

Nowcasting du produit intérieur brut de la Côte d'Ivoire

Franck Migone

ANSTAT

Agence Nationale de la Statistique de Côte d'Ivoire

Résumé

Dans le besoin de disposer d'un suivi quasi-temps réel de la conjoncture de l'économie ivoirienne, cette étude met en place un cadre de nowcasting du taux de croissance du Produit Intérieur Brut (PIB). Pour ce faire, nous construisons une base de données couvrant la période la période du troisième trimestre de 2013 au quatrième trimestre 2024 à partir d'indicateurs traditionnels et de données alternatives, et huit modèles d'apprentissage automatique dont les performances sont comparées à un au benchmark autorégressif selon des métriques standards. L'étude explore également les performances de ces modèles sur des sous-périodes de crise, reprise et stabilité. Les résultats mettent en évidence la meilleure performance des modèles de machine learning, avec une réduction des erreurs prédictives allant jusqu'à 30 %, la régression Ridge se distinguant par sa robustesse et sa stabilité et anticipant une croissance de 7,06 % au premier trimestre 2025. L'intégration de données alternatives permet en outre d'améliorer la précision de 10 % en RMSE et de 6 % en MAE, confirmant leur valeur ajoutée, tandis que l'analyse des valeurs de Shapley souligne le rôle déterminant des agrégats monétaires, des indices de prix et des données Google trends et satellitaires dans l'explication des dynamiques de croissance ivoirienne.

Most-clés : *nowcasting, PIB, Côte d'Ivoire, apprentissage automatique, données alternatives, valeurs de Shapley*.

1 Introduction

Le Produit Intérieur Brut (PIB) constitue l'un des indicateurs macroéconomiques les plus utilisés pour mesurer la performance économique d'un pays. Il traduit la valeur totale des biens et services produits au cours d'une période donnée (Eurostat, 2020). Selon l'OCDE, le taux de croissance du PIB représente un baromètre essentiel de la vitalité économique nationale, fournissant aux décideurs publics et privés des informations clés pour orienter leurs stratégies. Cependant, la publication des comptes nationaux suit un processus long et complexe, souvent assorti d'un délai de plusieurs semaines, voire de plusieurs mois, entre la période de référence et la diffusion des résultats. Ce décalage limite la réactivité des politiques économiques et la capacité d'anticipation des acteurs économiques.

C'est dans ce contexte qu'a émergé la pratique du *nowcasting* — littéralement, la « prévision du présent ». Introduite en économie par Giannone, Reichlin et Small (2008), cette approche vise à estimer en quasi temps réel des agrégats macroéconomiques, tels que la croissance du PIB, à partir d'indicateurs disponibles à une fréquence plus élevée. Initialement fondé sur des modèles économétriques traditionnels tels que les modèles à facteurs dynamiques (*Dynamic Factor Models*, DFM) ou les équations de pont (*Bridge Equations*), le nowcasting s'est heurté à certaines limites : dépendance à des hypothèses structurelles fortes, difficulté à intégrer des données massives et hétérogènes, et faible flexibilité face aux non-linéarités (Stock et Watson, 2016 ; Banbura et Modugno, 2010).

Les progrès récents en apprentissage automatique (*machine learning*) offrent une alternative prometteuse à ces contraintes. Ces méthodes permettent de modéliser des relations complexes, non linéaires et multidimensionnelles entre variables, renforçant ainsi la précision et la robustesse des prévisions (Cascaldi-Garcia et al., 2023). Leur application en contexte macroéconomique s'est intensifiée, notamment durant la crise de la COVID-19, où elles ont permis de suivre plus finement les fluctuations de l'activité (Buell et al., 2021).

Cependant, l'efficacité du nowcasting repose sur la disponibilité de données à haute fréquence, publiées avec régularité et faible délai. Dans de nombreux pays en développement, dont la Côte d'Ivoire, les systèmes statistiques officiels demeurent contraints par des retards de diffusion et une couverture partielle des indicateurs infra-annuels. Face à ces limites, l'usage de données dites « alternatives » s'impose progressivement. Issues de sources numériques ou satellitaires (Google Trends, mobilité, transactions électroniques, luminosité nocturne), ces données permettent d'enrichir l'information disponible et d'améliorer la qualité des prévisions. Bolivar (2024) montre par exemple que l'intégration de la luminosité nocturne améliore significativement le nowcasting de l'activité économique en Bolivie.

Dans ce contexte, le développement d'un modèle de nowcasting du taux de croissance du PIB ivoirien, combinant données traditionnelles et alternatives, constitue une innovation majeure pour la modernisation de la statistique macroéconomique nationale. L'objectif de cette étude est donc de proposer un cadre empirique mobilisant des modèles d'apprentissage automatique afin d'anticiper la croissance du PIB en temps quasi réel.

