ANALYSIS OF GREEN ENERGY IN EUROPE
(SOLAR & WIND)

Professor Vijayalakshmi Saravanan

Jose David Cortes (jose.d.cortes@coyotes.usd.edu)
Bhulakshmi Kunchala (bhulakshmi.kunchala@coyotes.usd.edu)

Data Analysis, Decision Making, and Visualization

CSC-557-U15-92793

December 1, 2021

Abstract
The climate change is a multi-dimensional concept, which can be measured at perceived at different levels. One of those dimension is about the energy that human kind have been using the last century. The global energy markets are changing due to the clean energy transition. Is crucial to focus the attention in the way we generate the energy to avoid keep destroying the planet and preserve the human kind on earth.

According to the International Energy Agency, renewable energy already exceeded coal as the main source of power capacity in 2015[1]. In the recent "Clean for Energy for All Europeans" package, the European Commission has defined three main goals to keep the European Union (EU) competitive and to lead this transition: putting energy efficiency first, achieving global leadership in renewable energies, and providing a fair deal for consumers. Thus, placing renewable energy in a central position, Europe has set itself a target to collectively reach a share of at least 27% renewables in the final energy consumption by 2030. This could be translated as about half of the EU's electricity generation will come from renewables. However, the rapidly growing share of electricity production from intermittent renewable sources (wind and solar) increases the stochastic nature of the power system introducing instability to the system and high uncertainties to the market design. As a consequence, planning and scheduling tools for the power sector have been updated and the study of power systems with a high share of intermittent RES-E has become an established field in Power System Analysis. In particular, the adequate modeling of high RES-E penetration systems crucially depends on the accurate representation of the spatial and temporal characterization of the wind and solar sources. RES-E data inherently bears the risk of being imperfect, inappropriate, or incomplete which might lead to errors in power system studies which could be either overstating or downplaying the possible role of solar and wind energy in the future energy mix.

Literature review/existing work

-There are many articles, reports, and news about The European Green Deal[2]. Actually is one third of the 1.8 trillion euro investments from the NextGenerationEU Recovery Plan, and the EU's seven-year budget.

-The EU has adopted targets to achieve a 20% share of renewable energy in energy consumption by 2020, and 32% by 2030. Experiences gained in the early 2000s demonstrated the importance of enabling frameworks for renewables, and such frameworks remain at the heart of the EU's policy process[3].

Description of the Data, tools, and methodology

In this study, the general approach applied to convert solar resources into power generation consists in converting satellite-based radiation data using the PVGIS model (Figure 1). The first step of the methodology is the meteorological data treatment; in this case, it is necessary to calculate solar radiation satellite-based data into radiation on inclined planes. Then, the radiation data is converted into theoretical potential; i.e. the solar electricity generation in each area given by kW generated from each kW peak of a typical PV System. In this approach, a sensitivity analysis is carried out to assess the impact of the spatial distributions of the PV-modules at country and regional levels. Thus, the quality of the assumptions is also gauged to estimate the locations of the PV farms for each region. Finally, to obtain the power generation the installed capacity of each region is calculated and then, the time series are corrected with the TSO(transmission system operator) actual generation and statistically validated for power system analysis, by assessing the power peaks and ramps, duration curves and capacity factors.

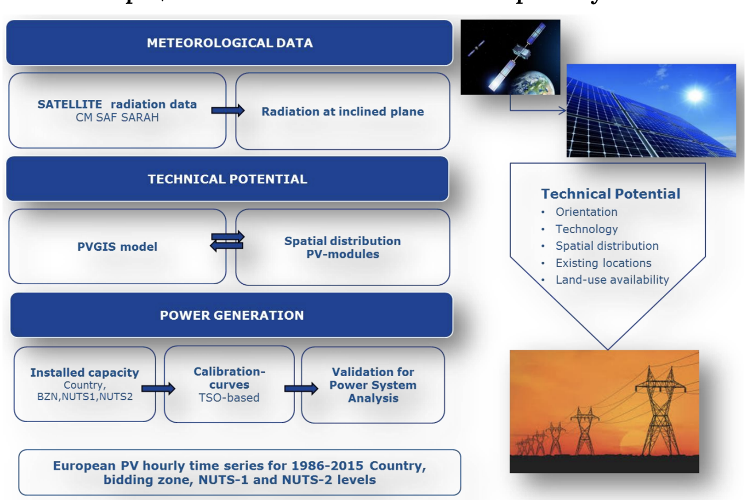

Figure 1. Summary of overall methodology to develop EMHIRES-PV dataset

EMHIRES dataset:

Solar Power generation.

