

Intelligence artificielle et prévision à moyen terme de l'inflation dans la CEMAC

Thierry MVONDO

**Docteur en sciences
économiques**

Direction des Études,
de la Recherche, des Statistiques
et des Relations Internationales
Mvondot@beac.int

**BANQUE DES ETATS DE
L'AFRIQUE CENTRALE**

736, Avenue Monseigneur Vogt
BP: 1917, Yaoundé-Cameroun
Tel: (237) 22234030 / 22234060
Fax: (237) 22233329
www.beac.int

Les opinions émises dans ce document de travail sont propres à leur(s) auteur(s) et ne représentent pas nécessairement la position de la Banque des Etats de l'Afrique Centrale.

The opinions expressed in this working paper are those of the author(s) and do not necessarily represent the views of the Bank of Central African States.

Intelligence artificielle et prévision à moyen terme de l'inflation dans la CEMAC

Artificial Intelligence and Medium-term Inflation Forecast in CEMAC

Dr. Thierry Mvondo

Novembre 2025

Résumé :

Cet article montre que la prévision à moyen terme de l'inflation dans la CEMAC peut-être améliorée par la combinaison des outils classiques aux algorithmes d'apprentissage automatique (ML) et de l'intelligence artificielle (IA). Pour cela, le résidu du SBVARX actuellement utilisé est entraîné suivant trois approches : le XGBoost, le Random Forest et le SVR. Les métriques calculées montrent que les approches hybrides (SBVARX-IA/ML) performant en général mieux que le modèle SBVARX standard. Le test DM conclut quant à lui à un surcroît de précision en prévision avec l'approche SBVARX-SVR, notamment lorsque l'horizon de prévision tend vers le moyen terme. Pour la BEAC dont la politique monétaire est confrontée à de nombreux chocs exogènes depuis l'ère Covid-19, cette approche hybride permettrait d'intégrer les non-linéarités et complexités qui écarteraient de plus en plus les prévisions à moyen terme de l'inflation de ses réalisations.

Mots-clés -----

Prévision – Inflation – SBVAR-Random Forest – SBVAR-XGBoost – SBVAR-SVR

Classification JEL : E31, E37, C22, C53, C55, O55

Abstract:

This paper shows that medium-term inflation forecasting in the CEMAC can be improved by combining classical tools with machine learning (ML) and artificial intelligence (AI) algorithms. For this purpose, the residual of the currently used SBVARX is trained using three approaches: XGBoost, Random Forest, and SVR. The calculated metrics show that the hybrid approaches (SBVARX-AI/ML) generally perform better than the standard SBVARX model. The DM test, for its part, concludes that there is an increase in forecasting precision with the SBVARX-SVR approach, particularly when the forecasting horizon tends towards the medium term. For the BEAC, whose monetary policy has been facing numerous exogenous shocks since the Covid-19 era, this hybrid approach would make it possible to integrate the nonlinearities and complexities that are increasingly moving medium-term inflation forecasts away from actual outcomes.

Key Words -----

Forecast – SBVAR-Random Forest – SBVAR-XGBoost – SBVAR-SVR

JEL Classification : E31, E37, C22, C53, C55, O55

1. Introduction

La conduite de la politique monétaire par la Banque des Etats de l'Afrique Centrale (BEAC) est, depuis 2019, confrontée à des défis inédits. Ces défis découlent de la succession de chocs exogènes majeurs qui ont bouleversé l'environnement macroéconomique mondial : la pandémie de Covid-19, les tensions géopolitiques issues de la guerre russo-ukrainienne et du conflit au Moyen-Orient, ainsi que les perturbations persistantes des chaînes d'approvisionnement mondiales [Parast et Shekarian (2019), Dolgui et Ivanov (2021)]. Ces crises ont induit une volatilité extrême des cours des matières premières et des denrées alimentaires, rendant l'exercice de prévision de l'inflation particulièrement ardu.

Dans ce contexte de risque global ou poly crise (Magnan et al, 2021), un écart croissant est observé entre les prévisions d'inflation établies par la BEAC et les réalisations effectives de cet indicateur. L'explication de cet écart réside probablement dans les complexités et les non-linéarités induites par ces chocs sur la dynamique des prix dans la zone CEMAC. Ces dynamiques remettent fondamentalement en question l'hypothèse de linéarité sur laquelle reposent les approches économétriques traditionnelles actuellement mobilisées par la Banque Centrale, notamment les modèles Vectoriels Auto Régressifs (VAR) standards. Cette situation pose avec acuité la nécessité d'améliorer les outils actuels de prévision de l'inflation à moyen terme, pour garantir l'atteinte de l'objectif statutaire de stabilité monétaire.

Conformément à l'article premier de ses Statuts, la BEAC a pour objectif principal la stabilité monétaire, qui se traduit sur le plan externe, par un taux de couverture extérieure de la monnaie suffisant et, sur le plan interne, par un taux d'inflation maintenu en deçà de 3% annuellement. Respecter cette cible d'inflation nécessite une anticipation précise des mouvements de prix, les actions de politique monétaire n'ayant un effet complet qu'entre six et huit trimestres. Pour ce faire, la BEAC s'appuie classiquement sur la modélisation VAR, privilégiant notamment l'approche bayésienne structurelle (SBVARX) avec restrictions de signes pour combiner les taux de croissance des variables macroéconomiques clés (PIB, masse monétaire, dépenses publiques, etc.).

Face à des problématiques similaires, de nombreuses banques centrales à travers le monde se sont tournées vers les approches d'intelligence artificielle (*AI*, *Artificial Intelligence*) et d'apprentissage automatique (*Machine Learning*, *ML*). Ces techniques sont particulièrement adaptées pour gérer de larges volumes de données, capter des relations non-linéaires et améliorer la précision des prévisions (*nowcasting* et horizons plus longs), surtout en périodes de forte volatilité économique. Des institutions comme la Banque Centrale Européenne (Lenza et al., 2023), la Banque d'Angleterre (BBIM, 2025), la Banque du Canada (Braanlund et al, 2023) et la Réserve Fédérale américaine (Fed/St-Louis, 2024) intègrent désormais le ML, souvent via des modèles hybrides, pour compléter ou affiner leurs cadres économétriques traditionnels.

Fort de ce constat et de l'expérience internationale, cet article se propose d'explorer une voie d'amélioration pour les prévisions d'inflation de la BEAC : la combinaison de la modélisation SBVARX actuelle avec des approches d'IA/ML. L'objectif est double : évaluer si ces modèles hybrides surpassent les performances prédictives du SBVARX classique, et démontrer leur

capacité supérieure à capter les non-linéarités issues des crises récentes (Covid-19, guerre en Ukraine, crise du fret).

