

# Rainfall Prediction in the Krueng Pase Watershed Using Support Vector Regression

**Ichwana Ramli**

Department of Agricultural Engineering, Universitas Syiah Kuala, Banda Aceh, Indonesia | Department of Environmental Master's Program, Universitas Syiah Kuala, Banda Aceh, Indonesia  
ichwana.ramli@usk.ac.id

**Teuku Ferijal**

Department of Agricultural Engineering, Universitas Syiah Kuala, Banda Aceh, Indonesia  
t.ferijal@usk.ac.id (corresponding author)

**Siti Rusdiana**

Department of Mathematics and Natural Science, Universitas Syiah Kuala, Banda Aceh, Indonesia  
siti.rusdiana@usk.ac.id

Received: 25 July 2025 | Revised: 24 September 2025 | Accepted: 9 October 2025

Licensed under a CC-BY 4.0 license | Copyright (c) by the authors | DOI: <https://doi.org/10.48084/etasr.13635>

## ABSTRACT

In the present work, the Support Vector Regression (SVR) model was employed to predict monthly rainfall amounts in the Krueng Pase Watershed, Indonesia, for the period 1992-2020. The meteorological factors considered were: the number of rainy days, temperature, humidity, solar exposure, and wind velocity. Three SVR kernel functions were applied (linear, Radial Basis Function (RBF), and polynomial). A comparison was conducted using an 80/20 training-testing split together with a grid search for optimization. The model performance was evaluated by means of the Root Mean Square Error (RMSE) and the coefficient of determination ( $R^2$ ). The results indicated that the polynomial kernel exhibited the best performance with an RMSE of 57.71 mm and  $R^2 = 0.25$ . The Kendall-Tau analysis revealed that the number of rainy days and humidity were the most significant positive predictors, whereas temperature and solar exposure had negative impacts. The moderate predictive skill indicated that large-scale climatic drivers were missing, yet it confirmed that SVR provides a strong, data-driven method for regional rainfall forecasting.

**Keywords-rainfall prediction; Support Vector Regression (SVR); Krueng Pase; Machine Learning (ML)**

## I. INTRODUCTION

Climate change is a global phenomenon that significantly affects both environmental and hydrological systems. High atmospheric moisture and precipitation levels are some of its consequences worldwide. In particular, global warming has contributed to the melting of polar ice caps, leading to rising sea levels and subsequent coastal flooding [1].

Regarding the frequency and intensity of extreme rainfall events, both the quantity and quality of available water is influenced [2]. The direct effects of climate change are caused by changes in precipitation and temperature, while changes in plant physiology represent indirect effects. For example, rising temperatures may reduce streamflow by increasing Evapotranspiration (ET), although in some cases this can be controlled by an increased annual precipitation.

Rainfall, one of the primary climatic variables, plays a key role in water resource management. Accurate rainfall prediction is essential for ensuring adequate water supply for

agriculture, human consumption, and flood prevention. In practice, it is useful for flood early warning, reservoir operation, and irrigation scheduling, thereby linking forecasting directly to hydrologic decision-making and disaster preparedness [3-5]. Various models have been developed to predict the rainfall patterns with successful results, such as the Seasonal Autoregressive Integrated Moving Average (SARIMA) [6]. However, traditional linear models like linear regression have usually been insufficient when incorporating variables like temperature, humidity, and solar exposure [7].

To address these limitations, Machine Learning (ML) models, including Support Vector Machines (SVMs), have been widely adopted. SVMs, introduced by Vapnik in the late 1960s, are Artificial Intelligence (AI) methods well-regarded for their effectiveness in pattern recognition, optimization, and generalization. A major strength is their training process, in which convex optimization problems need to be solved to obtain optimal solutions. Additionally, SVM employs the principle of inductive learning, which minimizes the structural

risk to achieve a balance between model complexity and generalization performance on unseen data.

SVM gained attention in 1995 with applications in both classification - known as Support Vector Classification (SVC) - and regression, referred to as SVR [8]. SVR is utilized for solving nonlinear regression problems with large datasets and continuous outputs [9]. It basically mitigates overfitting while producing accurate estimations. Recent applications of SVR in environmental modeling tasks, such as urban heat estimation, have revealed that when equipped with proper kernel functions and tuned hyperparameters, the model can achieve high performance in non-linear systems [10].

Watersheds are vital for maintaining the stability of regional water supplies. Any alterations within them, especially those driven by climate changes, can significantly influence the discharge patterns on both seasonal and annual timescales [11, 12]. Among these changes, inter- and intra-annual rainfall variability (e.g., ENSO/IOD/MJO-related) plays a major role in shaping the seasonal discharge characteristics [13]. For these reasons, it is important to analyze the hydrological behavior of specific watersheds, such as the Krueng Pase, in order to better understand these dynamics.

