IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received October 17, 2020, accepted October 25, 2020, date of publication October 29, 2020, date of current version November 12, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3034690

Forecasting Hourly Global Horizontal Solar Irradiance in South Africa Using Machine Learning Models

TENDANI MUTAVHATSINDI[1], CASTON SIGAUKE[1], (Member, IEEE), AND RENDANI MBUVHA[2]

[1]Department of Statistics, University of Venda, Thohoyandou 0950, South Africa
[2]School of Statistics and Actuarial Science, University of the Witwatersrand, Johannesburg 2050, South Africa

Corresponding author: Caston Sigauke (caston.sigauke@univen.ac.za)

This work was funded by Department of Science and Technology - Council for Scientific and Industrial Research (DST-CSIR) National e-Science Postgraduate Teaching and Training Platform (NEPTTP).

ABSTRACT Solar irradiance forecasting is essential in renewable energy grids amongst others for back-up programming, operational planning, and short-term power purchases. This study focuses on forecasting hourly solar irradiance using data obtained from the Southern African Universities Radiometric Network at the University of Pretoria radiometric station. The study compares the predictive performance of long short-term memory (LSTM) networks, support vector regression and feed forward neural networks (FFNN) models for forecasting short-term solar irradiance. While all the models outperform principal component regression model, a benchmark model in this study, the FFNN yields the lowest mean absolute error and root mean square error on the testing set. Empirical results show that the FFNN model produces the most accurate forecasts based on mean absolute error and root mean square error. Forecast combination of machine learning models' forecasts is done using convex combination and quantile regression averaging (QRA). The predictive performance we found is statistically significant on the Diebold Mariano and Giacomini-White tests. Based on all the forecast accuracy measures used in this study including the statistical tests, QRA is found to be the best forecast combination method. QRA was also the best forecasting model compared with the stand-alone machine learning models. The median method for combining interval limits gives the best results on prediction interval widths analysis. This is the first application of LSTM on South African and African solar irradiance data to the best of our knowledge. This study has shown that providing adequate and detailed evaluation metrics, including statistical tests in forecasting gives more insight into the developed forecasting models.

INDEX TERMS Forecast combination, machine learning, neural networks, solar irradiance forecasting, support vector regression.

NOMENCLATURE

AAKR	Auto Associative Kernel Regression
ANN	Artificial Neural Networks
ARIMA	Autoregressive Integrated Moving Average
DBN	Deep Belief Networks
ERBF	Exponential Radial Basis function
FFNN	Feed Forward Neural Networks
GCV	Generalised Cross validation
GHI	Global Horizontal Irradiance
KNN	K-Nearest Neighbour
Lasso	Least Absolute Shrinkage and Selection Operator
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
MAPE	Mean absolute Percentage Error
MLPNN	Multi-Layer Perceptron Neural Network
NWP	Numerical Weather Prediction
PCR	Principal Component Regression
PICP	Prediction Interval Coverage Probability
PINAW	Prediction Interval Normalised Width
PINC	Prediction Interval Nominal Confidence
PIW	Prediction Interval Width
PLAQR	Partially Linear Additive Quantile Regression
PV	Photovoltaic

The associate editor coordinating the review of this manuscript and approving it for publication was Long Wang.

QRA	Quantile Regression Averaging
RMSE	Root Mean Square Error
RNN	Recurrent Neural Network
SVM	Support Vector Machines
SVR	Support Vector Regression
SVRGA	Support Vector Regression based Genetic Algorithm
SVRPSO	Support Vector Regression based Particle Swarm Optimisation

I. INTRODUCTION

The growth of industrialisation globally is resulting in the depletion of fossil fuels which are used for electricity generation [1]. The continuous use of fossil fuels in the generation of electricity also continues to cause environmental problems such as global warming [2]. Renewable energy sources are inexhaustible, clean and environmentally friendly [3]. This calls for researchers to focus extensively on renewable energy modelling and probabilistic forecasting.

Solar energy is one of the most important forms of renewable energies that can contribute to the environmental and energy challenges [4]. However, the integration of solar energy requires accurate forecasts for the effective management of the electrical grid [5]. Electricity utility decision-makers face a challenge of balancing demand and supply of electricity in a cost-effective way which also favours future economic prosperity and environmental security. This paper compares the predictive performance of long short-term memory (LSTM) networks, support vector regression (SVR), and feed forward neural networks (FFNN) model for forecasting short-term solar irradiance.

A. AN OVERVIEW OF THE LITERATURE ON SOLAR IRRADIANCE FORECASTING

Studies on solar energy were first initiated by Liu and Jordan [6]. Since then, researchers have been giving attention to solar irradiance with many modelling techniques being employed. Solar irradiance forecasting mainly consists of physical models and statistical data-driven models [7]. Physical models are based on numerical weather predictions (NWP) for solar irradiance forecasting. According to Sun *et al.* [2], solar irradiance forecasting techniques can be classified into three groups; traditional mathematical statistics, numerical weather forecasting and machine learning. Researchers have recently paid attention to machine learning techniques such as artificial neural networks (ANNs) [8]–[11], and support vector machines (SVM) [12]–[14] in forecasting solar irradiance.

VanDeventer *et al.* [15] proposed a genetic algorithm-based support vector machine model which they used for short-term forecasting of photovoltaic power. Based on the mean absolute percentage error and root mean square error measures the proposed model was found to give more accurate forecasts compared to the traditional support vector machine model. In a recent study, [16] provided a detailed and comprehensive comparative analysis of hybrid models used in solar radiation forecasting. The authors grouped hybrid models into six classes. It is then argued that although the hybrid models are more complex than stand-alone models, forecasting results from these models are generally very accurate.

ANN has been applied extensively in forecasting solar irradiance. Paoli *et al.* [17] presented an integrated model to forecast daily solar irradiance using ANNs. In the study, the multi-layer perceptron neural network (MLPNN) is used to forecast daily solar irradiance. Initially, the seasonal index adjustment method is used to adjust the original solar irradiance sequence. Empirical results from the study showed that the mean absolute percentage error of the MLPNN model is the lowest compared to those from the autoregressive integrated moving average (ARIMA), Bayesian, Markov chain and K-nearest neighbour (KNN) models. Using the K-means clustering algorithm, [18] grouped input data into regions of the reconstructed phase-space which had similar attributes. The authors went on to model different groups using a non-linear autoregressive neural network and then forecasted the solar irradiance on the test data. The results of the study showed that the clustering of the input space is an important task to interpret the behaviour of the series.

Gensler *et al.* [19] in their study, discussed different ANN and deep neural network (DNN) architectures in the field of solar power forecasting. Different models such as physical forecasting model, MLPNN, LSTM, deep belief networks (DBN) and Auto-Encoders were employed and compared. The results of the study showed that deep learning algorithms have better solar power forecasting performance compared to ANN and physical models on the data from 21 solar power plants in Germany. In another study, [20] discussed an application of different machine learning algorithms for short-term solar irradiance forecasting. The authors used an SVR model with fixed parameter optimisation and two hyperparameters optimised SVR models, support vector regression with optimised hyperparameters using Genetic Algorithm (SVRGA) as well as Particle Swarm Optimisation (SVRPSO). Their study used solar irradiance data from Chicago, United States of America. The results showed that SVRPO outperforms SVR model with fixed parameter optimisation and SVRGA.

Solar irradiance forecasting in South Africa using different methods and techniques is discussed in the literature. A hybrid model for solar irradiance and photovoltaic (PV) power short-term forecasting is discussed in [5]. The prediction of solar irradiance is done using physical models known as clear-sky models which estimate solar irradiance in the absence of clouds. Short-term PV forecasting was then done using an auto-associative kernel regression (AAKR) technique which is usually used for fault detection. An application of the proposed model was done using a PV plant located in South Africa. The empirical results from this study showed that the developed model produces accurate solar irradiance forecasts.

Adeala *et al.* [21] presented a study on the prediction of global solar irradiance using multiple linear regression

with the inclusion of weather parameters in addition to the traditional extraterrestrial irradiance and sunshine hours. The study is done in all nine provinces of South Africa. It is shown that the inclusion of weather parameters improves the accuracy and performance of solar irradiance models for some locations. In another study, Mpfumali *et al.* [22] used partially linear additive quantile regression (PLAQR) models in predicting day-ahead GHI. Data from Tellerie radiometric station in the Northern Cape province was used. Two PLAQR models were developed, one without interactions and the other with hierarchical pair-wise interactions obtained using the least absolute shrinkage and selection operator (Lasso). The study also combined forecasts from individual models using the convex combination method and quantile regression averaging (QRA). The results of the study showed that the QRA forecast combination model was the best forecasting model compared with other individual models in the study. A more recent study on global horizontal irradiance using South African data is that of [23]. In this study, [23] used three methods, seasonal autoregressive fractionally integrated moving average (SARFIMA), harmonically Coupled SARIMA (HCSAFRIMA) and Regression model with SARFIMA error terms (SARFIMAX) to address the long-range dependence inherent in the solar irradiance data from three radiometric stations in South Africa. An additive quantile regression model was used for benchmarking with the three developed models. Results from this study showed that long memory is anti-persistent in all models except one that showed persistence.

B. RESEARCH HIGHLIGHTS

Based on the literature review discussed in I-A this study makes the following contributions on global horizontal irradiance forecasting.

