Machine Learning Algorithms for Photovoltaic System Power Output Prediction

Spyros Theocharides, George Makrides and George E. Georghiou
PV Technology Laboratory, FOSS Research Center for Sustainable Energy,
Department of Electrical and Computer Engineering,
University of Cyprus
75 Kallipoleos Avenue, P.O Box 20537, Nicosia, 1678, Cyprus

Andreas Kyprianou
FOSS Research Center for Sustainable Energy,
Department of Mechanical and Manufacturing Engineering,
University of Cyprus
75 Kallipoleos Avenue, P.O Box 20537, Nicosia, 1678, Cyprus

Abstract— **Accurate photovoltaic (PV) production forecasting is necessary for the optimal integration of this technology into existing power systems and is important for both grid and plant operators. The purpose of this work is to assess the performance of different machine learning models for predicting the power output of PV systems. Specifically, a variety of methods were explored including artificial neural networks (ANNs), support vector regression (SVR) and regression trees (RT), with varied hyper-parameters and features. The output power prediction performance of each model was tested on actual PV production data-sets acquired over a period of a year and compared against an existing persistence model (PM). The comparative analysis between the different optimally devised models showed that the ANN outperformed the other models, achieving the lowest prediction mean absolute percentage error (MAPE) and normalised root mean square error (nRMSE) of 0.6 % and 0.76 %, respectively. The prediction capabilities of the SVR and RT models can be considered equivalent for this case since the nRMSE results were 1.13 % and 1.33 %, respectively. Finally, all the models achieved higher relative improvement over the PM since the skill score (SS) ranged from 86 % to 92 %.**

Keywords: artificial neural networks, photovoltaic, prediction, regression tree, support vector regression.

I. INTRODUCTION

The high penetration level of photovoltaic (PV) systems to the utility grid and the intermittent nature of the generated power, introduce new challenges for the stability of electricity grids. In particular, utilities can sustain large shares of PV within the distribution network provided that the electrical quality of the supply is addressed through the use of ancillary services and modern control capabilities. In this aspect, accurate PV production forecasting can mitigate the power quality effects posed by large shares of distributed systems through active grid management and is, therefore, an important feature that can assist utilities and plant operators in the direction of energy management and dispatchability planning. More specifically, short-term PV production forecasts (intra-hour) is necessary for power ramp and voltage flicker prediction as well as control operations and dispatch management. On the other hand, mid-term PV production forecasting (intra-day and day-ahead) is used for load consumption and production monitoring to control voltage and frequency levels and reduce secondary reserve.

PV production forecasts of both point and aggregated systems, have only recently been introduced into the electricity system operational practices. Most commonly adopted approaches are based on parametric models that require detailed PV system parameters and characteristics [1-3]. Because this information is not always readily available in most cases, some simplifications and assumptions are made which subsequently affect the accuracy of the estimations. In this sense, parametric models are driven by the performance of the component models and parameters.

In the scope of improving the forecasting accuracy of the power produced, adaptive machine learning approaches that can capture system behaviour without the need of datasheet and installation parameters have been gaining ground. In this respect, machine learning models that do not require knowledge of the physical parameters but construct relationships between input and output variables based on data-driven approaches have already been developed and presented in the literature [4-9]. However, as most approaches are not fully tested and verified on a large amount of field data and different technologies, there is yet no complete universal forecasting model and methodology to ensure accurate forecasts according to the technology and location. In addition, a primary objective is to develop a methodology that extends the forecasting accuracy beyond the state-of-the-art pursuing in parallel minimal computational and data complexity.

