

Some important notes:

Renzo Miguel Caballero Rosas

January 27, 2020

With respect to all the wind data download:

We have all the files in the super folder:

1. Wind_Generation-Forecast_1.ipynb.
2. Wind_Generation-Forecast_2.ipynb.
3. Wind_Generation-Forecast_3.ipynb.
4. Wind_Generation-Forecast_All.ipynb.
5. Wind_Data_1.csv.
6. Wind_Data_2.csv.
7. Wind_Data_3.csv.

However, I will create new Python files to download all the needed data (I will start working on it the day 14/01/2020. The date is important because the webpage changes over time. It is <https://apps.ute.com.uy/SgePublico/ConsPrevGeneracioEolica.aspx>). The files have the main name **Wind_Generation-Forecast_Downloader.ipynb**. To the name, we add the data and source to differentiate between different files.

Making it automatic:

Each time we send the cURL, we receive the data for the next three days. For example, if we send a request for the day "03/12/2019" with the starting time "0700" (corresponding to 01:00 hrs), we would receive:

- 0700 - 03/12/2019
- 0800 - 03/12/2019
- 0900 - 03/12/2019
- ...
- 0500 - 06/12/2019

- 0600 - 06/12/2019

Notice that the last point of data is three days later, with one hour less. This is the date we will check to ensure we have or not some data points.

Initial document:

For the script to work , we need to create a **.csv** file with the next table inside:

empty	*empty*	Starting_Time	Min_ConfInt	Max_ConfInt	Forecast	Real
----------------	----------------	---------------	-------------	-------------	----------	------

All the data will be written below (in the next rows). If we open this file with a text editor, it should have written:

„Starting_Time,Min_ConfInt,Max_ConfInt,Forecast,Real***end of line***

Notice the end of line character ***end of line***. This means that it is necessary to press **Enter** after we finish writing the sentence.

What is the data?

We have real production from 3 of November 2016 until today, with a frequency of 6 measurements per hour.

For the forecasts, we have three companies: AWSTP, MTLOG, and UTEP5. Each one with a different starting day:

1. AWSTP \implies 23/04/2019 until today.
2. MTLOG \implies 08/01/2016 until today.
3. UTEP5 \implies 05/07/2017 until today.

Each company provides a 72 hrs forecast with a frequency of one measurement per hour. This forecasts are provided four times a day, at the times 0100, 0700, 1300, and 1900.

Something is wrong?

For MTLOG and UTEP5, all the data is from 0100, even the ones with different names. I have to download it all correctly later.

Matlab files:

- **dataManagment.m**: It loads all the .csv and .ods files, and creates lists and cells with the paths. All this information is stored in .m files.
- **dataConditioner.m**: It is to check the data, it has inside some verification tests.

What are the plots?

We have plots and data corresponding to MTLOG 0100 forecast and ADME real production. We have for 01:00 to 22:00 hrs, and for 01:00 to 07:00 hrs.

About **someResults**

In this folder, we have most of the plots and LaTeX files that explain visually what the data is. Inside we find:

- The PDF **windData.pdf**: In these slides, we explain some of the main issues of the raw data. In particular, we show the effect of the artificial intervention in the generation and a delay in the data.
- Folder **MTLOG_0100_2019**: In this folder, we have the plot for the 365 days for the year. We show how applying a delay in ADME, the real production from both, ADME and UTE, match. We plot from 01:00 to 22:00 hrs.
- Folder **histograms**: In this folder, we show how given some criteria, we remove the days where artificial intervention was done.
- Folder **dataToUse**: In this folder, we show for the 365 days the first 6 hours of real ADME production (with the delay correction) and the forecast.
- Folder **dataToUse_24**: In this folder, we have the plot for the 365 days for the year. We show how applying a delay in ADME, the real production from both, ADME and UTE, match. We plot from 01:00 to 01:00 hrs of the next day.
- Folder **dataToUse_24_corrected**: In this folder, we only show the days with no artificial effect. Also, all the plots are normalized. ADME is corrected.
- Folder **final**: In this folder, we show and explain precisely the data that will be used. The file allData.pdf contains all the information and statistical analysis of the data.