1) The Diebold-Mariano tests show that the forecasts in all pairwise comparisons are not equally accurate.
2) Based on the Giacomini-White test, the feed forward neural network has the highest predictive power.
3) Forecasts based on linear quantile regression averaging are unbiased, while those from feed forward neural networks and long short-term memory are all biased with more under-predictions compared to over-predictions.
4) As a stand-alone model, the feed forward neural network is an appropriate model for forecasting hourly solar irradiance.
5) Based on all the forecast accuracy measures and the statistical tests used in the study, the linear quantile regression averaging model produces the most accurate forecasts that are unbiased.
6) The median method for combining interval limits yields better results since the number of forecasts below and above limits decrease significantly.
7) To the best of our knowledge, this is the first paper to use LSTM in forecasting global horizontal irradiance using South African data.

The remainder of this paper is as follows. Section II provides a discussion about the models used in this study. The empirical results and discussion are presented in Section III. Finally, Section IV concludes this paper and discusses future work.

II. MODELS

A. FEED FORWARD NEURAL NETWORKS

A feed forward neural network (FFNN) is a type of artificial neural network (ANN) in which connections between the nodes do not form a cycle or a loop. The study of this technique was first initiated by McCulloch and Pitts [24] who created a computational model of the concept. Different researchers have further expanded the concept of ANN to cover many features [25]. FFNN works by feeding input data in one end which is then processed by the network and comes out as output in the other end. Information flows in the forward direction only. Two basic formulations of ANN are discussed in the next two sections.

1) SINGLE-LAYER PERCEPTRON

Single-layer perceptron is the simplest type of a neural network which has a single layer of output nodes [24]. A single-layer neural network can be described mathematically as follows:

$$y_k = g\Big(\sum_{i=0}^{D} \omega_i x_i\Big), \tag{1}$$

where y_k is the output, $g(\cdot)$ is an activation function, x_i is input and ω_i represents the corresponding weight for x_i. Single-layer neural networks are not usually used in practice, but help in understanding the basic concept of neural networks.

2) MULTI-LAYER PERCEPTRON

Multi-layer perceptron consists of multiple layers of computational units, usually interconnected in a feed-forward way [24]. A multi-layer neural network with one hidden layer can be written as

$$y_k = h\Big(\sum_{j=0}^{M} \omega_{kj}^{(2)} g(a_j)\Big), \tag{2}$$

where,

$$a_j = g\Big(\sum_{i=0}^{D} \omega_{ij}^{(1)} x_i\Big). \tag{3}$$

The multi-layer neural network is very similar to single-layer neural network except that multi-layer neural network's output of the inner layer is again multiplied by a new weight vector and wrapped in an activation function.

B. LONG SHORT-TERM MEMORY NETWORKS

The LSTM (long short-term memory) network is a type of recurrent neural network (RNN). RNN is used when dealing

with sequential data, like time series data. The LSTM network was originally introduced by Hochreiter and Schmidhuber [26] in a paper titled *Long Short-Term Memory*. The LSTM network is different from regular RNN due to the reason that it consists of LSTM blocks instead of nodes. One LSTM block can be represented by the following system of equations,

$$\begin{aligned} f_t &= g(W_t.[x_t, h_{t-1} + b_f]) \\ i_t &= g(W_i.[x_t, h_{t-1} + b_i]) \\ o_t &= g(W_o.[x_t, h_{t-1} + b_o]) \\ c_t &= f_t c_{t-1} + i_t \tanh(g(W_c.[x_t, h_{t-1} + b_c])) \\ h_t &= o_t \odot \tanh(c_t), \end{aligned} \tag{4}$$

where $g(\cdot)$ is the sigmoid function, $\tanh(\cdot)$ is the hyperbolic tangent, x_t is the input vector, h_t is the output vector, c_t is a cell state vector, W are weights and b are biases, f_t, i_t, and o_t are gates of the block.

C. SUPPORT VECTOR REGRESSION

Support vector regression (SVR) is based on support vector machine (SVM) which is a supervised machine learning technique which involves statistical learning theory and the principle of structural risk minimization. SVR was introduced by Drucker *et al.* [27] and was extended from the SVM model. Different basic kernel functions are used in SVM models, which can be classified as polynomial (Poly), Gaussian kernel, exponential radial basis function (ERBF), radial basis function (RBF), sigmoid and linear [1]. The SVR works by mapping the input space into a high-dimensional feature space and constructs the linear regression in it which can be expressed as

$$f(x) = \omega\phi(x) + b, \tag{5}$$

where ω is the weight vector, $\phi(x)$ maps inputs x into a high dimensional feature space that is nonlinearly mapped from the input space x, and b is the bias term.

To calculate the coefficients ω and b it is required to reduce the regularized risk function which can be expressed as

$$\frac{1}{2}\|\omega\|^2 + C\frac{1}{l}\sum_{i=1}^{l} L_\epsilon(y_i, f(x_i)), \tag{6}$$

where $\|\omega\|^2$ is a regularized term which maintains the function capacity. C is a cost error. The empirical term from the second term in equation 6 can be defined as

$$L_\epsilon(y_i, f(x_i)) = \{|y_i = f(x_i)| = \epsilon, |y_i = f(x_i)| \geq \epsilon\}. \tag{7}$$

Equation 6 expressed the transformation of the primal objective function in order to get the values of ω and b by introducing the positive slack variables $\xi_i(^*)$.

$$\begin{aligned} &\text{minimise } \frac{1}{2}\|\omega\|^2 + C\frac{1}{l}\sum_{i=1}^{l}(\epsilon_i + \xi_i^*) \\ &\text{subject to } \alpha(x) = \begin{cases} y_i - \langle\omega, x_i\rangle - b \leq \epsilon + \xi_i \\ \langle\omega, x_i\rangle + b - y_i \leq \epsilon + \xi_i^* \\ \xi_i, \quad \xi_i^* \geq 0. \end{cases} \end{aligned} \tag{8}$$

The optimization problem expressed in equation 8 has to be transformed into its dual formulation by using the Lagrange multipliers to solve it more efficiently.

D. BENCHMARK MODEL

1) PRINCIPAL COMPONENT REGRESSION

Principal Component Regression (PCR), a benchmark model in this study, is a regression analysis that is used to deal with multiple regression data that have multicollinearity [28]. Multicollinearity occurs when one predictor variable in a multiple regression model can be predicted with other variables linearly. It leads to least squares estimates which have large variances, which means they may be distant from their true values. A multiple regression model is given by

$$\mathbf{Y} = \mathbf{X}\beta + \varepsilon, \tag{9}$$

where $\mathbf{Y}$ is a vector of observed values, $\mathbf{X}$ is a matrix of explanatory variables, β is a parameter vector, and ε is vector of error terms. The least squares solution of equation 9 is given by

$$\hat{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}. \tag{10}$$

To get the first b Principal Components (PCs), we approximate the matrix $\mathbf{X}$ using singular value decomposition (SVD):

$$\mathbf{X} = \tilde{\mathbf{X}}_{(b)} + \varepsilon_{\mathbf{X}} = (U_{(b)}D_{(b)})V_b^T + \varepsilon_{\mathbf{X}} = \mathbf{T}_{(b)}\mathbf{P}_{(b)}^T + \varepsilon_{\mathbf{X}}, \tag{11}$$

where $\mathbf{T}$ represents orthogonal scores and $\mathbf{P}$ loadings. U and V are orthonormal and the matrix D is diagonal with positive real entries. Now regressing Y on the scores leads to

$$\hat{\beta} = \mathbf{P}(\mathbf{T}^T\mathbf{T})^{-1}\mathbf{T}^T\mathbf{Y}. \tag{12}$$

E. VARIABLE SELECTION

Variable Selection involves the selection of feature variables that explains a target variable thereby reducing the number of feature variables and complexity of a model. This process is beneficial in terms of avoiding over-fitting, making a model easier to interpret, and reduces in computational time. There are many variable selection methods, however, in this study, we use Lasso (least absolute shrinkage and selection operator) [29]. Let's assume a regression model with response variable Y and predictors $X_1,\ldots,X_p$. The Lasso formulation is given as

$$\min_\beta \frac{1}{h}(\mathbf{Y} - \mathbf{X}\beta) + \lambda\sum_{j=1}^{p}\|\beta_j\|, \tag{13}$$

where h is the number of data points and $\lambda \geq 0$ is the parameter which controls the strength of the penalty. The major advantage of Lasso is that it leads to simpler and more interpretable models that involve only a subset of the predictors [29].

F. FORECAST COMBINATION

Forecast combination is a method used to combine forecasts from different fitted models with a purpose of improving forecast accuracy [30]. There are many forecast combination methods, but this research will only focus on linear quantile regression averaging (LQRA) and convex combination method.

