

Article

# Suspended Sediment Yield Forecasting with Single and Multi-Objective Optimization Using Hybrid Artificial Intelligence Models

Arvind Yadav <sup>1</sup>, Premkumar Chithaluru <sup>1,2</sup>, Aman Singh <sup>3,4,\*</sup>, Marwan Ali Albahar <sup>5</sup>, Anca Jurcut <sup>6</sup>, Roberto Marcelo Álvarez <sup>2,3</sup>, Ramesh Kumar Mojjada <sup>7</sup> and Devendra Joshi <sup>1</sup>

<sup>1</sup> Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram 522502, Andhra Pradesh, India

<sup>2</sup> Department of Project Management, Universidad Internacional Iberoamericana, Campeche C.P. 24560, Mexico

<sup>3</sup> Higher Polytechnic School, Universidad Europea del Atlántico, C/Isabel Torres 21, 39011 Santander, Spain

<sup>4</sup> Department of Engineering, Universidad Internacional Iberoamericana, Arecibo, PR 00613, USA

<sup>5</sup> Department of Computer Science, Umm Al Qura University, Mecca P.O. Box 715, Saudi Arabia

<sup>6</sup> School of Computer Science, University College Dublin, Belfield, D04 V1W8 Dublin, Ireland

<sup>7</sup> Renault Nissan Technology & Business Centre, Chengalpattu 603002, Tamil Nadu, India

\* Correspondence: amansingh.x@gmail.com

**Abstract:** Rivers play a major role within ecosystems and society, including for domestic, industrial, and agricultural uses, and in power generation. Forecasting of suspended sediment yield (SSY) is critical for design, management, planning, and disaster prevention in river basin systems. It is difficult to forecast the SSY using conventional methods because these approaches cannot handle complicated non-stationarity and non-linearity. Artificial intelligence techniques have gained popularity in water resources due to handling complex problems of SSY. In this study, a fully automated generalized single hybrid intelligent artificial neural network (ANN)-based genetic algorithm (GA) forecasting model was developed using water discharge, temperature, rainfall, SSY, rock type, relief, and catchment area data of eleven gauging stations for forecasting the SSY. It is applied at individual gauging stations for SSY forecasting in the Mahanadi River which is one of India's largest peninsular rivers. All parameters of the ANN are optimized automatically and simultaneously using the GA. The multi-objective algorithm was applied to optimize the two conflicting objective functions (error variance and bias). The mean square error objective function was considered for the single-objective optimization model. Single and multi-objective GA-based ANN, autoregressive and multivariate autoregressive models were compared to each other. It was found that the single-objective GA-based ANN model provided the best accuracy among all comparative models, and it is the most suitable substitute for forecasting SSY. If the measurement of SSY is unavailable, then single-objective GA-based ANN modeling approaches can be recommended for forecasting SSY due to comparatively superior performance and simplicity of implementation.

**Keywords:** Mahanadi River; water discharge; genetic algorithm; ANN; SSY

**MSC:** 68T05



**Citation:** Yadav, A.; Chithaluru, P.; Singh, A.; Albahar, M.A.; Jurcut, A.; Álvarez, R.M.; Mojjada, R.K.; Joshi, D. Suspended Sediment Yield Forecasting with Single and Multi-Objective Optimization Using Hybrid Artificial Intelligence Models. *Mathematics* **2022**, *10*, 4263. <https://doi.org/10.3390/math10224263>

Received: 8 October 2022

Accepted: 11 November 2022

Published: 15 November 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Soil erosion is the most serious land degradation and water resource deterioration problem in river basins which is caused by human activities, topography, soil type, land use, natural processes, and rainfall characteristics [1,2]. Rivers are an important component of hydrology that acts as a primary dynamic geological transportation medium. It contains sediment and carries water which transfers the weathered materials of the continents to the ocean. Forecasting of suspended sediment yield (SSY) is critical for understanding the land–ocean mass balance. The SSY is defined as the amount of suspended sediment carried

by flowing water from a watershed to a point of reference over a specific period. The SSY is a key variable in a river due to its various effects on marine industries, hydroelectric equipment longevity, silting, water quality, geomorphology, watershed management, soil erosion and loss, reservoir sedimentation, and ecology [3–5]. Sedimentation can also reduce: the river's transport capacity, roadside ditches, navigation channels, streams, lake storage capacity, and reservoirs, resulting in more frequent flooding; and it produces numerous harmful impacts, which are demonstrated by various researches [4,6,7]. SSY forecasting also helps to prevent natural disasters [8], assists policymaking [9], and helps determine the service life of a reservoir [10]. Long-term and short-term forecasts of SSY are essential. Reservoir operations are commonly planned primarily based on month-to-month intervals; thus, monthly sediment yield forecasting is essential, which plays an important role in planning and dealing with the water resources domain [11]. Many researchers have developed artificial intelligence-based SSY forecasting models based on monthly data [12,13].

The study of hydrologic time series forecasting of SSY has attracted tremendous attention to improving forecasting accuracy quality. Different kinds of forecasting methods can be seen in the various literature [14,15]. Several researches have been carried out in the past to reduce the complexities of the forecasting problem to develop practical techniques. In the case of hydrological time series forecasting, traditional time series models, such as autoregressive moving average (ARMA), autoregressive integrated moving average (ARIMA), multivariate auto regressive (MAR) and autoregressive (AR), have been widely utilized [16]. However, in the context of the hydrological process of SSY, these traditional linear approaches are unable to capture complex non-stationarity and non-linearity. Traditional forecasting models have several drawbacks because most traditional modeling methods assume the data is stationary and linear, as well as the need for a high number of input parameters [17]. As a result, many researchers have concentrated on the development of those types of models which are capable of simulating non-stationary and nonlinear processes. The artificial intelligence (AI) approaches have shown promise in forecasting nonlinear hydrological processes and dealing with complicated and noisy datasets [12,13]. One of the most widely-used AI approaches is the ANN [18]. Due to cost-effectiveness, few data requirements, and simplicity, the ANN technique was well-suited to the forecasting and prediction modeling of a nonlinear and dynamic system in hydrology [18,19]. The ANN possesses the capability of learning complex nonlinear relationships between the input and output data, which works on the basis of a biological brain and its association with the nervous system [18]. The ANN model predicts future values by performing a nonlinear functional mapping of past observations. The ANN has been used broadly in hydrology for forecasting rainfall, runoff, flood, flow, and sediment yield modeling [5,20–23].

The primary drawbacks of the ANN are underfitting and overfitting issues [24]. Furthermore, choosing ANN characteristics such as network topologies, hidden layer nodes, starting weights and hidden layer numbers is critical for model fitting [25]. Incorrect selections of these parameters may result in a poorly generalized model. In general, the parameters of ANN models are chosen through grid search or trial-and-error techniques [26]. However, developing the parameters' values takes a significant amount of computational effort and this strategy is not guaranteed to provide a near optimal or optimal solution [27]. The simultaneous optimization of all related ANN's model parameters overcomes the limitations of trial-and-error methods, which was demonstrated in several applications of AI models [27,28]. The genetic algorithm (GA) is a population-based global search optimization algorithm based on Darwin's theory of evolution, which is capable of finding the optimum parameters of the AI models [27–29]. The GA's widely applicable optimization algorithms can be used to solve various discontinuous, non-differentiable, stochastic or highly nonlinear problems in a noisy environment [27]. The GA has been effectively utilized to optimize ANN predictive model parameters simultaneously [27,29], especially in hydrology [30,31] and forecasting models [32,33]. The GA can save the ANN from

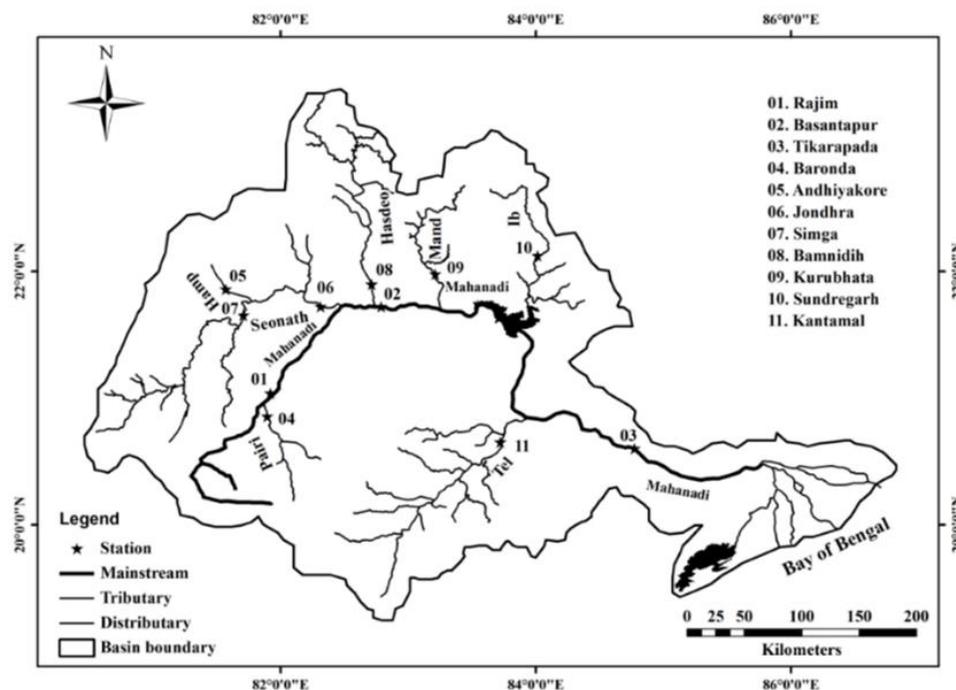
getting trapped in a local minimal area and pick out the global optimal parameters for the ANN. The GA can help the ANN avoid becoming stuck in a local minimum zone and choose the ANN model's optimal parameters [34]. Furthermore, the multi-objective optimization models (bias and variance) performed better than the single-objective (SO) (mean-squared-error) optimization for generating the ANN forecasting models of SSY [35]. Several AI sediment yield models were built for a specific geographical area by utilizing just rainfall (RF), suspended sediment yield (SSY), water discharge (Q), and temperature (T) as temporal data [36]. However, previous research has shown that relief (R), rock type (RT), and catchment area (CA) as spatial data also have a significant impact on sediment yield [37].

