

NUM-2.1 Data Analysis

Cécile Mallet & Richard Wilson

3 ECTS

This course consists of 2 parts: (1) statistical methods for data analysis and (2) Fourier methods.

Data analysis (Cécile Mallet)

- 1- Basic concepts in statistics: mean, variance, histogram, central limit theorem, probability densities, expectations and co-variance, data visualization.
- 2- Linear regression, deterministic and probabilistic modelisation. Confidence intervals and tests.
- 3- Principal component analysis (PCA), maximum variance formulation, application of PCA.
- 4- Tutorial on linear regression and PCA. Examples will be taken from climate databases (satellite observations and/or output of climate models)
- 5- Non supervised and supervised approaches in classification (k-means clustering, k-nearest neighbors algorithm)

Fourier methods (Richard Wilson)

1. Mathematical tools for signal processing: function spaces, Fourier series, Fourier transform, convolution. Power spectral density.
2. Introduction to the theory of distributions. Dirac distribution and Dirac comb. Fourier transform of distributions.
3. Digital signals. Sampling. Discrete Fourier Transform. Discrete correlation/convolution. Power spectral density.
4. Random signals. Stationarity and ergodicity. Moments of a random process. Estimation. Spectral analysis of random signals. Power spectrum estimate.

Cécile Mallet is Associate professor at the Paris-Saclay University and scientist at LATMOS. Research topics: Statistics, Artificial intelligence, Remote sensing, Atmosphere.

Richard Wilson is associate professor at Sorbonne University and scientist at LATMOS. Research topics: Meso and small scale dynamics of the atmosphere, turbulence, signal processing.

