

Article

Short-Term Solar Power Forecasting Using Genetic Algorithms: An Application Using South African Data

Mamphaga Ratshilengo [†], Caston Sigauke ^{*,†} and Alphonse Bere 

Department of Statistics, University of Venda, Private Bag X5050, Thohoyandou 0950, South Africa; 14003914@mvula.univen.ac.za (M.R.); alphonse.bere@univen.ac.za (A.B.)

* Correspondence: caston.sigauke@univen.ac.za; Tel.: +27-15-962-8135

† These authors contributed equally to this work.

Abstract: Renewable energy forecasts are critical to renewable energy grids and backup plans, operational plans, and short-term power purchases. This paper focused on short-term forecasting of high-frequency global horizontal irradiance data from one of South Africa's radiometric stations. The aim of the study was to compare the predictive performance of the genetic algorithm and recurrent neural network models with the K-nearest neighbour model, which was used as the benchmark model. Empirical results from the study showed that the genetic algorithm model has the best conditional predictive ability compared to the other two models, making this study a useful tool for decision-makers and system operators in power utility companies. To the best of our knowledge this is the first study which compares the genetic algorithm, the K-nearest neighbour method, and recurrent neural networks in short-term forecasting of global horizontal irradiance data from South Africa.

Keywords: Giacommini–White test; global horizontal irradiance; genetic algorithm; Lasso; machine learning; Murphy diagram



Citation: Ratshilengo, M.; Sigauke, C.; Bere, A. Short-Term Solar Power Forecasting Using Genetic Algorithms: An Application Using South African Data. *Appl. Sci.* **2021**, *11*, 4214. <https://doi.org/10.3390/app11094214>

Academic Editor: Yosoon Choi

Received: 18 April 2021

Accepted: 28 April 2021

Published: 6 May 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

1.1. Background

Renewable energy sources (RES) are rising rapidly in different countries, powered by the cost reduction of wind turbines and photovoltaic (PV) panels [1]. They are increasingly becoming the future's dominant energy source, but harvesting them requires an understanding of the mechanisms of their volatility and the ability to predict various environmental processes over a scale range [2]. Such understanding of the environment is the key to renewable energy processing, in particular solar and wind power. When the global population and industrialisation continue to rise exponentially, the fossil fuels used to generate electricity are also rapidly depleting [3]. The ongoing overuse of fossil fuels in the production of electricity continues to inflict environmental problems such as global warming. RESs are environmentally friendly and inexpensive [4].

Solar energy is one of the most critical sources of renewable energy that can contribute to addressing current environmental and energy issues on the electrical grid [5]. It is widely regarded to be the world's best-growing energy industry [6]. Nonetheless, for proper and effective management of the electrical grid, the integration of solar energy onto the electrical grid requires detailed forecasting [7]. Policy-makers for power utilities face the challenge of balancing demand and electricity supply in a cost-effective way which also benefits future economic growth and environmental sustainability. Solar irradiance forecasting is critical for backup programming, strategic planning, short-term power procurement, shifting other energy sources, backup utilisation planning, and peak load demand [8]. It is relevant for various other activities along with photovoltaics (PVs), agriculture, medical study applications, and desalination by seawater [9]. The steadily increasing integration of solar systems around the world is an indication of the need for accurate knowledge

and understanding of solar resources in the design of solar electric grids. Solar irradiance studies are therefore of great importance for the optimal project and power forecasting of PV grid-connected plants.

A modelling framework to determine the statistical distribution of the daily total solar radiation at different locations is discussed in [10]. Researchers have since been paying attention to solar irradiance, using many modelling techniques. Solar irradiance forecasts consist primarily of physical models and data driven models [11,12]. Physical models are based on numerical weather forecasts. The data driven models consist of statistical and machine learning models.

2. An Overview of the Literature on Solar Forecasting

Current research on solar irradiance forecasting is focussing to a very large extent on the use of machine learning techniques such as artificial neural networks (ANNs) [13] and support vector machines (SVMs) [14]. ANNs can map relationships between various variables given there is available information to learn from [15].

In a study by [16], hybrid neural networks and metaheuristic algorithms are applied to short-term forecasting of solar energy. Cadenas and Rivera [17] used ANN for short-term wind forecasting in La Venta, Oaxaca, and Mexico. Capizzi et al. [18] proposed the application of a recurrent neural network (RNN), which was based on the control strategy for storage of the battery energy in systems generation with intermittent renewable energy sources. Tsai et al. [19] applied a neural network to the short-term load and wind power forecasting based on prediction intervals.

Yang et al. [20] suggested a combined autoregressive and dynamic system (CARDS) for the 1 hour ahead global solar irradiance prediction. The suggested approach improves the global solar irradiance forecast accuracy by 30% compared to ANNs models. Andrade and Bessa [1] used a system of numerical weather predictions (NWP) to conduct their research on enhancing renewable energy forecasting. Throughout their report, they suggested a feature engineering approach to obtain further details for wind and solar energy forecasts from an NWP's grid. They considered that adequate processing of direct NWP data elements can boost the forecasting capability and that the renewable energy forecaster could invest time in this knowledge discovery process. The use of ANN as a new strategy for assessing future energy consumption levels in SA is suggested by [21]. In their study, particle swarm optimisation (PSO) was used to train ANNs. Estimates of the annual electricity demand were determined per scenario, and it was noted that the proposed ANN approach attains a comparatively better energy demand forecast within acceptable errors. Sigauke [22] discussed the implementation of generalised additive models (GAMs) to the prediction of medium-term demand for electricity utilising data from SA. In the study, the variable selection was carried out using Lasso (least absolute shrinkage and selection operator) via hierarchical interactions. Marwala and Twala [23] used autoregressive moving average (ARMA), Neural networks (NNs), and Neuro-Fuzzy (NF) systems to model and forecast non-linear processes. They found that NF has a better ability than ARMA and NNs to model the system, while NNs are better than ARMA.

King et al. [24] proposed a genetic algorithm (GA) framework using 'negative load' and the 'inclusive' strategy. They emphasized that, as wind power cost for the 'negative load' approach is presumed to be zero, the results do not show the actual cost. The actual cost of wind power should therefore be in the 'zero-fee' strategy. They also discovered that the use of wind power in the economic load dispatch (ELD) computation affects cost, load volatility pictured by some other plants, and reserve necessity because of the expected error. Mellit and Shaari [25] proposed an RNN-based approach for forecasting the regular production of electricity from a PV power system (PVPS). In the study, three RNN models were compared where the proposed third RNN-II provided more reliable daily forecasts for generating electricity compared with the other proposed MLP and RNN-I models. A combination approach technique for short-term forecasting of solar irradiance and photovoltaic (PV) power is presented in [18]. The solar irradiance prediction was

carried out with the help of physical techniques named clear-sky techniques. Short-term PV forecasting is then carried out with the help of an auto-associative kernel regression (AAKR) strategy that is typically implemented for the detection of faults. The findings of the proposed model showed increased efficacy in forecasting solar irradiance.

A GA approach is used by [26] to achieve optimal wind turbine location for higher production capacity when limiting the number of turbines fixed and land used for every wind farm. They considered two cases, which are: non-uniform wind with variable direction, and the uniform wind with direction. Khan and Byun [27] presented an application of GA rooted optimised feature engineering and machine learning to the forecasting of energy consumption. The work focussed on comparing various forecasting methods including XGBoost, support vector regression, and K-nearest neighbour regressor algorithms. The work looked at various meteorological features which are temperature, wind speed, rain, humidity, and time lags. In this work, it is concluded that combinations of prediction methods yield good results.

Table 1 presents a summary of some previous studies on modelling and forecasting solar irradiance using genetic algorithms, recurrent neural networks and K-nearest neighbour.

Table 1. Summary of some previous studies on modelling and forecasting solar irradiance using GA, RNN, and KNN.

