JISA (Jurnal Informatika dan Sains)
Vol. 06, No. 01, June 2023
e-ISSN: 2614-8404
p-ISSN:2776-3234

Global Horizontal Irradiance Prediction using the Algorithm of Moving Average and Exponential Smoothing

Alfin Syarifuddin Syahab[1*)], Arief Hermawan[2], Donny Avianto[3]
[1]Badan Meteorologi Klimatologi dan Geofisika, Stasiun Klimatologi D.I. Yogyakarta
[1,2]Program Studi Magister Teknologi Informasi, Universitas Teknologi Yogyakarta
[3]Program Studi Informatika, Universitas Teknologi Yogyakarta

Email: [1]6220211009.alfin@student.uty.ac.id, [2]ariefdb@uty.ac.id , [3]donny@uty.ac.id

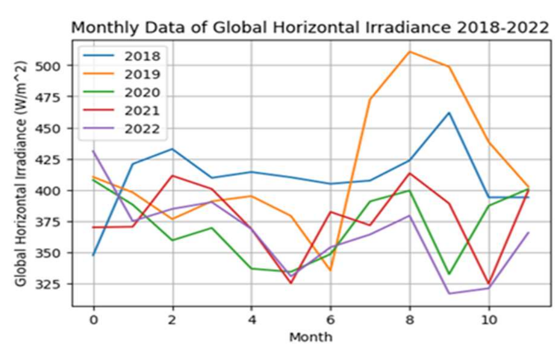
Abstract – To reduce the discrepancy between the results of the expected data and the actual data, prediction is a procedure that is calculated systematically based on owned historical and present information. For the creation of solar energy projects and for decision-making in other connected domains, solar radiation intensity prediction is essential. This study aims to create a predictive model on monthly global horizontal irradiance data. The method used is the Simple Moving Average algorithm, Exponentially Weighted Moving Average and Single Exponential Smoothing. The stages carried out in this study include data collection, data preprocessing, testing of predictive models, interpretation of data visualization, and performance evaluation. The results of calculating the error value and correlation produce an evaluation of the performance of the prediction model. The SES method, which obtained an MAE value of 7.13, a MAPE of 0.02%, an MSE of 88.07, an RMSE of 9.38, and an R^2 of 0.94, was determined to be the best prediction model by the calculation of the prediction model performance evaluation. A MAE value of 9.45, a MAPE of 0.02%, an MSE of 150.16, an RMSE of 12.25, and an R^2 of 0.91 were obtained by the EWMA method, which is also the method that produced the second-best result. A MAE value of 14.38, a MAPE of 0.04%, an MSE of 367.59, an RMSE of 19.17, and an R^2 of 0.77 were obtained by the SMA method, which is the third-best result.
Keywords – Prediction, Moving Average, Exponential Smoothing, Global Horizontal Irradiance

I. INTRODUCTION

Solar radiation is one of the most influential weather parameters in the climate system, where all weather and climate phenomena are initially caused by variations in the distribution of solar radiation reception. Fluctuations in the intensity of solar radiation received at the earth's surface form climate patterns on various time scales. In other hand, solar radiation patterns also provide important information in various sectors, such as agriculture, water resources, and energy [1]. It is important to make models and analyzes based on the availability of energy sources. It depends on the production of solar energy in a certain area to have a good design and analysis of solar energy [2]. Solar power is an unlimited source of energy because solar energy is the largest energy on Earth. For tropical countries like Indonesia, sunlight is very easy to find. Indonesia's average daily insolation ranges from 4.5 – 4.8 kWh/m2/day, so it has the potential to develop solar energy into a renewable energy source [3].

The value of the intensity of the sun changes throughout the year. The different graphical patterns of solar radiation intensity can also be caused by several factors, the distance from the sun to the earth which has changed from the previous year, the influence of gases, dust and water vapor which absorb sunlight and the length of time the sun rises and sets [4]. In order to see the pattern of radiation data in the future, an effective and efficient method is needed. One of them is by making a predictive model based on existing data from observations of the BMKG (Badan Meteorologi Klimatologi dan Geofisika) station as well as using several mathematical equations and data that has been published by previous solar radiation researchers [5]. Prediction of solar radiation intensity is a necessity for the establishment of solar energy projects and for decision making in other related fields [6]. Prediction is a process that is estimated systematically based on past and present information that is owned, so that the difference between the actual data and the results of the predicted data can be minimized. Predictions do not have to give a definite answer to what will happen, but try to find answers as close as possible to what will happen.

In order to have a good analysis of solar radiation data patterns in the future, in this paper we create a solar energy prediction model based on solar radiation data. This study discusses testing solar energy models using the Simple Moving Average (SMA), Exponentially Weighted Moving Average (EWMA), and Single Exponential Smoothing (SES) methods. The data collected from global horizontal irradiance (GHI) in instrument measurements at the Climatology Station Special Region of Yogyakarta. Global horizontal irradiation (GHI) measurements by ground-based are usually applied using pyranometers for solar energy and atmospheric applications [7]. This research used monthly average data for 5 years from 2018 – 2022 that shown in Figure 1.



e-ISSN: 2614-8404
p-ISSN:2776-3234

Figure 1. Monthly Data of Global Horizontal Irradiance.

