Fourier Series Approach for Bias Correction in Statistical Downscaling Methods

Tiani Wahyu Utami[1,*] Asrori Gufron[2] Fatkhurokhman Fauzi[3]
[1,2,3]Universitas Muhammadiyah Semarang, Semarang, Central Java 50273, Indonesia
tianiutami@unimus.ac.id

Abstract. Earth System Models (ESM) are model that can forecast, and simulate past, present, and future situations including climate change. The local climate has not been well represented by ESM results. The Statistical Downscaling (SD) approach is one attempt to solve this issue. Climate studies in high-latitude nations have made extensive use of Statistical Downscaling (SD) methodologies; nevertheless, there are still relatively few of these studies conducted in low-latitude regions (the Tropics, including Indonesia). We need a technique that works to lessen bias because the SD findings still contain a significant amount of bias. Fourier series is the bias reduction technique applied in this study. Fourier series is the bias reduction technique applied in this study. For the ESM RCP 4.5 scenario, this study corrects and downscales the bias of the relative humidity and temperature data. The analysis's conclusions, which are explained by SD, show that Merra-2 (local) dependant data has an impact on the RCP 4.5 downscaling process and that the graph becomes closer to Merra-2 data. The model that was produced was superior since the research for bias correction using the Fourier Series approach for temperature yielded 98% with MSE 0.0290 and for relative humidity yielded 97% with MSE 0.3223. From 2006 to 2057, Indonesia's Temperature Humidity Index (THI) fell within the acceptable range.

Keywords: Bias Correction, Fourier Series, Statistical Downscaling, Temperature Humidity Index (THI).

1. Introduction

Over time, the world's climate has been damaged by human activities. This occurs due to an increase in the greenhouse effect, meaning there is an increase in the concentration of gases that block the reflection of sunlight energy from the earth. Therefore, the planet Earth we live on is getting hotter [1]. The country of Indonesia is located at 6º LU – 11º LS and 95º E - 141º E is a country located on the equator latitude between the Pacific Ocean and the Indian Ocean which causes Indonesia to only have two seasons. Climate change in Indonesia is not only impacting the country's environment, but also profoundly affecting society and its development. Climate change in Indonesia has received world attention, this matter was discussed at the United Nations Climate Change Conference in Bali in

I. Yustar Afif and R. Nindyo Sumarno (eds.), *Proceedings of the 2nd Lawang Sewu International Symposium on Engineering and Applied Sciences (LEWIS-EAS 2023)*, Advances in Engineering Research 234,
https://doi.org/10.2991/978-94-6463-480-8_14

December 2007. Participants at the conference discussed solutions to climate change problems in Indonesia and the world which were attended by 10,000 participants [2].

Changes in air temperature have an impact on the level of human comfort, such as the relative humidity (RH) that changes in Indonesia. In Indonesia, the comfortable temperature is 28.5 + 1.5 °C Ta and the relative humidity (RH) is 60% [3]. The comfort level Temperature Humidity Index (THI) in this case also influences climate change in Indonesia because it has an impact on human activities. THI is comfort index that combines air temperature and wet bulb temperature factors which are then modified by combining air temperature and RH [4]. Increasing air humidity and air temperature causes a decrease in comfort in an area, and productivity. A model is needed that can simulate climate, predict past and present climate changes, and create future climate change scenarios by involving large-scale data information such as Earth System Models (ESM) [5].

There is a technique that is often used in the world of climatology to reduce the spatial scale to the regional level, namely Statistical Downscaling. Statistical downscaling is a method that can be used to determine regional scales, one of the statistical downscaling methods is climate-imprint. Statistical Downscaling method is an efficient and effective method, one of which is the climate-imprint method. Several studies on statistical downscaling include [6] conducted statistical downscaling research to predict monthly rainfall at Sembalun Station using a nonparametric kernel regression approach which provided relatively similar rainfall prediction patterns, research conducted by [7] regarding bias correction / Constructed Analogues with Quantile Mapping (BCCAQ) is a hybrid downscaling method that uses output from CA and quantile mapping at high scale resolution. Furthermore, research [8] using hybrid model Fourier Series-Polynomial Local to reduce the bias correction at Statistical Downscaling method.