Authors	Data	Models	Main Findings
Gardes and Girard [28]	France hourly rainfall data from 1993 to 2000	Nearest neighbour model.	Empirical results show that the nearest neighbour hill estimator gives the same weight to all largest observations.
VanDeventer et al. [29]	Hourly solar photovoltaic data	A genetic algorithm-support vector machine (GASVM) model.	Based on the RMSE and MAPE, GASVM had greater predictive ability compared to SVM.
Al-Iahham et al. [30]	Sky image data from 2004 to 2020	KNN and random forest models.	The results show that KNN achieves good computational complexity reduced by 30% of the state-of-the-art algorithms.
Pattanaik et al. [31]	Solar data from 2000 to 2017	Genetic algorithm and Artificial neural network models.	The results show that GA forecasting is much more convenient and also produces accurate results.
Benamrou et al. [32]	Hourly GHI data from 2015 to 2017	Xgboost, LSTM RNN and random forest network models.	Deep LSTM is found to be the best model in forecasting one hour ahead GHI.
Brahma and Wadhvani [33]	Daily solar irradiance data from 1983 to 2019	LSTM, Bidirectional LSTM, GRU, XGBoost, and CNN LSTM.	Results show that the forecasting tasks of shorter horizons give better accuracy while longer horizons require more complex models.

The one-hour and regular wind power generation forecasting using Bayesian neural networks is discussed in [34]. Results from this study show that the NNs based on Bayesian neural networks exhibit comparable predictive performance to maximum likelihood neural networks for both one hour and a day ahead forecasts. Mpfumali et al. [35] used data from Tellerie radiometric station, South Africa for the period between August 2009 to April 2010 for a probabilistic prediction of 24 h ahead of global horizontal irradiance (GHI). They used different techniques in forecasting, including quantile regression averaging (QRA) and machine learning techniques, and it is concluded that QRA gives higher accuracy than the machine learning techniques based on the probabilistic error measures. In a study of the classification of solar energy resources in SA using new index measured solar radiation resources, [5] used eight stations and five distinct groups. Results from the study showed that the solar utility index (SUI) has higher solar resources, which discriminate groups compared to other indices such as the probability of persistence (POPD) and the fractal dimension. Apart from typical interplanetary radiation and hours of

sunshine [36], presented a report on the prediction of global solar radiation using multiple linear regression models. The study was undertaken in all nine provinces of South Africa. The study revealed that the use of weather parameters for some locations increases the accuracy and efficiency of solar radiation models.

Research Highlights

This current study differs from previous studies in that it compares genetic algorithm (GA), recurrent neural networks (RNN), and the K-nearest neighbours (KNN) models in short-term forecasting of high-frequency global horizontal irradiance using South African data. To the best of our knowledge, this is the first paper to carry out such a comparative study of the models using South African data. South Africa receives sunlight throughout the year, so the natural advantage of South Africa is that it is included in the world's highest in renewable energy. The global average annual 24-h solar radiation for South Africa has been about 220 W/m^2 , 150 W/m^2 for parts of the US, and about 100 W/m^2 towards Europe and the United Kingdom [37].

The following is a summary of the highlights and findings of this study.

1. Based on the RMSE and rRMSE, GA was found to be the best model. However, based on MAE and rMAE, RNN was the best model.
2. From the Diebold–Mariano tests, the null hypothesis that the forecast accuracy between a pair of forecasts from the two methods is the same was rejected for all the three pairs.
3. Based on the Murphy diagrams, the GA dominated both RNN and KNN, meaning that it provides the greatest predictive ability.
4. Based on the Giacomini–White tests, GA was found to have the best conditional predictive ability compared to the other two models.
5. The results from this paper yield improved results from the previous papers.
6. To the best of our knowledge, this is the first paper to compare the genetic algorithm, the K-nearest neighbour method, and recurrent neural networks in short-term forecasting of global horizontal irradiance data from South Africa.

The rest of the following paper is as follows: Section 3 is the materials and methods used in this paper. The empirical results and discussion are in Section 4. The conclusion is provided in Section 5.

3. Models

3.1. Genetic Algorithm

The GAs follow the theory of Darwin of the survival of the fittest and links the natural selection and genetics experiment with random operators to form searching mechanisms for improved results [38]. It was mentioned in the previous chapter that, in the late 1960s, Holland proposed GAs, and Goldberg (1989) implemented the algorithm for the first time to address engineering optimisation problems. Goldberg (1989) later illustrated the utility of GAs for structural optimisation by solving the classic ten-bar truss issue. A small group of researchers then implemented the algorithm to explore its use in structural optimisation. In forecasting the evolution of one and two-dimensional non-linear systems over time, Darwin (1960) showed magnificent results.

Implementation of the Genetic Algorithm

- Initially, the time sequence $\{x(t_i), t_i = 1, \dots, N\}$ with the sequence of equations for candidate population is given at random.
- Normally, such equations are of the form $x(t) = (A \otimes B) \otimes (C \otimes D)$, with the parameter variables A, B, C, and D are the earlier state variables ($x(t - \tau), x(t - 2\tau), \dots, x(t - m\tau)$), with τ being the discrete time unit), or the actual values constant, so that the \otimes symbol is one of the known four basic arithmetic operators ($+$, $-$, \times and \div).
- Other mathematical operators may be achievable, but increasing the number of accessible operators makes the functional optimisation method difficult.

- A parameter that measures the results of equation strings in a training set is its fitness to the information described by:

$$R^2 = 1 - \frac{\Theta}{\sigma}, \tag{1}$$

where σ is the variance and where Θ is implied by:

$$\Theta^2 = \sum_{t=m+1}^N (x(t) - p(x(t - \tau), x(t - 2\tau), \dots, x(t - m\tau)))^2. \tag{2}$$

- Values of R^2 close to one show highly accurate forecasts, whereas low positive or negative values indicate the algorithm’s weak forecast capability.
- The equation strings with a higher number of R^2 would be taken to exchange the character string parts among them (reproduction and crossover) in discarding the few unsuitable individuals.
- The offspring is more difficult to produce than the parents.
- The total number of characters throughout the equation strings is upper bounded to avoid the generation of offspring with unreasonable frequency.
- A small proportion of the strong fundamental components, individual operators and variables, of the strings of the equation are progressively mutated at random.
- The process is performed several times to enhance the fitness of the developing population, and the empirical method to estimate function $p(\cdot)$ is obtained at the edge of the evolutionary process.

3.2. Recurrent Neural Networks

Mathematically, the RNNs are defined as:

$$y_j(t + 1) = \varphi\left(\sum_{i=1}^{m+n} w_{ji}z_i(t)\right), \tag{3}$$

$$z_i(t) = \begin{cases} y_i(t) & (i \leq n) \\ u_{i-n} & (i > n), \end{cases} \tag{4}$$

where m stand for the proportion of inputs, n for the proportion of hidden and output neurons, φ is for the arbitrary differential component, generally a sigmoid function, y_j determines the output of the j th neuron, and w_{ji} the relationship between the i th and the j th neurons. For simplicity, the external inputs u_i and recurrent inputs y_i are represented as z_i [39].

RNNs are defined as the networks with loops in them that enable the continuity of data [40]. They are used for modelling data reliant on time. The information is supplied one after the other to a network and at a single point, the network nodes save their state and then use it to alert the following step. Not the same way as multilayer perceptron (MLP), RNNs use input data temporally, making them more suitable for time series data. An RNN realises the ability through recurrent neuronal connections. A basic equation which provided an input sequence $x = (x_1, x_2, \dots, x_T)$ for the RNN hidden state h_t is:

$$h_t = \begin{cases} 0, & \text{if } (t = 0) \\ \phi(h_{t-1}, x_t) & \text{otherwise,} \end{cases} \tag{5}$$

where the Φ function is non-linear. Recurrent hidden state update is realised as follows:

$$h_t = g(wx_t + uh_{t-1}), \tag{6}$$

where g is a function of the hyperbolic tangent. Generally, this generic environment of RNN with no neurons often suffers from gradient issues that are going away.