The prediction model consists of several methods, including Moving Averages and Exponential Smoothing. Moving Average is a method obtained by adding and finding the average value of the period, removing the oldest value every time and adding a new value [8]. Technical analysis is needed because it is easier to implement, requiring only historical GHI data. Moving Average is one part of the time series prediction method. The Moving Average method has two variants, the Simple Moving Average and the Weighted Moving Average. The difference between the two is in the weighting technique. Single Moving Average method that supports the calculation of time series data [9]. In the Exponentially Weighted Moving Average that a higher weight is given to the period closer to the predicted period [10]. Then, Exponential smoothing is a prediction method for recent observations. This prediction method focuses on deprioritizing older data, which means it pays more attention to the most recent observed values [8].

Previous research related to solar radiation data resulted in an evaluation of model performance in estimating the value of solar radiation intensity in the study area using two different modeling approaches, namely the empirical model by Keiser, Arkansas (AR) and the deterministic model. The three main weather variables used as model data input are rainfall (mm), maximum temperature (°C), and minimum temperature (°C). The results of testing have an R^2 value of 0.72 [11]. Similar research related to solar radiation data is to estimate the potential intensity of solar radiation as a renewable energy. The data used are solar radiation intensity data from Automatic Weather Station (AWS) from January 2012 to December 2015. Solar radiation intensity data were analyzed using non-linear regression methods, namely exponential methods, logarithmic methods, and power methods. The calculation results show that the exponential method gives a smaller error value than the other two methods. The error value on MAPE obtained ranges from 0.06% to 11.98% [12]. Those researches above has not yet arrived at a predictive model. Our research developed prediction model and modify the time series analysis method in predicting solar radiation data with only one parameter in order to further simplify computational performance. Then, the Coefficient of Determination (R^2) and Mean Absolute Percentage Error (MAPE) would be increased to the optimum work with the different model of Moving Average and Exponential Smoothing.

In another study, the prediction of motor vehicle test retribution data was carried out using the method Exponential Smoothing and Moving Average which aims to compare the effectiveness of the two methods in predicting motor vehicle testing fees. The measure of the effectiveness of the method is seen based on the MSE and MAPE values [8]. In our research shown prediction model in GHI data use single exponential smoothing and various model of Moving Average, there are Simple Moving Average and Exponentially Weighted Moving Average. Then those model would be evaluated with error and correlation value using Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Coefficient of Determination (R^2) values.

Those model is implemented by Python to perform time series statistical data processing. As the final result, this study obtains a statistical model for predicting GHI compared to actual data with the lowest error value and the result have the good correlation between actual data and predicted data. It is hoped that this prediction model can be applied directly in predicting the availability of solar energy for planners of solar energy equipment and planners in fields including: climatology, hydrology, agriculture, and architecture.

II. RESEARCH METHODOLOGY

In the research methodology explained regarding the important steps carried out to carry out comparative methods in predicting GHI data, including data collection, data preprocessing, prediction model testing of dataset with SMA, EWMA, and SES by Python, interpreting the visualization of prediction model result, evaluating the performance of prediction models with MAE, MAPE, MSE, RMSE, and R^2. Finally, we have prediction models of time series data in Global Horizontal Irradiation. The flowchart of this research shown in Figure 2.

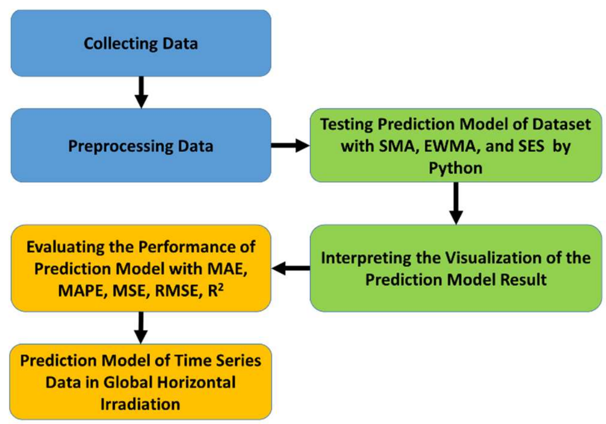


Figure 2. Flowchart of prediction model testing

After knowing the design work flow of this research, below will explain in more detail the processes involved in conducting data processing and data analysis on GHI predictions. Each process also explains the theoretical basis used to support the research ideas being carried out.

A. Collecting Data

The Automatic Solar Radiation System in Special Region of Yogyakarta, provided the pyranometer instrumentation for the ground measurements used to collect the GHI data, as illustrated in Figure 3.

e-ISSN: 2614-8404
p-ISSN:2776-3234

Figure 3. Measurement of GHI using a Pyranometer

The GHI data used are daily data recorded have ten minutes' interval from January 2018 to December 2022 at local time. This data can be processed to create a dataset that will be used to evaluate a time series prediction model.

