Photovoltaic power generation forecasting models analysis *

BERMUDEZ Yaiza
G2ELab and INP
Grenoble, France
yaiza.bermudez@inp-grenoble.org

Supervised by: MOUNIE Gregory.

I understand what plagiarism entails and I declare that this report is my own, original work.

BERMUDEZ Yaiza, 17/04/2023,

Abstract

Solar energy is a highly sustainable energy source, but its production fluctuates based on weather conditions, posing challenges for energy management. Hence, accurate solar power generation forecasting is crucial to successfully integrate solar energy sources and optimizing the electrical usage. This paper aims to compare the effectiveness of data-driven models and theory-driven models using physics formulas. We propose and evaluate multiple data-driven models, such as Extreme Gradient Boosting (XGBoost) and Long-Short Term Memory (LSTM), considering their Root Mean Squared Error (RMSE), their adaptability to different locations based on data requirements, and their carbon footprint. Furthermore, we compare these data-driven models with global horizontal irradiance (GHI) clear-sky models, such as the Ineichen-Perez and Simple Solis models. The data-driven model have demonstrated to have better performance than the theory driven once with 9.4% RMSE against 29.65%.

1 Introduction

In our society, due to our lifestyle and population growth, using as much fossil energies is becoming an issue. Therefore, to satisfy the world's growing energy demand, meanwhile reducing the carbon footprint of the electrical sector, one of our best alternatives is solar energy. Indeed, solar energy is becoming an increasingly popular source of renewable energy worldwide. Solar power is a clean and sustainable source of energy that can be used to generate electricity. Unlike fuel-based energy sources, such as coal or natural gas, it does not release harmful pollutants into the environment. As a sustainable and adaptable source of energy, solar energy offers a promising solution to the challenges of climate change and the transition to a more sustainable future in terms of energy.Therefore, being able to forecast solar power generation could help people better optimize their electrical consumption based to the daily generation, for example.

This energy source being so important for the planet the aim of my internship is to first analyse existing solar power generation forecasting models, then try to implement a sustainable model for the needs of the lab whom goal is to use the forecasting to prove physical formulas. As well as evaluating the advantages of employing a data-driven model for solar power generation forecasting over another model. By using historical data, a data-driven approach can capture intricate patterns and relationships that may not be easily taken into account by the physics formulas. This can potentially enhance the accuracy and reliability of the forecasts. The evaluation should also consider the benefits and limitations of both approaches to determine the most suitable method. An additional goal for the forecasting model is to ensure its adaptability to different locations. This means that the model should be capable of accommodating specific characteristics of each location, allowing for accurate solar power generation forecasts regardless of geographical differences. Photovoltaic (PV) forecast models vary based on the available data and the forecast horizon we want to predict (short term, long term,...). In our case, we have access to weather data coming from the roof of the GreEN-ER building including solar radiation, cloudiness, wind, etc and we would like to be able to forecast for the next hour so at t + 1, t being the actual time.

We will first expose the state of the art in order to see what as already been done and which models could be used on the GreEN-ER building base on the available data and the wanted forecast horizon.

1.1 State of the art

Data driven models

The forecast of solar power generation is an increasing topic of discussion nowadays. Therefore, a lot of models using different techniques of machine learning and deep learning have been studied trough out the years. We establish with the following Figure 1 some of the models that can be interesting for us base on the data and horizon forecast. The models were selected regarding the data used, their forecasting horizon (hourly) and the year of the article (most recent ones) and

*These match the formatting instructions of IJCAI-07. The support of IJCAI, Inc. is acknowledged.

the aim of the table is to compare the accuracy.

Authors	Algorithm	Description	Metrics
[Pedro and Coimbra, 2015]	ANN and KNN	ANN was combined with k-nearest-neighbours (kNN) optimization to forecast global solar irradiance for time horizons from 15 min to 2h.	RMSE(%)<15.00
[Sharma et al. 2016]	WNN	A feed-forward ANN with Wavelet model was developed with Morlet and Mexican hat type activation functions. Its predictions were compared with persistence, ETS, ARIMA and ANN for temporal horizons of 15 min and 24h	nRMSE (%) 9.42 to 15.41
[Qing and Nui, 2018]	LSTM	A LSTM model developed for day ahead prediction of solar irradiance from weather data. Its performance is better than persistence, LLSR and multilayered FFNN-BPNN models.	RMSE 18.34%
[Dimitropoulos et al, 2021]	LSTM	LSTM is an artificial recurrent neural network (RNN) architecture used in the field of deep learning. LSTM has feedback connections and can process entire sequences of data.	RMSE 1.920
	SVM	Support-vector machines are supervised learning models with associated learning algorithms that analyse data for classification and regression analysis.	RMSE 2.014
	MLR	Linear regression is a linear approach to modelling the relationship between a scalar response and one or more explanatory variables.	RMSE 1.938
	XGBoost	Gradient ML technique for regression and classification problems, which produces a prediction model in the form of an ensemble of weak prediction models, typically decision trees. When a decision tree is the weak learner, the resulting algorithm is called gradient boosted trees.	RMSE 1.871

