A Project Report

On

PREDICTION AND ANALYSIS OF SOILING LOSS OF SOLAR POWER SYSTEMS

BY

Under the supervision of

Dr. Ankur Bhattacharjee

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS OF

Design Projects (ECE F366 (F367) / EEE F366 (F367) / INSTR F376 (F367)

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE PILANI (RAJASTHAN)

HYDERABAD CAMPUS

(DECEMBER 2021)

ACKNOWLEDGMENTS

The success and final outcome of this project required a lot of guidance and assistance from many people and I am extremely privileged to have got this all along the completion of my project. I would like to express my gratitude to Dr. Ankur Bhattacharjee for providing us with the opportunity to work on this project and for his guidance throughout the project. I would also like to thank my team members for their support and hard work.

Birla Institute of Technology and Science-Pilani,

Hyderabad Campus

Certificate

This is to certify that the project report entitled "**PREDICTION AND ANALYSIS OF SOILING LOSS OF SOLAR POWER SYSTEMS**" submitted by Mr. KSHITIJ DEVIKAR (ID No. 2019A3PS0379H) in partial fulfillment of the requirements of the course Design Projects (ECE F366 (F367) / EEE F366 (F367) / INSTR F376 (F367), embodies the work done by them under my supervision and guidance.

Date: 10th December 2021

(Dr. Ankur Bhattacharjee)

BITS- Pilani, Hyderabad Campus

ABSTRACT

Solar power represents one of the most accessible and clean energy resources that can be found on the planet. Therefore, it is important to optimize the performance of the Photovoltaic plants so that maximum amount of power is generated from the incident solar rays. However, the generation of power from the PhotoVoltaic modules has various challenges like the deposition of soil and dust on top of the PV modules, ambient temperature, wind, etc. In this study, we aim to evaluate and compare various modern Deep Learning models such as Artificial Neural Networks (ANN), Recurrent Neural Networks (RNN), and Long Short-Term Memory (LSTM) in predicting the percentage transmittance of the modules for a given range of recorded wavelengths. Three PV modules were set up at different positions - vertical, horizontal and inclined at the local tilt angle, and data for the wavelength of incident light and the corresponding percentage solar transmittance was measured throughout the year 2020. Factors like wind and ambient temperature were not included in our calculations. The accuracy of the prediction models was determined by computing error statistics like mean absolute error and percentage error (MAE/ MAPE), root mean square error (RMSE), and the correlation coefficient (R, R^2).

The ANN prediction model is accomplished using nntool of MatLab, where the accuracy of the predicted data is around 99.01% - 99.19 for three different sets of input data, i.e., inclined, vertical, and horizontally positioned solar panels. LSTM models consist of a recurrent network where the data is looped repeatedly for higher accuracy, resulting in the predicted information having an error percentage less than ten.

CONTENTS

INTRODUCTION

In recent years, with pollution levels reaching new levels and resources like fossil fuels quickly running out of our hands, solar power is our safest option. It is not just environment friendly, it also has zero production cost and is a resource we can never run out of. The increasing market for solar power as an alternate form of energy prompts interest in developing and operating photovoltaic systems to tap solar power on a large scale to integrate it into society as a viable energy source.

The efficiency of solar power systems is so far influenced by many factors, such as weather conditions, topographic elevation, solar inclination, seasonal changes, and discontinuous production. It is more efficient to use solar energy production information ahead of time to counter the operating costs caused by requirements of energy reserves or shortage of electricity supplies from Photovoltaic systems. [1]. Therefore, solar transmittance forecasting is a crucial part of generating an optimal photovoltaic plant.

Powerful Deep Learning techniques like Neural Networks like ANN and LSTM are now being used in many studies these days to predict the power output of soiled PV modules [2,3,4,5] as they are accurate, interactive and flexible. Such works are essential to make efficient working solar power models, such as designing a module that can wipe itself if it gets dusted [6].

ANN (Artificial Neural Networks) is a widely used tool for data prediction wherein we actually create a network with a certain number of neurons that are interconnected and feed the data to the network so that the network is trained as per the data input given after which the network is simulated with the testing data to predict the outcome.

LSTMs are built with Recurrent Neural Network architecture but with extended memory. Unlike FeedForward Neural Networks, RNNs cycle the data in loops, making the algorithm more suitable for sequential data like time series such as weather records and stock prices.

