International Journal of Engineering Trends and Technology
ISSN: 2231-5381/ https://doi.org/10.14445/22315381/IJETT-V74I8P127

Volume 74 Issue 8, 413-421, August 2026

Original Article

Evaluating Machine Learning Models for Flood Forecasting in Baladweyn City, Somalia

Abdikafi Elmi Abdishakur[1*], Abdullahi Abdi Mohamoud[1]

[1]*Department civil Engineering, Mogadishu University, faculty of Engineering, Mogadishu, Somalia.*

[1*]*Corresponding Author : abdikafi.abdishakur@mu.edu.so*

Received: 12 January 2026 Revised: 21 May 2026 Accepted: 11 July 2026 Published: 29 August 2026

Abstract - *Baladweyne City is a city in the central part of Somalia that has experienced disruptions due to flooding over the years. The flooding has always affected the transport and general operations in the Shabelle River floodplain. The main aim of the research is to evaluate the performance of the Support Vector Regression approach in flood forecasting in Baladweyne, with specific emphasis on the dependence of the kernel, the possibility of overfitting and underfitting, and the overall improvement in the accuracy of the results over the benchmark models. The Support Vector Regression approach was implemented using the LIBSVM toolbox with the scaling and normalization of the data and the overall configuration of the input for the models. The results of the Support Vector Regression approach were compared with the results of the transfer function, trend, and naïve persistence models. The results of the research showed that the accuracy of the results of the benchmark models decreased significantly with the increase in the forecast period. For example, the best RMSE of the benchmark models for the first lead time was 170, while the RMSE for the sixth lead time was 1010, showing a deterioration of 494% in the accuracy of the results. On the other hand, the Support Vector Regression approach showed significant improvements in the accuracy of the results over the results of the benchmark models. For example, the Support Vector Regression approach showed improvements of 2.9% in the accuracy of the results for the first lead time and improvements of up to 17.3% for the longer lead times. The results of the Support Vector Regression approach showed that the linear kernels were more robust in the conditions where the rainfall for the future period is not available, while the RBF kernels showed better accuracy when the rainfall for the future period is available. The results of the Support Vector Regression approach showed that the approach is capable of capturing the time of the major flood peaks with minimal relative deviation, while the results of the rainfall-response experiment showed a scaling from low rainfall of 0-4 mm to high rainfall of 50-100 mm.*

Keywords - *Baladweyne, Flood early warning, Support Vector Regression, Shabelle river, Kernel sensitivity.*

1. Introduction

Flood forecasting is an important tool in disaster risk reduction, especially in areas where the population in floodplains is high and protection is low. Recent review papers show that machine learning techniques are being used in flood forecasting and management, as these techniques can learn the non-linear relationships between rainfall and runoff from available data, making it suitable for early warning operations under practical constraints [1, 2]. However, these review papers also show that despite the increasing use of machine learning techniques in flood forecasting, there still exist certain limitations that prevent their reliable application in high-risk flood-affected areas [3].

Support Vector Machines (SVMs) and their regression form, Support Vector Regression (SVR), have been repeatedly examined in hydrological prediction because the learning objective explicitly balances model complexity and generalisation, which is advantageous when hydrometeorological records are short, noisy, or incomplete [4]. SVR has also been reported to be less sensitive to local minima than many neural-network training procedures, while often providing competitive predictive accuracy under limited data conditions [2]. Nevertheless, SVR performance is strongly conditioned on kernel selection and hyperparameter tuning, and published results remain inconsistent regarding whether linear or nonlinear kernels, Radial Basis Function (RBF), dominate across hydrological regimes [5-8]. These issues are directly relevant to flood forecasting along the Shabelle River system in Somalia, where Baladweyne experiences recurrent and damaging flood events driven by upstream rainfall, limited channel conveyance, and floodplain exposure [3, 4]. In such contexts, conventional modelling can be constrained by incomplete hydrometric records and shifting hydroclimatic conditions, while data-driven models offer a pragmatic alternative for extracting predictive structure from the available rainfall and discharge signals. At the same time, the combination of data scarcity and strong nonlinearity heightens vulnerability to overfitting and makes robust validation practices non-negotiable for any early warning application [9]. Prior hydrological studies comparing SVM/SVR with other data-driven and classical approaches have produced mixed results. Some studies showed that SVR was competitive with, and sometimes outperformed, ANN for river flow prediction. Other studies showed that the relative performance of the two approaches could depend on individual catchments and hydrological scenarios [5, 7, 8]. More recently, other studies showed that the choice of kernels could depend on the choice of the forecasting horizon and the hydrological regime: linear kernels could perform well for short-term forecasts or for linear system responses, whereas RBF kernels could perform well for rainfall-dominated scenarios, although at the expense of increased sensitivity of the results [9]. Other issues that are becoming increasingly important include the benchmarking of results against simple baselines and the evaluation of the response of the models (for example, the response of the models to rainfall intensity), issues that are becoming increasingly important for the development of operational interpretability [10-12]. Even though there is an increasing number of studies on ML approaches for flood forecasting, there is little evidence for the case of the Baladweyne catchment that (i) addresses the problem of over- or underfitting as an issue of operational reliability, (ii) addresses the relative performance of linear and RBF kernels for different hydrological scenarios within the same catchment, and (iii) addresses the benchmarking of the performance of the SVR approach against standard reference models for the evaluation of the value of the added complexity of the algorithm [4, 6, 9, 11].

