

Analyse de la Production d'Électricité d'Origine Nucléaire aux États-Unis (2001-2024)

Sommaire

Introduction.....	
I. Analyse Descriptive et Tendance de la Série.....	
Premières observations de la série chronologique.....	1.1
Estimation et analyse de la tendance à long terme.....	1.2
II. Analyse de la Saisonnalité et Décomposition du Modèle.....	
Décomposition de la série selon un modèle additif.....	2.1
Analyse des coefficients saisonniers.....	2.2
Analyse de la série corrigée et des résidus.....	2.3
III. Prévisions et Évaluation des Modèles.....	
Méthodologie de prévision.....	3.1
Validation des modèles sur l'année 2024.....	3.2
Prévisions pour l'année 2025.....	3.3
Conclusion.....	
Summary.....	

Introduction

L'Energy Information Administration (EIA) est une agence fédérale américaine dont la mission est de collecter, analyser et diffuser des informations sur l'énergie. Ces données sont cruciales pour l'élaboration de politiques publiques, l'efficience des marchés et la compréhension des interactions entre l'énergie, l'économie et l'environnement.

Dans le cadre de cette étude, nous nous intéressons à l'une des séries temporelles publiées par l'EIA : la **production mensuelle d'électricité d'origine nucléaire aux États-Unis**, mesurée en milliers de mégawattheures. La période d'observation s'étend de janvier 2001 à décembre 2024, nous fournissant 24 années complètes de données.

L'objectif de ce rapport est de réaliser une analyse statistique complète de cette série. Dans un premier temps, nous effectuerons une analyse descriptive pour identifier ses principales composantes, notamment sa **tendance à long terme et ses variations saisonnières**. Pour ce faire, nous utiliserons des méthodes de lissage par moyennes mobiles et de régression.

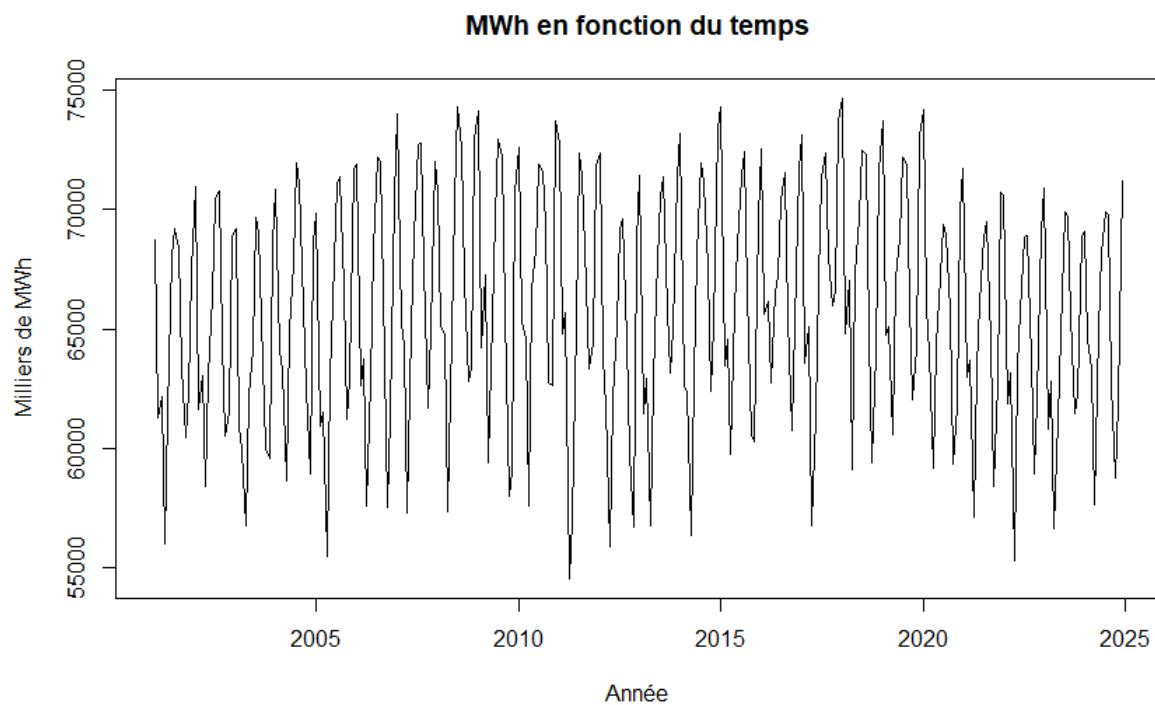
Dans un second temps, nous procéderons à la **décomposition** de la série afin d'isoler et d'analyser ses différentes composantes. L'étude des résidus nous permettra d'évaluer la qualité d'ajustement de notre modèle.

