Ajuntament de Barcelona

Anàlisi temporal amb Machine learning

Laia Subirats Maté

8 de juliol de 2021

Índex

1. Representació d'estats
2. Comparació de sèries temporals mateixa mida
3. Comparació de sèries temporals diferent mida
4. Cheat Sheet Python
5. Visualitzacions gràfiques de time series

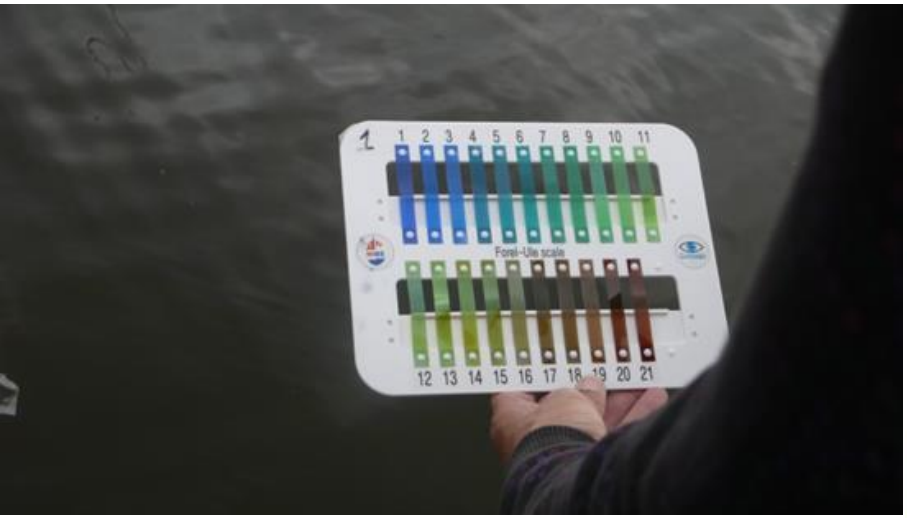
1. Representació d'estats

- Decrementa, estacionari i incrementa

Artificial intelligence and earth observation to explore water quality in the **Wadden Sea**

1. Representació d'estats

- Decrementa, estacionari i incrementa
- Incrementa fortament, incrementa suaument, etc.