European Meteorological derived High-resolution RES generation time series for present and future scenarios EMHIRES is the first publically available European solar power generation dataset derived from meteorological sources that are available at country, bidding zone, NUTS-1, and NUTS-2 level. It was generated applying using the validated and robust PVGIS model to estimate the solar electricity potential capturing local geographical information to generate meteorologically derived solar power time series at the high temporal and spatial resolution, validated with transmission system operators' data.

Win Power generation.

The general approach to convert wind resources into power generation consists in converting wind speed data from weather models or observations using power curves. The power curves, which are turbine-dependant, provide the value of electrical power output as a function of wind speeds at the hub height. The approach followed to develop EMHIRES converts the wind resources into power generation combining a high detailed wind farm database with a high spatial and temporal resolution wind speed dataset. The methodology used for EMHIRES is summarised in Figure 2.

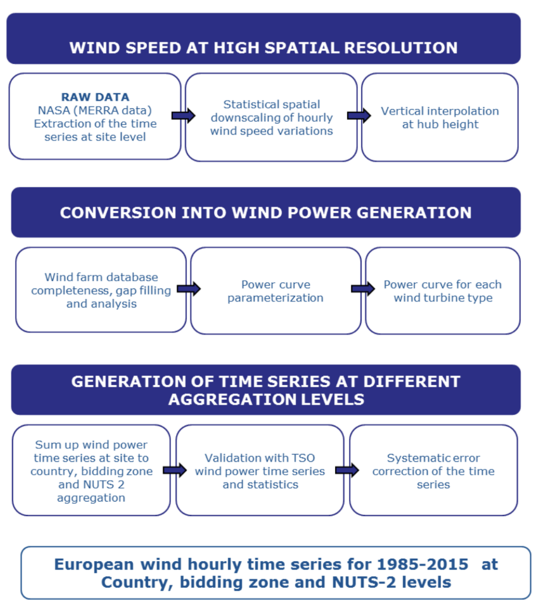

Figure 2. Summary of the steps followed to develop EMHIRES dataset

We find the dataset here: https://zenodo.org/record/4803353#.YaHtzvFKg-Q

We prepare an exploratory procedure divided in three stages(repository/pythonCode):

1. Creation of Cluster(FIPRODA_01.ipynb).
   We study the Solar Generation on a country level in order to make cluster[4] of country which present the same profile so that each group can be investigate in more details in a future project.
2. Data Analysis(FIPRODA_02.ipynb).
   In this second part, we have explored the data set in order to assess the impact of meteorological and climate variability on the generation of solar power. We've also shown the variation during the day, the months of the year and accross years. The dataset seems to be clean and a function to add date time informations is already implemented.
3. Power Plant Prediction for ES30(Madrid, Spain)(FIPRODA_03.ipynb)
   we have made an exploratory analysis of the data set. We were able to understand how solar generation is dependant of time.
   Now in this final step, we are going to training different types of machine learning models in order to make predictions. For that, we only keep one region 'ES30' corresponding to the Madrid(Spain) Area.
   As usual, we started with some data preparation: first we imported the needed libraries, then we import the .xlxs file into a data frame. After that we can reuse a previous function to add date-time information (year, month, week of the year, the day of the year, and the hour of the day)

Conclusions

* The Linear Regression is underfitting and doesn't perform well with our dataset.
* Using the Polynomial feature allows us to get better results with no underfitting this time but this is not enough compared to the baseline.

* Finally, the boosting family performs a little better but there isn't any real convincing difference compared to the baseline. It could have been interesting to tune hyperparameters with the RandomizedSearchCV or GridSearchCV methods, but I don't think we'll obtain a real gain.
* The prophet is also not suited here. Those models perform well when it comes to getting an overall shape but aren't suited when there is a short-term change due to the weather on the global tendency. That's why deep learning models are better here. Nevertheless, it should be noted to the R.N.N uses the history of the past two days, this can also explain why the last 3 models are more performant. The same history can probably be used with other models to obtain better forecasts.

From a general perspective:
* The RMSE is around 0.05 compared to values ranging from 0 to 1.
* Anyway, it is necessary to relativize because most of the values are under 0.7 and it should be put into perspective that many values are nulls.
* Finally, the RMSE with RNN is half the RMSE of the baseline which is quite optimistic.
* It could be interesting to tune the parameters of a few models to see if we gain a few percent(next time).

Bibliography

[1] T. Huld and A. M. Gracia Amillo, "Estimating PV module performance over large geographical regions: the role of irradiance, air temperature, wind speed, and solar spectrum," Energies, pp. 5159-5181, 2015.

[2]https://ec.europa.eu/info/strategy/priorities-2019-2024/european-green-deal_en Nov 29, 2021

[3]https://www.irena.org/publications/2018/Feb/Renewable-energy-prospects-for-the-EU Nov 27, 2021

[4] Patel. A, "Hands-On Unsupervised Learning Using Python" Machine Learning, pp. 124-147, 2019.