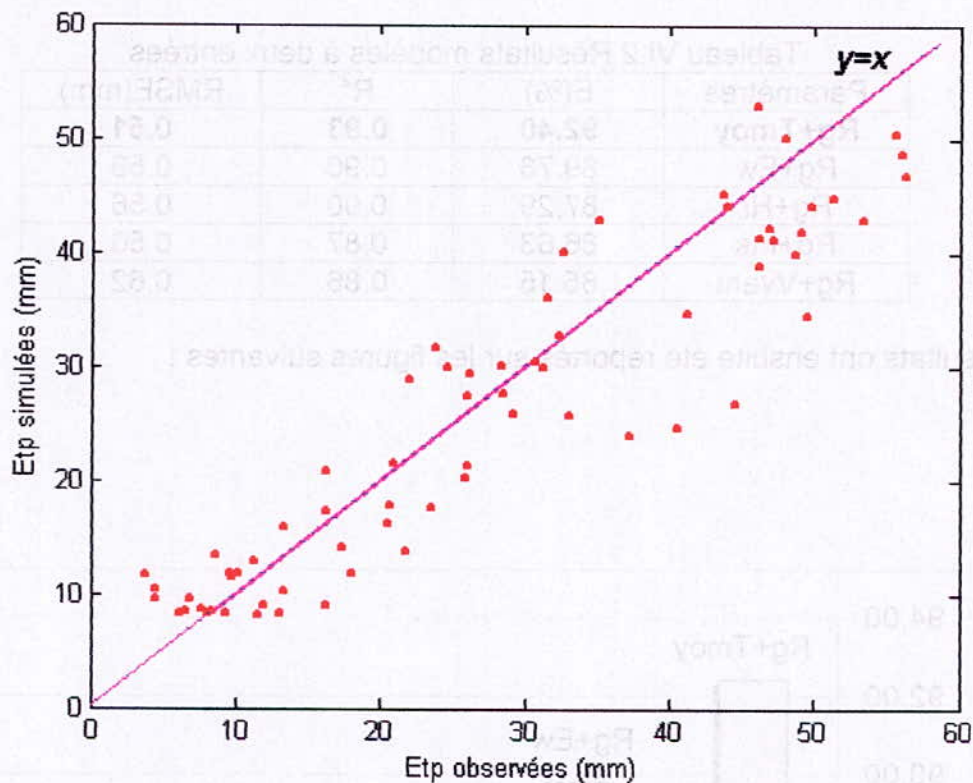


Figure VI.5 : Corrélation entre la série observée et la série simulée pour les deux années de validation.

L'observation de ce dernier graphe met en évidence la bonne corrélation existant entre la série observée et la série simulée car l'on remarque que le nuage de points est peu dispersé et qu'il s'étire assez bien le long de la première bissectrice.

Conclusion

A la lumière des résultats obtenus précédemment, on peut conclure que :

Le meilleur modèle au pas de temps décadaire à un seul paramètre en entrée est le modèle entraîné sur une période de **8 ans** de calage et prenant le **rayonnement global** comme entrée car obtenant les meilleurs résultats pour les critères de comparaison qui sont : $E=85.80\%$, $R^2=0.87$ et $RMSE=0.60$ mm.

VI.1.2. Modèles à deux entrées

La démarche entreprise a consisté en la recherche du meilleur modèle à deux entrées, en faisant varier un seul paramètre : Le type de **couple de paramètres** en entrée. Le nombre d'années d'apprentissage est par ailleurs fixé à **8** pour **2** années de validation. La constitution des couples de paramètres en entrée s'est effectuée en prenant en compte les résultats obtenus lors de la modélisation à un seul paramètre en entrée. En effet, les couples ont été constitués à partir du **rayonnement global**, qui est le paramètre ayant donné les meilleurs résultats précédemment, auquel on a ajouté les autres paramètres un à un dans l'ordre du classement du tableau VI.1.

Les différents essais ainsi que leurs résultats sont résumés et classés sur le tableau VI.2 :

Tableau VI.2 Résultats modèles à deux entrées

Paramètres	E(%)	R ²	RMSE(mm)
Rg+Tmoy	92.40	0.93	0.51
Rg+Ew	89.78	0.90	0.59
Rg+Hr	87.29	0.90	0.56
Rg+Ins	86.63	0.87	0.56
Rg+Vvent	85.15	0.86	0.62

Ces résultats ont ensuite été reportés sur les figures suivantes :

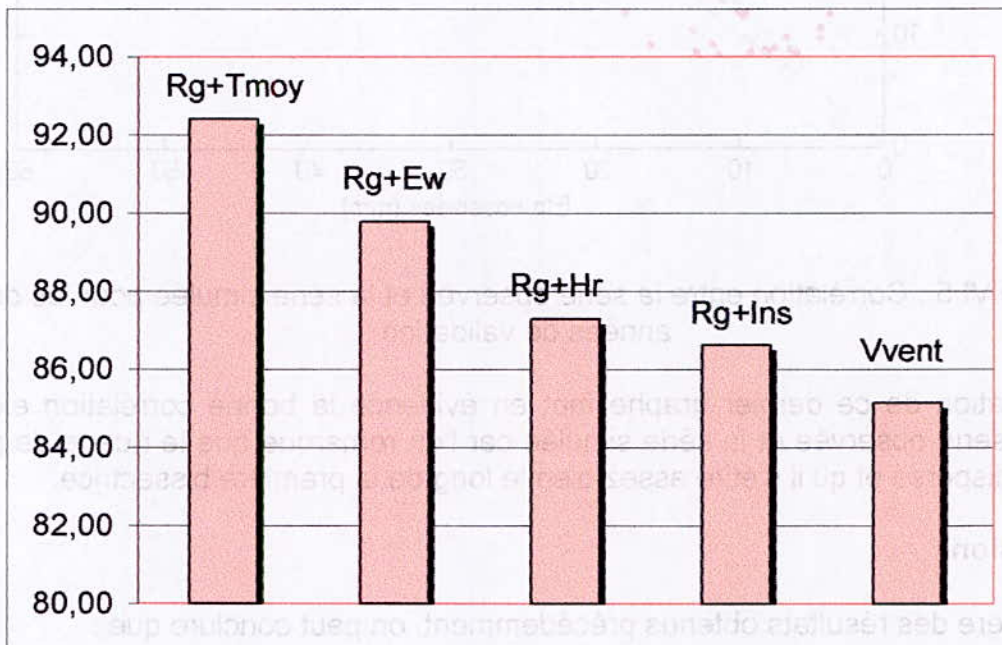


Figure VI.6 : Evolution du critère de Nash par type de couple en entrée

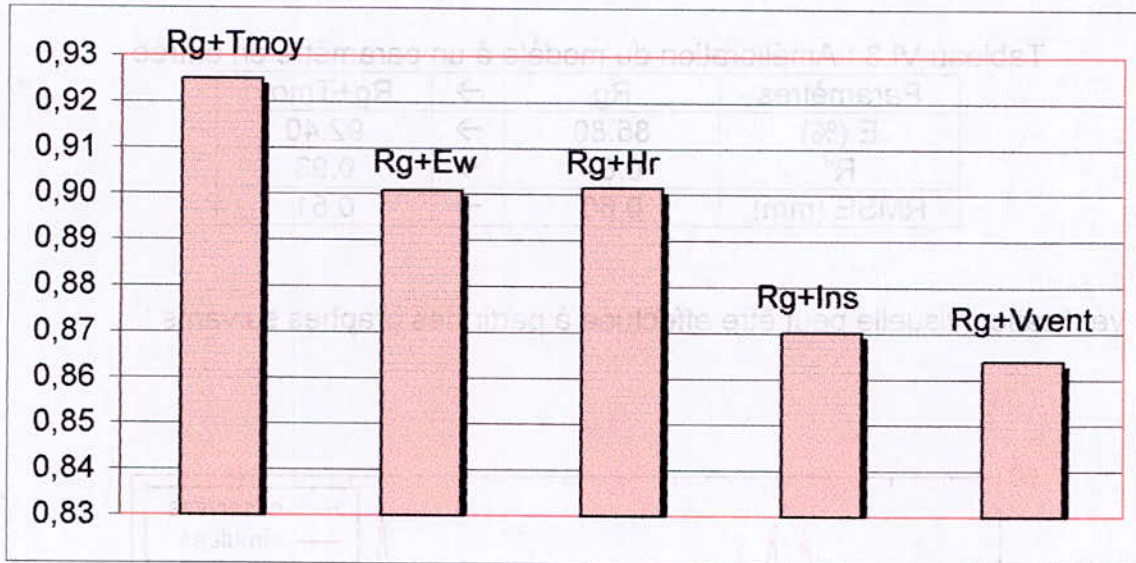


Figure VI.7 : Evolution du coefficient de détermination par type de couple en entrée.

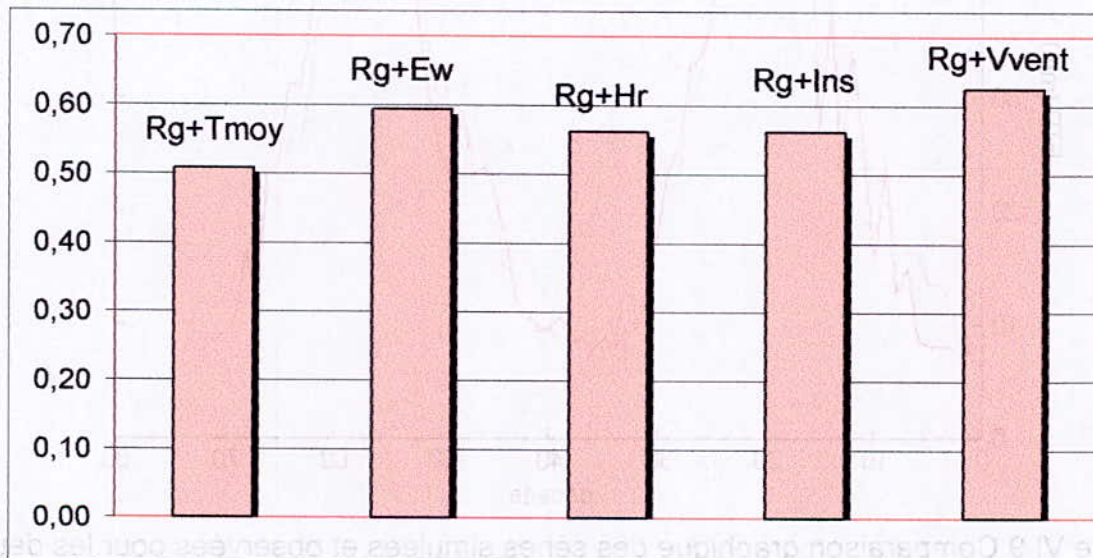


Figure VI.8 : Variation de la RMSE par type de couple en entrée.

