UKJAES

University of Kirkuk Journal For Administrative and Economic Science

ISSN:2222-2995 E-ISSN:3079-3521

University of Kirkuk Journal For Administrative and Economic Science

Ahmed Botan, Saeed Soran & Nihad Hozan. Evaluating Kernel-Specific Hyperparameter Tuning in Support Vector Regression for Time Series Forecasting. *University of Kirkuk Journal For Administrative and Economic Science* (2025) 15 (4) Part (2):363-375.

Evaluating Kernel-Specific Hyperparameter Tuning in Support Vector Regression for Time Series Forecasting

Botan Ahmed [1], Soran Saeed [2], Hozan Nihad [3]

[1,2,3] *Sulaimani Polytechnic University, College of Informatics, Sulaimani, Iraq*

Abstract: This study evaluates kernel-specific hyperparameter tuning in Support Vector Regression (SVR) for time series forecasting using a simulated dataset. Four SVR kernels Linear, Sigmoid, Radial Basis Function (RBF), and Polynomial were systematically tested under individually optimized configurations. For each kernel, key hyperparameters including C, epsilon, gamma (used in RBF, Polynomial, and Sigmoid), degree (used only in Polynomial), and coef0 were tuned to enhance predictive accuracy. Performance was assessed using RMSE, MSE, R^2, AIC, and BIC metrics. Results indicate that the Linear kernel yielded the best overall performance, demonstrating superior generalization and stability. The findings underscore that careful selection and tuning of kernel-specific hyperparameters significantly improve SVR model effectiveness in capturing temporal patterns and reducing forecast error.

Keywords: Support Vector Regression, Time Series Forecasting, Kernel Functions, Hyperparameter Tuning, Machine Learning.

تقييم ضبط المعاملات الفائقة الخاصة بالنواة في الانحدار باستخدام المتجهات الداعمة (SVR) لتنبؤ السلاسل الزمنية

م.م. بوتان أحمد[١]، أ.د. سوران سعيد[٢]، أ.م.د. هوزان نهاد[٣]

[١،٢،٣] *جامعة السليمانية التقنية، كلية المعلوماتية، السليمانية، العراق*

المستخلص: تُقيّم هذه الدراسة ضبط المعاملات الفائقة الخاصة بكل نوع من أنواع النواة في نماذج الانحدار باستخدام المتجهات الداعمة (SVR) لأغراض التنبؤ بالسلاسل الزمنية، وذلك باستخدام مجموعة بيانات محاكاة. تم اختبار أربع نوى من SVR بشكل منهجي: النواة الخطية، ونواة السيني، ونواة الدالة الأسية الشعاعية (RBF) ، والنواة متعددة الحدود، وذلك ضمن إعدادات محسّنة لكل منها على حدة. بالنسبة لكل نواة، تم ضبط المعاملات الأساسية مثل C ، epsilon، gamma (تُستخدم في RBF ومتعددة الحدود وSigmoid)، degree (تُستخدم فقط في متعددة الحدود)، و coef0 بهدف تحسين الدقة التنبؤية.

تم تقييم الأداء باستخدام مؤشرات RMSE، وMSE، وR^2، وAIC، و BIC. وتشير النتائج إلى أن النواة الخطية قدّمت أفضل أداء إجمالي، حيث أظهرت قدرة عالية على التعميم واستقراراً ملحوظاً. وتؤكد هذه النتائج أن الاختيار الدقيق وضبط المعاملات الفائقة الخاصة بالنواة يُسهم بشكل كبير في تحسين فعالية نموذج SVR في التقاط الأنماط الزمنية وتقليل خطأ التنبؤ.

IRAQI Academic Scientific Journals
University of Kirkuk Journal For Administrative and Economic Science (2025) 15 (4) Part (2): 363-375

الكلمات المفتاحية: الانحدار بدعم المتجهات، التنبؤ بالسلاسل الزمنية، دوال النواة، ضبط المعاملات الفائقة، تعلم الآلة.

Corresponding Author: E-mail:

INTRODUCTION

Support Vector Regression (SVR) is a powerful supervised learning technique derived from Support Vector Machines (SVM), designed to handle regression tasks by mapping input data into high-dimensional feature spaces using kernel functions [1]. Its strength lies in its ability to generalize well even with limited data, making it particularly suitable for time series forecasting where patterns may be nonlinear, seasonal, and noisy [2].

In time series modeling, kernel selection plays a pivotal role in SVR performance. Each kernel linear, radial basis function (RBF), polynomial, and sigmoid transforms the input space differently, influencing the model's ability to capture underlying trends and fluctuations. Linear kernels are often favored for their simplicity and interpretability, especially when the data exhibits smooth and structured behavior. In contrast, nonlinear kernels like RBF and polynomial are capable of modeling complex relationships but require careful tuning to avoid overfitting or instability [3].

Recent studies emphasize that while nonlinear kernels offer flexibility, their performance is highly sensitive to hyperparameter configurations such as gamma, degree, and coef0, which can significantly affect generalization and computational efficiency [1]. Moreover, sigmoid kernels, though theoretically versatile, have shown erratic behavior in practical applications, often leading to unreliable predictions unless tightly constrained [3].

This paper presents a comparative analysis of Support Vector Regression (SVR) models across four kernel types Linear, RBF, Polynomial, and Sigmoid applied to a simulated seasonal time series dataset. The objective is to evaluate each kernel's predictive accuracy, stability, and model complexity using a suite of performance metrics, including Root Mean Squared Error (RMSE), Mean Squared Error (MSE), Coefficient of Determination (R^2), Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC). By systematically tuning hyperparameters and analyzing trade-offs across these metrics, the study aims to provide actionable insights into kernel selection for structured forecasting tasks.

