

**JOURNAL OF INFORMATION
SYSTEM AND TECHNOLOGY
MANAGEMENT
(JISTM)**www.gaexcellence.com/jistm

THE DEVELOPMENT OF AN ARTIFICIAL NEURAL NETWORK MODEL FOR RAINFALL ESTIMATION ACCURACY USING PERSIANN SATELLITE AND METEOROLOGICAL VARIABLES AT SUNGAI PINANG CATCHMENT

Intan Shafeenar Ahmad Mohtar^{1*}, Anisatul Batrisyiah Asmuri², Raja Nor Husna Raja Mohd Noor³, Nuryazmeen Farhan Haron⁴, Nur Azwa Muhamad Bashar⁵, Juzailah Nur Yunus⁶

¹Civil Engineering Studies, UiTM Pulau Pinang, Campus Permatang Pauh, 13500 Permatang Pauh, Pulau Pinang, Malaysia

 intanshafeenar@uitm.edu.my

 <https://orcid.org/0009-0006-8183-7612>

² Civil Engineering Studies, UiTM Pulau Pinang, Campus Permatang Pauh, 13500 Permatang Pauh, Pulau Pinang, Malaysia

 anisasmuri1412@gmail.com

 <https://orcid.org/0009-0003-0891-0615>

³ Civil Engineering Studies, UiTM Pulau Pinang, Campus Permatang Pauh, 13500 Permatang Pauh, Pulau Pinang, Malaysia

 rajanorhusna@uitm.edu.my

 <https://orcid.org/0000-0002-8500-4259>

⁴ Faculty of Civil Engineering, UiTM Shah Alam, Selangor, Malaysia

 nuryazmeen@uitm.edu.my

 <https://orcid.org/0000-0002-0108-3982>

⁵ Civil Engineering Studies, UiTM Pulau Pinang, Campus Permatang Pauh, 13500 Permatang Pauh, Pulau Pinang, Malaysia

 nurazwa.bashar@uitm.edu.my

 <https://orcid.org/0000-0002-1518-7424>

⁵ Civil Engineering Studies, UiTM Pulau Pinang, Campus Permatang Pauh, 13500 Permatang Pauh, Pulau Pinang, Malaysia

 juzailah.nur@uitm.edu.my

 <https://orcid.org/0000-0002-7744-3921>

*Corresponding Author

Article Info:**Article history:**

Received date: 23.07.2026

Revised date: 20.08.2026

Accepted date: 27.09.2026

Published date: 30.09.2026

Abstract:

Accurate rainfall estimation is essential for flood forecasting, water resource management, and disaster risk reduction, particularly in highly urbanized tropical catchments. However, satellite-derived rainfall products often exhibit systematic estimation errors when applied at local scales. This study aims to enhance the accuracy of precipitation estimation from PERSIANN satellite rainfall estimates using an Artificial Neural Network (ANN) model at Sungai Pinang catchment,

To cite this document:

Mohtar, I. S. A., Asmuri, A. B., Noor, R. N. H. R. M., Haron, N. F., Bashar, N. A. M., & Yunus, J. N. (2026). The Development of An Artificial Neural Network Model for Rainfall Estimation Accuracy Using PERSIANN Satellite and Meteorological Variables at Sungai Pinang Catchment. *Journal of Information System and Technology Management*, 11 (44), 952-974.

Penang, Malaysia. Historical rainfall observations from five rain gauge stations were integrated with meteorological variables, including temperature, relative humidity and wind speed. The ANN model employed a three-layer feedforward Backpropagation Neural Network with a 5–10–1 architecture and was trained using data collected between October and December 2018, while an independent dataset from September 2018 was used for model testing. The results demonstrate that the proposed ANN model significantly improved the accuracy of the satellite rainfall estimates across all five rainfall stations. The result indicates that r increased from range 0.132~0.354 (before ANN post processing) to 0.557 ~0.821 (after ANN post processing). Station Kolam Takungan Air Itam indicate the highest percentage increase which is 67.3%. The result for R^2 improved from range 0.010~0.125 to 0.8769 ~0.9819 with station Kolam Bersih Pulau Pinang had the highest percentage increase of 93.83%. In addition, the RMSE decreased from range 22.634 ~53.3 to 9.85~37.643, and the MAE was reduced from range 12.929~ 26.734 to 6.281~23.247, indicating substantial reductions in rainfall estimation errors. The findings demonstrate that by integrating satellite-derived rainfall estimates with meteorological variables together with an ANN model effectively enhances rainfall estimation accuracy. The proposed approach provides a reliable tool for improving flood forecasting, hydrological modelling, and water resource management in highly urbanized tropical catchments.

DOI: 10.35631/JISTM.1144056

Keyword:

Artificial Neural Network (ANN) model, Flood Forecasting, PERSIANN, Rainfall Estimation, Satellite Precipitation, Sungai Pinang Catchment.



© The authors (2026). This is an Open Access article distributed under the terms of the Creative Commons Attribution (CC BY NC) (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact jistm@gaexcellence.com.

Introduction

Accurate rainfall estimation is fundamental for effective water resource management, flood forecasting, agricultural planning, and disaster risk reduction (Latif et al., 2023; Benziiane, 2024). In tropical countries such as Malaysia, rainfall variability is strongly influenced by monsoon systems, climate variability, and rapid urbanization, all of which contribute to increasing flood frequency and severity (Yee et al., 2022; Saleh et al., 2022). Among the urban catchments in Penang, the Sungai Pinang catchment has experienced recurrent flash floods and has been identified as one of the most identified flood-prone urban catchments in Penang river basins in the state (Sivaji, 2025). The basin supports dense residential, commercial, and industrial developments, where rapid land-use changes and inadequate drainage systems further increase flood vulnerability. Consequently, accurate rainfall estimation is essential for improving flood forecasting, supporting sustainable urban planning, and strengthening disaster resilience within the catchment.

Rainfall observations are conventionally obtained from ground-based rain gauges, which provide highly accurate point measurements. However, their sparse spatial distribution often limits their ability to represent the spatial variability of rainfall, particularly in highly urbanized catchments where rainfall patterns are heterogeneous. Satellite-based precipitation products, such as the PERSIANN family (PERSIANN-CDR, PERSIANN-CCS, and PDIR-Now), provide continuous rainfall observations with extensive spatial and temporal coverage and have therefore become valuable sources of hydrological information (Hsu et al., 1997; Ceferino-Hernández et al., 2024; Eini et al., 2022; Baig et al., 2023). Despite these advantages, satellite-derived rainfall estimates frequently exhibit systematic biases at local and daily scales because precipitation is inferred indirectly from satellite observations rather than measured directly at the ground surface (Papacharalampous et al., 2023). These limitations reduce their reliability for flood forecasting and hydrological modelling in tropical urban environments. In addition, studies by Parsons et al. (2026) also indicate that the performance of satellite rainfall products often varies across seasons, climates, and geographic locations where it may perform poorly in mountainous and tropical regions where precipitation processes are complex.

Artificial Neural Networks (ANNs) have emerged as effective machine learning techniques for modelling complex nonlinear relationships between meteorological variables and rainfall to address these limitations. ANN models have demonstrated promising performance in rainfall prediction, rainfall-runoff simulation, streamflow forecasting, and water resource management because the model can learn non-linear patterns from multiple input variables by integrating satellite precipitation data with meteorological parameters such as temperature, relative humidity, and wind speed. Therefore, ANN models can reduce estimation errors and improve rainfall prediction accuracy. Although ANN-based rainfall predictions have been widely investigated, relatively few studies have focused on enhancing satellite-derived rainfall estimates for small tropical urban catchments such as the Sungai Pinang catchment. This represents an important research gap, particularly for flood-prone urban areas where accurate rainfall information is essential for early warning systems and disaster management.

This study develops an ANN based framework to improve the accuracy of satellite-derived rainfall estimation in the Sungai Pinang catchment, Penang, Malaysia. Historical rainfall observations from five rain gauge stations were used as reference data, while PERSIANN satellite rainfall data together with meteorological variables, including temperature, relative humidity, and wind speed, were used as an input variable. The objectives of this study are to: (i) analyze rainfall characteristics using historical rain gauge observations; (ii) evaluate the accuracy of PERSIANN satellite rainfall estimates; and (iii) improve satellite rainfall estimation by integrating satellite precipitation and meteorological variables using an ANN model. The enhanced rainfall estimation model is expected to provide more reliable rainfall information for flood forecasting, water resource management, and decision-making by agencies such as the Department of Irrigation and Drainage (DID), thereby contributing to improved urban flood resilience.