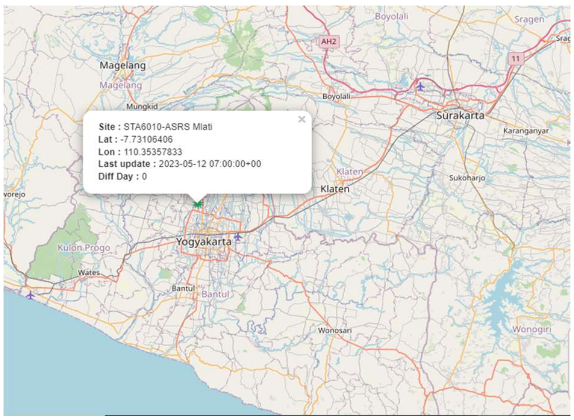


Figure 4. The location maps of GHI measurement in Climatology Station of Yogyakarta

Measurement GHI is product of Automatic Solar Radiation System (ASRS) that operated by Climatology Station Yogyakarta, Badan Meteorologi Klimatologi dan Geofisika (BMKG) located at Mlati, Sleman, Special Region of Yogyakarta. The location maps of this station shown in Figure 4.

B. Preprocessing Data

Preprocessing data is carried out quantitatively as data values in the form of counts or numbers where each dataset has a numerical value associated with data obtained from measurement sensors. This data is quantifiable information that can be used for mathematical calculations and statistical analysis.

Table 1. Raw Data of GHI Measurements

IDStation	Tanggal	Jam	GHI
STA6010	1/1/2019	11:00:00	739.3
STA6010	1/1/2019	11:10:00	483.4
STA6010	1/1/2019	11:20:00	464.6
STA6010	1/1/2019	11:30:00	397.3
STA6010	1/1/2019	11:40:00	751.8
STA6010	1/1/2019	11:50:00	1210.1
STA6010	1/1/2019	12:00:00	692.1

Table 1 depicts the unprocessed measuring equipment data, which included up to 4000 lines for each mount. The raw data is in local time, and we only use the values of the GHI parameters expressed in W/m^2. The next step is to filter the raw data, which must only include data from six in the morning to six in the evening. This time frame during which the GHI value could be obtained. Finally, the processing of daily data into monthly data was completed. The data output from preprocessing the data would be utilized as input for further calculations to choose the optimal prediction model.

Table 2. Monthly GHI Data 2018-2022

GHI	2018	2019	2020	2021	2022
January	347.64	410.40	407.93	370.00	431.06
February	420.80	398.03	388.17	370.49	374.98
March	432.79	376.63	359.59	411.39	384.77
April	409.70	390.90	369.52	400.80	390.22
May	414.34	395.03	336.87	368.70	368.70
June	410.12	379.18	334.35	325.25	330.71
July	404.90	335.35	348.53	382.38	354.03
August	407.49	472.53	390.79	371.64	364.28
September	423.65	510.87	399.49	413.40	379.37
October	461.95	498.74	332.44	389.10	316.87
November	394.04	438.44	387.30	324.75	321.11
December	394.04	402.64	400.73	399.48	365.67

Table 2 shows the results of the data preprocessing in the form of monthly GHI data. Monthly GHI data is obtained from 2018 to 2022. The time series data will be processed using Python for further analysis by visualization using data graphics and data statistical processing. Then using the SMA, EWMA, and SES methods to form the best predictive model.

C. Testing of Prediction Model

Testing using the SMA, EWMA and SES algorithms is a system development methodology after data preprocessing is carried out in analyzing data processing quantitatively. Quantitative data processing techniques use mathematical and statistical equations then are assisted by Python to calculate the algorithms of prediction model. The following are theoretical explanations of SMA, EWMA and SES.

Simple Moving Average (SMA) is the simplest moving average and does not use a value in predicting calculations. Although the SMA is quite effective in determining the current trend [13]. Despite its simplicity, SMA is very good in identifying the current trend [14]. This Equation (1) is the how to calculate the SMA.

$$Pred_{t+1} = \frac{Act_t + Act_{t-1} + \ldots\ Act_{t-n+1}}{n} \quad (1)$$

$Pred_{t+1}$ is the prediction for period $t+1$. x_t is the data for period t. Act_t is the data real from observation for period t. Then, n is the period of moving averages. In the

e-ISSN: 2614-8404
p-ISSN:2776-3234

SMA this is calculated by taking the average value of the GHI data observation values at two month of span back.

Exponentially Weighted Moving Average (EWMA) is one approach to dealing with unstable data volatility. This method gives weight to changes in value each period using the decay factor (α). The parameter α shows the weight scale of the latest data observations with previous data with a value of 0 < α < 1. The higher the α value, the greater the weight that will be applied to past data so that the time series data is smoother [15]. To calculate EWMA, this formula shown in Equation 2.

$$y_t = \frac{\sum_{i=1}^{n} w_i * x_{t-i+1}}{\sum_{i=1}^{n} w_i} \quad (2)$$

y_t is the weighted moving average at time t. And x_t is the value at time t. w_i is the weight assigned to each value. n means the number of values used to calculate the moving average. This function is calculated using weights $w_i = (1-\alpha)^2$ would be the new formula in Equation (3).

