

PRÉDICTION DE LA DEMANDE EN ÉLECTRICITÉ

Création d'un modèle permettant de déterminer la demande en électricité

SOMMAIRE

Partie I

Contexte et enjeux

Partie II

Présentation des données utilisées

Partie III

Analyse des données téléchargées

Partie IV

Création de modèles de prédiction

CONTEXTE ET ENJEUX

demande en électricité

CONTEXTE ET ENJEUX

Le projet ici nous permet de nous mettre dans la peau d'un employé d'Enercoop.

Société coopérative qui s'est développée grâce à la libéralisation du marché de l'électricité en France. Elle est spécialisée dans les énergies renouvelables.

MISSION Prédire la demande en électricité

DIFFICULTÉS La demande en électricité dépend de plusieurs paramètres
Éviter les coupures électriques

DÉMARCHE ET DONNÉES

présentation des données utilisées

DONNÉES

02

DONNÉES ET DÉMARCHES

Nous disposons de données téléchargées principalement via le site RTE et Cegibat.

RTE France (Gestionnaire du Réseau de Transport d'Électricité)

années de référence

2012 - 2021

+

CEGIBAT

région de référence

PARIS

VARIABLES

Territoire : France

Consommation : consommation totale
(GWh)

Mois : du 01/01/2012 au 01/06/2021

DJU chauffage : Degrés jours unifiés
(Paris)

DJU climatisation : Degrés jours unifiés
(Paris)

Échelle
régionale

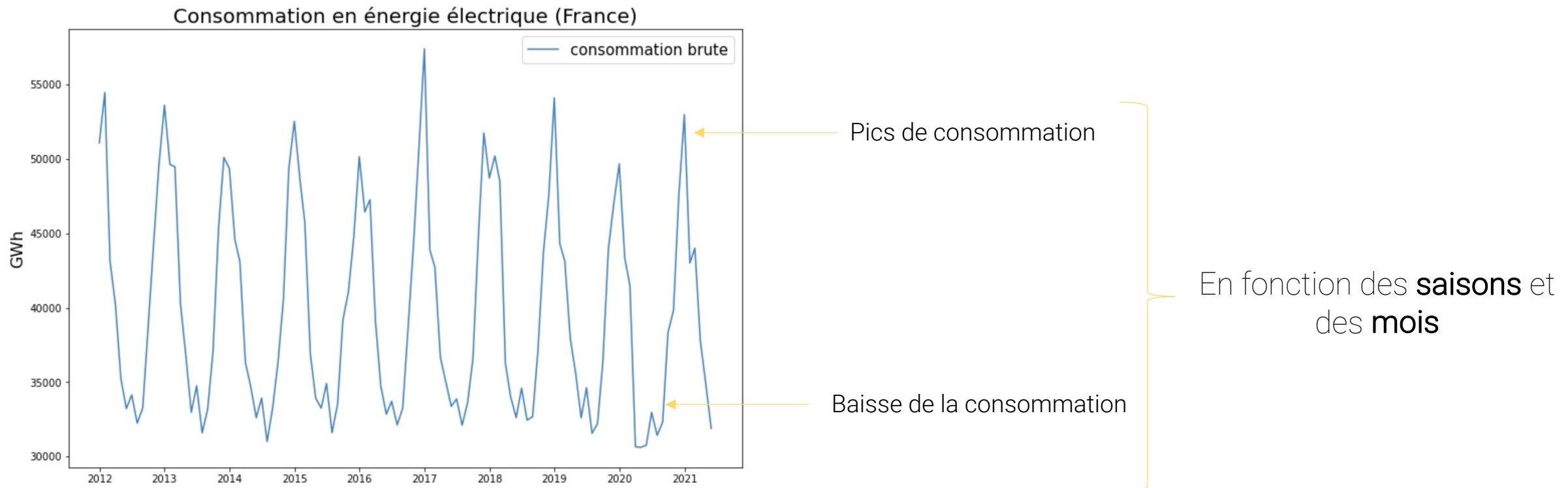
ANALYSE

analyse des variables

ANALYSE DES DONNÉES

Nous réalisons une analyse de la consommation en énergie électrique (chauffage) en France.

Consommation en énergie électrique consommation brute en France



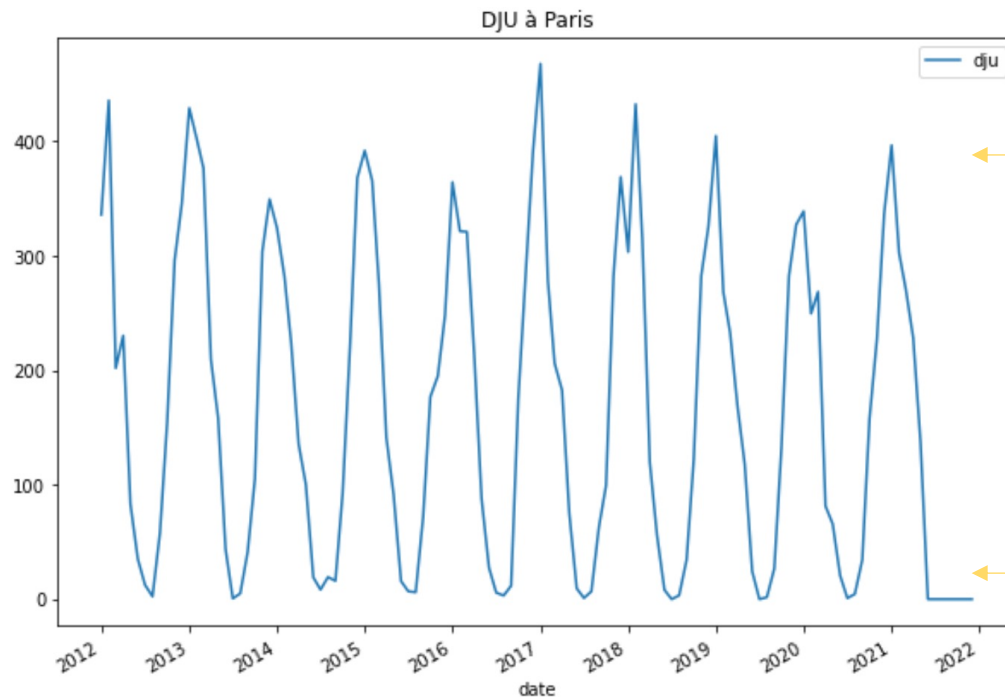
ANALYSE 03

ANALYSE DES DONNÉES

Nous réalisons une analyse des DJU à Paris.

Degrés Jours Unifiés consommation de chauffage à Paris

Le **degré jour unifié (DJU)** est la différence entre la température extérieure et une température de référence qui permet de réaliser des estimations de consommations d'énergie et de quantifier les dépenses d'énergie nécessaires.