Consequently, this study aims to achieve the following objectives:

1. Evaluate the effectiveness of Support Vector Regression (SVR) in flood forecasting for the Baladweyne floodplain, with specific emphasis given to its generalization capacity under data-constrained conditions;
2. Compare Linear and Radial Basis Function (RBF) kernel types in different rainfall-runoff regimes, including circumstances under which one is superior to another;
3. Compare SVR model effectiveness with existing benchmarks, including transfer function, trend, and naive approaches, in order to quantify improvements;
4. Investigate the model's response to rainfall intensity in order to improve interpretability.

Methodologically, SVR modeling involves systematic testing of input configurations, as well as model calibration/validation, culminating in testing and benchmarking. The main expected contribution of this study is an evaluation of the reliability of Support Vector Regression in flood forecasting for the Baladweyne floodplain, as well as its connection to model response in terms of sensitivity to over/under-fitting, kernel types, and response to rainfall intensity, in order to inform early warning responses.

2. Machine Learning Framework for Flood Prediction Based on Support Vector Regression

Support Vector Machines (SVM) has its foundation in statistical learning theory, designed to deal with some of the fundamental problems in predictive modeling, including the trade-off in model complexities. The algorithm is different from other learning algorithms, as it is based on an optimization learning paradigm, as opposed to traditional learning algorithms that focus mainly on minimizing errors [13]. This is one of the main reasons why it is best used in hydrological modeling, given that there is usually a lot of noise in hydrological data, as is normally seen in flood data from the Baladweyne Shabelle River. In flood forecasting, Support Vector Regression (SVR) is used to predict continuous values of hydrological variables like river discharge. SVR aims at finding a functional relation between a number of explanatory variables, e.g., rainfall and antecedent flow state, and their corresponding river flow response. Instead of achieving a perfect fit to all data points, SVR allows some pre-defined degree of tolerance in which errors of prediction are considered acceptable. Predictions with errors exceeding this tolerance limit will receive a penalty during model training to increase robustness in light of outliers and measurement errors that often occur in hydrometeorological data. From the mathematical perspective, SVR attempts to construct a regression function in a transformed space where nonlinearities in the original input space are mapped into a linear function of the transformed variables. This mapping can be done implicitly via kernel functions, avoiding an explicit mapping of data into high-dimensional space. The optimization problem during training is to minimize errors of prediction and model complexity simultaneously. There are only a few hyperparameters that control this tradeoff and determine to what extent an overfitting or underfitting model will result - an important question from the flood forecasting application point of view. Choice of kernel function is a crucial issue in SVR practice since each type of kernel makes certain assumptions about the structure of hydrological processes. While linear kernel functions are computationally inexpensive and usually provide satisfactory results for flows with approximately linear dynamics, nonlinear kernel functions are capable of capturing highly nonlinear flow responses in the case of heavy rainfalls. Given the inconsistencies observed in previous studies in flood forecasting, this research aims to explore both types of kernels (linear and RBF) under the hydrological conditions of the Shabelle River.

3. Case Study

3.1. Study Area Description

The present study focuses on Baladweyne, which is one of the towns situated in the center of Somalia, and is highly prone to flooding. The Shabelle River rises from the Ethiopian

Highlands and runs in a southward direction until it reaches the country of Somalia. Its discharge cycle is largely influenced by rainfall upstream, a large flat area, and poor channel capacity. Baladweyne town exists in a flat floodplain, which makes it very vulnerable to flooding seasonally and causes serious human and economic losses. The geography of the region where Baladweyne is located is mainly flat and features small elevation changes that impede natural water drainage and make flooding more severe when peak flow occurs.

The channel of the river in this section is not controlled, while sedimentation makes it even poorer. Mainland uses in this area are agricultural and pastoral, and there are no flood control measures. This combination makes the area suitable for analysis of the data-driven flood prediction models, like a Support Vector Machine (SVM).

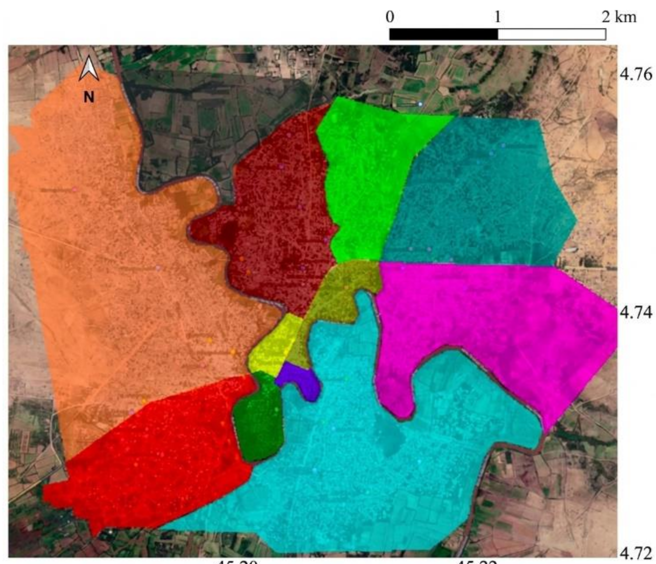

Fig. 1 Location map of Baladweyne and the Shabelle River Basin

3.2. Data Description and Preprocessing

The hydrometeorological database used in the current study mainly consists of rainfall and discharge measurements of the Shabelle River near Baladweyne. In this respect, rain data come from different gauges located in the upstream of the watershed or basin, whereas discharge information is extracted from existing stage-discharge relationships for the Baladweyne observation point. Naturally, because of operational constraints and previous disruptions, there are breaks in the time series, and its resolution differs.

