Robust statistics for signal processing

Master ATSI

outside cloud

within cloud

Mean

though, both seem
to have same
distance from Mean!

Slides available at https://alexandrerenaux.github.io/
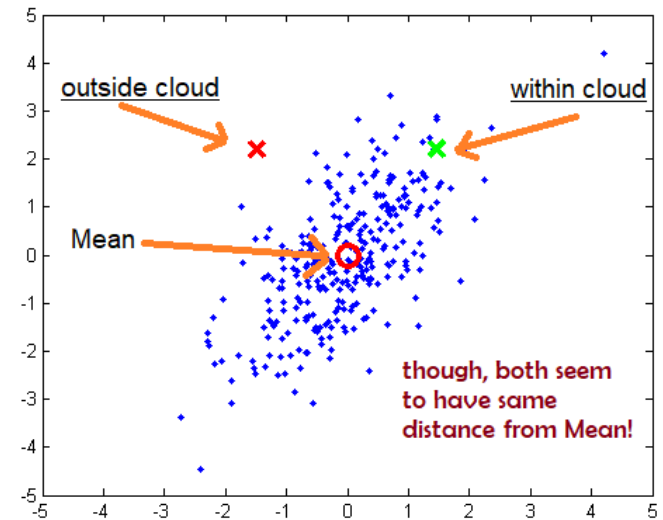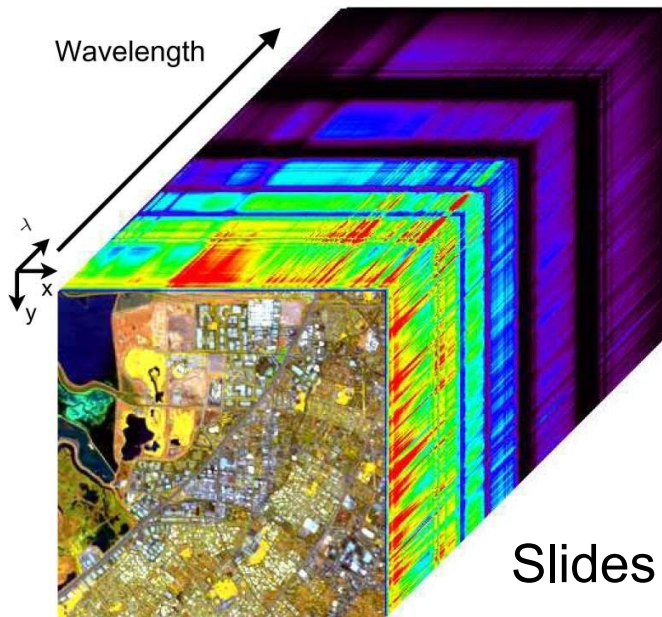
Faculty Professors

Frédéric Pascal, CentraleSupelec (Paris-Saclay Univ),
Laboratory of Signal and Systems

https://fredericpascal.blogspot.com/p/home.html

Alexandre Renaux, Faculty of Science (Paris-Saclay Univ),
Laboratory of Signal and Systems

https://l2s.centralesupelec.fr/u/renaux-alexandre/

Mohammed El Korso, Paris Nanterre University,
Laboratory for Energetic Mechanics and Electromagnetism

https://sites.google.com/site/nabkorso/

Stefano Fortunati, Institut of Polytechnic Science and Aeronautics,
Laboratory of Signal and Systems

https://l2s.centralesupelec.fr/u/fortunati-stefano/

Research keywords: data science, statistical signal processing, machine learning, big data, information geometry, robust statistics, estimation theory (parametric and semi-parametric), radar/satellite imaging, radioastronomy, MIMO/STAP processing, real estate price optimization, portfolio optimization, etc.

Master ATSI Robust statistics

Robust statistics: why ???

⚠️ In many cases, physical models and Gaussian assumption are well justified (thanks to the central limit theorem for example) and one can use:

➡️ Classical methods such as least square, maximum likelihood, etc
Classical performances metrics such as covariance, Cramér-Rao bound, etc
Classical result about Gaussian random vector (confidance intervals, etc)

These are the fundamental background of this course (the first lecture will give a quick reminder)
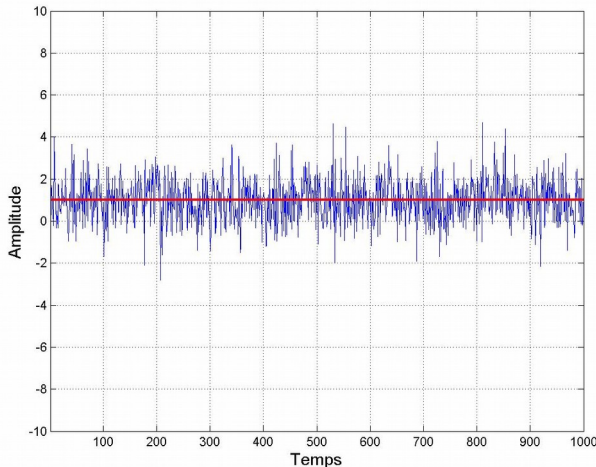
$$\widehat{\theta} = \frac{1}{N} \sum_{t=1}^{N} y(t)$$

No bias
Efficient (w.r.t. the CRB)

Robust statistics: why ???

