# Insights on Machine Learning Models for Solar and Wind Energy

**Amlan Chakrabarti**

**A.K. Choudhury School of Information Technology**
**University of Calcutta. India**

# Abstract

**The use of renewable energy (RE) has recently increased worldwide.**

**Solar and wind, notably are the most emerging source of alternative sources of energy. In regards to total energy consumption, India is ranked third after China and the United States of America (USA).**

**The key challenge is to develop machine learning models for accurately forecasting renewable energy, allowing grid operators and energy managers to adjust energy supply and demand accordingly to improve grid stability and reliability.**

# Learning Objective

- Indian Power Scenario
- Motivation
- Challenges
- Global Research Status
- Case Studies
- Future Scope
- References

# Indian Power Scenario

**403 GW Installed** | **86% from Wind & Solar Generation**


2nd Largest source of Power Generation in India is RE (28%)

Data is as on May 2022

Taken form ** National institute of Wind Energy (NIWE)

# Motivation

- India is world's **3rd largest** renewable energy producer with **40%** of energy capacity installed in the year 2022 (**160 GW of 400 GW**) coming from renewable sources [1][2].
- India aims to achieve **175 GW** of renewable energy capacity by 2022, with **100 GW** from solar, **60 GW** from wind.
- India has committed for a goal of **500 GW** renewable energy capacity by 2030 [3].
- Accurate forecasting of renewable energy allows grid operators and energy managers to adjust energy supply and demand accordingly.
- To improve grid stability and reliability.
- India is committed to reducing its carbon emissions and meeting its climate change goals. Renewable energy forecasting can help in achieving these goals.
- Collaboration with **National Institute of Wind Energy (NIWE)**.
- NIWE has interest in short-term (1 hr to 4 hrs ahead) wind forecasting.
- Green energy corridor to facilitate Grid integration of large scale renewables.

# Challenges

- **Weather variability:** India's weather is highly variable and unpredictable, which makes it challenging to accurately predict renewable energy generation.
- **Lack of Sufficient Data:** There are few weather stations and wind measuring devices, and the data collected is often of low quality and insufficient to develop accurate forecasting models.
- **Lack of Standardization:** lack of standardization in the data collection and forecasting methodologies, making it difficult to compare results across regions and technologies.
- **Inadequate Forecasting Models:** India lacks the technical expertise and tools to develop sophisticated forecasting models. The existing models are often outdated and not suitable for India's weather patterns and renewable energy sources.
- **Policy and regulatory challenges:** India's renewable energy sector is subject to complex policies and regulations.

# Global Research Staus

- **According to Table 1:**
- Recent papers adopted Bidirectional LSTM and reported better results.
- Most of the studies limited to one solar station.
- There are very limited studies have been conducted for India.
- According to [10], 42% of the analyzed articles developed hybrid approaches, 83% performed short-term prediction.
- According to [11], research topics such as spatial forecast verification or forecast downscaling are to be tackled.

**Table 1.** Studies based on currently implemented models on solar irradiation prediction.

| Citation | Data | Model Name | Forecasting Window | Country | Correctness | A/D |
|---|---|---|---|---|---|---|
| [8] | Time Series | LSTM (Unidirectional) | Hourly | Egypt | The claimed forecasting error is 82.15, and 136.87 in terms of RMSE | Perfromed better compared to MLR, BRT, and NN |
| [27] | Time Series | Bi-LSTM (Bidirectional) | Hourly | China | BI-LSTM produced correlation coefficient of 98%, and RMSE of 0.791 | Same past context used for both the forward and the backward mode |
| [28] | Time Series | PSO-LSTM (Bidirectional) | Multiple days | China | PSO-LSTM achieved the lowest MAE, and RMSE as 8.14, and 19.41 | Same past context used for both the forward and the backward mode |
| [36] | Time Series | CNN-LSTM (Unidirectional) | 1-Day, 1-Week, 2-Week and 1-Month | Australia | It achieved lower MAPE < 11%, and RRMSE < 15% compared to benchmark models | This study is limited to one solar station |
| [33] | Time Series | LSTM-CNN (Unidirectional) | Multiple days | China | LSTM-CNN achieved the best MAE, RMSE, and MAPE as 0.221, 0.621, and 0.042 | This study is limited to one solar station |
| [6] | Time Series | MLP (Unidirectional) | Monthly | UAE | MLP has shown the best MBE, ans RMSE as 0.0003, and 0.179 | The model is validated for three solar stations |
| [30] | Time Series | RF | 1 h to 6 h | France | RF achieved the lowest forecasting error as 19.65% to 27.78% in terms of RMSE | This study is restricted to one solar station |
| [10] | Aerosol Optical Depth (AOD) and the Angstrom Exponent data | MLP (Unidirectional) | 1 h | Saudi Arabia | MLP achieved lower RMSE under 4% and forecast skill of over 42% | The study is restricted to one solar site |
| [29] | Time Series | RF | 5 min to 3 h | Australia | RF achieved the lowest overall MAE, and MRE as 110.46, and 10.5% | The proposed model is univariate, and restricted to one solar site |