3.3. Benchmark Models

3.3.1. Implementation of the K-Nearest Neighbours

The KNN algorithm is a direct, simple, supervised machine learning algorithm that is used to solve problems of both classification and regression [41]. The supervised machine learning algorithm (opposed to an unsupervised machine learning algorithm) is an algorithm that depends on the labelled input data when training the function that will be used to produce acceptable results when the data that is not labelled is provided. Hence, supervised machine learning algorithms are also used to address regression or classification problems.

The strengths and weaknesses of the the proposed models, GA, RNN and KNN are presented in Table 2.

Table 2. Model comparisons.

Models	Strengths	Weaknesses
M1 (GA)	<ol style="list-style-type: none"> 1. The concept is simple to explain. 2. Works well with a combination of discrete/continuous problems. 3. It's ideal for a noisy environment. 4. Write robustly to local minima/maxima. 	<ol style="list-style-type: none"> 1. Implementation remains an art. 2. It takes a long time to compute. 3. It requires less knowledge about the problem, but it can be difficult to design an objective function and get the representation and operators correct.
M2 (RNN)	<ol style="list-style-type: none"> 1. It can process inputs of any length. 2. The model size does not increase as the input increases. 3. Weights can be shared between time steps. 4. It is designed to remember each piece of information over time, which is extremely useful in any time series predictor. 	<ol style="list-style-type: none"> 1. It can be challenging to train it. 2. The computation is sluggish because of its recurrent existence. 3. Problems like exploding and gradient vanishing are common. 4. When using relu or tahn as activation functions, processing long sequences becomes extremely difficult.
M3 (KNN)	<ol style="list-style-type: none"> 1. There is no need for a training period. 2. It is simple to implement. 3. Inference is based on the approximation of a large number of samples. 4. New data can be easily added. 	<ol style="list-style-type: none"> 1. It is ineffective when dealing with large dataset. 2. With high dimensions, it does not work well. 3. Sensitive to noisy data, missing values, and outliers.

3.3.2. The KNN Algorithm

1. load the solar data from USAid Venda,
2. initialise K to the specified number of neighbours,
3. for every example in the data,
 - calculate the distance between a query example and the current example,
 - add the distance and an example index to an ordered set,
4. order the set of distances and indices by distances (in ascending order) from the smallest to the largest,
5. choose the first K entries in the list which has been sorted,
6. get the labels for K entries which have been chosen,
7. returns the K label mean if a regression occurs,
8. returns the K label mode if it is classified.

3.4. Variable Selection, Parameter Estimation

3.4.1. Variable Selection

In the variable selection, we involve the selection of function variables that determines the target variables when limiting the number of variables from the system. The variable selection framework plays an important role in terms of avoiding over-fitting, promoting

analysis of the patterns, and computational time reduction. We have various variable selection methods, but in this paper, we use least absolute shrinkage and selection operator (Asso) [42].

3.4.2. The Prediction Intervals

A prediction interval (PI) by its nature is a useful tool for modelling uncertainty. It is composed of lower and upper boundaries which cover the unidentified target value of the future value with any probability $(1 - a)\%$ called confidence level. They are more appropriate and more valuable information than point forecasts for decision-makers [43]. The width of the prediction interval (PIW) is given as:

$$PIW_t = UL_t - LL_t, \tag{7}$$

where UL_t and LL_t are the upper and lower bounds, respectively. In this study, probability density plots, and box and whisker plots were used to find the model which yields narrower PIW.

3.4.3. Evaluation of the Prediction Intervals

For a prediction interval (PI) with nominal confidence (PINC) of $(1 - a)100\%$, it is formulated as the probability with which $\hat{y}_{t,\tau}$ belonging to the predictive interval (LL_t, UL_t) . We provide PINC as follows:

$$PINC = P(\hat{y}_{t,\tau} \in (UL_t, LL_t) = (1 - a)100\%). \tag{8}$$

This study uses the prediction interval normalised average width (PINAW) and the prediction interval coverage probability (PICP). The PICP is described by [43]:

$$PICP = \frac{1}{m} \sum_{t=1}^m I_t, \tag{9}$$

with m being the number of the forecasts and I being the binary variables given by:

$$I_t = \begin{cases} 1, & \text{if } y_t \in (UL_t, LL_t) \\ 0, & \text{if otherwise.} \end{cases} \tag{10}$$

PINAW is another measure used to assess the accuracy of forecast intervals and is provided as [43]:

$$PINAW = \frac{1}{m(\max(y_t) - \min(y_t))} \sum_{t=1}^m (UL_t - LL_t). \tag{11}$$

3.5. Evaluation Metrics

The Mean absolute error (MAE), Mean square error (MSE), Root mean square error (RMSE), Relative MAE (rMAE), and Relative RMSE (rRMSE) used to evaluate the performance of the models. They are defined as follows:

$$MAE = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t|, \tag{12}$$

$$rMAE = \frac{1}{n} \sum_{t=1}^n \left[\frac{\hat{y}_t - y_t}{y_t} \right], \tag{13}$$

$$MSE = \frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2, \tag{14}$$

$$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{n}}, \quad (15)$$

$$\text{rRMSE} = \frac{100}{\bar{y}} \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}}, \quad (16)$$

in which y_t and \hat{y}_t are the actual and predicted values, \bar{y} is the mean value of y_t , $t = 1, \dots, n$, k is the dummy variable time, and n is the number of data elements. The smaller value error is estimated closer to the true values.

3.6. The Tests for Predictive Accuracy

Diebold–Mariano and Giacomini–White Tests

Under the Diebold–Mariano test, the null hypothesis is that of equal predictive accuracy of two competing forecasts [44,45]. One of the main advantages of this test is that it takes into account the sampling variability in the average losses [46].

The Giacomini–White (GW) test which is a generalisation of the DM test, tests for equal conditional predictive ability of two competing models [47]. The test accounts for uncertainty in the estimation of parameters [47], making it a better test compared to the DM test.

3.7. Murphy Diagram

It is built on the concept that “if something wrong can happen, it’s going wrong”. It is close to other methods of analysis such as fault trees, as they evaluate errors based on the possible causes of such errors. The Murphy diagram works by comparing the mean forecasts. Hence, the mean of the forecast distribution is acquired by minimising the squared error loss function, $\mathbf{S}(x, y) = (x - y)^2$, where x is the point forecast and y is the actual observation [48]. With the equation:

$$E(Y) = \operatorname{argmin}_x \mathbf{S}(x, y), \quad (17)$$

it is given that any scoring function meeting the constraint can be defined as:

$$\mathbf{S}(x, y) = \int_{-\infty}^{\infty} \mathbf{S}_{\theta}(x, y) d\mathbf{H}(\theta) \quad (18)$$

with \mathbf{H} as a non-negative indicator and

$$\mathbf{S}_{\theta}(x, y) = \begin{cases} |y - \theta| & \text{if } \min(x, y) \leq \theta < \max(x, y) \\ 0 & \text{otherwise.} \end{cases}$$

Various \mathbf{H} assessments offer different scoring functions, but for all such scoring functions, $\mathbf{S}_{\theta}(x, y)$ is the same.

If the point forecasts for n events is given, then we are able to get the average value of $\mathbf{S}_{\theta}(x, y)$ for each θ :

$$s(\theta) = \frac{1}{n} \sum_{i=1}^n \mathbf{S}_{\theta}(x_i, y_i), \quad (19)$$

and plot this as a θ function. This is what [48] call the “Murphy diagram”. The same approach can be applied for quantile and expertise forecasts.

3.8. Data and Features

The GA, RNN, and the benchmark model, which is the k-nearest neighbour (KNN) models will be implemented using global horizontal irradiance (GHI) data from the Vuwani radiometric station in the Limpopo province of SA.

3.8.1. Data

The data used in this paper is obtained from the USAid Venda radiometric station measured at one-minute intervals accessible at Vuwani radiometric station (USAid Venda) (<https://sauran.ac.za/>, accessed 3 May 2020). Figure 1 shows the pyranometer at the USAid Venda radiometric station, which is on an inside enclosure at Vuwani in the Limpopo province.



Figure 1. Picture showing the location of the Vuwani radiometric station (USAid Venda). Source: <https://sauran.ac.za/>, accessed 3 May 2020.

