Prediction of Channel Sinuosity in Perennial Rivers Using Bayesian Mutual Information Theory and Support Vector Regression Coupled with Meta-Heuristic Algorithms

Masoud Haghbin[1], Ahmad Sharafati[2]*, Davide Motta[3]

[1] Department of Structural Mechanics & Hydraulics Engineering, University of Granada, 18001 Granada, Spain

[2] Department of Civil Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran.

[3] Department of Mechanical and Construction Engineering, Northumbria University, Wynne Jones Building, Newcastle upon Tyne NE1 8ST, United Kingdom

Corresponding author: Ahmad Sharafati

Corresponding email: asharafati@gmail.com, asharafati@srbiau.ac.ir

Abstract

Support Vector Regression (SVR) combined with Invasive Weeds Optimization (IWO), standalone SVR, and Radial Basis Function Neural Networks (RBFNNs) are applied to estimate channel sinuosity in perennial rivers. With this aim, a dataset with 132 sinuosity data and related geomorphologic data, corresponding to 119 perennial streams, is considered. Bayesian Mutual Information theory is used to determine the parameters affecting channel sinuosity to reveal that bankfull depth affects sinuosity the most. Seven input parameter combinations for sinuosity prediction are considered, and in both training and testing stages, the SVR-IWO model ($R_{Train} = 0.959$, $RMSE_{Train} = 0.072$, $MAE_{Train} = 0.037$, $R_{test} = 0.892$, $RMSE_{Test} = 0.103$, $MAE_{Test} = 0.065$) shows the best prediction performance while the standalone SVR model generated the results with performances of ($R_{Train} = 0.792$, $RMSE_{Train} = 0.158$, $MAE_{Train} = 0.141$, $R_{test} = 0.704$, $RMSE_{Test} = 0.163$, $MAE_{Test} = 0.151$) . Model prediction uncertainty is quantified in terms of entropy for the three models considered, further confirming that the sinuosity set predicted by the SVR-IWO model is the closest to the observed set.

Keywords: channel sinuosity, perennial rivers, prediction, Bayesian Mutual Information theory, Meta-Heuristic Algorithms

1. Introduction

Channel sinuosity measures the deviation of a stream route from a straight downstream route, calculated as the ratio between the stream length and the corresponding length measured along a straight route. For a straight channel, sinuosity is equal to one; meandering channels and, in general, channels with bends have a sinuosity larger than one. As meandering rivers migrate because of bank erosion, sinuosity changes in time (Ahmed et al., 2019; Dente et al., 2019; Van den Berg, 1995; Woolderink et al., 2021).

The sinuosity of alluvial channels depends on topographic, geologic, and hydraulic conditions. Studies on river sinuosity and, more generally, meander migration have been conducted using numerical or analytical models, laboratory experiments, and field observations. The specific focus of this paper is on sinuosity and the improvement of "classic" empirically-based regression equations expressing sinuosity as a function of parameters such discharge, channel geometry, and sediment characteristics, measured with field observations (Deike and White, 1969; Ferguson, 1977; Leopold and Wolman, 1957; Moody-Stuart, 1966; Schumm, 1963). Several attempts were made to determine the influential factors on sinuosity in alluvial meandering channels, i.e., flow conditions and stream power, bed material, and topography (Ghosh, 2000; Hooke, 1975; Le Roux, 1992; Lewin and Brewer, 2001; Nanson and Hickin, 1983; Smith, 1998; Van den Berg, 1995).

Nowadays, there is a strong demand for finding cost-effective approaches to solve different and complex problems in hydrology, geoscience, hydraulic, and various fields of earth sciences. (Haghbin et al., 2021, 2020; Jamei et al., 2020; Sharafati et al., 2021b, 2020). SC models are well-known alternatives that are quite practical in prediction and classification problems (Jamei and Ahmadianfar, 2020; Pourrajab et al., 2020; Sharafati et al., 2021a). However, in some complex issues, these types of approaches are trapped in local solutions, to solve this weakness, researchers are conducted different ways to improve the performance and reliability of existing modes by combining them with different evolutionary or stochastic models (Diop et al., 2020; Naganna et al., 2019; Tao et al., 2021b, 2021a).

In the past two decades, soft computing methods have been increasingly applied in engineering problems such as analyzing alluvial channels and, specifically, meandering rivers (Waszczyszyn, 2010).

One of the first studies was carried out by (Javaheri et al., 2008), who applied fuzzy clustering methods to determine the geomorphologic characteristics of the Karoon River in Iran. (Sahu et al., 2011) used a Back-Propagation Neural Network (BPNN) approach for estimating flow velocity in meandering rivers as a better alternative to "classic" regression models. (Riahi-Madvar et al., 2011) showed the potential of Artificial Neural Networks (ANNs) in predicting the geometry of alluvial channels, considering parameters such as discharge, width, depth, and median diameter of the bed material. Other soft computing models, namely Support Vector Machine (SVM) and Adaptive Neuro-Fuzzy Inference System (ANFIS), were used for the prediction of alluvial channel patterns and their roughness coefficients (Beechie and Imaki, 2014; Moharana and Khatua, 2014). Several studies adopting soft computing approaches in alluvial channels focused on prediction of the flow field in curved channels, the stable river profile, and width, and the threshold channel bank profile (Baghalian et al., 2012; Bonakdari et al., 2019; Gholami et al., 2019a, 2019b, 2018; Shaghaghi et al., 2017; Tahershamsi et al., 2012). Another study (Pham et al., 2019) assessed the variation of the river in Da Dien Estuary, Vietnam using different techniques such as Logistic Regression, Neural Networks, Bootstrap AdaBoost, LogitBoost, bagging, and random subspace. Their finding revealed that random subspace could provide more reliable results in comparison with others. In line with this study for applying SC models in river engineering, (Gholami et al., 2020) employed Feed Forward Neural Networks (FFNN), Extreme Learning Machine (ELM), and their combined types (FFNN-ELM) for assessing the bank profile morphology in different rivers, They reported that their new hybrid version could generate superior results in comparison of standalone models.

Together, all these studies show that soft computing models can reliably estimate alluvial channel properties and fluvial conditions. However, no research has been conducted so far regarding their performance in estimating channel sinuosity, which is the primary goal of this study. This investigation focuses on the application of Support Vector Regression (SVR) "tuned" by Invasive Weed Optimization (IWO), standalone SVR, and Radial Basis Function (RBF) algorithms. It identifies the best metaheuristic model for improving the performance of a standalone SVR model in predicting channel sinuosity and quantifies the relative uncertainty. It also presents a new model that can be applied in alluvial channels to estimate channel sinuosity as an alternative to the "classic" regression-based approaches.

2. Methodology

This section includes an overview of soft computing models, dataset, input combinations, performance indices, and uncertainty approaches, which were employed in our study.

2.1 Channel Sinuosity Prediction Models

Soft computing models can overcome the limitations (over or underestimated results) of classic empirical relations for complex or non-linear problems. The following sections provide a brief background on the soft computing models used in this study.

2.1.1 Support Vector Regression (SVR)

The concept of SVR originates from Support Vector Machine (SVM), which is a supervised learning approach used in classification or regression problems. SVM is widely used in different application areas, such as assessing scouring, sea surface temperature, and water quality (Liu et al., 2019).

The SVR method is a type of SVM suitable for regression problems. Unlike SVM, which is based on transforming the original finite-dimensional space into a higher-dimensional space for classification problems, the SVR provides a regression without the primary restrictions, for instance, generating inaccurate results when the number of input parameters is increased in the datasets, or there are noisy ones which linear models are usually suffered from them. The SVR regression expression that maps the input vector to a higher dimension of solution pace is as follows:

$$\mathrm{F}(\vec{\mathrm{x}}, \vec{\omega}) = \sum_{\mathrm{j}=1}^{\mathrm{m}} \omega_{\mathrm{j}}\, \mathrm{g}_{\mathrm{j}}(\vec{\mathrm{x}}) \quad (1)$$

where $\mathrm{g}(\vec{\mathrm{x}})$ is non-linear transformation function, $\vec{\omega}$ is norm vector, and m is the number of samples.

Figure 1 contains a schematic view of linear regression achieved after mapping the input vector to a high dimensional space through non-linear transformation. The interval of non-approximation error is limited with dashed lines. The initial step of SVR is to determine a loss function used to assess the error while the regression process is performed. There are different types of loss functions, such as quadratic, Huber, and ε-insensitive. The most common loss function is ε-insensitive, which can be determined as below:

$$L_s\left(y, F(\vec{x}, \vec{\omega})\right) = \begin{cases} 0 \quad \text{if } |y - F(x, \omega)| < \varepsilon \\ |y - F(x, \omega)| - \varepsilon \ \text{ otherwise} \end{cases} \tag{2}$$

The next step of SVR is an optimization problem, which consists of finding the most appropriate vector $\vec{\omega}$ which satisfies the specific error restriction given in Equation (2). This translates into the following expression:

$$\begin{aligned} &\mathit{Objective\ function} : \min \frac{1}{2} || \ \vec{\omega} \ ||^2 \\ &\text{Subjected to} \\ &y_i - F(\vec{x_i}, \vec{\omega}) \le \varepsilon \\ &F(\vec{x_i}, \vec{\omega}) - y_i \le \varepsilon \end{aligned} \tag{3}$$

By considering slack variables, the objective function and its associated constraints are modified as follows:

$$\begin{aligned} &\mathit{Objective\ function} : \min \frac{1}{2} || \ \vec{\omega} \ ||^2 + C \textstyle\sum_{i=1}^{n} (\ \varepsilon_i - \ \varepsilon_i^*) \\ &\text{Subjected to} \\ &y_i - F(\vec{x_i}, \vec{\omega}) \le \varepsilon_i - \ \varepsilon_i^* \\ &F(\vec{x_i}, \vec{\omega}) - y_i \le \varepsilon_i - \ \varepsilon_i^* \end{aligned} \tag{4}$$

These new constraints present that the value of ε as loss function should not exceed more than $\varepsilon_i - \ \varepsilon_i^*$. In addition, C is determined as a constant coefficient. Finally, the dual is achieved to solve the problem again, and the final result is determined as the solution of the following objective function:

$$\begin{aligned} &\text{Max } F(\vec{x}) = \sum_{i=1}^{i} (\alpha_i - \alpha_i^*) k(x_i, x_j) + b \\ &\text{Subjected to} \\ &0 \le \alpha_i^* \le C \\ &0 \le \alpha_i \le C \end{aligned} \tag{5}$$

where α_i^* *and* α_i are Lagrange multipliers, C is a coefficient measuring the trade-off between weights and approximation error, $k(x_i, x_j)$ is the kernel function, and b is a constant value. Kernel functions such as Radial Basis and polynomial are employed to calculate $g_j(x)$ and optimize the calculations in high dimension spaces. The obtained non-linear regression with kernel function is determined as bellows:

$$k(x_i, x_j) = \exp\,(-||x_i - x_j||/\sigma^2|| \qquad (6)$$

where σ denotes the kernel parameters.

[Figure 1]

2.1.2 Invasive Weed Optimization (IWO)

(Mehrabian and Lucas, 2006) presented Invasive Weed Optimization (IWO) as a metaheuristic soft computing approach for solving different engineering problems. It is inspired by the process of making colonies by weeds and finding a suitable place for growth and reproducing. IWO is made of four steps, and a flow chart of IWO is shown in Figure 2. The first step is the random initialization of the first population (primary solution):

$$X = [X_1 X_2, \ldots, X_m] \qquad (7)$$

The second step is reproduction, where a seed population is generated by each of the members of the initial population. The number of seeds produced by each member depends on the value of its fitness function, as follows

$$Weed = \text{floor}\left(\frac{f - f_{min}}{f_{max} - f_{min}}\right)(s_{max} - s_{min}) + s_{min} \qquad (8)$$

where the floor, f, f_{min},f_{max}, s_{min} *and* s_{max} are the round down operator, fitness, minimum and maximum fitness values, and minimum and maximum number of seeds, respectively.

The third step is spatial dispersal, where the newly generated seeds spread over and search space. The level of difference between each member of the population and its offspring is quantified as the distance between the parent plant and where the seed falls on the ground. The distance is described using a normal distribution with a certain standard deviation, which in each iteration is decreased as follows

$$\sigma_{iter} = \frac{(iter_{max} - iter)^n}{(iter_{max})^n}\left(\sigma_{inital} - \sigma_{final}\right) + \sigma_{final} \tag{9}$$

where $\sigma_{iter}, iter_{max}, \sigma_{inital}$ and σ_{final} are standard deviation for the present iteration, maximum number of iterations, and initial and final standard deviations.

The fourth step is competitive exclusion, where a new population is produced for the next iteration by selecting the best-adapted individual in the neighborhood of each solution member based on fitness.

[Figure 2]

Algorithm 1: Invasive Weed Optimization
Input: **Maximum iterations** **Number of population** **Minimum and maximum number of seeds** **Variance reduction exponent** **Initial and final values of standard deviation** **Output: IWO input structure** **1. (Step 1) Initialization** **2. (Step1.1) Initialize population members.** **3. (Step 1.2) Create, evaluate and sort the population** **4.(Step1.3) Initialize the best solution ever found** **5 (Step2) Update for i =1 ,…., nPop do;** **6.(Step2.1) compute standard deviation, Initialize off springs population, reproduction** **7.(Step 2.2) Evaluate, sort, deleting extra members, merge new population** **8.(Step 2.3) Update worst cost, Truncation, and store the best solution.** **9.(Step 3) Stopping Criterion. If the stopping criterion is satisfied, then stop and output. otherwise, go Step 2.**

2.1.3 Coupled SVR-IWO

Optimizing a SVR model by "tuning" (finding the best combination of) its parameters is essential (Fattahi and Babanouri, 2017). The IWO algorithm can be used to tune the SVR model parameters, namely the regularization parameter *C*, the error margin ε and the RBF kernel parameter σ. Figure 3 shows a flow chart of the optimization of the SVR parameters using IWO.