$$y_t = \frac{x_t + (1-\alpha)x_{t-1} + (1-\alpha)^2 x_{t-2} + \ldots + (1-\alpha)^t x_0}{1 + (1-\alpha) + (1-\alpha)^2 + \ldots + (1-\alpha)^t} \quad (3)$$

The Single Exponential Smoothing (SES) method is used for short-term predictions. The model assumes that the data fluctuates around a fixed mean value, without a consistent trend or pattern of growth. The Single Exponential Smoothing method takes into account the weight of the previous data by giving weight to each data period to distinguish the priority of a data [16]. The formula for SES is as follows [17] in Equation (4).

$$Pred_{t+1} = \alpha * Act_t + (1-\alpha) * Pred_{t-1} \quad (4)$$

$Pred_{t+1}$ is a new prediction for period to $t+1$. Act_t describes the observation value of t period. α is a weight value indicating smoothing constant ($0 < \alpha < 1$). $Pred_{t-1}$ is a prediction for period $t-1$. Meanwhile, to calculate the alpha value (α) using the formula $\alpha = 2/(n+1)$. Where α is the weight value indicating smoothing constant ($0 < \alpha < 1$). Then, n is the number of time periods.

D. Interpreting of Visualization Data

Data interpretation is a stage carried out with the aim of associating the relationship between various research variables. In this study, the data was processed using a time series chart to find out the pattern of the data. After analyzing several prediction methods, the results are shown using a graph that displays the data prediction pattern of each method. This helps in identifying data patterns resulting in determining the best model.

E. Evaluating Performance.

The model of prediction need to evaluate with the values of MAE, MAPE, MSE, RMSE, and R^2. The evaluation with the error and correlation values describe about the effectiveness the model work for this data case. It is important to measure the best model needed. The following are theoretical explanations of MAE, MAPE, MSE, RMSE and R^2.

Mean Absolute Error (MAE) is one of the methods used to measure the accuracy of the prediction model. The MAE value shows the average absolute error between the prediction ($Pred$) results and the actual (Act) value. Then n is the number of periods used for calculations The MAE formula is explained as follows in Equation (5).

$$MAE = \frac{1}{n} \sum_{t=1}^{n} |Act_t - Pred_t| \quad (5)$$

Mean Absolute Percentage Error (MAPE) is a measure of the accuracy of $Prediction$ result compared to the $Actual$ result. This MAPE measurement is usually used to measure the accuracy of the predicted value of the time series. Result of MAPE measurements are generally in the form of a percentage. The smaller the percentage value generated by MAPE, the better the prediction results [18]. This formula shown in Equation (6).

$$MAPE = \frac{\sum_{t=1}^{n} \left| \left(\frac{Act_t - Pred_t}{Actual_t} \right) x10 \right|}{n} \quad (6)$$

Mean Square Error (MSE) is a calculation used to calculate the average power of errors in GHI data. From this equation it can be interpreted that $\sum_{t=1}^{n}(Act - Pred)^2$ is the result of subtraction between the actual and prediction values that have been squared, then the results are summed. And n is the number of periods used for calculations [19]. To calculate the MSE value, this Equation (7) shown below.

$$MSE = \frac{\sum_{t=1}^{n}(Act_t - Pred_t)^2}{n} \quad (7)$$

Root Mean Square Error (RMSE) is calculated by calculating the root of the sum of all the squared prediction errors divided by the number of prediction time data. The smaller the RMSE value, the more accurate the model is [20]. The RMSE determination was carried out to determine the best GHI prediction model. To calculate the RMSE value, the following Equation (8) shown below.

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n}(Act - Pred)^2}{n}} \quad (8)$$

The coefficient of determination (R^2) is used to determine the proportion of the influence of all variables independent of the dependent variable [21]. The greater the R^2 of the independent variable, the more dominant the influence of the dependent variable is [22]. The formula of R^2 shown in Equation (9).

$$R^2 = 1 - \frac{\sum_{i=1}^{m}(Pred_i - Act_i)^2}{\sum_{i=1}^{m}(\overline{Act} - Act_i)^2} \quad (9)$$

$Pred_i$ is the prediction value to i , and the $Actual_i$ is the observation value to i. Then, $\overline{Act}$ is the mean of all observation data collected in dataset. The method predicts the Act_i for the corresponding $Pred_i$ of the dataset of GHI.

III. Results And Discussion

This part provides a comprehensive application and discussion of the research findings. Results are presented in tables, graphs, and figures. The analysis explained detail.

A. Testing of Prediction Model

The calculation of SMA using two span data executed by Python and convert to the table that consist of some attributes including period, GHI, and SMA. Table 3 shown the result of SMA calculation.

