

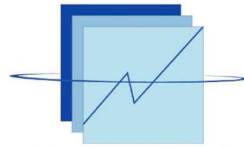
**RÉPUBLIQUE DU SÉNÉGAL**

Un peuple- un But- une Foi



\*\*\*\*\*

**Agence nationale de la Statistique et de la Démographie**



**ANSD**

Agence Nationale de  
la Statistique et de la Démographie

**École nationale de la Statistique et de l'Analyse Économique Pierre Ndiaye**



**Projet d'économétrie des séries temporelles**

---

Analyse des dynamiques de la relation entre les investissements directs étrangers, l'aide publique au développement, la formation brute du capital fixe et la croissance économique dans le contexte Camerounais

---

*Rédigé par :*

**Julien Parfait BIDIAS ASSALA**

ÉLÈVE INGÉNIEUR STATISTICIEN ÉCONOMISTES

*Sous la supervision de :*

**Dr. SOULEYMANE FOFANA**

ENSEIGNANT À L'ENSAE

**©Juillet-2023**

# Sommaire

|  |            |
|--|------------|
| <b>Résumé</b>  | <b>III</b> |
| <b>I Dynamique du PIB camerounais : approche par la méthode de Box-Jenkins</b> | <b>IV</b>  |
| I.1 Approche méthodologique . . . . .  | IV         |
| I.2 Présentation des données . . . . .   | V          |
| I.3 Analyse préliminaire : la stationnarité . . . . .                          | VI         |
| I.4 Identification du modèle . . . . .   | VIII       |
| I.5 Estimation des paramètres . . . . .  | IX         |
| I.6 Validation par tests ou tests d'adéquation du modèle . . . . .             | IX         |
| I.7 Vérification des hypothèses du modèle-analyses des résidus . . . . .       | X          |
| I.8 Prévisions dans le modèle . . . . .  | XIV        |

## Liste des figures

|   |  |      |
|---|--|------|
| 1 | Box-plot du taux de croissance . . . . .   | V    |
| 2 | Evolution du taux de croissance . . . . .  | VI   |
| 3 | Fonction d'autocorrélation partielle . . . . .                                       | VIII |
| 4 | Fonction d'autocorrélation . . . . .   | IX   |
| 5 | Fonction d'autocorrélation simple (gauche) et partielle des résidus (droite) . . . . | XII  |
| 6 | Courbe des résidus . . . . .   | XIII |
| 7 | Évolution continue du taux de croissance avec prévisions . . . . .                   | XIV  |
| 8 | Densité des résidus . . . . .  | XV   |

## Liste des tableaux

|   |   |     |
|---|---|-----|
| 1 | Tendances centrales, dispersion et forme . . . . .                  | V   |
| 2 | Résultats du test de Dickey-Fuller augmenté avec tendance . . . . . | VII |
| 3 | Résultats du test de Dickey-Fuller augmenté sans tendance . . . . . | VII |
| 4 | AIC et BIC . . . . .  | X   |
| 5 | Prévisions des taux de croissance . . . . .                         | XIV |

## Résumé

La présente étude a pour objectif de décrire la dynamique du taux de croissance du PIB réel du Cameroun. Pour ce faire, nous avons d'abord considéré la série des taux de croissance avec sa trajectoire passée et toutes les perturbations liées via une approche Box-Jenkins.

Il ressort que le taux de croissance du PIB du Cameroun évolue tel un modèle auto régressif moyenne mobile et que celui-ci n'est pas pur du fait de la présence de volatilité. Le taux de croissance du Cameroun est estimé pour l'année 2021 à 4,06% contre sa vraie valeur 3,64%<sup>a</sup>.

La modélisation des séries temporelles dans le cas univariée permet ainsi de définir des orientations en matière de politiques de croissance potentielle.

---

<sup>a</sup>INS Cameroun, 2021.

# I Dynamique du PIB camerounais : approche par la méthode de Box-Jenkins

Le recours aux techniques de traitement via l'économétrie des séries temporelles permet de retracer la dynamique de plusieurs phénomènes supposés se répéter dans les mêmes conditions. Les séries temporelles (ou séries chronologiques) portent donc sur l'analyse statistique et probabiliste d'observations ordonnées chronologiquement. Cela permet de justifier l'étude à travers le temps de l'évolution du PIB camerounais par ces mêmes techniques. L'objectif étant de pouvoir à travers une analyse, une modélisation rigoureuse caractériser ou définir le modèle qui permet aux mieux d'expliquer les périodes de fortes hausses et les périodes de fortes baisses de ce dernier. En outre, les modèles ne sont pas conçus pour refléter la réalité à 100% mais pour la décrire, prévoir le phénomène étudié dans le futur. C'est ainsi qu'il existe une kyrielle de modèles abondant dans la littérature. Dès lors, on peut donc s'interroger directement sur quel modèle choisir ?