Figure 1: Current models and their accuracy

Some models use solar radiance forecasting rather than the PV generation [Ahmed *et al.*, 2020] this could be due to the fact that the production is strongly correlated to the solar radiance [Dimitropoulos *et al.*, 2021]. We still consider those models since we have access to solar radiance data. Based on the evaluation using the RMSE metric in Figure 1, XGBoost and LSTM models show better accuracy compared to other models. It's notable that both XGBoost and LSTM models have displayed superior accuracy when compared to other models, demonstrating their effectiveness in capturing complex patterns and relationships within the data. My supervisor has also oriented me towards N-BEATS as a potential model, that can offer a valuable contribution to further improving the accuracy of solar power generation forecasting.

Data processing

In addition to the model itself, we also need to consider the data used and its pre-processing. During our research, we came across a model utilizing auto-encoders, which seemed promising for identifying the most significant features in the data. However, our findings regarding this approach were inconclusive. We also explored the use of wavelets in other models [Almaghrabi *et al.*, 2022] , but due to the numerous types of wavelets available, we have not reached any definitive conclusions. To identify suitable models for our specific case scenario, we looked into the available data. Within the laboratory, we had access to a range of data from 2018 to 2022, including solar radiance, temperature, wind speed and direction, rainfall, and sky images captured above the building.

In the case of XGBoost, the model uses features such as temperature, pressure, humidity, and solar radiance to predict solar power generation. On the other hand, LSTM utilizes historical power generation data to capture temporal dependencies and make forecasts. By considering different sets of input variables, XGBoost focuses on weather-related factors that impact solar power generation, while LSTM emphasizes the historical patterns of power generation itself. Both approaches have demonstrated their efficacy in achieving accurate predictions [Dimitropoulos *et al.*, 2021].

Physical models

Over the years, numerous physical models have been proposed to accurately model solar irradiance under clear sky, which plays a crucial role in solar power production. These models have significantly contributed to our understanding of solar irradiance and have open the way for advancements in solar energy forecasting. By establishing an accurate model of solar irradiance and combining it with information on our panels characteristics, we can try to compute our overall energy production.

According to the suggestions provided in [Reno *et al.*,], it is noted that complex models heavily rely on local measurements that can be challenging to obtain. Hence, simpler models seem accurate enough for practical purposes. Considering this, even though the Bird clear sky model has demonstrated high precision [Mabasa *et al.*, 2021] ($RMSE \leq 5\%$), we have chosen not to implement it. Instead, models such as the Ineichen-Perez clear sky model ($RMSE \leq 10\%$), which requires nine parameters including solar zenith, azimuth, latitude, longitude, altitude, etc., or the simplified Solis model ($RMSE \leq 10\%$), which has shown coherence between irradiance components and solar elevation angles [Ineichen, 2016] and requires six parameters. These two models seem to offer a good balance between accuracy and ease of implementation.

The pvlib Python library provides access the above models for solar irradiance estimation under clear sky. In particular, it offers the Hay Davies [Mesri-Merad *et al.*, 2012] model as the default for calculating solar irradiance . Additionally, the library includes the isotropic model, which has demonstrated satisfactory results according to [Loutzenhiser *et al.*, 2007].

1.2 Motivation

As previously mentioned, solar energy is a significant and sustainable energy source. Accurate forecasting of solar power generation can greatly enhance energy management and facilitate the adaptation of its origin to optimize overall energy utilization.

The ultimate objective is to transition to cleaner energy sources and align our consumption accordingly. Therefore, it is important to avoid using forecasting models that consume excessive time and energy. A key challenge lies in striking a balance between the energy impact of the forecast model. As well as, to establish reliable and accurate forecasts, due to its dependency to weather conditions, solar energy is susceptible to uncertainty, making it challenging.

Therefore, we will look into the cost of the chosen algorithms, the quality of the model is not only base on its accuracy but on its accuracy over its training cost [Tran *et al.*, 2022]. We need to be careful with that, if it's really costly to train the model but it only has to be done ones maybe this could be alright. On the contrary if the model has to be retrained before every prediction we might as well go for a less efficient but cheaper model to train.

