Jurnal Info Sains : Informatika dan Sains
Volume 15, Number 02, 2025, DOI 10.54209/infosains.v15i02
ESSN 2797-7889 (Online)
https://ejournal.seaninstitute.or.id/index.php/InfoSains

Comparative Validation of NASA POWER and ERA5 Satellite-Based Meteorological Data Using BMKG Observations in Bandar Lampung, Indonesia

Ayu Aprilia[1], Alka Budi Wahidin[2], Ahmad Faruq Abdurrahman[1], Surya Prihanto[1], Yusril Al-Fath[1]

[1]Pyhsics Department, Universitiy of Lampung, Bandar Lampung, Lampung, Indonesia
[2]Department of atmospheric and planetary sciences, Sumatra Institute of Technology, South Lampung, Lampung, Indonesia

Article Info	ABSTRACT
Keywords: NASA POWER, Copernicus ERA5, BMKG, Validation, meteorological data.	This study aims to evaluate the performance of satellite reanalysis data from NASA POWER and Copernicus ERA5. BMKG data was used as a reference to compare the accuracy of satellite reanalysis data. Data was specifically collected from Lampung Province for the years 2022 to 2024. The data compared includes temperature, humidity, wind speed, and rainfall. The temperature data from ERA5 provided consistent and accurate results with a MAE of 1.82 and an r of 0.59. POWER showed commendable performance in capturing relative humidity with a MAE of 5.05% and an r of 0.33. For the wind speed variable, both models showed underestimation for Copernicus and overestimation for NASA POWER. For rainfall (RR), both models failed to predict extreme weather. NASA POWER showed an MAE of 8.59 mm/day and Copernicus showed a value of 7.68 mm/day for the rainfall variable. Future research can focus on bias correction results and machine learning to overcome the challenges faced by satellite data in predicting rainfall.
	Corresponding Author: Ayu Aprilia Physics Department, University of Lampung Bandar Lampung, Lampung ayu.aprilia@fmipa.unila.ac.id

INTRODUCTION

Accuracy is increasingly important in various fields, including technology. Accurate meteorological data is essential for addressing issues such as climate change, food security, water resource management, and sufficient input data for early warning systems for disaster risk management (Dunn et al., 2020). Based on research, the development of early warning system accuracy and climate models is greatly influenced by the quality and completeness of data input. For climate, the data inputs that must be considered are temperature, humidity, rainfall, and wind speed (Wilks, 2011). These data are important parameters for measuring weather and climate prediction values.

The BMKG (Meteorology, Climatology, and Geophysics Agency) is a government agency in Indonesia that conducts data observation activities and prepares data and information related to meteorology (Arnah Ritonga et al., 2025; Prakasa & Utami, 2019). The diverse microclimate iconography and tropical climate of Indonesia require continuous observation data for analysis and study of regional climate dynamics. If land-based observations through observation stations are often complicated and result in observation errors, the data will be at risk of being lost (Shaharudin, 2020). The absence of data will affect and damage the validity of the lost data, existing data statistics, and climatological data.

Overcoming missing data requires a solution, one of which is through interpolation techniques. Interpolation is generally used to estimate continuous variables such as heat (temperature) and humidity, which are related to time. Conversely, for rainfall and other data, interpolation is more complicated and very risky due to temporal and spatial regression. Other solutions for missing data are expected to come from other data sources and land-based data observations, especially for areas with reduced data observations (McMahon et al., 2013; Wangwongchai et al., 2023). In validating this study, it is very important to evaluate global satellite data as well as reanalysis data, it must be ensured that the data can be traded as a substitute for BMKG data.

Satellites and reanalysis data offer alternatives and complements to ground data, especially when ground data are unavailable. With extensive spatial coverage and temporal consistency, C3S and POWER-NASA have been widely used in climate and energy research (Aprilia et al., n.d.; Aprilia, Wahidin, & Syafriadi, 2025b; Rodrigues & Braga, 2021a; Stackhouse Jr, 2002). POWER-NASA has also been widely used in agricultural and renewable energy research (Aprilia, Wahidin, & Abdurrahman, 2025; Cox et al., 2006). Copernicus ERA5 is currently the latest reanalysis (Bell et al., 2021). Regional data for validation needs to be obtained from observations. For observation purposes, direct observation data can be used data (Tan et al., 2023; Urraca et al., 2018).

