

RÉPUBLIQUE DU CAMEROUN

Paix-Travail-Patrie

UNIVERSITÉ DE YAOUNDÉ I

ÉCOLE NORMALE SUPÉRIEURE DE
YAOUNDÉ

DÉPARTEMENT DE MATHÉMATIQUES



REPUBLIC OF CAMEROON

Peace-Work-Fatherland

UNIVERSITY OF YAOUNDE I

HIGHER TEACHER'S TRAINING
COLLEGE OF YAOUNDE

DEPARTMENT OF MATHEMATICS

**MODÉLISATION DE LA PROLIFÉRATION DE LA JACINTHE
D'EAU A PARTIR DES SÉRIES TEMPORELLES**

**Mémoire rédigé en vue de l'obtention du Diplôme de
Professeur de l'Enseignement Secondaire deuxième grade
(D.I.P.E.S II) en mathématiques**

Par :

MPONO BABOYA Sabine

Matricule : 08Y600

Master en Mathématiques

Professeur des Collèges d'Enseignement Général

Sous l'encadrement de :

Pr. TEWA Jean Jules

Maître de conférences /UY1

Année académique : 2018-2019

Dédicace

Je dédie ce travail :
A mon tendre et cher époux
Car il a cru en moi et a toujours investi tant moralement que
financièrement à l'élaboration de ce travail.

Remerciements

Sous la conduite du Dieu Tout-puissant, ce travail est le concours du soutien de plusieurs personnes à qui je souhaiterai exprimer ma profonde gratitude :

- Tout d'abord à mon encadreur, le Professeur **TEWA Jean Jules** qui m'a fait honneur en acceptant de diriger ce travail et m'a par ailleurs guidé dans le choix de mon thème de recherche. Son admirable culture scientifique particulièrement variée et sa rigueur au travail ont fait de moi l'heureuse étudiante que je suis aujourd'hui.
- Je remercie également Dr. **GHOUE Jean-Paul** et Dr. **KENFACK Sonia Nadege** tous chercheurs à l'Herbier National pour leur soutien, leur aide dans la collecte des données et leur profonde conviction de l'intérêt de l'analyse de la dynamique des plantes envahissantes dans les cours d'eau à travers le Cameroun.
- Mes remerciements vont aussi à l'endroit du Professeur **TCHANTCHO Bertrand**, enseignant au Département de Mathématiques de l'École Normale Supérieure de Yaoundé pour sa disponibilité aux côtés des étudiants. Malgré de nombreux impondérables, vous avez toujours donné le meilleur de vous-même. J'avoue que sans vos encouragements, ce travail n'aurait pas certainement pas vu le jour...
- Mes remerciements vont aussi à l'endroit de tous les enseignants de l'École Normale Supérieure de Yaoundé, en particulier ceux du département de Mathématiques pour tous les efforts par eux consentis pour ma formation au premier cycle puis au second cycle. Vraiment, je puis vous dire que vos enseignements se sont bonifiés avec le temps.
- J'adresse un vibrant merci à tous mes camarades de promotion qui ont toujours su se mettre au service des autres à travers le travail de groupe et le sens du partage.
- Mes remerciements s'adressent également à toute la famille **BOYAYINI** pour son soutien durant les années passées à l'École Normale Supérieure de Yaoundé I. je voudrais particulièrement remercier Clarisse, Roméo, Guy, Cunégonde, Salomon et Ange pour toute l'assistance qu'ils n'ont jamais cessé de manifester à ma modeste personne.
- Je remercie aussi toutes personnes qui, de près ou de loin, m'ont aidé sans oublier ceux qui au quotidien se sacrifient pour que nous ayons un monde meilleur pour la future génération.

Déclaration sur l'honneur

Le présent document est une œuvre originale de la candidate et n'a été soumis nulle part ailleurs, en partie ou en totalité, pour une autre évaluation académique. Les contributions externes ont été dûment recensées et mentionnées en bibliographie.

Signature de la candidate

Mpono Baboya Sabine

Table des matières

Dédicace	i
Remerciements	ii
Déclaration sur l'honneur	iii
Liste des figures	vi
Résumé	vii
Abstract	viii
Introduction générale	1
1 Généralités sur les séries temporelles et outils biologiques nécessaires pour la modélisation	4
1.1 Rappels de Probabilité et de Statistiques	4
1.2 Série temporelle et processus aléatoire	5
1.2.1 Processus stationnaires	6
1.2.1.1 Stationnarité au sens strict ou stationnarité forte	6
1.2.1.2 Stationnarité au second ordre ou stationnarité faible	6
1.2.2 Caractéristiques d'une série temporelle	7
1.2.2.1 la fonction d'auto-covariance et l'auto-corrélation	7
1.2.2.2 La fonction d'auto-corrélaton partielle	8
1.2.3 Processus bruit blanc	8
1.2.4 Le théorème de décomposition de Wold	9
1.2.5 Processus non stationnaire	9
1.2.5.1 Caractéristiques des processus TS	9
1.2.5.2 Caractéristiques des processus DS	9
1.2.5.3 Tests de racine unitaire	10
1.3 Biologie sur la Jacinthe d'eau	11
1.3.1 Information taxonomique	11
1.3.2 Origine de la jacinthe d'eau	11

1.3.3	Mécanismes de dissémination	12
1.3.4	Valeur économique de la jacinthe d'eau	13
2	Processus ARMA	15
2.1	Généralités et notations	15
2.2	Définition des processus ARMA	15
2.2.1	Equations de Yule-Walker	16
2.2.2	Autocorrélations partielles et Autocorrélations	16
2.3	Définition des processus ARMA (p,q)	17
2.4	Validation des processus ARMA	18
2.5	Prévision des processus ARMA	20
3	Modélisation et prévision de la prolifération de la Jacinthe d'eau	21
3.1	Application à l'étude de la croissance de la jacinthe d'eau	22
3.2	conclusion	28
	Apport Pédagogique	29
	Conclusion et perspectives	31
	Bibliographie	32

Table des figures

1.1	Jacinthe d'eau recouvrant complètement une partie d'eau à petit BONANJO à proximité du fleuve Wouri	12
1.2	Organisation de l'appareil végétal de la jacinthe d'eau (1- inflorescence ; 2- Limbe foliaire ; 3-bractée supérieure ; 4-pétiole ; 5-pédoncule floral ; 6-nouvelle feuille ; 7-jeune pousse ; 8- une racine.)	12
1.3	Meubles et objets artisanaux obtenus à partir des cordes de jacinthe (a-chaise ; b-mannequin ; c- chapeau)	13
1.4	Photo d'une poignée du compost fabriqué lors de l'essai à Gbessou	14
3.1	Graphe de la série SURFACE	22
3.2	Corollégramme et corollégramme partiel de la variable surface	23
3.3	Graphe de la série en différence première	23
3.4	Corrélogramme de la série en différence première	24
3.5	Graphe de la série en différence première	24
3.6	Modèle ar (1)	25
3.7	Modèle ma (1)	25
3.8	Modèle arma (1,1)	25
3.9	Corrolégramme et corrélogramme partiels des résidus du modèle ar (1)	26
3.10	Corrolégramme et corrélogramme partiels des résidus du modèle ma (1)	27
3.11	Critère de choix des modèles	28
3.12	Prévisions de le série	28

Résumé

Dans ce travail, nous décrivons de manière précise la dynamique d'évolution de la Jacinthe d'eau dans le littoral camerounais pendant 132 semaines. A l'aide de la modélisation à partir des séries temporelles, nous fournissons une méthode utile et pratique permettant de décrire l'évolution des plantes et de prédire leur expansion. La méthode de Box et Jenkins est utilisée pour déterminer les paramètres du modèle de type ARIMA obtenu au cours de la modélisation.

Mots clés : Séries temporelles, Méthode de Box et Jenkins, ARIMA, processus stationnaire.

Abstract

In this dissertation, we describe in a precise way the evolution of Water hyacinth in the Region of littoral in Cameroon during 132 weeks. For our work we use times serials to describe and predicts the evolution of plant. The Box and Jenkins method is use to determine the parameters of the ARIMA model wich has been obtain during our study.

Keywords : Time serial, Box and Jenkins method, ARIMA, stationary process.

Introduction générale

Contexte

Les objectifs de développement durable (ODD) portent sur de nombreux sujets parmi lesquels : la lutte contre le changement climatique, la protection de la faune et de la flore aquatique et terrestre. Ces objectifs, au nombre de 17, font partie du nouveau programme de développement mondial post-2015 qui cadre avec la vision du Cameroun à l'horizon 2035. C'est suivant cette logique que s'inscrivent les nombreuses conférences et colloques organisés par le Ministère de l'Environnement de la Protection de la Nature et du Développement Durable (MINEPDED). En 2010, Le MINEPDED a mis sur pied le projet n° 00103/M/MINEPDED/SG/DAG/SDBMM/SM/2010 portant sur la collecte, la valorisation et la gestion des déchets de la jacinthe d'eau afin de lutter efficacement contre l'invasion de cette plante. Les 17 et 18 mars 2016, il s'est tenue à Douala la réunion d'évaluation du projet de lutte contre la jacinthe d'eau et les autres plantes aquatiques envahissantes au Cameroun. Le sujet occupe au premier vol le MINEPDED, car la jacinthe d'eau et les plantes aquatiques envahissantes, comportent de nombreux effets néfastes sur l'environnement. Elles favorisent l'ensablement des flans d'eau, leur rétrécissement, voire leur disparition, entraînant la détérioration de la qualité de l'eau, empêchant le passage des rayons lumineux et par conséquent, la disparition de certaines espèces piscicoles, ainsi que la navigation fluviale, entre autres.