1) QUANTILE REGRESSION AVERAGING

QRA was first initiated by Maciejowska *et al.* [31]. QRA treats forecasts from different models as independent variables and actual observations as a dependent variable (global horizontal irradiance). Let $\hat{y}_{t,\tau}$ be hourly global horizontal irradiance, K be methods used to forecast the next m observations, i.e. m is the total number of forecasts. The forecast combination, $\hat{y}_{t,\tau}^{QRA}$, is given by

$$\hat{y}_{t,\tau}^{QRA} = \beta_0 + \sum_{k=1}^{K} \beta_{t,k}\hat{y}_{t,k} + \varepsilon_{t,\tau}, \quad \tau \in (0,1),\ t = 1, \ldots, m, \tag{14}$$

where $\hat{y}_{t,k}$ represents predictions from method k, $\hat{y}_{t,\tau}^{QRA}$ is the combined forecasts, and $\varepsilon_{t,\tau}$ is the error term. QRA aims to minimise

$$\arg\min_{\beta} \sum_{t=1}^{n} \rho_\tau(\hat{y}_t^{QRA} - \beta_0 - \sum_{k=1}^{K} \beta_{t,k}\hat{y}_{t,k}), \tag{15}$$

where ρ_τ is the pinball loss function given as $\rho_{\tau(s)} = s(\tau - \mathbf{I}(s < 0)$ and $\mathbf{I}(.)$ is an indicator function. In this study we use linear quantile regression averaging.

2) CONVEX COMBINATION

Convex combination method computes the sequence of instantaneous losses suffered by the predictions from the experts (models) using a loss function [32]. The loss function can be based on square, absolute, percentage, or pinball loss. The combined forecasts will be compared with forecasts from each model using the equation given as

$$\hat{y}_{t,\tau}^{c} = \sum_{m=1}^{M} \omega_{mt}\hat{y}_{mt,\tau}, \tag{16}$$

where ω_{mt} is the weight given to forecast m.

G. PREDICTION INTERVALS

The prediction interval widths (PIWs) for every model, $M_j, j = 1, \ldots, k$, are denoted as $\text{PIW}_{ij}, i = 1, \ldots, m$, $j = 1, \ldots, k$. PIW_{ij} is calculated as

$$\text{PIW}_{ij} = \text{UL}_{ij} - \text{LL}_{ij}, \tag{17}$$

where UL_{ij} and LL_{ij} are the upper and lower limits of the prediction interval, respectively. Probability density plots and box and whisker plots will be used in this study to find the model which yields narrower PIWs.

The probability that the forecast $\hat{y}_{t,\tau}$ lies in the prediction interval $(\text{LL}_{ij}, \text{UL}_{ij})$ is known as a prediction interval with nominal confidence (PINC) of $100(1-\alpha)\%$. PINC is given by

$$\text{PINC} = P(\hat{y}_{t,\tau} \in (LL_{ij}, UL_{ij}) = 100(1-\alpha)\%. \tag{18}$$

There are many indices used to assess the reliability of prediction intervals, however, in this study we use the prediction interval coverage probability (PICP) and the prediction interval normalised average width (PINAW) [33]. PICP is given by

$$\text{PICP} = \frac{1}{m}\sum_{i=1}^{m} I_{ij}, \tag{19}$$

where m is the number of forecasts and I is a binary variable given by,

$$I_{ij} = \begin{cases} 1, & \text{if } y_i \in (LL_{ij}, UL_{ij}) \\ 0, & \text{otherwise.} \end{cases} \tag{20}$$

PINAW is another index used to evaluate the reliability of prediction intervals and is given by

$$\text{PINAW} = \frac{1}{m(\max(y_{ij}) - \min(y_{ij}))}\sum_{i=1}^{m}(UL_{ij} - LL_{ij}), \quad j = 1, \ldots, k. \tag{21}$$

H. EVALUATION METRICS

The mean absolute error (MAE) and root mean squared error (RMSE) are going to be used to evaluate the accuracy of our forecasts. The equations of the above measures are respectively given by

$$\text{MAE} = \frac{1}{m}\sum_{t=1}^{m}|y_t - \hat{y}_t|, \tag{22}$$

$$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^{m}(y_t - \hat{y}_t)^2}{m}}, \tag{23}$$

where $\hat{y}_t$ are predicted values by the model, y_t are the values actually observed, and m is the number of predictions.

I. TESTS FOR PREDICTIVE ACCURACY

1) DIEBOLD-MARIANO TEST

Testing the statistical significance in the accuracy of two forecasts is very important as this helps to determine a model with higher predictive ability. One such test for predictive accuracy is the Diebold-Mariano(DM) test [34], [35]. Under this test, the null hypothesis is that of equal predictive accuracy of two competing forecasts. One of the main advantages of this test is that it takes into account the sampling variability in the average losses [36].

Let $y_t, t = 1, \ldots, m$ be the GHI values and two forecasts $\hat{y}_{i,t}, \hat{y}_{j,t}, \forall i \neq j, i, j = 1, 2, \ldots, K$. Assuming the errors from the forecasts are defined as $\varepsilon_{i,t} = \hat{y}_{i,t} - y_t, i = 1, 2$. If $g(\varepsilon_{i,t})$ is an error loss function, then a loss function which penalises heavily underprediction than overprediction is given as [35]:

$$g(\varepsilon_{i,t,\tau}) = e^{\lambda\varepsilon_{i,t}} - 1 - \lambda\varepsilon_{i,t}. \tag{24}$$

2) GIACOMINI-WHITE TEST

A generalisation of the DM test known as the Giacomini-White (GW) test, tests the conditional predictive ability of two competing forecasting methods [37]. The GW test, tests for equal conditional predictive ability. The GW test has an additional advantage over the DM test in that it accounts for uncertainty in the estimation of parameters [37]. The GW test is carried out under two test approaches. One is based on the unconditional test which only focuses on the restriction to the forecast model. The other is the conditional approach which allows the evaluation of the model, the estimation method used and the choice of the forecasting horizon [36].

J. DATA AND VARIABLES

The data used in this study are from the Southern African Universities Radiometric Network (SAURAN) database, accessible at http://sauran.net [38]. The focus is on only one radiometric station found at an inland region in Pretoria, South Africa. The station of focus is the University of Pretoria radiometric (UPR) station which is on latitude -25.75308 and longitude 28.22859 and elevation 1410m. The pyranometer (See Fig. 1) is on the roof of a university building.

FIGURE 1. UPR-GIZ University of Pretoria pyranometer. Source: https://sauran.ac.za/.

This study seeks to model hourly solar irradiance (global horizontal irradiance (GHI)) using independent variables such as temperature, wind speed, relative humidity, barometric pressure, wind direction standard deviation, rainfall, hour, month, a non-linear trend, and a lagged hourly solar irradiance at lags 1 and 2.

The data used is hourly solar irradiance from 1 January 2014 to 31 December 2018 with 20056 observations with hours from 7:00 AM to 5:00 PM considered to be sunshine hours. The UPR data set is split into training data, 1 January 2014 to 31 December 2017, i.e., $n_1 = 16044$ and testing data, from 1 January 2017 to 31 December 2018, i.e., $n_2 = 4012$, which is 20% of the total number of observations.

K. COMPUTATIONAL TOOLS

The software packages that are used for data analysis in this study are R and Python. The FFNN and LSTM networks algorithm are implemented in this study using the Keras deep learning package (https://keras.io/). SVR is implemented using the sklearn python package.

III. EMPIRICAL RESULTS AND DISCUSSION

A. EXPLORATORY DATA ANALYSIS

The summary statistics of hourly solar irradiance for the sampling period January 2014 to December 2018 is given in Table 1 for the UPR radiometric station. The distribution of hourly solar irradiance is not normally distributed since it is skewed to the right and platykurtic as shown by the skewness value of 0.089 and a kurtosis value of -0.997 given in Table 1.

TABLE 1. Descriptive statistics for hourly solar irradiance (W/m^2).

Min	Median	Mean	Max	St.Dev.	Skewness	Kurtosis
0.0	496.7	496.6	1173.9	298.760	0.089	-0.997

Figure 2 shows the time series plot of hourly solar irradiance together with density, normal quantile to quantile (Q–Q) and box plots which all show that the data is not normally distributed.

The non-linear trend values used in this project are obtained or extracted by fitting a cubic smoothing spline function which is given by:

$$\pi(t) = \sum_{t=0}^{n} (y_t - f(t))^2 + \lambda \int \{f''(t)\}^2 dt, \qquad (25)$$

where λ is a smoothing parameter which is estimated using the generalised cross validation (GCV) criterion.

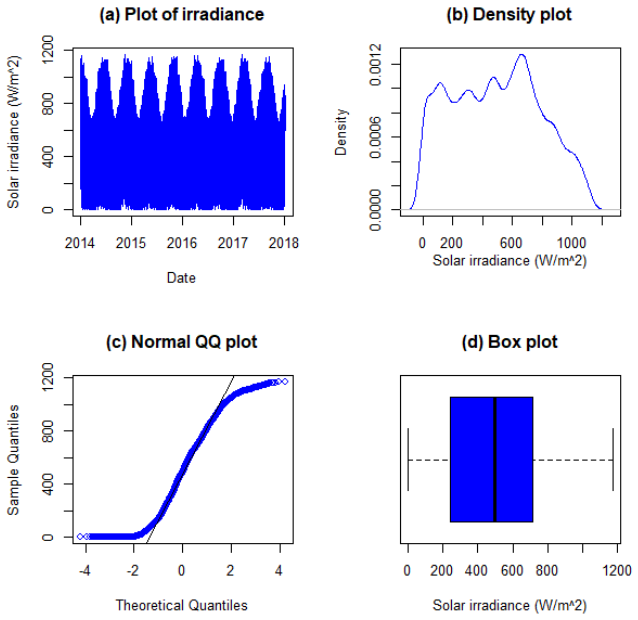

FIGURE 2. Diagnostic plots for hourly solar irradiance (W/m^2).