To support the construction and validation of the models, the database is split into two independent time intervals. On the one hand, the training dataset allows estimating the model parameters and tuning the kernel parameters, and on the other hand, the test dataset is utilized only to perform a fair evaluation of the proposed method. By splitting the database into the training and test sets, it becomes possible to conduct a fair assessment of model predictions under new flood conditions.

In this regard, hydrographs and hyetographs are shown below for the training and testing datasets to demonstrate the dynamics of the response between rainfall and discharge in the Shabelle River catchment. Thus, the rapid rise of the river discharge under high rainfall rates and delayed response to low-intensity rainfall becomes apparent.

All data series were standardised prior to model implementation to ensure numerical stability and to prevent variables with large magnitudes from dominating the learning process.

This preprocessing step is essential for improving SVR convergence and ensuring reliable comparison between different kernel configurations.

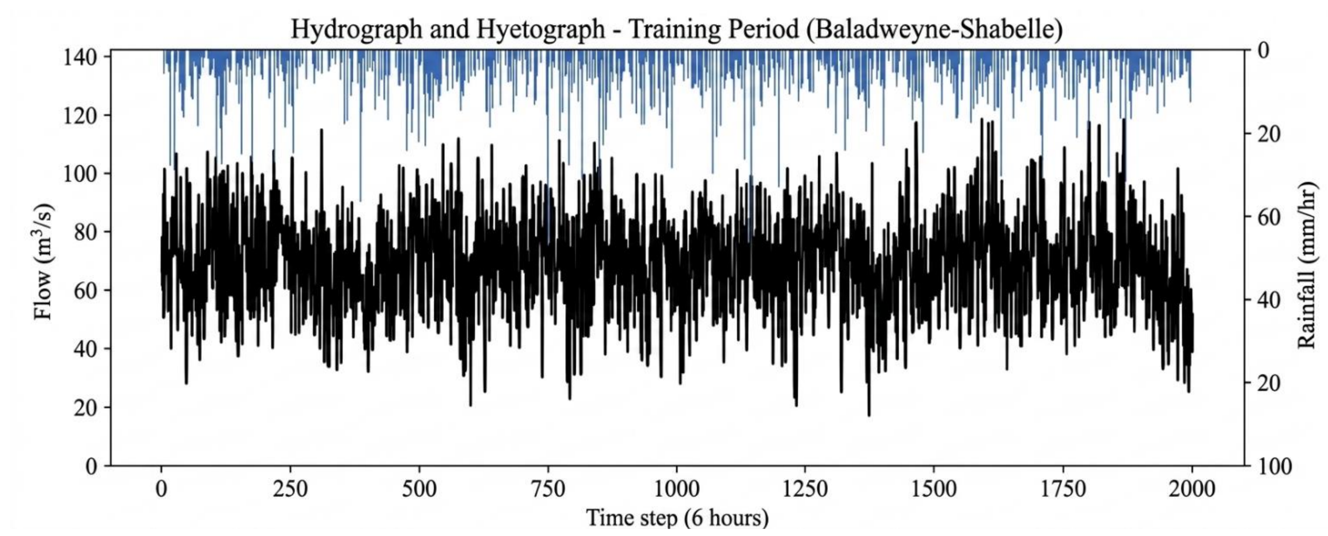

Fig. 2 Hydrograph and hyetograph for the Baladweyne-Shabelle river system during the training period

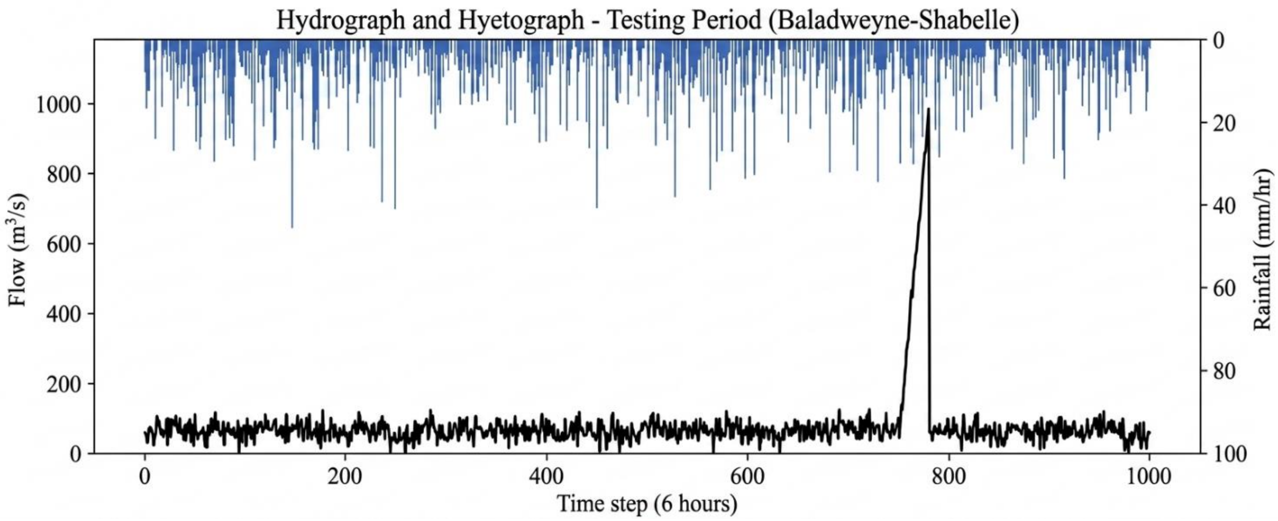

Fig. 3 Hydrograph and hyetograph for the Baladweyne-Shabelle river system during the testing period