Table 3. Simple Moving Average Result

Period	GHI	SMA
2018-01-01	347.639895	NaN
2018-02-01	420.797452	384.218674
2018-03-01	432.787111	426.792282
2018-04-01	409.699329	421.243220
2018-05-01	414.344918	412.022123
2018-06-01	410.121655	412.233286
2018-07-01	404.904114	407.512885
2018-08-01	407.490266	406.197190
2018-09-01	423.651089	415.570677
2018-10-01	461.945614	442.798351
2018-11-01	394.039211	427.992413
2018-12-01	394.039211	394.039211
2019-01-01	410.402197	402.220704
2019-02-01	398.030920	404.216559
2019-03-01	376.634749	387.332834
2019-04-01	390.901592	383.768170
2019-05-01	395.028568	392.965080
2019-06-01	379.183653	387.106110
2019-07-01	335.354438	357.269046
2019-08-01	472.530611	403.942524
2019-09-01	510.872491	491.701551
2019-10-01	498.741799	504.807145
2019-11-01	438.436004	468.588902
2019-12-01	402.644495	420.540250
2020-01-01	407.930904	405.287700
2020-02-01	388.174581	398.052743
2020-03-01	359.593566	373.884073
2020-04-01	369.520977	364.557272
2020-05-01	336.874319	353.197648
2020-06-01	334.351108	335.612714
2020-07-01	348.525823	341.438466
2020-08-01	390.786970	369.656397
2020-09-01	399.492266	395.139618
2020-10-01	332.436617	365.964442
2020-11-01	387.296813	359.866715
2020-12-01	400.729336	394.013075
2021-01-01	370.003121	385.366229
2021-02-01	370.491812	370.247466
2021-03-01	411.388600	390.940206
2021-04-01	400.798169	406.093384
2021-05-01	368.697186	384.747678
2021-06-01	325.250182	346.973684
2021-07-01	382.384502	353.817342
2021-08-01	371.635482	377.009992
2021-09-01	413.398784	392.517133
2021-10-01	389.095656	401.247220
2021-11-01	324.754933	356.925295
2021-12-01	399.483070	362.119001
2022-01-01	431.063684	415.273377
2022-02-01	374.975540	403.019612
2022-03-01	384.766990	379.871265
2022-04-01	390.218336	387.492663
2022-05-01	368.697186	379.457761
2022-06-01	330.708146	349.702666
2022-07-01	354.025390	342.366768
2022-08-01	364.279495	359.152443
2022-09-01	379.365520	371.822508
2022-10-01	316.873755	348.119638
2022-11-01	321.114181	318.993968
2022-12-01	365.668471	343.391326

Calculation Data in The calculation of EWMA using two span data executed by Python and convert to the table that consist of some attributes including period, GHI, and EWMA. Table 4 shown the result of EWMA calculation.

Table 4. Exponentially Weighted Moving Average Result

Period	GHI	EWMA
2018-01-01	347.639895	347.639895
2018-02-01	420.797452	402.508063
2018-03-01	432.787111	423.470481
2018-04-01	409.699329	414.174953
2018-05-01	414.344918	414.288731
2018-06-01	410.121655	411.506865
2018-07-01	404.904114	407.103017
2018-08-01	407.490266	407.361222
2018-09-01	423.651089	418.221685
2018-10-01	461.945614	447.371465
2018-11-01	394.039211	411.816428
2018-12-01	394.039211	399.964928
2019-01-01	410.402197	406.923112
2019-02-01	398.030920	400.994983
2019-03-01	376.634749	384.754826
2019-04-01	390.901592	388.852670
2019-05-01	395.028568	392.969935
2019-06-01	379.183653	383.779080
2019-07-01	335.354438	351.495985
2019-08-01	472.530611	432.185736
2019-09-01	510.872491	484.643573
2019-10-01	498.741799	494.042390
2019-11-01	438.436004	456.971466
2019-12-01	402.644495	420.753485
2020-01-01	407.930904	412.205098
2020-02-01	388.174581	396.184753
2020-03-01	359.593566	371.790628
2020-04-01	369.520977	370.277527
2020-05-01	336.874319	348.008722
2020-06-01	334.351108	338.903646
2020-07-01	348.525823	345.318431
2020-08-01	390.786970	375.630790
2020-09-01	399.492266	391.538441
2020-10-01	332.436617	352.137225
2020-11-01	387.296813	375.576950
2020-12-01	400.729336	392.345207
2021-01-01	370.003121	377.450483
2021-02-01	370.491812	372.811369
2021-03-01	411.388600	398.529523
2021-04-01	400.798169	400.041954
2021-05-01	368.697186	379.145442
2021-06-01	325.250182	343.215269
2021-07-01	382.384502	369.328091
2021-08-01	371.635482	370.866352
2021-09-01	413.398784	399.221307
2021-10-01	389.095656	392.470873
2021-11-01	324.754933	347.326913
2021-12-01	399.483070	382.097684
2022-01-01	431.063684	414.741684
2022-02-01	374.975540	388.230921
2022-03-01	384.766990	385.921634
2022-04-01	390.218336	388.786102
2022-05-01	368.697186	375.393491
2022-06-01	330.708146	345.603261
2022-07-01	354.025390	351.218014
2022-08-01	364.279495	359.925668
2022-09-01	379.365520	372.885569
2022-10-01	316.873755	335.544360
2022-11-01	321.114181	325.924241
2022-12-01	365.668471	352.420394

The calculation of SES using two span data and 0.75 alpha value executed by Python and convert to the table that consist of some attributes including period, GHI, and SES. Table 5 shown the result of SMA calculation.