The findings contribute to a broader understanding of SVR kernel behavior and offer practical guidance for researchers and practitioners seeking to balance model interpretability, predictive precision, and computational efficiency in real-world time series applications.

1st: MATERIALS AND METHODS:

1- SUPPORT VECTOR REGRESSION (SVR)

The main goal of *SVR* is to find a function that approximates the underlying relationship between input variables and continuous output values. This function should ideally fit the data within a specified margin of error, denoted by a threshold called ϵ. The regression function is generally represented as [4]:

$$f(x) = [w, \phi(x)] + b \qquad \dots (1)$$

where:

w is the weight vector that determines the importance of each feature.

ϕ(x) is a kernel function that maps the input features into a higher-dimensional space to facilitate the modeling of complex relationships.

b is the bias term that adjusts the function output.

A. THE REGULARIZATION PARAMETER C IN SUPPORT VECTOR REGRESSION (SVR)

The parameter C in SVR controls the balance between model complexity and the tolerance for errors beyond the ϵ-insensitive margin. It determines how strongly deviations outside the ϵ threshold are penalized. A larger value of C enforces stricter penalties on errors, leading the regression function to fit the training data more closely but with increased risk of overfitting. In contrast, a smaller C allows more flexibility within the margin, promoting smoother functions that may generalize better but at the cost of higher training error [5].

B. THE ϵ-INSENSITIVE LOSS IN SUPPORT VECTOR REGRESSION (SVR)

The parameter epsilonϵ defines the width of the ϵinsensitive tube, within which errors are not penalized. A small ϵ-makes the model sensitive to minor variations, improving accuracy but increasing the risk of overfitting. A larger ϵ ignores small deviations, leading to simpler models with better generalization but reduced precision [6].

C. THE KERNEL PARAMETER γ IN SUPPORT VECTOR REGRESSION (SVR)

The parameter γ controls the influence of individual training samples in kernel-based SVR, particularly when using the radial basis function (RBF) kernel. A small γ value produces a smoother decision function with broader influence, which may underfit the data. In contrast, a large γ value makes the model focus narrowly on specific training points, capturing complex patterns but with a higher risk of overfitting [7].

D. THE DEGREE PARAMETER IN POLYNOMIAL KERNEL SVR

The degree parameter is relevant when SVR uses a polynomial kernel. It specifies the degree of the polynomial function applied to the input features. A higher degree allows the model to capture more complex nonlinear relationships, but it may also lead to overfitting and increased computational cost. In contrast, lower degree values produce simpler decision functions that may underfit the data if the underlying relationship is highly nonlinear [8].

E. THE COEFFICIENT PARAMETER $COEF0$ IN SVR

The parameter $coef0$ is used in polynomial and sigmoid kernels to control the influence of higher-order versus lower-order terms in the kernel function. A larger $coef0$ increases the impact of higher-order terms, allowing the model to capture more complex patterns, while a smaller value emphasizes linear behavior. Proper tuning of $coef0$ is therefore important to balance model complexity and generalization performance [9].

2- LOSS FUNCTION

SVR employs the ϵ-insensitive loss function, defined as:

$$L_\epsilon\,(y_i, f(x_i)) = \begin{cases} 0 & if\ |y_i - f(x_i)| \leq \epsilon \\ |y_i - f(x_i)| - \epsilon & otherwise \end{cases} \quad \dots (2)$$

This loss function allows for a margin of tolerance, meaning that small deviations within the ϵ margin are not penalized, which helps in focusing on more significant errors [10].

3- SVR MODEL TUNING

In this study, four SVR kernels linear, RBF, polynomial, and sigmoid were evaluated. A grid search was applied to optimize the hyperparameters within the following ranges: $C \in \{0.1,1,10\}$, $\epsilon \in \{0.01,0.05,0.1\}, \gamma \in \{scale, 0.001,0.01,0.1\}$, degree $\in \{2,3,4\}$ for the polynomial kernel, and $coef0 \in \{0.0,0.1,0.5,1.0\}$ for the polynomial and sigmoid kernels. Model selection was based on the lowest Root Mean Squared Error (RMSE), with the linear kernel emerging as the best-performing SVR variant.

4- TRAINING THE MODEL

The experimental setup employed a consistent framework for all models to ensure fair comparison. Input sequences for both SVR and Transformer models were constructed using a window size of 10. The dataset was split into training and testing subsets in an 80/20 ratio. Model performance was evaluated using the Root Mean Square Error (RMSE), which served as the primary metric for assessing predictive accuracy and generalization.

Hyperparameter optimization was conducted via grid search for both SVR and Transformer models. For the hybrid ensemble, a stacking approach was adopted, with a linear regression model serving as the meta-learner to combine base model predictions. All models were trained and evaluated on the same dataset splits, with RMSE computed on the test set to ensure a consistent and unbiased performance comparison.

5- MAKING PREDICTIONS

After training, the model can predict new outputs using the learned parameters:

$$f(x) = \sum \quad i = ln\, ln\, (\alpha_i - \beta_i)\, K\left(x_i, x_j\right) + b \qquad \ldots (3)$$

This equation allows for the generation of predictions based on the input features and the learned support vectors.

6- EVALUATE PRECISION OF FORECASTING MODELS

To test the accuracy and the performance of the proposed model used, some statistical tests and measurements, including, mean square error, root of mean square error [11].

A. MEAN SQUARE ERROR (MSE)

Mean Squared Error (MSE) is a widely used metric for assessing the accuracy of a predictive model. It measures the average of the squares of the errors that is, the average squared difference between the actual values and the values predicted by the model [12]:

$$MSE = \frac{1}{n}\sum_{t=1}^{n} \quad (y_t - \hat{y}_t)^2 \qquad \ldots (4)$$

where:

n is the number of observations.

y_t is the actual value.