Pics

Baisse

En fonction des **saisons** et
des **mois**

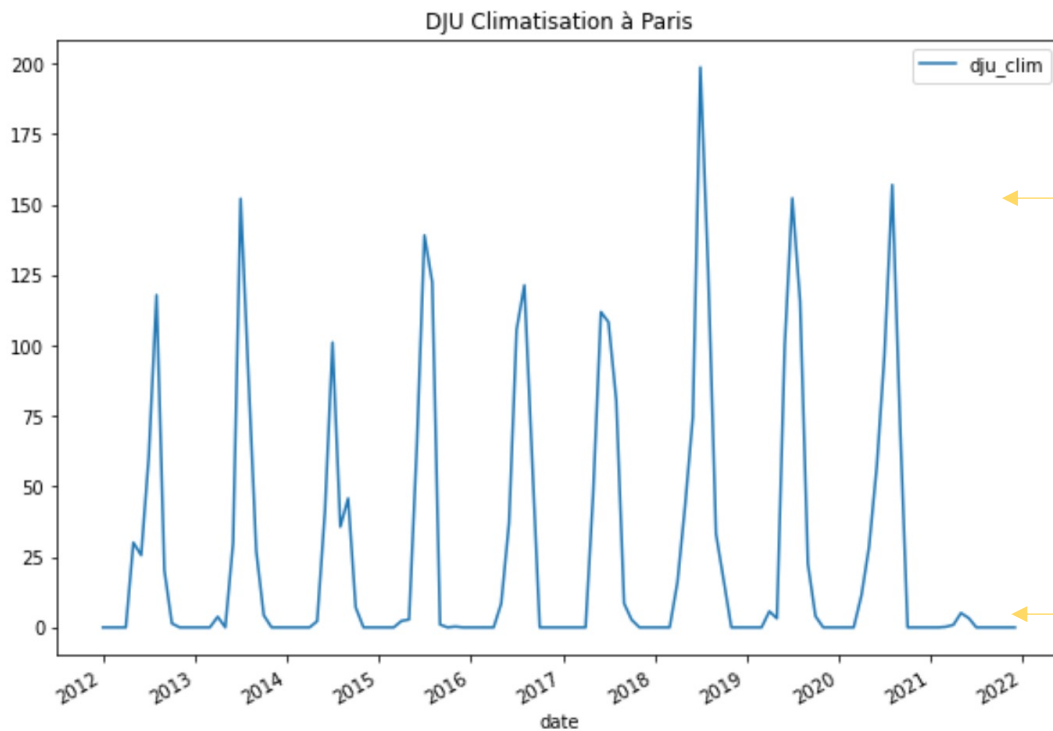
hiver

ANALYSE 03

ANALYSE DES DONNÉES

Nous réalisons une analyse des DJU à Paris.

Degrés Jours Unifiés consommation de climatisation à Paris



Pics

Baisse

En fonction des **saisons** et
des **mois**

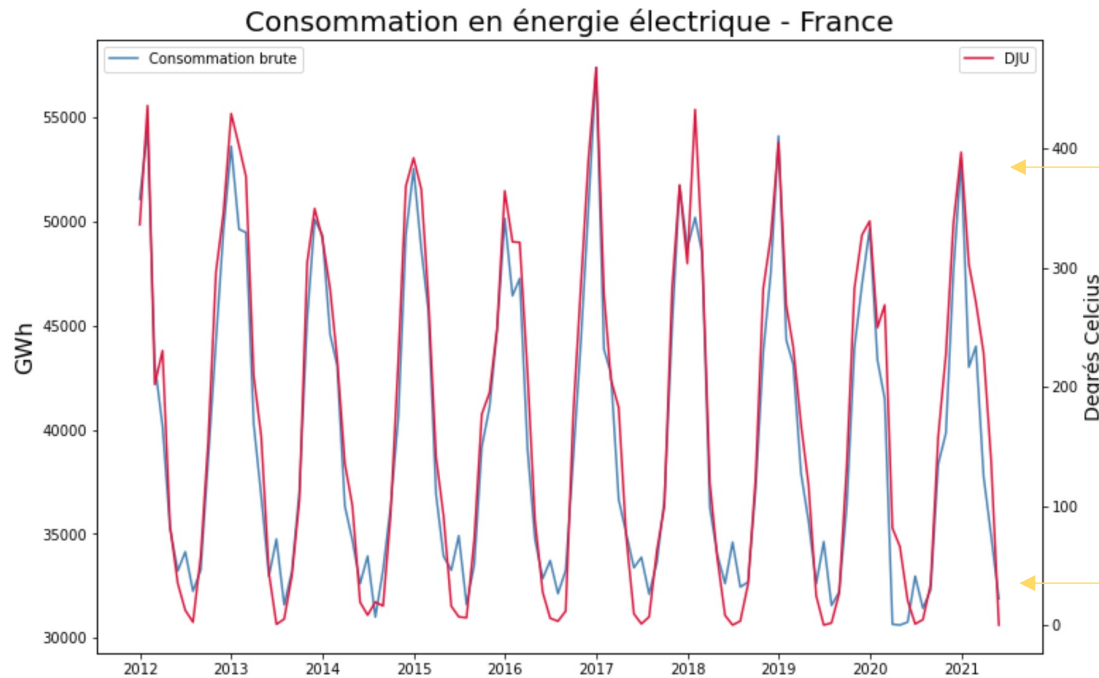
été

ANALYSE 03

ANALYSE DES DONNÉES

Nous réalisons une comparaison entre la consommation en énergie électrique en France et les DJU chauffage.

Comparaison consommation de chauffage à Paris et d'électricité



Pics de consommation

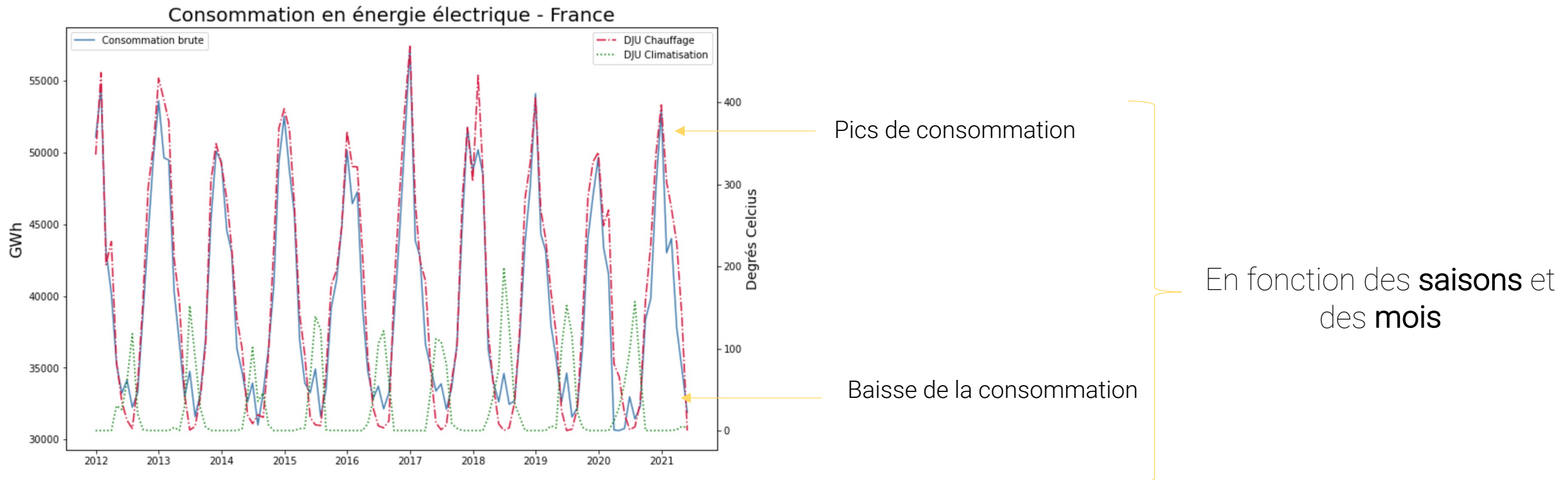
Baisse de la consommation

En fonction des **saisons** et
des **mois**

ANALYSE DES DONNÉES

Nous réalisons une comparaison entre la consommation en énergie électrique en France et les DJU chauffage et climatisation.

