



Munich Personal RePEc Archive

Inflation Forecasting in Côte D'Ivoire: A Comparative Analysis of the Arima, Holt-Winters, and Lstm Models

Koffi, Siméon

Conseiller Technique du Directeur Général de l'Economie (DGE),
Chargé de cours - Programme de Gestion des Politiques
Economiques (GPE) de l'Université FHB, Côte d'Ivoire,
Responsable du Pôle « Economie et Développement Durable » du
Groupe CONDAD Invest

1 August 2022

Online at <https://mpra.ub.uni-muenchen.de/113961/>
MPRA Paper No. 113961, posted 02 Aug 2022 13:57 UTC

Prévision de l'inflation en Côte D'ivoire : Analyse Comparée des Modèles ARIMA, HOLT-WINTERS, et LSTM

Koffi Siméon¹

Conseiller Technique du Directeur Général de l'Economie (DGE)

Chargé de cours - Programme de Gestion des Politiques Economiques (GPE) de l'Université FHB, Côte d'Ivoire

Responsable du Pôle « Economie et Développement Durable » du Groupe CONDAD Invest

(Courriel1 : simeon.koffi@toki.waseda.jp ; Courriel2 : si.koffi@dge.gouv.ci)

Abstract

This paper attempts to highlight the role of new short-term forecasting methods. It leads to the fact that artificial neural networks (LSTM) are more efficient than classical methods (ARIMA and HOLT-WINTERS) in forecasting the HICP of Côte d'Ivoire. The data are from the “Direction des Prévisions, des Politiques et des Statistiques Economiques (DPPSE)” and cover the period from January 2012 to May 2022. The root mean square error of the long-term memory recurrent neural network (LSTM) is the lowest compared to the other two techniques. Thus, one can assert that the LSTM method improves the prediction by more than 90%, ARIMA by 68%, and Holt-Winters by 61%. These results make machine learning techniques (LSTM) excellent forecasting tools.

Key words: LSTM, ARIMA, HOLT-WINTERS

Résumé

Le présent article tente de mettre en évidence l'importance des nouvelles techniques de prévision à court terme. Il aboutit au fait que les réseaux neuronaux artificiels sont plus performants que les méthodes classiques (ARIMA, HOLT-WINTERS) dans la prévision de l'IHPC en Côte d'Ivoire. Les données proviennent de la Direction des Prévisions, des Politiques et des Statistiques Economiques (DPPSE) et couvrent la période de janvier 2012 à mai 2022. L'erreur quadratique moyenne du réseau neuronal récurrent à mémoire à long terme (LSTM) est la plus faible comparativement aux deux autres techniques. La méthode LSTM améliore donc la prévision de plus de 90%, ARIMA de 68% et Holt-Winters de 61%. Ces résultats font des techniques d'apprentissage automatique (LSTM) d'excellents outils de prévision.

Mots clés : LSTM, ARIMA, HOLT-WINTERS

¹ Les opinions exprimées ici sont celles de l'auteur et n'engagent la responsabilité d'aucune autre personne ou institution.

Introduction

Le ciblage de l'inflation est l'un des objectifs prépondérants des Banques Centrales (Banque Centrale Européenne (BCE), Banque Centrales des Etats de l'Afrique de l'Ouest (BCEAO), Banque Etats de l'Afrique Centrale (BEAC), etc.) comme en témoignent les différentes politiques de chacune de ces Unions et les différents critères de convergence qui les régissent. La prévision de l'inflation a une importance particulière pour tous les pays, qu'ils aient ou non des politiques de ciblage de l'inflation clairement définies. Dans un passé récent, certaines Banques centrales (BCE, BCEAO) ont mené avec succès une politique de ciblage de l'inflation et n'ont pas connu d'évolution inflationniste élevée jusqu'au début de la crise ukrainienne le 24 février 2022. En Afrique de l'Ouest, des pays ayant leur propre monnaie (Nigéria et le Ghana) aspirant à créer une zone monétaire avec les pays de l'Union Economique et Monétaire Ouest Africaine (UEMOA) ont toujours eu des niveaux d'inflations élevés (à deux chiffres).

Compte tenu de l'importance que revêt l'inflation dans l'économie des différents Etats, il convient de mieux la prédire afin d'adopter des politiques susceptibles de la faire infléchir au cas où elle dépasserait les seuils prédéterminés.

La problématique de cette note s'articule autour des nouvelles techniques de prévision, sont-elles plus précises que les approches classiques ? En d'autres termes, les modèles de réseaux de neurones artificiels (RNA) (beaucoup usités aujourd'hui) sont-ils les mieux adaptés dans la prévision des séries temporelles ?

Pour répondre à cette préoccupation, la présente note examine l'application des modèles classiques - ARIMA (*Auto-Regressive Integrated Moving Average*) et Holt-Winters et d'un modèle RNA - LSTM (*Long Short-Term Memory*) dans la prévision de l'inflation en Côte d'Ivoire de janvier 2012 au mois de mai 2022. Les données proviennent de la Direction des Prévisions, des Politiques et des Statistiques Economiques (DPPSE). La double possibilité de comparer les modèles est envisagée. Tout d'abord, les modèles sont comparés au sein de leur propre classe, puis les meilleurs modèles de chaque classe sont comparés entre eux. Pour rendre cette note synthétique uniquement les modèles retenus pour chaque classe seront présentés.

Il existe de nombreux modèles qui peuvent être utilisés pour la modélisation et la prévision de l'inflation. Dans cette note, on teste la possibilité de modéliser l'inflation avec des modèles univariés. Dans ce cas, la prévision des valeurs futures est basée uniquement sur les données historiques de l'inflation. Ce type de modèles donne souvent une prévision plus précise par rapport aux modèles multivariés plus complexes.