Statistical Downscaling still produces results that are less than optimal. To reduce them, a bias correction is needed with the aim of reducing the bias of the downscaling results so that they are more representative of the local climate. To reduce the results of bias correction in order to minimize the results of bias, nonparametric regression approach with the Fourier series approach is needed. The advantage of the Fourier series nonparametric regression approach is that it is able to handle data with functions that have a trigonometric cosine distribution [9], therefore statistical downscaling with Fourier Series nonparametric regression can explain the temporal variability and extreme values of climate data.

2. Literature Review

2.1. Statistical Downscaling

Statistical Downscaling is defined as an effort to link global scale variables (explanatory variables) and local scale variables (response variables). Downscaling is based on the assumption that local climate is affected by the climate of the earth and continents [10].

Regional climate interacts between the atmosphere such as terrain, oceans, specific (topical) circulation, vegetation and land use distribution. Local scale models have high resolution with a grid size smaller than the grid of the ESM which can take into account local topography, vegetation and soil types, and translate the ESM prediction results at the local scale. There are two types of downscaling approaches, namely Dynamical Downscaling and Statistical Downscaling. Dynamical downscaling is carried out by assigning ESM to a higher spatial resolution, while SD is based on the functional relationship between large-scale predictors and small-scale response variables [11].

2.2. Series Fourier

Fourier Series function is as follows: Fourier series is an infinite series with terms containing trigonometric components, sine-cosine, which converge to a periodic function. Fourier Series function is as follows:

$$f(t) = \frac{1}{2}a_0 + yt + \sum_{k=1}^{k} a_k\, cos\left(\frac{2\pi kt}{2L}\right) + \hat{b}_k sin\left(\frac{2\pi kt}{2L}\right)$$

If $T_i \in [-L, L]$ dan $Y_i \in R$ and it is assumed that $f(t)$ uses the Fourier Series approximation defined:

$$f(t) = \frac{1}{2}a_0 + yt + \sum_{k=1}^{k} a_k\, cos\left(\frac{2\pi kt}{2L}\right) \tag{1}$$

With a_o, a_k and b_k are Fourier coefficients. The Fourier Series estimator is defined by the choice of the smoothing parameter K. The smaller the smoothing parameter K, the less smooth the estimate of f. Therefore, it is necessary to choose the optimal [12].

2.3. Temperature Humidity Index

The human comfort index is determined from the heat energy received by humans in carrying out all activities. Temperature Humidity Index (THI) is a comfort index that combines air temperature and wet bulb temperature factors. THI is part of the human comfort indicators which focus on temperature and humidity relatively [4].

Temperature is a climate parameter that greatly influences the comfort felt by humans. As the air temperature increases beyond a certain limit, the comfort felt by humans will decrease because the sensation of heat from the air temperature will be too great for humans to feel. On the other hand, as the temperature decreases to a certain limit, the comfort felt by humans will also decrease because it is too cold [13]. The comfort index can be determined using the THI method calculation which connects temperature and relative humidity conditions developed by Nieuwolt in the form of the following equation [8]:

$$THI = (0{,}8T) + \frac{RH \times T}{500} \tag{2}$$

T: Temperature (°C), RH: Relative Humidity (%).

The comfort index value used to determine the comfort category is obtained by correlating human respondents' assessments to obtain the following range:

- $21 \leq THI \leq 24$=100%of respondents feel comfortable
- $24 < THI \leq 27$=50% of respondents feel comfortable
- $THI > 27$ = 0% of respondents feel comfortable

The THI value for determining human comfort is derived from human physiology which is connected to the environmental conditions around the human [14].

3. Materials and Methods

3.1. Data Source

The data used in this study is secondary data which consists of two data, namely reanalysis data and scenario data.