2 Revue de littérature

2.1 Les modèles traditionnels

Les premières approches de *nowcasting* reposent sur des modèles économétriques destinés à pallier le décalage temporel de publication des agrégats macroéconomiques. Parmi eux, le modèle de pont (*bridge model*) établit une relation linéaire entre la variable cible, comme le PIB, et des indicateurs conjoncturels disponibles à haute fréquence, préalablement agrégés à une fréquence trimestrielle comparable (Cascaldi-Garcia et al., 2023). Ce modèle contourne ainsi le problème de fréquence mixte en projetant les valeurs manquantes à l'aide de modèles autorégressifs simples (Schumacher, 2014 ; Bañbura et al., 2014). Son estimation, fondée sur les moindres carrés ordinaires, reste accessible et efficace pour des dispositifs à faible nombre de variables (Baffigi et al., 2004).

Une amélioration notable provient des régressions MIDAS (*Mixed Data Sampling*) introduites par Ghysels et al. (2004). Ces modèles exploitent directement des séries à fréquences mixtes sans agrégation préalable, en pondérant les observations selon des retards distribués paramétrés. Cette approche conserve davantage d'information et réduit le biais d'approximation associé à la discrétisation temporelle.

Les modèles à facteurs dynamiques (DFM) constituent une autre avancée majeure. Ils reposent sur l'idée qu'un petit nombre de facteurs latents explique la covariance observée entre un grand ensemble de séries macroéconomiques (Stock et Watson, 2016 ; Giannone et al., 2008). L'estimation peut s'effectuer via une analyse en composantes principales (ACP) ou par filtrage de Kalman dans un cadre d'espace d'état (Doz et al., 2012). Les travaux de Bañbura et Modugno (2010) ont par ailleurs étendu l'algorithme EM afin de traiter les panels déséquilibrés et les séries à fréquences mixtes. Ces modèles offrent ainsi une flexibilité appréciable, mais restent limités dans leur capacité à capturer les non-linéarités et les interactions complexes entre indicateurs.

2.2 Les modèles d'apprentissage automatique

L'apprentissage automatique (*machine learning*) s'impose progressivement comme une alternative puissante aux approches traditionnelles. Il repose sur des algorithmes capables d'apprendre des relations complexes à partir des données sans hypothèses structurelles préalables (Samuel, 1959 ; Kuhn et Johnson, 2016). Le *nowcasting* s'inscrit naturellement dans l'apprentissage supervisé, où la relation entre les indicateurs conjoncturels et la croissance du PIB est apprise à partir de données historiques étiquetées.

Les modèles de régression pénalisée (Ridge, LASSO, Elastic Net) visent à contrôler la variance et à améliorer la stabilité des estimations grâce à des contraintes de régularisation (Barhoumi et al., 2022). Les modèles d'arbres de décision et leurs extensions — forêts aléatoires, bagging ou boosting — offrent une flexibilité accrue pour traiter les non-linéarités et les interactions entre variables (James et al., 2021). Les machines à vecteurs de support (SVM, SVR) de Vapnik et Cortes (1995) exploitent des noyaux (*kernel trick*) pour modéliser des relations non linéaires dans des espaces de grande dimension. Enfin, l'apprentissage profond (*deep learning*) et les réseaux neuronaux récurrents (LSTM) se distinguent par leur aptitude à capter les dépendances temporelles des séries économiques (Schmidhuber, 2015). Ces méthodes se révèlent particulièrement performantes en contexte de volatilité ou de chocs économiques (Cascaldi-Garcia et al., 2023 ; Buell et al., 2021).

Malgré leurs atouts, ces approches présentent des défis méthodologiques : risque de sur-apprentissage, difficulté d'interprétation des modèles dits « boîtes noires » et forte sensibilité au réglage des hyperparamètres (James et al., 2021). Des méthodes comme la validation croisée, le *Grid Search* ou les valeurs de Shapley contribuent toutefois à atténuer ces limites en améliorant la robustesse et l'interprétabilité des prédictions.

2.3 Revue empirique et comparaison

La littérature empirique met en évidence la supériorité des modèles d'apprentissage automatique par rapport aux approches traditionnelles dans plusieurs contextes. En Inde, Ghosh et Ranjan (2023) montrent que les forêts aléatoires surpassent les modèles ARIMA et Prophet pour le *nowcasting* du PIB. Richardson et al. (2020) obtiennent des résultats similaires en Nouvelle-Zélande, où le SVM et le réseau neuronal réduisent significativement l'erreur quadratique moyenne par rapport à une autorégression de référence. De même, Al-Rawashdeh (2024) démontre, dans le cas de la Jordanie, la supériorité du modèle XGBoost sur les modèles à facteurs dynamiques, tandis que Zhang et al. (2023) confirment la robustesse de la régression Ridge en Chine, y compris en période de crise. Dans le contexte des données alternatives, Suphaphiphat, Wang et Zhang (2022) montrent que l'intégration de données Google Trends et environnementales améliore la précision des prévisions pour les économies européennes, surtout en période de rupture conjoncturelle.