1- Etat de l'art

Les techniques d'*apprentissage automatique (machine learning)* ont introduit de nouvelles approches des problèmes de prévision où les relations entre les variables sont complexes. Ces algorithmes d'apprentissage, tels que les réseaux neuronaux récurrents (RNN) et les mémoires à long terme (LSTM), ont suscité beaucoup d'intérêt ces dernières années. Ils sont appliqués dans de nombreuses disciplines, dont l'économie. Ces méthodes d'*apprentissage automatique* sont capables d'identifier la structure des données et réaliser des prévisions avec une précision exceptionnelle.

Chen *et al.* (2001), Nakamura (2006) et McAdam et McNelis (2005) traitent de la prévision de l'inflation à l'aide de réseaux neuronaux. Leurs études montrent que la qualité de prévision de ces réseaux neuronaux surpasse les modèles linéaires classiques pour des horizons plus courts.

Ahmed *et al.* (2010) et Stock et Watson (1999) ont aussi comparé les méthodes linéaires et non linéaires pour la prévision macroéconomique en calculant la moyenne de leurs performances sur un grand nombre de séries chronologiques macroéconomiques. Ils ont abouti aux mêmes résultats que Chen *et al.* (*Op. Cit.*).

Suhartono (2005), quant à lui, compare les performances des réseaux neuronaux, ARIMA et ARIMAX dans la prévision de l'inflation en Indonésie. Il conclut que les réseaux neuronaux donnent une prévision d'inflation plus précise que les approches classiques (ARIMA et ARIMAX).

A côté des techniques d'*apprentissage automatique* subsistent les techniques classiques. Ainsi, Pufnik et Kunovac (2006) donnent une prévision de l'inflation à court terme basée sur le modèle ARIMA en observant l'indice des prix à la consommation en Croatie. D'autres articles traitent également de la prévision de l'inflation à l'aide du modèle ARIMA avec l'utilisation de la méthodologie Box-Jenkins dans l'évaluation du modèle (Alnaa et Ahiakpor, 2011 ; Okafor et Shaibu, 2013).

Dans la littérature économique, une attention particulière est accordée aux comparaisons de la performance des modèles ARMA, ARIMA et GARCH (Nyoni, 2018), la comparaison des modèles VAR et ARIMA dans la prévision de l'IHPC en Autriche (Fritzer *et al.*, 2002) et la comparaison des modèles ARIMA, VAR et ECM (Uko et Nkoro, 2012).

2- Méthodologie de la recherche²

2.1. Modèle ARIMA

Les résultats des travaux de Box et Jenkins (1976) ont permis de définir la méthodologie d'évaluation du modèle ARIMA et de conclure que ce type de modèle est particulièrement adapté aux prévisions à court terme. Les modèles ARIMA ont deux composantes, la composante autorégressive (AR) et la composante moyenne mobile (MA).

Dans un **processus AR d'ordre p**, une observation à l'instant t est exprimée comme une moyenne pondérée des observations passées plus un bruit blanc gaussien :

$$AR(p) : X_t = \sum_{k=1}^p \phi_k X_{t-k} + \varepsilon_t \quad (1)$$

Où, ϕ_k ($k=1,2,3,\dots,p$) sont les paramètres du modèle, $(X_t)_{t \in \mathbb{Z}}$ un processus stationnaire et ε_t une erreur gaussienne.

$(X_t)_{t \in \mathbb{Z}}$ est un **processus MA** (*Moving Average*) d'ordre q noté $MA(q)$, si :

$$MA(q) : X_t = \varepsilon_t + \sum_{i=1}^q \theta_i \varepsilon_{t-i} \quad (2)$$

Avec $(\theta_1, \dots, \theta_q) \in \mathbb{R}^q$ et $\theta_q \neq 0$

² Les détails sur les différents modèles présentés dans cette note pourront être consultés dans le livre de Guy Mélard : « Méthodes de prévision à court terme ». Editions Ellipses, 2007.

Soit $(\varepsilon_t)_{t \in \mathbb{Z}}$ un bruit blanc de variance σ^2 .

Un processus $(X_t)_{t \in \mathbb{Z}}$ est un **processus ARMA** d'ordre (p, q) , si : $(X_t)_{t \in \mathbb{Z}}$ est stationnaire et

$$\forall t \in \mathbb{Z} : X_t = \sum_{i=1}^p \varphi_i X_{t-i} + \varepsilon_t + \sum_{i=1}^q \theta_i \varepsilon_{t-i} \quad (3)$$

Avec $(\varphi_1, \dots, \varphi_p) \in \mathbb{R}^p$, $\varphi_p \neq 0$, $(\theta_1, \dots, \theta_q) \in \mathbb{R}^q$ et $\theta_q \neq 0$.

Le processus $(X_t)_{t \in \mathbb{Z}}$ est un **processus ARIMA** (*AutoRegressive Integrated Moving Average*) si la série $(X_t)_{t \in \mathbb{Z}}$ n'est pas stationnaire ; dans ce cas, il convient de la rendre stationnaire en la différenciant d fois. La série différenciée est représentée par le modèle $ARIMA(p, d, q)$.

2.2. Modèle HOLT-WINTERS

La méthode Holt-Winters prévoit des séries de la forme :

$$\hat{x}_{t+1} = a_t + b_t t \quad (4)$$

où \hat{x}_t est la prévision de la série x_t , a_t est une moyenne qui dérive dans le temps, et b_t est un coefficient tributaire du temps qui dérive également. La méthode de Holt-Winters consiste à choisir les paramètres de lissage de sorte à minimiser la somme des carrés des erreurs de prévision.

Etant donné la série x_t , les paramètres de lissage α et β , et les valeurs initiales a_0 et b_0 permettent d'écrire les équations suivantes :

$$\begin{cases} a_t = \alpha x_t + (1 - \alpha)(a_{t-1} + b_{t-1}) \\ b_t = \beta(a_t - a_{t-1}) + (1 - \beta)b_{t-1} \end{cases} \quad (5)$$

Après avoir calculé les séries de termes constants et linéaires, a_t et b_t respectivement, la prévision à la période τ est donnée par :

$$\hat{x}_{t+\tau} = a_t + b_t \tau \quad (6)$$

2.3. Modèle LSTM (*Long Short-Term Memory*)

Le LSTM est un type de réseau neuronal récurrent (RNN) capable de se souvenir des valeurs des étapes précédentes en vue d'une utilisation future. Avant de se pencher sur le LSTM, il est nécessaire d'avoir un aperçu de ce à quoi ressemble un réseau neuronal.

