Enercoop

Prédiction de consommation électrique

Plan

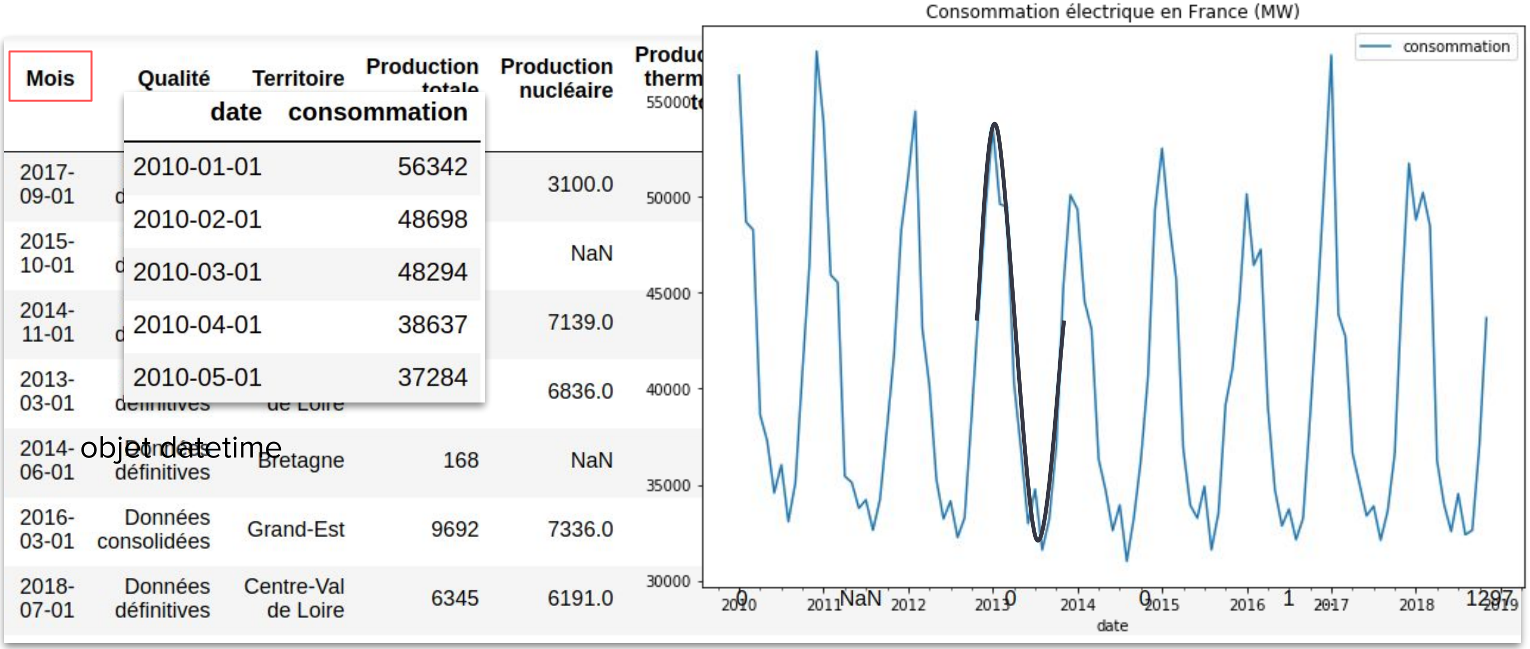
- Préparation des données
- Consommation corrigée de l’effet température
- Désaisonnalisation
- Prédiction
  - ➔ Méthode Holt-winter
  - ➔ Méthode SARIMA
- Conclusion

Plan

- **Préparation des données**
- Consommation corrigée de l’effet température
- Désaisonnalisation
- Prédiction
  - ➜ Méthode Holt-winter
  - ➜ Méthode SARIMA
- Conclusion

Préparation - Consommation d'électricité

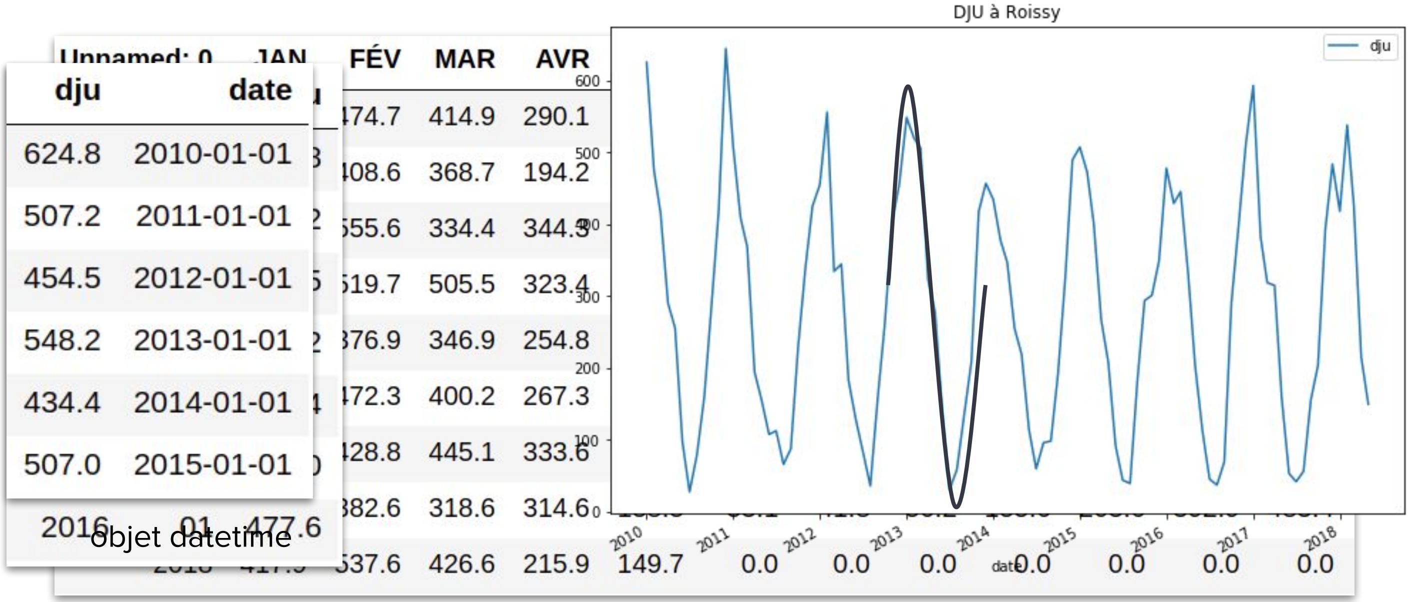
- Source : RTE - Du 01/01/10 au 01/11/18