Interprétation :

On remarque que la variation des différents couples de paramètres en entrée influe de manière significative sur les coefficients de Nash et de détermination alors que la RMSE y est moins sensible et reste globalement autour de 0.55 mm. On peut aussi dire que, de manière générale, le fait de rajouter un paramètre au rayonnement global permet une amélioration des résultats obtenus par le rayonnement global comme entrée seule, sauf pour le cas du couple (Rg+Vvent).

L'amélioration des résultats diffère d'un paramètre à un autre, et le paramètre dont la contribution a apportée la meilleure amélioration est la température. L'évolution des valeurs des critères statistiques s'est faite comme suit :

Tableau VI.3 : Amélioration du modèle à un paramètre en entrée

Paramètres	Rg	→	Rg+Tmoy
E (%)	85.80	→	92.40
R ²	0.87	→	0.93
RMSE (mm)	0.60	→	0.51

Une vérification visuelle peut être effectuée à partir des graphes suivants :

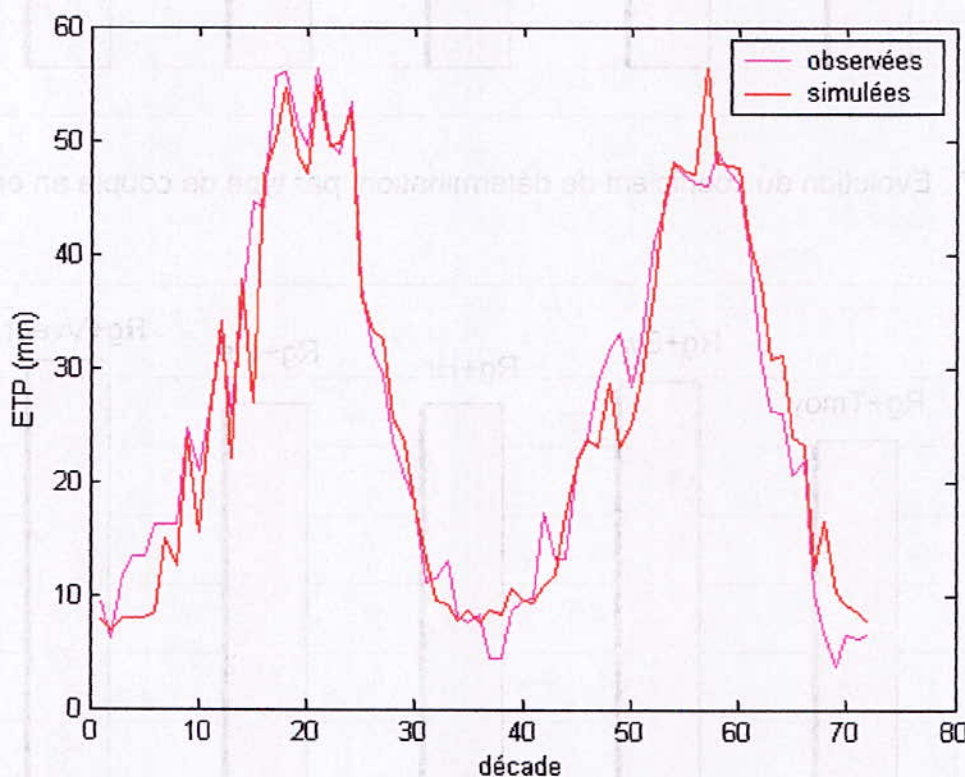


Figure VI.9 Comparaison graphique des séries simulées et observées pour les deux années de validation.

L'observation de cette figure nous permet de confirmer les très bons résultats obtenus par les critères statistiques de comparaison et que la série simulée suit très bien la tendance de la série observée. Les deux sont presque confondues, même si un léger décalage est systématiquement observé au niveau des pics.

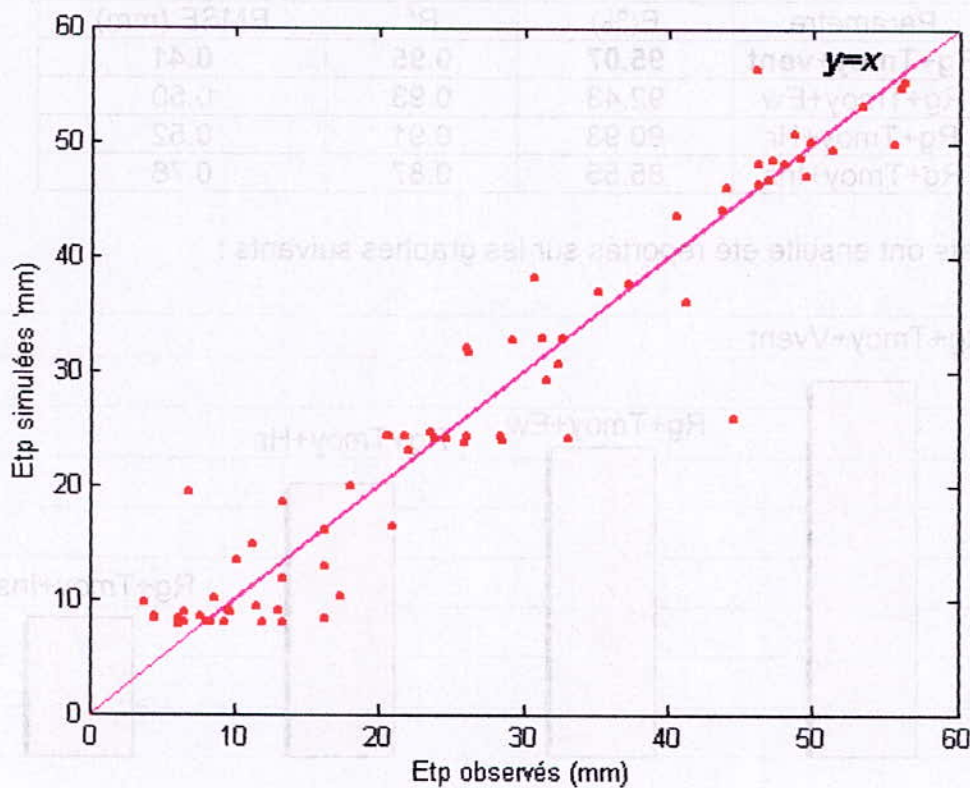


Figure VI.10 : Corrélation entre la série observée et la série simulée pour les deux années de validation.

La très bonne corrélation existant entre les séries observées et simulées se confirme encore sur la dernière figure. En effet, on remarque que le nuage de points est très peu dispersé et qu'il s'étire très bien le long de la première bissectrice.

Conclusion

La tendance à l'amélioration des résultats par l'augmentation des paramètres en entrée se confirme.

Le couple (Tmoy+Rg) offre l'amélioration la plus nette et donne les meilleurs résultats pour les critères statistiques et graphiques, il peut donc être considéré comme le meilleur couple pour le modèle à deux paramètres en entrée.

VI.1.3. Modèles à trois entrées

Les triplettes de données en entrée ont été constituées à partir des résultats obtenus par la modélisation à une seule entrée puis par la modélisation à deux entrées. La démarche a consisté à ajouter au couple (Tmoy++Rg), les différents paramètres restants selon le classement établi au tableau VI.1.

Les résultats des différents essais ont été classés et résumés dans le tableau VI.4:

Tableau VI.4 : Résultats des modèles à trois entrées

Paramètre	E(%)	R ²	RMSE (mm)
Rg+Tmoy+vent	95.07	0.95	0.41
Rg+Tmoy+Ew	92.43	0.93	0.50
Rg+Tmoy+Hr	90.93	0.91	0.52
Rg+Tmoy+Ins	85.55	0.87	0.76

Ces résultats ont ensuite été reportés sur les graphes suivants :

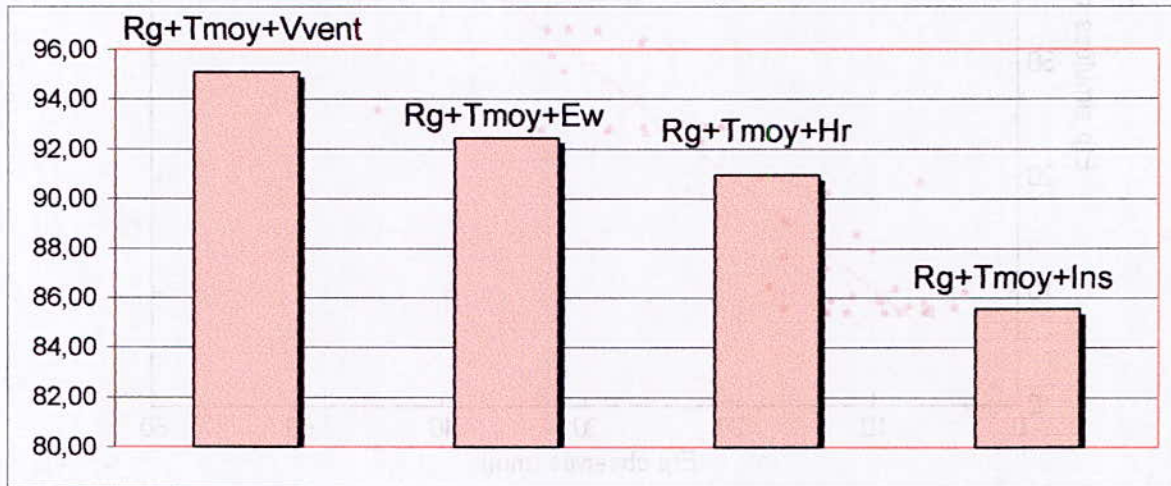


Figure VI.11 : Variation du critère de Nash par type de triplette de données en entrée

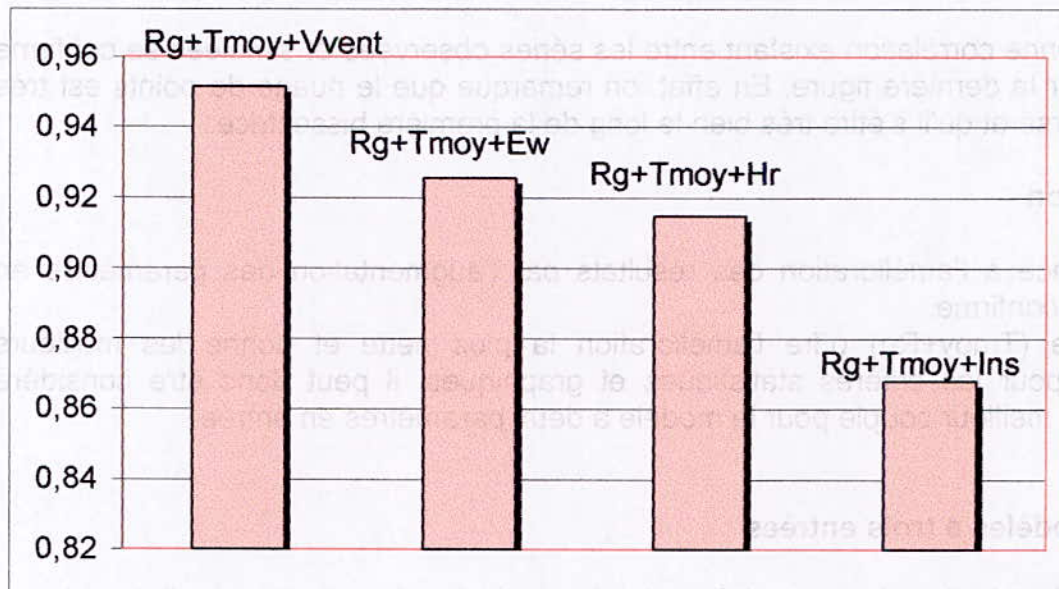


Figure VI.12 : Variation du coefficient de détermination par type de triplette de données en entrée.