Concernant les pays émergents, Galvez-Soriano (2023) observe au Mexique la performance des équations de pont dans un environnement de forte variance inter-séries. Muchisha et al. (2021) obtiennent pour l'Indonésie de meilleurs résultats avec les forêts aléatoires et la régression Elastic Net, tandis que Bolivar (2024) démontre que l'ajout de données satellitaires améliore significativement les prévisions du PIB bolivien. En Afrique subsaharienne, les travaux du FMI (Buell et al., 2021 ; Barhoumi et al., 2022) confirment la pertinence des modèles d'apprentissage automatique, en particulier lorsqu'ils sont enrichis de données alternatives telles que les paiements mobiles ou les recherches Google.

3 Données

3.1 Variable cible

Dans cet exercice de nowcasting, notre variable cible n'est autre que le taux de croissance trimestriel du Produit Intérieur Brut (PIB) réel de la Côte d'Ivoire en glissement annuel. La série utilisée provient du World Development Indicator (WDI) de la Banque Mondiale et s'étend sur la période du troisième trimestre 2013 (2013 T3) au dernier trimestre 2024 (2024 T4), soit la période de disponibilité des données. Sur cette période, la croissance du PIB de la Côte d'Ivoire a été d'abord marquée par une trajectoire soutenue avant 2020 avec un rythme moyen supérieur à 7% (Voir Graphique 1). Il s'en est suivi une rupture brutale en 2020 liée à la pandémie de COVID-19 avec une chute de la croissance à -1,5% au second trimestre de cette année. Enfin, une reprise progressive est perçue à partir de 2021 dans le contexte post-COVID suivie d'une période plutôt stable à partir de 2023. Cette dynamique justifie pleinement l'intérêt d'un dispositif de prévision en temps quasi réel, capable d'anticiper les retournements de cycle et de fournir des estimations fiables en période d'incertitude.



FIGURE 1 – Evolution du taux de croissance du PIB de la Côte d'Ivoire entre 2013 et 2024

3.2 Prédicteurs

Les prédicteurs proviennent de diverses sources selon la nature des données. Les principales sources de données traditionnelles incluent l'Agence Nationale de la Statistique (ANStat), le Fonds Monétaire International (FMI), la Banque mondiale, ainsi que les données financières de la Banque Centrale des États de l'Afrique de l'Ouest (BCEAO). Concernant les données alternatives, nous retenons des données satellitaires issues du World Development Indicator (WDI) de la Banque mondiale, des données sur la végétation de l'Organisation pour l'alimentation et l'agriculture (FAO) et des données Google Trends, extraites via l'API Python. La période globalement couverte par les données est celle de janvier 2010 à juin 2025, l'annexe A présente la période de disponibilité de chaque indicateur.

4 Méthodologie

4.1 Modèles

Pour prédire en temps quasi réel le taux de croissance du Produit Intérieur Brut (PIB) de la Côte d'Ivoire, nous retenons un panel de modèles d'apprentissage automatique et statistiques, éprouvés par la littérature et adaptés aux contraintes des pays en développement. Ce panel comprend six modèles de machine learning ainsi que deux références classiques (un modèle autorégressif et un modèle à facteurs dynamiques). Les caractéristiques et formulations mathématiques principales sont détaillées ci-dessous.

4.1.1 Benchmark autorégressif (AR)

Comme modèle de référence naïf et robuste, nous considérons un modèle autorégressif d'ordre p estimé sur le taux de croissance trimestriel du PIB. Ce benchmark permet d'évaluer les gains prédictifs apportés par les approches plus sophistiquées, notamment les modèles à facteurs dynamiques et les méthodes d'apprentissage automatique. Le modèle est spécifié comme suit :

$$y_t = \mu + \sum_{i=1}^p \beta_i y_{t-i} + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \sigma^2), \quad (1)$$

où y_t désigne le taux de croissance trimestriel du PIB, μ une constante, β_i les coefficients autorégressifs associés aux retards i , et ε_t un terme d'erreur assimilé à un bruit blanc gaussien.

Avant l'estimation, l'on s'assure de la stationnarité de la série par les tests de Dickey-Fuller augmenté (ADF), de Phillips-Perron (PP) et de Kwiatkowski-Phillips-Schmidt-Shin (KPSS). En cas de non-stationnarité, la série est rendue stationnaire par différenciation. Une fois cette condition vérifiée, l'ordre optimal p est déterminé par une recherche heuristique sur $p \in \{1, 2, 3, 4\}$, en retenant à chaque trimestre de nowcasting la valeur minimisant le critère d'information d'Akaike (AIC). La borne supérieure de quatre retards est motivée par la fréquence trimestrielle des données, permettant de couvrir au plus un cycle annuel tout en limitant le risque de sur-paramétrisation. Le modèle est finalement estimé par la méthode des moindres carrés ordinaires (MCO).

4.1.2 Modèle à facteurs dynamiques.

Le principe des modèles à facteurs dynamiques repose sur l'hypothèse qu'un petit nombre de facteurs latents non observables peut rendre compte de la dynamique commune sous-jacente à un ensemble de séries chronologiques observées, souvent corrélées entre elles.

