CoreID JOURNAL

Vol. 2 No. 3, November 2024 (86-92)
ISSN: 2987-6990
DOI: https://doi.org/10.60005/coreid.v2i3.87

Comparison of Rainfall Prediction Results in South Bangka Regency Using Support Vector Regression and SARIMA

Pitra Wati [1], Adriyansyah [2*], Ineu Sulistiana [3]
[1]Department of Mathematics, University of Bangka Belitung, Bangka, Indonesia
[2]Department of Civil Engineering, University of Bangka Belitung, Bangka, Indonesia
[3]Department of Mathematics, University of Bangka Belitung, Bangka, Indonesia

Article Info

Article history:

Received August 28, 2024
Revised Nov 12, 2024
Accepted Nov 25, 2024

Keywords:

Rainfall prediction
Support vector regression
SARIMA
Radial basis function
Machine learning

ABSTRACT

Climate variables such as rainfall play a crucial role in many sectors, including water resource management, disaster mitigation, agriculture, transportation, and others. Rainfall is influenced by various meteorological and geographical factors. These factors interact in complex ways, making rainfall amounts uncertain and difficult to predict. The uncertainty in rainfall intensity can significantly impact agricultural production, clean water availability, and the potential occurrence of other disasters. Therefore, accurate rainfall prediction is essential for effective planning and informed decision-making. South Bangka is one of the regencies with the highest rice production levels compared to other regencies in the Bangka Belitung Islands Province. This highlights the importance of accurate rainfall prediction to support the agricultural sector, which serves as a pillar of the economy in South Bangka. Based on this issue, the present study aims to identify a more accurate prediction model by comparing the performance of Support Vector Regression (SVR) and Seasonal Autoregressive Integrated Moving Average (SARIMA) models in forecasting rainfall in South Bangka. The findings of this study indicate that the SVR method using the Radial Basis Function (RBF) kernel yields more accurate predictions compared to SVR with linear, polynomial, and sigmoid kernels. Moreover, the SVR with RBF kernel also outperforms the SARIMA model in terms of prediction accuracy.

Corresponding Author:

Adriyansyah
Department of Civil Engineering, University of Bangka Belitung, Bangka, Indonesia
Email: adriyan_ubb@yahoo.com

1. INTRODUCTION

Climate variables such as rainfall play a crucial role in many sectors, including water resource management, disaster mitigation, agriculture, transportation, and others. Rainfall is influenced by a number of factors, both meteorological and geographical. Meteorological factors include air humidity, air pressure, air temperature, wind, and global climate phenomena. Meanwhile, geographical factors consist of elevation and topography, distance from the sea, monsoons, and ocean currents. These factors interact in a highly complex manner, resulting in uncertainty and making rainfall difficult to predict. The uncertainty in rainfall intensity can impact agricultural production, the availability of clean water, and the potential occurrence of disasters such as droughts, floods, and landslides. This highlights the importance of accurate rainfall prediction to support optimal planning and decision-making.

Over the past few decades, numerous predictive methods have been developed and applied in the field of climate science, including rainfall forecasting. Conventional statistical methods as well as machine learning-based approaches have been continuously explored to achieve higher prediction accuracy. One of the widely used conventional statistical methods is the Seasonal Autoregressive Integrated Moving Average (SARIMA), particularly for time series analysis, due to its capability to capture seasonal patterns and trends within the data. SARIMA is an extension of the ARIMA model, specifically designed to accommodate seasonality. The inclusion of seasonal components allows SARIMA to effectively identify recurring patterns in rainfall data, thereby producing more accurate forecasts, especially for datasets exhibiting seasonal trends.

The advancement of technology and machine learning has significantly accelerated the development of artificial intelligence based prediction methods. Predictive models such as Support Vector Regression (SVR) have received considerable attention due to their ability to handle non-linear data and capture more complex patterns. SVR is a machine learning method derived from Support Vector Machine (SVM) and is specifically designed for regression tasks [1], [2], [3]. It operates by mapping the input data into a higher-dimensional space using a kernel function and then identifying the optimal hyperplane that minimizes prediction errors [4], [5], [6], [7]. SVR offers advantages such as the ability to model non-linear relationships and strong resistance to overfitting. It is also flexible in handling noisy data, making it a suitable choice for rainfall prediction applications.