In this study, multi-objective (MO) and single objective GA-based ANN are utilized to construct an accurate and robust forecasting model in the Mahanadi River (MR), using T, R, SSY, Q, and RF as temporal data and CA, R and RT as spatial data. The generalized hybrid GA-based multi-objective optimization ANN (GA-MOO-ANN) with two different objective functions and single-objective GA-based ANN (GA-ANN) model with single objective function were developed by using combined data of eleven gauge stations for one-month-ahead forecasting of the SSY in MR basin, India. The GA-MOO-ANN version has been used to optimize the two conflicting objective functions, i.e., error variance and mean error (bias). The mean square error (MSE) is a nonlinear combination of error variance and bias. On the other hand, the MSE objective function has been optimized through the GA-ANN model. All ANN model parameters, such as inputs, transfer functions, initial network weights, bias terms, combinational coefficient ( $\mu$ ), and hidden neurons were optimized concurrently by utilizing a large quantity of Q, T, RF, and SSY as temporal data and RT, CA and R as spatial data. The proposed fully-automated parameter tuning and highly generalized SSY forecasting model for the MR reduced the need for human intervention, which is the research's main innovation. The proposed model was applied to each of the eleven individual gauging stations for forecasting SSY. As per the authors' knowledge, no attempt has been made to forecast the SSY in the MR basin using the multi-objective and single-objective GA-based ANN to optimize all ANN model parameters simultaneously. After the reliable hybrid GA and ANN-based forecasting model's development, the performance of the models was examined with the same test dataset. The comparison was performed amongst the hybrid AI (GA-ANN-51 and GA-MOO-ANN) and traditional regression (MAR and AR) models. The generalized GA-ANN-51 model provided the best results among all comparing models by considering the optimum input variables and associated parameters. This superiority was achieved due to the optimization of all ANN parameters simultaneously using the GA. Both hybrid AI (GA-ANN-51 and GA-MOO-ANN-51) models provided better results than the AR and MAR methods. The best-purpose model provided the best result at Tikarapara, which may be due to the highest Q, SSY, CA, and RF amongst all stations in the MR basin. The proposed forecasting models of SSY, which are climatic variables-dependent, are very helpful for planners and managers of water resources for good understanding of the problems and to find alternative solutions to handle problems in the future. If a measurement of SSY is unavailable, then this proposed hybrid GA-based ANN modeling approach can be recommended for the forecasting of SSY in MR, due to superior performance in comparison and ease of implementation. The rest of the paper is structured as follows. Section 2 presents descriptions of the study area. Section 3 provides detailed descriptions of the materials and proposed methods. Section 4 presents an analysis of results and discussion on various developed hybrid AI models and other comparative models. Section 5 describes the conclusions and future work of the paper.

## 2. Study Area

The MR system was chosen for the forecasting of the SSY. This river is a major peninsular river in east central India that flows east (Figure 1). Out of India's larger rivers, it is fourth ranked, covering 141,589 square kilometers, or about 4.3 percent of the country's

total land area [38]. The river contributes 53 percent of its basin area to Odisha, 46 percent to Chhattisgarh, and the remaining 1 percent to Maharashtra, Madhya Pradesh, and Jharkhand [38]. The MR is 851 km long and it begins at an elevation of approximately 442 m above mean sea level in Pharsiya village in the Dhamtari district of Chhattisgarh, India. The river flows through Odisha for 494 km and Chhattisgarh for the remaining 357 km. The geographical coordinates of the MR extend between the east longitudes of  $80^{\circ}30'$  to  $86^{\circ}50'$  and north latitudes of  $19^{\circ}20'$  to  $23^{\circ}35'$ . The world's largest earthen dam (Hirakud dam) is constructed in the MR. Figure 1 depicts a map of the locations of the basin with gauging stations.



**Figure 1.** Location map of the Mahanadi basin with hydro-meteorological stations adapted from [28].

Based on daily data from 1971 to 2004, the mean annual RF ranged from 1200 to 1400 mm [39]. During the monsoon season (June to October months), the basin of Mahanadi receives approximately 90 percent of its annual RF. The RF distribution in the MR basin is uneven. According to daily data from 1969 to 2004, during the warmer months of April and May, the maximum  $T$  varies from 39 to 45 degrees Celsius, and during the winter months of December and January, minimum  $T$  ranges from 4 to 12 degrees Celsius [39]. The highest and lowest relative humidity variations are 68–87 percent and 9–45 percent, respectively.

The MR basin's two main land uses are agriculture and forestry. Agriculture (54.27 percent), forestry (32.74 percent), wasteland (5.24 percent), construction (3.30 percent), and water bodies (14.45 percent) fill the MR's total basin area as per data of the 2005–2006 year [39]. Chilka Lake and Hirakud Dam are the larger water bodies in the MR.

The upper stream of the river is dominated by Proterozoic sedimentary rocks such as calcareous shale, sandstone, and limestone, whereas the lower stream is dominated by metamorphic silicate rocks. The basin's various lithologies include 34% granite, 7% khondalite, 15% charnockite, 17% Lower Gondwana limestone and shale, 5% coastal alluvium, and 22% Upper Gondwana sandstone and shale [40].

### 3. Materials and Methods

The behavior of SSY data is generally nonlinear and complex in nature with respect to its controlling factors. Variation in different statistical parameters can provide insight into the performance of the models. Monthly hydro-climatical data are used for developing the robust AI modelling. Data normalization and data division are also two important steps,

which need to be performed prior to the AI modelling. The comparison of the hybrid AI modelling with traditional regression models is required for checking the hybrid model's performance and forecasting capability. Therefore, this section aims to study the descriptive analysis of data and various models for SSY forecasting.

### 3.1. Materials

At all gauge stations, except Kantamal gauge station, monthly SSY, RF, Q, and T (Source: Central Water Commission (CWC), Mahanadi Bhawan, Bhubaneswar, Odisha) temporal data for the years 1990–2010, as well as spatial data including R, CA and RT from eleven MR gauging stations, were used to develop the proposed models. In this study, the sample size of the data used was 2616. There were six inputs parameters (water discharge, rainfall, temperature) and one output (suspended sediment yield) parameters. The monthly suspended sediment yield in the basin varied from 0 tons/month to 17,336,901 tons/month. The variation in monthly water discharge ranged from 0 cummec to 330,767 cummec. The rainfall in the basin varied from 0 cm to 1222.7 cm. The variation in monthly temperature ranged from 16 °C to 39 °C. The catchment area, rock type and relief value were mapped between 0 and 1. Unlike other temporal variables change (water discharge, rainfall, temperature and suspended sediment yield), spatial variables such as rock type, relief, and catchment area are fixed for each gauging station and do not change. The value of rock type was assigned a value between 0 and 1. If a gauge station had extremely hard rock with the lowest weatherability, the rock type value was set to zero. If the rock type was soft, such as clay or limestone, and easily decomposed and weathered, the rock type value was set to 1. Similarly, the relief and catchment area values were assigned values between 0 and 1. The highest relief values in the gauge station within the river basin were considered 1, the lowest relief values in the gauge station were considered 0, and the in between values were linearly interpolated between 0 and 1. The catchment area was also coded similarly to relief to obtain values between 0 and 1. Training data (70 percent of the dataset) were used to construct the models; testing data (15 percent) were utilized in the proposed model to assess the model's performance and validation data (15 percent) were used to avoid overfitting and underfitting issues of the developed models. At all gauge stations, except Kantamal gauge station, the training data range from 1 June 1990, to 31 May 2004, while the validation data range from 1 June 2004, to 31 May 2007, and the testing data range from 1 June 2007, to 31 May 2010. Furthermore, at the Kantamal gauge station, validation data range from 1 April 2002, to 31 December 2005; training data range from 1 June 1990, to 31 March 2002; and testing data range from 1 January 2005, to 30 September 2008. Testing data were unseen data that were not used in the model development process. The data from all 11-gauge stations of the MR basin were integrated to generate a single MR test, validation, and training dataset. The t-test was performed to check the similarity of distribution of divided testing, validation and the training data set. Property of these data sets should be similar for model development. The list of abbreviations and their definitions is shown in Table 1.

The different statistical parameters of hydro-climatic data (Q, RF, T, and SSY) are calculated at each station, which is given in Table 2. The SSY data sets are characterized by the highest value of KURT (13.89–78.03), Max/Mean (13.99–33), and COV (2.23–3.89), at each station among all hydro-climatic data sets. It may cause a difficult estimation of SSY as compared to Q, RF, and T data.

Data normalization is a pre-processing step that was used before developing the models. It is a technique for removing the impact of data ranges on the variables in a data set while maintaining the data range's uniqueness for model development. During training, it ensures quick processing and convergence, as well as reduced prediction error [24,28]. In this study, data normalization and data division of all inputs and output data set were performed before the development of all forecasting models. The normalization process is briefly described in various literature, with scales of all variables within the range of

0 to 1 [24,36]. The following equations are used to normalize the data between the a and b values:

$$c_{norm} = a + \frac{C_i - C_{min}}{C_{max} - C_{min}} \times (b - a) \tag{1}$$

where  $C_{min}$  is the minimum value,  $C_{max}$  is the maximum value,  $C_i$  is the  $i^{th}$  original value, and  $c_{norm}$  is the normalized value of the data set. The  $a$  and  $b$  are the minimum and maximum values of the data set.