En analyse univariée, l'on considère le plus souvent qu'un phénomène présent peut dépendre dans le futur de ses valeurs passées : il s'agit des processus autorégressifs ( $AR^1$ ). Ou encore, de ses perturbations antérieures : il s'agit des processus moyennes mobiles ( $MA^2$ ). Enfin, le processus peut être une combinaison des deux ( $ARMA$ ) et présenter dans un cas bien précis une saisonnalité on parlera de modèles  $ARIMA$ . Cette typologie de modèles fera donc l'objet de cette section dans laquelle nous choisirons à travers l'inférence statistique celui adéquat, permettant d'expliquer les fluctuations de croissance du PIB du Cameroun de la période 1970 à 2020 et aussi de pouvoir se prononcer en termes de perspectives.

## I.1 Approche méthodologique

L'approche retenue pour étudier notre série et choisir donc le bon modèle porte essentiellement sur celle de Box-Jenkins (1976) compléter par Anderson (1977), Ljung et Box (1978), Roy (1982). La méthode se décompose en 7 étapes (Anderson, 1977) :

- familiarisation avec les données,
- analyse préliminaire,
- identification ou spécification du modèle,
- estimation des paramètres,
- validation par tests ou test d'adéquation du modèle,
- prévision,
- interprétation des résultats.

---

<sup>1</sup>Auto regressif

<sup>2</sup>Mobile Avrage

## I.2 Présentation des données

Les données qui feront l'objet de notre analyse proviennent du site du groupe la Banque Mondiale. Il s'agit des taux de croissance du PIB réel de la période 1976 à 2020. Il s'agit donc de données quantitatives, positives ou négatives et indexées par le temps. Aussi, La série ne présentent aucune valeur manquante. Aussi, il important de préciser qu'à ce jour, les taux de croissance des années 2021 et 2022 sont déjà connus et valent respectivement 3,6% et 3,5% selon la Banque Mondiale. Ainsi, nous travaillerons avec nos données sur la période 1976 à 2020, ensuite nous fournirons des estimations du taux de croissance du PIB camerounais pour les années ultérieures.

### I.2.1 Analyse descriptive

Le tableau ci-dessous résume un ensemble de caractéristiques spécifique du taux de croissance du PIB.

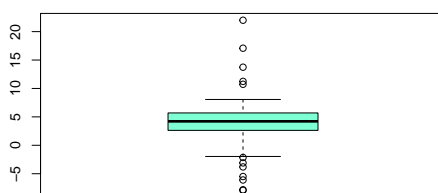
Table 1: Tendances centrales, dispersion et forme

| variable            | moyenne | mediane | écart-type | min   | max   | kurtosis | skewness | coefficient de variation |
|---------------------|---------|---------|------------|-------|-------|----------|----------|--------------------------|
| Taux croissance PIB | 3,90    | 4,22    | 5,47       | -7,93 | 22,00 | 0,43     | 1,81     | 0,71                     |

Source : Groupe de la Banque mondiale, calcul des auteurs

La valeur moyenne du taux de croissance du PIB se situe à 3,9%, c'est une statistique très importante car elle permet de caractériser la notion d'ergodicité d'une série temporelle c'est-à-dire Calculer l'ensemble des caractéristiques d'une série à partir d'une certaine trajectoire. Aussi, avec la présence de valeurs extrêmes, on ne peut apprécier ce résultat à sa juste valeur. La valeur médiane des taux de croissance du PIB révèle que pendant plus de la moitié de la période d'étude le taux de croissance été de 4,22%. Cette valeur est relativement proche de celle du PIB. On pourrait croire une présomption de symétrie de la série, ce qui traduirait un comportement similaire de la série autour de sa valeur moyenne. La valeur maximale de la série est égale à 22% contre 3,6% cela sous-entend une présence de valeurs aberrantes. C'est que le graphe ci-dessous permet de confirmer.

Figure 1: Box-plot du taux de croissance

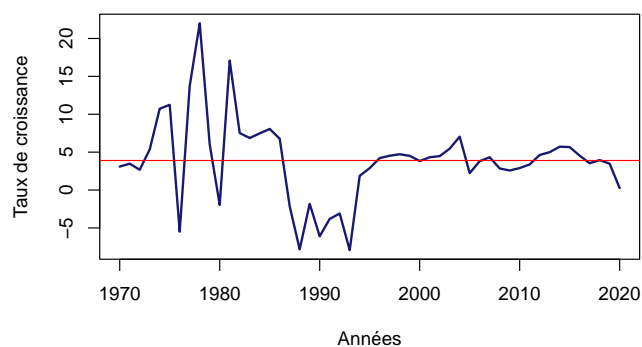


Source : Groupe de la Banque Mondiale, calcul de l'auteur

Nous avons des valeurs aberrantes mais en quantité insuffisante. Aussi, leur présence se justifie de part l'alternance des signes positifs et négatifs que prend notre série. Ainsi, la présence d'une valeur négative viendra rendre une valeur positive aberrante. La problème du traitement de ces valeurs ne se pose donc pas avec acuité, ce d'autant plus qu'elles ne sont pas nombreuses.

### I.2.2 Evolution de la croissance économique Camerounaise

Figure 2: Evolution du taux de croissance



source : Groupe de la Banque Mondiale, calcul de l'auteur

Pour apprécier la dynamique du taux de croissance nous allons tout d'abord commencer par une analyse graphique.