Several previous studies have employed SVR and SARIMA models to predict rainfall and other types of time series data. For instance [8] applied a hybrid Multi-Kernel SVR–SARIMA model to forecast rainfall in Manado, North Sulawesi. [9] forecasted long-term monthly average temperatures across various climate zones in Iran using SARIMA, SVR, and SVR-FA. [10] integrated Particle Swarm Optimization (PSO) and SVR into an enhanced ARIMA model for improved monthly rainfall forecasting. [11] conducted time series forecasting on industry sales of printing and writing paper as well as monthly shipments of pollution control equipment using a Seasonal SVR model. [12] implemented a hybrid SVR–Firefly algorithm to predict monthly rainfall. Additionally, [13] conducted a comparative study and analysis of SARIMA and Long Short-Term Memory (LSTM) models for time series forecasting.

South Bangka Regency, Bangka Belitung Islands Province, is a region characterized by a tropical type-A climate with low to moderate annual rainfall levels [14]. Areas with low rainfall tend to be more vulnerable to drought if there are no additional water sources such as rivers or reservoirs. This condition requires special attention since South Bangka is an agricultural and plantation area that relies heavily on rainfall for irrigation purposes. Accurate rainfall prediction is therefore essential to support the agricultural sector, which serves as a cornerstone of the economy in South Bangka. Farmers and policymakers can utilize the results of rainfall prediction to make more informed decisions during dry seasons or extreme rainy seasons.

Based on the above background, this study aims to identify a more accurate prediction model by comparing the performance of SVR and SARIMA methods in forecasting rainfall in South Bangka Regency. The rainfall data used in this study cover the period from January 2010 to December 2023. These data are employed to predict rainfall from January 2024 to December 2027. The findings of this study are expected to contribute to the development of more effective and applicable rainfall prediction methods, particularly in the context of water resource management and disaster risk mitigation. Additionally, the results are expected to provide recommendations on the most suitable prediction methods for regions with highly variable rainfall patterns such as South Bangka.

2. METHOD

2.1 Research Data

The data used in this study consist of monthly rainfall records in South Bangka Regency from January 2010 to December 2023. These are secondary data obtained from BPS – Statistics of South Bangka Regency. The data are presented in Figure 1.

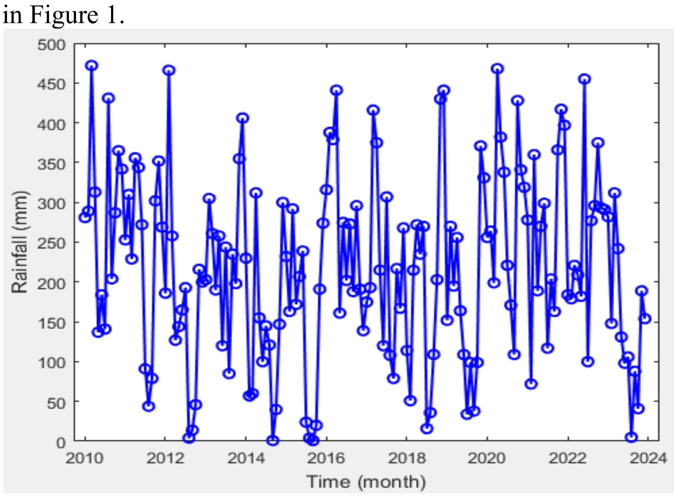

Figure 1. Monthly rainfall data in South Bangka Regency from January 2010 to December 2023.

2.2 Support Vector Regression (SVR)

Support Vector Regression (SVR) is an extension of the Support Vector Machine (SVM) designed specifically for regression problems. SVR works by finding a function f(x) that minimizes prediction errors while maintaining low model complexity. This approach is based on the principle of maximum margin. SVR is particularly effective for non-linear data due to the use of kernel functions. A kernel function projects data from a low-dimensional input space into a higher-dimensional feature space through a non-linear mapping, making the data more amenable to linear modeling. The selection of an appropriate kernel function and its parameters significantly affects the accuracy of the SVR model [15]. Improper selection may lead to underfitting or overfitting.