Un des désagréments généralement évoqués, lorsque l'on aborde la question de la prolifération de cette plante, est sa propension à recouvrir totalement la surface qu'elle colonise. De plus l'étude de la dynamique de cette plante est menée au sein des équipes pluridisciplinaires composées entre autre des mathématiciens, des biologistes, des environnementalistes etc.

Au-delà des résultats importants obtenus à l'issue des recherches menées dans les domaines de la biologie végétale et animale, les mathématiques contribuent de plus en plus à la compréhension du fonctionnement des écosystèmes et de prédiction de leurs évolutions. A cet effet, et ce malgré la complexité des systèmes écologiques étudiés, les modèles mathématiques élaborées n'ont cessé d'accroître leurs caractères explicatifs et prédictifs.

Comme l'atteste l'abondante littérature consacrée à l'évolution des systèmes agro-écologiques, la modélisation mathématique est devenue un élément indispensable de toute étude et recherche dans le domaine de la dynamique de population.

Les modèles, étant une formalisation du monde réel, décrivent l'évolution de manière peu précise car la détermination de paramètres peut s'avérer difficile. Par conséquent, le développement

des modèles requiert une connaissance approfondie du fonctionnement du système considéré. Toutefois, la construction et l'analyse des modèles dynamiques posent aussi des questions méthodologiques qui relèvent des domaines des mathématiques et de la statistique. Néanmoins, ces questions sont encore peu abordées dans le domaine de la modélisation en agronomie, ce qui génère des difficultés dans la manipulation et l'utilisation des modèles élaborés : les modélisateurs sont généralement des spécialistes du système à modéliser mais, manquent de méthodes mathématiques et statistiques pour travailler avec les modèles malgré l'existence de méthodes éprouvées et relativement accessibles, associées à la puissance de calcul des ordinateurs qui permet de les mettre en application par tous les modélisateurs.

L'objectif de ce mémoire est de présenter une méthode de modélisation univariée (le modèle ARMA :Auto Regressive Moving Average) et son application à l'étude de la dynamique des populations. En général, cette approche permet uniquement de décrire et non d'expliquer l'évolution ou le comportement d'une série étudiée, en raison du manque d'un ensemble d'information plus large. Une fois le modèle trouvé, nous l'utiliserons pour prédire les valeurs futures des observations. En particulier, ce travail porte sur la détermination d'un modèle de type ARMA pour appréhender la prolifération de la jacinthe d'eau et de prévoir sa dynamique future.

En vue d'améliorer les connaissances sur le processus mathématique qui gouverne la prolifération de cette plante, nous allons dans le cadre de ce mémoire utiliser les méthodes d'analyse des séries temporelles. L'utilisation de cette méthode scientifiquement robuste peut permettre de peaufiner le dispositif d'alerte et de contrôle existant.

Problématique

En biologie végétale, la modélisation cherche à expliquer les lois majeures qui gouvernent la vie d'une plante au niveau de l'individu ainsi qu'en communauté. En fonction de l'objectif de l'étude, différentes approches sont utilisées.

Une chose est de modéliser, une autre est d'appliquer ces modèles et de les confronter à des données expérimentales. A partir des modèles ARMA, il est possible de faire de la prévision ; mais, la justesse des prévisions dépend fortement de la qualité du modèle choisi. Il est donc primordial de trouver des modèles qui reflètent le mieux possible la réalité afin de minimiser les erreurs de prévision.

Objectif

L'objectif de ce mémoire est de proposer un modèle mathématique qui décrit mieux la prolifération de la jacinthe d'eau, afin de faire des prévisions sur sa capacité à envahir un milieu donné. Une méthode de prévision très populaire est basée sur l'étude rigoureuse de séries temporelles (chronologiques). Par conséquent, il s'agit de développer, d'adapter et de tester si cette méthode statistique est applicable à la croissance de la Jacinthe d'eau.

Organisation du travail

Le présent mémoire sur l'étude de la prolifération du Jacinthe d'eau sera subdivisé en trois chapitres :

- le premier chapitre présente l'ancrage théorique et définit les principales notions qui seront utilisées dans le mémoire ;
- le deuxième chapitre expose quelques notions théoriques utiles sur les processus ARMA. Cette présentation est assortie d'une présentation des différents tests et critères servant à la validation des paramètres obtenus ;
- le troisième chapitre dévoile les résultats que nous avons obtenus à partir des données collectées auprès des chercheurs de l'Herbier National et des simulations numériques qui s'en sont suivies.

GÉNÉRALITÉS SUR LES SÉRIES TEMPORELLES ET OUTILS BIOLOGIQUES NÉCESSAIRES POUR LA MODÉLISATION

Dans ce chapitre, nous présentons les notions nécessaires et préalables à une bonne compréhension des généralités sur les séries temporelles. L'étude des séries temporelles suppose que l'on fasse au préalable un certain nombre de rappels en probabilité et en statistiques. L'agronomie étant une discipline transversale, nous donnons aussi quelques notions utiles pour la connaissance biologique de la jacinthe d'eau.

1.1 Rappels de Probabilité et de Statistiques

Nous donnons les principales lois obtenues lors de manipulations de la loi normale. Ces résultats servent à établir les intervalles de confiance et les procédures de test dans le cas d'un échantillon issu d'une loi normale. Des tables statistiques et des logiciels permettent de les manipuler.

Définition 1.1.1. loi du Khi-deux

Soit X_1, \dots, X_n une suite de variables aléatoires indépendantes de même loi $\mathcal{N}(0, 1)$.

Alors la variable aléatoire $\sum_{i=1}^n X_i^2$ suit une loi appelée **loi du Khi-deux** à n degré de liberté.

On note : $\sum_{i=1}^n X_i^2 \sim \chi_n^2$.

Définition 1.1.2. loi de Student

Soient $U \sim \mathcal{N}(0, 1)$ et $Z \sim \chi_n^2$. Alors $\frac{U}{\sqrt{Z/n}}$ suit une loi appelée **loi de Student** à n degré de liberté. On note $\frac{U}{\sqrt{Z/n}} \sim St(n)$.

Propriété 1.1.1. loi de Fisher

Soient X_1, \dots, X_n des variables indépendantes et de même loi $\mathcal{N}(m, \sigma^2)$. Posons $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$

et $(S')_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$ Alors

$$\cdot \bar{X}_n \sim \mathcal{N}(m, \sigma^2/n)$$

$$\cdot \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - m)^2 \sim \chi_n^2$$

$$\cdot \frac{(n-1)(S')_n^2}{\sigma^2} \sim \chi_{n-1}^2$$

$$\cdot \frac{\bar{X}_n - m}{S'_n / \sqrt{n}} \sim St(n-1).$$

Démonstration. Hypothèse : X_1, \dots, X_n sont des variables indépendantes et de même loi $\mathcal{N}(m, \sigma^2)$

1) Montrons que $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i \sim \mathcal{N}(m, \sigma^2/n)$

X_1, \dots, X_n étant des variables gaussiennes, alors \bar{X}_n est gaussienne.

Déterminons les paramètres. On a :

$$\begin{aligned} E(\bar{X}_n) &= E\left(\frac{1}{n} \sum_{i=1}^n X_i\right) \\ &= \frac{1}{n} \sum_{i=1}^n E(X_i) \\ &= m \end{aligned}$$

$$\begin{aligned} \text{var}(\bar{X}_n) &= \text{var}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) \\ &= \frac{1}{n^2} \sum_{i=1}^n \text{var}(X_i) \\ &= \sigma^2/n \end{aligned}$$

Donc $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i \sim \mathcal{N}(m, \sigma^2/n)$

2) Montrons que $\frac{1}{\sigma^2} \sum_{i=1}^n (X_i - m)^2 \sim \chi_n^2$

On a : $\frac{1}{\sigma^2} \sum_{i=1}^n (X_i - m)^2 = \sum_{i=1}^n \left(\frac{X_i - m}{\sigma}\right)^2$

Il suffit de montrer que $\left(\frac{X_i - m}{\sigma}\right) \sim \mathcal{N}(0, 1)$

$$\begin{aligned} E\left(\frac{X_i - m}{\sigma}\right) &= \frac{E(X_i) - m}{\sigma} \\ &= 0 \end{aligned}$$

$$\begin{aligned} \text{var}\left(\frac{X_i - m}{\sigma}\right) &= \frac{1}{\sigma^2} \text{var}(X_i) \\ &= 1 \end{aligned}$$

Donc $\left(\frac{X_i - m}{\sigma}\right) \sim \mathcal{N}(0, 1)$ et d'après la définition 1.1.1, $\sum_{i=1}^n \left(\frac{X_i - m}{\sigma}\right)^2 \sim \chi_n^2$

Par conséquent, $\frac{1}{\sigma^2} \sum_{i=1}^n (X_i - m)^2 \sim \chi_n^2$ □

1.2 Série temporelle et processus aléatoire

Une série temporelle peut-être observée de manière discrète ou de manière continue. Dans le cadre de ce travail, nous nous limiterons à la présentation des séries temporelles discrètes observées à temps réguliers. Après avoir rappelé la définition d'une série temporelle, nous reviendrons sur le concept fondamental de stationnarité. Cette section s'achève par la présentation des principales caractéristiques des séries temporelles.

Définition 1.2.1. (*série temporelle*)

Une **série temporelle** (**série chronologique**) est une suite de nombres réels, indexés par les entiers tels que le temps. En d'autres termes, c'est un ensemble d'observations, X_t , chacune étant enregistrée à un instant spécifique t .

Le but de cette partie étant de présenter des techniques pour modéliser de telles observations, Nous proposons alors de les modéliser comme une réalisation d'un processus stochastique $\{X_t, t \in T\}$.

Définition 1.2.2. (*Processus stochastique*)

Un **Processus stochastique** est une famille de variables aléatoires X_t définies sur un espace probabilisés.

Une série temporelle est ainsi la réalisation d'un processus aléatoire.