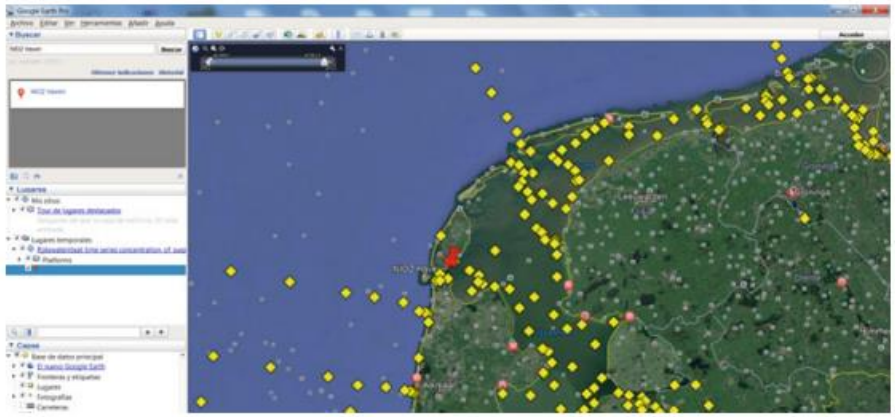
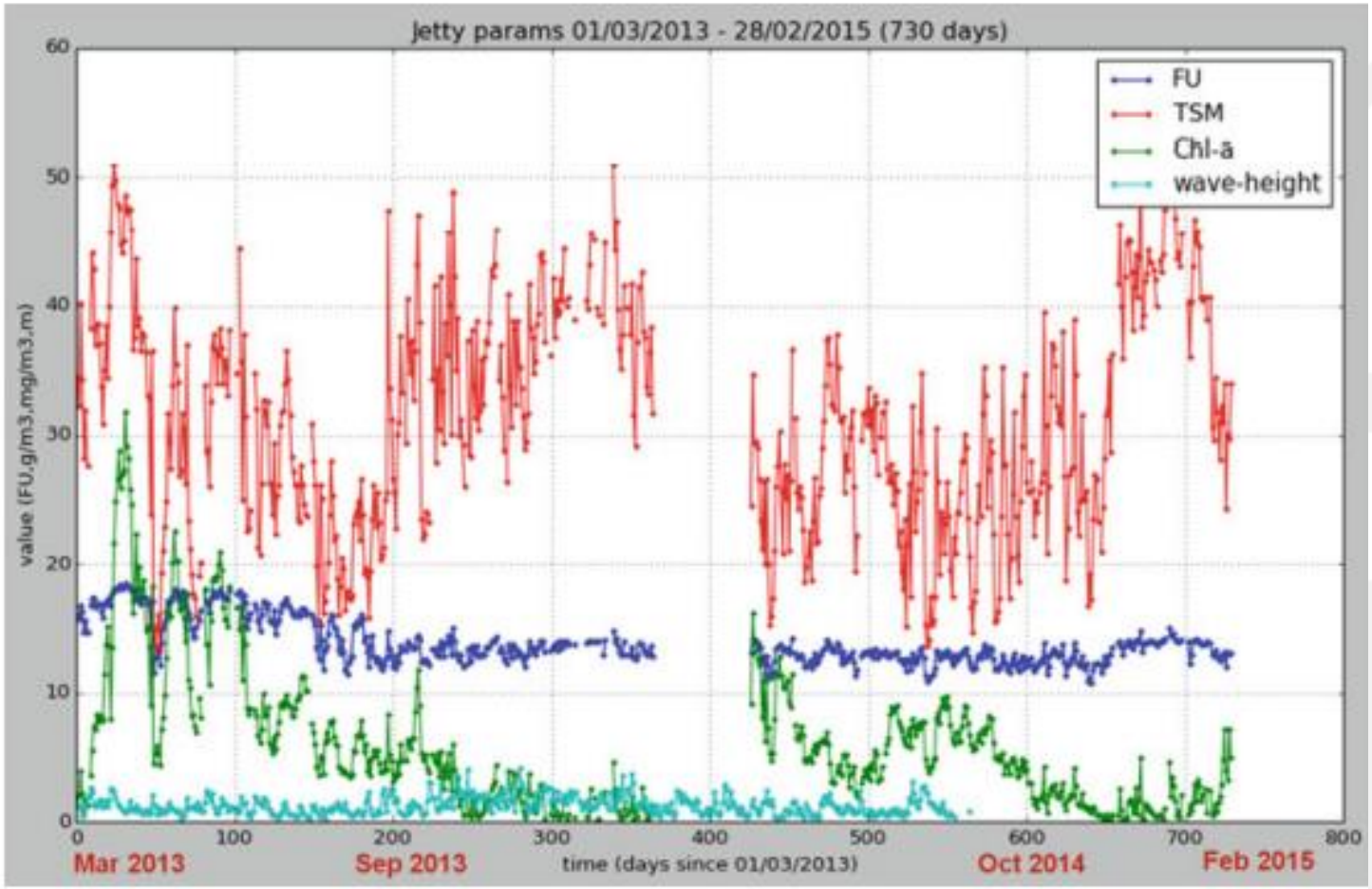