Figure 3 shows box plots of the monthly and hourly distribution of hourly irradiance data from 2014 to 2018, respectively. It can be seen that solar irradiance shows seasonality since its values are low during the winter season and are high during the summer season. The solar irradiance series is also highly correlated with the hour of the day since it tends to be low in the morning and increases towards the peak in the afternoon and decreases towards the evening.

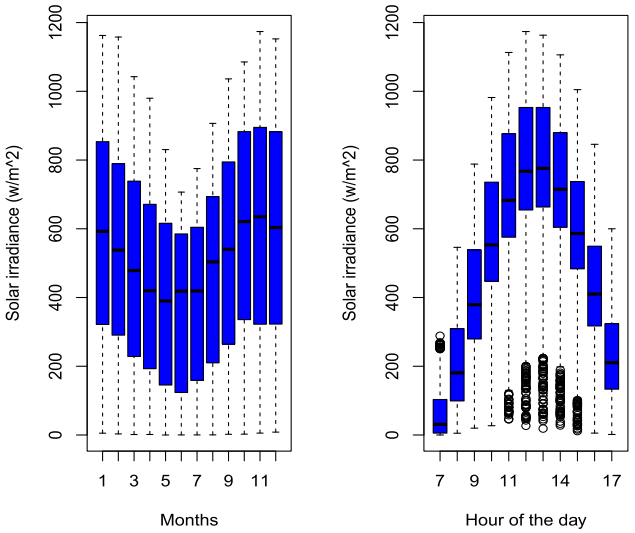

FIGURE 3. Distribution of monthly (left panel) and hourly (right panel) solar irradiance data.

VARIABLE SELECTION USING LASSO

Different weather variables are recorded by SAURAN at different radiometric stations. Variable selection is done in this study using the Least Absolute Shrinkage and Selection Operator (Lasso) to remain with significant variables in predicting global solar irradiance. Figure 4 shows the importance of 13 weather variables in predicting global solar irradiance as assessed by Lasso. Wind direction (Wd) and day of the month have the least significance in forecasting solar irradiance and are therefore not considered in the modelling stage.

B. MACHINE LEARNING RESULTS

The models considered are M2 (LSTM), M3 (SVR), and M4 (FFNN) which are all machine learning models and M1 (PCR) which is a benchmark model in this study. RMSE and MAE are normally used for forecast evaluation and are used in this study.

The parameters for the machine learning models were set as follows:

- LSTM: Hidden neurons were 100. The input layer of our trained LSTM network had 12 features and 1 timestep. Output layer with no activation function had one neuron. Maximum epochs were set to be 300. The performance was measured by minimizing mean square error.

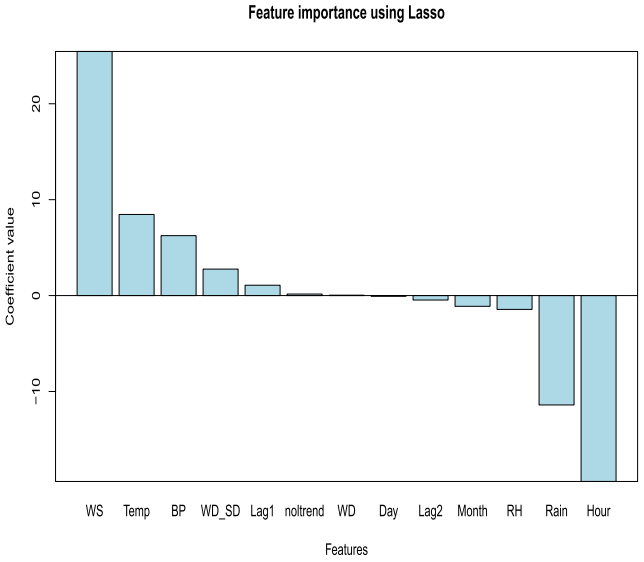

FIGURE 4. Variable selection using Lasso.

- SVR: The radial basis function (RBF) is found to be a better kernel than linear kernel according to the results of Grid Search.
- FFNN: The used FFNN consists of 32 input layer, 32 hidden layers and one output layer. Output layer with no activation function had one neuron. Maximum epochs were set to be 100. The first two layers used ReLU as an activation function. Several experiments for an optimal choice of maximum epochs is done using Grid Search.

Table 2 shows a comparative analysis of the fitted machine learning models, together with the benchmark model. All machine learning models outperform the benchmark model on both the training set and testing set. The M4 (FFNN) model has the least RMSE (83.087) and MAE (51.237), showing that it is the best fitting model according to the summary of the error measures for evaluation of the models on the training set.

TABLE 2. Comparative analysis of the fitted machine learning models.

Evaluation of the models on the training set				
	M1 (PCR)	M2 (LSTM)	M3 (SVR)	M4 (FFNN)
RMSE	130.241	91.884	95.251	83.087
MAE	94.346	55.878	68.705	51.237
Evaluation of the models on the testing set				
	M1 (PCR)	M2 (LSTM)	M3 (SVR)	M4 (FFNN)
RMSE	129.449	94.131	99.302	88.326
MAE	94.554	56.291	72.606	53.719

Table 2 also summarises the error measures for the evaluation of the models on the testing set. From Table 2 model M4 (FFNN) has the least RMSE (88.326) and MAE (53.719) among the three machine learning models. This means that model M4 is the best forecasting model. The results are contradictory to findings by [11] which

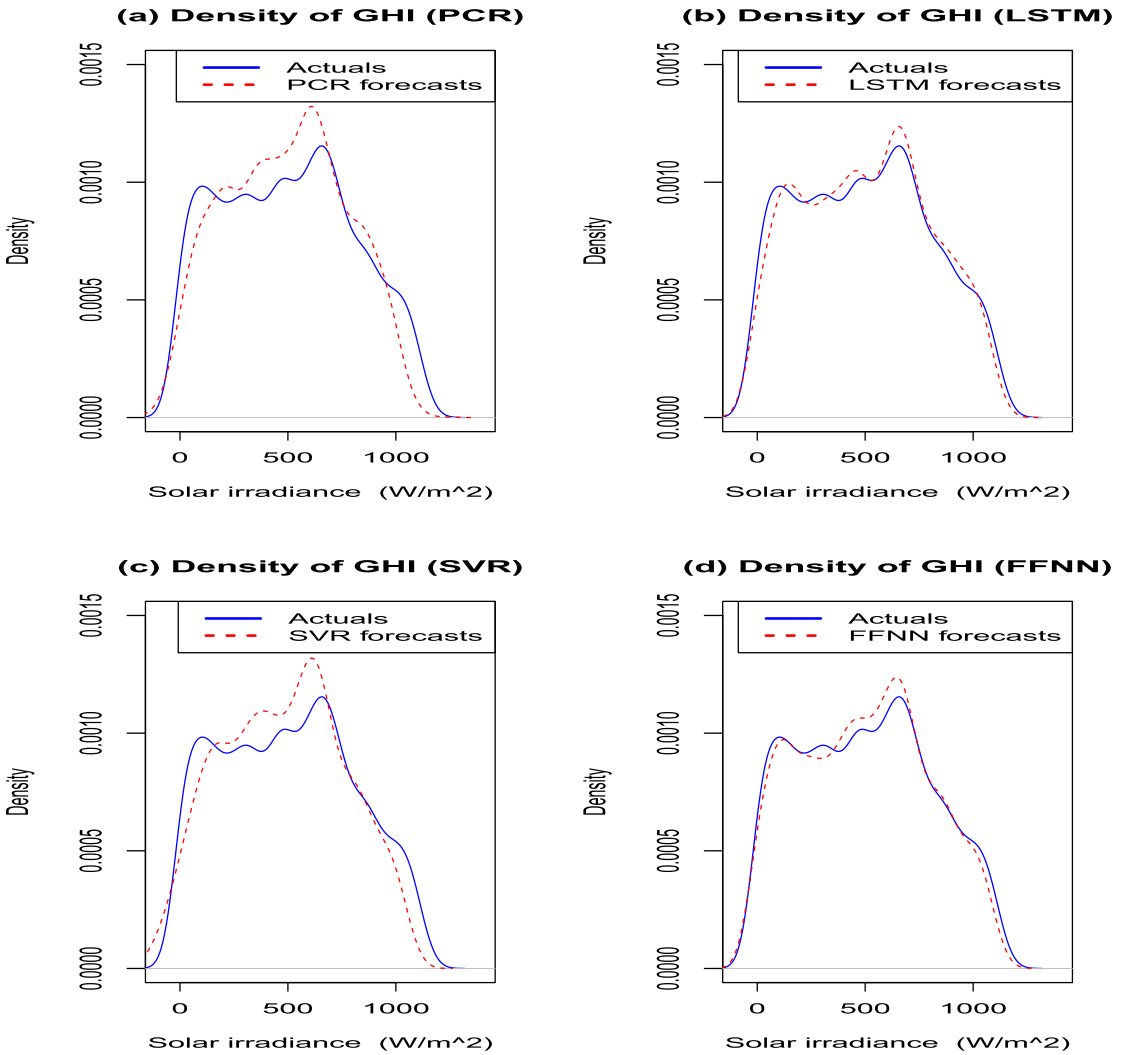

FIGURE 5. **Probability densities of (a) Top left panel: Forecasts (dashed line) using model M1 (PCR) and actual hourly irradiance (solid line) (b) Top right panel: Forecasts (dashed line) using model M2 (LSTM) and actual hourly irradiance (solid line) and (c) Bottom left panel: Forecasts (dashed line) using model M3 (SVR) and actual hourly irradiance (solid line) (d) Bottom right panel: Forecasts (dashed line) using model M4 (FFNN) and actual hourly irradiance (solid line).**

showed that the predictive performance of LSTM outperforms the multilayered FFNN using the backpropagation algorithm.

