

Rapport de la production d'électricité de l'ensemble des énergies

Rapport écrit par:

Enzo LERICHE, Samuel DARMALINGON, Franck TANKHAPANYA

BUT Science des données FI EMS

2023-11-13



IUT de Paris - Rives de Seine
Université de Paris

Sommaire

Introduction	3
Analyse	4
Importation des données	4
Importation du jeu de donnée	4
Transformation des données en série temporelle	4
Tendance de la série	5
Moyenne Mobile	5
Courbe de régression	6
Saisonnalité	8
Fonction decompose	8
Traçage de la saisonnalité	8
Série désaisonnalisée	9
CVS	9
Graphe des résidus	10
Boîte à moustaches des résidus	11
Prévision de la production d'électricité pour l'année 2022	12
Modèle polynomial	14
HoltWinters.	14
ARMA.	15
Représentation des 3 modèles en 1 graphique	17
Vérification des méthodes de prévisions pour 2021.	18
Modèle polynomial.	19
Holt-Winters.	19
ARMA.	19
Représentation des 3 modèles	20
Conclusion	21
Conclusion général	21
Résumer en anglais	21

Introduction

L'énergie occupe une place centrale dans les préoccupations mondiales, avec sa production, distribution et consommation jouant un rôle fondamental dans le développement économique, la stabilité environnementale, et la qualité de vie des populations. Dans ce contexte, l'Administration américaine de l'information sur l'énergie (EIA) émerge comme un acteur clé, s'engageant activement à collecter, analyser, et diffuser des informations cruciales sur le secteur énergétique. En tant que statisticien au sein de l'équipe dédiée à l'analyse des données de l'EIA, notre mission est d'explorer l'ensemble des séries temporelles associées à la production d'électricité, couvrant la période allant de 2001 à 2022.

Les données que nous explorons, contenues dans le fichier "AllFuels", offrent une vue détaillée des divers modes de production d'électricité, allant du charbon aux énergies renouvelables, couvrant une période de 21 années complètes d'observation.

Notre analyse se concentrera sur plusieurs aspects clés de ces séries temporelles. Tout d'abord, nous explorerons la tendance générale en utilisant des filtres de moyennes mobiles et des courbes de régression des moyennes annuelles. Ensuite, nous plongerons dans la composante saisonnière des données, calculerons les coefficients saisonniers et présenterons une décomposition complète de la série. Après avoir désaisonnalisé les séries, nous examinerons les résidus à travers des boîtes à moustaches.

La seconde phase de notre analyse sera dédiée à la prévision. Nous recourrons à trois méthodes distinctes pour anticiper la production d'électricité en 2022. Afin d'évaluer la robustesse de ces approches, nous les appliquerons également sur les données de l'année 2021, en utilisant les données recueillies entre 2001 et 2020. En conclusion, nous procéderons au calcul de l'erreur quadratique moyenne de prévision pour chacune de ces méthodes, offrant ainsi une mesure quantitative de leur performance respective.

Ce rapport vise à fournir une compréhension approfondie de la dynamique temporelle de la production d'électricité, offrant des éléments pertinents pour les décideurs, les chercheurs et le public. Chaque étape de notre analyse sera accompagnée de commentaires approfondis sur les graphiques générés et les indicateurs calculés, garantissant ainsi une interprétation claire et informée des résultats.

Analyse

Importation des données

Importation du jeu de donnée

```
data<- read.table("C:/Users/User/OneDrive/Documents/BUT/2eme annee/serie temporelle pour sae/AllFuels.t
```

Après cette importation, on a transformé ces données en série temporelle afin d'étudier la production d'électricité des énergies au cours du temps de 2001 à 2021 de manière mensuelle.

Transformation des données en série temporelle

Pour pouvoir transformer les données en série temporelle, nous avons utilisé la fonction “*ts*” en R.

-“*frequency*” : permet d'indiquer le nombre d'observation par unité de temps dans la série temporelle. Par exemple, pour les données mensuelles, la fréquence est de 12, pour des données trimestrielles la fréquence est 4.

-“*start*” : spécifie le point de départ de la série temporelle.

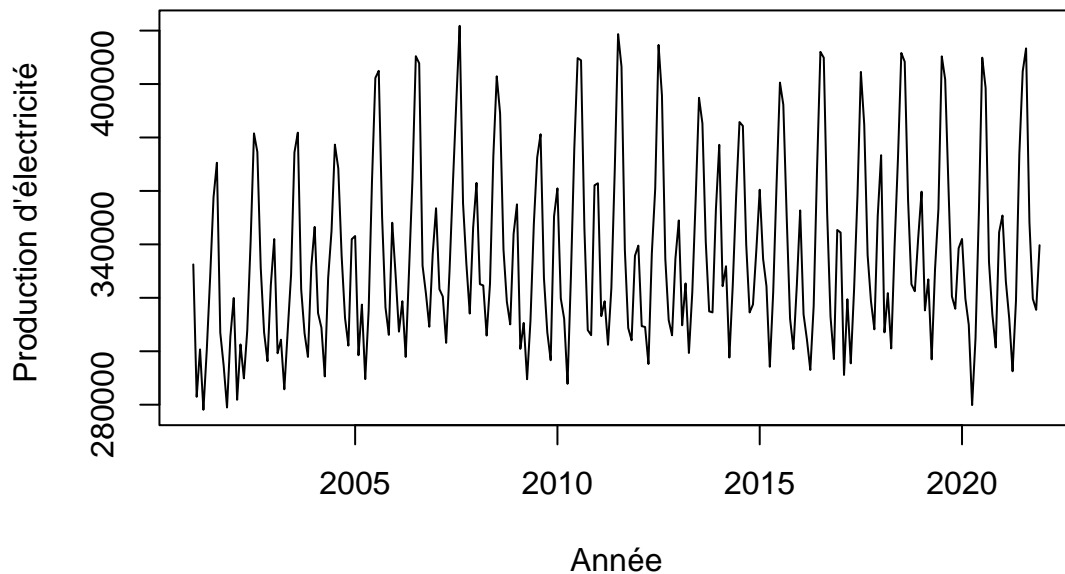
```
Data <- ts(data,frequency=12,start=c(2001,1))
```

Dans notre cas, nous voulions des données mensuelles, nous avons donc précisé “*frequency=12*” ; de plus, notre jeu de donnée commençait en 2001 avec un pas de 1 année par 1 année.

Vous pouvez observer ci-dessous la production d'électricité des énergies au fil du temps mesuré en mégawatt/heure.

```
plot(Data, main = "La production d'électricité en fonction du temps",  
      ylab = "Production d'électricité ", xlab = "Année")
```

La production d'électricité en fonction du temps



Nous observons qu'il y a une fluctuation de la production d'électricité au cours du temps.