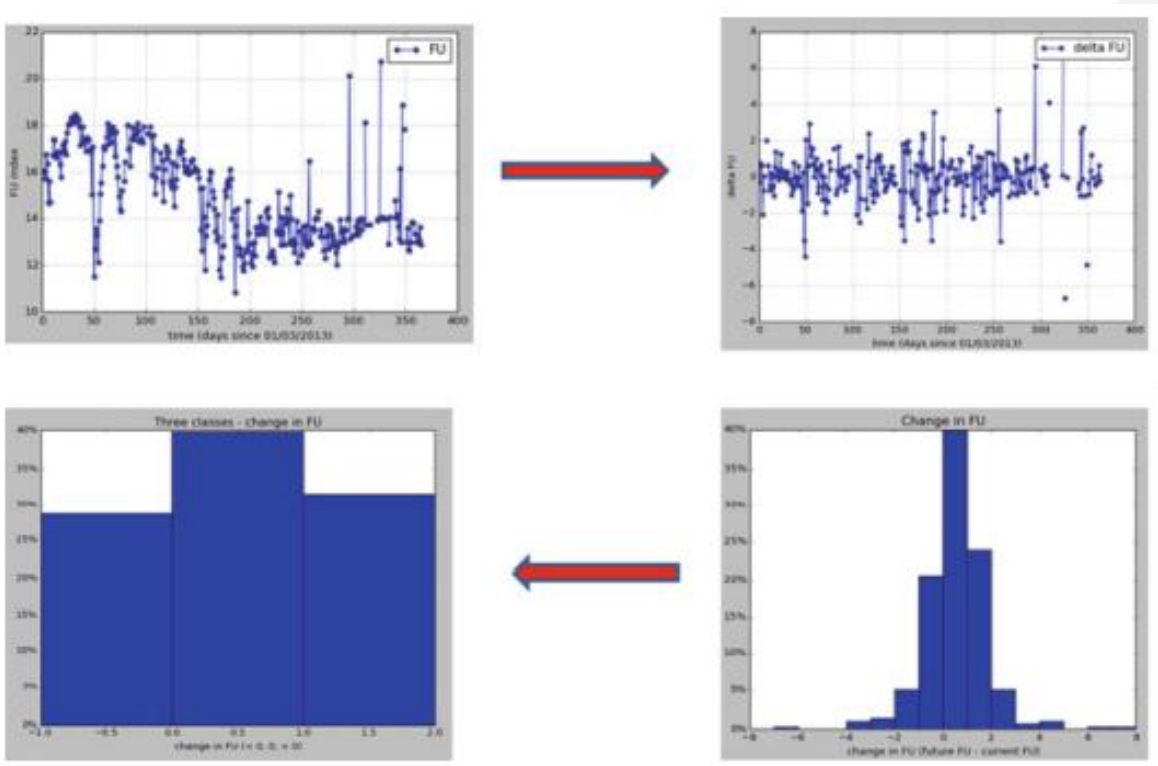
Fig. 1 In situ monitoring platform: *yellow markers* indicate the Dutch national water quality monitoring network (Rijkswaterstaat); the *red pin* (NIOZ jetty) indicates the location of the observation platform of the Royal Netherlands Institute of Sea Research (NIOZ). Source: http://kml.deltares.nl/kml/rijkswaterstaat/waterbase/concentration_of_suspended_matter_in_water.kml and Google Earth

1. Representació d'estats



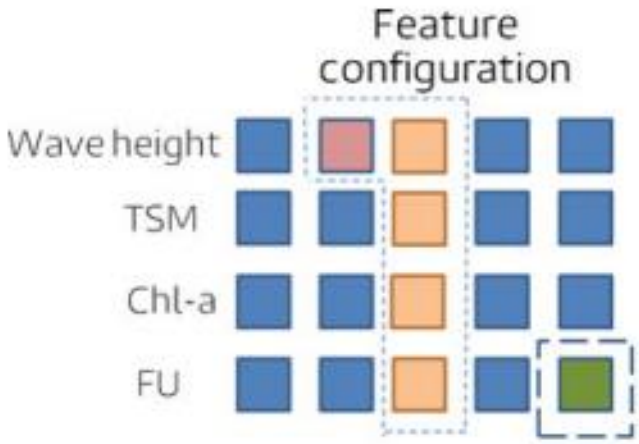
1. Representació d'estats

- A model of the target variable "FU colour" at future points (2 days, 4 days, 7 days) has been learned



1. Representació d'estats

- The target-value attribute is FU at 2 days into the future, and the input vector includes the following attributes: wave height, TSM, Chl-a, FU at the current time point, and wave height at 1 day in the past.