Enfin, la partie principale de notre mission consistera à **prévoir la production d'électricité pour l'année 2025**. Nous mettrons en œuvre trois approches distinctes : un modèle de régression sur la tendance désaisonnalisée, la méthode de lissage exponentiel de Holt-Winters, et un modèle économétrique SARIMA. Afin de déterminer la méthode la plus fiable, nous évaluerons leur performance respective en calculant leur erreur quadratique moyenne sur les données de l'année 2024.

I. Analyse Descriptive et Tendance de la Série

1.1. Premières observations de la série chronologique

L'analyse d'une série temporelle débute systématiquement par sa représentation graphique. Celle-ci permet de déceler visuellement les structures fondamentales qui la composent. Le graphique ci-dessous représente la production mensuelle d'électricité nucléaire aux États-Unis de 2001 à 2024.



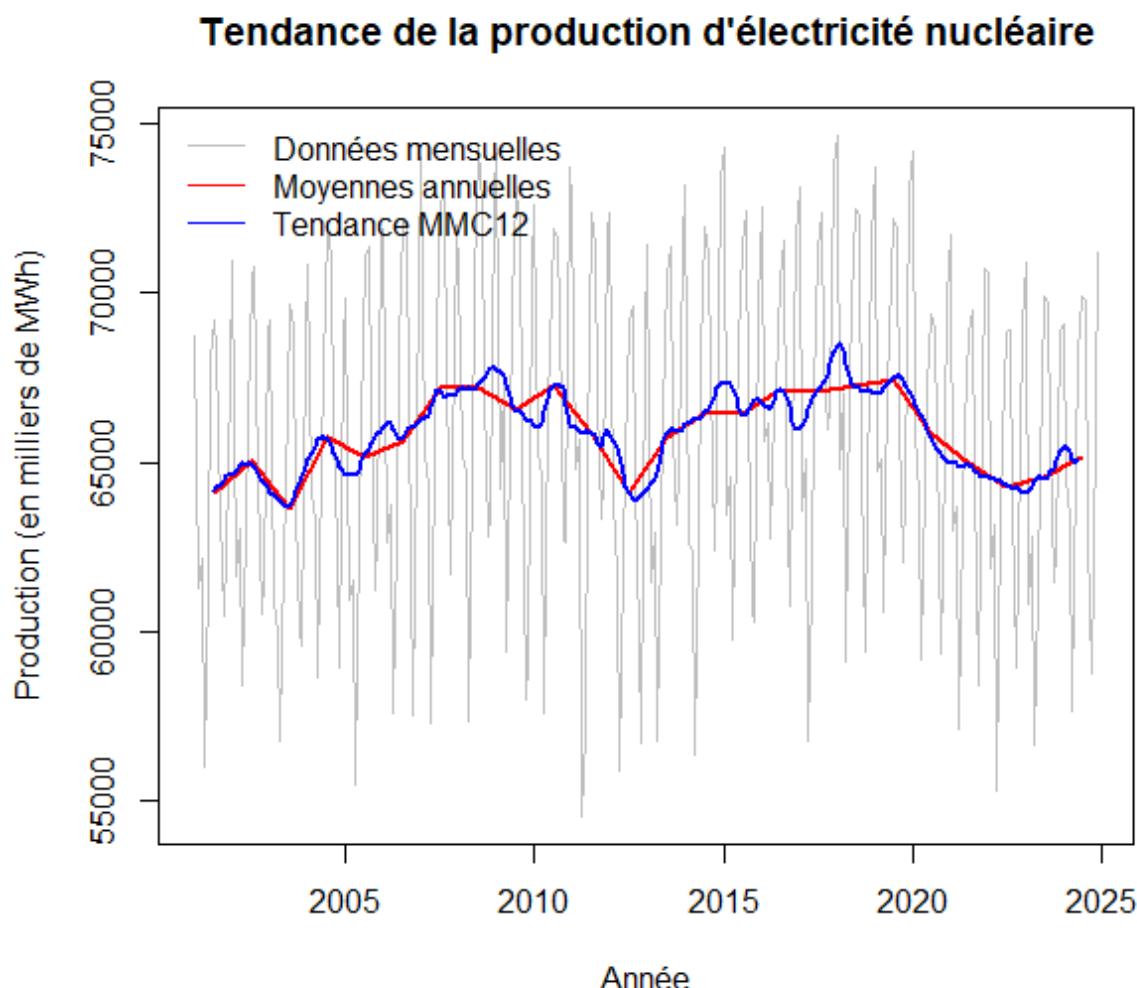
À première vue, la série ne présente pas une simple croissance ou décroissance monotone. Son évolution à long terme semble complexe. De plus, on distingue clairement des **oscillations régulières** qui se répètent chaque année. Ces fluctuations périodiques sont le signe d'une forte **composante saisonnière**, que nous analyserons plus en détail dans la deuxième partie.