Previous studies on rainfall distribution in Indonesia have indicated significant spatiotemporal clustering, reflecting a high variability in precipitation patterns that must be captured in predictive modeling [14]. As a result, a great amount of research focuses on the assessment of the climate change effects on river discharge. The turning of rainfall forecasts into useful watershed information, such as predicting months with high flows for flood preparedness, defining criteria for releasing water from reservoirs, and planning irrigation, could create a direct link between forecasting and management results [15, 16].

The Krueng Pase River, which serves as a critical water source for irrigation, is heavily influenced by the physical and

environmental characteristics of its surrounding watershed. During periods of high rainfall, it is prone to overflowing, causing damage to nearby residential areas. This flooding is further intensified by the river's naturally meandering course and its gentle slope, both of which slow the water flow and increase the flood risk during intense rain events. Water resource management remains a significant concern as the inadequate control of the water flow can disrupt agriculture and endanger local communities. Moreover, the area's agricultural sector is highly dependent on water availability, and mismanagement can lead to both drought and excessive runoff. Similar challenges have been observed in other flood-prone watersheds in Indonesia, where spatial risk assessments and hydrological modeling have proven essential for mitigation planning [15]. In this context, reliable monthly rainfall estimates are useful for predicting the risk of flooding, allocating water to agriculture, and informing community early-warning practices in the Krueng Pase watershed.

This study aims to analyze the characteristics of the rainfall patterns and the factors influencing these changes within the Krueng Pase Watershed, located in North Aceh Regency. SVR was selected to examine the relationship between rainfall and its influencing variables, including the number of rainy days, temperature, humidity, solar exposure, and wind velocity.

## II. RESEARCH METHODS

This study was conducted in the Krueng Pase Watershed, located between  $5^{\circ}09'12''$ - $4^{\circ}49'25''$  North Latitude and  $96^{\circ}51'27''$ - $97^{\circ}14'55''$  East Longitude, involving areas within Lhokseumawe City, North Aceh District, and Bener Meriah District (Figure 1). Meteorological data were obtained from the Meteorology, Climatology, and Geophysics Agency (BMKG), consisting of the monthly rainfall amount (mm), number of rainy days (days), and monthly mean values of temperature ( $^{\circ}\text{C}$ ), humidity (%), solar exposure (hours/day), and wind velocity (knots), for the period of January 1992 to December 2020.

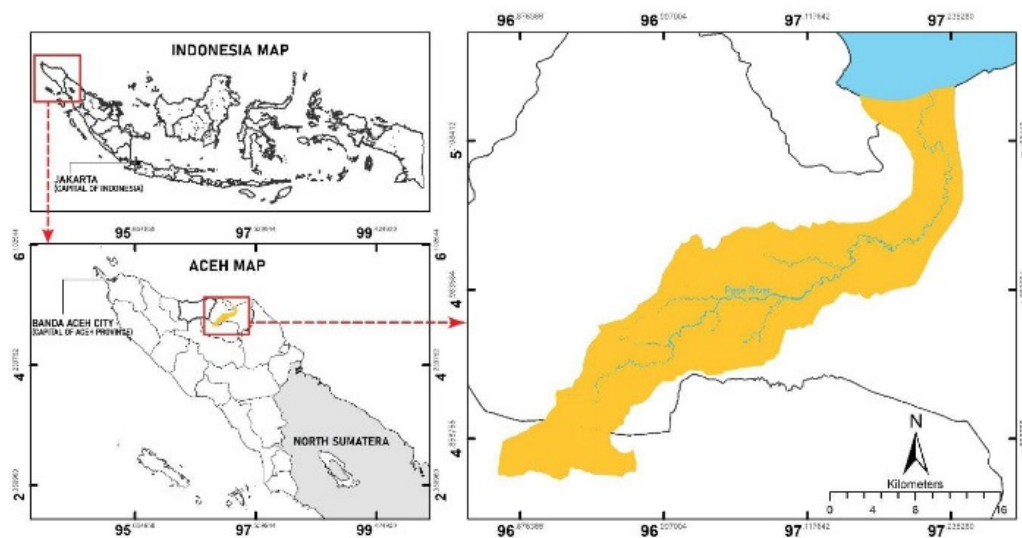


Fig. 1. Map of the Krueng Pase watershed area as the research location.

The research methodology comprised some key stages:

- Descriptive statistical analysis to examine the central tendency, dispersion, and distribution characteristics of each variable.
- Data preprocessing, including the handling of missing values and normalization.
- Correlation analysis using the Kendall-Tau method to identify the relationships between rainfall and climatic variables.
- Multicollinearity testing to ensure the independence of the predictors.
- Development of a predictive model using SVR to estimate rainfall based on the selected climatic predictors.

The dataset was divided into two parts: 80% for training and 20% for testing.

