

# Chapitre IV

## Estimation des paramètres du modèle autorégressif ARMA par les algorithmes génétiques

### IV.1 Introduction

L'Algorithme Génétique est une approche d'optimisation évolutive basée sur des algorithmes de recherche aléatoires, développée par Holland en 1975 [129]. L'AG est une méthode d'optimisation combinatoire très répandue en raison de sa robustesse pour les problèmes complexes et non linéaires. Elle présente de nombreux avantages par rapport aux autres méthodes d'optimisation classiques. Afin d'obtenir de meilleures solutions, différents étapes de calcul telles que la sélection, la mutation et le croisement ont été implémentés dans l'algorithme. L'algorithme peut facilement converger vers une bonne solution plus rapidement que les autres approches classiques.

Les étapes de base du principe de fonctionnement de l'AG sont les suivantes : 1) initialisation d'une population de solutions possibles, puis 2) application des opérateurs de sélection, de croisement et de mutation, respectivement. La fonction d'évaluation est calculée pour chaque solution candidate. Après élimination de la mauvaise solution de la population existante, une nouvelle population est créée à nouveau à l'aide des opérateurs d'AG et le mécanisme de travail est maintenu jusqu'à ce que le critère soit satisfait [130].

Dans ce chapitre, nous proposons une méthode de prédiction intégrant les algorithmes génétiques et la méthode de prédiction à moyenne mobile autorégressive (ARMA) afin de tirer profit de la force intrinsèque des deux modèles à savoir ARMA et les algorithmes génétiques. L'approche est proposée pour prédire l'irradiation solaire globale (Global Horizontal Irradiance GHI) dans deux sites de test qui sont (Brasilia et Ny-Alesund).

### IV.2 Principe de fonctionnement des algorithmes génétiques

Les algorithmes génétiques manipulent un ensemble de points dans l'espace de recherche, appelé population d'individus. Chaque individu ou chromosome représente une solution possible du problème posé. Il est constitué d'éléments, appelés gènes, dont les valeurs sont appelées allèles. Les algorithmes génétiques font évoluer cette population

d'individus par générations successives, en utilisant des opérateurs inspirés de la théorie de l'évolution qui sont : la sélection, le croisement et la mutation. De génération en génération, la force des individus de la population augmente et après un certain nombre d'itérations, la population est entièrement constituée d'individus tous forts, soit de solutions quasi-optimales du problème posé [131].

### IV.3. Vocabulaire des algorithmes génétiques

Dû au fait que les AG s'appuient à la fois sur la génétique naturelle et l'intelligence artificielle, les terminologies utilisées pour les AG dans la littérature sont un mélange de naturel et de l'artificiel. Les termes usuels appliqués dans les AG sont résumés dans le tableau ci-dessous.

TABLEAU IV.1 Vocabulaire des AG [132].

Algorithmes Génétiques	Significations
<b>Chromosome (individu)</b>	Solution
<b>Gènes (variable)</b>	Partie de la solution
<b>Lieu</b>	Position du gène
<b>Allèles</b>	Valeur du gène
<b>Phénotype</b>	Solution décodée
<b>Génotype</b>	Solution encodée

### IV.4 Structure de l'algorithme génétique

L'implémentation d'un AG est spécifique au problème à résoudre. Pour l'utiliser, il faut disposer des cinq éléments suivants [133]:

- Un principe de codage des éléments de la population, qui consiste à associer à chacun des points de l'espace d'état une structure de données, la qualité de ce codage des données conditionne le succès des algorithmes génétiques ; bien que le codage binaire a été très utilisé à l'origine, les codages réels sont désormais largement exploités, notamment dans les domaines applicatifs pour l'optimisation de problèmes à variables réelles.
- un mécanisme de génération de la population initiale qui doit être capable de produire une population d'individus non homogène servant de base pour les générations futures; le choix de la population initiale est important, car il influence la rapidité de la convergence vers l'optimum global; dans le cas où l'on ne dispose que de peu

d'informations sur le problème à résoudre, il est essentiel que la population initiale soit répartie sur tout le domaine de recherche.

- une fonction à optimiser : celle-ci retourne une valeur appelée *fitness* ou fonction d'évaluation de l'individu.
- des opérateurs permettant de diversifier la population au cours des générations et d'explorer l'espace d'état; l'opérateur de croisement recompose les gènes d'individus existant dans la population alors que l'opérateur de mutation garantit l'exploration de l'espace d'état.
- des paramètres de dimensionnement, représentés par la taille de la population, le nombre total de générations, ou le critère d'arrêt, ainsi que les probabilités d'application des opérateurs de croisement et de mutation.

Un algorithme génétique est défini par un cycle de population et fait intervenir trois facteurs importants : fitness, croisement, mutation (figure VI.1). Un cycle représente le passage d'une population à la génération suivante soit l'évolution génétique d'une population.

L'organigramme fonctionnel de la figure VI.2 illustre la structure de l'algorithme génétique. Les diverses phases et les mécanismes associés à chacune d'entre elles seront présentés dans les sections suivantes.

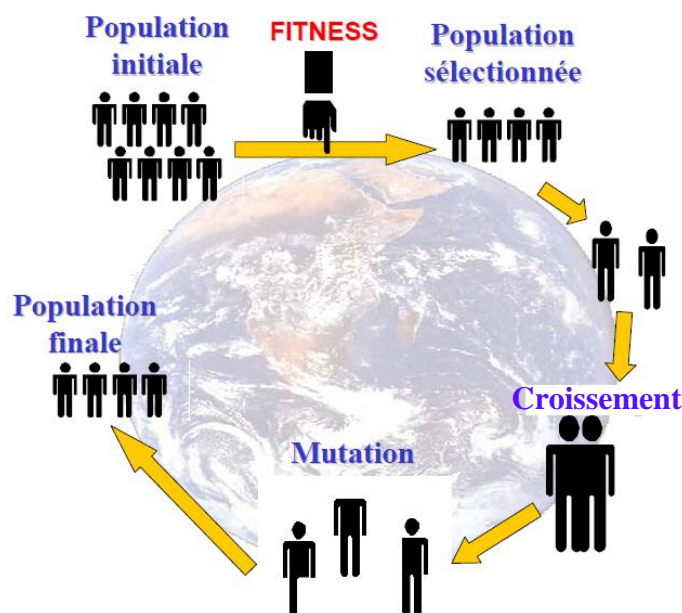


FIGURE VI.1 Illustration d'un cycle dans un algorithme génétique.