But several measurement campaigns confirmed impulsive (heavy-tailed) noise, e.g. in indoor mobile communication channels or biomedical sensor (array) measurements.

+ outliers appear (i.e. a data point that differs significantly from other observations) in the mesured data due to sensor failure, experiment errors, etc.
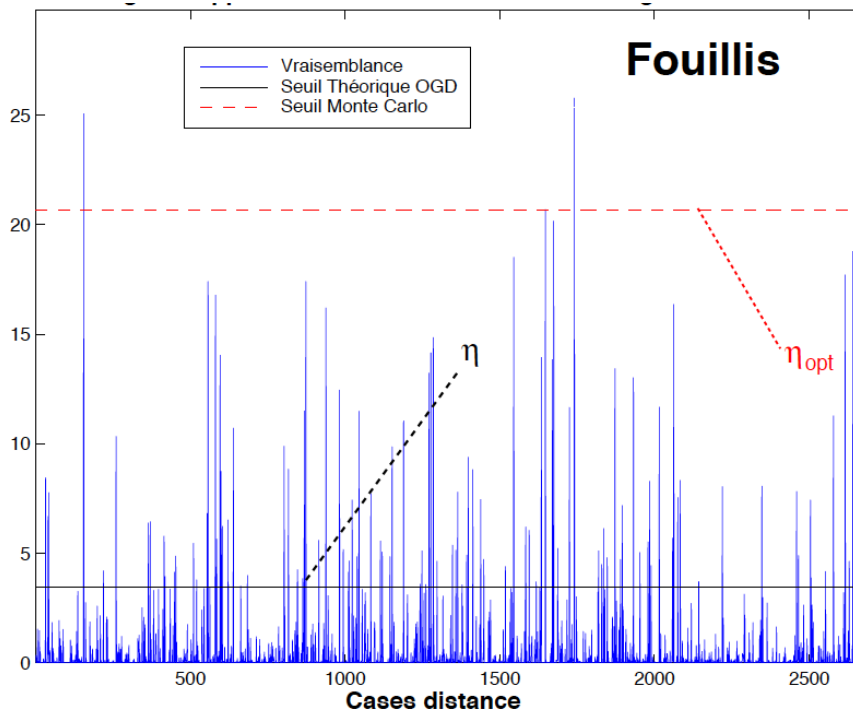
Sea radar measurement

Noise only (i.e. no sources in the data) !!!!

Classical theory fail

> 100 sources detected

Goal of the course

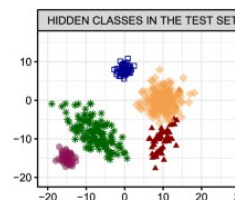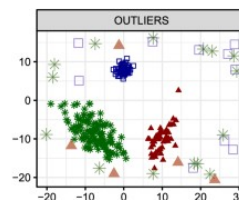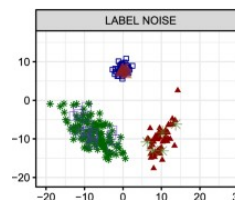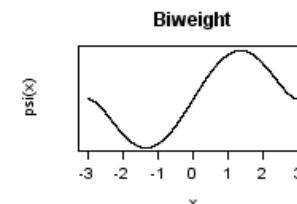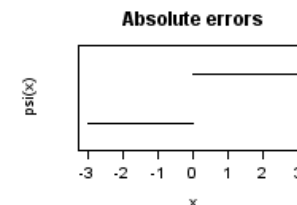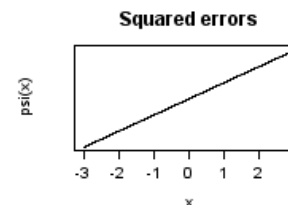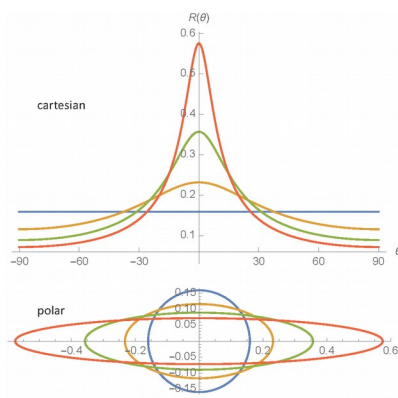Fundamental questions (and answers) to this course will be how to

- quantify robustness (and what are the mathematical definitions of statistical robustness)

- go beyond gaussianity

- design estimators able to deal with a certain amount of outliers with a reasonable efficiency

- quantify the quality of these estimators

- apply these robustified estimators to classical techniques (supervised and unsupervised learning methods) with real world applications

Organization

Outline of the course available at

https://www.universite-paris-saclay.fr/formation/master/electronique-energie-electriq
ue-automatique/m2-automatique-traitement-du-signal-et-des-images#programme

Keywords: robust regression, M-estimator, sensivity curve, influence function,
elliptical distributions, etc.

