



RAPPORT SUR LE NOWCASTING

SY OMAR SAIP
POULIZAC SAMANTHA



Introduction

Dans un monde économique en perpétuelle évolution, disposer de prévisions économiques précises et fiables est essentiel pour décideurs politiques, les entreprises et les investisseurs. Ces prévisions permettent d'anticiper les tendances, de formuler des politiques adaptées et de prendre des décisions stratégiques éclairées. Cependant, les méthodes traditionnelles de prévision, souvent basées sur des données historiques publiées avec un retard significatif, présentent des limitations importantes. C'est dans ce contexte que le nowcasting a émergé comme une technique révolutionnaire.

Le nowcasting, ou prévision instantanée, est une méthode innovante qui permet de produire des estimations économiques en temps réel en utilisant des données disponibles à haute fréquence. Cette approche offre une précision accrue en capturant les dynamiques économiques actuelles avec une latence minimale. Le nowcasting est particulièrement pertinent pour les indicateurs économiques qui sont publiés moins fréquemment, comme le Produit Intérieur Brut (PIB) trimestriel. En combinant des données à haute fréquence et des techniques avancées, le nowcasting permet d'estimer de manière plus immédiate l'état actuel de l'économie.

Nous nous appuyerons sur les données disponibles de la base FRED (Federal Reserve Economic Data), qui fournit des séries chronologiques économiques complètes et fiables. Les données relatives aux dépenses totales en construction en million de dollars, englobant toutes les activités de construction dans le pays, y compris les bâtiments résidentiels, commerciaux, industriels et les infrastructures publiques, seront intégrées dans un modèle de nowcasting pour produire des prévisions en temps réel du PIB en milliard de dollars. Cette approche permettra de démontrer comment les techniques de nowcasting peuvent améliorer la précision des prévisions économiques et fournir des informations opportunes pour la prise de décision.

Les données

Les dépenses totales en construction, communément appelées TTLCONS, représentent l'ensemble des investissements réalisés dans le secteur de la construction aux États-Unis. Ce secteur est un pilier de l'économie américaine, reflétant directement dans la santé économique et les niveaux

d'investissement à la fois privés et publics. Les données sur TTLCONS incluent les sous-secteurs suivants :

- **Constructions Résidentielle** qui comprend la construction de maison individuelles, d'appartements et d'autres types de logements mais aussi ce qui inclut les rénovations, les réparations et les extensions de logements existants.
- **Construction Non Résidentielle** telles que les bâtiments commerciaux qui couvrent les bureaux, les magasins, les hôtels et d'autres infrastructures commerciales. Il y a aussi les structures industrielles qui englobent les usines, les entrepôts et autres infrastructures industrielles. Et aussi les établissements institutionnels qui incluent les écoles, les hôpitaux, les bâtiments gouvernementaux et autres structures institutionnelles.
- **Construction Publique** comprend les infrastructures publiques telles que les routes, les ponts, les tunnels, les aéroports et autres infrastructures financées par le gouvernement mais aussi les projets publics non résidentiels qui incluent les bâtiments publics, les parcs, les installations sportives, etc.

Les données TTLCONS sont publiées mensuellement par le Bureau of the Census des États-Unis, fournissant des informations détaillées et à jour sur les tendances des dépenses dans le secteur de la construction. Une augmentation des dépenses totales en construction est généralement perçue comme un signe positif de croissance économique, tandis qu'une diminution peut indiquer un ralentissement économique.

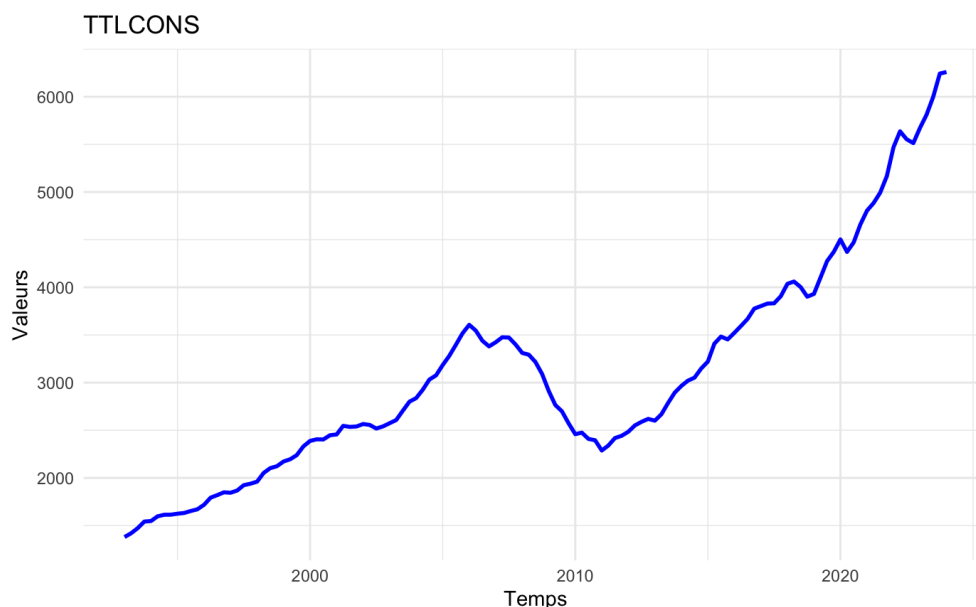


Figure 1 : Visualisation graphique des dépenses totales en construction des États-Unis selon les trimestres

Graphiquement, nous pouvons constater qu'il y'a eu une diminution des dépenses ce qui peut être expliqué par la crise des subprimes, qui a éclaté en 2007 – 2008, a été une période de turbulences économiques majeures, marquée par l'effondrement du marché immobilier et une récession mondiale. Le secteur de la construction a été l'un des plus durement touchés par la crise. En effet, la crise des subprimes a entraîné un effondrement des dépenses en construction résidentielles aux États-Unis, avec une chute significative des nouvelles constructions et des rénovations domiciliaires. Les

investissements dans les bâtiments commerciaux et industriels ont également diminué, tandis que les budgets pour les infrastructures publiques ont été réduits. Les mesures de relance économique ont partiellement atténué l'impact, mais la crise a laissé des séquelles durables sur le secteur de la construction.

Le Produit Intérieur Brut (PIB), ou GDP en anglais, est la valeur totale de tous les biens et services produits dans un pays sur une période donnée, généralement un trimestre ou une année. Le PIB est le principal indicateur de la performance économique d'un pays, mesurant la taille et la santé de son économie. Aux États-Unis, le PIB est calculé et publié par le Bureau of Economic Analysis (BEA) et se compose des éléments suivants :

- **Consommation des Ménages** qui inclue les biens durables, les biens non durables et les services.
- **Investissements** qui incluent les investissements résidentiels, les investissements non résidentiels et les investissements publics.
- **Dépenses Publiques** qui comprend les dépenses du gouvernement fédéral et les dépenses des États et des Collectivités Locales.
- **Commerce Extérieur** qui est les exportations nettes.