- Réseau de neurones artificiels (RNA)

Un réseau neuronal est constitué d'au moins trois couches, à savoir : une couche d'entrée, des couches cachées et une couche de sortie. Le nombre de caractéristiques de l'ensemble de données détermine la dimension ou le nombre de nœuds dans la couche d'entrée. Ces nœuds sont connectés par des liens appelés « synapses » aux nœuds créés dans la ou les couches cachées. Les synapses portent des poids pour chaque nœud de la couche d'entrée. Les poids jouent essentiellement le rôle de décideur qui choisit quel signal peut passer et lequel ne le peut

pas. Les poids indiquent également la force ou l'étendue de la couche cachée. Un réseau neuronal apprend essentiellement en ajustant le poids pour chaque synopsis.

Dans les couches cachées, les nœuds appliquent une fonction d'activation (par exemple, *sigmoïde* ou *tangente hyperbolique (tanh)*) sur la somme pondérée des entrées pour transformer les entrées en sorties, ou valeurs prédites. La couche de sortie génère un vecteur de probabilités pour les différentes sorties et sélectionne celles dont le *taux d'erreur* ou le *coût* est le plus faible, c'est-à-dire qui *minimise les différences entre les valeurs attendues et prédites*, à l'aide d'une fonction d'activation appelée *SoftMax*³.

(Formellement, la fonction *SoftMax* se présente comme suit : $\sigma(z)_j = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}}$ avec $j \in \{1, \dots, K\}$)

Les pondération des erreurs obtenues lors du premier « entraînement » du réseau peuvent ne pas être les meilleures. Pour trouver les valeurs optimales, les erreurs sont rétro-propagées dans le réseau de la couche de sortie vers les couches cachées et, par conséquent, les poids sont ajustés. La procédure est répétée plusieurs fois avec les mêmes observations et les poids sont réajustés jusqu'à ce qu'il y ait une amélioration des valeurs prédites et donc du coût. Lorsque la fonction de coût est minimisée, le modèle est formé.

- Réseau de neurones récurrents (RNN)

Un réseau neuronal récurrent (RNN) est un cas particulier de RNA dont l'objectif est de prédire l'étape suivante de la séquence d'observations par rapport aux étapes précédentes observées dans la séquence. Dans les RNN, les couches cachées servent de mémoire interne pour stocker les informations capturées lors des étapes précédentes de lecture des données séquentielles. Les RNN sont dits « récurrents » parce qu'ils effectuent la même tâche pour chaque élément de la séquence, avec la particularité d'utiliser les informations capturées précédemment pour prédire les futures données séquentielles non vues. Le principal défi que pose un RNN est dû au fait que ces réseaux ne se souviennent que de quelques étapes antérieures de la séquence et ne sont donc pas adaptés à la mémorisation de séquences de données plus longues. Ce problème apparemment insoluble est résolu en utilisant la « ligne de mémoire » introduite dans le réseau de mémoire à long-court terme (LSTM).

- Réseau de mémoire à long-court terme (LSTM)

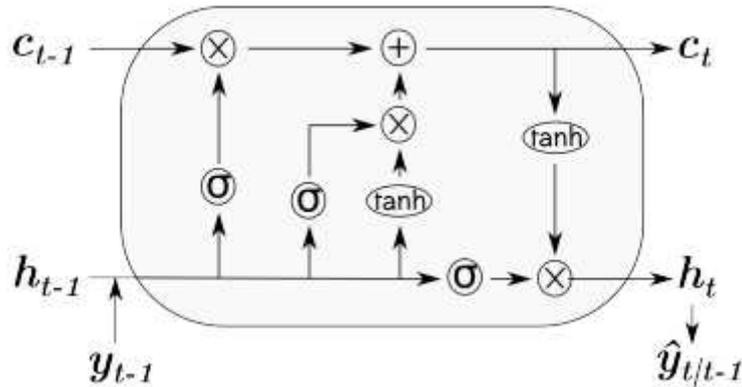
Le LSTM est un type spécial de RNN avec des caractéristiques supplémentaires pour mémoriser les séquences de données. La mémorisation de la tendance antérieure des données est possible grâce à certaines portes et à une ligne de mémoire incorporée dans les LSTM.

Chaque LSTM contient des cellules où les flux de données sont capturés et stockés. Les cellules ressemblent à une ligne de transport qui relie un module à un autre en transportant les données du passé et en les rassemblant pour le présent. Grâce à l'utilisation de certaines portes dans chaque cellule, les données peuvent être éliminées, filtrées ou ajoutées pour les cellules suivantes. Chaque couche sigmoïde produit des nombres dans la gamme de 0 et 1, décrivant la

³ La fonction *softmax* est utilisée en pour convertir un score en probabilité dans un contexte de classification multi-classe. Par exemple, si dans un contexte de classification d'objets nous obtenons : un score de 1200, 600, 200 et 0 pour que la probabilité que l'objet soit respectivement une pomme, une orange, une poire et pour les autres probabilités. L'application de la fonction d'activation *softmax* permettra de générer un score de probabilité normalisé, dont la somme totale des probabilités sera égale à 1. Les résultats de la fonction *softmax* donnent pour la pomme, l'orange et la poire les probabilités respectives 0,6 ; 0,3 et 0,1.

quantité de chaque segment de données qui devrait être laissée passer dans toute cellule. Plus précisément, une estimation de 0 implique de « ne rien laisser passer », tandis qu’une estimation de 1 indique de « laisser tout passer ».