date	consommation
2010-01-01	56342
2010-02-01	48698
2010-03-01	48294
2010-04-01	38637
2010-05-01	37284

Préparation - DJU

- DJU quotidien et mensuel *(définition)*
- Source : GRDF - Roissy - Du 01/01/10 au 01/05/18



dju	date
624.8	2010-01-01
507.2	2011-01-01
454.5	2012-01-01
548.2	2013-01-01
434.4	2014-01-01
507.0	2015-01-01

Préparation - Jointure

Jointure interne

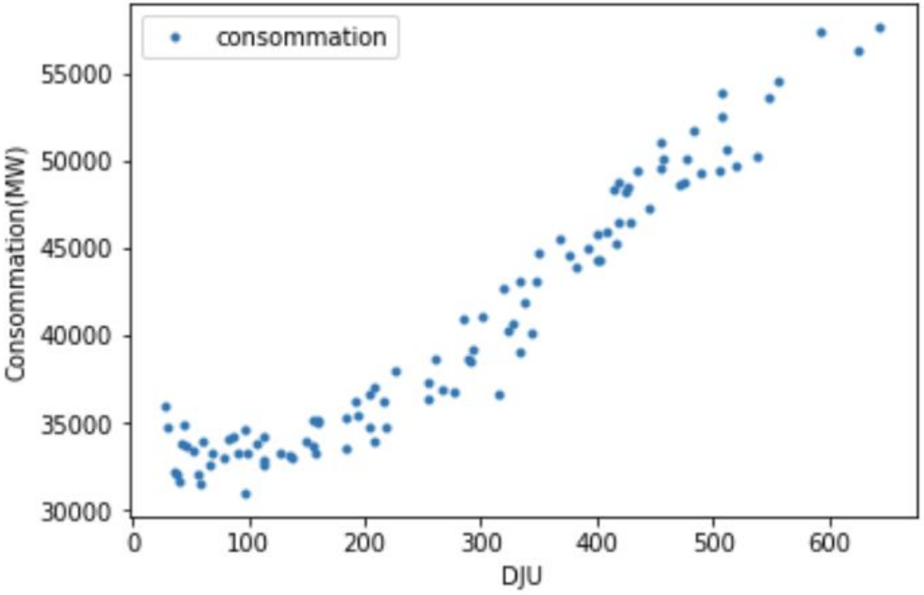
Seulement les **dates** contenant les **deux valeurs**

date	dju	consommation
2010-01-01	624.8	56342
2010-02-01	474.7	48698
2010-03-01	414.9	48294
2010-04-01	290.1	38637
2010-05-01	255.4	37284

Plan

- Préparation des données
- **Consommation corrigée de l’effet température**
- Désaisonnalisation
- Prédiction
  - ➔ Méthode Holt-winter
  - ➔ Méthode SARIMA
- Conclusion

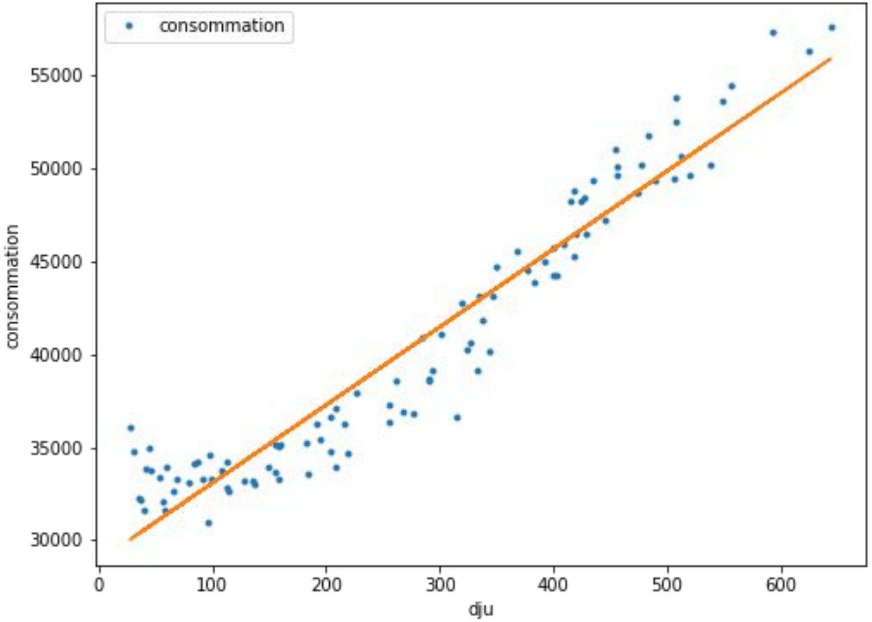
Consommation liée à l'effet température ?