One limitation that needs to be addressed is the availability of data, particularly because we aim to propose an adaptable model that can accommodate to different locations. Achieving higher accuracy in solar power forecasting often requires a significant amount of historical data, which may not be available for every site. Thus, ensuring the model's adaptability across multiple locations becomes a challenging aspect to consider when using data-driven models.

2 Models Description

Data Processing

The data used for our analysis comes from the weather station on the rooftop of the GreEN-ER building. We have used solar radiance, humidity, pressure, temperature as well as the power production data from 2020 until now. Before training, the data is cleared of any inconsistencies or outliers. Once cleaned, the data is normalized to a common scale between 0 and 1, using min-max scaling. This ensures that all features contribute equally to the model training process, preventing any particular feature from dominating the learning process. By cleaning and normalizing the data, we create a standardized and reliable input for training our model.

In our physical model, we rely on parameters such as latitude, longitude, altitude, angle, and power of reference of the panels. We do not require data measured by a weather station for this purpose. Instead, we utilize models from the PVLib library to calculate additional information, such as the solar zenith angle throughout the day.

Extreme Gradient Boosting

Extreme Gradient Boosting (XGBoost), is an open source machine learning library designed to perform gradient boosting for decision tree (GBDT) models. It provides parallel tree boosting and is the leading machine learning library for regression, classification, and ranking problems.

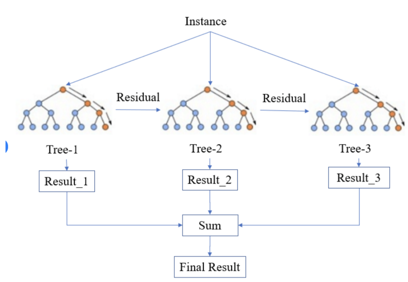

Figure 2: A general architecture of XBoost [1]

As we can see from Figure 2 in XGBoost every tree learns from the residual of the previous one therefore the results is construct based on the sum of the sum of all results, this is why this model is so efficient.

[1] Image from: https://www.researchgate.net/figure/Simplified-structure-of-XGBoost_fig2_348025909

Long-Short Term Memory

Long Short-Term Memory (LSTM) is a recurrent neural network (RNN) architecture that comprises four feed forward neural networks. Each of these neural networks is composed of an input layer and an output layer, with connections between input neurons and output neurons, as we can see in Figure 3. This arrangement leads to a fully connected structure with four layers in the LSTM unit. LSTM is widely employed for time series forecasting due to its effectiveness in capturing long-term dependencies and handling sequential data.

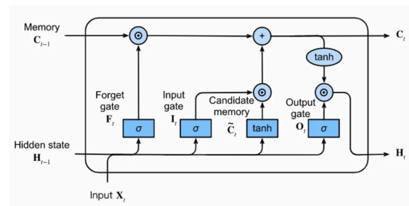

Figure 3: A general architecture of LSTM [2]

Neural Basis Expansion Analysis for Time Series

My supervisor oriented me towards the N-BEATS (Neural Basis Expansion Analysis for Time Series) [Oreshkin *et al.*, 2020] model which is a type of Neural network. N-BEATS is a fast Deep Learning model that recreates the mechanisms of statistical models using double residual stacks of fully connected layers as we can see in Figure4. This model has shown to be really efficient for time series forecasting.

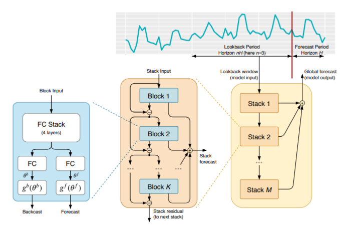

Figure 4: Architecture of the N-BEATS model[3]

The same way as XGBoost, N-BEATS uses the residuals from the previous block to learn on the next, making this model also really accurate and efficient.