$\hat{y}_t$ is the predicted value.

A lower MSE indicates a better fit of the model to the data, as it suggests that the predictions are closer to the actual values. however, MSE can be sensitive to outliers, because it squares the errors, which can disproportionately affect the overall score [12].

B. SQUARE ROOT OF MEAN SQUARE ERROR (RMSE)

The Root Mean Square Error ($RMSE$) is a commonly used metric to evaluate the accuracy of a predictive model, providing a measure of the model's prediction error. It represents the square root of the average squared differences between predicted and actual values, making it easier to interpret than the Mean Squared Error (MSE) because it is expressed in the same units as the original data [13]:

$$MSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n} \quad (y_t - \hat{y}_t)^2} \qquad \dots (5)$$

7- DATA DESCRIPTION

In this study, a simulated time series dataset was constructed to emulate the structural patterns commonly observed in real-world forecasting scenarios. The sample consists of 500 time points, each representing a sequential observation influenced by three key components: a deterministic trend, a seasonal cycle, and stochastic noise. These elements were deliberately chosen to reflect the complexity and variability often encountered in domains such as economics, environmental monitoring, and epidemiological modeling.

$$data(t) = 0.05t + 10 \cdot sin\left(\frac{2\pi t}{50}\right) + noise \qquad \dots (6)$$

Trend: linear growth (0.05 * t)
Seasonality: sinusoidal (nonlinear, periodic)
Noise: random Gaussian

The response variable in this dataset is the synthesized signal, which captures the combined effect of long-term progression, periodic fluctuations, and random disturbances. This signal is not derived from empirical measurements but is generated through mathematical modeling to ensure full control over its underlying structure.

- **Trend**: The trend component represents a gradual linear increase over time, simulating phenomena such as population growth, technological adoption, or cumulative disease spread. It is modeled as 0.05t, where t is the time index. This steady upward trajectory provides a baseline for long-term change.
- **Seasonality**: Superimposed on the trend is a sinusoidal seasonal pattern with a fixed period of 50 time steps. This component mimics cyclical behaviors such as seasonal demand, climate cycles, or biological rhythms. The amplitude of 10 units ensures that the seasonal effect is pronounced and easily distinguishable.
- **Noise**: To introduce realistic variability, Gaussian noise with zero mean and a standard deviation of 2 is added to the signal. This stochastic element reflects unpredictable fluctuations that may arise from measurement errors, external shocks, or latent variables not captured by the model.

Together, these components form a composite signal that challenges predictive models to disentangle and learn both the deterministic and random aspects of the data. The dataset is particularly suitable for benchmarking regression algorithms, time series models, and hybrid architectures that aim to capture temporal dependencies and structural patterns.

Although the dataset is synthetic, its design is grounded in principles of time series analysis and statistical modeling, making it a valuable tool for testing methodological robustness and generalization performance.

2nd: RESULTS AND DISCUSSIONS

To assess the predictive performance of Support Vector Regression (SVR) models, a series of experiments were conducted using four distinct kernel types: linear, radial basis function (RBF), polynomial, and sigmoid. The evaluation focused primarily on Root Mean Square Error (RMSE) as a measure of accuracy, with each model undergoing systematic hyperparameter tuning to identify its optimal configuration. The discussion below interprets the comparative results and highlights key insights into kernel behavior and model reliability. The tables that follow summarize the best-performing hyperparameter settings for each kernel, providing a clear view of how tuning influenced prediction outcomes.

1- LINEAR KERNEL

Figure 1 illustrates the performance of SVR models using the linear kernel across an extensive grid of hyperparameter combinations. The results show remarkable stability in RMSE values, particularly under low regularization (C = 0.1), where RMSE consistently remained at 2.8133 regardless of changes to gamma or coef0—parameters that are functionally irrelevant in linear kernel formulations. As the regularization strength increased to C = 10, the model achieved its best performance with C = 10, ε = 0.05, and γ = 'scale', yielding an RMSE of 2.7332, alongside improvements in MSE (7.4706), R^2 (0.8541), AIC (971.08), and BIC (1971.46). These results suggest that higher regularization combined with moderate error tolerance enables the SVR to better capture underlying temporal patterns while maintaining generalization. Coef0 variations had no measurable impact across trials, reaffirming their irrelevance in this context. Additionally, narrowing the ε range from 0.1 to 0.05 provided marginal but consistent gains across all metrics. Overall, the linear kernel demonstrated robust, interpretable, and computationally efficient behavior, making it a strong candidate for time series forecasting with smooth and structured signals.