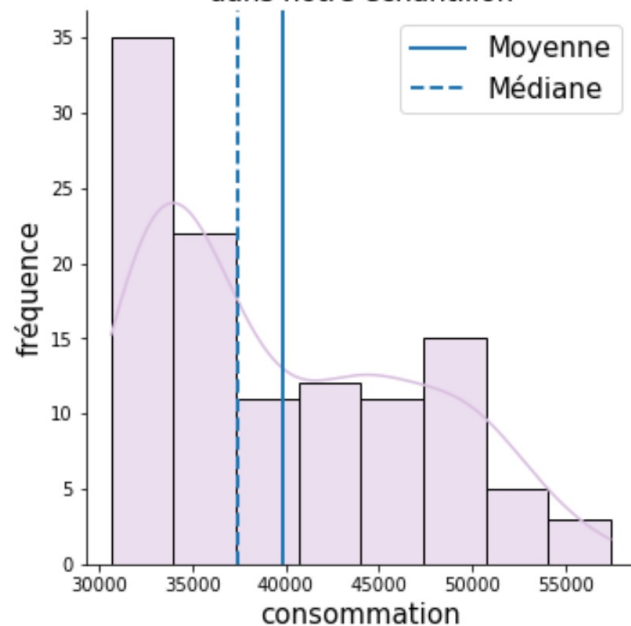
Comparaison consommation de chauffage, de climatisation et d'électricité



Nous réalisons une comparaison entre la consommation en énergie électrique en France et les DJU chauffage et climatisation.

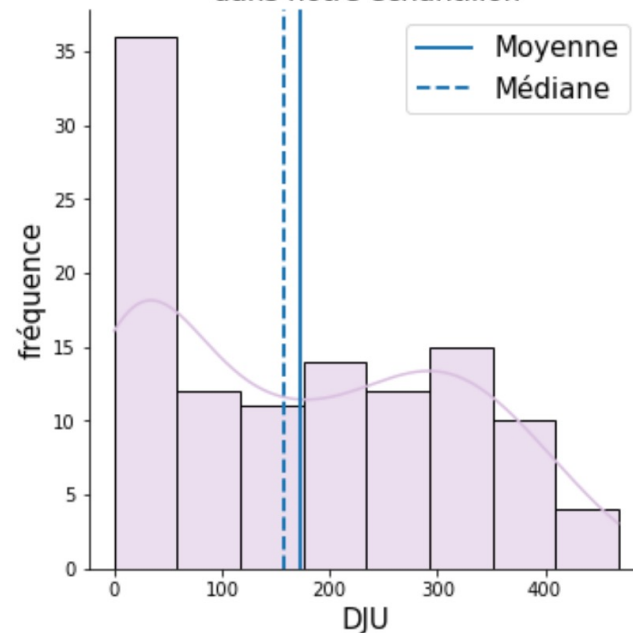
Répartition de nos données consommation de chauffage, de climatisation et d'électricité

Répartition de la variable consommation dans notre échantillon

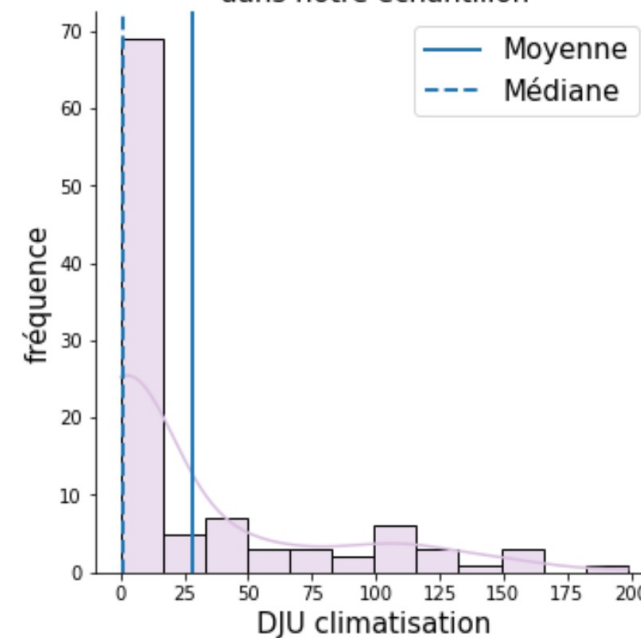


Les variables ne suivent pas
une loi normale
(Kolmogorov-Smirnov)

Répartition de la variable DJU dans notre échantillon



Répartition de la variable DJU clim dans notre échantillon

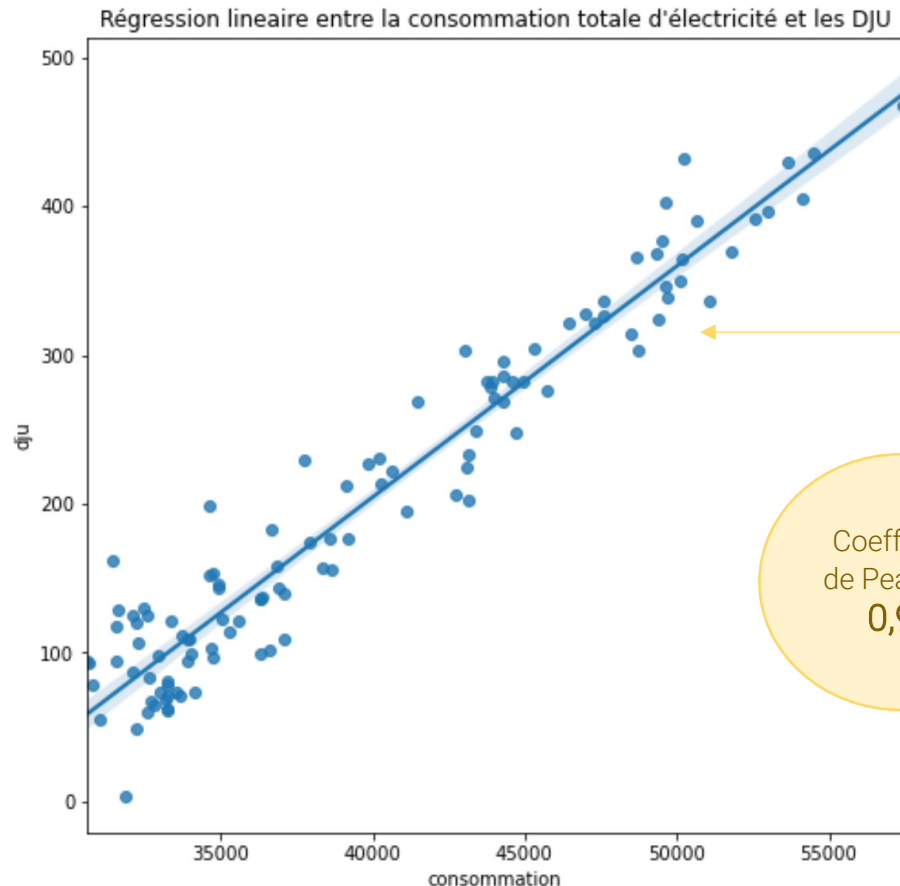


ANALYSE 03

ANALYSE DES DONNÉES

Nous réalisons une analyse de la relation entre la consommation en énergie électrique en France et les DJU chauffage.

Relation entre nos données dju chauffage à Paris et d'électricité



Coefficient
de Pearson :
0,97

On observe la droite
d'ajustement.

Les points *suivent* la **droite de régression** et couvrent bien l'étendue entière des valeurs. On constate donc une relation entre nos deux variables.

D'après notre modèle, les DJU expliquent environ **94%** de la variance de la consommation électrique.