Ineichen-Perez clear sky model

The Ineichen clear-sky model is used to estimate the Global Horizontal Irradiance (GHI) and the Direct Horizontal Irradiance (DHI) under clear-sky conditions. It takes into account the extraterrestrial solar radiation (DNI_{TOA}), and the position of the sun (θ_Z) . The model is based on empirical relationships and can be expressed as follows: [Reno *et al.*,]

$$GHI = a_1 \cdot DNI_{TOA} \cdot \cos(\theta_Z) \cdot e^b \tag{1}$$

[2] Image from: https://medium.com/@ottaviocalzone/an-intuitive-explanation-of-lstm-a035eb6ab42c

[3] Image from : [Oreshkin *et al.*, 2020]

Where:

$$b = -a_2 \cdot AM \cdot (fh_1 + fh_2 \cdot (T_L - 1)) \quad (2)$$

AM is the absolute atmospheric air mass,TL is the linke turbidity from [Ineichen and Perez, 2002], h represents the altitude above sea level and $a_1 = 5.09 \times 10^{-5} h + 0.868$, $a_2 = 3.92 \times 10^{-5} h + 0.0387$, $fh_1 = e^{-h/8000}$, $fh_2 = e^{-h/1250}$.

$$DHI = GHI - DNI \times \cos(\theta_Z) \quad (3)$$

We used the python library PVlib to compute all the parameters of equation 1, as well as the equation itself. The model in pvlib computes GHI, DNI and DHI from (1), (3) and DNI (direct normal irradiation) with DNI_{TOA} .

Simplified solis clear sky model

The simplified Solis model is a simplified version of the original Solis model, which is a well-known model used for solar irradiance estimation. It aims to provide a much easier computationally efficient approach while still maintaining reasonable accuracy.

$$GHI = DNI'_{TOA} \cdot e^{\frac{-\tau_g}{\cos(\theta_z^g)}} \cdot \cos(\theta_z) \quad (4)$$

Where DNI'_{TOA} is a modified version of DNI_{TOA} that takes into account different altitude and atmospheric conditions, τ_g corresponds to the total optical depth. We used PVlib to calculate all the needed parameters as well as the (4).

3 Methodology

Data driven models

We divided it into training and test set. For the XGBoost model, the training set consisted of the before mentioned weather data. As for LSTM, it gets the historical production of the solar panels. Since the historical power production data from the panels on the roof only begins in September 2022 (the month they were installed), we trained our model initially on smaller panels not located on the roof. These smaller panels provide historical power production data up until 2020.

We implemented the data-driven models, including XGBoost and LSTM, using the sklearn library in Python. The N-Beats model was implemented by my supervisors based on [Oreshkin *et al.*, 2020].

To optimize the hyperparameters of XGBoost, we utilized Grid Search. Through this process, we determined that the following hyperparameters were the most effective in our case: learning rate = 0.03, max depth = 5, and number of estimators = 50. The training lasted approximately 1 minute. The model is trained from 2020-2021 and predicts on 2022 with a one hour time step.

For the LSTM model, we utilized a single layer with 128 units. Additionally, we included a dense layer with a rectified linear unit (ReLU) activation function. The model was trained for 50 epochs with a batch size of 32. We also incorporated early callback with a patience value of either 5 or 10 to prevent over fitting. The training lasted approximately 1 to 2 minutes. The model is trained on data from 2021 and learns from the past 24 to predicts the next one.

The N-Beats model architecture includes a 1 block structure, consisting of a residual dense layer with 4 layers of 64 neurons each, followed by 2 dense output layers with 24 neurons each. A stack is formed by combining 5 blocks. The full N-Beats model is composed of 5 stacks. During the training process, the model was trained for 20 epochs with a batch size of 64. The Adam optimizer was used, and the mean squared error (MSE) was employed as the loss function. The training is done over the historical data of production, it looks at the past 24 hours and predicts the next 24hours, but we shift it by 1 hour so we actually only keep the next hour. The train took approximately 2-3 minutes to complete.

Theory driven models

As mentioned before we used the python library pvlib to implement the physical models. We first implemented a solar module similar to the panels installed on the roof of GreENER, to compute the power production based on the clear sky models.

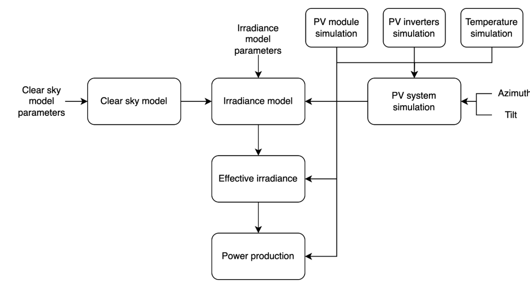

Figure 5: Diagram of the steps to compute solar power production using the physical formulas in pvlib

As shown in Figure 14 the clear sky model, either ineichen or simplified solis, is used to calculate the irradiance, since the model need GHI, DNI, and DHI for its computation. Then we compute the effective irradiance using pvlib and finally the power production.

