

Science des données III : cours 5



Séries spatio-temporelles (partie 3)

Philippe Grosjean & Guyliann Engels

Université de Mons, Belgique
Laboratoire d'Écologie numérique des Milieux aquatiques

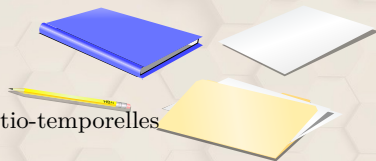


<http://biodatascience-course.sciviews.org>
sdd@sciviews.org

Décomposition des séries temporelles

Objectifs du cours

- Appréhender la décomposition des séries spatio-temporelles
- Comprendre le filtrage par fenêtre mobile
- Savoir utilisation le filtrage par les moyennes mobiles



Rappel : les composantes d'un signal spatio-temporel

- Une série spatio-temporelle peut être considérée comme la **superposition de plusieurs signaux plus simples** :
 - Une ou plusieurs **tendances générales** (augmentation ou diminution progressive sur le long terme),
 - Des **tendances locales**, responsables de la variation *locale* de la valeur moyenne des valeurs,
 - Des **cycles** qui correspondent à des variations périodiques (cycle circadien, effet saisonnier, lunaire ou des marées, ...)
 - Un **bruit blanc** représenté par une variation purement aléatoire (analogue aux **résidus** d'un modèle linéaire)
- Ces différentes composantes peuvent se combiner de **deux façons différentes** :
 - De manière **additive** (modèle additif)
 - De manière **multiplicative** (modèle multiplicatif issu d'un effet à caractère exponentiel)

Organisation de l'analyse dans le package **pastecs**

- Le package **pastecs** (“Package for the Analysis of Space-Time Ecological Series”) organise les différentes étapes du traitement d’une série spatio-temporelle
- A chaque étape correspond un **objet** particulier
- Pour la décomposition de série, c’est l’objet **tsd**, pour “Time Series Decomposition”
- Ces objets ont des **méthodes** qui leur sont propres: **print()**, **summary()**, **plot()**, **extract()** et **specs()** ici, voir **?tsd**

Calcul par fenêtre mobile

- Centrer une fenêtre autout d'une observation. L'**ordre k** de la fenêtre est le nombre d'observations de la série régulière que l'on y inclut à gauche, et à droite. Par exemple, $k = 3 \Rightarrow$ la fenêtre compte 7 observations : la valeur cible, ainsi que 3 observations à gauche et 3 observations à droite.
- Effectuer un **calcul** dans cette fenêtre, et **remplacer la valeur cible par ce résultat**.
- **Décaler** la fenêtre d'une observaton vers le droite.
- Naturellement, on perd des valeurs au début et à la fin, à moins d'utiliser des astuces de calcul pour ces cas particuliers (dupliquer ou cycler la série, par exemple).

Méthode des moyennes mobiles

- Le calcul effectué est la moyenne:

$$M_t = \frac{\sum_{\tau=-k}^k X_{t+\tau}}{2.k + 1}$$

- $2.k + 1$ est la taille de fenêtre aussi appelée ici **bande de lissage**
- Cette technique permet donc de **lisser** la série, mais surtout, d'**éliminer les cycles** de période proche à la bande de lissage.
- Dans pastecs, soit `decaverage()`, soit `tsd(method = "average")`, voir `?decaverage`

Démonstration

Décomposition des séries **nottem** et **co2** par moyennes mobiles. Ensuite, extraction des composantes issues de la décomposition et analyse de leurs propriétés (`acf()`, `spectrum()`, `trend.test()`, ...)