All meteorological variables were subjected to a preprocessing stage to ensure data quality and numerical stability. The missing values were handled by linear interpolation for gaps shorter than three consecutive months, while longer gaps were excluded from the analysis. To standardize the predictors and improve kernel computation, all input variables were normalized using Min-Max scaling to the range [0, 1], as defined by:

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

where  $X$  is the original observation,  $X_{min}$  and  $X_{max}$  are the minimum and maximum values of the variable, respectively, and  $X'$  is the normalized value. This transformation ensured that no single variable dominates the regression process due to differences in measurement scales.

The target variable (monthly rainfall amount in mm) was retained in its original scale during model training and evaluation to preserve physical interpretability. However, for the purpose of generating partial residual plots, a natural logarithmic transformation of the rainfall values was applied. This step was implemented only in the visualization stage to enhance interpretability by stabilizing variance and emphasizing nonlinear relationships between the predictors and the response. All analyses used Python 3.10, scikit-learn 1.2.2, NumPy 1.26, pandas 2.1, statsmodels 0.14, matplotlib 3.8. Seeds were fixed and scalars fit on training folds only.

#### A. Kendall-Tau Correlation Analysis

The Kendall rank correlation analysis was used to identify the relationships and test hypotheses between two or more variables, particularly when the data were ordinal or ranked [17]. This method is applicable for analyzing samples larger than 10 and can be extended to calculate partial correlation coefficients. The equation used for Kendall rank correlation  $\tau$  is [18]:

$$\tau = 2 \left( \frac{N_c - N_d}{N(N-1)} \right) \quad (2)$$

where  $N_c$  is the number of concordant pairs,  $N_d$  is the number of discordant pairs, and  $N$  is the sample size. For large sample

sizes ( $N > 10$ ),  $\tau$  is approximately normally distributed. The test statistic is calculated using:

$$Z = \frac{\tau \sqrt{2(2N+5)}}{3\sqrt{(N-1)}} \quad (3)$$

The null hypothesis ( $H_0$ ) of no relationship between the two variables is rejected if the  $p$ -value is less than a significant level of 5% ( $\alpha = 0.05$ ).

#### B. Support Vector Regression

SVR was employed to model the relationship between the response variable and the selected climatic predictors. It is an extension of SVM mainly used for regression tasks, by estimating a function that maps the predictors to the target variable within a specified margin of error while minimizing model complexity [19-21]:

$$f(x) = w^T \phi(x) + b \quad (4)$$

where  $\phi(x)$  maps the input features into a higher-dimensional space,  $w$  is the weight vector, and  $b$  is the bias term.

The residuals are defined as:

$$r = y - f(x) \quad (5)$$

SVR uses an  $\epsilon$ -insensitive loss function, where deviations within  $\pm \epsilon$  from the actual value are tolerated and not penalized [22]:

$$L_\epsilon(y, f(x)) = \begin{cases} 0, & \text{if } |y - f(x)| \leq \epsilon \\ |y - f(x)| - \epsilon, & \text{otherwise} \end{cases} \quad (6)$$

The optimization objective was to minimize the flatness of the regression function and penalize deviations outside the  $\epsilon$  margin. The SVR optimization problem was formulated as shown in (7)-(10), where the objective function and constraints defined the  $\epsilon$ -insensitive loss with slack variables:

$$\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l (\xi_i + \xi_i^*) \quad (7)$$

subject to the constraints:

$$y_i - w^T \phi(x_i) - b \leq \epsilon + \xi_i \quad (8)$$

$$w^T \phi(x_i) + b - y_i \leq \epsilon + \xi_i^* \quad (9)$$

$$\xi_i, \xi_i^* \geq 0 \quad (10)$$

where  $\xi_i$  and  $\xi_i^*$  are slack variables that allow for errors beyond the  $\epsilon$  margin, and  $C > 0$  is a regularization parameter that controls the trade-off between the model complexity and tolerance for error.

#### C. Kernel Function

The kernel method involved four key aspects. Initially, the data were mapped into a Euclidean feature space, followed by the search for linear relationships within this feature space. Third, when implementing the algorithm, only the inner products between the vectors in the feature space were required. Finally, these products were efficiently computed directly from the original data using what is known as the kernel trick [23]. In this study, three kernels were used to define  $\phi(x)$ :

- Linear Kernel: Used to model the linear relationships between variables:

$$K(x_i, x_j) = x_i^T x_j \quad (11)$$

- Polynomial Kernel: A non-linear kernel that can model more complex patterns with varying degrees:

$$K(x_i, x_j) = (\gamma x_i^T x_j + r)^d x_i^T x_j \quad (12)$$

- RBF Kernel: Kernel suitable for capturing the non-linear relationships by mapping data into higher-dimensional spaces:

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2) \quad (13)$$

These kernel functions allowed SVR to capture various types of relationships between the climatic variables and rainfall. Parameter tuning was conducted using a grid search to identify the optimal values for  $C$ ,  $\epsilon$ , and kernel-specific parameters, such as  $\gamma$  and polynomial degree  $d$ .