To accurately simulate the PV system, we considered the tilt and orientation of the solar panels. Since the panels on the roof do not lay horizontally and have different orientations, we divided them into groups based on the inverter they are linked to. We separately computed the solar production for the panels facing the south-west direction (139 degrees) and the panels facing the north-east direction (-41 degrees) This distinction allows us to account for the different sun exposure throughout the day. After calculating the solar production for each group, we sum up the results to obtain the total production of all the panels on the roof. This approach enables us to accurately estimate the overall solar power generation of the PV system.

Evaluation

We utilize the Root Mean Square Error (RMSE) metric, for model evaluation. The RMSE provides a measure of the average magnitude of the prediction errors, allowing us to assess the accuracy of our model's predictions.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \quad (5)$$

Where N is the number of data points, y_i is the true value and $\hat{y}_i$ is the predicted value.

To facilitate the comparison of prediction errors across tests on different solar panels production, we calculate the percentage of the RMSE. This involves dividing the RMSE value by the max of the true values and multiplying by 100.

Visualizing the model's predictions and comparing them with the actual data can provide valuable insights into its performance. Therefore, in addition to considering the RMSE, it is beneficial to analyze the results graphically to better understand the behavior and accuracy of the models.

As mentioned earlier, while accuracy is an essential consideration, it is not the only factor in evaluating the models. We also place significant importance on developing a sustainable and adaptable model. These criteria play a crucial role in our evaluation process to ensure that the selected model aligns with our goals of long-term environmental sustainability and flexibility in accommodating diverse conditions and locations.

4 Results

During our experimentation with data-driven models, we explored different approaches for training and testing. We tested scenarios such as training the models on data from 2020/2021 and testing on data from 2022, as well as training on data from 2021 and testing on data from 2022.

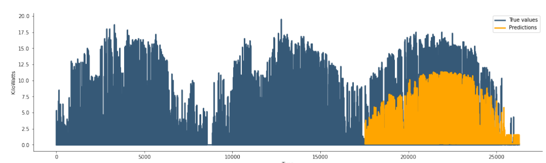

Figure 6: 1h time step predictions of XGBoost trained on 2020/2021

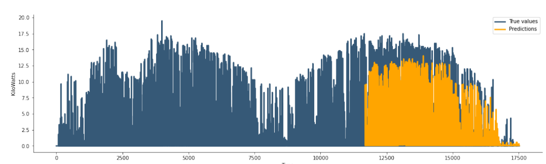

Figure 7: 1h time step predictions of XGBoost trained 2021

Although the computed RMSE may suggest certain trends (Figure 11), it is also important to visually analyze the results. In Figure 6 and Figure 7, we can observe the performance of the XGBoost model. By examining these figures, we can deduce that training the XGBoost model on the data from 2021 leads to better accuracy when predicting the solar power generation for the year 2022.

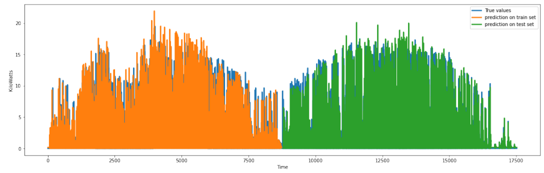

Figure 8: LSTM predictions on 2022 on the smaller panels

On Figure 8 and 11 we can see the results of the plotting and RMSE of the LSTM model, it is so far the most accurate model we have to test its adaptability and accuracy we train it on the smaller panel and test it on the once on the roof. We can visualise this results on Figure 9 and 11.

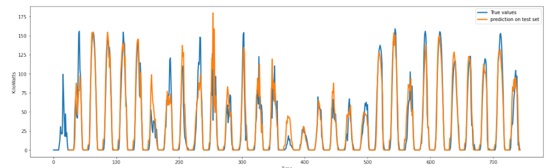

Figure 9: LSTM predictions on 2022-2023 on the roof panels

N-BEATS is the most complex model we have test as well as the most performant one, we can observe on Figure 10, the prediction of N-BEATS compared to the real values of the production of the roof panels.