2.2.1 Dataset

The dataset used in the SVR model is divided into two parts: training data and testing data. The training data typically comprise 70–80% of the total actual data, while the testing data account for the remaining 20–30%. The training data are used to train the SVR model, whereas the testing data are used to evaluate its performance. The mathematical formulation of the dataset can be expressed as follows:

$$\left\{\left(x_i, y_i\right) \middle| i = 1, 2, \ldots, n\right\} \tag{1}$$

where x_i is the input and y_i is the target output.

2.2.2 Model SVR

The objective of SVR is to find the best function that can predict the target output values from a given set of input data with minimal error. The function used in the SVR model is defined as follows:

$$f\left(x\right) = w^T \varphi\left(x\right) + b \tag{2}$$

where w is the weight vektor, $\varphi\left(x\right)$ is the function that maps x into a higher-dimensional space, and b is the bias. The difference between $f\left(x\right)$ dan y must be less than or equal to the predefined tolerance ε, which is expressed as:

$$\left|y_i - \left(w^T \varphi\left(x_i\right) + b\right)\right| \le \varepsilon \tag{3}$$

2.2.3 Kernel Function

In the SVR method, kernel functions are used to transform non-linear data in the input space into a high-dimensional feature space in order to find the optimal hyperplane. The performance of the SVR model depends on the use of several hyperparameters, including C, ε, degree, kernel parameters, and the type of kernel function employed. Several kernel functions that can be used in SVR include the linear kernel, polynomial kernel, sigmoid kernel, and radial basis function (RBF) kernel. The functions for these kernels are shown in Table 1.

Table 1. Kernel Function

Kernel	**Function**
Linear	$K\left(x_i, x_j\right) = x_i^T x_j$
Polynomial	$K\left(x_i, x_j\right) = \left(x_i^T x_j + c\right)^d$
Sigmoid	$K\left(x_i, x_j\right) = \tanh\left(\alpha x_i^T x_j + c\right)$
Radial Basis Function	$K\left(x_i, x_j\right) = \exp\left(-\gamma \left\Vert x_i - x_j \right\Vert^2\right)$

2.2.4 SVR Model

If any data points exceed the predetermined ε margin, a penalty is imposed by introducing slack variables ξ_i dan ξ_i^*. The optimization problem in the SVR method is formulated as follows:

$$\min_{w,b,\xi,\xi^*} \frac{1}{2}\left\Vert w \right\Vert^2 + C \sum_{i=1}^{n} \left(\xi_i + \xi_i^*\right) \tag{4}$$

constraints:

$$y_i - w^T \varphi\left(x_i\right) - b \le \varepsilon + \xi_i \; ; \; w^T \varphi\left(x_i\right) + b - y_i \le \varepsilon + \xi_i^* \; ; \; \text{and } \xi_i, \xi_i^* \ge 0 \qquad (5)$$

where C is the hyperparameter, ε is the margin of tolerance, ξ_i dan ξ_i^* are the slack variables.

2.2.5 Data Prediction

After the SVR model has been trained using the training data, the next step is to perform prediction on new data, referred to as the testing data. Testing data prediction is an evaluation process used to determine whether the trained SVR model can accurately predict unseen data. If the model performs well on training data but poorly on testing data, it indicates overfitting. Conversely, if the model is too simplistic and fails to capture the relationship patterns between input and output, it indicates underfitting. The prediction results on testing data can provide insight into the model's generalization ability, which is essential for forecasting future data.

2.2.6 Model Evaluation

The performance of the prediction model is evaluated using model evaluation metrics. Evaluation metrics are quantitative tools used to measure the level of accuracy between the model's prediction results and the actual data. The main objective of this evaluation is to assess how well the model represents patterns from historical data and its ability to predict new data. The evaluation metrics used in this study are Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE). RMSE is commonly used because it provides a direct indication of the magnitude of prediction error in the same scale as the data. However, RMSE is sensitive to extreme values or outliers. In contrast, MAE provides a linear assessment of error without excessively penalizing outliers. The smaller the values of RMSE and MAE, the more accurate the prediction results of the model used. The formulas for RMSE and MAE are presented in the following equations:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(y_i - \hat{y}_i\right)^2} \quad \text{dan} \quad \text{MAE} = \frac{1}{n} \sum_{i=1}^{n} \left|y_i - \hat{y}_i\right| \qquad (6)$$

where n is the number of data, y_i is the actual data, and $\hat{y}_i$ is the predicted results.