3.8.2. Features

Models developed in this chapter will use GHI as the response variable and the predictor variables are air temperature (Temp), barometric pressure (BP), rainfall (Rain), relative humidity (RH), wind direction (WD), wind direction standard deviation (WD StdDev), and wind speed (WS).

4. Results

The GA, RNN, and KNN algorithms implemented using the Keras deep learning package (<https://keras.io/>, accessed 3 May 2020). All models were implemented using Python packages. Both Python and R [49] are used in this paper.

4.1. Empirical Results and Discussion

This section presents the data analysis using the algorithms discussed in the methodology.

4.1.1. Exploratory Data Analysis

Data covering the period 4 January 2020 to 31 October 2020 were used in this study. The data were obtained from Vuwani radiometric station. We can see from Figure 2 and Table 3 that the GHI values are right-skewed and platykurtic.

In this work, a non-linear trend was extracted by fitting the cubic smoothing spline function provided by the equation:

$$\pi(t) = \sum_{t=0}^n (y_t - f(t))^2 + \lambda \int \{f''(t)\}^2 dt, \quad (20)$$

where λ is the smoothing parameter that is predicted using the generalised cross-validation (GCV) criterion. Figure 3 illustrates the cubic smoothing spline and non-linear trend fitted with the projected lambda value. The non-linear trend values extracted, used to model solar irradiance.

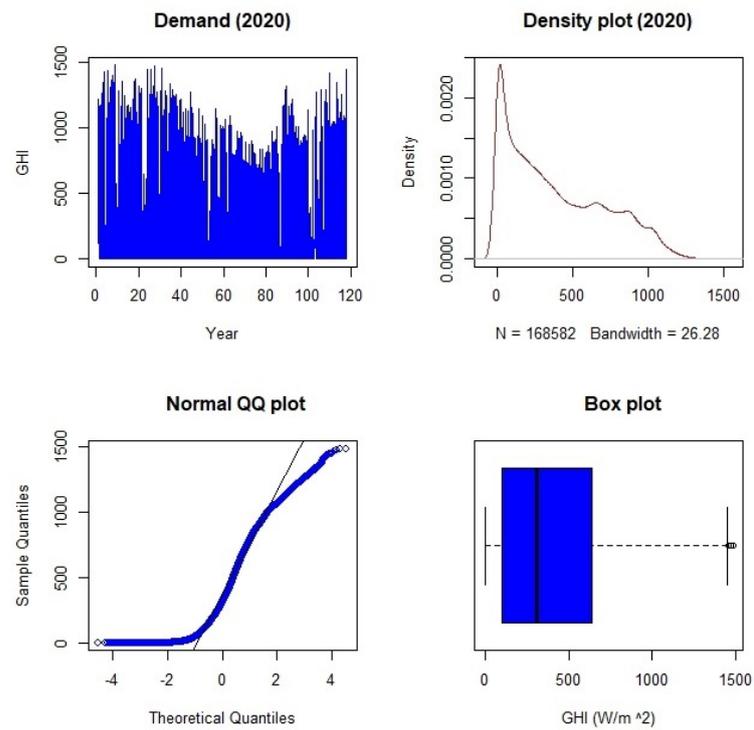


Figure 2. Time series plot of GHI, Density plot, Quantile-Quantile(qq) plot, and Box and whisker plot.

Table 3. The descriptive statistics of the GHI measured in W/m^2 .

Min	Max	Median	Mean	St.Div	Skewness	Kurtosis
0.0032	1481.6300	307.6562	388.0439	324.1666	0.6130	−0.7568

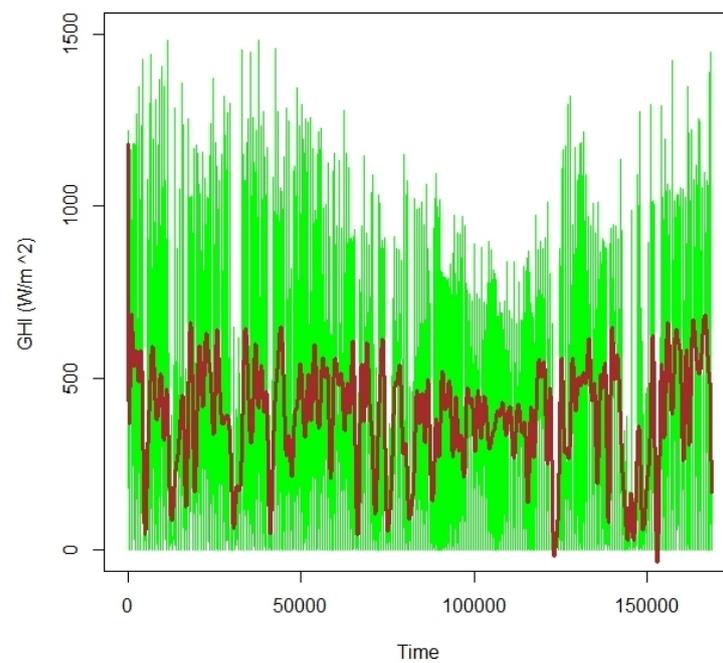


Figure 3. Plot of global horizontal irradiance from 04/01/2020 to 31/10/2020 (green) superimposed with a fitted cubic smoothing spline trend (red).

4.2. Variable Selection

In this work, variable selection is done using Lasso, which uses the ℓ loss function penalty:

$$\hat{\beta}_{Lasso}(\lambda) = \arg \min \|\bar{y} - \mathbf{X}\hat{\beta}\|_2^2 + \lambda \|\hat{\beta}\|_1. \quad (21)$$

Table 4 shows parametric coefficients of the Lasso regression. Based on these results, all variables except rainfall (Rain), which has a zero coefficient, are important predictor variables.

Table 4. Parametric coefficients.

Variables	Coefficients
Intercept	-2.158294×10^4
Temp	3.815809×10^1
RH	-1.326136×10^0
WD	1.645751×10^{-1}
Rain	0.000000
WS	1.866991×10^0
WD StdDev	7.173793×10^0
BP	2.215658×10^1

4.3. Forecasting Results

Similar to other works done previously, performance measures such as MAE and RMSE will be used to select the best model for short-term forecasting of GHI. Table 5 presents a summary of the evaluation metrics used in this study. Based on the RMSE and rRMSE, GA is found to be the best model. However, based on MAE and rMAE, RNN is the best model. With these results, it is then difficult to select the best model. More evaluation metrics will have to be used to come up with the best model out of the three presented in this study.

Table 5. Assessment of models.

	GA	RNN	KNN
RMSE	35.50	56.89	57.48
rRMSE	5.96	7.54	7.58
MAE	26.74	20.18	20.94
rMAE	5.17	4.49	4.58

Figure 4 shows a plot of GHI superimposed with forecasts from RNN, GA, and KNN models. It appears as if GA tends to over-predict GHI on sunny days compared to cloudy and rainy days (where the volatility is very high).

Figure 5 illustrates the density plot of the actual GHI (solid lines) and model's forecasts (dashed lines). The forecasts from the models appear to be different from the actual observations, with RNN being the one that is closer to the actual observations.

4.4. Models' Comparative Analysis

In this section, the evaluation of the fitted models centred on the empirical prediction intervals (PIs) together with the forecast error distributions from each model forecast discussed.

4.4.1. Evaluation of Predictive Interval

A comparison of the best models using PICP, PINAW, and PINAD is shown in Table 6. The PINC is taken to be a 95 percent level of confidence, where it is valid for all the models. All the models have valid PICPs. They are all greater than 95%. A model with the narrowest PINAW and smallest PINAD is recognised to be the best fitting model [50]. In this case, the best fitting model is RNN.