A study [7] with data from a solar power plant in Tiruchirapalli, India worked with ANN models to forecast a 24-hour (Day-ahead) solar power output; we are looking to predict the transmittance values using data recorded monthly throughout the year 2020. We use the data recorded across twelve months of 2020, containing transmittance percentages for just over eighteen hundred wavelengths ranging from three hundred to eleven hundred. The angle of inclination (of the PV modules) is a significant factor that determines the radiation passing through, depending on the region, each plant has a particular angle that maximizes the output, we work with three data sets; transmittance recorded for PV modules placed horizontally, vertically, and optimally inclined. ANN and LSTM are the models used here to predict the required data.

The growing popularity has led to increased research, comparing various forecasting techniques; auto-regressive models, neural networks, and optimized neural networks [8].

Even though the generation of power from the PV modules is affected by various factors like irradiance, wind speed, ambient temperature, dust, snow, the orientation of the PV modules, maintenance, etc., [1,9,10,11] we have not considered these factors.

Neural Net Fitting Tool using MATLAB

The data is formatted and given as input (using two variables wavelength and a number of the month), and the target data is also supplied (in terms of % Transmittance) to the ANN model which is further used for training and testing purposes.

The ANN model is created using the Neural Net Fitting Tool which is present in the MATLAB which allows us to select the training algorithm and the number of neurons to be used.

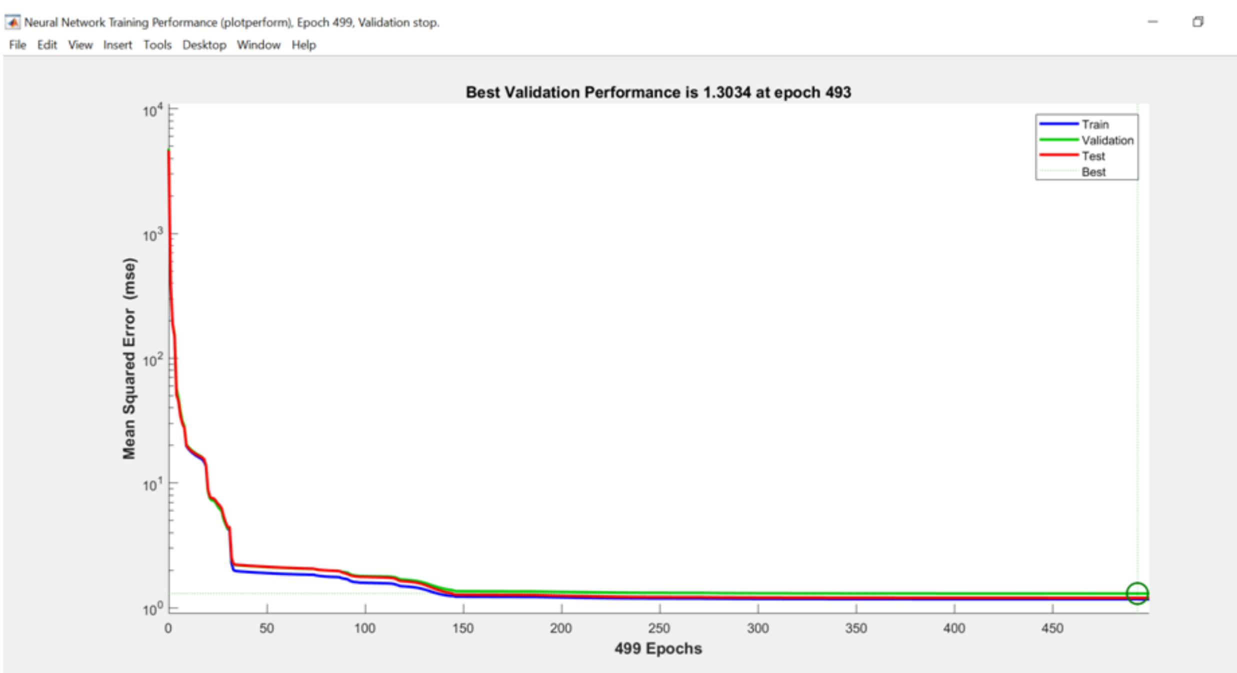
This ANN model automatically divides the data provided as 70% data for training, 15% data validation and 15% data for testing. Percentage of data for Validation and Training could be manually changed from 5% to 35% (in increments of 5%).

The training algorithm used here is “Levenberg-Marquardt” which is a built-in algorithm in this tool.

Note that the observed accuracy is better in this particular algorithm than the other two built-in algorithms that are present.