Au-delà de ces apports méthodologiques, l'étude analysera les implications pratiques de cette approche en termes de politique monétaire pour l'affinement du cadre décisionnel de la BEAC.

La suite de cet article est organisée comme suit : le point 2 effectue une revue de la littérature sur les modèles de prévision de l'inflation et l'apport des modèles hybrides ; le point 3 présente la méthodologie utilisée ; le point 4 discute de la portée des résultats obtenus ainsi que des implications de politique monétaire ; et le point 5 conclut l'analyse.

2. Les modèles de prévision de l'inflation : des approches classiques aux apports de l'IA/ML

Cette section dresse un état de l'art des cadres théoriques et méthodologiques mobilisés pour la prévision de l'inflation. Nous y présentons les fondements théoriques qui sous-tendent les modèles économétriques traditionnels, avant de détailler ces derniers et d'ouvrir sur la contribution récente des approches d'intelligence artificielle et d'apprentissage automatique (IA/ML).

2.1. Fondements théoriques de l'inflation : les sources des co-mouvements économiques

Les modèles économétriques classiques, en particulier les approches VAR, cherchent à capturer statistiquement les relations issues des grandes théories économiques sur l'inflation. Ces théories se regroupent principalement autour de trois mécanismes fondamentaux, auxquels s'ajoute l'approche structuraliste, particulièrement pertinente pour les économies en développement comme celles de la CEMAC.

2.1.1. L'Inflation par la demande et le cadre Néo-Keynésien

La théorie keynésienne ([Lavoie, 1985](#)) pose que l'inflation résulte d'un excès de demande globale par rapport à l'offre globale (capacité de production). Cette idée est affinée par la Nouvelle École Keynésienne ([Akerlof et Shiller, 2009](#)) qui intègre les rigidités de prix et de salaires, ainsi que l'importance des anticipations rationnelles. Les déterminants clés deviennent l'écart de production (*output gap*), les anticipations et les rigidités nominales. Dans ce cadre, la politique monétaire, via l'ajustement du taux d'intérêt, joue un rôle central pour ancrer les anticipations et stabiliser les prix.

2.1.2. L'inflation par les coûts (ou par l'offre)

Ce mécanisme explique l'augmentation des prix par une hausse des coûts de production répercutée sur les prix de vente. Contrairement à l'inflation par la demande, l'offre globale diminue face à une demande stable, créant des tensions sur les prix. Les causes principales incluent l'augmentation des coûts des matières premières (chocs pétroliers), la spirale prix-salaires, l'inflation importée et la fiscalité. Dans ce contexte, la politique monétaire se révèle

souvent moins efficace, car elle ne peut pas agir directement sur les goulots d'étranglement de l'offre.

2.1.3. L'Inflation monétaire et le monétarisme

Pour [Milton Friedman \(1968\)](#), "l'inflation est toujours et partout un phénomène monétaire", découlant d'une croissance de la masse monétaire supérieure à celle de la production réelle. La théorie quantitative de la monnaie résume ce postulat. Ses implications politiques sont claires : le contrôle de l'inflation passe par la limitation stricte de la croissance monétaire par la banque centrale, sauf en cas de dominance fiscale ([Sargent et Wallace, 1981](#)).

2.1.4. L'Inflation structurelle : une pertinence pour la CEMAC

Développée par les économistes Latino-américains (années 1950-1960), cette théorie attribue l'inflation des pays en développement à des déséquilibres structurels visibles dans la CEMAC : goulots d'étranglement agricoles, pénuries de devises, contraintes fiscales et rigidités institutionnelles. Pour les structuralistes, les politiques monétaires restrictives classiques sont inefficaces face à ces causes profondes ; elles nécessitent plutôt des réformes structurelles.

2.2. Les Méthodes économétriques classiques de prévision de l'inflation

Les banques centrales s'appuient sur un arsenal de méthodes économétriques pour traduire ces théories en prévisions quantitatives. Le choix de la méthode dépend de l'horizon de prévision et des données disponibles.

2.2.1. Les Modèles univariés (séries temporelles)

Ces modèles prédisent l'inflation uniquement sur la base de son historique. Ils incluent les modèles ARIMA ([Box et Jenkins, 1970](#)) pour les prévisions à court terme, et les modèles de type *Random Walk* (marche aléatoire), souvent utilisés comme modèles de référence (*benchmarks*).

2.2.2. Les Modèles multivariés (Vector AutoRegressive - VAR)

La modélisation VAR ([Sims, 1980](#)) est l'approche privilégiée par de nombreuses banques centrales, dont la BEAC. Elle modélise plusieurs variables macroéconomiques simultanément, sans imposer de restrictions théoriques strictes *a priori*. Toutes les variables dépendent de leurs propres valeurs passées et de celles des autres variables du système (inflation, taux d'intérêt, PIB, etc.), permettant l'analyse des fonctions de réponse impulsionnelle face aux chocs.

2.2.3. Les Modèles structurels et semi-structurels

Ces modèles intègrent directement les relations de causalité issues de la théorie économique. On y trouve :

- La Courbe de Phillips ([Phillips, 1950](#)) et sa version Néo-Keynésienne moderne, liant chômage/écart de production et inflation.
- Les modèles à grande échelle comme les DSGE (Dynamic Stochastic General Equilibrium) ou les QPM (Quarterly Projection Model), qui décrivent l'ensemble de l'économie via des systèmes d'équations complexes.

2.3. L'Apport des approches IA/ML à la prévision de l'inflation : revue de la littérature et applications en banques centrales

La littérature récente converge vers l'idée que les méthodes d'intelligence artificielle (IA) et d'apprentissage automatique (*Machine Learning*, ML) apportent des améliorations significatives par rapport aux modèles économétriques traditionnels. Ces gains de performance sont particulièrement notables dans les environnements riches en données (*data-rich environments*) et face aux défis posés par l'instabilité structurelle récente (Grande Modération, pandémie de COVID-19, chocs d'approvisionnement, tensions géopolitiques). Les modèles linéaires conventionnels (Courbe de Phillips, ARIMA, VAR) montrent en effet leurs limites face aux non-linéarités et aux ruptures de régimes induites par ces chocs.

2.2.1. Avantages méthodologiques et modèles clés

L'IA/ML offre un nouveau paradigme axé sur les données (*data-driven*), permettant de découvrir des motifs cachés, des dépendances à long terme, et d'exploiter des variables jusqu'alors inexploitable, comme les données textuelles via le Traitement Automatique du Langage (*Natural Language Processing*, NLP).