1. Reanalysis data were obtained from the Global Earth Sciences Data and Information Services Center (GES DISC), which is MERRA-2 data with website https://disc.gsfc.nasa.gov/datasets/M2SDNXSLV_5.12.4/summary.
2. Whereas the scenario data was obtained from the Federation Earth System Grid (ESGF) with the website https://esgfnode.lln.gov/search/esgf-llnl/ with the CSIRO-Mk3.6.0 model in the form of the RCP4 scenario

3.2. Data Variables

1. The CSIRO-Mk3.6.0 RCP4.5 scenario data variable in the form of daily average temperature data and daily Relative Humidity with coordinates of 6.528°-(-12.124°) S and 93.75°-142.5°E. This data has a coordinate grid of 1.875° × 1.875°.
2. MERRA-2 reanalysis data variables in the form of daily average temperature and daily Relative Humidity data in Indonesia with the coordinates of 6°-(-11°) S and 95°-141.25°E. This data has a grid size of 0.5°×0.625°2.1.

3.3. Analysis Method

1. Performing data extraction includes adjusting the location, removing missing-values, calculating the grid average and adjusting the matrix. CSIRO-MK3.6.0 model data and MERRA-2 reanalysis data extracted and stored.
2. Conduct data exploration by means of data exploration with time series data plots on the variable temperature and relative humidity of the CSIRO-MK3.6.0 model, the outcome of the RCP 4.5 scenario. MERRA-2 data reanalysis to determine the characteristics and relative humidity in Indonesia.

3. Performing partial downscaling on each variable temperature and relative humidity using the Climate Imprint (CI) method on the model in CSIRO-MK3.6.0 on the MERRA-2 data reanalysis.
4. Correcting the bias of CI downscaling with regression nonparametric Fourier Series
   a. Bias correction for temperature
   b. Make a plot of air temperature data in Indonesia from the results of downscaling (x, y)
   c. Modelling the air temperature level in Indonesia with the Fourier series approach as follows:
   d. Given nonparametric regression model y=f(x)+ε
   e. Running a program for determining optimal K with the GCV method at air temperature levels in Indonesia based on the GCV method
   f. After K is optimal, the next step is to run the nonparametric regression model estimation program with the Fourier series approach on air temperature data in Indonesia.
   g. Modelling air temperature and relative humidity in Indonesia with Fourier Series approach
5. Evaluate the original MERRA-2 data corrected model, Downscaling and Fourier Series Estimation
6. Performing of THI calculation

4. Result and Discussion

This chapter will discuss the results obtained in this study, regarding the results of the Climate Imprint (CI) downscaling method that will be used for climate projections using the RCP4.5 scenario. The descriptive analysis carried out in this study included time series plots and maps of Temperature and Relative Humidity that occurred in Indonesia in the period 2006-2100 for RCP4.5 data and 2006-2019 for MERRA-2 data.

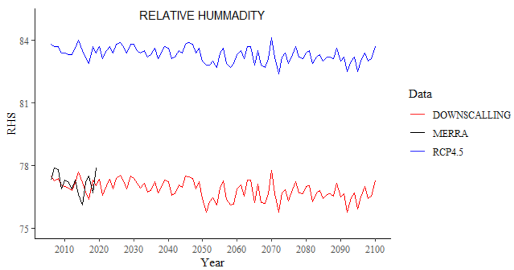


Fig 1. Time Series Plot RH

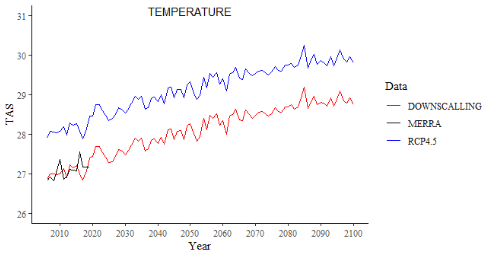
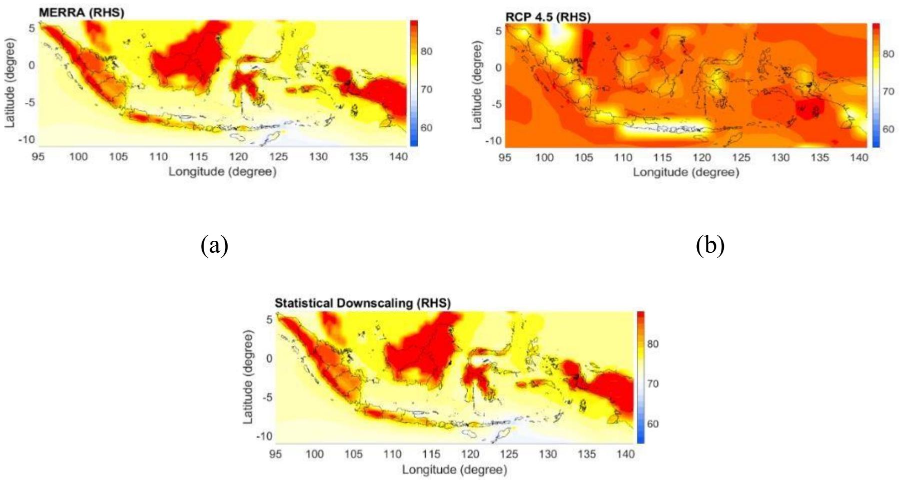