Tendance de la série

Moyenne Mobile

Présentation des moyennes mobiles :

Par la suite, nous avons calculé la moyenne mobile (MM) et la moyenne mobile centrée (MMC) de cette série chronologique. Cette démarche a été faite pour étudier la tendance de la série au cours du temps. Car la moyenne mobile a pour but de calculer la moyenne des valeurs adjacentes sur une fenêtre glissante de la série temporelle. Cela permet de lisser les fluctuations à court terme et de mettre en évidence des tendances ou des motifs plus généraux. Or que la moyenne mobile centrée réduit le décalage temporel introduit par la moyenne mobile et permet de mieux représenter la tendance centrale des données. Ces deux méthodes sont utilisées pour lisser les variations à court terme et identifier les tendances plus larges. En outre, nous avons calculé ces moyennes d'ordre 12 car elles permettent d'identifier les tendances annuelles.

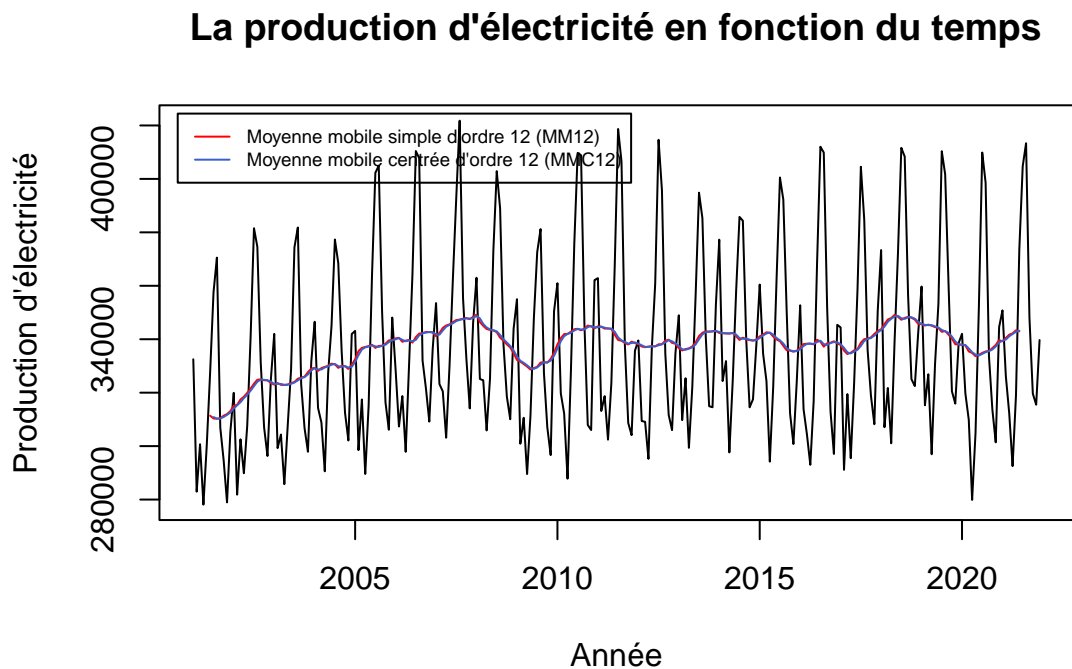
Calcul des moyennes mobiles

```
MM12 <- filter(Data, filter=rep(1/12,12))  
MMC12 <- filter(Data,filter=c(1/24,rep(1/12,11),1/24))
```

Nous avons représenté ces moyennes mobiles simples et centrées sur le graphique de la production d'électricité au fil du temps mesuré en mégawatt/heure.

Représentation graphique

```
plot(Data, main = "La production d'électricité en fonction du temps",
      ylab = "Production d'électricité", xlab = "Année")
lines(MM12, col = "red")
lines(MMC12, col = "#3A5FCD")
legend("topleft", lty=c(1,1), col=c("red", "#3A5FCD"),
      legend=c("Moyenne mobile simple d'ordre 12 (MM12)",
               "Moyenne mobile centrée d'ordre 12 (MMC12)"),
      inset=0.02, cex=0.6)
```



Par ce graphique, nous remarquons que les moyennes mobiles se superposent au cours du temps. Par conséquent, on peut supposer la présence d'un cycle ou d'une saisonnalité dans la série temporelle. Un motif cyclique malgré le lissage, cela pourrait être révélateur de composantes saisonnières.

Courbe de régression

Ensuite, pour l'étude de la tendance de la série, nous avons examiné la courbe de régression des moyennes annuelles. En effet, la courbe de régression offre une méthode pour modéliser et comprendre la tendance à long terme d'une série ou le comportement moyen d'une série au fil du temps. Elle peut être essentielle pour les prévisions d'analyse. Dans notre cas, nous avons calculé la courbe de régression des moyennes annuelles et nous l'avons comparé à la moyenne mobile centrée d'ordre 12 sur le graphique qui représente la production d'électricité des énergies au cours du temps.

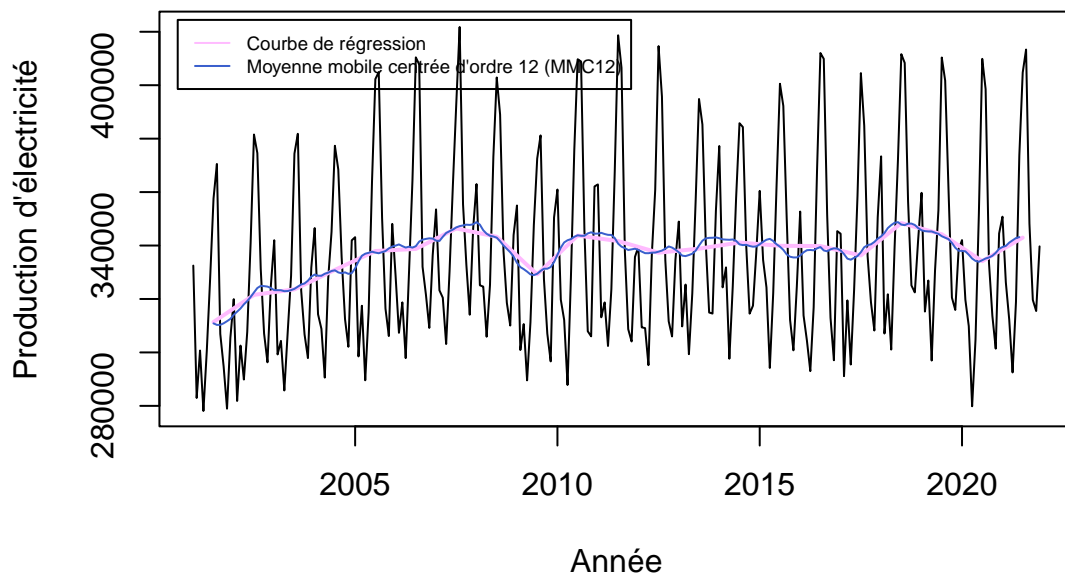
Calcul et représentation de la courbe de régression basée sur les moyennes annuelles.

```
x <- as.vector(time(Data))
y <- as.vector(Data)
length(x)/12
```

```
## [1] 21
```

```
xclass <- cut(x,21)
meanannual <- tapply(y,xclass,mean)
tt <- seq(2001.5, 2021.5, by=1)
plot(Data, main = "La production d'électricité en fonction du temps",
      ylab = "Production d'électricité", xlab = "Année")
lines(tt, meanannual, col="#FFBBFF", lwd=2)
lines(MMC12,col="#3A5FCD")
legend("topleft",lty=c(1,1),col=c("#FFBBFF", "#3A5FCD"),
      legend=c("Courbe de régression",
               "Moyenne mobile centrée d'ordre 12 (MMC12)"),
      inset=0.02,cex=0.6)
```

La production d'électricité en fonction du temps



De ce graphique, nous observons que la moyenne mobile centrée a la même forme que la courbe de régression des moyennes annuelles, cela indique que la tendance à long terme de la série temporelle est bien capturée par la moyenne mobile centrée, et que cette tendance suit une évolution relativement régulière ou prévisible. En d'autres termes, la moyenne mobile centrée offre une représentation lissée de la série, mettant en évidence la tendance à long terme, et la similitude avec la courbe de régression renforce la confiance dans la capacité du modèle à capturer et à représenter la dynamique de la série temporelle. Cependant, il est toujours important de vérifier d'autres aspects tels que les résidus du modèle pour évaluer la qualité globale de l'ajustement.