TABLE I. DESCRIPTIVE STATISTICS OF METEOROLOGICAL VARIABLES IN THE KRUENG PASE WATERSHED (1992–2020)

Variable	Mean	Standard deviation	Median	Skewness	Kurtosis
Rainfall (mm/month)	118.3	89.45	93.7	1.53	3.16
Number of rainy days (days/month)	14.557	5.691	14	0.15	-0.55
Temperature (°C)	26.782	1.125	26.7	2.66	13.28
Humidity (%)	81.338	3.876	81.19	-0.18	-0.14
Wind velocity (knots)	125.27	54.55	120	1.48	5.28
Solar exposure (hours/day)	6.2704	1.2514	6.4	-0.57	0.93

The number of rainy days was more stable, averaging 14.56 days per month with near-symmetric distribution, reflecting typical wet-season patterns. Temperature displayed high skewness (2.66) and kurtosis (13.28), pointing to rare but intense heat events. In contrast, humidity was more normally distributed, demonstrating year-round atmospheric moisture stability conducive to precipitation. The wind velocity and solar exposure varied from moderate to high. The average wind velocity was 125.27 knots, and solar exposure was 6.27 h/day. Both factors affected evaporation and convective processes.

### B. Correlation Analysis

Kendall Tau correlation coefficients (Figure 2) were computed to assess the monotonic relationships between rainfall and its influencing variables. Only the statistically significant Kendall's  $\tau$  values ( $p < 0.05$ ) were visualized. Rainfall was positively correlated with the number of rainy days ( $\tau = 0.401$ ) and humidity ( $\tau = 0.323$ ), suggesting that increased rainfall was generally associated with more frequent wet days and elevated moisture levels. Meanwhile, negative correlations were observed between rainfall and temperature ( $\tau = -0.204$ ) and solar exposure ( $\tau = -0.235$ ), meaning that drier and sunnier conditions typically reduced rainfall accumulation. The wind velocity showed no significant relationship with rainfall, implying a limited direct influence. These findings aligned with prior work in tropical monsoon regions, where the rainfall patterns were more closely influenced by the thermal uplift, humidity, and cloud cover rather than surface winds [25, 26].

### D. Model Selection

RMSE was used to evaluate SVR performance [24]:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (14)$$

The SVR model with the lowest RMSE on the test dataset was selected as the best model for rainfall prediction.

## III. RESULTS AND DISCUSSION

### A. Statistical Descriptive of Data

The descriptive statistics for the Krueg Pase Watershed (1992-2020) indicated large month-to-month variation in rainfall. As shown in Table I, the monthly rainfall amount exhibited a mean of 118.3 mm and a high standard deviation of 89.45 mm. The distribution was positively skewed (1.53) with leptokurtic characteristics (3.16), implying frequent extreme rainfall events.

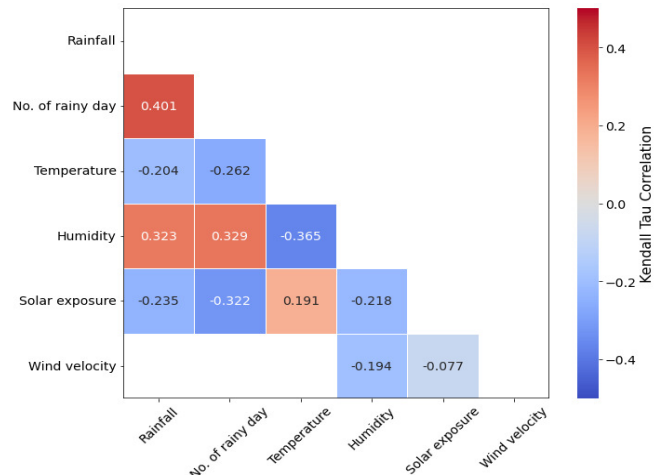


Fig. 2. Kendall's  $\tau$  correlation heatmap between monthly rainfall (mm) and predictors (1992–2020). Only statistically significant  $\tau$  values ( $p < 0.05$ ) were depicted.

### C. Model Diagnostics

To ensure reliable regression modeling, a multicollinearity test was performed. All variables exhibited Variance Inflation Factor (VIF) values below 10, confirming the absence of multicollinearity. Residual diagnostics were conducted using a Residuals versus the Fitted plot (Figure 3) revealed mild curvature and scattered outliers, suggesting potential nonlinearity in the data. The residual pattern implied a violation of the linearity assumption and the presence of heteroscedasticity.

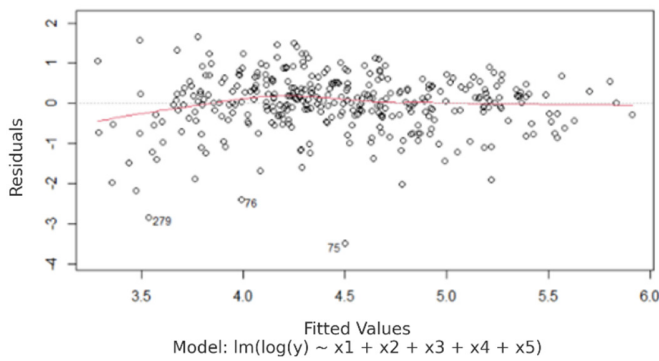


Fig. 3. Residuals versus fitted.