There have been many studies around the world that attempt to combine satellite data and reanalysis, with varying objectives. Some studies have seen good results in temperate regions (Tan et al., 2023). Regions with complex topography and tropical regions, on the other hand, have produced biased results (Tan & Armanuos, 2023; H. Wang et al., 2022; W. Wang & Hocke, 2022). This shows that the performance of reanalysis data is influenced by many factors (Urraca et al., 2018). In Indonesia, research on satellite data validation and reanalysis is still limited. Several studies have examined satellite data in Indonesia (Aprilia et al., n.d.; Aprilia, Wahidin, & Abdurrahman, 2025; Aprilia, Wahidin, & Syafriadi, 2025a). However, comprehensive analysis to evaluate the performance of NASA POWER and Copernicus data for multiple basic meteorological variables at the same time and location is still limited. Based on these issues, this study aims to evaluate the accuracy of satellite data. Satellite data from NASA POWER and Copernicus will be compared with BMKG field observation data. Daily data parameters such as temperature, humidity, rainfall, and wind speed will be evaluated to determine the performance of satellite data and reanalysis.

METHODS

The research aims to evaluate the accuracy of meteorological data from satellite sources by comparing it with terrestrial observation data. The research is quantitative in nature and uses a comparative observational method. The data is analyzed statistically to determine the accuracy of each dataset. The research was conducted using daily data for a period of 3 years from January 1, 2022 to December 31, 2024. Observation data from the BMKG station in Bandar Lampung at coordinates -5.45° Latitude and 105.31° Longitude. BMKG data was obtained through an integrated database system with the BMKG direct observation station. The Copernicus model was accessed through the Climate Data Store (CDS) with a spatial resolution of 0.25 x 0.25. NASA POWER data was accessed through the POWER project platform with a resolution of 0.5 x 0.5.

Figure 1. The BMKG Lampung Station at coordinates –5.4554° S, 105.3103° E as a ground station

The reference data (Ground Truth) in this study uses daily terrestrial data from BMKG stations, with observation locations shown in Figure 1. Data from BMKG (ground truth), Copernicus, and NASA Prediction of Worldwide Energy Resources (POWER) consists of temperature (TEM), relative humidity (RH2M), wind speed (WIND), and rainfall (RR). NASA POWER and Copernicus data are set to the same coordinates as BMKG. The time used for each data set is the same, namely at 07:00 WIB, 13:00 WIB, 19:00 WIB and 01:00 WIB every day. The research flow from the data collection stage to the evaluation is presented in Figure 2.

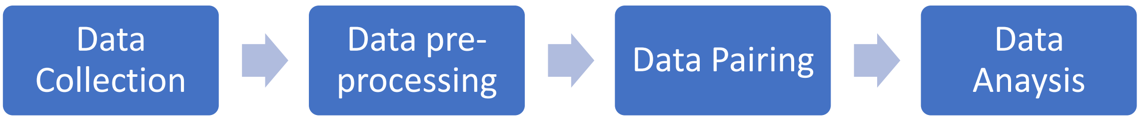

Figure 2. Research Flow Chart

The research data was collected from the platforms https://dataonline.bmkg.go.id, https://cds.climate.copernicus.eu/, and https://power.larc.nasa.gov/. The data was then compiled based on parameter and time compatibility. BMKG data as reference data was reviewed in advance to identify and handle missing data and invalid values (Shaharudin, 2020). The data will be reviewed for the percentage of missing data for each parameter. If the missing data is less than 6%, temporal linear interpolation will be performed to fill in the

Jurnal Info Sains : Informatika dan Sains
Volume 15, Number 02, 2025, DOI 10.54209/infosains.v15i02
ESSN 2797-7889 (Online)
https://ejournal.seaninstitute.or.id/index.php/InfoSains

empty data (Wangwongchai et al., 2023). For parameters with a high percentage of missing data, the daily data will be deleted (Duarte et al., 2022). Data from satellite sources will only be paired with BMKG data that is complete or has been interpolated.

Satellite data (Copernicus and NASA POWER) and BMKG data that have undergone pre-processing are matched for each date and time. The data will form pairs (paired data) for each parameter. Then, the data will be evaluated for suitability using a series of statistical metrics in accordance with climatological standards (Ruppert, 2004; Wilks, 2011). The statistics used in the evaluation are calculating bias, mean absolute error (MAE) in equation (1), bias in equation (2), root mean square error (RMSE) in equation (3), and Pearson's correlation coefficient (r) in equation (4).