The scope of this work is to present the methodology to derive accurate day-ahead hourly power predictions for PV systems, with the use of novel machine learning approaches optimised according to the selection of input features and architectural parameters. In particular, the machine learning methods designed for PV power prediction included Artificial Neural Networks (ANNs), Support Vector Regression (SVR) and Regression Trees (RT). The basic methodology followed was to train the models with acquired data-sets and construct relationships between the input features and the output which in this case is the power prediction for the new time. For this investigation PV operational and meteorological data-sets were acquired over a period of a year for a grid-connected poly-

978-1-5386-3669-5/18/$31.00 ©2018 IEEE

crystalline Silicon (poly-c Si) PV system installed at the outdoor test facility of the University of Cyprus (UCY) and used for both the development and benchmarking of the models. The acquired data-sets were utilised in order to train the different designed models by applying several combinations of training features while in parallel changing architectural parameters, in order to optimise the model and reduce the prediction accuracy. As such, a comparative assessment between the performances of each devised prediction method was directly performed based on commonly used metrics in the PV production forecasting research. Furthermore, the prediction accuracy of the advanced machine learning prediction models was compared to a baseline persistence model (PM) which is a naïve predictor that simply predicts that the output power will remain the same for the next forecast interval (in the case of day-ahead hourly data, the power output from the previous 24 hours is used as the prediction).

II. Experimental Setup

A. PV System Measurements

A grid-connected PV system that comprises of five poly-c Si PV modules, rated at 235 W_p each, as depicted in the manufacturer's datasheet, was used in this investigation for benchmarking testing. The modules of the system are connected in series to form a PV string at the input of a string inverter. The system was mounted in portrait arrangement on aluminium mountings, at the optimum annual energy yield plane-of-array (POA) angle for Cyprus of 27.5˚. The main technical specifications of the test PV system are summarised in Table I.

TABLE I. Installed PV system technical characteristics.

Technical characteristics	
Modules	5 × poly-c Si 235 W_p
System power (datasheet)	1175 W_p
Efficiency	14.40 %

Additionally, the PV system was connected to an advanced data-acquisition platform, used for the monitoring and storage of meteorological and PV system real outdoor operational data. The platform comprises of meteorological and PV operational measuring sensors connected to a central data acquisition system. The performance of the system and the prevailing meteorological conditions were recorded according to the requirements set by the IEC 61724 [10]. Specifically, the meteorological measurements include the incident global irradiance (G_I), relative humidity (RH), wind direction (W_a), wind speed (W_s), as well as ambient temperature (T_{amb}). The PV system operational measurements include maximum power current (I_{mp}), voltage (V_{mp}) and power (P_{mp}), as measured at the output of the PV system (dc side) [11, 12].

III. Methodology

The methodology followed to develop optimal machine learning models for day-ahead PV power predictions included a training phase (to effectively apply a learning technique to a performance function), a validation phase (to identify the important features and architectural parameters of each model) and a testing phase (to assess the prediction accuracy performance). More specifically, in order to capture the systematic behaviour of a PV system and to design optimal machine learning models a sequence of training and validation stages must be performed by varying the input parameters (features), training duration and architectural parameters of the devised models. It is important to mention that this work focuses on comparing and improving the accuracy of derived machine learning models (as the training and testing were performed on historical yearly data) that will subsequently operate on provided weather variable forecasts. The study does not, therefore, consider the Numerical Weather Prediction (NWP) associated errors in day-ahead PV production forecasts and only investigates the optimisation of the machine learning power prediction models (see Fig. 1).

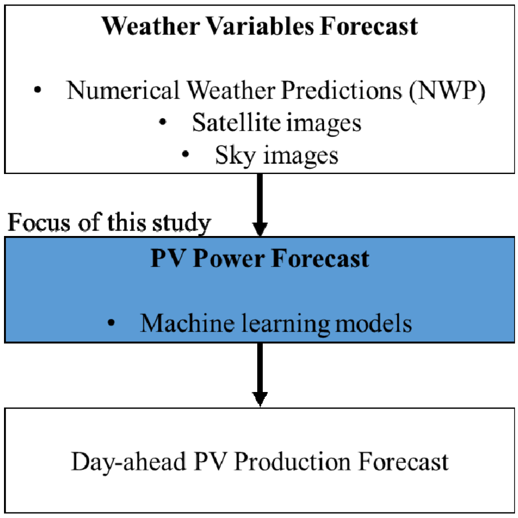

Fig. 1. Flow-chart of a typical approach for generating short-term day-ahead PV production forecasts from NWP and machine learning models. The focus of this study is to improve the PV power prediction stage of the PV production forecasting process).

