



CENTRE NATIONAL  
DE LA RECHERCHE  
SCIENTIFIQUE



UNIVERSITÀ  
DI CORSICA

PASQUALE  
PAOLI

# Bilan Intermédiaire: résultats quatrième trimestre



Etablissement Public de Santé

Sujet de thèse:  
Prédiction de séries temporelles et  
applications aux énergies  
renouvelables

**CHD Castelluccio, Unité Radiophysique Service Radiothérapie, 20176 Ajaccio**

**Laboratoire SPE, Université de Corse, CNRS-UMR 6134, 20250 Corte**

**Cyril Voyant**  
[cyrilvoyant@hotmail.com](mailto:cyrilvoyant@hotmail.com)

# synthèse de la dernière réunion

2/28

- *Concentrer les investigations sur la prédiction à horizon 24 Heures par pas horaire ou journalier dans le cas ARMA*
- *Continuer l'étude physique des phénomènes d'irradiation solaire*
- *Couplage avec des données Metars ou Métosat 7-8*
- *Commencer à essayer Matlab pour l'étude de séries chronologiques*
- *Dresser une comparaison par paramètres objectifs entre la prédiction stochastique ARMA et la prédiction ANN*
- *Finaliser la bibliographie par un document synthétisant les résultats*

# Sommaire

3/28

- Echelle journalière, optimisation des *ANN*
- Echelle journalière, comparaison *ANN* Vs *ARMA*
- Echelle horaire, comparaison *ANN* Vs *ARMA*
- Perspectives

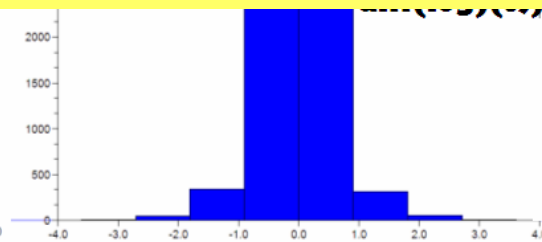
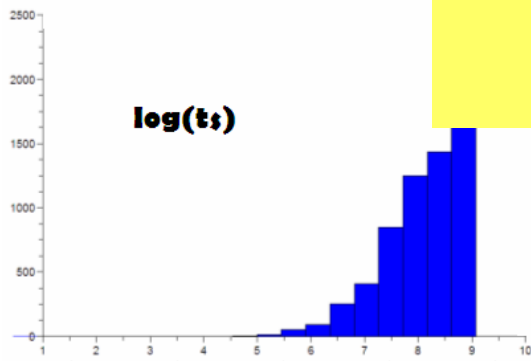
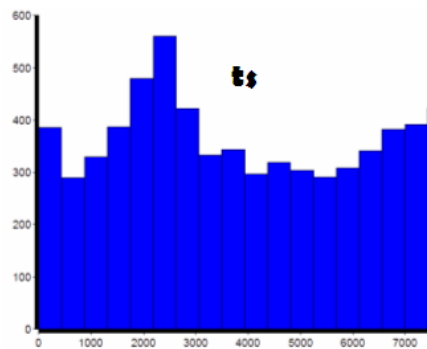
# Sommaire

4/28

- Echelle journalière, optimisation des *ANN*
- Echelle journalière, comparaison *ANN* Vs *ARMA*
- Echelle horaire, comparaison *ANN* Vs *ARMA*
- Perspectives

# La série temporelle étudiée : Graphique et distribution

## Histogramme de la série journalière



L'unité propre aux séries temporelles étudiées par la suite est l'énergie d'irradiation surfacique, elle ne sera pas répétée à chaque fois

ts\_soleil\_quotidien (Wh/m2)



Graphique de la série

# Optimisation du réseau choisi pour la prédiction de la ressource solaire

1. Type de réseau
2. Prétraitement des données d'entrée
3. Nombre de couches et de neurones cachés
4. Choix des fonctions de transfert
5. Nombre de neurones d'entrée
6. Algorithme d'apprentissage
7. Mode online ou batch
8. Weight decay
9. Moment
10. Pas d'apprentissage
11. Initialisation des poids
12. Early stopping
13. Fonction de coût
14. Fonction de comparaison
15. Duplication d'entrée
16. Bruitage des données d'entrée
17. Répartition des données (apprentissage validation et test)
18. Permutation du pool d'apprentissage, taille d'échantillon
19. Customisation => architecture, normalisation, fonction de transfert

Réseau de départ : PMC 8x10x1 ; purlin x tansig x purlin ; LM (**Cf Andranova 2006**)

A chaque étape on étudie un paramètre qui sert aux étapes suivantes

## Optimisation du réseau choisi pour la prédiction de la ressource solaire

Nous cherchons les paramètres ( $X_i$ ) du réseau ANN qui minimisent l'erreur de prédiction, soit :

$$dE = 0$$

On applique l'hypothèse **naïve** (**mais fausse**) que toutes les variables sont indépendantes, on obtient

$$dE = \sum_{i=1}^N \frac{\partial E}{\partial X_i} \cdot dX_i$$

On va étudier séparément tous les paramètres du réseau

$$\frac{\partial E}{\partial X_i} = 0$$

## Optimisation des ANN : le nombre de layer

8/28

La littérature nous enseigne que les prédictions par ANN utilisent 1 ou 2 couches cachées.

Multiplier le nombre de couches alourdi le système mais n'est pas forcément bénéfique => parcimonie

Layer1	0	1	5	10	20	25	50
Layer2							
0	21,25	21,26	21,26	21,20	21,19	21,65	21,06
1	21,25	21,06	21,20	21,01	21,21	20,98	21,35
5	21,26	21,12	21,29	21,45	20,99	21,88	25,45
10	21,20	21,19	21,30	21,5	22,22	21,85	21,76
20	21,19	21,28	21,46	21,08	22,2	21,84	22,77
25	21,65	21,28	20,86	22,01	22,45	22,37	22,31
50	21,06	24,00	21,18	22,13	21,87	20,95	22,37

