International Journal of Current Engineering and Technology
E-ISSN 2277 – 4106, P-ISSN 2347 – 5161
Available at http://inpressco.com/category/ijcet

Research Article

Solar Forecasting Analysis using Machine Learning

Yogita Mohanrao Shinde and Prof.Dipak Gupta

Department of ME Computer Siddhant College Of Engineering Pune, India

Received 10 Nov 2020, Accepted 10 Dec 2020, Available online 01 Feb 2021, **Special Issue-8 (Feb 2021)**

Abstract

The rapid increase in solar power plant installed capacity leads to considerable difficulties in terms of power system operation and control, resulting from highly stochastic nature of solar energy harvesting. The paper considers the problem of dayahead solar power plant output forecasting, based on the meteorological data. The improvement of solar power plant output prediction will significantly simplify power system operation mode planning taking into market procedures and active power generation reserves allocation. As a case study the authors use meteorological data for a real operated solar power plant. As a results of regression modelling the statistical significance of the meteorological parameters was analyzed. The optimal mathematical formulation of regression model was provided. In addition, the paper gives the idea of empirical cauterization approach, providing significant improvement of prediction accuracy. The results of the verification on real data allow deciding on the applicability of the pro posed methods in industrial operation.

Keywords: *Time series Analysis, Regression Analysis, Clustering, Forecasting of electricity generation, solar power plants, Machine Learning*

Introduction

Solar power forecasting involves knowledge of the Sun´s path, the atmosphere's condition, the scattering processes and the characteristics of a solar energy plant which utilizes the Sun's energy to create solar power. Solar photovoltaic systems transform solar energy into electric power. The power output depends on the incoming radiation and on the solar panel characteristics. Photovoltaic power production is increasing nowadays. Forecast information is essential for an efficient use, the management of the electricity grid and for solar energy trading. Common solar forecasting method include stochastic learning method, local and remote sensing method, and hybrid method. The energy generation forecasting problem is closely linked to the problem of weather variables forecasting. Indeed, this problem is usually split into two parts, on one hand focusing on the forecasting of solar PV or any other meteorological variable and on the other hand estimating the amount of energy that a concrete power plant will produce with the estimated meteorological resource. In general, the way to deal with this difficult problem is usually related to the spatial and temporal scales we are interested in, which yields to different approaches that can be found in the literature. In this sense, it is useful to classify these techniques depending on the forecasting horizon, so it is possible to distinguish between now-casting (forecasting 3–4 hours ahead), short-term forecasting (up to 7 days ahead) and longterm forecasting (months, years...).
To achieve the forecasting model for solar power generation we need to have large amount of data to be processed. This data can be obtained from various solar power plants and can be used as dataset for further processing.
Dataset should cover maximum parameters such as temperature, humidity, cloud cover, wind speed, power unit generated, date, time, etc. As there is no standard dataset for this, we will be using some dataset as collected data from solar power plants. This data will go through machine learning algorithms and can generate models on which current data and statistics are to be mapped to get predictions as forecasting for desired time period.

Literature Survey

Denis A. Snegirev, Stanislav A. Eroshenko, Alexandra I. Khalyasmaa, Valeria V. Dubailova, Alina I. Stepanova during this paper, The rapid increase in solar power plant installed capacity leads to considerable difficulties in terms of power system operation and control, resulting from highly stochastic nature of solar energy harvesting. The paper considers the problem of day-ahead solar power plant output forecasting, based on the meteorological data. The improvement of solar power plant output prediction will significantly simplify power system operation mode planning taking into market procedures and active power generation reserves allocation. As a case study the authors use meteorological data for a real operated solar power plant. As a results of regression modelling the statistical significance of the meteorological parameters was analyzed. The optimal mathematical formulation of regression model was provided. In addition, the paper gives the idea of empirical cauterization approach, providing significant improvement of prediction accuracy. The results of the verification on real data allow deciding on the applicability of the proposed methods in industrial operation.