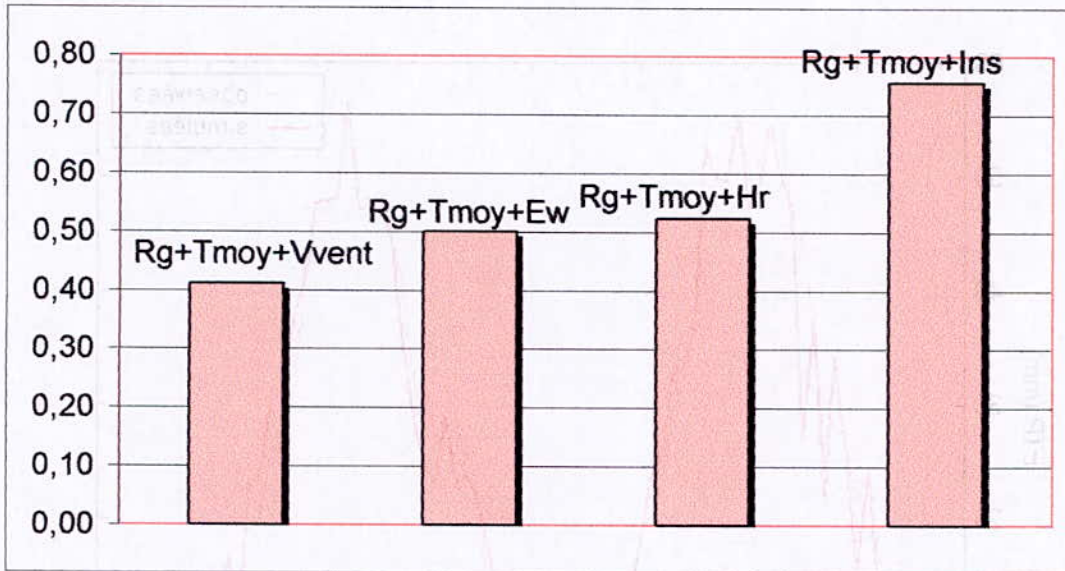


Figure VI.13 : Variation de la RMSE par type de triplette de données en entrée.

Interprétation

D'après les différentes figures représentées ci-dessus, on remarque que la variation des triplets de paramètres en entrée influence de manière significative les trois critères statistiques. Mais, de manière générale, le fait de rajouter un paramètre ne permet pas toujours une amélioration des résultats obtenus par le couple (Rg+Tmoy).

En effet, seul l'ajout du paramètre vitesse du vent offre une amélioration significative des résultats, l'ajout de la tension de vapeur ne les changent pas alors que les critères sont moins bons pour l'ajout de l'humidité relative et de l'insolation.

Seul le paramètre vitesse du vent a apporté un complément d'information au couple (Tmoy+Rg). L'amélioration des valeurs des critères statistiques s'est faite comme suit :

Tableau VI.5 : Amélioration du modèle à deux paramètres en entrée

Paramètres	Rg+Tmoy	→	Rg+Tmoy+Vvent
E (%)	92.40	→	95.07
R ²	0.93	→	0.95
RMSE (mm)	0.51	→	0.41

Une vérification visuelle peut être effectuée à partir des figures suivantes :

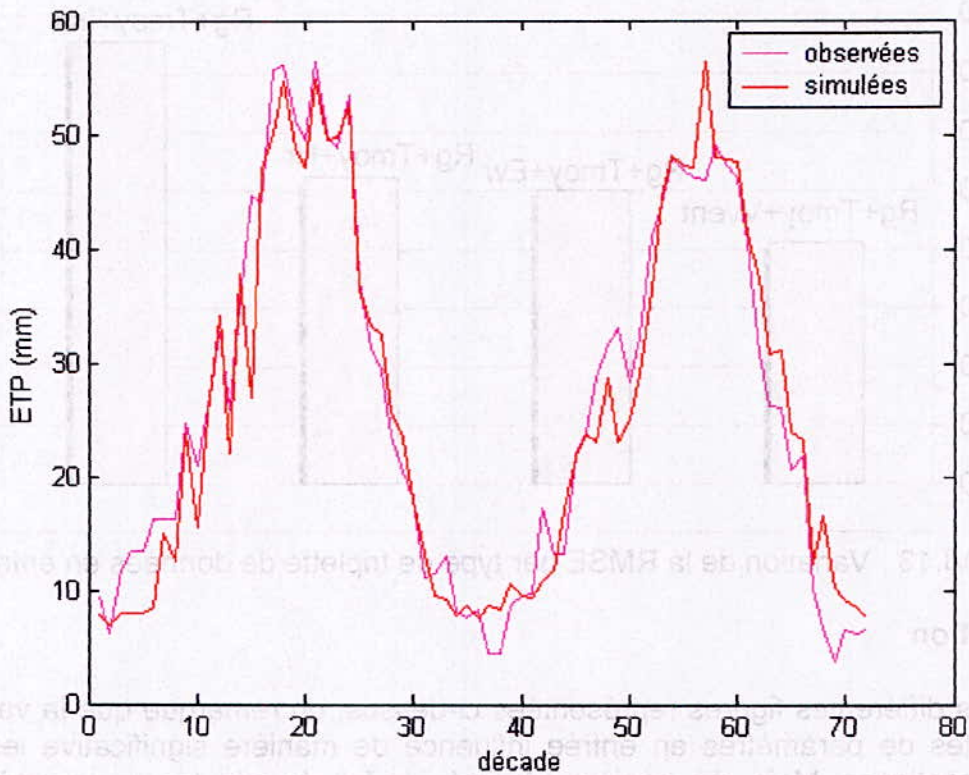


Figure VI.14 : Comparaison série simulée et série observée

L'observation de cette figure nous permet de confirmer les très bons résultats obtenus par les critères statistiques de comparaison. Cette figure nous permet de confirmer que la série simulée suit très bien la tendance et qu'elle ne rate presque aucune singularité : les deux séries se confondent presque entièrement.

Paramètres	Rp+Tmoy	Rp+Tmoy+Vent
E (M)	92.40	92.07
R ²	0.93	0.93
RMSE (mm)	0.51	0.41

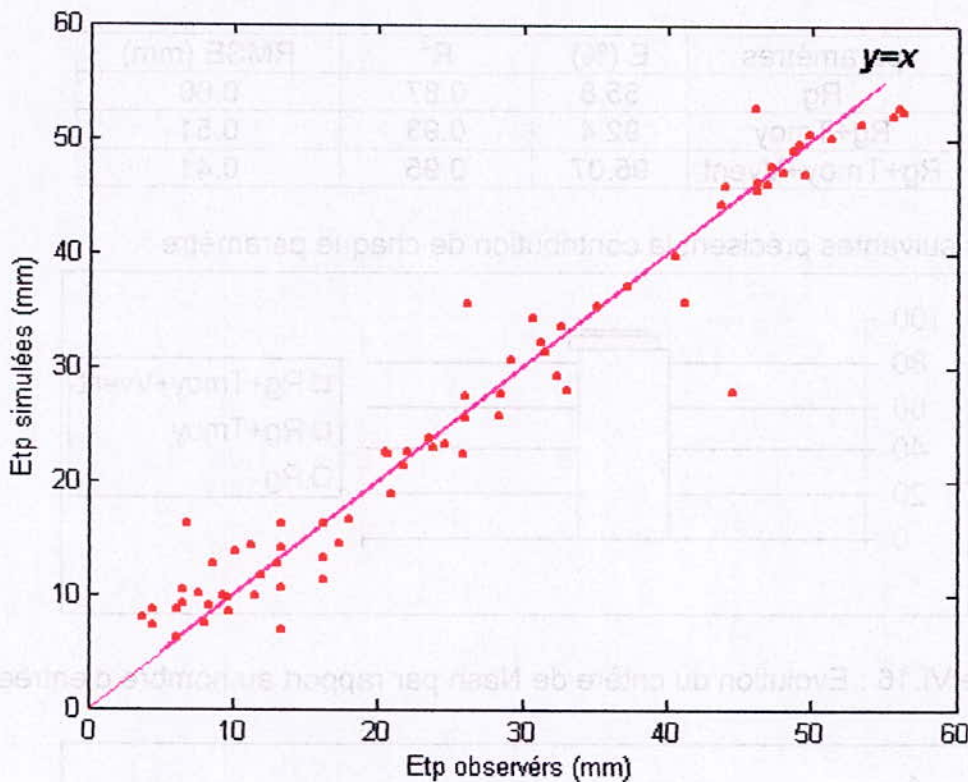


Figure VI.15 : Corrélation entre les séries observées et simulées

L'observation de cette dernière figure met en évidence la presque parfaite corrélation existant entre la série observée et la série simulée car l'on remarque que le nuage de points n'est presque pas dispersé et qu'il s'étire très bien le long de la première bissectrice.

Conclusion

La tendance à l'amélioration des résultats par l'augmentation du nombre de paramètres en entrée n'est plus générale.

La triplette (Tmoy, Rg, Vvent) offre l'amélioration la plus nette et donne les meilleurs résultats pour les critères statistiques et graphiques, elle peut donc être considérée comme la meilleure entrée pour le modèle à trois paramètres.