2.3 Seasonal Autoregressive Integrated Moving Average (SARIMA)

The Seasonal Autoregressive Integrated Moving Average (SARIMA) is an extension of the Autoregressive Integrated Moving Average (ARIMA) model designed for time series data that exhibit seasonal patterns. The SARIMA model is denoted as ARIMA (p,d,q) $(P, D, Q)^s$. The model consists of two components, represented using lowercase and uppercase letters. The notation *(p, d, q)* refers to the non-seasonal part of the model, while (*P, D, Q*) represents the seasonal component. The *s* indicates the number of periods per season. The general form of the SARIMA model is expressed as follows:

$$\phi_p\left(B\right)\Phi_P\left(B^s\right)\left(1-B\right)^d\left(1-B^s\right)^D Y_t = \theta_q\left(B\right)\Theta_Q\left(B^s\right)a_t \qquad (7)$$

where $\phi_p(B)$ is polynomial of order p representing the non-seasonal autoregressive component, $\theta_q(B)$ is a polynomial of order q representing the non-seasonal moving average component, and $\Phi_P(B^S)$ is a polynomial of order P representing the seasonal autoregressive component.

The SARIMA model assumes that the data used is stationary, meaning it has a constant mean and variance over time. Stationarity is tested using the Augmented Dickey-Fuller (ADF) method. If the data is not stationary, a transformation process is performed through differencing on both non-seasonal and seasonal components until stationarity is achieved. The initial values of the SARIMA parameters (*p*, *d*, *q*, *P*, *D*, *Q*) are determined by analyzing patterns in the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots. The ACF plot is used to identify the order of the Moving Average (MA), while the PACF plot is used to determine the order of the Autoregressive (AR). The best model is selected based on evaluation criteria such as the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Mean Absolute Percentage Error (MAPE), Root Mean Squared Error (RMSE), or other evaluation metrics.

3. RESULTS AND DISCUSSION

3.1 SVR Prediction Results

The SVR method begins by dividing the dataset into training and testing sets. In this study, the data was split with 70% used for training and 30% for testing. With a total of 168 data points, 118 were used for training and 50 for testing. Four kernel functions were utilized, namely the linear kernel, polynomial kernel, sigmoid kernel, and radial basis function (RBF) kernel. The kernel parameters used in the prediction are shown in Table 2.

Table 2. Kenel parameters used for prediction

Kernel parameters			
Linear	Polynomial	Sigmoid	RBF
ε = 0.0001	ε = 0.01	ε = 0.001	ε = 0.0001
C = 100	*C* = 10	*C* = 1	*C* = 1000
-	γ = 12	γ = 30	γ = 235
-	*Coef 0* = 6	*Coef 0* = 10	-
-	Degree = 1	-	-

Based on the analysis results of the kernel functions and parameters used, it can be concluded that the best SVR model, based on the smallest RMSE and MAE values, is the SVR with the Radial Basis Function (RBF) kernel. The following presents the prediction accuracy levels based on RMSE and MAE values for all kernel functions used.

Table 3. Accuracy level of kernel types based on RMSE and MAE

Kernel type	RMSE			MAE		
	Training	Testing	Average	Training	Testing	Average
Linear	118.35	103.5	110.93	94.25	82.31	88.28
Polynomial	115.83	117.46	116.65	93.19	94.67	93.93
Sigmoid	134.33	181.98	158.16	100.97	151.80	126.38
RBF	0.93	0.83	0.88	0.13	0.4	0.27

The prediction results of rainfall data in South Bangka Regency using SVR show that the Radial Basis Function (RBF) kernel provides excellent predictive performance. This is evidenced by the average RMSE and MAE values of 0.88 and 0.27, respectively. These values indicate a very low prediction error and a very high model accuracy. Other kernels produced relatively higher average RMSE and MAE values, indicating that those approaches are less suitable for capturing the data patterns. The average RMSE and MAE values in Table 6 clearly demonstrate that the RBF kernel significantly outperforms the others in modeling non-linear data, which tend to have fluctuating and complex characteristics.