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (1)$$

With n is number of data points, y_i is observed value and $\hat{y}_i$ is predicted value. The definition of $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ and $\bar{\hat{y}} = \frac{1}{n}\sum_{i=1}^{n} y_i$

$$Bias = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i) \quad (2)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (3)$$

$$r = \frac{\sum_{i=1}^{n}(y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}\sqrt{\sum_{i=1}^{n}(\hat{y}_i - \bar{\hat{y}}_i)^2}} \quad (4)$$

In bias evaluation, the model's systematic tendency to overestimate or underestimate actual values is assessed. MAE is used to measure the average absolute difference between predicted and observed values. RMSE is the root mean square error (Willmott & Matsuura, 2005). The r-value measures the strength and linear relationship between predicted and actual values (Sedgwick, 2012). Each metric has a specific use. Bias is used to detect over prediction or under prediction. RMSE is more sensitive to outliers, thus highlighting large errors (Zimmerman et al., 2003). The r value shows the extent to which the prediction pattern aligns with the observed values.

RESULTS AND DISCUSSION

This study aims to examine the quality of satellite meteorological data in Lampung Province. Satellite data products from NASA POWER and Copernicus have almost perfect data completeness. Meanwhile, BMKG data, which is used as ground truth reference data, has significant missing data challenges, especially in rainfall data. Given this, all data must undergo a preprocessing stage.

BMKG rainfall data (RR) collected over the last three years from 2022 to 2024 has a missing data rate of 32.2%. Missing data for all parameters is shown in Table 1. The results of the missing data calculation form the basis for deciding whether empty data will be interpolated or deleted. For temperature and humidity data, where the missing data rate is less than 6%, the missing data will be filled in using interpolation. For rainfall data with a BMKG missing data rate of 32.2%, the data (Copernicus, NASA POWER, and BMKG) will be deleted on the same date.

Jurnal Info Sains : Informatika dan Sains
Volume 15, Number 02, 2025, DOI 10.54209/infosains.v15i02
ESSN 2797-7889 (Online)
https://ejournal.seaninstitute.or.id/index.php/InfoSains

Table 1. Data availability

Variable	Missing Count	Total Data	Missing Percent	Summary
TEM1 (Copernicus)	0	8552	0	Continue
RH2M1 (Copernicus)	0	8552	0	Continue
WIND1 (Copernicus)	0	8552	0	Continue
RR1 (Copernicus)	4	8552	0.093545%	Continue
TEM2 (NASAPOWER)	0	8552	0	Continue
RH2M2 (NASAPOWER)	0	8552	0	Continue
WIND2 (NASAPOWER)	0	8552	0	Continue
RR2 (NASAPOWER)	0	8552	0	Continue
TEM3 (BMKG)	244	8552	5.706268%	Need Interpolation
RH2M3 (BMKG)	232	8552	5.425631%	Need Interpolation
WIND3 (BMKG)	0	8552	0	Continue
RR3 (BMKG)	1376	8552	32.1797	Avoid Interpolation

The evaluation results from 8552 daily data showed varying results in Table 2 for the Copernicus and NASA POWER models. The prediction results for air temperature (TEM) from the Copernicus model provided better accuracy with a MAE value of 1.815OC and a correlation (r) of 0.5888. NASA Power showed an MAE value of 1.953°C with a correlation (r) of 0.530. These results are in line that Copernicus data showed high consistency in surface temperature estimates with an RMSE between 1.8-2.2°C in tropical regions (Tan et al., 2023).

The relative humidity parameter (RH2M) shows that NASA POWER performs better with an MAE of 5.047% compared to Copernicus, which has an MAE of 8.738%. NASA POWER's correlation value is higher than Copernicus. These results are in line with the research by White et al. (2021), which found that NASA POWER data has good accuracy in estimating atmospheric humidity in areas with high cloud cover. This study found that Copernicus model results for relative humidity parameters (Cox et al., 2006; Rodrigues & Braga, 2021b).