**Table 1.** List of abbreviations with definitions.

Abbreviations	Meaning
SSY	Suspended sediment yield
ANN	Artificial neural network
GA	Genetic algorithm
ARMA	Autoregressive moving average
ARIMA	Autoregressive integrated moving average
AR	Autoregressive
AI	Artificial intelligence
SO	Single objective
Q	Water discharge
RF	Rainfall
T	Temperature
R	Relief
RT	Rock type
CA	Catchment area
MO	Multi-objective
MR	Mahanadi River
GA-MOO-ANN	Genetic algorithm based multi-objective optimization with artificial neural network
MSE	Mean square error
GA-ANN	Genetic algorithm based artificial neural network
MAR	Multivariate autoregressive
CWC	Central Water Commission
KURT	Coefficient of kurtosis
COV	Coefficient of variation
Max/mean	Maximum-mean ratio
SKEW	Coefficient of skewness
MLP	Multi-layer perceptron
LM	Levenberg Marquardt
CE-NSGA	Controlled elitist non-dominated sorting-based genetic algorithm
ACF	Autocorrelation function
CCF	Cross-correlation function
ADF	Augmented Dickey-Fuller
C	Critical value
GA-ANN-51	Genetic algorithm-based ANN models with 51 inputs
MAE	Mean absolute error
r	Coefficient of correlation
VAR	error variance

### 3.2. Methods

In this study, the SSY ( $S_{(t+1)}$ ) at the time  $t+1$  is forecasted using a given set of time-lagged past observation data. As a result, the following equation can be used to represent the relationship between the input and output variables, where time is represented by  $t$ ,  $n$  is the number of lags, and  $f$  represents the nonlinear forecasting function:

$$S_{t+1} = f(Q_t, RF_t, T_t, S_t, Q_{t-1}, RF_{t-1}, T_{t-1}, S_{t-1}, \dots, Q_{t-n}, RF_{t-n}, T_{t-n}, S_{t-n}, CA, RT, R) \tag{2}$$

This equation is used for the time series analysis of the proposed forecasting model, where  $S_t$  denotes the month-to-month SSY at time  $t$ ,  $S_{t-n}$  is the SSY at time  $t - n$ . The inputs of the forecasting model represent the previous  $n$  values of each input parameter,

such as SSY, Q, T, and RF, corresponding to time interval  $t, t - 1, t - 2, \dots, t - n$ , along with spatial parameters (CA, RT, and R) and output data is the one-step-ahead SSY ( $S_{t+1}$ ).

**Table 2.** The statistical parameters of hydro-climatic data in the Mahanadi River basin.

Stations	Statistics	SSY (Tons/Month)	Q (m <sup>3</sup> /s)	T (°C)	RF (mm)
Tikarapara	KURT	19.8585	11.1375	2.0617	10.0958
	SKEW	1.1512	1.3505	−0.6082	1.3555
	COV	2.3612	1.5331	0.1865	1.3955
	Max/Mean	16.0860	8.7893	1.3192	9.3705
Basantpur	KURT	33.3957	9.3052	1.8236	5.3033
	SKEW	1.1258	1.3735	−0.9199	1.5773
	COV	2.6518	1.6663	0.6592	1.5270
	Max/Mean	23.9063	10.1539	1.8981	6.5383
Andhiyarakhore	KURT	18.9710	11.5115	6.1357	8.3059
	SKEW	1.2076	1.2939	−0.3966	1.5339
	COV	2.3830	1.7993	0.2033	1.3353
	Max/Mean	18.7828	10.6993	1.3393	9.3392
Simga	KURT	35.6630	12.6576	7.2139	5.2091
	SKEW	1.0595	1.3592	−0.6160	1.3631
	COV	2.8285	2.0308	0.2869	1.3233
	Max/Mean	29.1062	13.5639	1.3182	6.7273
Sundargarh	KURT	18.8011	10.5151	7.3831	6.2788
	SKEW	1.3321	1.5659	−0.6638	1.5879
	COV	2.2330	1.6037	0.2380	1.3013
	Max/Mean	13.992	9.8012	1.3251	7.7768
Bamnidih	KURT	78.0323	28.2353	9.3761	3.2963
	SKEW	0.7559	0.9122	−0.5130	1.6335
	COV	3.8858	1.5338	0.2212	1.3352
	Max/Mean	33.0939	11.0378	1.3721	6.2068
Jondhra	KURT	38.0896	12.0792	7.1031	3.3099
	SKEW	1.1392	1.3593	−0.5590	1.5703
	COV	2.6035	1.8583	0.22801	1.3512
	Max/Mean	27.6338	13.1736	1.3899	6.0795
Kantamal	KURT	13.8865	9.3119	3.8359	7.2308
	SKEW	1.1672	1.3163	−0.9372	1.5125
	COV	2.5586	1.6966	0.3508	1.6086
	Max/Mean	17.0331	9.3033	2.3833	8.3330
Baronda	KURT	23.9991	9.0739	7.6910	5.7056
	SKEW	1.0789	1.3389	−0.7507	1.6850
	COV	2.7805	2.1363	0.2813	1.5596
	Max/Mean	22.8991	10.3937	1.3830	8.1952
Kurubhata	KURT	15.3906	9.3768	8.7586	3.8577
	SKEW	1.3132	1.3061	−0.5103	1.7005
	COV	2.2750	1.6677	0.2133	1.3598
	Max/Mean	15.0912	9.6309	1.3671	7.2033
Rajim	KURT	37.2735	10.1960	6.9175	5.7107
	SKEW	1.0012	1.3051	−0.3591	1.5580
	COV	2.9922	2.1717	0.2982	1.5529
	Max/Mean	26.2990	11.29511	1.5685	6.979329

COV: Coefficient of variation; SKEW: Coefficient of skewness; KURT: Coefficient of kurtosis; Max/mean: Maximum-mean ratio.

MATLAB software was used to create the customized code for GA-MOO-ANN and GA-ANN models. The ANN’s parameters were chosen using normalized data. The purpose of this study is to develop a multi-layer perceptron (MLP) ANN technique using a

Levenberg Marquardt (LM) training algorithm to forecast SSY. The MLP-based ANN can be trained using a variety of weight optimization strategies. The LM training algorithm was utilized in this work because of its capability to attain convergence faster and train quicker than the gradient descent training technique [41]. The MLP-based ANN typically consists of the hidden layer, output layer, and input layer. Each layer is interconnected with the weights, which are optimized during the training process. The error, which is the difference between the forecasted and the observed value, is iteratively backpropagated until the weights are optimized. The detailed description of MLP architecture and LM training algorithms of the ANN model is discussed by various researchers in many kinds of literature [24,28,36,41]. The five important ANN model parameters such as inputs, combinational coefficient ( $\mu$ ) of LM, hidden layer neurons, transfer function, and bias weights and connection weights are selected optimally using the GA.

The GA illustrates natural selection and genetics-based heuristic search techniques [42]. It has been demonstrated that the GA can find the global optimum solution to a variety of research challenges. It is a global optimization approach that is based on population and used to search for the best parameters for the ANN model. It is based on Charles Darwin's evolutionary theory. In a noisy environment, the GA is commonly utilized for optimization to tackle various differentiable, non-continuous stochastic, or highly nonlinear problems [42]. The GA was utilized in this study, in conjunction with ANN, for selecting the optimum parameters of the ANN. The GA-ANN model has only one hidden layer in the MLP. The LM algorithm is applied to train the ANN, and GA is applied for the selection of all of the optimum ANN parameters. To provide a more robust solution by offering a globally optimum solution, the ANN model parameters selection and its training were performed simultaneously, utilizing multi-objective and single-objective GA.

All five ANN parameters are represented as chromosomes, which are binary strings. The GA process usually begins with a randomly initialized chromosome that represents design or decision variables. The fitness value is then allocated by evaluating each chromosome (both objective and constraint conditions are checked). The ANN in the GA-ANN model has been trained for each chromosome of the GA and calculates the fitness value using the training data. The population is then updated by the mutation, reproduction, and crossover processes to get better results. These procedures are applied to generate a new population. The fitness function is analyzed and tested on the new population. Each chromosome is divided into five pieces, each of which represents a different ANN parameter. After selecting the set of 51 initial input variables for forecasting model, the final set of optimum input variables was selected by using the GA. The purpose of the GA is to select a subset of variables from the set of 51 input variables that can minimize the forecasting error. The forecasting model was developed using ANN models. Therefore, the GA was used to optimum input subset selection, also optimized the ANN parameters. The first part of the chromosome represents the input parameters. If the bit value is 1, then the concerned input is included in the specified subset; likewise, if the bit value is zero, then the corresponding input is not included in the subset. The transfer function for the output and hidden layers is represented by a three bit binary number in the second section of the chromosome. The transfer functions in the output layer and hidden layers are represented in this section. There are three different types of transfer functions: log sigmoidal, linear, and tan sigmoidal. There are nine alternative combinations of transfer functions for the output layer and hidden layers. The hidden layer neurons are represented by five bits in the third section of the chromosome. This binary number is translated to decimal during modeling to generate the hidden layer neurons. Due to processing cost and model complexity, the maximum hidden neurons are limited in the range of one to thirty-two. All decimal integers from 1 to 32 can be represented using the chromosomes of fivebit. The  $\mu$  is represented by the fourth section of the chromosome using an eight bit binary number. The  $\mu$  expressed the decimal values ranging from the lowest value 0 to the highest 255, which is normalized between 0.0010 and  $9 \times 10^9$  [24,36]. The normalization equation is given in equation 1. The connection weights and biases weights of the ANN

models are represented in the fifth section. Due to the changes in input quantities and the hidden neurons, the length of the chromosome varies.

To build a stable convergence model, the GA model parameters such as mutation probability, the highest number of generations (50), crossover probability (0.6), and the number of populations (50) were chosen using a trial-and-error process. The fitness value, which shows the measurement of chromosomes' success using validation data, has been calculated for each chromosome, which is based on training data. After each generation, bad chromosomes with low fitness function value are eliminated to maintain the population size of 50. The obtained population of chromosomes after one generation will be the starting solution for the next generation. The chromosome with the optimum fitness function value has a better probability of being selected for the next generation of GA. A roulette wheel reproduction operator was used in the research. Some chromosomes are chosen by elitism based on fitness values. The number of elites passed to the next generation of the GA is two. The GA should have a high probability of crossover and a low probability of mutation to work successfully [28,42]. Crossover and mutation operations are then performed for the selected chromosomes by the roulette wheel selection criterion which is based on the calculated fitness function. A new individual is found by the benefit from the parent fitness through the crossover operation. A mutation operator is responsible for obtaining the variety in the population. To keep the algorithm from going into a random search, a low value of mutation probability is commonly chosen. The root mean squared error of the training data as fitness function is used to estimate the fitness values of all chromosomes in the single objective ANN model. The maximum number of generations, minimum fitness function value, or minimal modification to the fitness function values in successive generations are employed as terminating conditions. The best solution value is obtained on the basis of minimum RMSE or MSE of fitness value after final generation. The chromosome from the population corresponding to that best solution is optimal learning parameters (transfer functions, number of hidden neurons,  $\mu$  value, and initial network weights and bias terms) for the ANN model. The flowchart of the proposed genetic algorithm-based ANN model is shown in Figure 2. The detailed description about operations of single objective GAs are described well by various researchers in much of the literature [25,28,34].