For the development of the machine learning models, their optimisation and benchmarking, the yearly acquired data-set of the PV system installed at the UCY was used. The yearly data-set was separated into the train, validation and test set with two different data split approaches. The first approach utilized successive samples that were split in a conventional 70:15:15 % portion. The second approach included samples taken randomly from the entire data-set at a 70:15:15 % portion. The 70:15:15 % data split portion was selected in order to include as much systematic information for the training but also keep enough days for validation and testing. Although there are many different ways to split the data-set (training/ testing set and cross-validation), in this work we used data splitting that involves partitioning the data into an explicit training set used to prepare the model, the validation set used to tune the parameters of the model and the test set necessary to evaluate the model performance on unseen data. The training, validation and testing set, comprised of the model inputs which included the measurements of G_I, T_{amb}, RH, W_S, W_a and the calculated parameters of the Azimuth (AzS) and Elevation (AlS) Angle of the Sun, calculated using solar position algorithms [13] and used to address the angular response of the PV system. Finally, the dc P_{mp} was the output of each model.

A. Artificial Neural Networks

ANNs use concepts borrowed from an understanding of human brains in order to model arbitrary functions and in spite of their complexity can be applied to model real-world problems. An ANN constructs relationships between a set of input features and the output using a model derived from our understanding of how a biological brain responds to stimuli from sensory inputs. In the same manner, as the brain uses a network of interconnected neurons, the ANN uses a network of artificial neurons or nodes to solve complex non-linear problems. A typical artificial neuron with n input dendrites can be represented by the following function:

$$y(x) = f\left(\sum_{i=1}^{n} w_i x_i\right) \quad (1)$$

where w_i are the weights that allow each of the n number of x inputs to contribute to the sum of input signals. The net total is used by the activation function $f(x)$ to produce the resulting signal $y(x)$ which is modelled as the output axon.

Training ANNs is generally performed by applying a learning model to a cost function. More specifically, the training phase in neural networks is formulated in terms of the minimisation of a loss function, which in this case comprised of the data-set and a regularisation term. The error term models the discrepancy between the produced output of the network and the desired output. This is measured utilising common indicators such as the mean square error (MSE). In a multi-layer network, minimising the error by training the network is achieved with the back-propagation (BP) algorithm, which is used to calculate the error contribution of each neuron after a batch of data by distributing the error back from the output through the network layers [13].

In addition, the regularisation term is used to prevent overfitting, by controlling the effective complexity of the neural network. The regularisation of the designed networks in this study was performed by adding a penalty equal to the L2-norm of the weights, to reduce the value of the weights by the same factor.

In order to achieve optimal prediction performance, different input feature combinations and hidden layer topologies (number of hidden units) were investigated while training the network models. Among the different forms of ANNs, a designed Feed Forward Neural Network (FFNN) is proposed in this research. An FFNN is defined as a simple type of neural network in which the information flow is in the forward direction from the input towards the hidden and output nodes. The initially designed Feed Forward Neural Network (FFNN), which was subsequently used in the tuning process, comprised of two inputs (G_I, T_{amb}), four hidden nodes and an output layer (P_{mp}) (see Fig. 1).

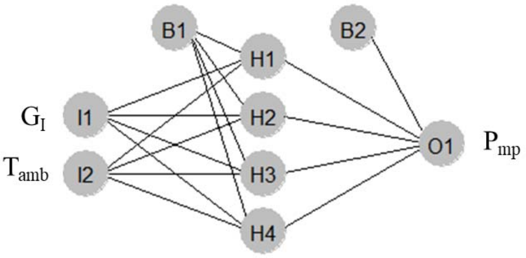

Fig. 2. Network interconnection diagram (NID) of an FFNN ANN.