The probability densities of hourly solar irradiance and the forecasted values for the period January 2018 to December 2018 for machine learning models M2, M3, M4, are shown in Figure 5, including that of M1. The forecasts from the three models appear to have some slight difference with the actual observations. Figure 9 in Appendix A1 shows plots of forecasts on the first day of each season in South Africa. The forecasts are fairly close to the actual observations, especially in spring and summer. Each of the figures shows forecasts for the first day of four seasons in South Africa which are: autumn (March-May), winter (June-August), spring (September-November), and summer (December-February). The selected days are March 1 (autumn), June 1 (winter), September 1 (spring), and December 1 (summer), all in 2018.

C. FORECAST COMBINATION

This section gives results of forecast combination of machine learning models forecasts. Two forecast combination methods are used, which are convex combination and quantile regression averaging.

1) CONVEX COMBINATION

Combining forecasts often leads to better forecast accuracy. The forecasts from different regression-based time-series which are fitted above may be improved by combining them using an R package called 'opera' developed by [39]. The models developed are also referred to as experts. To combine forecasts we find an aggregation of experts and analyse them by looking at the oracles. The opera package computes weights when combining the forecasts. Convex combination method works by computing the sequence of instantaneous losses suffered by the predictions from the experts (models) using loss function. The absolute loss

suffered by the experts are given in Figure 6. From Figure 6 the convex combination model is shown as the best forecasting mode1 followed by model M4 (fFFNN), uniform combination (Uniform), model M2 (fLSTM), and model M3 (fSVR), respectively.

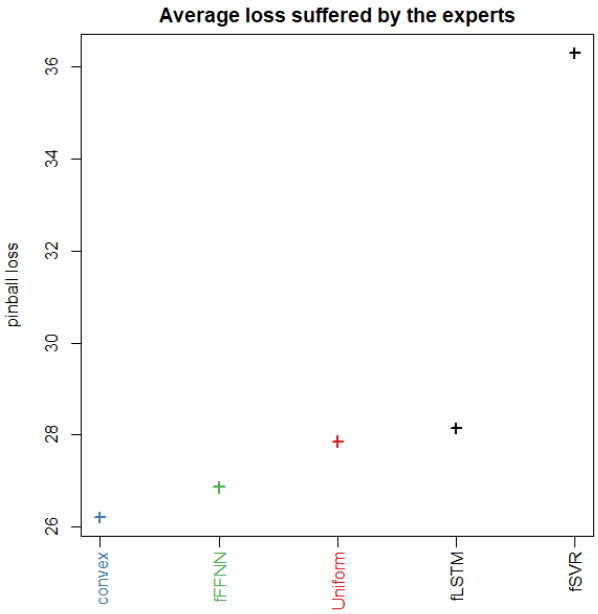

FIGURE 6. Average loss suffered by the models.

2) QUANTILE REGRESSION AVERAGING

Quantile regression averaging (QRA) is another technique normally used to combine forecasts by using forecasts from each model as independent variables. The three models M2 (LSTM), M3 (SVR), and M4 (FFNN) are combined based on QRA, resulting in model M6. Model M6 (QRA) is given by:

$$y_{t,\tau}(\text{QRA}) = \beta_0 + \beta_1 \text{fM2} + \beta_2 \text{fM3} + \beta_3 \text{fM4} + \varepsilon_t, \quad (26)$$

where fM2, fM3, and fM4 represent the forecasts from models M2, M3, and M4, respectively.

Table 3 gives a summary of the accuracy measures for the machine learning models, the M5 (Convex) model, and the M6 (QRA) model. Based on MAE, model M6 is the best forecasting model compared with the convex model and the machine learning models. MAE shows a slight improvement after forecast averaging. The results of absolute loss suffered by the models based on pinball losses show M6 as the best model since it has the smallest average pinball loss value (26.017).

Combining forecasts from individual models improves the accuracy of hourly solar irradiance predictions. Since model M6 is giving the best forecasts we plot its forecasts in Figure 7. Figure 10 in Appendix A1 shows plots of forecasts on the first day of each season in South Africa. The forecasts are fairly close to the actual observations, especially in spring and summer.

TABLE 3. Comparative analysis of the machine learning models, Convex model, and QRA model.

Measure	M2	M3	M4	M5 (Convex)	M6 (QRA)
RMSE	94.131	99.302	88.326	88.732	88.600
MAE	56.291	72.606	53.719	52.405	52.034
Pinball loss	28.146	36.303	26.859	26.202	26.017

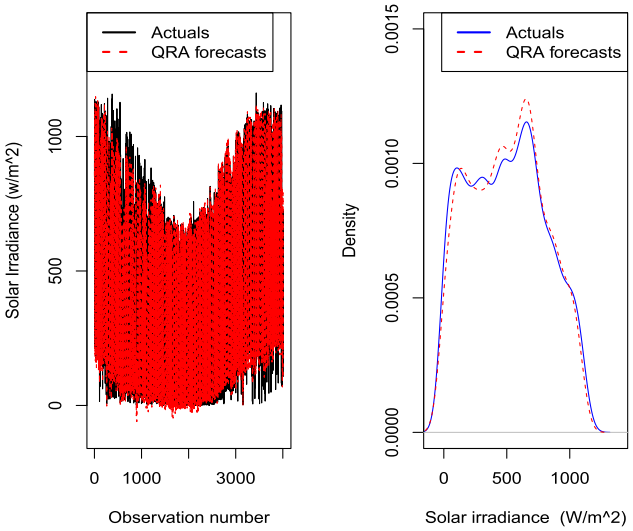

FIGURE 7. Model M6 (QRA) forecasts with density plot.

D. COMPARATIVE ANALYSIS OF THE MODELS

The evaluation of the fitted models based on the empirical prediction intervals (PIs) and forecast error distributions of each model forecasts are presented in this section.

1) EVALUATION OF PREDICTION INTERVALS

To select the best model based on the analysis of the PIWs, we need to calculate the PICPs and PINAWs including a count of the number of predictions below and above the PIs. This is done for various PINC values, which are 90%, 95% and 99%, respectively. A model with a PICP value close to the PINC value is preferred. A fitted model that has better PIWs has the lowest value of PINAW and is one which is preferred over other fitted models. Table 4 shows a comparative evaluation of the models using PI indices for different PINC values. All models have valid PICPs for the three PINC values. Model M2 and M3 have the same lowest PICPs at 90% and 95%, with model M3 having the lowest PICP at 99% compared to all other models. The best model based on PINAW at 90% is M6, M4 at 95%, and M3 at 99%. There is no consistency in the PICPs and PINAWs for different PINC values. However, due to results in Table 3, model M6 (QRA) is selected as the best model.

2) COMBINATION OF INTERVAL LIMITS

Suppose we have $100(1-\alpha)\%$ forecast intervals $[LL_{ij}, UL_{ij}]$, $i = 1, \ldots, m, j = 1, \ldots, k$, where m is the number of forecast point and k is the number of forecasters. In this study we use three basic methods for combining interval forecasts

TABLE 4. Model PINCs comparisons.

PINC	Model	PICP (%)	PINAW (%)	Below LL	Above UL
90%	M2	90.03	25.58	200	200
	M3	90.03	26.59	201	199
	M4	90.05	23.91	199	200
	M5	90.05	23.88	200	199
	M6	90.03	23.67	200	200
95%	M2	95.04	36.17	100	99
	M3	95.04	34.35	100	99
	M4	95.06	32.79	99	99
	M5	95.06	33.08	99	99
	M6	95.06	32.81	99	99
99%	M2	99.05	58.03	19	19
	M3	99.03	53.35	19	20
	M4	99.05	54.96	19	19
	M5	99.05	55.12	19	19
	M6	99.05	54.94	19	19

which are; average method, median method, and envelop method discussed by [30]. Average (Av) limits are given by $LL_{Av} = (1/k)\sum_{j=1}^{k} LL_{ij}$ and $UL_{Av} = (1/k)\sum_{j=1}^{k} UL_{ij}$. Median (Md) limits are given by $LL_{Md} = \{LL_{i1}, \ldots, LL_{ik}\}$ and $UL_{Md} = \{UL_{i1}, \ldots, UL_{ik}\}$. Envelop (En) limits are given by $LL_{En} = \min\{LL_{i1}, \ldots, LL_{ik}\}$ and $UL_{En} = \max\{UL_{i1}, \ldots, UL_{ik}\}$. Table 5 shows a comparative analysis of PIWs after combining the PIWs of individual models based on the average, median, and envelop methods. All combination methods have valid PICPs for the three PINC values with the median method having the lowest PINAWs for the three PINC values. Combining interval limits yield better results since the number of forecasts below and above limits decreases significantly.