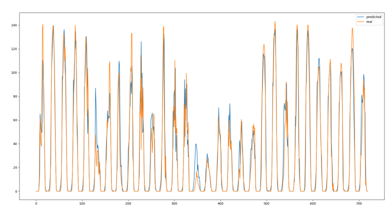

Figure 10: N-BEATS predictions 2022-2023 data for the roof panels

Model	Train set on small panels	Test set	RMSE
XGBoost	2020-2021	2022	2.63 15.04%
	2021	2022	2.86 16.35%
LSTM	2021	2022	2.51 14.36%
		roof 2022-2023	16.69 11.63%
N-Beats	2021	2022	1.74 9.94%
		roof 2022-2023	13.49 9.40%

Figure 11: Comparison of the data model based on their RMSE

From Figure12, Ineichen clear sky model with the isotropic irradiance model is the most accurate among the theory-

Clear sky	Get - irradiance	RMSE
Ineichen	haydavies	30.017% RMSE : 43.98
	isotropic	29.658% RMSE : 43.46
Simplified solis	haydavies	33.39% RMSE : 48.93
	isotropic	32.967% RMSE : 48.31

Figure 12: Comparison of the physical model based on their RMSE

driven models. This model requires few data inputs, that can all be computed using the pvlib library, making it a promising option for solar power generation forecasting.

5 Discussion

Overall, while XGBoost demonstrates good accuracy, its reliance on historical weather data and the need for storage, limit its adaptability to different locations. This model's predictions is specific to the solar plant and the weather conditions on which it was trained, making it less suitable for generalization. We have conducted tests on the roof panels, and the results were consistent with our expectations, as the model's performance was not good.

Comparing the results of our XGBoost model with the XGBoost model from [Dimitropoulos *et al.*, 2021] would not be fair or meaningful because the models have been trained on different datasets. The data here relies on the location we can not conclude on which model is more accurate.

LSTM and N-BEATS have demonstrated a strong capability for adaptation, which was anticipated due to its learning approach. By using the data from the previous 24 hours of production for the predictions, it can capture and learn the trends. Additionally, LSTM is relatively small with only one 128-layer making it a good fit for our needs, since it provides a balance between complexity, performance and adaptability. On the other hand N-BEATS is much bigger but its training time is comparable to LSTM's. However, to determine the "best" model, we would need to further analyze LSTM and N-BEATS' footprint. This includes assessing the computational resources and memory requirements of both models.

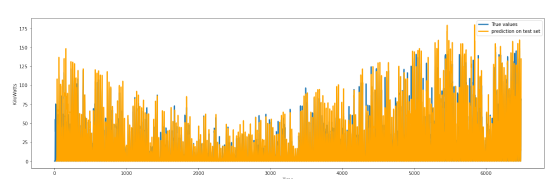

Figure 13: LSTM predictions over 2023 data

In Figure 13 and 14, we can compare the predictions of LSTM and Ineichen-Perez models on the same production data. As expected, LSTM shows better prediction performance. This is because the physical model relies on clear sky principles and does not consider cloudiness, thereby limiting

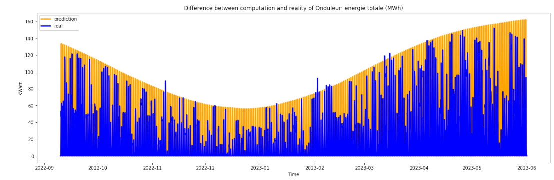

Figure 14: Ineichen-Perez with Isotropic irradiance model over 2023 data

its accuracy. On the other hand, LSTM takes into account the previous 24 hour production and can capture the variations of weather conditions more effectively.

Although the physical model Ineichen-Perez is less accurate compared to data-driven models, it offers the advantage of adaptability. The model only requires latitude, longitude, altitude, and the characteristics of the solar panels, making it suitable for various locations. Furthermore, since it does not rely on measured data like data-driven models, it has a smaller environmental footprint, contributing to its sustainability.

6 Conclusion

In this paper, we compare data driven model to physical model approach for solar power forecasting. Data-driven models have consistently demonstrated superior performance compared to theory-driven models in forecasting solar power generation. These models leverage historical data and advanced machine learning techniques to capture complex patterns and relationships, resulting in more accurate predictions. Their ability to adapt to changing conditions and learn from large datasets makes them highly effective in forecasting solar power generation. In contrast, theory-driven models have some limitations but also advantages. While they may not always achieve the same level of accuracy as data-driven models. These models can be particularly useful in situations where historical data is limited or unavailable. However, it is important to acknowledge that theory-driven models may have assumptions and simplifications that can introduce uncertainties, especially when considering factors like shading, cloudiness, wind speed, and air temperature. To mitigate these uncertainties, a hybrid approach that combines both data-driven and theory-driven models can be employed. By incorporating data-driven techniques to account for specific environmental factors, the hybrid model can potentially enhance accuracy and reliability in solar power generation forecasting. Furthermore, adopting a hybrid approach that combines both theory-driven and data-driven models would offer a balanced solution. This approach would leverage the adaptability of the physical model, the performance of the data-driven model, and minimize the data requirements, resulting in a smaller footprint. Given more time, implementing and testing this hybrid model would have been a valuable endeavor, as it could potentially provide more accurate and reliable solar power generation forecasts while considering the specific characteristics and limitations of each approach.