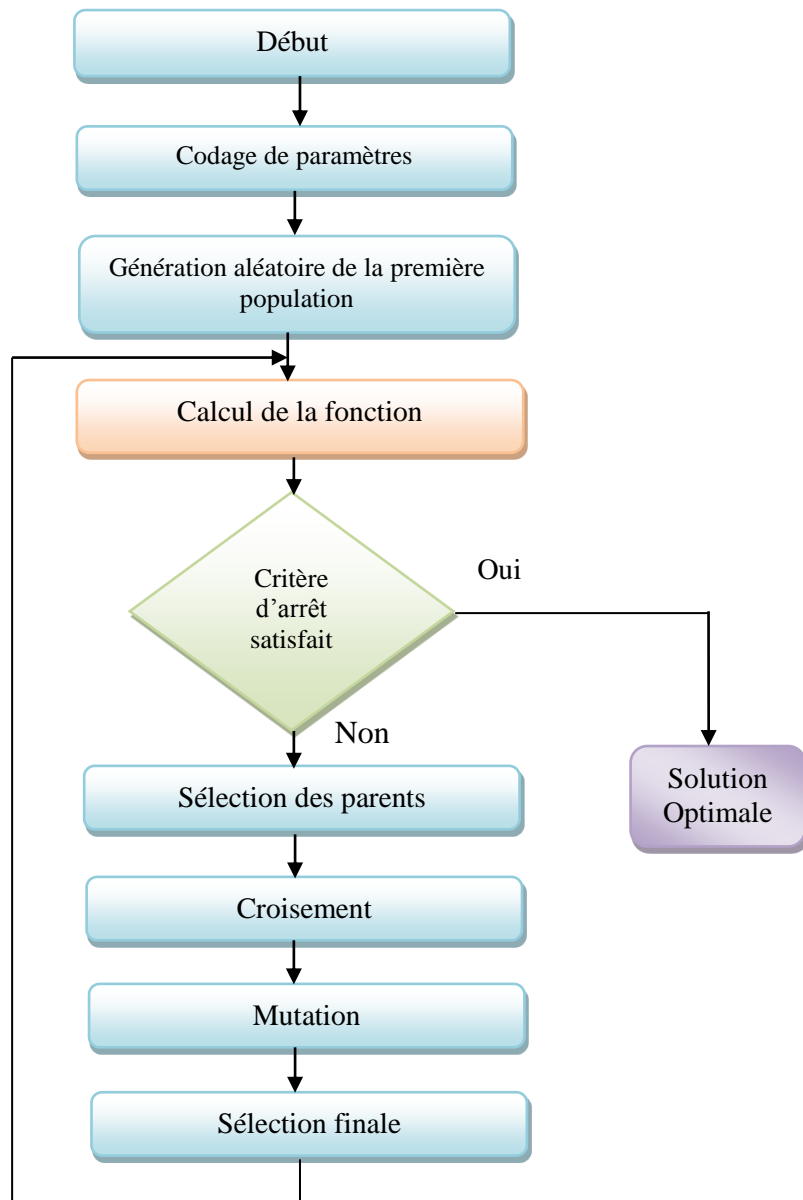


FIGURE IV.2 Organigramme général de l'AG [134].

#### IV.4.1 Le codage

Dans une première phase, il faut représenter les différents états possibles de la variable dont on cherche la valeur optimale sous une forme utilisable par un AG : c'est le codage. Chaque paramètre d'une solution est assimilé à un gène, toutes les valeurs qu'il peut prendre sont les allèles de ce gène. Un chromosome est une suite de gènes, on peut choisir de regrouper les paramètres similaires dans un même chromosome et chaque gène sera repérable par sa position. Chaque individu est représenté par un ensemble de chromosomes.

Il existe principalement deux types de codage : le codage binaire et le codage réel :

#### IV.4.1.1 Codage binaire

Chaque gène dispose du même alphabet binaire  $\{0, 1\}$ . Chaque paramètre  $x_i$  situé dans un intervalle  $\{x_{\min}, x_{\max}\}$ , est associé à une chaîne binaire  $b_0, b_1, \dots, b_{L_{x_i}-1}$  définie sur  $L_{x_i}$  bits. A chaque chaîne correspond une valeur entière naturelle :

$$g_i = \sum_{i=0}^{L_{x_i}-1} 2^i b_i \quad (\text{IV.1})$$

Cette chaîne doit être décodée pour pouvoir calculer la valeur de la fonction coût (fonction d'évaluation) qui lui est associée. Le paramètre réel  $x_i$  de l'espace de recherche relatif à  $g_i$  est obtenu par interpolation linéaire :

$$x_i = x_{\min} + \frac{x_{\max} - x_{\min}}{2^{L_{x_i}-1}} \cdot g_i \quad (\text{IV.2})$$

#### IV.4.1.2 Codage réel

Cela peut-être utile notamment dans le cas où l'on recherche le maximum d'une fonction réelle. La représentation des gènes par des nombres réels arrive naturellement dans le cas d'optimisation de paramètres avec des variables comprises dans des domaines continus. La taille des chromosomes est la même que la taille du vecteur qui sera la solution du problème. Suite à cette approche, chaque gène représente une variable du problème.

Dans le codage binaire, l'utilisation de domaines vastes amène à une perte de précision des solutions.

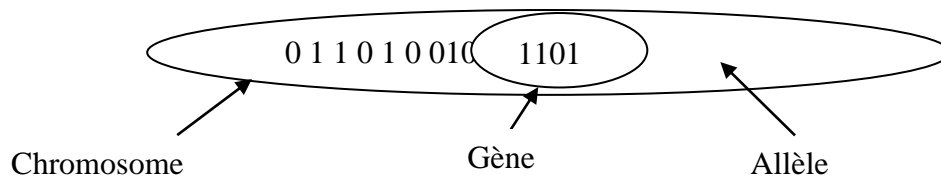


FIGURE IV.3 Structure d'un chromosome.

Autre avantage de l'utilisation des nombres réels est leur capacité à exploiter la gradualité (les changements survenus au niveau des variables impliquent des changements au niveau de la fonction) des fonctions qui ont des variables continues. Dans ce cas, un avantage très important est la capacité de réglage local de solutions, qui se produit beaucoup plus rapidement que dans le cas du codage binaire des algorithmes génétiques.

#### IV.4.2 Création de la population

La population initiale d'un AG représente la population de départ de l'algorithme. Elle doit contenir des chromosomes qui sont bien répartis dans l'espace des solutions pour fournir à l'AG un matériel génétique varié. Lors de la genèse de la population, l'ensemble dans lequel est effectuée la recherche de solutions doit être de dimension finie. La qualité et le choix du codage des chromosomes de la population est déterminant et conditionne le succès de la recherche pour la meilleure solution. Le choix de la taille de la population est aussi important que le codage, car une population trop grande entraînera une recherche exhaustive, tandis qu'une population trop petite ne permettra pas à l'algorithme d'atteindre une solution optimale.