Un constat saillant est celui de la stagnation du PIB du Cameroun de l'année 1990 à l'année 2010. Ce comportement est principalement dû aux PAS<sup>3</sup> dont l'objectif était l'efficacité économique, la rationalisation des dépenses publiques. Ensuite, Ces programmes ont eu un effet durant une longue période. D'aucun estime que ces effets étaient plutôt positifs et d'autre part négatifs. De plus, on constate que la série fluctue beaucoup autour de sa valeur moyenne représentée en rouge. Cette valeur vaut précisément 3,9%. Ce comportement est assimilable à celui d'un **processus stationnaire**. Mais pour confirmer cela il faudra effectuer tests statistiques.

### I.3 Analyse préliminaire : la stationnarité

L'analyse préliminaire portera sur l'étude de la stationnarité de la série des taux de croissances. Celle-ci va nous permettre d'identifier par la suite le(s) modèle(s) potentiels adaptés à son comportement.

**La stationnarité** est un concept clé pour la validité extrême des séries temporelles. Une série est dite stationnaires si ses réalisations sont issues d'un même processus stochastique dont les paramètres (moyenne, variance, asymétrie, autocorrélation...) restent constants au cours du temps. A travers la stationnarité le passé est comparable au présent et au futur.

Considérons une suite de données temporelles  $(X_t)_{t \in \mathbb{Z}}$ . On dira que  $(X_t)$  est stationnaire en moyenne si :

$$E(X_1) = E(X_2) = \dots = E(X_t) \quad (1)$$

La série sera alors dite stationnaire en variance si la variance est invariante pour tout changement d'horizon temporel. L'analyse descriptive nous a fourni une présomption de stationnarité. Toutefois, cela reste à vérifier.

Il existe plusieurs test permettant de tester la stationnarité ou de détecter la non stationnarité dans le comportement d'une série temporelle. Il en existe deux types, avec la stationnarité comme hypothèse nulle ou hypothèse alternative.

<sup>3</sup>Programmes d'ajustement structurels

### • Tests de non stationnarité : Tests de racines unitaires

On peut citer les tests de Dickey-Fuller (Dickey-Fuller simple et Dickey-Fuller Augmenté) et le test de Phillips-Perron. Les tests de racine unitaire permettent de tester l'hypothèse de présence ou non de racine unité. L'hypothèse nulle est la non-stationnarité.

### • Test augmenté de Dickey Fuller (ADF)

Le test de Dickey-Fuller simple est limité dans la mesure où l'on ne s'intéresse qu'à un modèle de type AR(1) (avec ou sans constante, avec ou sans trend) et que les perturbations sont i.i.d. Toutefois, il a été étendu, par Dickey-Fuller (1981), au cas où la série  $X_t$  admet une représentation AR(p) (les perturbations suivant alors un  $AR(p-1)$ ), puis par Saïd et Dickey (1984, 1985) au cas où l'erreur suit un processus  $ARMA(p; q)$ . Considérons le modèle suivant :

$$\Delta X_t = \phi X_{t-1} + \lambda + \delta t + \sum_{j=1}^p \gamma_j \delta_{X_{t-j}} + \eta_t \quad \text{avec} \quad \eta_t \rightarrow BB(0, \sigma_\eta^2) \quad (2)$$

Pour tester la présence de racine unitaire, on va tester la significativité du coefficient qui se trouve devant  $X_{t-1}$ . L'hypothèse initiale ou nulle est :  $H_0 : \text{Non stationnarité}, \phi = 0$ . Mais avant tout nous commençons par tester la significativité de la tendance.

Table 2: Résultats du test de Dickey-Fuller augmenté avec tendance

| lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)          | Estimate | Std. Error | t value | Pr(> t ) |     |
|---|----------|------------|---------|----------|-----|
| (Intercept)   | 3,83307  | 1,82619    | 2,099   | 0,041464 | *   |
| z.lag.1   | -0,67212 | 0,16886    | -3,98   | 0,000248 | *** |
| tt  | -0,04692 | 0,05434    | -0,863  | 0,39247  |     |
| z.diff.lag  | 0,0427   | 0,14901    | 0,287   | 0,775757 |     |
| Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 |          |            |         |          |     |

Source : Auteur

On commence par interpréter la p-value associée au paramètre  $\delta$  qui correspond au résultat sur la ligne **tt**. On constate qu'il n'est pas significatif. Alors le modèle de l'équation (2) n'est pas adapté. On va tester le modèle ci-dessous.

$$\Delta X_t = \phi X_{t-1} + \lambda + \sum_{j=1}^p \gamma_j \delta_{X_{t-j}} + \eta_t \quad \text{avec} \quad \eta_t \rightarrow BB(0, \sigma_\eta^2) \quad (3)$$

Table 3: Résultats du test de Dickey-Fuller augmenté sans tendance

| lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)               | Estimate | Std. Error | t value | Pr(> t ) |     |
|---|----------|------------|---------|----------|-----|
| (Intercept)   | 2.51783  | 1.00454    | 2.506   | 0.015793 | *   |
| z.lag.1   | -0.64820 | 0.16611    | -3.902  | 0.000309 | *** |
| z.diff.lag  | 0.03161  | 0.14805    | 0.214   | 0.831878 |     |
| Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 |          |            |         |          |     |

Source : Auteurs

On constate que la constante  $\lambda$  sur la ligne de **l'intercept** est significative, on peut donc tester la présence de racine unitaire. Le paramètre  $\phi$  sur la ligne écrit z.lag.1 a une p-value inférieure au seuil de 5%. Il est donc significatif. On peut donc dire que notre série est stationnaire. Ainsi, le test s'arrête et on peut travailler directement sur la série ( $X_t$ ). Ce qui est en conformité avec l'analyse descriptive.