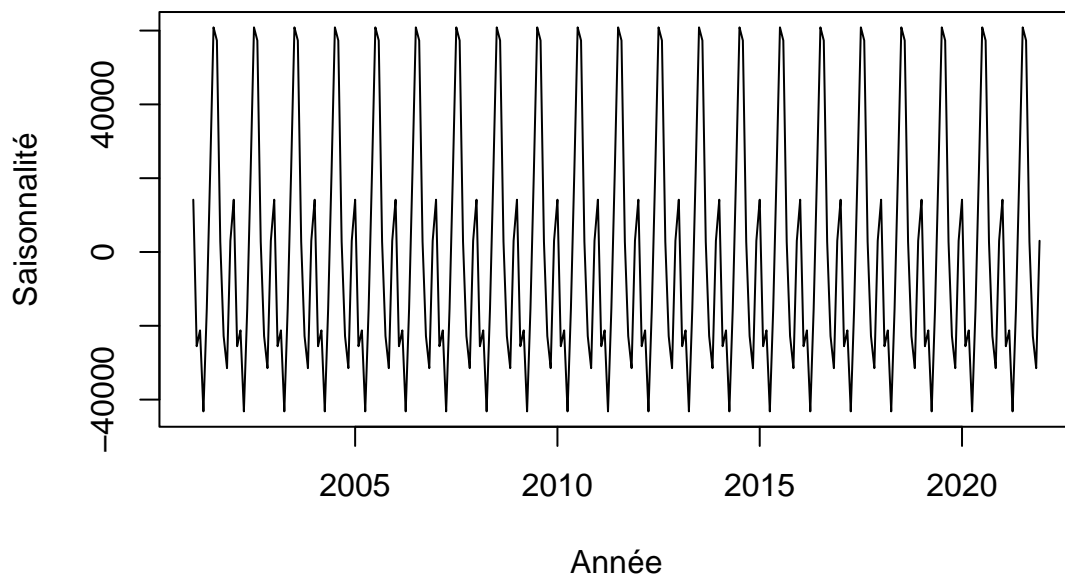
Saisonnalité

Fonction decompose

Par ailleurs, nous nous sommes préoccupé de la décomposition de la série parce qu'elle permet d'analyser ces différentes composantes de manière séparée, facilitant ainsi l'interprétation et la modélisation de la série.

```
dec <- decompose(Data)
plot(dec$seasonal,
     main= "Décomposition saisonnière de la production d'électricité",
     xlab="Année", ylab="Saisonnalité")
```

Décomposition saisonnière de la production d'électricité



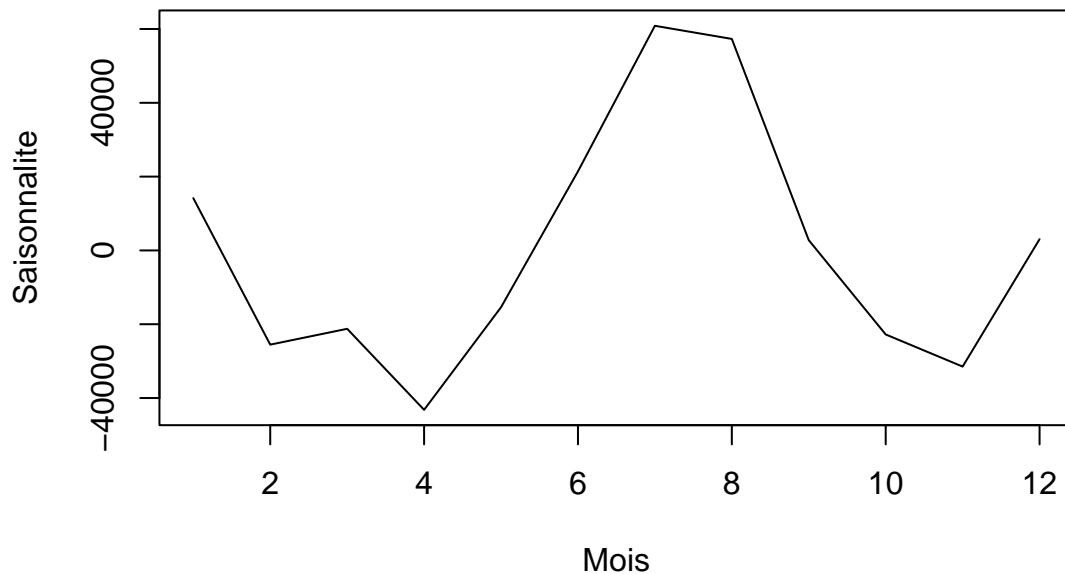
Sur ce graphique, nous remarquons par cette décomposition saisonnière, que la saisonnalité est répétitive chaque année. Par conséquent, nous avons porté notre intérêt sur une année.

Traçage de la saisonnalité

Traçage de la saisonnalité sur une année

```
plot(1:12,dec$seasonal[1:12],type="l",xlab="Mois",ylab="Saisonnalité",
     main="Saisonnalité sur une année")
```


Saisonnalité sur une année



Nous observons que la production d'électricité pendant l'hiver diminue, puis à partir du printemps, la production d'électricité augmentent fortement pour ensuite atteindre un pic en été. Ensuite, cette production diminue pendant l'automne. Donc nous pouvons dire que par une grande production des énergies durant l'été, la population consomme plus durant l'été contrairement à la période froide.

