Review for the Solar Radiation Forecasting Methods Based on Machine Learning Approaches

U Hemavathi\*, Ann C V Medona\textsuperscript{2}, V Dhipil Kumar\textsuperscript{1}, and R Raja Sekar\textsuperscript{1}

\textsuperscript{1}Department of Computer Science and Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, Tamil Nadu, India
\textsuperscript{2}Assistant Professor, Department of Artificial Intelligence and Data Science, Saveetha Engineering College, Chennai, Tamil Nadu, India
Email: \*uhemav@gmail.com

Abstract. Predictions of solar potential for these systems' production are important, whether they ensure sound activity or the perfect control of an energy discharge heading to the solar system. It is important to base the prediction on solar irradiance before predicting solar systems performance. The measurement of solar radiation elements is a very significant criterion for applications of solar energy. Several globalized solar radiation prediction modes can be done in the two major categories: cloud imagery with physical models and machine learning techniques are correlated. In this paper, the methods used to predict solar radiation are explained with machine learning algorithms.

Keywords: renewable prediction, solar system, ML methods.

1. Introduction
The large integration of renewable energies (especially non-visible energy sources as wind or solar) into an existing or future energy supply framework is one of the most significant challenges for the near future global energy supplies. An energy supplier should always strike a precise balance between the output of electricity and utilization. Solar electricity is the most abundant source of energy on Earth. It is soon expected to play an extremely important role, particularly in developing countries [1-4]. The advent of interest using this energy source has risen dramatically in recent years, primarily due to traditional sources' climbing costs.

The incorporation of renewable power into a network intensify network management uncertainty and the continuity of the equilibrium between output and use due to their unpredictability and intermittence [5-8]. There is practically no reliable way, except practise measuring, to reliably measure intermittent solar irradiance. Several experiments have measured the daily and hourly radiation data to capture the pattern, spontaneously shifting solar radiation over one month. The non-controllable intermittence and solar output features contribute to several additional load variations, local control efficiency, and reliability problems [9-13]. The power prediction solar systems outputs are important for efficient electricity grid service, or optimized control of energy flows into the solar system grid or improve the solar system's energy flux control [14-18].

For developing and testing solar energy utilization technologies, solar radiation data is important. The data are virtually calculated the most reliable but not always readily available measurement instruments and their respective registry processes, owing primarily to initial expenditure and maintenance costs.
2. Methods
The models discussed in this paper relate instead of being tilted for measuring solar radiation data on a horizontal surface on the ground surface area. Their predictions are not only for a few minutes or seconds but for global radiation over longer times of an hour and a day. There is practically no reliable way, except practice measuring, to reliably measure intermittent solar irradiance. Several experiments have measured the daily and hourly radiation data to capture the pattern, spontaneously shifting solar radiation over one month. The mean meaning of the global sunbeam, intercepted by horizontal surfaces for a day for a month, is the "monthly average global sunlight,". Other words have been identical if new parameters are given, equivalent meanings. The Prediction scale for energy management is shown in fig. 1.

The estimated time impacts the efficiency and usefulness of the model tremendously. From the data consistency point of view, a suggested framework for estimating hourly radiation is better and more efficient than the model for estimating everyday radiation. The data reduction approaches were closely analyzed in the classification of current studies because often contradictory arguments are the real procedures. For example, any report states that the calculation refers to everyday radiation but is simply the total monthly or long-term daily radiation, which makes scientific effects unsustainable by radiation.

![Figure 1. Prediction scale for energy management](image-url)

3. Ann Method
Some information on how to compare ANN models should be given first before implementing ANN (artificial neural network). In terms of computing, an ANN model is also more complicated than an analytical model, as shown in fig. 2. In an ANN model, there are various variables such as the data, number of layers and neurons, preparation, transition, etc. Any improvement will build a new ANN model, such that the necessary rules need to be created. The ANN models equate. First of all, the performance of an ANN model must be considered. The relation does not matter if the result is different, and we label the parts weekly, average daily dose, radiation every day and exposure every hour.

A computational simulation methodology that has advanced most in recent years is the artificial neural network. It is motivating and able to process non-linear interactions and data sorted from biological neural systems; detect patterns, refine, cluster and simulate patterns.
4. **Random Forest**

A Random Woods, consisting of a selection or ensemble of many decision-making booms constructed from each test collected with a substitution (a bootstrapper sample) from a training set is the proposed approach’s machine-learning technique. Also, just a random variable subset is used to separate a node when a tree is created. Therefore one or more findings can be presented at the final nodes (or leaves). In the regression problem, each tree can produce an answer if there is a prediction set, provided the state of the assumptions made on the resulting node.

Usually, the dependency mean is approximated by mean rating. The trees’ random arrangement usually raises the forest distortion marginally in contrast with the distortion of a single non-random tree. Variance declines, much more than the rise in bias, thereby creating a stronger overall model. Finally, all trees’ answers can then be combined to reach a single model solution vector, and a weighted mean is also employed. Important changes were made to the classification’s precision by increasing trees’ ensemble and encouraging them to vote for the most common class.