1.2.1 Processus stationnaires

La classe des processus aléatoires est très large, l'analyse des séries temporelles a donc été centrée initialement sur une classe particulière de processus : les processus aléatoires stationnaires. Il importe donc de définir les aspects liés à stationnarité.

Considérons une série temporelle X_t , $t=1, \dots, T$ (T est le nombre d'observations de la série). Avant d'effectuer des tests spécifiques sur cette série et de chercher à la modéliser, il convient d'étudier ses caractéristiques stochastiques, telles que son espérance et sa variance.

1.2.1.1 Stationnarité au sens strict ou stationnarité forte

Définition 1.2.3. Le processus X_t est stationnaire au sens strict si pour tout (t_1, t_2, \dots, t_n) avec $t_i \in T$, $i=1, \dots, n$ et si pour tout $\tau \in T$ avec $t_{i+\tau} \in T$, $\{X_{t_1}, X_{t_2}, \dots, X_{t_n}\}$ à la même distribution de probabilité jointe que $\{X_{t_1+\tau}, X_{t_2+\tau}, \dots, X_{t_n+\tau}\}$.

Un processus strictement stationnaire a toutes ses caractéristiques (c'est-à-dire tous ses moments) invariantes dans le temps.

1.2.1.2 Stationnarité au second ordre ou stationnarité faible

Définition 1.2.4. Le processus X_t est stationnaire au second ordre si :

1. $E(X_t^2) < +\infty, \forall t \in \mathbb{Z}$
2. $E(X_t) = m, \forall t \in \mathbb{Z}$
3. $Cov(X_t, X_{t+h}) = \gamma(h), \forall t, h \in \mathbb{Z}$ où γ est la fonction d'autocovariance du processus.

La condition (1) signifie que le processus est du second ordre c'est-à-dire que les moments d'ordre 2 sont finis.

Selon la condition 2., la moyenne du processus aléatoire est constante.

La condition 3. traduit le fait que la covariance entre deux périodes t et $t+h$ est uniquement fonction de la différence des temps h .

En résumé, X_t est stationnaire au second ordre si sa moyenne, sa variance et sa covariance sont indépendantes du temps et si sa variance est finie.

1.2.2 Caractéristiques d'une série temporelle

Les principales caractéristiques temporelles d'un processus sont données par l'auto-corrélation (simple) et l'auto-corrélation partielle.

1.2.2.1 la fonction d'auto-covariance et l'auto-corrélation

- La fonction d'auto-covariance $\{\gamma_h\}_{h \in \mathbb{Z}}$ mesure la covariance entre une variable et cette même variable a des dates différentes, pour un délai h :

$$\gamma_h = \text{cov}(X_t, X_{t+h}) = E[(X_t - E(X_t))(X_{t+h} - E(X_{t+h}))].$$

Ainsi $\gamma_0 = \text{var}(X_t) = E[(X_t - E(X_t))^2] = \sigma_X^2$.

Elle fournit une information sur la validité de la série et sur les liaisons temporelles qui existent entre les diverses composantes de la série X_t .

- la fonction d'auto-corrélation est définie par :

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)}, h \in \mathbb{Z}$$

On appelle coefficient d'auto-corrélation d'ordre 1 (resp. d'ordre k) le coefficient de corrélation linéaire $\rho(1)$ (resp. $\rho(k)$) calculé entre la série et cette série décalée d'une période (resp : de k périodes)

Théorème 1.2.1. Si $\gamma(\cdot)$ est une fonction d'auto-covariance d'un processus stationnaire $(X_t)_{t \in \mathbb{Z}}$ alors

- (i) $0 \leq \gamma(0)$
- (ii) $|\gamma(h)| \leq \gamma(0), \forall h \in \mathbb{Z}$
- (iii) $\gamma(-h) = \gamma(h), \forall h \in \mathbb{Z}$

Démonstration. (i) $\gamma(0) = \text{var}(X_t)$ d'où $0 \leq \gamma(0)$

(ii) : par l'inégalité de Cauchy-Schwarz. $E[(X_t - E(X_t))(X_{t+h} - E(X_{t+h}))] \leq \text{var}(X_t)^{\frac{1}{2}} \text{var}(X_{t+h})^{\frac{1}{2}}$

(iii) $\gamma(-h) = \text{cov}(X_{t-h}, X_t) = \text{cov}(X_t, X_{t+h}) = \gamma(h)$

1.2.2.2 La fonction d'auto-corrélation partielle

Elle mesure la liaison (linéaire) entre X_t et X_{t+h} une fois retirés les liens transitant par les variables $X_{t-1}, \dots, X_{t-h+1}$. Notons ρ_h et Φ_{hh} les fonctions respectives d'auto-corrélation et d'auto-corrélation partielle de X_t . Soit P_h la matrice symétrique formée des $(h-1)$ paramètres d'auto-corrélations de X_t :

$$P_h = \begin{bmatrix} 1 & \rho_1 & \cdot & \cdot & \cdot & \rho_{h-1} \\ \rho_1 & 1 & & & & \\ \cdot & & \cdot & & & \\ \cdot & & & \cdot & & \\ \cdot & & & & \cdot & \\ \rho_{h-1} & & & & \rho_1 & 1 \end{bmatrix}$$

La fonction d'auto-corrélation partielle est donnée par :

$$\Phi_{hh} = \frac{|P_h^*|}{|P_h|}$$

où $|P_h^*|$ est le déterminant de la matrice P_h et P_h^* est donnée par :

$$P_h^* = \begin{bmatrix} 1 & \rho_1 & \cdot & \cdot & \rho_{h-2} & \rho_1 \\ \rho_1 & 1 & & & & \\ \cdot & & \cdot & & & \\ \cdot & & & \cdot & & \\ \cdot & & & & 1 & \\ \rho_{h-1} & & & & \rho_1 & \rho_h \end{bmatrix}$$

P_h^* est ainsi la matrice de P_h dans laquelle on a remplacé la dernière colonne par le vecteur $[\rho_1 \dots \rho_h]^t$.

1.2.3 Processus bruit blanc

Définition 1.2.5. Un processus stationnaire ε_t , centré est un **bruit blanc** si :

- $E[\varepsilon_t] = 0; \forall t$
- $V[\varepsilon_t] = \sigma_\varepsilon^2; \forall t$
- $\text{cov}[\varepsilon_t, \varepsilon_{t'}] = 0; \text{pour } t \neq t'$

Un bruit blanc est ainsi un processus de moyenne nulle, de variance constante et non autocorrélé. On note : $\varepsilon_t \sim BB(0, \sigma_\varepsilon^2)$.

Remarque 1.2.1. Un bruit blanc n'est pas nécessairement gaussien.

1.2.4 Le théorème de décomposition de Wold

L'énoncé du théorème de Wold est le suivant :

Théorème 1.2.2. (Décomposition de Wold)

Tout processus stationnaire d'ordre deux $(x_t)_{t \in \mathbb{Z}}$ peut être représenté sous la forme :

$$x_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} + k_t$$

où les paramètres ψ_j satisfont : $\psi_0 = 1$, $\psi_j \in \mathbb{R}$, $\forall j \in \mathbb{N}^*$, $\sum_{j=0}^{\infty} \psi_j^2 < \infty$ et où ε_t est un bruit blanc i.i.d. (c'est-à-dire variables indépendantes et identiquement distribuées). On dit que la somme des chocs passés correspond à la composante linéaire stochastique de x_t . le terme k_t désigne la composante linéaire déterministe telle que $\text{cov}(k_t, \varepsilon_{t-j}) = 0$, $\forall j \in \mathbb{N}$.

1.2.5 Processus non stationnaire

Les séries observées montrent souvent, dans la réalité, un comportement non stationnaire, il convient alors d'identifier le type de non stationnarité auquel la série est soumise afin d'utiliser un modèle approprié de façon à la rendre stationnaire.

Deux grandes catégories de processus non stationnaires peuvent être distingués :

- les processus TS (Trend Stationary) présentant une non stationnarité de nature déterministe ;
- les processus DS (Difference Stationary) présentant une non stationnarité de nature stochastique.

1.2.5.1 Caractéristiques des processus TS

Si l'on suppose qu'un processus peut s'écrire comme la somme d'une fonction déterministe du temps et d'un élément stochastique stationnaire, alors un processus TS est donné par :

$$X_t = f_t + \varepsilon_t$$

Où f_t est une fonction déterministe du temps et ε_t est un processus stationnaire.

Remarque 1.2.2. Un processus TS est un processus que l'on peut rendre stationnaire par une régression sur une tendance déterministe.

1.2.5.2 Caractéristiques des processus DS

Un processus DS s'écrit comme suit :

$$X_t = \rho^\tau X_{t-\tau} + \beta \sum_{j=0}^{\tau-1} \rho^j + \sum_{j=0}^{\tau-1} \rho^j \varepsilon_{t-j}$$

Où ε_t est un processus stationnaire.

Remarque 1.2.3. un processus DS est un processus que l'on peut rendre stationnaire par l'utilisation d'un filtre aux différences :

$$(1 - L)^d X_t = \beta + \varepsilon_t$$

Où ε_t est un processus stationnaire, d est l'ordre de différenciation (ou d'intégration) et L est l'opérateur retard.

Plus généralement, on a la définition suivante due à Granger.

Définition 1.2.6. Une série X_t est intégrée d'ordre d , ce que l'on note $X_t \sim I(d)$, s'il est nécessaire de la différencier d fois pour la rendre stationnaire.

En d'autres termes, $X_t \sim I(d)$ si et seulement si $(1 - L)^d X_t \sim I(0)$

Eu égard à ce qui précède, il convient de faire clairement la distinction entre processus TS et processus DS. Cette distinction peut être effectuée par le biais des tests de racine unitaire.