L'équation d'état : Cette équation représente la dynamique des facteurs latents, variables inobservables qui capturent les mouvements sous-jacents communs aux données observées. L'équation factorielle est définie comme un système dynamique permettant aux facteurs inobservés \mathbf{F}_t d'évoluer selon un processus VAR(p) et peut s'écrire :

$$\mathbf{F}_t = A_1 \mathbf{F}_{t-1} + A_2 \mathbf{F}_{t-2} + \cdots + A_p \mathbf{F}_{t-p} + \mathbf{u}_t, \quad \mathbf{u}_t \sim \mathcal{N}(\mathbf{0}, Q).$$

On peut aussi écrire la forme d'espace d'état d'ordre 1 :

$$\mathbf{F}_{t+1} = A \mathbf{F}_t + \mathbf{u}_t, \quad \mathbf{u}_t \sim \mathcal{N}(\mathbf{0}, Q),$$

où la matrice A est la matrice de transition (ou la forme compagnon du VAR).

L'équation d'observation : Cette équation relie les facteurs latents aux variables observées. Elle précise comment les variables observées sont générées à partir des facteurs latents et peut s'écrire ainsi :

$$\mathbf{X}_t = \Lambda \mathbf{F}_t + \xi_t, \quad \xi_t \sim \mathcal{N}(\mathbf{0}, \Sigma_\xi).$$

Où :

- \mathbf{X}_t est le vecteur des variables observées à l'instant t ,
- Λ est la matrice de chargement factoriel (loadings) qui rend compte de l'influence des facteurs sur les variables observées,
- ξ_t est le vecteur des erreurs de mesure à l'instant t et Σ_ξ sa matrice de covariance.

Suivant Bańbura et Modugno (2010), on suppose souvent que Σ_ξ est diagonale autrement dit, que les erreurs idiosyncrasiques non corrélées.

Le nombre de facteurs statiques r est déterminé par les critères d'information proposés par Bai et Ng (2002). Le r optimal minimise ces critères ; ici, pour des raisons de parcimonie et de robustesse, nous limitons $r_{\max} = 2$. Quant à l'ordre p des facteurs dynamiques, il est choisi en minimisant le critère d'information d'Akaike (AIC) sur la grille $p \in \{1, 2, 3, 4\}$.

La croissance trimestrielle du PIB (en glissement annuel) est incorporée dans l'estimation des facteurs et des paramètres. Comme elle n'est pas observée mensuellement, la valeur trimestrielle est rattachée au troisième mois du trimestre tandis que les deux premiers mois sont traités comme manquants via une matrice de sélection S_t (pattern de disponibilité des observations).

Pour relier la mesure trimestrielle au contre-partie mensuelle latente, nous utilisons l'approximation de Mariano et Murasawa (2003), qui exprime la croissance trimestrielle comme une combinaison linéaire (opérateur polynôme en retard) des variations mensuelles latentes. Cette configuration est implémentée conformément à la pratique de Bańbura et Modugno (2010) dans les estimations d'un DFM avec données à fréquences mixtes et permet d'utiliser la méthode de ceux-ci pour l'estimation du modèle.

4.2 Méthodes basées sur l'apprentissage automatique

Nous considérons plusieurs familles de modèles supervisés pour le nowcasting du PIB : régressions pénalisées, méthodes d'arbres (Random Forest, Extra Trees) et Support Vector Regression. Les modèles sont évalués selon la RMSE et leurs hyperparamètres sont sélectionnés par validation croisée temporelle combinée à une Grid Search. Les prétraitement est ajusté uniquement sur l'échantillon d'entraînement afin d'éviter les fuites d'information.

Régressions pénalisées Les coefficients sont estimés en minimisant l'erreur quadratique augmentée d'une pénalité :

$$\min_{\beta_0, \beta} \sum_{t=1}^T \left(y_t - \beta_0 - \sum_{i=1}^n \beta_i x_{t,i} \right)^2 + \lambda \text{penalty}(\beta),$$

où $\lambda \geq 0$ règle l'intensité de la régularisation. Les variantes utilisées sont :

- **Ridge** : $\text{penalty}(\beta) = \sum_{i=1}^n \beta_i^2$ (norme L_2) — réduit la variance et atténue la multicolinéarité sans annulation des coefficients.
- **LASSO** : $\text{penalty}(\beta) = \sum_{i=1}^n |\beta_i|$ (norme L_1) — réalise aussi une sélection de variables en rendant certains coefficients nuls.
- **Elastic Net** : combinaison des deux pénalités,

$$\text{penalty}(\beta) = \alpha \sum_{i=1}^n |\beta_i| + (1 - \alpha) \sum_{i=1}^n \beta_i^2, \quad \alpha \in [0, 1],$$

permettant un compromis entre sélection et régularisation.