Fig 2. Time Series Plot TAS

Based on the time series graph, Temperature data in Indonesia for the three data, namely MERRA, RCP 4.5, and downscaling, experienced fluctuations but experienced a significant increase over a certain period of time. Relative Humidity data in Indonesia for the three data, namely MERRA, RCP 4.5 and downscaling, fluctuates over a period of years.

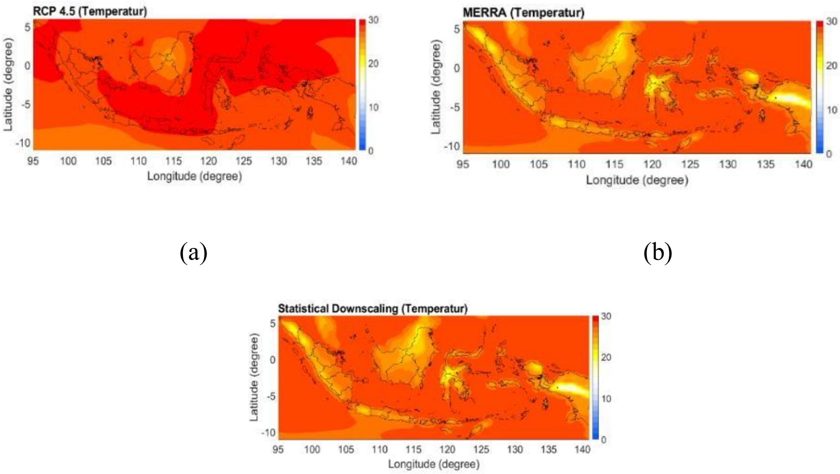


(a) (b)

(c)

Fig 3 Map of Relative Humidity (a) RH RCP 4.5 (b) RH MERRA-2 (c) RH After Statistical Downscaling.

Based on Figure 3, it can be seen that MERRA-2 and RCP 4.5 have different RH levels. In order to get effective results, Downscaling is carried out on the two data. downscaling succeeded in proving that it lowers the air humidity level to regional standard. Air humidity is relatively the same, between 69.1-76.8 degrees Celsius. The humidity level varies slightly between 92.3-100 degrees Celsius in the eastern part of Indonesia from Sulawesi to Papua.



(a) (b)

(c)

Fig 4. Temperature Map (a) TAS RCP 4.5 (b) TAS MERRA-2 (c) TAS After Statistical Downscaling

Based on Figure 4, it can be seen that MERRA-2 and RCP 4.5 have different temperature levels. MERRA-2 is the dependent variable which functions to carry out the downscaling process to provide results from a high grid scale to a local scale. Downscaling succeeded in proving that it lowers the air temperature to regional standards. Differences in temperature levels in Indonesia are caused by differences in latitude in an area and Indonesia which is in the equator area. Downscaling against MERRA-2 has proven to reduce the scale of the TAS and RHS grids. Therefore, model is needed to reduce the bias correction of the downscaling data to produce the best data.