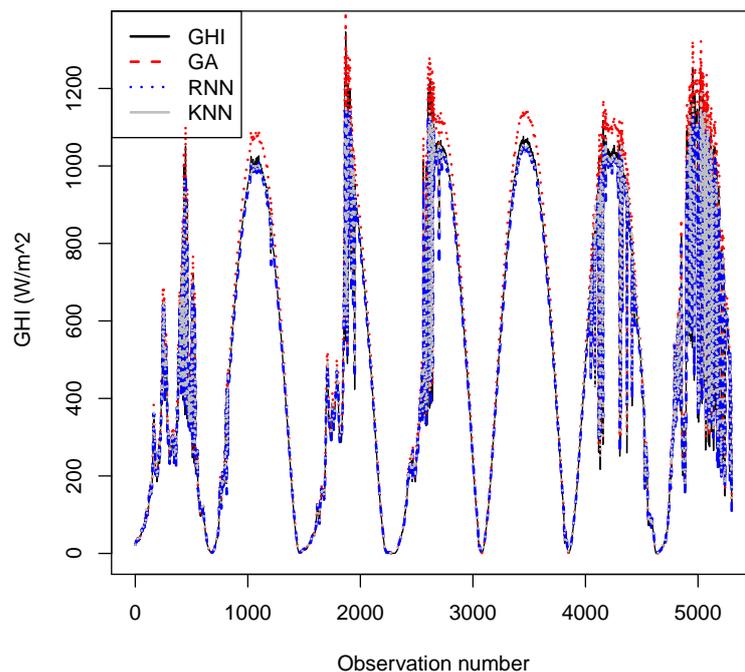


Figure 4. Plot of GHI superimposed with forecasts from RNN, GA, and KNN models.

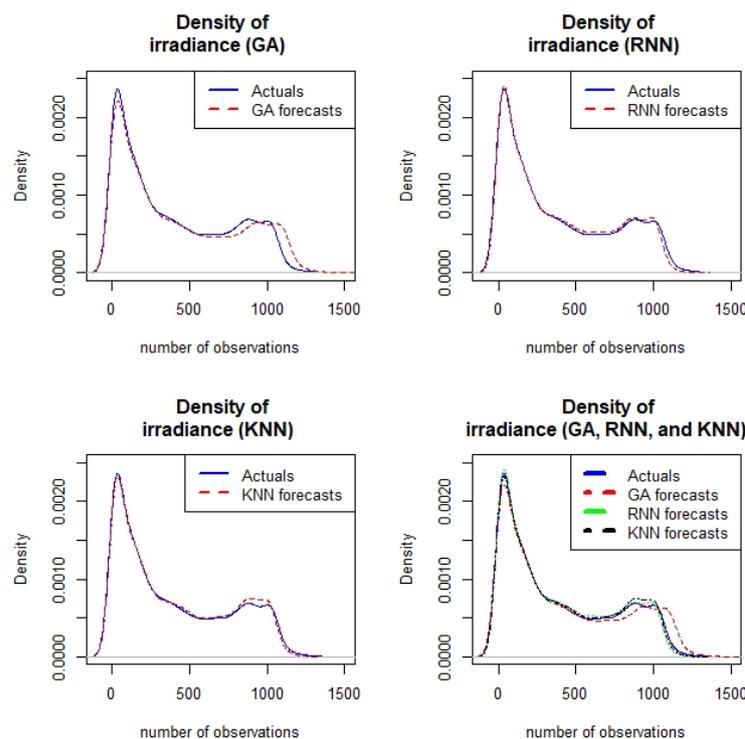


Figure 5. Density plots of the actual GHI (solid lines) and model’s forecasts (dashed lines) where **Top left panel:** Actual GHI and forecasts of GA **Top right panel:** Actual GHI and forecasts of RNN **Bottom left panel:** Actual GHI and forecasts of KNN and **Bottom right panel:** Actual GHI and forecasts of RNN, GA, and KNN where the actual GHI and forecasts of GA are shown by solid lines.

Table 6. Comparative analysis of the best models with the confidence interval (CI) (PICP, PINAW and PINAD) at 95%.

Models	PICP	PINAW	PINAD
GA	98.00%	11.81%	0.07%
RNN	98.60%	11.47%	0.05%
KNN	98.12%	15.09%	0.08%

4.4.2. Residual Analysis

Table 7 provides descriptive statistics of the residuals for models GA, RNN, and KNN with a confidence level of 95% for the PINC value. From the table, we can see that GA model is the one with the smallest standard deviation compared to other models, which tells us that it has a smaller error distribution. So, GA makes it the best model compared to others. For the GA, RNN, and KNN, the error distributions are approximately normal because their skewness is close to zero. The values for kurtosis are greater than 3 for all the models, except for GA model.

Figure 6 shows the box plots of the forecast errors for the entire fitted models ResGA, ResRNN, and ResKNN, where ResGA, ResRNN, and ResKNN are the residuals from GA, RNN, and KNN, respectively. The box plot for the prediction errors from the GA model shows that the distribution has shorter tails an indication of a smaller error distribution compared to those of the other two models.

Table 7. Residual comparison of the models.

	Mean	Median	Min	Max	StDev	Skewness	Kurtosis
GA	-26.74	-19.25	-83.12	-0.10	23.36	-0.54	-1.14
RNN	3.78	-0.24	-688.63	764.30	56.76	0.48	41.14
KNN	-5.44	-4.33	-747.01	796.95	57.22	0.08	43.48

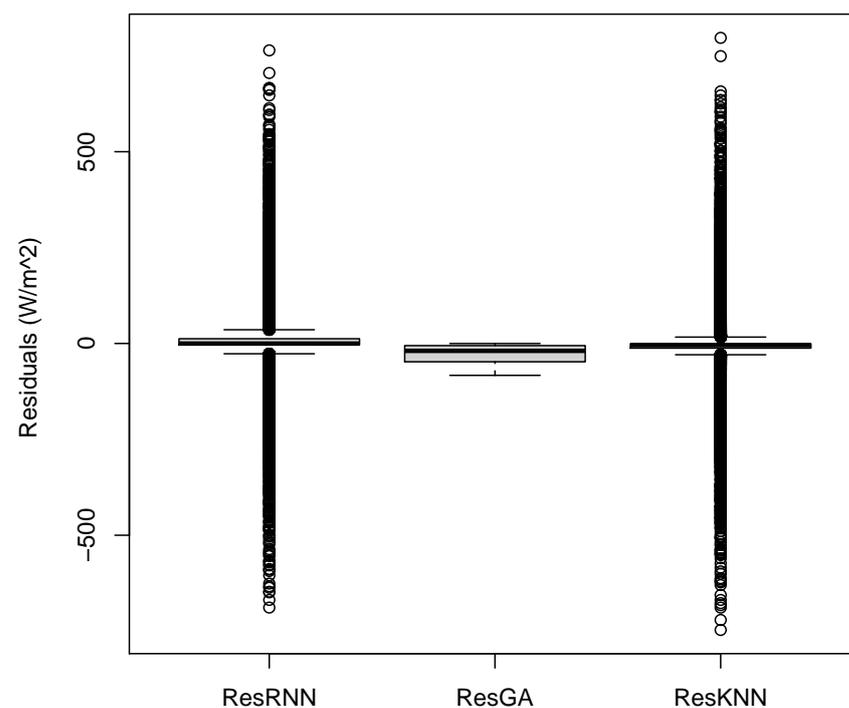


Figure 6. Box plots of the residuals from ResGA, ResRNN, and ResKNN.

4.4.3. Diebold–Mariano Test

The null hypothesis is that the forecast accuracy between a pair of forecasts from the two methods is the same. Results presented in Table 8 show that we reject the null hypothesis of equal forecast accuracy between pairs of the models.

Table 8. Diebold-Mariano test.

Statistic	p -Value	
GA	−37.795	<0.00001
RNN	−47.789	<0.00001
KNN	−47.312	<0.00001

4.4.4. Murphy Diagrams

Murphy diagrams (MDs) are comparable to several other methods of analysis, such as fault trees, because they analyse errors based on the possible causes of those errors. It is not essential to identify a particular scoring function before forecast evaluation when using MDs to compare forecast results [51]. Figures 7–9 are the Murphy diagrams where, for the difference between the models, the shaded area displays 95% point wise confidence intervals.