VI.1.4. Conclusion pour le pas de temps décadaire

Des différents modèles testés, il apparaît que plus le nombre d'entrées augmente plus les résultats des différents critères statistiques et graphiques, utilisés pour la vérification, s'améliorent. L'évolution de la valeur prise par les différents critères par rapport au nombre d'entrées est reportée sur le tableau suivant :

Tableau VI.6 : Variation des résultats par rapport au nombre d'entrées

paramètres	E (%)	R ²	RMSE (mm)
Rg	85.8	0.87	0.60
Rg+Tmoy	92.4	0.93	0.51
Rg+Tmoy+Vvent	95.07	0.95	0.41

Les figures suivantes précisent la contribution de chaque paramètre :

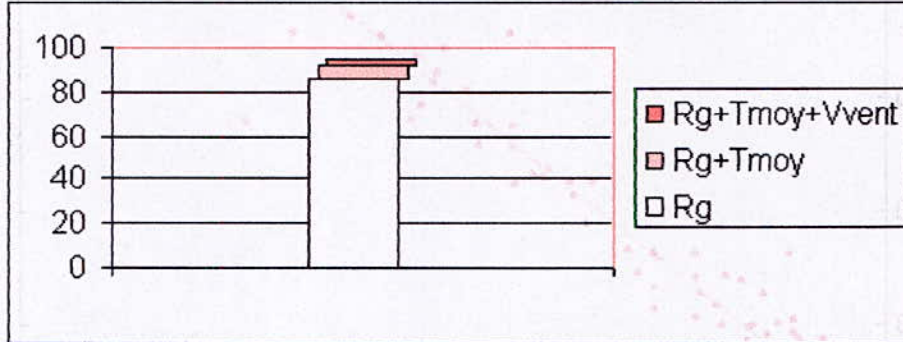


Figure VI.16 : Evolution du critère de Nash par rapport au nombre d'entrées

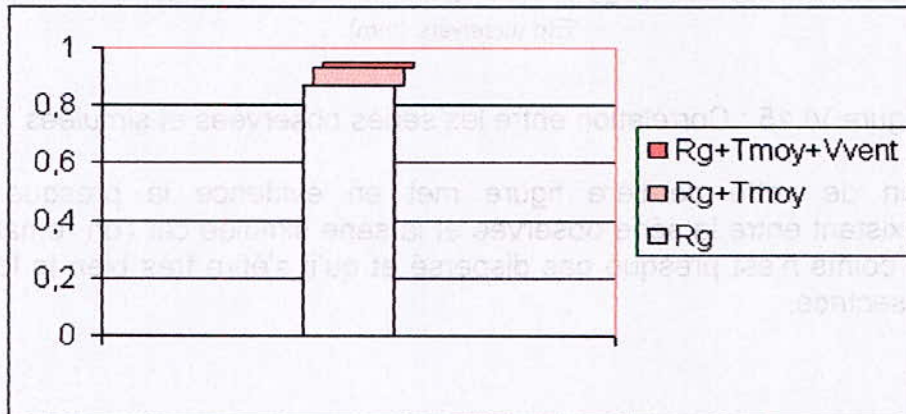


Figure VI.17 : Evolution du coefficient de détermination par rapport au nombre d'entrées

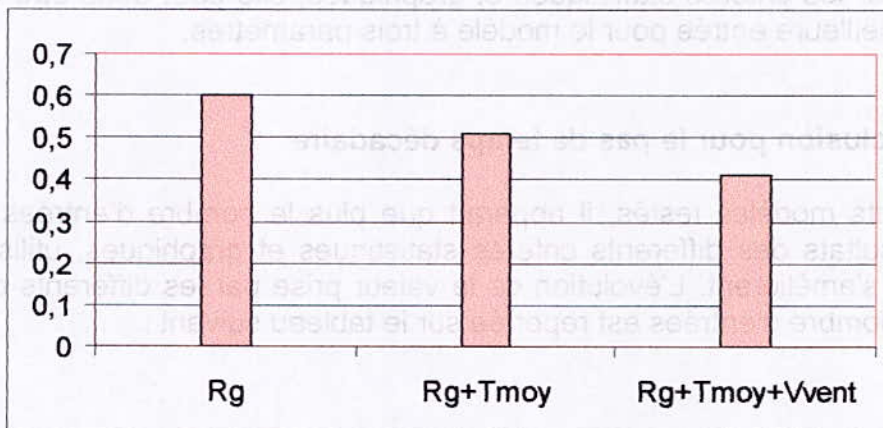


Figure VI.18 : Evolution de la RMSE par rapport au nombre d'entrées

Interprétation

Il est aisé de remarquer sur les trois précédents graphes que le gain en efficacité des critères statistiques est bien là, mais il est tout de même utile de le nuancer. Le passage d'un modèle à une entrée (Rayonnement global) à celui de deux entrées (Rayonnement global+Température moyenne) a permis une amélioration nette des critères statistiques qui justifie amplement cette démarche. Par contre, le gain apporté par le passage au modèle à trois entrées, à partir du modèle à deux entrées, reste discutable, car il reste faible par rapport à la difficulté de mise en oeuvre qu'engendre cette démarche. Il apparaît que le modèle à trois entrées est plus exigeant en termes d'informations et de temps de calcul, sans pour autant apporter une amélioration concrète des résultats.

L'on peut conclure ainsi que pour ce qui est du pas de temps décadaire, le modèle le plus efficace est le perceptron multicouches modélisant l'évapotranspiration potentielle à partir du couple Rayonnement global et Température moyenne en entrée.

VI.2. Modèles au pas de temps mensuel

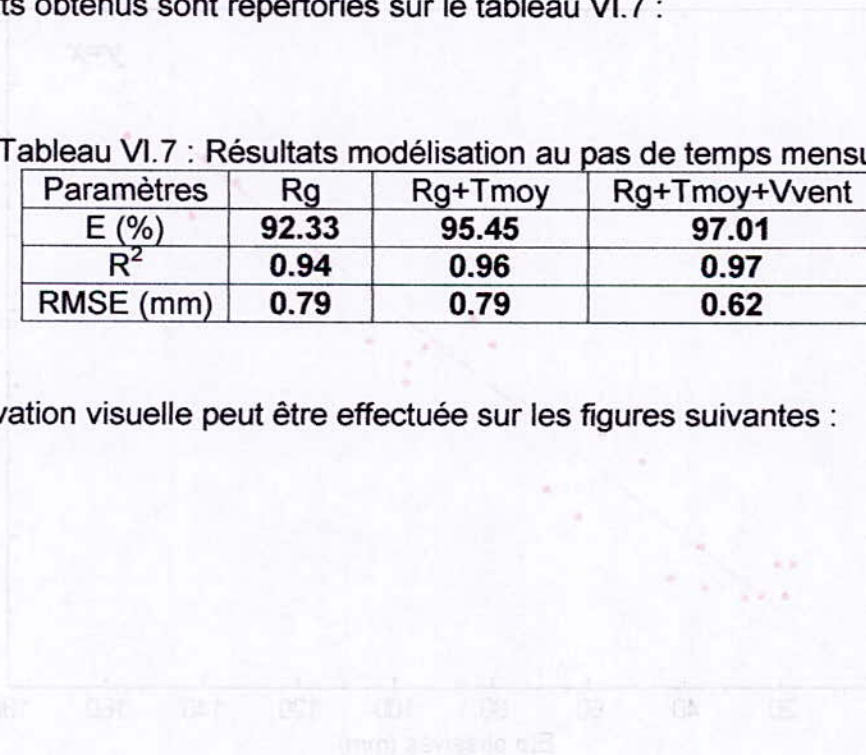
Les modèles ayant donné les meilleurs résultats au pas de temps décadaire ont été repris et testés au pas de temps mensuel. Les caractéristiques des modèles par type d'entrée ont été entièrement reprises et les conditions d'apprentissage restent les mêmes.

Les résultats obtenus sont répertoriés sur le tableau VI.7 :

Tableau VI.7 : Résultats modélisation au pas de temps mensuel

Paramètres	Rg	Rg+Tmoy	Rg+Tmoy+Vvent
E (%)	92.33	95.45	97.01
R ²	0.94	0.96	0.97
RMSE (mm)	0.79	0.79	0.62

Une observation visuelle peut être effectuée sur les figures suivantes :