Figure 1 : représentation des cellules LSTM



Source : Extrait de Anna et Andresen (2019)

Sur la figure 1, la variable y_{t-1} représente l'input à l'étape t (par exemple la valeur retardée de l'inflation). c est l'état du réseau qui représente sa mémoire du passé. h_t représente l'output du LSTM à l'étape t . σ et \tanh dénotent les portes qui sont de petits réseaux neuronaux ayant la fonction sigmoïde ou la fonction tangente hyperbolique comme activations à la sortie. A gauche du diagramme, la prédiction du réseau pour la période t qui a été calculée à l'étape précédente (\hat{h}_{t-1}) et la valeur d'entrée à cette étape (y_{t-1}) sont combinées puis filtrées par les différentes portes. Ces dernières déterminent quelle partie de l'état c_{t-1} doit être oubliée, quelles informations et quelle quantité doivent être ajoutées et de quelle manière la prédiction à l'étape t doit être ajustée. A droite du diagramme, la sortie de la cellule est présentée : il s'agit d'une nouvelle valeur actualisée h_t et de \hat{y} .

Les LSTM contiennent principalement trois types de portes dans le but de contrôler l'état de chaque cellule : la porte d'oubli (*forget gate*) sort un nombre entre 0 et 1, où 1 indique de « conserver complètement l'information » ; tandis que, 0 implique d'« ignorer complètement l'information » ; la porte d'entrée (*input gate*) choisit quelles nouvelles données doivent être stockées dans la cellule. Tout d'abord, une *couche sigmoïde*, appelée « *couche porte d'entrée* », choisit les valeurs qui seront modifiées. Ensuite, une *couche tanh* fabrique un vecteur de nouvelles valeurs normalisées comprises entre -1 et 1 et (enfin) la porte de sortie (*output gate*) décide de ce qui sera produit par chaque cellule. La valeur produite sera basée sur l'état de la cellule ainsi que sur les données filtrées et nouvellement ajoutées. (Cf. Annexe 1, Mécanisme des trois portes).

2.4. Comparaison de la précision des prévisions

La comparaison de la précision des différents modèles (ARIMA, Holt-Winters, LSTM) sera effectuée sur la base de 3 critères :

Erreur quadratique moyenne (Root mean square error) :

$$RMSE = \sqrt{\frac{1}{g} \sum_{j=1}^g (x_{m+j} - \hat{x}_m(j))^2}$$

Erreur absolue moyenne (Mean absolute error) :

$$MAE = \frac{1}{g} \sum_{j=1}^g |x_{m+j} - \hat{x}_m(j)|$$

Pourcentage d'erreur absolue moyenne (Mean absolute percentage error) :

$$MAPE = \frac{100}{g} \sum_{j=1}^g \left| \frac{x_{m+j} - \hat{x}_m(j)}{x_{m+j}} \right|$$

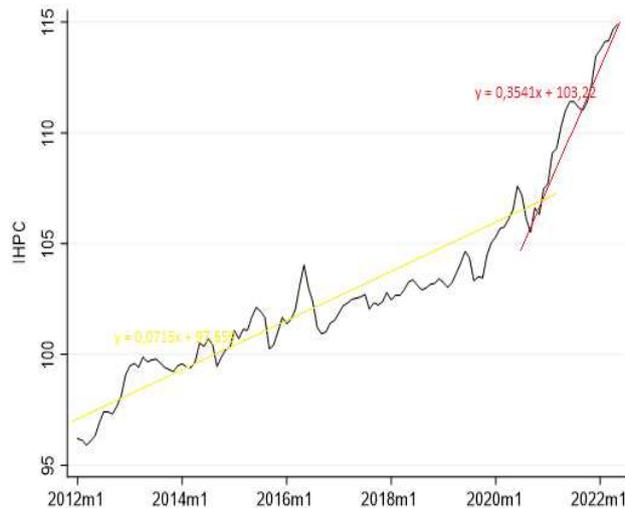
Avec $x_{m+j} - \hat{x}_m(j)$ qui dénote l'erreur de prévision, c'est-à-dire la différence entre les valeurs observées et les valeurs prédites. Un bon modèle de prévision est celui pour lequel les statistiques de précision des modèles sont petites. Mais, il n'existe pas de seuil à partir duquel une statistique pourrait être qualifiée de petite. C'est pourquoi pour choisir le meilleur modèle de prévision, les valeurs statistiques seront comparées entre elles. Le meilleur modèle sera celui qui minimise les trois (03) critères : RMSE, MAE et MAPE.

3- Résultats et discussion

L'évolution de l'indice harmonisé des prix à la consommation (IHPC) est représentée dans la figure 1 sur la période janvier 2012 à mai 2022. Les données proviennent de la Direction des Prévisions, des Politiques et des Statistiques Economiques (DPPSE).

La courbe de l'IHPC semble⁴ présenter une rupture structurelle en octobre 2019. En effet de 2012 à 2019, l'IHPC augmentait en moyenne mensuelle de 7,15%. Mais depuis le mois d'octobre 2019, la hausse est passée à 35,41% soit 5 fois la moyenne sur la première période.

Figure 1 : Evolution de l'IHPC en Côte d'Ivoire



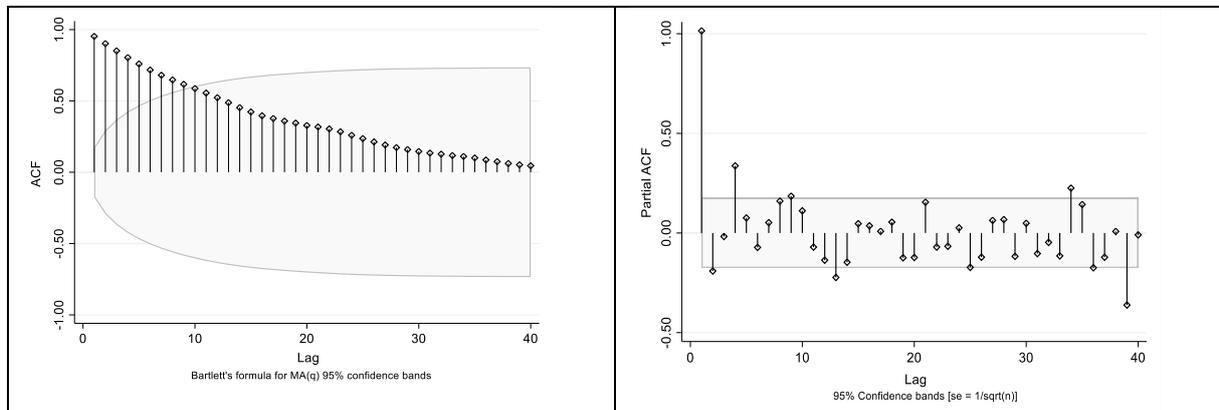
Source : Auteur à partir des données de la DPPSE

⁴ Pour confirmer cette rupture structurelle, il est nécessaire de réaliser des tests appropriés (Zivot et Andrews, Clemao et Clemio).