[Figure 3]

2.1.4 Radial Basis Functions Neural Networks (RBFNNs)

Radial Basis Function Neural Networks (RBFNNs) are neural networks of the feed-forward type and are often utilized for multi-variable regression in complex problems (Chen et al., 2016). A flow chart of RBFNN is shown in Figure 4. In a RBFNN model, neurons are arranged in three layers. The first layer contains the input nodes; the second layer is a hidden layer, and it includes tuned units using a radial basis function, and the third layer is the output layer that can be computed as follows

$$y = f(x) = \sum_{k=1}^{N} w_k \, \varphi_k(||x - c_k||) \tag{10}$$

where y, $x, N, w_k, \varphi_k \; and \; c_k$ are output, input vector, number of neurons in the hidden layer, weights of output obtained from the previous layer, radial basis function, and center of the radial basis function, respectively. The term $||x - c_k||$ expresses the Euclidean distance between inputs and the center of the radial basis function.

Multi-quadratic and Gaussian functions are widely used as the radial basis function. Previous studies reported that outputs are generally not sensitive to the type of radial basis function selected (Chen et al., 2016; Li et al., 2008; Yang and Paindavoine, 2003). In this study, the Gaussian function is used. It is expressed as follows

$$\varphi(||x-c||)=\exp\left(-\frac{||x-c||^2}{r^2}\right) \quad (11)$$

where *c* and *r* represent the center and radius of the Gaussian function, respectively.

Choosing appropriate centers is the main challenge in the application of this method. In this regard, an Orthogonal Least Squares (OLS) approach can be used to explore the optimal number of centers during the training phase. An overview of this approach is discussed in (Ham and Kostanic, 2001).

[Figure 4]

2.2 Dataset and Input Parameters

To carry out this study, 132 sinuosity data, corresponding to 119 perennial streams (Table 1), were obtained from (Van den Berg, 1995). The main features of the streams selected are described as follows (Kleinhans and van den Berg, 2011; Van den Berg, 1995): the stream planform characteristics are not affected by man-made activity such as dams, dredging, and jetties; no artificial cut-offs; perennial flow regime; and no roads near the streams.

[Table 1]

Seven parameters were considered for channel sinuosity prediction: valley slope (S_v), bankfull depth (d), bankfull width (w), bankfull discharge (Q_{bf}), median bed sediment grain size (D_{50}), potential specific stream power (ω_v) and potential specific stream power at the transition between single and thread channel ($\omega_{v,t}$). The parameters $S_v,\ d,\ w,\ Q_{bf}\ and\ D_{50}$ were directly obtained from the field observations gathered by (Van den Berg, 1995); the remaining two parameters, ω_v and $\omega_{v,t}$, were computed with the formulations suggested by (Van den Berg, 1995) as follows:

$$\text{Sand bed river: } \omega_v = 2.1 S_v\sqrt{Q_{bf}} \quad (12)$$
$$\text{Gravel bed river: } \omega_v = 3.3 S_v\sqrt{Q_{bf}} \quad (13)$$
$$\omega_{v,t} = 900 D_{50}^{0.42} \quad (14)$$

The dataset was divided randomly into the training data (60% of data) and the remaining (40% of data) for the testing data. The range of values for the seven parameters in the dataset considered is shown in Table 2.

[Table 2]

2.3 Mutual Information Theory for Input Variable Combination Selection

Information theory or entropy of information is recognized as a tool for measuring uncertainties due to lack of knowledge and chaos and was first presented by (Shannon and Weaver, 1949). Mutual Information theory is a useful statistical tool in engineering for measuring the degree of dependency of an event on certain parameters (Archer et al., 2013; Hlaváčková-Schindler et al., 2007; Rényi, 1959) and is a widely used approach for the selection of variables and determination of flow of information (Singh, 2014). The Mutual Information between two random variables, X and Y, is expressed as follow

$$I(X,Y) = H(X) - H(X|Y) \tag{15}$$

where $I(X,Y)$ represents the mutual information (entropy) or dependency degree of the variables X and Y. $H(X)$ is the entropy of the random variable X, with possible values $\{X_1, X_2, \dots, X_N\}$. It is about the information in the variable X and the uncertainty associated with it, and it is expressed as follows

$$H(X) = -\sum_{i=1}^{N} P(X_i) \log P(X_i) \tag{16}$$

where $P(X_i)$ are the probability values associated to the values X_i. The term $H(X|Y)$ denotes the conditional entropy for X given Y and is computed as follows

$$H(X|Y) = -\sum_{i=1}^{N}\sum_{j=1}^{M} P(X_i, Y_j) \log P(X_i|Y_j) \tag{17}$$

In the above relation, $P(X_i|Y_j)$ expresses the conditional probability of X on Y. Equation (15) is affected bias for small datasets (Archer et al., 2013). To tackle this problem, researchers have been using Bayesian inference to measure H (Hutter, 2002; Hutter and Zaffalon, 2002). Bayesian inference employs Dirichlet prior and posterior probabilities to fit a multinomial distribution to variables. The mathematical expressions of the Dirichlet prior and posterior probabilities are shown below

$$\mathrm{Dir}(\alpha) \triangleq \mathrm{Dir}(\alpha_1, \alpha_2, \dots, a_K) = \frac{\Gamma(Ka)}{\Gamma(\alpha)^K} \prod_{i=1}^{K} \pi_i^{\alpha-1}$$
$$\mathrm{Dir}(\alpha) \triangleq \mathrm{Dir}(\alpha_1 + n_1, \dots, \alpha + n_K) = \Gamma(Ka + N) = \prod_{i=1}^{K} \frac{\pi_i^{n_{i+\alpha-1}}}{\Gamma(a + n_i)} \quad (18)$$

where α, K, π_i, N and n_K denote the Dirichlet concentration coefficient, numbers of the defined bin in the distribution, the probability that data sample X is placed in the i_{th} bin, number of all samples, and number of samples places in the i_{th} bin. By applying the Dirichlet distribution to the Mutual Information expression, its new expression is obtained:

$$\text{Bayesian Mutual Information} = \frac{1}{ln2}(\psi_0(N + Ka + 1) - \sum_{i,j} \frac{n_{i,j}+\alpha}{(N+\alpha K)} [\psi_0\left(n_{x_i} + \alpha K_y + 1\right) + \psi_0\left(n_{yi} + \alpha K_x + 1\right) - \psi_0\left(n_{ij} + \alpha + 1\right)) \quad (19)$$

In the above relation, all parameters are the same as for Equation (18) and ψ_0 denotes the digamma function. A detailed overview of this approach is discussed in (Hutter, 2002; Hutter and Zaffalon, 2002).

The formulation in Equation (19) is used in this study to compute the Bayesian Mutual Information associated with sinuosity for each of the seven input variables considered. This informs the selection of the input variable combinations for the prediction of sinuosity.

2.4 Prediction Performance Indices

The sinuosity prediction performance of the models considered in this study was quantified using three indices, Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Coefficient of Correlation (R). They are computed as follows

Root Mean Square Error (RMSE)

$$\sqrt{\frac{\sum_{i=1}^{n}(X_{Observed} - X_{Predicted})^2}{n}} \tag{20}$$

Mean Absolute Error (MAE)

$$\frac{\sum_{i=1}^{n}|X_{Predicted} - X_{Observed}|}{n} \tag{21}$$

Coefficient of Correlation (R)

$$\frac{n(\sum X_{Observed}X_{Predicted}) - (\sum X_{Observed})(\sum X_{Predicted})}{\sqrt{[n\sum(X_{Observed}^2) - (\sum X_{Observed})^2][n\sum(X_{Predicted}^2) - (\sum X_{Predicted})^2]}} \tag{22}$$

where $X_{Observed}$ and $X_{Predicted}$ are observed and predicted values, and n is the number of observed/predicted values.

2.5 Model Prediction Uncertainty

There are two types of uncertainty, aleatory and epistemic. The aleatory uncertainty is due to the inherent randomness in physical phenomena, with the variable values that a probability distribution can describe. The epistemic uncertainty is the uncertainty in the modeling of the physical processes. The idealization on which models or mathematical expressions rely is linked to epistemic uncertainty. The epistemic (model) uncertainty is the subject of this section.

As seen in Section 2.3, in information theory, the "randomness" or "uncertainty" of a variable, such as a sinuosity, can be evaluated in terms of entropy. Five methods are considered here to compute it, one "classic" - Maximum Likelihood (ML) – and the other four of Bayesian type - Jeffrey, Laplace, Schurmann-Grassberger (SG), and Minimax. The difference between these methods resides in the way the value of α

in the Dirichlet priors is computed. Its value is zero in the ML method, which reduces this method to measure entropy to the "classic" entropy definition seen in Equation (16). The value of α is 0.5, 1, $\frac{1}{Size\ of\ vector\ of\ generated\ results}$, and $\sqrt{\frac{Summation\ of\ generated\ results}{Size\ of\ vector\ of\ generated\ results}}$ for Jeffrey, Laplace, SG, and Minimax methods, respectively.

3. Results and Discussion

3.1 Bayesian Mutual Information and Input Variable Combination

The Mutual Information associated with sinuosity for each of the seven input variables considered for the dataset used in this study is shown in Table 3 and Figure 5. The variables affecting sinuosity the most and the least are bankfull depth and valley slope, respectively.

[Table 3]

[Figure 5]

Based on the analysis above, seven input variable combinations for the prediction of sinuosity are selected, as shown in Table 4.

[Table 4]

3.2 Model Prediction Performance

The sinuosity prediction performance of SVR-IWO, standalone SVR, and RBF models was compared, based on the performance indices RMSE, MAE, and R. Table 5 summarizes the best-obtained results for each of the estimators (i.e., the results obtained for the best performing input variable combination for each model), for training and testing stages. In the training stage, the SVR-IWO-M2 model sinuosity prediction are the most accurate, producing the minimum MAE (0.037) and the maximum R (0.959). The standalone SVR-M2 still generates acceptable results, while RBF-M1 produces the highest prediction error. Similar findings are obtained in the testing stage, where the SVR-IWO-M2 model generates the minimum sinuosity prediction error (RMSE = 0.103, MAE = 0.065, and R=0.892). The standalone SVR-M2 (RMSE = 0.163, MAE = 0.151 and R = 0.704) performs slightly better than RBF-M1 (RMSE = 0.236, MAE = 0.177 and R = 0.462).

[Table 5]

The prediction performance of the different models is visualized for training and testing stages in the form of heat maps (Figure 6), scatter plots (Figure 7), and boxplots (Figure 8). In all plots, the best input parameter combination is considered for each model.

In the heat map diagrams, the prediction performance indices are normalized (standardized) with respect to the difference between the maximum and minimum value of each performance index. The larger the value of the normalized index, the better the prediction performance. The heat maps in Figure 6 clearly show that the SVR-IWO model has the best prediction performance among the three models considered for both the training and testing stages.

[Figure 6]

From the scatter plots in Figure 7, the RBF model produces the results with the minimum linear correlation between observed and estimated sinuosity ($R_{test} = 0.462$), while the SVR-IWO model provides the highest correlation ($R_{test} = 0.892$).

[Figure 7]

Boxplots (Figure 8) illustrate the variability of observed and predicted sinuosity values. The SVR-IWO model shows the range (IQR = 0.336) that is the closest to the range of the observed data (IQR = 0.301) in the training stage. The same applies in the testing stage, with modeled IQR of 0.287 against an observed IQR of 0.308.

[Figure 8]

A sensitive analysis is conducted to evaluate the influence of each input variable on the variation of the output variable. In this regard, a new index named relative coefficient variation (RCV) presented by (Tafarojnoruz and Sharafati, 2020) is used in this study. The RCV is expressed as follows:

$$RCV(x_i, y) = \frac{CV_y}{CV_{x_i}} \tag{23}$$

Where x_i is the i[th] input variable, y is the output variable, and CV is the coefficient of determination. The input variable with the highest RCV is the impactful variable. Results show that the RCV values of S_v, d, w, Q_{bf}, D_{50}, ω_v and $\omega_{v,t}$ are 0.16, 0.18, 0.04, 0.01, 0.15, 0.15, 0.42, respectively. Hence, the potential specific stream power at the transition between single and thread channel (RCV=0.425) is most impactful variable.

3.3 Uncertainty Analysis

The model uncertainty (epistemic uncertainty) is assessed for the three different models considered (SVR-IWO, standalone SVR, and RBF), each for their best performing input variable combination. The Entropy Package within the R software is used to compute entropy using the ML, Jeffrey, Laplace, SG, and Minimax methods (Hausser et al., 2015).