Le PIB trimestriel des États-Unis est un indicateur suivi de près par les économistes, les décideurs politiques et les investisseurs, car il offre une vue d'ensemble de l'état de l'économie. Une croissance positive du PIB indique une expansion économique, tandis qu'une contraction du PIB peut signaler une récession.

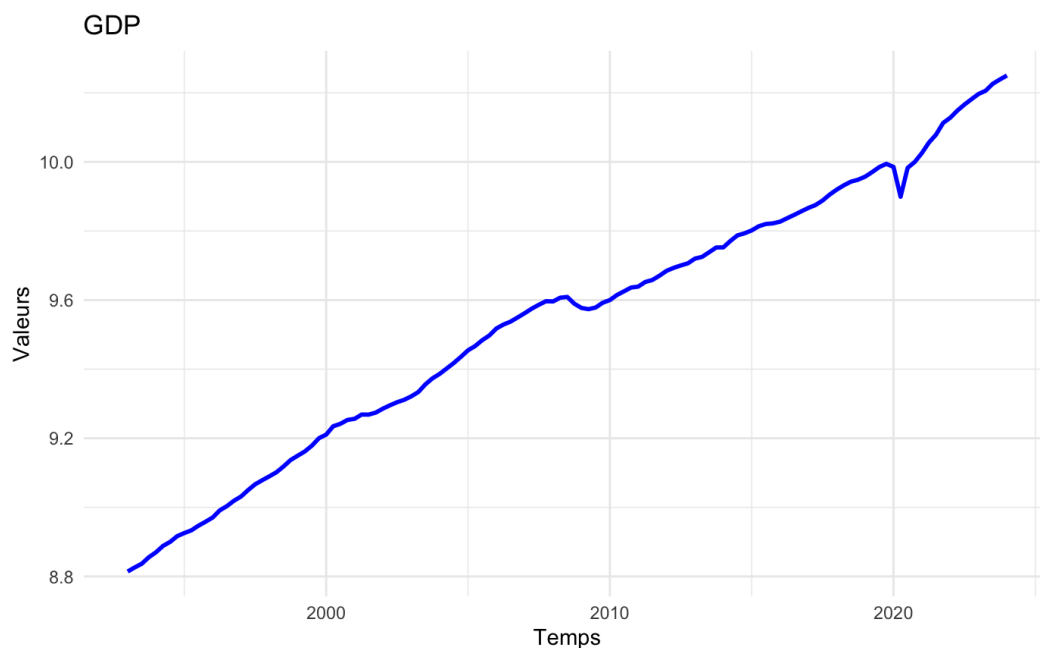


Figure 2 : Visualisation graphique du PIB des États-Unis selon les trimestres

On peut voir un ralentissement du PIB pendant la crise des subprimes de 2008. En effet, la crise des subprimes était principalement centrée sur le marché immobilier et financier, avec des répercussions

sur les institutions financières et le marché du logement. En revanche, la pandémie de COVID-19 a entraîné une perturbation généralisée de l'économie, avec des fermetures d'entreprises, des pertes d'emplois massives et des perturbations de la chaîne d'approvisionnement à l'échelle mondiale. Cela a eu un impact significatif sur la production, la consommation et l'investissement.

Les données sur les dépenses totales en construction (TTLCONS) et le PIB (GDP) sont étroitement liés et essentiels pour l'analyse économique. En effet les dépenses en construction peuvent être un indicateur avancé de la performance économique. Une augmentation des investissements dans la construction peut indiquer une confiance accrue des entreprises et des ménages, signalant une croissance économique future. Les dépenses en constructions contribuent directement au PIB à travers les investissements résidentiels et non résidentiels. Par conséquent les fluctuations des dépenses en construction affectent directement les estimations du PIB. Les décideurs politiques utilisent les données de TTLCONS et du GDP pour formuler des politiques fiscales et monétaires. Une baisse significative des dépenses en construction pourrait inciter le gouvernement à augmenter les dépenses publiques pour stimuler l'économie.

En utilisant les données de TTLCONS pour faire du nowcasting, nous pouvons fournir des prévisions en temps réel du PIB. Ces prévisions peuvent aider à anticiper les tournants économiques et à adapter les stratégies en conséquence. Notre jeu de données couvre la période allant du quatrième trimestre de 1993 au premier trimestre de 2024.

Avant de commencer les différents modèles de prédiction pour le trimestre en cours, nous allons convertir la valeur de TTLCONS de millions dollar en milliard de dollars. Cette conversion garantit la cohérence des unités de mesure, facilitant ainsi la comparaison des résultats et l'interprétation des relations entre les variables, notamment avec le PIB, qui est également exprimé en milliards de dollars.

Dans notre jeu de données, les valeurs de TTLCONS sont exprimées mensuellement. Pour les transformer en données trimestrielles, nous allons les agréger en effectuant la somme des valeurs mensuelles pour chaque trimestre.

Les modèles

Les prévisions WS

Les prévisions WS (Within Sample) font référence aux prédictions faites par un modèle sur les données qu'il a utilisées lors de son entraînement. En d'autres termes, il s'agit des prévisions pour les observations qui étaient déjà présentes dans l'échantillon de données initial sur lequel le modèle a été construit. Ces prévisions permettent d'évaluer la performance du modèle sur les données qu'il connaît déjà.

Le modèle BRIDGE

Le modèle bridge est une technique utilisée en nowcasting pour prédire des variables économiques trimestrielles, comme le PIB, en utilisant des données disponibles à une fréquence plus élevée. Ce modèle établit une liaison entre les variables à haute fréquence et celles à basse fréquence, permettant ainsi de produire des prévisions en temps réel avant la publication des données trimestrielles complètes. Le modèle utilise les informations les plus récentes des indicateurs mensuels pour estimer les valeurs trimestrielles de manière plus précise et rapide.

Pour notre modèle nous créons une variable retard du PIB pour capturer l'inertie et les dynamiques temporelles, améliorer la précision des prédictions et faciliter l'intégration des données haute et basse fréquence, rendant ainsi le modèle plus robuste et fiable. Une variable retard, également appelée variable laggée, est une version décalée dans le temps d'une variable économique ou financière.

Nous allons donc expliquer le PIB par la variable retard et la variable TTLCONS. Les résultats de ce modèle indiquent que l'effet de l'augmentation d'une unité des dépenses totales en construction sur PIB, toutes choses étant égales par ailleurs. Et montre aussi l'effet d'une augmentation d'une unité de la variable retard du PIB sur le PIB actuel. Les résultats de l'analyse révèlent que les variables TTLCONS et la variable retard sont significatives au niveau de 5%. De plus, le modèle dans son ensemble est globalement satisfaisant, car la valeur de la p-valeur du test de Fisher est inférieure à 5%. Avec un coefficient de détermination de 99.8%, il est évident que les variables explicatives capturent efficacement la variance du PIB, renforçant ainsi la robustesse du modèle. Le modèle global est très satisfaisant.