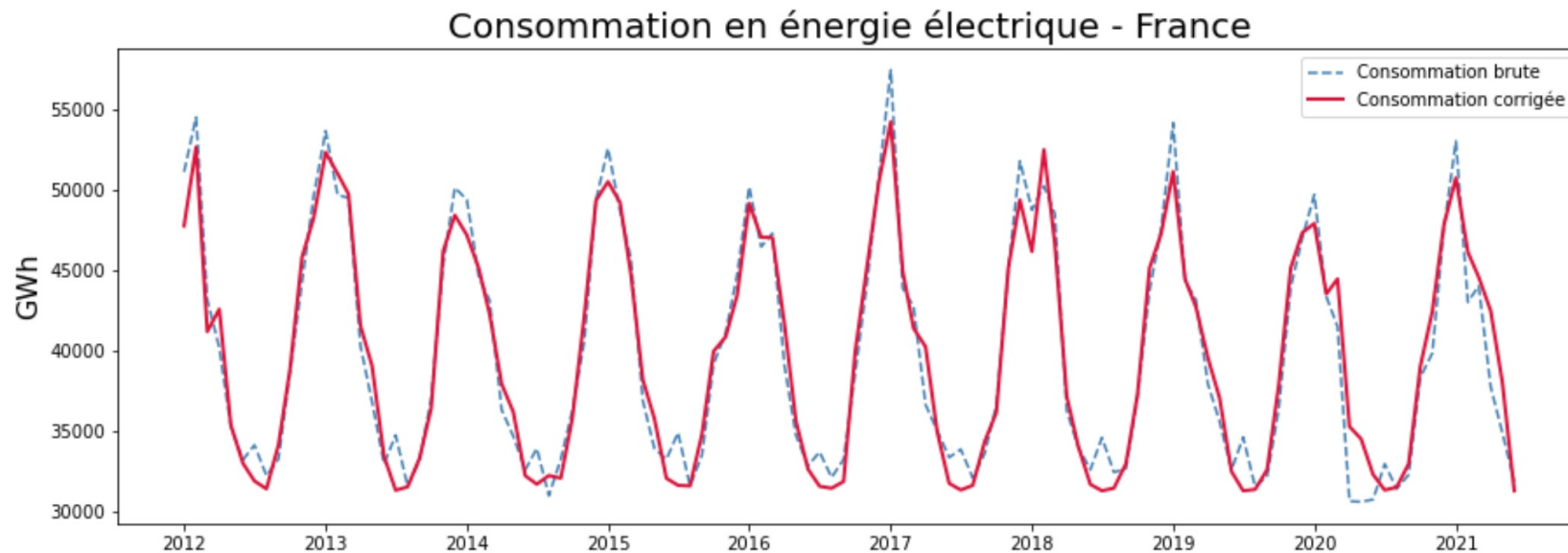
ANALYSE 03

ANALYSE DES DONNÉES

Nous réalisons une correction de consommation de l'effet température en utilisant une régression linéaire.

Correction de l'effet température grâce au DJU chauffage

Consommation corrigée =>
obtenue grâce à la régression linéaire « *.predict* »



On constate que la consommation est **améliorée**. Cela a permis d'enlever du bruit dans le signal initial.

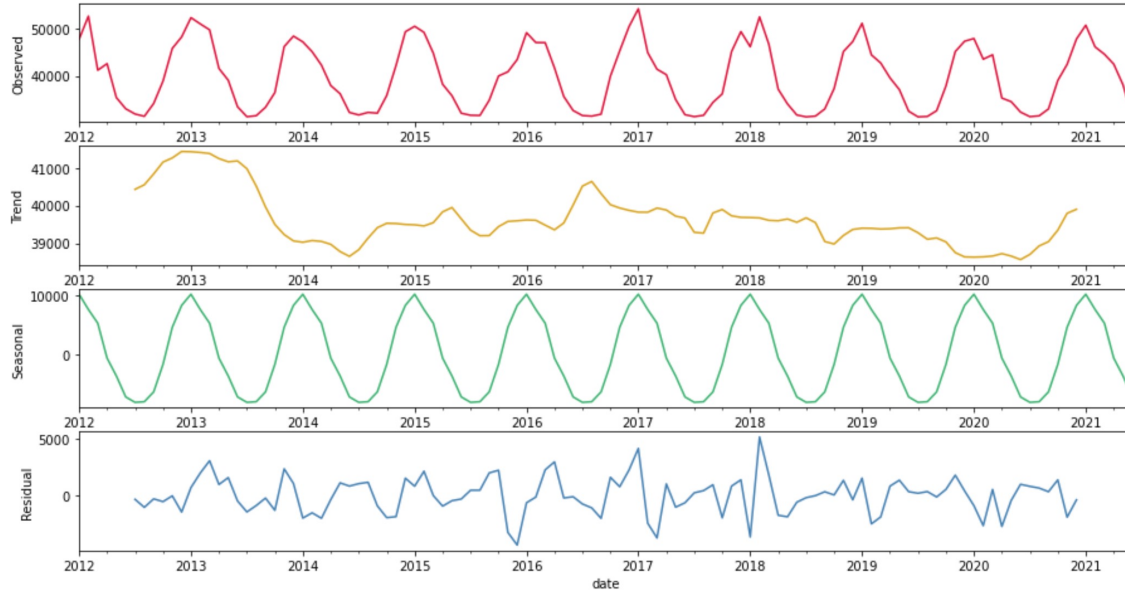
ANALYSE 03

ANALYSE DES DONNÉES

Nous réalisons une désaisonnalisation de la consommation grâce aux moyennes mobiles.

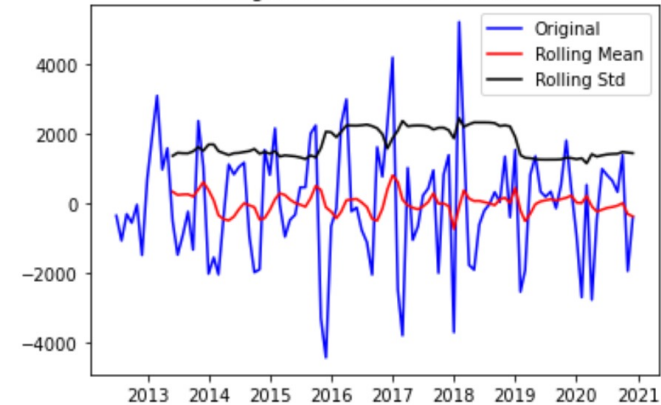
I. Désaisonnalisation de la consommation décomposition de la série temporelle

Décomposition saisonnière de la consommation corrigée



Consommation désaisonnalisée =
consommation corrigée – saisonnalité

Rolling Mean & Standard Deviation



Results of Dickey-Fuller Test:
Test Statistic -6.397172e+00
p-value 2.037264e-08

Notre série est
stationnaire.

ANALYSE 03

ANALYSE DES DONNÉES

Nous réalisons une dessaisonnalisation de la consommation grâce aux moyennes mobiles.

II. Désaisonnalisation de la consommation soustraction des moyennes mobiles

```
# On soustrait les moyennes mobiles de notre consommation corrigée
data["diff"] = data['conso_corr'] - data['MM']

# Calcul des moyennes mobiles sur notre consommation corrigée
data['MM'] = data['conso_corr'].rolling(window=12).mean() # 12 mois
```

ADF Statistic: -7.808870583541389
p-value: 7.158549139339587e-12

Notre série est
stationnaire.

III. Désaisonnalisation de la consommation décroissance exponentielle

```
# On soustrait les moyennes mobiles exponentielles de la consommation corrigée
data['diff_ema'] = data['conso_corr'] - data['conso_corr'].ewm(span=12).mean() # 12 mois
```

ADF Statistic: -7.618862984154269
p-value: 2.1551054428282957e-11

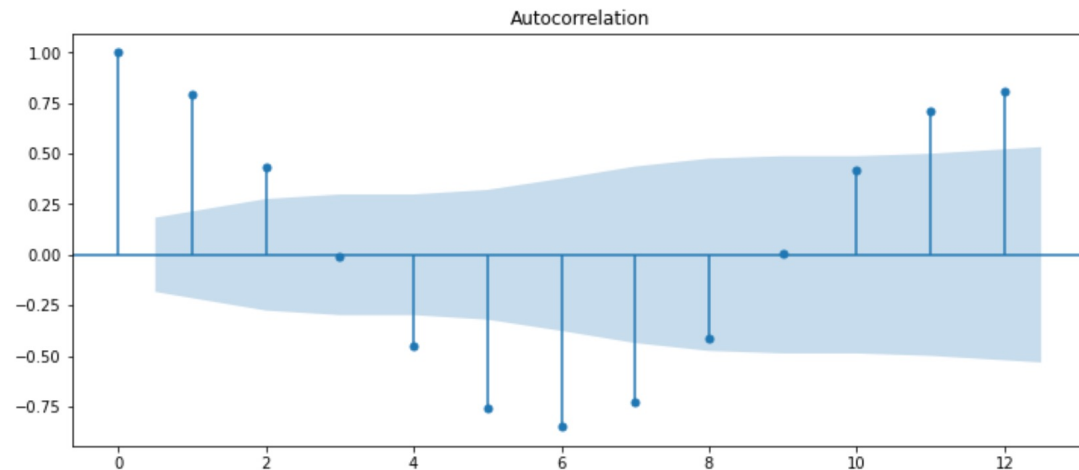
Notre série est
stationnaire.