Table 2. Statistical Analysis of The Data

Variable	Satellite	Mean Observation	Mean Prediction	Bias	MAE	RMSE	R	r^2
TEM1	Copernicus	28.515	26.719	-1.796	1.815	2.001	0.587	0.345
TEM2	NASA POWER	28.515	26.579	-1.936	1.952	2.134	0.530	0.280
RH2M	Copernicus	81.110	78.012	-3.097	8.737	17.183	0.237	0.056

Variable	Satellite	Mean Observation	Mean Prediction	Bias	MAE	RMSE	R	r^2
RH2M	NASA POWER	81.110	83.951	2.8413	5.047	6.244	0.332	0.110
WIND	Copernicus	1.605	0.7122	-0.892	1.098	1.464	0.5087	0.258
WIND	NASA POWER	1.605	2.1348	0.5296	1.120	1.349	0.389	0.151
RR	Copernicus	7.927	0.4696	-7.457	7.679	18.633	0.164	0.027
RR	NASA POWER	7.927	6.2323	-1.695	8.589	17.471	0.221	0.048

Figure 3 shows a comparison of the MAE of two satellite models against observation data (BMKG). The MAE values of four variables were observed for four conditions: dry, heavy rain, dry season, and wet season. The results show that most of NASA POWER's predictions have lower MAE values than Copernicus. NASA POWER's values are closer to BMKG observation data. The error values are higher during heavy rain conditions (RR> 20mm/day). In extreme conditions, rainfall variation shows a decrease in model accuracy. This is in line with previous research that satellite-based models often have difficulty capturing the spatial and temporal variability of rainfall. Satellite models have limitations in temporal resolution and dependence on cloud-based estimation algorithms (Kumar et al., 2022; Urraca et al., 2018).

Figure 4 shows the bias distribution (Predicted-Actual) for each variable in the weather condition group. Based on the results, it can be seen that TEM and RH2M tend to underestimate (negative bias) and NASA POWER tends to slightly overestimate. This is because Copernicus ERA5 uses data assimilation systems with radiative components that are conservative towards surface temperature. NASA POWER is obtained from MERRA2 (Modern-Era Retrospective Analysis for Research and Applications), which is more adaptive to surface radiation.

Jurnal Info Sains : Informatika dan Sains
Volume 15, Number 02, 2025, DOI 10.54209/infosains.v15i02
ESSN 2797-7889 (Online)
https://ejournal.seaninstitute.or.id/index.php/InfoSains

