KSCE Journal of Civil Engineering (2019) 23(2):923-934

DOI 10.1007/s12205-018-0128-1

Water Resources and Hydrologic Engineering

pISSN 1226-7988, eISSN 1976-3808
www.springer.com/12205

Application of Support Vector Regression for Modeling Low Flow Time Series

Bibhuti Bhusan Sahoo*, Ramakar Jha, Anshuman Singh***, and Deepak Kumar******

Received January 24, 2018/Revised May 29, 2018/Accepted September 3, 2018/Published Online December 17, 2018

Abstract

Hydrologic time series modeling using historical records plays a crucial role in forecasting different hydrological processes. The objective of this study is to analyze the suitability of Support Vector Regression (SVR) for modeling monthly low flows time series for three stations in Mahanadi river basin, India. The 'low flow' threshold was taken as the $Q75$ discharge, i.e., the flow is equal to or surpassed for the duration of 75% of the observation period which was obtained from the daily discharge data. The potential applicability of SVR model is assessed with two different framework models (ANN-ELM, GPR) based on various statistical measures (r^2, RMSE, MAE, Nash-Sutcliffe coefficient, objective function (OBJ), Scatter Index (SI) and BIAS). The model selection was based on lowest OBJ value for each station amongst three models (SVR, ANN-ELM, GPR). The SVR model was trained using the Radial Basis Function (RBF). Using the same inputs, the other two models (ANN-ELM and GPR) was also tested. From results, among all the stations, the SVR outperformed GPR and ANN-ELM with lowest OBJ value for the three stations a (1.378, 1.202, 1.570). In addition, the accuracy of the three models were checked using mean forecasting error which were (0.474, 0.421, 0.509) for SVR, (0.507, 0.489 0.500) for GPR and (0.564, 0.603, 0.772) ANN-ELM for the three stations. The results confirm that SVR can be used satisfactorily for modeling monthly low flows in the Mahanadi river basin, India. Hence, the SVR model could be employed as a new reliable and accurate data intelligent approach for predicting the 'low flow' ($Q75$ discharge) based on precedent data in water resources and its allied field.

Keywords: *artificial intelligence, forecasting, hydrologic time series, low flows, support vector regression, predictive modeling*

1. Introduction

Hydrological extremes (flood and droughts) occur frequently now a day across various parts of the world as a result of climate change. It's a very challenging task for the water resources managers to manage the available water resources in any river basin. Therefore, it is highly essential to have the knowledge of various flow conditions such as low flows and high flows in the river. Forecasting of the low flow is essential for a large range of water resources management strategies such as operation of water supply utilities, water resources planning and management, admonitions of pollution level in the water, environment protection and sustainable development of water resources, etc. The identification of appropriate forecasting model to predict the future monthly low flow is an essential requirement for effective water resources planning and management (Haghiabi *et al.*, 2018; Parsaie and Haghiabi, 2017a; Qishlaqi *et al*., 2017).

Precise forecasting of the hydrological process is one of the important issues, which provides reliable and accurate applications in water resources activities (Haghiabi, 2016, 2017; Parsaie *et al*., 2016; Rezaie-Balf and Kisi, 2017). Nowadays, the AI methods have been applied for prediction of flow in streams based on the compound open channel (Parsaie and Haghiabi, 2017c; Parsaie *et al*., 2017b). In operational hydrology, forecasting hydrological time series plays a significant role (Wang *et al.*, 2009). In recent years, AI technology has become more profoundly used tool for hydrological forecasting as for their better decisions making. The major advantage of these AI techniques is that they hardly bother about the complicated nature of the different hydrological phenomena (Atiquzzaman and Kandasamy, 2016). Furthermore, such techniques provide satisfactory forecasts based on historical data (Abbot and Marohasy, 2012; Acharya *et al.*, 2014; Deo *et al*., 2016; Wu and Chau, 2010). These AI techniques engage data-driven models that depend on the building of a mathematical relationship between the output and the input variables (Deo and Şahin, 2015; Tiwari and Adamowski, 2013). Traditionally Autoregressive Moving-Average (ARMA) models developed by (Box and Jenkins, 1970) are used in hydrologic times series forecasting. However, a wide number of different AI models have been used by a number of researchers for the modeling of hydrological time series (Arena *et al.*, 2006; Chen and Rao, 2003; Hipel and McLeod, 1994; Komorník *et al*., 2006; Salas,

*Ph.D. Research Scholar, Dept. of Civil Engineering, National Institute of Technology, Patna 800005, India (Corresponding Author, E-mail: bibhuti5000@gmail.com)
**Professor, Dept. of Civil Engineering, National Institute of Technology, Patna 800005, India (E-mail: rjha43@gmail.com)
***Assistant Professor, Dept. of Civil Engineering, National Institute of Technology, Patna 800005, India (E-mail: asingh@nitp.ac.in)
****Ph.D. Research Scholar, Dept. of Civil Engineering, National Institute of Technology, Patna 800005, India (E-mail: decage007@gmail.com)

1993; Srikanthan and McMahon, 2001; Toth *et al*., 2000).

The AI techniques are considerably able to explore lengthy time series and become gradually more prevalent in operational hydrology across the globe by many researchers (Wang *et al*., 2009). Some of the frequently used AI technique includes Artificial Neural Networks (ANNs) (Alvisi *et al.*, 2006; Campolo *et al.*, 2003; Sudheer and Jain, 2004; Tao *et al*., 2008), the Adaptive Neural-Based Fuzzy Inference System (ANFIS) (Alvisi *et al*., 2006; Nayak *et al.*, 2004), Genetic Programming (GP) (Deo and Samui, 2017; Shiri and Kişi, 2011; Sivapragasam *et al*., 2008), Gaussian Processes Regression (GPR) (Deo and Samui, 2017), Support Vector Machine (SVM) (Deka, 2014; Khan and Coulibaly, 2006; Yu *et al.*, 2006), Relevance Vector Machine (RVM) (Deo and Şahin, 2015; Okkan and Inan, 2014), Extreme Learning Machines (ELM) (Deo and Şahin, 2015, 2016; Rezaie-Balf and Kisi, 2017). These models perform reasonably well in predicting hydrological time series.

Some studies have also showen the development of data-driven models for forecasting flood (Deo and Şahin, 2015; Han *et al.*, 2007; Kumar *et al*., 2015; Seo *et al*., 2015; Yu *et al.*, 2006), drought (Belayneh and Adamowski, 2012; Kim and Valdés, 2003; Nikbakht *et al*., 2012) and various hydrologic and hydaulic phenomena by various researchers across the globe (Najafzadeh *et al.*, 2016; Najafzadeh *et al*., 2018; Najafzadeh and Saberi-Movahed, 2018; Najafzadeh *et al.*, 2017; Parsaie *et al*., 2017a; Parsaie and Haghiabi, 2017b; Parsaie *et al*., 2018b; Parsaie *et al*., 2017b; Zahiri and Najafzadeh, 2018) but to the best of the authors' knowledge, a forecasting model for low flow has not been explored especially in this study basin.

In a developing country like India there is increasing pressure to improve the reliability of the water resources scheme and enrich ecosystems degraded by over abstraction and pollution. Both surface water and groundwater resources are under greatest pressure during the low flow periods and with population growth and climate and land use changes these pressures will increase (Gustard and Demuth, 2009). Many river basins of the country are facing vegetation/crops stress due to low flow situation in rivers. In a country like India, where most of the irrigation is done through canal systems. Therefore, it is essential to manage the available water resources efficiently so that water is available during the lean period.

The main purpose of the study is to examine the suitability of SVR in forecasting monthly low flow time series and a comparative study with other AI technique (ANN-ELM and GPR) to check its robustness. Previously many researchers have successfully applied the SVR model in modeling the hydrology phenomenon (Azamathulla *et al.,* 2016; Haghiabi *et al*., 2017; Najafzadeh *et al*., 2016; Parsaie *et al.*, 2018a). The SVM model was chosen because of its high ability in pattern recognition (Azamathulla *et al*., 2016).