Forêts d'arbres (Random Forest & Extra Trees)

- **Random Forest** : agrégat d'arbres de décision construits sur des échantillons bootstrap ; à chaque nœud, un sous-ensemble aléatoire de prédicteurs est considéré pour sélectionner le meilleur seuil de séparation. Les prévisions sont la moyenne des prédictions des arbres. Hyperparamètres clés : nombre d'arbres (*n_estimators*), profondeur maximale (*max_depth*), taille minimale des feuilles (*min_samples_leaf*), et nombre de variables candidates par split (*max_features*).
- **Extra Trees** : similaire au Random Forest mais les seuils de coupure sont choisis aléatoirement (en plus de la sélection aléatoire des variables), ce qui augmente la diversité des arbres et accélère l'apprentissage tout en réduisant le surapprentissage.

Support Vector Regression (SVR) Le SVR cherche une fonction aussi plate que possible tout en limitant les écarts supérieurs à une marge ε . Le problème primal s'écrit :

$$\min_{w, b, \xi, \xi^*} \frac{1}{2} \|w\|^2 + C \sum_{t=1}^T (\xi_t + \xi_t^*)$$

sous les contraintes

$$\begin{cases} y_t - \langle w, x_t \rangle - b \leq \varepsilon + \xi_t, \\ \langle w, x_t \rangle + b - y_t \leq \varepsilon + \xi_t^*, \\ \xi_t, \xi_t^* \geq 0, \end{cases}$$

où $C > 0$ contrôle le compromis biais/variance. La prédiction prend la forme

$$\hat{y}_t = \sum_{i=1}^N \alpha_i K(x_t, x_i) + b,$$

avec $K(\cdot, \cdot)$ le noyau et N le nombre de vecteurs de support. Les noyaux considérés sont :

- linéaire : $K(x_i, x_t) = x_i^\top x_t$;
- polynomial : $K(x_i, x_t) = (\gamma x_i^\top x_t + r)^d$ (avec γ, r, d) ;
- gaussien (RBF) : $K(x_i, x_t) = \exp(-\gamma \|x_i - x_t\|^2)$.

Sélection des hyperparamètres et évaluation La sélection combine une validation croisée temporelle à 5 plis de type fenêtre extensible et une Grid Search sur l'espace d'hyperparamètres spécifique à chaque modèle. La performance retenue est la RMSE moyenne sur les plis :

$$\text{RMSE} = \sqrt{\frac{1}{T} \sum_{t=1}^T (y_t - \hat{y}_t)^2},$$

où y_t est l'observation et \hat{y}_t sa prédiction ; l'hyper-paramétrage minimisant la RMSE moyenne est retenu, puis le modèle est réentraîné sur l'échantillon d'apprentissage complet avant production des nowcasts. Le tableau 1 précise l'espace de recherche utilisé pour chaque méthode.

TABLE 1 – Espace de recherche des hyperparamètres par modèle

Modèle	Champs des hyperparamètres explorés
SVR (RBF)	$C \in [10^{-1}, 10^2]$ $\varepsilon = \text{IQR}(Y)/13,9 \pm 20\%$ $\gamma \in [10^{-2}, 10^2]$
SVR (linéaire)	$C \in [10^{-1}, 10^2]$ $\varepsilon = \text{IQR}(Y)/13,9 \pm 20\%$
SVR (polynomial)	$C \in [10^{-1}, 10^2]$ $\varepsilon = \text{IQR}(Y)/13,9 \pm 20\%$ $r \in \{0, 1\}$ $d \in \{1, 2, 3\}$
Random Forest	<code>max_depth</code> $\in \{\text{None}, 3, 6, 10\}$ <code>max_features</code> $\in \{\sqrt{.}, \log_2\}$ <code>min_samples_leaf</code> $\in \{1, 2, 3\}$ <code>n_estimators</code> $\in \{300, 600, 1000\}$
Extra Trees	<code>max_depth</code> $\in \{3, 6, 10\}$ <code>max_features</code> $\in \{\sqrt{.}, \log_2\}$ <code>min_samples_leaf</code> $\in \{1, 2, 3\}$ <code>n_estimators</code> = 300
LASSO	$\lambda \in [10^{-1}, 10^2]$
Ridge	$\lambda \in [10^{-1}, 10^2]$
Elastic Net	$\lambda \in [10^{-1}, 10^2]$ $\alpha \in [0, 1]$