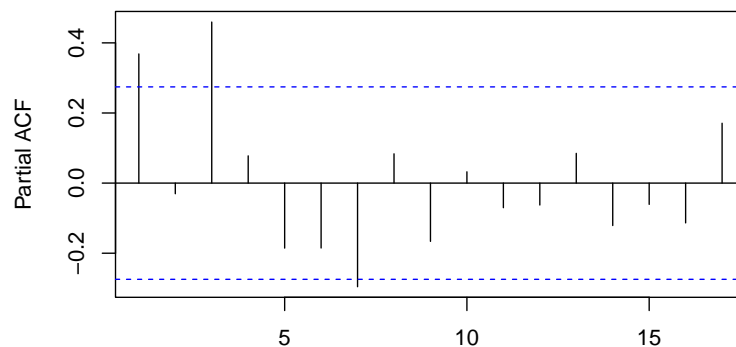
## I.4 Identification du modèle

L'étape d'identification ou de la spécification d'un processus consiste à observer la nature des fonctions empiriques d'autocorrélation et d'autocorrélation partielle, de la série elle-même, pour choisir les ordres des parties AR (choix de l'entier  $p$ ) et MA (choix de l'entier  $q$ ), ainsi que l'ordre d'intégration. Ces ordres correspondent aux valeurs de  $h$  pour lesquelles l'ACF empirique et le PACF empirique se trouvent à l'extérieur des intervalles de confiance. Par ailleurs, la série étant déjà stationnaire, elle est donc  $I(0)$  i.e intégré d'ordre zéro.

### I.4.1 Identification de l'ordre $p$ pour la partie AR

Pour déterminer l'ordre  $p$  nous allons représenter la fonction d'autocorrélation partielle (PACF<sup>4</sup>) de la série.

Figure 3: Fonction d'autocorrélation partielle



Source : Groupe de la Banque Mondiale, calcul de l'auteur

L'axe des abscisses donne les valeurs des différents ordres. On constate qu'à l'ordre zéro et à l'ordre 3 nous avons des valeurs du coefficient de corrélation significatives. Ainsi,  $p = 1, 2, 3$ .

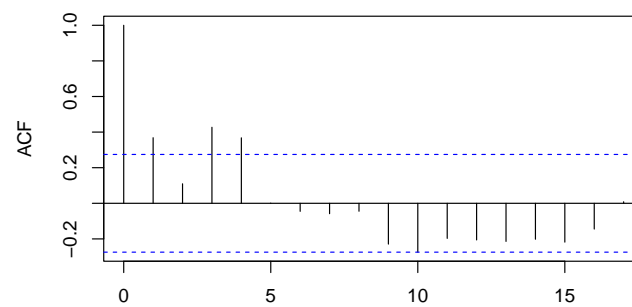
### I.4.2 Identification de l'ordre $q$ pour la partie MA

Pour identifier l'ordre  $q$  de la partie MA, nous allons représenter la fonction d'auto corrélation (ACF<sup>5</sup>) de la série. Comme nous pouvons le voir à travers la figure ci-dessous, l'ordre  $q$  sera égal à 1, 2, 3 et 4 car ce sont les ordres significatifs pour la fonction d'auto corrélation.

<sup>4</sup>Partial Auto correlation function

<sup>5</sup>Auto correlation function

Figure 4: Fonction d'autocorrélation



Source : Groupe de la Banque Mondiale, calcul de l'auteur

Ainsi, les modèles candidats sont :

- AR(1) ; AR(2); AR(3);
- MA(1) ; MA(2); MA(3); MA(4);
- ARMA(1,1) ; ARMA(1,2); ARMA(1,3); ARMA(1,4);
- ARMA(2,1) ; ARMA(2,2); ARMA(2,3), ARMA(2,4);
- ARMA(3,1) ; ARMA(3,2); ARMA(3,3), ARMA(3,4)

## I.5 Estimation des paramètres

L'estimation des paramètres est effectuée en tenant compte de tous les potentiels modèles suscités. La méthode du maximum de vraisemblance est utilisée pour l'estimation des paramètres. L'intérêt de cette méthode est qu'elle utilise toute l'information à court comme à long terme concernant les comportements des séries puisque sont estimés simultanément les paramètres autorégressif, moyenne mobile et le paramètre de différenciation fractionnaire  $d$ <sup>6</sup>. La méthode s'inscrit dans le domaine temporel.

## I.6 Validation par tests ou tests d'adéquation du modèle

L'ensemble des informations sur la modélisation du modèle optimal à la description du taux de croissance du PIB tels que les paramètres, les valeurs estimées, le test de significativité des paramètres et les métriques d'évaluation de la qualité du modèle sont fournies dans le tout premier tableau figurant en annexe.