Graeme Vanderstar, Petr Musilek, Alexandre Nassif, in this work it is the need to accurately forecast available sola irradiance is a significant issue for the power industry and poses special challenges for utilities who serve customers in isolated regions where weather forecast data may not be abundant. This paper proposes a method to forecast two hour ahead solar irradiance levels at a site in Northwestern Alberta, Canada using real-time solar irradiance measured both locally and at remote monitoring stations. This paper makes use of an Artificial Neural Network (ANN) to forecast the solar irradiance levels and uses the genetic algorithm to determine the optimal array size and positioning of solar monitoring stations to obtain the most accurate forecast from the ANN. The findings of this paper are that it is possible to use as few as five remote monitoring stations to obtain a near-peak forecasting accuracy from the algorithm and that providing adequate geospatial separation of the remote monitoring sites around the target site is more desirable than clustering the sites in the strictly upwind directions.

Fatih Serttas, Fatih Onur Hocaoglu, Emre Akarslan,
during this study, Photovoltaics' (PV's) are widely preferred in electricity generation market in recent years. However many parameters effect solar power generation such as irradiance, temperature, humidity etc. Therefore, solar power generation forecasting is quite significant to plan and manage energy distribution. In this study, a novel methodology called Mycielski-Markov is utilized to forecast solar power generation for short term period. This novel hybrid method is developed based on two different techniques; Mycielski signal processing technique and probabilistic Markov chain. Mycielski investigates the data history and finds the recurrence of the solar radiation data. It predicts the next data due to the recurrence in a deterministic way. On the other hand, Markov produces the transition probabilities of the solar energy states and forecast new state according to these probabilities. It is obtained that, the methods in proposed hybrid hierarchy; provide a good forecasting accuracy with a 0.87 correlation of determination value.

Hossein Panamtash, Qun Zhou, Solar power has been growing rapidly in recent years. Many countries have invested in solar energy technology, especially in Photovoltaic (PV) power generation. With the increased penetration level, solar power forecasting becomes more challenging. To cope with solar power uncertainties, probabilistic forecasting provides more information than traditional point forecasting. Moreover, multiple PV sites with spatial-temporal correlations need to be taken into account. To produce probabilistic forecasts, this paper applies quantile regression on top of time series models. Considering the coherency among multiple PV sites, a reconciliation is applied using a copula-based bottom-up method or proportion-based top-down method. Numerical results show that the proposed methods efficiently produce accurate and coherent probabilistic solar power forecasts.

M. Z. Hassan, K.M.E. Ali, ABM Shawkat Ali, Jashnil Kumar, in this work, Unpredictability of solar resource poses difficulties in grid management as solar diffusion rates rise continuously. One of the big challenges with integrating renewables into the grid is that their power generation is intermittent and unruly. Thus, the task of solar power forecasting becomes vital to ensure grid constancy and to enable an optimal unit commitment and cost-effective dispatch. Latest techniques and approaches arise worldwide each year to progress accuracy of models with the vital aim of reducing uncertainty in the predictions. This paper appears with the aim of compiling a big part of the knowledge about solar power forecasting, focusing on the most recent advancements and future trends. Firstly, the inspiration to achieve an accurate forecast is presented with the analysis of the economic implications it may have.

Xin Huang, Huaning Wang, Long Xu, Wenqing Sun, In this paper, Solar flare is one type of violent eruptions from the Sun. Its effects almost immediately arrive to the near-Earth environment, so it is crucial to forecast solar flares in space weather. So far, the physical mechanisms of solar flares are not yet clear, hence we learn a solar flare forecasting model from the historical observational magnetograms by using the deep learning method. Instead of designing the feature extractor by the solar physicist in the traditional solar flare forecasting model, the proposed forecasting model can automatically learn features from input raw data, and followed by a classifier for foretasting from the learned features. The experimental results demonstrate that the proposed model can achieve better performance of solar flare forecasting comparing to traditional solar flare forecasting models.