Taken form ** **Malakar, Sourav, et al. "A novel feature representation for prediction of global horizontal irradiance using a bidirectional model." Machine Learning and Knowledge Extraction 3.4 (2021): 946-965**

# Problem Space

- **Problem 1:** The current literature does not provide clear guidelines for design choices of LSTM based models in solar forecasting [4-7].
- **Problem 2:** As renewable energy has a strong local weather dependence. Hence, individual forecasting model for each station-season combination, which results in a considerable number of models across India.
- **Problem 3:** For a country like India evidently, no such short-term solar forecasting models have been found based on cloud-dependent clustering.
- **Problem 4:** Prior research provides no information on variations in model performance for different terrains and seasons in wind forecasting.

# Case Studies

- **Solution (problem 3):** Deep-Learning-Based Adaptive Model for Solar Forecasting Using Regime Dependent Clustering [9]

# Deep-Learning-Based Adaptive Model for Solar Forecasting Using Regime Dependent Clustering

# Deep-Learning-Based Adaptive Model for Solar Forecasting Using Regime Dependent Clustering (Contd.)

## Sky conditions

(a) Clear

(b) Isolated


(c) Scattered

(d) Broken

(d) Overcast

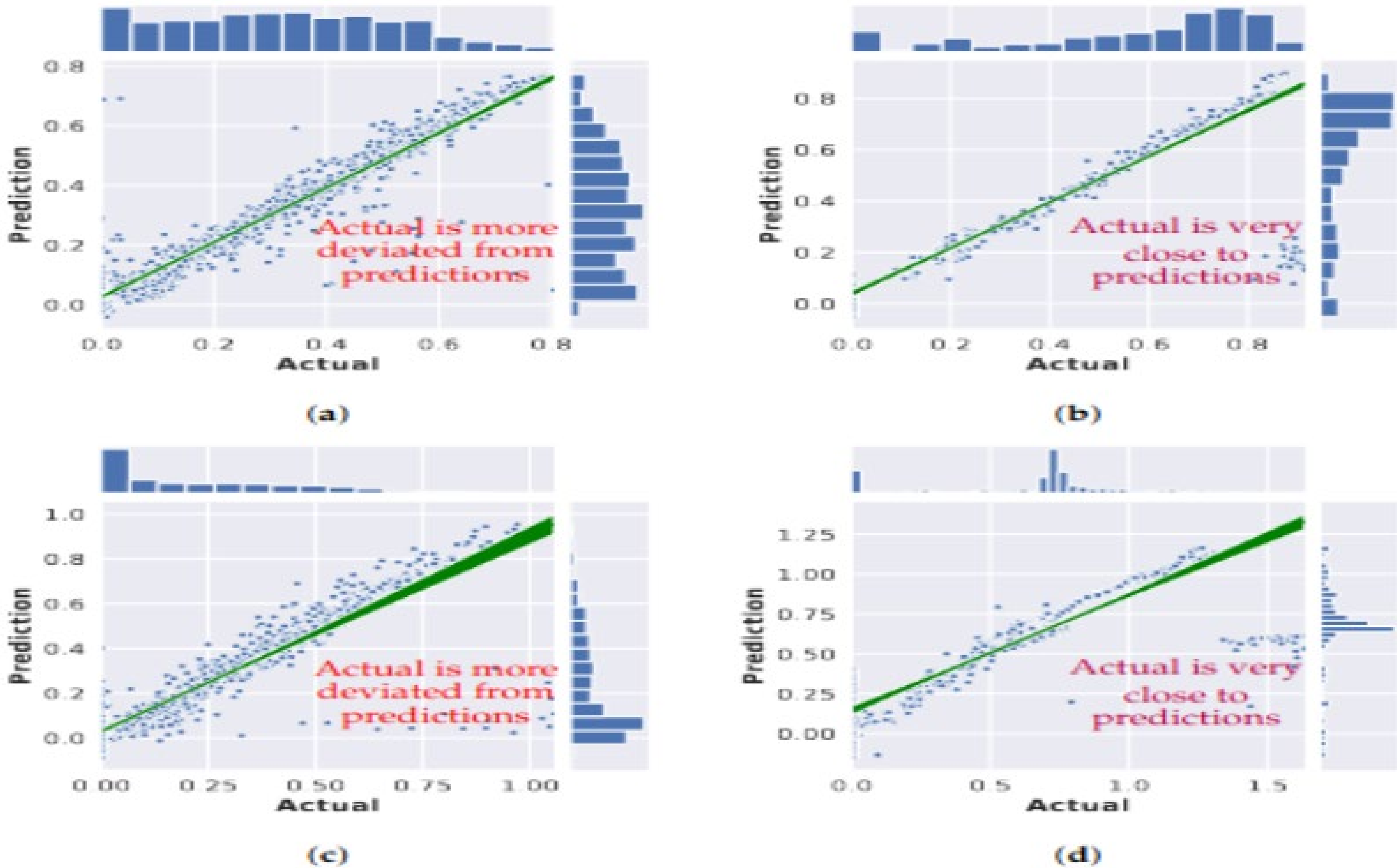
# Deep-Learning-Based Adaptive Model for Solar Forecasting Using Regime Dependent Clustering (Contd.)

