

# Ondelettes pour l'analyse de séries temporelles ou spatiales

*Lise Vaudor, ISIG, UMR 5600*

## Contents

<b>Introduction</b>	<b>1</b>
Objectifs . . . . .	1
Présentation rapide de l'application TheWaveletToolKat . . . . .	1
<b>Chargement ou simulation de données (Onglet "Data")</b>	<b>2</b>
Chargement de données . . . . .	2
Simulation de données . . . . .	2
<b>Analyse des données (Onglet "Analysis")</b>	<b>2</b>
Transformée de Fourier . . . . .	2
Transformée en ondelettes continues (CWT) . . . . .	3
Transformée en ondelettes croisées (XWT) . . . . .	4
Transformée en ondelettes discrètes (DWT) . . . . .	4

## Introduction

### Objectifs

Les ondelettes= une incursion dans le monde impitoyable des maths et de la physique. MAIS des applications dans beaucoup de domaines, et entre autres dans le domaine du traitement de séries. . .

**Donner à voir** plutôt que raisonner sur les équations. . . Se concentrer sur l'**analyse** plutôt que sur la **prise en main de l'outil**: utilisation d'une **application** "clés en main" plutôt que recours à la ligne de commande R (elle aussi impitoyable).

### Présentation rapide de l'application TheWaveletToolKat

Une application 'shiny', basée sur le langage R.

Instructions d'installation dans HowToInstall.txt, Lancement de l'application par double-clic sur "TheWavelet-ToolKat.bat".

- Onglet de présentation, encarts d'info marqués par 
- Onglet "Data"
- Onglet "Analyses"

## Chargement ou simulation de données (Onglet “Data”)

Possibilité de **charger** un jeu de données ou de **simuler** des données (choix de “Type of data” dans l’appli).

Dans les deux cas, possibilité d’analyser une seule série ( $y_1$ ) ou deux séries ( $y_1$  et  $y_2$ ) (activation de “Add variable  $y_2$ ” dans l’appli).

### Chargement de données

- **Charger le jeu de données.** Renseigner (le cas échéant) le séparateur de colonnes, le format des données manquantes, etc. (cf )
- **Choisir les variables** à analyser. Si la variable  $x$  est une date, le préciser, et préciser son format. (cf )
- Préciser si vous souhaitez analyser une deuxième série,  $y_2$
- **Choisir le pas** auquel les données seront interpolées (cf )

Le jeu de données est dès lors **formaté, et nettoyé de ses valeurs manquantes**. C’est ce jeu de données (**téléchargeable** dans le panneau latéral) qui sera utilisé pour les analyses.

**Exercice** : Charger le jeu de données “donnees\_meteo\_Bron.txt”.

### Simulation de données

Caractéristiques des données simulées : cf .

- Choisir la longueur de la série
- Choisir si vous souhaitez analyser une ( $y_1$ ) ou deux ( $y_1$  et  $y_2$ ) séries.
- Choisir les sous-signaux en réglant les valeurs de  $\sigma$  (pour le bruit) et l’amplitude  $A$  (pour les sinusoides).
- Choisir les caractéristiques des sous-signaux en réglant les valeurs de période  $T$ , phase  $\phi$ , et la portion du signal affectée par les variations ( $x_{min}$  et  $x_{max}$ ).

Le jeu de données est dès lors **génééré**. C’est ce jeu de données (**téléchargeable** dans le panneau latéral) qui sera utilisé pour les analyses.

## Analyse des données (Onglet “Analysis”)

### Transformée de Fourier

Cf .