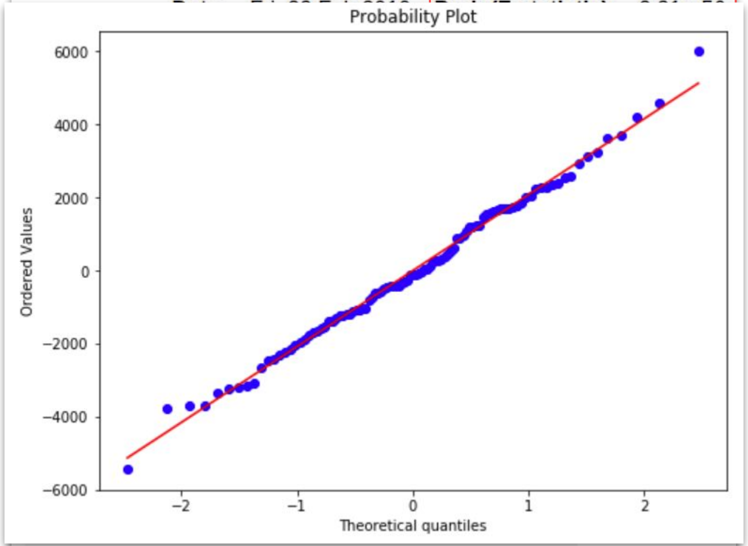
→ On peut retrancher l'effet température

Régression linéaire

→ **OLS de statsmodel**



Dep. Variable:	consommation	**R-squared:**	0.921
Model:	OLS	**Adj. R-squared:**	0.920
Method:	Least Squares	**F-statistic:**	1154.



Skew:	0.141	**Prob(JB):**	0.845
Kurtosis:	3.017	**Cond. No.**	635.

Régression linéaire - Variation expliquée

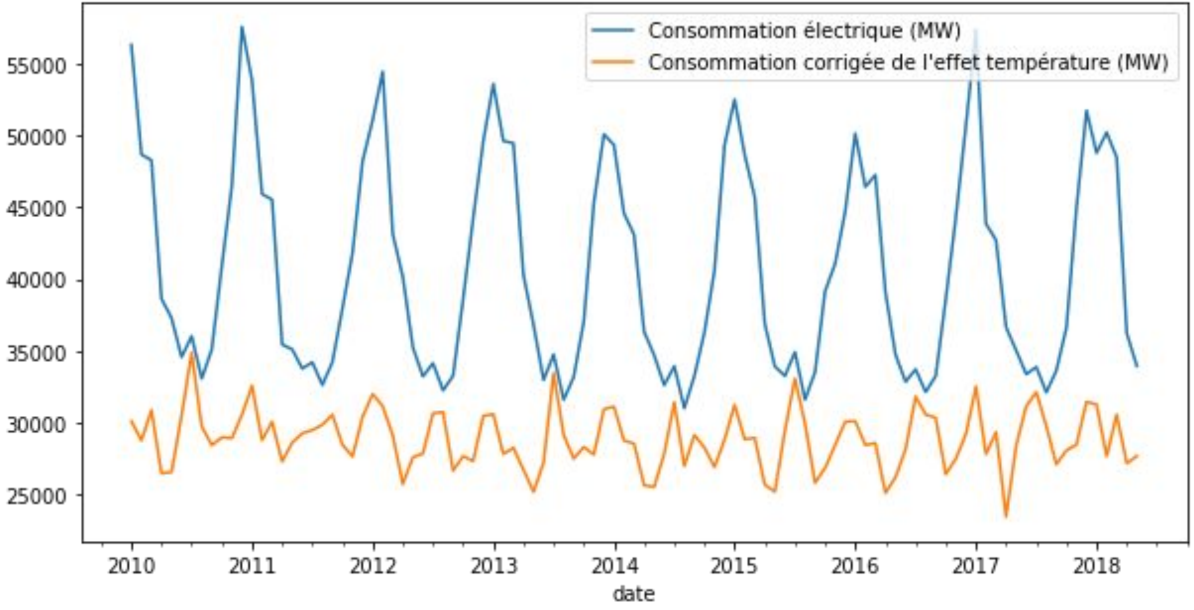
$$R^2 = 0.92$$

L'**effet température** explique
la **majeur partie** de la **consommation électrique**