4.3 Evaluation et comparaison des modèles

4.3.1 Evaluation de la performance des modèles

Dans ce papier, pour évaluer et sélectionner le modèle de nowcasting du PIB le plus performant, nous procédons à une division temporelle de notre jeu de données en deux parties distinctes. La première, utilisée pour l'entraînement des modèles, couvre la période allant du troisième trimestre de l'année 2013 au dernier trimestre 2019. Sur cette période, nous entraînons nos modèles d'apprentissage automatique et optimisons leurs hyperparamètres en utilisant une validation croisée temporelle à 5 plis comme défini ci-dessus dans la sous-section dédiée aux choix des hyperparamètres. Cette méthode permet d'alterner entre un jeu d'entraînement et un jeu de validation à chaque itération, garantissant ainsi une évaluation plus robuste du modèle. Parallèlement, nous estimons nos modèles de référence sur cette même période afin d'effectuer des comparaisons de performance. C'est également sur cette période qu'est étudiée la stationnarité des variables. Une fois l'entraînement effectué, nous évaluons chaque modèle en le testant sur la seconde partie de notre jeu de données, s'étendant du premier trimestre de l'année 2020 au quatrième trimestre de l'année 2024. Il est à noter que cette période couvre également la crise économique liée à la pandémie de COVID-19 et la phase de reprise, ce qui permet de tester la capacité des modèles à généraliser et à prévoir dans un contexte économique versatile. Les performances des modèles sont quantifiées à l'aide de l'erreur moyenne absolue (Mean Absolute Error, MAE) et de la racine carrée de l'erreur quadratique moyenne (RMSE).

4.3.2 Test de Diebold–Mariano proposé par Harvey, Leybourne, et Newbold (1997)

Au-delà des métriques usuelles, nous procédons à des comparaisons deux à deux des modèles à l'aide du test de Diebold–Mariano (1995), corrigé par Harvey–Leybourne–Newbold (1997) afin de corriger les biais liés aux petits échantillons. Ce test permet d'identifier le modèle statistiquement le plus précis en comparant l'exactitude prédictive de deux modèles pour une même variable cible et un même horizon de prévision.

Lorsqu'un modèle est retenu à l'issue de ce test, cela indique qu'il présente une perte moyenne significativement plus faible que l'autre. Le test peut être spécifié à partir de différentes fonctions de perte, telles que l'erreur absolue ou l'erreur quadratique. Dans cette étude, nous retenons l'erreur quadratique afin d'assurer une cohérence avec notre métrique principale d'évaluation.

Les hypothèses du test sont définies comme suit :

$$\begin{aligned} H_0 : \mathbb{E}(d_t) &= 0 \quad (\text{absence de différence significative}), \\ H_1 : \mathbb{E}(d_t) &\neq 0 \quad (\text{différence significative}). \end{aligned}$$

où d_t désigne la perte différentielle à l'instant t , définie comme la différence entre les erreurs quadratiques des deux modèles considérés.

Sous l'hypothèse nulle H_0 , la statistique de Diebold–Mariano corrigée selon Harvey–Leybourne–Newbold suit approximativement une loi de Student à $p - 1$ degrés de liberté, où p représente le nombre de prévisions comparées. En cas de rejet de l'hypothèse nulle, le signe de la moyenne de la perte différentielle \bar{d} permet d'identifier le modèle le plus précis. Pour deux modèles 1 et 2, dont les erreurs quadratiques à l'instant t sont notées $\varepsilon_{1,t}^2$ et $\varepsilon_{2,t}^2$, on a :

$$\bar{d} = \frac{1}{p} \sum_{t=1}^p d_t = \frac{1}{p} \sum_{t=1}^p (\varepsilon_{1,t}^2 - \varepsilon_{2,t}^2).$$

Ainsi, si $\bar{d} > 0$, le modèle 2 est statistiquement plus précis, tandis que si $\bar{d} < 0$, le modèle 1 est statistiquement plus précis.

4.4 Evaluation de la robustesse des modèles sur les sous périodes de la dynamique du PIB ivoirien

Comme le graphique 1 le montre, la croissance économique de la Côte d'Ivoire a connu une période de récession correspondant à celle de la pandémie de COVID-19. Ainsi, afin de s'assurer que les modèles choisis soient suffisamment robustes pour gérer les périodes telles que les crises économiques, nous procédons à une évaluation supplémentaire de leurs performances durant la crise COVID-19, la reprise rapide et la période de stabilité qui s'en est suivi. La première phase de cette évaluation consiste à entraîner les modèles sur les données du troisième trimestre 2013 au dernier trimestre 2019. Cette période est choisie car elle inclut à la fois des données économiques normales avant la crise COVID-19, ce qui permet d'analyser la réaction du modèle face aux perturbations économiques majeures et sans précédent. Une fois les modèles entraînés sur cette période, nous évaluons leur performance sur la période COVID-19 en calculant les mêmes métriques définies précédemment. Cette étape nous permet de tester si les modèles peuvent encore produire des prévisions fiables pendant un choc économique important. Ensuite, la période d'entraînement des modèles est étendue à la période de la crise et les modèles sont réentraînés, afin de tenir compte des changements structurels introduits par la crise COVID-19. Cette phase est importante, car la pandémie a pu provoquer des changements dans la structure de l'économie et donc prendre cette période en compte le modèle permet de tester l'adaptabilité du modèle aux nouvelles dynamiques économiques post-COVID. La performance est ensuite évaluée sur la période de reprise. Enfin, pour la dernière période, celle de stabilité, les modèles sont entraînés sur le panel couvrant également la période de reprise, puis évalué sur la dernière période. Ainsi, l'évaluation de la robustesse des modèles repose sur leur capacité à maintenir une bonne précision de prévision pendant les différentes sous-périodes notamment celles de crises et leur aptitude à s'adapter aux changements économiques post-crise.