- Hourly Variation of Cloud Types (Taken from: https://www.waff.com/weather/

| | | | |
|---|---|---|---|
| 6 PM 7/6 | 90° | Mostly Cloudy | 18% |
| 7 PM 7/6 | 88° | Mostly Cloudy | 10% |
| 8 PM 7/6 | 85° | Partly Cloudy | 12% |
| 9 PM 7/6 | 83° | Partly Cloudy | 14% |
| 10 PM 7/6 | 82° | Partly Cloudy | 9% |
| 11 PM 7/6 | 81° | Partly Cloudy | 9% |
| 12 AM 7/7 | 80° | Mostly Clear | 10% |
| 1 AM 7/7 | 79° | Mostly Clear | 10% |

# Deep-Learning-Based Adaptive Model for Solar Forecasting Using Regime Dependent Clustering (Contd.)

- Parity plot showing forecast and actual clearness indices.



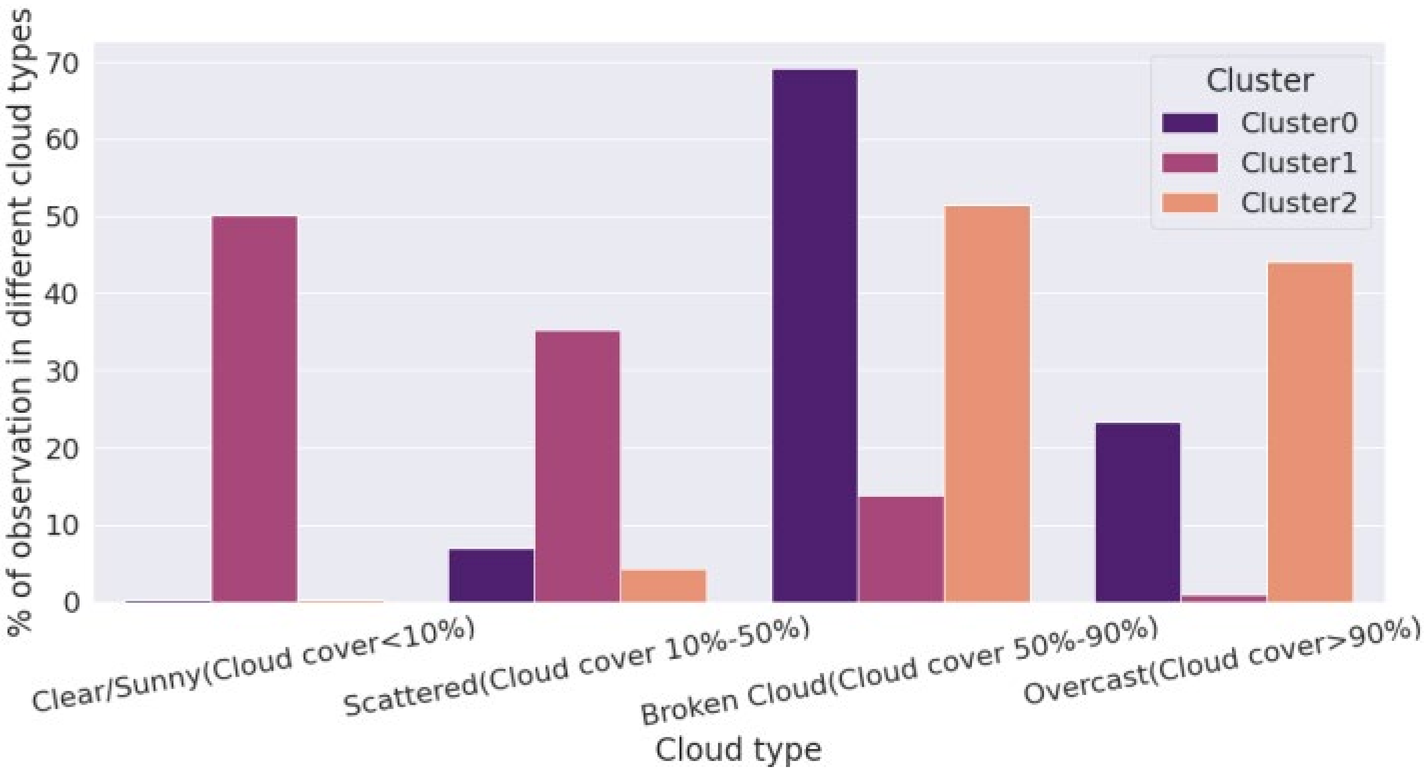
# Deep-Learning-Based Adaptive Model for Solar Forecasting Using Regime Dependent Clustering (Contd.)

- Variables used for K-medioid clustering.
- Clearness index (Kt): Ratio of surface radiation divided by the extraterrestrial radiation

| Predictors | Description |
|---|---|
| **Locally derived variables** | |
| $kt_{trend}$ | Recent trend in Kt. |
| Kt temporal variability (Stdev 1-h) | The temporal variability in Kt for past one hour. |
| Kt Slope (1-h) | The slope of Kt over the past hour. |