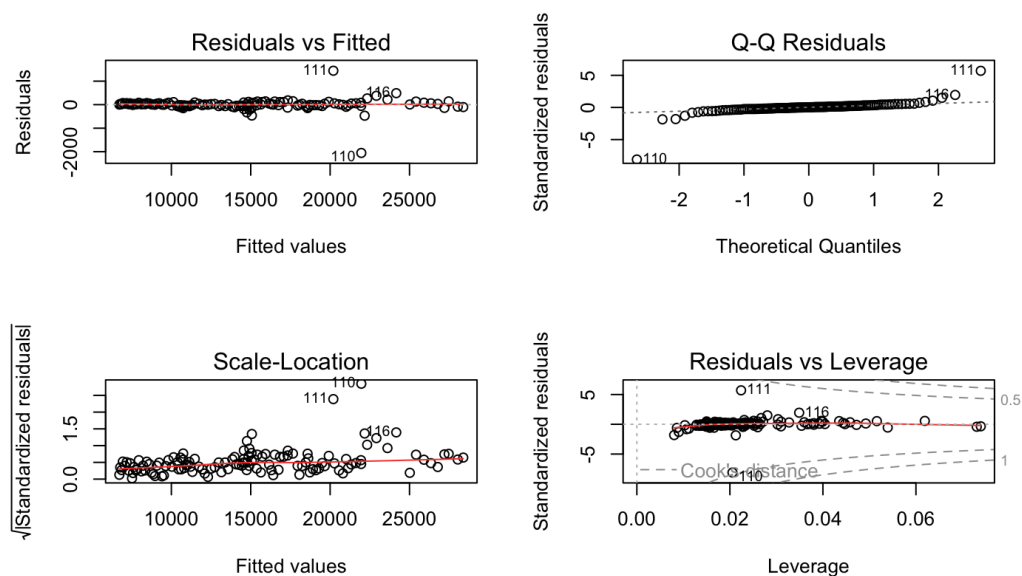


Figure 3 : Graphiques pour évaluer la validité du modèle de régression linéaire, en termes de normalité des résidus, d'hétéroscédasticité et d'influence des observations individuelles

L'examen du QQ-plot indique que les résidus suivent une distribution normale, confirmant ainsi l'adéquation du modèle aux hypothèses de normalité des résidus. Les résidus sont généralement proches de zéro, à l'exception d'une valeur extrême qui peut être considérée comme une observation aberrante. En outre, l'examen de la variance des résidus par rapport aux valeurs ajustées montre une stabilité, ce qui suggère que les résidus sont homoscedastiques.

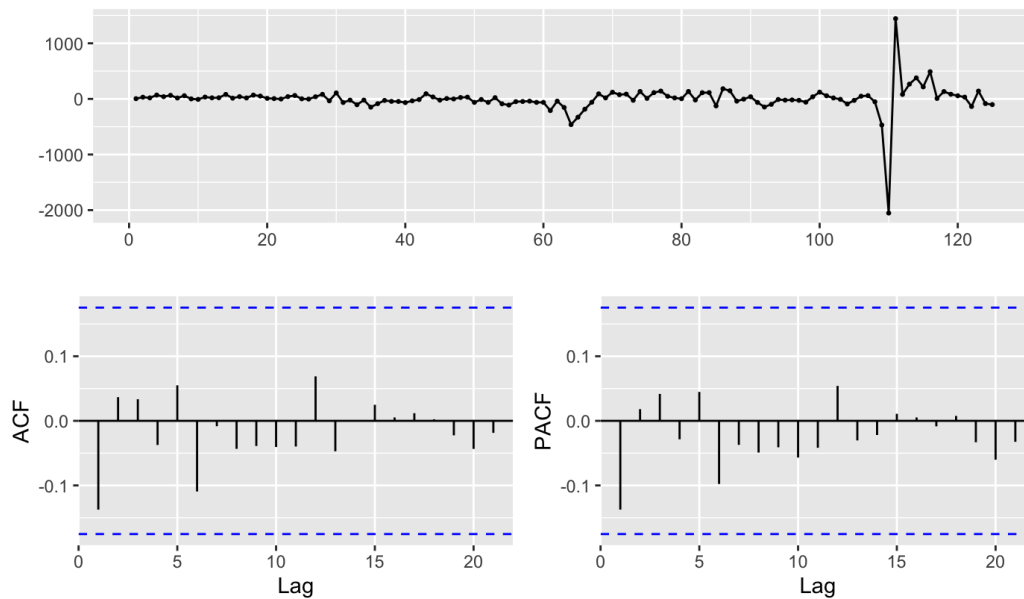


Figure 4 : Graphique d'affichage des résidus pour l'évaluation temporelle de la régression linéaire

De plus, l'absence de corrélation entre les résidus est confirmée, ce qui indique l'absence d'autocorrélation. Ces résultats démontrent que les hypothèses fondamentales de la régression linéaire multiple sont vérifiées et que le modèle est statistiquement fiable.

A partir de ce modèle, nous allons maintenant tracer les prédictions avec leurs intervalles de confiance.