Série désaisonnalisée

CVS

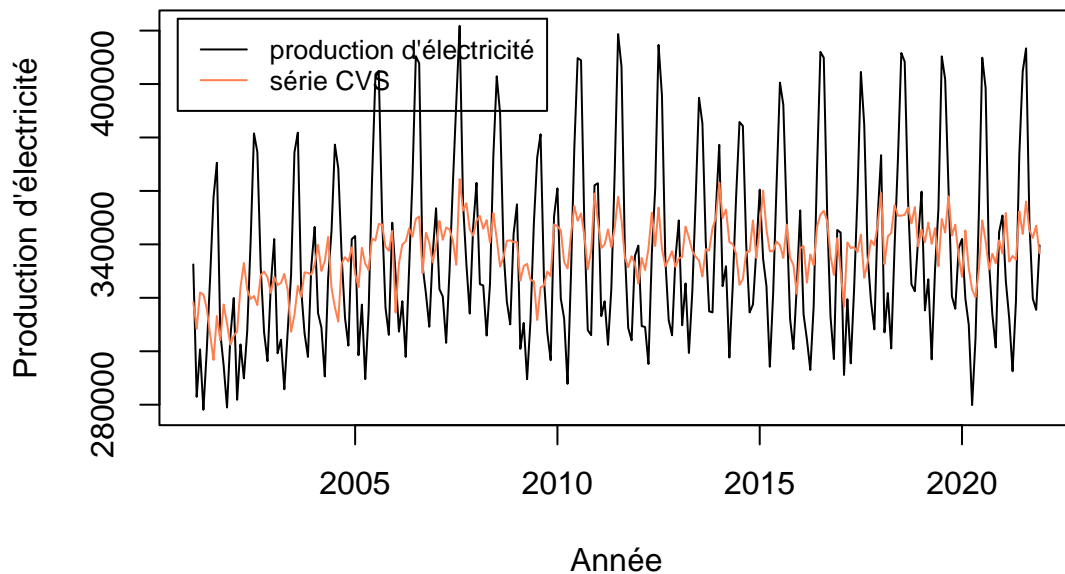
Après avoir analysé les variations saisonnières, nous nous sommes penchées sur la tendance de la série et les composantes irrégulières. L'objectif de la désaisonnalisation est de mieux mettre en évidence la tendance sous-jacente et de rendre les données plus comparables d'une période à l'autre en éliminant les variations cycliques associées aux saisons. Afin de voir les possibles erreurs sur cette série. De ce fait, nous avons calculé la série corrigée des variations saisonnières (CVS) qui est une série temporelle où les composantes saisonnières ont été supprimées.

Calcul de la CVS et représentation graphique

```
CVS <- y - dec$seasonal

plot(Data, main="La production d'électricité en fonction du temps",
      ylab="Production d'électricité", xlab="Année")
lines(x, CVS, type="l", col="coral")
legend("topleft", lty=c(1,1), col=c("black", "coral"),
      legend=c("production d'électricité", "série CVS"), inset=0.02, cex=0.8)
```

La production d'électricité en fonction du temps



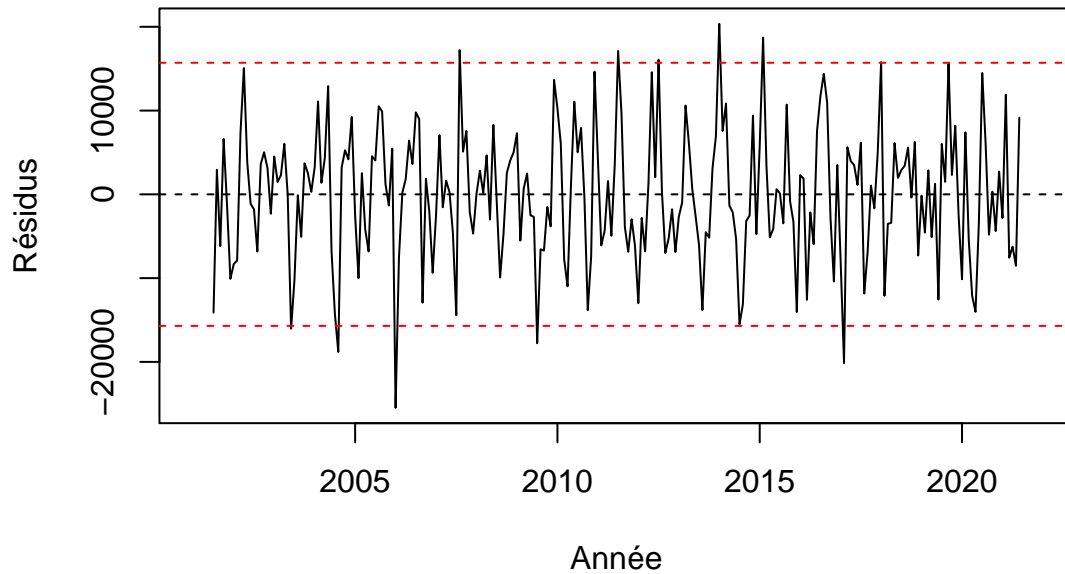
Nous apercevons que la série CVS n'est pas confondue à la série temporelle de la production d'électricité d'origine. Cela signifie que la composante saisonnière joue un rôle significatif dans la série temporelle. Effectivement, la série originale est fortement influencée par des variations régulières liées à la saison.

Nous nous sommes penchés sur les résidus de la série, car les résidus représentent la différence entre les valeurs observées d'une variable et les valeurs prédites par la courbe de régression tracée précédemment. Ce qui nous permet ainsi d'évaluer l'efficacité du modèle.

Graphe des résidus

```
plot(dec$random,main="Le graphe des résidus de la production d'électricité",
      xlab= "Année", ylab="Résidus")
sd_resid<- sd(dec$random,na.rm=TRUE)
abline (h= 2*sd_resid,col="red", lty=2)
abline (h=- 2*sd_resid,col="red", lty=2)
abline (h= 0,col="black", lty=2)
```

Le graphe des résidus de la production d'électricité

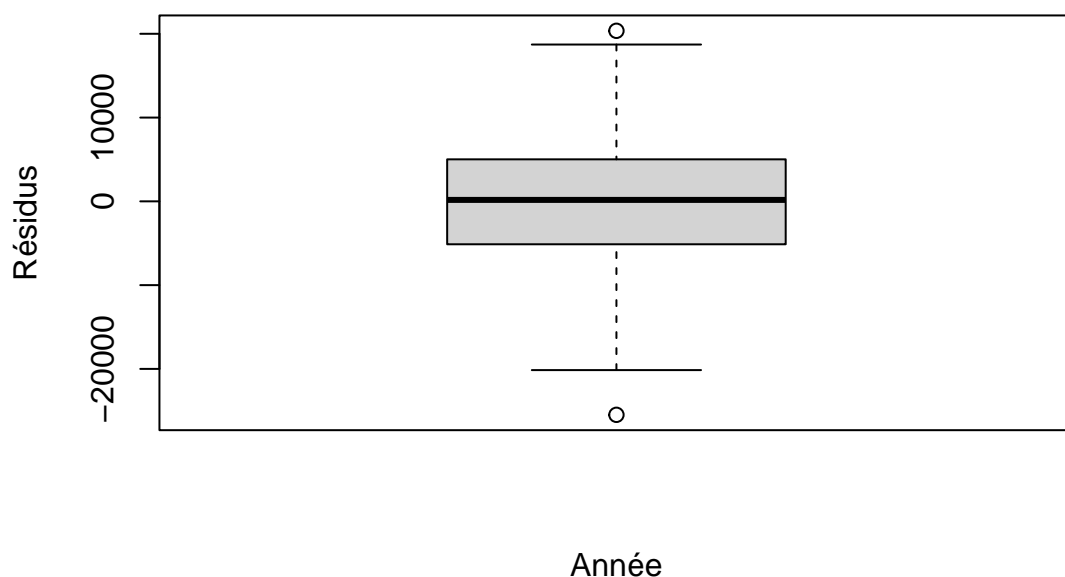


Nous remarquons qu'il y a 10 valeurs erronées de prédiction. Le nombre d'erreurs n'est pas conséquent donc nous pouvons dire que la performance du modèle n'est pas mauvaise.