D. Partial Residual Analysis

Partial residual plots (Figure 4) were used to explore the marginal relationship between each predictor and the log-transformed rainfall response, adjusting for the effects of other variables. This transformation was applied only for visualization purposes to improve interpretability and reveal potential nonlinear patterns.

The plot for the number of rainy days demonstrated a strongly upward curve, suggesting a nonlinear positive influence. Temperature showed a flatter trendline, indicating a relatively weaker effect. Humidity and wind velocity exhibited slightly nonlinear patterns, with the former showing a consistent positive trend and the latter displaying a modest nonlinear effect despite its negative overall correlation. Solar exposure displayed a downward trend with mild curvature, reinforcing its negative relationship with the response. These patterns confirmed the existence of nonlinear effects.

The correlation structure, multicollinearity test, and residual diagnostics collectively revealed that, although the selected

predictors were statistically valid and relevant, their relationships with rainfall deviated from strict linearity. These findings supported the use of SVR to explore the nonlinear predictor-response patterns through kernel functions.

E. Support Vector Regression Analysis

SVR models were developed using three kernel functions for the prediction of the monthly rainfall amount (mm). Hyperparameter optimization for each function was conducted using a grid search approach, focusing on minimizing the RMSE and maximizing the coefficient of determination ( $R^2$ ) on the training dataset.

The SVR model with a linear kernel, using a cost parameter ( $C$ ) of 1, produced an RMSE of 60.78 mm and an  $R^2$  value of 0.22. This indicated a limited ability to capture the nonlinear characteristics evident in the data. In contrast, the RBF kernel achieved better performance with a minimum RMSE of 59.21 mm and an  $R^2$  of 0.21 when the gamma parameter was set to 0.07 and the cost parameter to 0.25. This improvement reflected the RBF's ability to map data into higher-dimensional space, thus better modeling the underlying nonlinear patterns [27].

Among all tested configurations, the polynomial kernel performed better. The optimal configuration - gamma of 0.10, cost of 1, and a polynomial degree of 1 - yielded an RMSE of 58.71 mm and an  $R^2$  value of 0.25. This superior behavior confirmed the polynomial kernel's flexibility in capturing both the linear and nonlinear relationships, which aligned with the diagnostic findings from the residual analysis and partial residual plots. Prior applications of SVR agreed with these findings as they showed improved predictive performance, including flash-flood forecasting [15], environmental monitoring [19], and time series forecasting in commodity markets [23].

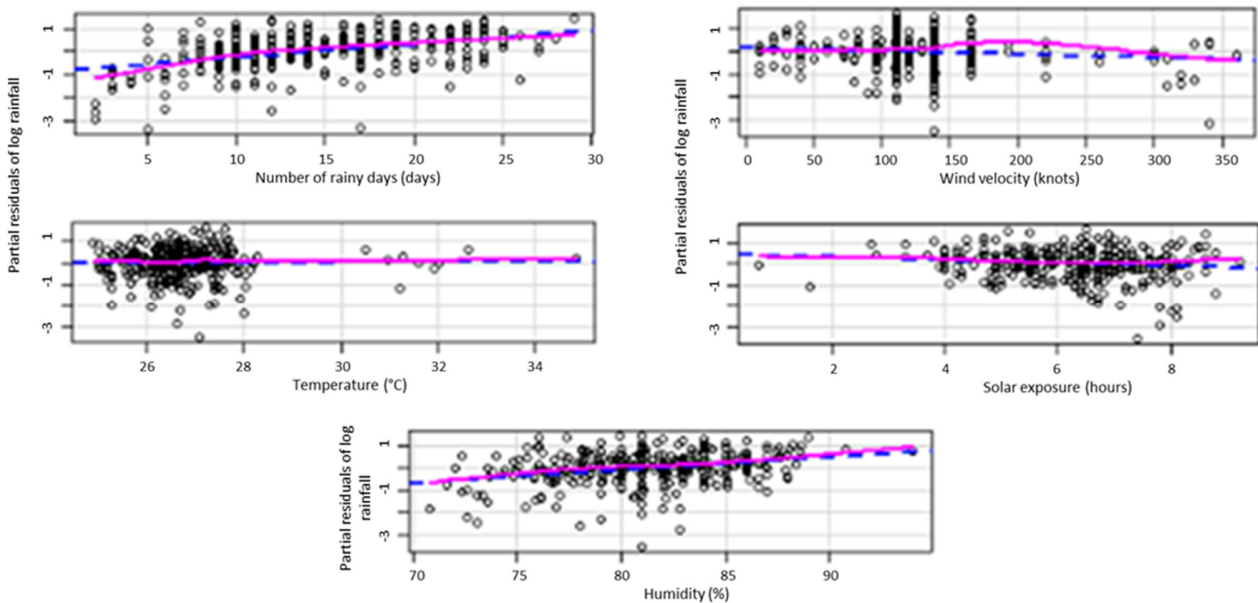


Fig. 4. Partial residual plots for each predictor using log-transformed rainfall.