ANALYSE 03

ANALYSE DES DONNÉES

Nous réalisons une dessaisonnalisation de la consommation grâce aux moyennes mobiles.

IV. Désaisonnalisation de la consommation décalage temporel



Série différenciée I-B¹²

Pvalue du test de Dickey-Fuller: $5.192108824766355e-07$
Statistique de test: -5.778482475877935

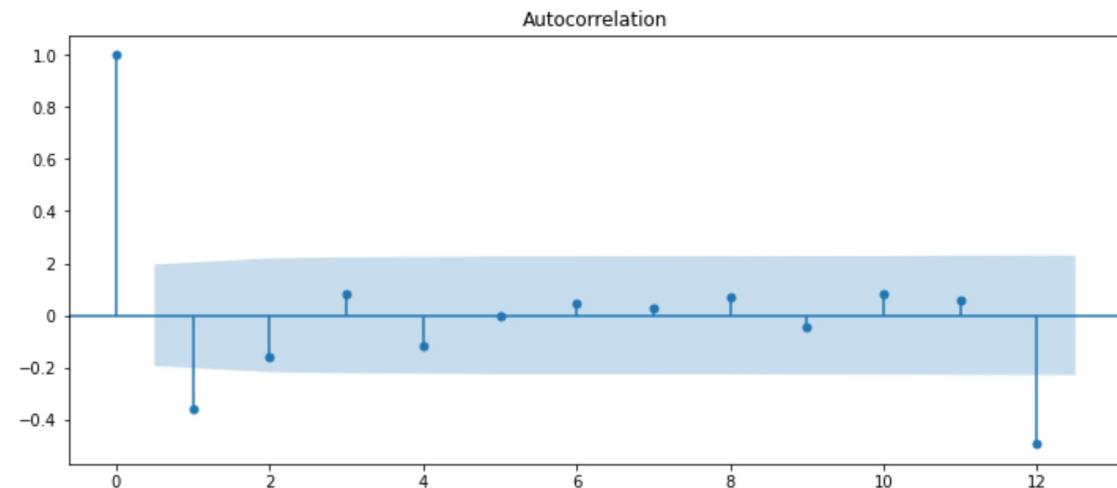
Notre série est
stationnaire.

Série non différenciée

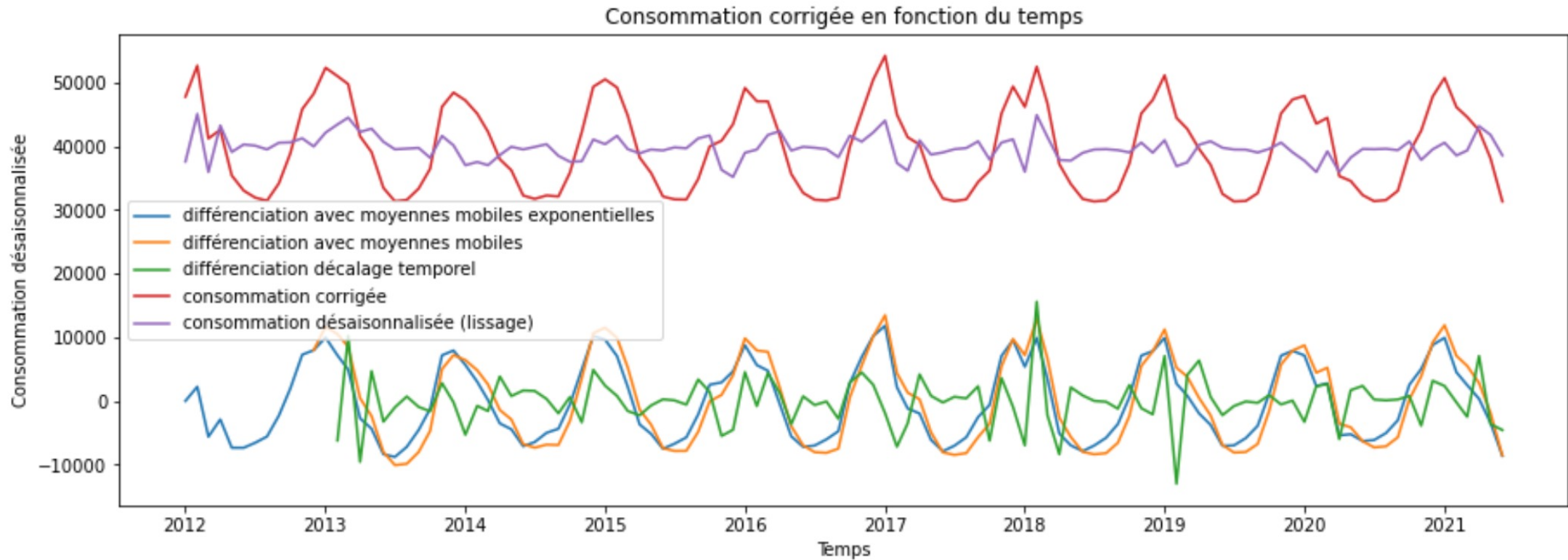
Test Statistic
p-value

-2.147588
 0.225815

Notre série n'est
pas stationnaire.