4.1. Fourier Series Non-Parametric Regression Bias Correction

The first thing is to determine the optimal K value of the nonparametric regression model. The parameter K is assumed to be a positive integer. The GCV method is the method chosen to determine optimal K. The following are the results obtained for each K using the GCV method:

Table 1. The GCV value for each K is optimal of Relative Humidity Modelling

K	GCV
80	1.183×10^{-02}
90	9.464×10^{-04}
93	$\mathbf{1.317 \times 10^{-04}}$

Based on Table 1, the smallest GCV value at K=93 is the optimal K. If the value of K = 93 then the number of parameters that must be estimated is 95 parameters of Relative Humidity Modelling.

Table 2. The R^2 and MSE values for each K are optimal

Nilai K	R^2	MSE
85	**0,97 (97%)**	**0.0041**
90	0,99 (99%)	0.0008
93	1(100%)	1.290×10^{-18}

The best model was chosen based on the larger R^2 results and is a parsimony (simple) model so that the model taken when the K value = 85 consists of 87 parameters. The results of Relative Humidity (RH) modelling are presented by the following formula:

$\hat{y}$ = 77,759 - 0,0136 t +-0,126 cos t – 0,0511 cos 2t – 0,0509 cos 3t +... -0,0237 cos 85t

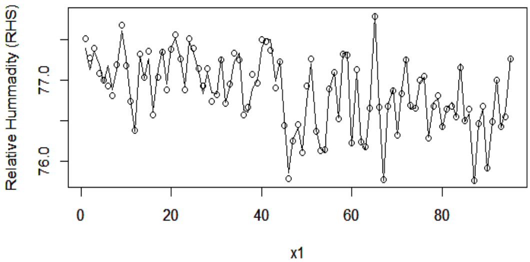


Fig 5. Relative Humidity (RH) Data Plot with Fourier Series

Based on Figure 5, shows that the highest Relative Humidity for Indonesia occur in 2070 which is equal to 77.80% while for Relative Humidity the lowest occurs in 2092 with Relative Humidity of 75.73%. The graph is close to the actual value after downscaling.

Table 3. The GCV value for each K is optimal of Temperature Modelling

K Value	GCV
80	3.625×10^{-03}
90	3.014×10^{-04}
94	$\mathbf{5.947 \times 10^{-06}}$

Based on Table 3, the smallest GCV value at K=94 is the optimal K. If the value of K = 94 then the number of parameters that must be estimated is 96 parameters.

Table 4. The R^2 and MSE values for each K are optimal

K Value	R^2	MSE
50	0.97 (97%)	0.0106
70	**0.98 (98%)**	**0.0050**
90	0.99 (99%)	0.0002

The best model was chosen based on the larger R^2 results and is a parsimony (simple) model so that the model taken when the K value = 70 consists of 72 parameters. The results of Temperature modelling are presented by the following formula:

$\hat{y}$ = 24,751 + 0,0289 t + 0,00773 cos t – 0,0182 cos 2t – -0,0197 cos 3t +... – 0,0312 cos 350 t

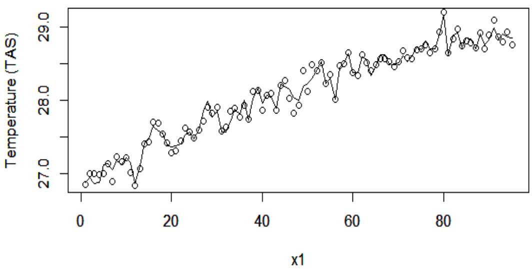
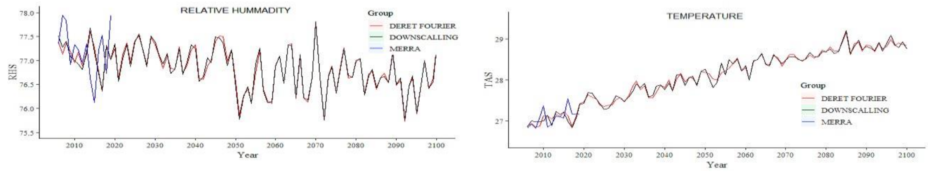


Fig 6. Temperature Data Plot with Fourier Series

Base on Figure 6, shows that the highest Temperature ($\hat{y}$) for Indonesia occurred in 2085 which was 29.15° C while for the lowest Temperature occurred in 2008 with a temperature of 26.86° C. The graph is close to the actual value after downscaling.