Ainsi, nous avons d'abord après modélisation choisis les modèles ayant leurs coefficients tous significatifs. Dans notre cas, il s'agit des modèles MA(1), AR(1) et ARMA(3,2). Ils se formulent ainsi qu'il suit :

$$MA(1) \quad : \quad X_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} \quad (4)$$

$$AR(1) \quad : \quad X_t = \phi_1 X_{t-1} + \varepsilon_t \quad (5)$$

<sup>6</sup>Le nombre de différenciation

$$ARMA(3,2) : X_t = \sum_{i=1}^3 \phi_i X_{t-i} + \varepsilon_t = \sum_{j=1}^2 \theta_j \varepsilon_{t-j} \quad (6)$$

Le choix du modèle optimal dépendra des valeurs des critères d'informations AIC et BIC. Nous choisirons le modèle pour lequel l'AIC et le BIC sont très petits.

Table 4: AIC et BIC

| Modèles   | AIC      | BIC      |
|-----------|----------|----------|
| MA(1)     | 319,1399 | 319,5828 |
| AR(1)     | 320,9895 | 321,4324 |
| ARMA(3,2) | 318,3772 | 319,4107 |

Source : Auteurs

Le tableau ci-dessus nous montre que le meilleur modèle est donc le modèle ARMA(3,2) en dernière ligne car il contient les valeurs les plus petites suivant les critères AIC et BIC. Il a le plus petit AIC (339,06) et le plus petit BIC (339,895). Dans toute la suite de notre modélisation, nous travaillerons avec le modèle décrit par l'équation (6).

## I.7 Vérification des hypothèses du modèle-analyses des résidus

Avant toute tentative de prévision dans un modèle, il est important au préalable de valider un ensemble d'hypothèses permettant de garantir la qualité du modèle et donc par conséquent sa validité et la fiabilité des prévisions.

Pour ce faire, on passe par l'analyse des résidus. Il s'agit en fait de la différence entre la vraie valeur et la valeur estimée. Le comportement des résidus est très important dans la validation d'un modèle. Il permet en effet de se prononcer quant à la qualité des erreurs du modèle (est-ce qu'on se trompe moins ou alors on se trompe complètement ?). On voit dès lors que c'est un aspect très important pour la prévision.

Pour mener à bien cette section, nous analyserons d'abord la nullité des résidus car en moyenne les erreurs se compensent et donc on n'est supposé se tromper moins, des tests de non auto-corrélation des erreurs seront effectués pour confirmer l'hypothèse selon laquelle les erreurs d'une période à l'autre ne sont pas liées. Ensuite, nous poursuivrons avec le test d'homoscédasticité de la variance des erreurs car il permet de confirmer qu'elle n'est pas volatile, c'est-à-dire qu'elle est constante au cours du temps. Pour être plus explicite, cette hypothèse permet de dire qu'à différentes périodes nous n'avons pas de fortes perturbations tout comme nous n'avons pas de perturbations élevées. Dans le même temps, nous poursuivrons avec le test de stabilité des coefficients (test de stabilité structurel de Chow), viendra enfin les tests de normalités. Une fois ces hypothèses vérifiées, nous pourrions passer à la prévision.

### I.7.1 Test de nullité de la moyenne des résidus

Les hypothèses de ce test sont :

$$\begin{aligned} H_0 : \bar{e} &\equiv 0 \\ H_1 : \bar{e} &\neq 0 \end{aligned}$$

La statistique de test vaut :  $t_{\bar{e}} = \frac{\bar{e}}{\sigma_{\bar{e}}}$  qui suit sous  $H_0$  une loi de Student à  $T_{(T-(p+q+1))}$ . Cela revient à comparer  $\bar{e}$  à  $1,06 \times \sigma_{\bar{e}}$ . Or  $\bar{e} = 0,07645462$  et  $\sigma_{\bar{e}} = 4,027465$  ce qui fait que

$\bar{e} < 1,96 \times \sigma_{\bar{e}}$ . On accepte alors l'hypothèse  $H_0$ , la moyenne des résidus n'est pas significative. En moyenne les erreurs se compensent, on ne se trompe donc pas.

### I.7.2 Tests de non auto corrélation des erreurs

La littérature confère une typologie de tests assez variées pour effectuer les tests de non auto corrélation des perturbations. On distingue les tests Portmanteau<sup>7</sup>, le test de Durbin Watson, etc. Pour le cas d'espèce, nous utiliserons les tests Portmanteau car ce sont les plus utilisés.