Boîte à moustaches des résidus

```
boxplot(dec$random,  
  main= "Boîte à moustaches des résidus de la production d'électricité",  
  xlab= "Année", ylab="Résidus")
```

Boîte à moustaches des résidus de la production d'électricité



Par cette boîte à moustaches, les erreurs sont représentées sur ce graphique de la répartition des observations prédites.

Prévision de la production d'électricité pour l'année 2022

Suite à l'examination de la série, nous avons effectué des prévisions pour l'année 2022 à partir de différents modèles.

Tout d'abord, nous avons hésité entre le choix de deux modèles, linéaire ou polynomial. De ce fait, nous avons calculé ces modèles et nous les avons comparés à la série temporelle des observations entre 2001 à 2021.

Le modèle linéaire est une méthode statistique qui cherche à modéliser la relation entre une variable de sortie et des variables d'entrée en supposant que cette relation est linéaire. L'équation du modèle prend la forme d'une combinaison linéaire des variables d'entrée, pondérées par des coefficients. L'objectif est de trouver les coefficients qui minimisent la différence entre les valeurs prédites par le modèle et les valeurs réelles observées. Bien que simple et interprétable, le modèle linéaire peut ne pas convenir aux relations complexes et non linéaires.

Modèle linéaire

```
modele_lin <- lm(CVS~x)
modele_lin$coefficients#les coefficients estimés
```

```
## (Intercept)          x
## -1473403.227      899.912
```

```
tend_lin <- modele_lin$coefficients[1]+modele_lin$coefficients[2]*x
```

Or, le modèle polynomial est une extension qui permet de capturer des relations non-linéaires en utilisant des polynômes. Le modèle polynomial offre une flexibilité accrue, mais il faut choisir judicieusement l'ordre du polynôme pour éviter le surajustement. En général, le modèle linéaire est plus simple et moins sujet aux surajustement, mais il peut ne pas être adapté à des relations plus complexes dans les données.

Modèle polynomial

```
modele_poly <- lm(CVS~x+I(x^2))
modele_poly$coefficients#les coefficients estimés
```

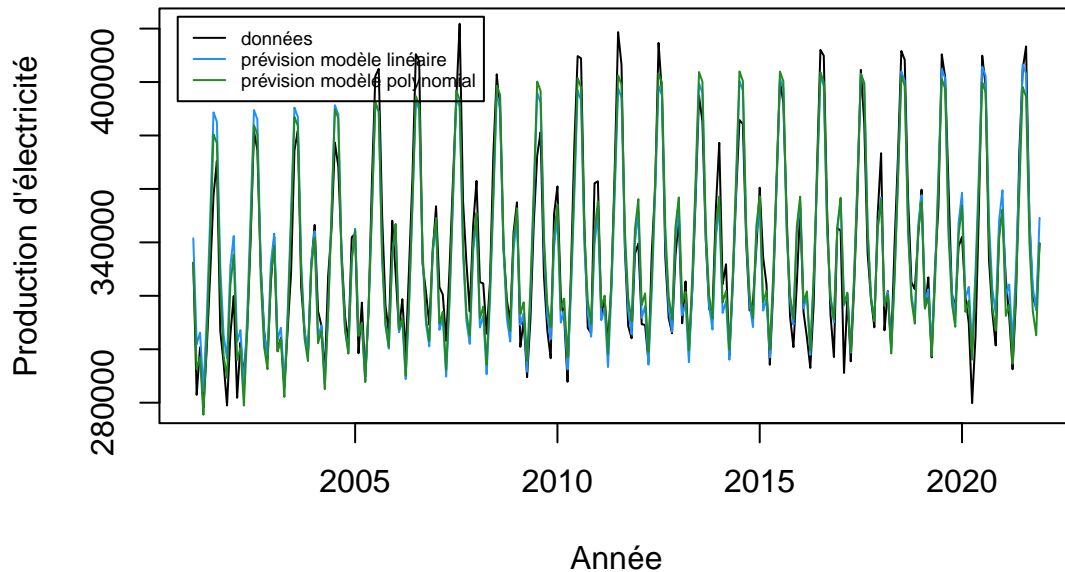
```
## (Intercept)          x          I(x^2)
## -5.352477e+08  5.316383e+05 -1.319288e+02
```

```
tend_poly <- modele_poly$coefficients[1]+modele_poly$coefficients[2]*x+
  modele_poly$coefficients[3]*x^2
```

Donc nous les avons représentés sur la série temporelle et puis observer le meilleur modèle.

```
plot(Data, main="Production Electrique 2001-2022:\n Comparaison des modèles",
      ylab="Production d'électricité", xlab="Année")
lines(x,tend_lin+dec$seasonal,type="l",col="dodgerblue")
lines(x,tend_poly+dec$seasonal,type="l",col="forestgreen")
legend("topleft",lty=c(1,1),col=c("black","dodgerblue","forestgreen"),
      legend=c("données",
               "prévision modèle linéaire","prévision modèle polynomial"),
      cex=0.55,inset=0.02)
```

Production Electrique 2001–2022: Comparaison des modèles



Nous constatons que le modèle polynomial est plus efficace, car nous pouvons voir qu'il y a une grande similarité entre ce modèle et les données temporelles que nous possédons. Dès lors, nous avons utilisé ce modèle pour faire ce modèle pour les prévisions de 2022.

Modèle polynomial

Prévision avec le modele polynomial

```
x_fut <- seq(2022,2023,length.out=13)[-13]
modele_poly <- lm(CVS~x+I(x^2))
tend_poly_fut <- modele_poly$coefficients[1]+
  modele_poly$coefficients[2]*x_fut+modele_poly$coefficients[3]*x_fut^2
tend_poly_fut+dec$seasonal[1:12]
```

```
## [1] 350580.4 310718.7 314847.3 292733.1 320373.4 357022.9 396302.1 392578.4
## [9] 337877.5 312150.9 303264.8 337613.4
```

HoltWinters.

Présentation du modèle :

Le modèle Holt-Winters, une méthode de prévision utilisée pour les séries temporelles, qui prend en compte la tendance, la saisonnalité et les fluctuations aléatoires. Ce modèle est particulièrement utile lorsque les données présentent des motifs saisonniers. Elle utilise des équations de mise à jour pour estimer ces composantes au fil du temps, puis utilise ces estimations pour faire des prédictions pour l'année 2022.

Prévision avec le modele HoltWinters

```
hw <- HoltWinters(Data)
prev_hw <- predict(hw, n.ahead=12)
```

ARMA.

Présentation du modèle :

Le modèle ARMA (AutoRegressive Moving Average) est un modèle statistique utilisé pour modéliser les séries temporelles. Il combine deux composants principaux : l'autorégression (AR) et la moyenne mobile (MA).

-*MA* : les valeurs acceptables pour q et Q de la partie MA de la série SARMA sont déduit du graphique de l'autocorrélation qu'on obtient à l'aide de la commande *acf*.