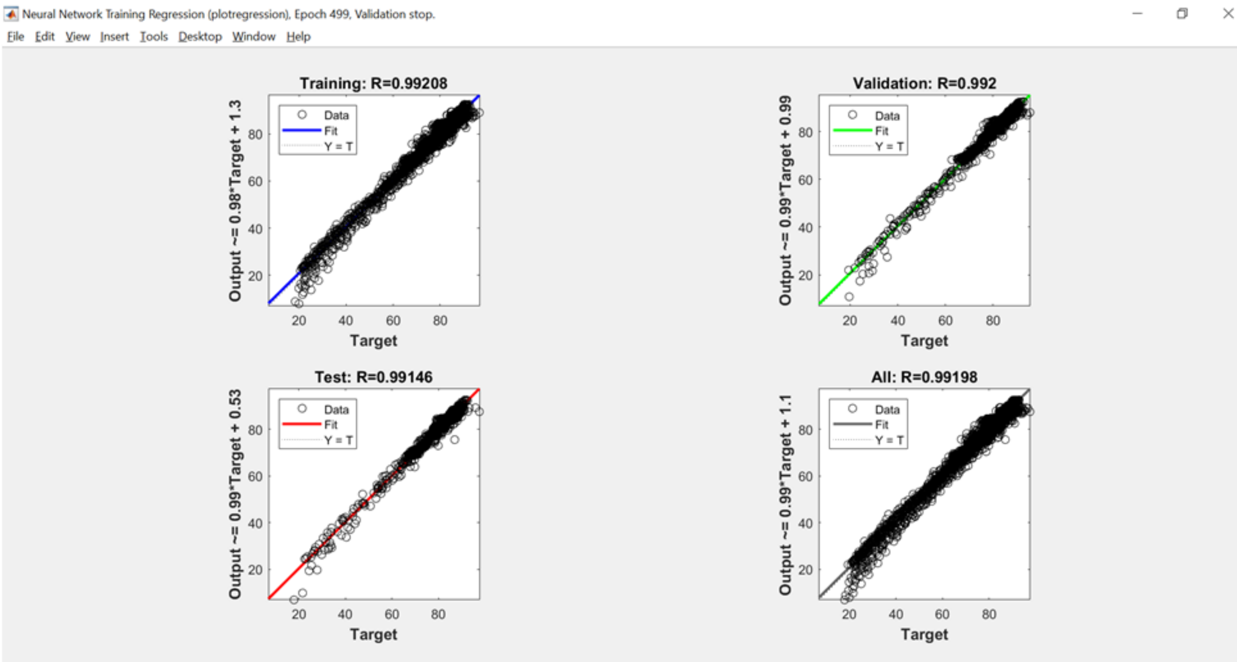
Horizontal Output

Mean Squared Error Plot



The error has reduced as the number of data points used to train the ANN model increased.