Results are summarized in Table 6 and visualized in Figure 9. For both training and testing stages, the entropy of the sinuosity values predicted by the SWR-IWO model is the closest to the entropy of the observed values (percent difference of $ML = 0.0443, Jeffery = 0, SG, Minimax\ and\ Laplace 0.0221$) for training stage and ($ML = 0.246, Jeffery = 0.0273, Laplace = 0.0546, SG\ and\ Minimax = 0$)for testing stage). In other words, the information content of the sinuosity set produced by the SVR-IWO model is the most similar to that of the observed sinuosity data. This further confirms that the SVR-IWO model is the most reliable in predicting sinuosity among the three models considered.

[Figure 9]

[Table 6]

Overall, prior studies attempt to determine the sinuosity of channels by dividing thalweg length to length of valley length (Flor et al., 2010). However, this way of determining sinuosity is flawed due to ignoring the impacts of several essential parameters such as river bed material properties and river discharge (Van den Berg, 1995). These parameters play vital roles in the sinuosity of channels. However, this study attempts to present a novel model for determining channel sinuosity, including the parameters mentioned above and applicable for different rivers worldwide. Moreover, we employed new methods for assessing channel sinuosity in various case studies with different flow and bed material conditions. The results revealed the employed models are reliable for evaluating channel sinuosity in other case studies.

4. Conclusion

In this paper, a Support Vector Regression (SVR) model is combined with a metaheuristic model, the Invasive Weed Optimization (IWO), to estimate sinuosity in perennial rivers and compared with standalone SVR and RBF models. Seven parameters for prediction are considered $(S_v, d, w, Q_{bf}, D_{50}, \omega_v, \omega_{v,t})$, with seven possible input combinations (M1 to M7), obtained via Bayesian Mutual Information analysis, which identifies the degree of dependency channel sinuosity on the seven aforementioned parameters. The dataset from Van den Berg (1995) is used for the analysis.

Three prediction performance indices (RMSE, MAE, and R) are employed to assess the accuracy of the estimators in both training and testing stages, in addition to diagrams (heat map, scatter plots, and boxplots).

Overall, the SVR-IWO model, with M2 input parameter combination ($d, w, Q_{bf}, D_{50}, \omega_v, \omega_{v,t}$) shows the best prediction performance, both in training stage (MAE = 0.037, RMSE = 0.072 and R = 0.959) and testing stage (MAE = 0.103, RMSE = 0.103 and R = 0.892) among the models considered in this study. In addition, the uncertainty analysis was conducted by using different types of entropies. The results revealed that the entropy values obtained by the SWR-IWO model are close to the entropy of the observations in both training and testing stages.

5. Declarations

Funding : No funding.

Competing interests: The authors declare that they have no competing interests.

Availability of data and materials : Please contact the corresponding author for data requests.

Code availability: Please contact the corresponding author for code requests.

Ethics approval: Not applicable**.**

Consent to participate: Not applicable**.**

Consent for publication: Not applicable**.**

Authors' contributions:

Masoud Haghbin proposed the topic, carried out the investigation, modeling and participated in drafting the manuscript. **Ahmad Sharafati** participated in coordination, aided in the interpretation of results, and paper editing. **Davide Motta** helped in data gathering**,** carried out the visualization and paper editing. All authors read and approved the final manuscript.

Acknowledgements : Not applicable

References

Ahmed, J., Constantine, J.A., Dunne, T., 2019. The role of sediment supply in the adjustment of channel sinuosity across the Amazon Basin. Geology 47, 807–810.

Archer, E., Park, I.M., Pillow, J.W., 2013. Bayesian and quasi-Bayesian estimators for mutual information from discrete data. Entropy 15, 1738–1755.

Baghalian, S., Bonakdari, H., Nazari, F., Fazli, M., 2012. Closed-form solution for flow field in curved channels in comparison with experimental and numerical analyses and artificial neural network. Eng. Appl. Comput. Fluid Mech. 6, 514–526.

Beechie, T., Imaki, H., 2014. Predicting natural channel patterns based on landscape and geomorphic controls in the Columbia River basin, USA. Water Resour. Res. 50, 39–57.

Bonakdari, H., Gholami, A., Gharabaghi, B., 2019. Modelling Stable Alluvial River Profiles Using Back Propagation-Based Multilayer Neural Networks, in: Intelligent Computing-Proceedings of the Computing Conference. Springer, pp. 607–624.

Chen, W., Liu, W., Huang, W., Liu, H., 2016. Prediction of salinity variations in a tidal estuary using artificial neural network and three-dimensional hydrodynamic models. Comput. Water, Energy, Environ. Eng. 6, 107–128.

Deike, G.H., White, W.B., 1969. Sinuosity in limestone solution conduits. Am. J. Sci. 267, 230–241.

Dente, E., Lensky, N.G., Morin, E., Dunne, T., Enzel, Y., 2019. Sinuosity evolution along an incising channel: New insights from the Jordan River response to the Dead Sea level fall. Earth Surf. Process. Landforms 44, 781–795.

Diop, L., Samadianfard, S., Bodian, A., Yaseen, Z.M., Ghorbani, M.A., Salimi, H., 2020. Annual rainfall forecasting using hybrid artificial intelligence model: integration of multilayer perceptron with whale optimization algorithm. Water Resour. Manag. 34, 733–746.

Fattahi, H., Babanouri, N., 2017. Applying optimized support vector regression models for prediction of tunnel boring machine performance. Geotech. Geol. Eng. 35, 2205–2217.

Ferguson, R.I., 1977. Meander sinuosity and direction variance. Geol. Soc. Am. Bull. 88, 212–214.

Flor, A., Pinter, N., Remo, J.W.F., 2010. Evaluating levee failure susceptibility on the Mississippi River using logistic regression analysis. Eng. Geol. 116, 139–148.

Gholami, A., Bonakdari, H., Akhtari, A.A., Ebtehaj, I., 2019a. A combination of computational fluid dynamics, artificial neural network and support vectors machines model to predict flow variables in curved channel. Sci. Iran. 26, 726–741.

Gholami, A., Bonakdari, H., Ebtehaj, I., Gharabaghi, B., Khodashenas, S.R., Talesh, S.H.A., Jamali, A., 2018. A methodological approach of predicting threshold channel bank profile by multi-objective evolutionary optimization of ANFIS. Eng. Geol. 239, 298–309.

Gholami, A., Bonakdari, H., Ebtehaj, I., Khodashenas, S.R., 2020. Reliability and sensitivity analysis of robust learning machine in prediction of bank profile morphology of threshold sand rivers. Measurement 153, 107411.

Gholami, A., Bonakdari, H., Samui, P., Mohammadian, M., Gharabaghi, B., 2019b. Predicting stable alluvial channel profiles using emotional artificial neural networks. Appl. Soft Comput. 78, 420–437.

Ghosh, P., 2000. Estimation of channel sinuosity from paleocurrent data: a method using fractal geometry. J. Sediment. Res. 70, 449–455.

Haghbin, M., Sharafati, A., Dixon, B., Kumar, V., 2020. Application of Soft Computing Models for Simulating Nitrate Contamination in Groundwater: Comprehensive Review, Assessment and Future Opportunities. Arch. Comput. Methods Eng. 1–23.

Haghbin, M., Sharafati, A., Motta, D., Al-Ansari, N., Noghani, M.H.M., 2021. Applications of soft computing models for predicting sea surface temperature: a comprehensive review and assessment. Prog. Earth Planet. Sci. 8, 1–19.

Ham, F., Kostanic, I., 2001. Fundamental neurocomputing concepts. Principles of Neurocomputing for Science and Engineering.

Hausser, J., Strimmer, K., Strimmer, M.K., 2015. Package 'entropy.'

Hlaváčková-Schindler, K., Paluš, M., Vejmelka, M., Bhattacharya, J., 2007. Causality detection based on information-theoretic approaches in time series analysis. Phys. Rep. 441, 1–46.

Hooke, R.L.B., 1975. Distribution of sediment transport and shear stress in a meander bend. J. Geol. 83, 543–565.

Hutter, M., 2002. Distribution of mutual information, in: Advances in Neural Information Processing Systems. pp. 399–406.

Hutter, M., Zaffalon, M., 2002. Distribution of mutual information for robust feature selection.

Jamei, M., Ahmadianfar, I., 2020. A rigorous model for prediction of viscosity of oil-based hybrid nanofluids. Phys. A Stat. Mech. Its Appl. 556, 124827.

Jamei, M., Ahmadianfar, I., Chu, X., Yaseen, Z.M., 2020. Prediction of surface water total dissolved solids using hybridized wavelet-multigene genetic programming: New approach. J. Hydrol. 589, 125335.

Javaheri, N., Ghomeshi, M., Kashefipour, S.M., 2008. Use of the fuzzy method for determination of sediment balance and its role on the morphological changes in meandering rivers. Asian J. Sci. Res. 4, 32–40.

Kleinhans, M.G., van den Berg, J.H., 2011. River channel and bar patterns explained and predicted by an empirical and a physics-based method. Earth Surf. Process. Landforms 36, 721–738.

Le Roux, J.P., 1992. Determining the channel sinuosity of ancient fluvial systems from paleocurrent data. J. Sediment. Res. 62, 283–291.

Leopold, L.B., Wolman, M.G., 1957. River channel patterns: braided, meandering, and straight. US Government Printing Office.

Lewin, J., Brewer, P.A., 2001. Predicting channel patterns. Geomorphology 40, 329–339.

Li, M.M., Verma, B., Fan, X., Tickle, K., 2008. RBF neural networks for solving the inverse problem of backscattering spectra. Neural Comput. Appl. 17, 391–397.

Liu, L., Xin, J., Feng, Y., Zhang, B., Song, K.-I., 2019. Effect of the Cement–Tailing Ratio on the Hydration Products and Microstructure Characteristics of Cemented Paste Backfill. Arab. J. Sci. Eng. 1–10.

Mehrabian, A.R., Lucas, C., 2006. A novel numerical optimization algorithm inspired from weed colonization. Ecol. Inform. 1, 355–366.

Moharana, S., Khatua, K.K., 2014. Prediction of roughness coefficient of a meandering open channel flow using Neuro-Fuzzy Inference System. Measurement 51, 112–123.

Moody-Stuart, M., 1966. High-and low-sinuosity stream deposits, with examples from the Devonian of Spitsbergen. J. Sediment. Res. 36, 1102–1117.

Naganna, S.R., Deka, P.C., Ghorbani, M.A., Biazar, S.M., Al-Ansari, N., Yaseen, Z.M., 2019. Dew point temperature estimation: application of artificial intelligence model integrated with nature-inspired optimization algorithms. Water 11, 742.

Nanson, G.C., Hickin, E.J., 1983. Channel migration and incision on the Beatton River. J. Hydraul. Eng. 109, 327–337.

Pourrajab, R., Ahmadianfar, I., Jamei, M., Behbahani, M., 2020. A meticulous intelligent approach to predict thermal conductivity ratio of hybrid nanofluids for heat transfer applications. J. Therm. Anal. Calorim. 1–18.

Rényi, A., 1959. On measures of dependence. Acta Math. Acad. Sci. Hungarica 10, 441–451.

Riahi-Madvar, H., Ayyoubzadeh, S.A., Atani, M.G., 2011. Developing an expert system for predicting alluvial channel geometry using ANN. Expert Syst. Appl. 38, 215–222.

Sahu, M., Jana, S., Agarwal, S., Khatua, K.K., Mohapatra, S., 2011. Point form velocity prediction in meandering open channel using artificial neural network, in: 2nd International Conference on Environmental Science and Technology. pp. 209–212.

Schumm, S.A., 1963. Sinuosity of alluvial rivers on the Great Plains. Geol. Soc. Am. Bull. 74, 1089–1100.

Shaghaghi, S., Bonakdari, H., Gholami, A., Ebtehaj, I., Zeinolabedini, M., 2017. Comparative analysis of GMDH neural network based on genetic algorithm and particle swarm optimization in stable channel design. Appl. Math. Comput. 313, 271–286. https://doi.org/10.1016/j.amc.2017.06.012

Shannon, C.E., Weaver, W., 1949. A mathematical model of communication. Urbana, Univ. Illinois Press 11.

Sharafati, A., Haghbin, M., Haji Seyed Asadollah, S.B., Tiwari, N.K., Al-Ansari, N., Yaseen, Z.M., 2020. Scouring Depth Assessment Downstream of Weirs Using Hybrid Intelligence Models. Appl. Sci. 10, 3714.

Sharafati, A., Haghbin, M., Torabi, M., Yaseen, Z.M., 2021a. Assessment of novel nature-inspired fuzzy models for predicting long contraction scouring and related uncertainties. Front. Struct. Civ. Eng. https://doi.org/10.1007/s11709-021-0713-0

Sharafati, A., Masoud, H., Tiwari, N.K., Bhagat, S.K., Al-Ansari, N., Chau, K.-W., Yaseen, Z.M., 2021b. Performance evaluation of sediment ejector efficiency using hybrid neuro-fuzzy models. Eng. Appl.

Comput. Fluid Mech. 15, 627–643.

Singh, V.P., 2014. Entropy theory in hydraulic engineering: An introduction. American Society of Civil Engineers.

Smith, C.E., 1998. Modeling high sinuosity meanders in a small flume. Geomorphology 25, 19–30.

Tafarojnoruz, A., Sharafati, A., 2020. New formulations for prediction of velocity at limit of deposition in storm sewers based on a stochastic technique. Water Sci. Technol. 81, 2634–2649.

Tahershamsi, A., Tabatabai, M.R.M., Shirkhani, R., 2012. An evaluation model of artificial neural network to predict stable width in gravel bed rivers. Int. J. Environ. Sci. Technol. 9, 333–342.