Literature Review

Accurate rainfall estimation is fundamental to hydrological modelling, flood forecasting, and water resource management, particularly in highly urbanized and flood-prone catchments. Traditional rainfall observations are primarily based on rain gauge networks that provide accurate point-based measurements but often fail to represent the spatial variability of rainfall due to their sparse spatial distribution (Latif et al., 2023). This limitation is of particular

importance in tropical urban environment where the rainfall intensity may vary significantly over small distances.

Satellite-based precipitation products have become increasingly important in hydrological applications in order to overcome these limitations. Among these products, the PERSIANN family including PERSIANN-CDR, PERSIANN-CCS, and PDIR-Now provides continuous rainfall observations with broad spatial and temporal coverage using remote sensing techniques (Ceferino-Hernández et al., 2024; Eini et al., 2022). The output from this product is widely used for flood forecasting, drought monitoring, and water resource assessment, particularly in regions with sparse rain gauge networks. However, satellite-derived rainfall estimates frequently exhibit systematic overestimation or underestimation at local scales because precipitation is inferred indirectly from cloud-top infrared observations rather than measured directly at the ground surface (Papacharalampous et al., 2023). Consequently, additional bias correction or calibration techniques are often required before these outputs can be reliably applied in operational hydrological studies.

ANN has been found to be one of the most successful machine learning techniques for modelling complex nonlinear relationships among meteorological variables and rainfall. Inspired by the learning mechanism of the human brain, ANNs can extract hidden patterns from multidimensional datasets without explicitly assuming data distribution (Aprianto et al., 2023). Therefore, the ANN models have better prediction accuracy than the conventional statistical approaches because they can capture the nonlinear interactions among the meteorological variables (Nandakumar et al., 2021; Thakur et al., 2021). The feedforward neural networks trained with the backpropagation algorithm are still the most popular among the available ANN architectures due to their simplicity, computational efficiency and strong generalization capability (Moghim & Bras, 2017; Qamar & Ali Zardari, 2023).

Recent research has shown that the combination of ANN models with multiple meteorological variables significantly enhances the accuracy of rainfall estimation. Mohtar et al. (2025a) suggested the combination of meteorological variables such as relative humidity, temperature and forecast rainfall from Weather Research Forecast (WRF) model, and integration with ANN model had able to enhance the accuracy of the rainfall estimation. Pratomo et al. (2022) and Ridwan et al. (2021) reported that the addition of temperature, relative humidity, and wind speed as additional inputs has significantly reduced the prediction errors over the models that only use rainfall observations. Similarly, Dao et al. (2025) used ANN to correct the biases in satellite-derived precipitation estimates under tropical climatic conditions and obtained significant improvements in RMSE and MAE.

Furthermore, research by Kunnath-Poovakka and Eldho (2023) showed that the correction of satellite precipitation products with rain-gauge observations significantly improved the accuracy of the products for estimating rainfall and made the products more applicable for hydrological modelling. Additionally, the results of Fuentes-Barrios and Sierra-Lorenzo (2021) showed that ANN models enhance the accuracy of the Quantitative Precipitation Forecast (QPF) by decreasing prediction errors and increasing the correlation coefficients. Papacharalampous et al. (2023) demonstrated the potential of machine learning algorithms to integrate satellite-derived precipitation data with ground observations, resulting in a significant enhancement in the accuracy of rainfall estimation. Similarly, Sadeghi et al. (2021) and Lim et al. (2024) found that satellite rainfall estimates that were consistently improved by ANN produce more reliable predictions than the original satellite products.

Despite these advances, relatively few studies have investigated ANN-based enhancement of satellite rainfall estimates in small tropical urban catchments, where rainfall variability is strongly influenced by rapid urbanization and localized convective storms. Most previous studies have focused on regional-scale applications or large river basins, with limited attention given to densely urbanized catchments such as the Sungai Pinang catchment. Furthermore, the combined use of PERSIANN satellite precipitation data with local meteorological variables (temperature, relative humidity, and wind speed) for rainfall estimation remains insufficiently explored in Malaysia. The identification of this research gap is important because improvement of satellite rainfall estimates can provide more reliable rainfall information for flood forecasting, water resource management, and urban disaster preparedness.

Previous studies have demonstrated that satellite-based precipitation products often require bias correction and validation against ground observations to improve reliability. Artificial Neural Networks (ANNs) offer a promising solution because they can model complex non-linear relationships between satellite observations and actual rainfall measurements, thereby reducing estimation errors and improving prediction accuracy. Machine learning-based correction approaches have shown superior performance compared with conventional statistical methods in rainfall estimation applications (Sumith,2025;). Therefore, this study develops an ANN-based rainfall estimation model by integrating PERSIANN satellite precipitation data with meteorological variables to improve rainfall estimation accuracy in the Sungai Pinang catchment. The proposed framework aims to reduce the uncertainty associated with satellite rainfall products and provide a more reliable rainfall dataset for hydrological and flood management applications.

Methodology

This study aims to enhance the accuracy of rainfall data estimation using the Artificial Neural Network (ANN) model by focusing on three objectives. The objectives which are to analyze the rainfall patterns from rain gauges, to evaluate the accuracy of satellite rainfall data (PERSIANN) compared to ground observations, and to enhance estimation accuracy using an ANN model. The flow of the methodology is summarized in Figure 1.

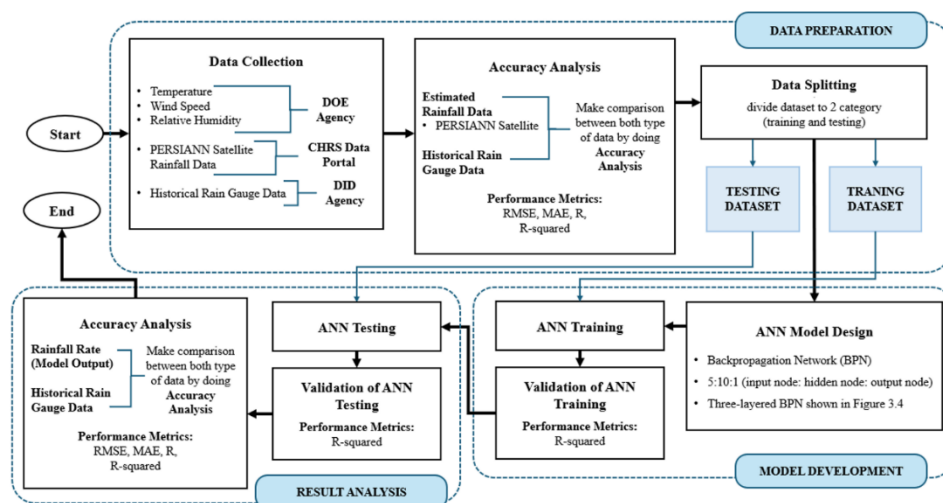


Figure 1: Flow Chart of Methodology

Catchment Area

The study area is located at the Sungai Pinang catchment in Penang, Malaysia (Figure 2). The catchment is the largest river (Saleh et al., 2022) an approximately 51 km² of basin size and comprises highest population density from George Town, Air Hitam and Paya Terubong urban areas (Yusoff & Rindam, 2012). Historically, a severe flood event in 2018 occurs in the study area. Two weather stations and five rainfall stations within the catchment were used to collect data, presented in Figure 3 and listed in Table 1. Rainfall stations were under Department of Irrigation and Drainage (DID), while for weather stations, it is operated under Department of Environment (DOE).