4.2. Mera-2 Data Plots, Downscaling RCP 4.5 and After Estimating the Fourier Series

The result of the estimation of the Fourier series is (y̑) which is then compared with data from downscaling and actual data from MERRA-2 as the dependent variable. The results of these comparisons are in the form of plot data.



(a) (b)

Fig 7. (a) Plot Data Relative Humidity (b) Plot Data Temperature

Base on Figure 7, shows that the Merra-2 data affects the results of the downscaling seen in the graph above the Merra data line with the downscaling line which shows that there is a decrease in the grid scale that occurs before and after the data is downscaled. For estimated data of the Fourier Series the graph is close to the actual value of the downscaling.

Based on result model evaluation to determine the best model of Downscaling Climate Imprint (CI) and nonparametric Regression, it is necessary to calculate the two data with the dependent variable MERRA-2 by calculating the Mean Squared Error (MSE) value.

Table 5. MSE Value Downscaling and Fourier Series

MSE	Relative Hummadity	Temperature
Downscalling	0,4031	0,0406
Fourier Series	0,3223	0,0290

Calculate THI the formula used Equation 2. THI in Indonesia has increased every year but not too significant between the average range of 0.0125° C per year with the highest number occurring in 2085 at 27.78002° C. Meanwhile, the lowest THI in Indonesia occurred in 2017 at 25.59337°C.

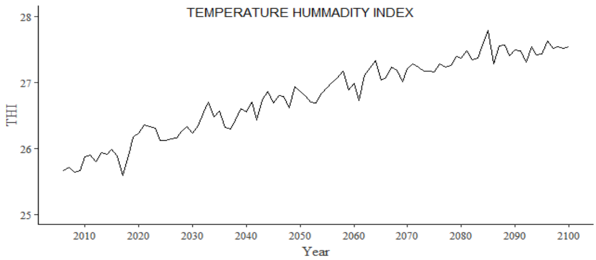


Fig 8. Plot of Indonesian THI

Based on Figure 8, shows that THI in Indonesia has increased every year with an average of being in fairly comfortable conditions of 25.5934°C-26.9998°C which occurred in 2006 - 2057. Then for discomfort it occurred in 2058 - 2100 with a THI value of 27 .07301° C - 27.78002° C.

5. Conclusion

The Statistical Downscaling model with the Fourier Series Nonparametric Regression method approach for projecting Indonesia's comfort level under the climate change scenario is as follows:

1. Model data from RH obtained K =85 model resulting in model with R^2 = 97%. The model formed is as follows:
2. Temperature obtained K = 70 model resulting in a model with R^2 = 98%. The model formed is as follows:
3. Based on the results obtained from the bias correction with Fourier Series Nonparametric Regression using climate change scenario data that the highest Relative Humidity occurred in 2070 which showed value of 77.80% and the lowest Relative Humidity in 2092 showed a value of 75.73% with an MSE value of 0.3223. Whereas in the Temperature variable (TAS), the highest temperature occurred in 2085, which was 29.15° C and the lowest temperature occurred in 2008 showing the number 26.86° C with MSE value of 0.0290.
4. The description of THI under the climate scenario obtained is that the highest THI occurred in 2085 at 27.78° C. Meanwhile for Indonesia the lowest THI occurred in 2017 at 25.59° C. On average, THI is in quite comfortable conditions 25.59 ° C - 26.99 ° C which occurred in 2006 –2057. Then for discomfort occurred in 2058 – 2100 with a THI value of 27.07° C - 27.780° C

Acknowledgments

Thank you to Universitas Muhammadiyah Semarang for support and funding through LPPM (Lembaga Penelitian dan Pengabdian Masyarakat) based on the “Internal Research".