Irani Majumder, Manoja Kumar Behera, Niranjan Nayak, In present scenario the energy system face various challenges as the demand for energy is increasing significantly and the resources in terms of fossil fuels are limited, need for renewable resources have become very much vital at present. Accurate and reliable solar power forecasting is essential for the legitimate functioning of the power system. Given momentous uncertainties involved in solar power generation due to variation of temperature and irradiance, forecasting provides a unique solution for the uncertainties and variability's in solar data. In this paper a forecasting method has been mentioned that is contingent on a hybrid empirical mode decomposition (EMD) and Extreme Learning Machine (ELM). The nonstationary time series is further decomposed into distinct intrinsic mode functions (IMF). A short term forecasting is also carried out in this work to prove the accuracy of the given model. This model is implemented in MATLAB/SCRIPT environment.

D. A. Snegirev, S. A. Eroshenko, R. T. Valiev, A. I. Khalyasmaa,The rapid growth of solar power plants, integrated into United Power System of Russia (UPS), keeps stable over the last several years. The high percentage of solar power plants in the power system inevitably leads to significant increase of power generation output uncertainty, affecting the high-voltage power network. The reserve capacities, which are allocated on conventional power plants to cover the active power imbalance, should be as less as possible to improve economic efficiency of solar power plants implementation. For this reason, solar power plant operation forecasting becomes an urgent issue in power system and electric energy markets operation. The authors present the structure and mathematical description of the model of solar power plant operation short-term forecasting, which will give the possibility to assess the mean hourly value of solar power plant generation capacity. All the calculations are made for real solar power plant.

Gokhan Mert Yagli, Dazhi Yang, Dipti Srinivasan, and Monikae, Forecasting of solar PV generation plays an important role in power system operations. Forecasts are required on various geographical and temporal scales, which can be modeled as hierarchies. In a geographical hierarchy, the overall forecast for the region can either be obtained by directly forecasting the regional time series or by aggregating the individual forecasts generated for the sub-regions. This leads to a problem known as aggregate inconsistency as the two sets of forecasts are most likely different due to modeling uncertainties. Hence, practice is not optimal. Statistically optimal aggregation known as reconciliation, has been proven to provide aggregate consistent forecasts. Reconciliation helps system operators to have a superior foresight in a region-wise level, which eventually results in efficient system planning. The focus of this paper is on improving reconciliation accuracy. In addition, the effects of more accurate disaggregated and aggregated forecasts on the final reconciled predictions have been analyzed.

Yashar Sahraei Manjili, Rolando Vega, Mo M. Jamshidi, An adaptive framework for day-ahead forecasting of available solar energy is proposed based on a combination of data analytic approaches consisting of artificial intelligence and statistical techniques. Models are developed and validated utilizing a large dataset from the National Renewable Energy Laboratory (NREL) archive, the Automated Surface Observing System, and the solar position and intensity calculator (i.e., NREL-SOLPOS) sampled at 1-min intervals during eight years (2005–2012) for a site in Golden, CO, USA. The methodology is now ready for testing and validation in San Antonio, TX, USA, with data collected in the largest solar photovoltaic plant in TX, Alamo 1, which is the first solar plant in TX connected to the transmission grid allowing solar energy bidding into the market. A uniqueness of the methodology developed is that an integrated serial time-domain analysis coupled with multivariate analysis was used for preprocessing. The resulting enhanced dataset is used for adaptive training of the neural network- based forecast engine.

Proposed Methodology

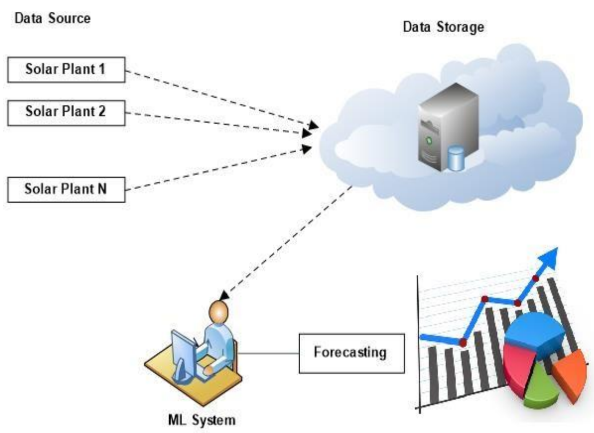

Fig.1. System Architecture

3.1 Methodologies:

3.1.1 Regression Analysis:

Regression analysis is used in stats to find trends in data. For example, you might guess that there's a connection between how much you eat and how much you weigh; regression analysis can help you quantify that. Regression analysis will provide you with an equation for a graph so that you can make predictions about your data. For example, if you've been putting on weight over the last few years, it can predict how much you'll weigh in ten years' time if you continue to put on weight at the same rate. It will also give you a slew of statistics (including a p-value and a correlation coefficient) to tell you how accurate your model is. Most elementary stats courses cover very basic techniques, like making scatter plots and performing linear regression. However, you may come across more advanced techniques like multiple regressions.