8% de la **consommation électrique** est utilisé pour **autre chose**

Correction de l’effet température

consommation corrigée **=** *consommation* **-** *DJU* **x** *coefficient(DJU)*

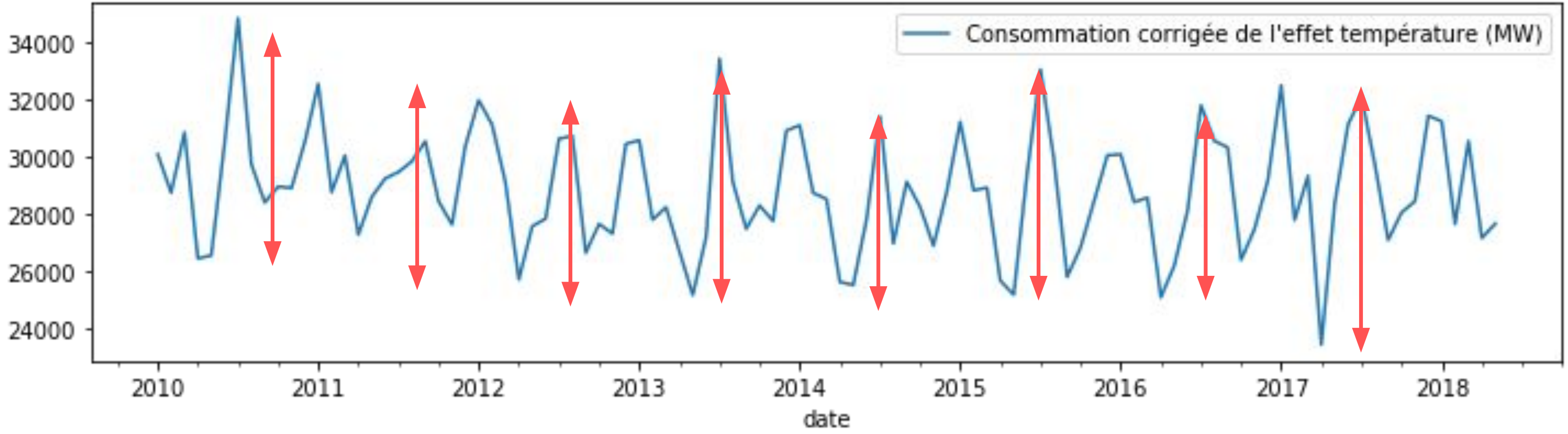


Plan

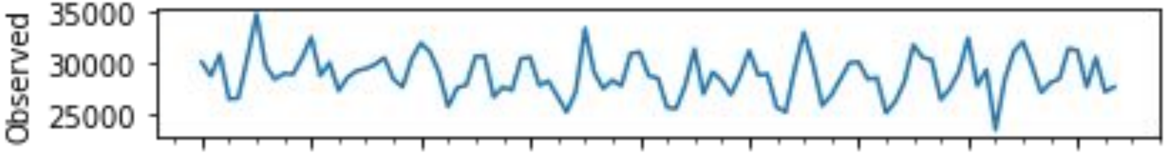
- Préparation des données
- Consommation corrigée de l’effet température
- **Désaisonnalisation**
- Prédiction
    - ➔ Méthode Holt-winter
    - ➔ Méthode SARIMA
- Conclusion

Désaisonnalisation

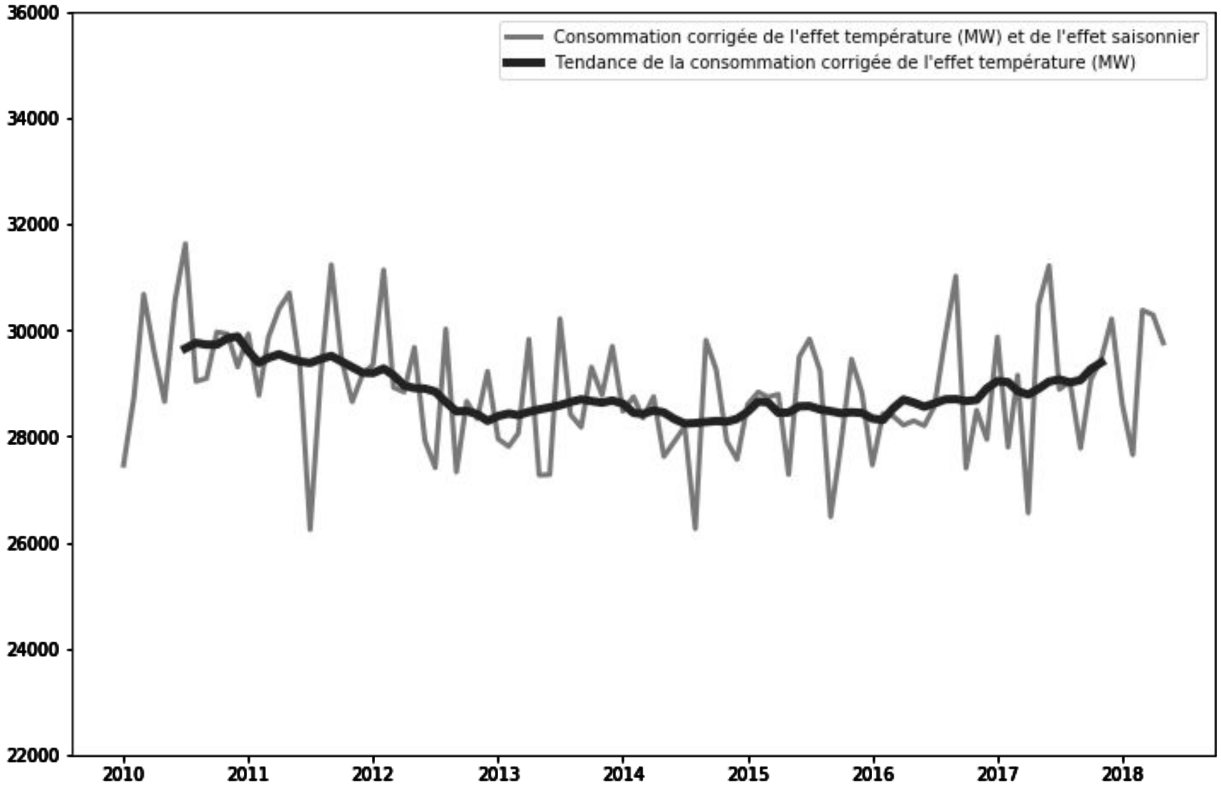
```
seasonal_decompose(data['consommation_corrigee'], model='additive')
```