Table 5 Single Exponential Smoothing Result

Period	GHI	SES
2018-01-01	347.639895	347.639895
2018-02-01	420.797452	402.508063
2018-03-01	432.787111	425.217349
2018-04-01	409.699329	413.578834
2018-05-01	414.344918	414.153397
2018-06-01	410.121655	411.129591
2018-07-01	404.904114	406.460483
2018-08-01	407.490266	407.232820
2018-09-01	423.651089	419.546522
2018-10-01	461.945614	451.345841
2018-11-01	394.039211	408.365868
2018-12-01	394.039211	397.620875
2019-01-01	410.402197	407.206867
2019-02-01	398.030920	400.324907
2019-03-01	376.634749	382.557288
2019-04-01	390.901592	388.815516
2019-05-01	395.028568	393.475305
2019-06-01	379.183653	382.756566
2019-07-01	335.354438	347.204970
2019-08-01	472.530611	441.199201
2019-09-01	510.872491	493.454168
2019-10-01	498.741799	497.419891
2019-11-01	438.436004	453.181976
2019-12-01	402.644495	415.278865
2020-01-01	407.930904	409.767894
2020-02-01	388.174581	393.572909
2020-03-01	359.593566	368.088402
2020-04-01	369.520977	369.162833
2020-05-01	336.874319	344.946448
2020-06-01	334.351108	336.999943
2020-07-01	348.525823	345.644353
2020-08-01	390.786970	379.501316
2020-09-01	399.492266	394.494528
2020-10-01	332.436617	347.951095
2020-11-01	387.296813	377.460383
2020-12-01	400.729336	394.912098
2021-01-01	370.003121	376.230365
2021-02-01	370.491812	371.926450
2021-03-01	411.388600	401.523063
2021-04-01	400.798169	400.979392
2021-05-01	368.697186	376.767738
2021-06-01	325.250182	338.129571
2021-07-01	382.384502	371.320769
2021-08-01	371.635482	371.556804
2021-09-01	413.398784	402.938289
2021-10-01	389.095656	392.556314
2021-11-01	324.754933	341.705278
2021-12-01	399.483070	385.038622
2022-01-01	431.063684	419.557419
2022-02-01	374.975540	386.121010
2022-03-01	384.766990	385.105495
2022-04-01	390.218336	388.940126
2022-05-01	368.697186	373.757921
2022-06-01	330.708146	341.470590
2022-07-01	354.025390	350.886690
2022-08-01	364.279495	360.931294
2022-09-01	379.365520	374.756963
2022-10-01	316.873755	331.344557
2022-11-01	321.114181	323.671775
2022-12-01	365.668471	NaN

B. Interpreting of Visualization Data

Data visualization on GHI time series data can be done periodically. Figure 5 shown the graphical pattern of GHI data from 2018-2022.

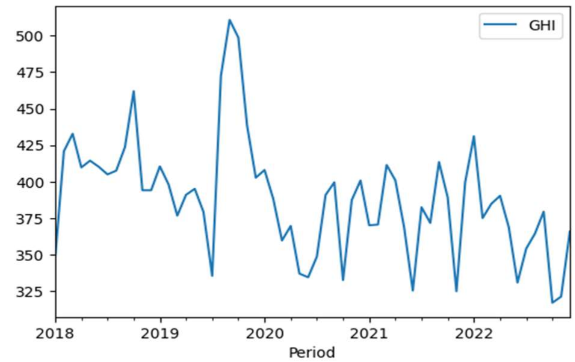


Figure 5. The Graphical Pattern of GHI Data 2018-2022

Patterns can be observed visually where the data is lowest and highest. Furthermore, the range of values can be observed through the graph.

The results of SMA, EWMA, and SES calculations are visualized in graphs. The graph can explain the comparison between the SMA, EWMA, and SES prediction models visually and makes it easier to see the pattern of predictions made. Figure 6 shows a comparison graph of the SMA, EWMA, and SES prediction models.

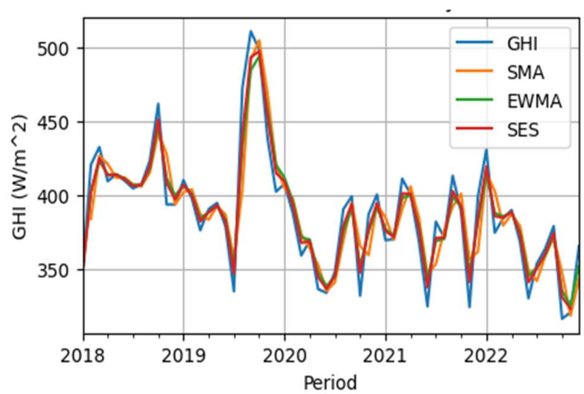


Figure 6. Comparison of SMA, EWMA, and SES Model

From the comparison of models based on graphs, it will be explained in more detail in the next section related to the evaluation values in the form of errors and the correlation of each prediction model with actual observation data.