#### IV.4.3 Évaluation de la population

Il s'agit d'un processus au cours duquel, les chromosomes de la population sont évalués par la fonction sélective (*fitness function*). C'est une fonction qui permet une évaluation de la population selon le critère de performance choisi. Elle permet ainsi de choisir les candidats les plus performants. Le critère de performance est lié à la nature du problème. Dans certains cas, la fonction sélective peut être établie comme une fonction de l'erreur. La plupart des algorithmes ont pour action de maximiser la valeur de la fonction sélective. Etant donné une fonction  $f$  réelle à une ou plusieurs variables, le problème d'optimisation sur l'espace de recherche  $E$  s'écrit de la manière suivante :

$$\max_{x \in E} f(x) \quad (\text{VI.5})$$

Dans beaucoup de problèmes, l'objectif est exprimé sous forme de minimisation d'une fonction coût  $h$  :

$$\min_{x \in E} h(x) \quad (\text{VI.6})$$

Le passage du problème de minimisation à un problème de maximisation est obtenu par transformation de la fonction  $h$  selon la relation suivante :

$$f(x) = \frac{1}{(1+h(x))} \quad (\text{VI.7})$$

### IV.5 Les mécanismes d'un AG

A partir d'une première population d'individus créée aléatoirement, les AG génèrent de nouveaux individus plus performants que leurs prédécesseurs en effectuant des opérations génétiques. Les AG utilisent des outils tels que la reproduction, le croisement et la mutation.

Ces outils sont basés sur des processus aléatoires. La reproduction est une version artificielle de la sélection naturelle, c'est un processus dans lequel chaque individu est copié en fonction des valeurs de la fonction d'évaluation. Le croisement est l'opérateur le plus dominant dans un AG, il permet à deux chaînes d'échanger des portions de leurs structures produisant ainsi de nouvelles chaînes. La mutation est un opérateur local qui est appliqué avec une très faible probabilité. Dans ce qui suit, nous allons discuter le rôle de ces paramètres.

#### IV.5.1 La sélection

Cet opérateur est peut-être le plus important puisqu'il permet aux individus d'une population de survivre, de se reproduire ou de mourir. En règle générale, la probabilité de survie d'un individu sera directement reliée à son efficacité relative au sein de la population.

Il existe plusieurs méthodes pour la reproduction. La méthode la plus connue et utilisée est la roue de loterie biaisée (*roulette wheel*) de [135]. Selon cette méthode, chaque chromosome sera dupliqué dans une nouvelle population proportionnellement à sa valeur d'adaptation. L'inconvénient majeur de cette méthode repose sur le fait qu'un individu n'étant pas le meilleur, peut tout de même dominer la sélection. Elle peut aussi engendrer une perte de diversité par la domination d'un super individu. Un autre inconvénient c'est sa faible performance vers la fin quand l'ensemble des individus se ressemblent.

Brièvement, il existe d'autres méthodes, la plus connue étant celle du tournoi (*tournament selection*) : on tire deux individus aléatoirement dans la population et on reproduit le meilleur des deux dans la nouvelle population. On refait cette procédure jusqu'à ce que la nouvelle population soit complète. Cette méthode donne de bons résultats. Toutefois, aussi important que soit la phase de sélection, elle ne crée pas de nouveaux individus dans la population. Ceci est le rôle des opérateurs de croisement et de mutation.

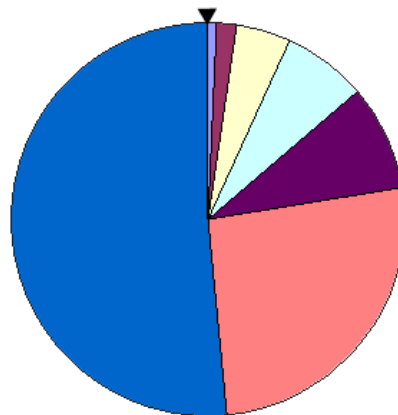


FIGURE IV.4 Schéma d'un exemple d'une roue de loterie.

### IV.5.2 Fonction d'adaptation

Ce sont les principes de survie et de reproduction décrits par Charles Darwin [136], où une population ayant une grande variété engendrera, de génération en génération, une contenance des individus dont le génotype se traduit par une meilleure adaptation. Pour cela, il nous faut une fonction dite d'adaptation qui va être dans notre cas la fonction objectif elle-même [137].

### IV.5.3 Croisement

L'opérateur de croisement permet la création de nouveaux individus selon un processus très simple. Il permet donc l'échange d'information entre les chromosomes (individus). Tout d'abord, deux individus, qui forment alors un couple, sont tirés au sein de la nouvelle population issue de la sélection. Puis un (potentiellement plusieurs) site de croisement est tiré aléatoirement.

Il existe différents types de croisements pour un algorithme génétique classique. Sur la figure IV.4, les trois principaux types de croisement sont présentés (en un point, en deux points et uniforme). Cette opération est contrôlée par une probabilité  $p_c$  souvent supérieure à 0.5 [138].

#### IV.5.3.1 Croisement en un point

Dans ce type de recombinaison, un point de croisement est choisi au hasard. À partir de ce point de coupure, les données sont échangées entre les individus afin de créer des enfants (voir figure IV.5.a).

#### IV.5.3.2 Croisement en deux points

Dans ce cas, on choisit au hasard deux points de coupure à partir desquels on procède à l'échange de données, afin de créer des enfants comme le montre la figure IV.4.

#### IV.5.3.3 Croisement uniforme

La mise en œuvre de ce procédé est très simple, elle consiste à définir de manière aléatoire un "masque", c'est-à-dire une chaîne de bits de même longueur que les chromosomes des parents sur lesquels il sera appliqué. Ce masque est destiné à savoir, pour chaque *locus* (position sur le chromosome), de quel parent le premier fils devra hériter le gène.

Si face à un *locus*, le masque présente un 0, le fils héritera le gène du parent n°1, s'il présente un 1 il en héritera du parent n°2. La création du fils n°2 se fait de manière symétrique: si pour un gène donné, le masque indique que le fils n°1 devra recevoir celui du

parent n° 1 alors le fils n° 2 le recevra du parent n°2, et si le fils n°1 le reçoit du parent n°2 alors le fils 2 le recevra du parent n°1, comme illustrée sur la figure IV.4.c [139].