| **Remotely derived variables** | |
| KtPrev15 nearby mean | The spatial mean of Kt. |
| KtPrev15 nearby std | The spatial variability of Kt. |
| Cloud-cover variability (Stdev) | The spatial variability of cloud cover. |
| Cloud Cover Squired | Thickness of cloud cover. |

# Deep-Learning-Based Adaptive Model for Solar Forecasting Using Regime Dependent Clustering (Contd.)

- Understanding cloud patterns via cluster-specific distribution of cloud type.



# Deep-Learning-Based Adaptive Model for Solar Forecasting Using Regime Dependent Clustering (Contd.)

- Forecasting performance of CB-LSTM compared to multivariate and spatiotemporal LSTM

| Forecasting Sites | CB-LSTM NRMSE | CB-ANN NRMSE | ST-LSTM NRMSE |
|---|---|---|---|
| Composite | | | |
| Bhainsdehi | 0.1936 | 0.2762 | 0.2515 |
| Begamganj | 0.2016 | 0.3118 | 0.3410 |
| Dindori | 0.2530 | 0.3440 | 0.3352 |
| Hot and dry | | | |
| Tiruchirappalli | 0.2903 | 0.3414 | 0.5419 |
| Idukki | 0.4934 | 0.5288 | 0.6037 |
| Madurai | 0.3208 | 0.3288 | 0.5447 |
| Warm and humid | | | |
| Khaga | 0.2641 | 0.3195 | 0.3300 |
| Vaibhavwadi | 0.3599 | 0.4980 | 0.4423 |
| Osmanabad | 0.3757 | 0.4760 | 0.4408 |