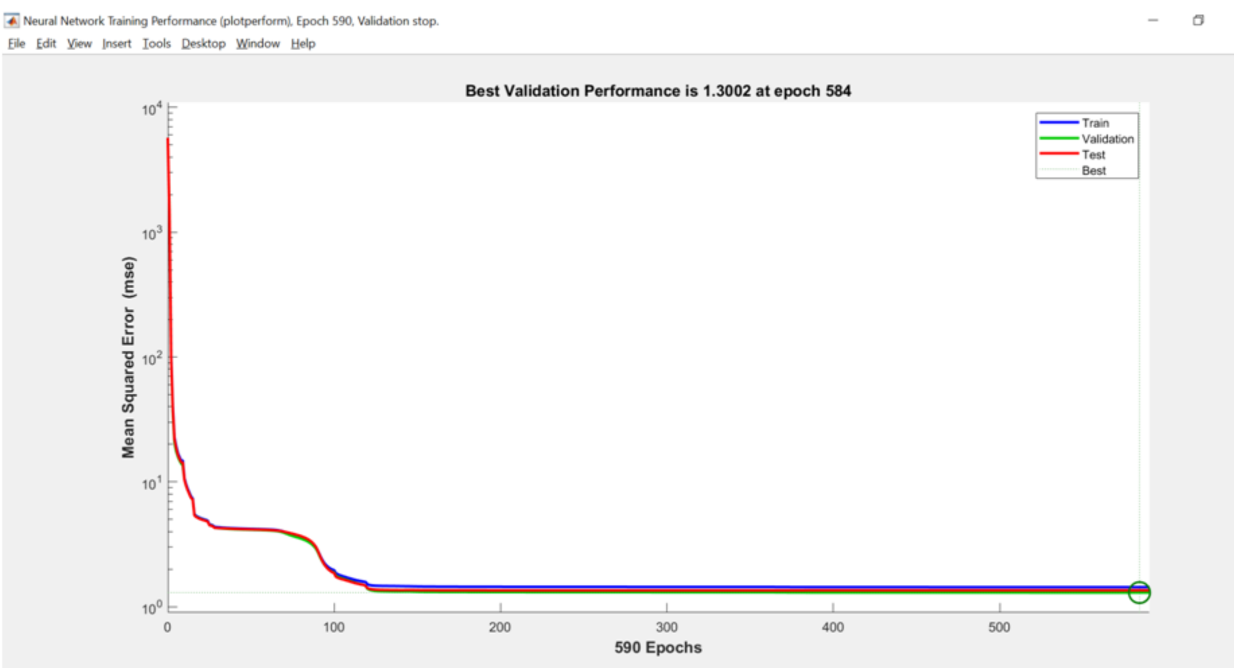
Regression Plot



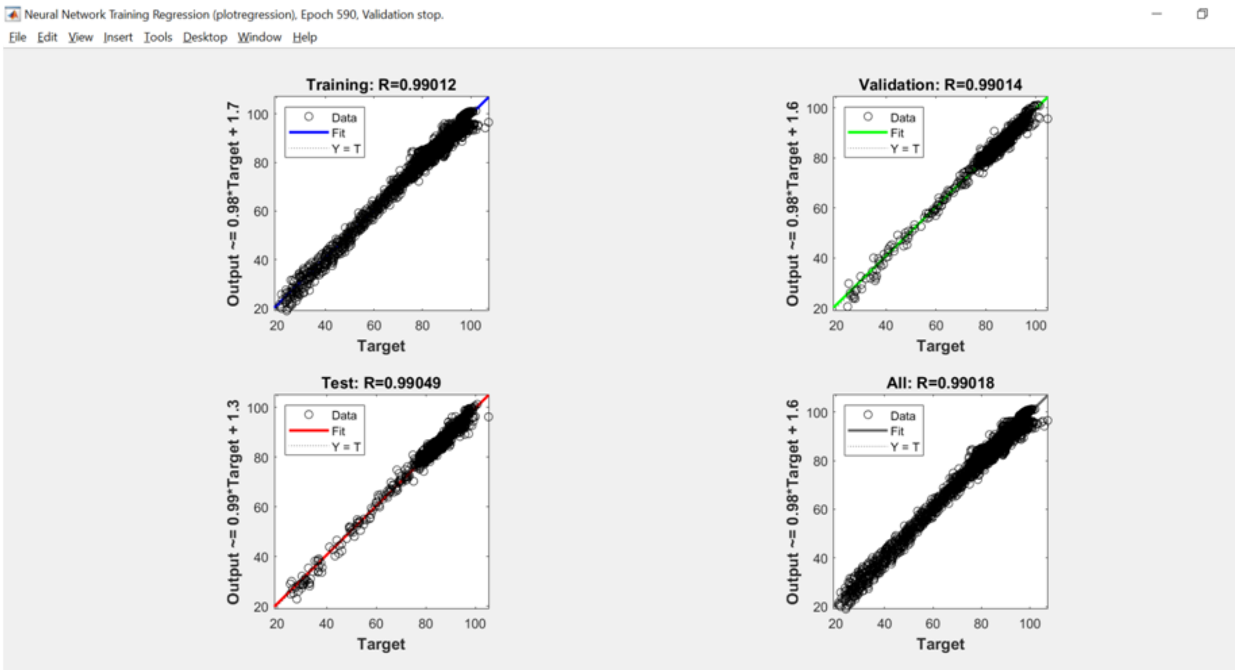
This indicates that the model has provided the horizontal output's predicted values with accuracy of 99.19 %.