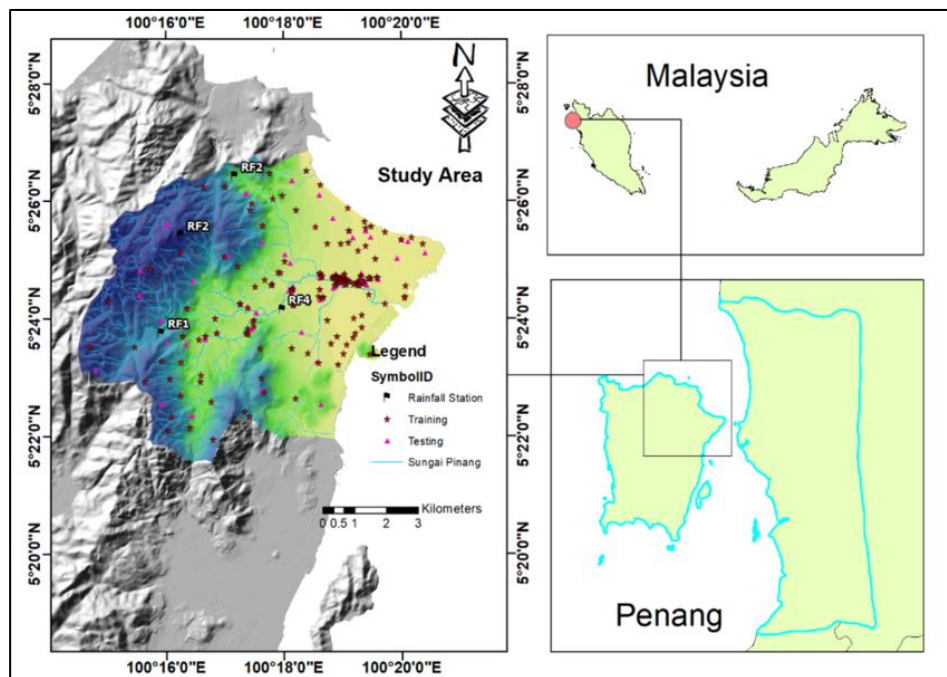


Figure 2: Location of Sungai Pinang Catchment

(Source: Saleh et al., 2022)

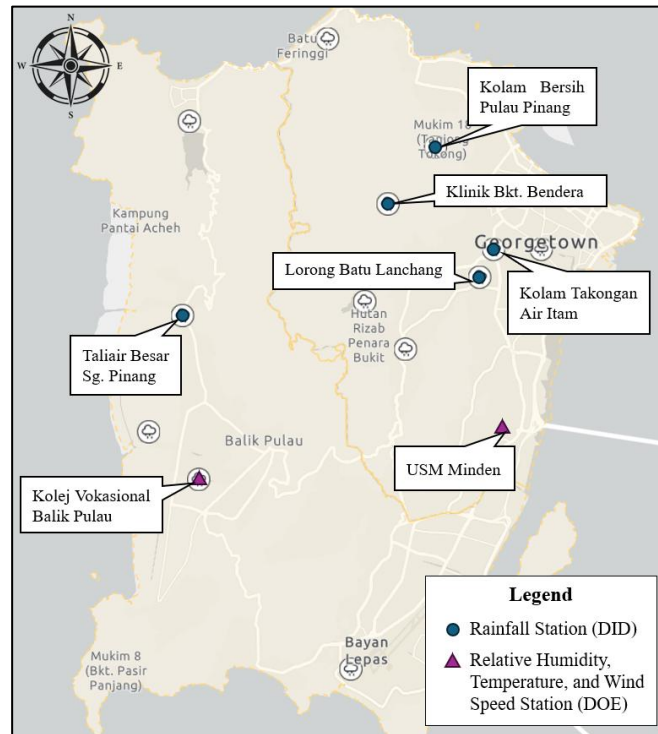


Figure 3: Location of Rainfall Stations and Weather Stations.

Table 1: List of Sungai Pinang Catchment Rainfall and Weather Stations

Station	No.	Station ID	Station Name	Latitude	Longitude
Rainfall Station	1	5402002	Kolam Bersih Pulau Pinang	5.440	100.286
	2	5402001	Klinik Bkt. Bendera	5.424	100.271
	3	5403001	Lorong Batu Lanchang	5.403	100.299
	4	5302003	Kolam Takongan Air Itam	5.396	100.265
	5	5302001	Tali Air Besar Sg. Pinang	5.392	100.213
Weather Station	6	CA08P	USM Minden	5.357	100.305
	7	CA09P	Kolej Vokasional Balik Pulau	5.337	100.214

Data Collection

Three reliable sources were used to obtain meteorological data. Table 2 displays the source of data collection and its parameter. This includes rainfall gauge data provided by the Department of Irrigation and Drainage (DID), meteorological parameters such as ambient temperature, relative humidity, and wind speed from the Department of Environment (DOE), and satellite-based rainfall estimates from the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) satellite dataset that accessed through the Centre for Hydrometeorology and Remote Sensing (CHRS) data portal. The Artificial Neural

Network (ANN) model was trained using rainfall and meteorological data from October to December of 2018.

Data from September 2018 were used for accuracy evaluation and model testing. Year 2018 was chosen for this study because during this time, it was a time when the Sungai Pinang catchment was experiencing a lot of hydrological stress. The area experienced frequent, intense flash floods throughout the year 2018, especially in July and during the monsoon season (October until February), following the record-breaking catastrophic events in late year 2017. These events provide a reliable dataset for testing the ANN model's capability in enhancing satellite precipitation estimates during critical flood-prone periods in a highly urbanized tropical basin. New input data which are from 1st September 2018 until 30th September 2018 was incorporated into the model for the purpose of validation. The selection of this data period primarily focused case study on extreme flood events during a high-stress hydrological period instead of using a general long-term climatological model.

Table 2: The Source of Data Collection and Its Parameter

Agencies/Source	Parameters
CHRS data portal	PERSIANN satellite rainfall data
Department of Environment (DOE)	Ambient temperature, relative humidity, and wind speed data
Department of Irrigation and Drainage (DID)	Data from rainfall gauges

Artificial Neural Network (ANN) Design

Artificial Neural Networks (ANNs) are computational models inspired by the biological nervous system functioning as the interconnected processing elements (neurons) are collaborated to solve complex problems (Grosan & Abraham, 2011). Each neuron is associated with weights and thresholds that govern its contribution to the network. In this study, a three-layer feedforward ANN with a 5:10:1 configuration (five input nodes, ten hidden nodes, and one output node) was used in this analysis (Figure 4). The research did not focus on different neurons. Thus, the variations in the quantity of neurons are not made in this study. The network was trained using a Backpropagation Network (BPN) algorithm. There are 5 input variables that were inserted into the ANN model. The input variables include temperature, relative humidity, wind speed, rainfall gauge data, and satellite rainfall estimate from the PERSIANN satellite, while the target output was the observed rainfall rate from gauges.

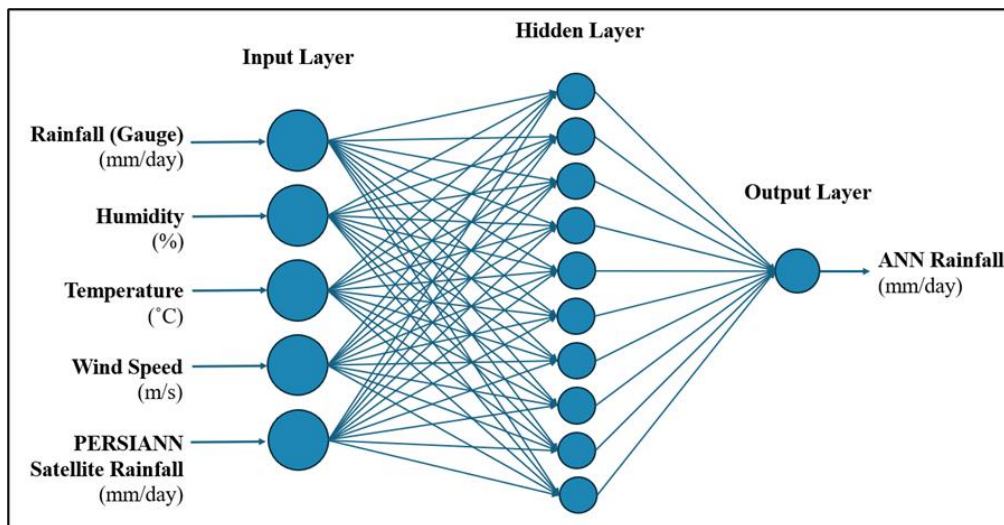


Figure 4: Three-layered Backpropagation in ANN Model

The training phase of the ANN model was conducted using three-month period of 1st October 2018 until 31st December 2018 equivalent to 92 days. The model was trained using a Backpropagation Network (BPN) algorithm, with validation datasets incorporated during training to minimize overfitting. The optimization techniques are also applied to reduce estimation errors. Subsequently, model outputs were validated by comparing ANN-predicted rainfall values against observed rainfall from the rain gauges. Model performance was assessed using root mean squared error (RMSE), mean absolute error (MAE), coefficient of determination (R^2), and correlation coefficient (r). This phase was performed to fulfil the second objective of the study, which was to assess the accuracy of satellite-based rainfall estimations in comparison to ground-based measurements.