Plusieurs modèles de ML se distinguent dans la littérature sur la prévision de l'inflation à moyen terme (horizons cruciaux pour les banques centrales, jusqu'à deux ans) :

- Régressions Pénalisées (LASSO, Ridge, Elastic Net) : efficaces pour gérer de grands ensembles de prédicteurs (*big data*) via la régularisation. Des études, notamment celles de [Liu, Pan et Xu \(2024\)](#) pour le FMI, démontrent qu'elles surpassent les modèles de référence pour la prévision de l'inflation de base au Japon, grâce à l'intégration de variables étendues (anticipations ménages, taux de change). [Joseph et al. \(2021\)](#) soulignent leur performance compétitive et leur meilleure interprétabilité par rapport aux modèles non linéaires complexes ;
- Modèles Ensemblistes (Forêts Aléatoires, *Boosting* - XGBoost, LightGBM) : ces méthodes non paramétriques excellent dans la capture des non-linéarités et des interactions complexes. [Medeiros et al \(2021\)](#) montrent que les *Random Forests* dominent souvent les autres approches pour la prévision de l'inflation américaine, notamment grâce à leur capacité de sélection automatique de variables ;
- Réseaux de Neurones Profonds (*Deep Learning*) : Bien que plus complexes, des architectures avancées comme les RNN (*Recurrent Neural Networks*) et les transformateurs sont explorées pour leur capacité à capturer les dépendances temporelles de long terme dans les séries économiques.

2.2.2. Intégration dans la pratique des banques centrales

L'instabilité inflationniste récente a relancé l'intérêt des banques centrales pour ces outils flexibles. L'approche privilégiée consiste souvent à développer des modèles hybrides, qui combinent la rigueur structurelle des modèles traditionnels (comme le SBVAR) à la flexibilité adaptative de l'IA/ML.

Plusieurs institutions monétaires de premier plan intègrent désormais ces méthodes :

- Banque Centrale Européenne (BCE) : la BCE explore activement les modèles ML (Random Forest, XGBoost, LSTM) pour prévoir l'inflation de la zone euro (Lenza *et al.*, 2023). Les gains de précision sont significatifs sur les horizons courts (1 à 3 mois), surtout lors de chocs exogènes majeurs. La BCE insiste sur l'importance de l'explicabilité (*explainability*) via des outils comme SHAP pour éclairer les décisions ;
- Banque d'Angleterre : elle a développé un modèle de référence, le *Blockwise Boosted Inflation Model* (BBIM), qui repose sur un empilement d'algorithmes *boosting* appliqués à différents blocs analytiques (énergie, marché du travail, etc.). Le BBIM surpasse régulièrement les modèles semi-structurés internes en période de forte volatilité des prix ;
- Réserve Fédérale Américaine (Fed) : les travaux se concentrent sur l'intégration du NLP et des Grands Modèles de Langage (LLMs) pour analyser l'hétérogénéité des données textuelles (discours, rapports d'entreprises) afin de mesurer en temps réel le ton économique et les anticipations des agents (Fed/St-Louis, 2024).

2.2.3. Défis et limites

Malgré ces avancées, des défis persistent. Le principal obstacle pour les décideurs politiques reste le problème de la « boîte noire » (manque d'interprétabilité) des modèles complexes d'IA, un point crucial pour justifier les décisions de politique monétaire devant le public et les gouvernements.

En conclusion, la littérature confirme le potentiel transformateur de l'IA et du ML dans la prévision de l'inflation. Les modèles basés sur ces techniques offrent une précision supérieure, notamment à moyen terme, en exploitant la richesse des données modernes et en modélisant des dynamiques complexes. Ils sont destinés à devenir un complément indispensable aux méthodes économétriques traditionnelles.

3. Méthodologie

Conformément aux outils de prévision à moyen terme utilisés par la BEAC et à la revue de la littérature présentée précédemment, l'approche Structurale Bayésienne Vectorielle Auto Régressive avec variables exogènes (SBVARX) est retenue comme modèle de base pour cette étude. Notre approche repose sur un modèle hybride : nous supposons que le SBVARX capture la dynamique linéaire des données, tandis que les résidus (erreurs de prévision)

comportent des non-linéarités et des complexités susceptibles d'être mieux modélisées par une approche IA/ML.

Le modèle global que nous retenons se présente sous la forme suivante :

$$Y_t = \hat{Y}_t\{\text{SBVAR}\} + \mathcal{F}_{AI/ML}(X_t) + \epsilon_t^{\text{SBVARX}} \quad (1)$$

Ainsi, ce modèle hybride modélise l'inflation réelle (Inf_t) comme la somme de la prévision linéaire structurelle ($Inf_{t+h|t}^{\text{SBVARX}}$) et d'une composante non-linéaire ou résiduelle ($RÉS_{t+h|t}$), qui est estimée par un modèle IA/ML ($\mathcal{FIA/ML}$): $Inf_t = Inf_{t+h|t}^{\text{SBVARX}} + RÉS_{t+h|t}$.

3.1 Estimation du modèle de base SBVARX et obtention des résidus

Le modèle SBVARX (Structural Bayesian Vector Autoregression with exogenous variables) nous permet de modéliser l'interdépendance linéaire entre un ensemble de variables endogènes (Y_t) et incorpore des variables exogènes (X_t) qui peuvent influencer le système (nous avons retenu le cours du baril dans notre cas).

Le modèle SBVARX sous-jacent s'écrit par conséquent :

$$\Gamma_0 \hat{Y}_t\{\text{SBVAR}\} = C + \Gamma_1 \hat{Y}_{t-1}\{\text{SBVAR}\} + \dots + \Gamma_p \hat{Y}_{t-p}\{\text{SBVAR}\} + \Psi_0 X_0 + \dots + \Psi_q X_q + U_t \quad (2)$$

Où :

- $\hat{Y}_t\{\text{SBVAR}\}$ est un vecteur colonne de N variables macroéconomiques (endogènes) prépondérantes à la prévision de l'inflation, y compris cette dernière au temps t (incluant l'inflation) ;
- X_t est un vecteur colonne de variables exogènes ;
- C est un vecteur d'intercepts ;
- Γ_i et Ψ_j sont des matrices de coefficients ;
- U_t est un vecteur de chocs ou résidus non corrélés, avec une matrice de covariance Σ_u .

