Mémoire de Séries Temporelles

Auteurs:

Lesage Laurent - Maillart Arthur - Montagne Pierre

Responsable:

Christian Robert

1 avril 2016

Plan

Introduction

En bleu : Laurent

L'objectif de ce mémoire est de mettre en oeuvre les techniques d'estimation de tendance, de saisonnalité et de composantes aléatoires, pour ensuite réaliser des prédictions sur deux jeux de données. Les données concernent un historique de la production hydraulique d'électricité et de consommation d'énergie, de 1981 jusque 2015. Les deux prédictions porent sur l'année 2016.

Proposition modification plan.
I-Série 1

a) méthode 1
b) méthode 2

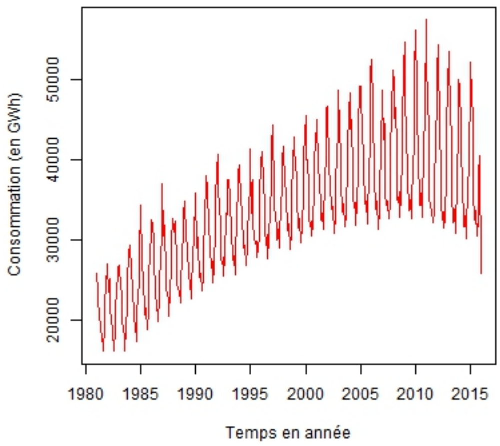
II-Série 2

a) méthode 1
b) méthode 2

Permet de bien dérouler les méthodes.

1 Modélisation de la série de la consommations réelle

Nous commençons par quelques remarques sur le chronogramme de la série de la consommation réelle. Sur la représentation graphique ci-dessous de la série nous pouvons observer un trend qui a priori est quadratique ainsi qu'une saisonnalité d'ordre 12. Notons que la variance semble augmenter après 2006. **Mettre un graphique propre**



Dans un premier temps, nous avons utilisé la méthode de Buys-Ballot (Attention aux spec du modèle) sur le logarithme de la série (pour atténuer les effets de la variance après 2006). Nous avons effectué un décomposition avec un trend quadratique et une saisonnalité d'ordre 12. Cependant, nous n'avons pas réussi à modéliser la composante aléatoire de manière satisfaisante. En observant bien les ACF et PACF obtenus, nous avons préféré éliminer uniquement le trend et modéliser avec un SARIMA la saisonnalité et le bruit (a reformuler). Nous détaillons maintenant notre approche.

1.1 Buys-Ballot

- **Objectif de la méthode :**
  La méthode de Buys-Ballot généralisée consiste à proposer une décomposition du type tendance + saisonnalité + composante aléatoire.

$$X_t = m_t + S_t + Y_t \text{ avec t=1,..., T.}$$

- **Conditions :**
  Les conditions à respecter sont les suivantes :

  – La tendance m_t se décompose de la sorte :

$$m_t = \sum_{i=1}^{k} b_i m_t^{(i)}$$

  - les b_i des constantes à déterminer
  - k l'ordre de la tendance.

– La tendance S_t se décompose de la sorte :

$$S_t = \sum_{j=1}^{l} c_j S_t^{(j)}$$

- avec les c_i des constantes à déterminer
- k la saisonnalité à modéliser.

– La composante aléatoire Y_t est caractérisée par :

- Son Espérance : $\mathbb{E}[Y_t] = 0$
- Sa Variance : $\mathrm{V}[Y_t] = \text{constante}$

- **Intérêt principal :**
  Pouvoir faire des prévisions en utilisant la tendance et la saisonnalité (la composante aléatoire est d'espérance nulle).

- **Résumé des éléments :**

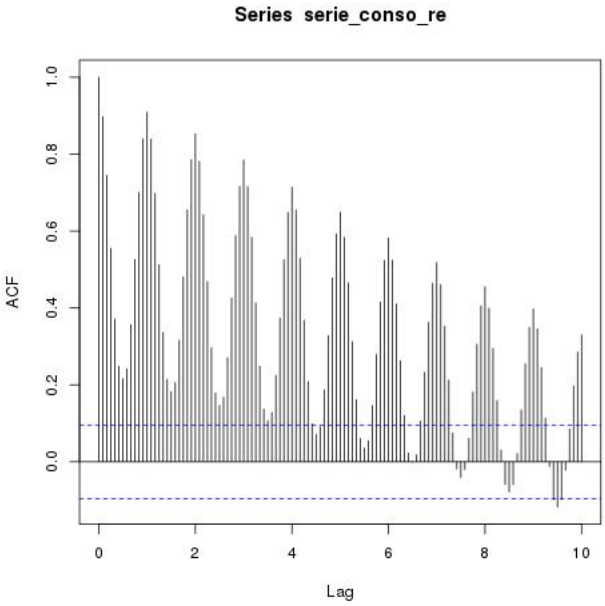
  Décrire le tableau avec les éléments

1.2 Modéliser un SARIMA sur la série entière

Grand schéma synthétique avec les différentes étapes successives : méthodologie de Box et Jenkins