Tao, H., Al-Bedyry, N.K., Khedher, K.M., Shahid, S., Yaseen, Z.M., 2021a. River water level prediction in coastal catchment using hybridized relevance vector machine model with improved grasshopper optimization. J. Hydrol. 598, 126477.

Tao, H., Habib, M., Aljarah, I., Faris, H., Afan, H.A., Yaseen, Z.M., 2021b. An intelligent evolutionary extreme gradient boosting algorithm development for modeling scour depths under submerged weir. Inf. Sci. (Ny). 570, 172–184.

Van den Berg, J.H., 1995. Prediction of alluvial channel pattern of perennial rivers. Geomorphology 12, 259–279.

Waszczyszyn, Z., 2010. Advances of Soft Computing in Engineering. Springer Science & Business Media.

Woolderink, H.A.G., Cohen, K.M., Kasse, C., Kleinhans, M.G., Van Balen, R.T., 2021. Patterns in river channel sinuosity of the Meuse, Roer and Rhine rivers in the Lower Rhine Embayment rift-system, are they tectonically forced? Geomorphology 375, 107550.

Yang, F., Paindavoine, M., 2003. Implementation of an RBF neural network on embedded systems: realtime face tracking and identity verification. IEEE Trans. Neural Networks 14, 1162–1175.

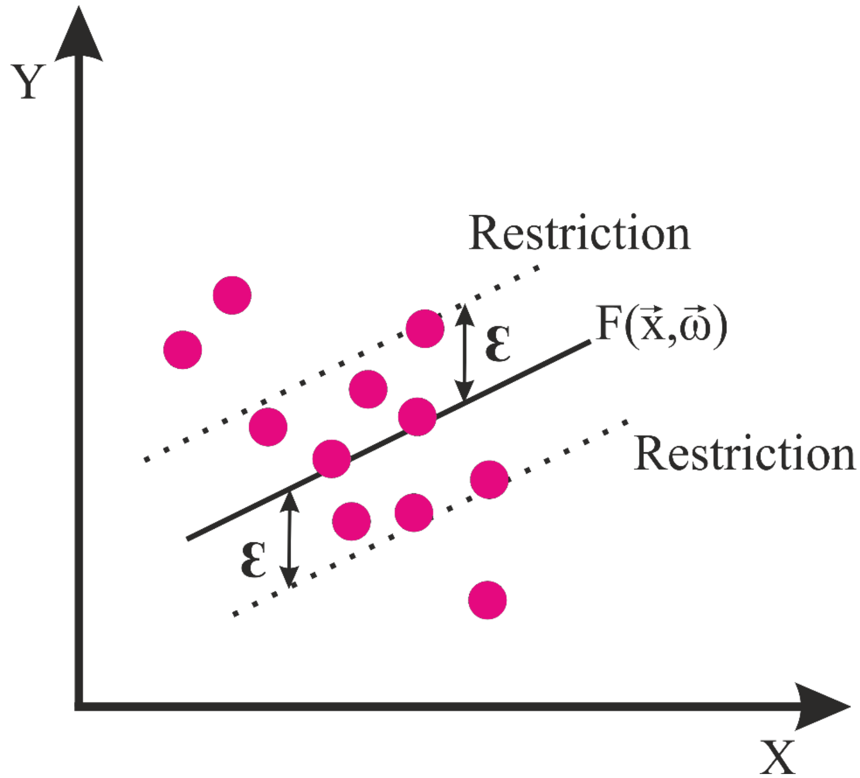


Figure 1: Schematic view of linear regression in Support Vector Regression (SVR).

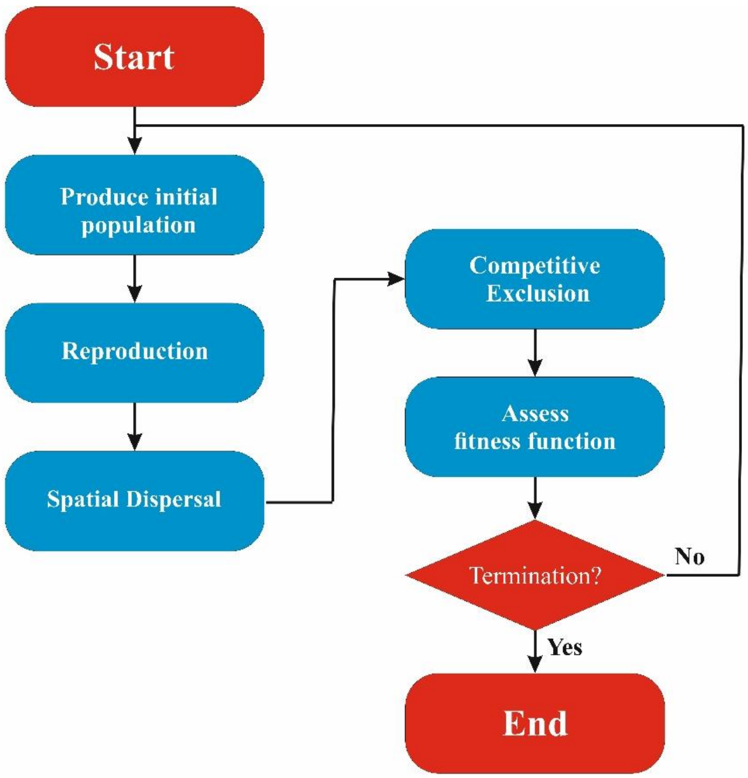


Figure 2: Flow chart of Invasive Weed Optimization (IWO).

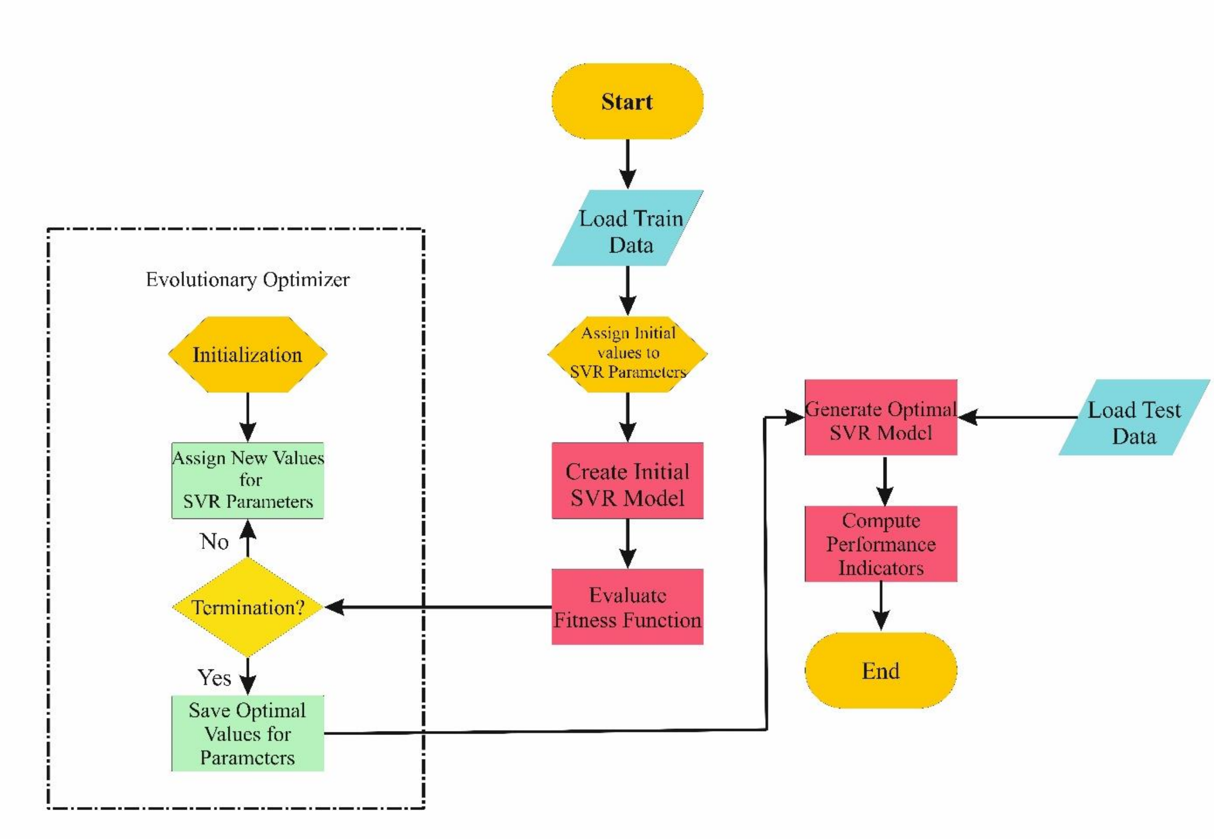


Figure 3: Flow chart of the optimization of the SVR model parameters.

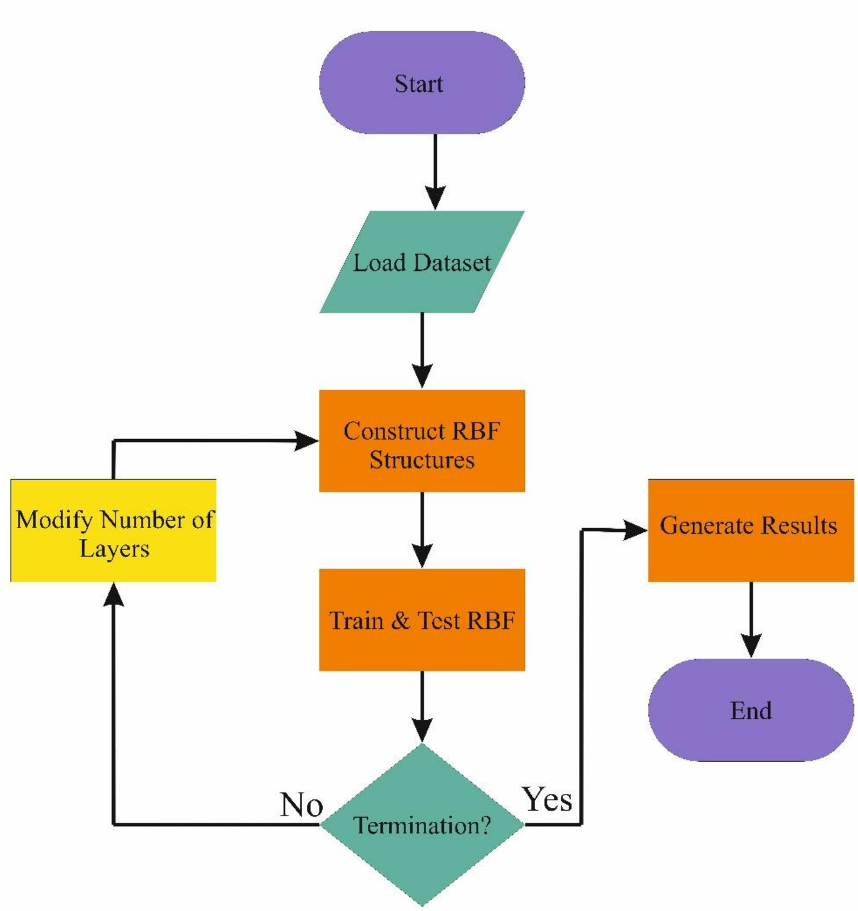


Figure 4: Flow chart of a Radial Basis Functions Neural Network (RBFNN).

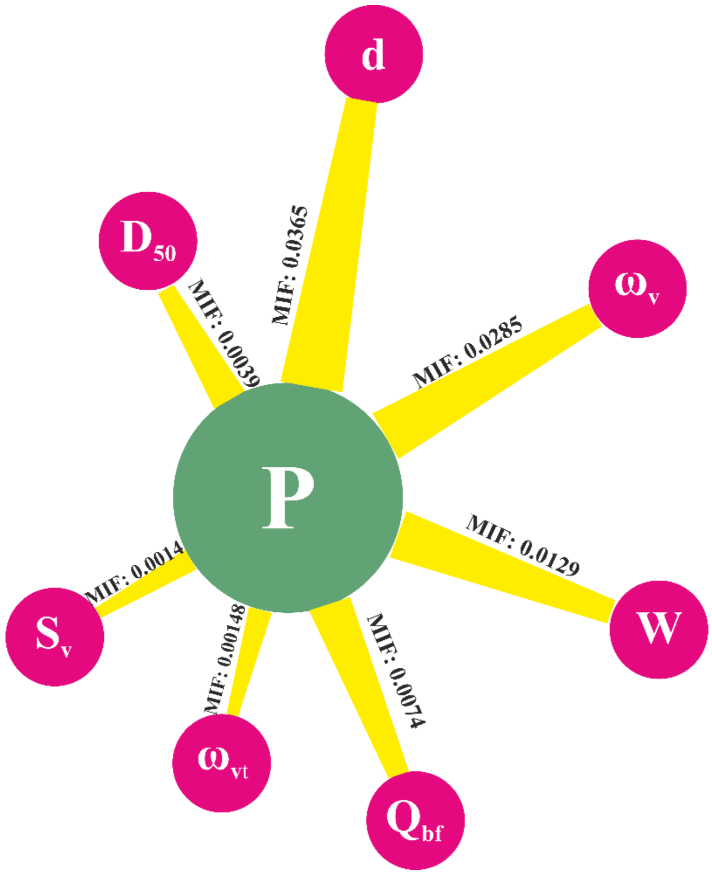


Figure 5: Mutual Information (MIF) between inputs and output (sinuosity).