2. Defining Low Flow

An appropriate definition of low flow differs according to the need of the study (Pyrce, 2004). Low flow situation is defined by the determination of a certain percentiles of discharge (Ahn and

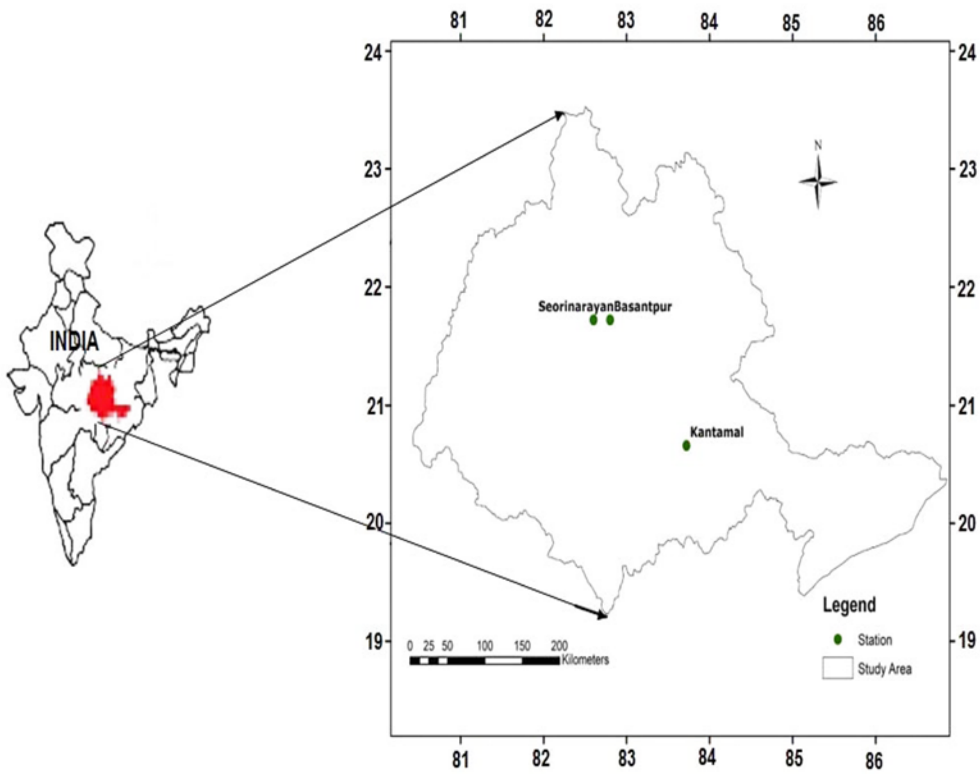

Fig. 1. Location Map of Selected Stations in Mahanadi Basin

Palmer, 2016) or a truncation level. Many past studies have reported a number of low flow indices such $Q95$ (Laaha and Blöschl, 2005), $Q85$ (Giuntoli *et al*., 2013), $Q75$ (Demirel *et al*., 2013; Jha and Smakhtin, 2008; Pyrce, 2004), used for low flow study, where $Q95$, $Q85$, $Q75$ are the discharge equaled or surpassed for the duration of 95%,85%, and 75% of the observation period respectively. In this study, 'low flow' threshold is taken as the $Q75$ discharge, i.e., the flow is equal to or surpassed for the duration of 75% of the observation period which was obtained from the daily discharge data. It is to be noted that the hydrological, topographical and climatic conditions of rivers of India, in general, are quite different and the approach suggested by $Q75$ (Jha *et al*., 2008; Pyrce, 2004) may be applicable for Mahanadi river based on actual field conditions. The low flow assessment plays a crucial role in low flow managing (Dracup *et al*., 1980) along with many environmental purposes related to water resources management.

3. Study Area and Data Collection

The Mahanadi basin (Fig. 1) situated between 80°28' and 86°43' E and 19°8' and 23°32' N. It travels a distance of 851 km from the origin before falling into the Bay of Bengal. The daily discharge data were collected from Central Water Commission (CWC) Bhubaneswar for the three stations namely Basantapur, Kantamal and Seorinarayan used for the study. (Fig. 2) demonstrates the time series plot of daily discharge data for the three stations.

4. Methodology

4.1 Gaussian Process Regression (GPR)

Gaussian process regression is a non-parametric method which defines a prior over functions, which can be converted into a posterior over functions for prediction problems (Williams and Rasmussen, 1996). The advantage of the Gaussian process framework is that without prior knowledge about the data and functional dependencies noise models can be formulated using matrix operations. In the case of maximum likelihood or Bayesian approach, the hyper-parameters which control the form of the Gaussian process can be estimated from the data and which ultimately leads to a formation of "Automatic Relevance Determination" (MacKay, 1996).

In GPR, the main assumption is that for every independent variable $u(k)$, there exists a dependent variable $v(k)$, which provides the value of the function f at that location such as;

$$v(k) = f[u(k)] + \zeta(k) \tag{1}$$

Where

v(k)= Objective variable

ζ(k)~N(0, σ^2)= Gaussian noise with variance

σ^2 and u(k)= Regression in space R^D for every input vector u, there is an associated random variable f(u).

The error ζ is normally and identically distributed with zero mean and variance σ^2 and $f(u)$ is drawn by a Gaussian process, u specified by kernel K such that

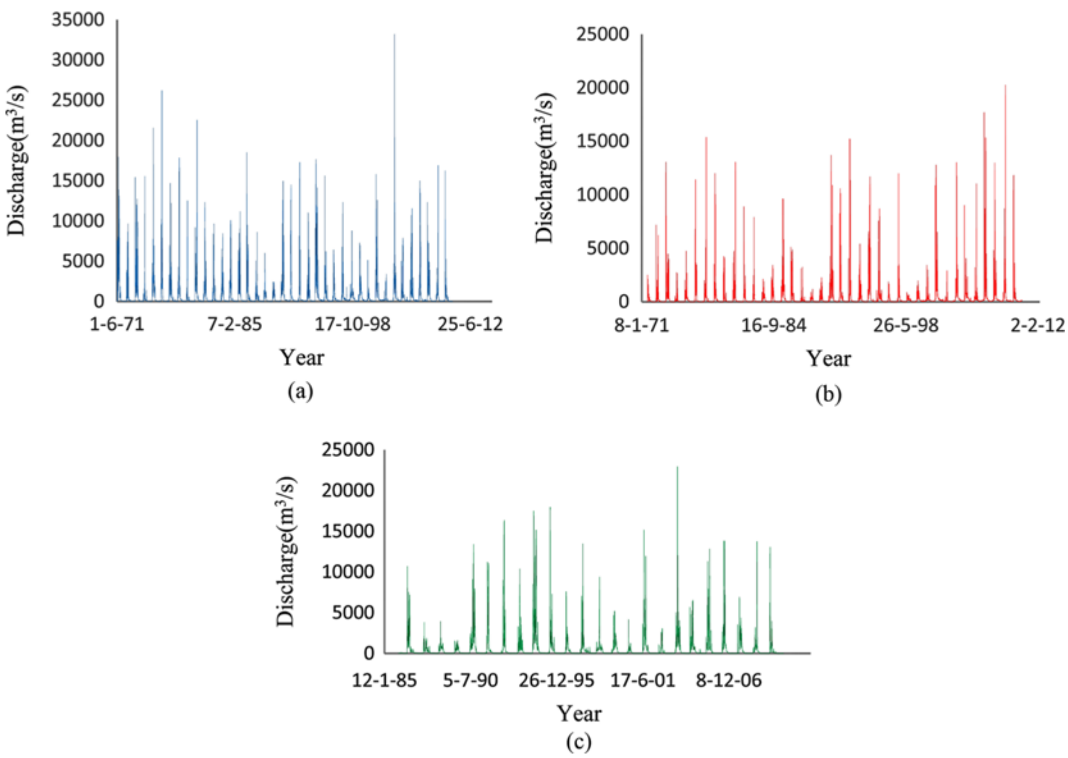

Fig. 2. Daily Discharge Time Series of the Three Stations: (a) Basantapur, (b) Kantamal, (c) Seorinarayan

$$y(u) = (u_1, u_2, \cdots\cdots u_n) \sim N(0, K + \sigma^2 I) \tag{2}$$

Where $K_{i,j} = C(u_i, u_j)$ is the covariance matrix and I is the identity matrix.

In training of GPR model, one needs to select an appropriate covariance function along with its parameters. After fixing the value of Gaussian noise, a GPR model can be trained using Bayesian inference based on maximizing the marginal likelihood (Pal and Deswal, 2010).

$$p(\sigma^2, k) = \frac{1}{2} y^T (K + \sigma^2 I)^{-1} y + \frac{1}{2} \log|K + \sigma^2 I| - \log p(\sigma^2) - \log p(k) \tag{3}$$

In training model, to get the best-fitted hyper-parameters, the partial derivative of Eq. (3) cosidering σ^2 and k has to be performed with optimal minimization gradient descent.

4.2 Artificial Neural Network -Extreme Learning Machines (ANN-ELM)

ANN-ELM is a fast training AI technique used for prediction, using Single Layer Feedforward Neural Network (SLFN) to establish a relationship between complex nonlinear dependent and independent variable (Tokar and Johnson, 1999). The advantage of this technique that it does not need any information about the complexity of the hydrological process. Among the other generalized feedforward neural network, Extreme Learning Machine (ELM) is an efficient technique for its fast learning computational time than other conventional gradiant methods. The ability of randomly learning input weights and choice of non-linear activation function made this technique more popular in the research community. In this study, ANN-ELM hidden node of the layer is independent of hidden layer while training input. This means hidden nodes were independent of input training set (Lei *et al*., 2015). The extremely fast execution algorithm of ANN-ELM distinguishes it from other AI techniques, for example, ANN or SVM (Rajesh and Prakash, 2011).