Inclined Output

Mean Squared Error Plot



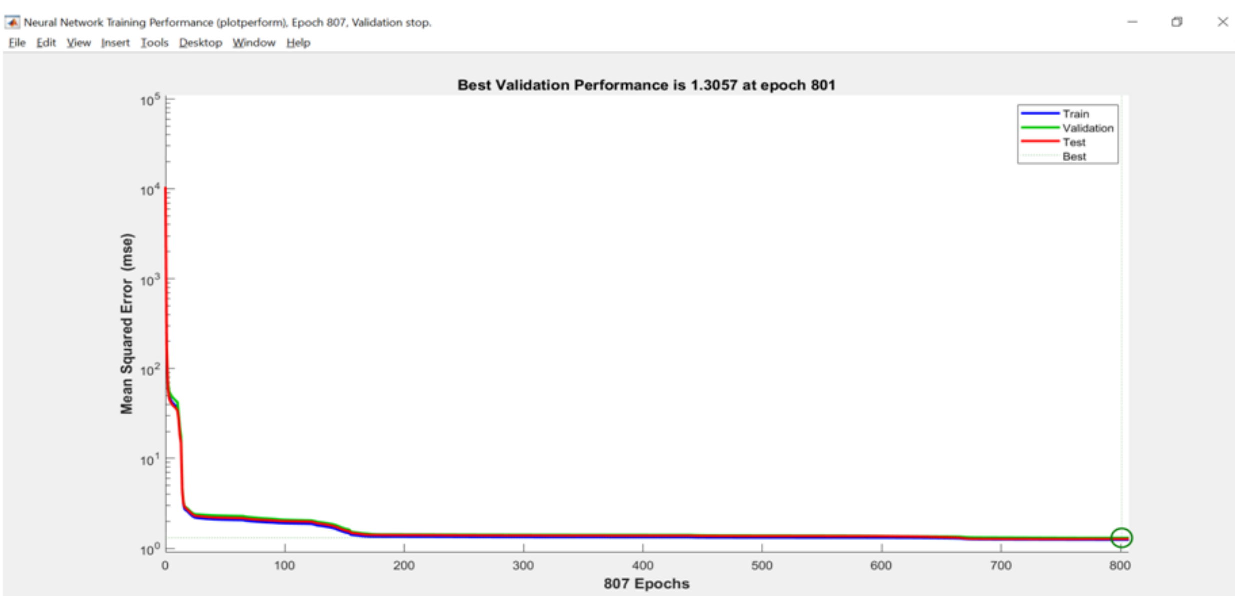
Regression Plot



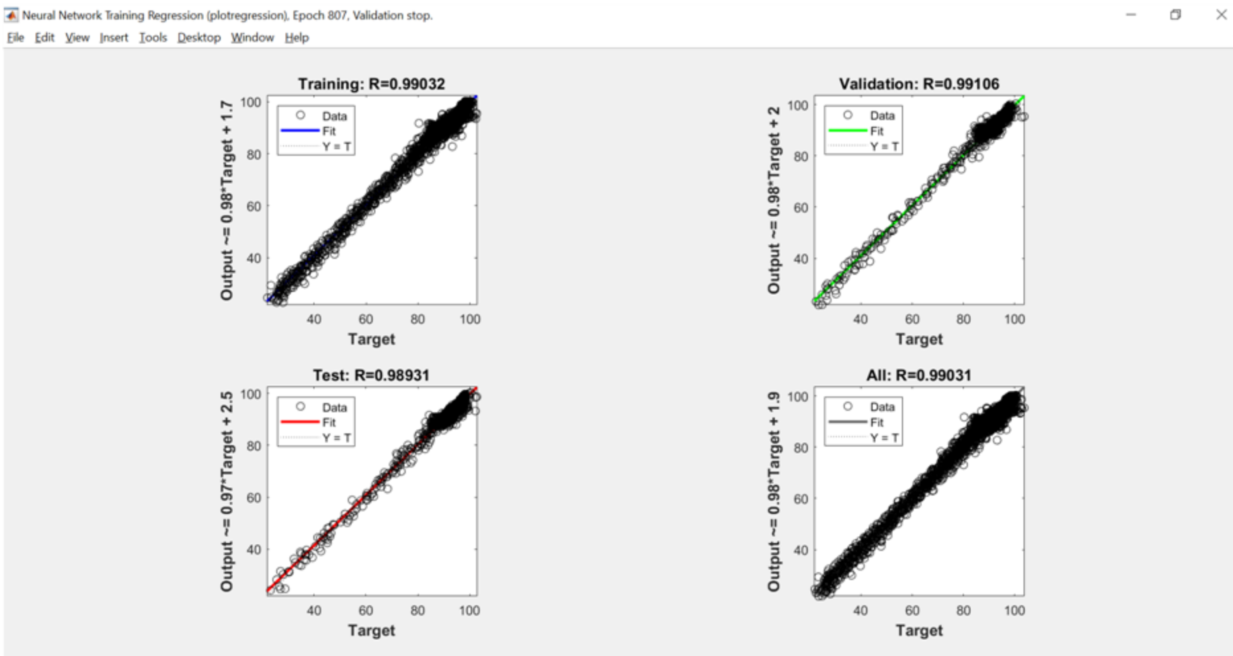
This graph shows that the ANN model has provided the inclined output's predicted values with accuracy of 99.01 %.