3.1. Prédiction de l'IHPC à l'aide du modèle ARIMA

La figure 2 (Fonction d'autocorrélation et la fonction d'autocorrélation partielle) de l'IHPC indique clairement que cette variable n'est pas stationnaire. Ces observations graphiques sont confirmées par les tests de racines unitaires de Dickey-Fuller Augmenté (DFA), de Phillips-Perron (PP) et de Kwiatkowski–Phillips–Schmidt–Shin (KPSS) (Cf. Tableau 1 en Annexe 1).

Figure 2 : Fonction d'autocorrélation (ACF) et la fonction d'autocorrélation partielle (Partielle ACF)

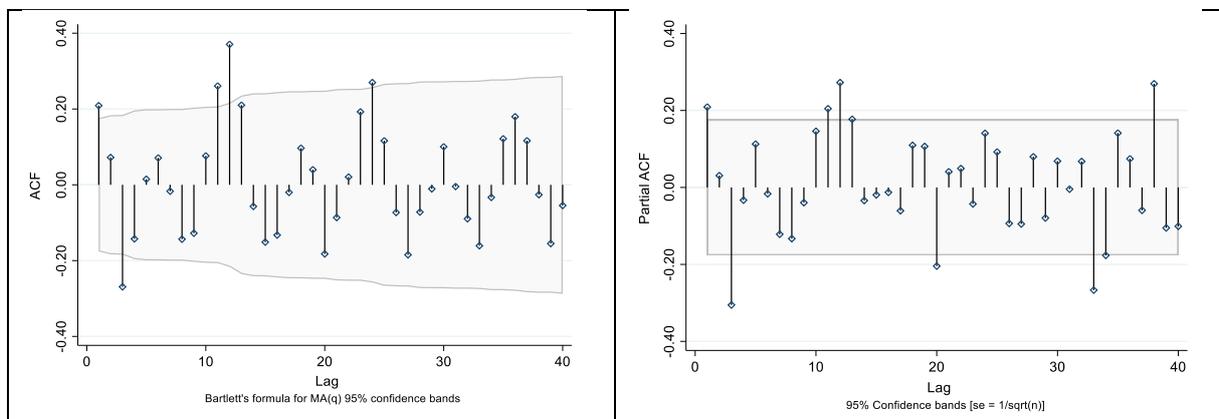


Source : Auteur

Selon les valeurs des statistiques des tests DFA et PP et des probabilités p calculées pour la série, l'IHPC n'est pas stationnaire. Les résultats des tests de racines unitaires indiquent que les séries sont $I(1)$ c'est-à-dire stationnaires en différence première. C'est donc la série transformée qui sera utilisée dans la procédure Box-Jenkins.

La figure 2 sur les ACF et PACF permet d'identifier l'ordre des modèles AR et MA. A partir des différents corrélogrammes, l'ordre des différents modèles peut être établi comme suit : AR (1, 3, 11, 12) et MA (1, 3, 11, 12). A partir de cette information, plusieurs modèles ARIMA sont susceptibles d'être utilisés. Le meilleur modèle est celui qui minimisera les valeurs de tous les critères d'information : Le critère d'information d'Akaike (AIC) et le critère d'information bayésien (BIC). En se basant sur ces critères, le modèle ARIMA (1, 1, 11) a été sélectionné comme le meilleur des modèles.

Figure 3 : Fonction d'autocorrélation (ACF) et la fonction d'autocorrélation partielle (Partielle ACF)

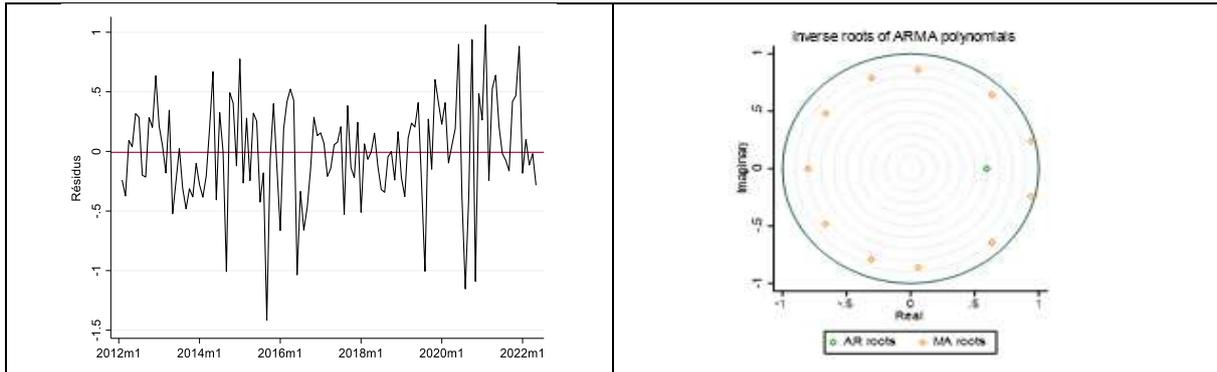


Source : Auteur

Après avoir sélectionné le meilleur modèle, en minimisant les valeurs des critères d'information, il est nécessaire de vérifier que les erreurs sont des bruits blancs et que les racines

unitaires du modèle ARIMA choisi sont incluses (strictement) dans le cercle de racines unitaires. Sur la base de la valeur de la statistique du test de Portmanteau (33,237) et de la p value correspondante (0,767) avec un risque d'erreur de 5%, on peut confirmer l'hypothèse nulle que la distribution des résidus peut être décrite par la distribution normale. Par ailleurs, les racines unitaires du modèles ARMA choisi sont tous compris dans le cercle de racines unitaires (Cf. figure 4). Ainsi, le modèle répond aux deux caractéristiques requises de la validité (stabilité du modèle et normalité de la distribution des résidus), et peut donc être utilisé dans des analyses ultérieures comme référence dans la prévision de l'IHPC.