C. Evaluating Performance.

In this study, evaluation of error values was used in the form of MAE, MAPE, MSE, RMSE, and evaluation of correlation values using R2. The Table 6 shows the results of a comparison of the performance evaluation of the SMA, EWMA, and SES prediction models.

Table 6. Comparison the Evaluation of SMA, EWMA, and SES Model

METHODS	MAE	MAPE	MSE	RMSE	R^2
SMA	14.38	0.04%	367.59	19.17	0.77
EWMA	9.45	0.02%	150.16	12.25	0.91
SES	7.13	0.02%	88.07	9.38	0.94

D. Prediction Model of GHI

Application of the SMA algorithm. EWMA and SES are comparative analysis of performance tests on predicted solar radiation (GHI) data. The test results show that SMA, EWMA and SES have evaluation values in the form of error and correlation values. The range of each value can be considered to make decisions with the best model of prediction. The tested models could improve the methods of previous research in estimation of solar radiation with some other parameters that result of testing has an R^2 value of 0.72. In our research used only single parameter to support the more effective computation and our research produced better correlation values for prediction obtained 0.77 in SMA, 0.91 in EWMA, and 0.94 in SES. Then, our research also could improve the MAPE value in other previous research about estimate the potential intensity of solar radiation as a renewable energy. The error value on MAPE obtained ranges from 0.06% to 11.98%. In our research could modify the data solar radiation (GHI) into time series prediction model that produced the better MAPE value obtained 0.04% in SMA, 0.02% in EWMA, and 0.02% in SES. Based on the evaluation of model prediction, the SES model could work better than EWMA and also EWMA could work better than SMA in GHI prediction data.

IV. CONCLUSION

The results of this study indicate a comparative analysis of the performance levels of the SMA, EWMA, and SES methods in predicting GHI data. Evaluation of the GHI data prediction method can be carried out based on the calculation of error values and correlations in MAE, MAPE, MSE, RMSE, and R2. The results of the calculation of the prediction model performance evaluation produced the best prediction model is the SES method that obtained an MAE value of 7.13, a MAPE of 0.02%, an MSE of 88.07, an RMSE of 9.38, and an R2 of 0.94. The second best result is the EWMA method that obtained an MAE value of 9.45, a MAPE of 0.02%, an MSE of 150.16, an RMSE of 12.25, and an R2 of 0.91. The third best result is the SMA method that obtained an MAE value of 14.38, a MAPE of 0.04%, an MSE of 367.59, an RMSE of 19.17, and an R2 of 0.77. In this study using data and methods that are still limited, therefore suggestions for further research can use a more diverse method and increase the amount of data to improve higher predictive performance.

ACKNOWLEDGMENT

In this research there are certainly challenges and obstacles. Therefore, the researcher says thank you to the lecturers of the Data Mining and Analysis course, Master of Information Technology, Yogyakarta University of Technology, Mr. Arif Hermawan and Mr. Donny Avianto who have provided scientific, material and motivational support for this research and all parties who have assisted in this research.

REFERENCES

[1] Y. Sianturi and C. M. Simbolon, "Pengukuran dan Analisa Data Radiasi Matahari di Stasiun Klimatologi Muaro Jambi," *Megasains,* vol. 12 No. 1, pp. 40-47, 2021. https://doi.org/10.46824/megasains.v12i1.45.

[2] I. Morokia and A. S. Wauran, "Model Statistika Prediksi Energi Sur ya Dengan Menggunakan Autoregresif Integrated Moving Average (ARIMA)," *JURNAL MIPA ,* vol. 8 No. 3, pp. 181-183, 2019. https://doi.org/10.35799/jmuo.8.3.2019.26193.

[3] J. S. Setyono, F. H. Mardiansjah and M. F. K. Astuti, "Potensi Pengembangan Energi Baru dan Energi Terbarukan di Kota Semarang," *Jurnal Riptek,* vol. 13 No. 2, p. 177–186, 2019. https://doi.org/10.35475/riptek.v13i2.68.

[4] A. Octavianti, Muliadia and Apriansyah, "Estimasi Intensitas Radiasi Matahari di Wilayah Kota Makassar," *PRISMA FISIKA,* vol. 6 No. 3, pp. 152-159, 2018. http://dx.doi.org/10.26418/pf.v6i3.28711.

[5] Y. S. Utomo, "Prediksi Radiasi Surya Global Bulanan Kota Bandung Menggunakan Data LPM (Lama Penyinaran Matahari)," *Jurnal Material dan Energi Indonesia,* vol. 07 No. 2, p. 21–27, 2017. https://doi.org/10.24198/jmei.v7i02.15534.g7501.

[6] A. A. Teyabeen, N. B. Elhatmi, A. A. Essnid and F.Mohamed, "Estimation of monthly global solar radiation over twelve major cities of Libya," *Energy and Built Environment,* pp. 1-12, 2022. https://doi.org/10.1016/j.enbenv.2022.07.006.

[7] M. C. Brito, S. R. A. e and M. Po, "Geographical distribution of the angle of incidence uncertainty on the measurement of global horizontal irradiance," *Measurement: Sensors,* vol. 22 no. 100382, pp. 1-8, 2022. https://doi.org/10.1016/j.measen.2022.100382.