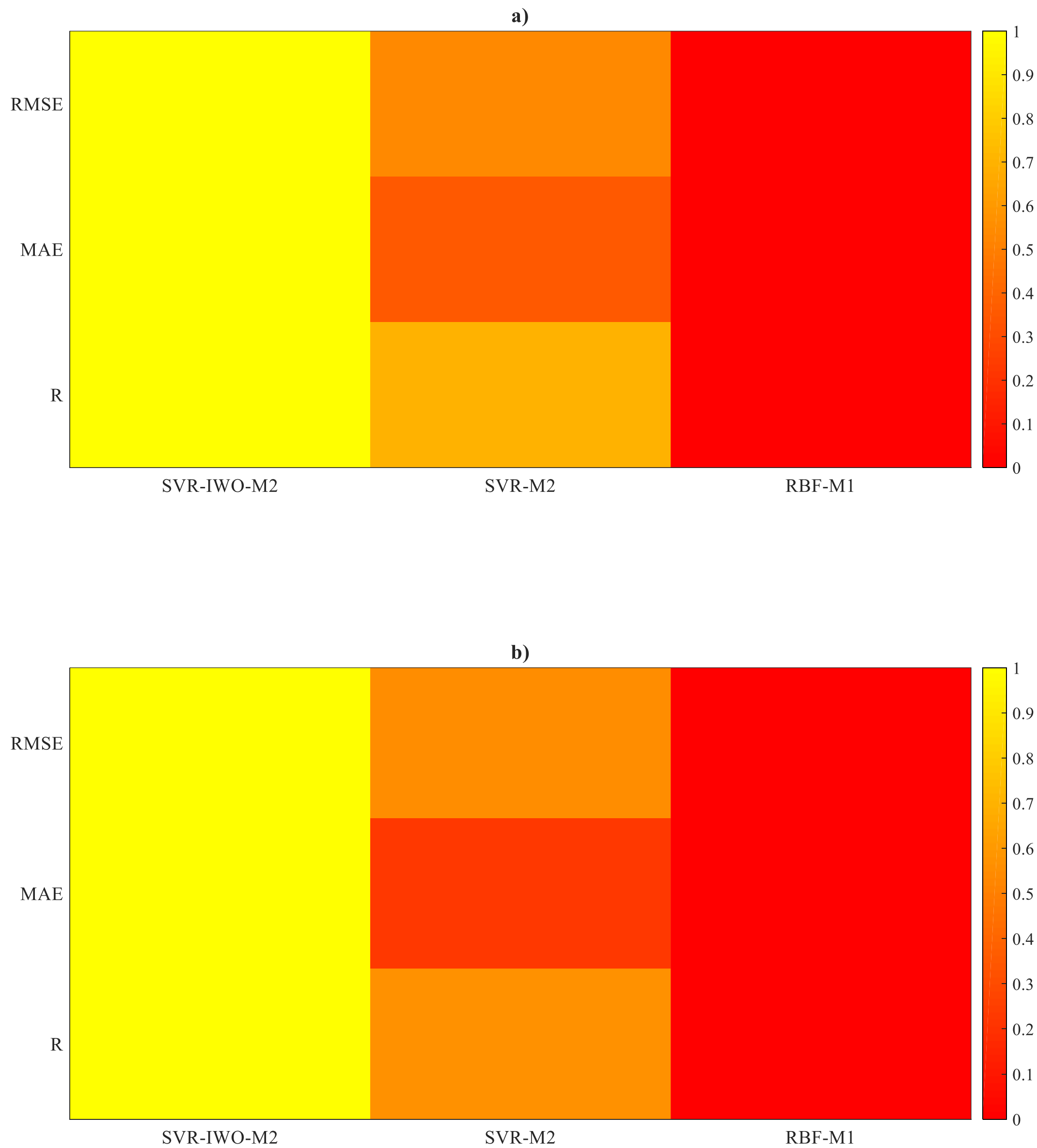


Figure 6: Performance heat map for (a) training stage and (b) testing stage.

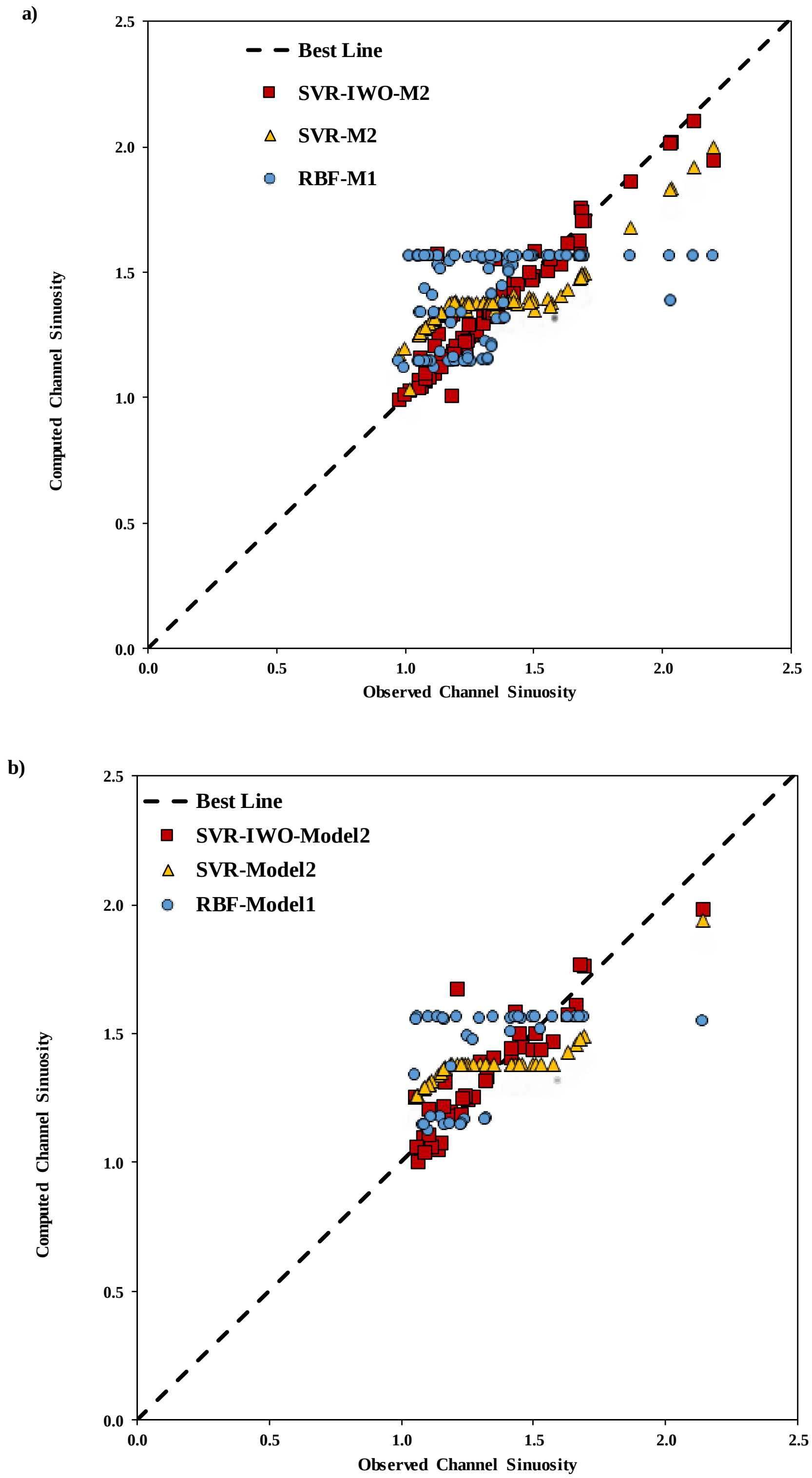


Figure 7: Scatter plot for (a) training stage and (b) testing stages.

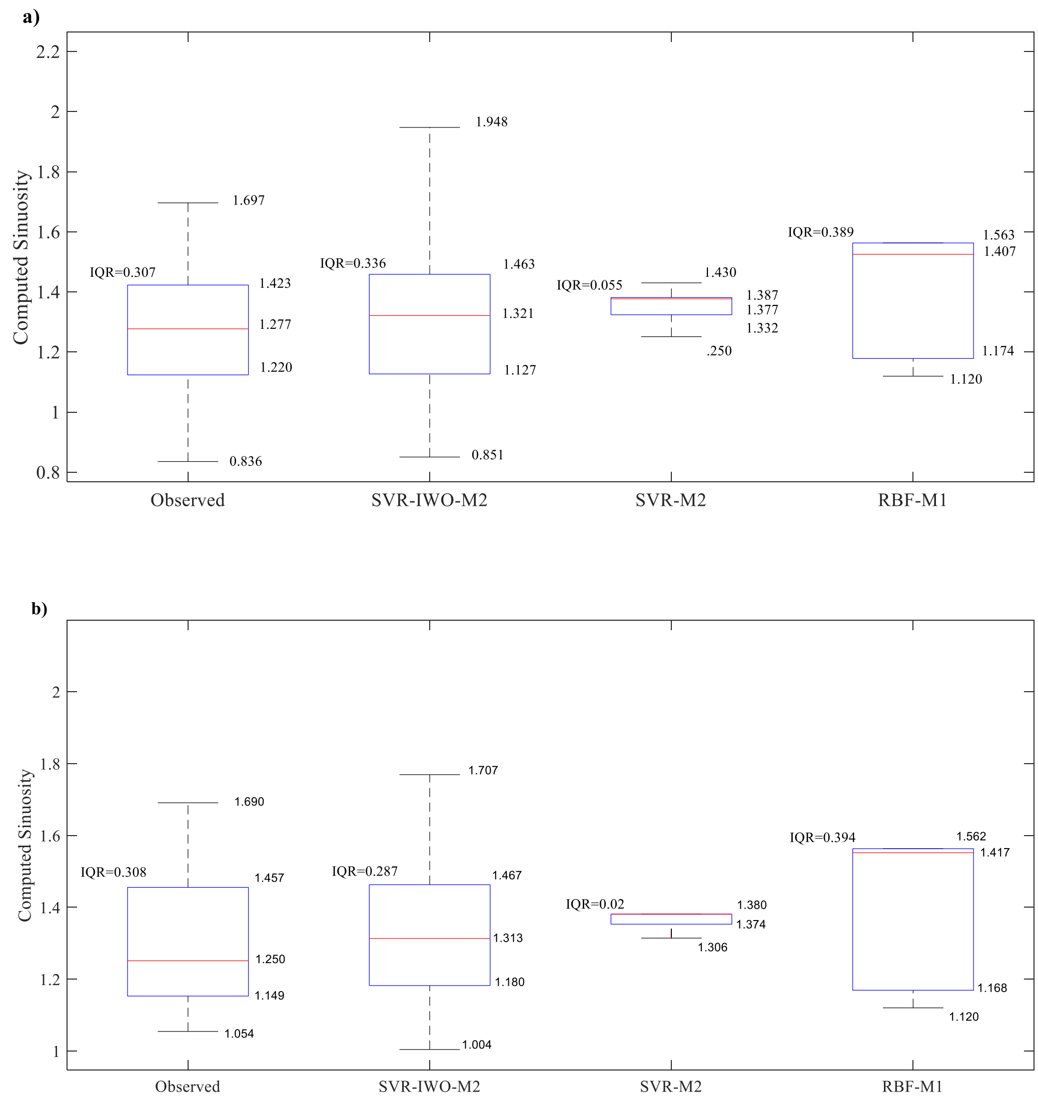


Figure 8: Boxplot for (a) training stage and (b) testing stage.

a)

b)