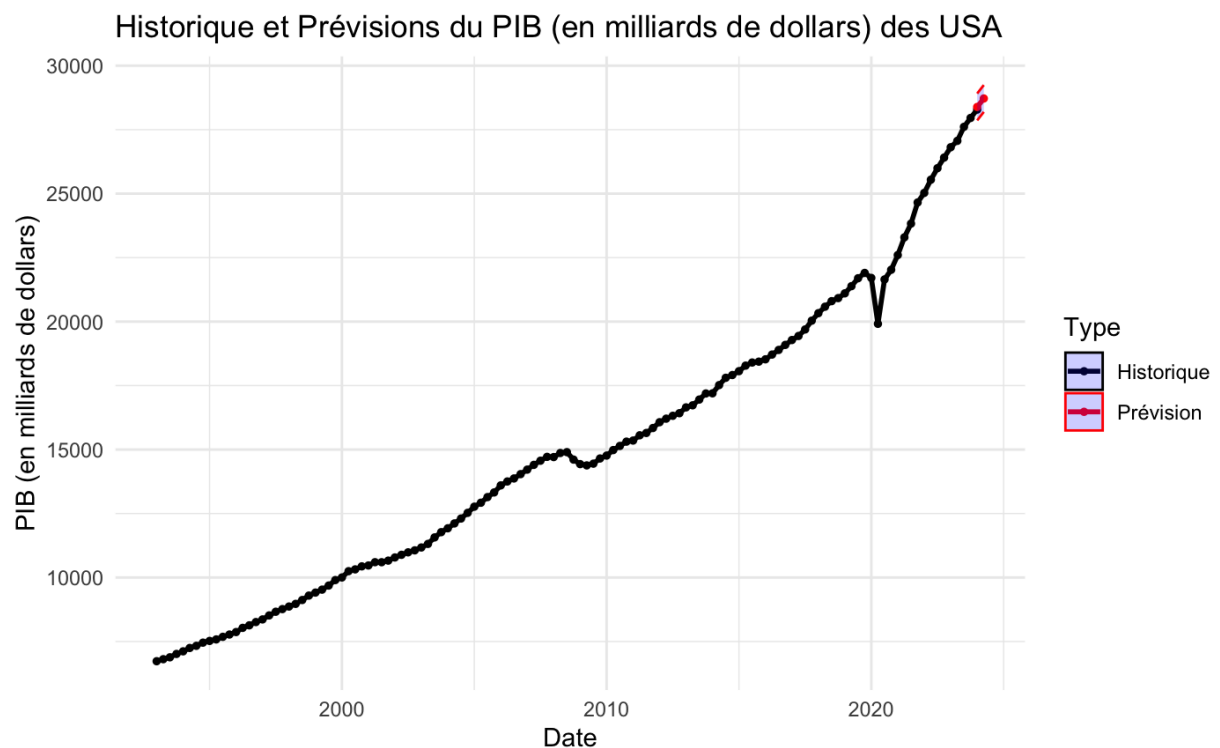


Figure 5 : Graphique des historiques et prévisions du PIB des Etats-Unis

Pour le deuxième trimestre, le PIB prédit est de 28 717,93 milliards de dollars, avec un intervalle de confiance allant de 28 192,53 à 29 243,33 milliards de dollars.

Le modèle MIDAS Zero

Le modèle MIDAS Zéro (Mixed Data Sampling) est une méthode utilisée pour faire des prévisions économiques en combinant des données collectées à différentes fréquences. Le modèle utilise des coefficients fixes pour chaque période de retard, sans prendre en compte les valeurs passées des variables explicatives.

Pour le modèle MIDAS Zéro, nous allons utiliser seulement les données des dépenses totales en construction pour prédire le PIB. Le PIB sera prédit en fonction de 12 retards des données des dépenses totales en construction avec un intervalle de 3 mois entre chaque retard. Pour les résultats du modèle, les coefficients des variables retardées de TTLCONS varient en signe et en magnitude, indiquant des influences mixtes sur le PIB. Aucun des coefficients n'est statistiquement significatif, ce qui suggère que les valeurs passées de TTLCONS, telles qu'elles sont incluses dans ce modèle, n'ont pas de relation claire avec le PIB. D'après ce modèle, les dépenses totales en construction, avec les retards spécifiés, n'apportent pas d'information statistiquement significative pour prédire le PIB.

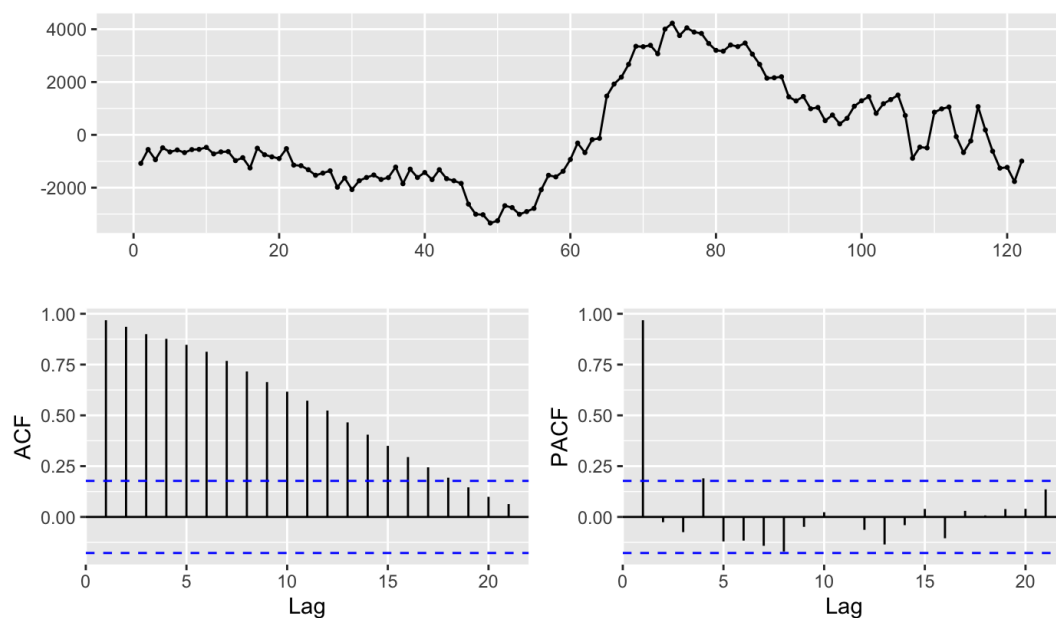


Figure 6 : Graphique d'affichage des résidus pour l'évaluation temporelle du modèle MIDAS zéro

Les résidus du modèle présentent des signes d'hétéroscédasticité et d'autocorrélation. En conséquence la capacité prédictive du modèle laisse à désirer.

Forecasts from MIDAS regression forecast (static)

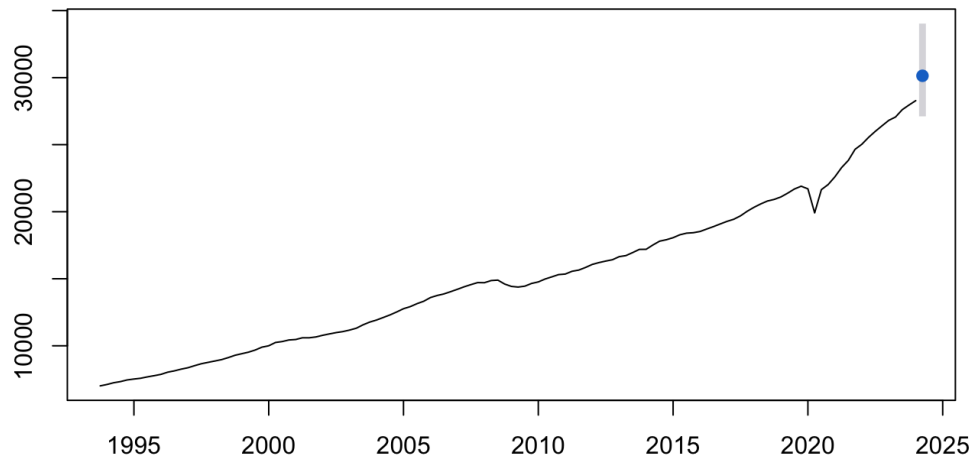


Figure 7 : Graphique des historiques et prévisions du PIB des Etats-Unis à l'aide du modèle MIDAS zéro

Pour le deuxième trimestre de 2024, la prévision ponctuelle du PIB est de 30 137,69 milliards de dollars, avec un intervalle de confiance à 95% allant de 27 112,87 à 34 029,79 milliards de dollars.

Le modèle MIDAS non Zéro

Le modèle MIDAS non-zéro est une méthode statistique sophistiquée qui permet de combiner des données économiques collectées à différentes fréquences pour améliorer la précision des prévisions économiques. Contrairement aux modèles traditionnels, le modèle MIDAS non-zéro prend en compte les valeurs passées des variables explicatives et utilise des coefficients variables pour chaque période de retard.