Artificial Neural Network (ANN) Model Testing and Validation

The ANN model was tested to evaluate its generalization capability. An independent dataset covers the period from 1st to 30th September 2018 was used as an input. This dataset provided an additional basis for assessing the model's effectiveness in enhancing rainfall estimation accuracy. The testing procedure mirrored the training process, and the model outputs were subsequently visualized through predicted versus observed rainfall plots. Then, model validation was performed by comparing ANN-predicted rainfall with gauge observations from the Department of Irrigation and Drainage (DID).

Standard Evaluation Metrics for Validation

Several error metrics are used to assess Artificial Neural Networks (ANN) model performance. Performance was quantified using RMSE, MAE, R^2 , and r . These evaluations confirmed the model's accuracy and reliability, thereby demonstrating its potential for application in improved rainfall estimation.

Root Mean Square Error

Root mean square error (RMSE) provides a measure of how concentrated the data is around the line of best fit, where lower values indicate better model performance. The statistical

metrics calculation formula for RMSE is shown below (Dao et al., 2025), labelled as Equation 3.1.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Pred_i - Obs_i)^2} \quad (3.1)$$

Where:

n = the number of databases,
 Obs_i = the observed value at time i ,
 $Pred_i$ = the prediction value at time i .

Mean Absolute Error

Mean absolute error (MAE) provides a linear score of average error but less sensitive to outliers than RMSE. Equation 3.2 below shows the formula of MAE from the study by Dao et al. (2025).

$$MAE = \frac{1}{n} \sum_{i=1}^n |Obs - Pred| \quad (3.2)$$

Where:

n = the number of databases,
 Obs = the actual value at time i ,
 $Pred$ = the prediction value at time i .

Coefficient of Determination (R^2)

The Coefficient of determination and also known as R-squared (R^2), measures the percentage of variance in the observed (rain gauge) data that is explained by the Artificial Neural Network (ANN) model. The formula for R^2 is shown in equation 3.3 below.

$$R^2 = 1 - \frac{\sum_{i=1}^n (Obs - Pred)^2}{\sum_{i=1}^n (Obs - \underline{Obs})^2} \quad (3.3)$$

Where:

n = the number of databases,
 Obs = the observed value,
 $\underline{Obs}$ = the average observed value,
 $Pred$ = the predicted value.

Values R^2 is range from -1 to 1, with 1 indicating perfect positive correlation and -1 indicating perfect negative correlation. R^2 value that closer to 1 indicates a better fit between the predicted and the actual data.

Correlation Coefficients (r)

The function of correlation coefficient is to assesses the linear relationship between the estimated and observed rainfall. Equation 3.4 shows the formula of correlation coefficients (r) stated in a study from Dao et al. (2025).

$$r = \frac{\sum_{i=1}^n ((Pred - \underline{Pred})(Obs - \underline{Obs}))}{\sqrt{\sum_{i=1}^n (Pred_i - \underline{Pred})^2} \sqrt{\sum_{i=1}^n (Obs - \underline{Obs})^2}} \quad (3.4)$$

Where:

n = the number of databases,

Obs = the observed value,

$\underline{Obs}$ = the average actual value,

$Pred$ = the predicted value,

$\underline{Pred}$ = the average predicted value.

Values of r is range from -1 to 1, with 1 indicating perfect positive correlation and -1 indicating perfect negative correlation.

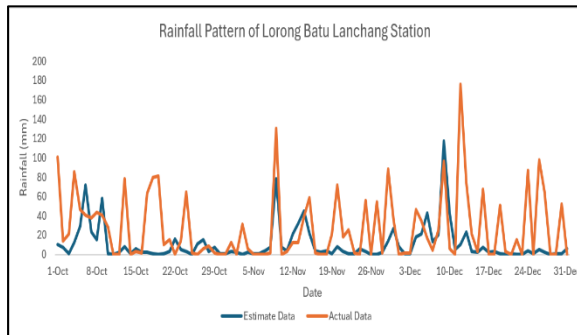
Results and Discussion

This section presents the findings from the rainfall data analysis, accuracy evaluations of satellite estimates, and ANN model performance. The discussion follows the structure of the study's objectives: (i) rainfall pattern analysis, (ii) evaluation of satellite-based estimation accuracy, and (iii) enhancement using ANN. The results are interpreted using statistical measures: root mean squared error (RMSE), mean absolute error (MAE), coefficient of determination (R^2), and correlation coefficient (r).

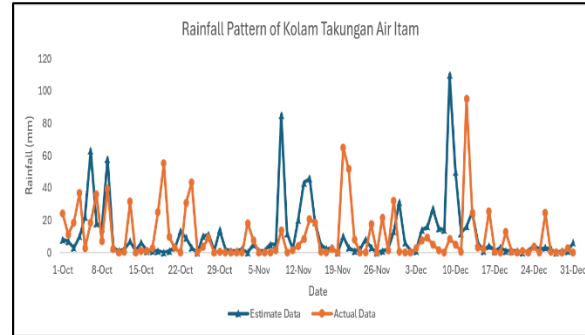
Rainfall Pattern Analysis

Rainfall data from five rain gauge stations in the Sungai Pinang catchment were analyzed to understand temporal and spatial rainfall distribution during the study period. The analysis revealed that October and November 2018 experienced the highest total rainfall, consistent with the region's wet monsoon season. The station at Kolam Bersih recorded the highest monthly rainfall, exceeding 370 mm in November, while Tali Air Besar Sg. Pinang station showed the lowest values, indicating spatial variability within the catchment. The rainfall intensity varied significantly between stations due to localized weather conditions and urban topography. These variations highlight the need for spatially distributed data in hydrological modelling and flood forecasting. Figure 5 demonstrates the rainfall pattern at different locations of rainfall stations.

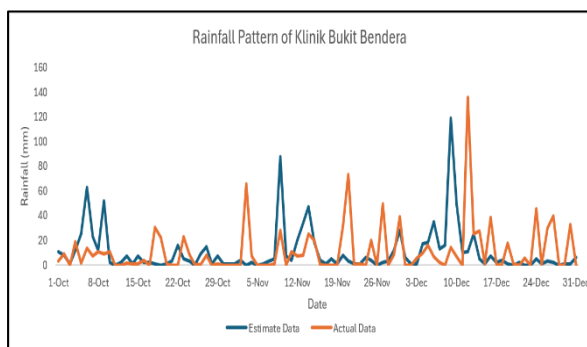
In general, the PERSIANN data have a similar pattern compared to the actual rainfall data with minor differences. Hence, the findings from Baig et al. (2023) can be verified which stated that PDIR-Now can be considered as possible alternatives to rain gauges since it has outperformed other PERSIANN models in terms of rainfall detection, correlation, and spatial distribution. However, the accuracy of the data value is critical where overestimation and underestimation occurred during the rainfall estimation. These variations highlight the limitations of the PDIR-Now data and the importance of data enhancement before it could be utilized in the future.