The controlled elitist non-dominated sorting-based GA (CE-NSGA) is utilized to create the multi-objective-based GA with two conflicting objectives, i.e., error variance and bias to optimize the parameters of the ANN [43,44]. Both objective functions evaluated all chromosomes' fitness values for the initial population. The chromosomes are sorted using a controlled non-dominant sorting strategy. In this sorting, the population was divided into many fronts (referred to as levels) based on the non-domination level. For determining non-dominated rankings, the CE-NSGA is combined with ANN by error variance and mean error objectives. The ANN model may be used to forecast the SSY with known input variables using these optimum parameters. Various kinds of literature provide extensive descriptions of multi-objective GAs or multi-objective GA-based ANN [43–45]. Rosales-Perez [46] and Yadav et al. [45] demonstrated detailed theoretical knowledge on multi-objective GAs variance optimization with variance and bias optimization.

In the autoregressive (AR) model, the SSY (S) was forecasted using a linear combination of time series SSY data. The equation of the AR model is given below [16].

$$S_{t+1} = a_0 S_t + a_1 S_{t-1} + a_2 S_{t-2} + \dots + a_n S_{t-n}, \quad (3)$$

where  $n$  is the AR model's order number, and  $a_i$  ( $i = 0, 1, 2, \dots, n$ ) denotes the regression model's coefficients. The AR model uses weighted sums of past SSY values to forecast future SSY values. Different input parameter selections developed various AR models with different temporal lags. In this AR model, the maximum lag was selected using an autocorrelation function (ACF).

Multivariate autoregressive modeling (MAR) is also employed for SSY forecasting. It is based on numerous factors which have a linear relationship. The MAR models are created by combining past time-step data from numerous variables in a linear fashion,

which may be thought of as a series of linear regressions [47]. The MAR equation is given below [47].

$$S_{t+1} = a_1Q_t + b_1RF_t + c_1T_t + d_1S_t + a_2Q_{t-1} + b_2RF_{t-1} + c_2T_{t-1} + d_2S_{t-1} + \dots + a_nQ_{t-n} + b_nRF_{t-n} + c_nT_{t-n} + d_nS_{t-n} + eCA + fRT + gR \tag{4}$$

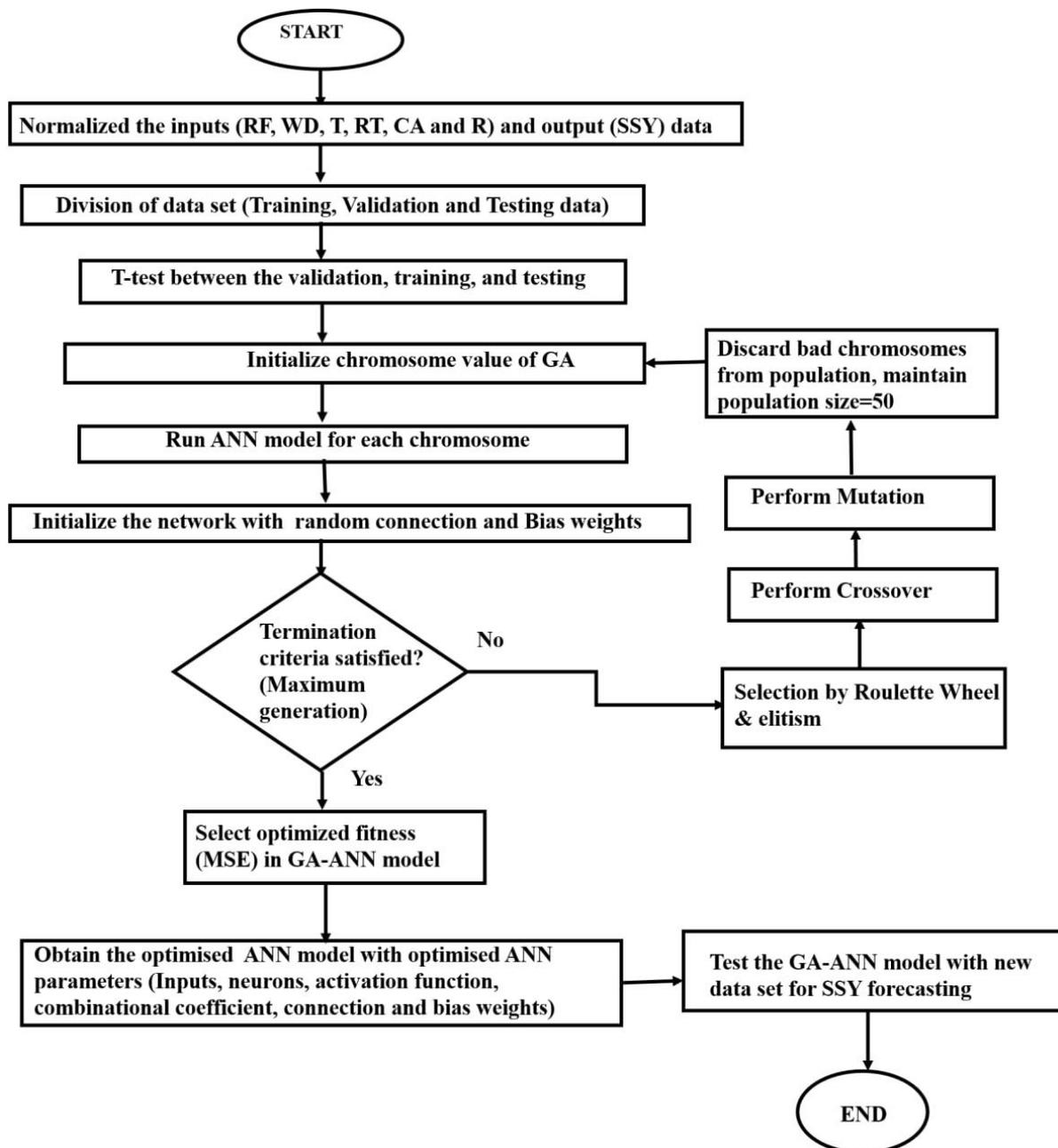


Figure 2. Flow chart of hybrid GA-based ANN model.

This equation represents the linear MAR forecasting model up to n lags. The  $a_{ij}$ ,  $b_i$ ,  $c_i$ ,  $d_i$ ,  $e$ ,  $f$ , and  $g$  ( $i = 0, 1, 2, \dots, n$ ) represent the coefficients of the MAR model. The  $a_i$ ,  $b_i$ ,  $c_i$ , and  $d_i$  represents the coefficients of the Q, RF, T, and S temporal variable, respectively. The  $e$ ,  $f$ , and  $g$  values are the coefficients of the CA, RT, and R, respectively. The spatial variables are the CA, RT, and R. Ives et al. [47] provide a thorough overview of the MAR's underlying theory and assumptions, as well as methods for estimating the parameters of the model. The MAR modeling has been used successfully in hydrology.

The MAR method was conducted in this research using the training dataset. The maximum lag was selected using the cross-correlation function (CCF). The MAR model used 51 input variables (12 each from each of the four temporal variables and 3 spatial variables) to forecast a one-step SSY value. A validation dataset is not necessary because the linear model does not overfit. The testing of MAR and AR methods was carried out using the same data as the ANN models.

#### 4. Results and Discussions

In this study, multi-objective GA and single-objective GA-based ANN hybrid models were developed for forecasting SSY in the MR. To evaluate the forecasting abilities of the models, the proposed hybrid model's performance was compared to the MAR and AR models.

##### 4.1. Testing the Stationary Data

All required data for developing the time series forecasting models should be stationary [48]. The Philips–Perron test, Augmented Dickey–Fuller (ADF) test and Ljung–Box Q test are all available to verify the stationary status of data [49]. The ADF test is employed in this research. In the ADF test, the null hypothesis demonstrates that a time series has a unit root, while the alternative hypothesis represents the stationarity of the time series. The ADF statistic is a negative number and the lower this number which indicates the greater the null hypothesis rejection [50]. The ADF test has a significance level that ranges from 0 to 0.999. The hypothesis value, i.e.,  $h = 0$  implies that this test does not reject the unit root's null hypothesis. The test of ADF has rejected the null hypothesis while accepting the alternate hypothesis if the  $h$  value is equal to 1. The low  $p$  values, i.e., low probability of acceptance of the null hypothesis helps to reject the null hypothesis. Table 3 presents the ADF test statistics, critical values,  $p$ -values, and  $h$ -values for all four parameters like S, T, RF, and Q. The results of the ADF test represent that the test statistics are more negative than the critical values, and the  $p$ -values are significantly very low. These results demonstrated that the ADF test rejects the null hypothesis. Thus, the test results indicate that the data are stationary.

**Table 3.** ADF test for checking the stationary trend of data.

Hydro-Climatical Data	Test Statistic	Critical Value (C)	$h$	$p$
SSY	−27.487	−1.942	1	0.001
Q	−22.189	−1.942	1	0.001
RF	−6.436	−1.942	1	0.001
T	−1.942	−1.942	1	0.001

##### 4.2. Statistical Analysis for Input Lags Selection

Linear and nonlinear correlation analyses were used to analyze the impact of hydro-climatic variables such as T, RF, and Q on SSY data in the MR basin. Q and RF have a good relationship with SSY, and Q is more favorably connected with SSY than RF and T [45].

Since the sediment yield data are seasonal and available monthly and affected by the previous month's data, the maximum lag needs to be selected for the forecasting model. An autocorrelation function (ACF) is utilized to evaluate the temporal correlations of sediment yield, whereas a cross-correlation function (CCF) was used for temporal cross-correlation between sediment yield and other controlling variables. The ACF of the SSY with various time lags is presented in Figure 3. The highest correlation was obtained at lag 1, and the correlation decreases with increasing time lag (Figure 3).