For N arbitrary distinct input samples (u_i, t_i), where $u_i = [u_{i1}, u_{i2}, \cdots\cdots u_{in}]^T \in R^n$ and $t_i = [t_{i1}, t_{i2}, \cdots\cdots u_{im}]^T \in R^m$. The mathematical equation for SLFNs with N hidden neuron can be formularized as (Huang *et al*., 2006).

$$\sum_{i=1}^{\hat{N}} \beta_i g_i(u_j) = \sum_{i=1}^{\hat{N}} \beta_i g(w_i \cdot u_j + b_i) = o_j \quad j = 1, \ldots, N \tag{4}$$

where, $w_i = [w_{il}, w_{i2}, \cdots\cdots\cdots, w_{in}]^T$ is the weight vector connecting the i^{th} hidden node and the input nodes, $\beta = [\beta_{i1}, \beta_{i2}, \cdots\cdots \beta_{in}]$ is the weight vector connecting the i^{th} hidden node and the output nodes and b_i is the threshold of the i^{th} hidden node. $w_i \cdot u_i$ denotes the inner product of w_i and u_j. The standard SLFNs with N hidden nodes with activation function $g(u)$ can be approximate these N samples with zero error which means $\sum_{j=1}^{\tilde{N}} \|O_j - T_j\| = 0$. There exist b_i, w_i. The term b_i of equation (4) is estimated as such:

$$\sum_{i=1}^{\tilde{N}} \beta_i g(w_i \cdot u_j + b_i) = t_j, \; j = 1, \ldots\ldots, N. \tag{5}$$

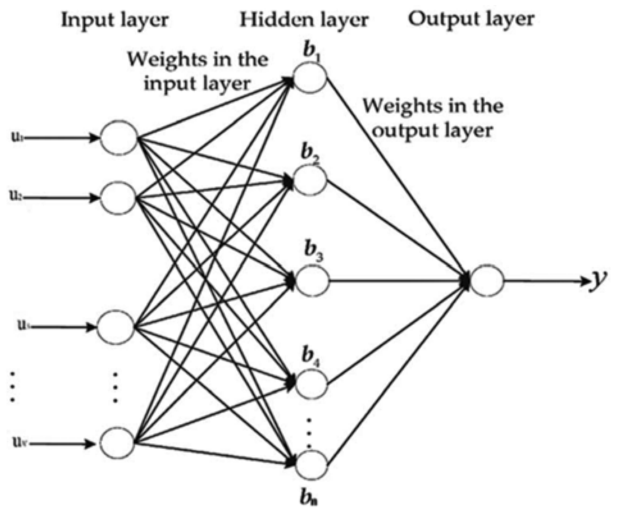

Fig. 3. Structure of Artificial Neural Network-Extreme Learning Machine

The Eq. (5) can also be written as,

$$H\beta = T \tag{6}$$

Where,

$$H = (w_1 \cdots w_n,\ b_1 \cdots b_n,\ u_1 \cdots\cdots u_n) = \begin{bmatrix} g(w_1.u_1 + b_1) & \cdots & g(w_{\tilde{n}}.u_1 + b_{\tilde{N}}) \\ \vdots & \cdots & \vdots \\ g(w_1.u_N + b_1) & \cdots & g(w_{\tilde{N}}..u_N + b_{\tilde{N}}) \end{bmatrix}_{N \times \tilde{N}} \tag{7}$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_{\tilde{N}}^T \end{bmatrix}_{\tilde{N} \times M} \text{ and } T = \begin{bmatrix} t_1^T \\ \vdots \\ t_N^T \end{bmatrix}_{N \times M} \tag{8}$$

Where H_0 is called Moore–Penrose generalized inverse of matrix H. SLFNs can approximate the training samples with zero error when there is an equal number of both the hidden neurons and training sample. H_0 may be calculated using several methods including orthogonal projection method, orthogonalization method, iterative method, Singular Value Decomposition (SVD), etc. The SVD method is used to calculate H_0, which was shown that SLFNs with randomly generated hidden nodes and with a widespread piecewise continuous activation function that could universally approximate any continuous target function. Furthermore, details regarding ANN-ELM can be found in (Huang *et al.*, 2006). An illustrative architecture of ANN-Extreme Learning Machine is shown in (Fig. 3).

4.3 Support Vector Regression (SVR)

SVR a robust and efficient algorithm developed by Vapnik (Vapnik, 1998) based on Statistical learning theory. It became more popular due to its successful application in classification (Burges, 1998; Osuna *et al.*, 1997) and regression tasks (Burges, 1998) to get minimum regression error, especially in case time series forecast (Müller *et al.*, 1997), therefore for practitioners mainly in water resource using SVR for forecasting time series usually overlook the choices of the margin setting which varying cost function to achieve lesser RMSE error. The effectiveness the nonlinear SVR depend upon soft margin constant parameter C, ε-insensitive loss function and gamma. As these are highly interdependent, therefore altering the one parameter affects the other. The cost parameter inspects the smoothness of the approximation function keeping fixed kernel parameters. A greater cost value yields lower bias and higher variance because of penalizing the cost of misclassification a lot and vice versa. Whereas, kernel parameter ε–tube relates smoothening the complexity of fitting the training data. Finding proper values of C and ε is often a tradeoff process. Therefore, this method is non-adaptive as well as insensitive to the input data. Hence in most of the cases, its result is less in the case of testing data whereas optimal performance on the training data. In order to train SVR model, the input predictors consist of a set of training data $(u_1, v_1), \cdots\cdots, (u_n, v_n)$, $u \in R^n$ and $v \in R$. where u, v represents the predictors variable and the output value respectively and R^n represents vector space dimensionality and R denotes the one-dimensional vector space where n represent the total number of predictors for model. In this SVR model five input variables $Q75_{t-1}$, Q_{75t-2}, Q_{75t-3}, $Q75_{t-4}$ and $Q75_{t-5}$ were taken whereas $Q75_t$ as output. The ε-insensitive loss function can be defined as follows.

$$L(v) = 0 \ \text{ for } |f(u) - v| < \varepsilon \text{ otherwise } L(v) = |f(u) - v| - \varepsilon \tag{9}$$

Equation (9) defines a tube which is represented by ε in (Fig. 4). The forecasted value has no loss when all the forecasted value within the tube (ε) otherwise forecasted loss is estimated by modulus of their deviation (forecast value-actual value) minus epsilon (ε).

The main objective during training SVR to get an optimal function that provides a minimum absolute deviation from ε with minimum flatness. While training the SVR, the kernel trick facilitates to project all inputs data (u_i) to high dimensional space through non-linear function $f(u)$. The high dimensional projected data is estimated based on Eq. (10)

$$f(w,b) = (w.\phi(u)) + b \tag{10}$$

Where
b = Constant
w = Vector coefficient

This nonlinear regression problem can be expressed as in (Fig. 4) that shows the generalized concept of SVR corresponding to Eq. (15)

$$\min_{w,b,\xi,\xi^*} \quad \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}\left(\xi_i + \xi_i^*\right) \tag{11}$$

$$v_i - (w \cdot \phi(u_i + b)) \le \varepsilon + \xi_i \tag{12}$$

Subjected to

$$(w \cdot \phi(u_i + b)) - v_i \le \varepsilon + \xi_i^* \tag{13}$$

$$\xi_i, \xi_i^* \ge 0 \quad i = 1,2,3,........n \tag{14}$$

where, ξ_i and ξ_i^* are slack variables introduced to evaluate the deviation of training samples outside the ε-insensitive zone, hence the modulus distance from the training data, where the errors less than ε are ignored and index i labels the n training cases with u_i is the independent variable.

Hence, the dual form of nonlinear SVR can be formulated as using the kernel trick is expressed as follows:

$$f(u) = \sum_{i,j=1}^{n}\left(\alpha_i - \alpha_i^*\right)\left(\phi(u_i) \cdot \phi(v_j)\right) + b \tag{15}$$

Where α_i, α_i^* are Lagrange multipliers variables constraints which lead to the construction of the dual optimization problem.

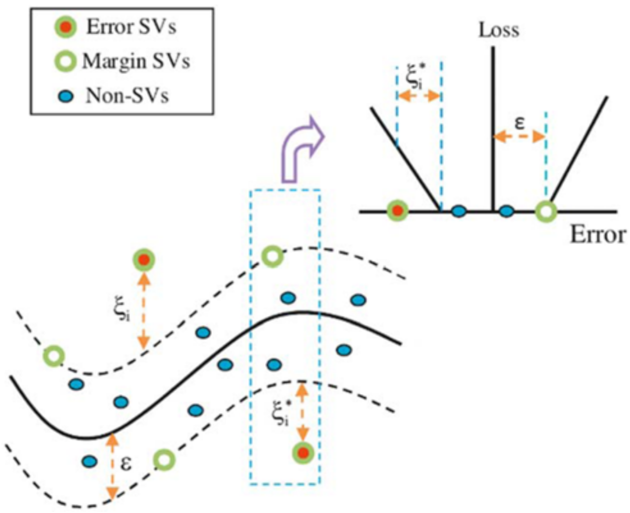

Fig. 4. Nonlinear ε-insensitive Loss Function. Adapted from (Deka, 2014)

5. Input Variable Selection and Model Setup

In hydrological time series model, the predicted value is computed from the predictors based on historical data, which generally consists of different time lags of inputs and their combinations (Nayak *et al*., 2004; Wang *et al*., 2009). For developing these AI models based on inputs, there were no existing universally accepted guidelines (Sudheer *et al*., 2002; Wang *et al*., 2009). Preciously many researchers have adopted different time lags of input and its combination as predictors to develop the forecasting model (Cheng *et al.*, 2005; Nayak *et al*., 2004; Sudheer *et al*., 2002; Wang *et al*., 2009). The focus of the study was to predict $Q75$ discharges, i.e., the monthly low flow using different time lags values to build up a model of the following form:

$$A^m = f\left(B^m\right) \tag{16}$$

Where B^m is an m-dimensional input vector consisting of variables $b_1 \cdots, b_i \cdots b_m$, and A^m is the output variable, consisting of the subsequent variables of interest $a_1 \ldots, a_i \ldots a_m$.