1.2.5.3 Tests de racine unitaire

Il existe un grand nombre de tests de racine unitaire. Dans le cadre de ce travail, nous allons présenter les tests de Dickey-Fuller qui sont des tests paramétriques qui reposent sur l'estimation d'un processus autorégressif. Dickey et Fuller considèrent trois modèles de base pour la série $X_{t,t=1,\dots,T}$.

– Modèle 1 : modèle sans constante ni tendance déterministe :

$$(1 - \rho L) X_t = \varepsilon_t$$

– Modèle 2 : modèle avec constante sans tendance déterministe :

$$(1 - \rho L) (X_t - \mu) = \varepsilon_t$$

– Modèle 3 : modèle avec constante et tendance déterministe :

$$(1 - \rho L) (X_t - \alpha - \beta t) = \varepsilon_t$$

Dans chacun des cas, on suppose que $\varepsilon_t \sim BB(0, \sigma_\varepsilon^2)$

Lors de l'implémentation d'un test de Dickey et Fuller, on teste l'hypothèse nulle de racine unitaire (X_t est intégré d'ordre 1, c'est-à-dire non stationnaire) contre l'hypothèse alternative d'absence de racine unitaire (X_t est intégré d'ordre 0, c'est-à-dire stationnaire). La formalisation des tests des modèles 1, 2 et 3 présentés ci-dessus sont les suivantes :

-Modèle 1

$$\begin{cases} H_0 : \rho = 1 \Leftrightarrow X_t = X_{t-1} + \varepsilon_t \\ H_1 : |\rho| < 1 \Leftrightarrow X_t = \rho X_{t-1} + \varepsilon_t \end{cases}$$

Ainsi, sous l'hypothèse nulle, X_t suit un processus de marche aléatoire sans dérive. Sous l'hypothèse alternative, X_t suit un processus AR (1).

-Modèle 2

$$\begin{cases} H_0 : \rho = 1 \Leftrightarrow X_t = X_{t-1} + \varepsilon_t \\ H_1 : |\rho| < 1 \Leftrightarrow X_t = \rho X_{t-1} + y + \varepsilon_t \end{cases}$$

Avec $y = \mu(1 - \rho)$

L'hypothèse nulle correspond à un processus de marche aléatoire sans dérive. Sous l'hypothèse alternative, X_t suit un processus AR (1) avec dérive.

-Modèle 3

$$\begin{cases} H_0 : \rho = 1 \Leftrightarrow X_t = X_{t-1} + \beta + \varepsilon_t \\ H_1 : |\rho| < 1 \Leftrightarrow X_t = \rho X_{t-1} + \lambda + \delta t + \varepsilon_t \end{cases}$$

Avec $\lambda = \alpha(1 - \rho) + \rho\beta$ et $\delta = \beta(1 - \rho)$

Sous l'hypothèse nulle, X_t suit une marche aléatoire avec dérive. Sous l'hypothèse alternative, nous remarquons qu'on peut écrire :

$$X_t = \alpha + \beta t + \eta_t$$

Avec $\eta_t = X_t - \alpha - \beta t = \frac{1}{1-\rho L} * \varepsilon_t$

Ainsi, sous l'hypothèse alternative, X_t est un processus TS avec erreurs ARMA. On peut le rendre stationnaire en calculant les écarts par rapport à la tendance estimée par les MCO.

1.3 Biologie sur la Jacinthe d'eau

1.3.1 Information taxonomique

La jacinthe d'eau appartient à la famille Pontederiaceae, du genre Eichhornia et de l'espèce : *Eichhornia crassipes* (Martius) Solms-Laubach. (TSAGUE, 2003)[9]. Il existe six (06) autres espèces du genre *Eichhornia* (Dagno, 2007)[2]. Seule *E. natans* est endémique en Afrique (Bénin, Sénégal, Egypte, Soudan, Nigeria, Madagascar et Mali) (TSAGUE, 2013)[9].

1.3.2 Origine de la jacinthe d'eau

Les résultats de l'enquête relative à l'évaluation socio-économique et culturelle de l'impact de la plante dans la région du Littoral font état de ce que cette plante aurait été introduite dans la région à partir d'une zone située vraisemblablement en amont du Wouri notamment à Bonamoussadi. La légende selon les riverains est qu'une dame venue de l'occident, résidant le quartier, aurait apporté cette espèce qu'elle cultivait comme plante ornementale dans des bacs ; la plante a proliféré si bien qu'elle est devenue incontrôlable, et de ce fait a été déversée dans le Wouri où elle s'est répandue dans tout l'estuaire par la faveur du mouvement des marées associé à l'organisation de son appareil végétatif.

1.3.3 Mécanismes de dissémination

La jacinthe d'eau peut se reproduire aussi bien de façon sexuée que par multiplication végétative, elle va généralement favoriser la reproduction par voie végétative (TSAGUE, 2013)[9]. La reproduction sexuée est un mode de multiplication secondaire pour la jacinthe d'eau (*E. crassipes*), qui est une espèce monoïque (fleurs mâles et fleurs femelles sont portées sur le même pied), de forts taux d'autofécondation (TSAGUE, 2013)[9]. La reproduction sexuée conduit à la formation des graines dont la longévité est de 15 ans (Dagno, 2007)[2]. Un plant de jacinthe d'eau produit environ 5.000 graines, qui sont dispersées par l'eau ou le vent, et vont coloniser les berges où le niveau d'eau est faible.

La multiplication végétative est le mode presque exclusif d'extension de la plupart des espèces aquatiques proliférantes telle que la jacinthe d'eau. Elle est fondamentale pour le maintien de la plante dans le milieu et ses capacités d'extension, ce qui lui permet d'augmenter rapidement sa biomasse. Les nouvelles plantes sont produites à partir de l'élongation des stolons due à la division des méristèmes axillaires de la plante mère (Dagno, 2007)[2]. Les clones très fragiles restent fixés à la plante mère par le stolon, puis se détachent sous la pression des courants d'eau, permettant à de nouveaux individus de coloniser d'autres zones (TSAGUE, 2013)[9]. D'après certaines études, deux (2) plants de jacinthe d'eau peuvent par clonage, produire 1200 plants en quatre (4) mois. Certains auteurs ont démontré que la durée du clonage de la jacinthe d'eau est d'environ six (6) à dix-huit (18) jours et dépend des conditions écologiques.



FIGURE 1.1 – Jacinthe d'eau recouvrant complètement une partie d'eau à petit BONANJO à proximité du fleuve Wouri



FIGURE 1.2 – Organisation de l'appareil végétal de la jacinthe d'eau (1- inflorescence ; 2- Limbe foliaire ; 3-bractée supérieure ; 4-pétiole ; 5-pédoncule floral ; 6-nouvelle feuille ; 7-jeune pousse ; 8- une racine.)

1.3.4 Valeur économique de la jacinthe d'eau

Plusieurs tentatives de valorisation de la jacinthe d'eau sont en cours d'étude. Jusqu'à présent, aucune de ces activités n'a réellement abouti à un développement industriel (Dagno, 2007)[2]. Son utilisation semble cependant ne pas être viable du fait qu'elle nécessite une grande quantité d'eau pour sa mise en culture. Néanmoins la plante est utilisée :

- En phytoremédiation et le traitement des eaux usées (Chine, USA, Sri Lanka, Brésil, Égypte, Inde, Argentine, France) (Dagno, 2007)[2] ;
- Dans la production de biogaz (TSAGUE, 2013)[9] ;
- Dans l'alimentation animale (TSAGUE, 2013)[9] ;
- En tant qu'engrais organique (PRADELLE, 2013) :
- Dans l'industrie à papier (Bangladesh, Ouganda) et pour la confection de particules de panneau de construction (Inde) et le briquetage de charbon de bois (Kenya) (TSAGUE, 2013)[9] ;
- Dans l'industrie des meubles dans les pays asiatiques Cambodge, vannerie (Inde, Fidji, Philippines, Ouganda) et pour la réalisation de filets de pêche et de cordes (Bangladesh) (TSAGUE, 2013)[9] ;
- Pour la fabrication du papier à Sodiko (Douala).

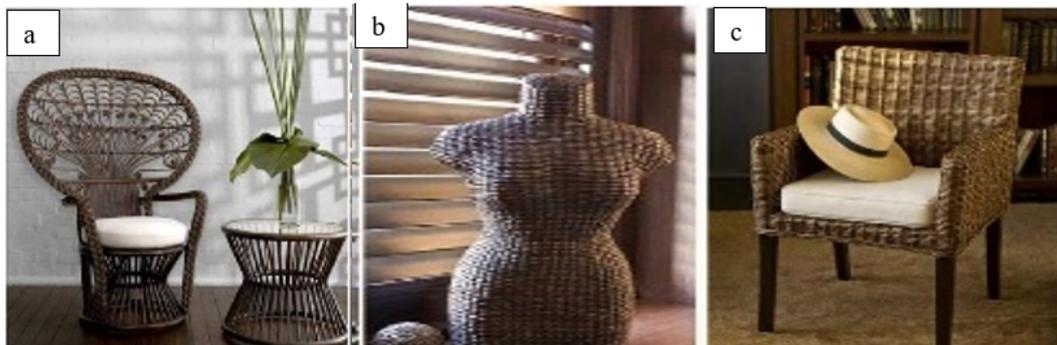


FIGURE 1.3 – Meubles et objets artisanaux obtenus à partir des cordes de jacinthe (a-chaïse ; b-mannequin ; c- chapeau)



FIGURE 1.4 – Photo d'une poignée du compost fabriqué lors de l'essai à Gbessou

PROCESSUS ARMA

Les processus ARMA (Auto Regressive Moving Average) ont été introduits par Box et Jenkins. Le but est de modéliser une série temporelle en fonction de ses valeurs passées, mais aussi en fonction des valeurs présentes et passées d'un bruit. Afin de déterminer le processus ARMA adéquat, Box et Jenkins ont suggéré une procédure à quatre étapes :

- identification du modèle
- estimation des paramètres
- validation du modèle (test de diagnostic)
- prévision à l'aide du modèle.