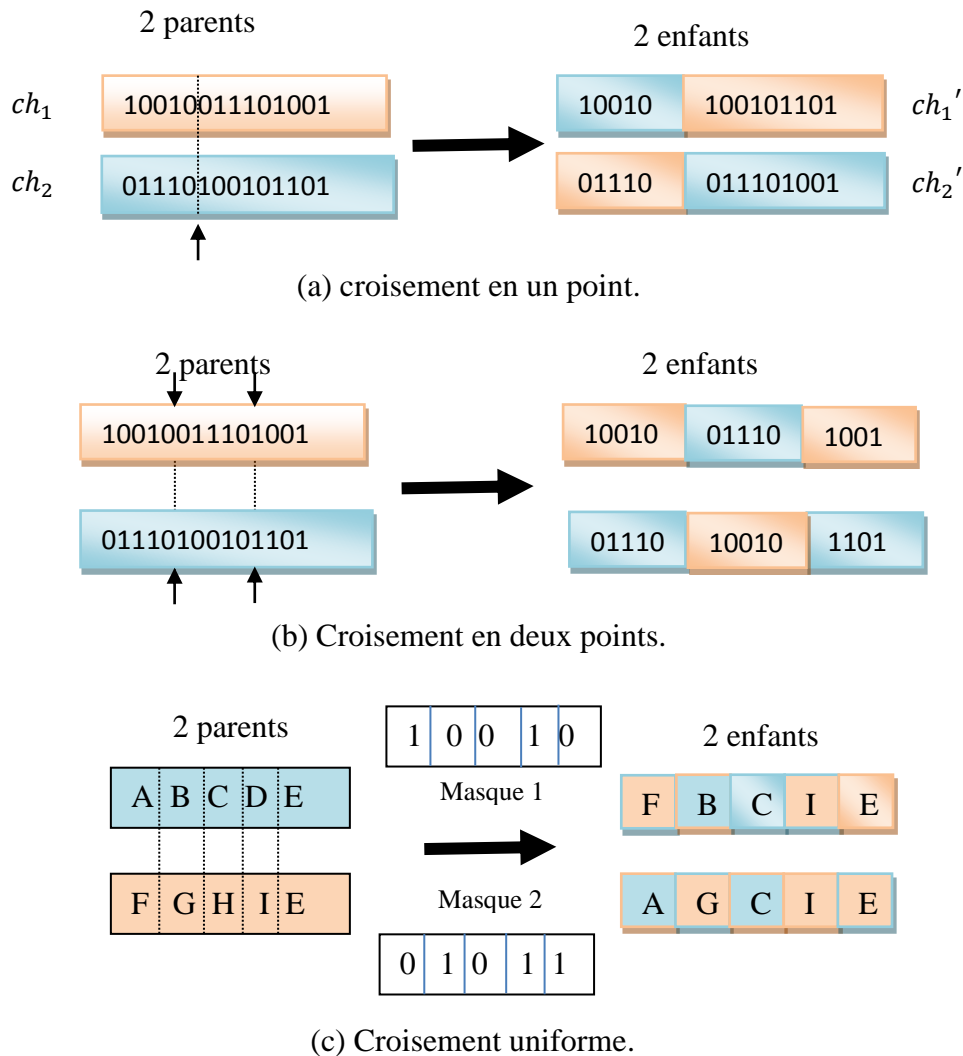


FIGURE IV.5 Exemples d'opérations de croisement [140].

#### IV.5.4 Mutation

Durant la mutation, des individus sont altérés de façon aléatoire. Ces altérations sont la plupart du temps petites. Selon la représentation des variables, différentes méthodes de mutation peuvent être employées. On a ainsi les méthodes pour des variables à valeurs réelles et les méthodes pour des variables binaires.

##### IV.5.4.1 Mutation à valeurs réelles

Ici la mutation est réalisée en rajoutant aux individus des réels de très petites valeurs créées de manière aléatoire.

#### IV.5.4.2 Mutation binaire

La mutation est définie ici comme étant l'inversion d'un bit dans un chromosome (voir figure IV.6). Le choix du bit à inverser est fait de manière aléatoire. Les mutations empêchent la recherche de la solution de stagner, en apportant une légère variation. Elles permettent aussi d'assurer une recherche globale plutôt que locale.

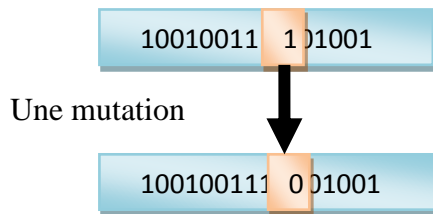


FIGURE IV.6 Mutation binaire.

#### IV.5.5 Insertion

La population issue des opérations génétiques sera insérée dans l'ancienne population d'une manière à garder les individus ayant la fonction fitness la plus grande.

#### IV.5.6 Critère d'arrêt

Les opérations génétiques précédentes seront exécutées autant de fois que nécessaire, et à chaque fois on obtient une nouvelle population ayant des individus de plus en plus adaptés. Le critère d'arrêt peut être choisi par plusieurs façons, soit par le test de la fonction coût, soit par le nombre de générations, soit par le test de changements dans la population. Si le nombre de générations est fixé, on obtient un temps d'exécution fixe mais le programme peut ne pas s'arrêter à la précision souhaitée. Si l'on connaît la valeur maximale de la fonction fitness, le programme s'arrêtera une fois il atteint cette valeur, on obtient la précision souhaitée, mais cela peut exiger un temps de calcul non contrôlé.

### IV.6 Influence des paramètres des AG

Les opérations de l'algorithme génétique sont guidées par un certain nombre de paramètres fixés à l'avance. La valeur de ces paramètres influe sur la réussite d'un algorithme génétique. Ces paramètres sont les suivants :

#### IV.6.1 La taille de la population $N$

Si  $N$  est trop grand le temps de calcul de l'algorithme peut s'avérer très important, et si  $N$  est trop petit, il peut converger trop rapidement vers un mauvais chromosome (pas de diversité).

**IV.6.2 La longueur des individus S**

La longueur de chaque individu représente la précision des solutions. Plus  $S$  est grand, plus on gagne de précision, et on perd dans le temps de calcul et la taille mémoire nécessaire.

**IV.6.3 La probabilité de croisement  $p_c$** 

Plus elle est élevée, plus la population subit de changements importants. Les valeurs généralement admises sont comprises entre 0,5 et 0,9.

**IV.6.4 La probabilité de mutation  $p_m$** 

Ce taux est généralement faible, puisque un taux élevé risque de conduire à une solution sous optimale. Nous avons examiné les principales caractéristiques des algorithmes génétiques qui les rendent des outils simples et efficaces pour explorer un espace de performance donné.

Dans ce qui suit, les AG seront utilisés pour estimer les paramètres d'un modèle ARMA et améliorer ainsi sa performance pour la prédiction de l'irradiation solaire globale (GHI).

**IV.7 Optimisation des modèles ARMA**

Les modèles Box-Jenkins (modèles Autoregressifs) sont des méthodes de prédiction très populaires en raison de leurs bonnes capacités de prédiction pour différents types de données. [141], ont mis au point un nouveau modèle appelé ARMA (modèle de moyenne mobile autorégressif) en 1994, qui est appliqué aux problèmes de prédictions dans le cas des séries temporelles stationnaires.