Parce que la relation (2) ci-dessus est difficilement estimable du fait de l'endogénéité des variables, nous faisons usage de la forme réduite donnée par :

$$\hat{Y}_t\{\text{SBVAR}\} = \Gamma_0^{-1} \Gamma_1 \hat{Y}_{t-1}\{\text{SBVAR}\} + \dots + \Gamma_0^{-1} \Gamma_p \hat{Y}_{t-p}\{\text{SBVAR}\} + \Gamma_0^{-1} \Psi_0 X_0 + \dots + \Gamma_0^{-1} \Psi_q X_q + V_t \quad (3)$$

Sous cette forme où l'estimation est possible même par les moindres carrés, la corrélation entre les résidus ne permet pas l'analyse impulsionnelle. Nous revenons à la relation (2) qui comporte toutefois plus de paramètres ($\frac{k(k-1)}{2}$) que la relation (3) par l'imposition de restrictions de signes que permet le VAR bayésien. De ce fait, les coefficients (Γ_i et Ψ_j) sont estimés non pas seulement par maximum de vraisemblance, mais en intégrant des *priors* de Minnesota pour éviter le surajustement et améliorer la stabilité des prévisions.

Les restrictions de signes retenues, par ailleurs compatibles avec la pratique dans les banques centrales sont les suivantes :

Tableau n°1 : restrictions de signes pour le SBVARX

Variable	Choc d'offre	Choc de demande	Choc de demande de monnaie	Choc de politique monétaire
Inflation	+	+	0	-
Output gap	-	+	0	-
Croissance M2	0	0	+	0
Taux d'intérêt	0	+	+	+

La forme réduite du modèle ci-dessus (3) est réécrite sous une format compacte tel que :

$$Y_t = X_t \Phi + u_t \quad (4)$$

Avec :

$X_t = [I, Y_{t-1}, \dots, Y_{t-p}]$, la matrice des régresseurs
 Φ , la matrice contenant tous les coefficients.

Suivant la logique bayésienne nous traitons les coefficients de la matrice Φ comme des variables aléatoires et donc, nous leur assignons des distributions a priori. L'inférence a posteriori est fait par la méthode des chaînes de Markov et les simulations de Monte Carlo (MCMC).

Les distributions a posteriori de Φ , sont obtenues suivant la relation :

$$P(\Phi|Y) \propto P(Y|\Phi) \times P(\Phi) \quad (5)$$

Avec : $P(Y|\Phi)$ la fonction de vraisemblance supposée Gaussienne et $P(\Phi)$ La distribution a priori.

L'estimation du modèle SBVAR par la méthode bayésienne fournit respectivement la matrice de coefficients $\hat{\Phi}$ et le vecteur des résidus \hat{u}_t obtenu tel que :

$$\hat{\epsilon}_t^{SBVARX} = Y_t - X_t \hat{\Phi} \quad (6)$$

En définitive, la prévision de l'inflation à l'horizon h , $\hat{Y}_{t+h|t}$ {SBVAR}, est obtenue par itérations récursives du système (3) et, l'erreur de prévision du modèle SBVARX pur est notée $\epsilon_{t+h|t}^{SBVARX}$.

3.2 Modélisation des résidus par les méthodes IA/ML

Le modèle IA/ML ($\mathcal{F}_{IA/ML}$) est entraîné pour prédire ou affiner la composante résiduelle ($RÉS_{t+h|t}$) en utilisant un ensemble étendu de prédicteurs Z_t (qui peuvent inclure les variables Y_t, X_t et d'autres données à haute fréquence) : $RÉS_{t+h|t} \approx \mathcal{F}_{IA/ML}(Z_t)$.

Nous supposons ici que l'erreur de prévision du modèle SBVARX standard contient encore des informations inexploitées, souvent liées à des non-linéarités, des interactions complexes ou des changements de régime que le modèle linéaire ne peut pas capturer. Par conséquent, le vecteur des résidus de la relation (1) peut s'écrire tel que :

$$\epsilon_t^{\text{SBVARX}} = \mathcal{F}_{\text{AI/ML}}(X_t) + \eta_t \quad (7)$$

Nous le modélisons ici suivant trois approches (Random Forest, XGBoost et SVR), confrontées par la suite avec des métriques puis le test de DM (pour [Diebold et Mariano, 1995](#)) pour plus de précisions. Il s'agit des algorithmes XGBoost, Random Forest et SVR.

3.2.1 Le modèle de forêt aléatoire

Il a été proposé par [Ton Kam Ho \(1995\)](#) puis étendu par [Leo Breiman et Adèle Cutler \(2000\)](#) à partir de la technique des arbres de décision, en rendant ceux-ci aléatoires. Son avantage réside dans sa capacité à fournir des prévisions précises à court terme notamment des prix ou de l'inflation, même en environnement complexe d'une part. D'autre part, il limite les problèmes de surajustement par la flexibilité de ses hyperparamètres.

Lorsqu'il est utilisé en régression, la prédiction finale est la moyenne des prédictions de tous les arbres et son principe mathématique repose sur la minimisation de l'erreur quadratique moyenne tel que :

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (f_i - y_i)^2 \quad (8)$$

En régression, Le *Random Forest* agrège les résultats de plusieurs arbres de décision construits sur des sous-échantillons aléatoires des données et des caractéristiques :

$$\mathcal{F}_{\text{RF}}(Z_t) = \frac{1}{M} \sum_{m=1}^M h_m(Z_t) \quad (9)$$

Où h_m est la prédiction du m -ième arbre, et M est le nombre total d'arbres dans la forêt. Le résultat est la moyenne (ou la médiane) des prédictions individuelles, offrant une bonne robustesse et une variance réduite par rapport à un arbre unique.

3.2.2 Le modèle XGBoost (eXtreme Gradient Boosting)

Il est l'un des algorithmes d'apprentissage automatique les plus populaires et les plus performants, particulièrement pour les données tabulaires ou structurées. C'est une mise en œuvre optimisée et distribuée de l'algorithme de Gradient Boosting (ou *renforcement de gradient*). Comme le Random Forest, c'est une méthode d'apprentissage d'ensemble qui combine une série de modèles faibles (généralement des arbres de décision) pour créer un modèle prédictif fort. La différence clé avec le *Random Forest* (qui construit des arbres indépendamment en parallèle), est que XGBoost construit les arbres de manière séquentielle et itérative. Chaque nouvel arbre cherche à corriger les erreurs (résidus) commises par l'ensemble des arbres précédents. Ainsi : (i) une prédiction initiale est faite ; (ii) un premier arbre de décision est construit pour prédire les erreurs (résidus) de la prédiction initiale ; (iii) les prédictions sont mises à jour en combinant le résultat initial et l'erreur prédite

par le nouvel arbre et ; (iv) ce processus se répète, chaque nouvel arbre minimisant une fonction de perte à l'aide de la descente de gradient, jusqu'à ce que les erreurs soient minimisées ou qu'un critère d'arrêt soit atteint.