Avant de présenter ces quatre étapes, nous donnons brièvement quelques définitions.

2.1 Généralités et notations

Définition 2.1.1. (*Opérateur retard*)

L'opérateur retard noté L , est tel que :

$$L^n X_t = X_{t-n}$$

On constate ainsi que l'opérateur retard transforme une X_t en sa valeur passée.

Définition 2.1.2. (*Le polynôme retard*)

Le polynôme retard $\phi(L)$ est défini comme suit :

$$\Phi(L) = 1 - \Phi_1 L - \Phi_2 L^2 - \dots - \Phi_p L^p$$

Si l'on applique le polynôme retard $\phi(L)$ à une série X_t , on a :

$$\Phi(L) X_t = X_t - \Phi_1 X_{t-1} - \Phi_2 X_{t-2} - \dots - \Phi_p X_{t-p}$$

2.2 Définition des processus ARMA

Définition 2.2.1. On appelle **processus autorégressif d'ordre p** , noté $AR(p)$, un processus stationnaire X_t vérifiant une relation du type :

$$X_t - \Phi_1 X_{t-1} - \Phi_2 X_{t-2} - \dots - \Phi_p X_{t-p} = \varepsilon_t$$

Où les $\Phi_i (i = 1, 2, \dots, p)$ sont des réels et $\varepsilon_t \sim BB(0, \sigma_\varepsilon^2)$.

En introduisant l'opérateur retard, la relation peut encore écrire :

$$(1 - \Phi_1 L - \Phi_2 L^2 - \dots - \Phi_p L^p) X_t = \varepsilon_t$$

Soit $\Phi(L) X_t = \varepsilon_t$

Avec : $\Phi(L) = 1 - \Phi_1 L - \Phi_2 L^2 - \dots - \Phi_p L^p$.

Définition 2.2.2. (Autocorrélations)

Les autocorrélations d'un processus AR(p) peuvent être calculées en multipliant l'équation de la définition par $X^{t-k} (k > 0)$.

En notant $\rho_k = \frac{y_k}{y_0}$ on obtient la fonction d'autocorrelation d'un processus AR(p) donnée par :

$$\rho_k = \sum_{i=1}^p \Phi_i \rho_{k-i} \quad \forall k > 0$$

2.2.1 Equations de Yule-Walker

Comme on l'a montré précédemment, les autocorrélations d'un processus AR(p) sont décrites par une équation de récurrence linéaire d'ordre p. en écrivant cette relation pour différentes valeurs de k (k=1,2,...,p) on obtient les équations de Yule-Walker :

$$\begin{pmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_{p-1} \\ \rho_p \end{pmatrix} = \begin{pmatrix} 1 & \rho_1 & \cdots & \rho_{p-1} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p-1} & \cdots & \cdots & 1 \end{pmatrix} \begin{pmatrix} \Phi_1 \\ \Phi_2 \\ \vdots \\ \Phi_{p-1} \\ \Phi_p \end{pmatrix}$$

Ces équations permettent ainsi d'obtenir les coefficients d'autocorrélation en fonction des coefficients autorégressifs et inversement.

2.2.2 Autocorrélations partielles et Autocorrélations

- Autocorrélations partielles

Il est possible de calculer les autocorrélations partielles d'un processus AR(p) à partir de équations de Yule-Walker et des autocorrélations. On utilise pour cela l'algorithme de Durbin :

$$\left\{ \begin{array}{l} \Phi_{11} = \rho_1 \quad \text{initialisation de l'algorithme} \\ \Phi_{kk} = \frac{\rho_k - \sum_{j=1}^{k-1} \Phi_{k-1,j} \rho_{k-j}}{1 - \sum_{j=1}^{k-1} \Phi_{k-1,j} \rho_j} \quad \text{pour } k = 2, 3, \dots \\ \Phi_{kj} = \Phi_{k-1,j} - \Phi_{kk} \Phi_{k-1,k-j} \quad \text{pour } k = 2, 3, \dots \text{ et } j = 1, \dots, k-1 \end{array} \right.$$

Propriété 2.2.1. Pour un processus $AR(p)$, $\Phi_{kk} = 0 \forall k > p$. En d'autres termes, pour un processus $AR(p)$, s'annulent à partir du rang $p+1$.

Définition 2.2.3. (processus moyenne mobile d'ordre q) On appelle **processus moyenne mobile d'ordre q** , noté $MA(q)$, un processus stationnaire X_t vérifiant une relation du type :

$$X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$$

Où les θ_i ($i=1, \dots, q$) sont des réels et $\varepsilon_t \sim BB(0, \sigma_\varepsilon^2)$.

En introduisant l'opérateur retard, la relation peut encore s'écrire :

$$(1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q) \varepsilon_t = X_t$$

Soit : $\Theta(L) \varepsilon_t = X_t$ avec $\Theta(L) = 1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q$

Remarque 2.2.1. Par définition, un processus $MA(q)$ est toujours stationnaire.

- Autocovariance

La fonction d'autocovariance d'un processus $MA(q)$ est donnée par :

$$\gamma_k = E [X_t X_{t-k}]$$

Quelques calculs conduisent à l'expression :

$$\gamma_k = \begin{cases} (-\theta_k + \theta_1 \theta_{k+1} + \dots - \theta_{q-k} \theta_k) \sigma_\varepsilon^2 & \text{si } k = 1, \dots, q \\ 0 & \text{si } k > q \end{cases}$$

- Autocorrélation

En notant $\rho_k = \frac{\gamma_k}{\gamma_0}$, la fonction d'autocorrélation, on a :

$$\rho_k = \begin{cases} \frac{-\theta_k + \theta_1 \theta_{k+1} + \dots - \theta_{q-k} \theta_k}{1 + \theta_1^2 + \dots + \theta_q^2} & \text{si } 1 \leq k \leq q \\ 0 & \text{si } k > q \end{cases}$$

Propriété 2.2.2. Pour un processus $MA(q)$, $\rho_k = 0 \forall k > q$. En d'autres termes, pour un processus $MA(q)$, les autocorrélations s'annulent à partir du rang $q+1$.

2.3 Définition des processus ARMA (p,q)

Ces processus constituent une extension naturelle des processus AR et MA. Ce sont des processus mixtes au sens où ils incorporent simultanément des composantes AR et MA.

Définition 2.3.1. Un processus stationnaire X_t suit un ARMA (p,q) s'il vérifie la relation suivante :

$$X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \dots - \phi_p X_{t-p} = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$$

Où les coefficients ϕ_i ($i=1, \dots, p$) et θ_j ($j=1, \dots, q$) sont des réels et $\varepsilon_t \sim BB(0, \sigma_\varepsilon^2)$.

En introduisant l'opérateur retard, la relation précédente s'écrit :

$$\text{Soit } \Phi(L)X_t = \Theta(L)\varepsilon_t$$

$$\text{Avec : } \Phi(L) = 1 - \phi_1L - \phi_2L^2 - \dots - \phi_qL^q \text{ et } \Theta(L) = 1 - \theta_1L - \theta_2L^2 - \dots - \theta_qL^q$$

Une extension des processus ARMA : les processus ARMA intégrés Les processus ARMA (p,q) ne peuvent être estimés que sur des séries stationnaires. Si X_t est une série à non stationnarité stochastique, il convient d'utiliser un processus ARIMA (p,d,q) où d est l'ordre de différenciation (ou d'intégration).

$$\Phi(L)(1 - L)^d X_t = \Theta(L)\varepsilon_t$$

Où d est un entier et ε^t est le bruit blanc.

Si la série $\gamma_t = (1 - L)^d X_t$ stationnaire, alors estimer un processus ARIMA (p,d,q) sur X_t est équivalent à estimer un processus ARMA (p,q) sur γ_t .

2.4 Validation des processus ARMA

La modélisation est un procédé itératif. Après avoir estimé plusieurs fois un modèle ARMA, il convient de valider les paramètres obtenus afin de départager les modèles obtenus. Pour cela, on applique les tests sur les paramètres et sur les résidus. Si plusieurs modèles sont validés, les critères d'information servent à peaufiner la comparaison.

Tests sur les paramètres

Le premier test que l'on peut mener consiste à tester l'hypothèse nulle $p' = p - 1$ et $q' = q$. On regarde si l'on peut diminuer d'une unité le nombre de retards intervenant dans la partie AR.

En d'autres termes, on teste l'hypothèse nulle de processus ARMA (p-1,q) ($\phi_p = 0$) contre l'hypothèse alternative de processus ARMA (p,q) ($\phi_p \neq 0$).

Pour opérationnaliser le test, on calcule la statistique

$$t_{\hat{\phi}_p} = \frac{\hat{\phi}_p}{\hat{\sigma}_{\hat{\phi}_p}}$$

La statistique $t_{\hat{\phi}_p}$ ainsi définie suit une loi de Student à (T-h) degrés de liberté, h étant le nombre de paramètres estimés et T le nombre d'observation de X_t .

Tests sur les résidus

Ces tests ont pour objet de vérifier que les résidus estimés suivent bien un processus de bruit blanc. A cette fin, on peut appliquer des tests d'absence d'autocorrélation et des tests d'homoscédasticité.

Tests d'absence d'autocorrélation

Il existe un grand nombre de test d'absence d'autocorrélation. Cependant, dans le cadre de ce travail, nous présentons un des plus connus d'entre eux à savoir le test de Box et Pierce.