Using permutation importance, the number of rainy days and humidity ranked as the most influential predictors, while temperature and solar exposure showed smaller negative contributions; wind velocity exhibited minimal importance. This ranking was consistent with the Kendall-Tau correlation analysis emphasizing the dominant role of humidity and rainfall frequency in shaping the precipitation outcomes.

#### F. Final Model Selection and Testing Accuracy

Model generalizability was assessed using independent test data. Table II presents a comparison between the SVR kernel functions based on RMSE. The polynomial kernel achieved the lowest RMSE (57.71 mm), followed by the linear (57.85 mm) and the RBF (58.09), confirming its suitability for predicting the monthly rainfall amount in the Krueng Pase Watershed.

TABLE II. COMPARISON OF SVR KERNELS BASED ON TEST SET RMSE

SVR kernels	RMSE (mm)
Linear	57.85
RBF	58.09
Polynomial	57.71

To further ensure the validity of these results, 5-fold cross-validation tests were conducted. The cross-validation yielded a relatively small standard deviation of RMSE ( $\pm 1.2$  mm), confirming the consistency of the model's performance. The residual analysis on the test dataset showed a random dispersion around zero without a systematic pattern, although several outliers were observed during extreme rainfall months. This was consistent with the moderate  $R^2$  values ( $\sim 0.21$ - $0.25$ ), which suggest that additional explanatory variables beyond those included in this study may matter. Furthermore, a paired  $t$ -test on the RMSE differences across kernels indicated no statistically significant differences at  $\alpha = 0.05$ , implying that the advantage of the polynomial kernel is relatively small in practical terms.

Even though a complex SVR model with a polynomial kernel was used, the low predictive skill ( $R^2 = 0.2$ ) for the monthly rainfall arose because monthly-averaged local meteorological variables are not reliable indicators of the sub-monthly convective thresholds and large-scale climate drivers that fundamentally govern precipitation accumulation in the region [28, 29]. The model inputs may not include key signals from major processes like the Madden-Julian Oscillation and the Indian Ocean Dipole. As a result, the model struggles to capture the complex and changing relationships needed for accurate prediction.

#### G. Summary, Implications, and Limitations

This study concluded that SVR could be an effective method for predicting the monthly rainfall amount in the Krueng Pase Watershed. Among the three kernel types evaluated, the polynomial kernel consistently produced the lowest RMSE values in both the training and testing phases, indicating superior model performance in capturing the nonlinear relationships among meteorological variables.

The ability of SVR to accommodate both linear and nonlinear structures made it a valuable tool for regional water

resource planning, early warning systems, and climate-resilient agricultural decision-making. Comparable SVR applications in tropical/monsoonal settings reported similarly modest explanatory power for the monthly rainfall amount. For example, in northwest India, it was found that an RBF-kernel SVR performed best, whereas in the current case, the polynomial kernel yielded the lowest error [30]. This consistency across regions underscored the inherent difficulty of predicting month-aggregated rainfall in the tropics.

However, there are several limitations. First, the model's performance still yielded relatively moderate  $R^2$  values, suggesting that additional unexplained variance remains. This could be attributed to unmeasured environmental variables, such as ET, land cover, or sea surface temperature, which were not included in the analysis. Second, while SVR is effective for small- to medium-sized datasets, its computational cost increases significantly with larger datasets, which may limit scalability in operational settings. At the same time, this study used static hyperparameter tuning via grid search. Adaptive or automated optimization methods, such as Bayesian optimization, may further enhance performance. Finally, model interpretability remains a challenge in SVR, particularly for stakeholders who require transparency in prediction logic for policy or disaster mitigation planning.

#### IV. CONCLUSION

This study applied Support Vector Regression (SVR) to predict the monthly rainfall amounts in the Krueng Pase Watershed. The dataset consisted of meteorological data, such as the number of rainy days, temperature, humidity, solar exposure, and wind velocity for the period of January 1992 to December 2020. The modeling framework included data preprocessing, Kendall-Tau correlation analysis, multicollinearity tests, as well as model testing with three kernel functions (linear, Radial Basis Function (RBF), and polynomial).

The findings revealed that rainfall distribution had significant variations over time and nonlinear relationships with climate variables. Among the different SVR models, the polynomial kernel exhibited the lowest Root Mean Square Error (RMSE) of 57.71 mm and an  $R^2$  of 0.25, thus surpassing both linear and RBF kernels in terms of performance. Besides, model diagnostics and partial residual plots indicated that nonlinear relationships - especially the one between rainfall and humidity - were much more effectively captured by the polynomial kernel.