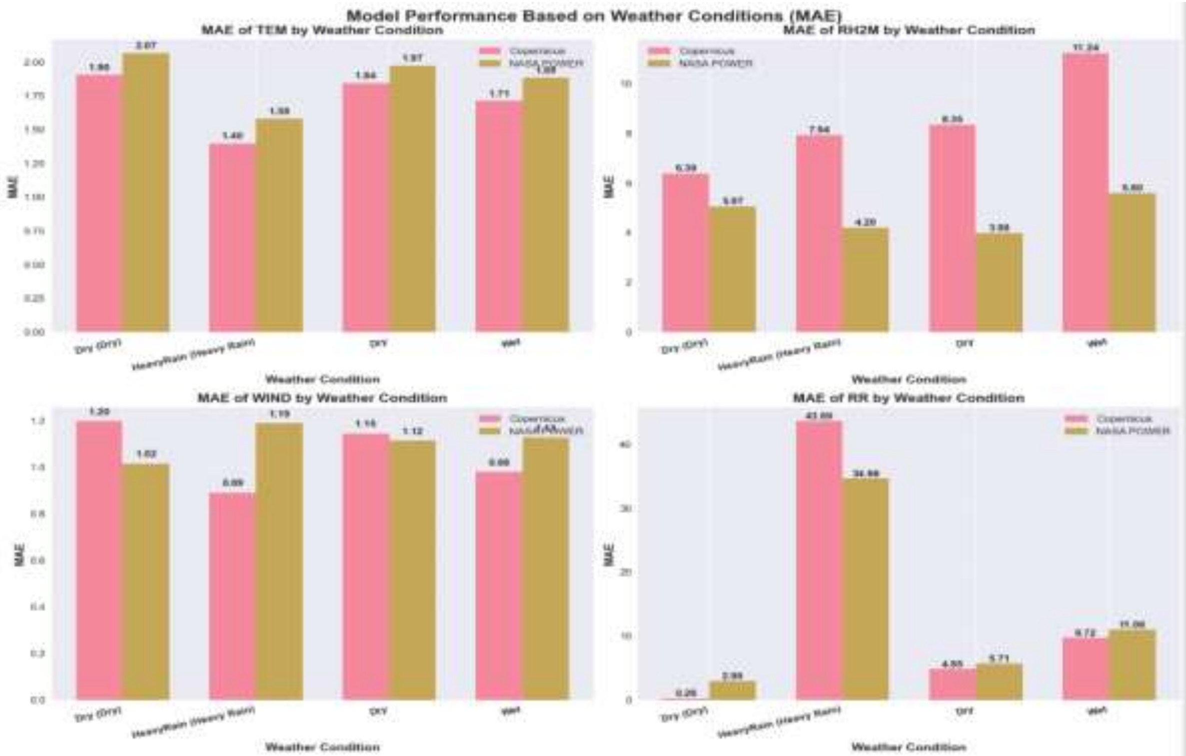

Figure 3. Model Performance Based on Weather Conditions (MAE)

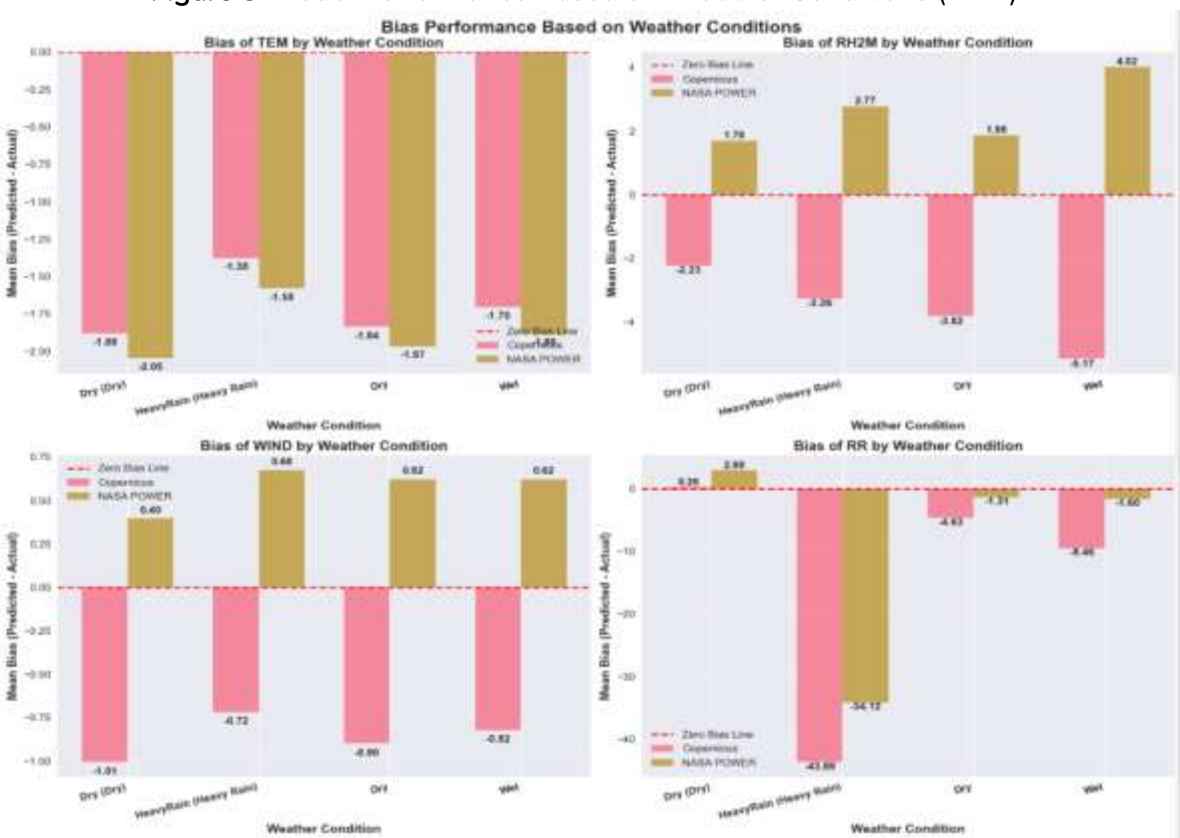

Figure 4. Model Performance Based in Weather Conditions (bias)

The rainfall variable in both satellite models shows a negative bias, especially in heavy rain conditions. This indicates that satellite data tends to underestimate extreme rainfall. This is in line with previous studies that passive satellite sensors often fail to detect high intensity due to microwave signal saturation and limited spatial resolution.