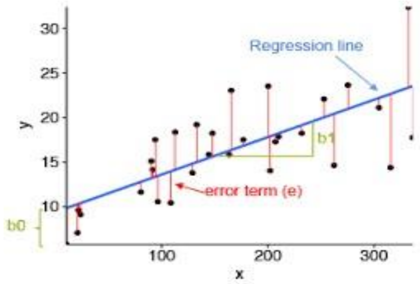

Fig.2. Regression Analysis

3.1.2 Time Series Analysis:

A time series is a series of data points indexed (or listed or graphed) in time order. Most commonly, a time series is a sequence taken at successive equally spaced points in time. Thus it is a sequence of discrete-time data. Examples of time series are heights of ocean tides, counts of sunspots, and the daily closing value of the Dow Jones Industrial Average.Time series are very frequently plotted via line charts. Time series are used in statistics, signal processing, pattern recognition, econometrics, mathematical finance, weather forecasting, earthquake prediction, electroencephalography, control engineering, astronomy, communications engineering, and largely in any domain of applied science and engineering which involves temporal measurements. Time series analysis comprises methods for analysing time series data in order to extract meaningful statistics and other characteristics of the data. Time series forecasting is the use of a model to predict future values based on previously observed values.

3.1.3 Support Vector Machine:

Super Vector Machine is a machine learning technique used in recent studies to forecast stock prices. Our goal is to use SVM at time t to predict whether a given stock's price is higher or lower on day t +m. We look at the technology sector and 34 technology stocks in particular. We input four parameters to the model - the recent price volatility and momentum of the individual stock and of the technology sector. These parameters are calculated using daily closing prices for each stock from the years 2007 through 2014. We analyse whether this historical data can help us predict price direction. If the Efficient Markets Hypothesis (EMH) holds true, prices should follow a random walk and be unpredictable based on historical data. We find that in the short-term this holds true, but in the long-term we are able to reach prediction accuracies between 55% and 60%. We conclude that our model is able to achieve significant prediction accuracies with some parameters in the long-term, but that we must look at more granular intra-day trading data to achieve prediction accuracies in the short-term.

SVM Algorithm:

1. Collect data of soler power plants in CSV file format
2. Read features of data such as temperature, humidity, cloud cover and wind speed
3. Read date-times for all and divide it in seasons
4. Save features along with power generated at that particular feature as label in a model file
5. Get current data
6. Map with model file
7. Show forecasting predictions

3.2 Mathematical Model:

Let S be the closed system defined as,

S = {Ip, Op, A, Ss, Su, Fi}

Where, Ip=Set of Input, Op=Set of Output, Su= Success State,

Fi= Failure State and A= Set of actions, Ss= Set of user's states.

- Set of input=Ip={username, password}
- Set of actions =A={F1,F2,F3} Where, o F1= Authentication of user o F2 = Input the dataset o F3= This result show and stored the database
- Set of user's states=Ss={registration state, login state, selection dataset, prediction logout}
- Set of output=Op={Show results}
- Su=Success state={Registration Success, Login Success }
- Fi=Failure State={Registration failed, Login failed}
- Set of Exceptions= Ex ={Null Pointer Exception while registration state, Record Not Found (Invalid Password) while login state , Null Values Exception while showing state}

Result and Discussion

In the proposed system, we propose a solar energy forecasting model with Machine Learning algorithm. In order to train the machine learning model, regression analysis and time series analysis is used. This system is useful in to improve the efficacy of solar plant.