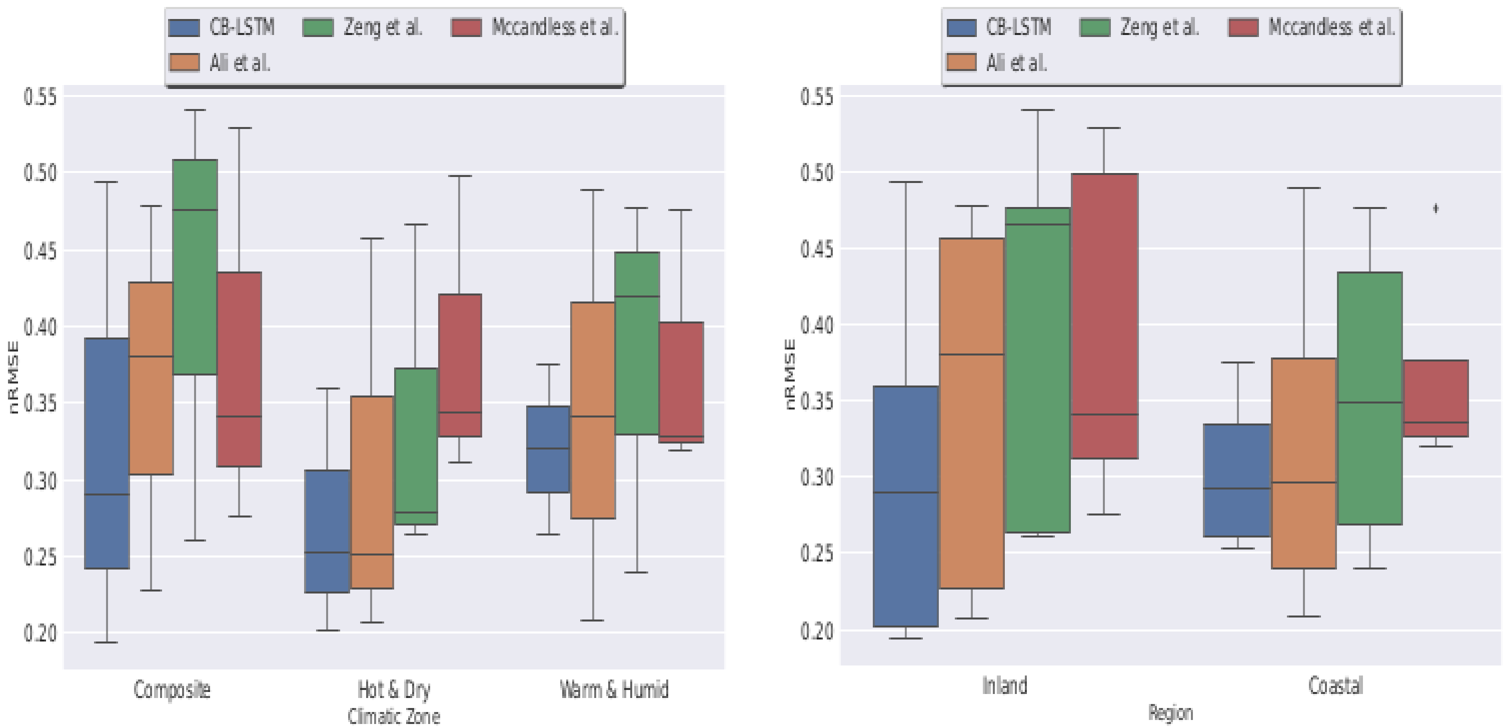
# Deep-Learning-Based Adaptive Model for Solar Forecasting Using Regime Dependent Clustering (Contd.)

- Forecasting performance of CB-LSTM compared with benchmark models in terms of NRMSE (%).

| Forecasting Sites | CB-LSTM | [4] | [5] | [6] |
|---|---|---|---|---|
| Composite | | | | |
| Bhainsdehi | 19.36% | 22.75% | 26.09% | 27.62% |
| Begamganj | 20.16% | 20.74% | 26.39% | 31.18% |
| Dindori | 25.30% | 25.09% | 27.93% | 34.40% |
| Hot and dry | | | | |
| Tiruchirappalli | 29.03% | 38.06% | 47.62% | 34.14% |
| Idukki | 49.34% | 49.82% | 54.04% | 52.88% |
| Madurai | 32.08% | 34.14% | 47.67% | 32.88% |
| Warm and humid | | | | |
| Khaga | 26.41% | 20.90% | 24.02% | 31.95% |
| Vaibhavwadi | 35.99% | 45.70% | 46.61% | 49.80% |
| Osmanabad | 37.57% | 48.89% | 41.98% | 47.60% |

# Deep-Learning-Based Adaptive Model for Solar Forecasting Using Regime Dependent Clustering (Contd.)

- Climatic-zone and region specific variability in predictions




(a) (b)

The symbol "†" indicates an outlier.