Ajout de sous titres correspondant aux différentes étapes

On commence par représenter l'autocorrélogramme de la série brute. Cet autocorrélogramme montre une décroissance lente ainsi qu'une saisonnalité d'ordre 12. Pour stationnariser la série nous allons donc différencier la série une fois. Ceci a pour effet de faire disparaître le trend mais pas la saisonnalité. Maintenant, nous faisons une différence avec un lag de 12 suivie d'une différenciation classique. Cette fois ci le trend et la saisonnalité on été éliminés. L'autocorrélogramme que nous obtenons semble stationnaire.



Nous effectuons donc un test pour confirmer notre hypothèse.

Test bruit blanc

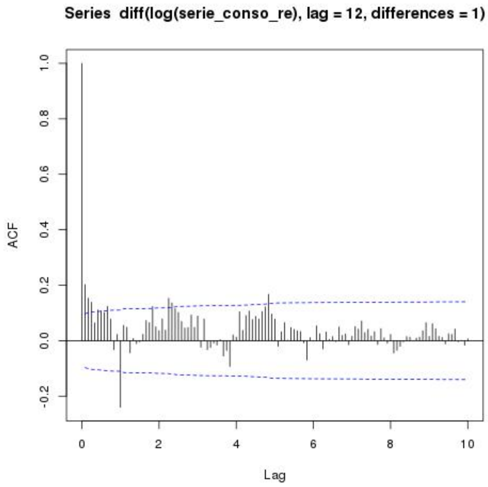
Pas Dickey Fuller qui donne de la merde

Les résultats du test confirment notre intuition. Nous pouvons donc continuer l'analyse.

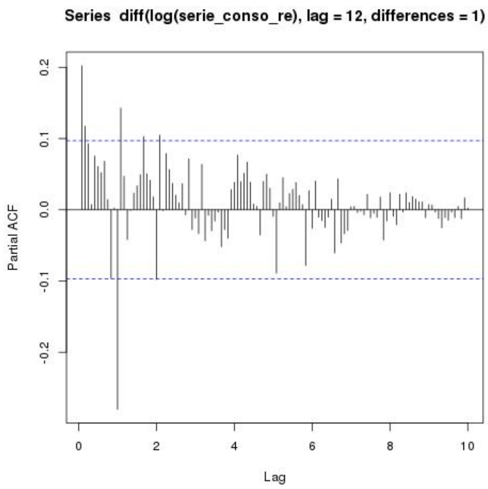
Hétéroscédasticité

Notre idée ici est d'ajuster un modèle SARIMA sur le logarithme de la série. Nous venons de déterminer en stationnarisant la série que le paramètre d vaut 1, le paramètre de saisonnalité vaut 12, et que $D = 1$. Reste maintenant à déterminer p,q,P,D,Q. Pour obtenir p_{max}, q_{max},P_{max}, Q_{max} nous allons observer plus en détail l'autocorrélogramme et l'autocorrélogramme partiel. Les graphiques sont dégueux mais c'est pour illustrer

Posons $s = 12$, alors on observe un pic significatif en s puis, il n'y en a plus. Par ailleurs, il y a un pic significatif qui suit les pics en $0 \times s$ et $1 \times s$, comme nous pouvons le voir ci-dessous.



Ainsi nous posons $p_{max} = 1$ et $P_max = 1$. Sur l'autocorrélogramme partiel, on observe 2 pics saisonniers ainsi qu'un pic significatif en dehors de ces pics saisonniers.



Ceci nous incite à poser $Q_{max} = 2$ et $q_{max} = 2$

Rappeler le modèle théorique.

Désormais, nous avons tous les paramètres pour ajuster notre modèle. Après quelques essais infructueux, nous observons un modèle qui se détache.

A développer ...

Comparaison des modèles acceptables avec plusieurs critères : AIC, BIC

Conclusion

Contents