1.2. Estimation et analyse de la tendance à long terme

Pour isoler la tendance de fond des variations saisonnières, nous avons recours à des techniques de lissage. Conformément à la mission, nous avons employé deux méthodes :

1. **La courbe des moyennes annuelles** : une méthode simple qui consiste à calculer la moyenne de production pour chaque année.
2. **La moyenne mobile centrée d'ordre 12 (MMC12)** : un filtre qui lisse les données en calculant la moyenne sur une fenêtre glissante de 12 mois. Cette méthode est particulièrement adaptée aux données mensuelles car elle neutralise les effets saisonniers.

Le graphique suivant superpose ces deux estimations de la tendance à la série originale.



L'analyse de ce graphique nous apprend plusieurs choses :

- La **MMC12 (en bleu)** offre une vision plus fine et réactive de la tendance que la courbe des moyennes annuelles (en rouge). Elle suit les évolutions de la série tout en gommant efficacement les variations saisonnières.
- L'observation de la tendance révèle trois phases distinctes :
 - **2001 - 2010** : Une période de **stabilité à un niveau de production élevé**, avec une légère croissance culminant autour de 2010.
 - **2011 - 2018** : Une phase de **déclin progressif** de la production.
 - **2019 - 2024** : Une période de **relative stabilisation** à un niveau inférieur à celui de la décennie précédente.

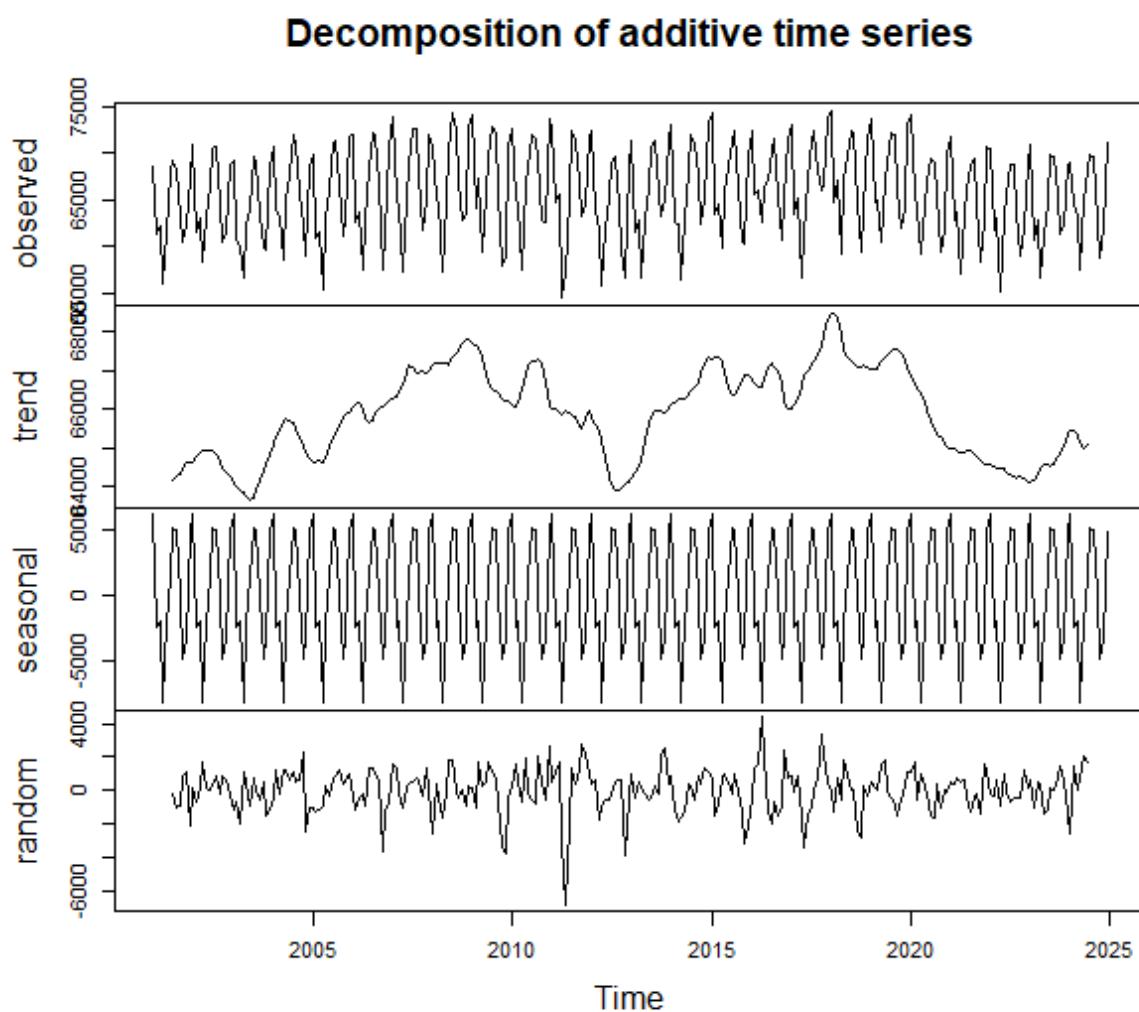
En conclusion, la tendance de la production nucléaire américaine n'est pas linéaire. Elle reflète des changements structurels sur le long terme, potentiellement liés à des politiques énergétiques, à la mise en service ou à l'arrêt de centrales, ou à la concurrence avec d'autres sources d'énergie.