La fonction-objectif caractérisant ce problème découle d'un arbitrage entre la précision des prévisions (mesurée à travers une fonction de perte) et la complexité du modèle (contrôlée au travers de la régularisation). Elle est donnée par :

$$\text{Obj} = \sum_{i=1}^n L(y_i, \hat{y}_i) + \Omega(f_t) \quad (10)$$

Où :

$L(y_i, \hat{y}_i)$: mesure l'erreur entre la valeur actuel y_i et la valeur prévue \hat{y}_i .

$\Omega(f_t)$: pénalise les arbres trop complexes pour éviter le surapprentissage.

Dans le cas de la régression comme le nôtre, l'on a :

$$L(y_i, \hat{y}_i) = L(y_i, \hat{y}_i)^2 \quad (11)$$

Lors de la construction de l'arbre, XGBoost calcule le gain pour chaque division possible tel que :

$$\text{Gain} = \frac{1}{2} \left(\frac{(\sum_{\text{gauche}} g_i)^2}{\sum_{\text{gauche}} h_i + \lambda} + \frac{(\sum_{\text{droite}} g_i)^2}{\sum_{\text{gauche}} h_i + \lambda} \right) + \gamma \quad (12)$$

Ainsi, une division n'est acceptée que si son gain est positif, assurant qu'elle améliore le modèle au regard de la régularisation. Les poids des feuilles sont calculés tel que :

$$w_j = - \frac{\sum_{i \in j} g_j}{\sum_{i \in j} h_i + \lambda} \quad (13)$$

Ceci montre comment la régularisation L_2 (λ) réduit les poids des feuilles et comment la régularisation L_1 (α) favorise davantage les poids nuls.

3.23 Le modèle de Régression à Vecteurs de Support (SVR)

C'est un algorithme d'apprentissage automatique supervisé utilisé pour les tâches de régression, c'est-à-dire pour prédire des valeurs numériques continues. Contrairement aux méthodes de régression traditionnelles qui minimisent l'erreur entre les valeurs prédites et réelles, la SVR cherche à trouver la ligne ou l'hyperplan optimal qui se situe dans une « marge de tolérance » spécifiée (appelée epsilon) autour des points de données. Les points de données qui se trouvent à l'intérieur de cette marge ne sont pas pris en compte dans le calcul de l'erreur du modèle. Le modèle se concentre sur la minimisation des erreurs uniquement pour les points situés à l'extérieur de cette zone. Les points de données qui se trouvent le plus près de la frontière de la marge, ou à l'extérieur de celle-ci, sont appelés « vecteurs de support ». Ce sont ces points qui déterminent la position et la structure du modèle de régression. La SVR peut être adaptée aux relations linéaires et non linéaires en utilisant différentes fonctions kernel, ce qui la rend très flexible. Sous l'hypothèse de non

linéarités comme postulé, la fonction Kernel transforme les points aberrants pour permettre une transformation linéaire, soit :

$$y = \sum_{i=1}^N (\alpha_i - \alpha_i^*) \cdot \langle \varphi(x_i), \varphi(x) \rangle + b \quad (14)$$

$$y = \sum_{i=1}^N (\alpha_i - \alpha_i^*) \cdot K(x_i, x) + b \quad (15)$$

Où la fonction Kernel peut avoir une forme polynomiale $K(x_i, x) = (x_i; x_j)^d$ ou de base radiale Gaussienne $K(x_i, x) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right)$.

Dans notre cas, La SVR cherche à trouver une fonction de régression $(w \cdot \Phi(Zt) + b)$ qui s'écarte de la cible ($\text{RÉS}_{t+h|t}$) d'une marge ϵ donnée, tout en minimisant la complexité du modèle : $\mathcal{F}_{\text{SVR}}(Z_t) = w \cdot \Phi(Zt) + b$

Où $\Phi(\cdot)$ est une fonction de kernel (e.g., RBF ou Gaussien) qui mappe les données dans un espace de dimension supérieure pour gérer les non-linéarités. La SVR est efficace dans des contextes de données limitées ou de grande dimensionnalité.

3.3 Combinaison des prévisions

La prévision à l'horizon (h) est obtenue comme la somme des prévisions des deux composantes :

$$\hat{Y}_{T+h} = \hat{Y}_{T+h}\{\text{SBVAR}\} + \mathcal{F}_{\text{AI/ML}}(\hat{X}_{T+h}) \quad (16)$$

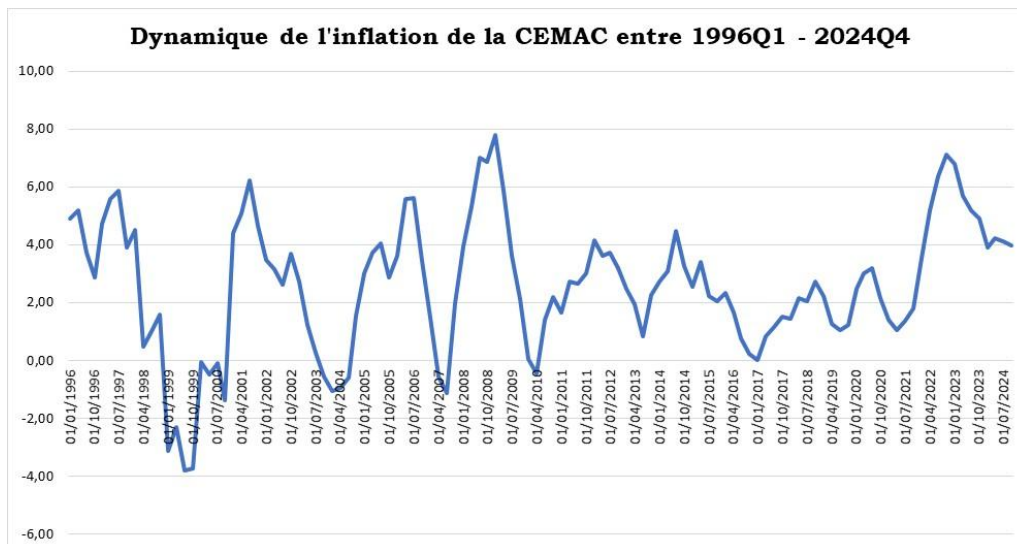
4. Résultats empiriques et discussions

4.1 Données et évaluation

Conformément aux mécanismes de transmission monétaire, nous mettons en relation quatre variable à savoir : (i) le taux d'inflation calculé sur l'indice harmonisé des prix à la consommation (IHPC) ; (ii) l'écart de production obtenu après trimestrialisation et application du filtre HP au PIB annuel ; (iii) le taux de croissance de la masse monétaire et ; (iv) le taux directeur de la banque centrale. Le cours du baril est utilisé comme variable exogène. Ces séries sont obtenues de la BEAC.