Modèle à une entrée (Rg) au pas de temps mensuel

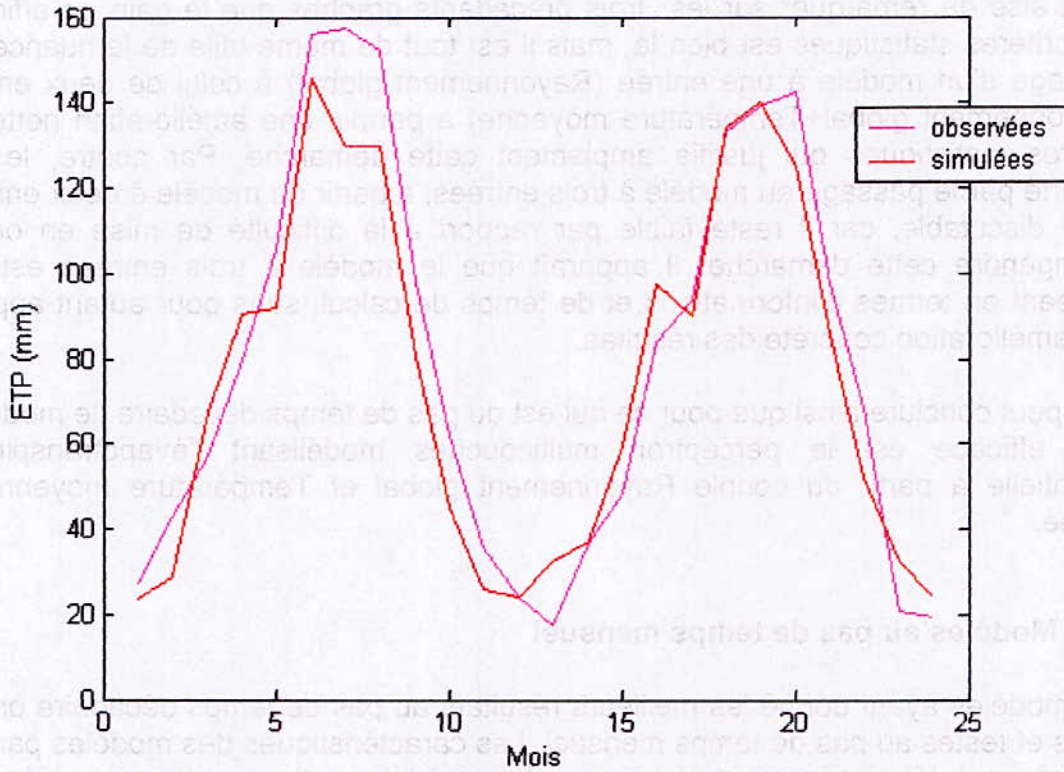


Figure VI.19 Représentation des séries simulées et observées

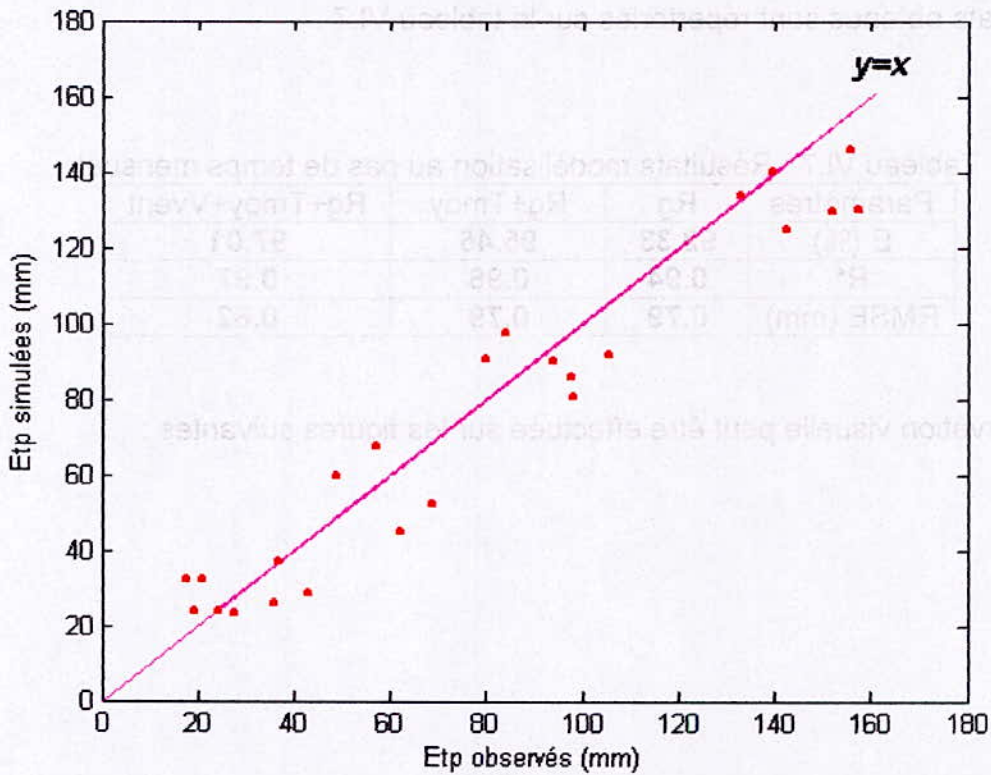


Figure VI.20 : Corrélation des séries simulées et observées

Modèle à deux entrées (Rg+Tmoy) au pas de temps mensuel

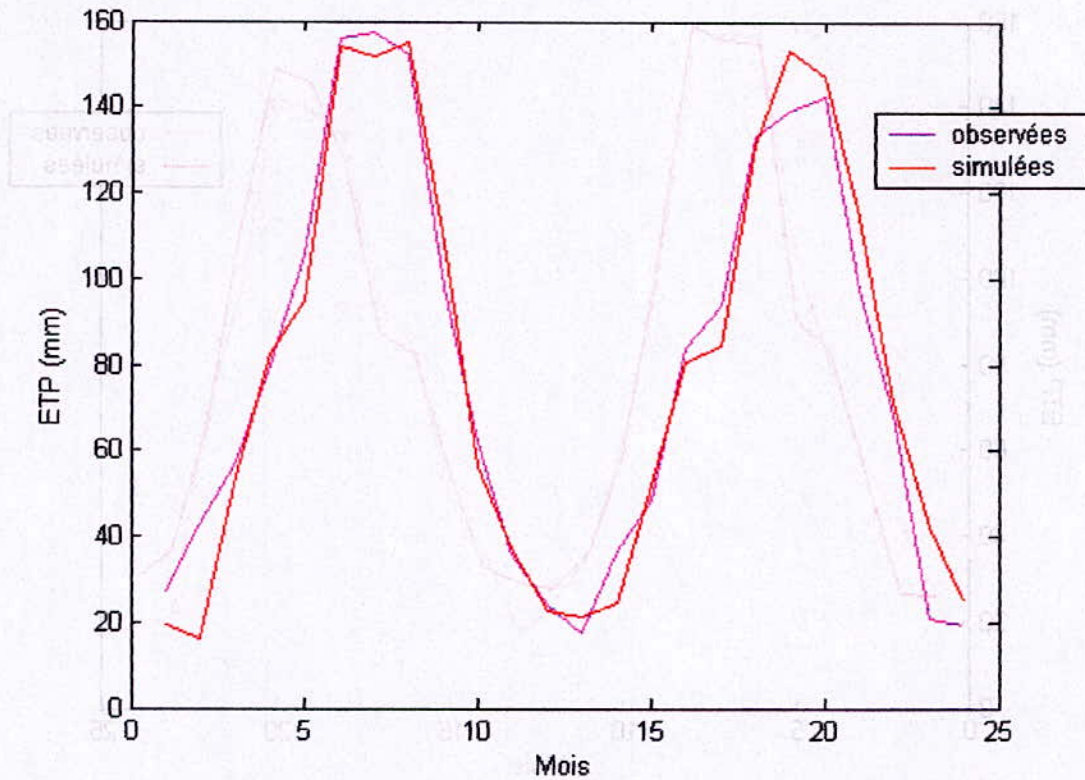


Figure VI.21 Représentation des séries simulées et observées

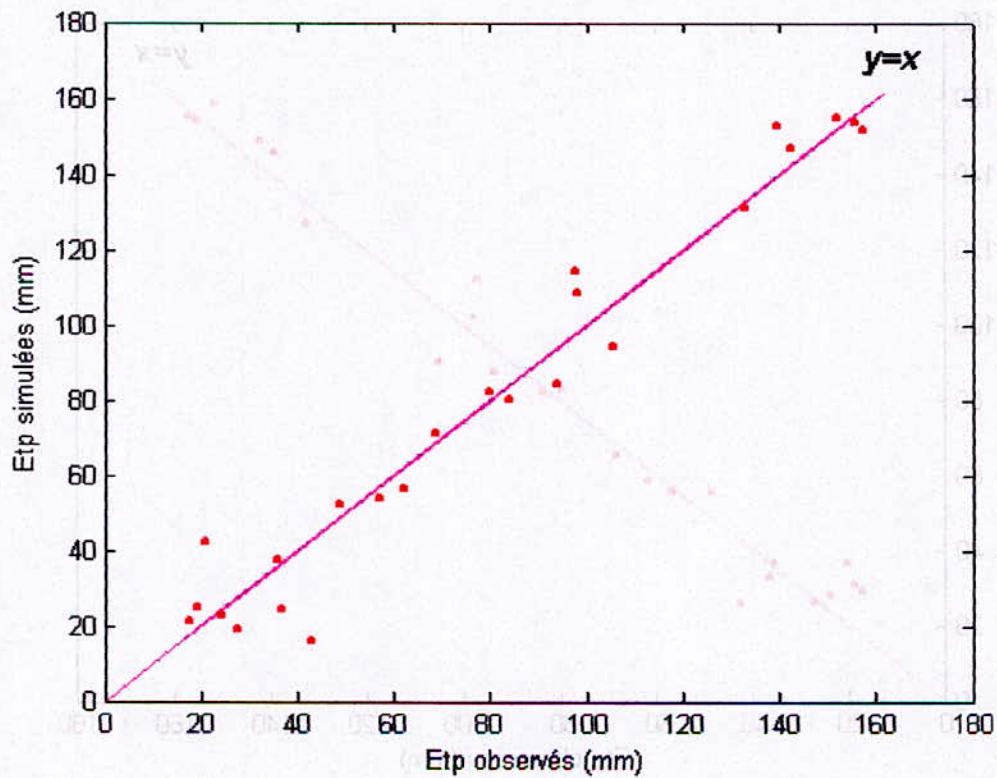


Figure VI.22 : Corrélation entre les séries observées et simulées

Modèle à trois entrées (Rg+Tmoy+Vvent) au pas de temps mensuel

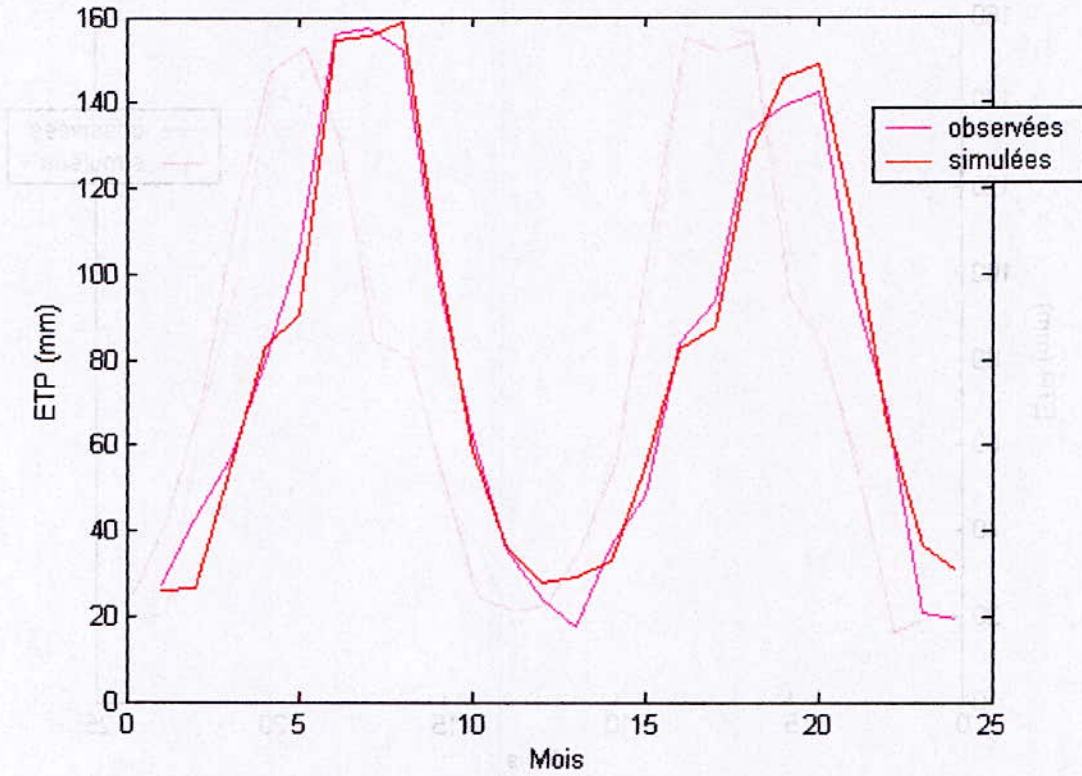


Figure VI.23 : Représentation des séries simulées et observées

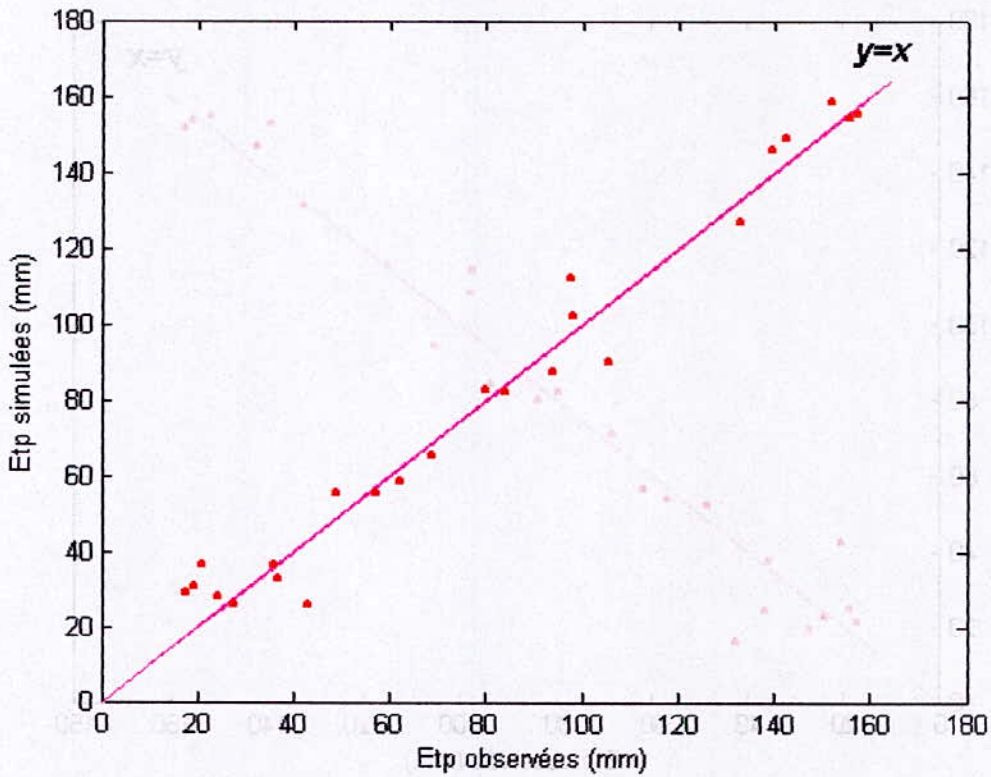


Figure VI.24 : Corrélation les séries observées et simulées

Interprétation

Du tableau VI.7 et des 6 figures précédentes, on remarque que dès le premier test, en prenant le Rayonnement global comme seule entrée, les résultats s'avèrent très concluants. Les séries simulées et observées se confondent presque, et la représentation de la corrélation des deux séries confirme cette tendance. On remarque aussi que ces résultats déjà très encourageants, tendent à s'améliorer par l'augmentation du nombre de paramètres en entrée. Il faut cependant noter que cette amélioration n'est pas très significative comme le montre les deux figures suivantes :

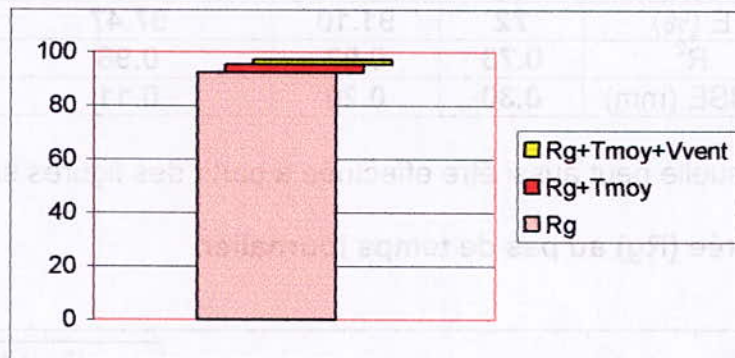


Figure VI.25 : Evolution du critère de Nash par rapport au nombre d'entrées