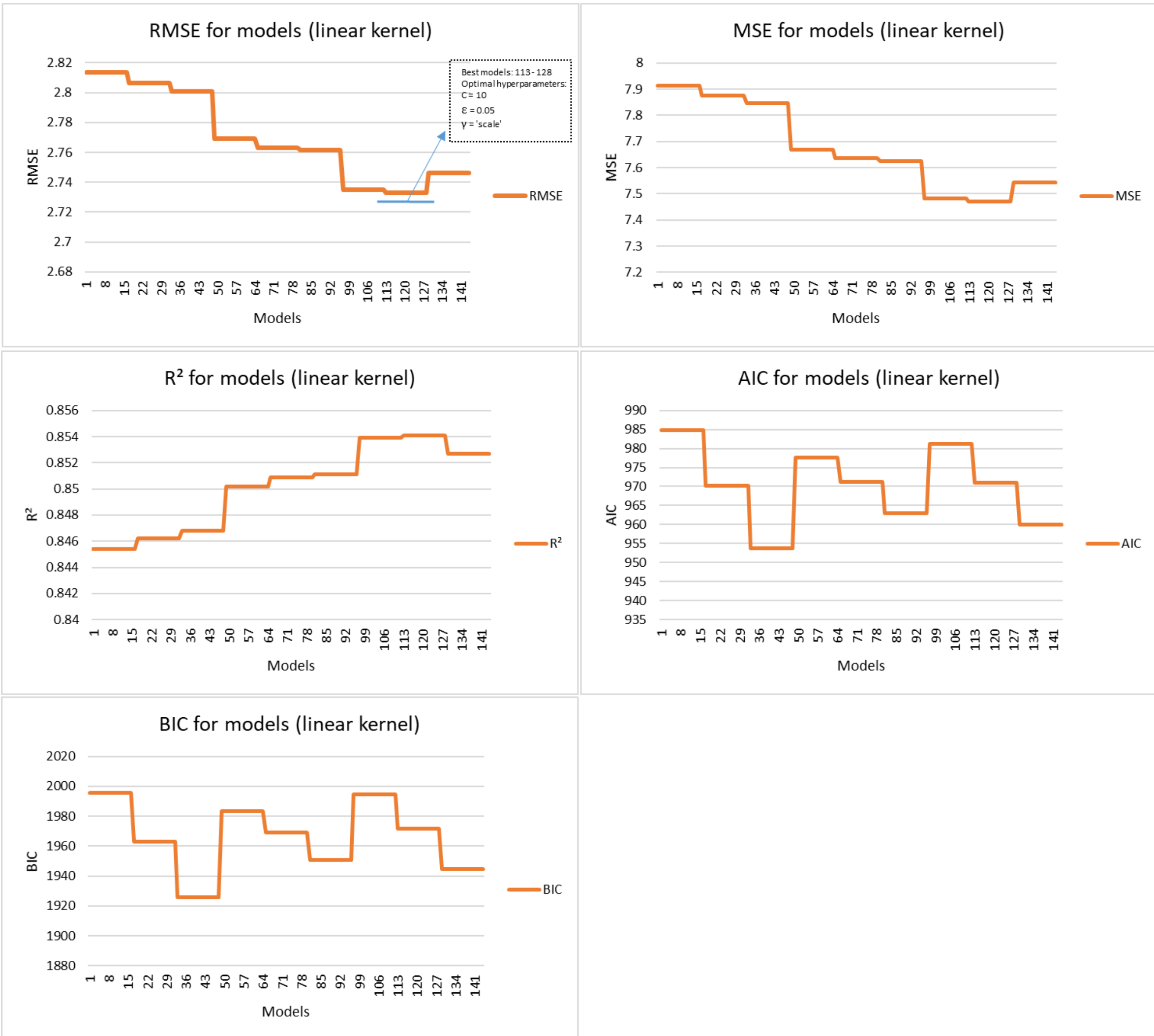

Figure (1): Metric results for models (linear Kernel)

2- RADIAL BASIS FUNCTION (RBF) KERNEL

Figure 2 presents the performance of SVR models using the RBF kernel across a broad hyperparameter grid. Unlike the linear kernel, the RBF kernel exhibited substantial variability in RMSE values, ranging from 13.65 in the least effective configuration to 2.9924 in the best. This wide spread underscores the RBF kernel's sensitivity to hyperparameter tuning—particularly the gamma parameter, which governs the influence of individual training samples. The optimal configuration, achieved with C = 10, ε = 0.1, and γ = 0.01, yielded an RMSE of 2.9924, MSE of 8.9547, R^2 of 0.8251, and favorable AIC (980.83) and BIC (1970.88) scores, indicating a well-balanced model with strong generalization. Lower gamma values (e.g., 0.001) consistently produced higher error metrics, suggesting that overly broad kernels failed to capture localized structure. Similarly, small regularization values (C = 0.1) led to underfitting, with RMSEs exceeding 11 and negative R^2 values. Overall, while the RBF kernel demonstrated the capacity to model nonlinear patterns, its performance was highly contingent on precise hyperparameter selection. Compared to the linear kernel (Figure 1), the RBF kernel required more careful tuning and still fell short in predictive accuracy, reinforcing the linear kernel's robustness for this dataset.

Figure (2): Metric results for models (RBF kernel)

3- POLYNOMIAL KERNEL

Figure 3 summarizes the performance of SVR models using the polynomial kernel across a diverse grid of hyperparameter combinations. This kernel introduces added complexity through the degree and coef0 parameters, which shape the curvature and offset of the decision boundary. Despite its flexibility, model performance varied considerably depending on the tuning configuration. The best-performing setup C = 10, ε = 0.1, γ = 0.01, degree = 2, and coef0 = 1.0 achieved an RMSE of 2.9924, MSE of 8.9547, R^2 of 0.8251, AIC of 980.83, and BIC of 1970.88, closely matching the optimal RBF configuration (Figure 2). This suggests that both kernels can effectively capture nonlinear relationships when precisely tuned. Across the grid, lower values of C (e.g., 0.1) consistently led to poor performance, with RMSEs exceeding 11.0 and negative R^2 values, indicating underfitting due to excessive regularization. Similarly, extreme gamma values (e.g., 0.001) produced RMSEs above 13.0 and inflated AIC/BIC scores, reflecting poor generalization. The degree parameter was also pivotal—quadratic configurations (degree = 2) consistently outperformed higher-degree polynomials, which likely introduced unnecessary complexity and overfitting. Compared to the linear kernel (Figure 1), the polynomial kernel demonstrated competitive performance only under optimal conditions. Its sensitivity to multiple interacting hyperparameters makes it less robust for general use. Nonetheless, it remains a viable option when the underlying data exhibits structured nonlinearities that simpler kernels may fail to capture.