Table 1. Five SVR Models Used in the Course of the Analysis with the Corresponding Predictor Variable as Input

M1	$Q75_t = f(Q75_{t-1})$
M2	$Q75_t = f(Q75_{t-1}, Q75_{t-2})$
M3	$Q75_t = f(Q75_{t-1}, Q75_{t-2}, Q75_{t-3})$
M4	$Q75_t = f(Q75_{t-1}, Q75_{t-2}, Q75_{t-3}, Q75_{t-4})$
M5	$Q75_t = f(Q75_{t-1}, Q75_{t-2}, Q75_{t-3}, Q75_{t-4}, Q75_{t-5})$

The analysis was carried out with one time lag of the low flow in the input vector, and the SVR model was built. The input vector was then reformed by consecutively adding low flow at one more time lag and so on, and a new SVR model was established every time respectively. Before the development of the models, the data was transformed using Eq. (18). Moreover, in the training period, 70% randomly sampled data were used and the rest 30% data was used as testing the model. Five SVR models (M1, M2, M3, M4, and M5) were built for the course of the analysis with the corresponding input vectors as listed the Table 1.

where $Q75_t$ denotes the low flow at time t.

The autoregressive processes were studied using, the Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF), which can also be engaged in prediction modeling (Lin *et al*., 2006). The values of ACF and PACF of monthly low flow sequence is calculated for lag 0 to 12 for all the three stations, which is presented in (Fig. 5), respectively. Also, a significant correlation of PACF at 95% confidence level interval up to 12 months of flow lag were observed. This showed that the twelve antecedent low flow values have the most information to predict future flow and can be considered as input for developing predictive models.

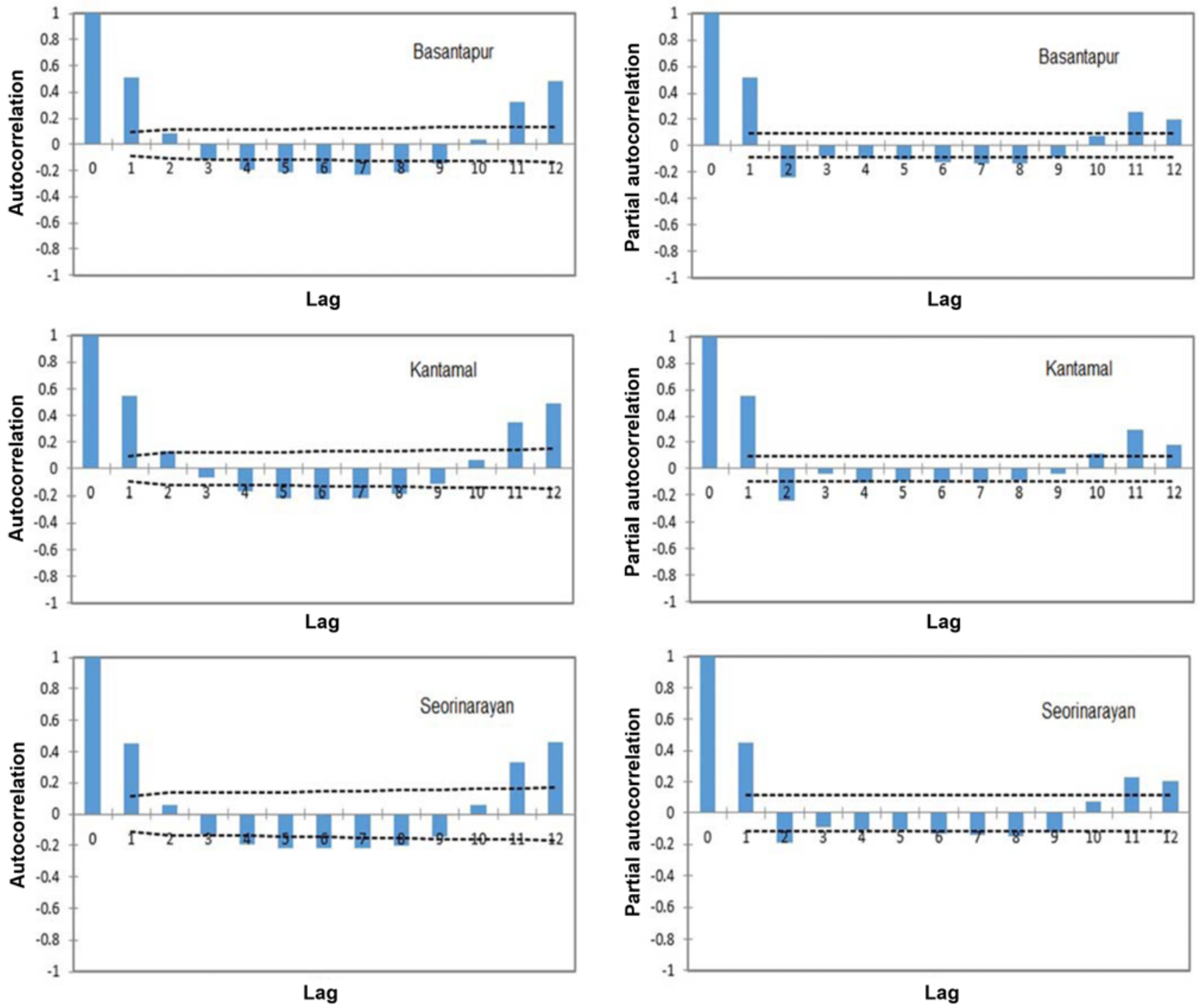

Fig. 5. Autocorrelation Function Partial Autocorrelation Function of Low Flow Time Series for the Three Stations

Table 2. SVR training and Testing Values of Coefficient of Correlation (r) for M1 to M5

Model	Basantapur		Kantamal		Seorinarayan	
	Training r	Testing r	Training r	Testing r	Training r	Testing r
M1	0.740	0.736	0.793	0.731	0.802	0.597
M2	0.870	0.847	0.878	0.786	0.865	0.806
M3	0.894	0.830	0.905	0.845	0.826	0.901
M4	0.910	0.829	0.92	0.872	0.884	0.803
M5	0.932	0.851	0.926	0.882	0.892	0.858

Further, the coefficient of correlation (r) was determined using the Eq. (17), during the training and testing to select the best model. The r-value of training and testing period is given in Table 2 for the respective stations. The, M5 is chosen as the best model as compared to other counterparts in all the stations.

$$r = \frac{\sum_{i=1}^{N}\left(Q75_{obs} - \overline{Q75}_{obs}\right)\left(Q75_{for} - \overline{Q75}_{for}\right)}{\sqrt{\sum_{i=1}^{N}\left(Q75_{obs} - \overline{Q75}_{obs}\right)^2}\sqrt{\sum_{i=1}^{N}\left(Q75_{for} - \overline{Q75}_{for}\right)^2}} \quad (17)$$

6. Results and Discussion

6.1 Development of Forecasting Models

Different AI algorithms (SVR, GPR, and ANN-ELM) were used for forecasting the $Q75$ discharge using the different package of the R software open source program. The objective of this study was to inspect and compare the viability of the SVR, GPR, and ANN-ELM modeling methodologies for forecasting monthly Q75 discharge in Mahanadi river basin for the selected stations. In data-driven predictive modeling, a major obstacle to determining a good subset of data to build an appropriate predictive model and to choose an appropriate testing subset to evaluate the forecasting model using available data. In past, researchers have used different data portion based on their problem of interest. Thus, there is no well-defined procedure or thumb rule for selecting the appropriate subset for training and testing which varies from the problem of interest.