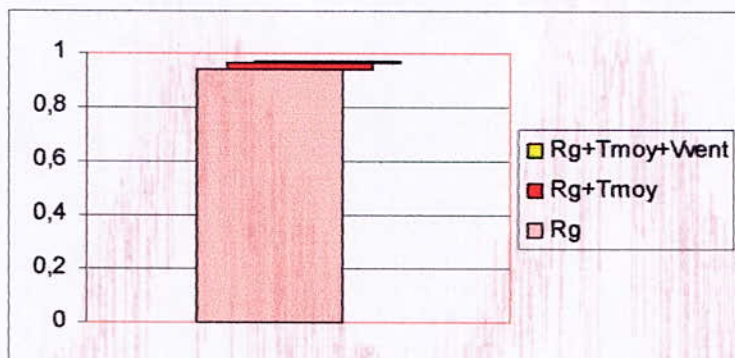


Figure VI.26 : Evolution du coefficient de détermination par rapport au nombre d'entrées

Ces figures mettent en évidence que le gain en efficacité des critères statistiques qui a été permis en passant d'un seul paramètre en entrée à deux puis à trois reste très limité.

Conclusion

L'utilisation au pas de temps mensuel, des modèles ayant donné les meilleurs résultats au pas de temps décadaire, se justifie car les résultats sont très bons. L'évolution des critères statistiques et graphiques en fonction de l'augmentation des paramètres en entrée n'est pas très significative, et le gain apporté par les modèles à deux et à trois entrées ne suffit pas à justifier cette complication du modèle en terme de données en entrée et en temps de calcul.

Nous concluons donc que le meilleur modèle au pas de temps mensuel est le modèle prenant le **Rayonnement global** comme entrée seule.

VI.3. Modèles au pas de temps journalier

Les modèles ayant donné les meilleurs résultats au pas de temps décadaire, et par la suite au pas de temps mensuel, ont été repris et testés au pas de temps journalier. Les résultats sont regroupés sur le tableau suivant :

Tableau VI.8 : Résultats de la modélisation au pas de temps journalier

Paramètres	Rg	Rg+Tmoy	Rg+Tmoy+Vvent
E (%)	72	91.10	97.47
R ²	0.73	0.92	0.98
RMSE (mm)	0.30	0.20	0.11

Une vérification visuelle peut aussi être effectuée à partir des figures suivantes :