Figure 4 : Validité du modèle ARIMA (1, 1, 11)



Source : Auteur

A partir du modèle sélectionné ci-dessus, la prévision de l'IHPC est réalisée sur la période de l'étude (Cf. figure 5). Les statistiques sur la précision du modèle sont :

$$RMSE = 0,429$$

$$MAE = 0,333$$

$$MAPE = 0,323$$

En moyenne, le modèle ARIMA a un taux de précision dans la prévision de l'indice harmonisé des prix à la consommation de 68%. Ces résultats obtenus seront par la suite comparés à ceux obtenus avec les deux autres approches afin de choisir le modèle qui prédit le mieux l'IHPC.

Figure 5 : Prévision de l'IHPC à l'aide du modèle ARIMA (1, 1, 11)



Source : Auteur à partir des données de la DPPSE (2022)

3.2. Prédiction de l'IHPC à l'aide du modèle HOLT-WINTERS

Dans la prédiction des séries temporelles à l'aide de la méthode Holt-Winters, le type de série temporelle est particulièrement importante. Comme dans l'approche ci-dessus, le choix a été fait de ne pas prendre en compte l'effet saisonnier jugé non significatif, la méthode Holt-Winters sera donc appliquée dans le cas d'un lissage non saisonnier.

Les résultats de la prédiction de l'IHPC sont présentés dans la figure 6 et les statistiques calculées afférentes à la précision du modèle sont :

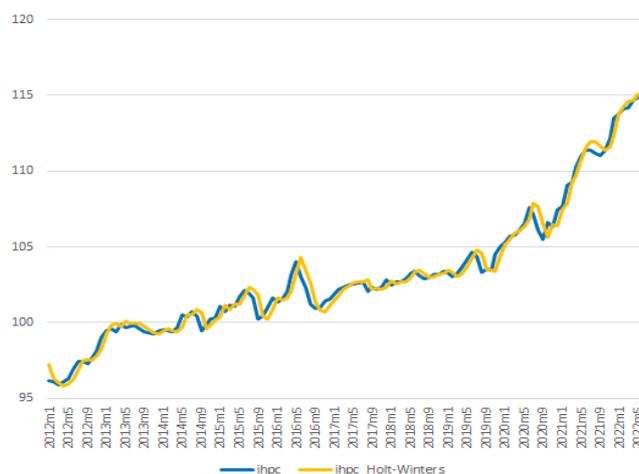
$$\text{RMSE} = 0,550$$

$$\text{MAE} = 0,403$$

$$\text{MAPE} = 0,390$$

En moyenne, le modèle Holt-Winters a un taux de précision dans la prédiction de l'indice harmonisé des prix à la consommation de 61%. Ces résultats obtenus indiquent que le modèle ARIMA performe mieux dans la prédiction de l'IHPC en Côte d'Ivoire que l'approche proposée par Holt-Winters.

Figure 6 : Prédiction de l'IHPC à l'aide de la méthode Holt-Winters non saisonnière



Source : Auteur à partir des données de la DPPSE (2022)

3.3. Prédiction de l'IHPC à l'aide de la méthode LSTM

Pour réaliser la prédiction avec la méthode LSTM, il convient de séparer les données en deux⁵ sous-ensembles. Le sous-ensemble qui servira d'entraînement (*training*) et le sous-ensemble qui sera utilisé comme *test*. Dans cette note, les données mensuelles qui couvrent la période de janvier 2012 à octobre 2019 (représentant 75% des données) serviront de sous-ensemble d'entraînement et les données de la période novembre 2019 à mai 2022 (25% des données)

⁵ Certaines études retiennent trois : sous ensemble d'entraînement, de test et de validation

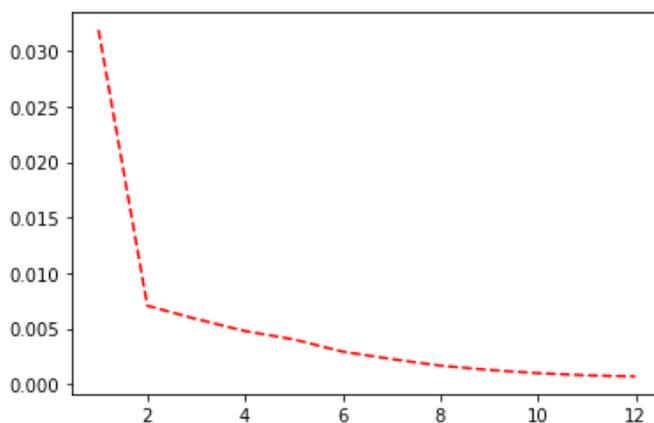
seront celles de sous-ensemble *test*. La procédure LSTM requiert que les données soient normalisées de sorte qu'elles aient une moyenne et une variance nulles.

Les méthodes d'*apprentissage automatique* disposent de différentes mesures qui permettent de savoir si notre modèle est performant. Cependant, ces mesures peuvent prêter à confusion quant à leur signification, à la manière dont elles peuvent être interprétées ou à ce qu'elles sont exactement. Dans cette note, seule l'erreur de prévision sera utilisée pour apprécier la qualité du modèle.