Pour illustrer l'application du modèle MIDAS non-zéro, nous avons choisi de prévoir le PIB trimestriel en utilisant les dépenses en construction disponibles mensuellement. Nous avons collecté les données pertinentes sur une période donnée pour construire notre modèle et effectuer nos prévisions.

Pour ce dernier modèle, le PIB sera prédit en fonction de 12 retards mensuels des données de TTLCONS avec intervalle de 3 mois entre chaque retard. Les coefficients qui représentent l'effet des valeurs retardées de TTLCONS sur le PIB varient en signe et en magnitude. Aucun de coefficients n'est statistiquement significatif, ce qui indique que les valeurs passées de TTLCONS n'ont pas de relation claire et significative avec le PIB dans ce modèle. Ce modèle n'apporte pas de preuves suffisantes pour établir une relation significative entre les dépenses en construction et le PIB. Les coefficients ne sont pas significatifs et la variabilité des résidus est relativement élevée.

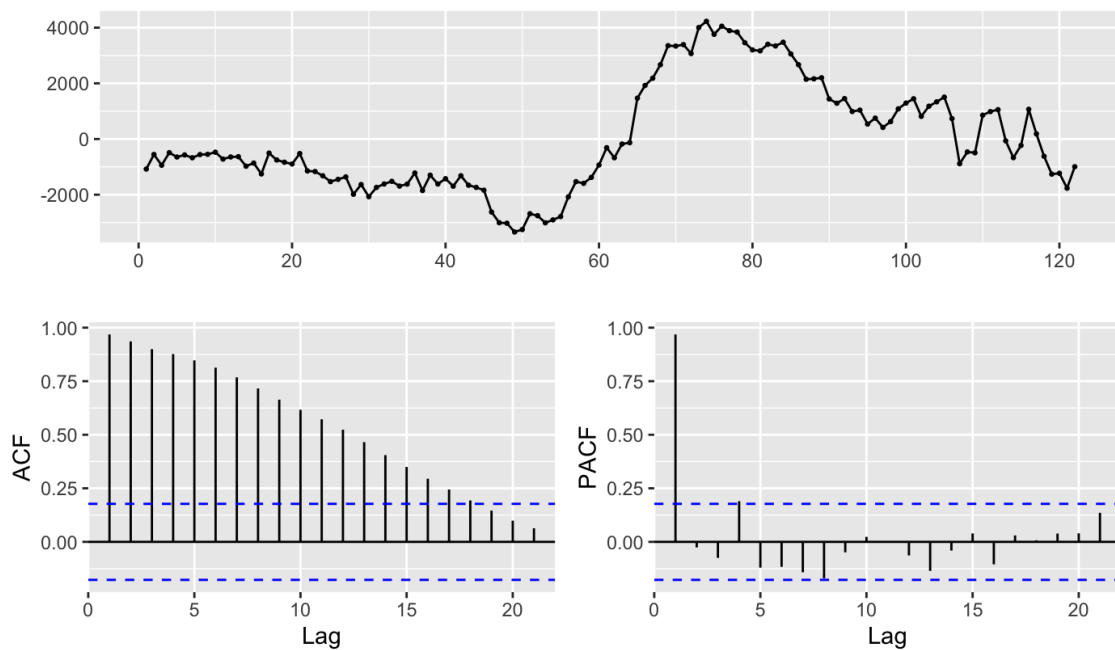


Figure 8 : Graphique d'affichage des résidus pour l'évaluation temporelle du modèle MIDAS non-zéro

Les résidus présentent des signes d'hétéroscédasticité et d'autocorrélation, ce qui indique une variabilité non constante des erreurs de prédiction et une corrélation entre les résidus successifs. En conséquence, la capacité prédictive du modèle laisse à désirer, ce qui signifie qu'il ne parvient pas à fournir des prévisions précises et fiables.

Forecasts from MIDAS regression forecast (static)

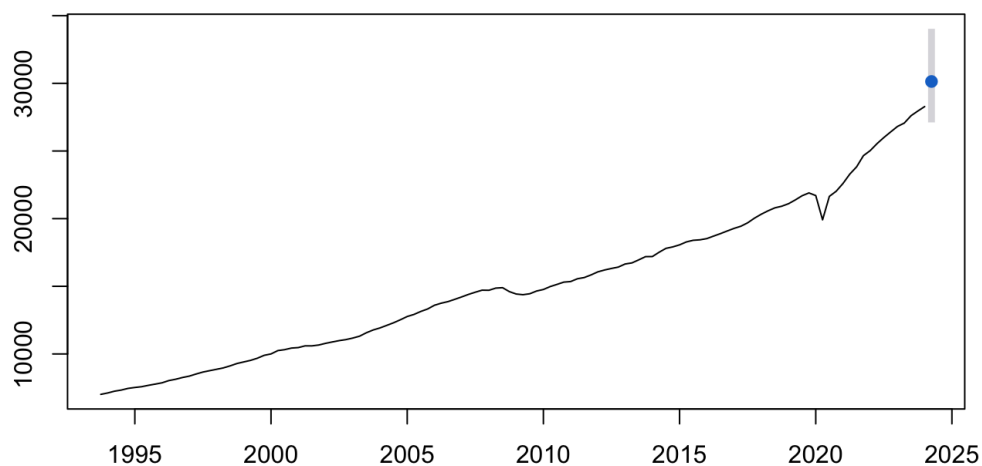


Figure 9 : Graphique des historiques et prévisions du PIB des Etats-Unis à l'aide du modèle MIDAS non-zéro

Pour ce deuxième trimestre de 2024, la prévision ponctuelle du modèle est de 30 137,69 milliards de dollars. L'intervalle de confiance à 95% s'étend de 27 112,87 à 34 029,79 milliards de dollars.

Les prévisions OOS

Les prévisions OOS (Out-of-Sample) se réfèrent à des prévisions réalisées en utilisant des données qui n'ont pas été utilisées pour estimer le modèle. Elles sont essentielles pour tester la capacité de généralisation et la robustesse d'un modèle prédictif. En évaluant les performances du modèle sur des données externes (hors échantillon), on peut obtenir une mesure plus fiable de son efficacité dans des conditions réelles.

Le modèle BRIDGE

Pour le modèle BRIDGE, nous avons vu auparavant qu'il fallait créer une variable retard du PIB. Nous allons expliquer le PIB par le PIB retardé et les dépenses totales en construction. Les résultats ont montré que pour cette prévisions OOS avec le modèle BRIDGE, nous avons un retard du PIB qui est la seule variable significative au niveau de 5%. TTLCONS n'est pas significatif alors qu'elle l'était dans la prévision WS. En ce qui concerne la qualité du modèle, le R^2 est de 99,95%, donc le modèle est bien ajusté. Le test de Fisher est significatif au niveau de 5%, d'où le modèle est globalement satisfaisant. Les résidus sont distribués selon une loi normale, mais des valeurs aberrantes sont observées aux extrémités du graphique quantile-quantile. Il semble que les résidus présentent une homoscedasticité car la variabilité du nuage de points des résidus par rapport aux valeurs réelles reste constante, puis augmente avant de se stabiliser à nouveau. Aucune autocorrélation n'est détectée dans les résidus.