Modèle à une entrée (Rg) au pas de temps journalier

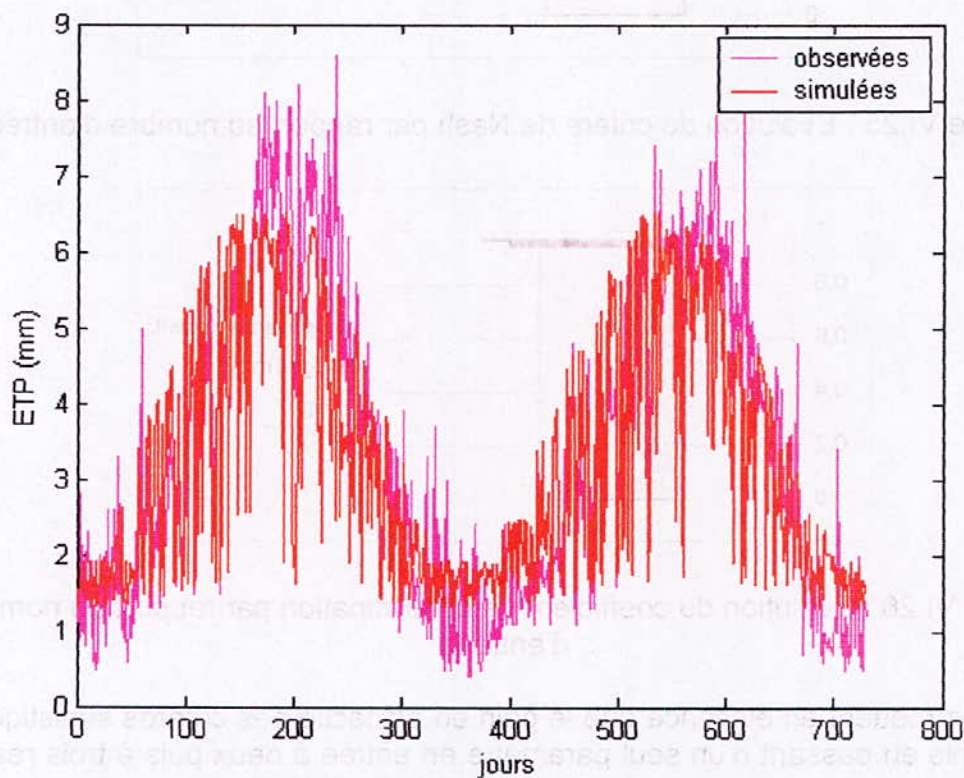


Figure VI.27 : Représentation des séries simulées et observées

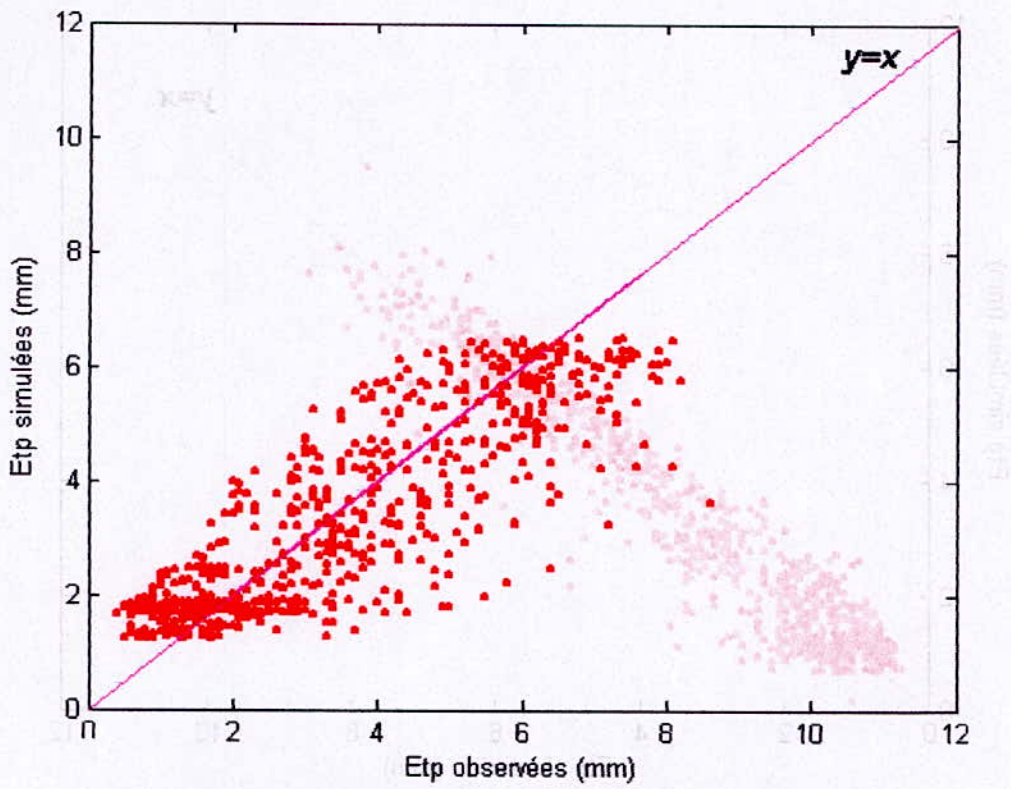


Figure VI.28 : Corrélation des séries simulées et observées

Modèle à deux entrées (Rg+Tmoy) au pas de temps journalier

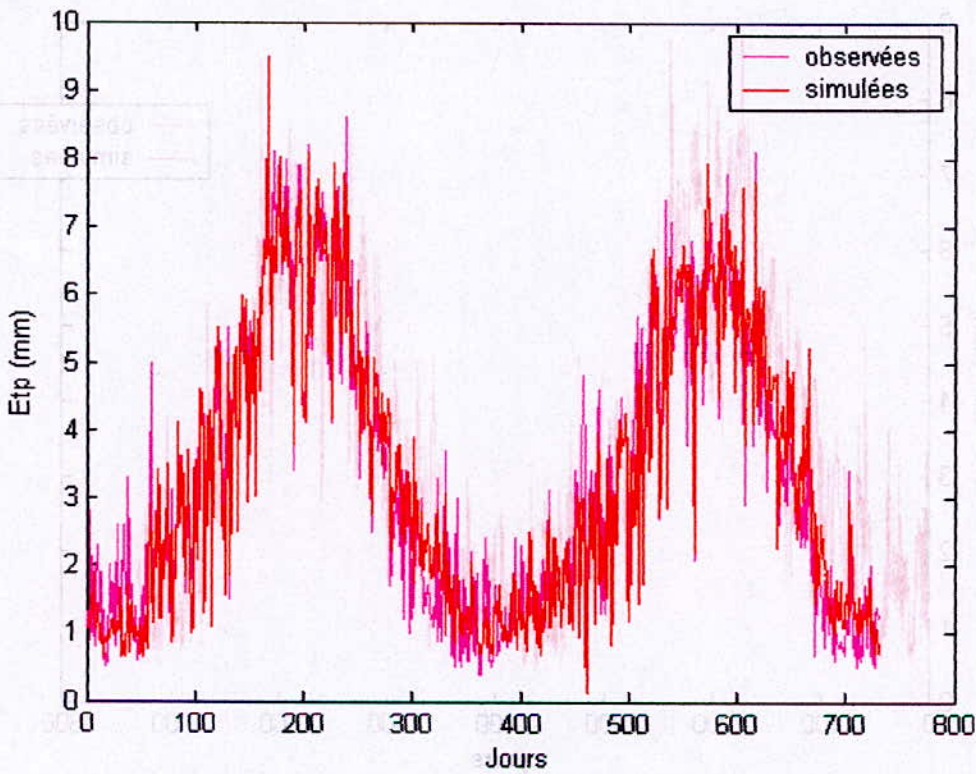


Figure VI.29 : Représentation des séries simulées et observées

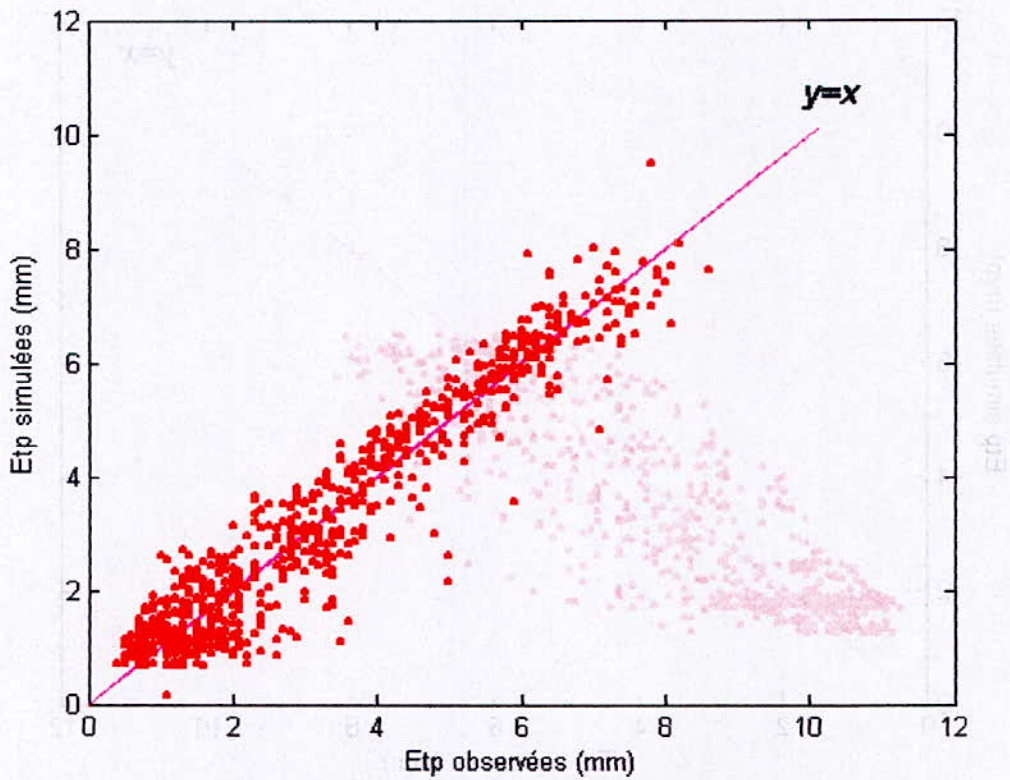


figure VI.30 : Corrélation entre les séries simulées et observées

Modèle à trois entrées (Rg+Tmoy+Vvent) au pas de temps journalier

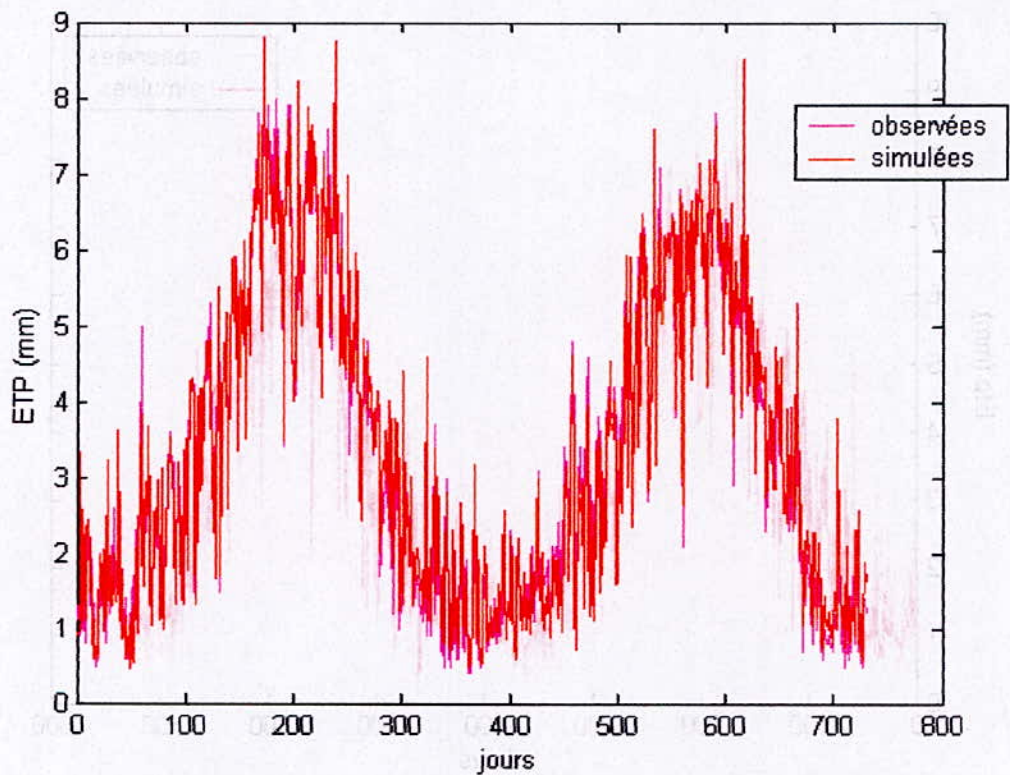


Figure VI.31 : Représentation des séries simulées et observées

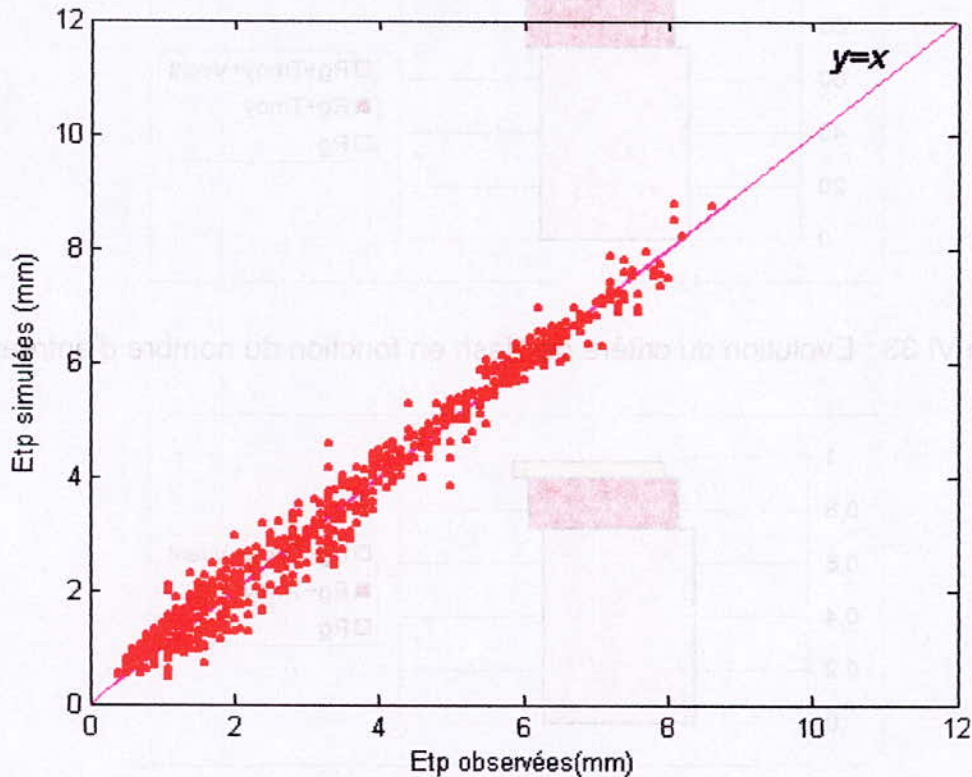


Figure VI.32 : Corrélation des séries simulées et observées

Interprétation

Les valeurs des critères statistiques du tableau VI.8, montrent que lors de la première modélisation à un seul paramètre en entrée (R_g), les résultats étaient assez médiocres. Par contre, une très nette amélioration des critères est observée lors du passage à deux paramètres en entrées. Cette amélioration est encore constatée de manière significative, même si elle est moins évidente que la première, en prenant trois paramètres en entrée.

Pour les critères graphiques, la figure VI.27 montre que la série simulée n'est pas en accord avec la série observée. On constate un certain décalage, alors que les extrêmes sont ratés de manière systématique. La figure VI.28 nous montre un nuage de points plus tôt étalé, et dispersé. Les observations visuelles confirment les résultats assez médiocres des critères statistiques obtenus en prenant le rayonnement global en entrée seule.

A l'inverse, les figures VI.29 et VI.30 indiquent une très bonne corrélation entre les séries observées et simulées par le modèle prenant le couple ($R_g + T_{moy}$) comme entrée. Cette corrélation s'améliore encore plus au vu des figures VI.31 et VI.32 qui correspondent aux séries observées et simulées par le modèle prenant en entrée la triplette ($R_g + T_{moy} + V_{vent}$). L'évolution des critères de Nash et du coefficient de détermination en fonction du nombre de paramètres en entrée est représentée sur les figures suivantes :



Figure VI.33 : Evolution du critère de Nash en fonction du nombre d'entrées.



Figure VI.34 : Evolution du coefficient de détermination en fonction du nombre d'entrées

On constate que le modèle à deux entrées permet une amélioration très importante des résultats du modèle à une entrée, et que cette amélioration se poursuit encore avec le passage au modèle à trois entrées, même si cette dernière n'est pas aussi nette.

Conclusion

La tendance à l'amélioration des résultats par l'augmentation des paramètres en entrée du modèle se confirme de manière très nette pour le pas de temps journalier. La modélisation par le rayonnement global en entrée n'étant pas très concluante, le passage au modèle prenant le couple rayonnement global et température moyenne comme entrée se justifie pleinement, car apportant une amélioration très nette de la modélisation. Le même raisonnement peut être adopté pour le passage du modèle à deux entrées ($Rg+Tmoy$) au modèle à trois entrées ($Rg+Tmoy+Vvent$), mais dans une moindre mesure.

Les meilleurs modèles au pas de temps journalier sont au choix les modèles à deux et trois entrées, selon la disponibilité de l'information vitesse du vent.

Conclusion

CONCLUSION GENERALE

L'évapotranspiration potentielle, ou ETP, est une composante très importante du bilan hydrique. Elle se définit comme étant la somme des volumes d'eau, perdus dans l'atmosphère, à travers le sol, les étendues d'eau et la végétation sous la pression de la demande climatique. Il est nécessaire de quantifier précisément cette demande, car elle joue un rôle important dans le calcul des bilans des retenues, et représente les besoins des cultures en eau. Elle est donc un élément clé, sans lequel aucune intensification de l'agriculture n'est possible. De part l'importance de son rôle dans la gestion des ressources en eau, différentes approches de calcul et de modélisation ont été tentées durant toute la deuxième moitié du XX^e siècle. Ces méthodes se sont basées sur des formules empiriques, obtenues à la suite d'ajustements statistiques, ou alors, sur des formules physiques, issue de la quantification du bilan de l'énergie. Lors du présent travail, nous avons modélisé

Le but du travail présenté était de proposer une nouvelle méthode de calcul de l'ETP, et d'optimiser le nombre de paramètres entrant e compte dans son calcul, en utilisant une technique d'intelligence artificielle : Les réseaux de neurones artificiels ou RNA. Comme leur nom l'indique, les RNA tentent de reproduire le fonctionnement du cerveau humain, en imitant sa composante de base : le neurone. Le but étant d'adopter un comportement dit intelligent, afin de pouvoir appréhender des phénomènes complexes.

