

Séries temporelles

Données

Le jeu de données oak_seeds.csv indique le nombre de graines du chêne *Quercus crispula* recueillies annuellement (1980-2017) par 16 capteurs (*trap*) situés dans un peuplement de cette espèce au Japon.

```
seed <- read.csv("../donnees/oak_seeds.csv")
head(seed)
```

```
##   year trap seeds
## 1 1980    1   13
## 2 1980    2  131
## 3 1980    3   44
## 4 1980    4   44
## 5 1980    5   47
## 6 1980    6   27
```

Le fichier oak_weather.csv contient quant à lui des données météo annuelles pour ce même site:

```
weather <- read.csv("../donnees/oak_weather.csv")
head(weather)
```

```
##   year temp_fl temp_gr rain_fl rain_gr
## 1 1980   14.9   15.2    75    437
## 2 1981    9.3   15.4    40    766
## 3 1982   11.5   15.8   109    487
## 4 1983   11.5   15.9    49    657
## 5 1984   13.4   17.1    49    622
## 6 1985   11.5   16.9    63    501
```

- *temp_fl*: Température (°C) moyenne durant la période de floraison de l'arbre.
- *temp_gr*: Température (°C) moyenne durant la saison de croissance
- *rain_fl*: Quantité totale de pluie (en mm) durant la période de floraison de l'arbre.
- *rain_gr*: Quantité totale de pluie (en mm) durant la saison de croissance.

Ces données proviennent de l'étude suivante:

Shibata, M., Masaki, T., Yagihashi, T., Shimada, T., & Saitoh, T. (2019). Data from: Decadal changes in masting behaviour of oak trees with rising temperature. Dryad Digital Repository. <https://doi.org/10.5061/dryad.v6wwpzgrb>

1. Série temporelle à l'échelle du site

- (a) Calculez le nombre total de graines recueillies par année (tous capteurs confondus) et appliquez une transformation racine carrée au résultat. Convertissez le résultat en tableau de données temporel (*tsibble*) et visualisez la série temporelle obtenue.

Note: Puisque nous utiliserons des modèles linéaires plutôt que des modèles généralisés dans cet exercice, la transformation racine carrée vise à stabiliser la variance des données de comptage.

- (b) Visualisez les corrélations temporelles pour cette série. Quel type de modèle ARIMA (AR et/ou MA, ainsi que leur ordre) pourrait être approprié ici?
- (c) Ajustez un modèle ARIMA en laissant la fonction choisir automatiquement le type et l'ordre du modèle. Que signifient les coefficients obtenus?
- (d) Joignez le jeu de données `weather` et ajustez un modèle ARIMA incluant les quatre variables météorologiques comme prédicteurs externes. Est-ce que ces variables contribuent à mieux prédire le nombre de graines produites par année?
- (e) Quel type de modèle est choisi par `ARIMA()` si vous ne considérez que la sous-série débutant en l'an 2000, sans prédicteur externe? Expliquez ce choix à partir du graphique en (a) et des corrélations temporelles pour cette sous-série.
- (f) Calculez les prévisions des modèles en (c) et (e) pour les cinq années suivantes. Comment ces prévisions diffèrent-elles?

2. Série temporelle par capteur

- (a) Reprenez le tableau original indiquant le nombre de graines par année et capteur, puis appliquez la transformation racine carrée au nombre de graines. Utilisez ensuite la fonction `lme` du package `nlme` pour ajuster un modèle linéaire mixte incluant: l'effet fixe des variables météo, l'effet aléatoire du capteur et les corrélations temporelles d'une année à l'autre.

Voici un exemple de la façon de spécifier un effet aléatoire d'une variable `GROUPE` sur l'ordonnée à l'origine d'un modèle `lme`, ainsi qu'une corrélation de type ARMA entre éléments successifs d'un même `GROUPE`:

```
library(nlme)
mod_lme <- lme(..., data = ...,
               random = list(GROUPE = ~1),
               correlation = corARMA(p = ..., q = ..., form = ~ 1 | GROUPE))
```

- (b) Comparez la précision des effets fixes dans ce modèle en (a) par rapport au modèle en 1(d). Quelle est la raison de cette différence?