II. Analyse de la Saisonnalité et Décomposition du Modèle

2.1. Décomposition de la série selon un modèle additif

L'analyse visuelle ayant montré des oscillations d'amplitude relativement constante au fil du temps, nous avons choisi un **modèle de décomposition additif**. Ce modèle postule que la valeur observée (y_i) est la somme de trois composantes : la tendance (f_i), la saisonnalité (s_i) et un résidu aléatoire (e_i). $y_i = f_i + s_i + e_i$

Nous avons utilisé la fonction `decompose()` de R pour extraire ces trois composantes de la série brute.



2.2. Analyse des coefficients saisonniers

Le panneau "seasonal" du graphique de décomposition met en évidence un schéma saisonnier très marqué et stable dans le temps. L'analyse des coefficients saisonniers nous permet de quantifier ce phénomène :

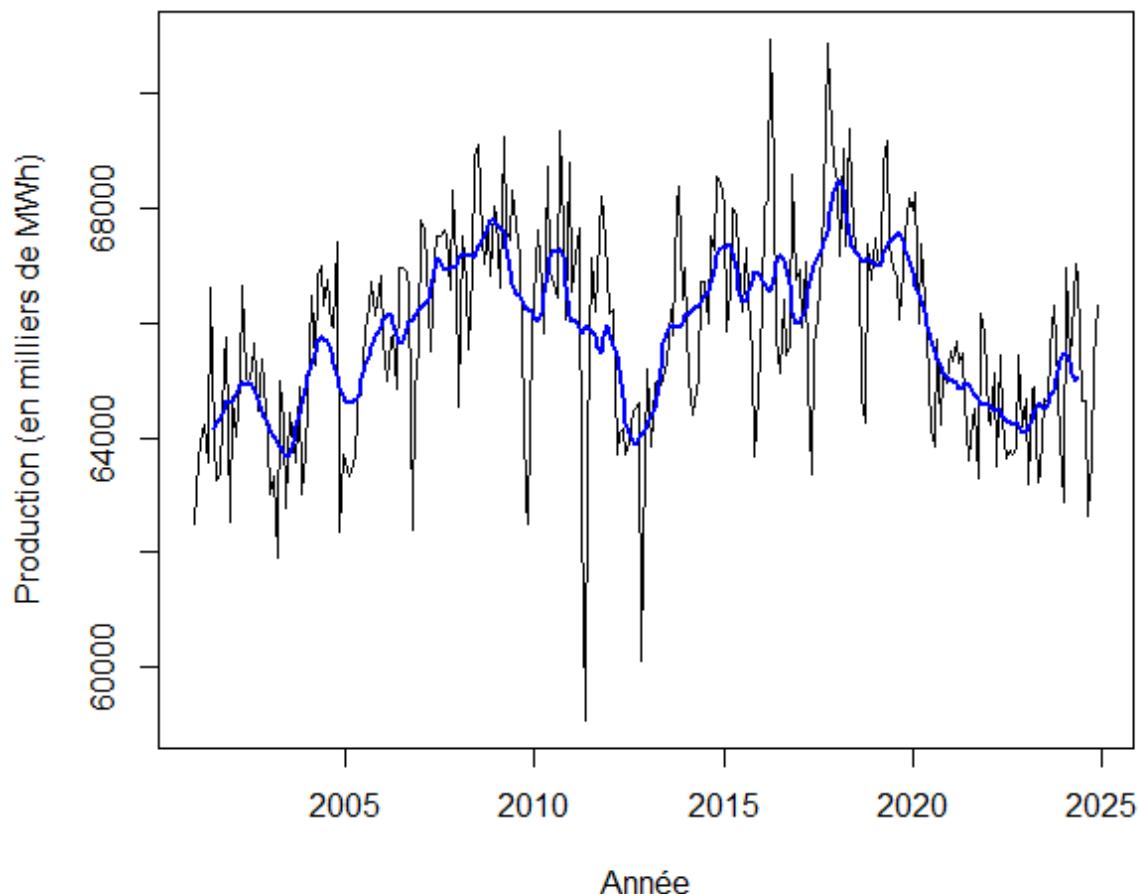
- **Pics de production** : On observe deux pics de production chaque année. Le plus important a lieu en **janvier**, et le second, légèrement moins prononcé, en **juillet et août**. Ces pics coïncident avec les périodes de forte demande en électricité, pour le chauffage en hiver et la climatisation en été.