Figure (3): Metric results for models (POLY kernel)

4- SIGMOID KERNEL

Figure 4 presents the performance of SVR models using the Sigmoid kernel, which introduces nonlinearity via a hyperbolic tangent transformation. The results reveal highly erratic behavior across hyperparameter combinations, with RMSE values ranging from 3.0109 to over 479.0, indicating extreme sensitivity and instability. The best-performing configuration C = 10, ε = 0.1, γ = 0.01, and coef0 = 0.0 achieved an RMSE of 3.0109, MSE of 9.0656, R^2 of 0.8229, AIC of 980.04, and BIC of 1967.50, placing it competitively alongside the optimal RBF and polynomial setups. However, this result was achieved under narrowly tuned conditions. Most other configurations particularly those involving large C values and gamma = 'scale' produced RMSEs exceeding 300, with corresponding MSEs and AIC/BIC scores indicating severe overfitting or numerical divergence. The kernel’s performance was especially volatile when gamma was set to 'scale' or 0.1, often resulting in RMSEs above 200 or even 400, and R^2 values plunging deep into the negative range. This behavior reflects the Sigmoid kernel’s tendency to saturate or misrepresent the input space when parameters are misaligned with the data distribution. Compared to the other kernels, the Sigmoid kernel demonstrated the least robustness and highest variance in performance. While theoretically capable of capturing complex nonlinearities, its practical utility in this time series context is limited by its sensitivity to hyperparameter interactions and lack of generalization stability.

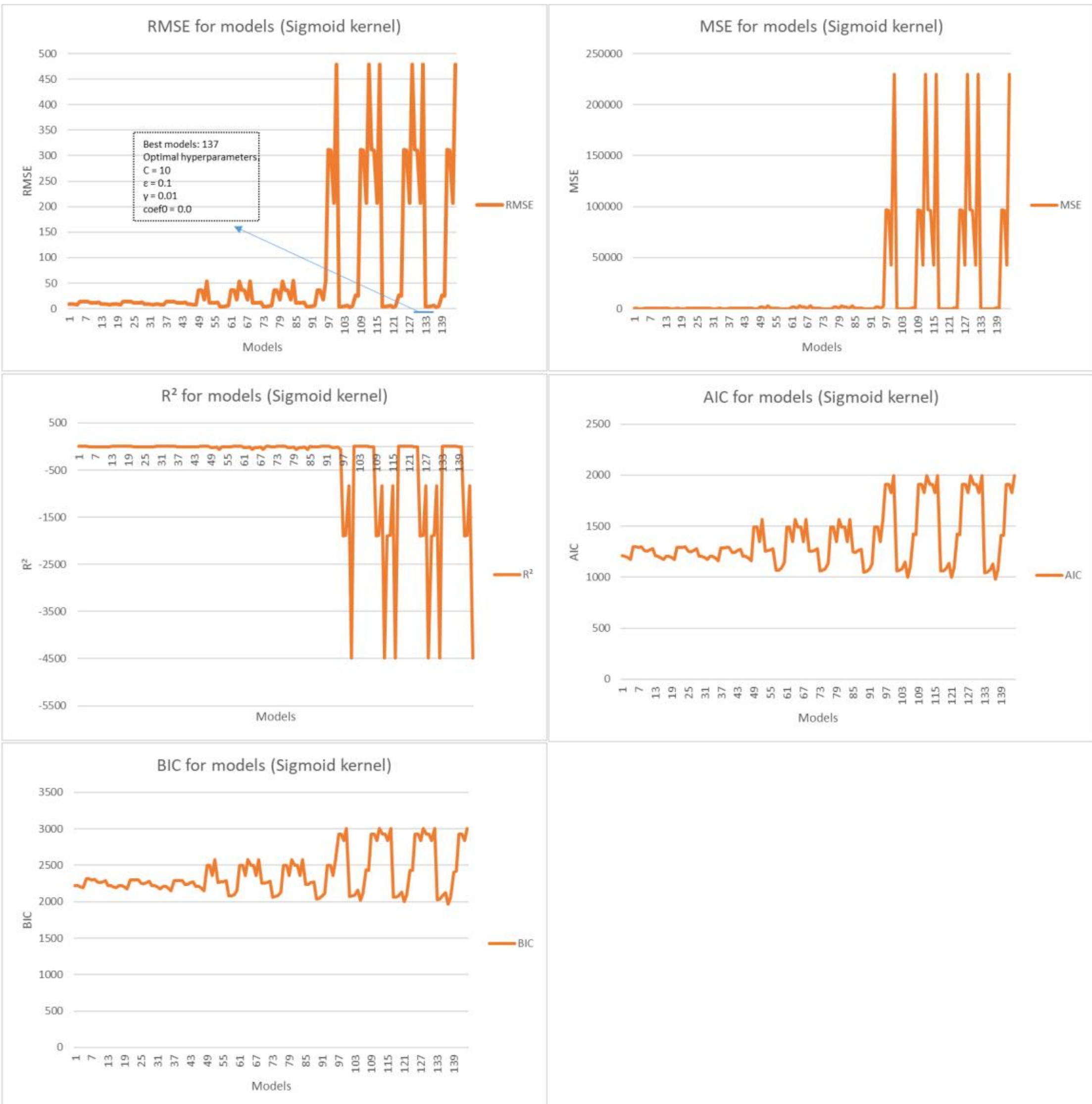

Figure (4): Metric results for models (Sigmoid kernel)

5- COMPARATIVE ANALYSIS AND SYNTHESIS

Across all four kernels, the linear kernel (Figure 1) emerged as the most stable and accurate, achieving the lowest RMSE of 2.7332 with minimal tuning complexity. The RBF kernel (Figure 2) showed strong performance under optimal conditions but required precise tuning of gamma and C to avoid underfitting. The polynomial kernel (Figure 3) offered competitive results, particularly with quadratic configurations, but was more sensitive to interactions among degree, gamma, and coef0. In contrast, the sigmoid kernel (Figure 4) exhibited erratic behavior and extreme sensitivity, making it the least reliable choice for this dataset.