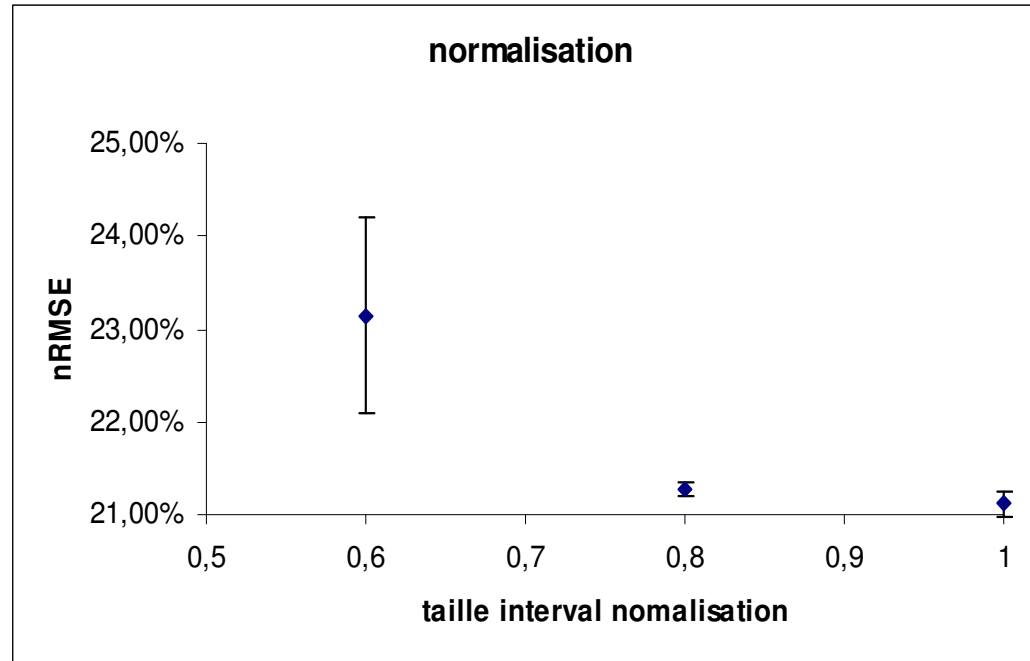
Mediane = 21,4 en gras et bleu les valeurs inférieures. nRMSE en fonction du nombre de neurones cachés sur deux layers

**Un seul layer choisi, dont le nombre de neurones est à affiner**



## Optimisation des ANN : la normalisation des données

9/28

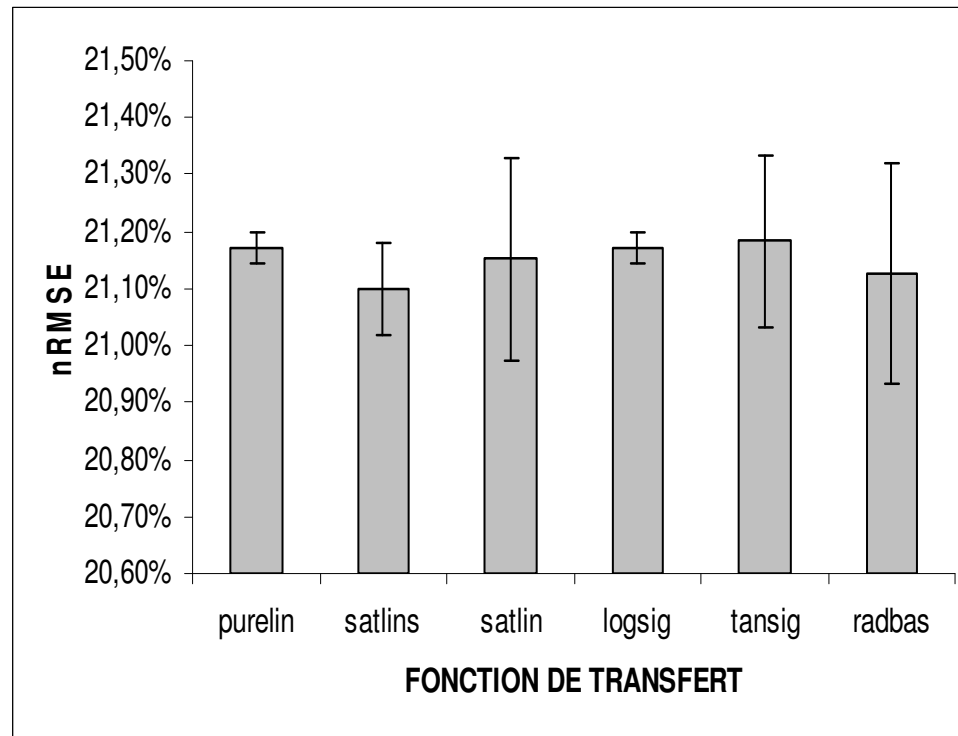


La normalisation des données d'entrée peut permettre d'améliorer les résultats en négligeant les effets de bord de la fonction de transfert (sigmoïde)

**La normalisation sur l'intervalle [0 1] est retenue**

## Optimisation des ANN : la fonction de transfert

10/28



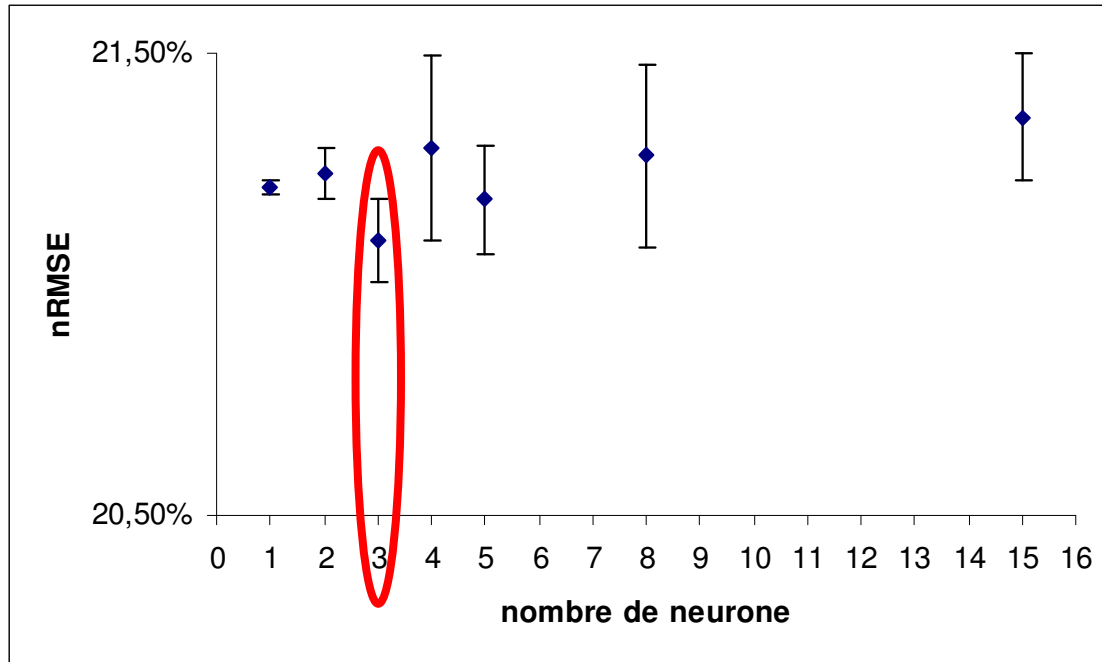
La fonction de transfert est le paramètre qui permet la prédiction de phénomène non linéaire.