The optimal network parameters were identified by comparing the performance of the network created based on the training set, with the validation set. Re-iteration of the process (epochs) was performed until the results from the network became acceptable while avoiding overtraining the network, that may result in overfitting.

Lastly, the tuned ANNs were then evaluated by feeding the test set input variables and assessing the actual measured power against the predicted power according to commonly used performance assessment metrics.

B. Support Vector Machine

Support Vector Machines (SVMs) use multidimensional surfaces to define the relationship between features and outcomes. The SVM approach was first introduced for classification purposes [14] and later was abstracted to SVR for regression applications [15]. SVR is a variant of SVM that is based on the principle of deriving a function $f(x)$, that maps patterns of the inputs $x_i \in \mathbb{R}$ to the output label $y_i \in \mathbb{R}$, based on a given train set and solving an optimisation problem. The SVR model implemented for PV power predictions was based on a probabilistic model that assumes (zero-mean) Laplace-distributed errors for the predictions and estimates the scale parameter using maximum probability. In this process, the input vector x is mapped into a high-dimensional feature space by using a non-linear mapping process and then linear regression is performed in the feature space. The developed model applied ε-SVR (ε is a parameter that indicates the tolerance margin of the error within which no penalty is associated with the training loss function with points predicted within a distance ε from the actual value) with total control on the cost of constraints C (error allowed for the model to have with anything beyond the specified ε will be penalised in proportion to C which is the regularization parameter). The kernel basis function used to transform the non-linear data into a higher dimensional feature space to make it possible to perform the linear separation is a sigmoid kernel:

$$K(X,Y) = \tanh(\gamma \cdot X^{\mathrm{T}} Y + r) \quad (2)$$

where X and Y are the training vectors, X^{T} is the transposed input vector, γ is a scaling parameter of the input data, and r is a shifting parameter that controls the threshold of mapping.

Furthermore, the L2 regularisation term was used to prevent overfitting, by controlling the effective complexity of the SVR model. An identical approach of splitting the data-set to 70:15:15 % (train, validation and test set) was followed in order to examine the optimal architectural parameters of the SVR model. The optimal hyper parameters of the SVR model were found through a grid search tuning process of the cost of constraints C and the kernel scaling γ parameter.

C. Regression Trees

Regression trees are methods that recursively partition the data to a simple prediction model within a partition of the split data, using the Analysis of Variance method (ANOVA) in order to analyse the differences or variations among the data partitions. A RT used for numeric prediction is built beginning at the root node where the data-set is partitioned using a divide-and-conquer strategy according to the feature that will result in the greatest increase in homogeneity in the outcome after a split is performed. Homogeneity is measured by statistics such as variance, standard deviation, or absolute deviation from the mean and depends primarily on the tree-growing algorithm used. A common splitting criterion is called the standard deviation reduction (SDR) given as:

$$SDR = sd(T) - \sum_i \frac{|T_i|}{|T|} \times sd(T_i) \tag{3}$$

where the $sd(T)$ function refers to the standard deviation of the values in the data-set T, while T_i are sets of values resulting from a split on a feature. The $|T|$ term refers to the number of observations in the data-set T. This splitting criterion essentially measures the reduction in standard deviation from the original value to the weighted standard deviation post-split. The RT developed was optimally pruned by applying a complexity parameter *CP* in a cross-validation approach on the validation set. The *CP* specifies how the cost of a tree is penalised by the number of terminal nodes. Small *CP* results in larger trees and potential overfitting, large *CP* in small trees and potential under-fitting.

D. Persistence Model

The PM is a data structure algorithm that preserves its previous version when it is reformed. In the PV power forecasting field, PM is the simplest type of forecasting method and hence commonly employed as a baseline for benchmarking. The persistence method assumes that the conditions at the time of the forecast will not change and is therefore functional when weather patterns have minor fluctuations, at climates that present stable behaviour.