Vertical output

Mean Squared Error Plot



Regression Plot



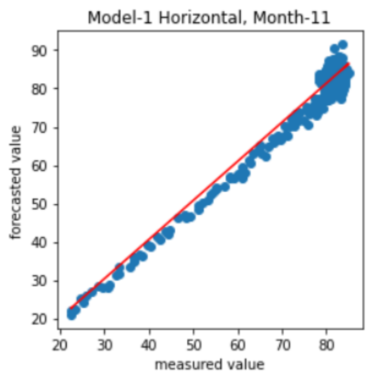
This regression plot shows that the ANN model has given the vertical output's predicted values with an accuracy of 99.03 %.

LONG SHORT-TERM MEMORY MODELS:
We design two LSTM models and compare the accuracy percentages; they work by recognizing a pattern, computing the data till time ‘t-1’to predict the value at a time ‘t’. Both the models are coded in Python (programming language).

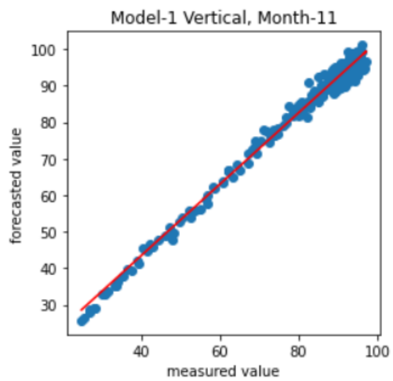
In both the models, we predict transmittance data for months 11 and 12 and compare the predicted values with the measured data to check its accuracy.
Most publications use either MAPE or MAE and R^2 with a graph between forecasted and measured values.
Error statistics MAPE, MAE, RMSE (close to each other in value) are computed, and we check for the correlation coefficients (R, R^2) and plot a linear regression graph with predicted values against measured transmittance, the closer the slope of the line is to one, the better the accuracy.

Model- 1
In the first algorithm, we train the model with the data of the first eight months. It learns how to predict the transmittance value for a month using data from the previous two months and uses the last four months for testing and predictions. After training, the model takes the data of months 9, 10 to predict a value for month 11 and months 10, 11 to predict a value for month 12. By varying the Epoch value, we determine the number of times the data runs through the algorithm. It chooses the run which has the least mean absolute error to forecast values of transmittance.
It is Run with an Epochs value of fifty

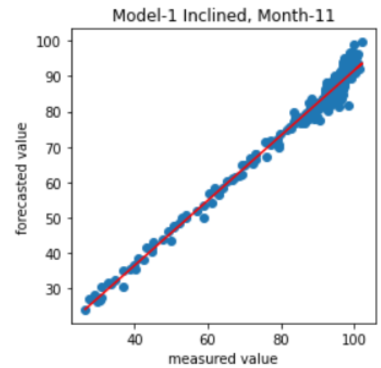
	MAPE	MAE	RMSE	R^2 Month 11	R^2 Month 12
Horizontal	2.488	1.977	2.308	0.907	0.918
Vertical	2.534	2.235	2.557	0.97	0.897
Inclined	5.386	4.92	5.81	0.97	0.966