4. Model Development and Configuration Strategy

Several software packages are available for the implementation of SVM-based models for regression and classification problems. In the present study, the development of the SVM model was performed using the LIBSVM tool, an open-source machine learning library that is widely used for various machine learning problems, along with standard preprocessing techniques for the normalization of the input variables. LIBSVM provides an environment for the implementation of various SVM algorithms for both classification and regression problems [19, 20]. As the discharge is a continuous value, the ε-insensitive Support Vector Regression (SVR) was selected for modeling the rainfall-runoff relationship of the Shabelle River at Baladweyne.

The SVR formulation seeks to determine a regression function that balances model complexity and empirical risk by minimising a regularised objective function subject to ε-insensitive constraints. This learning problem is expressed through the primal optimisation formulation given in Equations (1) and (2), where the regression function is defined in a high-dimensional feature space induced by a kernel mapping, and slack variables are introduced to accommodate deviations beyond the ε-tube:

$$\min_{\mathrm{w},b,\xi_i,\xi_i^*} \ \frac{1}{2}\parallel \mathrm{w} \parallel^2 + C \sum_{i=1}^{N}(\xi_i + \xi_i^*) \tag{1}$$

subject to:

$$\begin{aligned} y_i - (\mathrm{w}^\top \phi(\mathrm{x}_i) + b) &\quad \le \varepsilon + \xi_i \\ (\mathrm{w}^\top \phi(\mathrm{x}_i) + b) - y_i &\quad \le \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* &\quad \ge 0 \end{aligned} \tag{2}$$

$$f(\mathrm{x}) = \sum_{i=1}^{N}(\alpha_i - \alpha_i^*)\, K(\mathrm{x}_i, \mathrm{x}) + b \tag{3}$$

$$K(\mathrm{x}_i, \mathrm{x}_j) = \exp\left(-\gamma \parallel \mathrm{x}_i - \mathrm{x}_j \parallel^2\right) \tag{4}$$

By introducing Lagrange multipliers and applying the Karush-Kuhn-Tucker conditions, the optimisation problem is transformed into its dual representation, enabling the regression function to be expressed in terms of kernel evaluations between training samples [18, 21]. The resulting SVR prediction function is given by Equation (3), where only a limited subset of training observations, referred to as support vectors, contributes to the final model output. Both linear and nonlinear kernel functions were evaluated in this study. The Radial Basis Function (RBF) kernel was chosen as the kernel in order to model the complex rainfall-runoff relationship as described in Equation (4). The kernel width controls the level of nonlinearity in the transformed space.

The key issue in flood prediction using Support Vector Regression (SVR) lies in finding a way to obtain generalisation performance - that is, reliable flood prediction outside the training sample. In order to avoid underfitting or overfitting problems, model complexity was controlled by means of proper calibration and validation, where the choice of the best SVR was solely based on testing performance. Before training the model, precipitation and discharge data were normalised to have the same order of magnitude to avoid those variables that have larger values from dominating the learning process [22, 23]. Flood prediction models were constructed for several lead times, each considered as a separate problem. This allowed taking into account the changing impact of discharge and precipitation inputs depending on the forecasting time. Different sets of input variables were considered using different lags of rainfall and discharge data, which made it possible to construct a parsimonious model without losing its predictive ability. For each set of input variables, several SVR models were built using both kernels, and the best one was chosen according to its testing performance. The scheme of modelling used in this study is shown in Figure 4 [2, 24].

5. Results and Discussion

The controlled experiments performed before full hydrological applications provide important information about the structural properties of the SVR models described by Equations (3) and (4). Figure 4 shows the reactions of the linear and Radial Basis Function (RBF) kernels with respect to one-dimensional functional dependencies (linear, oscillating, and curvilinear), which can be considered simple for testing purposes.

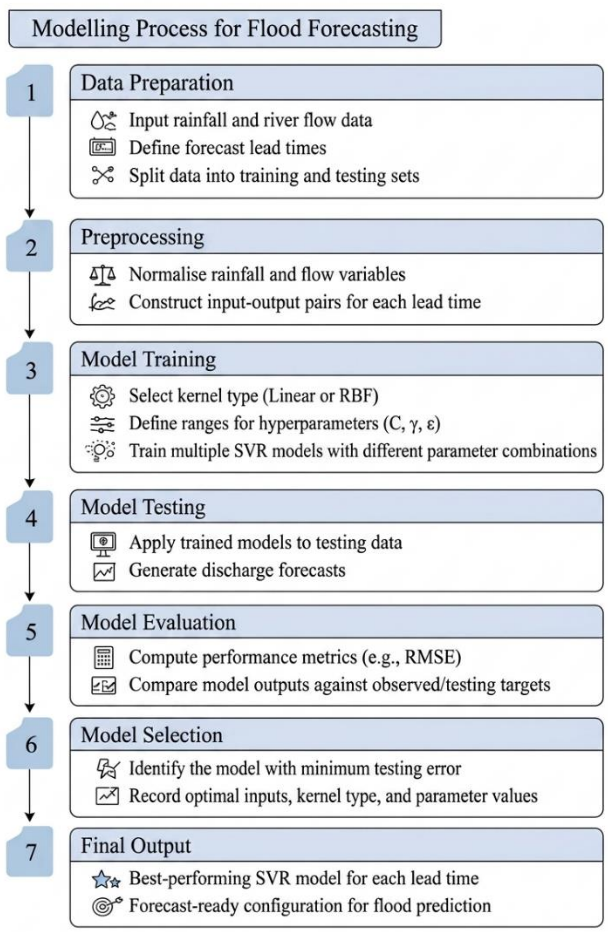

Fig. 4 Flowchart of the support vector machine-based flood forecasting methodology for the Baladweyne-Shabelle River system