L'échantillon porte sur la période 1996Q1-2024Q4. L'évolution de l'inflation est présentée ci-dessous, en prenant d'écarter la zone aberrante :

Graphique 1 : Taux d'inflation de la CEMAC en moyenne annuelle



Source : auteur

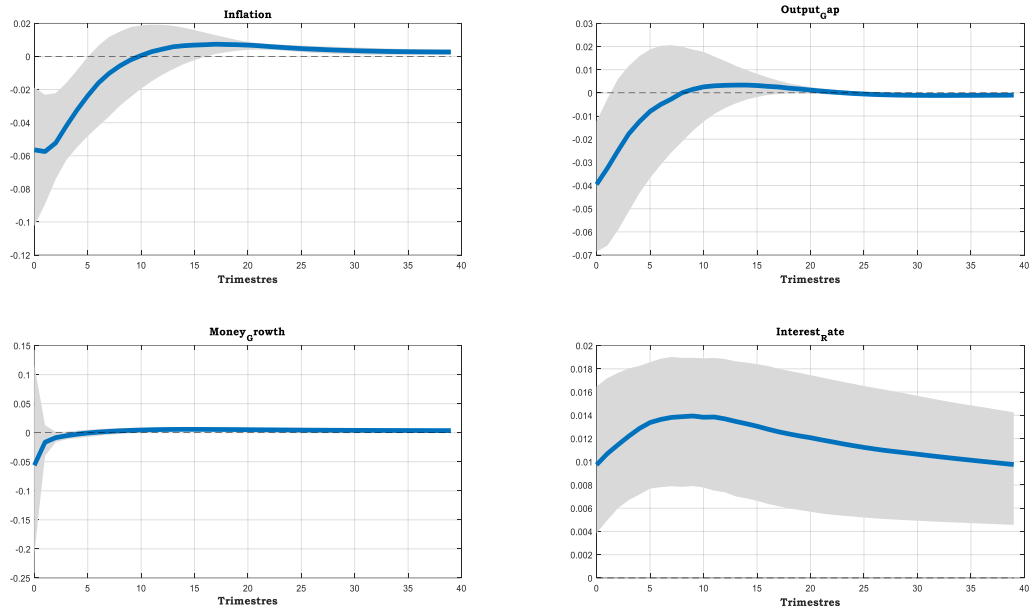
Ce graphique révèle deux périodes qui permettent d'examiner la meilleure adéquation du modèle hybride à savoir 2021Q1 et 2023Q1. Ces périodes correspondent à la crise énergétique de fin 2019, à la période Covid-19 et à la guerre en Ukraine C'est ce qui sera fait dans la suite du document.

Conformément à la méthodologie des méthodes d'apprentissage, notre échantillon a été partitionné à 80% pour l'entraînement et 20% pour les tests/validation. Comme critère d'évaluation, nous utilisons dans un premier temps les métriques ci-après pour comparer les performances hors échantillon d'entraînement : RMSE, MAPE, MAE, et RRMSE. Pour affiner la décision, nous avons recours au test DM qui permet d'examiner l'écart de perte entre les prévisions des modèles hybrides concurrents au SBVAR, sous l'hypothèse nulle d'une perte attendue égale. Son avantage réside dans sa capacité à traiter les erreurs de prévision qui sont autocorrélées et non normalement distribuées, ce qui est fréquent dans les séries temporelles financières.

4.2 Résultats de l'estimation du SBVARX

Les résultats du SBVARX sont examinés ici à partir des fonctions de réponses impulsionnelles liées à un choc de politique monétaire récessif. Ainsi suivant les faits stylisés communément admis par la communauté scientifique et les banques centrales, un choc monétaire récessif réduit le produit et l'inflation, la baisse de cette dernière étant graduelle pour caractériser sa persistance. Les fonctions de réponses impulsionnelles ci-dessous permettent cet examen.

Graphique n°2 : Faits stylisés d'un choc de politique monétaire récessif



L'examen de la figure n°1 ci-dessus montre que les faits stylisés d'un choc monétaire sont dans l'ensemble respectés. Ainsi, l'inflation chute jusqu'à la 10 période, puis remonte légèrement mais, avec une persistance des effets du choc. Parallèlement, le produit chute légèrement. Enfin, le taux directeur responsable de ces évolutions tarde à revenir à son état stationnaire, témoignant du souci de lisser celui-ci par la BEAC.

4.3 Comparaison des performances en prévision sur fenêtres glissantes

Les performances sont comparées suivant que la fenêtre de prévision glissante est fixée à une (tableau 2 et 3), quatre (tableau 4 et 5), puis huit (tableau 6 et 7) périodes, pour être conforme avec l'horizon des prévisions à moyen terme.

Tableau 2 : Indicateurs de performance sur fenêtre sur une période

Modèles	RMSE	MAE	MAPE	RRMSE
SBVARX	1,4598	1,0978	38,706	0,75198
SBVARX-XGBoost	1,3689	1,0582	41,219	0,70518
SBVARX-RF	1,3491	1,0324	40,086	0,69499
SBVARX-SVR	1,3823	1,0415	38,608	0,71209

Tableau 3 : Résultats du test de DM sur une période

Comparaisons	Statistique DM	P-value (5%)	Conclusion
SBVARX/SBVARX+XGBoost	1,2851	0,1987	Aucune différence significative dans la précision des prévisions
SBVARX/SBVAR+RF	1,4803	0,1388	Aucune différence significative dans la précision des prévisions
SBVARX/SBVAR+SVR	2,0490	0,0405	SBVARX significativement MOINS précis que SBVARX+SVR

Le tableau 2 ci-dessus montre que le modèle SBVARX a la métrique la plus élevée en termes de RMSE (1,4598), et de MAE (1,0978). Parallèlement, la meilleure performance entre les

modèles hybrides est à attribuer au modèle SBVARX-RF, avec un RMSE de 1,3491 et un Mae de 1,0324. Toutefois, le test de DM au tableau 3 montre que le modèle hybride SBVAR-RF est significativement plus précis que tous les autres à 5%.