- Tests portmanteau de Box et Pierce (1970)

Il s'agit de tester la significativité globale des auto corrélations à l'aide d'une statistique. Cette statistique est la suivante :

$$BP_K = T \sum_{k=1}^K \hat{\rho}_{\varepsilon}^2(k) \quad (7)$$

Où  $\hat{\rho}_{\varepsilon}^2(k)$  est la fonction d'auto corrélation d'ordre  $k$  des résidus estimées sur l'échantillon de la série,  $T$  est la longueur de la série et  $K$  le nombre de retard choisis pour calculer les auto corrélations.  $K = \left\lfloor \frac{T}{3} \right\rfloor = \left\lfloor \frac{51}{3} \right\rfloor = 17$ . Cette statistique de test suit sous  $H_0$  une loi de khi deux à  $K - (p + q)$  degrés de liberté. Avec  $p$  l'ordre du AR et  $q$  l'ordre du MA. Les hypothèses du test sont les suivantes :

$$\begin{aligned} H_0 : \hat{\rho}_{\varepsilon}^2(1) = \hat{\rho}_{\varepsilon}^2(2) = \dots = \hat{\rho}_{\varepsilon}^2(K) = 0 \text{ avec } K = 15 \\ \text{contre} \\ H_1 : \text{il } \exists \text{ au moins un } h \text{ telle que } \hat{\rho}_{\varepsilon}^2(h) \neq 0 \end{aligned}$$

Après mise en œuvre du test, les résultats sont tels que la p-value vaut 0.8173. Elle est largement supérieure au seuil critique de 5%. On accepte donc l'hypothèse nulle. Les coefficients de corrélations des erreurs sont non significatifs quelques soit les ordres. Il y a donc absence d'auto corrélation des erreurs.

- Ljung et Box

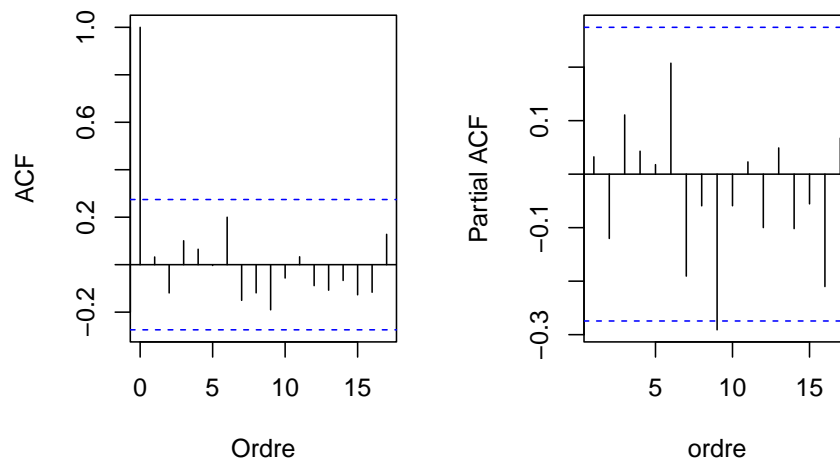
Il s'agit d'une version améliorée du test précédent lorsque la taille de l'échantillon n'est pas assez grande. Les hypothèses étant les mêmes. La statistique du test s'écrit :

$$T(T+2) \sum_{k=1}^K \frac{\hat{\rho}_{\varepsilon}^2(k)}{T-k} \quad (8)$$

Après calcul la p-value associée au test vaut 0.812. Elle est assez proche du test précédent. On conclut donc de la même façon en acceptant l'hypothèse de non auto corrélation des résidus. Par ailleurs, tous ces résultats peuvent aussi s'observer à travers les fonctions d'auto corrélations des résidus. Les résidus étant supposés par hypothèses comme des bruits blancs.

<sup>7</sup>Tests four tout

Figure 5: Fonction d'autocorrélation simple (gauche) et partielle des résidus (droite)



Source de données : Banque Mondiale, calcul des auteurs.

Ces deux figures permettent de confirmer que les résidus sont bel et bien des bruits blancs car toutes les auto corrélations sont non significatives. Elles sont à l'intérieur des intervalles de confiance qui contiennent la valeur zéro. Par ailleurs, la figure de droite celle de l'auto-corrélation partielle laisse entrevoir un ordre en dehors de cet intervalle, ce qui nous amène à penser à rejeter l'hypothèse probable d'un ARMA pur. Nous vérifions cette hypothèse plus bas lorsque nous ferons le test d'hétéroscédasticité conditionnelle.

En outre, les erreurs sont non auto corrélées comme nous l'avons montré avec les deux tests ci-dessous et leur moyenne est nulle alors c'est un bruit blanc centré qui est par définition stationnaire. Le problème soulevé ci-dessous s'explique par le fait que les résidus peuvent être non auto corrélés mais pas indépendants.

Dans toute la suite nous vérifions donc les hypothèses de normalité des résidus et d'homogénéité de la variance.

### 1.7.3 Tests de normalité des résidus

Pour tester la normalité des résidus nous utiliserons les tests paramétriques : test de Skewness et test de Jarque-Bera. Les tests non paramétriques seront aussi utilisés. Il s'agit des tests de Kolmogorov Smirnov et de Shapiro-Wilk.

Le test de Jarque-Bera a pour hypothèse nulle la normalité et ce test vérifie si le kurtosis et le skewness de la série des résidus sont proche de celui d'une loi normale. Après exécution de ce test nous avons une p-value de 0,366, ce qui nous amène donc à accepter l'hypothèse nulle de normalité des résidus. Dans ce cas le test de Skewness n'est plus nécessaire.

Le test de Shapiro-Wilk conclut au même résultat que le précédent avec une p-value de 0.08279. De même que pour le test de Kolmogorov qui prend en compte deux distributions et vérifie si elles sont en adéquation. Pour le réaliser nous avons considéré la série des résidus et celle d'une loi normale. Le test a conclu quant à la similarité des fonctions de répartition des deux lois.