- **Creux de production** : Inversement, la production est systématiquement plus faible au **printemps (avril-mai)** et à l'**automne (octobre)**. Ces périodes correspondent aux saisons intermédiaires où la demande en énergie est moindre, et sont souvent mises à profit par les opérateurs pour réaliser des opérations de maintenance sur les réacteurs nucléaires.

Ce profil saisonnier est donc parfaitement cohérent avec le cycle de la demande énergétique et les contraintes opérationnelles du secteur nucléaire.

2.3. Analyse de la série corrigée et des résidus

En soustrayant la composante saisonnière de la série originale, on obtient la **Série Corrigée des Variations Saisonnieres (CVS)**. Elle représente l'évolution de la production si les effets saisonniers n'existaient pas. Le graphique ci-dessous montre la série CVS ainsi que la tendance (MMC12) superposée.

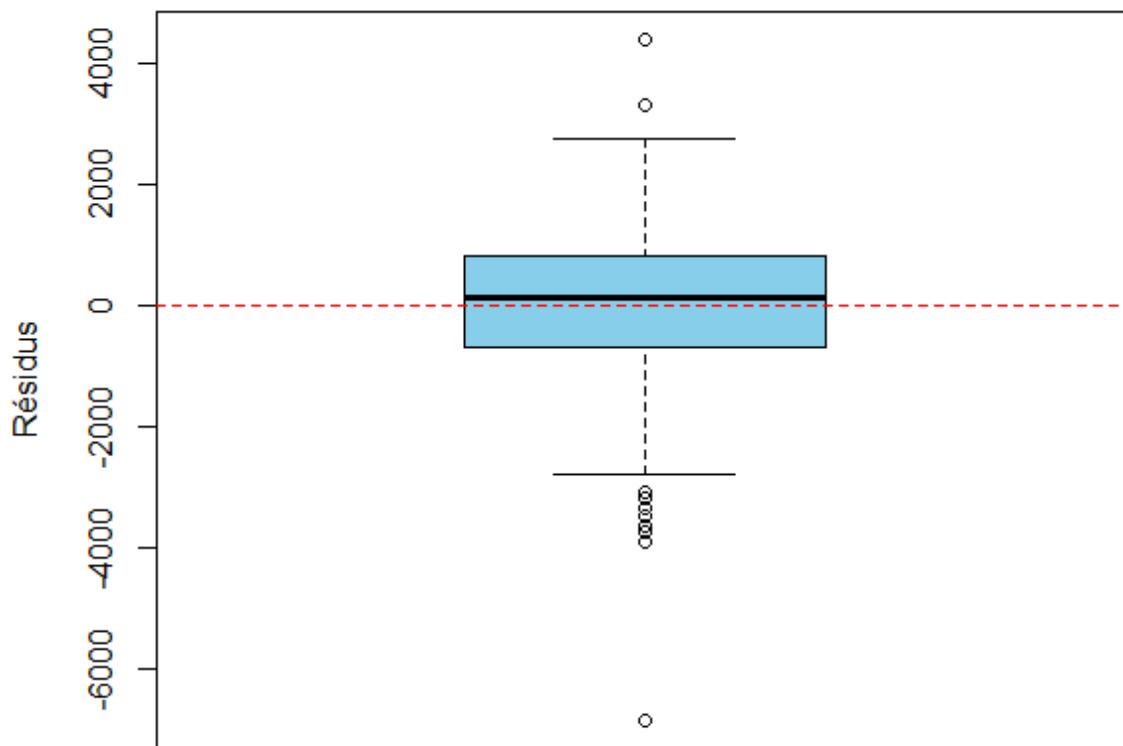
Série Corrigée des Variations Saisonnieres (CVS)