Test de Box et Pierce

Ce test est encore appelé test de portmanteau. Il a pour objectif de tester le caractère non auto-corrélé des résidus. La statistique de ce test s'écrit :

$$BP(K) = T \sum_{k=1}^K \hat{\rho}_{k(\hat{\varepsilon}_t)}^2$$

Où $\hat{\rho}_{k(\hat{\varepsilon}_t)}^2$ est le coefficient d'autocorrélation d'ordre k des résidus estimés et K est le nombre maximal de retards.

La statistique BP(K) suit une loi de Khi-deux à (K-p-q) degrés de liberté.

L'hypothèse nulle de ce test est $\hat{\rho}_{1(\hat{\varepsilon}_t)} = \hat{\rho}_{2(\hat{\varepsilon}_t)} = \dots = \hat{\rho}_{k(\hat{\varepsilon}_t)} = 0$.

Tests d'homoscédasticité

Les tests d'homoscédasticité sont multiples et dépendent dans la plupart des cas de la cause de l'hétéroscédasticité des résidus. Dans le cadre de notre travail, nous présentons uniquement le test de White.

Test de White

Ce test est fondé sur l'existence d'une relation entre le carré du résidu et une ou plusieurs variables explicatives en niveau et au carré. Cette relation se traduit par l'équation ci-après :

$$\varepsilon_t = a_0 + a_1 X_{t-1} + b_1 X_{t-1}^2 + a_2 X_{t-1} + b_2 X_{t-2}^2 + \dots + a_p X_{t-p} + b_p X_{t-p}^2 + v_t$$

Si au moins un des coefficients de régression est significatif, on rejette l'hypothèse nulle d'homoscédasticité en faveur de l'hypothèse d'hétéroscédasticité.

Pour effectuer ce test, on utilise la statistique du multiplicateur de Lagrange TR^2

Où T est le nombre d'observation et R^2 est le coefficient de détermination associé à la régression ci-dessus. Cette statistique suit une loi du Khi-deux à 2p degrés de liberté.

Les critères de choix de modèles

Lorsque les tests sur les paramètres et les résidus sont validés, il est possible que plusieurs modèles soient retenus. Les critères de choix des modèles permettent alors de ne retenir que le meilleur d'entre eux.

En statistique, les critères de choix des modèles sont plus connus sous le vocable de critère d'information. Nous présentons ici les plus utilisés d'entre eux.

- Le critère d'information d'Akaike

$$AIC = \log \hat{\sigma}_{\hat{\varepsilon}}^2 + \frac{2(p+q)}{T}$$

- Le critère d'information bayésien d'Akaike ou de Schwarz

$$SIC = \log \hat{\sigma}_{\hat{\varepsilon}}^2 + (p + q) \frac{\log T}{T}$$

- Le critère d'information de Hannan-Quinn

$$HQ = \log \hat{\sigma}_{\hat{\varepsilon}}^2 + \alpha(p + q) \log\left(\frac{\log T}{T}\right)$$

Où α ($\alpha > 2$) est une constante, p est l'ordre de la partie AR et q l'ordre de la partie MA. On cherche à minimiser ces différents critères. Leur application nous permet de retenir un modèle parmi les divers processus ARMA validés.

2.5 Prédiction des processus ARMA

Considérons un processus ARMA (p, q) :

Soit : $\Phi(L)X_t = \Theta(L)\varepsilon_t$

Et notons \hat{X}_{t+k} la prédiction faite en t pour la date $t+k$, k désignant l'horizon de prédiction de prédiction. Par définition, on a l'expression suivante :

$$\hat{X}_{t+k} = E(X_{t+k}|I_t)$$

Où I_t est l'information disponible à la période t soit $I_t = (X_1, X_2, \dots, X_t, \varepsilon_1, \varepsilon_2, \dots, \varepsilon_t)$.

Remarque 2.5.1. Dans le cas linéaire, l'espérance conditionnelle conduit à une fonction de régression.

MODÉLISATION ET PRÉVISION DE LA PROLIFÉRATION DE LA JACINTHE D'EAU

Introduction

La jacinthe d'eau est une plante pérenne (à croissance indéterminée.) (TSAGUE, 2013)[9]. Son taux de croissance varie en fonction de la densité du tapis végétal développé et de certaines conditions environnementales, telles que la température de l'eau et sa concentration en éléments nutritifs (TSAGUE, 2013). Cependant la croissance de la jacinthe d'eau peut être influencée par la densité végétale (TSAGUE, 2013)[9]. Les résultats obtenus par Reddy et Debusk (1984) montrent une baisse de la vitesse de croissance relative de la jacinthe d'eau ceci faisant suite à une augmentation de la densité végétale. Lorsque les plants sont jeunes et en faible densité, la plante privilégie le développement de son système racinaire au détriment de la croissance de son appareil végétatif. Puis, au fur et à mesure de sa croissance, la plante réduit sa production de racines et augmente sa biomasse en se reproduisant par voie végétative (production de stolons et de bourgeons) et par la croissance des plants existants. Enfin, lorsqu'elle atteint une densité maximale, la plante stoppe sa reproduction par voie végétative et continue à augmenter sa biomasse par la croissance des plants existants (TSAGUE, 2013)[9]. A terme, si la croissance aboutit à de trop fortes densités, la mortalité des plants s'ensuit. Cette espèce est capable de doubler sa biomasse tous les 6 à 18 jours en fonction des conditions environnementales (TSAGUE, 2013)[9]. Ces caractéristiques biologiques de *E. crassipes* constituent une menace non pas seulement pour les autres espèces mais également pour la physicochimie du biotope. L'intensité de cette croissance s'évalue soit par la surface de l'eau couverte durant une période donnée, soit par la densité ou le poids des plantes par unité de surface (TSAGUE, 2013)[9]. Dans les sites naturels de la jacinthe d'eau, cette densité est relativement faible (10 Kg/m² en poids frais), tandis que dans les régions d'infestation, elle atteint 60 Kg/m² (TSAGUE, 2013)[9]. Dans ce chapitre, nous vous présentons, à partir des résultats expérimentaux, l'intensité de la croissance de la jacinthe d'eau évalué par la surface de l'eau couverte durant une période donnée.

3.1 Application à l'étude de la croissance de la jacinthe d'eau

Les données utilisées dans cette partie ont été collectées par le Dr. KENFACK Sonia Nadège chercheur à l'Herbier National du Cameroun.

logiciel utilisé : EWIEWS

Considérons les données sur l'évolution de la jacinthe d'eau au cours de 132 semaines allant du 30/juin/2014 au 02/janvier/2017. Cette série est notée SURFACE et est représentée sur la figure 3.1 ci dessous. Nous allons utiliser la méthode de Box et Jenkins présentée dans le chapitre précédent pour décrire le processus qui gouverne cette dynamique.

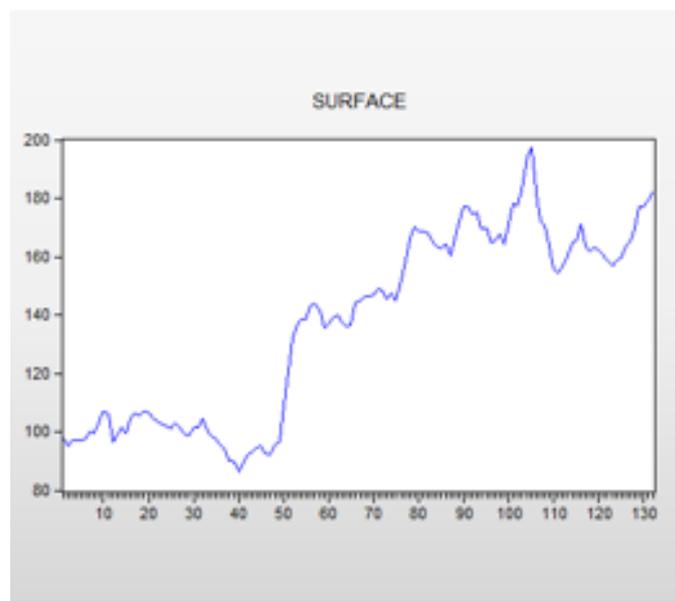


FIGURE 3.1 – Graphe de la série SURFACE

Le graphe ci-dessus fait ressortir une tendance globale à la hausse avec une possibilité de rupture de tendance après la cinquantième semaine d'observation. La série SURFACE semble non stationnaire. Cette intuition devra être confirmée ou infirmée par l'étude du corrélogramme de la série et des tests appropriés.

Sur la figure 3.2, la fonction d'autocorrélation est notée AC et la fonction d'autocorrélation partielle PAC. Q-Stat est la valeur de la statistique de test d'absence d'autocorrélation de Ljung-Box. Enfin, Prob est la probabilité d'accepter l'hypothèse nulle de nullité des autocorrélations.

Le corrélogramme de la série brute montre que les autocorrélations sont toutes significatives et décroissent très lentement. Cette remarque laisse présager que la série des observations obtenues à l'issue de notre collecte est non stationnaire.

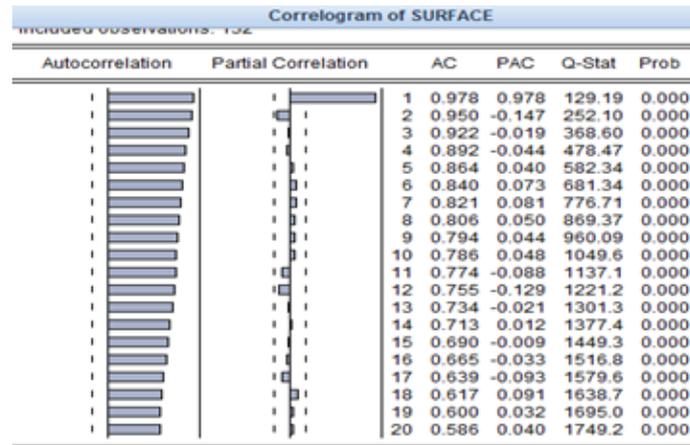


FIGURE 3.2 – Corollégramme et corollégramme partiel de la variable surface

Par ailleurs, la courbe représentant l'évolution de la série en différence première semble être stationnaire. En effet, l'observation du corrélogramme de la série en différence première ne présente plus aucune structure particulière.