Une fonction linéaire équivaut à faire du AR

**La fonction gaussienne (radbas) est retenue. C'est celle avec le meilleur potentiel de prédiction**

## Optimisation des ANN : le nombre de neurone cachées

11/28

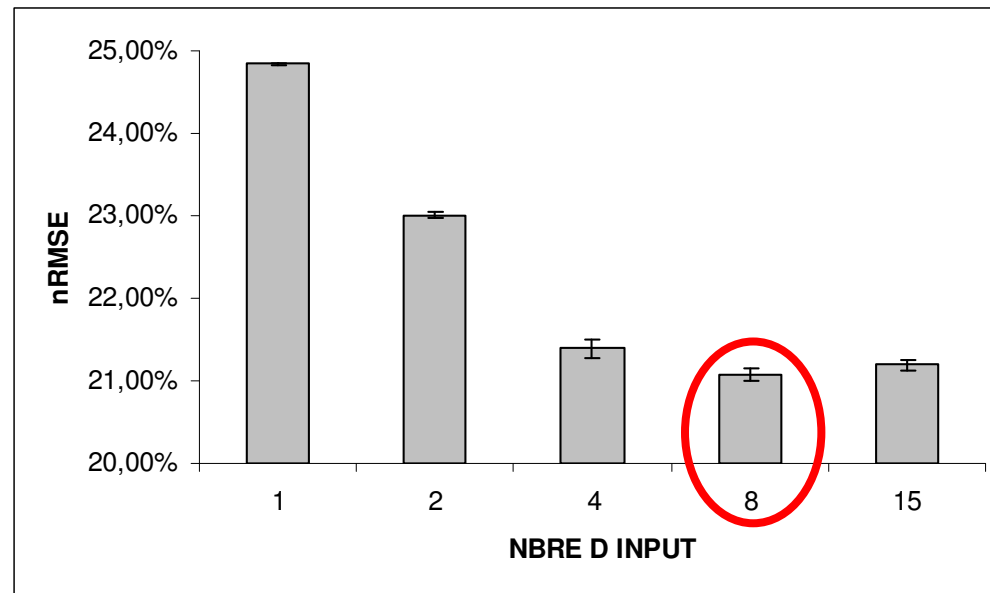


Maintenant qu'il a été montré qu'il n'était pas primordial de multiplier les layer cachés, voyons ce qu'il en est du nombre de neurones.

**Trois neurones semblent suffire pour la prédiction de l'irradiation solaire**

## Optimisation des ANN : le nombre de neurone d'entrée

12/28

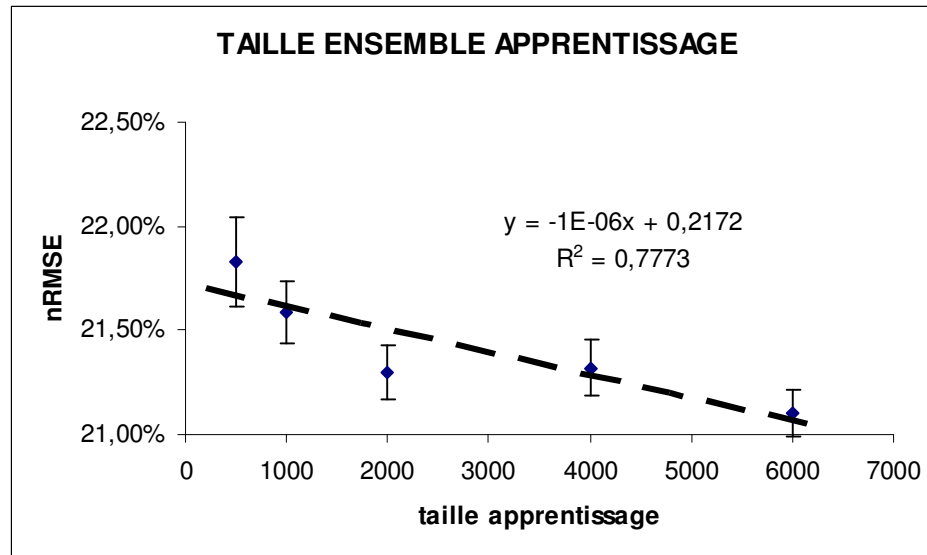


Il est important de corrélérer le nombre de neurones cachés au nombre de donnée d'entrée (input)