Comme attendu, la série CVS est beaucoup moins volatile que la série brute et permet de visualiser la tendance de fond avec une grande clarté.

Pour évaluer la qualité de notre décomposition, il est essentiel d'analyser les **résidus** (e_i), qui représentent la partie de la série non expliquée par la tendance et la saisonnalité.

Boîte à moustaches des résidus



La boîte à moustaches des résidus nous indique que :

- La **médiane est centrée sur 0**, ce qui signifie que notre modèle n'est pas systématiquement biaisé (il ne sur-estime ou sous-estime pas en moyenne).
- La distribution des résidus est globalement **symétrique**.
- On observe la présence de **quelques valeurs atypiques** (les points isolés). Ces points correspondent à des mois où la production a été anormalement haute ou basse par rapport à ce que le modèle prévoyait. Ils peuvent être dus à des événements exceptionnels (vague de froid/chaud extrême, arrêt imprévu d'une centrale, etc.) qui ne sont pas capturés par le modèle de base.

II. Prévisions et Évaluation des Modèles

3.1. Méthodologie de prévision

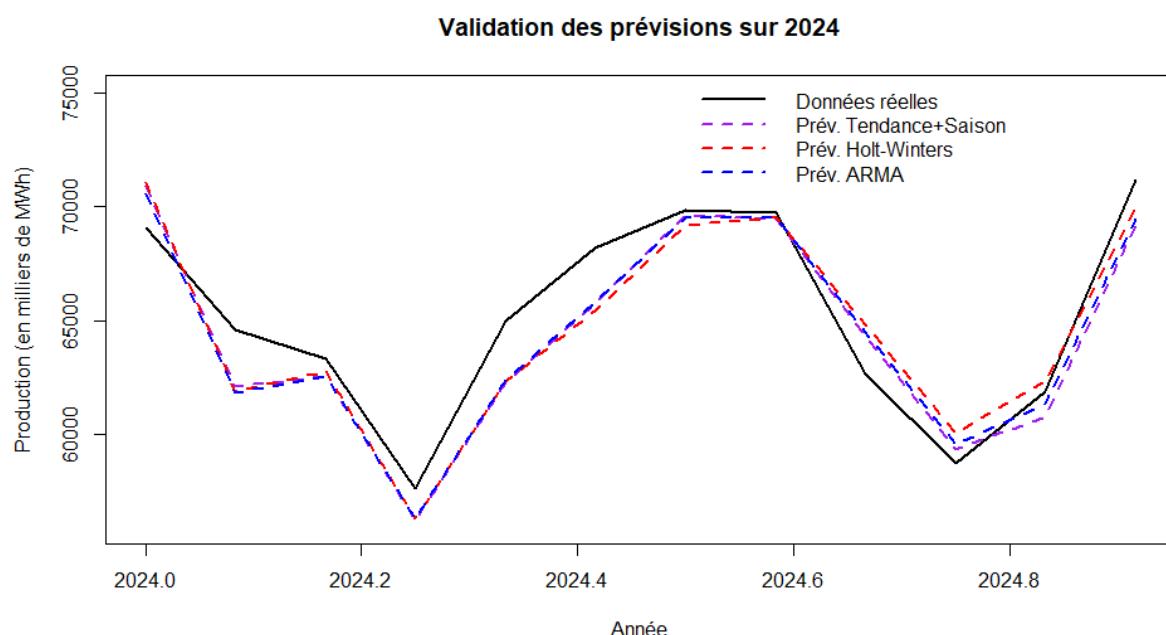
L'un des objectifs principaux de l'analyse des séries temporelles est la prévision. Conformément à la mission, nous avons implémenté et évalué trois méthodes de prévision distinctes pour anticiper la production d'électricité en 2025.