V. Comparaison des méthodes de désaisonnalisation de la consommation



MODÈLES

création de modèles de prédiction : Holt-Winters et SARIMA

MODÈLE 04

PRÉDICTION : MÉTHODE HOLT-WINTERS

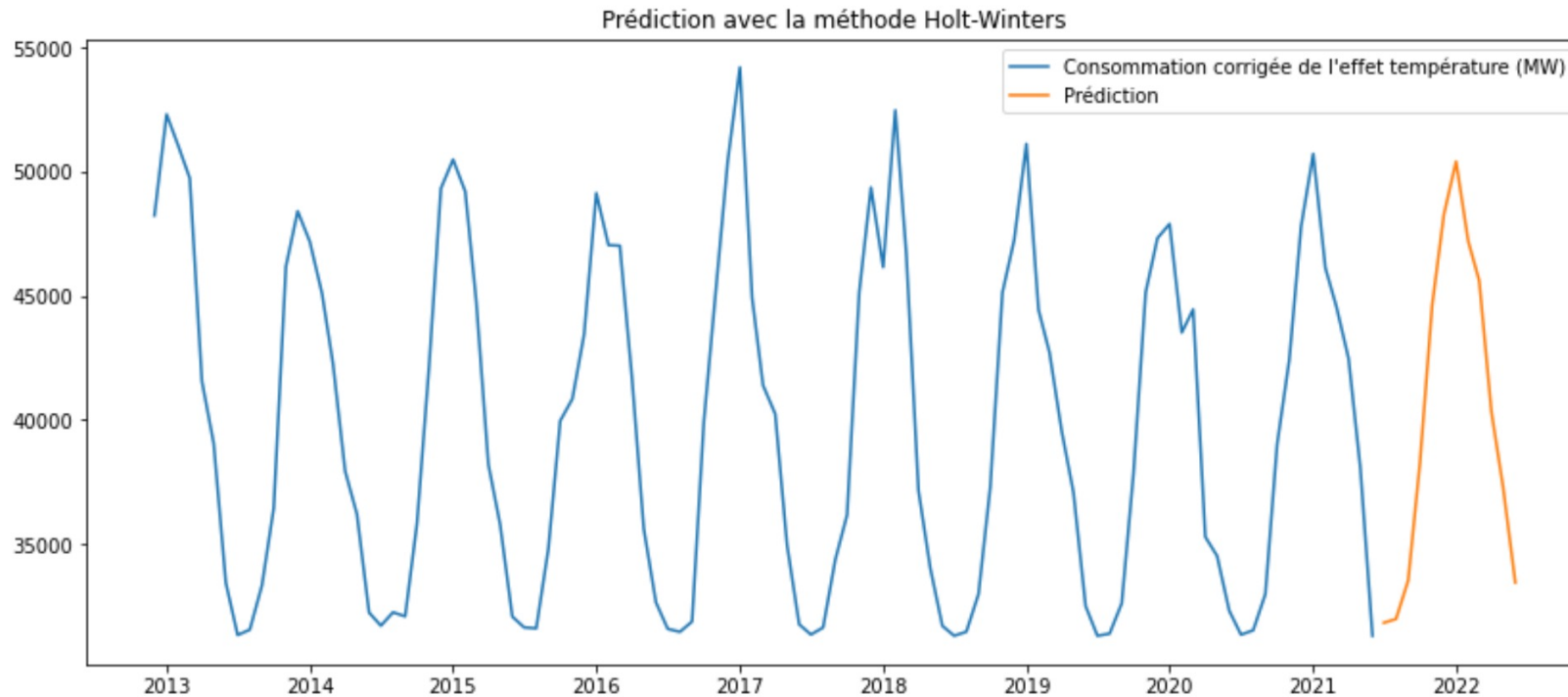
Nous allons prédire la demande en électricité sur les 12 prochains mois grâce à la méthode Holt-Winters.

```
hw = ExponentialSmoothing(np.asarray(data["consommation_corrige"]),  
                           seasonal_periods=12, trend="add", seasonal="add").fit()  
  
hw_pred = hw.forecast(12)
```

← modèle

← prédiction

Prédiction sur
12 mois

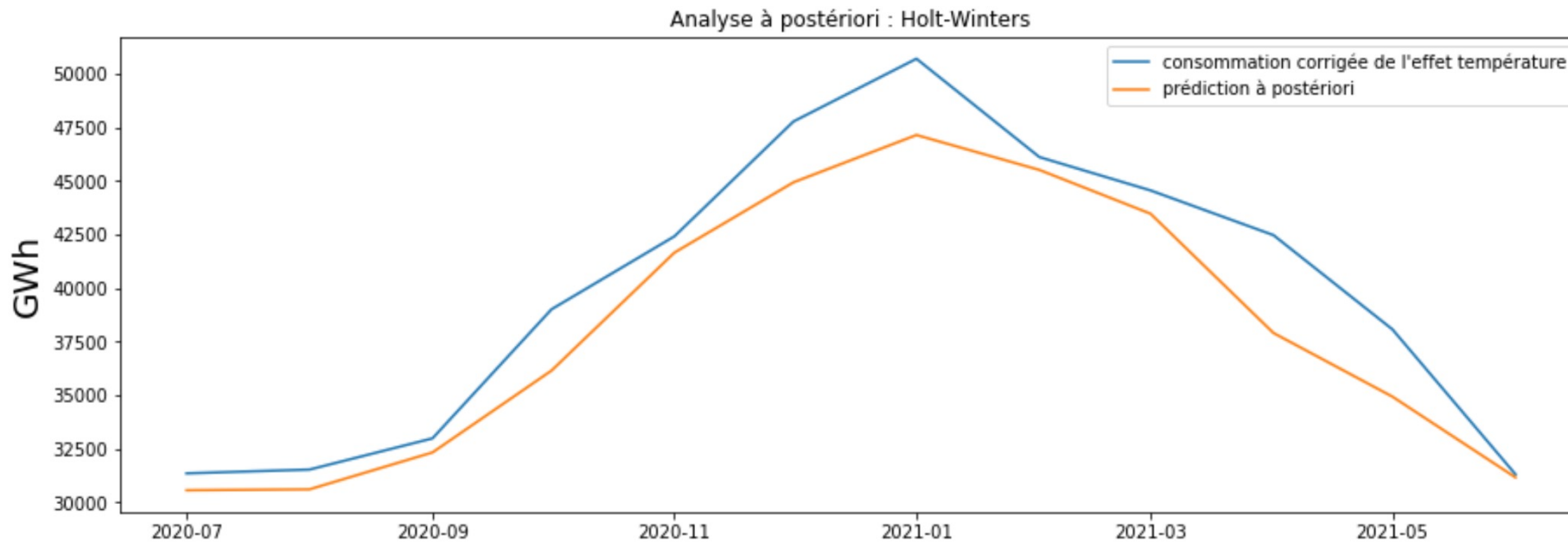


MODÈLE 04

PRÉDICTION : MÉTHODE HOLT-WINTERS

Nous allons prédire la demande en électricité sur les 12 prochains mois grâce à la méthode Holt-Winters.

Analyse à postériori comparaison entre nos données et la prédiction que le modèle aurait effectuée



RMSE :
2304.17 GWh
MAPE :
4,39%

MODÈLE 04

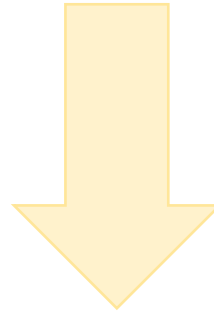
PRÉDICTION : MÉTHODE SARIMA

Nous allons prédire la demande en électricité sur les 12 prochains mois grâce à la méthode SARIMA.

Meilleur modèle SARIMA recherche automatique des paramètres pour obtenir le meilleur modèle

```
# on détermine les meilleurs paramètres
```

```
from pmdarima import auto_arima
stepwise_model=auto_arima(data["consommation_corrige"], m=12, D=1, seasonal=True, trace=True, error_action="ignore",
                           suppress_warnings=True, stepwise=True, information_criterion="aic")
print(stepwise_model.aic())
```



Meilleur modèle
SARIMA (3,0,2)(2,1,0)[12]