Désaisonnalisation



Désaisonnalisation



Plan

- Préparation des données
- Consommation corrigée de l’effet température
- Désaisonnalisation
- **Prédiction**
    - ➔ **Méthode Holt-winter**
    - ➔ Méthode SARIMA
- Conclusion

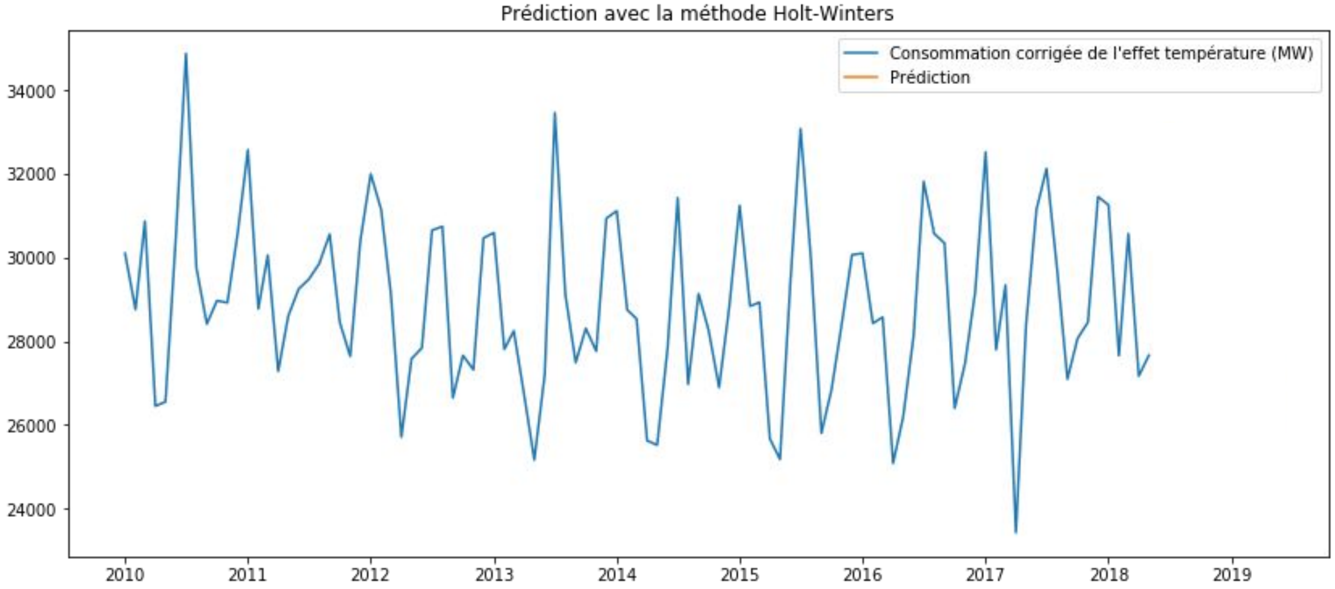
Holt-Winters - Modèle

Création du modèle

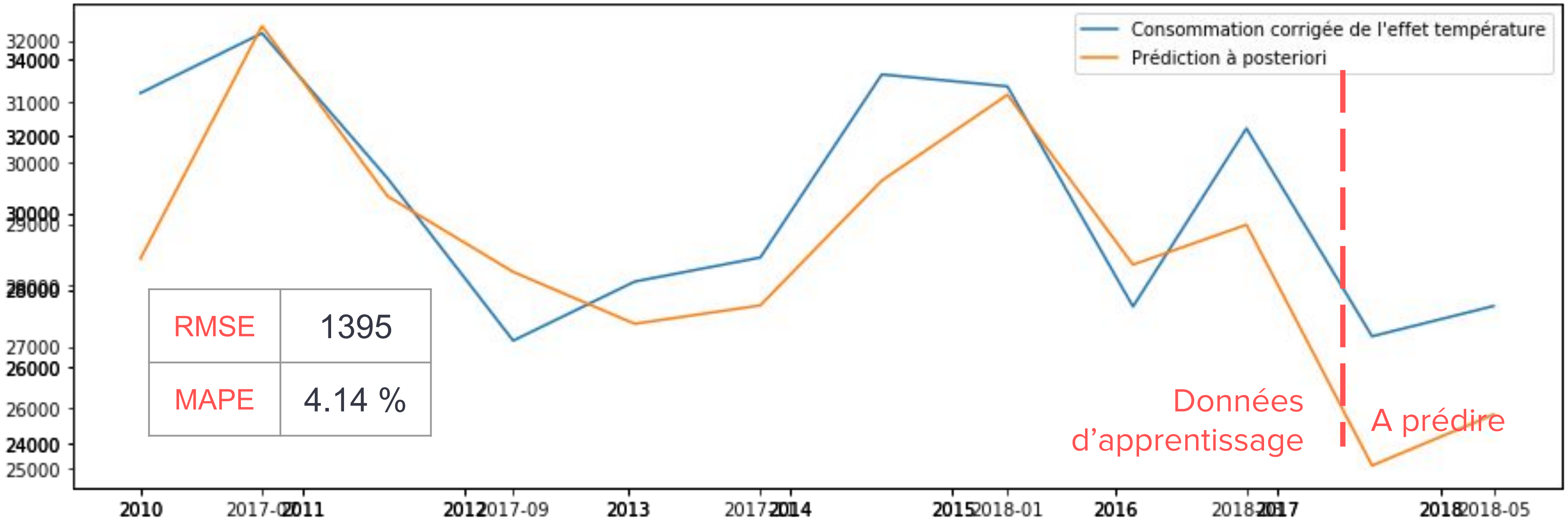
```
hw1_model = ExponentialSmoothing(np.asarray(y), seasonal_periods=12, trend='additive', seasonal='additive').fit()
```