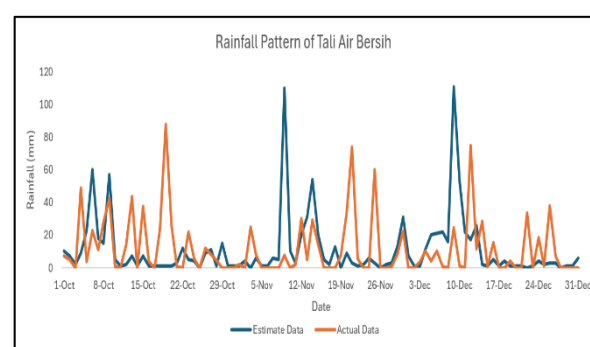
(i)



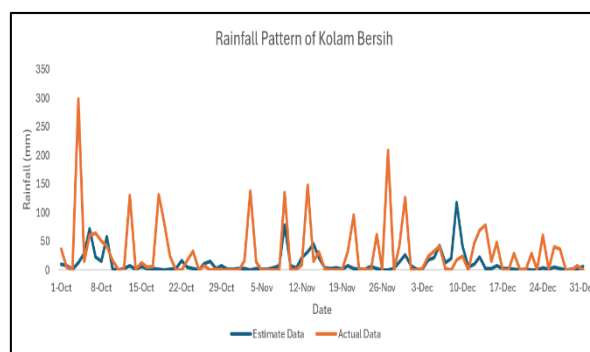
(ii)



(iii)



(iv)



(v)

Figure 5: Rainfall Pattern At (I) Lorong Batu Lanchang Station; (Ii) Kolam Takungan Air Itam Station; (Iii) Klinik Bkt. Bendera Station; (Iv) Tali Air Besar Sg. Pinang Station And (V) Kolam Bersih Pulau Pinang Station

Analysis of Accuracy of the ANN Based PERSIANN Rainfall

An accuracy analysis of the output of the ANN model was conducted to analyze the accuracy improvement of the estimation rainfall data. This allows for the comparison of initial and final accuracy analysis. The result for the accuracy analysis for the R^2 value was plotted for all following stations. Figure 6a and 6b demonstrate the result for R^2 value for station Lorong Batu Lanchang. The R^2 value increased from 0.125 (before ANN post-processing) to 0.9819 (after ANN post-processing). It is a tremendous increase for the R^2 value. Following Kolam

Takungan Air Itam station also exhibit a significant rise, where the R^2 value is from 0.022 to 0.8769 (figure 7a and 7b). This is also supported by the following station (figure 8a and 8b) which is at station Klinik Bkt Bendera, also shows a sharp rise from 0.01 (before ANN post-processing) to 0.8858 (after ANN post-processing). Next, for station Tali Air Besar Sg. Pinang, the R^2 value had also increased from 0.039 to 0.8959 (figure 9a and 9b). Lastly, for station Kolam Bersih indicated the highest percentage increase which is 93.83 % where the R^2 value from 0.017 to 0.8959 (Figure 10a and 10b).

The result of the statistical analysis before and after the utilization of the ANN model post-processing was summarized in Table 3. There is an improvement in the correlation for estimated rainfall data after ANN post processing where all correlations are higher for all rainfall stations. The findings show that the r value for each of the following stations (Lorong Batu Lanchang, Kolam Takungan Air Itam, Klinik Bkt. Bendera, Tali Air Besar Sg. Pinang, and Kolam Bersih Pulau Pinang) increased significantly from 0.354, 0.148, 0.1, 0.132, and 0.197 to 0.814, 0.821, 0.75, 0.742, and 0.557, respectively. Station Kolam Takungan Air Itam has the highest percentage increase which is 67.3%.

Additionally, the RMSE values for the following stations (Lorong Batu Lanchang, Kolam Takungan Air Itam, Klinik Bkt. Bendera, Tali Air Besar Sg. Pinang, and Kolam Bersih Pulau Pinang) decreased from 37.332, 22.634, 25.863, 24.786, and 53.3 to 20.049, 9.85, 11.846, and 37.643, respectively). The maximum reduction is observed at Lorong Batu Lanchang station. The MAE value decreased from 22.897, 12.929, 14.071, 14.239, and 26.734 to 14.579, 6.281, 8.598, 7.802, and 23.247 for each of the stations (Lorong Batu Lanchang, Kolam Takungan Air Itam, Klinik Bkt. Bendera, Tali air Besar Sg. Pinang, and Kolam Bersih Pulau Pinang). The Kolam Bersih Pulau Pinang station has the biggest reduction for the MAE value.

Summary of The Results

The result before ANN post processing analysis confirms that the untrained model was inadequate for accurate rainfall estimation. High RMSE and MAE values suggest large average errors, while the low r and R^2 values indicate a weak correlation between estimated and actual rainfall before training. The preprocessing step highlighted the inadequacy of raw or untrained model outputs and served as a baseline to measure improvements after training and post-processing. Hence, the result does support the past research's findings that the PERSIANN family products are unreliable in daily precipitation estimates and runoff simulations (Eini et al., 2022). A study from Ceferino-Hernández et al. (2024), declared that although the high temporal resolution is the main advantage of PERSIANN-CCS and PDIR-Now, the performance is limited because it does not measure rainfall but estimate it indirectly. Thus, it is confirmed that the potential of the PERSIANN data alone is insufficient to provide high accuracy of rainfall estimation.

The ANN post-processing successfully improved model performance and prediction accuracy across all stations. Hence, by completing this step, the ANN's model accuracy and reliability is improved. As a result, the model produces the estimated rainfall data with a better alignment with actual observations. Hence, the study from Pratomo et al. (2022) can be verified since the finding observed that rainfall prediction using an artificial neural network (ANN) with a combination of weather parameters had been proved to produce a smaller mean square error (MSE) value, which indicates a higher accuracy.

The findings demonstrate that by integrating satellite-derived rainfall estimates with meteorological variables together with an ANN model effectively enhances rainfall estimation accuracy. This aligned with the findings of Fuentes-Barrios & Sierra-Lorenzo (2021) where the researcher successfully presented the results of the application of a bias correction method QPF based ANN. The researcher also concludes that bias correction has a good ability to correct the precipitation forecast since some positive results were obtained in terms of the reduction of the MAE and MSE.

Furthermore, there are numerous studies that have indicated that the use of ANN model does improve the performance of the precipitation estimation with the observed data when tested for the reliability of the model. Research by Mohtar et al. (2025b) and Dao et al. (2025) had suggested that recent advances in machine learning (ML) techniques provide a great path possibility for improving the accuracy of precipitation estimation. In addition, Jabbari & Bae (2018) have carried out the bias correction of the climate modelled temperature and precipitation using ANN model since it had a superior performance compared with linear regression. Based on the findings from these studies, it can be concluded that ANN post-processing does enhance the accuracy of the predicted rainfall data.

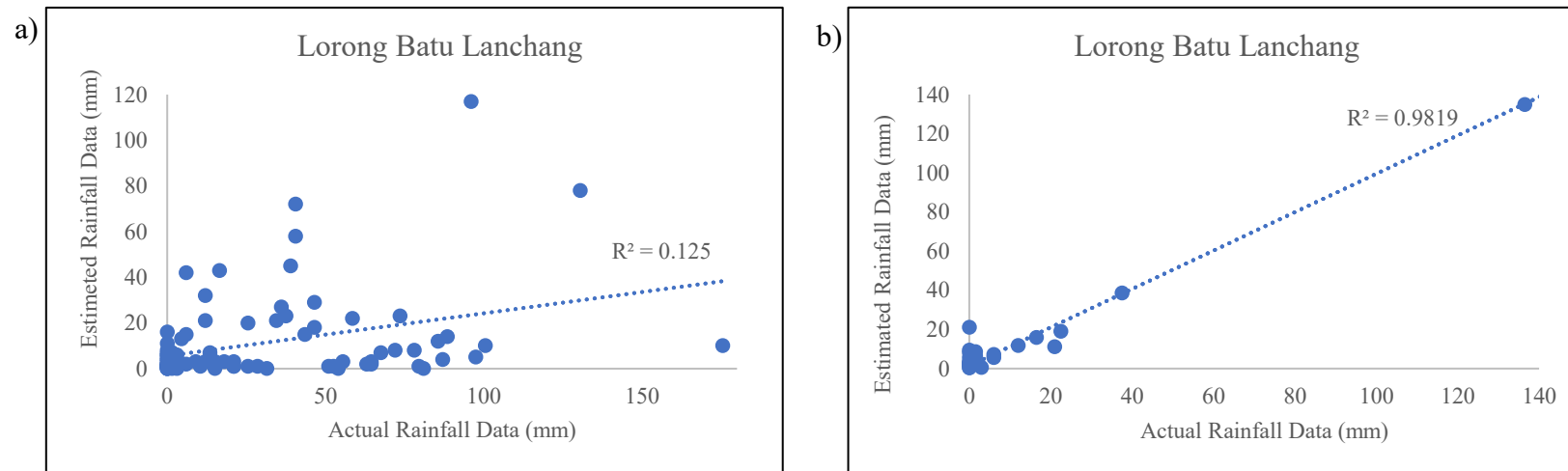


Figure 6: a) R^2 Value Before ANN Post-Processing b) R^2 Value After ANN Post-Processing