### I.7.4 Tests d'homoscédasticité de la variance des erreurs

Nous proposons les tests de White et d'hétéroscédasticité conditionnelle (test de ARCH). Aussi, n'importe quel autre test adapté conviendrait à la série des résidus.

- Test de White

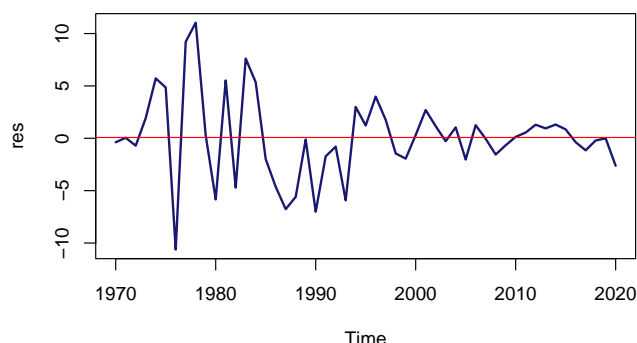
C'est le test le plus utilisé. C'est un test asymptotique, on ne peut donc l'utiliser que si n'est assez grand ( $>50$ ). White part de l'idée qu'en présence d'hétéroscédasticité la variance des erreurs est fonction des variables du modèle, de leurs carrés et des doubles produits et cela indépendamment de la forme de l'hétéroscédasticité. On en a déduit le test suivant : dans le cas d'hétéroscédasticité les résidus au carré du modèle estimé sont fonction des variables explicatives, de leurs carrés et des produits deux à deux.

$H_0$  : les erreurs sont homoscédastiques

$H_1$  : les erreurs sont hétéroscédastiques sans que l'on connaisse la forme exacte de l'hétéroscédasticité.

Après réalisation du test nous avons une p-value de 0.0603. Alors on accepte l'hypothèse nulle d'homogénéité de la variance des résidus.

Figure 6: Courbe des résidus



Source de données : Banque Mondiale, calcul des auteurs.

### I.7.5 Tests de stabilité des coefficients -Test de Chow

Pour réaliser ce test, on découpe la série décrivant l'évolution des taux de croissance en deux sous périodes. Les hypothèses du test sont formulées de telle sorte que l'on puisse connaître si oui ou non d'une période à l'autre les coefficients sont les mêmes.

$$H_0 : \hat{X}_t = \sum_{i=1}^3 \hat{\alpha}_i X_{t-i} + \sum_{j=1}^2 \hat{\gamma}_j \varepsilon_{t-j} \text{ si } t \leq 25$$

$$H_1 : \hat{X}_t = \sum_{i=1}^3 \hat{\beta}_i X_{t-i} + \sum_{j=1}^2 \hat{\lambda}_j \varepsilon_{t-j} \text{ si } t > 25$$

La statistique du test s'écrit comme suit :

$$F = \frac{\frac{(SCR - (SCR_1 + SCR_2))}{k}}{\frac{SCR_1 + SCR_2}{n_1 + n_2 - 2k}} \quad (9)$$

Elle suit sous  $H_0$  une loi de Fischer à  $(k, n-2k)$  degrés de libertés. Avec  $n_1 + n_2 = n$  (taille d'échantillon globale) et  $k$  le nombre de paramètres du modèle ARMA(3,2). Après réalisation du test, on trouve une statistique de test  $F = -3.577239$  et  $F_{\alpha=5\%}(6, 51 - 2 \times 6) = F(6, 39) = 2.342262$ . Ainsi, la valeur de la statistique de test est inférieure à la valeur lue dans la table de la loi de Fischer. On accepte donc l'hypothèse nulle de stabilité des coefficients ( $H_0 : \hat{\alpha}_i = \hat{\beta}_i$  et  $\hat{\gamma}_j = \hat{\lambda}_j$ ).

## I.8 Prévisions dans le modèle

Les hypothèses de la modélisation étant toutes vérifiées, on peut dès lors passer à la prévision. La série du taux de croissance du PIB du Cameroun va de l'année 1976 à 2020. Nous allons prédire le taux de croissance du PIB de l'année 2021 et 2022 et ensuite comparer aux valeurs réelles du PIB de ces années là car elles sont déjà disponibles dans le site de l'INS<sup>8</sup> du Cameroun et celui de la banque mondiale. En 2021 et 2022 le taux de croissance du PIB du camerounais a été estimé respectivement à 3,64% et 3,54% selon l'INS. Ci-dessous nous donnons la série de valeurs prédites ainsi qu'une représentation graphique de la série initiale augmentée des valeurs prédites.