-*AR*: les valeurs acceptables pour p et P de la partie AR de la série SARMA sont déduit du graphique de l'autocorrélation partielle qu'on obtient à l'aide de la commande *pacf*.

ACF et PACF

D'abord, nous allons constituer *Data_diff* qui est la série différenciée deux fois (à l'ordre 12 puis à l'ordre 1). Puis nous allons vérifier que cette série est bien stationnaire grâce à la commande "*adf_test*".

Nous tracerons ensuite les graphiques ACF et PACF pour déterminer les valeurs acceptables pour q , Q et p , P .

```
Data_diff <- diff(Data, lag=12, differences=1)
Data_diff <- diff(Data_diff, lag=1, differences=1)
require(tseries)
```

```
## Le chargement a nécessité le package : tseries
```

```
## Registered S3 method overwritten by 'quantmod':
##   method           from
##   as.zoo.data.frame zoo
```

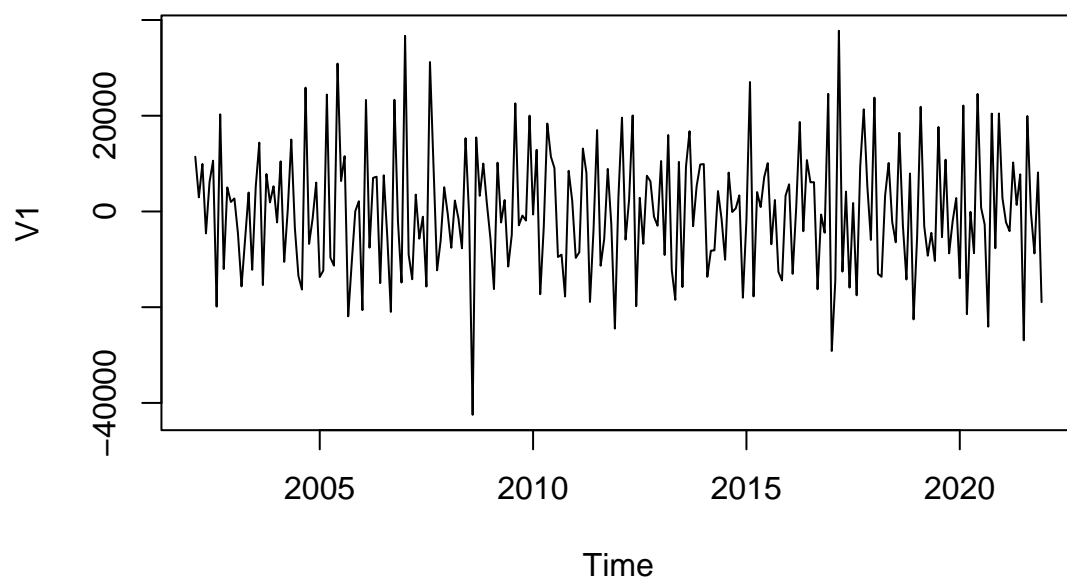
```
adf.test(Data_diff, alternative=c("stationary"), 12)
```

```
## Warning in adf.test(Data_diff, alternative = c("stationary"), 12): p-value
## smaller than printed p-value
```

```
##
##   Augmented Dickey-Fuller Test
##
## data:   Data_diff
## Dickey-Fuller = -6.1767, Lag order = 12, p-value = 0.01
## alternative hypothesis: stationary
```

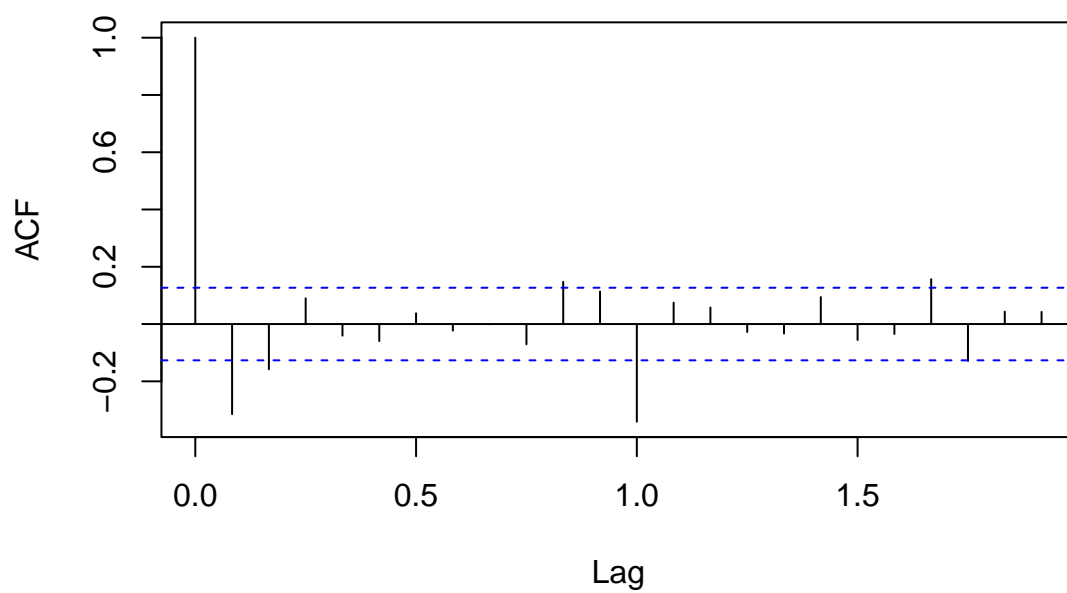
```
plot(Data_diff, main="Série différenciée")
```