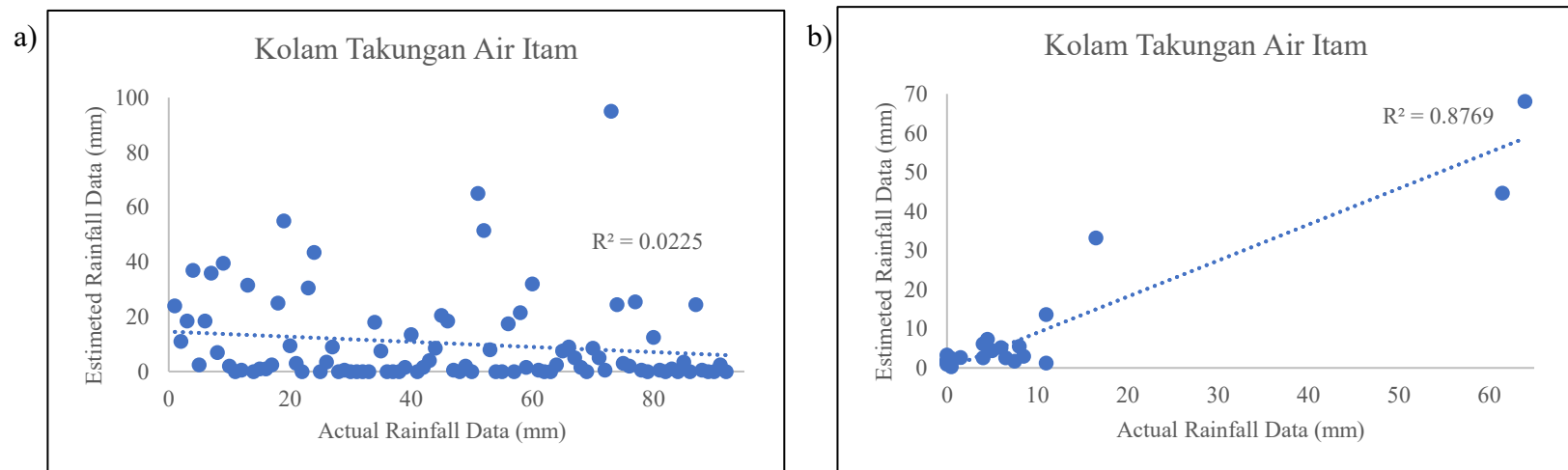


Figure 7: a) R^2 Value Before ANN Post-Processing b) R^2 Value After ANN Post-Processing

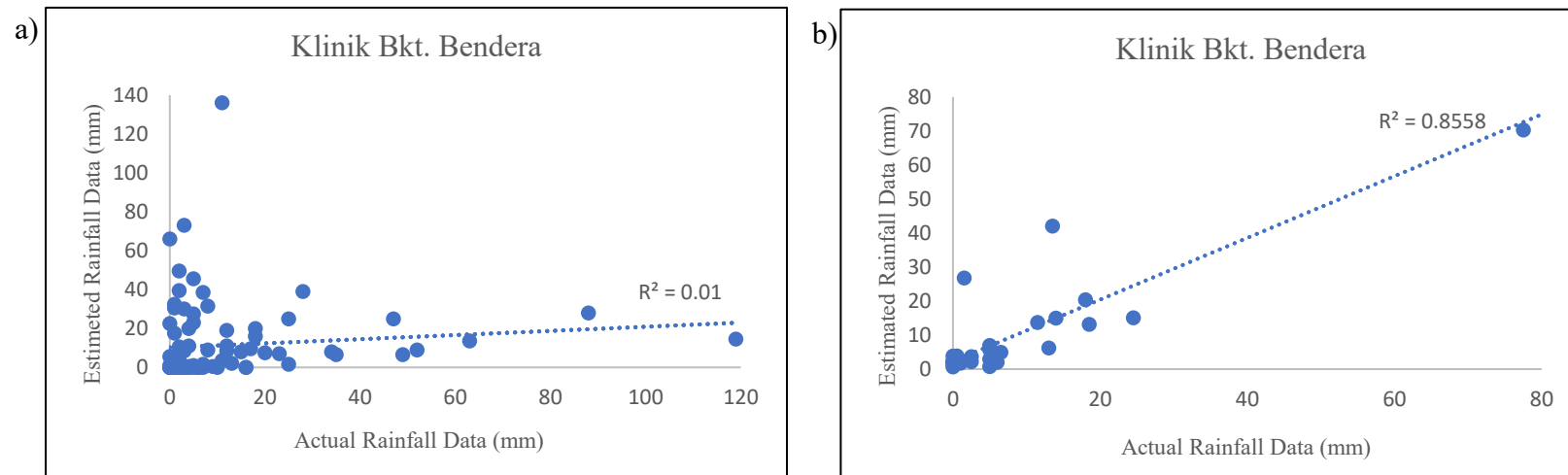


Figure 8: a) R^2 Value Before ANN Post-Processing b) R^2 Value After ANN Post-Processing

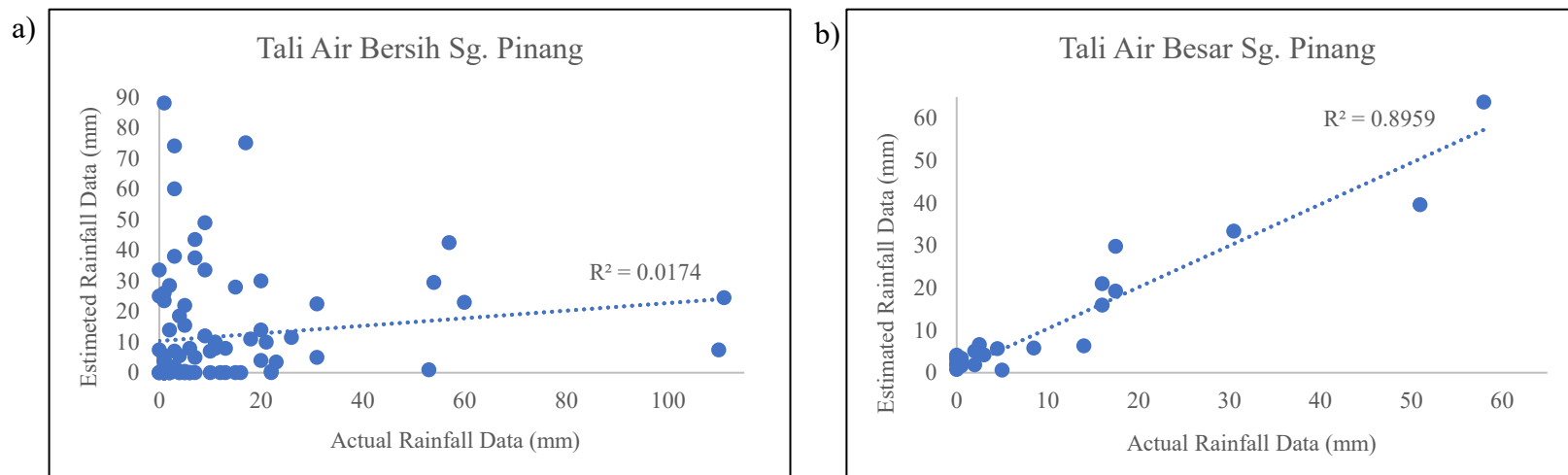


Figure 9: a) R^2 Value Before ANN Post-Processing b) R^2 Value After ANN Post-Processing