E. Performance Metrics

The prediction performance accuracy was assessed based on several predefined metrics when the test set was applied to the developed algorithms. The metrics commonly used in PV production forecasting applications include the mean absolute error (MAE) which measures the difference between two actual and predicted data, the mean absolute percentage error (MAPE) which measures the prediction accuracy of a forecasting method, the root mean square error (RMSE) which describes the standard deviation of the prediction errors, the normalised root mean square error which is the RMSE normalised to the nominal capacity of the PV system and the skill score (SS) which describes the accuracy and degree of association of the prediction model to a reference model (SS = 100 % indicates perfect prediction while SS = 0 indicates that the prediction model equals to the reference model). More specifically, the metrics used to analyse the performance of the algorithms are based on the difference between the predicted and actual dc power and are calculated as follows:

$$MAE = \frac{1}{n} \times \sum_{i=1}^{n} \left| y_{\text{actual,i}} - y_{\text{predicted,i}} \right| \tag{4}$$

$$MAPE = \frac{100}{n} \times \sum_{i=1}^{n} \left| \frac{y_{\text{actual,i}} - y_{\text{predicted,i}}}{y_{\text{actual,i}}} \right| \tag{5}$$

$$RMSE = \sqrt{\frac{1}{n} \times \sum_{i=1}^{n} (y_{\text{actual,i}} - y_{\text{predicted,i}})} \tag{6}$$

$$nRMSE = \frac{100}{P_{\text{nominal}}} \times \sqrt{\frac{1}{n} \times \sum_{i=1}^{n} (y_{\text{actual,i}} - y_{\text{predicted,i}})} \tag{7}$$

$$SS = 100 \times \left(1 - \frac{RMSE_{\text{predicted}}}{RMSE_{\text{reference}}} \right) \tag{8}$$

where $y_{\text{actual,i}}$ and $y_{\text{predicted,}}$ is the actual and predicted power respectively, P_{nominal} is the nominal peak power of the PV system (≈1365 W), $RMSE_{\text{predicted}}$ and $RMSE_{\text{reference}}$ is the root mean square error of the predicted and reference model (predictions of the PM), respectively.

IV. Results

The FFNN, SVR and RT models comprised of seven input features (G_{I}, T_{amb}, RH, T_{amb}, AlS, AzS, W_{S}, W_{a}) and were all trained with the same train set, that comprised of 70 % of data extracted randomly from the yearly poly-c Si PV system data-set [16]. In the scope of further improving the accuracy of the predictions obtained from the developed models, the effect of varying the number of hidden units of the FFNN, performing an automated grid search by altering the cost of constraints and gamma parameters for the SVR and optimising the complexity parameter (CP) of the RT model, was also investigated. The results presented in Table II showed that the best accuracy for the validation set nRMSE was obtained when the selected number of hidden neurons for the developed FFNN were in the range of 12 to 16. For the RT model, the best cross-validation showed that the highest prediction accuracy was obtained when the *CP* was set in the range of -0.10 to – 0.06. In order to minimise the error by decreasing the bias with more hidden units but at the same time not making the FFNN and RT models too complicated and prone to overfitting the number of nodes and *CP* were set to 12 and -0.10, respectively.

TABLE II. Cross-validation nRMSE for the ANN and RT hyper parameters.

Artificial Neural Network		*Regression Tree*	
Hidden Units	***nRMSE (%)***	***Complexity Parameter***	***nRMSE (%)***

2	0.95	0.22	1.39
4	0.90	0.20	1.39
6	0.80	0.18	1.39
8	0.79	0.16	1.38
10	0.78	0.14	1.37
12	0.76	0.12	1.37
14	0.76	0.10	1.33
16	0.76	0.08	1.33
18	0.81	0.06	1.33
20	0.85	0.04	1.36

The SVR model cross-validation results presented in Table III showed that the best accuracy based on the calculated validation set nRMSE was obtained when the SVR grid search tuning parameters C and γ were set to 1 and 0.1, respectively.

TABLE III. TEST SET VALIDATION NRMSE RESULTS OF THE TRAINING PERIOD AND SAMPLING TYPE OF THE TRAINING DATA-SET.