Acknowledgments

We would like to express our sincere gratitude to Vincent Imard, Engineer at G2ELab, and Benoit Declichant, Researcher at G2ELab, for their continuous support, guidance, and mentorship throughout the internship. Their knowledge and expertise in the field of solar power generation forecasting were instrumental in the development and implementation of the project.

We would also like to extend our thanks to Gregory Mounié, my supervisor from INP, for his support and guidance during this internship.

Furthermore, we would like to acknowledge the PhD students who provided valuable input and assistance during the course of this internship. Their collaboration and discussions enriched the project and helped shape its outcomes.

Lastly, we would like to thank the welcoming laboratory for providing the necessary resources and support for conducting this research.

A GreEN-ER Building

During the course of my internship, we utilized the data collected from the weather station at GreEN-ER (Figure 16). This data served as the basis for our analysis and model testing. Additionally, we evaluated our models using the production data obtained from various solar panels within the building.

The building itself features a rooftop solar panel installation with a capacity of 183 kilowatts peak, as depicted in Figure 15. This installation plays a significant role in our research and analysis. Furthermore, there is an additional solar panel installation adjacent to the building, with a capacity of 20 kilowatts peak, as shown in Figure 17. These different panel installations allowed us to evaluate our models under varying conditions and capacities.

Due to their location, the smaller panels are more susceptible to shadowing caused by the surrounding buildings. This shading phenomenon can have a significant impact on the performance and efficiency of the smaller panels. Therefore, because of this and the lack of historical data from the rooftop panels we used the smaller panels for training and both panels for testing.

Figure 15: The panels from the roof of GreEN-ER[4]

[4]Image from : https://www.univ-grenoble-alpes.fr/inauguration-de-la-centrale-photovoltaique-de-green-er-1re-experience-d-auto-consommation-collective-dans-une-universite-1226307.kjsp

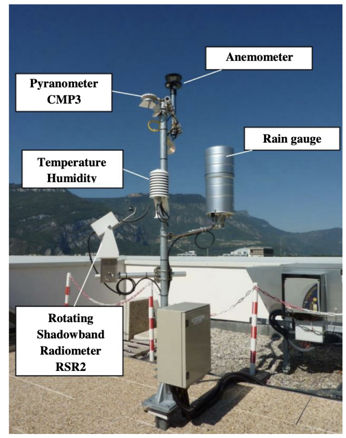

Figure 16: Weather staion from the roof of GreEN-ER

All the rooftop PV cells have the same inclination, 6 degrees and are oriented towards two different directions. One side is north-est and the other is south-west.

Figure 17: The smaller panels next to GreEN-ER

B Some physical explanations

In our theory-driven model, we incorporate several physical variables to accurately estimate solar power generation. These variables include Global Horizontal Irradiance (GHI), Direct Normal Irradiance (DNI), and Diffuse Horizontal Irradiance (DHI).To better understand the significance of these variables, we can visualize them in Figure 18. This figure provides a graphical representation of the GHI, DNI, and DHI.

The Python library pvlib uses various inputs such as the azimuth and tilt of the solar panel, altitude, and other parameters to calculate the solar position throughout the day, such as the zenith angle. To better understand the concept of the zenith angle, we can refer to Figure 19. This figure provides a visual representation of the sun's position in the sky, with the zenith angle being the angle between the vertical line and the line connecting the panel to the sun.

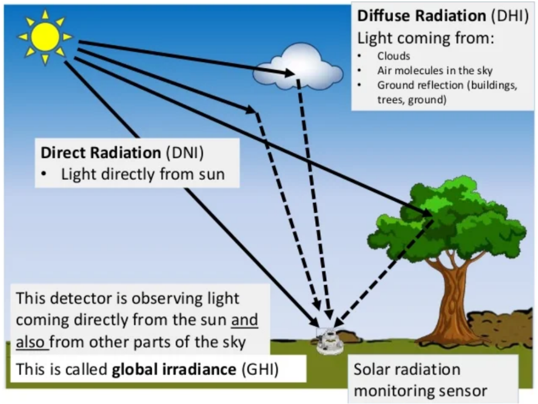

Figure 18: Diagram of solar irradiation partition [5]