It was also observed that ACF exhibit the next highest peak value at lag 12, which supports the SSY data set's seasonality. ACF plot reveals the maximum lag is 12, after which the cyclicity begins due to the seasonal behavior of the data. It is noted that the maximum number of lag that is selected using ACF is not going to be used directly for the

forecasting model [5]. Cross-correlation coefficient analysis was performed for monthly T, Q, and RF data with SSY (Figure 4a–c).

It was observed that the highest CCF of sediment yield was observed with Q (0.880) at time lag 0. This finding is in agreement with the value of the Pearson correlation coefficient. The CCF plots show that SSY exhibited a cyclic relationship with T, RF, and Q. As a result, the most essential information for SSY forecasting comes from the 12 antecedents of hydro-climatic monthly variables, including Q, RF, SSY, and T. Based on the autocorrelation coefficient of this investigation, 12 antecedents of SSY, T, RF, and Q were chosen as the input vector.

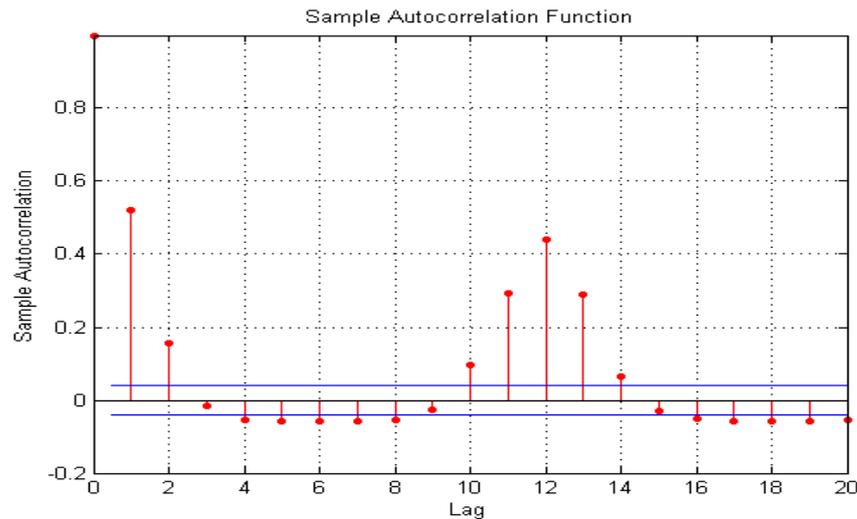


Figure 3. ACF of SSY for time series data with various time lags in MR basin.

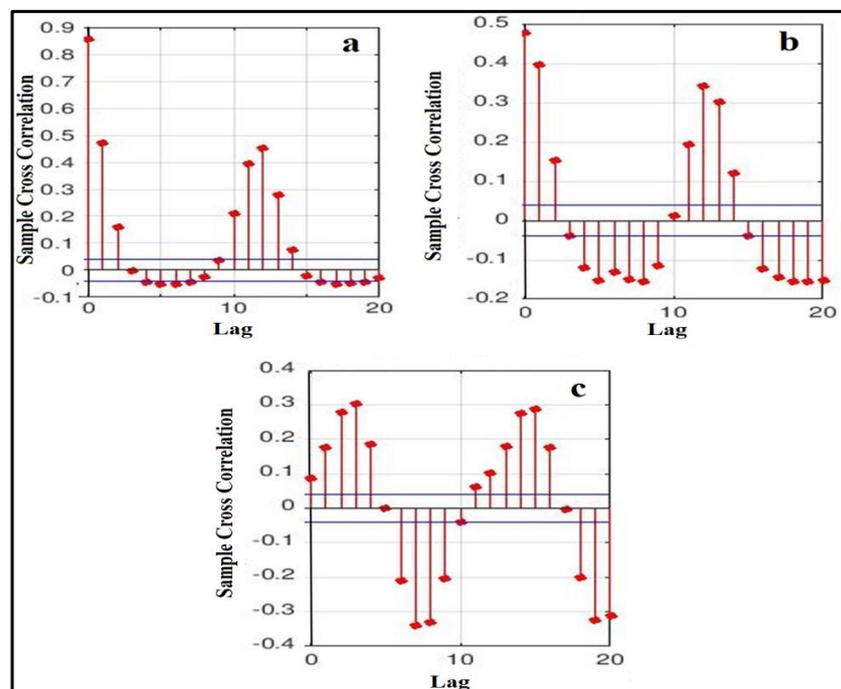


Figure 4. CCF plot of the hydro-climatical data between (a) Q and SSY (b) RF and SSY (c) T and SSY.

In this study, forecasting models were built for forecasting one-step-ahead SSY using all 48 temporal data (Q, T, RF, and SSY) with a 12-month lag and three spatial data (RT, CA, and R) as inputs. CA, R, and RT were fixed for each gauging station and not changed, unlike other temporal variables (Q, RF, T, and SSY). RT, R, and CA values were also plotted

within 0 and 1 in the forecasting model [45]. The selected maximum lag number for all models (AR MAR and hybrid artificial intelligence models) using ACF and CCF was used as input for the hybrid artificial intelligence models, where the inputs were selected using GA. Finally, the forecasting models have 51 input variables initially to forecast one step ahead of the SSY value. After selecting the set of 51 initial input variables for the forecasting model, the final set of optimum inputs is chosen using the GA.

#### 4.3. Artificial Intelligence Models for Forecasting the SSY

In this study, single-objective genetic-based ANN models with 51 initial inputs (GA-ANN-51) and single-objective genetic-based ANN models with 48 initial inputs (GA-ANN-48) forecasting models were developed by a conjunction of GA and ANN intelligence-based techniques. Multi-objective genetic-based ANN models with 51 initial inputs (GA-MOO-ANN-51) and multi-objective genetic-based ANN models with 48 initial inputs (GA-MOO-ANN-48) forecasting models were also developed by hybridization of ANN and multi-objective GA techniques. In these models, GA-ANN-51 and GA-MOO-ANN-51 models are developed by considering previous time series temporal data (Q, SSY, T, and RF) and spatial data (R, RT, and CA). The GA-ANN-48 and GA-MOO-ANN-48 models are developed using previous time series temporal data (T, Q, SSY, and RF) without considering the spatial data (RT, R, and CA). The GA-ANN-12 and GA-ANN-MOO-12 models were also built by considering the previous time series sediment data without considering the spatial data (RT, R, and CA) and without time series data of other variables (T, Q, and RF) with the specific lag selection of 12. The GA-ANN-15 and GA-MOO-ANN-15 models were developed by 15 input variables (12 from SSY temporal variables and three spatial variables) to forecast a one-step-ahead SSY value. The different types of models with lag input variables are presented in Table 4. All these models' performances are compared for evaluating the predictive capability of models based on statistical error analysis. Different models' performances are compared using the evaluation criteria (statistical error), and the best model is chosen. The testing, validation, and training datasets of the forecasting models were used to produce error statistics such as mean squared error (MSE), mean absolute error (MAE), coefficient of correlation ( $r$ ), error variance (VAR) and root mean squared error (RMSE).

**Table 4.** Different types of models use various input parameters with lags.

Models	Input Parameters	Initial Lags of Temporal Data
GA-ANN-12	SSY	12
GA-ANN-15	SSY, RT, R, and CA	15
GA-ANN-48	SSY, Q, RF, and T	48
GA-ANN-51	SSY, Q, RF, T, RT, R, and CA	51
GA-MOO-ANN-12	SSY	12
GA-MOO-ANN-15	SSY, RT, R, and CA	15
GA-MOO-ANN-48	SSY, Q, RF, and T	48
GA-MOO-ANN-51	SSY, Q, RF, T, RT, R, and CA	51

The optimal solution value is determined by the fitness value's least RMSE after the final generation. The fitness function value over different generations of the GA-ANN-51 model is shown in Figure 5.

The minimum fitness value is equal to 0.006, along with mean fitness of 0.019. It is also observed that after five generations, the best fitness function of each generation of genetic learning does not change. Table 5 presents optimally selected parameters by GA for developed GA-ANN and GA-MOO-ANN models. The optimal combinational coefficient and optimum neurons are one and eight, respectively. In the GA-ANN-51 model, the

tan-sigmoid transfer function was ideally selected at both the output and hidden layers. Similarly, the GA-based models' parameters were also determined.

The testing dataset's observed and model forecasted SSY was utilized to assess the generalization potential and performance of generated forecasting models. The statistical errors were derived from the validation, testing, and training data sets of all models to forecast one-step SSY value, which is shown in Table 6.

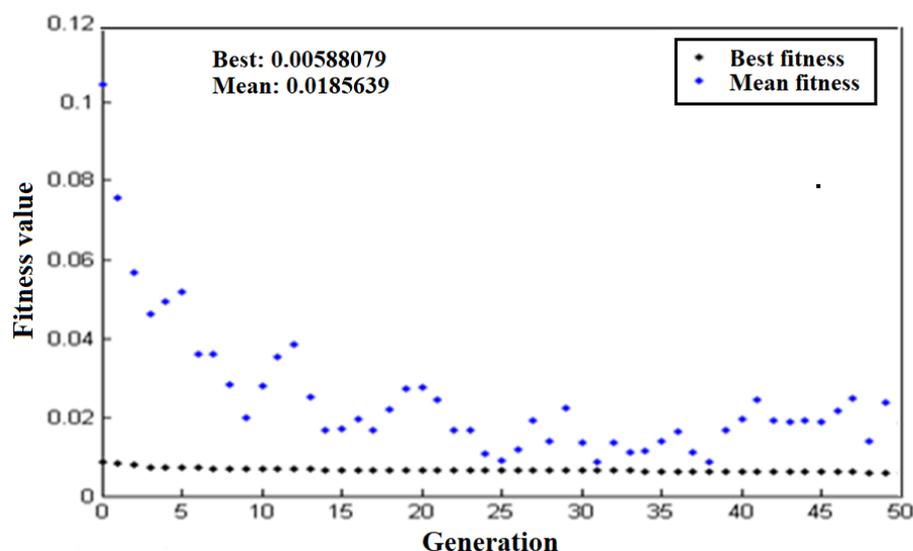


Figure 5. The variation of fitness value of each generation during the learning of the GA-ANN model.

Table 5. Optimally selected parameters of GA-MOO-ANN and GA-ANN models.