Prédiction

```
hw1_pred = hw1_model.forecast(DUREE_PREDITE)
```



Holt-Winters - Analyse à posteriori



RMSE	1395
MAPE	4.14 %

Plan

- Préparation des données
- Consommation corrigée de l’effet température
- Désaisonnalisation
- **Prédiction**
  - ➔ Méthode Holt-winter
  - ➔ **Méthode SARIMA**
- Conclusion

SARIMA - Description

$$SARIMA(p,d,q)(P,D,Q)_s$$

Définie une relation linéaire entre un instant t et les **p** instants précédents

Définie une relation entre une perturbation décorélée entre un instant t est les **q** instants précédents

Cette décomposition est aussi faite pour la partie saisonnière, selon les paramètres **P** et **Q**.

Stationnarisation incluse

SARIMA - Stationnarité

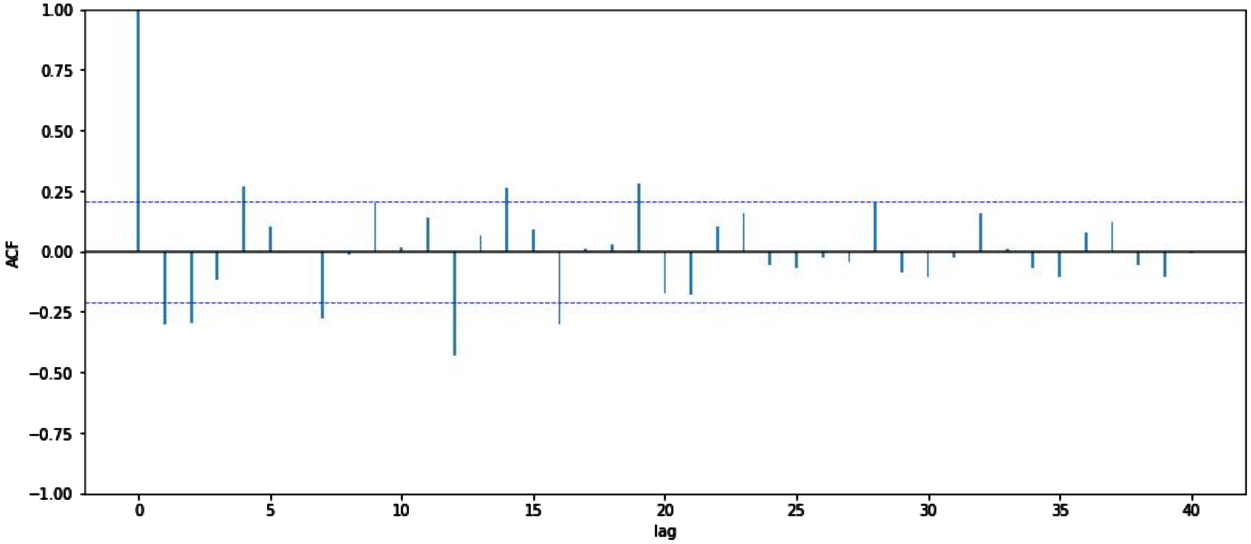
Un processus est stationnaire si :

- ➔ Son **espérance** est **constante** à travers le temps
- ➔ Sa **variance** est **constante** à travers le temps
- ➔ Les **autocorrélations entre deux moments séparés** dans le temps sont **constantes**