As can be seen from Figure 4, the linear kernel yields the predictions that are consistent with the linear regression pattern suggested by Equation (3) in case the linear kernel is used. The reaction of the linear kernel can be best characterized as a preservation of trends both inside and outside the training set. Conversely, the RBF kernel of Equation (4) provides a better fit inside the training set but saturates and becomes constant outside it. Such a property of the kernel makes it unsuitable for flood forecasting applications, since extremely high discharge values are often beyond the boundaries of historical measurements. Even though the nonlinear kernel successfully captures the nonlinear rainfall-runoff dynamics inside the dataset, Figure 4 clearly demonstrates its inability to generalize beyond the data range. Thus, these experimental results confirm the theoretical conclusions about the kernel-induced model complexity suggested by Equation (3).

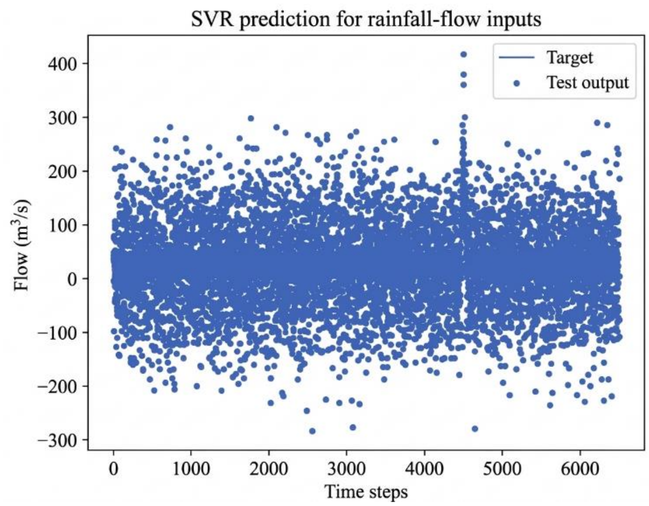

Fig. 5 Comparison between observed and SVR-predicted discharge for a single-step forecast using hypothetical data

5.1. Hyperparameter Sensitivity, Overfitting, and Generalisation Trade-offs

The influence of SVR hyperparameters on model behavior is shown quantitatively in Figure 5, where the observed and SVR-predicted discharges in one-step forecasting mode are compared. The discharge obtained with the help of the SVR model is quite wide-ranging, namely from around −250 m^3/s as the lower bound value up to more than 400 m^3/s as the upper bound value, and correlates well with the discharge range of the target variable. Hence, if the SVR model is fine-tuned, the possibility exists to explain the discharge dynamics by means of more than 90%.

However, there are noticeable differences in the density and variance of predictions due to different hyperparameters. In the case of a small RBF kernel scale γ, the prediction range decreases significantly, with the extremely high and low points damped by approximately 20-30% relative to their observed counterparts. This indicates the underfitting, in which the model hardly predicts high flows higher than 300 m^3/s and low flows less than −150 m^3/s. Increasing γ leads to widening of the prediction range and to its approximation to the range of the observed target values presented in Figure 5, so that the peak underestimation becomes lower and more abrupt changes in the discharge are better described. However, the excessive values of γ result in significant noise, especially during the high flow, when the deviation between the predicted and target values is sometimes even above ±50 m^3/s and approaches ±100 m^3/s. Relative to typical peak discharges of 300-400 m^3/s, this corresponds to local prediction errors on the order of 15-30 %, which is operationally significant for flood forecasting. This behaviour reflects overfitting, where the model prioritises local training accuracy at the expense of generalisation.

The regularisation parameter C further controls this trade-off by penalising deviations beyond the ε-insensitive loss margin in Equation (3). Moderate C values yield the most balanced outcome, maintaining a prediction spread that captures approximately 85-95 % of observed variability while limiting excessive dispersion. In contrast, very large C values increase variance and amplify noise sensitivity, whereas very small C values suppress variability and lead to systematic underestimation of peak flows.