**huit données d'entrée semblent suffire pour la prédiction de l'irradiation solaire  $[x(t)..x(t-7)]$**

## Optimisation des ANN : la taille de l'échantillon d'apprentissage

13/28

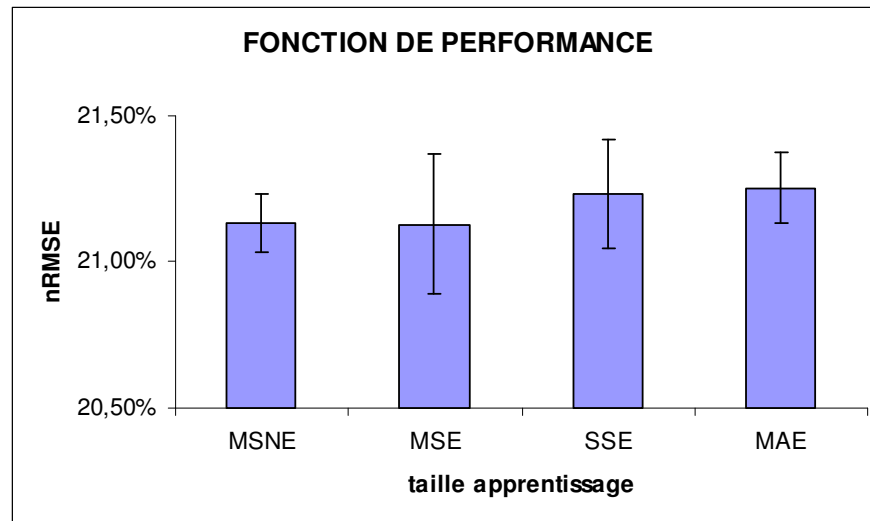


Plus la taille de l'ensemble d'apprentissage est importante, plus le sur-apprentissage est minimisé

**Il semble important de garder tout l'historique de la série pour entraîner le réseau. Ce résultat rend le test de déplacement de l'ensemble d'apprentissage caduque**

## Optimisation des ANN : la fonction de performance

14/28

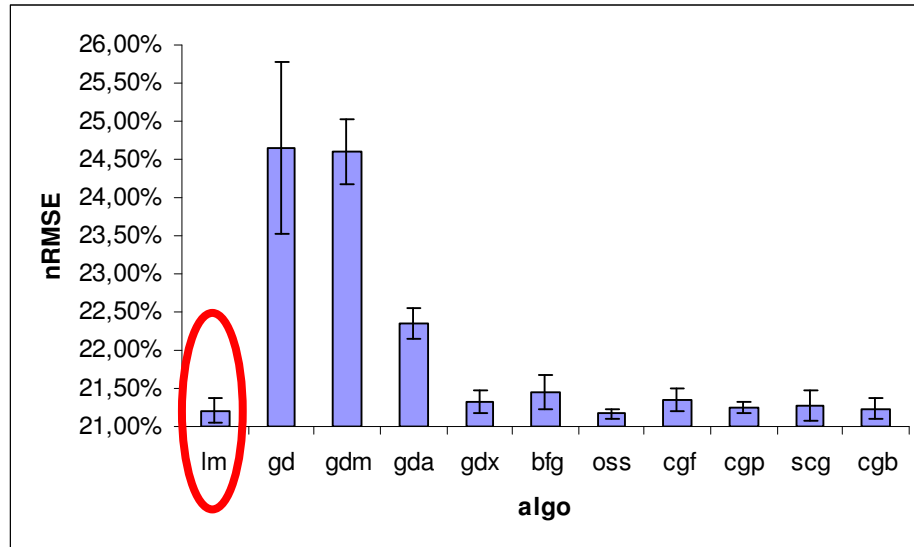


La fonction de coût utilisée par les algorithmes d'apprentissage est sans doute un paramètre secondaire, mais il est nécessaire de l'étudier

**La fonction MSE est choisie comme étant la meilleure**

## Optimisation des ANN : l'algorithme d'apprentissage

15/28

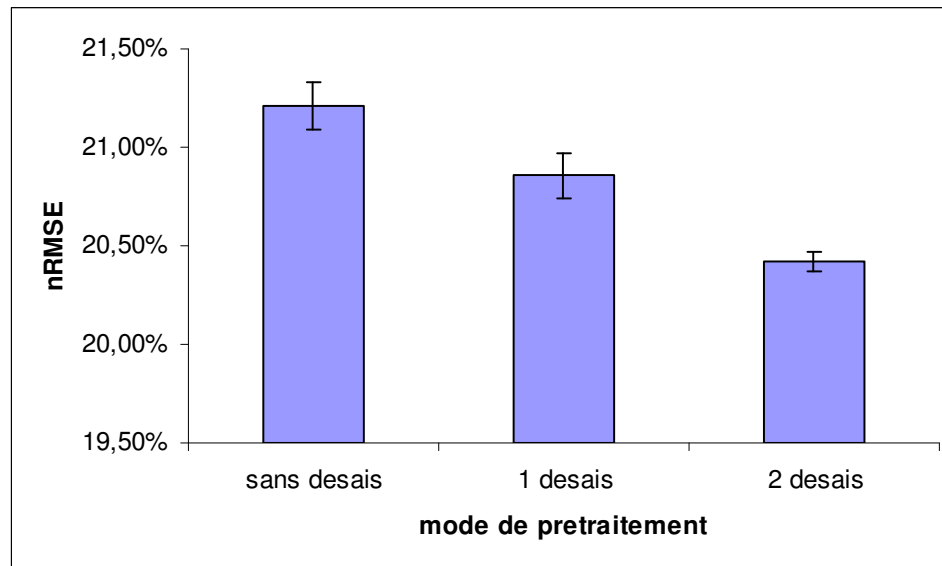


Tous les algorithmes présentés ci-contre sont paramétrables par de nombreux éléments, seul l'étude macroscopique, après optimisation intra-algo est présentée.

**L'algorithme de Levenberg-Marquardt est choisi pour entraîner le réseau**

## Optimisation des ANN : la stationnarisation des données

16/28



Comme lors de l'étude de processus ARMA, il semble intéressant d'étudier la stationnarisation des données d'entrée. La première dessaisonalisation est dite « physique » et prend en compte l'irradiation hors atmosphère. La deuxième est dite « rigide » et est purement statistique.