In this study, randomly sampled 70% of data was used in training period and remained 30% of the dataset was used for testing the developed SVR, GPR, ANN-ELM models. Before that, the data set was transformed using the formula.

$$d' = \ln(d + 10) \quad (18)$$

To develop the SVR, GPR, ANN-ELM model for the prediction of $Q75$, we have used the Radial Basis Function (RBF) Eq. (19) for both SVR, GPR as a kernel function while the ANN- ELM model used sigmoid function as its activation function for attaining higher accuracy.

$$K(x, x') = \exp\left(-\frac{(x_i - x)(x_i - x)^T}{2\sigma^2}\right) \quad (19)$$

6.2 Optimum Model Selection

In previous studies, the researchers had adopted trial and error procedure for the optimization of machine learning models (Khan and Coulibaly, 2006; Sivapragasam and Muttil, 2005). In reality, the existing SVR model accuracy strictly depends upon the choice of kernel, regularization parameter C, and ε-insensitive loss function (Khan and Coulibaly, 2006). This study provides an equation for predicting $Q75$ based on the developed SVM model. The developed SVR gives the following equation by putting

$$K(u, u') = \exp\left(-\frac{(u_i - u)(u_i - u)^T}{2\sigma^2}\right) \quad (20)$$

and $b = 0$ in Eq. (7) for prediction of $Q75$

$$Q75_{for}(u) = \sum_{i=1}^{n}\left(\alpha_i - \alpha_i^*\right)\cdot\exp\left(-\frac{(u_i - u)(u_i - u)^T}{0.2}\right) \quad (21)$$

Several-layer network was used to build the ANN-ELM framework for the forecasting the $Q75$ on trial and error based on different randomly selected training (70%) and test (30%) datasets. Five input neurons were selected based on predictor variable $x = [Q75_{t-1}, Q75_{t-2}, Q75_{t-3}, Q75_{t-4}, Q75_{t-5}]$, where one neuron represents the simulated $Q75$, but in the case of hidden layer neuron, maximum 20 neurons were trailed randomly with five subsequent increments to achieve the optimum case with sigmoid as the activation function. Fig. 4 shows the complete architecture of overall ANN-ELM.

The predictableness of SVR model for forecasting monthly $Q75$ for three stations namely Basantapur, Kantamal, Seorinarayan of the Mahanadi river basin, India were examined where the models used the five lag of $Q75$ time series data.

6.3 Performance Assessment of the Developed Models

The performance of the SVR, GPR and ANN-ELM models was assessed by the comparison of the observed $Q75$($Q75_{obs}$) and the forecasted $Q75$($Q75_{for}$) in the training and testing period using

1. Nash-Sutcliffe coefficient (E_{NS}):

$$E_{NS} = 1 - \left[\frac{\sum_{i=1}^{N}\left(Q75_{obs} - Q75_{for}\right)^2}{\sqrt{\sum_{i=1}^{N}\left(Q75_{obs} - \overline{Q75}_{for}\right)^2}}\right], \quad -\infty \le E_{NS} \le 1 \quad (22)$$

2. Coefficient of determination (r^2)

$$r^2 = \left[\frac{\sum_{i=1}^{N}\left(Q75_{obs} - \overline{Q75}_{obs}\right)\left(Q75_{for} - \overline{Q75}_{for}\right)}{\sqrt{\sum_{i=1}^{N}\left(Q75_{obs} - \overline{Q75}_{obs}\right)^2}\sqrt{\sum_{i=1}^{N}\left(Q75_{for} - \overline{Q75}_{for}\right)^2}}\right]^2 \quad (23)$$

3. Root-Mean-Square Error (RMSE):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}\left(Q75_{obs} - Q75_{for}\right)^2}{N}} \quad (24)$$

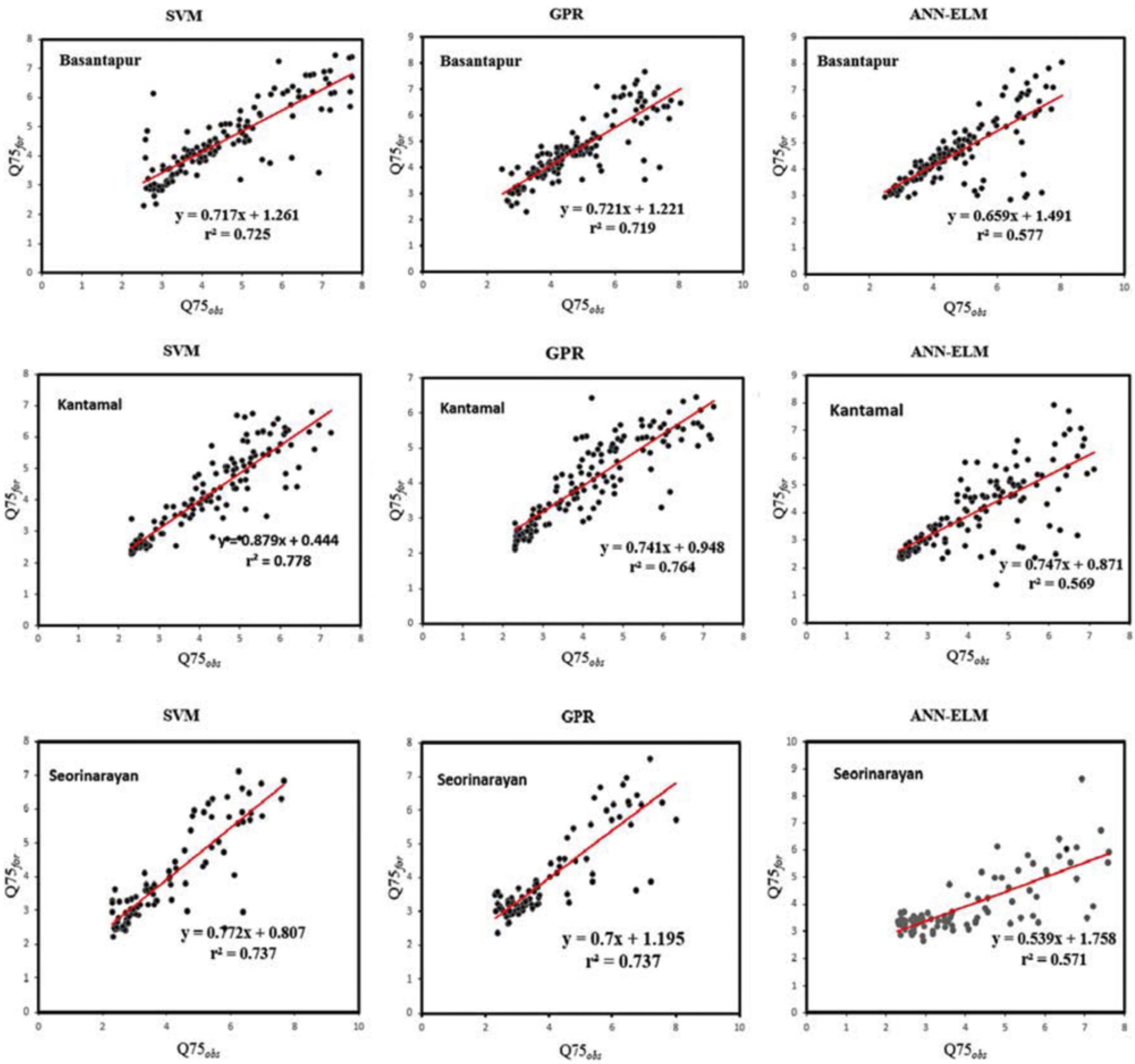

Fig. 6. Scatterplot of $Q75_{for}$ Versus $Q75_{obs}$ During the Testing Period for: (a) Basantapur, (b) Kantamal, (c) Seorinarayan for SVR Model Compared with GPR, and ANN-ELM Models

4. Mean Absolute Error (MAE)

$$MAE = \frac{\sum_{i=1}^{N} \left| Q75_{obs} - Q75_{for} \right|}{N} \tag{25}$$

Where

N = Number of data points (70% for training and 30% for testing of the data)

$Q75_{obs}$ = Observed $Q75$

$Q75_{for}$ = Forecasted $Q75$

$\overline{Q75_{obs}}$ = Average observed

$\overline{Q75_{for}}$ = Average forecasted Q_{75}

Figure 6 shows a scatterplot of forecasted ($Q75_{for}$) versus the observed value ($Q75_{obs}$) for the data analyzed in the testing period from SVR, GPR and ANN-ELM models along with linear regression equation of the form ($Q75_{for}$) = m($Q75_{obs}$) + C where the accuracy of ($Q75_{obs}$) judge in accordance with the gradient (m) and intercept Y of the regression. Perfect match achieved when ($Q75_{for}$) and ($Q75_{obs}$) data. For ideal, m = 1 and C = 0 should be sought whereas r^2 value should be close to 1.

Based on r^2, SVR attained the highest coefficient of determination ($r^2 \approx 0.725$) followed by relatively lower value of 0.719 for GPR and 0.577 for ANN-ELM for the station Basantapur. However, for the station Kantamal, SVR attained the highest coefficient of determination ($r^2 \approx 0.778$) followed by GPR ($r^2 \approx 0.764$) and ANN- ELM ($r^2 \approx 0.569$) during the testing period. In case of

Table 3. Model Performance Evaluation in the Training and Testing Period

Forecasting model	r^2		E_{NS}		*RNSE*		*MAE*	
Basantapur								
	Training	Testing	Training	Testing	Training	Testing	Training	Testing
SVR	0.869	0.725	0.868	0.724	0.498	0.764	0.269	0.474
GPR	0.851	0.719	0.851	0.714	0.535	0.760	0.358	0.507
ANN-ELM	0.598	0.577	0.577	0.557	0.902	0.946	0.573	0.564
Kantamal								
SVR	0.857	0.778	0.856	0.763	0.507	0.650	0.291	0.421
GPR	0.850	0.764	0.850	0.753	0.504	0.700	0.334	0.489
ANN-ELM	0.686	0.569	0.776	0.971	0.776	0.971	0.513	0.603
Seorinarayan								
SVR	0.795	0.737	0.789	0.728	0.700	0.789	0.317	0.509
GPR	0.738	0.737	0.738	0.735	0.779	0.778	0.505	0.500
ANN-ELM	0.667	0.571	0.666	0.561	0.872	1.022	0.637	0.772

station Seorinarayan, the SVR and GPR both attained a coefficient of determination ($r^2 \approx 0.737$) and ANN-ELM ($r^2 \approx 0.571$). For the station Seorinarayan both SVR and GPR has performed well while in case of Basantapur and Kantamal station SVR outperformed as compared to GPR and ANN-ELM. Other statistical metrics included Nash-Sutcliffe coefficient (E_{NS}), Root-Mean-Square Error (RMSE), Mean Absolute Error (MAE) along with the Coefficient of determination (r^2) for the training and testing period has been listed in Table 3.