The GA curve (in red colour) in Figure 7 dominates the RNN curve (in blue colour) suggesting that GA provides a greater predictive ability. The curves in Figure 8 are very close to each other. This suggests that neither of the two methods (RNN and KNN) dominate each other. The score differences fluctuate between positive and negative values and include zero for all values of θ . As for GA and KNN models, it is seen in Figure 9 that the forecasts from the GA model dominate those from the KNN model. This is an indication that the forecast accuracy from GA is better. Since GA dominates both RNN and KNN, we can conclude that the GA provides the greatest predictive ability.

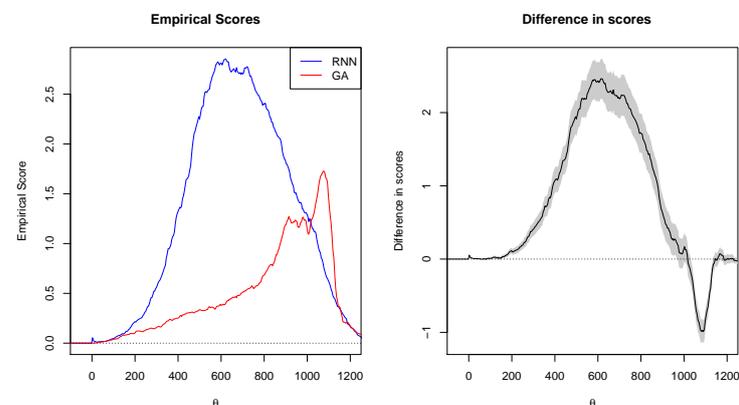


Figure 7. Plots of the RNN and GA forecasts showing the relationship between their empirical scores (**left panel**) and the differences in empirical scores with 95% confidence intervals (**right panel**).

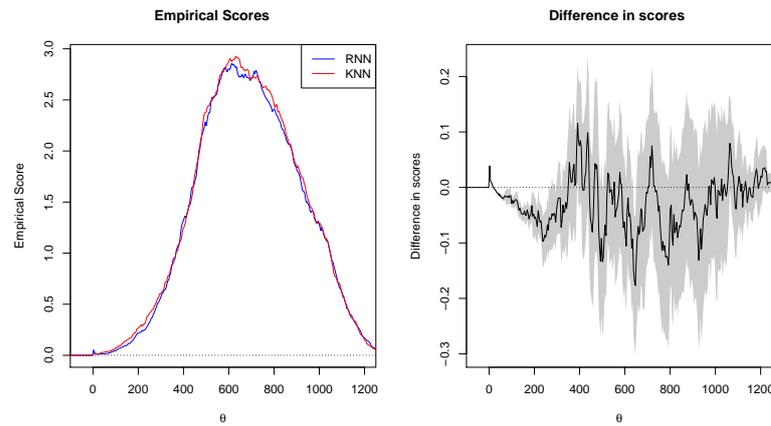


Figure 8. Plots of the RNN and KNN forecasts showing the relationship between their empirical scores (left panel) and the differences in empirical scores with 95% confidence intervals (right panel).

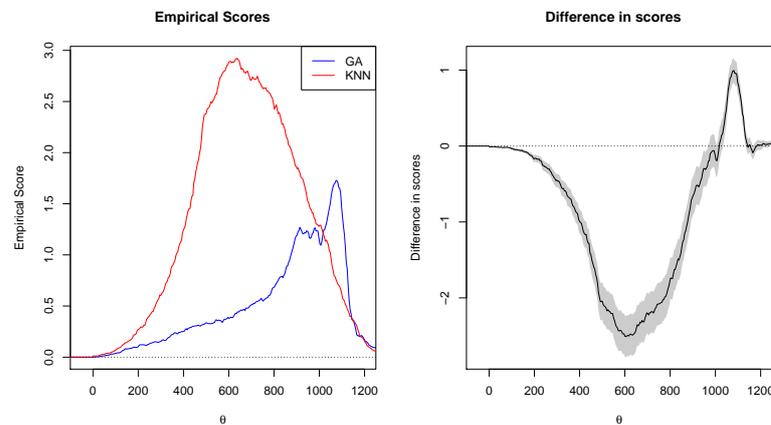


Figure 9. Plots of the GA and KNN forecasts showing the relationship between their empirical (left panel) and the differences in empirical scores with 95% confidence intervals (right panel).

4.4.5. Giacommini–White Test

Giacomini-White (GW) test is a test of equal conditional predictive ability.

From Table 9, GA dominates RNN, meaning that GA has better predictive ability. It is also seen that RNN has a better predictive ability compared to KNN and that GA dominates KNN. From these results, it can be deduced that GA dominates RNN which dominates KNN. Therefore, GA has the best conditional predictive ability compared to the other two models. These results are consistent with those from the Murphy diagrams.

Table 9. Model comparisons: Giacommini–White test.

Models	Test Statistic	p-Value	Result
RNN = GA	323.925	<0.0001	Sign of mean loss is (+). GA dominates RNN
RNN = KNN	4.569	0.1018	Sign of mean loss is (-). RNN dominates KNN
GA = KNN	294.676	<0.0001	Sign of mean loss is (-). GA dominates KNN

Figure 10 shows GHI superimposed with GA, RNN, and KNN forecasts during clear sky days (top panels) and during cloudy days (bottom panels) on four selected days during the testing period. For the selected days, the GA tends to over-predict GHI during sunny days compared to the other two models, RNN and KNN. However during cloudy days GA appears to perform better.

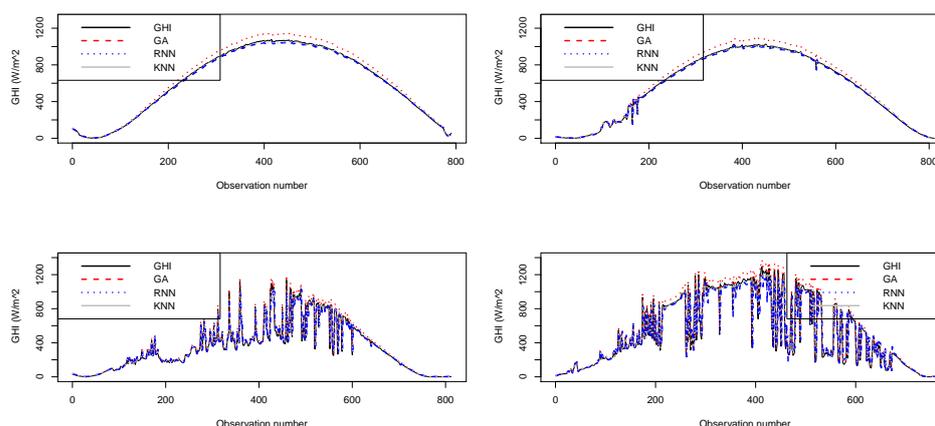


Figure 10. GHI superimposed with GA, RNN, and KNN forecasts **Top panels:** Clear sky days **Bottom panels:** Cloudy days.

4.4.6. Discussion of Results

This paper focused on forecasting GHI at one radiometric station in South Africa using high-frequency data (measured at one-minute intervals) obtained from the Vuwani radiometric station (USAid Venda). The data is from January 2020 to October 2020. Based on the RMSE and rRMSE, the GA model was found to be the best model. However, based on MAE and rMAE, RNN was the best model. With these results, it was then difficult to select the best model. Further evaluation of the forecasts were then done based on the Diebold–Mariano tests, Giacomini–White tests, and Murphy diagrams. From the Diebold–Mariano tests, the null hypothesis that the forecast accuracy between a pair of forecasts from the two methods is the same was rejected for all the three pairs. Based on the Murphy diagrams, the GA dominated both RNN and KNN, meaning that it provides the greatest predictive ability. Similarly, with the Giacomini–White tests, GA was found to have the best conditional predictive ability compared to the other two models.

Motivated by previous research by other authors such as [13,52], among others, a GA model was developed and compared to RNN and KNN models. Using South African hourly GHI data from University of Pretoria radiometric station, [52] used the long short-term memory (LSTM) network, feed forward neural network (FFNN), support vector regression (SVR), and principal component regression (PCR). Based on the RMSE, the GA model used in the current paper has a smaller evaluation metric value compared to those of the models used in [52].