Models	Neurons	Combinational Coefficient ( $\mu$ )	Transfer Function	Inputs
GA-ANN-12	3	2	Tan-sigmoid, Pure linear	6
GA-ANN-15	4	3	Log-sigmoid, Tan-sigmoid	11
GA-MOO-ANN-12	4	6	Tan-sigmoid, Pure linear	6
GA-MOO-ANN-15	15	10	Pure linear, Tan-sigmoid	10
GA-ANN-48	17	8	Log-sigmoid, Tan-sigmoid	21
GA-ANN-51	8	1	Tan-sigmoid, Tan-sigmoid	19
GA-MOO-ANN-48	19	9	Tan-sigmoid, Pure linear	21
GA-MOO-ANN-51	29	2	Log-sigmoid, Pure linear	22

In GA-ANN-12, GA-MOO-ANN-12, GA-ANN-15, GA-MOO-ANN-15, GA-ANN-48, GA-MOO-ANN-48, GA-ANN-51, and GA-MOO-ANN-51 models, the error statistics of all validation, testing and training data show that error variance, MAE, RMSE and MSE are low, while r is relatively high. These models, it may be inferred, have a high level of accuracy in the forecasting of SSY. All three datasets had a high level of consistency, which indicates that the constructed forecasting models had neither been overfitted nor under fitted, and that generalized models had been produced.

Based on the RMSE in Table 6, the hybrid GA-MOO-ANN-51 and GA-ANN-51 models produce more accurate results than other comparative models. It was also observed that the single objective GA-based ANN model with 19 selected input variables from the 51

initial inputs (GA-ANN-51) model contains the lowest VAR, RMSE, MSE, and the highest  $r$  among all comparative models. For forecasting of SSY, the GA-ANN-51 model outperforms the GA-ANN-12, GA-MOO-ANN-12, GA-ANN-15, GA-MOO-ANN-15, GA-ANN-48, GA-MOO-ANN-48, and GA-MOO-ANN-51 models. It may be due to using the more suitable informative variable of time series lag data of temporal Q, RF, T, and SSY, along with spatially controlling factors (RT, R, and CA), with simultaneous optimization of the model's parameters by hybridization techniques of single objective GA and ANN. The proposed method proved that by taking into account previous Q, RF, T, and SSY, the sediment yield at time  $t+1$  can be predicted. So, even if we do not know what time  $t + 1$  is, the SSY can be predicted.

**Table 6.** Testing, validation, and training data error statistics for various forecasting models used for one-step-ahead forecasting of SSY.

Models	Data Set	RMSE	MSE	MAE	VAR	$r$
GA-ANN-12	Training	0.017	$2.800 \times 10^{-4}$	0.00533	$2.900 \times 10^{-4}$	0.577
	Validation	0.007	$5.040 \times 10^{-5}$	0.00274	$5.05 \times 10^{-5}$	0.778
	Testing	0.012	$1.410 \times 10^{-4}$	0.00389	$1.400 \times 10^{-4}$	0.619
GA-ANN-15	Training	0.012	$1.360 \times 10^{-4}$	0.00413	$1.400 \times 10^{-4}$	0.825
	Validation	0.006	$3.550 \times 10^{-5}$	0.00261	$3.540 \times 10^{-5}$	0.851
	Testing	0.012	$1.440 \times 10^{-4}$	0.00401	$1.400 \times 10^{-4}$	0.654
GA-MOO-ANN-12	Training	0.017	$2.800 \times 10^{-4}$	0.00504	$2.800 \times 10^{-4}$	0.588
	Validation	0.007	$5.530 \times 10^{-5}$	0.00260	$5.55 \times 10^{-5}$	0.753
	Testing	0.012	$1.400 \times 10^{-4}$	0.00363	$1.400 \times 10^{-4}$	0.623
GA-MOO-ANN-15	Training	0.009	$8.530 \times 10^{-5}$	0.00335	$8.53 \times 10^{-5}$	0.895
	Validation	0.006	$3.540 \times 10^{-5}$	0.00247	$3.55 \times 10^{-5}$	0.851
	Testing	0.014	$1.900 \times 10^{-3}$	0.00440	$1.900 \times 10^{-4}$	0.513
GA-ANN-48	Training	0.015	$2.100 \times 10^{-4}$	0.00484	$2.100 \times 10^{-4}$	0.716
	Validation	0.007	$4.890 \times 10^{-5}$	0.00287	$4.87 \times 10^{-5}$	0.787
	Testing	0.012	$1.400 \times 10^{-4}$	0.00422	$1.400 \times 10^{-4}$	0.617
GA-ANN-51	Training	0.015	$2.350 \times 10^{-4}$	0.00457	$2.400 \times 10^{-4}$	0.674
	Validation	0.006	$3.460 \times 10^{-5}$	0.00221	$3.47 \times 10^{-5}$	0.855
	Testing	0.011	$1.200 \times 10^{-4}$	0.00375	$1.200 \times 10^{-4}$	0.689
GA-MOO-ANN-48	Training	0.015	$2.200 \times 10^{-4}$	0.00426	$2.200 \times 10^{-4}$	0.704
	Validation	0.007	$4.77 \times 10^{-5}$	0.00246	$4.780 \times 10^{-5}$	0.798
	Testing	0.013	$1.800 \times 10^{-4}$	0.00381	$1.800 \times 10^{-4}$	0.568
GA-MOO-ANN-51	Training	0.016	$2.400 \times 10^{-3}$	0.00468	$2.410 \times 10^{-4}$	0.669
	Validation	0.007	$5.250 \times 10^{-5}$	0.00287	$5.260 \times 10^{-5}$	0.773
	Testing	0.012	$1.400 \times 10^{-4}$	0.00380	$1.400 \times 10^{-4}$	0.643

After developing the forecasting models, these developed models were applied in every gauge station to evaluate their performance. Table 7 shows the error statistics of the proposed best GA-ANN-51 model at each gauge station during the testing phase. The GA-ANN-51 model was applied at individual gauge stations for SSY forecasting.

It is also observed from Table 7 that the observed and forecasted values are showing mixed kinds of performance, with some stations delivering satisfactory performance, whereas a few stations are showing poor performance. The forecasting model cannot forecast SSY accurately at Andhiyarakhore and Baronda stations due to the erratic and complex nonlinear nature of SSY. Due to the low CA and flat terrain forecasting model, it cannot provide accurate results. The Simga gauging station is the first starting gauging station of MR after originating from the starting point, which is near Nagri town and Pharsiya village in the Raipur district of Chhattisgarh in India. This station has low water flow and SSY. The forecasting model provides low accuracy at Simga station. The proposed models did not give good results at some stations due to the relatively high coefficient of

variation (COV), skewness, max/mean ratio, and kurtosis value of influential variables (Table 2). The  $r$  value is greater than 0.7 at Tikarapara, Kurubhata, and Basantpur gauging stations, which have a strong correlation between the forecasted and observed SSY. Very high  $Q$  is found at these stations. Thus, the model is providing satisfactory results for SSY forecasting in these stations. The remaining gauging stations show moderate correlation, i.e., the  $r$  value lies between 0.5 and 0.7.

**Table 7.** Performances evaluation of the proposed best GA-ANN-51 forecasting model at each gauging station during testing phase.

GA-ANN-51	MAE	MSE	RMSE	$r$	VAR
Tikarapara	0.009	$3.000 \times 10^{-4}$	0.018	0.783	$3.000 \times 10^{-4}$
Simga	0.002	$2.080 \times 10^{-5}$	0.005	0.454	$1.660 \times 10^{-5}$
Andhiyarakhore	0.001	$1.680 \times 10^{-6}$	0.001	0.283	$9.200 \times 10^{-7}$
Sundargarh	0.003	$4.120 \times 10^{-5}$	0.006	0.524	$4.080 \times 10^{-5}$
Bamnidih	0.003	$4.560 \times 10^{-5}$	0.007	0.595	$4.240 \times 10^{-5}$
Jondhara	0.003	$4.590 \times 10^{-5}$	0.007	0.542	$4.710 \times 10^{-5}$
Kurubhata	0.001	$4.220 \times 10^{-6}$	0.002	0.771	$4.270 \times 10^{-6}$
Basantpur	0.006	$1.000 \times 10^{-4}$	0.011	0.711	$1.000 \times 10^{-4}$
Baronda	0.001	$3.500 \times 10^{-6}$	0.002	0.482	$3.330 \times 10^{-6}$
Rajim	0.001	$4.010 \times 10^{-6}$	0.002	0.623	$3.630 \times 10^{-6}$
Kantamal	0.014	$8.180 \times 10^{-4}$	0.029	0.585	$8.300 \times 10^{-4}$

Figures 6 and 7 show the hydrologic graph and scatter plot, respectively, of the GA-ANN-51 model.

Except for Kantamal, Andhiyarakore, Simga, and Bamnidih stations, the hydrologic graph showed that the forecasted SSY respects the fluctuation in the observed SSY data. Similarly, except for the four gauging stations indicated above, the GA-ANN-51 results are closer to the bisector line (1:1 line) (Figure 7). The magnitudes of high, low, and medium SSY forecasted values using the best forecasting GA-ANN-51 model are also closer to the corresponding actual SSY which is seen in the scatter and hydrograph plots. At all 11 gauge stations in the MR basin, the GA-ANN-51 model showed a positive SSY value even when the SSY output was zero or close to zero (Figures 6 and 7). This finding demonstrated that the most accurate way to calculate SSY in the MR basin system is to use an ANN in conjunction with a GA. Among all 11 gauging stations, the proposed forecasting model provided the best accuracy at Tikarapara. It could be because Tikarapara is situated at the far downstream end of the MR basin which has the maximum CA, RF,  $Q$ , and SSY among all the gauging stations [45].

#### 4.4. Comparison Results of Forecasting Models

The models' results were compared to other methods using the same test dataset after the development of a reliable hybrid GA and ANN-based forecasting model. The best multi-objective ANN model (GA-MOO-ANN), best single-objective ANN model (GA-ANN-51), and classic AR and MAR models were compared. Table 8 shows the comparative findings of all models in test data of all combined gauging stations.

**Table 8.** The comparison of best performances of various types of forecasting models based on MAE,  $r$ , VAR, and RMSE.