References

1. S. Keman, “Global Climate, Human Health, and Sustainability Development,” *Jurnal Kesehatan Lingkungan Unair*, 3(2), 195–204, 2007
2. M. Measey. “Indonesia: A Vulnerable Country in the Face of Climate Change.” *Global Majority E-Journal*, 1(1), 31–45. 2010.
3. T. H. Karyono, “Predicting comfort temperature in Indonesia, an initial step to reduce cooling energy consumption,” *Buildings*, 5(3), 802–813. 2015. https://doi.org/10.3390/buildings5030802
4. A. N., Kakon, M. Nobuo, S. Kojima, & T. Yoko, “Assessment of Thermal Comfort in Respect to Building Height in a High-Density City in the Tropics,” *American Journal of Engineering and Applied Sciences*, 3(3), 545–551. 2010. https://doi.org/10.3844/ajeassp.2010.545.551
5. J. W. Hurrell, M. M. Holland, P.R. Gent, S. Ghan, J.E. Kay, P. J. Kushner, J.F. Lamarque, W. G. Large, D. Lawrence, K. Lindsay, W. H. Lipscomb, M.C. Long, N. Mahowald, D. R. Marsh, R. B. Neale, P. Rasch, S. Vavrus, M. Vertenstein, D. Bader, … S. Marshall, “The community earth system model: A framework for collaborative research,” *Bulletin of the American Meteorological Society*, 94(9), 1339–1360. 2013. https://doi.org/10.1175/BAMS-D-12-00121.1
6. M. Hadijati., D. Komalasari, N. Fitriyani, “Statistical Downscaling Regresi Nonparametrik Kernel untuk Prediksi Curah Hujan Bulanan Stasiun Sembalun,” *Prosiding SEMINAR NASIONAL MATEMATIKA II,* Bali, pp: 186 – 196, 2016.
7. F. Fauzi, H. Kuswanto, & R. M. Atok. Bias correction and statistical downscaling of earth system models using quantile delta mapping (QDM) and bias correction constructed analogues with quantile mapping reordering (BCCAQ). *Journal of Physics: Conference Series*, 1538(1). 2020. https://doi.org/10.1088/1742-6596/1538/1/012050
8. T. W. Utami, F. Fauzi, and E. Yuliyanto, “STATISTICAL DOWNSCALING USING REGRESSION NONPARAMETRIC OF FOURIER SERIES-POLYNOMIAL LOCAL OF CLIMATE CHANGE”, BAREKENG: J. Math. & App., vol. 17, no. 3, pp. 1411-1418, Sep. 2023.
9. A. Prahutama, Suparti, and T. W. Utami, “Modelling fourier regression for time series data - A case study: Modelling inflation in foods sector in Indonesia,” *Jounal*

of Physics: Conference Series, vol. 974, no. 1, 2018, doi: 10.1088/1742-6596/974/1/012067.

10. F. Fauzi, H. Kuswanto, R. M. Atok, and M. Salamah, "Statistical Downscaling and Bias Correction of the Earth System Models (ESM) Outputs for Future Climate Projection under Solar Geoengineering: Case Study Indonesia," *Journal of Mechanical Engineering*, vol. 10, no. 1, pp. 33–49, 2021

    K. Farzan et al., "Statistical downscaling and bias correction of climate model outputs for climate change impact assessment in the U. S. northeast," *Global and Planetary Change*, vol. 100, pp. 320–332, 2013, doi: 10.1016/j.gloplacha.2012.11.003.

11. L. J. Asrini and N. Budiantara, "FOURIER SERIES SEMIPARAMETRIC REGRESSION MODELS (CASE STUDY: THE PRODUCTION OF LOWLAND RICE IRRIGATION IN CENTRAL JAVA)," *Journal of Engineering and Applied Sciences*, vol. 9, no. 9, pp. 1501–1506, 2014, [Online]. Available: www.arpnjournals.com

12. Supari, F. Tangang, L. Juneng, and E. Aldrian, "Observed changes in extreme temperature and precipitation over Indonesia," *International Journal of Climatology*, vol. 37, no. 4, pp. 1979–1997, 2017, doi: https://doi.org/10.1002/joc.4829.
13. F. Marsitha, W. J. Pattipeilohy, and R. H. Virgianto, "Kenyamanan Termal Klimatologis Kota-Kota Besar Di Pulau Sulawesi Berdasarkan Temperature Humidity Index (THI)", *Jurnal Sains dan Matematika Unpam*, 1(2), 202-211, 2019. doi:10.32493/jsmu.v1i2.2384