ML technique

Support Vector Machine
- C (soft-margin) : 1.0
- Kernel : radial basis

Number of examples for training

415

Blind predictor benchmark

48%

Distribution of target FU classes

decrease	stable	increase
27%	48%	25%

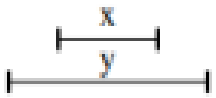
Accuracy (10-fold cross-validation)

53%

Barcelona Activa

1. Representació d'estats

- Incrementa suaument, decrementa suaument, etc.
- Intervals d'Allen: 13 possibles relacions

Relation	Symbol	Inverse	Meaning
x before y	b	bi	
x meets y	m	mi	
x overlaps y	o	oi	
x during y	d	di	
x starts y	s	si	
x finishes y	f	fi	
x equal y	eq	eq	

1. Representació d'estats

- Estandarització
- Diferents aproximacions temporals:
  - Estat anterior
  - Llarg termini
  - Canvi d'estats
- Es pot indicar l'estat inicial

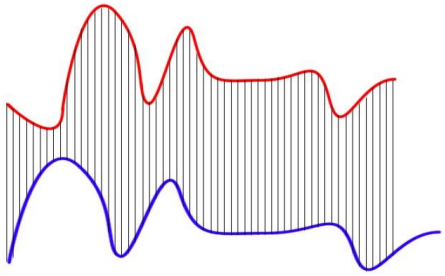
2. Comparació de sèries temporals mateixa mida

- Distància euclidea i altres distàncies

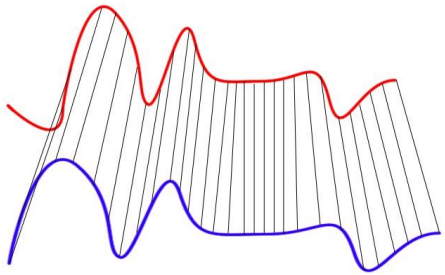
https://docs.scipy.org/doc/scipy/reference/generated/scipy.spatial.distance.euclidean.html
https://docs.scipy.org/doc/scipy/reference/spatial.distance.html

3. Comparació de sèries temporals amb diferent mida

- Dynamic Time Warping is used to compare the similarity or calculate the distance between two arrays or time series with different length.