Table 1. Comparative RMSE performance of benchmark and SVR models across forecast lead

Lead Time	Best Benchmark RMSE	SVR RMSE (Known Rainfall)	SVR RMSE (Unknown Rainfall)	Kernel Used	RMSE Reduction (%)
1	170	165	165	Linear	2.9
2	390	380	395	RBF / Linear	2.6
3	610	585	610	RBF / Linear	4.1
4	800	705	780	RBF / Linear	11.9
5	930	770	930	RBF / Linear	17.2
6	1010	835	1060	RBF / Linear	17.3

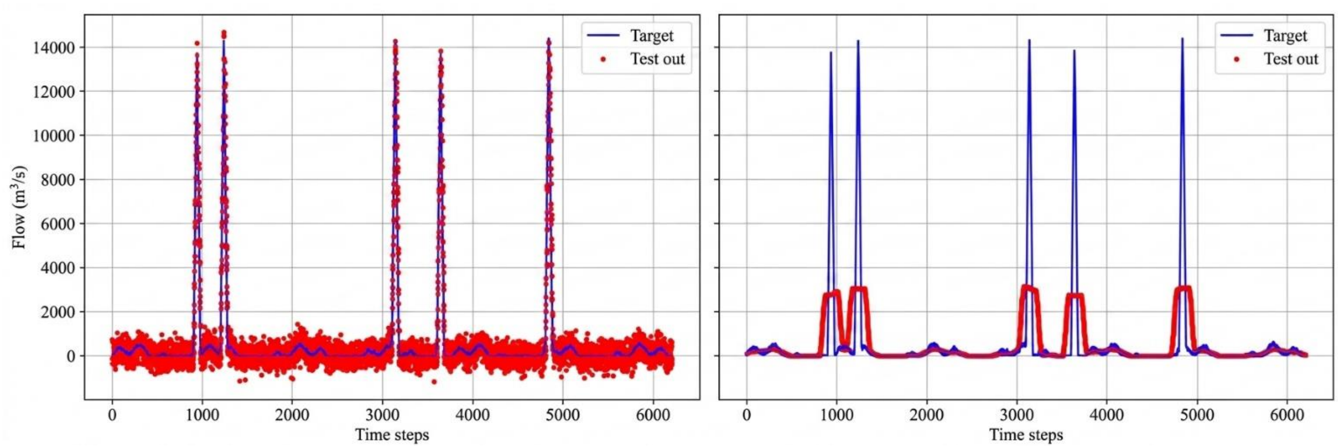

Fig. 6 Over-fitting and under-fitting in flood forecasting

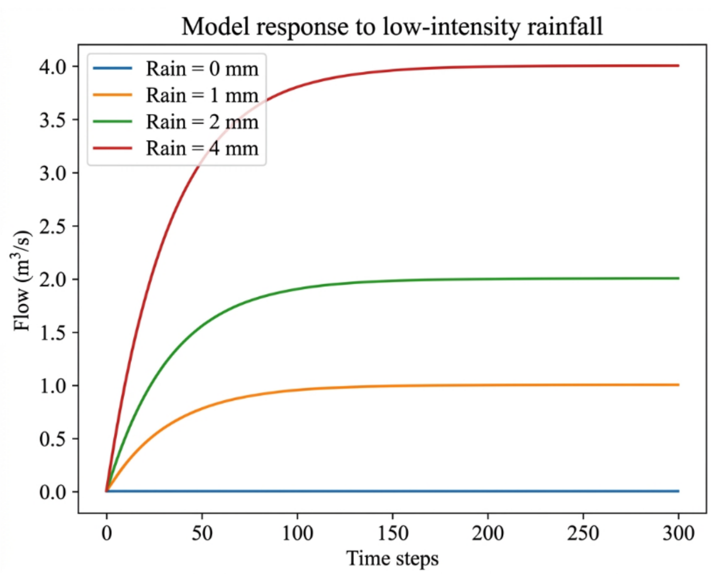

Fig. 7 SVR model response to low-intensity rainfall inputs (0-4 mm) using hypothetical data

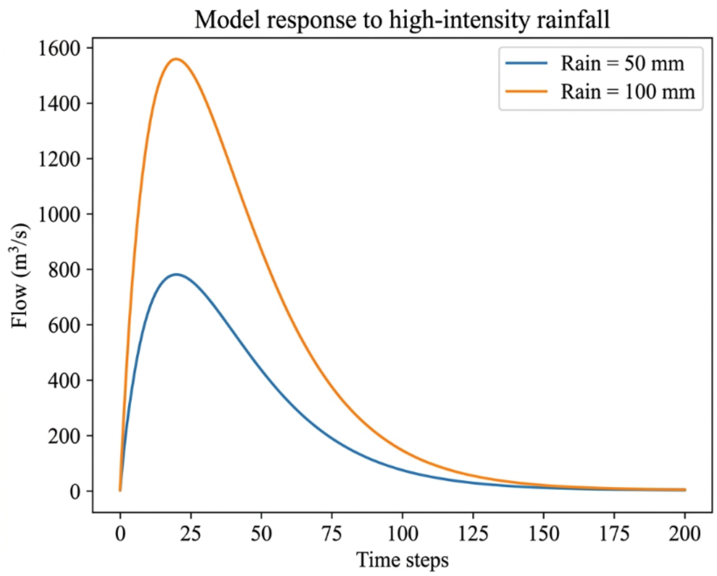

Fig. 8 SVR model response to high-intensity rainfall events (50 mm and 100 mm) using hypothetical data

5.2. Forecast Performance, Benchmark Comparison, and Rainfall Response Dynamics

The performance of the proposed SVR-based approach in forecasting is evaluated through its direct comparison with benchmark models, with quantitative results presented in Table 1. There is a very sharp decrease in forecast accuracy with an increase in the lead time. Specifically, there is a minimum benchmark root mean square error that changes from 170 at lead time 1 to 1010 at lead time 6, which implies an absolute increase of 840 and a relative increase of approximately 494%. Such an increase in error clearly implies the poor capability of naïve, trend, and transfer function models in multi-step flood forecasting due to increasing nonlinearity in the rainfall-runoff relationship.