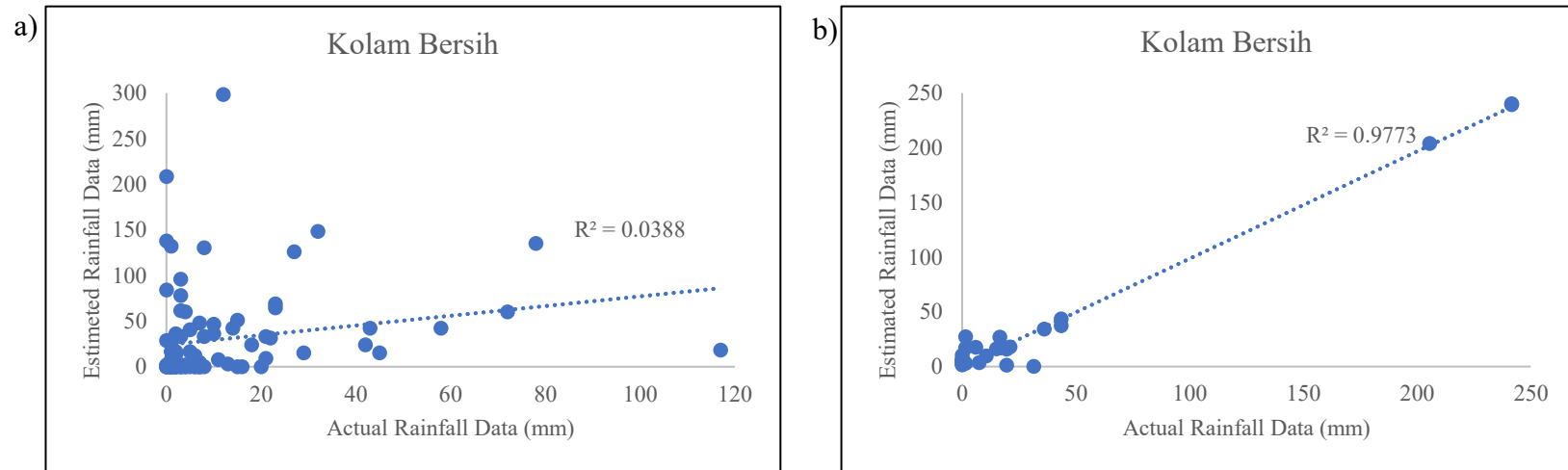


Figure 10: a) R^2 Value Before ANN Post-Processing b) R^2 Value After ANN Post-Processing

Table 3: Summary of Statistical Analysis Results for ANN Model Utilization

Rainfall station	r		R ²		RMSE		MAE	
	Before ANN	After ANN	Before ANN	After ANN	Before ANN	After ANN	Before ANN	After ANN
Lorong Batu Lanchang	0.354	0.814	0.125	0.9819	37.332	20.049	22.897	14.579
Kolam Takungan Air Itam	0.148	0.821	0.022	0.8769	22.634	9.85	12.929	6.281
Klinik Bkt. Bendera	0.1	0.75	0.01	0.8858	25.863	13.13	14.071	8.598
Tali Air Besar Sg. Pinang	0.132	0.742	0.017	0.8959	24.786	11.846	14.239	7.802
Kolam Bersih Pulau Pinang	0.197	0.557	0.039	0.9773	53.3	37.643	26.734	23.247

Limitation of Study

Although the ANN post-processing model significantly enhances PERSIANN satellite estimations, certain limitations must be addressed. Initially, the model conducted training and evaluation during high-stress hydrological phases from late 2018; extending this to multi-year continuous datasets will enhance the understanding of long-term climate variability and dry season dynamics. Secondly, inconsistencies in spatial resolution remain between satellite grid pixels and localized point rain gauges during brief, intense convective storms. Thus, although the 5–10–1 backpropagation demonstrate effective result, subsequent investigations can explore other deep learning methodologies to examine model adaption in different tropical urban catchments.

Conclusion

This study was conducted to improve the accuracy of rainfall estimation at the Sungai Pinang catchment by evaluating PERSIAN satellite data and enhance it using the ANN model. Five local rain gauge stations was analyzed and there is a significant spatial variation with peak rainfall recorded in October and November 2018. Initial comparison revealed that before ANN post-processing, the data correlation is poor ($r = 0.132 \sim 0.354$, $R^2 = 0.010 \sim 0.125$) and exhibits high error values (RMSE= 22.634 ~53.3, MAE=12.929~ 26.734). confirm the needs of post-processing. The integration satellite data with local meteorological parameters (temperature, humidity, and wind speed), the ANN model significantly improved performance metrics across all stations: r increased to $0.557 \sim 0.821$ after ANN post-processing, R^2 value improved to $0.8769 \sim 0.98$, RMSE decreased to $9.85 \sim 37.643$ and MAE was reduced to $6.281 \sim 23.247$. Notably, Station Kolam Takungan Air Itam achieved the highest r increase (+67.3%), while Station Kolam Bersih indicate the greatest R^2 increase (+93.83%) and MAE reduction. Station Lorong Batu Lanchang recorded the largest RMSE reduction.

The results demonstrated a significant accuracy enhancement with the application of ANN post-processing can minimize mistakes and increase data reliability. The proposed approach provides a reliable tool for improving flood forecasting, hydrological modelling, and water resource management in highly urbanized tropical catchments. This enhanced model can assist engineers, planners, and meteorological authorities in optimizing decisions about flood management, drainage design, and early warning systems in metropolitan locales such as Sungai Pinang.

-
- Acknowledgements:** The authors would like to express their sincere gratitude to Universiti Teknologi Mara (UiTM) for providing the necessary resources and support throughout the course of this research. Special appreciation is extended to colleagues and peers who contributed valuable insights and constructive feedback, which greatly enhanced the quality of this paper.
- Funding Statement:** No Funding
- Conflict of Interest Statement:** The authors declare that there is no conflict of interest regarding the publication of this paper. All authors have contributed to this work and approved the final version of the manuscript for submission to the International Journal of Information System and Technology Management (JISTM).
- Ethics Statement:** This study did not involve any human participants, animals, or sensitive data requiring ethical approval. The authors confirm that the research was conducted in accordance with accepted academic integrity and ethical publishing standards.
- Author Contribution Statement:** All authors contributed significantly to the development of this manuscript. Intan Shafeenar Ahmad Mohtar & Nur Azwa Binti Muhamad Bashar was responsible for the conceptualization, methodology, and overall supervision of the study. Anisatul Batrisyiah Binti Asmuri & Raja Nor Husna Raja Mohd Noor handled data collection, analysis, and interpretation of results. Nuryazmeen Farhan Haron & Juzailah Nur Md Yunus contributed to the literature review, drafting, and critical revision of the manuscript. All authors read and approved the final version of the manuscript prior to submission.
-