**Comme dans le cas des processus ARMA la stationnarisation semble apporter une diminution de l'erreur de prédiction**



# Sommaire

17/28

- Echelle journalière, optimisation des ANN
- Echelle journalière, comparaison ANN Vs ARMA
- Echelle horaire, comparaison ANN Vs ARMA
- Perspectives

# Echelle journalière : trois prédicteurs naïfs

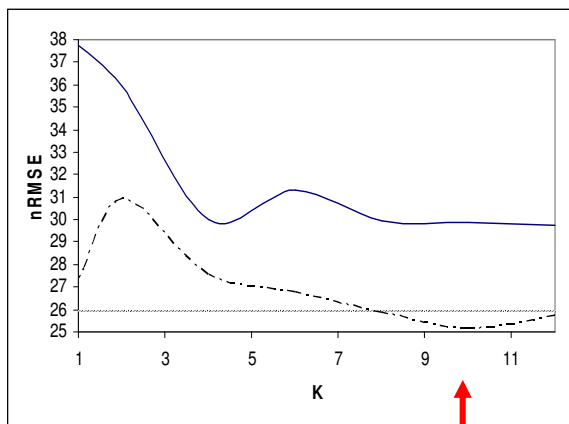
18/28

## Knn

On cherche parmi les  $n$  informations celles qui minimisent la quantité distance

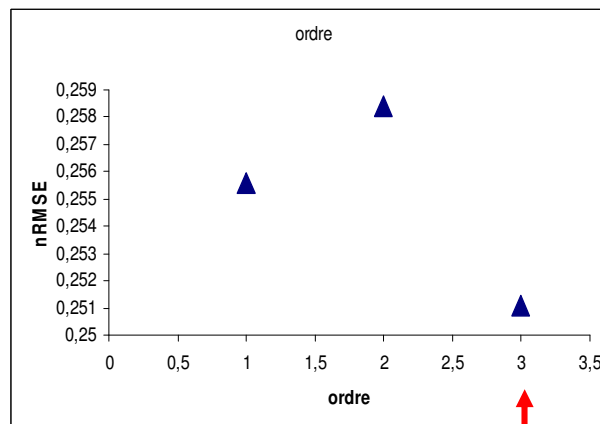
$$\text{Min}_r (d) = d(X_n, X_{n-r}) + d(X_{n-1}, X_{n-r-1}) + \dots + d(X_{n-k}, X_{n-r-k})$$

La grandeur  $r_0$  satisfaisant la relation ci-dessus nous donne le résultat suivant :  $X_{p+1} = X_{n-r_0+1}$



## Chaine de Markov

Avec l'hypothèse markovienne et l'acquisition de matrice de transfert basée sur les probabilités conditionnelles de la série, il est possible par produit matriciel de connaître la classe de l'élément  $x(t+1)$



## inférences Bayésiennes

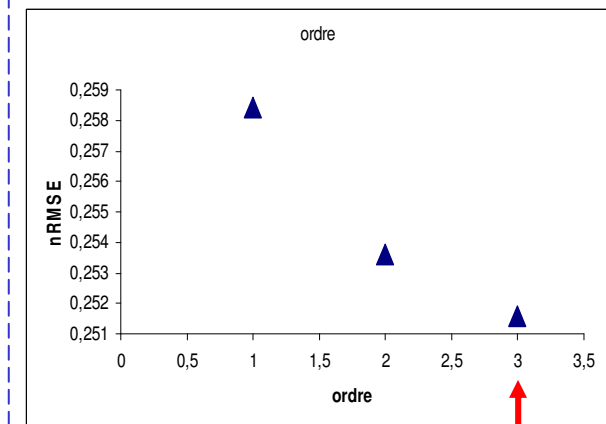
Comment classer un nouvel exemple en fonction d'un ensemble d'exemples pour lesquels on connaît la classe ?

Soit un exemple  $d = (d_1, \dots, d_n)$  et  $c$  classes  $k = 1, \dots, c$

$$\text{Classe}(d) = \underset{k}{\text{argmax}} \prod_i \hat{P}(d_i|k) \cdot \hat{P}(k)$$

proportion d'exemples  $d_i$  parmi ceux de la classe  $k$

proportion d'exemples de la classe  $k$



# Caractéristiques du réseau choisi pour la prédiction de la ressource solaire

1. Type de réseau : PMC dans un premier temps
2. Prétraitement des données d'entrée : normalisation sur  $[-1,1]$  avec stationnarisation
3. Nombre de couches et de neurones cachés : 1 couche cachée avec trois neurones
4. Choix des fonctions de transfert : purlin=>radbas=>purlin
5. Nombre de neurones d'entrée huit
6. Algorithme d'apprentissage Levenberg-Marquardt
7. Mode online ou batch batch
8. Weight decay étudié avec algo
9. Moment étudié avec algo
10. Pas d'apprentissage étudié avec algo
11. Initialisation des poids pris aléatoire
12. Early stopping étudié avec algo
13. Fonction de coût MSE
14. Fonction de comparaison nRMSE
15. Duplication d'entrée non fait
16. Bruitage des données d'entrées déjà très bruitées
17. Répartition des données (apprentissage validation et test) non fait
18. Permutation du pool d'apprentissage, taille d'échantillon échantillon total (6000 événements)
19. Customisation => architecture, normalisation, fonction de transfert pas de changement

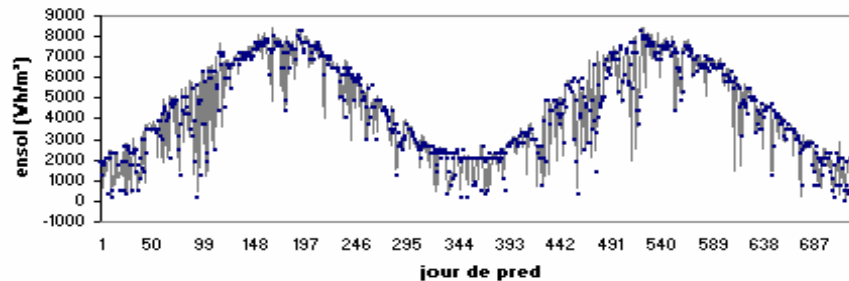
Réseau de départ : PMC 8x10x1 ; purlin  
x tansig x purlin ; LM (Cf Andranova)