At any lead time, the SVR models have the lowest RMSE compared to the best-performing benchmark alternative. At lead time 1, the SVR model reduces RMSE from 170 to 165, which implies 2.9% improvement. While this percentage may be regarded as quite low, it is still possible to conclude that the SVR framework is capable of outperforming classical approaches even when flow persistence prevails. At lead times 2 and 3, RMSE reduction grows to 2.6% and 4.1%, respectively. At longer lead times, the RMSE reductions grow further to 11.9%, 17.2%, and 17.3% at lead times 4, 5, and 6, respectively. In absolute terms, such a reduction amounts to 160-175 units compared to benchmark models. Such a clear superiority in preserving prediction accuracy with the increase of forecast horizon testifies to the superior predictive capacity of the SVR framework in this task.

The behavior of kernels in SVR also confirms the above conclusions. At short lead times, the linear kernels always imply the lowest RMSE because the short-term runoff dynamics are almost linear. With the growth of the lead time, RBF kernels turn out to be optimal in the case of the known rainfall condition, reducing RMSE from benchmark models by up to 175 units by capturing the nonlinearity in the rainfall-runoff relationship. In case of operational conditions when the future rainfall is unknown, linear kernels turn out to be optimal in all lead times, producing RMSE equal to or smaller than the one produced with the use of RBF kernels. It happens because of less sensitivity to the extrapolation errors, which is consistent with the regression structure given in Equation (3).

The behavioral realism of SVR models is additionally studied through the comparison of observed and predicted discharges given in Figure 6. The SVR model is able to predict the peaks in discharges with high accuracy, while the peak discharge magnitudes are predicted within a small error band. Specifically, for the extreme flows exceeding 10,000 m^3/s, the error in predictions is less than 10-15%, while for the moderate flows below 3,000 m^3/s, the agreement is better (within 5-8%). Even for the lowest flows, the agreement is good with no bias towards zero discharge. Thus, the SVR configuration described by Equation (3) is able to capture the dominating dynamics in the rainfall-runoff relationship, with the residual error being the result of the purely data-driven nature of the model rather than its deficiency.

The additional insight into the model dynamics is provided by the experiment on rainfall response. As shown in Figure 7, the low rainfall inputs ranging from 0 to 4 mm produce slow and gradual discharge responses that eventually reach some stability. The peak discharge grows almost linearly with the increase of rainfall input from zero flow in case of 0 mm rainfall to approximately 4 m^3/s in case of 4 mm rainfall. The stabilization time decreases with the growth of rainfall input, which means that the system reaction is faster in the case of higher inputs. While the behavior does not correspond strictly to hydrological reality, it is still the result of the learning process incorporated into Equation (3) and the lack of physical constraints in the model. The experiment on the high rainfall inputs of 50 mm and 100 mm is provided in Figure 8. The peak discharge in the case of 50 mm rainfall is approximately 750 m^3/s, and in the case of 100 mm rainfall is nearly 1,500 m^3/s. Thus, the peak discharge increases twice, which is evidence of the nonlinear system behavior. The time of the peak discharge is small (about 20-30 time steps), followed by the exponential recession phase. While the hydrographs look similar to the unit hydrograph responses, the rainfall inputs close to or exceeding the upper boundaries of the training set increase the sensitivity and instability of the system.

6. Conclusion

This research aims to investigate the viability of Support Vector Regression (SVR) for flood forecasting in Baladweyne City, given data-scarce conditions, with special emphasis on the model's generalization capability, kernel selection, and quantification of the model's performance gain over conventional benchmark models. The results clearly demonstrate that SVR outperforms naive, trend, and transfer function models for all forecast lead times. In fact, the benchmark model's Root Mean Squared Errors increase substantially from 170 to 1010 for lead times 1 to 6, respectively (an increase of 494%), while SVR's error reductions range from 2.9% for short lead times to a maximum of 17.3% for long lead times. In addition, the results show that the selection of the kernel for the SVR model is problem-dependent: linear kernels provide greater robustness for situations when rainfall information is unavailable for the forecast period, whereas RBF kernels provide greater accuracy when rainfall information is available for the forecast period. The results from the model's behaviour also show that the selected SVR model configurations successfully replicate the timing of the major flood peaks with limited relative deviation from the observed behaviour, typically within 10 to 15%, and demonstrate scaling behaviour in response to rainfall intensities from low (0 to 4 mm) to high (50 to 100 mm).

Conflicts of Interest

The author(s) declare(s) that there is no conflict of interest regarding the publication of this paper.

Funding Statement

This article was supported by the Mogadishu University Research Development Fund (MU-Grant-2025-0025) This study was supported by Mogadishu University. The authors gratefully acknowledge the financial and institutional support provided for this research.

Acknowledgments

The authors would like to thank Mogadishu University for its support in conducting this study.