Models	RMSE	MSE	VAR	$r$
GA-ANN-51	$1.103 \times 10^{-2}$	$1.220 \times 10^{-4}$	$1.220 \times 10^{-4}$	0.689
GA-MOO-ANN-51	$1.160 \times 10^{-2}$	$1.350 \times 10^{-4}$	$1.350 \times 10^{-4}$	0.643
AR	$1.335 \times 10^{-2}$	$1.780 \times 10^{-4}$	$1.780 \times 10^{-4}$	0.497
MAR	$1.296 \times 10^{-2}$	$1.680 \times 10^{-4}$	$1.680 \times 10^{-4}$	0.556

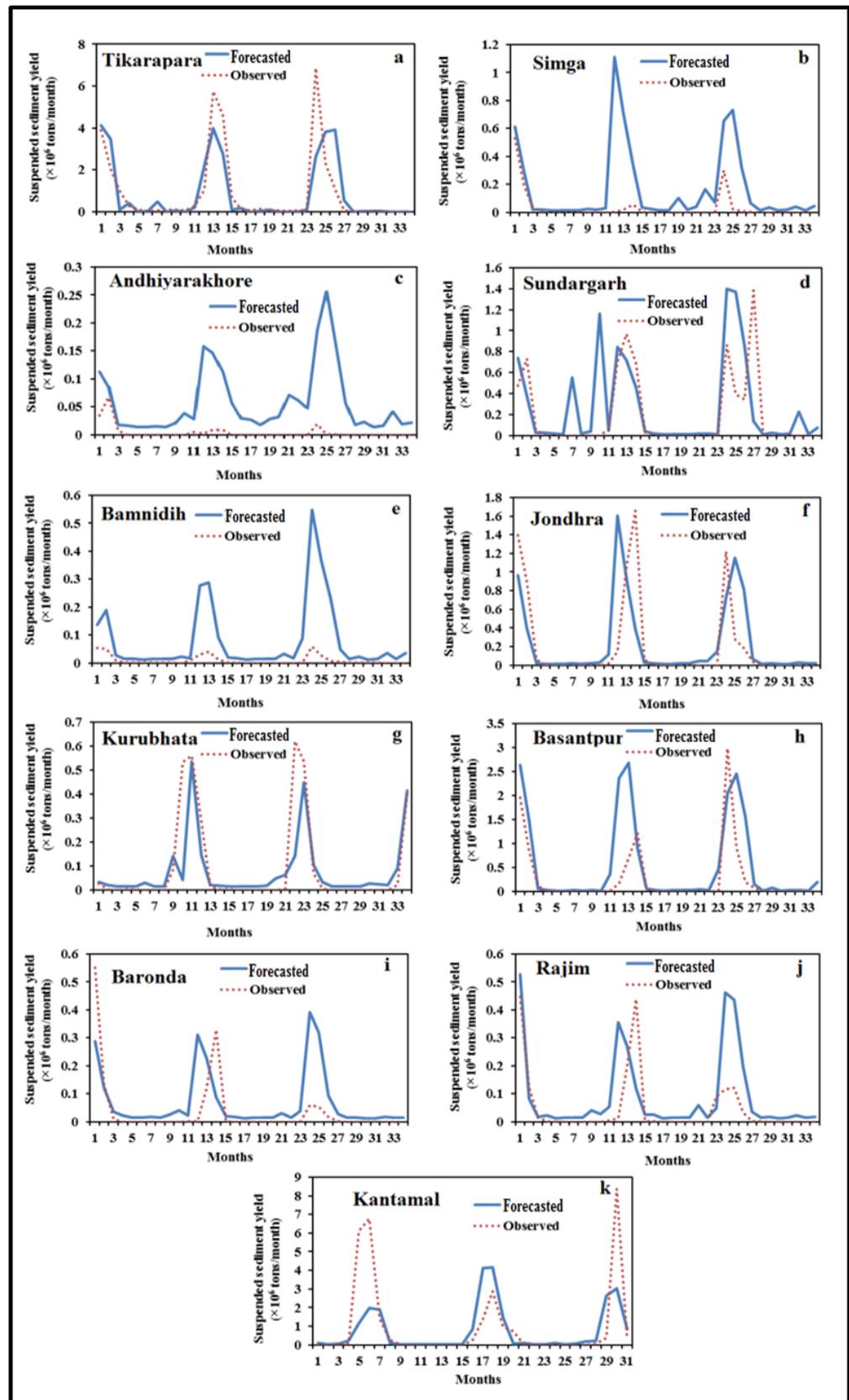


Figure 6. Comparison of the forecasted and observed SSY of the GA-ANN-51 forecasting model in a testing phase (a–k).

The proposed hybrid GA-ANN-51 model has the smallest variance, MSE, RMSE, and  $r$  among all comparative models (Table 8). The GA-ANN-51 model was shown to be the most effective in this statistical analysis. The GA-ANN-51 model has slightly lower RMSE, MSE, and VAR and slightly higher  $r$  than the GA-MOO-ANN-51 model. These findings also reveal that the RMSE values of the GA-MOO-ANN-51 and GA-ANN-51 models are lesser, as compared to the MAR and AR models. By determining the best input parameters and other associated ANN parameters, the GA-ANN-51 model outperforms the GA-MOO-ANN, AR, and MAR models. This supremacy is due to the GA's simultaneous optimization of all parameters of the ANNs. Traditional AR and MAR models performed worse than AI-based hybrid models (GA-ANN and GA-MOO-ANN). Furthermore, the magnitude of the high, medium, and low values of SSY predictions by this GA-ANN-51 forecasting model were closer to the corresponding actual values of the SSY, which is seen in the scatter and hydrograph plots. The best forecasting GA-ANN-51 model showed a positive value of SSY, even though SSY was nil or approximately zero at all 11 gauge stations in the MR basin, which is observed in the hydrograph and scatter plots. However, because of space limits, the MAR and AR models produced negative SSY estimates, as seen in the hydrograph and scatter plots, which are not reported in this article. The fact that the data displays considerable non-linearity in the context of small, valued samples of SSY was observed as a result of these negative values. The linear MAR model fails to account for this non-linearity, resulting in certain negative SSY estimates. This is completely impractical because SSY cannot have a negative value. The GA-ANN-51 model can deal with the nonlinearities of SSY. Traditional forecasting approaches (MAR and AR) fail to capture the complex nonlinear behavior of SSY. In comparison to the other models, the GA-ANN-51 appears to be the best model with the most generalized capability.

Only a linear combination of prior time-step SSY data is used in the autoregressive (AR) approach. Table 8 shows that the AR model has the worst performance among all models since it has the largest values of variance, MSE, RMSE, and the lowest value of  $r$ . The poor performance of the AR model is also owing to the inclusion of only SSY as an input time lag variable and the omission of data from T, RF, and Q. The T, particularly in the previous month, is a measure of soil moisture, which has a big impact on soil erosion processes and the river's sediment supply [26]. The GA-ANN-51 model's high performance is due to the use of more informative variables such as time series lag data of T, RF, Q and SSY combined with spatial data (CA, R, RT). Finally, it is found that the single-objective function-based GA-ANN model outperforms the multi-objective function-based GA-MOO-ANN model, AR, and MAR models in forecasting SSY in India's MR basin.

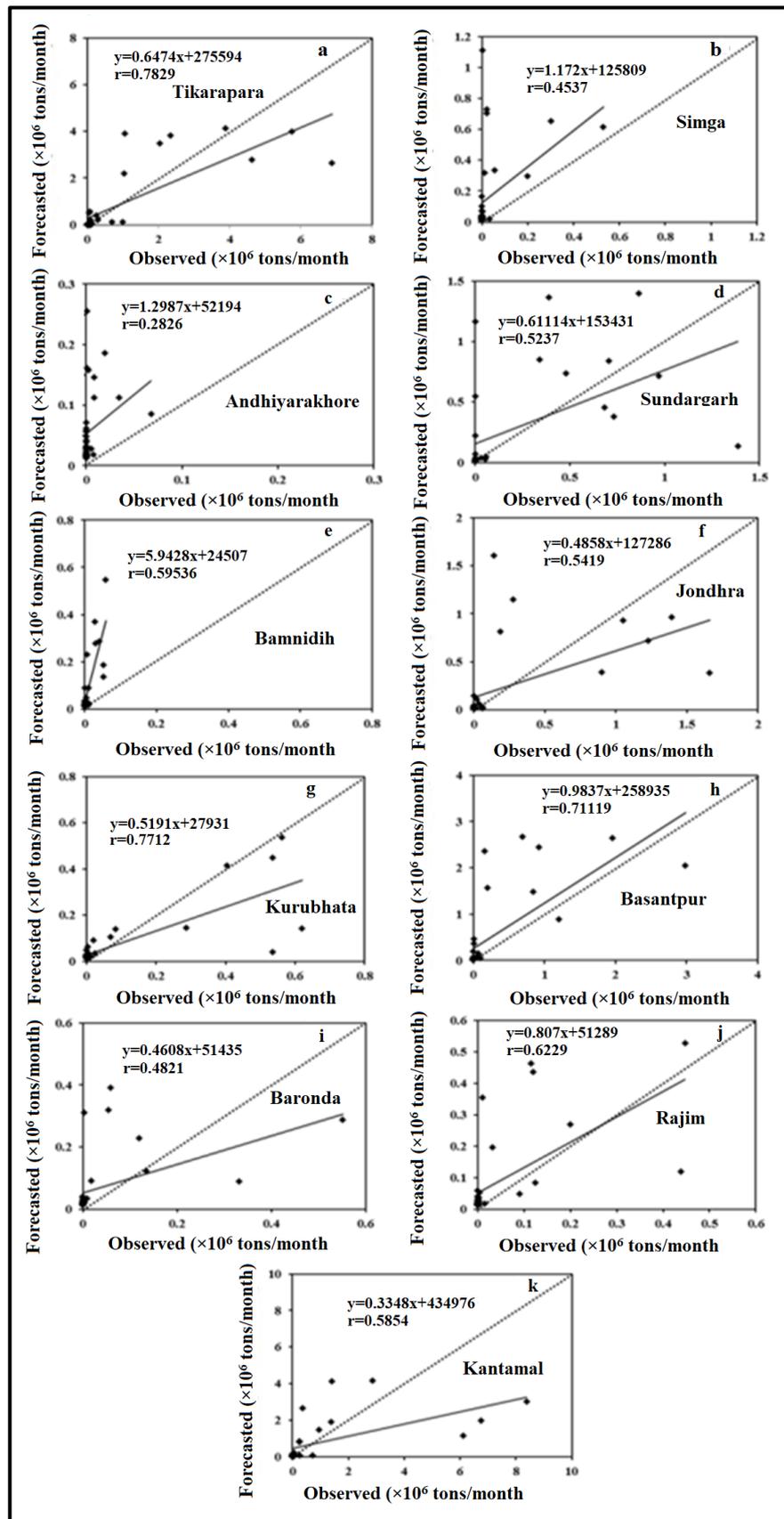


Figure 7. Scatter plot of forecasted and observed SSY from GA-ANN-51 forecasting model in a testing phase (a–k).