Gamma Parameter, γ (γ)	**Cost of Constraints Parameter, *C* (C)**			
	1	***100***	***1000***	***10000***
0.0001	1.1395	1.1453	1.1411	1.1388
0.001	1.1405	1.1512	1.1427	1.1762
0.01	1.1387	1.1485	1.1659	1.1793
0.1	1.1321	1.499	1.1439	1.1873

The power prediction results showed that all three machine learning models are relatively accurate exhibiting good agreement with the actual power produced for both clear sky and overcast days (see Fig. 3).

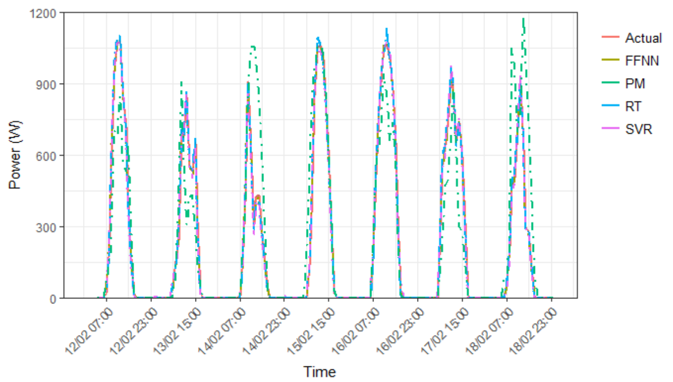

Fig. 3. Daily curves of the actual and the predicted power of all models over a period of a week for the poly-c Si PV system.

Amongst the proposed methods the ANN power predictions outperformed the other methods while the PM provided accurate predictions only for consecutive clear sky days (see Fig. 3).

The daily nRMSE results obtained for the test set period demonstrated that the optimally designed FFNN outperformed the other models based on the lower daily nRMSE compared to the SVR and RT models (see Fig. 4). More specifically, the FFNN showed an average daily nRMSE of 0.76 %, whereas the average daily nRMSE calculated for the SVR and RT models was 1.13 % and 1.33 %, respectively. For some days, the machine learning models exhibited nRMSE accuracies close to 0.50 %. It also apparent that the highest percentile of test set days with a daily nRMSE less than 0.85 % is 80 % for the FFNN and this signifies that the ANN model could capture the behaviour more efficiently compared to the other investigated machine learning models. The simple PM showed an average nRMSE of 9.73 % with some days further demonstrating nRMSE values up to 50 % (see Fig. 4d).