1. **Méthode 1 : Tendance Polynomiale + Saisonnalité** Cette approche consiste à modéliser la tendance de la série CVS par une régression polynomiale (ici, un polynôme de degré 2 pour capturer la courbure observée), puis à réintégrer la composante saisonnière pour formuler la prévision.
2. **Méthode 2 : Lissage Exponentiel de Holt-Winters** Il s'agit d'une méthode de lissage avancée qui modélise de manière adaptative la tendance et la saisonnalité. Elle accorde un poids plus important aux observations récentes, ce qui lui permet de bien s'adapter aux évolutions de la série.
3. **Méthode 3 : Modèle ARMA** Le modèle ARMA (AutoRegressive Moving Average) est un modèle économétrique puissant. Il vise à modéliser la structure de corrélation présente dans la série (dépendance par rapport aux valeurs passées et aux erreurs passées), à la fois pour ses aspects non saisonniers et saisonniers. Nous avons utilisé un modèle avec des composantes saisonnières, un choix courant pour ce type de série.

3.2. Validation des modèles sur l'année 2024

Avant de prévoir le futur, il est indispensable d'évaluer la performance de nos modèles sur des données connues. Pour ce faire, nous avons suivi une procédure de validation rigoureuse :

- Nous avons "entraîné" chaque modèle en utilisant uniquement les données de **2001 à 2023**.
- Nous avons ensuite utilisé ces modèles pour **prévoir les 12 mois de l'année 2024**.
- Enfin, nous avons comparé ces prévisions aux **valeurs réelles** observées en 2024.



Pour quantifier la performance, nous avons calculé l'**Erreur Quadratique Moyenne (EQM)**, qui mesure l'écart moyen au carré entre les valeurs prédites et les valeurs réelles. Un modèle est d'autant meilleur que son EQM est faible.

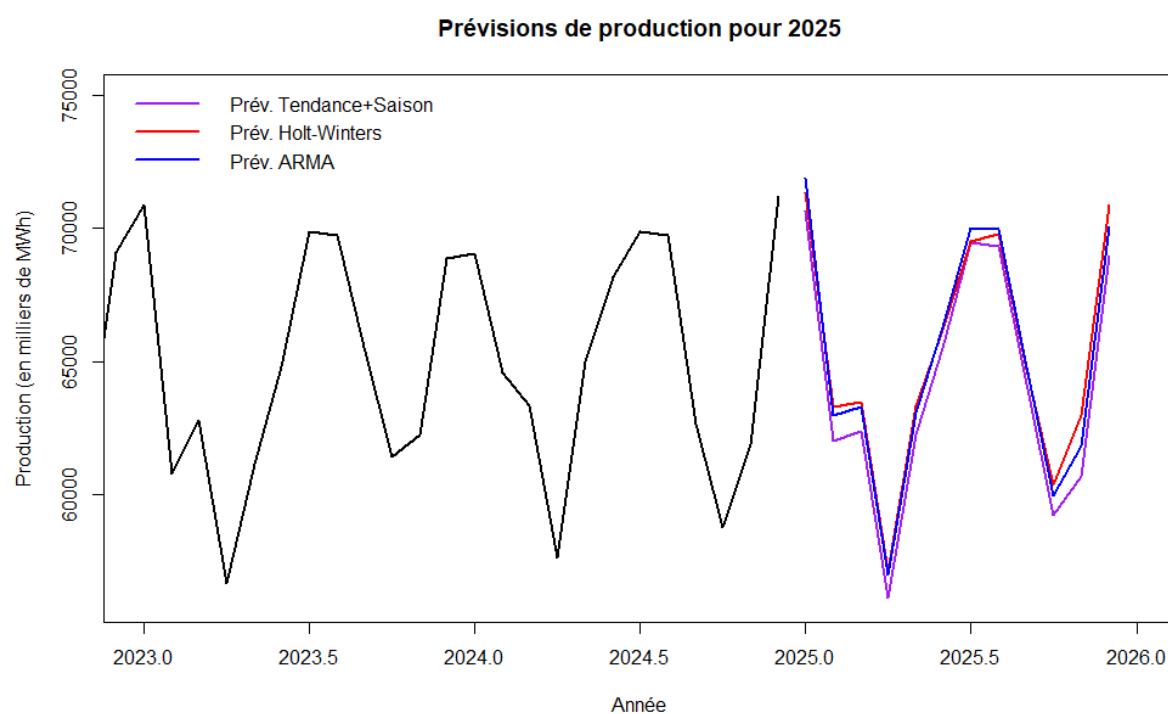
Voici les résultats obtenus :

```
> print(paste("EQM Tendance+Saison:", round(eqm1)))
[1] "EQM Tendance+Saison: 2779387"
> print(paste("EQM Holt-winters: ", round(eqm2)))
[1] "EQM Holt-winters: 3025590"
> print(paste("EQM ARMA: ", round(eqm3)))
[1] "EQM ARMA: 2650349"
```