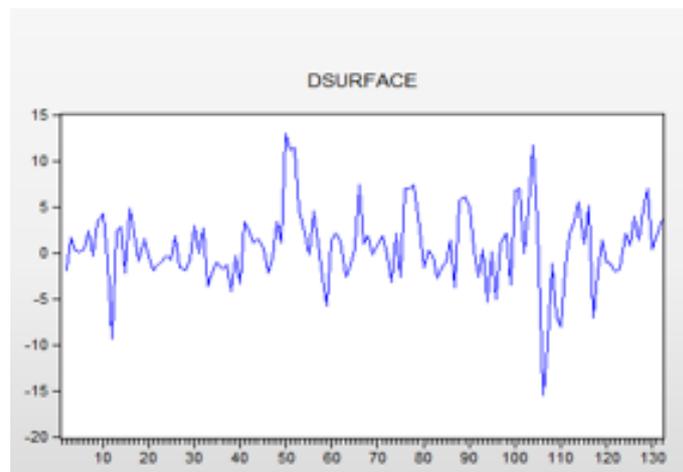


FIGURE 3.3 – Graphe de la série en différence première

Pour vérifier l'hypothèse de stationnarité nous allons utiliser la stratégie de Dickey-fuller consistant à tester en premier lieu l'hypothèse nulle de racine unitaire sur la série différenciée. Si l'hypothèse nulle est rejetée, nous testons, en second lieu, l'hypothèse nulle de racine unitaire sur la série non différenciée.

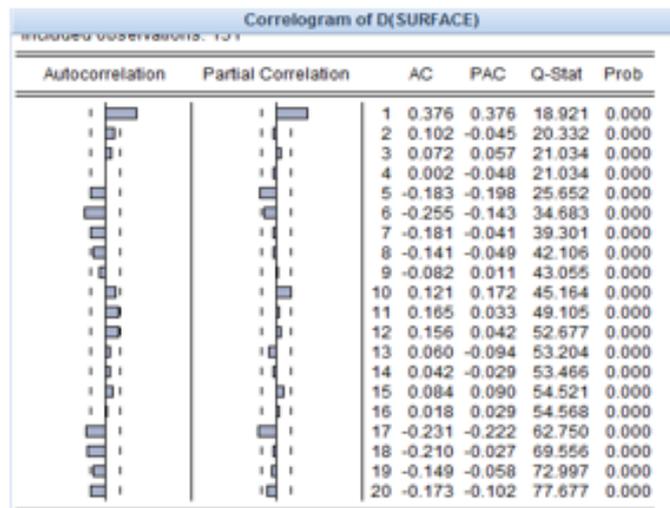


FIGURE 3.4 – Correlogramme de la série en différence première

Null Hypothesis: DSURFACE_OCCUPEE has a unit root
 Exogenous: Constant, Linear Trend
 Lag Length: 0 (Fixed)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-7.575198	0.0000
Test critical values:		
1% level	-4.030157	
5% level	-3.444756	
10% level	-3.147221	

*MacKinnon (1996) one-sided p-values.

FIGURE 3.5 – Graphe de la série en différence première

La statistique de Dickey-Fuller en valeur absolue est supérieure à la statistique de MacKinnon en valeur absolue au seuil de cinq pour cent. Ce résultat nous emmène à conclure que la série est stationnaire.

À présent que nous avons établi que la série des différences premières est stationnaire, on se propose de la modéliser par un processus de type ARMA. Pour cela, on reprend les quatre étapes de la méthodologie de Box et Jenkins.

Étape 1 : identification du modèle

Cette première étape s'effectue par le biais de l'étude des fonctions d'autocorrélation et autocorrélation partielle de la série différenciée.

Le graphique des fonctions d'autocorrélation et autocorrélation partielle de la série différenciée indique que :

- la première autocorrélation partielle de la série différenciée est significativement différente de zéro : on déduit que $p=1$
- la première autocorrélation de la série différenciée est significativement différente de zéro : on déduit que $q=1$

On vient donc d'identifier trois processus : AR (1), MA (1) et la combinaison des deux, à savoir un ARMA (1,1). L'estimation de ces trois modèles est déclinée ci-après.

Etape 2 : estimation du modèle

On procède à l'estimation des trois processus précédemment identifiés : le processus AR (1), le processus MA (1) et le processus ARMA (1,1).

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	0.393108	0.081148	4.844309	0.0000
R-squared	0.131285	Mean dependent var		0.670769
Adjusted R-squared	0.131285	S.D. dependent var		4.117131
S.E. of regression	3.837367	Akaike info criterion		5.535112
Sum squared resid	1899.575	Schwarz criterion		5.557170
Log likelihood	-358.7823	Hannan-Quinn criter.		5.544075
Durbin-Watson stat	1.960332			

FIGURE 3.6 – Modèle ar (1)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
MA(1)	0.383485	0.080928	4.738601	0.0000
R-squared	0.125959	Mean dependent var		0.650382
Adjusted R-squared	0.125959	S.D. dependent var		4.107898
S.E. of regression	3.840481	Akaike info criterion		5.536677
Sum squared resid	1917.408	Schwarz criterion		5.558625
Log likelihood	-361.6523	Hannan-Quinn criter.		5.545595
Durbin-Watson stat	1.927750			

FIGURE 3.7 – Modèle ma (1)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	0.258427	0.212387	1.216774	0.2259
MA(1)	0.157208	0.217088	0.724166	0.4703
R-squared	0.133578	Mean dependent var		0.670769
Adjusted R-squared	0.126810	S.D. dependent var		4.117131
S.E. of regression	3.847239	Akaike info criterion		5.547854
Sum squared resid	1894.560	Schwarz criterion		5.591969
Log likelihood	-358.6105	Hannan-Quinn criter.		5.565779
Durbin-Watson stat	1.997870			

FIGURE 3.8 – Modèle arma (1,1)

Au regard des résultats d'estimation, on constate que :

- le modèle AR (1) reste candidat dans la mesure où le coefficient autorégressif d'ordre 1 est significativement différent de zéro. En effet, le t de Student est supérieur à 1,96 ;

- le modèle MA (1) est également candidat puisque le coefficient est significativement différent de zéro ;
- les coefficients autorégressif et moyenne mobile du processus ARMA (1,1) n'étant pas significatifs, ce modèle est rejeté.

En conséquence, à l'issue de l'étape d'estimation, deux modèles sont encore candidats à l'explication de la dynamique de la Jacinthe d'eau dans le littoral camerounais : les processus AR (1) et MA (1).

Etape 3 : validation du modèle

Nous appliquons deux types de test sur les résidus : le test d'absence d'autocorrélation de Ljung-Box et le test d'homoscédasticité de White. Les autocorrélations et les autocorrélations partielles associées aux résidus des processus AR (1) et MA (1) sont reportés respectivement dans les figures ci-dessous.

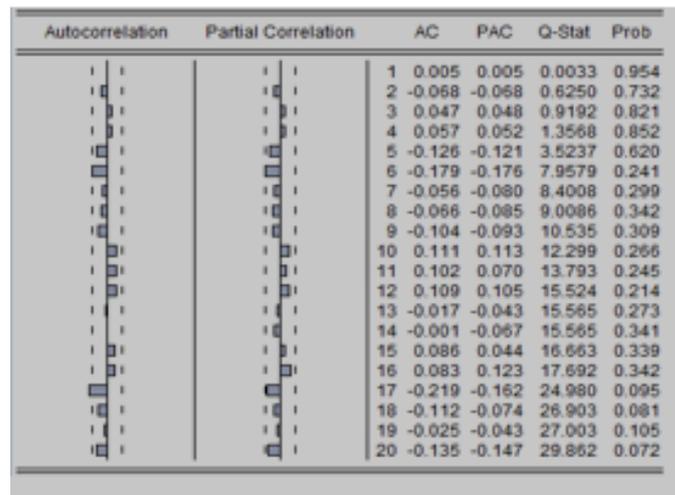


FIGURE 3.9 – Corrélogramme et corrélogramme partiels des résidus du modèle ar (1)

Tests d'absence d'autocorrélation des résidus Nous effectuons le test de Ljung-Box d'absence d'autocorrélation des résidus pour un nombre de retard maximal K égal à 20.

On peut déjà remarquer graphiquement que les autocorrélations des deux processus estimés se situent à l'intérieur de l'intervalle de confiance, ce qui semble indiquer que les résidus ne sont pas autocorrélés. Vérifions cette observation graphique à l'aide du test de Ljung-Box.

- Pour les résidus du processus AR (1) : la statistique du test de Ljung-Box vaut 29,862 pour un nombre de retards maximal égal à 20. On sait que cette statistique suit une loi du Khi-deux à (K-p-q) degré de liberté, soit ici 19. La valeur théorique du Khi-deux à 19 degrés de liberté est 30,14 au seuil statistique de cinq pourcent. En conséquence, on accepte l'hypothèse nulle d'absence d'autocorrélation des résidus ;

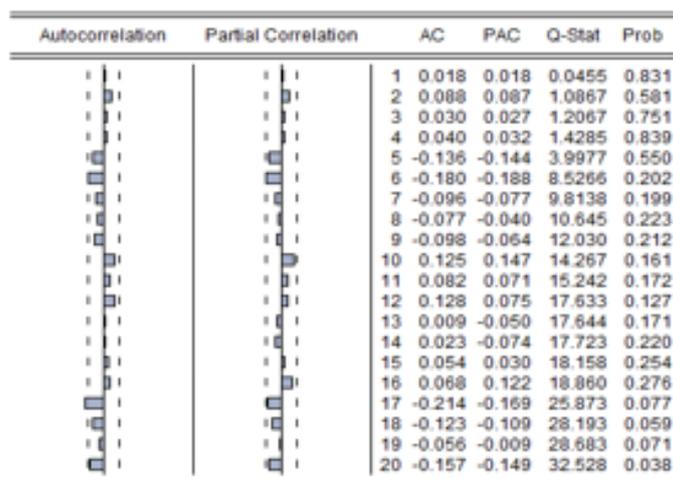


FIGURE 3.10 – Corrolégramme et corrélogramme partiels des résidus du modèle ma (1)

- Pour les résidus du processus MA (1), on effectue un raisonnement similaire. On a alors 32,528 qui est supérieur à 30,14 : les résidus du processus MA (1) sont autocorrélés.