A chaque étape on étudie un paramètre  
qui sert aux étapes suivantes

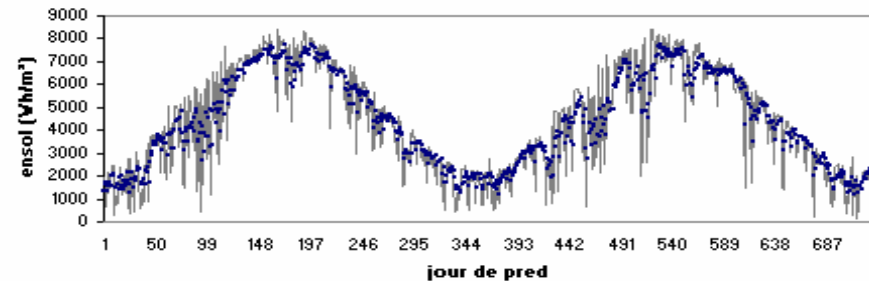
# Echelle journalière : ANN Vs ARMA

20/28

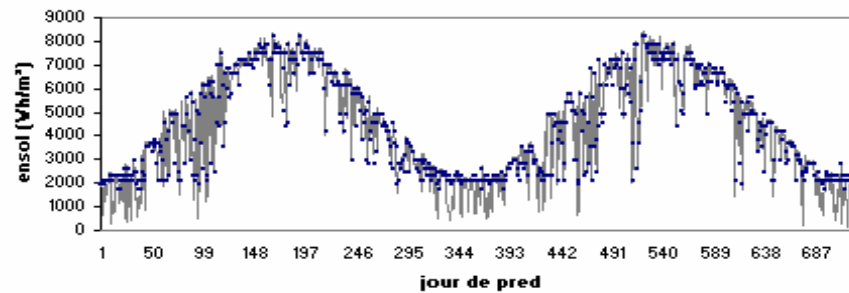
Bayes (ordre 3)



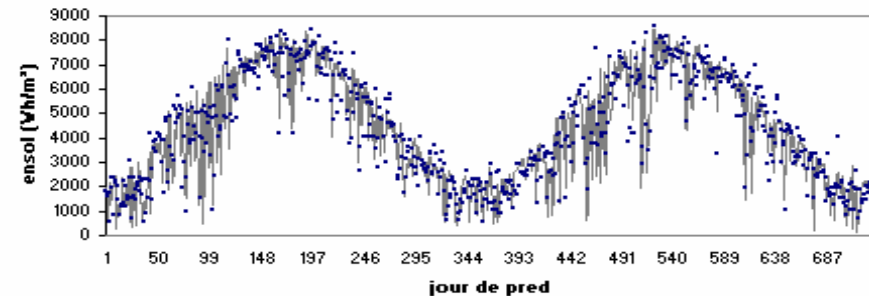
AR(8) sans stationnarsation



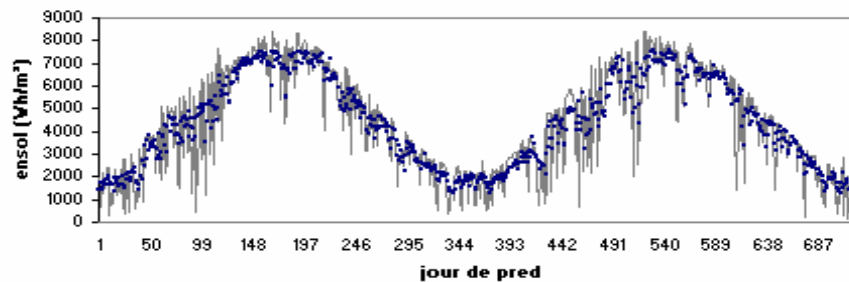
Markov (ordre 3)



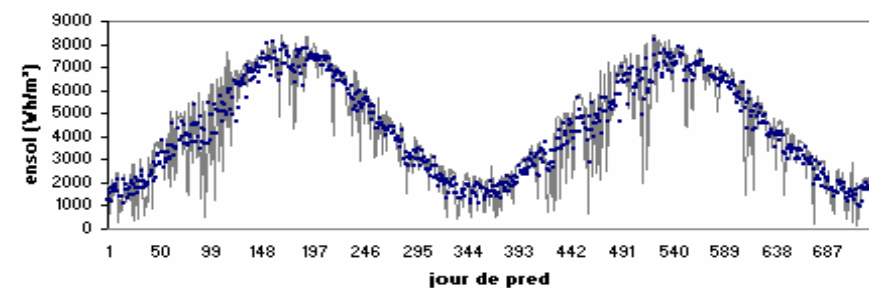
knn (ordre 10)



ANN



ARMA(2,2) avec stationnarsation



# Echelle journalière : ANN Vs ARMA:

21/28

	nRMSE (%)	
<b>Campo dell Oro (Ajaccio)</b> <b>apprentissage (01/01/71-31/12/87)</b> <b>prédiction (01/01/88-31/12/89)</b> <b>horizon 1 (pas journalier)</b>	Persistence	26
	Markov (ordre 3)	25,11
	Bayes (ordre 3)	25,16
	Knn (ordre 10)	25,2
	AR(8) sans desais	21,18
	ARMA(2,2) avec desais	20,31
	ANN [8-3-1] sans desais	20,97
	ANN [8-3-1] avec desais	20,17

la désaisonnalisation permet la prise en compte de la périodicité du phénomène

Les résultats obtenus dans le cas journaliers  
vont servir de base à notre étude horaire

# Sommaire

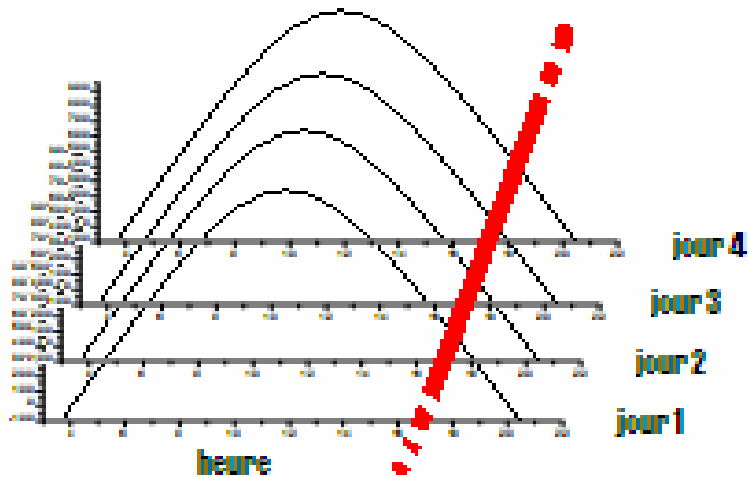
22/28

- Echelle journalière, optimisation des ANN
- Echelle journalière, comparaison ANN Vs ARMA
- Echelle horaire, comparaison ANN Vs ARMA
- Perspectives