L'erreur de prévision est une valeur qui représente la somme des erreurs de notre modèle. Elle mesure la qualité (ou la faiblesse) de notre modèle. Si les erreurs sont élevées, la perte sera élevée, ce qui signifie que le modèle ne réalise pas une bonne précision. Sinon, plus elle est faible, plus notre modèle fonctionne bien.

Le réseau utilisé est composé de 40 neurones, 12 cycles d'apprentissage et un seuil de gradient de 1 (pour éviter l'implosion du réseau). Selon la figure 7, l'erreur de prévision initiale est d'environ 3%, après le 12ème cycle cette erreur est inférieure à 0,5%. On peut adonc affirmer que le processus d'apprentissage du réseau neuronal est excellent et peut valablement servir à la prévision de l'IHPC.

Figure 7 : Evolution de l'erreur de prévision



Source : Auteur à partir des données de la DPPSE (2022)

Les résultats de la prévision de l'IHPC sont présentés dans la figure 8 et les statistiques calculées afférentes à la précision du modèle sont :

RMSE = 0,013

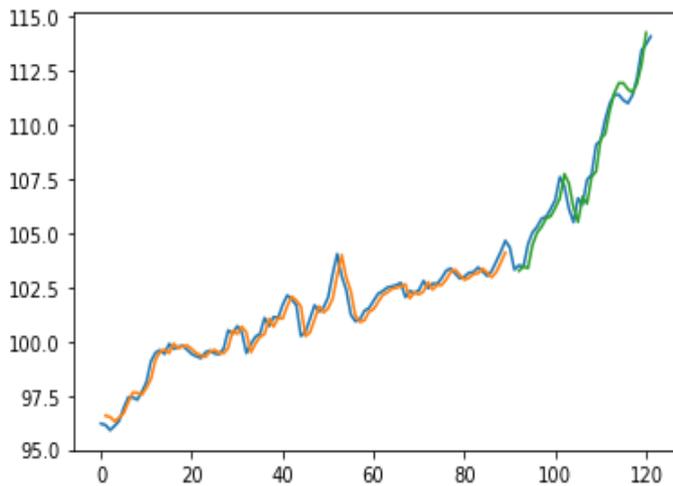
MAE = 0,074

MAPE = 0,075

En moyenne, le modèle LSTM a un taux de précision dans la prévision de l'indice harmonisé des prix à la consommation de 92,5%.

Note : La courbe en bleue représente l'IHPC, celle en orange le sous-ensemble training et enfin celle en vert le sous-ensemble test.

Figure 8 : Prédiction de l'IHPC à l'aide de l'approche LSTM

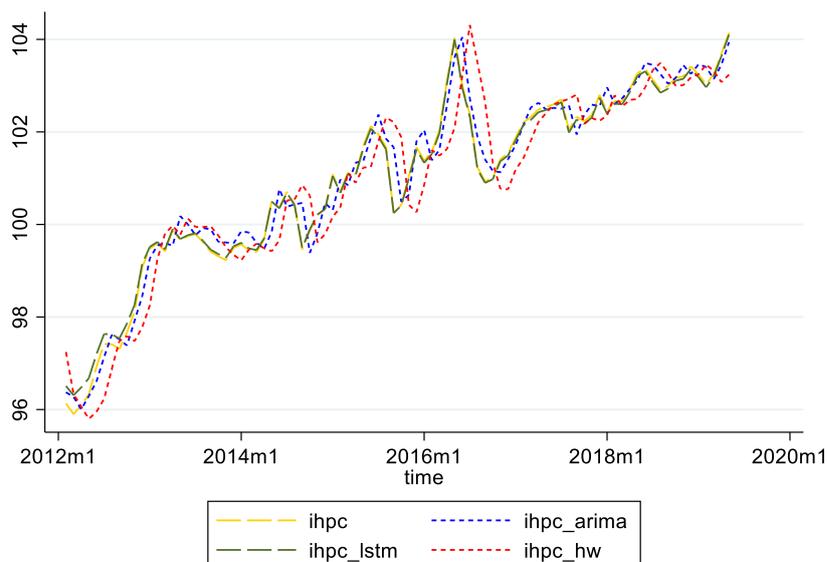


Source : Auteur à partir des données de la DPPSE (2022)

3.4. Comparaison de la précision des prévisions des modèles ARIMA, Holt-Winters et LSTM

La prédiction de l'IHPC réalisée en utilisant les trois approches est présentée dans la figure 8. De prime abord, il est donné de constater que l'estimation réalisée avec le modèle LSTM performe mieux que les deux autres. Après le modèle LSTM, c'est l'approche ARIMA qui réalise une performance acceptable. En conclusion, en se basant sur le graphique, il semblerait que des trois modèles, c'est le modèle LSTM qui est le meilleur dans la prédiction de l'IHPC.

Figure 9 : Prédiction de l'IHPC à l'aide des approches ARIMA, Holt-Winters et LSTM



Source : Auteur à partir des données de la DPPSE (2022)

Pour confirmer les conclusions issues de l'inspection graphique, il est utile de calculer les erreurs de prévision. Ces dernières ont été calculées en comparant les valeurs réelles (observées) de la série chronologique de l'indice des prix à la consommation harmonisé (IHPC) et les valeurs prédites sur la base de trois types de modèles. A ce niveau, il est important de rappeler que l'objectif de cette note est de déterminer la classe de modèles qui donne la prévision la plus précise. A cet effet, les valeurs des statistiques de prévision calculées sont présentées dans le tableau 2.

Tableau 2. Comparaison des statistiques de mesure de la précision des prévisions

	RMSE	MAE	MAPE
ARIMA	0,429	0,333	0,323
HW (non saisonnier)	0,550	0,403	0,390
LSTM	0,013	0,074	0,075

Source : Auteur

Dans le cas de la prévision de l'inflation en Côte d'Ivoire, le modèle LSTM donnent les prévisions les plus précises. L'application de moyennes mobiles autorégressives (ARIMA) donne des résultats légèrement moins bons, tandis que le modèle de prévision Holt-Winters (non-saisonnier) donne les prévisions d'inflation les plus imprécises (Cf. tableau 2).