TABLE 5. PINCs comparisons based on average, median, and envelop of lower and upper PIs.

PINC	Model	PICP (%)	PINAW (%)	Below LL	Above UL
90%	Average	90.68	24.73	194	180
	Median	90.18	23.89	200	194
	Envelop	93.82	30.03	136	112
95%	Mean	95.26	33.84	99	91
	Median	95.09	33.05	98	99
	Envelop	96.68	39.09	72	61
99%	Mean	99.10	55.28	20	16
	Median	99.10	55.08	19	17
	Envelop	99.33	59.79	14	13

3) RESIDUAL ANALYSIS

Summary statistics of the residuals of the models are given in Table 6. It can be seen that model M4 has the smallest standard deviation which indicates that it has the narrowest error distribution compared to models M3 to M6, this implies that M4 is the best compared to other models, followed by model M6. All the error distributions are skewed to the left since the values of their skewness are all negative. The values for kurtosis are all greater than 3, showing that the distributions are all leptokurtic.

Figure 8 shows box plots of the forecast errors for all the fitted models. From the figure, M4 has the narrowest error distribution compared to models M3 to M6, implying that M4 is the best model compared to other models.

TABLE 6. Model residuals comparisons.

	Mean	Median	Min	Max	St.Dev.	Skew.	Kur.
M2	-6.37	-2.97	-711.64	634.96	93.93	-1.01	8.44
M3	19.05	33.09	-637.20	600.87	97.47	-1.26	5.44
M4	3.05	8.22	-662.40	624.52	88.29	-1.18	8.83
M5	-0.44	4.17	-680.11	628.38	88.74	-1.16	9.15
M6	-5.45	0.00	-681.13	619.63	88.44	-1.21	9.19

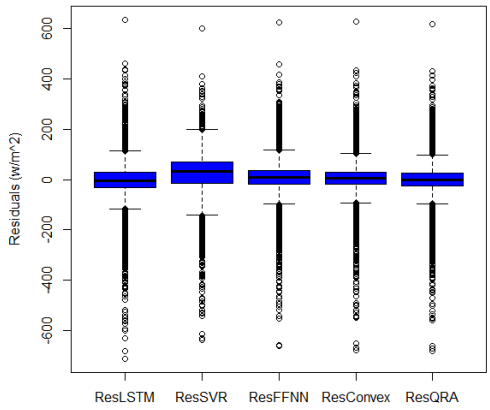

FIGURE 8. Box plots of residuals from models M2 (ResLSTM), M3 (ResSVR), M4 (ResFFNN), M5 (ResConvex), and M6 (ResQRA).

4) PERCENTAGE IMPROVEMENT

Table 7 gives the percentage of improvement rates of the best model over other models. The percentage improvement of M6 over M2, M3, and M4 are found to be 7.56%, 28.33%, and 1.30%, respectively. This means M4 (FFNN) is the second-best model after M6 (QRA) and their predictive performances are very close to one another.

TABLE 7. Percentage improvement rates.

Models	MAE (best)	MAE (other)	% improvement
LQRA & LSTM	52.034	56.291	7.56%
LQRA & SVR	52.034	72.606	28.33%
LQRA & FFNN	52.034	53.719	1.30%

5) EMPIRICAL RESULTS ON EVALUATION AND COMPARISON OF MULTIPLE FORECASTS

We present the results from the DM and GW tests. In all cases the null hypothesis is: H_0 : forecasts are equally accurate.

The three best models considered are M4 (feed forward neural network), M5 (convex combination based on the pinball loss) and M6 (quantile regression averaging). The null hypotheses are:

1) H_0 : forecasts from M4 and M5 are equally accurate.
2) H_0 : forecasts from M4 and M6 are equally accurate.
3) H_0 : forecasts from M5 and M6 are equally accurate.

Based on the results from Table 8 the DM tests show that in all cases the forecasts are not equally accurate. For the GW test we shall use the following to denote dominance, $M_i < M_j, \forall i \neq j$ to mean that M_j dominates M_i. From Table 8, M5 < M4, M6 < M4, M5 < M6 implies that M5

TABLE 8. Model comparisons.

Diebold-Mariano test			
Null hypothesis	Test statistic	p-value	Result
M4 = M5	6.5149	0.0000	Not equally accurate
M4 = M6	8.0586	0.0000	Not equally accurate
M5 = M6	3.9311	0.0001	Not equally accurate
Giacommini-White test			
Null hypothesis	Test statistic		Result
M4 = M5	6.9489		Sign of mean loss is (-). M4 dominates M5
M4 = M6	6.8686		Sign of mean loss is (-). M4 dominates M6
M5 = M6	2.2477		Sign of mean loss is (+). M6 dominates M5
Testing for unbiasedness (Mincer-Zarnowitz test)			
Model	Test statistic	p-value	Result: Unbiased
M4	10.5381	0.0145	No
M5	4.3821	0.0160	No
M6	8.8210	0.4235	Yes
Testing for unbiasedness using residuals ($\varepsilon_t = y_t - f_t$)			
Model	Number of $y_t - f_t > 0$	Number of $y_t - f_t < 0$	Number of $y_t - f_t = 0$
M4	2379 (59.30%)	1633 (40.70%)	0
M5	2240 (55.80%)	1772 (44.20%)	0
M6	2004 (49.95%)	2005 (50.05%)	3

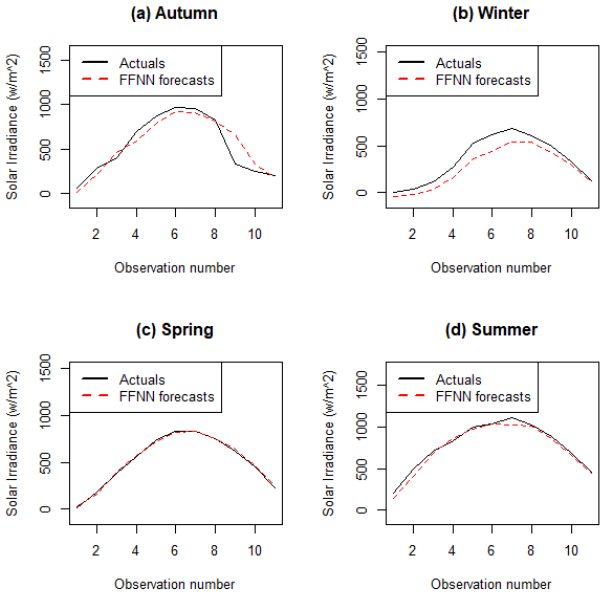

FIGURE 9. Graphical plot of forecasts (dashed line) using model M4 (FFNN) and actual hourly irradiance (solid line) for the first day of season in (a) **Top left panel:** Autumn (1 March 2018) and (b) **Top right panel:** Winter (1 June 2018) (c) **Bottom left panel:** Spring (1 September 2019) and (d) **Bottom right panel:** Summer (1 December 2019).

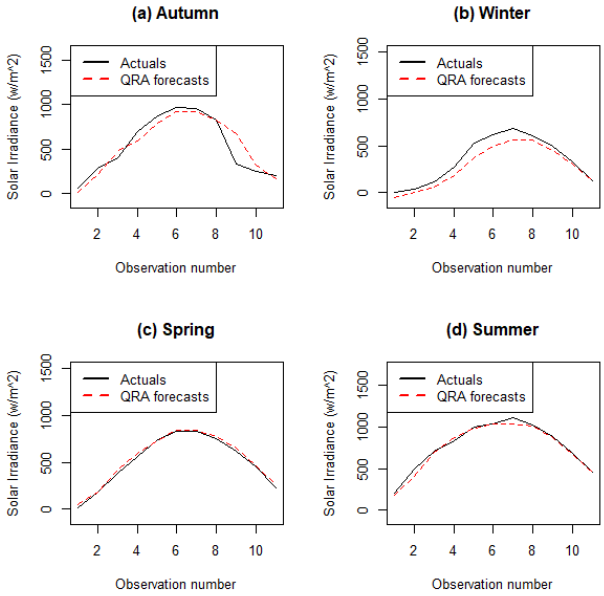

FIGURE 10. Graphical plot of forecasts (dashed line) using model M6 (QRA) and actual hourly irradiance (solid line) for the first day of season in (a) **Top left panel:** Autumn (1 March 2018) and (b) **Top right panel:** Winter (1 June 2018) (c) **Bottom left panel:** Spring (1 September 2019) and (d) **Bottom right panel:** Summer (1 December 2019).

< M6 < M4. Therefore M4 has the highest predictive power since it dominates the other two models.

On testing for unbiasedness in the forecast errors we use the Mincer-Zarnowitz test based on the HAC robust VCV approach. The null hypothesis is that the forecasts are unbiased. From Table 8 the null hypothesis is rejected for models M4 and M5 and we fail to reject for model M6. This means that forecasts from M4 and M5 are biased and those from M6 are unbiased all at the 5% level of significance. Based on these results we then compute the residuals and count the number of under-predictions and over-predictions. Positive residuals ($\varepsilon_t = y_t - f_t > 0$) means we have under-predictions and negative residuals ($\varepsilon_t = y_t - f_t < 0$) implies over-predictions. Based on the results approximately 59.30%, 55.80% and 49.95% of the forecasts from models M4, M5 and M6, respectively are under-predictions. This implies that forecasts from model M6 are unbiased. These results are consistent with the Mincer-Zarnowitz test.