Le modèle ARMA est une combinaison des modèles AR et MA. Dans le cas des séries non stationnaires, les données doivent d'abord être transformées en une forme stationnaire en introduisant l'opérateur de différentiation. C'est la première condition pour construire un modèle à moyenne mobile autorégressif (ARMA). Ici, la valeur future d'une variable est supposée être une fonction linéaire de plusieurs facteurs, plus précisément, les observations passées et les erreurs aléatoires. Généralement, un modèle ARMA est défini par l'équation suivante comme il a été détaillé dans le chapitre III:

$$y_t = \varepsilon_t + (a_1 y_{t-1} + a_2 y_{t-2} + \dots + a_p y_{t-p}) + (b_0 \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}). \quad (IV.8)$$

où  $y_t$  est la valeur réelle,  $\varepsilon_t$  est une erreur aléatoire à la période  $t$ .  $a_i$  ( $i = 1, 2, \dots, p$ ) et  $b_j$  ( $j =$

0,1,2, ... q) sont les paramètres du modèle. L'ordre d'un modèle ARMA est généralement désigné par la notation ARMA (p, q), où p est l'ordre de la partie autorégressive (AR), q est l'ordre du processus de moyenne mobile (MA).

Le choix du meilleur modèle ARMA peut être effectué en trois étapes principales: premièrement, l'étape d'identification qui choisit le modèle le mieux adapté à la fonction d'auto-corrélation (ACF) et la fonction d'auto-corrélation partielle (PACF) de la série chronologique. Si les ACF de la série temporelle disparaissent rapidement, nous considérons la série temporelle comme étant stationnaire. Inversement, si la courbe ACF décline très lentement, la série chronologique peut être considérée comme non stationnaire.

La figure IV.10 illustre la sélection du meilleur modèle ARMA basé sur les comportements d'ACF et de PACF. La deuxième étape consiste à estimer les paramètres du modèle et enfin appliquer la méthode de Box-Jenkins pour vérifier la validité statistique du modèle estimé.

En règle générale, cela ne suffit pas pour déterminer l'ordre du modèle en utilisant uniquement le PACF. Ainsi, nous utilisons différents critères pour choisir l'ordre optimal. Ici, nous utilisons uniquement le critère d'information Akaike (AIC) [142], considéré comme le plus utilisé, qui est défini comme suit:

$$AIC = \log(\nu) + 2m/N \quad (VI.11)$$

Où  $m = p + q$  et  $\nu$  représente la fonction de vraisemblance. Les résultats des calculs de l'AIC sont illustrés sur la figure IV.10.

#### **IV.7.1 Formulation d'estimation des paramètres de l'ordre d'ARMA à l'aide d'AG**

L'algorithme génétique est un outil de recherche puissant pour résoudre les problèmes d'optimisation de manière optimale. Il estime l'ordre et les paramètres du modèle ARMA comme illustré dans la figure IV.7. Pour commencer la procédure d'estimation, il faut d'abord définir une fonction objective et effectuer la recherche conformément à cette fonction objective.

Comme notre objectif est de produire un modèle ARMA optimisé, la fonction objective vise à réduire la différence entre les modèles réels et optimisés. Par conséquent, la fonction objective, appelée fonction de fitness, peut être définie comme [143]:

$$\text{Fitness} = \frac{1}{1 + \text{RMSE}} \quad (IV.9)$$

Où

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_0^N (y_t - \hat{y}_t)^2} \quad (\text{IV.10})$$

Ici, RMSE est l'erreur quadratique moyenne,  $y_t$  est la sortie réelle,  $\hat{y}_t$  est la sortie estimée et N est le nombre d'observations. Pour le modèle ARMA donné dans l'équation (1), le signal solaire est généralement enregistré avec un bruit blanc additif gaussien.

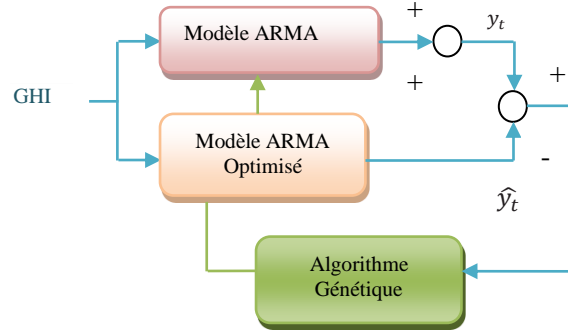


FIGURE IV.7 Architecture de l'AG pour l'estimation du modèle ARMA [139].

Les étapes de construction des modèles ARMA sont illustrées dans la figure VI.8.

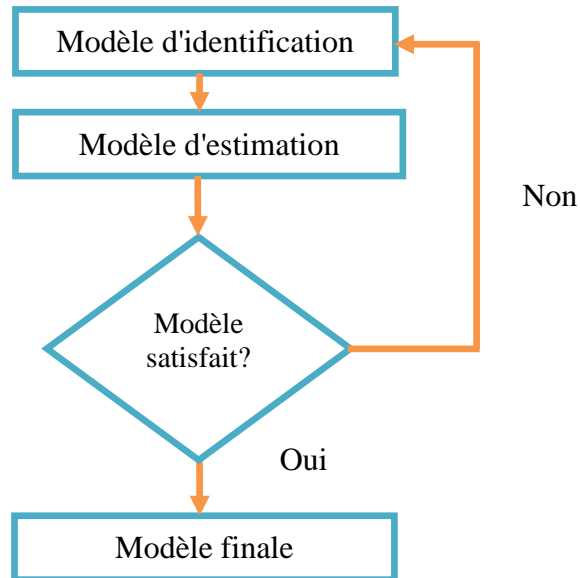


FIGURE IV.8 Les étapes de construction du modèle ARMA.

## IV.8 Résultats d'identification

Les données d'entrée de l'algorithme proposé sont divisées en deux parties: les paramètres liés à l'AG et les paramètres relatifs au modèle ARMA. Les paramètres liés à l'AG inclus dans notre simulation sont les suivants:

La taille de la population,  $N_p=80$ , la probabilité de croisement,  $P_c=0,65$ , la probabilité

de mutation,  $P_m=0,3$  pour le site de Ny\_Ålesund. En ce qui concerne le site de Brasilia, la taille de la population a été fixée à  $N_p = 65$ , la probabilité de croisement  $P_c = 0,7$ , la probabilité de mutation  $P_m = 0,3$ . Ici, l'analyse met l'accent sur le RMSE moyen et le nombre moyen de générations nécessaires pour la convergence.

En raison de la nature stochastique des AG, différentes exécutions sont effectuées avec un nombre arbitraire de générations aléatoires.

L'analyse liée au modèle ARMA concerne la détection d'ordre et l'optimisation des paramètres. L'analyse d'optimisation des paramètres est axée sur la convergence des paramètres du modèle ARMA. Dans l'algorithme proposé, la première étape de l'opération consiste à détecter le véritable ordre. Pour examiner l'efficacité de la méthode, différents tests ont été effectués pour détecter l'ordre réel avant d'estimer les paramètres opérationnels.

Le tableau IV.2 donne les résultats de prévision des modèles, selon les critères d'erreurs suivant: MAPE, RMSE, NMSE et  $R^2$ , le modèle AG-ARMA fonctionne mieux que l'ARMA classique.

