Prediction of Parking Area Solar Panel Electricity Generation at the University of Liège

A. Birtles, G. Delporte, R. Lambermont & A. Louis

Université de Liège

March 2024

Table of Contents

Introduction

We aim to predict the daily production of solar panels of two parking at the University of Liège.

Context

- The objective is to obtain solar panels predictions for the upcoming days to facilitate scheduling purpose (cost, savings, maintenance, ...).
- The users of our project are the administrators at the University in charge of the solar panels.
- Challenges reside in the lack of awareness regarding electricity production and the absence of a connection with the weather station.

The concept is to incorporate electricity production forecasts into a user-friendly calendar, streamlining user experience and facilitating convenient access to solar panel maintenance and usage.

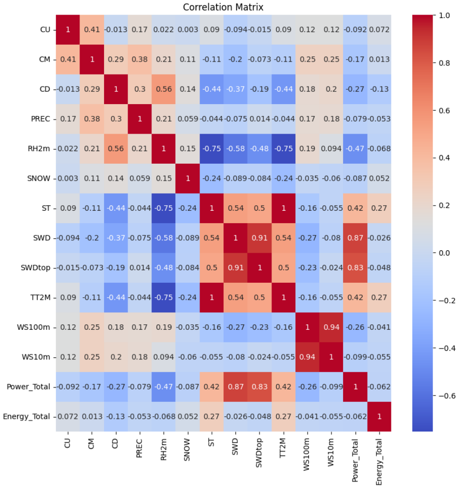
Data

The data for our project consist of two main components:

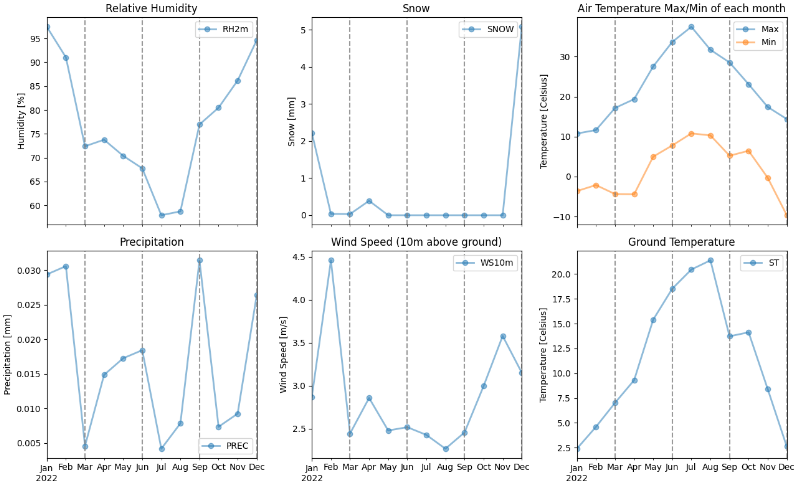
- Solar power production
- Forecast data from the laboratory of climatology of the university
    - This includes information on precipitations, relative humidity, surface temperature, snow, etc.

The feasibility of the project is ensured by the ample availability of labeled data. This abundance allows us to implement a straightforward machine learning model for accurate predictions.

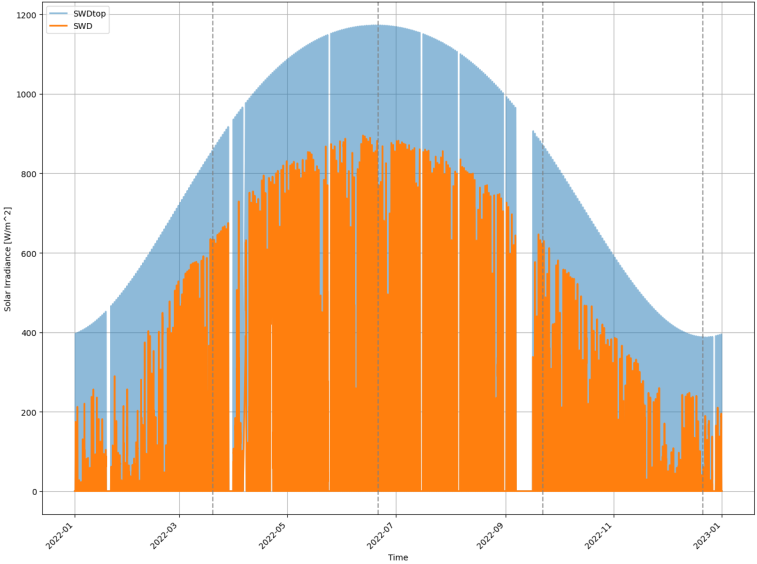
Exploratory and Data Analysis (1/4)