E. DISCUSSION

From the comparative analysis based on MAE and average pinball losses, M6 (QRA) is the best fitting model and can be

used for predicting hourly solar irradiance. From the PIWs analysis at 95% level of confidence, M2 (LSTM) has the narrowest PI compared to models M3 to M6, implying that M2 is the best compared to other models. Further analysis of the PIWs based on the PICPs and PINAWs including a count of the number of predictions below and above the PIs shows no consistency. The best model based on PINAW at 90% is M6 (QRA), M4 (FFNN) at 95%, and M3 (SVR) at 99%. There is no consistency in the PICPs and PINAWs for different PINC values. The residual analysis shows M4 (FFNN) as the best model with narrowest error distribution compared to other models, followed by model M6.

Table 5 shows a comparative analysis of PIWs after combining the PIWs of individual models based on the average, median, and envelop methods. All combination methods have valid PICPs for the three PINC values with the median method having the lowest PINAWs for the three PINC values. Combining interval limits yield better results since the number of forecasts below and above limits decreases significantly.

Empirical results on the evaluation and comparison of multiple forecasts given in Table 8 show that the forecasts are not equally accurate in all pairwise comparisons using the Diebold-Mariano test. Using the Giacomini-White test model M4 (FFNN) dominates all the other two, models M5 (convex) and M6 (QRA) meaning that it has the highest predictive power, followed by M6 and lastly M5. On testing for unbiasedness in the forecasts, it is seen that M4 and M5 produce forecasts which are biased and most of the forecasts are under-predictions. Model M6 produces unbiased forecasts. Based on all these analyses of the forecast accuracy measures and the statistical tests, M6 is considered as the best forecasting model.

TABLE 9. Forecast accuracy measures root mean square error (RMSE), mean absolute error (MAE) for the forecasts of January to December 2018.

Month	RMSE	MAE
January	103.389	69.195
February	124.483	91.413
March	103.091	75.659
April	92.043	58.864
May	59.154	33.180
June	34.847	20.039
July	44.066	24.256
August	40.814	23.729
September	53.305	30.064
October	98.832	51.750
November	118.346	71.924
December	121.898	77.909

Table 9 in Appendix A1 summarises the error measures for out of sample evaluation using model M6 for the month January to December 2018. The accuracy measures are significantly lower from April to August. Figures 11 and 12 show hourly solar irradiance superimposed with forecasts together with their respective densities. The forecasts follow actual hourly solar irradiance very well, especially in June and September 2018.

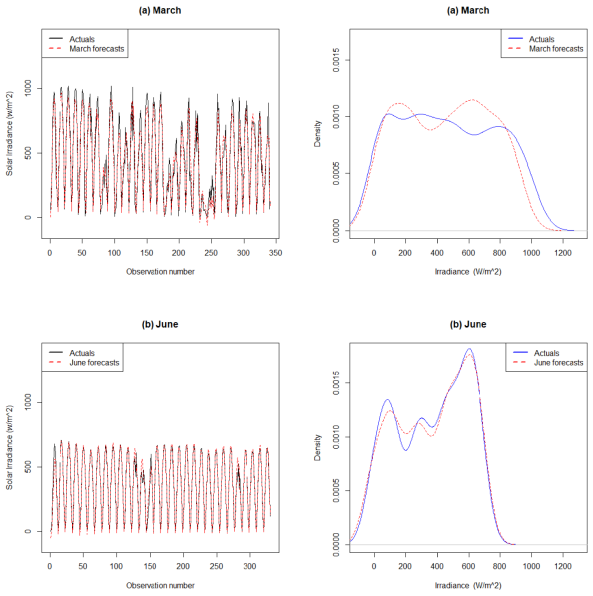

FIGURE 11. Hourly irradiance with forecasts for March and June 2018.

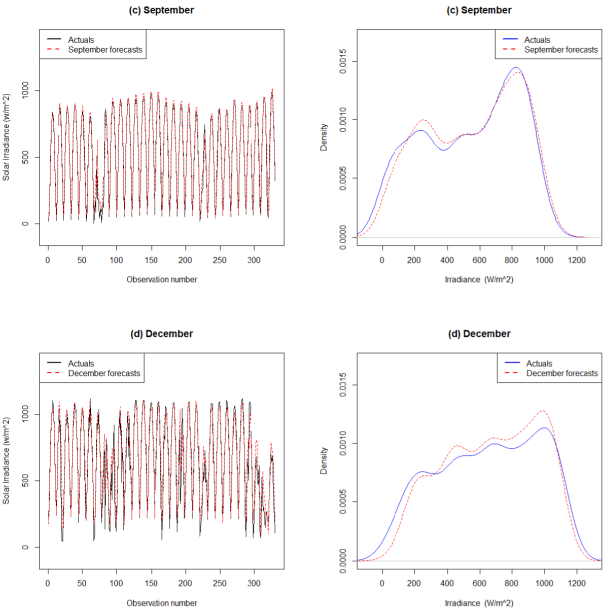

FIGURE 12. Hourly irradiance with forecasts for September and December 2018.

IV. CONCLUSION

The paper presented an application of modelling hourly solar irradiance using long short-term memory (LSTM) networks, support vector regression (SVR) and feed forward neural networks (FFNN) models. Variable selection was done using the least absolute shrinkage and selection operator (Lasso). According to findings, among the three fitted machine learning models, the FFNN model produced the best forecast accuracy based on the MAE and RMSE. Later, the forecasts from the machine learning models were

combined using the convex combination method and quantile regression averaging (QRA). Based on all the forecast accuracy measures used in this study including the statistical tests, the QRA model was found to be the best forecast combination method, and also the best forecasting model compared with the machine learning models. The median method for combining interval limits gave the best results on PIWs analysis.

The information derived from this research study on forecasting solar irradiance is important to electricity utility decision-makers in South Africa in balancing demand and supply of electricity in an effective way which favours future economic prosperity and environmental security. Our results in forecast combination are consistent with what is discussed in literature ([16], [40], among others). The main contribution of this paper is in providing adequate and detailed evaluation metrics including statistical tests in forecasting solar irradiance. Future work could focus on including more radiometric stations based in South Africa and also consider different weather conditions at coastal and inland regions by doing spatial analysis.

ACKNOWLEDGMENT

Opinions expressed and conclusions arrived at, are those of the authors and are not necessarily to be attributed to the NEPTTP.

AUTHOR CONTRIBUTIONS

Tendani Mutavhatsindi: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Project administration; Resources; Software; Validation; Visualization; Roles/Writing - original draft; Writing - review & editing.

Caston Sigauke: Conceptualization; Formal analysis; Investigation; Methodology; Resources; Software; Supervision; Validation; Visualization; Roles/Writing - original draft; Writing - review & editing.

Rendani Mbuvha: Conceptualization; Formal analysis; Investigation; Methodology; Resources; Software; Supervision; Validation; Visualization; Roles/Writing - original draft; Writing - review & editing.

CONFLICT OF INTEREST

On behalf of all authors, the corresponding author states that there is no conflict of interest.

APPENDIX A1: FORECASTS FOR THE MONTHS JANUARY TO DECEMBER 2018

See Table 9 and Figs 9-12.

REFERENCES

[1] A. Zendehboudi, Z. Huan, and C. C. Enweremadu, "Evaluation of global solar radiation using multiple weather parameters as predictors for South Africa provinces," *J. Cleaner Prod.*, vol. 199, pp. 272–285, Oct. 2018.

[2] S. Sun, S. Wang, G. Zhang, and J. Zheng, "A decomposition-clustering-ensemble learning approach for solar radiation forecasting," *Sol. Energy*, vol. 163, pp. 189–199, Mar. 2018.

[3] K. Mohammadi, S. Shamshirband, A. S. Danesh, M. S. Abdullah, and M. Zamani, "Temperature-based estimation of global solar radiation using soft computing methodologies," *Theor. Appl. Climatol.*, vol. 125, nos. 1–2, 2016, pp. 101–112.

[4] E. Zhandire, "Solar resource classification in South Africa using a new index," *J. Energy Southern Afr.*, vol. 28, no. 2, 2016, pp. 61–70.

[5] L. Cristaldi, G. Leone, and R. Ottoboni, "A hybrid approach for solar radiation and photovoltaic power short-term forecast," in *Proc. IEEE Int. Instrum. Meas. Technol. Conf.*, May 2017, pp. 1–6.

[6] B. Y. H. Liu and R. C. Jordan, "The interrelationship and characteristic distribution of direct, diffuse and total solar radiation," *Sol. Energy*, vol. 4, no. 3, pp. 1–19, Jul. 1960.

[7] X. Yang, F. Jiang, and H. Liu, "Short-term solar radiation prediction based on SVM with similar data," *Solar Energy*, vol. 4, no. 3, pp. 1–19, 2013.

[8] M. Ozgoren, M. Bilgili, and B. Sahin, "Estimation of global solar radiation using ANN over turkey," *Expert Syst. Appl.*, vol. 39, no. 5, pp. 5043–5051, Apr. 2012.