## Echelle horaire : Décomposition en série journalière

nouvelle série construite avec le même créneau  
 horaire sur les différents jours

$$X^*(t) = \dots, X(\text{Jour1})_{15h00}, X(\text{Jour2})_{15h00}, X(\text{Jour3})_{15h00}, X(\text{Jour4})_{15h00}, \dots$$

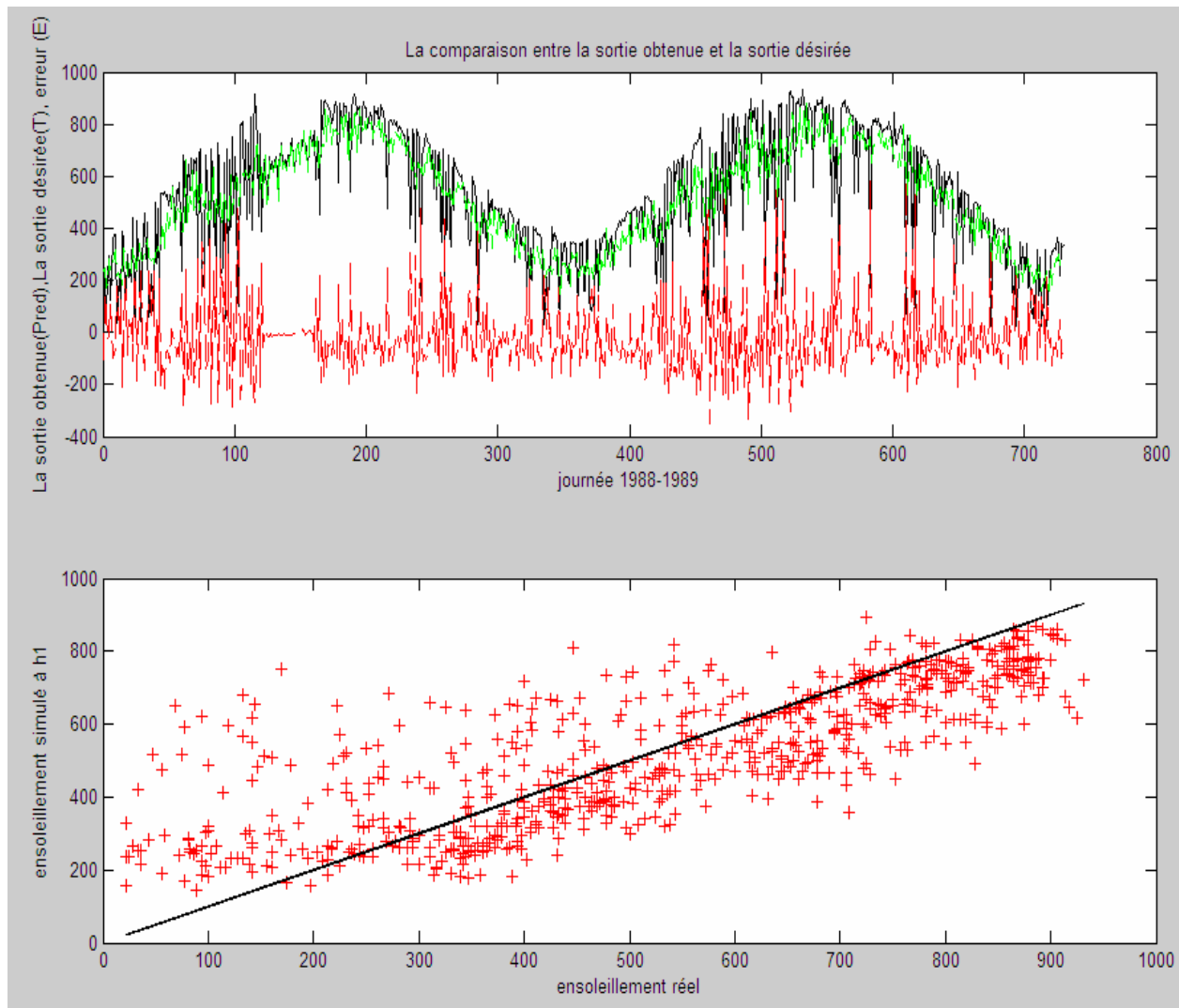


Comme lors de l'étude des  
 processus ARMA, nous effectuons  
 la décomposition en 9 prédicteurs  
 indépendants pour chaque créneau  
 horaire d'ensoleillement.

Les simulations à horizon 24h, concernant l'estimation du meilleur  
 réseau ANN, n'ont pas été effectuées dans le cas horaire. Comme il y a  
 transformation de la série horaire en 9 séries journalières, les résultats de  
 l'étude quotidienne ont été réutilisés