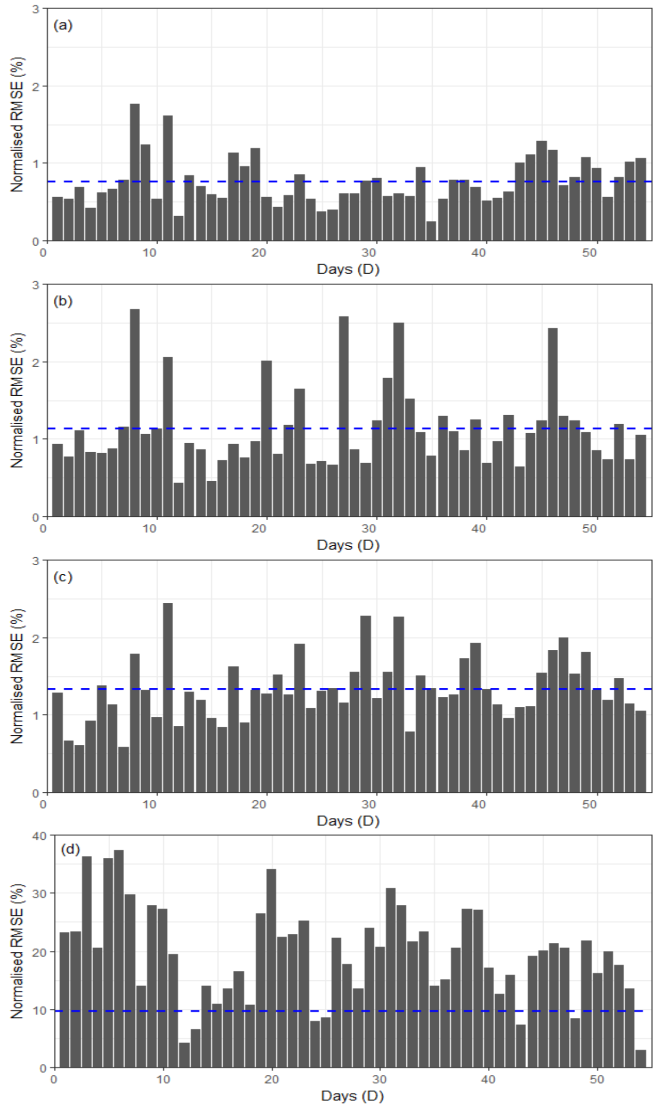

Fig. 4. Daily nRMSE over the test set period for the (a) FFNN, (b) SVR, (c) RT and (d) PM models. The horizontal line shows the average daily nRMSE.

Finally, Table IV summarises the results of all the performance metrics used to assess the developed machine learning models. As a whole, the prediction performance of the FFNN model achieved the lowest MAPE and nRMSE values of 0.6 % and 0.76 %, respectively. This signifies that the FFNN model could predict the power more accurately compared to the SVR and RT models. Based on the similar nRMSE results of the SVR and RT model, 1.13 % and 1.33 %, respectively, the prediction capabilities of the two data-driven approaches can be considered substantially equivalent for this case. The performance of each model was further compared with a persistence method commonly used as a reference in predictions associated with PV production. The performance of each model was evaluated using the SS (a measure of the relative improvement of the prediction over the persistence model, which is less site and year dependent and allows evaluating which model outperforms). All the models achieved higher relative improvement over the reference model since the SS results ranged from 86 % to 92 %.

TABLE IV. TEST SET MAPE, RMSE, NRMSE AND SS OF THE DIFFERENT MACHINE LEARNING MODELS.

	Performance Metrics			
Models	***MAPE (%)***	***RMSE (W)***	***nRMSE (%)***	***SS (%)***
ANN	0.61	10.37	0.76	92.22
SVR	0.75	15.42	1.13	88.34
RT	0.98	18.15	1.33	86.33

V. CONCLUSIONS

The continuous integration of grid-connected PV systems into existing power systems introduces new challenges for the stability of electricity grids and necessitates the requirement for energy management techniques such as accurate PV production forecasting.

The aim of the paper is to compare the prediction performance of several machine learning approaches including ANNs, SVR and RT. A year of measured weather and production data from a poly-c Si system installed at the UCY was employed to train, validate and test the models. For the development of the optimal FFNN, SVR and PM models, a procedure with subsequent stages of training and testing steps was followed by tuning the input features and hyper-parameters of each model. The final ANN, SVR and RT models included 7 input variables ($G_{I,}$ $T_{amb,}$ RH, $T_{amb,}$ AlS, AzS, $W_{S,}$ W_a) and were trained with a random sample 70:15:15 % train, validation and test set approach. The prediction results obtained when the test set was applied to the optimally designed models, demonstrated that the average daily nRMSE was 0.76 %, 1.13 %, 1.33 % and 9.73% for the FFNN, SVR, RT and PM, respectively. Furthermore, some days the machine learning models exhibited nRMSE accuracies close to 0.50 %. The prediction performance of the FFNN outperformed the other machine learning models, achieving the lowest MAPE and nRMSE of 0.6 % and 0.76 %, respectively. This signifies that the FFNN model could predict the power more accurately compared to the SVR and RT models. Based on the similar nRMSE results of the SVR and RT model, 1.13 % and 1.33 %, respectively, the prediction capabilities of the two approaches can be considered equivalent for this case. Finally, all the models achieved higher relative improvement over the reference model, since the SS results ranged from 86 % to 92 %.