Euclidean Matching

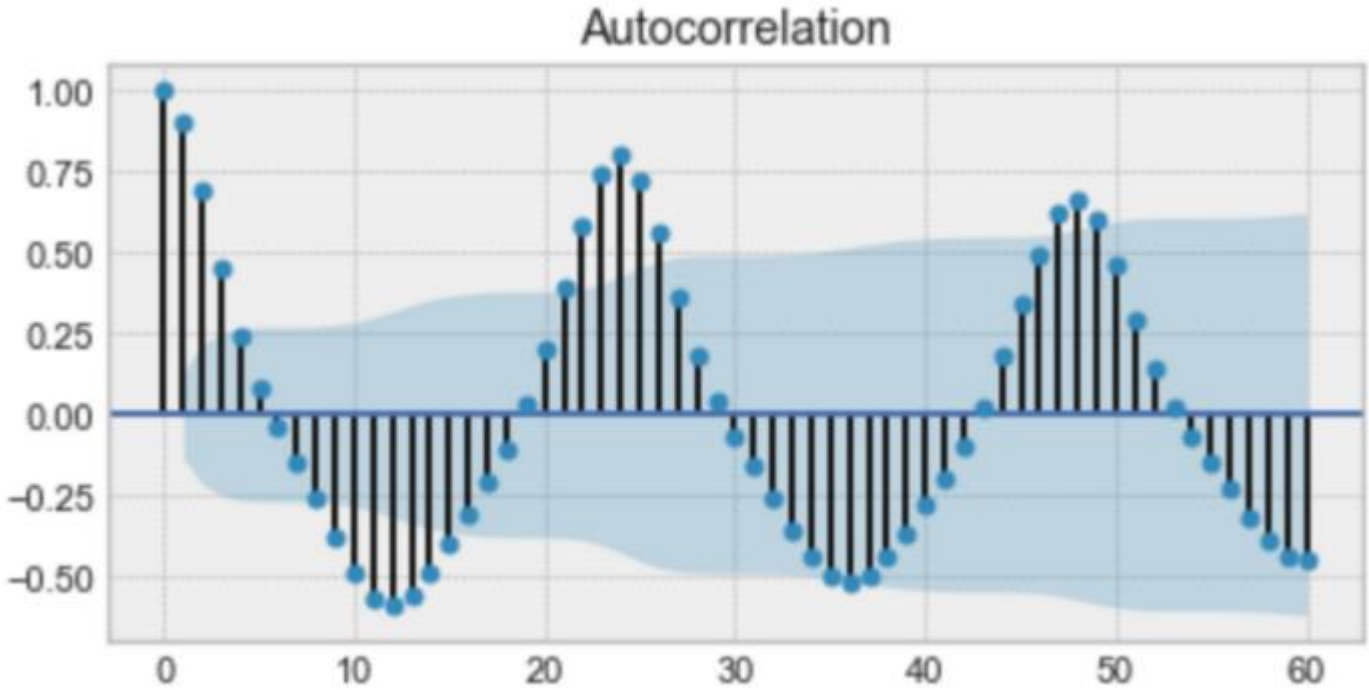


Dynamic Time Warping Matching

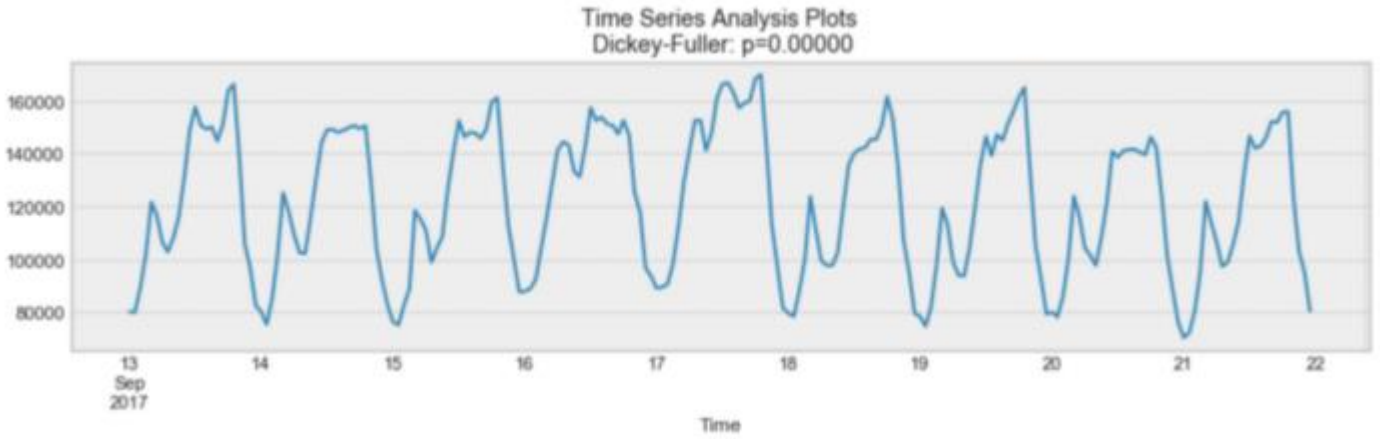
https://towardsdatascience.com/dynamic-time-warping-3933f25fcdd