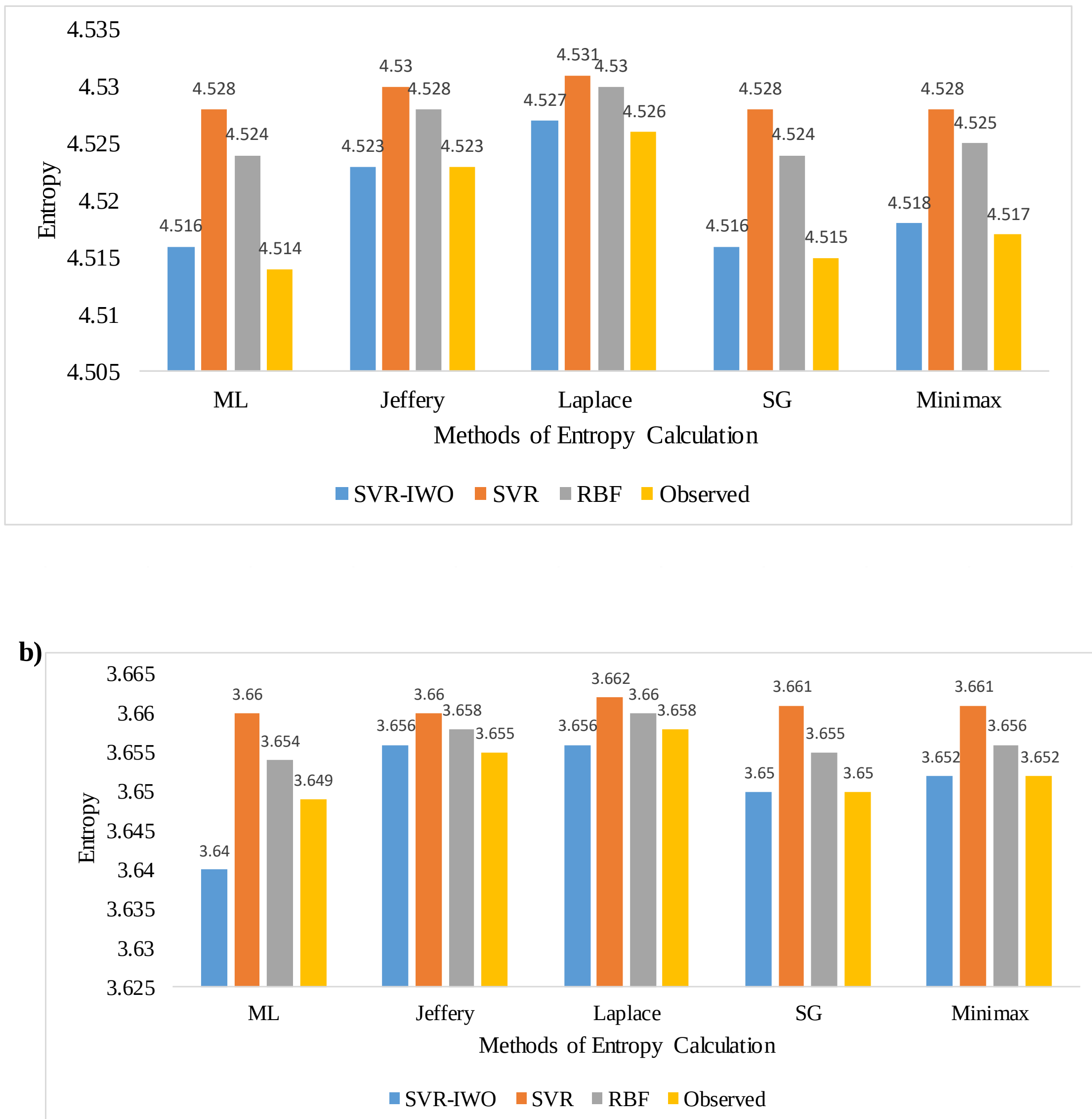


Figure 9: Model uncertainty for (a) training stage and (b) testing stage.

Table 1: Streams included in the channel sinuosity dataset.

Country	Number of Streams	Stream Names
Australia	8	Barwon, Butchers, Clarence, Derwert, S. Ashburton, Stanley, Swan, Wye
Canada	23	Dry Wood, East Prairie, Forks, Fraser, Grey, Assiniboine, Athabasca, Bow, Chalk, Crows Nest, Klondike, Knee Hills, Lesser Slave, Little Paddle, Little Red Deer, Little Smokey, Medicine, N. Saskatchewan, Paddle, Pembina, Pipestone, Vermillion, West Prairie
Colombia	1	Sinu
England	16	Asker, Ceiriog, Chittern, Churnet, Eden, Esk, Hamps Kielderburn, Lugg, Pinsley, Ribble, Rookhope Severn, Teign, Wylye, Wyre
European Countries	1	Danube
Ireland	1	Irishman
New Zealand	41	Ahuriri, Arrow, Buller, Cobb, Gimmerburn, Glen Roy, Gowan, Hakataramea, Hokitika, Hurunui, Inangahua, Kakanui, Kyeburn, Loganburn, Maerwhenua, Makarewa, Mangles, Mataura, Maruia, Maryburn, Matakitaki, Mataura, Motueka, Nobles, Ohau, Opihi, Opuha, Orari, Otapiri, Otekaieke, Otematata, Pomahaka, Rangitata, Riwaka, Selwyn, Taieri, Twizel, Waiau-uha, Waihopai, Waimakiriri, Wairau
Scotland	1	Eachaig
South Africa	1	Thomas
United states	23	Beaver, Blacks Fork, Branch, Learwater, Crystal, East Inlet, Elk, Glen, Halfmoon, Lake Fork, Little Grizzly, Little Muddy, Little Snake, Lleddam, Mississippi, North Platte, Otter, Stillwater, Taylor, Tomi Chi, White, William Fork, Willow
Wales	3	Afon, Lwyd, Irfon, Usk

Table 2: Range of values for the seven parameters considered for the estimation of sinuosity.

Parameter	Training Stage	Testing Stage
$w\ (m)$	5.21-375	4.2-390
$d\ (m)$	0.29-5	0.33-7.2
$Q_{bf}\ (\frac{m^3}{S})$	1.87-1372	1.71-1200
$D_{50}\ (mm)$	0.27-145	0.15-118.7
$S_v \times 1000$	0.22-22	0.05-28.3
$\omega_v\ (\frac{kW}{m^2})$	0.004-0.915	0.00155-0.940
$\omega_{v,t}(\frac{kW}{m^2})$	28.537-399.96	22.294-446.752
P	1-2.196	1.0545-2.142

Table 3: Mutual Information between inputs and output (sinuosity).

Input Variable	Mutual Information
S_V	0.0014
D_{50}	0.0039
ω_v	0.0285
d	0.0365
Q_{bf}	0.0074
$\omega_{v,t}$	0.00148
w	0.0129

Table 4: Input variable combinations considered for channel sinuosity prediction, selected based on Bayesian Mutual Information theory.

	S_V	D_{50}	ω_v	d	Q_{bf}	$\omega_{v,t}$	w
M1	✓	✓	✓	✓	✓	✓	✓
M2	-	✓	✓	✓	✓	✓	✓
M3	-	✓	✓	✓	✓	-	✓
M4	-	-	✓	✓	✓	-	✓
M5	-	-	✓	✓	-	-	✓
M6	-	-	✓	✓	-	-	-
M7	-	-	-	✓	-	-	-

Table 5: Summary of the prediction performance indices for the model considered.

Training Stage			
Model	**RMSE**	**MAE**	**R**
SVR-IWO-M2	**0.072**	**0.037**	**0.959**
SVR-M2	0.158	0.141	0.792
RBF-M1	0.260	0.198	0.404
Testing Stage			
Model	**RMSE**	**MAE**	**R**
SVR-IWO-M2	**0.103**	**0.065**	**0.892**
SVR-M2	0.163	0.151	0.704
RBF-M1	0.236	0.177	0.462

Table 6: Model uncertainty for training and testing stages. Entropy values are shown.

Training Stage					
Model	**ML**	**Jeffrey**	**Laplace**	**SG**	**Minimax**
SVR-IWO	4.516	4.523	4.527	4.516	4.518
SVR	4.528	4.53	4.531	4.528	4.528
RBF	4.524	4.528	4.53	4.524	4.525
Observed	4.514	4.523	4.526	4.515	4.517
Testing Stage					
Model	**ML**	**Jeffrey**	**Laplace**	**SG**	**Minimax**
SVR-IWO	3.640	3.656	3.656	3.65	3.652
SVR	3.660	3.660	3.662	3.661	3.661
RBF	3.654	3.658	3.660	3.655	3.656
Observed	3.649	3.655	3.658	3.650	3.652

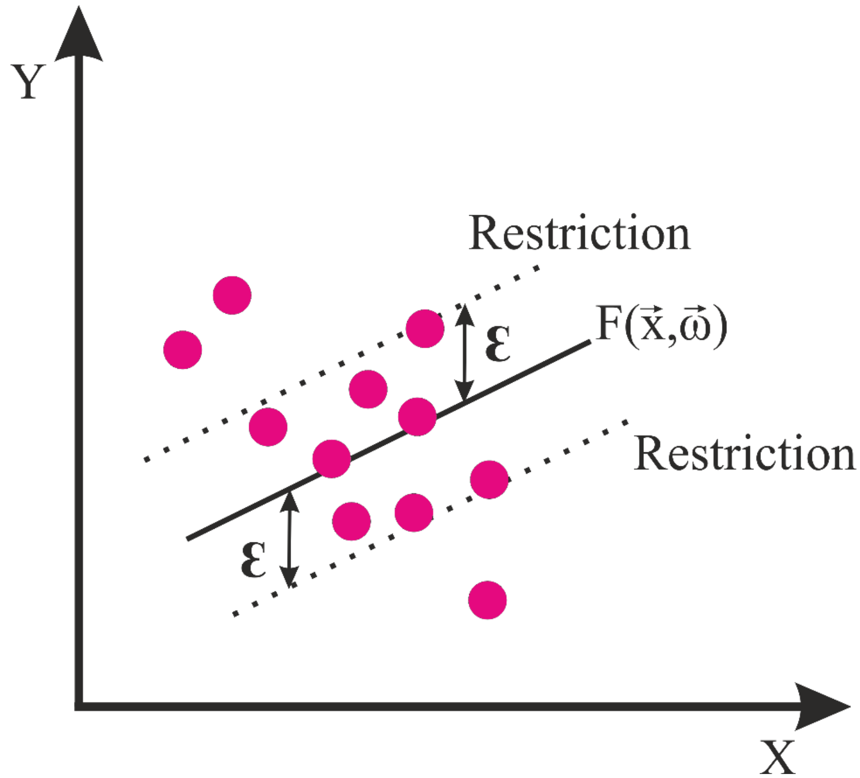


Figure 1: Schematic view of linear regression in Support Vector Regression (SVR).

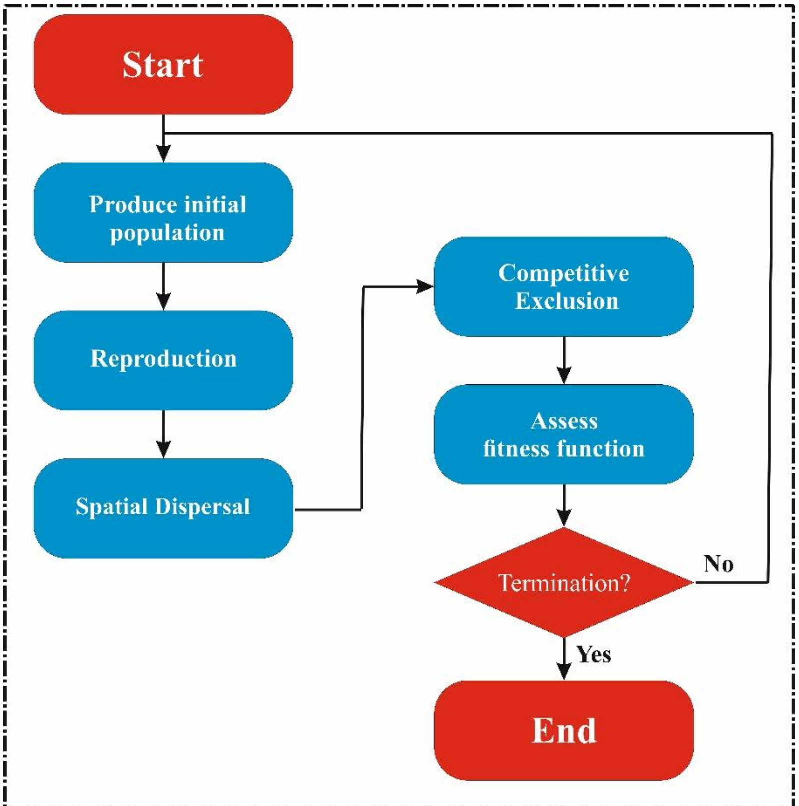


Figure 2: Flow chart of Invasive Weed Optimization (IWO).

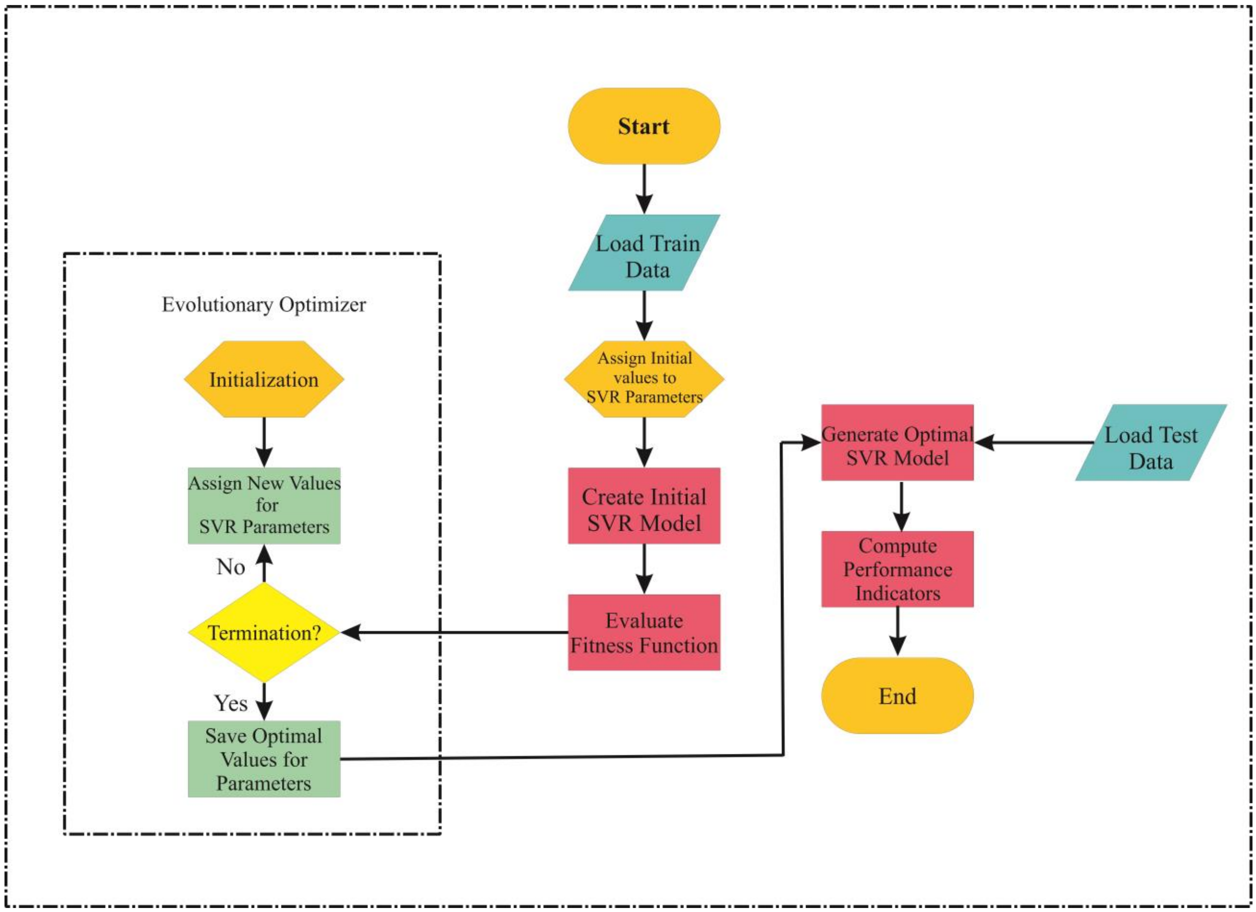


Figure 3: Flow chart of the optimization of the SVR model parameters.