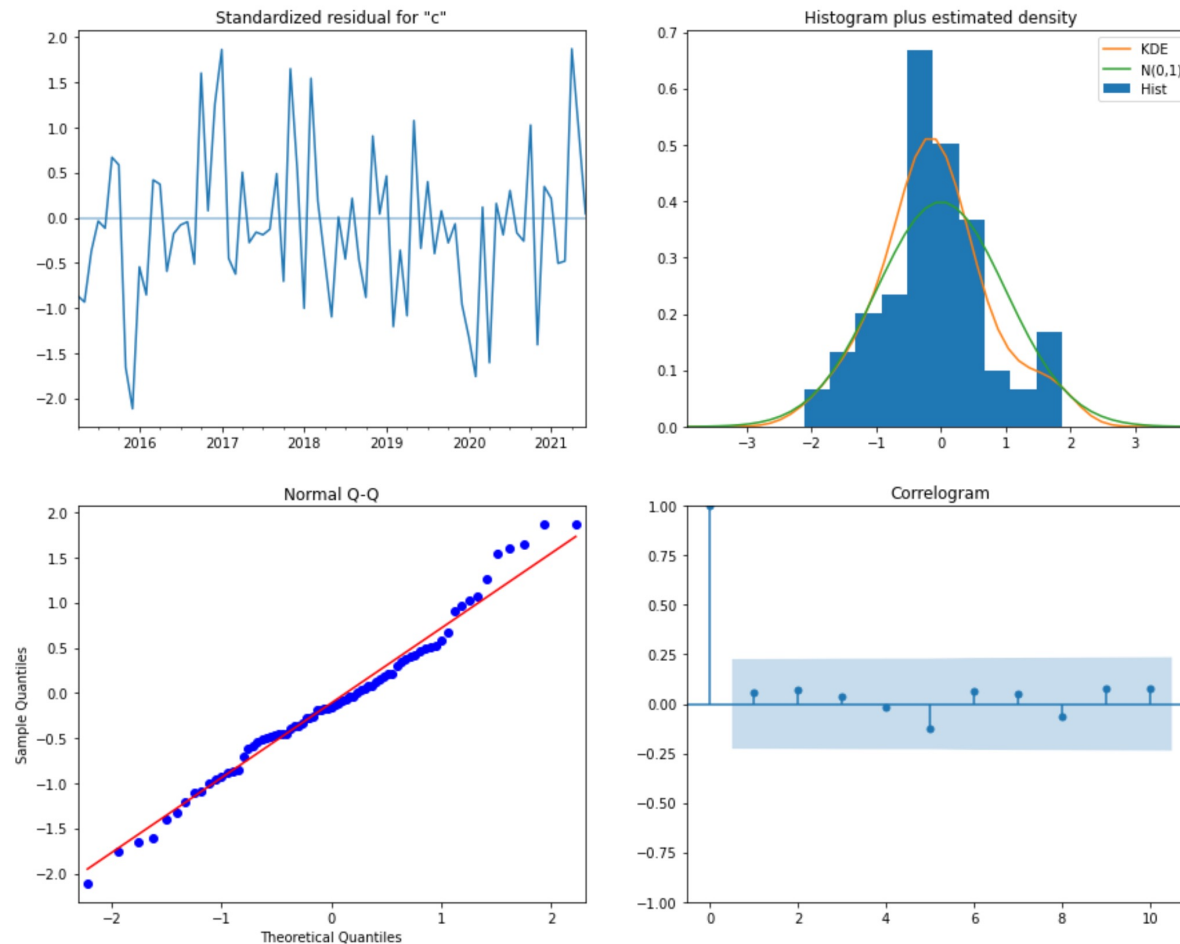
Le score A/C doit être le plus faible.

MODÈLE 04

PRÉDICTION : MÉTHODE SARIMA

Nous allons prédire la demande en électricité sur les 12 prochains mois grâce à la méthode SARIMA.

Modèle SARIMA



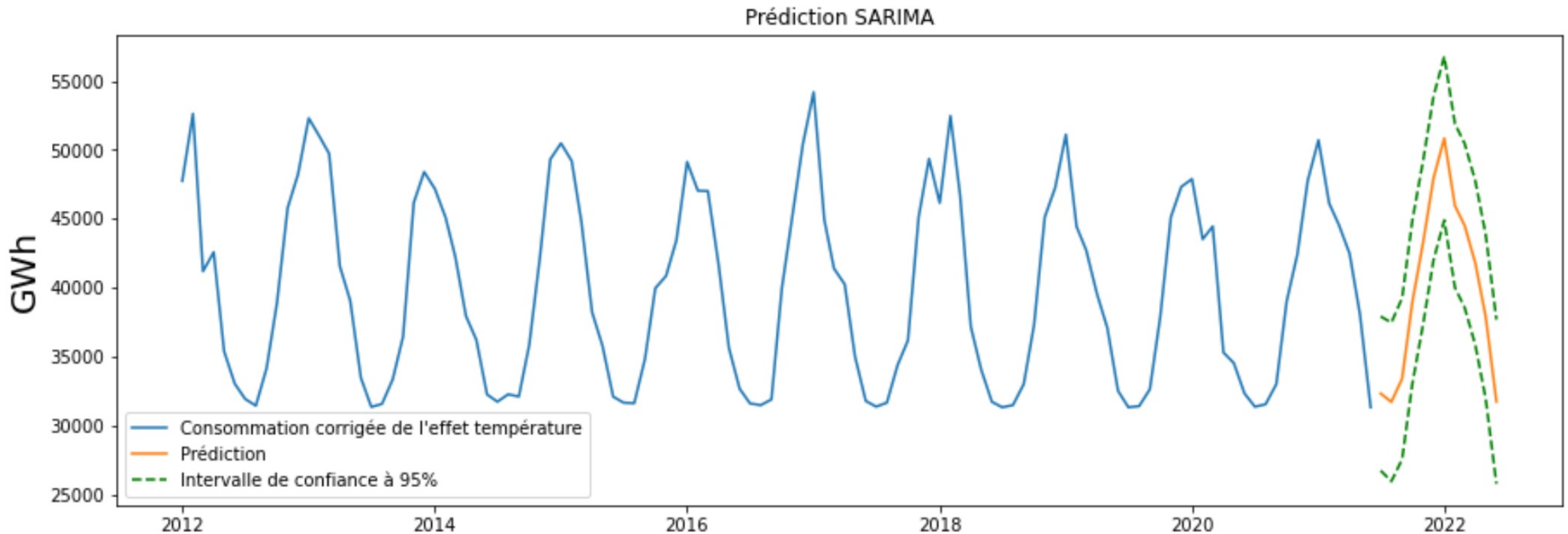
On constate qu'il n'y a **pas d'autocorrélation** dans les résidus et qu'ils **suivent une loi normale** ce qui *valide* notre modèle de régression.

MODÈLE 04

PRÉDICTION : MÉTHODE SARIMA

Nous allons prédire la demande en électricité sur les 12 prochains mois grâce à la méthode SARIMA.

Prédiction SARIMA

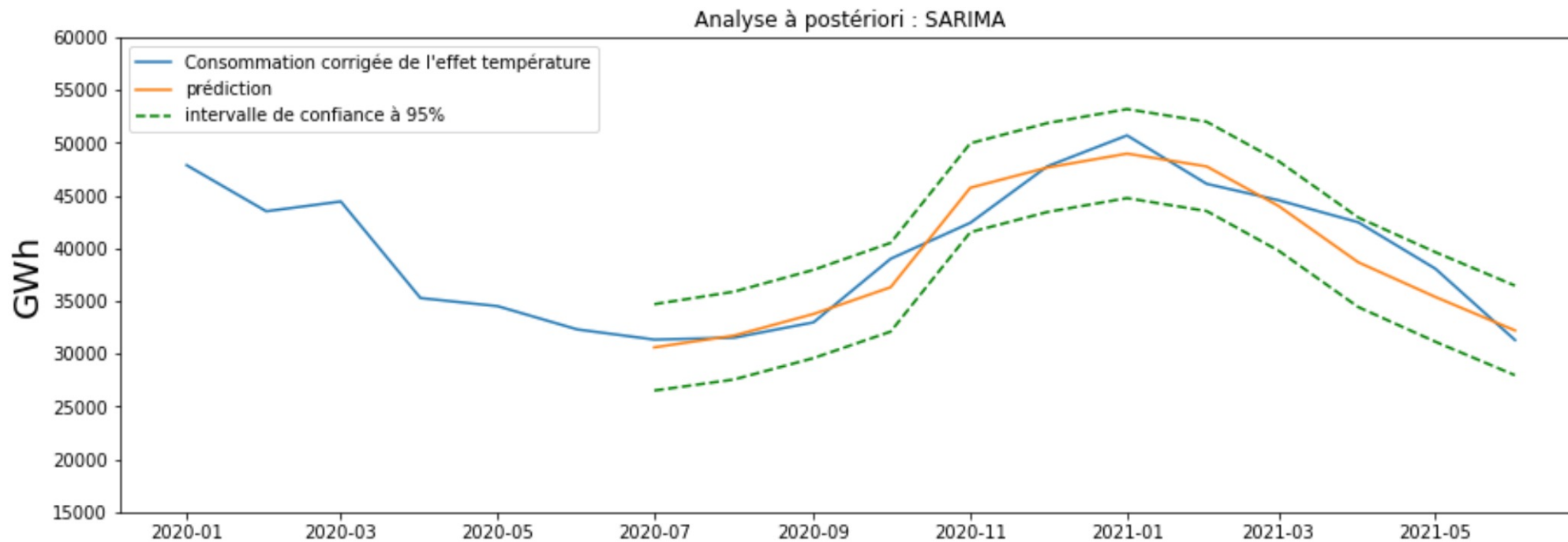


MODÈLE 04

PRÉDICTION : MÉTHODE SARIMA

Nous allons prédire la demande en électricité sur les 12 prochains mois grâce à la méthode SARIMA.