To check the accuracy of the SVR model over monthly forecasting horizons, the magnitude of the model's forecasting error $|FE|$ is analyzed during the testing period. Where $|FE| = |Q75_{for} - Q75_{obs}|$. The modulus symbol equals the error magnitude. Also, the mean forecasting errors and its standard deviation (σ) encountered by the comparative GPR, and the ANN-ELM models were shown in Table 4 for all the stations. According to this result, it was observed that the SVR model predicts more accurate overall tested months as compared with the other two counterparts for the station Basantapur and Kantamal whereas in case of the station Seorinarayan GPR predicts slightly better result than SVR.

Further some other statistical measures such as BIAS, objective function (OBJ) (Alavi and Gandomi, 2011; Najafzadeh and Azamathulla, 2013) and Scatter Index (SI) (Najafzadeh *et al.,* 2017) were calculated to check the performance of the developed models (Table 5). Finally, in this study the OBJ is selected to choose the more efficient model.

Table 4. Summary of the Error Statistics Regarding the Mean Forecasting Error, Standard Deviation (σ)

		SVR	GPR	ANN-ELM
Basantapur	Mean forecasting error	0.474	0.507	0.564
	Standard deviation(σ)	0.599	0.566	0.759
Kantamal	Mean forecasting error	0.421	0.489	0.603
	Standard deviation(σ)	0.496	0.501	0.761
Seorinarayan	Mean forecasting error	0.509	0.500	0.772
	Standard deviation(σ)	0.602	0.596	0.669

$$\mathrm{OBJ} = \left(\frac{\mathrm{No}_{\cdot\mathrm{Train}} - \mathrm{No}_{\cdot\mathrm{Test}}}{\mathrm{No}_{\cdot\mathrm{All}}}\right) \times \frac{\mathrm{RMSE}_{\mathrm{Train}} + \mathrm{MAE}_{\mathrm{Train}}}{\mathrm{r}^2_{\mathrm{Train}}} + 2\frac{\mathrm{No}_{\cdot\mathrm{Test}}}{\mathrm{No}_{\cdot\mathrm{All}}} \times \frac{\mathrm{RMSE}_{\mathrm{Test}} + \mathrm{MAE}_{\mathrm{Test}}}{\mathrm{r}^2_{\mathrm{Test}}} \tag{26}$$

$$BIAS = \frac{\sum_{i=1}^{N} Q75_{for(i)} - Q75_{obs(i)}}{N} \tag{27}$$

$$SI = \frac{RMSE}{\frac{1}{N}\sum_{i=1}^{N} Q75_{obs(i)}} \tag{28}$$

From the Table 5 it is observed that a minimum OBJ value (1.378) for the SVR model with SI = 0.168 and BIAS = -0.036 is obtained for the station Basantapur as compared to GPR (OBJ = 1.477) and ANN-ELM (OBJ = 2.557). A similar result followed in the case of Kantamal and Seorinarayan with SVR having lowest OBJ value as compared to GPR and ANN-ELM model. In fact, proposed SVR model for can be used satisfactorily for modeling

Table 5. Model Performance Evaluation using BIAS, SI and OBJ

Forecasting model	*BIAS*	*SI*	*OBJ*
Basantapur			
SVR	-0.036	0.168	1.378
GPR	-0.101	0.164	1.477
ANN-ELM	-0.129	0.205	2.557
Kantamal			
SVR	-0.061	0.158	1.202
GPR	-0.143	0.172	1.332
ANN-ELM	-0.183	0.244	2.418
Seorinarayan			
SVR	-0.126	0.199	1.570
GPR	-0.016	0.194	1.736
ANN-ELM	-0.134	0.258	2.794

low flow in the selected sites of the Mahanadi River basin.

7. Conclusions

In this study, the suitability of Support Vector Regression (SVR) techniques for modeling low flow is explored by building an SVR model for the selected stations for low flow analysis of Mahanadi river basin using the historical discharge data. A suitable GPR and ANN-ELM model is also tested for the same stations to compare the performance of the SVR, GPR and ANN-ELM models. It was witnessed that the SVR model perform well as compared to GPR and ANN-ELM models in all stations. The developed models were fine-tuned for better performance. To construct an optimum SVR, GPR and ANN-ELM model, the predictor data set is sampled randomly into 70% (training) and 30% (testing) subsets with five-fold cross-validation both in training as well as the testing phase. The results confirm that SVR can be used for satisfactorily among three models to forecast the monthly low flow data with a good level of accuracy. Therefore, it was evidenced by the relatively low forecasting errors that were registered with high correlation and the low OBJ in the case of SVR model. Hence, the outcomes of this study are greatly promising and advocate the applicability of SVR methodology for forecasting monthly low flow time series, and this possibly delivers valued reference for hydrologists and water resources professionals who rely on AI methods for modeling different hydrological time series forecasting and their better decisions making.

Notations

ANN-ELM = Artificial Neural Network -Extreme Learning Machines
AI = Artificial Intelligence
E_{NS} = Nash-Sutcliffe coefficient
GPR = Gaussian Processes Regression
MAE = Mean Absolute Error
$No_{.Train}$ = Number of training data
$No_{.Test}$ = Number of testing data
$No_{.All}$ = Total number of data
OBJ = Objective function
$Q75_{for}$ = $Q75$ forecasted
$Q75_{obs}$ = $Q75$ observed
r^2 = Coefficient of determination
RMSE = Root Mean Squared Error
Scatter Index = SI
SVR = Support Vector Regression
σ = Standard deviation

References

Abbot, J. and Marohasy, J. (2012). "Application of artificial neural networks to rainfall forecasting in Queensland, Australia." *Advances in Atmospheric Sciences*, Vol. 29, No. 4, pp. 717-730, DOI: 10.1007/s00376-012-1259-9.

Acharya, N., Shrivastava, N. A., Panigrahi, B., and Mohanty, U. (2014). "Development of an artificial neural network based multi-model ensemble to estimate the northeast monsoon rainfall over south peninsular India: An application of extreme learning machine." *Climate dynamics*,Vol. 43, Nos. 5-6, pp. 1303-1310, DOI: 10.1007/s00382-013-1942-2.

Ahn, K. H. and Palmer, R. N. (2016). "Use of a nonstationary copula to predict future bivariate low flow frequency in the connecticut river basin." *Hydrological Processes*, Vol. 30, No. 19, pp. 3518-3532, DOI: 10.1002/hyp.10876.

Alavi, A. H. and Gandomi, A. H. (2011). "Prediction of principal ground-motion parameters using a hybrid method coupling artificial neural networks and simulated annealing." *Computers & Structures*, Vol. 89, Nos. 23-24, pp. 2176-2194, DOI: 10.1016/j.compstruc.2011.08.019.

Alvisi, S., Mascellani, G., Franchini, M., and Bardossy, A. (2006). "Water level forecasting through fuzzy logic and artificial neural network approaches." *Hydrology and Earth System Sciences Discussions*, Vol. 10, No. 1, pp. 1-17, DOI: 10.5194/hess-10-1-2006.

Arena, C., Cannarozzo, M., and Mazzola, M. R. (2006). "Multi-year drought frequency analysis at multiple sites by operational hydrology– A comparison of methods." *Physics and Chemistry of the Earth, Parts A/B/C*, Vol. 31, No. 18, pp. 1146-1163, DOI: 10.1016/j.pce.2006.03.021.

Atiquzzaman, M. and Kandasamy, J. (2016). "Prediction of hydrological time-series using extreme learning machine." *Journal of Hydroinformatics*, Vol. 18, No. 2, pp. 345-353, DOI: 10.2166/hydro.2015.020.

Azamathulla, H. M., Haghiabi, A. H., and Parsaie, A. (2016). "Prediction of side weir discharge coefficient by support vector machine technique." *Water Science and Technology: Water Supply*, Vol. 16, No. 4, pp. 1002-1016, DOI: 10.2166/ws.2016.014.

Belayneh, A. and Adamowski, J. (2012). "Standard precipitation index drought forecasting using neural networks, wavelet neural networks, and support vector regression." *Applied Computational Intelligence and Soft Computing*, Vol. 2012, No. 6, DOI: 10.1155/2012/794061.