Barcelona Activa

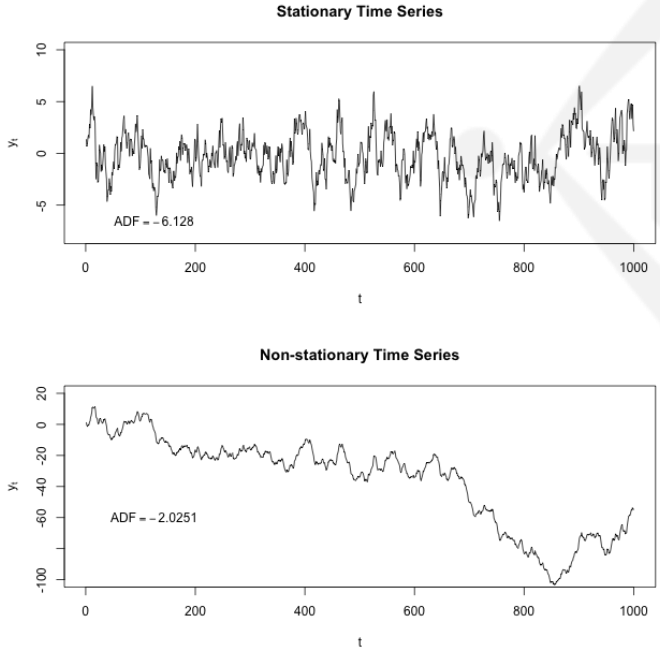
- Autocorrelation: Informally, **autocorrelation** is the similarity between observations as a function of the time lag between them.



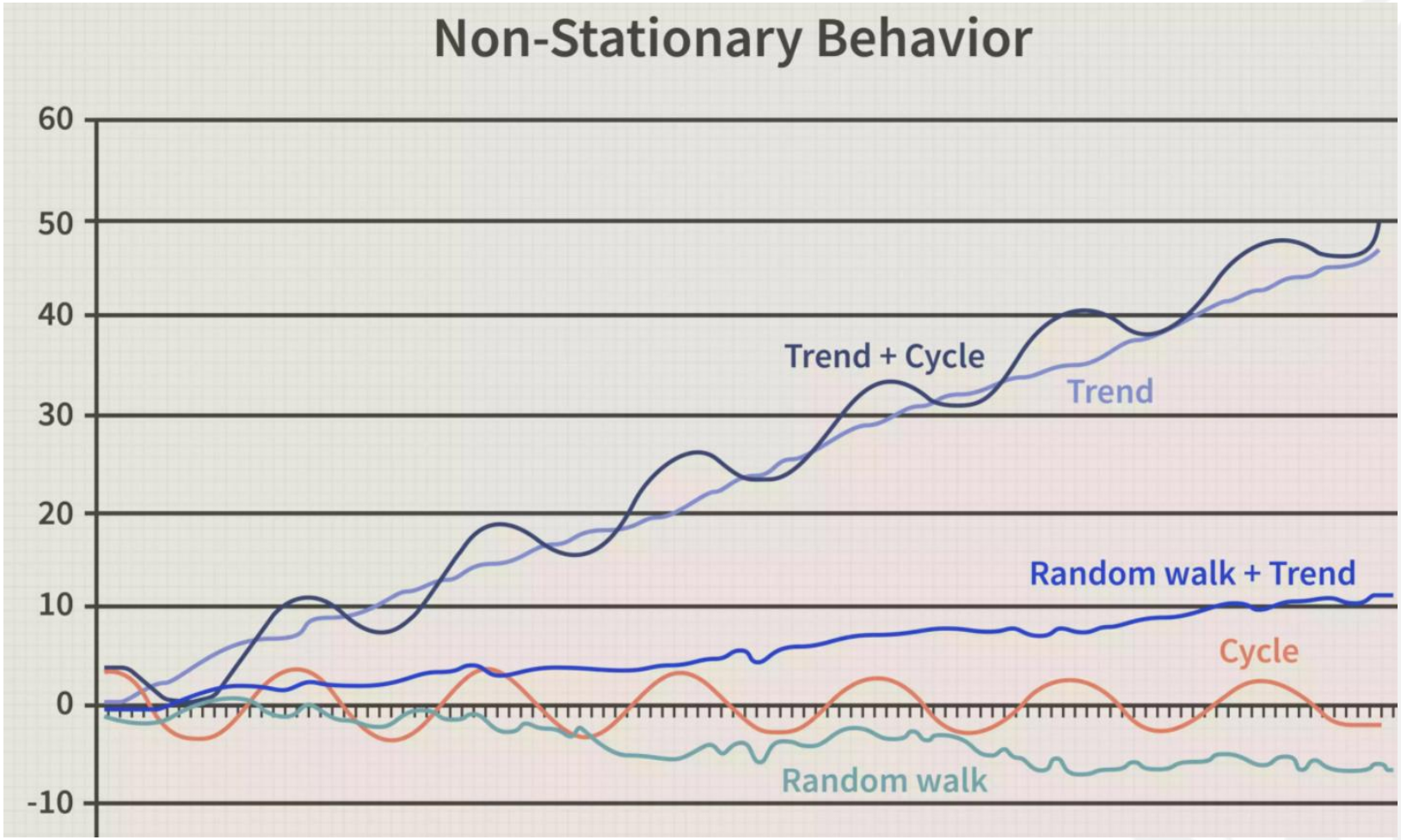
- Seasonality: **Seasonality** refers to **periodic fluctuations**. For example, electricity consumption is high during the day and low during night, or online sales increase during Christmas before slowing down again.