A partir de ce modèle, nous allons pouvoir effectuer les prédictions suivantes :

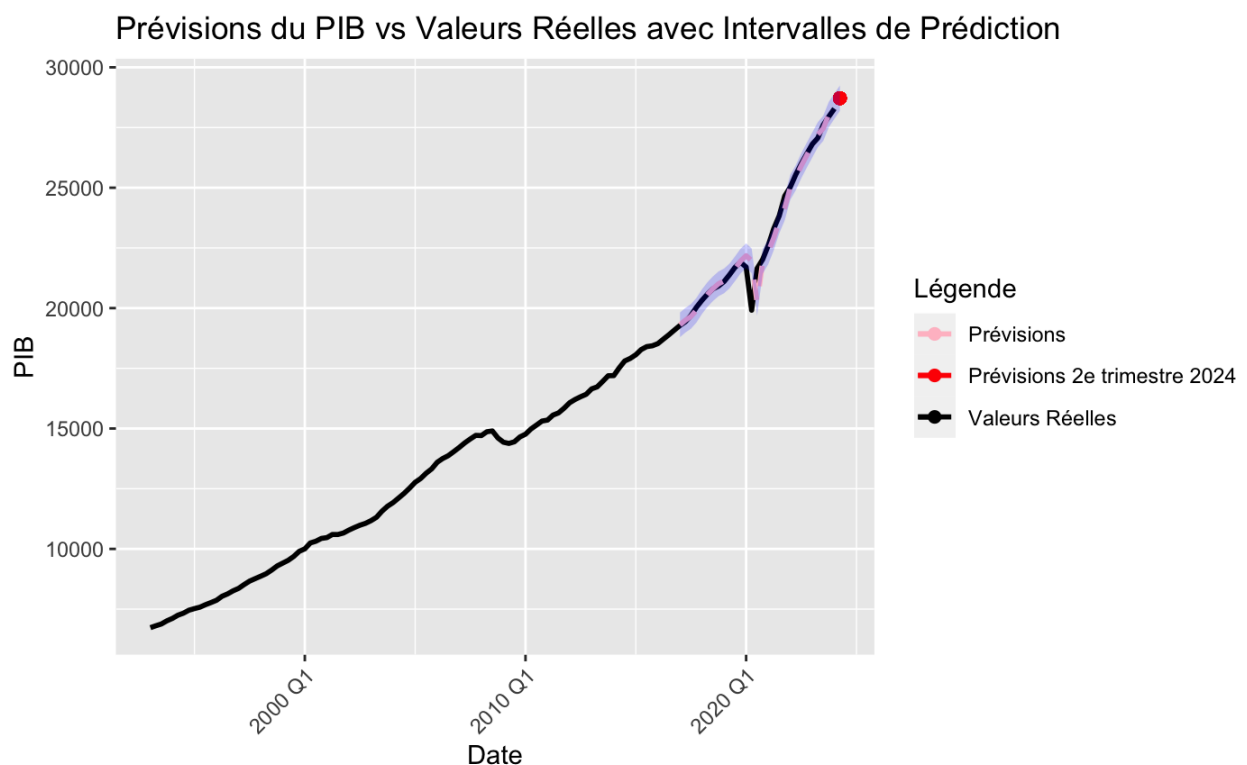


Figure 10 : Graphique des historiques et prévisions du PIB des Etats-Unis du Modèle BRIDGE OOS

Sur l'intégralité de la période de test, le modèle démontre une capacité précise à prédire le PIB, même pendant la chute survenue durant la pandémie de Covid-19 en 2020. Pour le deuxième trimestre de 2024, il prédit un PIB de 28 717,93 milliards de dollars, avec un intervalle de confiance allant de 28 192,53 jusqu'à 29 243,33 milliards de dollars.

Le modèle MIDAS Zero

Maintenant, nous allons procéder au modèle MIDAS Zero OOS. Les résultats du modèle présentent les mêmes résultats que le modèle sous WS, nous n'avons aucun coefficient qui n'est significatif au niveau de 5%. Et les résidus continuent de présenter des signes d'hétéroscédasticité et d'autocorrélation.

Cependant, pour les prévisions il y'a des différences.

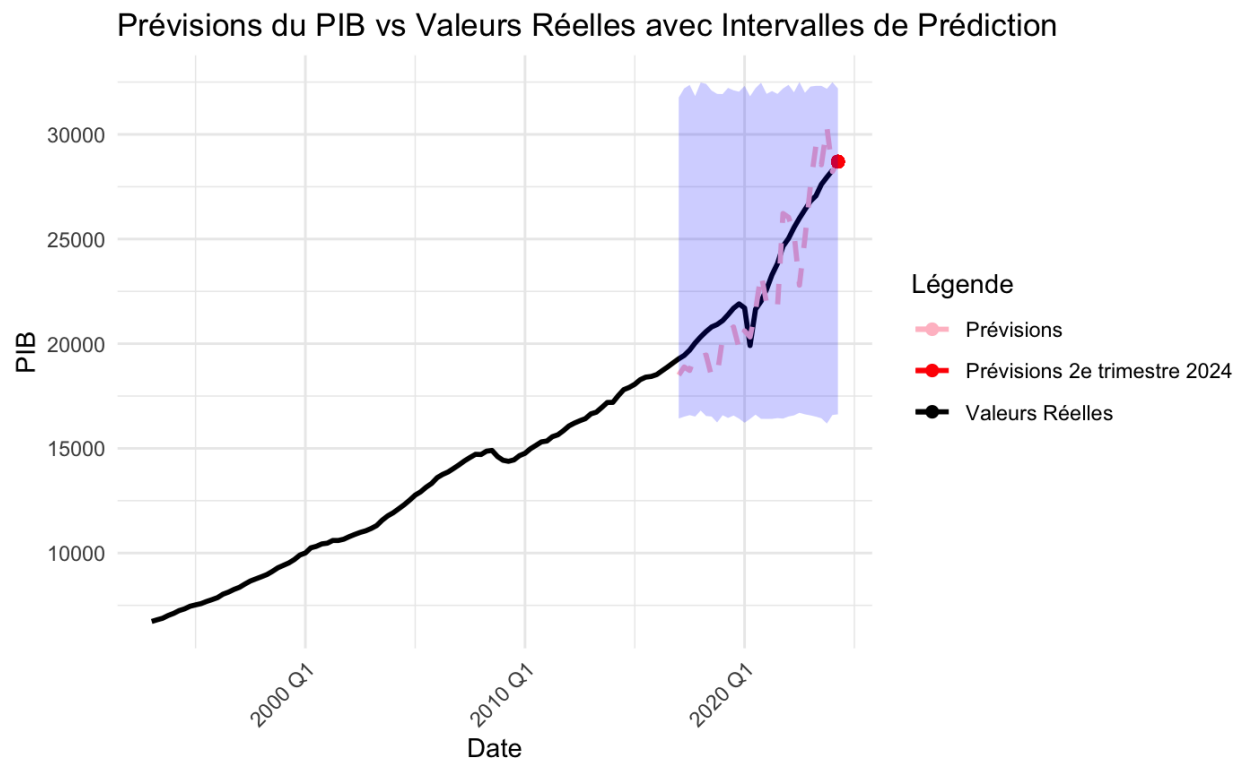


Figure 11 : Graphique des historiques et prévisions OOS du PIB des Etats-Unis à l'aide du modèle MIDAS zéro

Il est évident qu'il existe une disparité significative entre les valeurs réelles et les prévisions pour prévisions OOS du modèle MIDAS zéro, ce qui est insatisfaisant. La prédiction est de 28 696,91 milliards de dollars avec un intervalle de prédiction allant de 16 416,86 à 32 324,22 milliards de dollars, ce qui est beaucoup plus large que celui du modèle BRIDGE.