# Future Scope

- Regime-dependent short-range solar forecasting needs real-time forecasting.
- Propose new time-series imputation model for continuous missing data in wind forecasting.
- Deploy the proposed algorithms in real-life scenarios in India. Also extension of the testbed by covering more climatic zones, and solar and wind stations across India.

# Research Staff

**Sourav Malakar** is currently pursuing a PhD. degree in Information Technology from the University of Calcutta, Kolkata, India. He is working as a State Aided College Teacher (category-I) with Maharani Kasiswari College, Kolkata.

**Saptarsi Goswami** received his M.Tech degree in Computer Science & Applications from A. K. Choudhury School of IT, University of Calcutta, Kolkata, India, and Ph.D. from A. K. Choudhury School of IT, University of Calcutta, Kolkata, India. He is currently an Assistant Professor at the Computer Science Department, Bangabasi Morning College, University of Calcutta, Kolkata, India.

**Bhaswati Ganguli** received a PhD degree in biostatistics from Harvard University, Cambridge, MA, USA, in 2002. She is a Faculty Member of the Department of Statistics, Calcutta University, Kolkata, India. She has authored the R package SemiPar and was the Principal Investigator of the e-PG Pathshala project in statistics of the MHRD. Her research interests include smoothing, mixed models, and spatial data analysis.

**Amlan Chakrabarti** is a Full Professor in the A.K.Choudhury School of Information Technology at the University of Calcutta. He is an M.Tech. from the University of Calcutta and did his Doctoral research at Indian Statistical Institute, Kolkata. He was a Post-Doctoral fellow at the School of Engineering, Princeton University, USA during 2011-2012. He is a Sr. Member of IEEE and ACM, Distinguished Speaker of ACM, Secretary of IEEE CEDA India Chapter, Vice President of Data Science Society and Life Member of CSI India.

# References


- 1. "2021 Renewable Energy Country Attractiveness Index (RECAI)". www.ey.com.
- 2. Koundal, Aarushi (26 November 2020). "India's renewable power capacity is the fourth largest in the world, says PM Modi". ETEnergyworld. Retrieved 16 May 2021.
- 3. "India's 450 GW renewable energy goal by 2030 doable, says John Kerry". The Hindu. PTI. 20 October 2021. ISSN 0971-751X. Retrieved 29 January 2022.
- 4. Mohamed Abdel-Nasser and Karar Mahmoud. Accurate photovoltaic power forecasting models using deep lstm-rnn. Neural Computing andApplications, 31(7):2727–2740, 2019.
- 5. Rafael Caballero Roldán, Luis Fernando Zarzalejo Tirado, Álvaro Otero Martín, Luis Piñuel Moreno, and Stefan Wilbert. Short term cloud nowcasting for a solar power plant based on irradiance historical data. Journal of computer science & technology, 18(3):186–192, 2018

# References


- 6. E Nikitidou, A Zagouras, V Salamalikis, and A Kazantzidis. Short-term cloudiness forecasting for solar energy purposes in greece, based on satellite-derived information. Meteorology and Atmospheric Physics, 131(2):175–182, 2019.
- 7. Gangqiang Li, Huaizhi Wang, Shengli Zhang, Jiantao Xin, and Huichuan Liu. Recurrent neural networks based photovoltaic power forecasting approach. Energies, 12(13):2538, 2019.
- 8. Goswami, Saptarsi, et al. "A novel transfer learning-based short-term solar forecasting approach for India." Neural Computing and Applications 34.19 (2022): 16829-16843.
- 9. Malakar, Sourav, et al. "Deep-Learning-Based Adaptive Model for Solar Forecasting Using Clustering." Energies 15.10 (2022): 3568.
- 10. Carneiro, Tatiane Carolyne, et al. "Review on photovoltaic power and solar resource forecasting: current status and trends." Journal of Solar Energy Engineering 144.1 (2022).
- 11. Yang, Dazhi, et al. "A review of solar forecasting, its dependence on atmospheric sciences and implications for grid integration: Towards carbon neutrality." Renewable and Sustainable Energy Reviews 161 (2022): 112348.