La formule de Penman, qui est la relation liant l'ETP à un certain nombre de paramètres climatiques, à savoir : la température moyenne, l'insolation, la vitesse du vent et l'humidité relative, de par son caractère non linéaire et complexe, a été prise en considération pour ce type de modélisation.

Nous disposons d'une base de données comprenant, en plus des paramètres cités plus haut, le rayonnement global, la tension de vapeur et l'ETP calculée par la formule de Penman. Nous avons développé plusieurs modèles, en fixant l'ETP comme sortie, et en variant les combinaisons en entrée à partir des six paramètres climatiques disponibles, et ce, dans les trois pas de temps : journalier, décadaire et mensuel.

Les résultats obtenus ont été très satisfaisants, pour les trois pas de temps adoptés, et se répartissent comme suit :

Au pas de temps journalier : Les modèles à une entrée (le rayonnement global) ont donné des résultats médiocres, mais une amélioration très nette des performance est constatée lors de l'augmentation du nombre de paramètres en entrée à deux puis à trois. Cette amélioration est très significative dans les deux cas et elle justifie cette complication du modèle.

Au pas de temps décadaire : Les résultats du modèle à une entrée (le rayonnement global) ont été satisfaisants et se sont améliorer de manière significative lors du passage à deux paramètres en entrée formant le couple : (rayonnement global + température moyenne), justifiant ainsi pleinement cette complication du modèle. Par

contre l'ajout du paramètre vitesse du vent au précédent couple, n'a offert qu'un gain très limité, rendant ainsi le modèle à deux entrées le plus intéressant.

Au pas de temps mensuels : D'emblée le modèle à une entrée est considéré comme le plus efficace, car donnant d'excellents résultats, qui ne sont que très peu améliorés par l'augmentation du nombre de paramètres en entrée.

A la lumière des résultats obtenus, il apparaît que pour les trois pas de temps, l'ETP a été reconstituée, ou calculée, avec une erreur très limitée. De plus, le nombre de paramètres nécessaires au calcul a été optimisé, car les meilleurs modèles n'en demandent qu'un ou deux en entrée, nombre qu'il faut comparer aux 4 paramètres nécessaires pour le calcul par la formule classique.

Les résultats obtenus, lors de ce travail, peuvent ouvrir une somme de perspectives dans ce domaine : beaucoup de combinaisons de paramètres en entrée n'ont pas été testées, car la démarche de recherche des meilleurs modèles n'a pas été exhaustive. De plus, les modèles développés peuvent être adaptés pour la prévision et d'autres types de réseaux de neurones formels ou d'autres techniques d'intelligence artificielle, telle que la logique floue, peuvent être testés.

Le travail présenté n'a fait que défricher, très sommairement, une forêt de possibilités qui restent encore à explorer.

Bibliographie

- ALLEN RG., PEREIRA LS., RACES D., SMITH M., 1998. Crop evapotranspiration guidelines for computing cropwater requirements. FAO Irri.drain, 56.
- ANRH., 2004. Approche de l'équation de Penman pour le nord de l'Algérie. ANRH
- BERMAD A., 2004. élaboration d'une méthodologie de simulation et de prévision des variables climatologiques aléatoires et cycliques. Th doctorat. Ecole Nationale Polytechnique d'El Harrach.
- BOUCHET ., 1964. Evapotranspiration réelle, évapotranspiration potentielle et productivité végétale. L'eau et la productivité végétale, INRA, Paris.
- BROCHET P., GERBIER N., 1972. Une méthode pratique de calcul de l'évapotranspiration potentielle. Ann. Agron, 23.
- CALVET R., 1988. Etude sur les transferts d'eau dans le système sol-plante-atmosphère. La découverte, France.
- CHOISEL E., 1988. Estimation de l'évapotranspiration potentielle à partir des données météorologiques. La météorologie, 7 serie24,19-26.
- DAGNELIE P., 1992. Statistique théorique appliquée tome1. Les presses agronomiques de Grenoble.
- DAVALO E., NAIM P., 1991. Des réseaux de neurones. Eyrolles.
- DEBBAR H., 1993. Application des réseaux de neurones artificiels à la détection d'intrusion sur les systèmes informatiques. Th doctora. U Paris IV.
- DOORENBOS J., Pruit WQ., 1977. Guidelines for predicting cropwater requirements. FAO Irri.drain, 24, 156.
- DOWLA FV., ROGERS LL., 1995. Solving problems in environmental engeneering and geoscience with artificial neural net work, MIT Press, Cambridge.
- ENCARTA ., 2004. Encyclopédie multimédia de la connaissance.
- LEHRE A., Penman approach to evaporation and evapotranspiration. Dept of geology, humboldt state.
- PARIZEAU M., Réseaux de neurones. 2004. U Laval.
- PATTERSON DW., 1996. Artificial neural networks: theory and applications. Simonon and & Schuster, Singapore.
- SEYNI S., 2000. Approche de l'évapotranspiration potentielle par le rayonnement net (station de Dar El Beida)

TOUMI F., 2001. Conception d'un modèle d'évapotranspiration basée sur le rayonnement. Pfe. Institut national agronomique el harrach.

TOUZET C., 1992. Les réseaux de neurones. EERIE.

Ventura F., Spano D., DiceP., Snyder Rl.,1999. An evaluation of common evapotranspiration equations. Irrigation science, 18,163-170.

YAHIAOUI T., 1982. Rayonnement solaire sur l'Algérie. Office national de la météorologie.