Le modèle MIDAS non Zéro

Pour la prévision OOS à l'aide du modèle MIDAS non Zero, nous constatons que nous obtenons les mêmes conclusions que le modèle MIDAS Zéro, cela indique que les données utilisées pour le modèle peuvent être assez informatives pour orienter les coefficients du modèle vers des valeurs similaires, indépendamment de la valeur initiale. Ici encore, nous avons les résidus qui sont hétéroscédastiques et autocorrélés. Pour la prédiction, nous avons les résultats suivants :

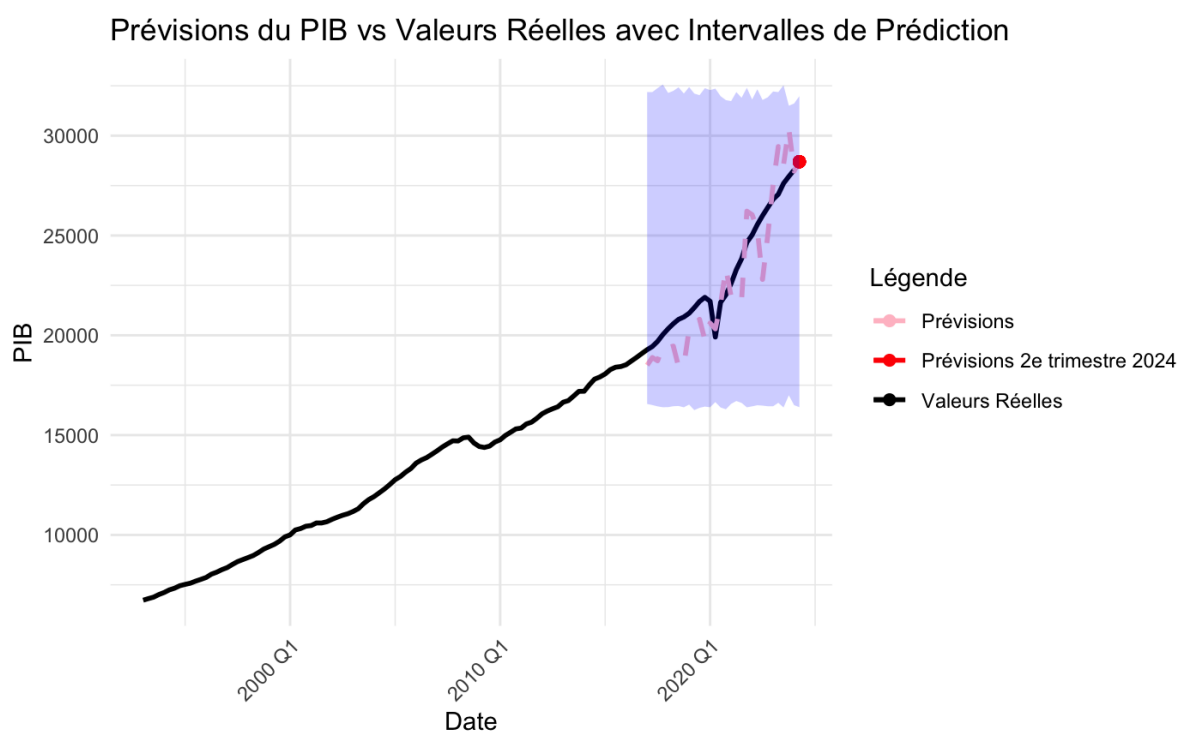


Figure 12 : Graphique des historiques et prévisions OOS du PIB des Etats-Unis à l'aide du modèle MIDAS non-zéro

La prédiction du modèle est de 38 696,91 milliards de dollars avec un intervalle de prédiction allant de 16 477,59 à 32 424,31 milliards de dollars.

Conclusion

	Modèle	MAE	RMSE
Within Sample	Bridge	103.34	252.87
Within Sample	MIDAS zéro	1 631.65	1 938.93
Within Sample	MIDAS non-zéro	1 631.65	1 938.923
Out Of Sample	Bridge	233.61	496 .57
Out Of Sample	MIDAS zéro	1 229.64	1 459.72
Out Of Sample	MIDAS non-zéro	1 229.64	1 459.72

Tableau 1 : Performance des Modèles Nowcasting : Évaluation In-Sample et Out-of-Sample

Le modèle bridge affiche une erreur absolu moyenne (MAE) de 103,34 et une racine de l'erreur quadratique moyenne (RMSE) de 252,87. En comparaison, le modèle MIDAS zéro et le modèle MIDAS non zéro ont des MAE et des RMSE plus élevées. Ainsi, le modèle bridge présente les meilleures métriques, démontrant une meilleure capacité prédictive. De plus, les hypothèses statistiques relatives aux résidus sont vérifiées pour le modèle bridge, renforçant sa fiabilité. En outre, l'écart-type du PIB est de 5 623 milliards de dollars, tandis que le RSME du modèle bridge est de 252,87, soulignant la précision relative de ce modèle par rapport à la variabilité des données sous-jacentes pour les prévisions WS. En ce qui concerne les performances hors échantillon, le modèle Bridge affiche une MAE de 233,61 et une RMSE de 496,57. Comparativement, les modèles MIDAS Zéro et MIDAS Non-Zéro présentent des valeurs de MAE et RMSE plus élevées. Ainsi, le modèle Bridge maintient sa supériorité même lorsqu'il est évalué sur des données hors échantillon. Pour le modèle Bridge, les erreurs augmentent de manière modérée en passant de l'évaluation in-sample à out-of-sample, ce qui indique une bonne capacité de généralisation. En revanche, pour les modèles MIDAS, les erreurs

diminuent en out-of-sample, ce qui est inhabituel et pourrait suggérer des anomalies dans les données ou les processus de validation croisée.

Dans cette étude, nous avons examiné et comparé trois modèles de nowcasting pour la prévision du produit intérieur brut (PIB) : le modèle Bridge, le modèle MIDAS Zéro et le modèle MIDAS Non-Zéro. Le nowcasting, ou prévision à court terme, est crucial pour les décideurs politiques et les acteurs du marché, car il fournit des estimations en temps réel de l'évolution économique, permettant ainsi de prendre des décisions informées.