Tableau 4 : Indicateurs de performance sur fenêtre à quatre périodes

Modèles	RMSE	MAE	MAPE	RRMSE
SBVARX	2,9869	2,4562	58,345	1,6270
SBVARX-XGBoost	2,9664	2,4473	58,756	1,6159
SBVARX-RF	2,9582	2,4377	59,304	1,6114
SBVARX-SVR	2,9084	2,3708	56,755	1,5842

Tableau 5 : Résultats du test de DM sur quatre périodes

Comparaisons	Statistique DM	P-value	Conclusion
SBVARX/SBVARX+XGBoost	0,3705	0,7110	Aucune différence significative dans la précision des prévisions
SBVARX/SBVAR+RF	0,4820	0,6298	Aucune différence significative dans la précision des prévisions
SBVARX/SBVAR+SVR	3,4043	0,0007	SBVARX significativement MOINS précis que SBVARX+SVR

Le tableau 4 ci-dessus montre que le modèle SBVARX standard obtient la métrique la plus élevée en termes de RMSE (1,4598), de MAE (1,0978) et de RRMSE (1,6270). Parallèlement, la meilleure performance entre les modèles hybrides est à attribuer au modèle SBVARX-SVR, avec un RMSE de 1,3491, un MAE de 2,3708, un MAPE de 56,755 et un RRMSE de 1,5842. Comme résultante, le test de DM montre que le modèle hybride SBVAR-SVR est significativement plus précis que tous les autres à 5%.

Tableau 6 : Indicateurs de performance à six périodes

Modèles	RMSE	MAE	MAPE	RRMSE
SBVAR	3,6188	3,1711	78,6370	1,9094
SBVAR-XGBoost	3,6297	3,2336	83,1990	1,9152
SBVAR-RF	3,6640	3,2694	84,3550	1,9333
SBVAR-SVR	3,5314	3,1040	78,5490	1,8634

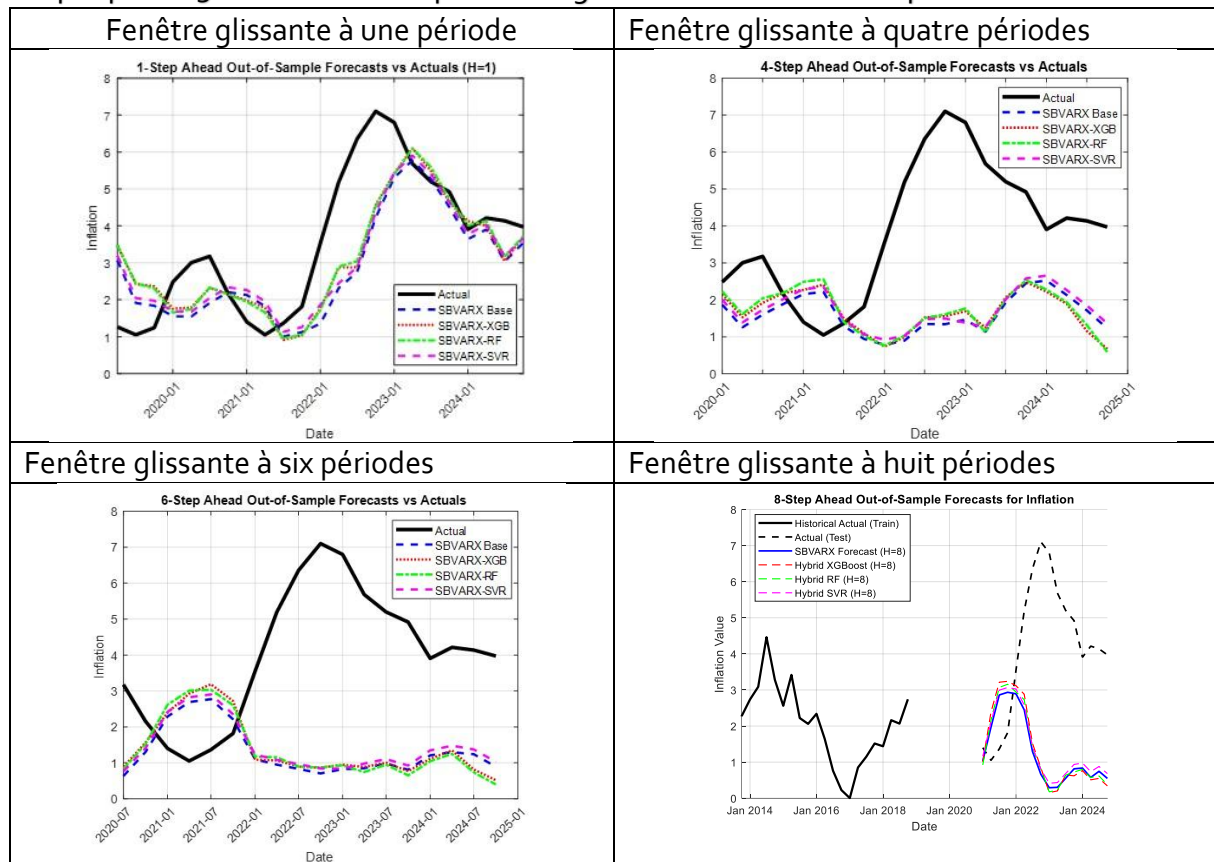
Tableau 7 : Résultats du test de DM sur six périodes

Comparaisons	Statistique DM	P-value	Conclusion
SBVARX/SBVARX+XGBoost	0,2708	0,7865	Aucune différence significative dans la précision des prévisions
SBVARX/SBVARX-RF	0,9574	0,3384	SBVARX significativement MOINS précis que SBVAR+RF
SBVARX/SBVARX-SVR	4,1312	0,0000	SBVARX significativement MOINS précis que SBVAR+SVR

Le tableau 6 ci-dessus montre que le modèle SBVARX standard est moins performant en termes de métrique, particulièrement pour le RMSE (3,618), le MAE (3,171), et le RRMSE (1,8634). Parallèlement, la meilleure performance entre les modèles hybrides est à attribuer au modèle SBVARX-SVR, avec un RMSE de 3,531, un MAE de 3,104, un MAPE de 78,549 et un RRMSE de 1,8634. Comme résultante, le test de DM montre que le modèle hybride SBVAR-SVR est significativement plus précis que tous les autres à 5%.