Box, G. and Jenkins, G. (1970). *Time series analysis; Forecasting and control*, Holden-Day, San Francisco(CA).

Burges, C. J. (1998). "A tutorial on support vector machines for pattern recognition." *Data mining and Knowledge Discovery*, Vol. 2, No. 2, pp. 121-167, DOI: 10.1023/A:1009715923555.

Campolo, M., Soldati, A., and Andreussi, P. (2003). "Artificial neural network approach to flood forecasting in the River Arno." *Hydrological Sciences Journal*, Vol. 48, No. 3, pp. 381-398, DOI: 10.1623/hysj.48.3.381.45286.

Chen, H.-L. and Rao, A. R. (2003). "Linearity analysis on stationary segments of hydrologic time series." *Journal of Hydrology*, Vol. 277, Nos. 1-2, pp. 89-99, DOI: 10.1016/S0022-1694(03)00086-6.

Cheng, C.-T., Lin, J.-Y., Sun, Y.-G., and Chau, K. (2005). "Long-term prediction of discharges in Manwan Hydropower using adaptive-network-based fuzzy inference systems models." *Advances in Natural Computation*, Vol. 3612, pp. 1152-1161, DOI: 10.1007/11539902_145.

Deka, P. C. (2014). "Support vector machine applications in the field of hydrology: A review." *Applied Soft Computing*, Vol. 19, pp. 372-386, DOI: 10.1016/j.asoc.2014.02.002.

Demirel, M. C., Booij, M. J., and Hoekstra, A. Y. (2013). "Identification of appropriate lags and temporal resolutions for low flow indicators in the River Rhine to forecast low flows with different lead times." *Hydrological Processes*, Vol. 27, No. 19, pp. 2742-2758, DOI: 10.1002/hyp.9402.

Deo, R. C. and Şahin, M. (2015). "Application of the artificial neural network model for prediction of monthly standardized precipitation and evapotranspiration index using hydrometeorological parameters and climate indices in eastern Australia." *Atmospheric Research*, Vol. 161, pp. 65-81, DOI: 10.1016/j.atmosres.2015.03.018.

Deo, R. C. and Şahin, M. (2016). "An extreme learning machine model for the simulation of monthly mean streamflow water level in eastern Queensland." *Environmental Monitoring and Assessment*, Vol. 188, No. 2, pp. 1-24, DOI: 10.1007/s10661-016-5094-9.

Deo, R. C. and Samui, P. (2017). "Forecasting evaporative loss by least-square support-vector regression and evaluation with genetic programming, Gaussian process, and minimax probability machine regression: Case study of Brisbane City." *Journal of Hydrologic Engineering*, Vol. 22, No. 6, pp. 05017003, DOI: 10.1061/(ASCE)HE.1943-5584.0001506.

Deo, R. C., Samui, P., and Kim, D. (2016). "Estimation of monthly evaporative loss using relevance vector machine, extreme learning machine and multivariate adaptive regression spline models." *Stochastic Environmental Research and Risk Assessment*, Vol. 30, No. 6, pp. 1769-1784, DOI: 10.1007/s00477-015-1153-y.

Dracup, J. A., Lee, K. S., and Paulson, E. G. (1980). "On the definition of droughts." *Water Resources Research*, Vol. 16, No. 2, pp. 297-302, DOI: 10.1029/WR016i002p00297.

Giuntoli, I., Renard, B., Vidal, J.-P., and Bard, A. (2013). "Low flows in France and their relationship to large-scale climate indices." *Journal of Hydrology*, Vol. 482, pp. 105-118, DOI: 10.1016/j.jhydrol.2012.12.038.

Gustard, A. and Demuth, S. (2009). *Manual on low-flow estimation and prediction*, Opera.

Haghiabi, A. H. (2016). "Prediction of longitudinal dispersion coefficient using multivariate adaptive regression splines." *Journal of Earth System Science*, Vol. 125, No. 5, pp. 985-995, DOI: 10.1007/s12040-016-0708-8.

Haghiabi, A. H. (2017). "Modeling river mixing mechanism using data driven model." *Water Resources Management*, Vol. 31, No. 3, pp. 811-824, DOI: 10.1007/s11269-016-1475-7.

Haghiabi, A. H., Azamathulla, H. M., and Parsaie, A. (2017). "Prediction of head loss on cascade weir using ANN and SVM." *ISH Journal of Hydraulic Engineering*, Vol. 23, No. 1, pp. 102-110, DOI: 10.1080/09715010.2016.1241724.

Haghiabi, A. H., Nasrolahi, A. H., and Parsaie, A. (2018). "Water quality prediction using machine learning methods." *Water Quality Research Journal*, Vol. 53, No. 1, pp. 3-13, DOI: 10.2166/wqrj.2018.025.

Han, D., Chan, L., and Zhu, N. (2007). "Flood forecasting using support vector machines." *Journal of Hydroinformatics*, Vol. 9, No. 4, pp. 267-276, DOI: 10.2166/hydro.2007.027.

Hipel, K. W. and McLeod, A. I. (1994). *Time series modelling of water resources and environmental systems*, Elsevier.

Huang, G.-B., Zhu, Q.-Y., and Siew, C.-K. (2006). "Extreme learning machine: Theory and applications." *Neurocomputing*, Vol. 70, Nos. 1-3, pp. 489-501, DOI: 10.1016/j.neucom.2005.12.126

Jha, R. and Smakhtin, V. (2008). *A review of methods of hydrological estimation at ungauged sites in India*, IWMI.

Jha, R., Sharma, K., and Singh, V. (2008). "Critical appraisal of methods for the assessment of environmental flows and their application in two river systems of India." *KSCE Journal of Civil Engineering*, Vol. 12, No. 3, pp. 213-219, DOI: 10.1007/s12205-008-0213-y.

Khan, M. S. and Coulibaly, P. (2006). "Application of support vector machine in lake water level prediction." *Journal of Hydrologic Engineering*, Vol. 11, No. 3, pp. 199-205, DOI: 10.1061/(ASCE)1084-0699(2006)11:3(199).

Kim, T.-W. and Valdés, J. B. (2003). "Nonlinear model for drought forecasting based on a conjunction of wavelet transforms and neural networks." *Journal of Hydrologic Engineering*, Vol. 8, No. 6, pp. 319-328, DOI: 10.1061/(ASCE)1084-0699(2003)8:6(319).

Komorník, J., Komorníková, M., Mesiar, R., Szökeová, D., and Szolgay, J. (2006). "Comparison of forecasting performance of nonlinear models of hydrological time series." *Physics and Chemistry of the Earth, Parts A/B/C*, Vol. 31, No. 18, pp. 1127-1145, DOI: 10.1016/j.pce.2006.05.006.

Kumar, R., Goel, N. K., Chatterjee, C., and Nayak, P. C. (2015). "Regional flood frequency analysis using soft computing techniques." *Water Resources Management*, Vol. 29, No. 6, pp. 1965-1978, DOI: 10.1007/s11269-015-0922-1.

Laaha, G. and Blöschl, G. (2005). "Low flow estimates from short stream flow records—a comparison of methods." *Journal of Hydrology*, Vol. 306, Nos. 1-4, pp. 264-286, DOI: 10.1016/j.jhydrol.2004.09.012.

Lei, Y., Zhao, D., and Cai, H. (2015). "Prediction of length-of-day using extreme learning machine." *Geodesy and Geodynamics*, Vol. 6, No. 2, pp. 151-159, DOI: 10.1016/j.geog.2014.12.007.

Lin, J.-Y., Cheng, C.-T., and Chau, K.-W. (2006). "Using support vector machines for long-term discharge prediction." *Hydrological Sciences Journal*, Vol. 51, No. 4, pp. 599-612, DOI: 10.1623/hysj.51.4.599.

MacKay, D. J. (1996). "Bayesian methods for backpropagation networks." *Models of neural networks III. Springer*, pp. 211-254, DOI: 10.1007/978-1-4612-0723-8_6.

Müller, K.-R., Smola, A. J., Rätsch, G., Schölkopf, B., Kohlmorgen, J., and Vapnik, V. (1997). "Predicting time series with support vector machines." *International Conference on Artificial Neural Networks*. Springer, pp. 999-1004, DOI: 10.1007/BFb0020283.

Najafzadeh, M. and Azamathulla, H. M. (2013). "Group method of data handling to predict scour depth around bridge piers." *Neural Computing and Applications*, Vol. 23, Nos. 7-8, pp. 2107-2112, DOI: 10.1007/s00521-012-1160-6.

Najafzadeh, M. and Saberi-Movahed, F. (2018). "GMDH-GEP to predict free span expansion rates below pipelines under waves." *Marine Georesources & Geotechnology*, pp. 1-18, DOI: 10.1080/1064119X.2018.1443355.

Najafzadeh, M., Etemad-Shahidi, A., and Lim, S. Y. (2016). "Scour prediction in long contractions using ANFIS and SVM." *Ocean Engineering*, Vol. 111, pp. 128-135, DOI: 10.1016/j.oceaneng.2015.10.053.