Tests d’homoscédasticité des résidus

Pour effectuer le test d’homoscédasticité, nous allons utiliser le test de White qui a été décrit dans le chapitre précédent.

- Pour les résidus du processus AR (1) : la statistique du test associée à la régression des résidus estimés $\hat{\varepsilon}_t^2$ sur une constante, sur la série différenciée $DSurface_{t-1}$ et sur la série $DSurface_{t-1}^2$ est donnée par : $TR^2 = 4,6782$;
- Pour le processus MA (1), la statistique de test est donnée par : $TR^2 = 4,4221$.

Dans le chapitre précédent, nous avons montré que sous l’hypothèse nulle d’homoscédasticité, la statistique TR^2 suit une loi du Khi-deux à $2p$ degrés de liberté, p étant le nombre de régresseurs (hors la constante) dans l’équation du test.

Sachant qu’au seuil statistique de 5 pourcent on a $\chi_2^2=5,99$, on accepte l’hypothèse nulle d’homoscédasticité des résidus pour les deux modèles estimés.

Il ressort des résultats que les résidus des deux processus estimés sont non autocorrélés et homoscédastiques. Afin de retenir le meilleur modèle, nous aurons recours aux critères d’information.

Utilisation des critères d’information

Pour choisir un modèle dont les paramètres ont été validés au cours des étapes précédentes, il convient de retenir le modèle pour lequel les critères d’information sont les plus faibles

Au regard du tableau ci-dessus, tous les critères d’information nous conduisent à choisir le processus AR (1) car c’est lui qui présente les critères d’information les plus faibles.

Au terme de ce processus ayant conduit à la validation du modèle, la dynamique d’évolution de

	Modèle AR (1)	Modèle MA (1)
Akaike info criterion	5.535112	5.536677
Schwarz criterion	5.557170	5.558625
Hannan-Quinn criter	5.544075	5.545595

FIGURE 3.11 – Critère de choix des modèles

la Jacynthe d'eau dans le Littoral camerounais est donnée par la relation :

$$X_t = 0,393108X_{t-1} + \epsilon_t$$

Avec X_t qui représente la surface occupée par la Jacynthe d'eau au cours de la semaine t.

Etape 4 : prévision

La dernière étape de la méthodologie de Box et Jenkins a pour objet la prévision. Dans le cas de notre étude, la prévision sur la dynamique de la croissance de la Jacynthe d'eau sera effectuée à partir du processus AR (1) dont l'équation est :

$$X_t = 0,393108X_{t-1} + \epsilon_t$$

Au cours des 133, 134, 135 et 136 semaines, les prévisions sur la surface occupée par la Jacynthe d'eau, estimées par l'espérance conditionnelle décrite dans le chapitre précédent seront : 184,6 mètres carrés, 186,1 mètres carrés, 188,7 mètres carrés et 189,9 mètres carrés respectivement.

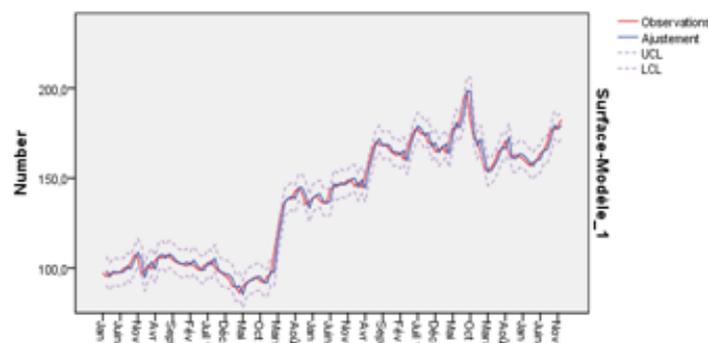


FIGURE 3.12 – Prévisions de la série

3.2 conclusion

Dans ce chapitre, il était question de trouver un modèle qui reflète le mieux possible la jacinthe d'eau. Pour ce faire, nous avons considéré un modèle de type AR(1). Les prévisions obtenues sur la dynamique de la Jacynthe d'eau sont bonnes car les courbes des prévisions et les observations initiales se situent à l'intérieur de l'intervalle de prévision (au seuil statistique de 5 pourcent).

Implication pédagogique

La préservation de l'environnement est considérée comme un enjeu majeur pour tout être vivant de notre ère. Aujourd'hui plus que par le passé, l'arbitrage entre « développement économique » et le « développement durable » est le défi des pays à l'échelle mondiale. A l'échelle micro-économique, l'élève ou l'étudiant en tant qu'être rationnel est confronté à de nombreux choix ; ce qui lui impose d'avoir une connaissance sur la dynamique d'évolution des êtres vivants animaux et végétaux qui l'entourent.

L'élève dans sa vie scolaire est ainsi appelé à étudier l'évolution des phénomènes démographiques (naissance, croissance, migration) chez les animaux et végétaux qui l'entourent. Face à la complexité des facteurs explicatifs de la dynamique des êtres vivants, les outils mathématiques enseignés aux élèves du secondaire et qui permettent d'établir un lien entre l'ampleur d'un phénomène entre deux périodes restent les fonctions numériques et les suites numériques.

Une petite enquête menée auprès de certains élèves des classes de première et de terminale du secondaire et portant sur la question de savoir quel(s) outil(s) peut-on utiliser pour prévoir la population du Cameroun en 2020 sachant que l'on possède l'effectif de la population camerounaise entre 1960 et 2015 a conduit aux résultats suivants : les fonctions numériques ou les suites numériques.

En interprétant ce résultat, nous comprenons que les seules méthodes que nos élèves connaissent sont les fonctions ou les suites numériques. Cela nous amène à nous poser certaines questions : les fonctions ou les suites numériques qui sont mentionnées dans les programmes d'études du secondaire peuvent-elles être critiquées ? Certainement non, car on les compareraient à quoi ? En effet, dans la pratique, la population d'une espèce à une période $n+1$ dépend de la population de cette espèce à la période n , des naissances, des morts et du solde migratoire de cette espèce à la période $n+1$. Par conséquent, l'utilisation des fonctions numériques peu dans certains cas conduire à des prévisions « grossières ». Il importe donc dans ces cas d'affiner la méthode utilisée.

Pour éviter d'être confronté à ce type de difficulté, nous proposons de :

- Élargir les programmes scolaires portant sur la dynamique des populations en insérant une leçon portant sur les notions élémentaires de séries temporelles ;
- Insérer cette leçon en mathématiques et imposer ce chapitre à toutes les filières scientifiques ;
- Insérer dans le programme d'informatique des cours d'initiation à la manipulation des

logiciels.

Conclusion et perspectives

Il était question dans ce travail d'étudier la dynamique d'évolution de la jacinthe d'eau dans le littoral camerounais. Pour le faire, nous avons utilisé la méthode de Box et Jenkins pour estimer le modèle ARMA qui a été jugé adéquat pour décrire le phénomène étudié. Cette technique de modélisation se décline en quatre étapes : (i) l'identification du modèle, (ii) l'estimation des paramètres du modèle, (iii) la validation du modèle retenu et (iv) la prévision.

Au terme de notre travail, il ressort que la dynamique d'évolution de la jacinthe d'eau dans le littoral camerounais peut être modélisée par un descriptif de type AR(1). Dans nos futurs travaux, nous proposons d'utiliser un modèle explicatif à l'instar du modèle VAR pour déterminer les facteurs qui influencent l'évolution de cette plante dans les côtes camerounaises.

Bibliographie

- [1] Arthur Charpentier. Modèles de prévision-séries temporelles. *Paris, France*, 2012.
- [2] Karim Dagno, Rachid Lahlali, Damien Friel, Mohammed Bajji, and M Haïssam Jijakli. Synthèse bibliographique : problématique de la jacinthe d'eau, eichhornia crassipes, dans les régions tropicales et subtropicales du monde, notamment son éradication par la lutte biologique au moyen des phytopathogènes. *Biotechnologie, Agronomie, Société et Environnement*, 11(4) :299–311, 2007.
- [3] Jean-Yves DAUXOIS. Master recherche de mathématiques probabilités.
- [4] I GANNAZ. Introduction à la statistique.
- [5] Sandrine Lardic and Valérie Mignon. *Econométrie des séries temporelles macroéconomiques et financières*. Economica, 2002.
- [6] Michel LUBRANO. Introduction à la modélisation des séries temporelles univariées. *Université de la méditerranée, Air-Marseille*, 2, 2008.
- [7] Corinne Perraudin. Séries chronologiques. *Quelques éléments du cours, Université Paris*, 1 :2004–2005, 2004.
- [8] Ariane PRADELLE. Etude des faisabilités d'un projet de compostage de la jacinthe d'eau dans la communauté de sô-ava au benin.
- [9] Hugue romuald TSAGUE. croissance et écologie de eichhorgnia crassipes (martius) solms-laubach dans quelques plans d'eau de douala-oest.