La **structure** du processus stationnaire **n'évolue pas** avec le temps,

ce qui le rend plus **prévisible**

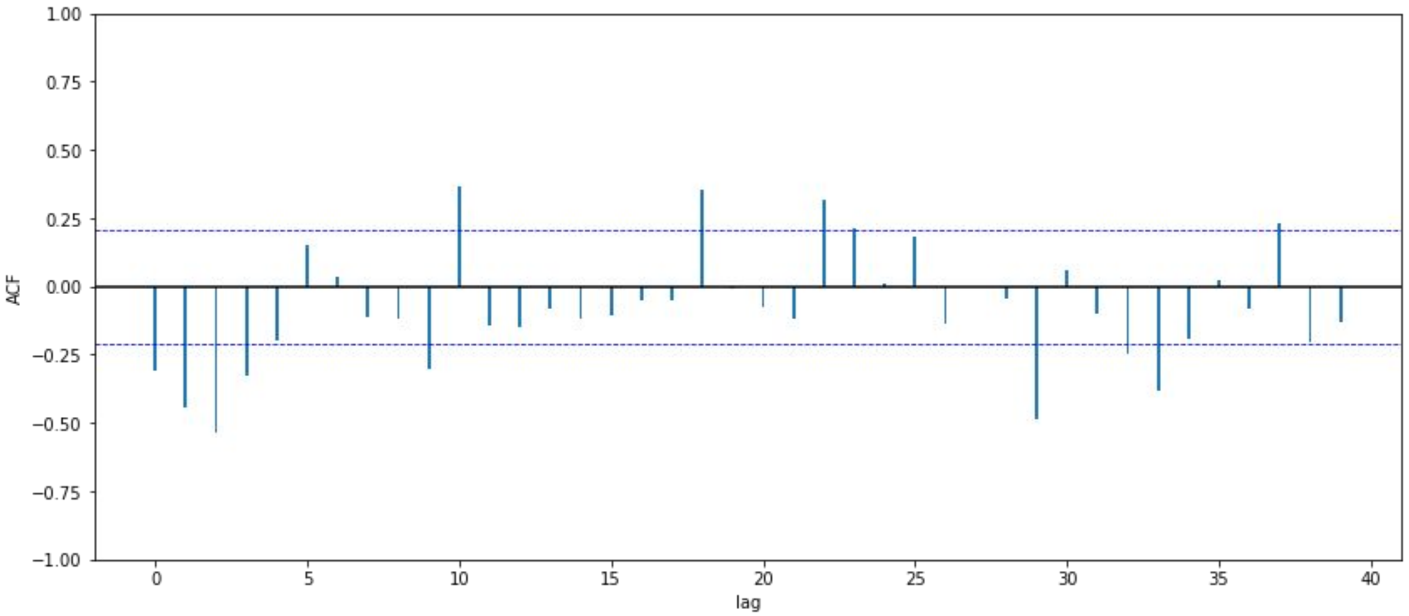
SARIMA - Autocorrélograme simple



Différenciation : on transforme la série $X_t \leftarrow X_t - X_{t-1}$

La série **semble** stationnaire $X_t \leftarrow X_t - X_{t-12}$

SARIMA - Autocorrélograme multiples

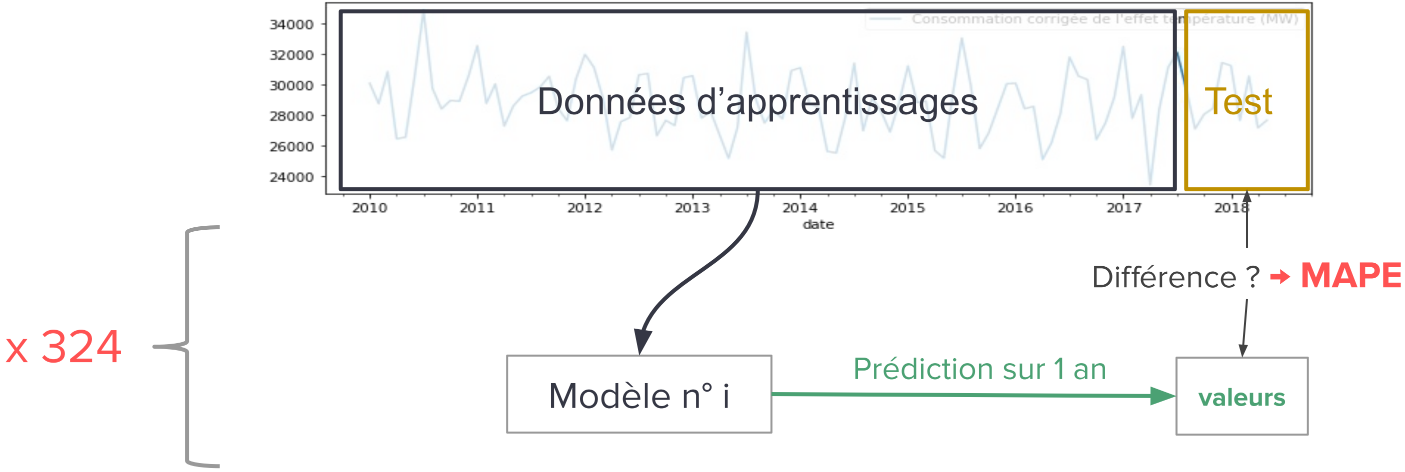


SARIMA - Recherche automatique des paramètres

On teste **324** modèles différents

$$\text{SARIMA}(p: 0 \rightarrow 2,\ d: 0 \rightarrow 1,\ q: 0 \rightarrow 2)(P: 0 \rightarrow 2,\ D: 0 \rightarrow 1,\ Q: 0 \rightarrow 2)_{12}$$

Analyse à posteriori



SARIMA - Recherche automatique des paramètres

MAPE croissant :

paramètres non-significatifs

modele	mape
(1, 0, 1)x(2, 0, 1, 12)	3.413167e+00
(2, 1, 2)x(2, 0, 0, 12)	3.418456e+00
(2, 1, 2)x(2, 0, 1, 12)	3.448125e+00
(2, 1, 0)x(2, 1, 0, 12)	3.485421e+00
(1, 1, 0)x(1, 0, 1, 12)	3.529038e+00
(1, 1, 2)x(2, 0, 0, 12)	3.553575e+00
(2, 1, 0)x(2, 0, 1, 12)	3.575302e+00
(2, 1, 0)x(1, 1, 1, 12)	3.604991e+00