4.5 Choix final du modèle central de nowcasting

Le choix final du modèle de nowcasting repose sur une évaluation approfondie de ses performances prédictives. La sélection se fait principalement sur la base des résultats obtenus sur la période de test allant de 2020T1 à 2024T4. Toutefois, le modèle idéal serait celui qui minimise l'erreur quadratique moyenne (RMSE) sur l'ensemble de la période de test, tout en présentant une erreur moyenne absolue (MAE) concurrente par rapport aux autres approches. Il devrait également dominer ses alternatives dans le cadre du test de Diebold–Mariano corrigé HLN, appliqué aux pertes quadratiques, ce qui atteste de sa supériorité statistique. Enfin, il doit maintenir des erreurs faibles lorsqu'il est confronté à des sous-périodes particulièrement instables, telles que la crise de la COVID-19, ainsi que lors des phases de reprise et de stabilisation. Cette combinaison de critères permettrait d'identifier le modèle le plus adapté à tous les contextes pour le nowcasting de la croissance du PIB ivoirien. Cependant, un tel modèle demeure difficile à trouver, la littérature souligne que certains sont meilleurs dans des contextes particuliers tel que les périodes normales. Ainsi, nous nous proposons d'identifier les meilleurs modèles sur les différentes sous-périodes afin de

mettre en évidence leur forces respectives pour de futures perspectives de nowcasting. Conformément aux études précédentes tel que celle de Bolivar (2024) ou encore de Richardson et al. (2020), nous continuons l'analyse avec le meilleur modèle sur l'ensemble de test. Ce modèle est ensuite réentraîné sur l'ensemble des données disponibles afin de consolider une base historique et de tirer parti de toute l'information disponible avant la production des prévisions futures.

5 Evaluation de l'importance des prédicteurs

Comprendre comment les prédictions sont obtenues et quelles sont les variables qui les conduisent, est pertinent dans la mesure où cette analyse peut permettre de déterminer les variables auxquelles le modèle est sensible donnant lieu par exemple à de possibles ajustements pour la maintenance lorsque le contexte change. Ainsi, dans ce dernier point de ce chapitre, nous présentons comment est évaluée la contribution des variables, ainsi que l'impact des données alternatives.

5.1 Évaluation de la contribution des variables

Les modèles de *machine learning* sont souvent considérés comme des « boîtes noires », ce qui signifie qu'il est difficile de comprendre comment les prédictions sont obtenues. Pour surmonter cette limite d'interprétabilité, nous recourons à la *décomposition de Shapley*, un outil issu de la théorie des jeux coopératifs qui permet d'attribuer équitablement la contribution de chaque variable explicative aux prédictions du modèle.

La valeur de Shapley ϕ_{x_i} pour une variable x_i est donnée par :

$$\phi_{x_i}(y) = \sum_{S \subseteq N \setminus \{x_i\}} \frac{s!(n-s-1)!}{n!} [y(S \cup \{x_i\}) - y(S)] \quad (2)$$

où :

- N est l'ensemble des variables explicatives et n sa taille ;
- S est un sous-ensemble ou coalition de variables explicatives de taille s ;
- $y(S)$ est la prédiction du modèle lorsque seules les variables de S sont considérées ;
- $y(S \cup \{x_i\}) - y(S)$ représente la contribution marginale de la variable x_i à la prédiction finale du modèle y lorsqu'elle est incluse dans la coalition S .

Le facteur de pondération $\frac{s!(n-s-1)!}{n!}$ correspond à la probabilité que la variable x_i se joigne à la coalition S , en supposant que toutes les permutations des variables sont équiprobables. Cela garantit que chacune est prise en compte de manière juste dans le calcul de sa contribution marginale.

Ainsi, la valeur de Shapley ϕ_{x_i} mesure la moyenne des contributions marginales du prédicteur x_i à toutes les coalitions auxquelles il est susceptible d'adhérer. En d'autres termes, elle estime l'impact moyen de cette variable sur la prédiction finale du modèle, en tenant compte de toutes les configurations possibles des variables dans le modèle. Elle nous permettra de déterminer les variables les plus importantes en termes de contribution dans le modèle.

5.2 Évaluation de l'impact des données alternatives

Au-delà des variables traditionnelles, notre cadre de nowcasting inclut également des prédicteurs alternatifs. Il est donc essentiel d'évaluer leur pertinence en analysant leur impact dans le dispositif de prévision. Pour ce faire, nous entraînons d'abord le modèle de nowcasting uniquement avec les données traditionnelles, puis nous calculons la MAE et la RMSE correspondantes. Ces métriques sont ensuite comparées à celles obtenues lorsque l'ensemble des prédicteurs (traditionnels et alternatifs) est utilisé.