Analyse à postériori comparaison entre nos données et la prédiction que le modèle aurait effectuée



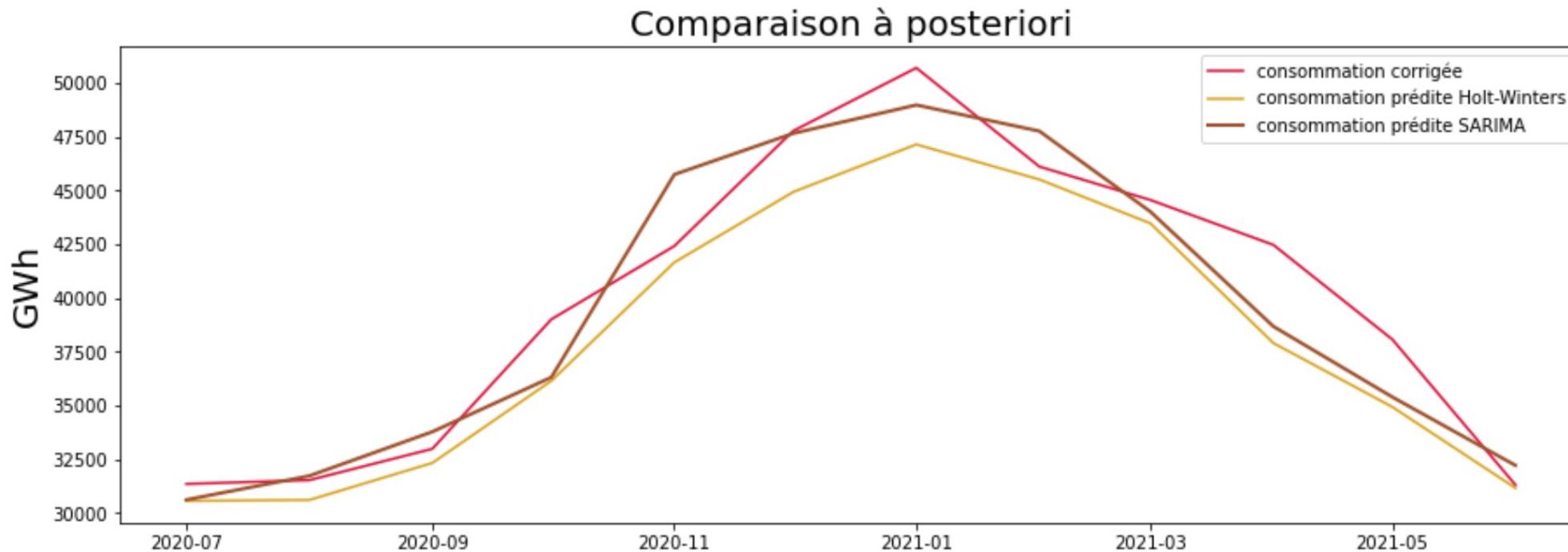
RMSE :
1999.12 GWh
MAPE :
3,95%

MODÈLE 04

PRÉDICTION : MÉTHODE HOLT-WINTERS ET SARIMA

Nous avons prédit la demande en électricité sur les 12 prochains mois grâce aux méthodes Holt-Winters et SARIMA.

Comparaison des analyses à postériori comparaison entre nos données et la prédiction que le modèle aurait effectuée



RMSE pour Holt-Winters :
2304.16 GWh

RMSE pour SARIMA :
1999.13 GWh

MODÈLE

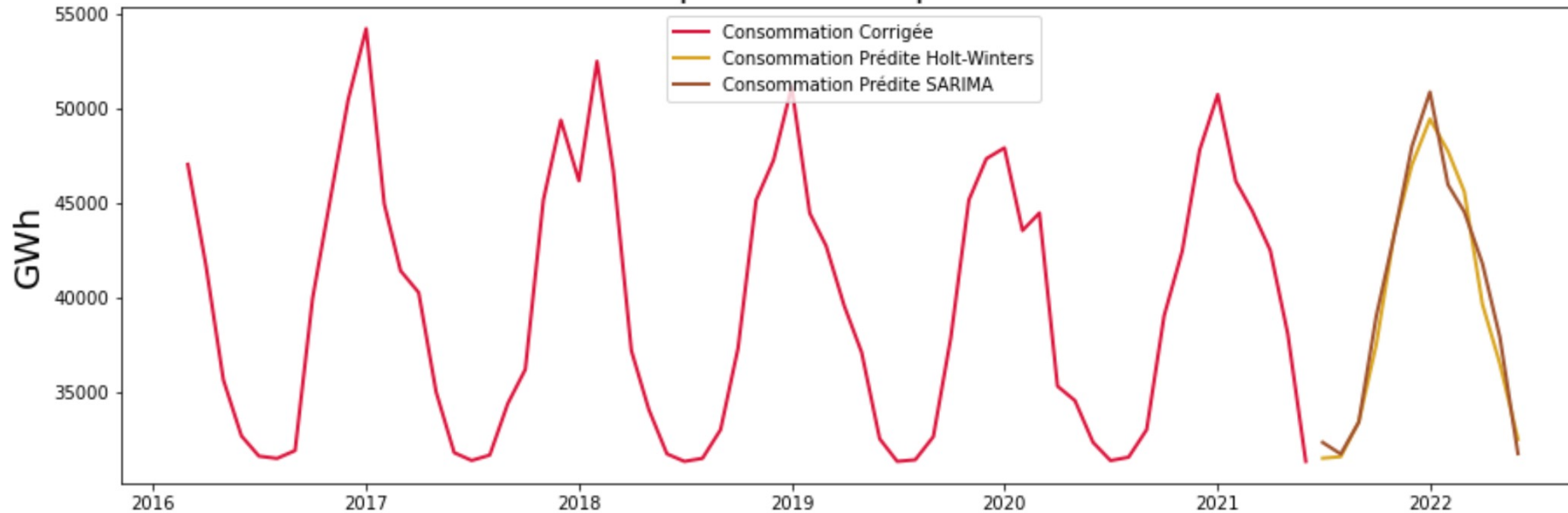
04

PRÉDICTION : MÉTHODE HOLT-WINTERS

Nous allons donc prédire la demande en électricité sur les 12 prochains mois grâce à la méthode Holt-Winters ainsi que SARIMA.

Comparaison des prédictions comparaison entre les modèles Holt-Winters et SARIMA

Comparaison des prédictions



La méthode SARIMA sur-estime la consommation.

BILAN 04

BILAN

Nous avons pu prédire la demande en électricité sur les 12 prochains mois grâce à la méthode la plus adéquate.

La consommation d'énergie dépend de la **température extérieure** : chauffage quand il fait froid, climatisation quand il fait chaud. Lorsque l'on analyse les évolutions annuelles de la consommation, on peut souhaiter neutraliser ce facteur exogène et donc effectuer une correction de l'effet température. Cette correction est fondée sur la notion de « *degrés-jours unifiés* » (DJU).

Le choix du modèle de prédiction s'est porté sur **SARIMA** : un taux RMSE *moins élevé* mais également la méthode qui correspond au mieux à une réflexion business. En effet, SARIMA sur-estime les GWh à prévoir ce qui évite *d'être confronté à d'éventuelles coupures électriques* au cours de l'année et surtout à la saison la plus importante, **l'hiver**.

MERCI

Anissa MANSOUR

parcours Data Analyst (2020/2021)