Comparative results of existing and proposed system is as follow,

Table 1: Comparative Results

Parameters	Existing System	Proposed System
Forecasting	Yes	Yes
Supervised Learning	Somewhat	Yes
SVM	No	Yes
Improved speed	No	Yes
Light Weight	No	Yes
Accuracy	Less	More

With reference to Table 1 it is clear that we overcome various problems in existing system and our approach works efficiently. Further results are to be shown in graph format as shown in figure 3 as follows,

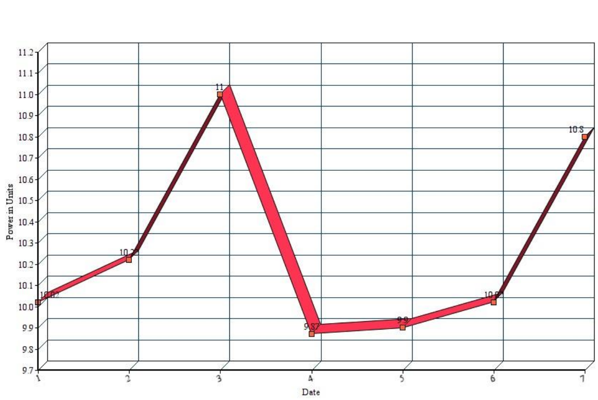

Fig.3. Power Forecasting Results

As shown in above figure 3 it is the forecasting result for seven days from current day. With respect to given data system analyses and produces seven different values in watts or in power units and plot a graphical representation for user to make it easily understandable.

Conclusion

Machine learning has demonstrated their success in many applications due to their ability to solve some problems with relative ease of use and the model-free property they enjoy. One of the main features, which can be attributed to support vector machine, is its ability to learn nonlinear problem offline with selective training, which can lead to sufficiently accurate response. Application of Support Vector Machine (SVM) to the above mentioned problem has attained increasing importance mainly due to the efficiency of present day computers. In addition, the times of simulation and testing in the Machine Learning application are minimal. As discussed in this paper it is clear that we can take predictions of power generation through solar power generation. This ultimately helps power manufacturers to analyze the performance of their solar power plant. Also with the help of this system one should be aware if extra power needs to be managed in earlier stages so that power outages can be mitigated. Further government organizations can use this system to predict power consumption and production ration and manage accordingly.

References

[1] Day-ahead Solar Power Plant Forecasting Accuracy Improvement on the Hourly Basis- Denis A. Snegirev, Stanislav A. Eroshenko, Alexandra I. Khalyasmaa, Valeria V. Dubailova, Alina I. Stepanova-2019

[2] Solar Forecasting using Remote Solar Monitoring Stations and Artificial Neural Networks - Graeme Vanderstar, Petr Musilek, Alexandre Nassif- 2018

[3] Short Term Solar Power Generation Forecasting: A Novel Approach- Fatih Serttas, Fatih Onur Hocaoglu, Emre Akarslan- 2018

[4] Coherent Probabilistic Solar Power Forecasting- Hossein Panamtash, Qun Zhou- 2018

[5] Forecasting Day-ahead Solar Radiation Using Machine Learning Approach- M. Z. Hassan, K.M.E. Ali, ABM Shawkat Ali, Jashnil Kumar- 2017

[6]Learning Solar Flare Forecasting Model from Magnetograms- Xin Huang, Huaning Wang, Long Xu, Wenqing Sun- 2017

[7] Solar Power Forecasting Using a Hybrid EMD-ELM Method- Irani Majumder, Manoja Kumar Behera, Niranjan Nayak- 2017

[8] Algorithmic Realization of Short-term Solar Power Plant Output Forecasting- D. A. Snegirev, S. A. Eroshenko, R. T. Valiev, A. I. Khalyasmaa- 2017

[9] Solar Forecast Reconciliation and Effects of Improved Base Forecasts- Gokhan Mert Yagli, Dazhi Yang, Dipti Srinivasan, and Monika- 2017

[10] Data-Analytic-Based Adaptive Solar Energy Forecasting Framework- Yashar Sahraei Manjili, Rolando Vega, Mo M. Jamshidi- 2017