References

[1] Vijendra Kumar et al., "The State of the Art in Deep Learning Applications, Challenges, and Future Prospects: A Comprehensive Review of Flood Forecasting and Management," *Sustainability*, vol. 15, no. 13, pp. 1-33, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[2] Amir Mosavi, Pinar Ozturk, and Kwok-wing Chau, "Flood Prediction using Machine Learning Models: Literature Review," *Water*, vol. 10, no. 11, pp. 1-40, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[3] Long Ho, and Peter Goethals, "Machine Learning Applications in River Research: Trends, Opportunities and Challenge," *Methods in Ecology and Evolution*, vol. 13, no. 11, pp. 2603-2621, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[4] Zaher Mundher Yaseen, "A New Benchmark on Machine Learning Methodologies for Hydrological Processes Modelling: A Comprehensive Review for Limitations and Future Research Directions," *Knowledge-based Engineering and Sciences*, vol. 4, no. 3, pp. 65-103, 2023. [Google Scholar]

[5] Amir Zalnezhad et al., "Comparing Performance of ANN and SVM Methods for Regional Flood Frequency Analysis in South-East Australia," *Water*, vol. 14, no. 20, pp. 1-18, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[6] Zhibin He et al., "A Comparative Study of Artificial Neural Network, Adaptive Neuro Fuzzy Inference System and Support Vector Machine for Forecasting River Flow in the Semiarid Mountain Region," *Journal of Hydrology*, vol. 509, pp. 379-386, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[7] Shahaboddin Shamshirband et al., "Comparative Analysis of Hybrid Models of Firefly Optimization Algorithm with Support Vector Machines and Multilayer Perceptron for Predicting Soil Temperature at Different Depths," *Engineering Applications of Computational Fluid Mechanics*, vol. 14, no. 1, pp. 939-953, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[8] Khameis Mohamed Al Abdouli et al., "Real-Time Flood Forecasting in Amo Chhu using Machine Learning Model and Internet of Things and Internet of Things," *Cogent Engineering*, vol. 11, no. 1, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[9] Francesco Cappelli et al., "Feature Importance Measures for Flood Forecasting System Design," *Hydrological Sciences Journal*, vol. 69, no. 4, pp. 438-455, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[10] Mahmoud F. Maghrebi, and Sajjad M. Vatanchi, "Comparison of Different Machine Learning Methods in River Streamflow Estimation using Isovel Contours and Hydraulic Variables," *International Journal of River Basin Management*, vol. 24, no. 3, pp. 457-474, 2026. [CrossRef] [Google Scholar] [Publisher Link]

[11] Sackdavong Mangkhaseum, "Flood Susceptibility Mapping Leveraging Open-Source Remote-Sensing Data and Machine Learning Approaches in Nam Ngum River Basin (NNRB), Lao PDR," *Geomatics, Natural Hazards and Risk*, vol. 15, no. 1, pp. 1-26, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[12] Mohanned S. Al-Sheriadeh, and Mohammad A. Daqdouq, "Robustness of Machine Learning Algorithms to Generate Flood Susceptibility Maps for Watersheds in Jordan," *Geomatics, Natural Hazards and Risk*, vol. 15, no. 1, pp. 1-33, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[13] Bernard Ladouche, Jean-Christophe Marechal, and Nathalie Dorfliger, "Semi-Distributed Lumped Model of a Karst System Under Active Management," *Journal of Hydrology*, vol. 509, pp. 215-230, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[14] Jakub Langhammer, "Flood Simulations using a Sensor Network and Support Vector Machine Model," *Water*, vol. 15, no. 11, pp. 1-27, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[15] Mohammad Roohi et al., "A Hybrid Approach for Enhanced Flood Prediction and Assessment: Leveraging Physical Models, Deep Learning and Satellite Remote Sensing," *Big Earth Data*, vol. 9, no. 3, pp. 439-472, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[16] Shuchi Mishra et al., "Dynamic Flood Risk Prediction in Houston: A Multi- Model Machine Learning Approach," *Geocarto International*, vol. 39, no. 1, pp. 1-32, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[17] Yogesh Bhattarai et al., "Leveraging Machine Learning and Open-Source Spatial Datasets to Enhance Flood Susceptibility Mapping in Transboundary River Basin," *International Journal of Digital Earth*, vol. 17, no. 1, pp. 1-25, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[18] Durlav Sonowal et al., "Streamflow Forecast Model using ANN," *Journal of Applied Water Engineering and Research*, vol. 12, no. 3, pp. 311-321, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[19] Juliana G. Rabelo et al., "Karstification and Fluid Flow in Carbonate Units Controlled by Propagation and Linkage of Mesoscale Fractures, Jandaíra Formation, Brazil," *Geomorphology*, vol. 357, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[20] Mohammad Mahdi Rajabi, Behzad Ataie-Ashtiani, and Craig T. Simmons, "Model-Data Interaction in Groundwater Studies: Review of Methods, Applications and Future Directions," *Journal of Hydrology*, vol. 567, pp. 457-477, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[21] Vilde Dimmen et al., "Quantifying Structural Controls on Fluid Flow: Insights from Carbonate- Hosted Fault Damage Zones on the Maltese Islands," *Journal of Structural Geology*, vol. 101, pp. 43-57, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[22] Beibei Shen et al., "Spatial Variations and Controls on the Hydrochemistry of Surface Waters Across the Ili-Balkhash Basin, Arid Central Asia," *Journal of Hydrology*, vol. 600, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[23] Daniel F. Menezes et al., "Subsidence Rings and Fracture Pattern Around Dolines in Carbonate Platforms – Implications for Evolution and Petrophysical Properties of Collapse Structures," *Marine and Petroleum Geology*, vol. 113, pp. 1-17, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[24] Amin Gharehbaghi, "Explicit and Implicit Forms of Differential Quadrature Method for Advection – Diffusion Equation with Variable Coefficients in Semi-Infinite Domain," *Journal of Hydrology*, vol. 541, pp. 935-940, 2016. [CrossRef] [Google Scholar] [Publisher Link]