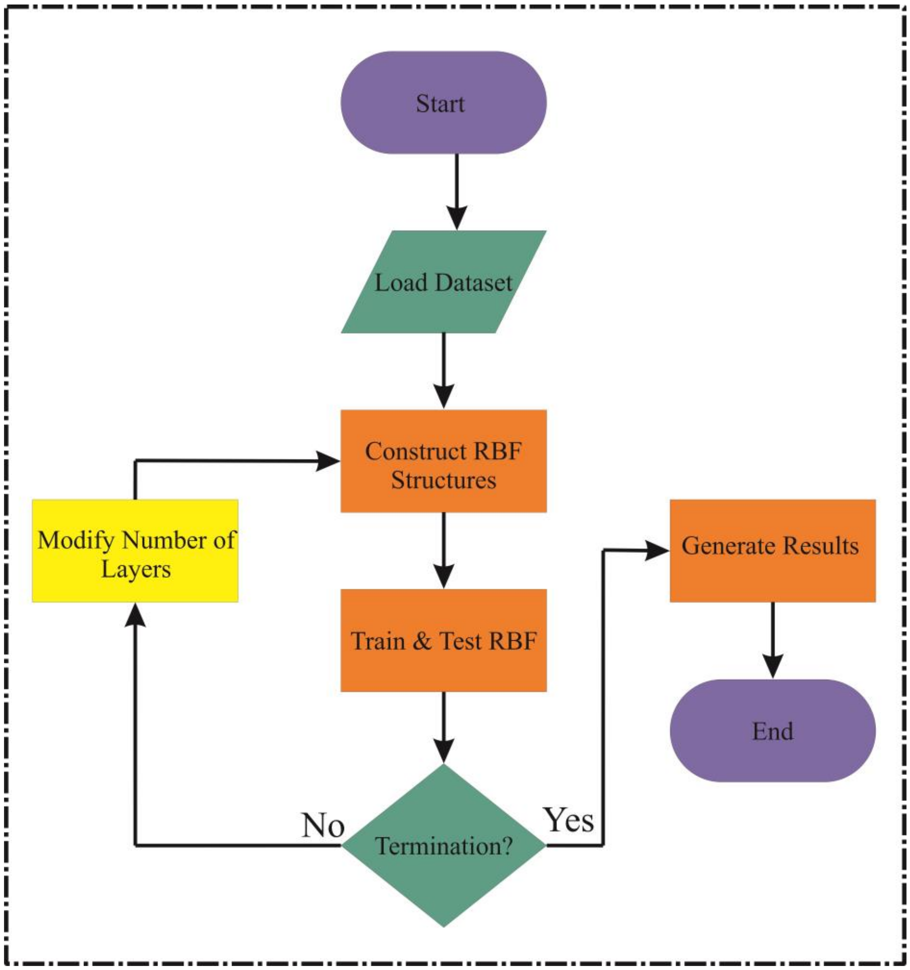


Figure 4: Flow chart of a Radial Basis Functions Neural Network (RBFNN).

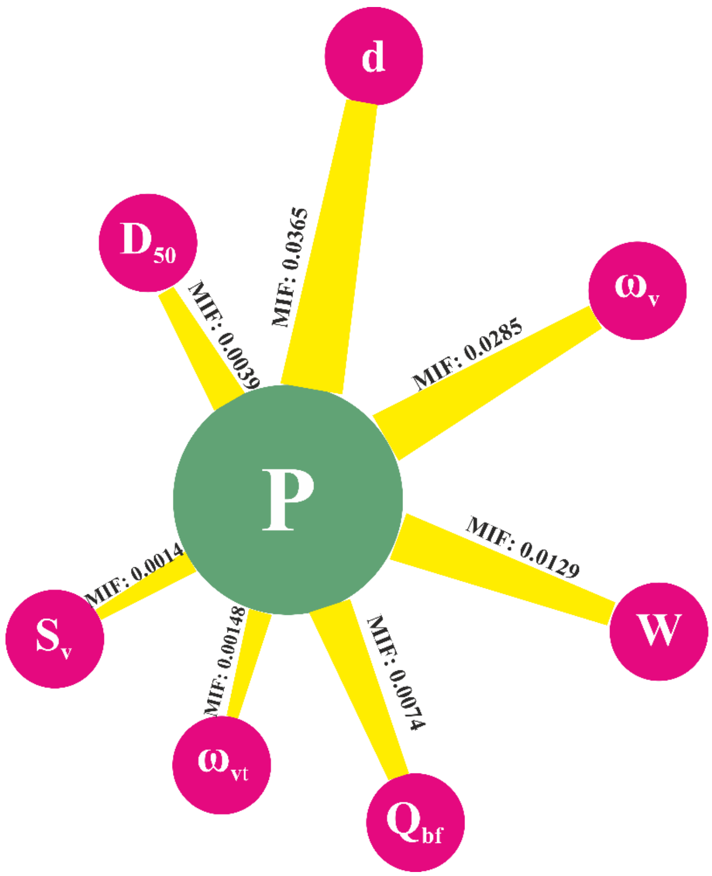


Figure 5: Mutual Information (MIF) between inputs and output (sinuosity).

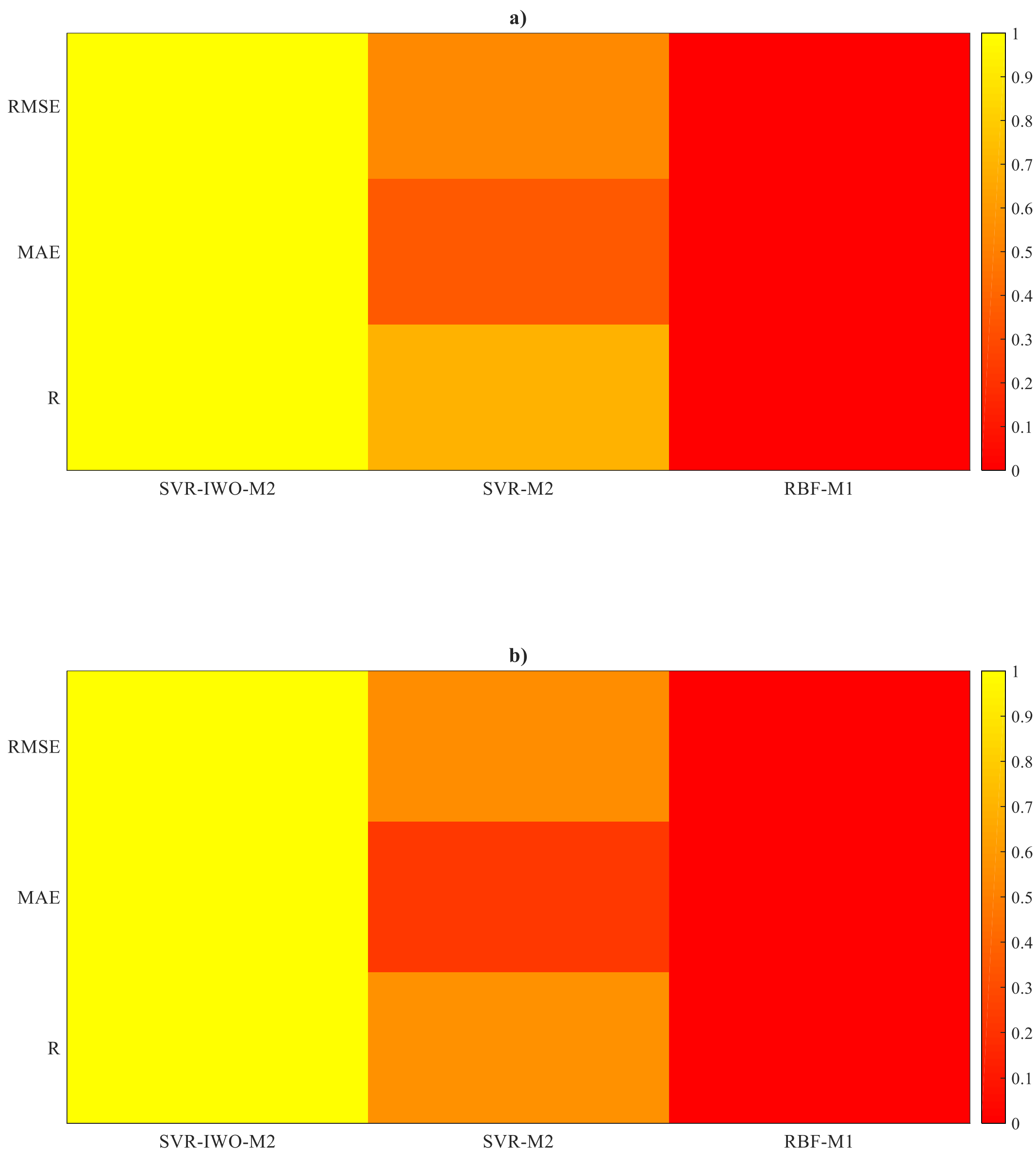


Figure 6: Performance heat map for (a) training stage and (b) testing stage.

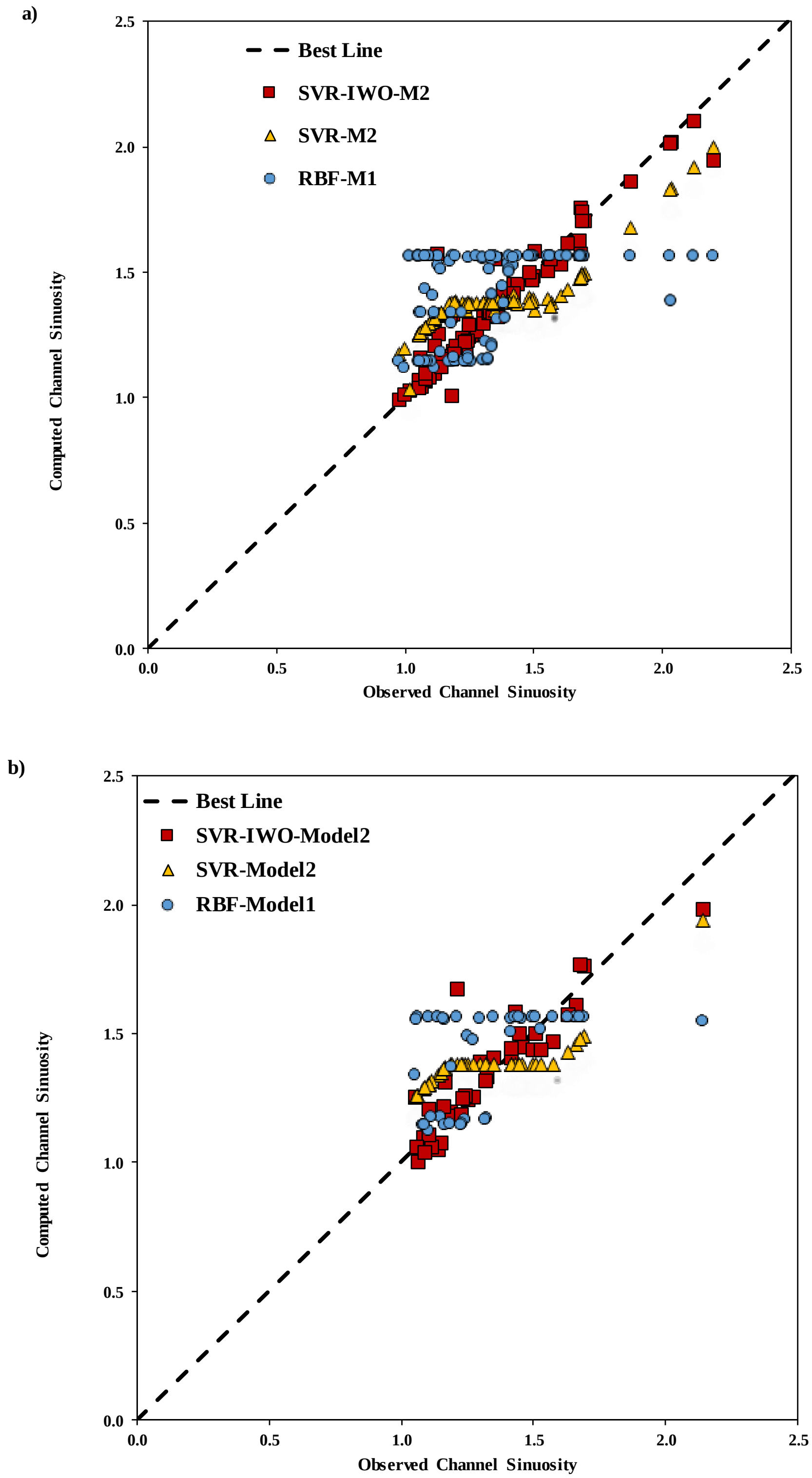
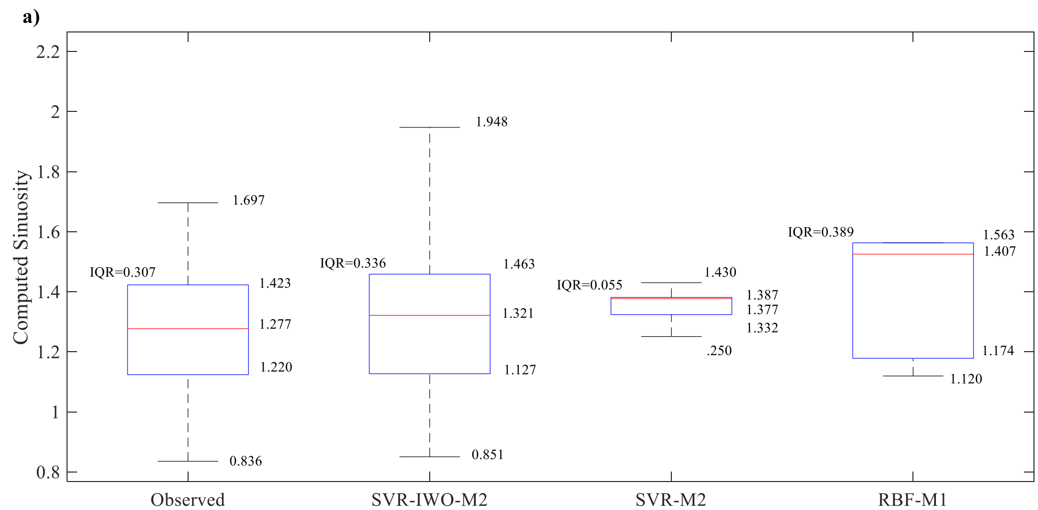