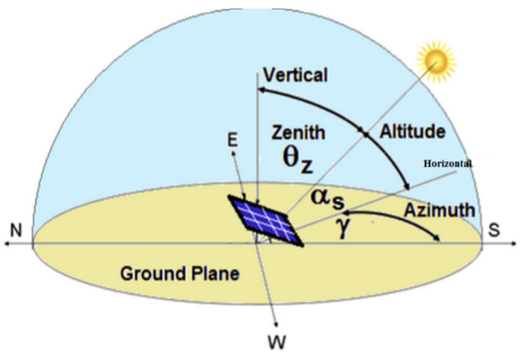

Figure 19: Variables computed from the solar position [6]

References

[Ahmed *et al.*, 2020] R. Ahmed, V. Sreeram, Y. Mishra, and M.D. Arif. A review and evaluation of the state-of-the-art in pv solar power forecasting: Techniques and optimization. *Renewable and Sustainable Energy Reviews*, 124(26), February 2020.

[Almaghrabi *et al.*, 2022] Sarah Almaghrabi, Mashud Rana, Margaret Hamilton, and Mohammad Saiedur Rahaman. Solar power time series forecasting utilising wavelet coefficients. *Neurocomputing*, 508:182–207, October 2022.

[Dimitropoulos *et al.*, 2021] Nikos Dimitropoulos, Nikolaos Sofias, Panagiotis Kapsalis, Zoi Mylona, Vangelis Marinakis, Niccolo Primo, and Haris Doukas. Forecasting of short-term pv production in energy communities through machine learning and deep learning algorithms. In *Proceedings of the 12th International Conference on Information, Intelligence, Systems Applications (IISA)*, 2021.

[Ineichen and Perez, 2002] Pierre Ineichen and Richard Perez. A new airmass independent formulation for the linke turbidity coefficient. *Solar Energy*, 73(3):151–157, September 2002.

[Ineichen, 2016] Pierre Ineichen. Validation of models that estimate the clear sky global and beam solar irradiance. *Solar Energy*, 132:332–344, 2016.

[Loutzenhiser *et al.*, 2007] P.G. Loutzenhiser, H. Manz, C. Felsmann, P.A. Strachan, T. Frank, and G.M. Maxwell. Empirical validation of models to compute solar irradiance on inclined surfaces for building energy simulation. *Solar Energy*, 81:254–267, 2007.

[Mabasa *et al.*, 2021] Brighton Mabasa, Meena D. Lysko, Henerica Tazvinga, Nosipho Zwane, and Sabata J. Moloi. The performance assessment of six global horizontal irradiance clear sky models in six climatological regions in south africa. *Energies*, 2021.

[Mesri-Merad *et al.*, 2012] M. Mesri-Merad, A. Cheknane I. Rougab, and N.I. Bachari. Estimation du rayonnement solaire au sol par des modèles semi-empiriques. *Revue des Energies Renouvelables*, 15:451 – 463, 2012.

[Oreshkin *et al.*, 2020] Boris N. Oreshkin, Dmitri Carpov, Nicolas Chapados, and Yoshua Bengio. N-beats: Neural basis expansion analysis for interpretable time series forecasting. In *Proceedings of a conference paper at ICLR*, 2020.

[Pedro and Coimbra, 2015] Hugo T.C. Pedro and Carlos F.M. Coimbra. Short-term irradiance forecastability for various solar micro-climates. *Solar Energy*, 122:587–602, December 2015.

[Qing and Niu, 2018] X. Qing and Y. Niu. Hourly day-ahead solar irradiance prediction using weather forecasts by lstm. *Energy*, 148:461–468, April 2018.

[Reno *et al.*,] Matthew J. Reno, Clifford W. Hansen, and Joshua S. Stein. Global horizontal irradiance clear sky models: Implementation and analysis. https://www.osti.gov/servlets/purl/1039404.

[Sharma *et al.*, 2016] Vishal Sharma, Dazhi Yang, Wilfred Walsh, and Thomas Reindl. Short term solar irradiance forecasting using a mixed wavelet neural network. *Renew Energy*, 90:481–492, May 2016.

[Tran *et al.*, 2022] Duc Hoang Tran, Van Linh Nguyen, Huy Nguyen, and Yeong Min Jang. Short-term solar power generation forecasting using edge ai. In *Proceedings of the 13th International Conference on Information and Communication Technology Convergence (ICTC)*, 2022.

[5] Image from : https://www.vebuso.com/2022/04/forecasting-solar-radiation-using-datarobot-to-optimize-power-generation/

[6] Image from : https://www.researchgate.net/figure/Azimuthal-zenith-and-hour-angles-for-solar-radiation-evaluation_fig4_304628964