Conclusions

Les progrès récents dans le développement de techniques sophistiquées basées sur l'informatique ont permis aux techniques d'*apprentissage automatique* de gagner en popularité parmi les chercheurs de diverses disciplines. La question principale est alors de savoir quelle est la précision et la puissance de ces approches nouvellement introduites dans les méthodes de prévision par rapport aux méthodes traditionnelles. La présente note compare la précision de trois modèles : ARIMA, Holt-Winters et LSTM.

Ces techniques ont été appliquées sur l'IHPC de la Côte d'Ivoire. Les résultats indiquent que le modèle LSTM améliore la prévision de plus de 90%, ARIMA de 68% et Holt-Winters de 61%. Les techniques d'*apprentissage automatique* (LSTM) sont d'excellents outils de prévision.

Pour confirmer ces résultats, il conviendrait dans les différentes Directions de pratiquer le même exercice avec d'autres variables macroéconomiques comme le produit intérieur brut (PIB).

References

- Ahmed, N. K., Atiya, A. F., Gayar, N. E., and El-Shishiny, H. (2010). An empirical comparison of machine learning models for time series forecasting. *Econometric Reviews* 29 (5-6), 594–621.
- Alnaa, S. E., and Ahiakpor, F (2011). ARIMA (Autoregressive Integrated Moving Average) approach to predicting inflation in Ghana, *Journal of Economics and International Finance*, 3 (5): 328 – 336.
- Chen, X., Racine, J., and Swanson, N. R. (2001). Semiparametric arx neural-network models with an application to forecasting inflation. *IEEE Transactions on neural networks* 12 (4), 674–683.
- Fritzer, F., Moser, G., and Scharler, J. (2002). Forecasting Austrian HICP and its Components Using VAR and ARIMA models, Working Paper OENB, 73.
- McAdam, P., and McNelis, P. (2005). Forecasting inflation with thick models and neural networks, *Economic Modeling*, 22, 548 – 567.
- Nakamura, E. (2006). Inflation forecasting using a neural network, *Economics Letter*, 86,373-378.
- Nyoni, T (2018I). Modelling and Forecasting Naira / USD Exchange Rate In Nigeria: a Box – Jenkins ARIMA approach, Munich University Library – Munich Personal RePEc Archive (MPRA), Paper No. 88622.
- Okafor, C., and Shaibu, I. (2013). Application of ARIMA models to Nigerian inflation dynamics, *Research Journal of Finance and Accounting*, 4(3), 138-150.
- Pufnik, A., and Kunovac, D. (2006). Short-Run Inflation Forecasting in Croatia Using Seasonal ARIMA Processes. Istrazivanja 1-18, Croatian National Bank, Zagreb.
- Stock, J. H., and Watson, M. W. (1999). Forecasting inflation. *Journal of Monetary Economics* 44 (2), 293 – 335.
- Suhartono, S. (2005). Neural Networks, ARIMA and ARIMAX Models for Forecasting Indonesian Inflation. *Widya Journal of Management and Accounting*, vol. 5, no. 3, 2005.
- Uko, A. K., and Nkoro, E. (2012). Inflation forecasts with ARIMA, vector autoregressive and error correction models in Nigeria. *European Journal of Economics, Finance & Administrative Science*, 50, 71-87.

ANNEXE 1 : Tableau 1. Tests de stationnarité de IHPC

Série	Tests		
	ADF	PP	KPSS
IHPC	-0.144	-0.278	0,214**
D.IHPC	-9.008***	-8.860***	0,188**

Note : mettre la signification des étoiles

ANNEXE 2 Formellement, le mécanisme des trois portes peut être décrit comme suit :

Porte d'entrée :

$$a_l^t = \sum_{i=1}^I w_{il} x_i^t + \sum_{h=1}^H w_{hl} z_h^{t-1} + \sum_{c=1}^C w_{cl} s_c^{t-1}$$

$$b_l^t = f(a_l^t)$$

Porte d'oubli

$$a_\phi^t = \sum_{i=1}^I w_{i\phi} x_i^t + \sum_{h=1}^H w_{h\phi} z_h^{t-1} + \sum_{c=1}^C w_{c\phi} s_c^{t-1}$$

$$b_\phi^t = f(a_\phi^t)$$

Porte de sortie

$$a_w^t = \sum_{i=1}^I w_{iw} x_i^t + \sum_{h=1}^H w_{hw} z_h^{t-1} + \sum_{c=1}^C w_{cw} s_c^t$$

$$b_w^t = f(a_w^t)$$

Où a et b sont des signaux, x_i et z_h les entrées et les sorties de la cellule, w la pondération de chaque connexion entre les différents signaux, s_c la valeur de la cellule et l , w et c les indices décrivant respectivement les signaux issus de la cellule et des portes d'entrée et d'oubli, $w_{h\phi}$ la pondération entre les sorties et la porte d'oubli et w_{il} celles reliant les portes d'entrée à la cellule d'entrée...

Il est utile de rappeler que le contenu de la cellule mémoire (s_c) peut être actualisé ou remis à 0. Il se présente comme suit :

$$s_c^t = b_\phi^t s_c^{t-1} + b_l^t g(a_c^t)$$

Où,

- le premier terme reconduit la valeur précédente de la cellule s_c^{t-1} , qui peut être annulée par la valeur de la porte d'oubli b_ϕ^t ;
- a_c^t somme pondérée des entrées et des sorties est dirigée par la valeur de la porte d'entrée b_l^t et s'écrit comme suit : $a_c^t = \sum_{i=1}^I w_{ic} x_i^t + \sum_{h=1}^H w_{hc} z_h^{t-1}$.

Enfin, la valeur à la sortie du neurone LSTM est calculée en utilisant une fonction d'activation h appliquée à la valeur de la cellule mémoire s_c^t , orientée par la valeur de la porte de sortie b_w^t telle que :

$$b_c^t = b_w^t h(s_c^t)$$