TABLEAU IV.2 Résultats de la prédiction pour les deux sites de tests obtenus avec les modèles ARMA, AG-ARMA et leur combinaison avec la méthode OP. Les meilleurs résultats sont en gras.

<i>Méthodes</i>	<i>Sites</i>	<i>R<sup>2</sup></i>	<i>RMSE</i>	<i>MAPE</i>	<i>NMSE</i>
<b>ARMA</b>	Ny_Ålesund	85.57	47.46	19.56	0.13
	Brasilia	60.44	211.19	35.36	0.42
<b>AG_ARMA</b>	Ny_Ålesund	85.90	39.140	18.47	0.12
	Brasilia	61.92	199.71	33.94	0.38
<b>OP</b>	Ny_Ålesund	<b>85.92</b>	<b>38.24</b>	<b>18.15</b>	<b>0.11</b>
	Brasilia	<b>62.35</b>	<b>198.99</b>	<b>33.65</b>	<b>0.37</b>

#### IV.8.1 Interprétation des résultats

D'après les résultats expérimentaux obtenus, il est clair que le modèle ARMA estimé par AG donne de meilleurs résultats par rapport au modèle ARMA classique, ceci est en raison de sa structure dynamique, adaptative et objective. Séparément, les modèles ARMA n'ont pas été en mesure de montrer l'efficacité des prévisions GHI en raison de structures traditionnelles et de jugements subjectifs sur la détermination des paramètres ARMA. Les modèles combinés AG-ARMA répondent à ces lacunes en termes d'obtention de meilleures solutions et de prévisions cohérentes. Comme le montre la figure IV.9.

De plus, nous pouvons améliorer davantage les résultats de prévision par la

proposition d'une autre méthode basée sur la combinaison des sorties des deux modèles obtenus. Nous pouvons ainsi clairement observer que, pour chaque série temporelle de GHI, les méthodes de combinaison génèrent les plus petites erreurs, et par conséquent, les meilleures précisions en termes de mesures d'erreur. Nous trouvons donc que les méthodes combinées ont une précision de prédiction globalement supérieure à celle des méthodes individuelles basées sur la prédiction de GHI par ARMA et AG-ARMA.

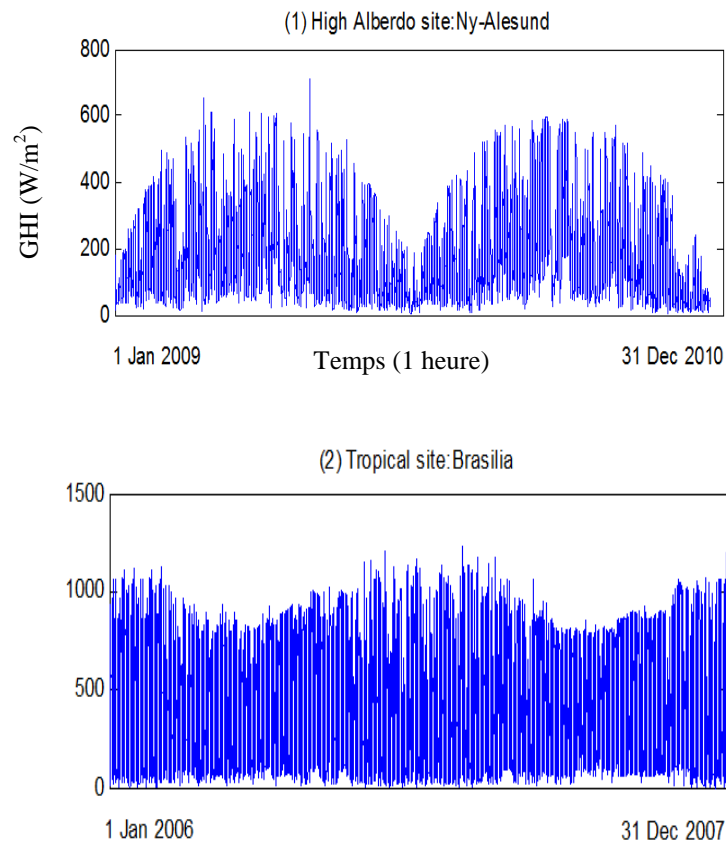
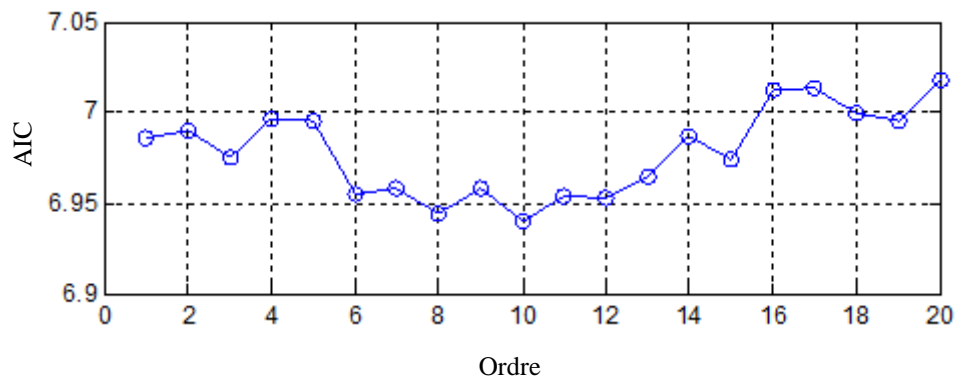
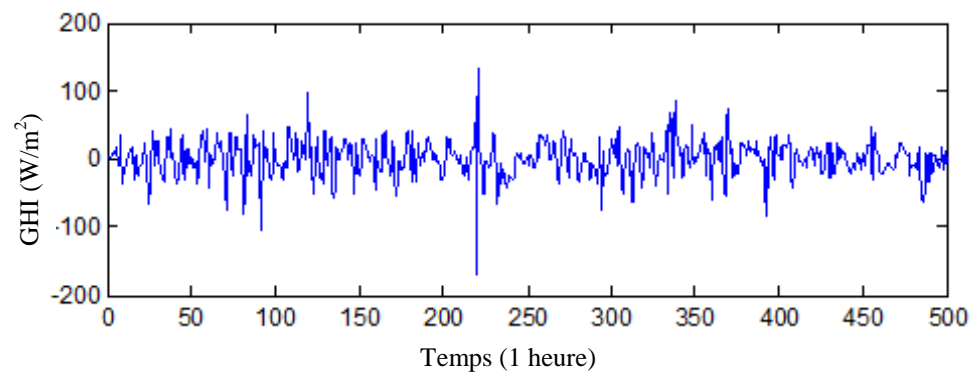
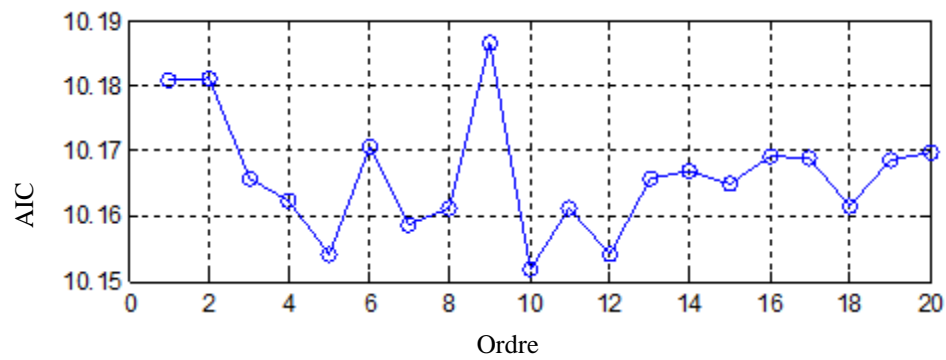
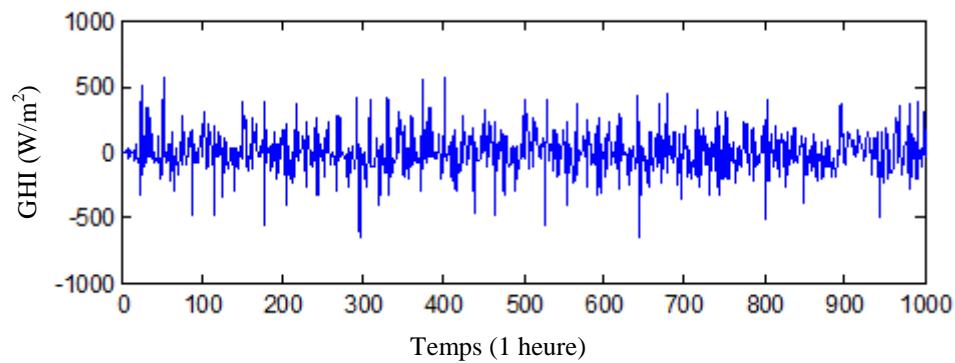