[8] E. Setyowati, "Perbandingan Metode Exponential Smoothing dan Moving Average dalam Peramalan Retribusi Pengujian Kendaraan Bermotor di Dinas Perhubungan Kota Blitar," *Jurnal Sains Dasar,* vol. 11 no. 1, pp. 35-38, 2022. doi 10.21831/jsd.v11i1.44391.

[9] A. Nurfadilah, W. Budi, E. Kurniati and D. Suhaedi, "Penerapan Metode Moving Average untuk Prediksi Indeks Harga Konsumen," *Jurnal Matematika ,* vol. 21 no. 1 , pp. 19-25, 2022.

[10] E. A. Gultom, N. Eltivia and N. I. Riwajanti, "Shares Price Forecasting using Simple Moving Average Method and Web Scrapping," *Journal of Applied Business,Taxation and Economics Research (JABTER),* vol. 2 no. 3, pp. 288-297, 2023. doi:10.54408/jabter.v2i3.164.

[11] Munawar, A. Mulsandi and A. M. Hidayat, "Model Estimasi Data Intensitas Radiasi Matahari untuk Wilayah Banten," *Jurnal Sains & Teknologi Modifikasi Cuaca,* vol. 21 no. 2 , pp. 53-61, 2020. https://doi.org/10.29122/jstmc.v21i2.4171.

[12] A. Octaviantia, Muliadia and Apriansyah, "Estimasi Intensitas Radiasi Matahari di Wilayah Kota Makassar," *PRISMA FISIKA,* vol. 6 no. 3, pp. 152-159, 2018. http://dx.doi.org/10.26418/pf.v6i3.28711.

[13] N. L. Saria and T. Hasanuddina, "Analisis Performa Metode Moving AverageModel untuk Prediksi Jumlah Penderita Covid-19," *Indonesian Journal of Data and Science(IJODAS),* vol. 1 no. 3, pp. 87-95, 2020. https://doi.org/10.33096/ijodas.v1i3.19.

[14] R. Ramadhan, R. Yusman and G. Pranoto, "Comparison of Holt Winters and Simple Moving Average Models to Identify The Best Model for Predicting Flood Potential Based on The Normalized Difference Water Index," *JISA (Jurnal Informatika dan Sains),* vol. 5 no. 2, pp. 99-105, 2022.

[15] H. D. Wulandari, Mustafid and H. Yasin, "Penerapan Metode Exponentially Weighted Moving Average (EWMA) dalam Pengukuran Risiko Inevstasi Saham Portofolio untuk Volatilitas Heterogen," *JURNAL GAUSSIAN,* vol. 7 no. 3, pp. 248-259, 2018. https://doi.org/10.14710/j.gauss.7.3.248-259.

[16] R. Yuniarti, "Analisa Metode Single Exponential Smoothing Sebagai Peramalan Penjualan Terhadap Penyalur Makanan (Studi Kasus : Lokatara Dimsum)," *Jurnal Manajemen & Bisnis Aliansi,* vol. 15 no. 2, pp. 29-34, 2020. https://doi.org/10.46975/aliansi.v15i2.63.

[17] Risqiati, "Penerapan Metode Single Exponential Smoothing dalam Peramalan Penjualan Benang," *Smart Comp ,* vol. 10 no. 3 , pp. 154-159, 2021. http://dx.doi.org/10.30591/smartcomp.v10i3.2887.

[18] D. Gunawan and W. J. Kurniawan, "Perancangan Sistem Informasi Purchase Order Menggunakan Metode Single Exponential Smoothing," *Jurnal Mahasiswa Aplikasi Teknologi Komputer dan Informasi,* vol. 2 no. 1, pp. 13-18, 2020.

[19] M. A. Maricar, "Analisa Perbandingan Nilai Akurasi Moving Average dan Exponential Smoothing untuk Sistem Peramalan Pendapatan pada Perusahaan XYZ," *JURNAL SISTEM DAN INFORMATIKA,* vol. 13 no. 2, pp. 36-45, 2019.

[20] E. K. Hendarwati, P. Lepong and Suyitno, "Pemilihan Semivariogram Terbaik Berdasarkan Root Mean Square Error(Rmse) Pada Data Spasial Eksplorasi Emas Awak Mas," *Jurnal Geosains Kutai Basin ,* vol. 6 no. 1, pp. 1-6, 2023.

[21] D. Setiawan, E. Shafira, K. Ukar, D. B. Ginting and A. Rini, "Analisis Faktor-Faktor yang Mempengaruhi Perilaku Pengguna Aplikasi Tokopedia," *Media Informatika,* vol. 21 no.1 , pp. 75-93, 2022.

[22] C. A. Rahmat, Kurniabudi and Y. Novianto, "Penerapan Metode Regresi Linier BergandaUntuk Mengestimasi Laju Pertumbuhan Penduduk Kabupaten Musi Banyuasin," *Jurnal Informatika dan Rekayasa Komputer (JAKAKOM),* vol. 3 no. 1, pp. 359-369, 2023.