Série différenciée

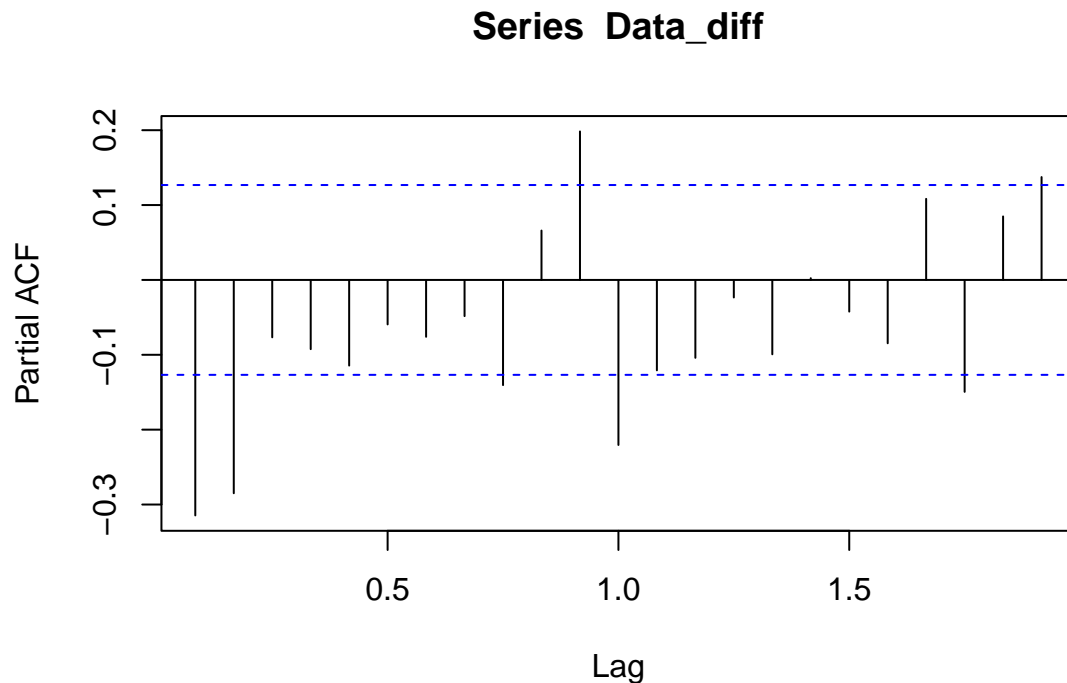


```
acf(Data_diff)
```

V1




```
pacf(Data_diff)
```



Prévision avec le modele ARMA

```
z1 <- arima(Data, order=c(1,1,1), seasonal =list(order=c(2,1,1), period=12))
AIC(z1)
```

```
## [1] 5049.732
```

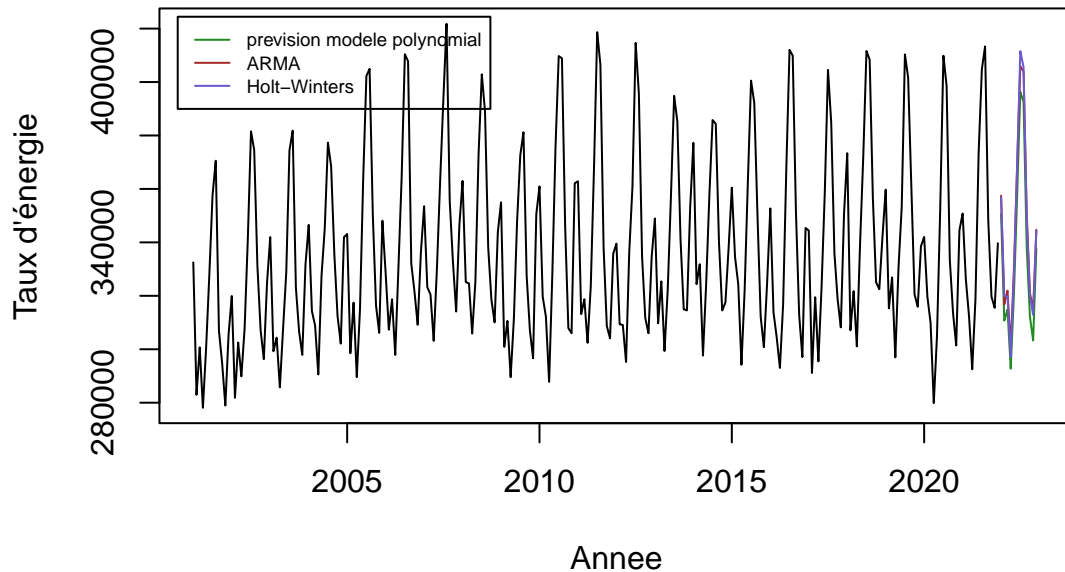
```
prev <- predict(z1, n.ahead=12)
```

Représentation des 3 modèles en 1 graphique

Après avoir calculé les prédictions avec les 3 modèles : polynomial, Holt-Winters, ARMA. Nous allons les représenter graphiquement ensemble sur le graphique de la production d'électricité.

```
plot(Data, main="Production d'électricité entre 2001 et 2022
(avec previsions sur l'annee 2022)",
      ylab="Taux d'énergie", xlab="Annee",xlim=c(2001,2023))
lines(x_fut,tend_poly_fut+dec$seasonal[1:12],type="l",col="forestgreen")
lines(prev$pred, col="brown",lwd=1)
lines(prev_hw, col="slateblue",lwd=1)
legend("topleft",lty=c(1,1),col=c("forestgreen","brown","slateblue"),
      legend=c("prevision modele polynomial",
               "ARMA",
               "Holt-Winters"),inset=0.02,cex=0.6)
```

Production d'électricité entre 2001 et 2022 (avec previsions sur l'annee 2022)



Vérification des méthodes de prévisions pour 2021.

Afin de déterminer le modèle le plus approprié pour effectuer des prévisions sur cette série temporelle, nous avons entrepris une vérification en utilisant ces modèles spécifiques au cours de l'année 2021. Cette étape s'avère cruciale pour évaluer la précision et l'efficacité de chacun des modèles dans la prédiction des données de la série. En analysant leur performance sur une période récente, nous sommes en mesure de prendre des décisions éclairées quant au choix du modèle le plus fiable pour anticiper la tendance de la série au fil du temps. Cette approche méthodique contribue à garantir la validité des prévisions futures en se basant sur des résultats récents et pertinents. Par ailleurs, pour chaque modèle, nous avons calculé l'erreur quadratique moyenne (MSE). C'est une mesure couramment utilisée pour évaluer les performances d'un modèle de prévision. L'erreur quadratique moyenne mesure la moyenne des carrés des différences entre les valeurs prédites par le modèle et les valeurs réelles observées. Entre plusieurs modèles, elle aide à comparer leur performance. Le modèle avec l'erreur quadratique moyenne le plus bas est généralement considéré comme le meilleur.

Création d'un échantillon d'apprentissage et d'un échantillon de test

```
data_train <- ts(Data[1:240],frequency=12,start=c(2001,1))
data_test  <- ts(Data[241:252],frequency=12,start=c(2021,1))
```

Modèle polynomial.

Calcul des prévisions et de l'erreur quadratique moyenne

```
dec <- decompose(data_train)
x <- as.vector(time(data_train))
y <- as.vector(data_train)
x_fut <- as.vector(time(data_test))
CVS <- y-dec$seasonal

modele_poly <- lm(CVS~x+I(x^2))
tend_poly <- modele_poly$coefficients[1]+
  modele_poly$coefficients[2]*x_fut+modele_poly$coefficients[3]*x_fut^2
pred_poly <- tend_poly+dec$seasonal[1:12]

mse_poly <- mean((pred_poly-data_test)^2)
mse_poly
```

```
## [1] 137194536
```

Holt-Winters.

Calcul des prévisions et de l'erreur quadratique moyenne

```
hw <- HoltWinters(data_train)
prev_hw <- predict(hw, n.ahead=12)

mse_hw <- mean((prev_hw-data_test)^2)
mse_hw
```

```
## [1] 58636322
```

ARMA.