5. Conclusions

The paper presented the application of forecasting high frequency solar irradiance data using genetic algorithm (GA), recurrent neural network (RNN), and k-nearest neighbour (KNN) models. The least absolute shrinkage and selection operator was used for variable selection. Based on the evaluation metrics used in the study GA was found to have the greatest conditional predictive ability compared to RNN and KNN. The study could be useful to decision makers in power utility companies such as Eskom in aligning electricity demand and its supply in an efficient way that promotes potential economic growth and environmental sustainability.

Author Contributions: M.R.: Conceptualisation: Data curation: Formal analysis: Investigation: Methodology: Paper administration: Resources: Software: Validation: Visualisation: Writing original draft: Writing—review and editing. C.S.: Conceptualisation: Formal analysis: Investigation: Methodology: Resources: Software: Supervision: Validation: Visualisation: Writing—review and editing. A.B.: Investigation: Resources: Software: Supervision: Validation: Visualisation: Writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by the National E-Science Postgraduate Teaching and Training Platform (NEPTTP) at <https://www.wits.ac.za/nepttp/>, accessed 1 February 2020.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used in this paper is obtained from the USAid Venda radiometric station measured at one-minute intervals accessible at <https://sauran.ac.za/>, accessed on 3 May 2020.

Acknowledgments: The authors are grateful to the numerous people for helpful comments on this paper. The support of the DST-CSIR National e-Science Postgraduate Teaching and Training Platform (NEPTTP) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at are those of the authors and are not necessarily to be attributed to the NEPTTP.

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

Abbreviations

The following abbreviations are used in this manuscript:

GA	Genetic Algorithm
RNN	Recurrent Neural Network
KNN	K-Nearest Neighbour
MAE	Mean Absolute Error
RMSE	Root Mean Square Error
MSE	Mean Square Error
rMAE	relative Mean Square Error
rRMSE	relative Root Mean Square Error
QRA	Quantile Regression Averaging
DM	Diebold-Mariano
RES	Renewable Energy Source
PV	Photovoltaic
NWP	Numerical Weather Prediction
ANN	Artificial Neural Network
SVM	Support Vector Machine
CARDS	Combined Autoregressive Furthermore, Dynamic System
PSO	Particle Swarm Optimisation
GAM	Generalised Additive Model
LASSO	Least Absolute Shrinkage Furthermore, Selection Operator
NF	Neuro-Fuzzy
ARMA	Autoregressive Moving Averaging
NN	Neural Network
GHI	Global Horizontal Irradiance
SUI	Solar Utility Index
AAKR	Auto-Associative Kernel Regression
MLP	Multilayer Perceptron
PI	Prediction Interval
PINC	Prediction Interval With Nominal Confidence
PIW	Prediction Interval Width
PICP	Prediction Interval Coverage Probability
PINAW	Prediction Interval Normalised Average Width
GCV	Generalised Cross Validation
MD	Murphy Diagram

References

1. Andrade, J.R.; Bessa, R.J. Improving renewable energy forecasting with a grid of numerical weather predictions. *IEEE Trans. Sustain. Energy* **2017**, *8*, 1571–1580. [[CrossRef](#)]
2. Kariniotakis, G. *Renewable Energy Forecasting: From Models to Applications*; Woodhead Publishing: Cambridge, UK, 2017.

3. Zendeheboudi, A.; Baseer, M.A.; Saidur, R. Application of support vector machine models for forecasting solar and wind energy resources: A review. *J. Clean. Prod.* **2018**, *199*, 272–285. [\[CrossRef\]](#)
4. Mohammadi, K.; Shamshirband, S.; Danesh, A.S.; Abdullah, M.S.; Zamani, M. Temperature-based estimation of global solar radiation using soft computing methodologies. *Theor. Appl. Climatol.* **2016**, *125*, 101–112. [\[CrossRef\]](#)
5. Zhandire, E. Solar resource classification in South Africa using a new index. *J. Energy South. Afr.* **2017**, *28*, 61–70. [\[CrossRef\]](#)
6. Kleissl, J. *Solar Energy Forecasting and Resource Assessment*; Academic Press: Cambridge, MA, USA, **2013**.
7. Cristaldi, L.; Leone, G.; Ottoboni, R. A hybrid approach for solar radiation and photovoltaic power short-term forecast. In Proceedings of the 2017 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), Turin, Italy, 22–25 May 2017; pp. 1–6.
8. Kostylev, V.; Pavlovski, A. Solar power forecasting performance towards industry standards. In *1st International Workshop on the Integration of Solar Power into Power Systems, Aarhus, Denmark*; 2011. Available online: <http://www.greenpowerlabs.com/gpl/wp-content/uploads/2013/12/wp-sol-pow-forecast-kostylev-pavlovski.pdf> (accessed on 1 February 2020).
9. Rezrazi, A.; Hanini, S.; Laidi, M. An optimisation methodology of artificial neural network models for predicting solar radiation: A case study. *Theor. Appl. Climatol.* **2016**, *123*, 769–783. [\[CrossRef\]](#)
10. Liu, B.Y.H.; Jordan, R.C. The interrelationship and characteristic distribution of direct, diffuse and total solar radiation. *Sol. Energy* **1960**, *4*, 1–19. [\[CrossRef\]](#)
11. Yang, X.; Jiang, F.; Liu, H. Short-term solar radiation prediction based on SVM with similar data. In Proceedings of the 2013 IEEE RPG (2nd IET Renewable Power Generation Conference), Beijing, China, 9–11 September 2013.
12. Sun, S.; Wang, S.; Zhang, G.; Zheng, J. A decomposition-clustering ensemble learning approach for solar radiation forecasting. *Sol. Energy* **2018**, *163*, 189–199. [\[CrossRef\]](#)
13. Reyes-Belmonte, M.A. Quo Vadis Solar Energy Research? *Appl. Sci.* **2021**, *11*, 3015. [\[CrossRef\]](#)
14. Fan, J.; Wang, X.; Wu, L.; Zhou, H.I.; Zhang, F.; Yu, X.; Lu, X.; Xiang, Y. 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. Manag.* **2018**, *164*, 102–111. [\[CrossRef\]](#)
15. Haykin, S.S. *Neural Networks and Learning Machines*; Pearson: Upper Saddle River, NJ, USA, 2009; Volume 3.
16. Abedinia, O.; Amjady, N.; Ghadimi, N. Solar energy forecasting based on hybrid neural network and improved metaheuristic algorithm. *Comput. Intell.* **2018**, *34*, 241–260. [\[CrossRef\]](#)
17. Cadenas, E.; Rivera, W. Short term wind speed forecasting in La Venta, Oaxaca, Mexico, using artificial neural networks. *Renew. Energy* **2009**, *34*, 274–278. [\[CrossRef\]](#)
18. Capizzi, G.; Bonanno, F.; Napoli, C. Recurrent neural network-based control strategy for battery energy storage in generation systems with intermittent renewable energy sources. In Proceedings of the 2011 International Conference on Clean Electrical Power (ICCEP), Ischia, Italy, 14–16 June 2011; pp. 336–340.
19. Tsai, S.B.; Xue, Y.; Zhang, J.; Chen, Q.; Liu, Y.; Zhou, J.; Dong, W. Models for forecasting growth trends in renewable energy. *Renew. Sustain. Energy Rev.* **2017**, *77*, 1169–1178. [\[CrossRef\]](#)
20. Peng, Z.; Yoo, S.; Yu, D.; Huang, D. Solar irradiance forecast system based on geostationary satellite. In Proceedings of the 2013 IEEE International Conference on Smart Grid Communications (SmartGridComm), Vancouver, BC, Canada, 21–24 October 2013; pp. 708–713.
21. Tartibu, L.K.; Kabengele, K.T. Forecasting net energy consumption of South Africa using artificial neural network. In Proceedings of the 2018 International Conference on the Industrial and Commercial Use of Energy (ICUE), Cape Town, South Africa, 13–15 August 2018; pp. 1–7.
22. Sigauke, C. Forecasting medium-term electricity demand in a South African electric power supply system. *J. Energy S. Afr.* **2017**, *28*, 54–67. [\[CrossRef\]](#)
23. Marwala, L.; Twala, B. Forecasting electricity consumption in South Africa: ARMA, neural networks and neuro-fuzzy systems. In Proceedings of the 2014 International Joint Conference on Neural Networks (IJCNN), Beijing, China, 6–11 July 2014; pp. 3049–3055.
24. Warsono, D.J.K.; Özveren, C.S.; Bradley, D.A. Economic load dispatch optimization of renewable energy in power system using genetic algorithm. *Proc. PowerTech* **2007**, 2174–2179. [\[CrossRef\]](#)
25. Mellit, A.; Shaari, S. Recurrent neural network-based forecasting of the daily electricity generation of a Photovoltaic power system. *Ecol. Renew. Energy (EVER) Monaco March* **2009**, 26–29. Available online: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.533.9274&rep=rep1&type=pdf> (accessed on 1 February 2020).
26. Grady, S.A.; Hussaini, M.Y.; Abdullah, M.M. Placement of wind turbines using genetic algorithms. *Renew. Energy* **2005**, *30*, 259–270. [\[CrossRef\]](#)
27. Khan, P.W.; Byun, Y. Genetic Algorithm Based Optimized Feature Engineering and Hybrid Machine Learning for Effective Energy Consumption Prediction. *IEEE Access* **2020**, *8*, 196274–196286. [\[CrossRef\]](#)
28. Gardes, L.; Girard, S. Conditional extremes from heavy tailed distributions: An application to the estimation of extreme rainfall return levels. *Extremes* **2010**, *13*, 177–204. [\[CrossRef\]](#)
29. VanDeventer, W.; Jamei, E.; Gokul, G.; Thirunavukkarasu, S.; Seyedmahmoudian, M.; Soon, T.K.; Horan, B.; Mekhilef, S.; Stojcevski, A. Short-term PV power forecasting using hybrid GASVM technique. *Renew. Energy* **2019**, *140*, 367–379. [\[CrossRef\]](#)
30. Al-lahham, A.; Theeb, O.; Elalem, K.; Alshawi, T.A.; Alshebeili, S.A. Sky imager-based forecast of solar irradiance using machine learning. *Electronics* **2020**, *9*, 1700. [\[CrossRef\]](#)