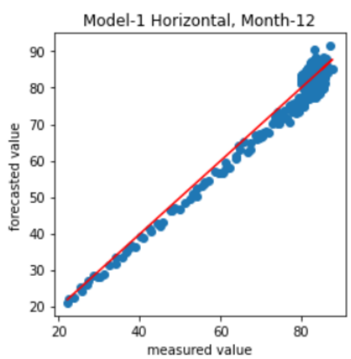
slope= 1.02



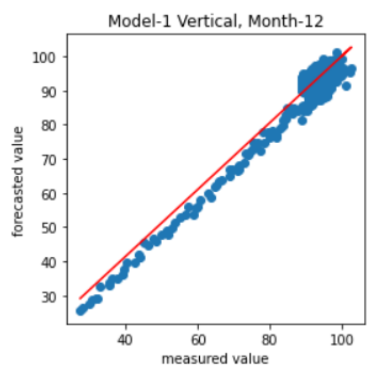
slope= 0.98



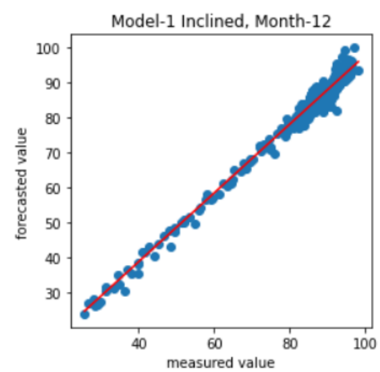
slope= 0.92



slope= 1.004



slope= 0.976



slope= 0.98

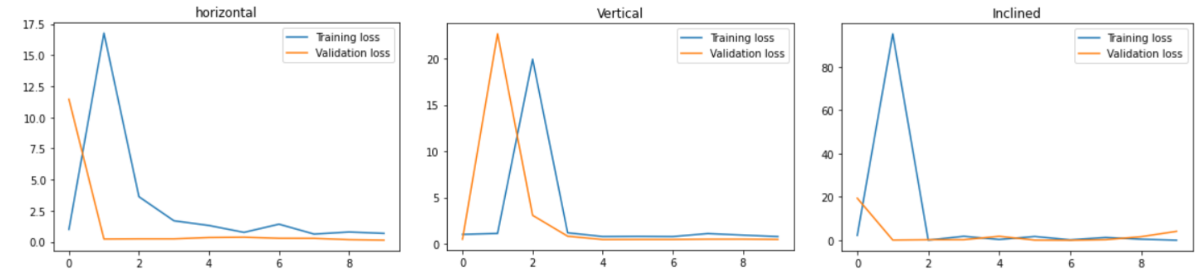
Model- 2

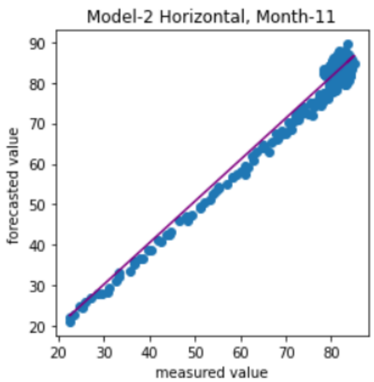
The second algorithm works by splitting the data into ten months for training and validation, two months for testing and prediction. It uses the transmittance data of the first ten months to predict the value for the eleventh month and uses the data of months two to eleven to predict the value for the twelfth month. We run the code multiple times and choose the run with minimal training and validation losses.

The algorithm is run twenty times, and the one with the least MAE is chosen.

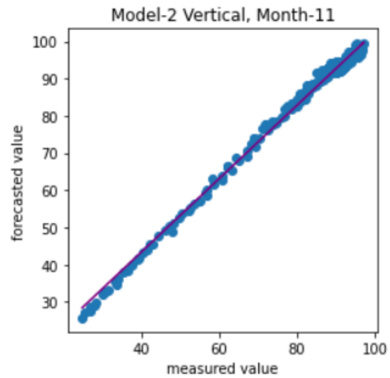
	MAPE	MAE	RMSE	R^2 Month 11	R^2 Month 12
Horizontal	2.325	1.85	2.12	0.928	0.938
Vertical	2.498	2.195	2.47	0.985	0.925
Inclined	5.736	5.03	6.41	0.988	0.986