These findings suggest that for structured, moderately noisy time series data, simpler kernels like the linear SVR not only outperform more complex alternatives but also offer greater interpretability and robustness. While nonlinear kernels can capture intricate patterns, their performance is highly contingent on careful hyperparameter selection and may not justify the added complexity in scenarios where the underlying signal is smooth and well-behaved.

6- KERNEL PERFORMANCE EVALUATIONS

Figure 5 shows a comparative visualization of SVR kernel performance, it reveals that the Linear kernel achieved the most favorable results across all evaluated metrics. With the lowest RMSE (2.7332) and MSE (7.4706), it demonstrated superior predictive accuracy and minimal error variance. Its R^2 value of 0.8541 further confirms strong explanatory power, indicating that the model captures a substantial portion of the variance in the time series data. Additionally, the Linear kernel yielded the lowest AIC (971.08) and BIC (1971.46), suggesting optimal model parsimony and generalization. In contrast, the Sigmoid kernel exhibited the highest RMSE (3.0109) and MSE (9.0656), along with the weakest R^2 (0.8229), reflecting reduced predictive reliability. The Polynomial and RBF kernels performed moderately, with slightly higher error metrics and complexity penalties. Overall, the results underscore the effectiveness of the Linear kernel when paired with targeted hyperparameter tuning, making it the most robust choice for this simulated time series forecasting task.

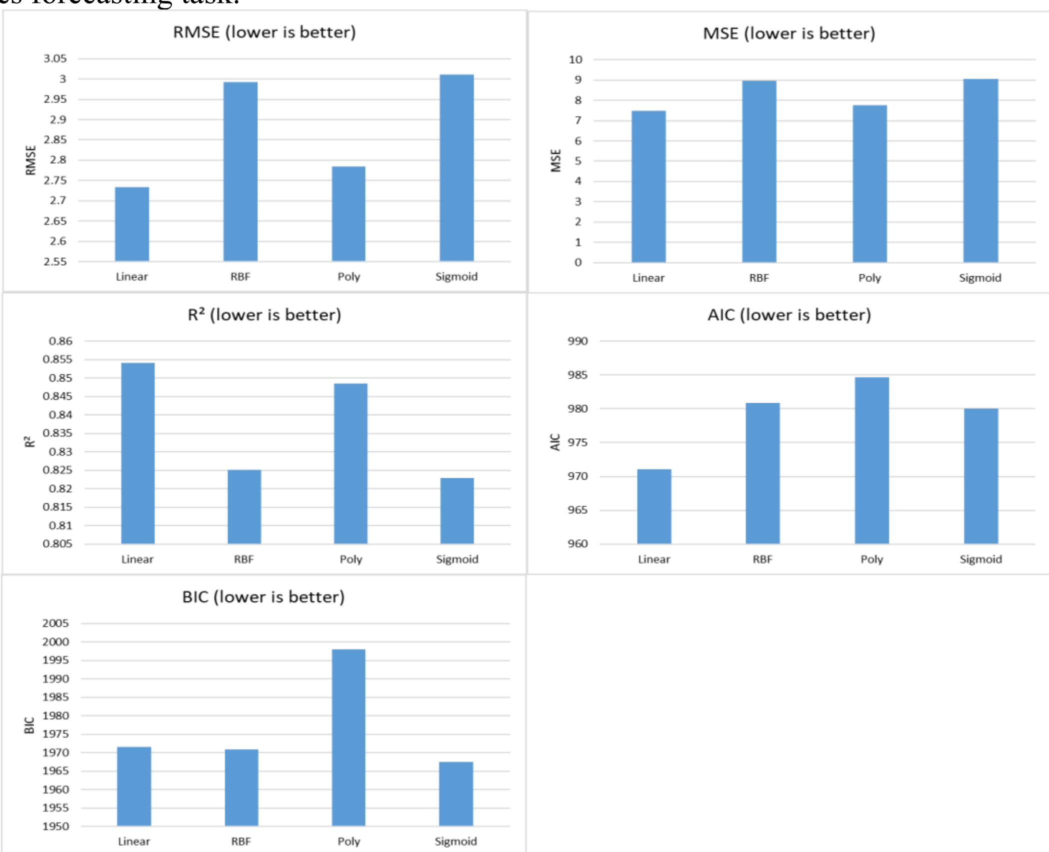

Figure (5): Kernel performance evaluations

7- PERFORMANCE OF THE BEST KERNEL (LINEAR) MODEL WITH THE BEST HYPERPARAMETER

The best-performing SVR model used a linear kernel with optimized hyperparameters: C = 10, epsilon = 0.05, and gamma = 'scale'. As shown in Figure 6 the graph, the model's predictions align closely with the actual time series across both training and test sets, capturing trend and seasonality without overfitting. The smooth overlap especially in the test segment demonstrates strong generalization and predictive stability. These hyperparameters struck a balance between bias and variance: C = 10 allowed flexible fitting without excessive sensitivity to noise, epsilon = 0.05 encouraged precision, and gamma = 'scale' ensured consistent scaling. Together, they enabled the linear SVR to model the structured signal effectively while maintaining interpretability.