31. Pattanaik, J.K.; Basu, M.; Dash, D.P. Improved real-coded genetic algorithm for fixed head hydrothermal power system. *IETE J. Res.* **2020**, *1*–10. [[CrossRef](#)]
32. Benamrou, B.; Ouardouz, M.; Allaouzi, I.; Ahmed, M.B. A proposed model to forecast hourly global solar irradiation based on satellite derived data, deep learning and machine learning approaches. *J. Ecol. Eng.* **2020**, *21*, 26–38. [[CrossRef](#)]
33. Brahma, B.; Wadhvani, R. Solar irradiance forecasting based on deep learning methodologies and multi-site data. *Symmetry* **2020**, *12* 1830. [[CrossRef](#)]
34. Mbuva, R.; Jonsson, M.; Ehn, N.; Herman, P. Bayesian neural networks for one-hour ahead wind power forecasting. In Proceedings of the 2017 IEEE 6th International Conference on Renewable Energy Research and Applications (ICRERA), San Diego, CA, USA, 5–8 November 2017; pp. 591–596.
35. Mpfumali, P.; Sigauke, C.; Bere, A.; Mulaudzi, S. Probabilistic solar power forecasting using partially linear additive quantile regression models: An application to South African data. *Energies* **2019**, *12*, 3569. [[CrossRef](#)]
36. Adeala, A.A.; Huan, Z.; Enweremadu, C.C. Evaluation of global solar radiation using multiple weather parameters as predictors for South Africa provinces. *Therm. Sci.* **2015**, *19*, 495–509. [[CrossRef](#)]
37. Al-Karaghoul, A.; Kazmerski, L.L. Energy consumption and water production cost of conventional and renewable-energy-powered desalination processes. *Renew. Sustain. Energy Rev.* **2013**, *24*, 343–356. [[CrossRef](#)]
38. Rajeev, S.; Krishnamoorthy, C.S. Genetic algorithm—Based methodology for design optimization of reinforced concrete frames. *Comput. Aided Civ. Infrastruct. Eng.* **1998**, *13*, 63–74. [[CrossRef](#)]
39. Garcia-Pedrero, A.; Gomez-Gil, P. Time series forecasting using recurrent neural networks and wavelet reconstructed signals. In Proceedings of the 2010 20th International Conference on Electronics Communications and Computers (CONIELECOMP), Cholula, Puebla, Mexico, 22–24 February 2010; pp. 169–173.
40. Ugurlu, U.; Oksuz, I.; Tas, O. Electricity price forecasting using recurrent neural networks. *Energies* **2018**, *11*, 1255. [[CrossRef](#)]
41. Horton, P.; Nakai, K. Better Prediction of Protein Cellular Localization Sites with the it k-Nearest Neighbors Classifier. In Proceedings of the International Conference on Intelligent Systems for Molecular Biology, Halkidiki, Greece, 21–25 June 1997; Volume 5, pp. 147–152.
42. Bien, J.; Taylor, J.; Tibshirani, R. A lasso for hierarchical interactions. *Ann. Stat.* **2013**, *41*, 1111–1141. [[CrossRef](#)]
43. Quan, H.; Srinivasan, D.; Khosravi, A. Uncertainty handling using neural network-based prediction intervals for electrical load forecasting. *Energy* **2014**, *73*, 916–925. [[CrossRef](#)]
44. Diebold, F.X.; Mariano, R. Comparing predictive accuracy. *J. Bus. Econ. Statist.* **1995**, *13*, 253–265.
45. Triacca, U. Comparing Predictive Accuracy of Two Forecasts. 2018. Available online: <http://www.phdeconomics.sssup.it/documents/Lesson19.pdf> (accessed on 17 January 2021).
46. Tarassow, A.; Schreiber, S. FEP—The Forecast Evaluation Package for Gretl; Version 2.41. 2020. Available online: http://ricardo.ecn.wfu.edu/gretl/cgi-bin/current_fnfiles/unzipped/FEP.pdf (accessed on 9 January 2021).
47. Lago, J.; Marcjasz, G.; De Schutter, B.; Weron, R. Forecasting day-ahead electricity prices: A review of state-of-the-art algorithms, best practices and an open-access benchmark. *Renew. Sustain. Energy Rev.* to be published. Available online: <https://arxiv.org/abs/2008.08004v1> (accessed on 22 August 2020).
48. Werner, E.; Gneiting, T.; Jordan, A.; Krüger, F. Of quantiles and expectiles: Consistent scoring functions, Choquet representations, and forecast rankings. *arXiv* **2015**. Available online: <https://arxiv.org/abs/1503.08195v2> (accessed on 23 December 2020).
49. R Core Team. R: A Language and Environment for Statistical Computing. 2021. Available online: <https://www.R-project.org/> (accessed on 15 January 2021).
50. Sun, X.; Wang, Z.; Hu, J. Prediction interval construction for byproduct gas flow forecasting using optimized twin extreme learning machine. *Math. Probl. Eng.* **2017**, 1–12. [[CrossRef](#)]
51. Ziegel, J.F.; Krüger, F.; Jordan, A.; Fasciati, F. Murphy Diagrams: Forecast Evaluation of Expected Shortfall. *arXiv* **2017**. Available online: <https://arxiv.org/abs/1705.04537v1> (accessed on 4 December 2020).
52. Mutavhatsindi, T.; Sigauke, C.; Mbuva, R. Forecasting Hourly Global Horizontal Solar Irradiance in South Africa Using Machine Learning Models. *IEEE Access* **2020**, *8*, 198872–198885. [[CrossRef](#)]