Training and Validation losses

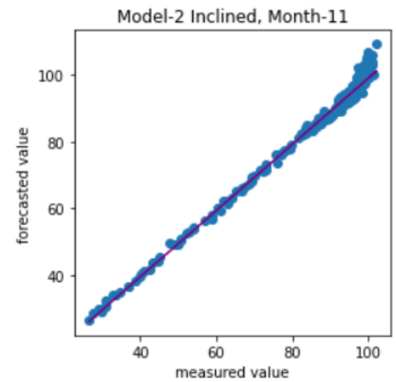




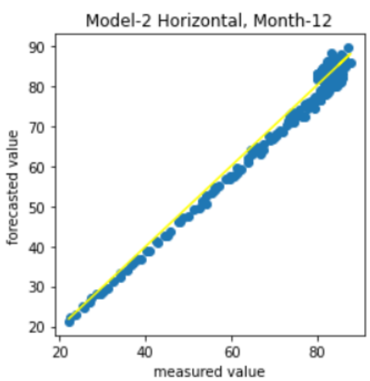
slope= 1.023



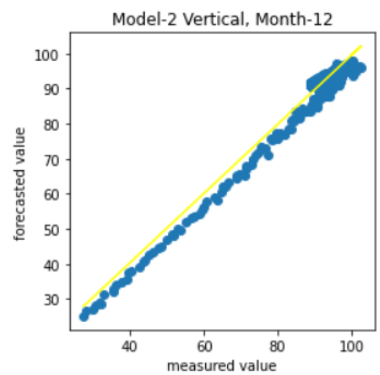
slope= 0.984



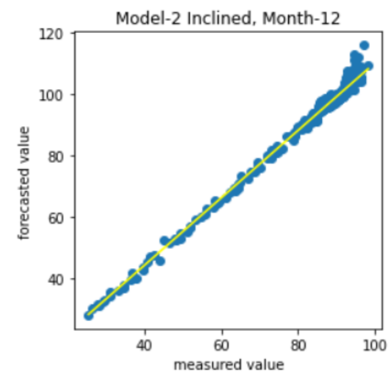
slope= 0.988



slope= 1.008



slope= 0.985



slope= 0.987

As we use most of the data to train the models and predict values only for the last two months, the accuracy is over ninety percent.

CONCLUSION

Analyzing both the models we can see that for the present data sets, both the models are very efficient as they both provide us with results with over 90% accuracy. The ANN model provides an accuracy of over 99% whereas the LTSM model provides us with an accuracy over 92%. Both models provide the best accuracy for the Horizontal inclination data. These calculations helped to predict the power outcome of the solar power plants that helps us to optimize the efficiency of the PV modules. The data from the three inclinations also provide us with the information to understand the effect of inclination on the power output, which also helps in improving the efficiency of the PV module.

REFERENCES

[1] Ibrahim Reda, A. A. (2003). *Solar position algorithm for solar radiation applications.* USA: Solar Energy 76 (2004) 577–589.

[2] Arbab, M. A., & Shapsough, S. (2020). *A Minimal Neural Network Model to Predict Power Loss due to Soiling in Stable Environments*. Department of Computer Science and Engineering American University of Sharjah.

[3] Shapsough, S., Dhaouadi, R., & Zualkernan, I. (2019). *Using Linear Regression and Back Propagation Neural Networks to Predict Performance of Soiled PV Modules*. The 9th International Conference on Sustainable Energy Information Technology (SEIT) August 19-21.

[4] Jong-Min Yeom, K.-S. H. (2010). *Improved estimation of surface solar insolation using a neural network and MTSAT-1R data.* Busan, Korea: Elsevier Ltd.

[5] Zhaoxuan Li, S. M. (2016). *A Hierarchical Approach Using Machine Learning Methods in Solar Photovoltaic Energy.* Energies.

[6] Kumar, S., & Kamal, N. (2020). *A Radial Basis Function Neural Network Based Approach to Mitigate Soiling from PV Module*. Journal of Physics: Conference Series.

[7] R. M. Ehsan, S. P. Simon and P. R. Venkateswaran, "Day-ahead prediction of solar power output for grid-connected solar photovoltaic installations using Artificial Neural Networks," 2014 IEEE 2nd International Conference on Emerging Electronics (ICEE), 2014, pp. 1-4, doi: 10.1109/ICEmElec.2014.7151201.

[8] Pedro, H. T. C., & Coimbra, C. F. M. (2012). *Assessment of forecasting techniques for solar power production with no exogenous inputs. Solar Energy, 86(7), 2017–2028.* doi:10.1016/j.solener.2012.04.004

[9] Refaat, S. S., Nounou, H., & Abu-Rub, O. H. (2018). *ANN based Prognostication of the PV Panel Output Power Under Various Environmental Conditions*. Texas A&M University.

[10] Babatunde, A. A., Abbasoglu, S., & Senol, M. (2017). *Analysis of the impact of dust, tilt angle and orientation on performance of PV Plants*. Renewable and sustainable energy reviews, volume 90.

[11] Catelani, M., Ciani, L., Cristaldi, L., Faifer, M., Lazzaroni, M., & Rossi, M. (2012). *CHARACTERIZATION OF PHOTOVOLTAIC PANELS: THE EFFECTS OF DUST*. 2nd IEEE ENERGYCON Conference & Exhibition, 2012 / Advances in Energy Conversion Symp.