3.2 SARIMA Prediction Results

After performing predictions using SVR with four different kernels, the prediction process was continued using the SARIMA method. The results obtained from SARIMA are used as a comparison to the SVR model. The initial step in the SARIMA method is to visualize the time series data through graphical analysis. If the mean value of the data fluctuates or is not constant over time, the data is considered non-stationary. The graph shown in Figure 3 indicates that the data exhibits a non-constant mean, leading to the conclusion that the data is non-stationary. Non-stationary data is transformed into stationary data using differencing. Once differencing is applied, the stationarity of the data is re-evaluated using the Augmented Dickey-Fuller (ADF) test. The results of the ADF test are presented in Table 4.

Table 4. ADF test results

Variable	Dickey-Fuller	P-Value
Rainfall	-4.4932	0.01

The ADF test results presented in Table 4 show that the p-value is less than 0.05. This indicates that the data became stationary after first-order differencing. Once stationarity is achieved, autocorrelation identification is carried out using the ACF and PACF plots. The purpose of this step is to determine the strength of the relationship between the current data and previous observations. The ACF and PACF plots are presented in Figure 2.

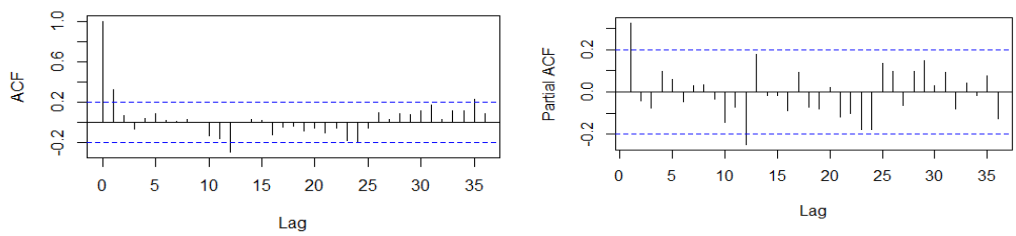

Figure 2. ACF and PACF

Figure 2 shows that lag 1 in the ACF provides a high autocorrelation value, close to 1. This indicates that the data is very similar to its previous values, suggesting that the time series data can be used to predict future values. The ACF and PACF plots also show a seasonal pattern at lag 12, indicating the presence of seasonality in the data. Therefore, the appropriate model is SARIMA $(p,d,q)\ (P,D,Q)^{12}$.

The prediction results using SARIMA are evaluated based on the smallest values of the Akaike Information Criterion (AIC), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE). These criteria are used to assess model suitability and to prevent overfitting. A smaller AIC value indicates a better-fitting model with a simpler structure compared to models with higher AIC values. Meanwhile, smaller RMSE and MAE values indicate higher prediction accuracy. The SARIMA modeling results based on these three criteria are presented in Table 5.

Table 5. SARIMA models based on AIC values

SARIMA	SARIMA (p, d, q) $(P, D, Q)^s$	AIC	RMSE	MAE
Model 1	$(1,1,1)$ $(2,1,1)^{12}$	2.56	165.34	133.90
Model 2	$(1,1,1)$ $(2,1,2)^{12}$	2.62	195.69	160.27
Model 3	$(0,1,1)$ $(2,1,2)^{12}$	2.68	232.05	201.41
Model 4	$(1,1,12)$ $(0,1,0)^{12}$	2.67	164.80	132.66
Model 5	$(12,1,24)$ $(0,1,0)^{12}$	2.76	188.55	155.15
Model 6	$(24,1,12)$ $(0,1,0)^{12}$	2.84	132.96	103.62
Model 7	$(0,0,0)$ $(1,1,2)^{12}$	2.65	104.92	84.41
Model 8	$(0,0,0)$ $(2,1,1)^{12}$	2.69	113.39	88.84

The modeling results in Table 5 show that the lowest AIC value is found in Model 1. Although the RMSE and MAE values for Model 1 are not the lowest, they remain fairly competitive compared to the other models. The smallest RMSE and MAE values are achieved by Model 7, which also presents a reasonably competitive AIC value. Models 5 and 6 have very high orders of p and q, which may lead to overfitting and result in larger AIC values. Models 7 and 8 indicate that annual seasonality is highly dominant, while the trend or non-seasonal components are relatively insignificant.