References

- Aprianto, R., Fitriyanto, S., & Nurul Walidain, S. (2023). Artificial neural network backpropagation for predicting rainfall (Case study in Sultan Muhammad Kaharuddin meteorological station). *Jurnal Ilmiah Multi Sciences*, 15(1), 63–70. <https://doi.org/10.30599/jti.v15i1.2110>
- Baig, F., Abrar, M., Chen, H., & Sherif, M. (2023). Evaluation of Precipitation Estimates from Remote Sensing and Artificial Neural Network Based Products (PERSIANN) Family in an Arid Region. *Remote Sensing*, 15(4). <https://doi.org/10.3390/rs15041078>
- Benziane, S. (2024). Survey: Rainfall Prediction precipitation, review of statistical methods. *WSEAS Transactions on Systems*, 23, 47–59. <https://doi.org/10.37394/23202.2024.23.5>
- Ceferino-Hernández, L., Magaña-Hernández, F., Campos-Campos, E., Morosan, G. A., Torres- Aguilar, C. E., Mora-Ortiz, R. S., & Díaz, S. A. (2024). Assessment of PERSIANN satellite products over the Tulijá River Basin, Mexico. *Remote Sensing*, 16(14). <https://doi.org/10.3390/rs16142596>
- Dao, V., Arellano, C. J., Nguyen, P., Almutlaq, F., Hsu, K., & Sorooshian, S. (2025). Bias correction of satellite precipitation estimation using deep neural networks and topographic information over the Western U.S. *Journal of Geophysical Research: Atmospheres*, 130(4). <https://doi.org/10.1029/2024JD042181>
- Eini, M. R., Rahmati, A., & Piniewski, M. (2022). Hydrological application and accuracy evaluation of PERSIANN satellite-based precipitation estimates over a humid continental climate catchment. *Journal of Hydrology:Regional Studies*, 41, 101109. <https://doi.org/10.1016/J.EJRH.2022.101109>
- Fuentes-Barrios, A., & Sierra-Lorenzo, M. (2021). Bias Correction Method Based on Artificial Neural Networks for Quantitative Precipitation Forecast. *Environmental Sciences Proceedings*, 8(1), 38. <https://doi.org/10.3390/ecas2021-10356>
- Grosan, C., Abraham, A., Jain, L. C., & Kacprzyk, J. (2011). *Intelligent systems* (Vol. 17, pp. 261-268). Berlin: Springer. <https://doi.org/10.1007/978-3-642-21004-4>
- Hsu, K. L., Gao, X., Sorooshian, S., & Gupta, H. V. (1997). Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks. *Journal of Applied Meteorology*, 36(9), 1176-1190. [https://doi.org/10.1175/1520-0450\(1997\)036<1176:PEFRSI>2.0.CO;2](https://doi.org/10.1175/1520-0450(1997)036<1176:PEFRSI>2.0.CO;2)
- Jabbari, A., & Bae, D. H. (2018). Application of Artificial Neural Networks for Accuracy Enhancements of Real-Time Flood Forecasting in the Imjin basin. *Water (Switzerland)*, 10(11). <https://doi.org/10.3390/w10111626>
- Kunnath-Poovakka, A., & Eldho, T. I. (2023). Bias correction of satellite precipitation products for hydrologic modeling in Western Ghats Region, India. *Journal of Hydrologic Engineering*, 28(4), 04023010. <https://doi.org/10.1061/JHYEFF.HEENG-5699>
- Latif, S. D., Alyaa Binti Hazrin, N., Hoon Koo, C., Lin Ng, J., Chaplot, B., Feng Huang, Y., El-Shafie, A., & Najah Ahmed, A. (2023). Assessing rainfall prediction models: Exploring the Advantages of Machine Learning and Remote Sensing Approaches. *Alexandria Engineering Journal*, 82, 16–25. <https://doi.org/10.1016/j.aej.2023.09.060>
- Lim, E. J., Misiran, S. E., Sapiri, M., & Yusof, H. (2024). Forecasting the monthly temperature and rainfall in Chuping Perlis. *EDUCATUM Journal of Science, Mathematics and Technology*, 11(2), 20. <https://doi.org/10.37134/ejsmt.vol11.2.3.2024>
- Moghim, S., & Bras, R. L. (2017). Bias correction of climate modeled temperature and precipitation using artificial neural networks. *Journal of Hydrometeorology*, 18(7), 1867–1884. <https://doi.org/10.1175/JHM-D-16-0247.1>

- Mohtar, I. S. A., Tahir, W., & Yusof, M. S. M. (2025a). Enhanced Flood Forecasting using Quantitative Precipitation Forecast from Weather Research Forecast-Artificial Neural Network (WRF-ANN) Model Post Processing. *Journal of Water Resources Management*, 3(1).
- Mohtar, I. S. A., Tahir, W., Jani, J., Mohamad, I. N., Deni, S. M., & Haron, N. F. (2025b). ANN Post-Processing of WRF Model Outputs for Enhanced Rainfall and Flood Forecasting. In 2025 IEEE International Conference on Communication, Networks and Satellite (COMNETSAT) (pp. 409-415). <https://doi.org/10.1109/COMNETSAT68601.2025.11324473>
- Nandakumar, S. D., Valarmathi, R., Sudha Juliet, P., & Brindha, G. (2021). Artificial Neural Network for Rainfall Analysis using Deep Learning Techniques. *Journal of Physics: Conference Series*, 1964(4). <https://doi.org/10.1088/1742-6596/1964/4/042022>
- Papacharalampous, G., Tyralis, H., Doulamis, A., & Doulamis, N. (2023). Comparison of Machine Learning Algorithms for Merging Gridded Satellite and Earth-Observed Precipitation Data. *Water (Switzerland)*, 15(4). <https://doi.org/10.3390/w15040634>
- Parsons, D., Stern, D., Ndanguza, D., Sylla, M. B., Musyoka, J., Bagiliko, J., Clarkson, G., Dorward, P. (2026). Evaluating Satellite and Reanalysis Rainfall Estimates For Climate Services in Agriculture: A Comprehensive Methodology. *Meteorological Applications*, 33(2), e70170.
- Pratomo, A. H., Santosa, B., Tahalea, S. P., Paripurno, E. T., Peasetyo, J. D., Jayadianti, H., & Pitayandanu, M. F. (2022). Rainfall Prediction using Artificial Neural Network with Historical Weather Data as Supporting Parameters. *Jurnal Informatika*, 16(2), 63. <https://doi.org/10.26555/jifo.v16i2.a25422>
- Qamar, R., & Ali Zardari, B. (2023). Artificial Neural Networks: An Overview. *Mesopotamian Journal of Computer Science*, 2023, 130–139. <https://doi.org/10.58496/MJCSC/2023/015>
- Rajput, N., Kasimahanthi, A., Reddy, V., Fatima, H., Devarajan, P. K., & Prasad, V. S. (2026). Evaluation and Error Correction of Satellite Rainfall Using Dense Network of Rain Gauges Over Mumbai. *Journal of the Indian Society of Remote Sensing*, 1-14.
- Ridwan, W. M., Sapitang, M., Aziz, A., Kushiar, K. F., Ahmed, A. N., & El-Shafie, A. (2021). Rainfall Forecasting Model using Machine Learning Methods: Case Study Terengganu, Malaysia. *Ain Shams Engineering Journal*, 12(2), 1651–1663. <https://doi.org/10.1016/j.asej.2020.09.011>
- Sadeghi, M., Nguyen, P., Naeini, M. R., Hsu, K., Braithwaite, D., & Sorooshian, S. (2021). PERSIANN- CCS-CDR, a 3-hourly 0.04° Global Precipitation Climate Data Record for Heavy Precipitation Studies. *Scientific Data*, 8(1). <https://doi.org/10.1038/s41597-021-00940-9>
- Saleh, A., Sabtu, N., & Bunmi, M. R. (2022). Flash flood Susceptibility Mapping in Sungai Pinang Catchment Using Weight of Evidence. *IOP Conference Series: Earth and Environmental Science*, 1091(1). <https://doi.org/10.1088/1755-1315/1091/1/012017>
- Sivaji, V. (2025). Sungai Pinang flood mitigation project 27% complete, on track for 2027 completion. <https://penanghyperlocal.com/sungai-pinang-flood-mitigation-project-27-complete-on-track-for-2027-completion/>. Accessed on 20 June 2026.
- Sumith, K. V. (2025). Comprehensive Evaluation of Satellite-Based Rainfall Measurements Through Rain Gauge Validation Using Advanced Statistical Regression and Machine Learning Models by Using Python: KV Sumith. *Water Resources Management*, 39(9), 4563-4587.
- Thakur, N., Karmakar, S., & Soni, S. (2021). Rainfall forecasting using various artificial neural network techniques. A review. *International Journal of Scientific Research in Computer*

Science, Engineering and Information Technology, 403–409.
<https://doi.org/10.32628/cseit2173159>.

- Yee, L. Z., Tan, K. W., & Huang, Y. F. (2022). Statistical downscaling of regional climate model: A case study in Penang Island, Malaysia. *Global Nest Journal*, 24(2), 267–275.
<https://doi.org/10.30955/gnj.003429>
- Yusoff, I. M., & Rindam, M. (2012). 3D GIS Urban Runoff Mechanism: A New Perspective using Volumetric Soft Geo-Object. *Geografia*, 8(5). *Geografia OnlineTM Malaysia Journal of Society and Space* 8 issue 5 (124- 139). Themed Issue: Masyarakat, Ruang dan Alam Sekitar (MATRA), ISSN 2180-2491.