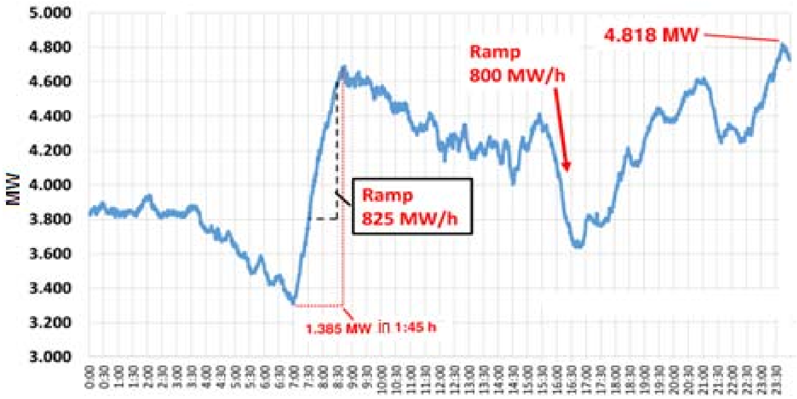
TIME SERIES LAB VERSION 1.0

Introduction

Wind and solar power plants are generally known as Variable Renewable Energy (VRE) sources, due to the strong volatility of the production in time. The intermittency of these sources is illustrated in the following picture, which shows wind power variability observed in the Brazilian northeast region on June 26th, 2016. As can be seen, the production can change significantly from one hour to the next.



Wind power variability observed in the Brazilian northeast region on June 26th, 2016.

The **Time Series Lab** (**TSL**) is a tool developed by PSR to produce future scenarios of intermittent renewable generation. The methodology consists in a non-parametric estimation method, called kernel density estimation, which aims at creating a statistical model that does not assume any specific distribution. Besides that, a Bayesian network is built so that the spatial and temporal correlation effects between VREs and hydro inflows may be captured.

In order to produce those future scenarios, the statistical model needs the historical renewable generation record. So, in the cases where the historical data is not available, the user would need to create it somehow, which is a very challenging task. Because of that, the TSL has a methodology that creates a synthetic hourly historical generation record using the information available at a global reanalysis database.

Before you dive into this document, let us introduce TSL's user interface:

Time Series Lab's user interface.

The TSL's installation setup is available in the following link (password: **apollo**).

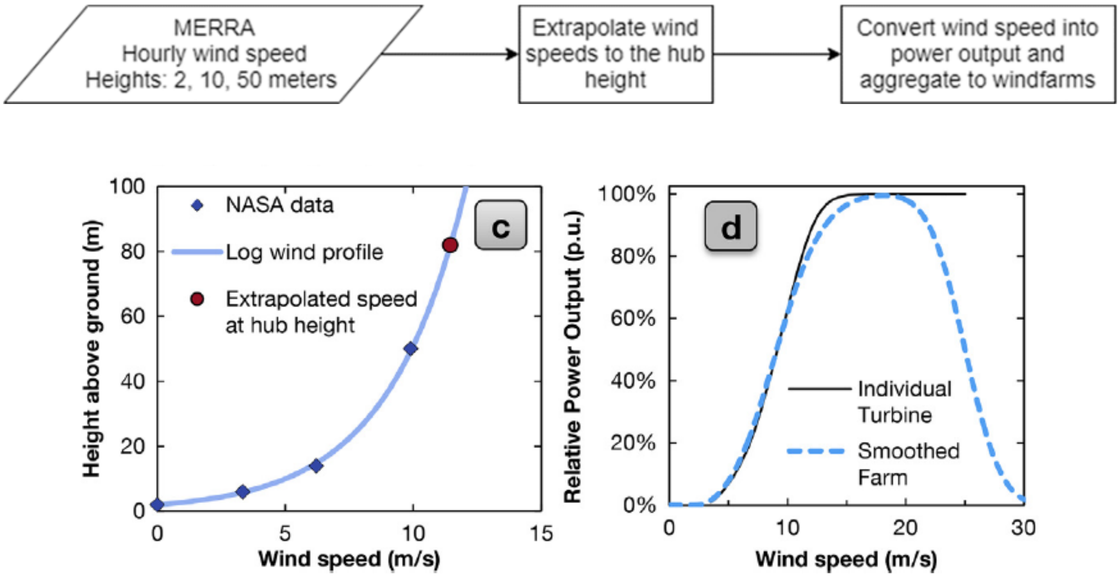
https://www.psr-inc.com/app/link/?t=d&f=timeserieslab-1.0-setup.zip

Creating a historical renewable record

Generally speaking, mesoscale climate models are used to produce wind velocity maps and time series with high spatial resolution, which in turn are suitable for assessing the penetration of renewable sources. However, if such data is not available, global reanalysis databases may be used as a reasonable alternative. The MERRA-2 database, for example, contains chronological data for wind speed, irradiation, and temperature from 1980 until nowadays [1].