Calcul des prévisions et de l'erreur quadratique moyenne

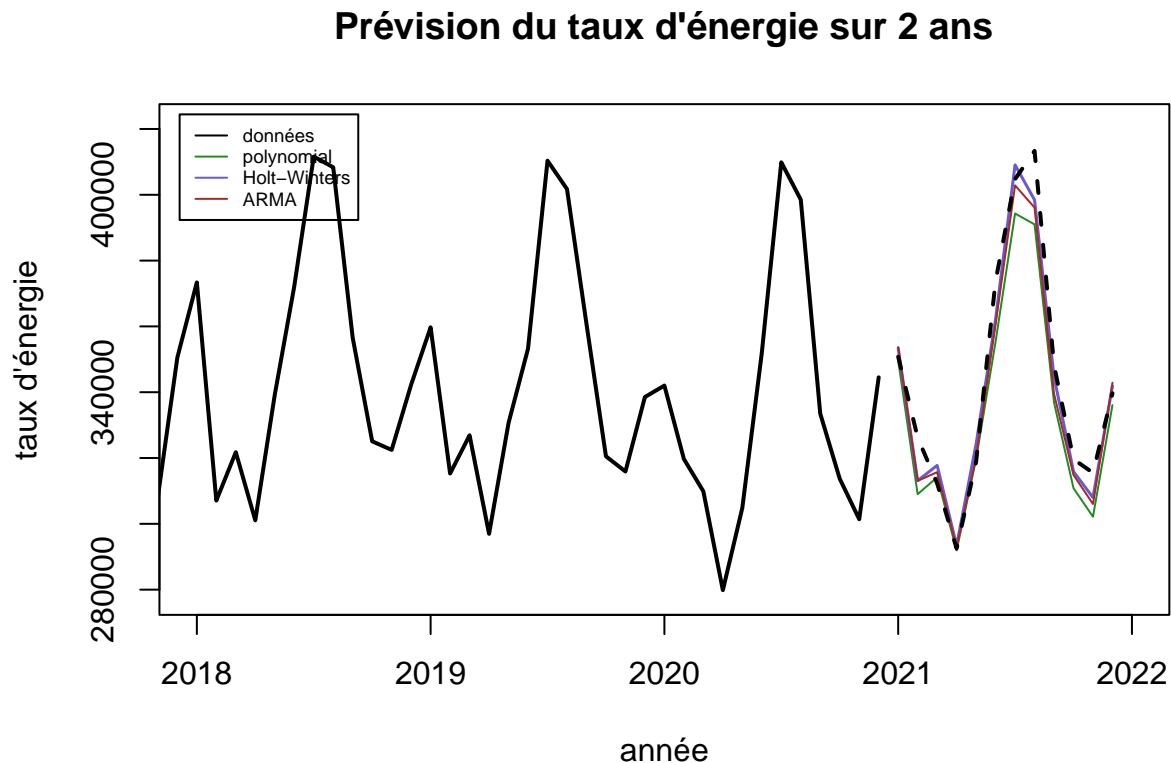
```
z <- arima(data_train, order=c(2,1,1), seasonal=list(order=c(2,1,1), period=12))
prev <- predict(z, n.ahead=12)

mse_arima <- mean((prev$pred-data_test)^2)
mse_arima
```

```
## [1] 72201074
```

Représentation des 3 modèles

```
plot(data_train, main="Prévision du taux d'énergie sur 2 ans", cex.main=1.2,
      xlab="année", ylab="taux d'énergie", xlim=c(2018,2022),lwd=2)
lines(x_fut,pred_poly, col="forestgreen",lwd=1,type="l")
lines(prev_hw, col="slateblue",lwd=1.5)
lines(prev$pred, col="brown",lwd=1)
lines(x_fut,data_test, col="black",lwd=2,type="l",lty=2)
legend("topleft", legend=c("données","polynomial","Holt-Winters","ARMA"),
      lty=c(1,1,1,1),
      col=c("black", "forestgreen","slateblue","brown"),
      cex=0.55,inset=0.02)
```



L'analyse du graphique dépeignant l'efficacité des modèles de prévision pour l'année 2021 révèle clairement que le modèle Holt-Winters se distingue comme le plus performant pour anticiper la production d'électricité globale. Cette conclusion est étayée par l'observation minutieuse des résultats, indiquant que l'erreur quadratique moyenne associée au modèle Holt-Winters est la plus basse, s'établissant à 58 636 322. Cette valeur réduite de l'erreur quadratique moyenne, démontre la capacité du modèle à minimiser les écarts entre les prévisions et les données réelles, validant ainsi son efficacité prédictive. En considérant cette performance remarquable, le choix du modèle Holt-Winters apparaît comme une option privilégiée pour des prévisions précises sur la série temporelle de production d'électricité.

Conclusion

Conclusion général

L'exploration des données, provenant du fichier "AllFuels", a permis de plonger dans la diversité des modes de production d'électricité, allant du charbon aux énergies renouvelables. Notre analyse s'est articulée autour de l'identification des tendances générales, de la décomposition saisonnière, de la désaisonnalisation, et enfin, de la prévision pour l'année 2022.

Les outils statistiques, telles que les moyennes mobiles, la courbe de régression, la décomposition saisonnière, et les modèles de prévision (polynomial, Holt-Winters, ARMA), ont été déployés pour mieux comprendre la dynamique temporelle de la production d'électricité. Les résultats ont révélé des tendances significatives, des variations saisonnières marquées, et des modèles de prévision prometteurs.

La série temporelle a révélé une tendance à long terme, efficacement saisie par la moyenne mobile centrée et confirmée par la courbe de régression des moyennes annuelles. La présence de saisonnalité a été clairement mise en évidence, démontrant des variations cycliques annuelles dans la production d'électricité. La désaisonnalisation a permis de distinguer que la composante saisonnière joue un rôle significatif dans la série temporelle, indiquant une forte influence des variations saisonnières régulières sur la série d'origine. Par ailleurs, une observation importante est la présence de très peu de résidus, suggérant une bonne adéquation du modèle de régression.

Concernant les prévisions pour 2022, les modèles polynomiaux, Holt-Winters et ARMA ont été évalués. Cependant, la détermination du modèle le plus efficace pour les prévisions à long terme sur la production d'électricité demeurerait incertaine. Afin de résoudre cette incertitude, des prévisions ont été réalisées pour l'année 2021, permettant une comparaison avec les observations réelles. Il en est ressorti que le modèle Holt-Winters s'est avéré le plus conforme et performant pour anticiper la production d'électricité globale. Cette conclusion est étayée par une erreur quadratique moyenne plus faible par rapport aux autres modèles, justifiant ainsi son adaptation optimale à la série temporelle étudiée.

En conclusion, cette analyse offre une perspective approfondie sur la production d'électricité, fournissant des informations cruciales pour les décideurs, les chercheurs, et le public. La combinaison de méthodes statistiques avancées et de modèles de prévision a permis de mieux appréhender les tendances passées et d'anticiper de manière informée les évolutions futures du secteur énergétique.

Résumer en anglais

In-depth analysis of electricity production data extracted from the "AllFuels" file has provided interesting insights into the diversity of production methods. Using various statistical tools such as moving averages, regression curves, seasonal decomposition, and the evaluation of forecasting models such as Holt-Winters, we were able to reveal general trends and anticipate future developments in the sector.

The emphasis on the Holt-Winters model, with a significantly lower mean square error, underlines its usefulness for accurate forecasts of electricity production. These results provide essential insights for stakeholders in the energy sector.

In summary, this study highlights the effectiveness of advanced time series analysis in the field of power generation. It provides valuable information on past trends while positioning itself as a crucial tool for anticipating future developments. The application of statistical methods and forecasting models are used to understand the complexity of electricity generation and guide future actions towards sustainability and energy efficiency.