Le modèle Bridge se démarque par sa capacité à intégrer efficacement les variables retardées et les dépenses totales en construction pour expliquer le PIB. Dans nos prévisions dans l'échantillon (WS), le modèle Bridge a démontré une grande précision, avec un MAE de 103,34 et un RMSE de 252,87. Cependant, cette précision élevée pourrait être le résultat d'un surapprentissage, où le modèle s'adapte trop étroitement aux données d'entraînement, risquant ainsi de ne pas généraliser efficacement aux nouvelles données.

Le modèle MIDAS Zéro, bien qu'il offre une implémentation simple avec des coefficients fixes pour chaque période de retard, a montré des performances inférieures aux attentes. Avec un MAE de 1 631,65 et un RMSE de 1 938,93 dans les prévisions WS, ce modèle a été moins précis que le modèle Bridge. De plus, aucun coefficient n'a été statistiquement significatif, remettant en question la validité des prédictions. Les problèmes persistants d'hétéroscédasticité et d'autocorrélation des résidus ont également affecté la fiabilité des prévisions.

Le modèle MIDAS Non-Zéro, malgré son utilisation des valeurs passées des variables explicatives pour améliorer la précision, a montré des performances similaires à celles du modèle MIDAS Zéro. Les prévisions WS ont présenté des MAE et des RMSE élevés, et aucun coefficient n'a été statistiquement significatif. Les problèmes de résidus ont également persisté, mettant en doute la validité des prévisions.

Le modèle Bridge se démarque comme le meilleur choix en raison de ses performances supérieures, démontrant ainsi une capacité prédictive plus efficace, tant pour les prévisions dans l'échantillon (WS) que pour celles hors échantillon (OOS). De plus, les tests statistiques appliqués aux résidus confirment leur conformité aux hypothèses requises, renforçant ainsi la fiabilité et la validité du modèle. En conclusion, le modèle Bridge se démarque comme étant le plus précis et fiable pour les tâches de nowcasting. Bien que les modèles MIDAS offrent une alternative intéressante, ils nécessitent des améliorations substantielles pour rivaliser avec le modèle Bridge. Pour garantir des prévisions économiques précises et fiables, il est essentiel de continuer à développer et à affiner les techniques de nowcasting, en tenant compte des défis et des limitations de chaque modèle.

Annexes

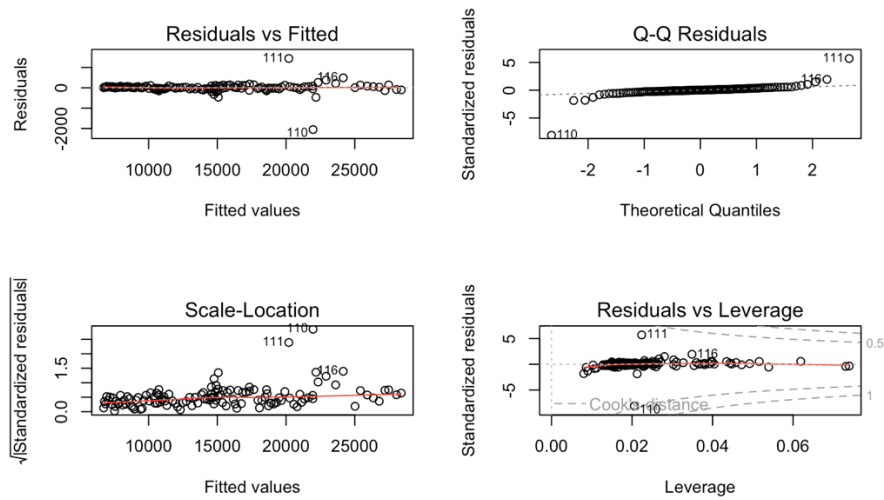


Figure 13 : Analyse des résidus du modèle BRIDGE (OOS)

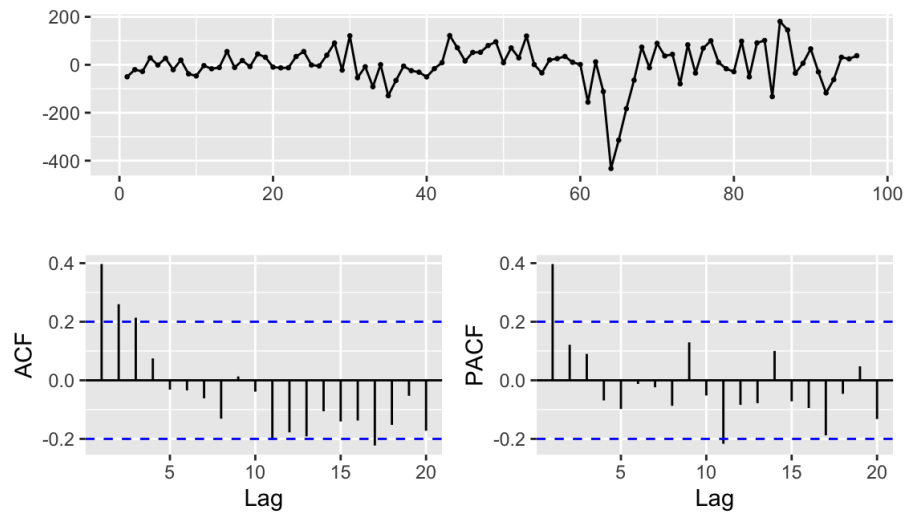


Figure 14 : Analyse autocorrélation du modèle BRIDGE (OOS)

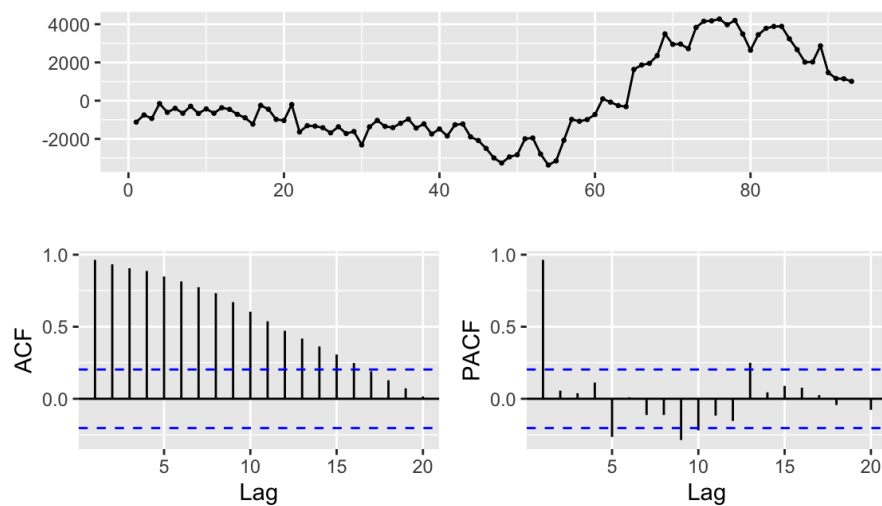


Figure 15 : Analyse autocorrélation du modèle MIDAS Zéro (OOS) et modèle MIDAS non-Zéro (OOS)