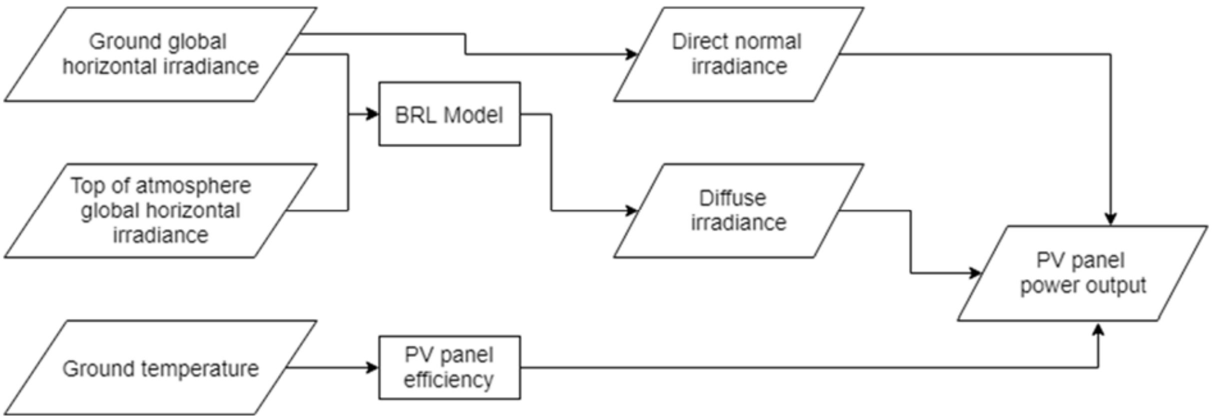
Reanalysis combines (i) historical meteorological observations with (ii) an atmospheric circulation model to infer the state of the global climate system. The model is programmed to replicate historical observations of satellites, field observations, ships, aircrafts, etc., simulating the past (hindcast), rather than a future forecast. The reanalysis process assimilates observed data and applies automated quality control to generate a standardized database with uniform spatial and temporal resolution and global coverage. Due to these characteristics, these bases are widely used in the energy sector to understand the availability of renewable wind and solar resources.

The **TSL-Data** module calculates the wind production through a model based on the Virtual Wind Farm (VWF) [2], developed by Renewables Ninja's team [3]. Once the wind speed data is extracted for the desired points, an additional function extrapolates the velocities to the height of the hub. The extrapolated wind speed is then converted into power generation based on capacity curves which vary according to the turbine model. Optionally, a "smoothing" factor can be applied to the capacity curve of a turbine to represent the effect of several turbines in a wind park, where there is variability in the production of the individual turbines.



Overview of the Virtual Wind Farm (VWF) model [2].

The solar production is based on the data of Global Horizontal Irradiation, i.e., the irradiation at the top of the atmosphere and the temperature extracted from MERRA-2. Taking this information into account the GSEE (Global Solar Energy Estimator) method [4] is then applied. GSEE is also developed by Renewables Ninja's team. It is possible to define the angles related to the panel inclination, as well as the tracking systems of one or two axes.



Overview of the Global Solar Energy Estimator (GSEE) model [4].

Although MERRA-2 is globally available at hourly basis since 1980 until today, it can contain systematic errors due to underlying weather forecast model errors. In addition, the spatial coarseness means the model is unable to resolve the detailed topography of a particular region, loosing speed-up and blockage effects. Finally, it is important to notice that the wind speed observations being used are primarily inferred by satellite data and ground observations made by short met masts which may not be precise to emulate the primary resource available for wind farms at the same locations. Therefore, correcting bias in reanalysis data is a fundamental step in modelling renewable energy output.

TSL-Data offers three choices for bias correction:

- **Single profile:** a constant capacity factor value for all the hours of all months;
- **Monthly profile:** one capacity factor value for each month; or
- **Hourly profile:** one typical day per month presenting an hourly profile (24 hours x 12 months = 288 capacity factor values).

For renewable plants which are already operating, observed historical data should be used to define these profiles that **TSL-Data** will apply in the bias correction phase. For future projects, the user may use historical data from the closest existing plant or just use the synthetic historical data (that came from MERRA-2 database) for bias correction. **TSL-Data** will fit two calibration parameters through a binary search approach in which it will match the renewable record simulated with MERRA-2 with the user-defined profiles.