If we consider the model with the smallest AIC value, Model 1 exhibits the most statistically optimal performance. However, in the context of this study, Model 7 was selected as the best model among all alternatives. This is because Model 7 produces a prediction pattern that closely resembles the forecast results generated by the SVR method using the RBF kernel. Additionally, Model 7 exhibits the smallest RMSE and MAE values, indicating that it provides predictions that are closest to the actual values. Therefore, the similarity in pattern with the SVR-RBF predictions and the lowest RMSE and MAE values make Model 7 the most representative SARIMA model for this case.

Based on the prediction results of the SVR model using the RBF kernel and the SARIMA model, it can be concluded that the SVR model with the RBF kernel outperforms SARIMA in terms of accuracy. This is evident from the RMSE and MAE values for each model. The SVR model with the RBF kernel yields RMSE and MAE values of 0.88 and 0.27, respectively, while the SARIMA $(0,0,0)$ $(1,1,2)^{12}$ produces values of 104.92 and 84.41. After obtaining the best prediction models from SVR and SARIMA, a comparison is carried out between the plots of the SVR model with RBF kernel and the SARIMA $(0,0,0)$ $(1,1,2)^{12}$. The following figure presents the comparison of SVR using RBF Kernel and SARIMA $(0,0,0)$ $(1,1,2)^{12}$ in predicting rainfall in South Bangka Regency from January 2024 to December 2027.

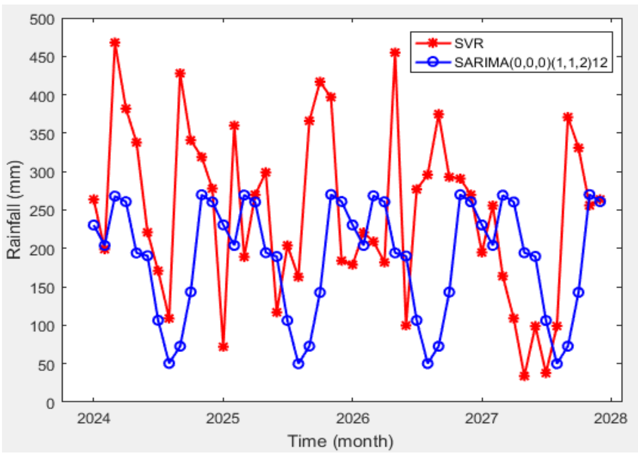

Figure 3. Comparison of prediction results between SVR and SARIMA $(0,0,0)(1,1,2)^{12}$

4. CONCLUSION

The analysis results indicate that the SVR method with the RBF kernel provides highly accurate predictions. This is evidenced by the average RMSE and MAE values of 0.88 and 0.27, respectively. The SVR method with the RBF kernel also produces prediction plots that closely align with the actual data. These findings lead to the conclusion that the appropriate choice of kernel type can significantly influence the accuracy of predictions. The RBF kernel is capable of capturing the complexity and non-linear patterns, including seasonal components and extreme fluctuations, in the rainfall data of South Bangka Regency. Therefore, the RBF kernel is recommended as the primary approach for rainfall data modeling and forecasting using SVR in similar case studies. In comparison, the SARIMA method yields average RMSE and MAE values of 104.92 and 84.41, respectively. This indicates that SARIMA also provides good prediction results. However, in terms of prediction accuracy, the SVR method with the RBF kernel demonstrates better performance than SARIMA.