Future studies should consider incorporating additional spatial and temporal predictors, testing ensemble or hybrid models, and evaluating model interpretability to enhance the usability of Machine Learning (ML) methods in environmental planning and disaster resilience efforts.

#### DATA AVAILABILITY

The monthly BMKG data (1992–2020) used in this study were obtained under BMKG's data-use policy. The dataset is available from the authors upon reasonable request and is subject to BMKG's data-use terms.

## ACKNOWLEDGMENT

The authors gratefully acknowledge funding from the Ministry of Education, Culture, Research, and Technology of Indonesia, facilitated through the Research and Community Service Institute (LPPM) of Universitas Syiah Kuala, under Grant No. 91/UN11.2.1/PG.01.03/SPK/PTNBH/2024. The authors also express their appreciation to the Indonesian Agency for Meteorology, Climatology, and Geophysics (BMKG) for generously supplying the rainfall data that formed the basis of this research.

## REFERENCES

- [1] R. Lindsey, "Climate Change: Global Sea Level." Climate.gov, <https://www.climate.gov/news-features/understanding-climate/climate-change-global-sea-level>.
- [2] G. Sun, S. G. McNulty, J. A. Moore Myers, and E. C. Cohen, "Impacts of Multiple Stresses on Water Demand and Supply Across the Southeastern United States," *JAWRA Journal of the American Water Resources Association*, vol. 44, no. 6, pp. 1441–1457, 2008, <https://doi.org/10.1111/j.1752-1688.2008.00250.x>.
- [3] J. I. Kuwajima, F. M. Fan, D. Schwanenberg, A. Assis Dos Reis, A. Niemann, and F. F. Mauad, "Climate change, water-related disasters, flood control and rainfall forecasting: a case study of the São Francisco River, Brazil," *Geological Society, London, Special Publications*, vol. 488, no. 1, pp. 259–276, Apr. 2019, <https://doi.org/10.1144/SP488-2018-128>.
- [4] D. Chitwatkulsiri and H. Miyamoto, "Real-Time Urban Flood Forecasting Systems for Southeast Asia—A Review of Present Modelling and Its Future Prospects," *Water*, vol. 15, no. 1, Jan. 2023, Art. no. 178, <https://doi.org/10.3390/w1510178>.
- [5] Md. M. Hassan *et al.*, "Machine Learning-Based Rainfall Prediction: Unveiling Insights and Forecasting for Improved Preparedness," *IEEE Access*, vol. 11, pp. 132196–132222, Nov. 2023, <https://doi.org/10.1109/ACCESS.2023.3333876>.
- [6] I. Ramli, S. Rusdiana, Ayulianur, and A. Achmad, "Comparisons among rainfall prediction of monthly rainfall basis data in Aceh using an autoregressive moving average," in *International Conference on Agricultural Technology, Engineering and Environmental Sciences*, Banda Aceh, Indonesia, Aug. 2019, <https://doi.org/10.1088/1755-1315/365/1/012008>.
- [7] I. Ramli, H. Basri, A. Achmad, R. Basuki, and M. A. Nafis, "Linear regression analysis using log transformation model for rainfall data in water resources management Krueng Pase, Aceh, Indonesia," *International Journal of Design & Nature and Ecodynamics*, vol. 17, no. 1, pp. 79–86, Feb. 2022.
- [8] A. Fauziah, D. Angreni, A. Pradjaningsih, A. Riski, and A. F. Hadi, "Support vector regression in statistical downscaling for rainfall forecasting," *International journal of scientific and technology research*, vol. 9, no. 03, pp. 5665–5669, Mar. 2020.
- [9] H. Byun and S.-W. Lee, "A survey on pattern recognition applications of support vector machines," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 17, no. 03, pp. 459–486, May 2003, <https://doi.org/10.1142/S0218001403002460>.
- [10] T. Kusumadewi *et al.*, "Synthesizing Environmental, Social, and Urban Density Metrics to Predict Urban Heat Island Dynamics using Remote Sensing and Support Vector Regression," *Engineering, Technology & Applied Science Research*, vol. 15, no. 3, pp. 23141–23148, June 2025, <https://doi.org/10.48084/etasr.9791>.
- [11] A. Wahyu, A. A. Kuntoro, and T. Yamashita, "Annual and Seasonal Discharge Responses to Forest/Land Cover Changes and Climate Variations in Kapuas River Basin, Indonesia," *Journal of International Cooperation Studies*, vol. 16, no. 2, pp. 81–100, Dec. 2010.
- [12] X. Ma, J. Xu, Y. Luo, S. Prasad Aggarwal, and J. Li, "Response of hydrological processes to land-cover and climate changes in Kejie watershed, south-west China," *Hydrological Processes*, vol. 23, no. 8, pp. 1179–1191, 2009, <https://doi.org/10.1002/hyp.7233>.