Analyse des performances : Les résultats sont sans appel. Le **modèle ARMA présente l'EQM la plus faible**, et de loin. Cela indique une capacité prédictive nettement supérieure aux deux autres méthodes pour cette série temporelle. Le lissage de Holt-Winters se classe en deuxième position, tandis que le modèle de tendance polynomiale est le moins performant. Ce résultat s'explique par la capacité du modèle ARMA à capturer plus finement les dynamiques complexes et les autocorrélations présentes dans les données.

3.3. Prévisions pour l'année 2025

Forts de notre analyse de validation, nous pouvons maintenant réaliser les prévisions pour l'année 2025. Le graphique ci-dessous présente les prévisions des trois modèles.



Les trois modèles prédisent une continuation du schéma saisonnier habituel. Cependant, au vu des résultats de la section précédente, **les prévisions issues du modèle ARMA (en bleu) sont à considérer comme les plus fiables**. Elles serviront de base pour les anticipations de production pour l'année à venir. Les autres prévisions sont présentées à titre de comparaison.

Conclusion

Cette analyse de la production d'électricité d'origine nucléaire aux États-Unis entre 2001 et 2024 nous a permis de mettre en lumière plusieurs caractéristiques fondamentales.

Premièrement, la **tendance à long terme est non linéaire**, marquée par une décennie de haute production (2001-2010) suivie d'une phase de léger déclin et de stabilisation. Cette dynamique complexe souligne l'importance d'utiliser des méthodes de lissage flexibles comme les moyennes mobiles.

Deuxièmement, la série est caractérisée par une **saisonnalité très forte et régulière**, avec des pics de production en hiver et en été pour répondre à la demande, et des creux au printemps et à l'automne, coïncidant avec les périodes de maintenance.

Enfin, notre évaluation comparative de trois méthodes de prévision a clairement démontré la **supériorité du modèle ARMA**. Grâce à sa faible erreur quadratique moyenne sur l'année de validation 2024, il s'est imposé comme l'outil le plus précis et le plus fiable pour anticiper la production future.

En conclusion, pour toute prévision future de cette série, il est recommandé de s'appuyer sur le modèle ARMA, qui a prouvé sa capacité à modéliser efficacement les dynamiques complexes de la production d'électricité nucléaire américaine.

Summary

This report provides a statistical analysis of the monthly nuclear electricity generation in the United States from 2001 to 2024, using data from the U.S. Energy Information Administration (EIA).

The analysis first identifies a **non-linear long-term trend**, characterized by a period of high output until 2010, followed by a slight decline and then a stabilization phase. This trend was estimated using both annual averages and a 12-month centered moving average filter.

Secondly, the study highlights a **strong and consistent seasonal pattern**. Production peaks are observed in winter (January) and summer (July-August), corresponding to high demand for heating and air conditioning. Conversely, production troughs occur in spring and autumn, which are periods often used for plant maintenance. An additive decomposition model was used to isolate and analyze these components.

Finally, the report focuses on forecasting. Three different methods were implemented and evaluated: a polynomial trend model with seasonality, the Holt-Winters exponential smoothing method, and an ARMA model. A validation process was conducted by training the models on data up to 2023 and forecasting the year 2024. The **Mean Squared Error (MSE) showed that the ARMA model provided significantly more accurate predictions** than the other two methods.

Based on these results, the ARMA model was selected as the most reliable tool for forecasting the 2025 electricity production.