REFERENCES

[1] Z. Guo and G. Bai, "Application of Least Squares Support Vector Machine for Regression to Reliability Analysis," *Chinese Journal of Aeronautics*, vol. 22, no. 2, pp. 160–166, Apr. 2009, doi: 10.1016/S1000-9361(08)60082-5.

[2] F. Zhang and L. J. O'Donnell, "Support vector regression," in *Machine Learning: Methods and Applications to Brain Disorders*, Elsevier, 2019, pp. 123–140. doi: 10.1016/B978-0-12-815739-8.00007-9.

[3] M. Sabzekar and S. M. H. Hasheminejad, "Robust regression using support vector regressions," *Chaos Solitons Fractals*, vol. 144, Mar. 2021, doi: 10.1016/j.chaos.2021.110738.

[4] M. Awad and R. Khanna, "Support vector regression," in *Efficient learning machines: Theories, concepts, and applications for engineers and system designers*, Springer, 2015, pp. 67–80.

[5] Y. Ding, L. Cheng, W. Pedrycz, and K. Hao, "Global nonlinear kernel prediction for large data set with a particle swarm-optimized interval support vector regression," *IEEE Trans Neural Netw Learn Syst*, vol. 26, no. 10, pp. 2521–2534, 2015.

[6] Z. Zhong and T. R. Carr, "Application of mixed kernels function (MKF) based support vector regression model (SVR) for CO2 – Reservoir oil minimum miscibility pressure prediction," *Fuel*, vol. 184, pp. 590–603, Nov. 2016, doi: 10.1016/j.fuel.2016.07.030.

[7] A. A. Akinpelu *et al.*, "A support vector regression model for the prediction of total polyaromatic hydrocarbons in soil: an artificial intelligent system for mapping environmental pollution," *Neural Comput Appl*, vol. 32, no. 18, pp. 14899–14908, Sep. 2020, doi: 10.1007/s00521-020-04845-3.

[8] R. E. Caraka, S. A. Bakar, and M. Tahmid, "Rainfall forecasting multi kernel support vector regression seasonal autoregressive integrated moving average (MKSVR-SARIMA)," in *AIP Conference Proceedings*, American Institute of Physics Inc., Jun. 2019. doi: 10.1063/1.5111221.

[9] P. Aghelpour, B. Mohammadi, and S. M. Biazar, "Long-term monthly average temperature forecasting in some climate types of Iran, using the models SARIMA, SVR, and SVR-FA," *Theor Appl Climatol*, vol. 138, no. 3–4, pp. 1471–1480, Nov. 2019, doi: 10.1007/s00704-019-02905-w.

[10] L. Parviz and M. Ghorbanpour, "Assimilation of PSO and SVR into an improved ARIMA model for monthly precipitation forecasting," *Sci Rep*, vol. 14, no. 1, Dec. 2024, doi: 10.1038/s41598-024-63046-3.

[11] P. F. Pai, K. P. Lin, C. S. Lin, and P. T. Chang, "Time series forecasting by a seasonal support vector regression model," *Expert Syst Appl*, vol. 37, no. 6, pp. 4261–4265, Jun. 2010, doi: 10.1016/j.eswa.2009.11.076.

[12] A. Danandeh Mehr, V. Nourani, V. Karimi Khosrowshahi, and M. A. Ghorbani, "A hybrid support vector regression–firefly model for monthly rainfall forecasting," *International Journal of Environmental Science and Technology*, vol. 16, no. 1, pp. 335–346, Jan. 2019, doi: 10.1007/s13762-018-1674-2.

[13] A. Kumar Dubey, A. Kumar, V. García-Díaz, A. Kumar Sharma, and K. Kanhaiya, "Study and analysis of SARIMA and LSTM in forecasting time series data," *Sustainable Energy Technologies and Assessments*, vol. 47, Oct. 2021, doi: 10.1016/j.seta.2021.101474.

[14] BPS-Statistics of Bangka Selatan Regency. "Bangka Selatan Regency in Figure 2024. Volume 22, 2024.

[15] H. Yang, Z. Wang, and K. Song, "A new hybrid grey wolf optimizer-feature weighted-multiple kernel-support vector regression technique to predict TBM performance," *Eng Comput*, pp. 1–17, 2022.