Overall, NASA POWER shows more consistent performance and less bias compared to Copernicus. This is particularly evident in humidity and rainfall variables. However, Copernicus still has an advantage in surface temperature estimation, especially during dry conditions. The accuracy of both models does face problems in handling extreme conditions. This confirms the need to apply bias correction and integrate local observation data (BMKG) in spatial representation in tropical regions such as Indonesia.

CONCLUSION

The research aims to evaluate the performance of satellite data reanalysis from NASA POWER and Copernicus ERA5. For BMKG data, the accuracy of satellite reanalysis data was tested. Data was specifically obtained for Lampung province from 2022 to 2024. The parameters compared were temperature, humidity, wind speed, and rainfall. For temperature data, ERA5 was quite consistent and accurate with an MAE of 1.82 and an r of 0.59. NASA POWER's humidity estimates were accurate with an MAE of 5.05% and an r of 0.33. For wind speed, both models were inaccurate, with Copernicus greatly underestimating and NASA POWER greatly overestimating. For precipitation, both models fail to account for extreme weather. For precipitation, NASA POWER records 8.69 mm/day and Copernicus 7.68 mm/day. For future studies, the focus can be directed towards bias correction and the use of machine learning to solve the challenges faced by satellite data in predicting precipitation.

ACKNOWLEDGEMENT

Aprilia, A., Faruq Abdurrahman, A., & Ananda, A. (n.d.). Analysis of Daily and Annual Variations of Solar Radiation, Air Temperature, Precipitation, Wind Speed, Humidity, and Air Pressure in Lampung as Indicators of Renewable Energy. *Indonesian Journal of Science Education and Applied Science (IJSEAS)*. https://ppjp.ulm.ac.id/journals/index.php/ijseas/index

Aprilia, A., Wahidin, A. B., & Abdurrahman, A. F. (2025). Integration of Machine Learning and NASA POWER Dataset for Predicting Coffee Production in Lampung. *Jurnal Fisika Flux: Jurnal Ilmiah Fisika FMIPA Universitas Lambung Mangkurat*, *22*(1), 44. https://doi.org/10.20527/flux.v22i1.20980

Aprilia, A., Wahidin, A. B., & Syafriadi. (2025a). Seismic Activity Analysis in Indonesia: Integrating Machine Learning, Geospatial Data, and Environmental Factors for Risk Assessment. *JPF (Jurnal Pendidikan Fisika) Universitas Islam Negeri Alauddin Makassar*, *13*(1), 12–26. https://doi.org/10.24252/jpf.v13i1.54081

Arnah Ritonga, Asni Al Amini, Livia Mutianda, Riamonda Singarimbun, Aiman Hidayat Baeha, Glensius Rayhane Pasaribu, & Juanda Arief Darmawan Damanik. (2025). Analisis Probabilitas Hujan Menggunakan Data Historis Dari BMKG Wilayah I Tahun

2013-2015. *JURNAL RISET RUMPUN MATEMATIKA DAN ILMU PENGETAHUAN ALAM*, *4*(1), 01–20. https://doi.org/10.55606/jurrimipa.v4i1.4367

Bell, B., Hersbach, H., Simmons, A., Berrisford, P., Dahlgren, P., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Radu, R., Schepers, D., Soci, C., Villaume, S., Bidlot, J. R., Haimberger, L., Woollen, J., Buontempo, C., & Thépaut, J. N. (2021). The ERA5 global reanalysis: Preliminary extension to 1950. *Quarterly Journal of the Royal Meteorological Society*, *147*(741), 4186–4227. https://doi.org/10.1002/qj.4174

Cox, S. J., Stackhouse Jr, P. W., Gupta, S. K., Mikovitz, J. C., Zhang, T., Hinkelman, L. M., Wild, M., & Ohmura, A. (2006). The NASA/GEWEX surface radiation budget project: Overview and analysis. *Extended Abstracts, 12th Conf. on Atmospheric Radiation*.