Language: English

7 courses sessions (3h each) + 2 lab sessions + 1 exam

Lab sessions: Matlab
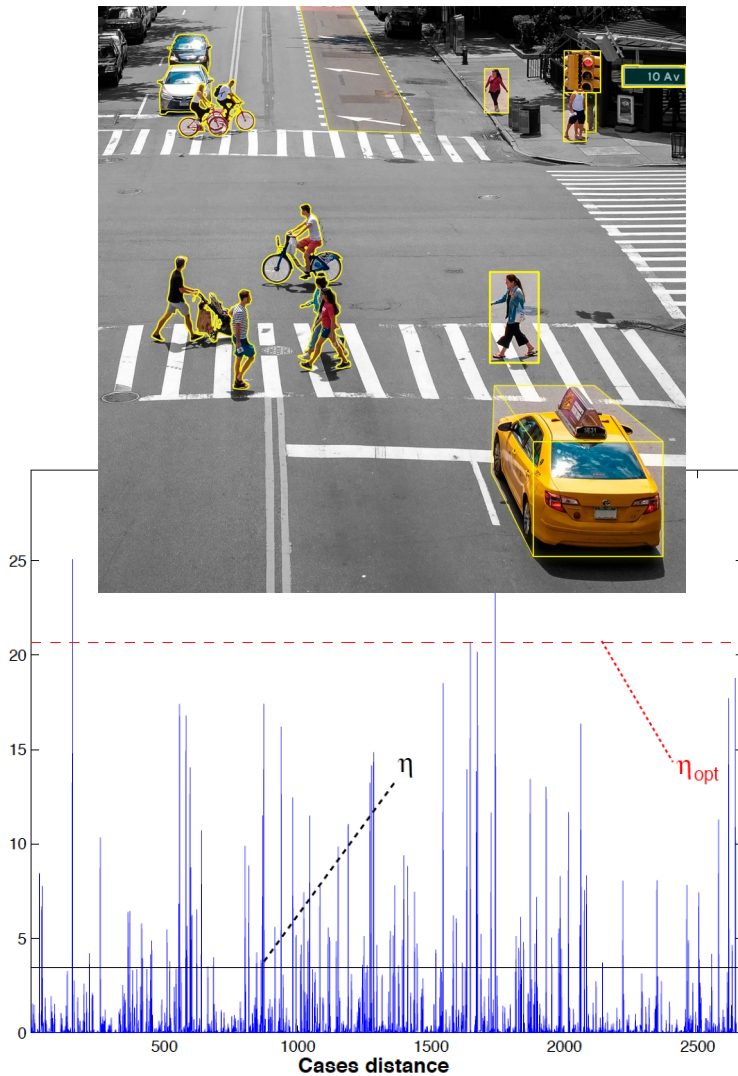
Grading: 2/3 exam + 1/3 lab session (reports)

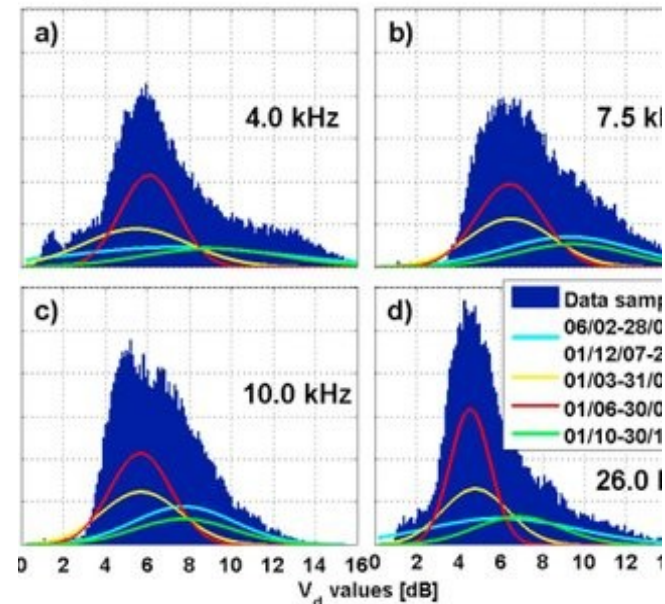Exam:

1 article to read (available one week before)
3 hrs to answer questions related to this article

Master ATSI Robust statistics

Covered applications

Radar detection and source localization in radioastronomy/autonomous navigation

Histograms for SB data; BW=75Hz; 06/02/2007-26/02/200

Natural atmospheric noise statistics

Hyperspectral imaging classification (environmental monitoring)