- **Stationarity** is an important characteristic of time series. A time series is said to be stationary if its statistical properties do not change over time. In other words, it has **constant mean and variance**, and covariance is independent of time.



4. Complete guide



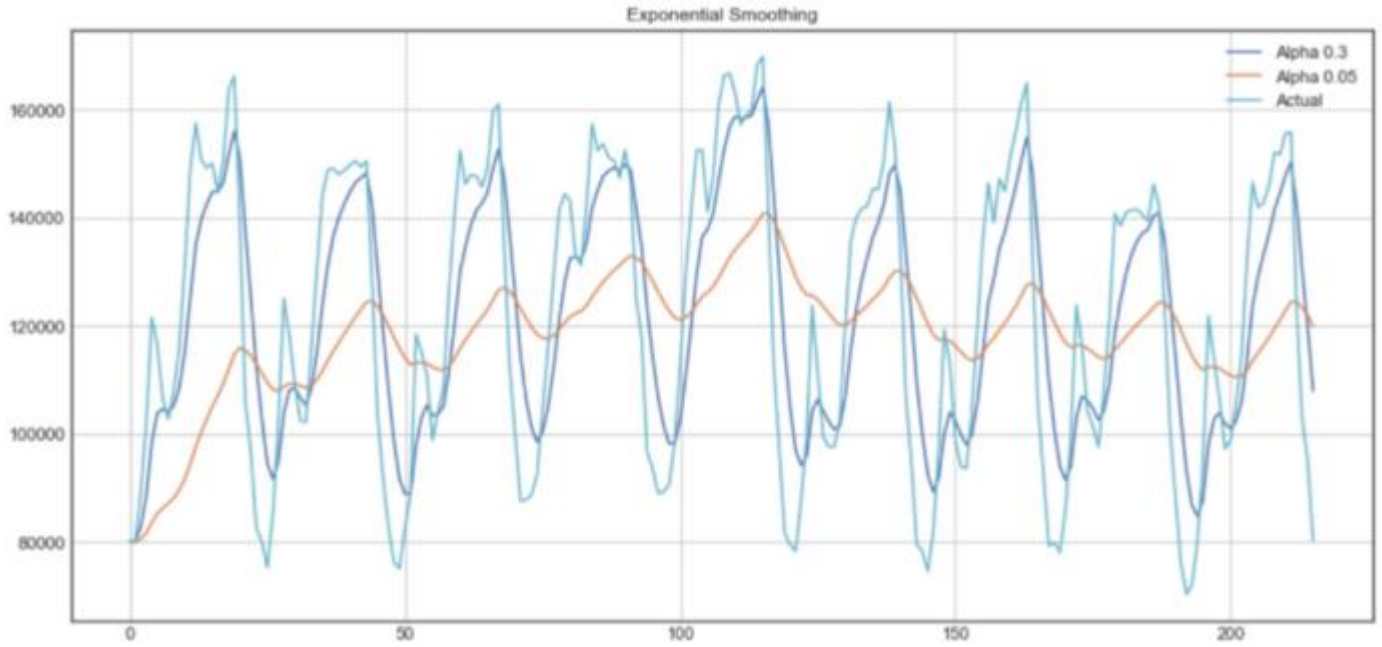
4. Complete guide

- In this website there is a complete guide of time analysis
  - moving average

https://towardsdatascience.com/the-complete-guide-to-time-series-analysis-and-forecasting-70d476bfe775