[9] A. Rezrazi, S. Hanini, and M. Laidi, "An optimisation methodology of artificial neural network models for predicting solar radiation: A case study," *Theor. Appl. Climatol.*, vol. 123, nos. 3–4, pp. 769–783, Feb. 2016.

[10] O. Nait Mensour, S. Bouaddi, B. Abnay, B. Hlimi, and A. Ihlal, "Mapping and estimation of monthly global solar irradiation in different zones in souss-massa area, morocco, using artificial neural networks," *Int. J. Photoenergy*, vol. 2017, pp. 1–19, May 2017.

[11] X. Qing and Y. Niu, "Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM," *Energy*, vol. 148, pp. 461–468, Apr. 2018.

[12] Z. Dong, D. Yang, T. Reindl, and W. M. Walsh, "A novel hybrid approach based on self-organizing maps, support vector regression and particle swarm optimization to forecast solar irradiance," *Energy*, vol. 82, pp. 570–577, Mar. 2015.

[13] R. C. Deo, X. Wen, and F. Qi, "A wavelet-coupled support vector machine model for forecasting global incident solar radiation using limited meteorological dataset," *Appl. Energy*, vol. 168, pp. 568–593, Apr. 2016.

[14] J. Fan, X. Wang, L. Wu, H. Zhou, F. Zhang, X. Yu, X. Lu, and Y. Xiang, "Comparison of support vector machine and extreme gradient boosting for predicting daily global solar radiation using temperature and precipitation in humid subtropical climates: A case study in China," *Energy Convers. Manage.*, vol. 164, pp. 102–111, May 2018.

[15] W. VanDeventer, E. Jamei, G. S. Thirunavukkarasu, M. Seyedmahmoudian, T. K. Soon, B. Horan, S. Mekhilef, and A. Stojcevski, "Short-term PV power forecasting using hybrid GASVM technique," *Renew. Energy*, vol. 140, pp. 367–379, Sep. 2019.

[16] M. Guermoui, F. Melgani, K. Gairaa, and M. L. Mekhalfi, "A comprehensive review of hybrid models for solar radiation forecasting," *J. Cleaner Prod.*, vol. 258, Jun. 2020, Art. no. 120357.

[17] C. Paoli, C. Voyant, M. Muselli, and M.-L. Nivet, "Forecasting of preprocessed daily solar radiation time series using neural networks," *Sol. Energy*, vol. 84, no. 12, pp. 2146–2160, Dec. 2010.

[18] K. Benmouiza and A. Cheknane, "Forecasting hourly global solar radiation using hybrid k-means and nonlinear autoregressive neural network models," *Energy Convers. Manage.*, vol. 75, pp. 561–569, Nov. 2013.

[19] A. Gensler, J. Henze, B. Sick, and N. Raabe, "Deep Learning for solar power forecasting—An approach using auto encoder LSTM neural networks" in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Dec. 2016, pp. 2858–2865.

[20] S. Sreekumar, K. C. Sharma, and R. Bhakar, "Optimized support vector regression models for short term solar radiation forecasting in smart environment," in *Proc. Region Conf.*, 2016, pp. 1929–1932.

[21] A. Adeala, Z. Huan, and C. Enweremadu, "Evaluation of global solar radiation using multiple weather parameters as predictors for South Africa provinces," *Thermal Sci.*, vol. 19, no. 2, pp. 495–509, 2015.

[22] P. Mpfumali, C. Sigauke, A. Bere, and S. Mulaudzi, "Day ahead hourly global horizontal irradiance forecasting—Application to South African data," *Energies*, vol. 12, no. 18, p. 3569, 2019.

[23] E. Ranganai and C. Sigauke, "Capturing long-range dependence and harmonic phenomena in 24-hour solar irradiance forecasting: A quantile regression robustification via forecasts combination approach," *IEEE Access*, vol. 8, pp. 172204–172218, 2020, doi: 10.1109/ACCESS.2020.3024661.

[24] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *Bull. Math. Biophys.*, vol. 5, no. 4, pp. 115–133, Dec. 1943.

[25] F. Rosenblatt, "The perceptron: A probabilistic model for information storage and organization in the brain," *Psychol. Rev.*, vol. 65, no. 6, p. 386, 1958.
[26] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
[27] H. Drucker, C. J. Burges, L. Kaufman, A. J. Smola, and V. Vapnik, "Support vector regression machines," in *Proc. Adv. Neural Inf. Process. Syst.*, 1997, pp. 155–161.
[28] R. X. Liu, J. Kuang, Q. Gong, and X. L. Hou, "Principal component regression analysis with spss," *Comput. Methods Prog. Biomed.*, vol. 71, no. 2, pp. 141–147, Jun. 2003.
[29] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. Roy. Stat. Soc., Ser. B Methodol.*, vol. 58, no. 1, pp. 267–288, Jan. 1996.
[30] A. Gaba, I. Tsetlin, and R. L. Winkler, "Combining interval forecasts," *Decis. Anal.*, vol. 14, no. 1, pp. 1–20, Mar. 2017.
[31] K. Maciejowska, J. Nowotarski, and R. Weron, "Probabilistic forecasting of electricity spot prices using factor quantile regression averaging," *Int. J. Forecasting*, vol. 32, no. 3, pp. 957–965, Jul. 2016.
[32] P. Gaillard and Y. Goude, "Forecasting electricity consumption by aggregating experts; How to design a good set of experts," in *Modeling and Stochastic Learning for Forecasting in High Dimensions* (Lecture Notes in Statistics), vol. 217. Cham, Switzerland: Springer, 2015, pp. 95–115.
[33] X. Sun, Z. Wang, and J. Hu, "Prediction interval construction for byproduct gas flow forecasting using optimized twin extreme learning machine," *Math. Problems Eng.*, vol. 2017, pp. 1–12, Dec. 2017.
[34] F. X. Diebold and R. Mariano, "Comparing predictive accuracy," *J. Bus. Econ. Statist.*, vol. 13, no. 1, pp. 253–265, 1995.
[35] U. Triacca. (2018). *Comparing Predictive Accuracy of Two Forecasts*. [Online]. Available: http://www.phdeconomics.sssup.it/documents/ Lesson19.pdf
[36] A. Tarassow and S. Schreiber. (2020). *FEP–The Forecast Evaluation Package for Gretl*. [Online]. Available: http://ricardo.ecn.wfu.edu/gretl/cgi-bin/current_fnfiles/unzipped/FEP.pdf
[37] J. Lago, G. Marcjasz, B. De Schutter, and R. Weron, "Forecasting day-ahead electricity prices: A review of state-of-the-art algorithms, best practices and an open-access benchmark," 2020, *arXiv:2008.08004*. [Online]. Available: http://arxiv.org/abs/2008.08004
[38] M. J. Brooks, S. Du Clou, W. L. Van Niekerk, P. Gauché, C. Leonard, M. J. Mouzouris, R. Meyer, N. Van der Westhuizen, E. E. Van Dyk, and F. J. Vorster, "SAURAN: A new resource for solar radiometric data in southern africa," *J. Energy Southern Afr.*, vol. 26, no. 1, pp. 2–10, Apr. 2017.
[39] P. Gaillard and Y. Goude. (2016). *Opera: Online Prediction by Expert Aggregation*. [Online]. Available: https://CRAN.R-project.org/package=opera
[40] M. N. Akhter, S. Mekhilef, H. Mokhlis, and N. Mohamed Shah, "Review on forecasting of photovoltaic power generation based on machine learning and Metaheuristic techniques," *IET Renew. Power Gener.*, vol. 13, no. 7, pp. 1009–1023, May 2019.

TENDANI MUTAVHATSINDI was born in Thohoyandou, Limpopo, South Africa, in 1994. He received the B.Sc. degree in statistics and mathematics from the University of Venda, South Africa, and the joint M.Sc. degree in statistics (e-Science) from the University of Venda, which is provided in collaboration with the University of the Witwatersrand. His research interests include machine learning, time series forecasting, and renewable energy.

CASTON SIGAUKE (Member, IEEE) received the M.Sc. degree in operations research from the National University of Science and Technology, Zimbabwe, and the Ph.D. degree in statistics from the University of the Free State, South Africa. He is currently a Senior Lecturer with the Department of Statistics, University of Venda, South Africa. His fields of expertise include forecasting and time series, statistics of extremes, and statistical learning and modeling. His research is on probabilistic load and renewable energy (solar and wind) forecasting, including optimization of grid integration of renewable energies. He is a member of the International Statistical Institute (ISI), International Institute of Forecasters (IIF), Operations Research Society of South Africa (ORSSA), and South African Statistical Association (SASA). Since 2013, he has been a Chartered Statistician (13ChM012) of the Institute of Certificated and Chartered Statisticians of South Africa (ICCSSA).

RENDANI MBUVHA was born in Venda, Limpopo, South Africa. He received the B.Sc. degree (Hons.) in actuarial science and statistics from the University of Cape Town and the M.Sc. degree in machine learning from the KTH Royal Institute of Technology, Sweden. He is currently a Lecturer of Statistics and Actuarial Science with the University of the Witwatersrand. He is a qualified Actuary and a holder of the Chartered Enterprise Risk Actuary designation. He was a recipient of the Google Ph.D. Fellowship, which supports his current Ph.D. research at the University of Johannesburg.