Cette démarche permet de quantifier la valeur ajoutée des données alternatives et de valider leur rôle en tant que compléments aux indicateurs macroéconomiques classiques dans le nowcasting du PIB ivoirien.

6 Résultats

6.1 Comparaison des modèles

Le graphique 4 compare les performances hors-échantillon, en fenêtre extensible, des modèles en RMSE et MAE. Les modèles de régressions pénalisées (Ridge, Elastic Net, Lasso) affichent les plus faibles erreurs, avec des RMSE comprises entre 2,4 et 2,6 et des MAE avoisinant 1,5. Ces résultats traduisent une nette amélioration par rapport au modèle autorégressif (AR), qui affiche une RMSE de 3,245 et une MAE de 2,150 soit des erreurs

quadratiques et erreurs absolues réduites respectivement d'environ 25% et de 30% en faveur des modèles de machine learning. Cela illustre leur aptitude à capter la dynamique globale de la croissance du PIB sur la période mieux que le modèle AR. Le SVR avec noyau RBF obtient des résultats comparables, mais légèrement moins performants en termes de MAE (1,738), ce qui signifie que ses prédictions s'éloignent plus en moyenne de la valeur observée. Par ailleurs, le SVR linéaire est plus stable que les modèles basés sur arbre de décision avec une RMSE de 2,516 points contre 2,774 et 2,779 respectivement pour Random Forest et Extra Trees. Cependant, ses prédictions ont un écart absolu moyen de 2,147 points contre 1,767 et 1,665 pour Random Forest et Extra Trees soit un écart de plus de 20% en sa défaveur. Ainsi voit-on que bien qu'il soit relativement stable en termes de variance des erreurs (RMSE), il se révèle moins précis en moyenne que les méthodes arborescentes lorsqu'on considère simultanément les deux métriques (Graphique 4).

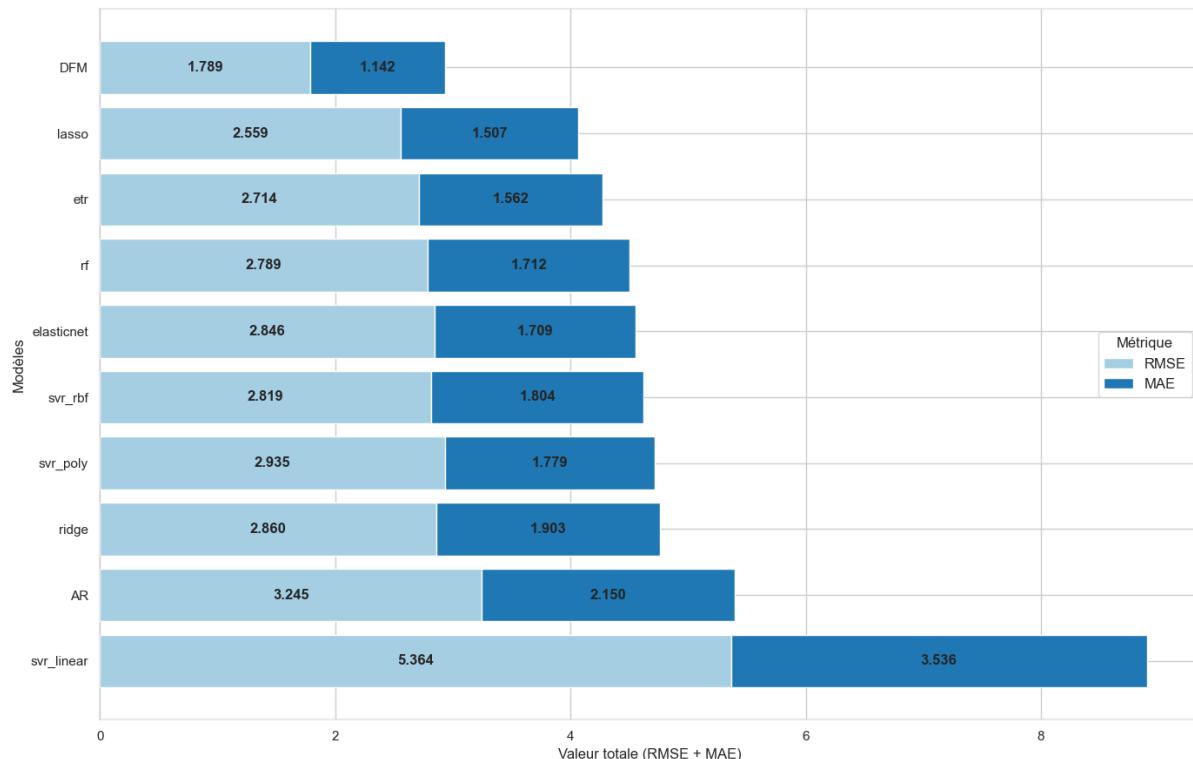


FIGURE 2 – Comparaison des modèles sur la période de test