- [13] J.-H. Qian, "Multi-scale climate processes and rainfall variability in Sumatra and Malay Peninsula associated with ENSO in boreal fall and winter," *International Journal of Climatology*, vol. 40, no. 9, pp. 4171–4188, 2020, <https://doi.org/10.1002/joc.6450>.
- [14] H. Kuswanto, D. Setiawan, and A. Sopaheluwakan, "Clustering of Precipitation Pattern in Indonesia Using TRMM Satellite Data," *Engineering, Technology & Applied Science Research*, vol. 9, no. 4, pp. 4484–4489, Aug. 2019, <https://doi.org/10.48084/etasr.2950>.
- [15] R. Kamma, S. Badaruddin, M. R. Mustamin, and M. I. Mukrim, "Flood Risk Assessment and Mitigation Strategies for the Sinjai and Tangka River Catchments in Indonesia using Hydraulic Modeling and Spatial Analysis," *Engineering, Technology & Applied Science Research*, vol. 15, no. 2, pp. 20623–20634, Apr. 2025, <https://doi.org/10.48084/etasr.9837>.
- [16] J. Wu, H. Liu, G. Wei, T. Song, C. Zhang, and H. Zhou, "Flash Flood Forecasting Using Support Vector Regression Model in a Small Mountainous Catchment," *Water*, vol. 11, no. 7, July 2019, Art. no. 1327, <https://doi.org/10.3390/w11071327>.
- [17] W. Xu, Y. Hou, Y. S. Hung, and Y. Zou, "A comparative analysis of Spearman's rho and Kendall's tau in normal and contaminated normal models," *Signal Processing*, vol. 93, no. 1, pp. 261–276, Jan. 2013, <https://doi.org/10.1016/j.sigpro.2012.08.005>.
- [18] W. J. Conover, *Practical Nonparametric Statistics*, 3rd ed. New York, NY, USA: John Wiley & Sons, 1971.
- [19] I. Sadidan, G. L. Sari, F. I. Alifin, and A. R. Budiarto, "Utilization of machine learning to predict the correlation between color of river water and other water quality characters," *BIS Information Technology and Computer Science*, vol. 1, Oct. 2024, <https://doi.org/10.31603/bistycs.136>.
- [20] B. Schölkopf and A. J. Smola, *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. Cambridge, MA, USA: MIT Press, 2002.
- [21] S. Abe, *Support Vector Machines for Pattern Classification*. London, UK: Springer, 2005.
- [22] S. Gunn, "Support Vector Machines for Pattern Classification," ISIS Research Group, School of Electronics and Computer Science, University of Southampton, Southampton, UK, Technical Report, 1998.
- [23] R. E. Caraka, H. Yasin, and A. W. Basyiruddin, "Forecasting crude palm oil (CPO) using support vector regression with radial basis kernel," *Journal of Mathematics*, vol. 7, no. 1, pp. 43–57, June 2017.
- [24] W. W. S. Wei, *Time Series Analysis: Univariate and Multivariate Methods*, 2nd ed. New York, NY, USA: Pearson Addison Wesley, 2006.
- [25] M. Bergemann and C. Jakob, "How important is tropospheric humidity for coastal rainfall in the tropics?," *Geophysical Research Letters*, vol. 43, no. 11, pp. 5860–5868, May 2016, <https://doi.org/10.1002/2016GL069255>.
- [26] A. Sharma, A. Mitra, V. Vasan, and R. Govindarajan, "Spatio-temporal relationships between rainfall and convective clouds during Indian Monsoon through a discrete lens," *International Journal of Climatology*, vol. 41, no. 2, pp. 1351–1368, Aug. 2020, <https://doi.org/10.1002/joc.6812>.
- [27] O. Chapelle, P. Haffner, and V. N. Vapnik, "Support vector machines for histogram-based image classification," *IEEE Transactions on Neural Networks*, vol. 10, no. 5, pp. 1055–1064, Sept. 1999, <https://doi.org/10.1109/72.788646>.
- [28] K. E. Trenberth, A. Dai, R. M. Rasmussen, and D. B. Parsons, "The Changing Character of Precipitation," *Bulletin of American Meteorological Society*, vol. 84, no. 9, pp. 1205–1218, Sept. 2003, <https://doi.org/10.1175/BAMS-84-9-1205>.
- [29] S. Lestari, A. King, C. Vincent, D. Karoly, and A. Protat, "Seasonal dependence of rainfall extremes in and around Jakarta, Indonesia," *Weather and Climate Extremes*, vol. 24, June 2019, Art. no. 100202, <https://doi.org/10.1016/j.wace.2019.100202>.
- [30] K. Polisetty, K. K. Paidipati, and J. D. Bodapati, "Modelling of Monthly Rainfall Patterns in the North-West India Using SVM," *Ingénierie des Systèmes d'Information*, vol. 24, no. 4, Aug. 2019, Art. no. 391, <https://doi.org/10.18280/isi.240405>.