

# Séries Chronologiques

Université d'Angers – M2 Data Science

## Guide pour le TP 2

Le TP 2 s'appuie sur le jeu de données `strates.txt`, issu du package `astsa` (*Applied Statistical Time Series Analysis*), dont la description proposée dans l'aide est : *Sedimentary deposits from one location in Massachusetts for 634 years, beginning nearly 12,000 years ago*. On se fixe comme objectif la modélisation de la série chronologique (au sens de l'erreur de prédiction one-step minimale, par exemple).

Quelques pistes suggérées :

- Justifier (visuellement) que la transformation logarithmique est raisonnable. Que peut-on dire de la stationnarité de la série transformée ? Pour s'en convaincre, proposer des critères visuels (comme des fenêtres glissantes) et des tests statistiques.
- Estimer par la méthode de votre choix une tendance cubique, de la forme

$$m_t = a_0 + a_1 t + a_2 t^2 + a_3 t^3.$$

Critiquer ce choix en termes d'extrapolation du phénomène.

- Estimer une tendance sinusoidale, de la forme

$$m_t = a_0 + a_1 \cos\left(\frac{2\pi t}{\theta}\right).$$

Pour le choix de  $\theta$ , on pourra par exemple tester plusieurs valeurs (entre 300 et 500), et garder celle qui minimise les carrés résiduels de la régression. Quels sont les avantages et inconvénients de cette méthode par rapport à la précédente ?

- Récupérer le résidu et tester la stationnarité.
- Mettre en place des modèles ARMA sur ce résidu. On validera (ou non) les choix effectués par :
  - la significativité des paramètres estimés,
  - la blancheur du résidu de l'ARMA (ACF/PACF empiriques, tests),
  - la normalité du résidu de l'ARMA (forme de histogramme, majorité des valeurs standardisées dans  $[-2, 2]$ , tests),
  - des critères généraux tels que l'AIC.
- Superposer les données initiales et les données reconstruites par le modèle (donc de la forme  $\hat{m}_t + \hat{Z}_t$  où  $Z_t$  est la fluctuation récupérée après l'estimation de la tendance et modélisée par un ARMA). Calculer le MSE pour la prédiction one-step.