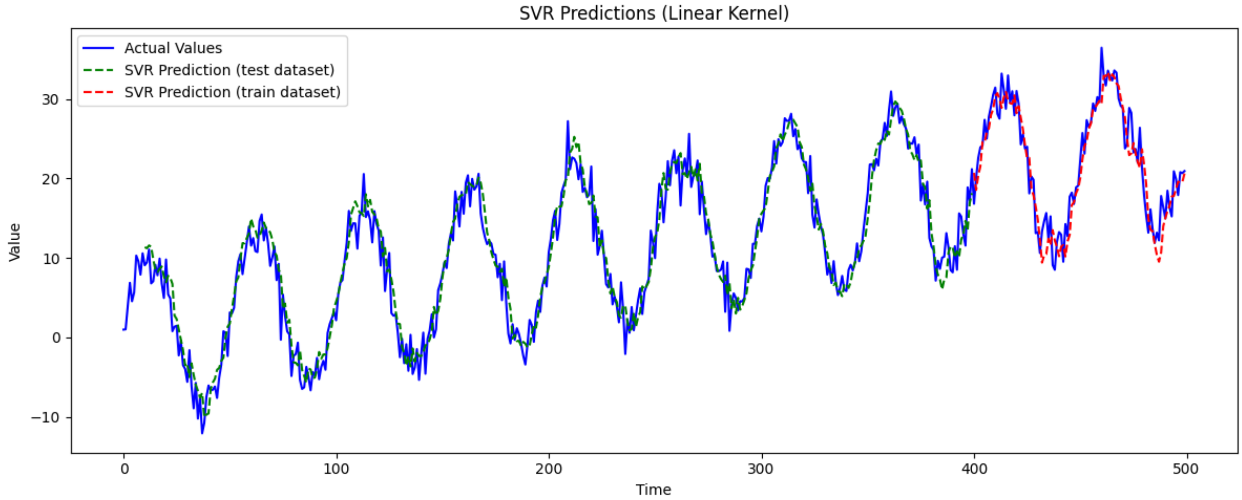

Figure (6): SVR Predictions (Linear Kernel)

Figure 7 illustrates the residuals versus predicted values for the linear SVR model, offering a diagnostic view of its performance. The residuals defined as the difference between actual and predicted values are scattered symmetrically around the zero line, with no discernible pattern or curvature. This indicates that the model does not systematically overestimate or underestimate across the prediction range, and that it has captured the underlying structure of the time series without introducing bias. The absence of funnel shapes or clustered errors suggests homoscedasticity, meaning the variance of the residuals remains consistent across different output levels. Overall, the plot confirms that the linear SVR model, with its tuned hyperparameters (C = 10, epsilon = 0.05, gamma = 'scale'), generalizes well and maintains predictive stability across the full signal.

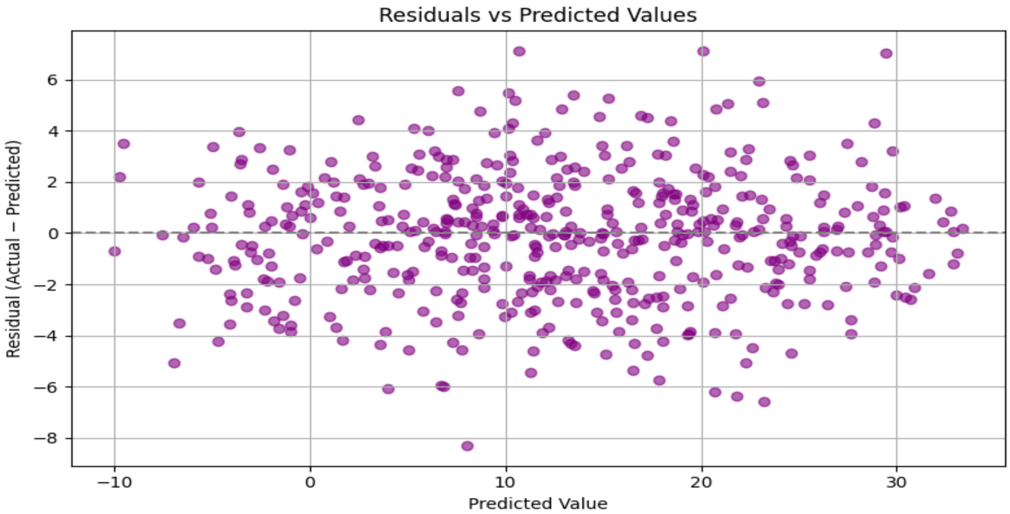

Figure (7): Scatter Plot

Figure 8 illustrates the RMSE comparison across four SVR kernels Linear, RBF, Polynomial, and Sigmoid. The Linear kernel achieved the lowest RMSE (2.7332), indicating the best predictive accuracy among the models tested. The Sigmoid kernel showed the highest RMSE (3.0109), suggesting weaker performance. Overall, the chart highlights how kernel choice significantly impacts SVR model precision