Najafzadeh, M., Rezaie-Balf, M., and Tafarojnoruz, A. (2018). "Prediction of riprap stone size under overtopping flow using data-driven models." *International Journal of River Basin Management*, pp. 1-8, DOI: 10.1080/15715124.2018.1437738.

Najafzadeh, M., Saberi-Movahed, F., and Sarkamaryan, S. (2017). "NF-GMDH-Based self-organized systems to predict bridge pier scour depth under debris flow effects." *Marine Georesources & Geotechnology*, Vol. 36, No. 5, pp. 589-602, DOI: 10.1080/1064119X. 2017.1355944.

Nayak, P. C., Sudheer, K., Rangan, D., and Ramasastri, K. (2004). "A neuro-fuzzy computing technique for modeling hydrological time series." *Journal of Hydrology*, Vol. 291, Nos. 1-2, pp. 52-66, DOI: 10.1016/j.jhydrol.2003.12.010.

Nikbakht, S. A., Zahraie, B., and Nasseri, M. (2012). "Seasonal meteorological drought prediction using support vector machine." Vol. 23, No. 2, pp. 72-84.

Okkan, U. and Inan, G. (2014). "Bayesian learning and relevance vector machines approach for downscaling of monthly precipitation." *Journal of Hydrologic Engineering*, Vol. 20, No. 4, pp. 04014051,

DOI: 10.1061/(ASCE)HE.1943-5584.0001024.

Osuna, E., Freund, R., and Girosi, F. (1997). *Support vector machines: Training and applications.*

Parsaie, A. and Haghiabi, A. H. (2017a). "Computational modeling of pollution transmission in rivers." *Applied water science*, Vol. 7, No. 3, pp. 1213-1222, DOI: 10.1007/s13201-015-0319-6.

Parsaie, A. and Haghiabi, A. H. (2017b). "Improving modelling of discharge coefficient of triangular labyrinth lateral weirs using SVM, GMDH and MARS Techniques." *Irrigation and Drainage*, Vol. 66, No. 4, pp. 636-654, DOI: 10.1002/ird.2125.

Parsaie, A. and Haghiabi, A. H. (2017c). "Mathematical expression of discharge capacity of compound open channels using MARS technique." *Journal of Earth System Science*, Vol. 126, No. 2, p. 20, DOI: 10.1007/s12040-017-0807-1.

Parsaie, A., Azamathulla, H. M., and Haghiabi, A. H. (2017a). "Physical and numerical modeling of performance of detention dams." *Journal of Hydrology*, DOI: 10.1016/j.jhydrol.2017.01.018.

Parsaie, A., Ememgholizadeh, S., Haghiabi, A. H., and Moradinejad, A. (2018a). "Investigation of trap efficiency of retention dams." *Water Science and Technology: Water Supply*, Vol. 18, No. 2, pp. 450-459, DOI:10.2166/ws.2017.109.

Parsaie, A., Haghiabi, A. H., Saneie, M., and Torabi, H. (2018b). "Prediction of energy dissipation of flow over stepped spillways using data-driven models." *Iranian Journal of Science and Technology, Transactions of Civil Engineering*, Vol. 42, No. 1, pp. 39-53, DOI: 10.1007/s40996-017-0060-5.

Parsaie, A., Haghiabi, A. H., Saneie, M., and Torabi, H. (2016). "Applications of soft computing techniques for prediction of energy dissipation on stepped spillways." *Neural Computing and Applications*, Vol. 29, No. 12, pp. 1393-1409, DOI: 10.1007/s00521-016-2667-z.

Parsaie, A., Yonesi, H., and Najafian, S. (2017b). "Prediction of flow discharge in compound open channels using adaptive neuro fuzzy inference system method." *Flow Measurement and Instrumentation*, Vol. 54, pp. 288-297, DOI: 10.1016/j.flowmeasinst.2016.08.013.

Pyrce, R. (2004). *Hydrological low flow indices and their uses*, Watershed Science Centre (WSC) Report.

Qishlaqi, A., Kordian, S., and Parsaie, A. (2017). "Hydrochemical evaluation of river water quality—a case study." *Applied Water Science*, Vol. 7, No. 5, pp. 2337-2342, DOI: 10.1007/s13201-016-0409-0.

Rajesh, R. and Prakash, J. S. (2011). "Extreme learning machines-a review and state-of-the-art." *International journal of wisdom based computing*, Vol. 1, No. 1, pp. 35-49.

Rezaie-Balf, M. and Kisi, O. (2017). "New formulation for forecasting streamflow: Evolutionary polynomial regression vs. extreme learning machine." *Hydrology Research*, Vol. 49, No. 3, pp. 939-953, DOI: 10.2166/nh.2017.283.

Salas, J. D. (1993). "Analysis and modeling of hydrologic time series." *Handbook of Hydrology*, Vol. 19, pp. 1-72.

Seo, Y., Kim, S., and Singh, V. P. (2015). "Multistep-ahead flood forecasting using wavelet and data-driven methods." *KSCE Journal of Civil Engineering*, Vol. 19, No. 2, pp. 401-417, DOI: 10.1007/s12205-015-1483-9.

Shiri, J. and Kişi, Ö. (2011). "Application of artificial intelligence to estimate daily pan evaporation using available and estimated climatic data in the Khozestan Province (South Western Iran)." *Journal of Irrigation and Drainage Engineering*, Vol. 137, No. 7, pp. 412-425, DOI: 10.1061/(ASCE)IR.1943-4774.0000315.

Sivapragasam, C. and Muttil, N. (2005). "Discharge rating curve extension–a new approach." *Water Resources Management*, Vol. 19, No. 5, pp. 505-520, DOI: 10.1007/s11269-005-6811-2.

Sivapragasam, C., Maheswaran, R., and Venkatesh, V. (2008). "Genetic programming approach for flood routing in natural channels." *Hydrological Processes*, Vol. 22, No. 5, pp. 623-628, DOI: 10.1002/hyp.6628.

Srikanthan, R. and McMahon, T. (2001). "Stochastic generation of annual, monthly and daily climate data: A review." *Hydrology and Earth System Sciences Discussions*, Vol. 5, No. 4, pp. 653-670, DOI: 10.5194/hess-5-653-2001.

Sudheer, K. and Jain, A. (2004). "Explaining the internal behaviour of artificial neural network river flow models." *Hydrological Processes*, Vol. 18, No. 4, pp. 833-844, DOI: 10.1002/hyp.5517.

Sudheer, K., Gosain, A., and Ramasastri, K. (2002). "A data-driven algorithm for constructing artificial neural network rainfall-runoff models." *Hydrological Processes*, Vol. 16, No. 6, pp. 1325-1330, DOI: 10.1002/hyp.554.

Tao, W., Kailin, Y., and Yongxin, G. (2008). "Application of artificial neural networks to forecasting ice conditions of the Yellow River in the Inner Mongolia Reach." *Journal of Hydrologic Engineering*, Vol. 13, No. 9, pp. 811-816, DOI: 10.1061/(ASCE)1084-0699(2008)13:9(811).

Tiwari, M. K. and Adamowski, J. (2013). "Urban water demand forecasting and uncertainty assessment using ensemble wavelet-bootstrap-neural network models." *Water Resources Research*, Vol. 49, No. 10, pp. 6486-6507, DOI: 10.1002/wrcr.20517.

Tokar, A. S. and Johnson, P. A. (1999). "Rainfall-runoff modeling using artificial neural networks." *Journal of Hydrologic Engineering*, Vol. 4, No. 10, pp. 232-239, DOI: 10.1061/(ASCE)1084-0699(1999)4:3(232).

Toth, E., Brath, A., and Montanari, A. (2000). "Comparison of short-term rainfall prediction models for real-time flood forecasting." *Journal of Hydrology*, Vol. 239, Nos. 1-4, pp. 132-147, DOI: 10.1016/S0022-1694(00)00344-9.

Vapnik, V. (1998). *Statistical learning theory*, Wiley, New York.

Wang, W.-C., Chau, K.-W., Cheng, C.-T., and Qiu, L. (2009). "A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series." *Journal of Hydrology*, Vol. 374, Nos. 3-4, pp. 294-306, DOI: 10.1016/j.jhydrol.2009. 06.019.

Williams, C. K. and Rasmussen, C. E. (1996). "Gaussian processes for regression." *Advances in Neural Information Processing Systems*, pp. 514-520.

Wu, C. and Chau, K.-W. (2010). "Data-driven models for monthly streamflow time series prediction." *Engineering Applications of Artificial Intelligence*, Vol. 23, No. 8, pp. 1350-1367, DOI: 10.1016/j.engappai.2010.04.003.

Yu, P.-S., Chen, S.-T., and Chang, I.-F. (2006). "Support vector regression for real-time flood stage forecasting." *Journal of Hydrology*, Vol. 328, Nos. 3-4, pp.704-716, DOI: 10.1016/j.jhydrol.2006.01.021.

Zahiri, A. and Najafzadeh, M. (2018). "Optimized expressions to evaluate the flow discharge in main channels and floodplains using evolutionary computing and model classification." *International Journal of River Basin Management*, Vol. 16, No. 1, pp. 123-132, DOI: 10.1080/15715124.2017.1372448.