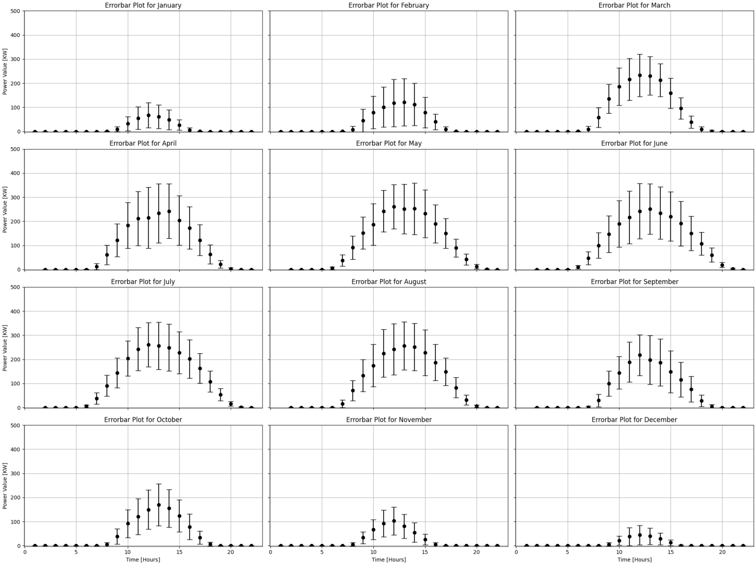
Exploratory and Data Analysis (2/4)



Exploratory and Data Analysis (3/4)



Exploratory and Data Analysis (4/4)



Models

We first tried to train some simple models from the scikit-learn library using a Train-Validation-Test split with ratios of 0.75 - 0.125 - 0.125:

- LinearRegression
- KNeighborsRegressor
- DecisionTreeRegressor
- RandomForestRegressor
- GradientBoostingRegressor
- SVR

Models (cont'd)

We used a 5-fold cross validation grid search to look for the best possible parameters for the KNeighborsRegressor and DecisionTreeRegressor as they are simple to train.

For the more complex problems such as RandomForestRegressor, GradientBoostingRegressor and GradientBoostingRegressor we simply used the basic parameters from the models defined in the scikit-learn library.

For the next milestone, we will try the RandomizedSearchCV to try reinforce the performance of the models that performed well (see next slide)

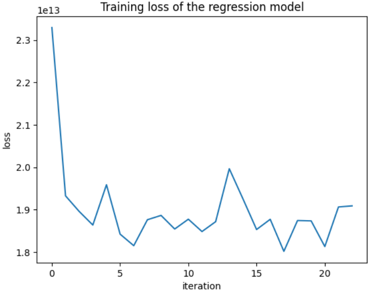
Models with Evaluation Metric

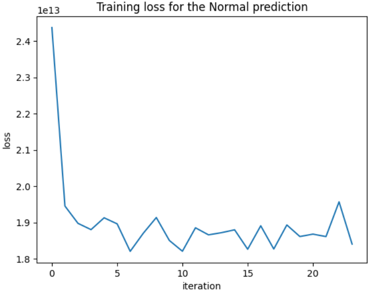
The Mean Square Error (MSE) metric is used to assess the effectiveness of models employed in predicting electricity production from the given dataset. We report the performance of the models fitted with the train data, chosen with the validation data and tested with the test data in the following table:

Solar Panel	Power 1	Power 2	Power 3	Power 4
Model	RF	RF	RF	RF
MSE	3.78×10^7	3.78×10^7	3.63×10^7	3.27×10^7

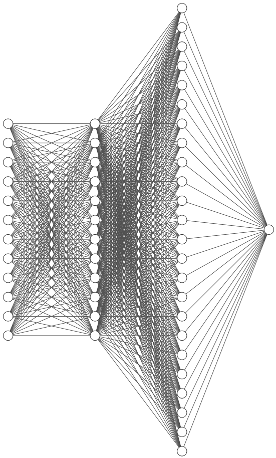
Solar Panel	Power 5	Power 6	Power 7	Power 8
Model	RF	RF	RF	RF
MSE	5.40×10^6	5.46×10^6	5.57×10^6	5.47×10^6

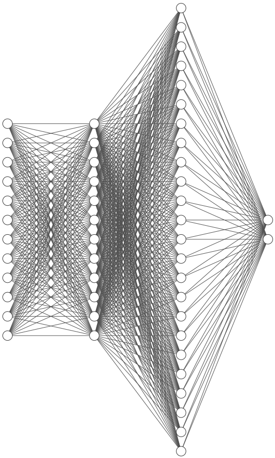
Neural Networks To test an other type of model, we have implemented 2 NN, the first one is a simple regression model with 2 hidden layers. The second one is the same model but the output models the mean and the standard deviation of a normal distribution.
Looking at the losses, we see that the model is not capable to learn enough.





NN Architecture





Conclusion

In conclusion, our project delves into the realm of probabilistic forecasting, offering the capability to showcase a range of potential solar energy production scenarios. This empowers decision-makers to:

- Make informed decisions
- Prepare for fluctuations (Ensure smooth energy flow by anticipating potential variations)
- Gain deeper understanding of the dynamics of solar energy production

This project goes beyond prediction, unlocking the full potential of solar energy through a data-driven, probabilistic approach.