Figure 7: Scatter plot for (a) training stage and (b) testing stages.



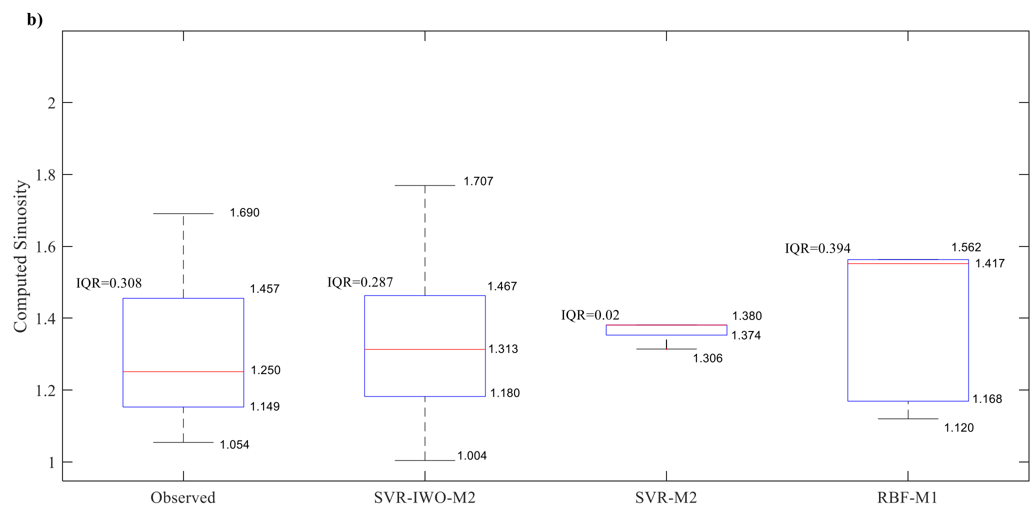


Figure 8: Boxplot for (a) training stage and (b) testing stage.

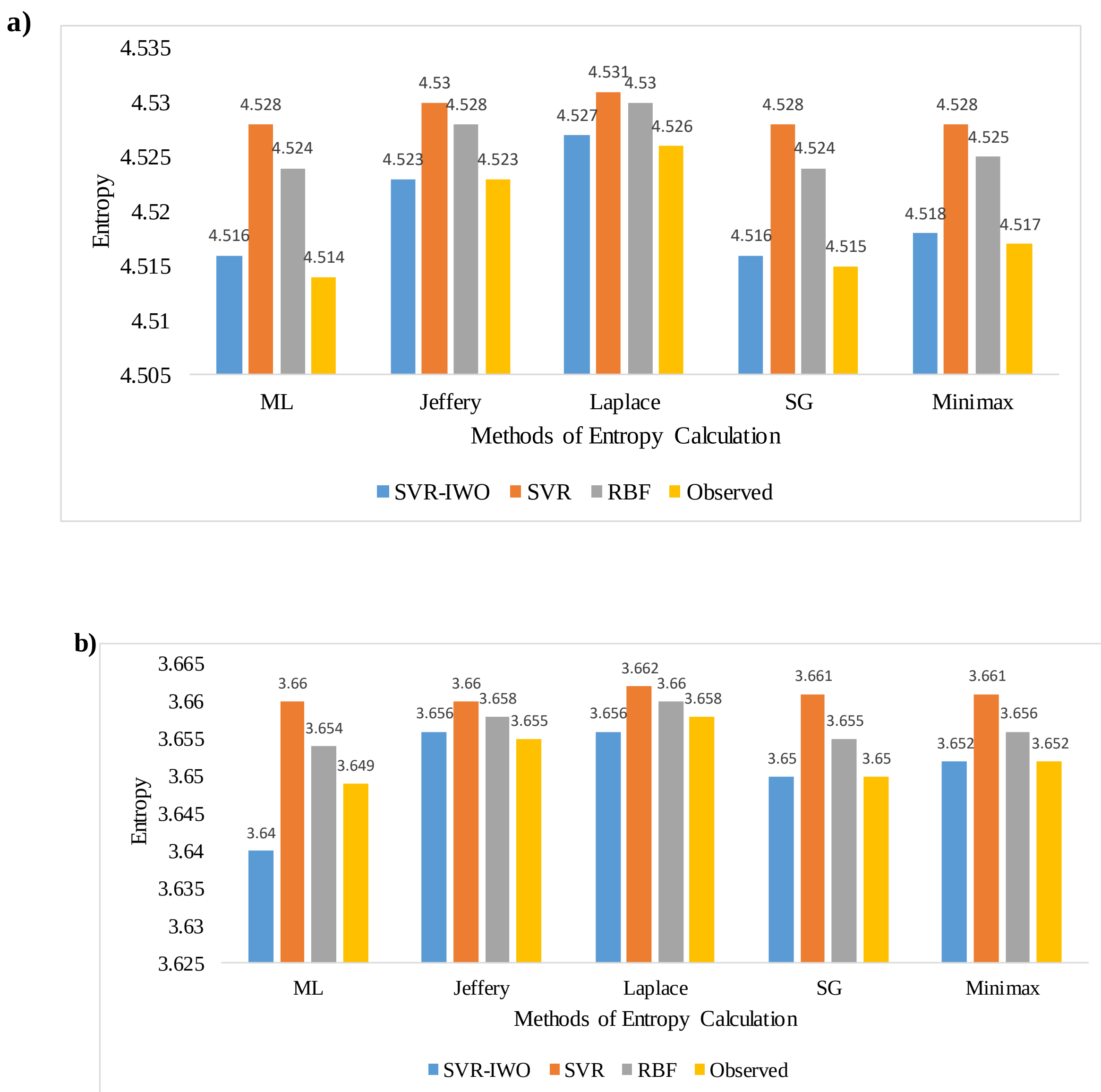


Figure 9: Model uncertainty for (a) training stage and (b) testing stage.

Table 1: Streams included in the channel sinuosity dataset.

Country	Number of Streams	Stream Names
Australia	8	Barwon, Butchers, Clarence, Derwert, S. Ashburton, Stanley, Swan, Wye
Canada	23	Dry Wood, East Prairie, Forks, Fraser, Grey, Assiniboine, Athabasca, Bow, Chalk, Crows Nest, Klondike, Knee Hills, Lesser Slave, Little Paddle, Little Red Deer, Little Smokey, Medicine, N. Saskatchewan, Paddle, Pembina, Pipestone, Vermillion, West Prairie
Colombia	1	Sinu
England	16	Asker, Ceiriog, Chittern, Churnet, Eden, Esk, Hamps Kielderburn, Lugg, Pinsley, Ribble, Rookhope Severn, Teign, Wylye, Wyre
European Countries	1	Danube
Ireland	1	Irishman
New Zealand	41	Ahuriri, Arrow, Buller, Cobb, Gimmerburn, Glen Roy, Gowan, Hakataramea, Hokitika, Hurunui, Inangahua, Kakanui, Kyeburn, Loganburn, Maerwhenua, Makarewa, Mangles, Mataura, Maruia, Maryburn, Matakitaki, Mataura, Motueka, Nobles, Ohau, Opihi, Opuha, Orari, Otapiri, Otekaieke, Otematata, Pomahaka, Rangitata, Riwaka, Selwyn, Taieri, Twizel, Waiau-uha, Waihopai, Waimakiriri, Wairau
Scotland	1	Eachaig
South Africa	1	Thomas
United states	23	Beaver, Blacks Fork, Branch, Learwater, Crystal, East Inlet, Elk, Glen, Halfmoon, Lake Fork, Little Grizzly, Little Muddy, Little Snake, Lleddam, Mississippi, North Platte, Otter, Stillwater, Taylor, Tomi Chi, White, William Fork, Willow
Wales	3	Afon, Lwyd, Irfon, Usk

Table 2: Range of values for the seven parameters considered for the estimation of sinuosity.

Parameter	Training Stage	Testing Stage
$w\ (m)$	5.21-375	4.2-390
$d\ (m)$	0.29-5	0.33-7.2
$Q_{bf}\ (\frac{m^3}{S})$	1.87-1372	1.71-1200
$D_{50}\ (mm)$	0.27-145	0.15-118.7
$S_v \times 1000$	0.22-22	0.05-28.3
$\omega_v\ (\frac{kW}{m^2})$	0.004-0.915	0.00155-0.940
$\omega_{v,t}(\frac{kW}{m^2})$	28.537-399.96	22.294-446.752
P	1-2.196	1.0545-2.142

Table 3: Mutual Information between inputs and output (sinuosity).

Input Variable	**Mutual Information**
$\boldsymbol{S_V}$	0.0014
$\boldsymbol{D_{50}}$	0.0039
$\boldsymbol{\omega_v}$	0.0285
$\boldsymbol{d}$	0.0365
$\boldsymbol{Q_{bf}}$	0.0074
$\boldsymbol{\omega_{v,t}}$	0.00148
$\boldsymbol{w}$	0.0129

Table 4: Input variable combinations considered for channel sinuosity prediction, selected based on Bayesian Mutual Information theory.

	S_V	D_{50}	ω_v	d	Q_{bf}	$\omega_{v,t}$	w
M1	✓	✓	✓	✓	✓	✓	✓
M2	-	✓	✓	✓	✓	✓	✓
M3	-	✓	✓	✓	✓	-	✓
M4	-	-	✓	✓	✓	-	✓
M5	-	-	✓	✓	-	-	✓
M6	-	-	✓	✓	-	-	-
M7	-	-	-	✓	-	-	-

Table 5: Summary of the prediction performance indices for the model considered.

Training Stage			
Model	**RMSE**	**MAE**	**R**
SVR-IWO-M2	**0.072**	**0.037**	**0.959**
SVR-M2	0.158	0.141	0.792
RBF-M1	0.260	0.198	0.404
Testing Stage			
Model	**RMSE**	**MAE**	**R**
SVR-IWO-M2	**0.103**	**0.065**	**0.892**
SVR-M2	0.163	0.151	0.704
RBF-M1	0.236	0.177	0.462

Table 6: Model uncertainty for training and testing stages. Entropy values are shown.

Training Stage					
Model	**ML**	**Jeffrey**	**Laplace**	**SG**	**Minimax**
SVR-IWO	4.516	4.523	4.527	4.516	4.518
SVR	4.528	4.53	4.531	4.528	4.528
RBF	4.524	4.528	4.53	4.524	4.525
Observed	4.514	4.523	4.526	4.515	4.517
Testing Stage					
Model	**ML**	**Jeffrey**	**Laplace**	**SG**	**Minimax**
SVR-IWO	3.640	3.656	3.656	3.65	3.652
SVR	3.660	3.660	3.662	3.661	3.661
RBF	3.654	3.658	3.660	3.655	3.656
Observed	3.649	3.655	3.658	3.650	3.652