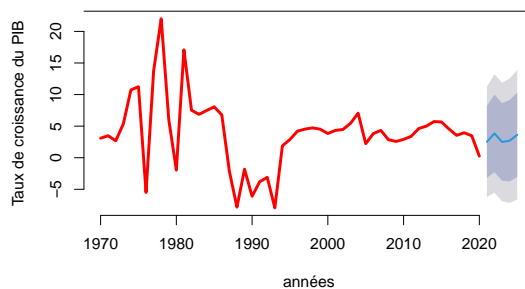
Table 5: Prévisions des taux de croissance

| Années                        | 2021 | 2022 |
|-------------------------------|------|------|
| Taux de croissance du PIB (%) | 4,06 | 4,19 |

Calcul des auteurs

La prévision n'est pas exacte à 100%. En outre, elle est une façade de la réalité et est à but purement descriptif et analytique car elle permet d'orienter les politiques économiques édictées par les gouvernements et les institutions internationales. L'idée étant de pouvoir dire aux populations que si les conditions présentes à l'année 2020 restent inchangées jusqu'en 2021 alors la croissance devrait augmenter de  $(4,06 - 3,54)$ , c'est-à-dire de 0,52 point. Cela permet donc d'espérer à une expansion économique et non une récession. Une façon de dire que les signes attendus sont plus important que les valeurs estimées. En conclusion, les marges de manœuvres d'accroissement de richesses seront caractérisées par moins d'austérité.

Figure 7: Évolution continue du taux de croissance avec prévisions



Source de données : Banque Mondiale, calcul des auteurs.

<sup>8</sup>Institut national de la statistique

## Caractérisation et analyse de la volatilité

Il peut arriver dans un modèle que les résidus soient non auto corrélés, des bruits blancs mais que ces résidus ne sont pas indépendants. Aussi, qu'ils soient caractérisés par de fortes volatilité pour une période et de basses pour une autre période. D'où la nécessité d'un test de volatilité. Nous présentons ici le test d'hétéroscédasticité de ARCH.

### • Test d'hétéroscédasticité conditionnelle de ARCH

Les hypothèses du test sont :

$$\begin{aligned} H_0 : V(\varepsilon_t) &= \hat{\sigma}_\varepsilon^2 \\ \text{contre} \\ H_1 : V(\varepsilon_t) &= \hat{\sigma}_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \dots + \alpha_p \varepsilon_{t-p}^2 \end{aligned}$$

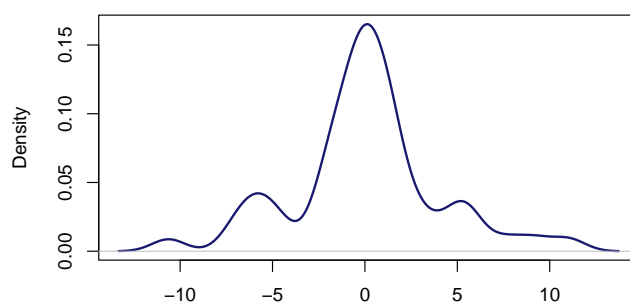
La statistique du test :

$$\chi_{ARCH}^2 = TR^2 \quad (10)$$

Nous prenons une fois de plus la p-value associée au test. Elle vaut 0.0008509. Elle est inférieure au seuil de 5%. Ainsi, nous rejetons l'hypothèse nulle. La variance est conditionnelle est non constante.

Les erreurs sont des bruits blancs centrés, c'est-à-dire que la série des résidus est stationnaire et de moyenne nulle. Mais bien qu'elles soient non auto corrélées le problème de volatilité peut se poser. C'est-à-dire l'alternance entre les périodes de fortes hausses et de forte baisse. En outre, si on élève les résidus au carré et que le test de corrélation s'avère positif alors nos résidus sont non auto corrélés mais pas indépendants. On serait amené à rejeter l'hypothèse d'un ARMA pur. De ce fait, il faudra modéliser la variance des erreurs. Le modèle le plus populaire est le modèle GARCH(p,q). Cette forte volatilité peut être caractérisée par une distribution à queues épaisses comme le montre le graphe ci-dessous.

Figure 8: Densité des résidus



Source de données : Banque Mondiale, calcul des auteurs.

Considérons donc la série des résidus au carrés ( $\hat{\varepsilon}_t^2$ ). On refait le test de Ljung-Box ou de Box-Pierce pour vérifier la non auto corrélation des erreurs. Après exécution du test sur les résidus au carré le test de Ljung-Box permet de conclure avec une p-value de 0,0003903 que l'on rejette l'hypothèse d'absence d'auto corrélation des erreurs. Nous sommes obligés de modéliser la variance des résidus. Pour cela, les modèles les plus utilisés dans la littérature sont les modèles GARCH.



## Conclusion

Le présent chapitre avait pour objectif d'analyser la dynamique du taux de croissance du PIB du Cameroun grâce à des techniques d'analyse de séries temporelles. Nous avons procédé de par la nature de la série à une modélisation suivant l'approche Box-Jenkins. Il ressort que le PIB du Cameroun évolue en suivant un processus de type ARMA mais que celui-ci présente le défaut de ne pas être un ARMA pur car la volatilité est présente. Toutefois, le modèle est adapté et a permis de faire des prévisions pouvant orienter la décision économique pour les années 2021 et 2022. In fine, la proposition d'un modèle de type ARCH ou de façon extensible un modèle GARCH ne vise pas à changer le modèle mais vient seulement comme limite de notre étude car ce sont des modèles prenant en compte cette volatilité et donc par conséquent ils auront l'avantage d'être plus précis.