5. **Support Vector Machines Algorithm**

The SVM support vector machines are based on a theory of inductive structural risk minimization (SRM) to reduce the top limit to the error generalizing the total amount of training errors and the trust level. It is the deviation from the commonly accepted ERM theory, which reduces just the training error. SVM typically performs more broadly than conventional neural networks that apply the ERM theory to solve many machine learning issues, depending on this induction theory. Compared to many other network programmings, which involves non-linear optimization and is in threat of staying in the local minima, the SVM's solution is often special and generally efficient. SVM relies on a subset of training points called supportive vectors to solve the issue. SVMs are highly robust frameworks for non-linear resolution issues and regression applications in science and industry and classification purposes. Parameters of the SVM is shown in fig. 3.
6. Features Of SVM For Regression Estimation

Several characteristics of SVM are described below according to the theoretical statement of SVM. Firstly, SVM determines regression using a collection of linear functions specified in a broad feature space, whilst inputs do not work in linear. The kernel functions are named. Second, SVM carried out the regression evaluation based on the concept of data mining, when using e-insensitive lower bound, depending on the risk evaluation.

7. Results

The knowledge is considered noise based on the intrinsic nature of the data collection, which influences interpretation by various external aspects. The results of the SVM and RF are based on solar radiation is predicted. The obtained results are verified that shows the distribution of energy consumption based on the hour. Figure 3, the Actual versus SVM with ML predicted results for solar radiations. Reference models depending on the forecast horizon are shown in fig. 4.
8. Conclusion
The reliability of these techniques usually depends on the consistency of the data obtained. The SVM and the random forests are the three approaches to be employed in the coming years. The findings are very positive, and there will be some fascinating experiments in the next few years. The results of the SVM, RF and machine learning are presented and validated. In practice, these approaches generate comparable error statistics, taking into account published articles. Methods may be more important to enforce the procedures themselves are with the errors mentioned in the literature.

References
[1] Liakos, K. G., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine learning in agriculture: A review. *Sensors*, 18(8), 2674.
[2] Kashyap, Y., Bansal, A., & Sao, A. K. (2015). Solar radiation forecasting with multiple parameters neural networks. *Renewable and Sustainable Energy Reviews*, 49, 825-835.
[3] Basaran, K., Özçift, A., & Kılınç, D. (2019). A new approach for prediction of solar radiation with using ensemble learning algorithm. *Arabian Journal for Science and Engineering*, 44(8), 7159-7171.
[4] Pani, A. K., & Nayak, N. (2019). A Short Term Forecasting of Photovoltaic Power Generation Using Coupled Based Particle Swarm Optimization Pruned Extreme Learning Machine. *International Journal of Renewable Energy Research (IJRER)*, 9(3), 1190-1202.
[5] Manie, N., & Pattanaik, B. (2019). Zeta DC-DC converter based on MPPT technique for BLDC application. *International Journal of MC Square Scientific Research*, 11(2), 1-12.
[6] Kumar, D. N. P., Murugesan, K., Kumar, K. A., & Subadra, M. (2012). Fuzzy based Priority Scheduler for WiMAX with Improved QoS constraints. *International Journal of Advanced Networking and Applications*, 4(3), 1616.
[7] Sumathy, B., Shiva, P. D., Mugundhan, P., Rakesh, R., & Prasath, S. S. (2019, November). Virtual Friendly Device for Women Security. In *Journal of Physics: Conference Series* (Vol. 1362, No. 1, p. 012042). IOP Publishing.
[8] Jayavel, K., & Venusamy, K. (2020, October). Design and Implementation of IoT Testbed with Improved Reliability using Conditional Probability Techniques. In 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC) (pp. 116-121). IEEE.
[9] Wang, S., Janabi, A., & Wang, B. (2020, October). Generalized Optimal SVPWM for the Switched-Capacitor Voltage Boost Converter, ECCE, (pp. 2708-2711). IEEE.
[10] Gibbons MS, Gilbert CL, Deshpande MS, inventors; Lear Corp, assignee. Methodology of maximizing charging power transfer for an electric vehicle when ac voltage sags. United States patent application US 15/801,772. 2019 May 2.
[11] Anitha, S., Keerthana, A., Kaviya, S. and JayaRajan, R., 2017. QUADRIPLEGIC AMBULATING ADDENDUM. *International Journal of MC Square Scientific Research*, 9(2), pp.15-23.
[12] Anjana, R. (2019, Feb). Fuzzy and PI Based Speed Control of BLDC Motor using Bidirectional Converter for Electric Vehicle Application. *Trends in Electrical Engineering*, 8(3), 35-45.
[13] Yadav, D., Parekh, U., Patel, K., Parmar, R., & Prakash, M. (2019). Application of Modified Three Phase Conduction Method to Minimize Torque Ripple in BLDC Motor. *IJRAR*, vol. 6(1), 155-156.
[14] Gibbons, M. S., Gilbert, C. L., & Deshpande, M. S. (2019). US Patent Application No. 15/801,772.
[15] Rajkumar, S., Sundaram, C. S., Sedhuraman, K., & Muruganandhan, D. (2019, March). Performance Analysis of Hub BLDC Motor Using Finite Element Analysis. In 2019 ICSCAN, (pp. 1-5). IEEE.

[16] Nasiriani, K., & Pasandi, M. (2020). Dynamic Voltage Restorer (DVR) For Protecting Hybrid Grids. arXiv preprint arXiv:2006.16452.

[17] Gupta, G., & Kumar, D. (2020). Power Quality Mitigation using DSTATCOM and DVR.

[18] Macinnis, AG and Walls, F.G., Avago Technologies General IP Singapore Pte Ltd, 2018. Bounded rate near-lossless and lossless image compression. US Patent 9,883,180.