ACKNOWLEDGMENT

This work was funded through the INFORPV project which is co-financed by the SOLAR-ERA.NET Transnational Calls of the European Union under the grant agreement number SOLAR-ERA.NET/1215/02. The authors would also like to acknowledge the financial support from the Cyprus Research Promotion Foundation.

REFERENCES

[1] S. Pelland, J. Remund, J. Kleissl, T. Oozeki and K. D. Brabandere, "Photovoltaic and solar forecasting: state of the art", IEA, Technical Report 5-40, 2013.

[2] E. Lorenz, T. Scheidsteger, J. Hurka, D. Heinemann and C. Kurz, "Regional PV power prediction for improved grid integration", Progress in Photovoltaics: Research and Applications, 2011; 19: 757–771, doi:10.1002/pip.1033.

[3] S. Pelland, G. Galanis, and G. Kallos, "Solar and photovoltaic forecasting through post-processing of the Global Environmental Multiscale numerical weather prediction model", Progress in Photovoltaics: Research and Applications, 2013; 21: 284–296.

[4] I. Ashraf and A. Chandra, "Artificial neural network based models for forecasting electricity generation of grid connected solar PV power plant", International Journal of Global Energy, 2004; 21: 119-130.

[5] D. Caputo, F. Grimaccia, M. Mussetta and R. E Zich, "Photovoltaic plants predictive model by means of ANN trained by a hybrid evolutionary algorithm", in: Proceedings of the 2010 International Joint Conference on Neural Networks, 2010, pp. 1-6.

[6] S. A. Kalogirou, "Applications of artificial neural-networks for energy systems", Applied Energy, 2000; 67: 17-35.

[7] U. K. Das, K. S. Tey, M. Seyedmahmoudian, , M. Y. Idna Idris, S. Mekhilef, B. Horan, and A. Stojcevski, "SVR-based model to forecast PV power generation under different weather conditions", Energies, 2017, vol. 10, no. 7, pp. 1-17.

[8] A. Fentis, L. Bahatti, M. Mestari, M. Tabaa, A. Jarrou and B. Chouri, "Short-term PV power forecasting using support vector regression and local monitoring data", 2016 International Renewable and Sustainable Energy Conference (IRSEC), Marrakech, 2016, pp. 1092-1097. doi: 10.1109/IRSEC.2016.7983968

[9] M. Pinho Almeida, O. Perpiñán Lamigueirob, L. Narvartec, "PV power forecast using a nonparametric PV model", In Solar Energy, vol. 115, 2015, doi:10.1016/j.solener.2015.03.006, pp. 354-368.

[10] IEC 61724:1998, "Photovoltaic system performance monitoring — guidelines for measurement, data exchange and analysis", 1st Ed. Geneva, Switz. IEC, 1998.

[11] G. Makrides, B. Zinsser, M. Schubert and G. E. Georghiou, "Energy yield prediction errors and uncertainties of different photovoltaic models", Progress in Photovoltaics: Research and Applications, 2013; 21: 500-516.

[12] G. Makrides, B. Zinsser, M. Norton and G. E. Georghiou, M. Schubert, and J. H. Werner, "Potential of photovoltaic systems in countries with high solar irradiation", Renewable and Sustainable Energy Reviews, 2010; 14: 754-762.

[13] R. Ibrahim and A. Afshin, "Solar position algorithm for solar radiation applications", Solar Energy, 2004; 6: 577-589.

[14] V. N. Vapnik, "The nature of statistical learning theory", Inc., Statistics for Engineering and Information Technology, 2000, 10.1007/978-1-4757-3264-1, pp. 1-299.

[15] A.J. Smola and B. Schölkopf, "A tutorial on support vector regression", Statistics and Computing, 2004, 14: 199–222.

[16] S. Theocharides, G. Makrides, V. Venizelou, P. Kaimakis, A. Kyprianou and G. E. Georghiou, "PV production forecasting model based on artificial neural networks", In Proceedings of 33rd European Photovoltaic Solar Energy Conference and Exhibition (EUPVSEC), Amsterdam, Netherlands, 2017, pp. 1-5.