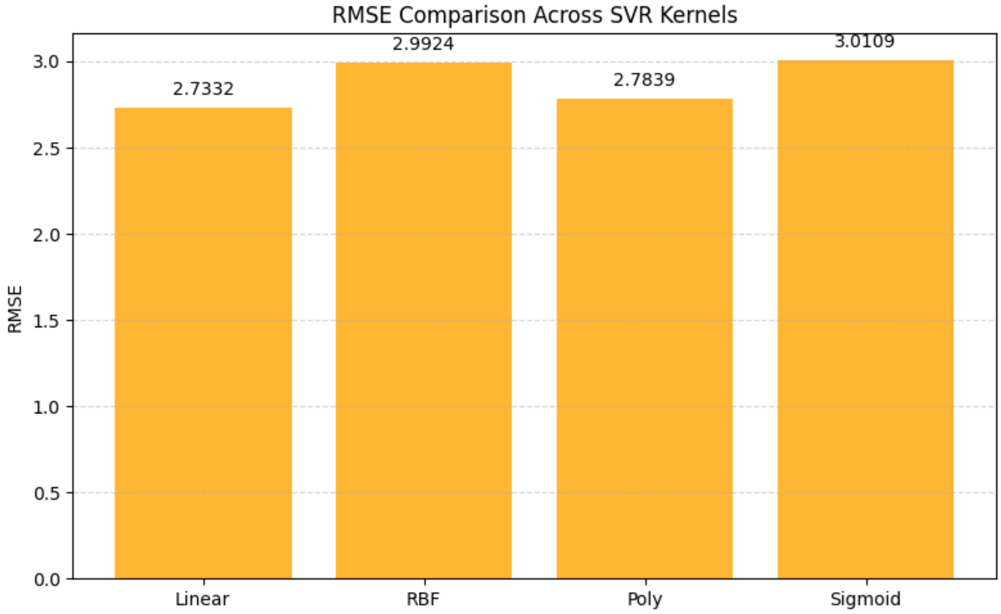

Figure (8): RMSE Comparison Across SVR Kernels

CONCLUSION

This study conducted a comprehensive benchmarking of Support Vector Regression (SVR) models across four kernel functions linear, radial basis function (RBF), polynomial, and sigmoid using a simulated seasonal time series dataset characterized by trend, periodicity, and moderate Gaussian noise. Through exhaustive hyperparameter tuning, each kernel was evaluated based on its predictive accuracy and sensitivity to configuration changes.

The linear kernel (Figure 1) consistently outperformed its nonlinear counterparts, achieving the lowest RMSE of 2.7332. Its simplicity and robustness across tuning scenarios underscore its suitability for structured time series data, especially when the underlying signal exhibits smooth and interpretable dynamics.

The RBF kernel (Figure 2) demonstrated strong performance under optimal conditions, with an RMSE of 2.9924, but required precise calibration of gamma and C to avoid underfitting or overfitting. While capable of modeling nonlinear relationships, its sensitivity to kernel width and regularization made it less reliable in broader applications.

The polynomial kernel (Figure 3) offered competitive results, particularly with quadratic configurations (degree = 2), yielding an RMSE of 2.7839. However, its performance was highly dependent on the interplay between degree, gamma, and coef0, making it more complex to tune and less robust across varying signal structures.

In contrast, the sigmoid kernel (Figure 4) exhibited the greatest volatility, with RMSE values ranging from 3.0109 to over 479.0, depending on the hyperparameter combination. Although it achieved moderate accuracy under specific conditions, its erratic behavior and extreme sensitivity to gamma and coef0 rendered it the least reliable choice for this dataset.

Taken together, these findings highlight the importance of kernel selection and hyperparameter tuning in SVR modeling. For time series data with structured patterns and moderate noise, the linear kernel offers a compelling balance of accuracy, interpretability, and stability. While nonlinear kernels can capture more complex relationships, their added complexity and tuning sensitivity may not yield proportional gains in performance. Future work may explore hybrid approaches or adaptive kernel strategies, but for foundational forecasting tasks, linear SVR remains a robust and efficient baseline.

FUNDING

None.

ACKNOWLEDGEMENT

None.

CONFLICT OF INTEREST

The author declares no conflict of interest.

References

1- Abdulraheem SA, Aliyu S, Abdullahi FB. Hyper-parameter tuning for support vector machine using an improved cat swarm optimization algorithm. Journal of the Nigerian Society of Physical Sciences. 2023 Sep 8:1007-.
2- Anand P, Rastogi R, Chandra S. Generalized—Loss Function-Based Regression. InMachine intelligence and signal analysis 2018 Aug 8 (pp. 395-409). Singapore: Springer Singapore.
3- Awad M, Khanna R. Support vector regression. InEfficient learning machines: Theories, concepts, and applications for engineers and system designers 2015 Apr 27 (pp. 67-80). Berkeley, CA: Apress.
4- Aziz AA, Mahmood HO, Rahim SA, Maaroof RS, Taher HA. Using Optimizing Parameters Support Vector Regression Model to Predict Potassium Ratio in Carb Fish. Journal of Survey in Fisheries Sciences. 2023;10(3S):4931-7.
5- He W, Wang Z, Jiang H. Model optimizing and feature selecting for support vector regression in time series forecasting. Neurocomputing. 2008 Dec 1;72(1-3):600-11.
6- Hodson TO. Root mean square error (RMSE) or mean absolute error (MAE): When to use them or not. Geoscientific Model Development Discussions. 2022 Mar 11;2022:1-0.
7- Kurani A, Doshi P, Vakharia A, Shah M. A comprehensive comparative study of artificial neural network (ANN) and support vector machines (SVM) on stock forecasting. Annals of Data Science. 2023 Feb;10(1):183-208.
8- Montesinos López OA, Montesinos López A, Crossa J. Overfitting, model tuning, and evaluation of prediction performance. InMultivariate statistical machine learning methods for genomic prediction 2022 Jan 14 (pp. 109-139). Cham: Springer International Publishing.
9- Nandi S, Badhe Y, Lonari J, Sridevi U, Rao BS, Tambe SS, Kulkarni BD. Hybrid process modeling and optimization strategies integrating neural networks/support vector regression and genetic algorithms: study of benzene isopropylation on Hbeta catalyst. Chemical Engineering Journal. 2004 Feb 15;97(2-3):115-29.
10- Schölkopf, B., & Smola, A. J. Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond. Cambridge, MA: MIT Press, 2002
11- Smola AJ, Schölkopf B. A tutorial on support vector regression. Statistics and computing. 2004 Aug;14(3):199-222.
12- Vinge R, McKelvey T. Understanding support vector machines with polynomial kernels. In2019 27th European signal processing conference (EUSIPCO) 2019 Sep 2 (pp. 1-5). IEEE.
13- Wei J, He X. Support vector regression model with variant tolerance. Measurement and Control. 2023 Nov;56(9-10):1705-19.