4.3 Analyse de la performance des modèles durant des périodes de chocs

Graphiques n°3 à 6 : fenêtres de prévisions glissantes sur différentes périodes



Source : auteur

L'histoire de l'inflation dans la CEMAC nous a révélé quelques événements sur lesquels pouvaient porter l'examen de la performance de nos modèles durant des périodes de chocs. Il s'agit de : (i) la crise énergétique de fin 2019 et l'accélération de l'inflation qui a suivi ; (ii) la décélération de l'inflation de la période Covid-19 et ; (iii) l'accélération de l'inflation suite à la guerre en Ukraine. Toutefois, du fait de la division en partie de l'échantillon en partie entraînement (80%) et partie test (20%), les deux premiers événements devant nous servir de repère ont été perdus. Par conséquent, l'analyse va se concentrer sur la guerre russo-ukrainienne. Pour ce faire et compte tenu de l'objectif lié à l'amélioration des prévisions à moyen terme, quatre fenêtres glissantes allant vers cet horizon ont été retenues : une période, deux périodes, quatre périodes et huit périodes.

La première fenêtre glissante est d'une période (H=1) et se prête davantage à une prévision à très court terme. Toutefois, la mise à jour de l'échantillon y étant quasi instantanée, elle devrait nous servir de référence par rapport aux autres fenêtres. Ainsi, il apparaît que l'accélération de l'inflation est bien reprise, avec toutefois un décalage. Pour les autres fenêtres où l'horizon des prévisions tend de plus en plus vers le moyen terme, les mouvements des prévisions avec l'historique de l'inflation sont moins marquées, sinon difficiles à identifier. L'on comprend ici pourquoi malgré la conclusion qui semble apparaître à l'observation des métriques, le test de DM ne retient que la supériorité du modèle SBVARX-SVR par rapport au SBVARX. Il faut souligner ici l'algorithme SVR est sensé performer dans les petits échantillons et donc, qu'avec l'allongement ultérieur de notre échantillon, les conclusions de ce test pour les autres algorithmes pourraient changer.

De tout ce qui précède, il apparaît que les approche hybride SBVARX-IA/ML performant autant, sinon plus que le SBVARX, notamment lorsque la fenêtre de prévision tend vers le moyen terme. Un tel résultat confirme la présence des non-linéarités et des complexités dans la dynamique de l'inflation dans la CEMAC, depuis de l'ère Covid-19.

5. Conclusion et implications de politique monétaire

Cet article s'est intéressé à la possibilité d'améliorer les prévisions de l'inflation à moyen terme dans la CEMAC, par la prise en compte de l'intelligence artificielle et des méthodes d'apprentissage automatique dans le modèle SBVARX actuellement utilisé. Le modèle hybride SBVAR-IA/ML construit a permis de capter aussi bien la dynamique linéaire de l'inflation (SBVARX) que ses dynamiques non-linéaires et complexes, à travers l'entraînement des résidus du SBVARX. Trois approches ont été retenues, du fait de leur adéquation avec la problématique : le Random Forest, le XGBoost, et le SVR. Les précisions prédictives des modèles hybrides ont par la suite comparé par le test DM sur des fenêtres glissantes d'une, deux, six et huit périodes. Sur la base des résultats obtenus, le modèle SBVARX est significativement moins précis que les modèles hybrides (SBVAR-Random Forest, SBVAR-XGBoost et SBVAR-SVR), et l'approche SBVARX-SVR est nettement meilleure. Pour la BEAC dont la politique monétaire est confrontée à de nombreux chocs exogènes depuis l'ère Covid-19, cette approche hybride permettrait d'intégrer les non-linéarités et complexités qui écartent de plus en plus les prévisions à moyen terme de l'inflation de ses réalisations.

Les résultats ci-dessus pourraient toutefois être améliorés avec : (i) l'allongement progressif de la base de données ; (ii) l'utilisation d'autres algorithmes et ; (iii) la disponibilité de données de certaines variables à l'instar du PIB sur des dimensions infra-annuelles. Une piste d'évolution pour la prévision à moyen terme serait l'incorporation des méthodes IA/ML dans les scripts actuels de prévision de l'inflation à moyen terme et l'automatisation de celui-ci.

Références Bibliographiques

1. Akerlof G., Shiller R. (2009) : *Les esprits animaux. Comment les forces psychologiques mènent la finance et l'économie*, Paris, Pearson.
2. Bank for International Settlements (2024): « IA and Text Embeddings for Inflation Monitoring », *BIS Working Paper*.
3. Bank of England (2025): « Blockwise Boosted Inflation Model (BBIM) », *BoE Research Paper*.
4. D'Agostino et al. (2021): « Combining Machine Learning and Econometrics for Inflation forecasting », *IMF WP*.

5. De Nederlandsche Bank (2023): « Nowcasting Inflation with High-Frequency Data and Machine Learning », *DNB Occasional Paper*.
6. Dolgui, A., & Ivanov, D. (2021): « Ripple effect in supply chains: risk propagation, quantification and mitigation », *International Journal of Production Research*.
7. Federal Reserve Bank of St-Louis (2024): « Inflation Forecasting with Large Language Models », *Fed Working Paper*.
8. International Monetary Fund (2024): « AI for Macro Forecasting: Opportunities and Risks », *IMF Policy Paper*.
9. Johan Brannlund, Helen Lao, Maureen MacIsaac, Jing Yang (2023): « Predicting Changes in Canadian Housing Markets with Machine Learning », *Document d'analyse du personnel 2023-21*, Banque du Canada.
10. Lavoie Marc (1985) : « Inflation, chômage et la planification des récessions : la Théorie générale de Keynes et après », *L'Actualité économique*, Volume 61, numéro 2, juin 1985.
11. Lenza M., Primiceri G., & L. Reichlin (2023): "Forecasting Euro Area Inflation with Machine learning Models", *ECB Working Paper Series*.
12. Magnan, A.K., Pörtner, H.O., Duvat, V.K.E. et al. (2021): « Estimating the global risk of anthropogenic climate change. *Nature Climate Change* ». 11, 879–885.
13. Medeiros et al (2022): « Forecasting Inflation with Machine Learning », *Journal of Business & Economic Statistics*.
14. Mehryar Mohri, Afshin Rostamizadeh, Ameet Talwalkar (2012): *Foundations of Machine Learning*, The MIT Press ISBN 9780262018258.
15. Parast, M. M., & Shekarian, M. (2019): « The impact of supply chain disruptions on organizational performance: A literature review », *The International Journal of Logistics Management*.
16. Pichler & Maier (2023): « AI-based Inflation Nowcasting in the Euro Area », *ECB Working Paper*.
17. Stock & Watson (2020): « Macroeconomic forecasting in the Age of Big Data ».
18. Thomas J. Sargent & Neil Wallace (2001): « Some Unpleasant Monetarist Arithmetic », *Federal Reserve Bank of Minneapolis, Quarterly Review* vol. 5, no 3.