FIGURE IV.9 Série chronologique de données GHI sur deux années successives, pour les sites Brasilia et Ny-Ålesund.



(a)



(b)

FIGURE IV.10 Critère AIC pour les sites : (a) Brasilia, (b) Ny-Ålesund.

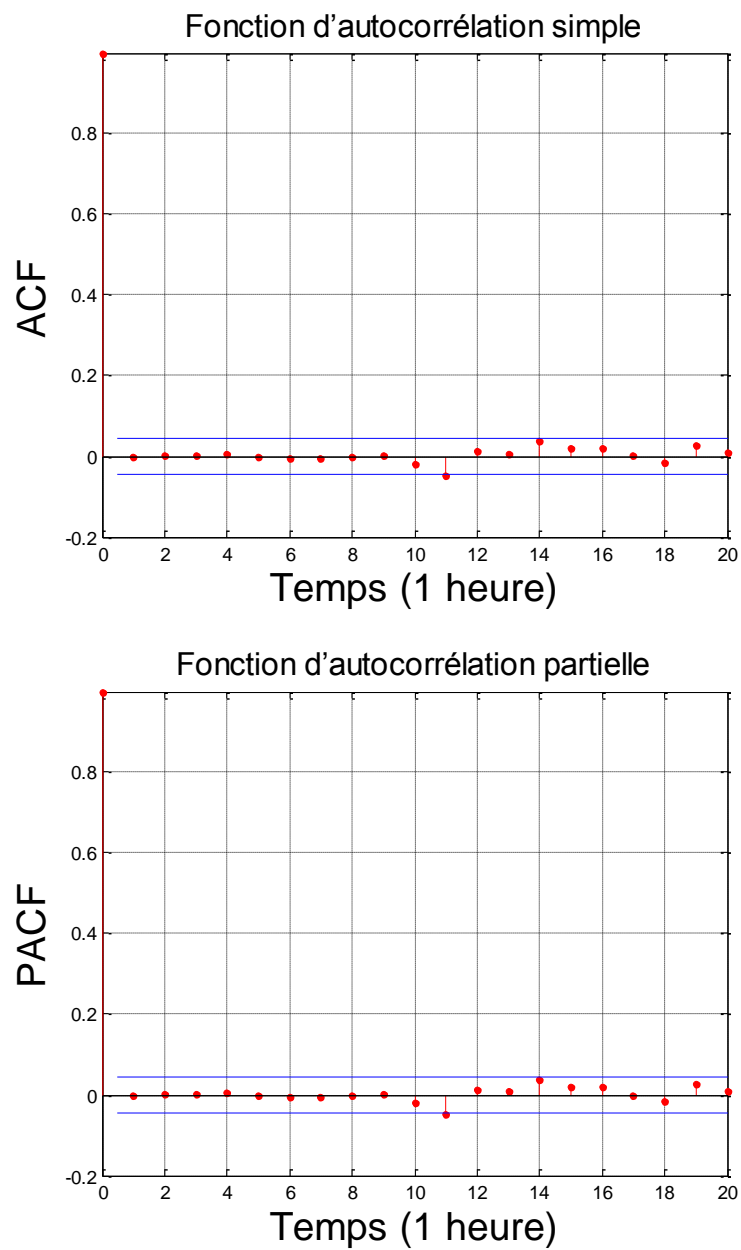


FIGURE IV.11 ACF et PACF des données solaire du site Ny-Ålesund.