**Exercice** : considérer les données simulées selon les caractéristiques suivantes et analyser leurs transformées de Fourier:

- **Cas 1** signal  $y_1$ :
  - $A_a = 1, T_a = 100$
- **Cas 2** signal  $y_1$ :
  - bruit blanc  $\sigma = 1$
- **Cas 3** signal  $y_1$ :

- bruit blanc  $\sigma = 1$
- $A_a = 1, T_a = 100$
- **Cas 4** signal  $y_1$ :
  - $A_a = 5, T_a = 100$
  - $A_b = 2, T_b = 250$
- **Cas 5** signal  $y_1$ :
  - $A_a = 1, T_a = 100, x_{min\ a} = 200, x_{max\ a} = 800$
- **Cas 6** signal  $y_1$ :
  - bruit blanc  $\sigma = 1,$
  - $A_a = 1, T_a = 100, x_{min\ a} = 200, x_{max\ a} = 800$
  - $A_b = 2, T_b = 250, x_{min\ b} = 400, x_{max\ b} = 900$
  - $A_c = 2, T_c = 40, x_{min\ c} = 1, x_{max\ c} = 500$

## Transformée en ondelettes continues (CWT)

Cf CWT\_Intro.html

Cf 🍷.

Attention ici distinction période (période du signal) et échelle (largeur de la fenêtre)...

### Différents types d'info

- **wavelet**: coefficients d'ondelettes, correspond aux variations pour différentes échelles (ou périodes)
- **power**: puissance (ou énergie), correspond au carré des coefficients d'ondelettes
- **power.corr.norm** : puissance corrigée, normalisée

### Différents types de graphiques

- **Type Info=f(x,T)**: on affiche l'info en fonction de la localisation (x) et de la période considérée (T). C'est la représentation la plus classique.
- **Type Info=f(x)**: on affiche l'info pour une seule période T, choisie par l'utilisateur. Cela permet une meilleure visibilité (pour une période choisie) de l'info affichée sur le 1er graphique (qui par construction doit représenter 3 dimensions).
- **Type Power=f(T)**: représente le spectre du signal, c'est à dire la quantité d'énergie dans le signal en fonction de la période considérée. On renonce ainsi à représenter l'info relative aux localisations (x) pour ne voir que l'info relative aux périodes (T).

**Exercice** : considérer les données simulées selon les caractéristiques suivantes:

- Reprendre les cas 1 à 6 et analyser leurs transformées en ondelettes continues
- Analyser le résultat graphique de type Info=f(x,T) en faisant varier la nature de l'information affichée.

## Transformée en ondelettes croisées (XWT)

Même principe que CWT sauf qu'on multiplie les coefficients d'ondelettes de deux séries. On obtient ainsi les périodes et localisations où le **produit des coefficients est fort**.

Le but de cette analyse est d'identifier les périodes et localisations pour lesquelles les deux séries auraient des **variations communes** (produit positif) ou opposées (produit négatif).

Problème: le produit pourrait être élevé parce que les coefficients sont très élevés pour l'une des deux séries, bien qu'ils ne le soient pas particulièrement pour l'autre...

D'où le calcul de la **cohérence d'ondelettes** (choix de l'info 'rsq' dans l'appli) qui s'apparente davantage au calcul d'un "coefficient de corrélation". En effet, dans le calcul de cohérence d'ondelettes, le produit des coefficients est divisé par les valeurs "moyennes" des coefficients d'ondelettes pour chaque série.

L'information de **phase** aide également à l'interprétation de ce résultat.

Cf 

### Exercice

Comparer les signaux suivants (vous pouvez les analyser individuellement -CWT-, puis conjointement -XWT-).

- signal  $y_1$ :
  - bruit blanc  $\sigma = 1$ ,
  - $A_a = 1, T_a = 100, x_{min\ a} = 200, x_{max\ a} = 800$
  - $A_b = 2, T_b = 250, x_{min\ b} = 400, x_{max\ b} = 900$
  - $A_c = 2, T_c = 40$
- signal  $y_2$ 
  - bruit blanc  $\sigma = 1$ ,
  - $A_a = 1, T_a = 100$
  - $A_c = 1, T_c = 40, x_{min\ c} = 100, x_{max\ c} = 700$

## Transformée en ondelettes discrètes (DWT)

Cf DWT\_Intro.html

Cf 

**Décomposition** du signal en **smooths** (moyennes locales de moyennes) et **détails** (moyennes locales de variations) à diverses échelles.

### Exercice

Charger le jeu de données 'data\_mystere\_2', et décomposer les signaux  $y_1$  et  $y_2$  à l'aide de la DWT pour "isoler" les variations aux échelles affectées par des variations.