Data sources

TSL-Data is based on the Renewables Ninja methodology: solar irradiance data is converted into power output using the GSEE model (Global Solar Energy Estimator) written by Stefan Pfenninger [4], and wind speeds are converted into power output using an adapted model based on VWF model (Virtual Wind Farm) written by Iain Staffell [2].

TSL-Data uses wind speeds, temperature and irradiance data from the NASA MERRA-2 reanalysis database [5].

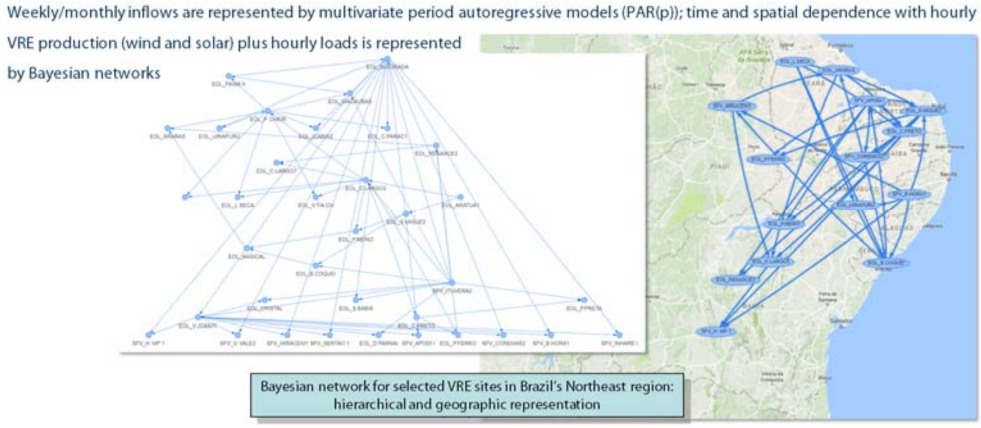
The wind speed map is obtained from the Global Wind Atlas 2.0, a free, web-based application developed, owned and operated by the Technical University of Denmark (DTU) in partnership with the World Bank Group, utilizing data provided by Vortex with funding provided by the Energy Sector Management Assistance Program (ESMAP). For additional information, please refer to [6].

Solar resource data obtained from the Global Solar Atlas, owned by the World Bank Group and provided by Solargis. For additional information, please refer to [6].

Generating synthetic renewable generation scenarios correlated with hydro inflows

Due to the spatial correlation of wind and solar production in different regions, as well as the spatial correlation between hydro inflows and wind speed in some regions, it is not adequate to model the generation scenario of each plant independently; it is necessary to represent the *joint probability distribution* of all intermittent renewable and hydro resources for both existing and future plants. In addition, this joint representation must be multiscale, that is, wind and solar are represented with hourly resolution, whereas hydro inflows are typically represented on a monthly or weekly basis.

The **TSL-Scenarios** module uses a Bayesian Network to produce those integrated multiscale wind, solar and hydro inflow scenarios. Bayesian Network is a statistical model that represents a set of variables and their *conditional dependencies* through a graph.



Bayesian network for selected VRE sites in Brazil's Northeast region.

These *spatial correlations* are measured in a monthly (or weekly) basis since hydro inflow data are usually available on these resolutions. Therefore, a Bayesian network per month (or per week) is built, so that the VRE scenarios for each month (or week) may be generated considering the spatial correlation.

Filling missing historical hydro inflow data

TSL-Scenarios uses a statistical method called Multivariate Imputation by Chained Equation (MICE) to fill missing data in the historical inflow record. This option is obligatory when users want to represent the correlation between inflows and intermittent renewable generation and for some specific gauging station(s) some historical data is missing.

It is worth mentioning that this option may also be used even if the user does not want to capture the aforementioned correlation effects since the inflow statistical model will have more samples to perform the parameters' estimation.

For more information, please check the TSL User Manual.

References

[1] https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/

[2] Staffell, Iain and Pfenninger, Stefan (2016). Using Bias-Corrected Reanalysis to Simulate Current and Future Wind Power Output. Energy 114, pp. 1224-1239. doi: 10.1016/j.energy.2016.08.068

[3] https://www.renewables.ninja/

[4] Pfenninger, Stefan and Staffell, Iain (2016). Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data. Energy 114, pp. 1251-1265. doi: 10.1016/j.energy.2016.08.060

[5] Rienecker MM, Suarez MJ, Gelaro R, Todling R, et al. (2011). MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. Journal of Climate, 24(14): 3624-3648. doi: 10.1175/JCLI-D-11-00015.1

[6] https://globalwindatlas.info