# Echelle horaire : ANN Vs ARMA

24/28



Exemple de simulation de  
la série 14h00

*En haut :*

*Noir=réel*

*Vert=prédiction*

*Rouge=erreur commise*

*En bas :*

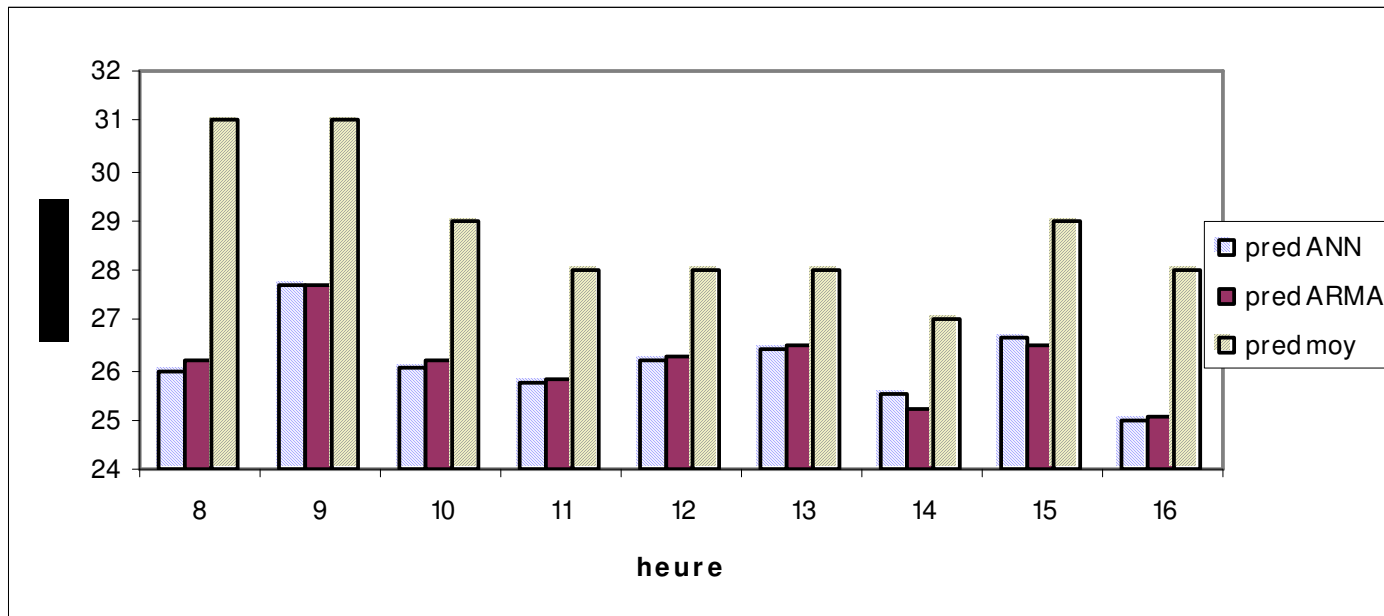
*Abscisse= réel*

*Ordonné= prédit*



# Echelle horaire : ANN Vs ARMA

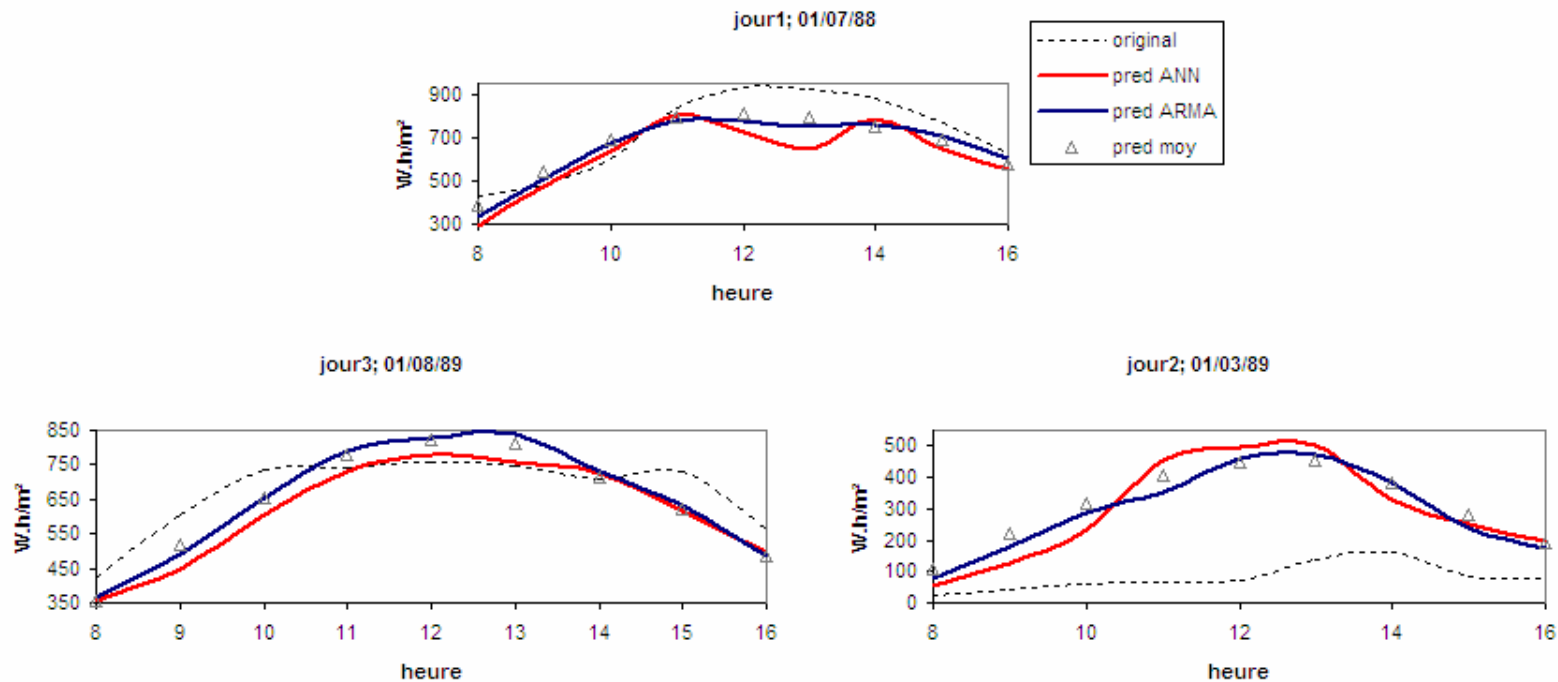
25/28



La prédiction moyenne est la moyenne du jour et de l'heure choisie sur les 17 ans d'apprentissage de la série

# Echelle horaire : ANN Vs ARMA

26/28



Difficulté de prédiction de la ressource solaire :

Trois cas de figure:

- jour1* : pas d'entrée nuageuse à 12h00
- jour3* : journée couverte imprévisible ou erreur de mesure ?
- jour2* : cas favorable en plein été

# Sommaire

27/28

- Echelle journalière, optimisation des ANN
- Echelle journalière, comparaison ANN Vs ARMA
- Echelle horaire, comparaison ANN Vs ARMA
- Perspectives

# Perspectives et conclusion

28/28

- Finir l'étude de tous les paramètres ANN dans le cas journalier
- Faire une étude d'optimisation des 9 ANN dans le cas horaire
- Prédire dans le cas horaire sans décomposition
- Quantifier la différence entre les RBFN et l'ANN retenu lors de l'étude
- Essayer d'autres types d'ANN que perceptron multicouche (SOM, Opfield....)
- Couplage avec données géographiques (SIG)
- Couplage avec données exogènes de type pression et température (metars, météosat)
- Bibliographie à documenter
- Nouveaux livres à analyser
- Valider le code Matlab
- Continuer l'étude des trois méthodes naïves
- on a presque le même nRMSE pour horizon 1 ou 31 => la stationnarisation induit une résultante quasi « bruit blanc ».
- .....