Second Assignment
2.5 points

Wind energy production

Wind energy production

- Wind blows, wind generators rotate, then electricity is generated
- Energy goes to the electricity network, and it is used by customers (heating, light, fridges, …)

Wind energy is non-operable

- The main issue with wind energy (and also photovoltaic solar) is that it is not under the control of the operator.
- It depends on the weather.
- At this point in time, wind energy cannot be stored.
- This is a problem because of the way the electricity market works

The electricity market

- Every day at, let's say 12:00, energy providers must give a forecast about how much energy they are going to provide for the next day, for every hour (i.e. 0:00h, 1:00h, …, 23:00h).
- This is not a problem for traditional energy sources (gas, oil, …). For instance, if at time 3:00h the provider forecasted it is going to produce x energy units, all it needs to do is to burn the appropriate amount of gas.
- But this cannot be done for wind, because it depends on the weather at 3:00h.
- All the wind energy provider can do is to forecast the weather for the next day at 3:00h.

Weather forecasts

- Nowadays, there are advanced mathematical / physical / computational models (called Numerical Weather Prediction models), which are able to forecast the weather several days in advance.
- The *Global Forecast System* (GFS, USA) and the *European Centre for Medium-Range Weather Forecasts* (ECMWF) are two of the most important NWPs.
  - http://www.ecmwf.int/

ECMWF Meteorological variables

- Some of the variables predicted by ECMWF:
  - 2 metre temperature
  - 10 metre U wind component; 10 metre V wind component
  - **100 metre U wind component; 100 metre V wind component**
  - Convective available potential energy
  - Forecast logarithm of surface roughness for heat
  - Forecast surface roughness
  - Instantaneous eastward turbulent surface stress
  - Instantaneous northward turbulent surface
  - Leaf area index, high vegetation
  - Leaf area index, low vegetation
  - …
- However, the relation between those variables and the electricity actually produced is not straightforward. Machine Learning models can be used for this task

From meteo to energy

- We intend to train a machine learning model *f*, so that:
  - Given the 00:00am ECMWF forecast for variables $A_{6:00}$, $B_{6:00}$, $C_{6:00}$, … at 6:00 am (i.e. six hours in advance)
  - $f(A_{6:00}, B_{6:00}, C_{6:00}, \ldots)$ = electricity generated at **Sotavento** at 6:00

Sotavento (http://www.sotaventogalicia.com/en)

- Sotavento is a wind farm at Galicia (North West Spain)

- We will use two sources of data:
  - Meteorological variables come from ECMWF
  - Electricity production data comes from Sotavento

The data

- It is common practice to use meteorological variables in a grid around the desired location
- In this case, we will use a 5x5 grid (Sotavento is actually located at the center = 13)
- Therefore, there are 22 ECMWF variables, forecasted at 5*5=25 locations = 550 variables (input attributes). Quite a lot.
- The **energy** column in the dataset contains the outcome to be predicted (energy generated).