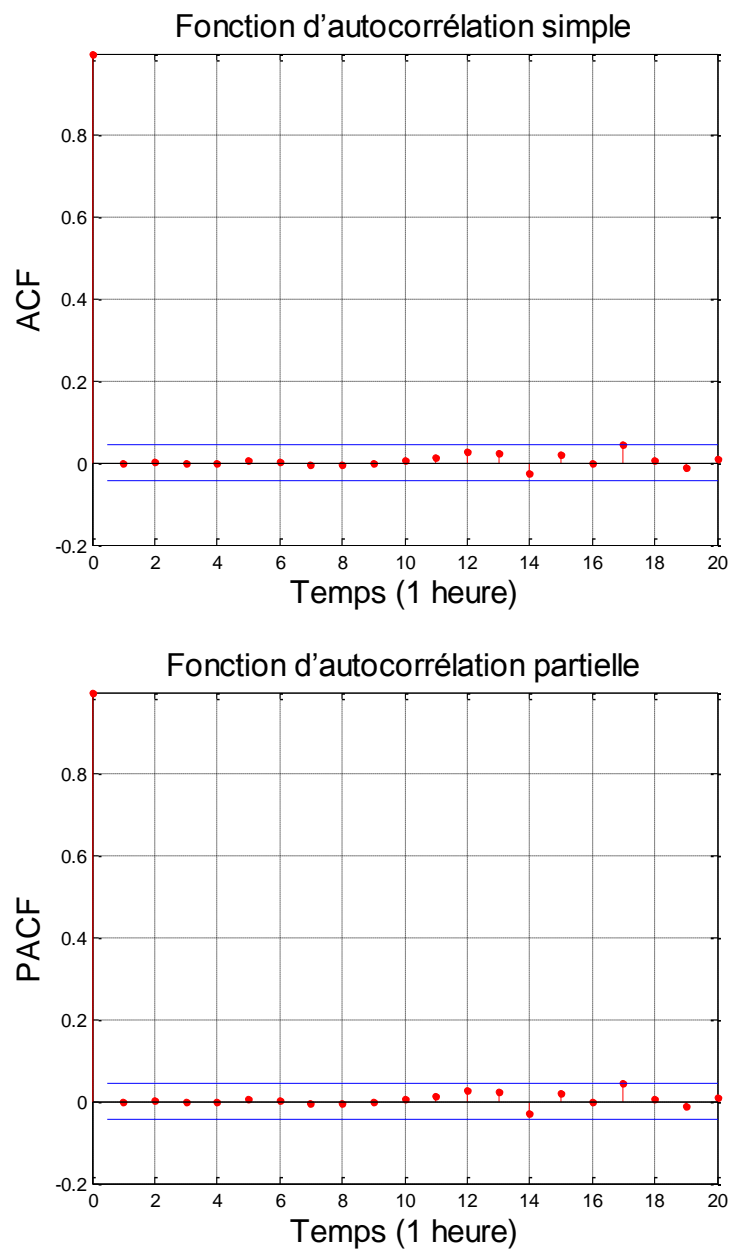


FIGURE IV.12 ACF et PACF des données solaire du site Brasilia

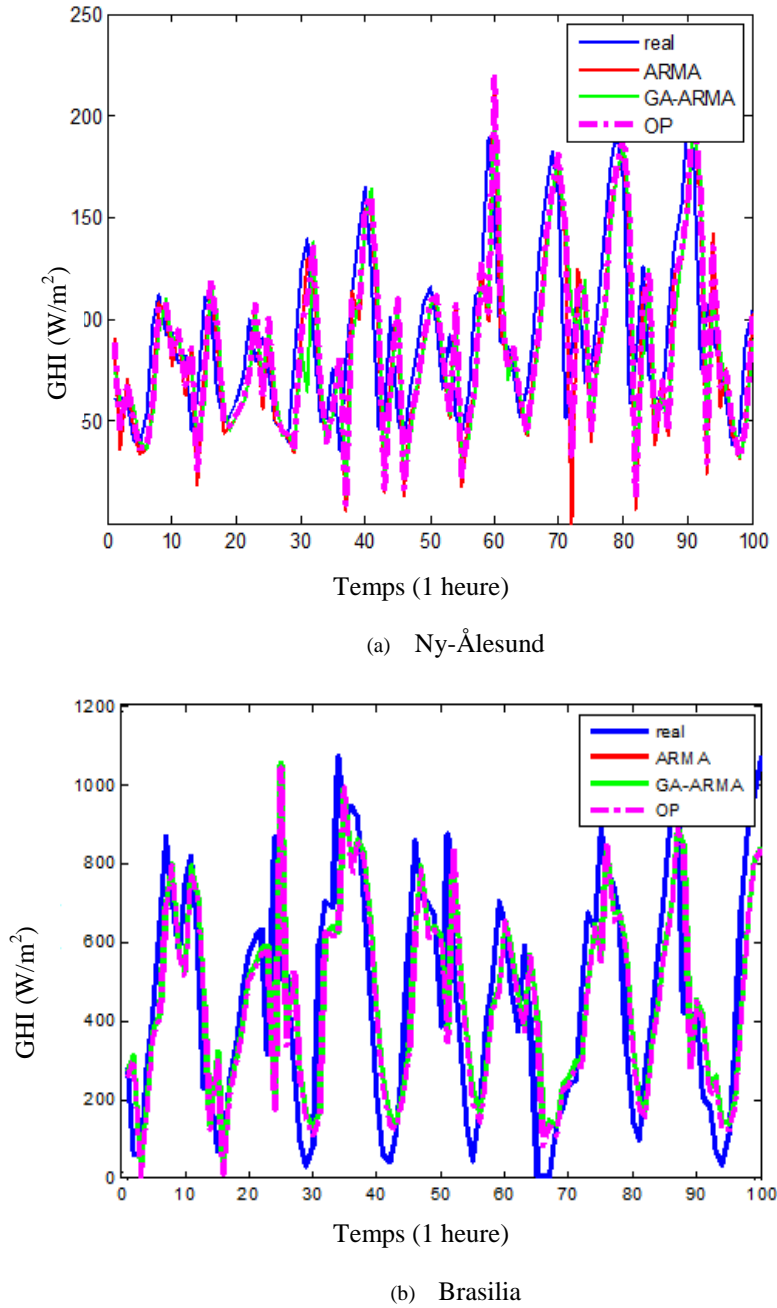


FIGURE IV.13 Comparaison des différents modèles sur les deux stations de test.

#### IV.8.2 Interprétations des courbes

Les tests effectués sur les deux sites (Brasilia et Ny-Alesund), montrent que les algorithmes génétiques constituent un outil puissant pour l'amélioration des résultats via l'optimisation des paramètres du modèle ARMA. Les critères utilisés, dans notre cas, pour juger la performance des AG sont : RMSE, NRMSE,  $R^2$  et MAPE.

Les algorithmes génétiques ont donné la possibilité de déterminer l'ordre et le nombre des paramètres à inclure dans un modèle ARMA automatiquement par la sélection du meilleur chromosome adéquat pour le signal étudié, ainsi que l'estimation supplémentaire de la valeur des paramètres par ajustement.

Enfin nous déduisons que les AG promettent d'être un moyen efficace et robuste au service d'autres algorithmes de prédiction classiques.

## **VI.9. Conclusion**

Dans ce chapitre, une approche AG pour l'optimisation des paramètres du modèle ARMA est introduite comme nouvelle méthodologie de prédiction pour les séries temporelles solaires de GHI. L'algorithme proposé joue un rôle efficace dans l'optimisation de la technique autorégressive ARMA. La fonction fitness calcule uniquement le RMSE, ce qui signifie qu'il n'y a pas de calculs complexes (dérivant des équations pour le processus d'estimation) dans le processus d'évaluation. L'algorithme minimise l'erreur dans l'estimation des paramètres et fournit de meilleurs résultats dans la détection de l'ordre du modèle par rapport à l'ARMA classique. L'algorithme proposé peut tout d'abord, détecter le vrai ordre dans un premier temps, puis estimer les paramètres en fonction de la valeur d'ordre obtenue dans un deuxième temps.

Il a été constaté que la technique de choix de l'ordre du modèle AG-ARMA et l'estimation des paramètres peuvent donner de bons résultats par rapport aux méthodes ARMA classiques. En outre, l'ajout d'une autre étape de calcul en combinant les deux modèles de prédiction peut améliorer davantage les performances de prédiction.