4. Complete guide

- In this website there is a complete guide of time analysis
  - exponential smoothing



- Autoregressive model i ARIMA

Example: An AR(1) process [edit]

An AR(1) process is given by:

$$X_t = c + \varphi X_{t-1} + \varepsilon_t$$

where ε_t is a white noise process with zero mean and constant variance σ_ε^2. (Note: The subscript on φ_1 has been dropped.) The process is wide-sense stationary if $|\varphi| < 1$ since it is obtained as the output of a stable filter whose input is white noise. (If $\varphi = 1$ then the variance of X_t depends on time lag t, so that the variance of the series diverges to infinity as t goes to infinity, and is therefore not wide sense stationary.) Assuming $|\varphi| < 1$, the mean $\mathrm{E}(X_t)$ is identical for all values of t by the very definition of wide sense stationarity. If the mean is denoted by μ, it follows from

$$\mathrm{E}(X_t) = \mathrm{E}(c) + \varphi\,\mathrm{E}(X_{t-1}) + \mathrm{E}(\varepsilon_t),$$

The notation $AR(p)$ indicates an autoregressive model of order p. The AR(p) model is defined as

$$X_t = c + \sum_{i=1}^{p} \varphi_i X_{t-i} + \varepsilon_t$$

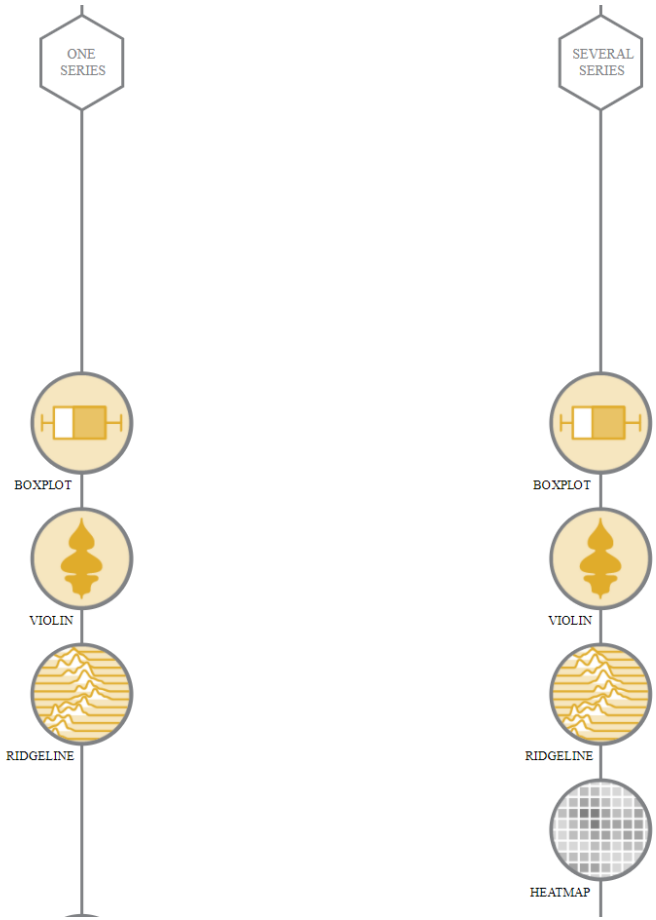
https://towardsdatascience.com/understanding-arima-time-series-modeling-d99cd11be3f8

4. Cheat Sheet

- https://machinelearningmastery.com/time-series-forecasting-methods-in-python-cheat-sheet/
- Podeu calcular noves sèries temporals de sèries ja existents i aplicar els mètodes a aquestes darreres

5. Gràfiques time series

- https://www.data-to-viz.com/



Gràcies per la vostra col·laboració!

Laia.Subirats@eurecat.org