Duarte, L. V., Formiga, K. T. M., & Costa, V. A. F. (2022). Comparison of Methods for Filling Daily and Monthly Rainfall Missing Data: Statistical Models or Imputation of Satellite Retrievals? *Water (Switzerland)*, *14*(19). https://doi.org/10.3390/w14193144

Dunn, R. J. H., Stanitski, D. M., Gobron, N., Willett, K. M., Ades, M., Adler, R., Allan, R., Allan, R. P., Anderson, J., & Argüez, A. (2020). Global climate. *Bulletin of the American Meteorological Society*, *101*(101 (8)), S9–S127.

Kumar, S., Amarnath, G., Ghosh, S., Park, E., Baghel, T., Wang, J., Pramanik, M., & Belbase, D. (2022). Assessing the Performance of the Satellite-Based Precipitation Products (SPP) in the Data-Sparse Himalayan Terrain. *Remote Sensing*, *14*(19). https://doi.org/10.3390/rs14194810

McMahon, T. A., Peel, M. C., Lowe, L., Srikanthan, R., & McVicar, T. R. (2013). Estimating actual, potential, reference crop and pan evaporation using standard meteorological data: a pragmatic synthesis. *Hydrology and Earth System Sciences*, *17*(4), 1331–1363.

Prakasa, A., & Utami, F. D. (2019). Sistem Informasi Radar Cuaca Terintegrasi BMKG. *Journal of Telecommunication, Electronics, and Control Engineering (JTECE)*, *1*(02), 78–87. https://doi.org/10.20895/jtece.v1i02.89

Rodrigues, G. C., & Braga, R. P. (2021). Evaluation of Nasa Power Reanalysis Products to Estimate Daily Weather Variables in a Hot Summer Mediterranean Climate. *Agronomy*, *11*(6). https://doi.org/10.3390/agronomy11061207

Ruppert, D. (2004). *The elements of statistical learning: data mining, inference, and prediction*. Taylor & Francis.

Sedgwick, P. (2012). Pearson's correlation coefficient. *Bmj*, *345*.

Shaharudin, S. M. (2020). Imputation methods for addressing missing data of monthly rainfall in Yogyakarta, Indonesia. *International Journal of Advanced Trends in Computer Science and Engineering*, *9*(1.4), 646–651. https://doi.org/10.30534/ijatcse/2020/9091.42020

Stackhouse Jr, P. W. (2002). New results from NASA/GEWEX Surface Radiation Budget Project: Evaluating El Nino effects at Different Scales. *11th Conference on Atmospheric Radiation*.

Tan, M. L., Armanuos, A. M., Ahmadianfar, I., Demir, V., Heddam, S., Al-Areeq, A. M., Abba, S. I., Halder, B., Kilinc, H. C., & Yaseen, Z. M. (2023). Evaluation of NASA POWER and

ERA5-Land for estimating tropical precipitation and temperature extremes. *Journal of Hydrology*, *624*, 129940.

Urraca, R., Huld, T., Gracia-Amillo, A., Martinez-de-Pison, F. J., Kaspar, F., & Sanz-Garcia, A. (2018). Evaluation of global horizontal irradiance estimates from ERA5 and COSMO-REA6 reanalyses using ground and satellite-based data. *Solar Energy*, *164*, 339–354. https://doi.org/10.1016/j.solener.2018.02.059

Wang, H., Zhang, J., Chen, L., & Li, D. (2022). Relationship between summer extreme precipitation anomaly in Central Asia and surface sensible heat variation on the Central-Eastern Tibetan Plateau. *Climate Dynamics*, *59*(3–4), 685–700. https://doi.org/10.1007/s00382-022-06148-w

Wang, W., & Hocke, K. (2022). Atmospheric Effects and Precursors of Rainfall over the Swiss Plateau. *Remote Sensing*, *14*(12). https://doi.org/10.3390/rs14122938

Wangwongchai, A., Waqas, M., Dechpichai, P., Hlaing, P. T., Ahmad, S., & Humphries, U. W. (2023). Imputation of missing daily rainfall data; A comparison between artificial intelligence and statistical techniques. *MethodsX*, *11*, 102459. https://doi.org/10.1016/j.mex.2023.102459

Wilks, D. S. (2011). *Statistical methods in the atmospheric sciences* (Vol. 100). Academic press.

Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, *30*(1), 79–82.

Zimmerman, D. W., Zumbo, B. D., & Williams, R. H. (2003). Bias in estimation and hypothesis testing of correlation. *Psicológica*, *24*(1).