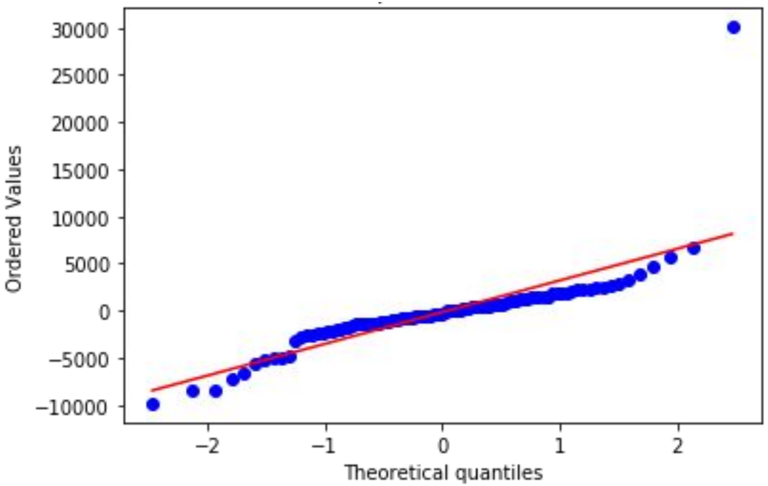
324 modèles ...

	coef	std err	z	P>|z|
ar.L1	-0.7092	0.142	-5.002	0.000
ar.L2	-0.4693	0.149	-3.142	0.002
ar.S.L12	-0.9228	0.117	-7.888	0.000
ar.S.L24	-0.4689	0.105	-4.456	0.000
sigma2	2.047e+06	4.32e+05	4.734	0.000

Candidat : SARIMA$(2,1,0)(2,1,0)_{12}$

SARIMA$(2,1,0)(2,1,0)_{12}$ - Analyse des résidus

Diagramme quantile-quantile

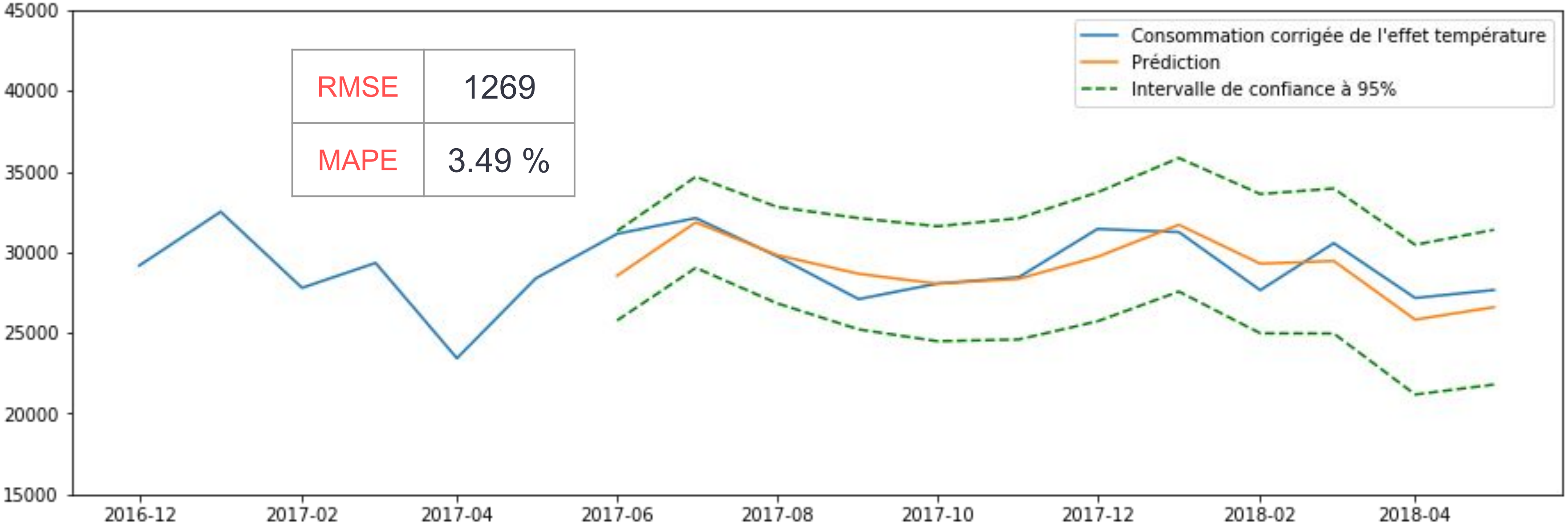


Test de Ljung-Box *(définition)*

Retard	p-valeur
6	0.05
12	0.09
18	0.16
24	0.28
30	0.36
36	0.44

Les résidus sont un bruit blanc → **On valide le modèle**

SARIMA - Analyse à posteriori



RMSE	1269
MAPE	3.49 %

Plan

- Préparation des données
- Consommation corrigée de l'effet température
- Désaisonnalisation
- Prédiction
  - ➜ Méthode Holt-winter
  - ➜ Méthode SARIMA
- **Conclusion**

Conclusion - **Prédiction** avec SARIMA