## 5. Conclusions and Future Scope

The study used single- and multi-objective neural networks, AR, and MAR models to forecast the SSY in the MR in India. In the case of all 11 gauging stations, Q, RF, T, RT, R, and CA are included as input factors in the developed models. Using single and multi-objective GAs, the hybrid forecasting models optimized various model parameters at the same time. The GA-ANN-51 and GA-MOO-ANN-51 forecasting modes, which are based on 11 gauging stations and 12 lags, have higher predictability than the AR and MAR models due to the selection of optimum input parameters and associated all ANN's parameters. This supremacy is due to the GA's simultaneous optimization of all parameters of the ANN model. The findings showed that the proposed hybrid GA-MOO-ANN-51 and GA-ANN-51 forecasting models performed well for forecasting SSY and have a stronger generalization power than other comparable MAR and AR forecasting models. The GA-ANN-51 model produces slightly better outcomes than the GA-MOO-ANN-51, but the differences are minor. It is also worth noting that all of the GA-ANN-51 models accurately forecast SSY in the sub-basin with the largest CA with the best result coming from Tikarapara, which has the highest CA. The GA was used to optimize all ANN parameters simultaneously, resulting in the superiority of hybrid forecasting models. The GA-ANN model forecasts low, medium, and high SSY values that are closer to the corresponding actual values of SSY. It is interesting to observe that even when SSY is zero or approximately zero at all eleven gauging stations in the MR basin, the best GA-ANN-51 forecasting model always produces a positive sediment value for low, high, and medium SSY. It was also concluded that, although the hybrid AI model was performing best for the sediment yield forecasting in the Mahanadi River, there is still some variability existing in the dataset that has not been captured by the model. The model variability or uncertainty may be due to human activities, runoff, or other unknown factors, which were not considered input parameters in this research but will be addressed in future research.

**Author Contributions:** Conceptualization, A.Y., P.C. and A.S.; methodology, M.A.A. and A.J.; software, R.M.Á., R.K.M. and D.J.; validation, A.Y. and A.S.; formal analysis, P.C. and M.A.A.; investigation, A.J. and R.M.Á.; resources, R.K.M.; writing—review and editing, A.S., M.A.A. and A.J.; writing—original draft preparation, A.Y. and P.C.; supervision, A.S. and D.J. All authors have read and agreed to the published version of the manuscript.

**Funding:** The authors would like to thank the Deanship of Scientific Research at Umm Al-Qura University for supporting this work by Grant Code: 22UQU4400257DSR03.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The authors would like to express their thanks to the Central Water Commission, Odisha for providing the data. The author honorable thanks the National Institute of Technology, Rourkela, India, for offering the necessary facilities for this research.

**Acknowledgments:** The authors would like to thank the Deanship of Scientific Research at Umm Al-Qura University for supporting this work by Grant Code: 22UQU4400257DSR03.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Yang, D.; Kanae, S.; Oki, T.; Koike, T.; Musiak, K. Global potential soil erosion with reference to land use and climate changes. *Hydrol. Process.* **2003**, *17*, 2913–2928. [[CrossRef](#)]
2. Kabir, M.A.; Dutta, D.; Hironaka, S. Estimating Sediment Budget at a River Basin Scale Using a Process-Based Distributed Modelling Approach. *Water Resour. Manag.* **2014**, *28*, 4143–4160. [[CrossRef](#)]
3. Gohil, P.P.; Saini, R. Coalesced effect of cavitation and silt erosion in hydro turbines—A review. *Renew. Sustain. Energy Rev.* **2014**, *33*, 280–289.
4. Kafle, K.R.; Khanal, S.N.; Dahal, R.K. Dynamics of the Koshi River on the Perspective of Morphology and Sedimentation with Emphasis on Post Disaster Impact of the 2008 Koshi Flood, Kathmandu Uni. *J. Sci. Eng. Tech.* **2015**, *11*, 71–92.
5. Pektas, A.O.; Cigizoglu, H.K. Long-range forecasting of suspended sediment. *Hydrol. Sci. J.* **2017**, *62*, 2415–2425. [[CrossRef](#)]

6. Gaurav, K.; Sinha, R.; Panda, P.K. The Indus flood of 2010 in Pakistan: A perspective analysis using remote sensing data. *Nat. Hazards* **2011**, *59*, 1815–1826. [[CrossRef](#)]
7. ZICL; Risk Nexus. *Urgent Case for Recovery: What We Can Learn from the August 2014 Karnali River Floods in Nepal*; Zurich Insurance Group Ltd.: Zurich, Switzerland, 2014.
8. Frémion, F.; Bordas, F.; Mourier, B.; Lenain, J.; Kestens, T.; Courtin-Nomade, A. Influence of dams on sediment continuity: A study case of a natural metallic contamination. *Sci. Total Environ.* **2016**, *547*, 282–294. [[CrossRef](#)]
9. Loucks, D.P.; Van Beek, E.; Stedinger, J.R.; Dijkman, J.P.; Villars, M.T. *Water Resources Systems Planning and Management: An Introduction to Methods, Models and Applications*; UNESCO: Paris, France, 2005.
10. Dutta, S. Soil erosion, sediment yield and sedimentation of reservoir: A review. *Model. Earth Syst. Environ.* **2016**, *2*, 123. [[CrossRef](#)]
11. Honorato, A.G.; Silva, G.B.; Guimarães Santos, C.A. Monthly streamflow forecasting using neuro-wavelet techniques and input analysis. *Hydrol. Sci. J.* **2018**, *63*, 2060–2075. [[CrossRef](#)]
12. Firat, M.; Güngör, M. Monthly total sediment forecasting using adaptive neuro fuzzy inference system. *Stoch. Environ. Res. Risk Assess.* **2010**, *24*, 259–270. [[CrossRef](#)]
13. Meshram, S.G.; Ghorbani, M.A.; Deo, R.C.; Kashani, M.H.; Meshram, C.; Karimi, V. New Approach for Sediment Yield Forecasting with a Two-Phase Feedforward Neuron Network-Particle Swarm Optimization Model Integrated with the Gravitational Search Algorithm. *Water Resour. Manag.* **2019**, *33*, 2335–2356. [[CrossRef](#)]
14. Terzi, Ö.; Ergin, G. Forecasting of monthly river flow with autoregressive modeling and data-driven techniques. *Neural Comput. Appl.* **2013**, *25*, 179–188. [[CrossRef](#)]
15. Seneviratna, D.M.K.N.; Rathnayaka, R.M.K.T. Rainfall data forecasting by SARIMA and BPNN model. *IOSR J. Math* **2017**, *6*, 57–63.
16. Salas, J.D.; Delleur, J.W.; Yevjevich, V.; Lane, W.L. *Applied Modeling of Hydrologic Time Series*; Water Resources Publications: Littleton, CO, USA, 1980; 484p.
17. Adamowski, J.F.; Fung Chan, H.; Prasher, S.O.; Ozga-Zieliński, B.; Sliusarieva, A.V. Comparison of multiple linear and nonlinear regression, autoregressive integrated moving average, artificial neural network, and wavelet artificial neural network methods for urban water demand forecasting in Montreal, Canada. *Water Resour. Res.* **2012**, *48*. [[CrossRef](#)]
18. ASCE. Task Committee on Application of Artificial Neural Networks in Hydrology Artificial neural networks in hydrology. I: Preliminary concepts. *J. Hydro. Eng.* **2000**, *5*, 115–123. [[CrossRef](#)]
19. Maier, H.R.; Dandy, G.C. Neural networks for the prediction and forecasting of water resource variables: A review of modelling issues and applications. *Environ. Model. Softw.* **2000**, *15*, 101–124. [[CrossRef](#)]
20. Dawson, C.W.; Harpham, C.; Wilby, R.L.; Chen, Y. Evaluation of artificial neural network techniques for flow forecasting in the River Yangtze, China. *Hydrol. Earth Syst. Sci.* **2002**, *6*, 619–626. [[CrossRef](#)]
21. Kar, A.K.; Lohani, A.K. Development of Flood Forecasting System Using Statistical and ANN Techniques in the Downstream Catchment of Mahanadi Basin, India. *J. Water Resour. Prot.* **2010**, *2*, 880–887.
22. Yang, S.; Yang, X.; Jiang, R.; Zhang, Y. New Optimal Weight Combination Model for Forecasting Precipitation. *Math. Probl. Eng.* **2012**, *2012*, 376010. [[CrossRef](#)]
23. Agarwal, A.; Rai, R.K.; Upadhyay, A. Forecasting of runoff and sediment yield using artificial neural networks. *J. Water Resour. Prot.* **2009**, *1*, 368–375. [[CrossRef](#)]
24. Bishop, M. *Neural Networks for Pattern Recognition*; Clarendon Press: Oxford, UK, 1998.
25. Chatterjee, S.; Bandopadhyay, S. Reliability estimation using a genetic algorithm-based artificial neural network: An application to a load-haul-dump machine. *Expert Syst. Appl.* **2012**, *39*, 10943–10951. [[CrossRef](#)]
26. Zhu, Y.; Lu, X.; Zhou, Y. Suspended sediment flux modeling with artificial neural network: An example of the Longchuanjiang River in the Upper Yangtze Catchment, China. *Geomorphology* **2007**, *84*, 111–125. [[CrossRef](#)]
27. Chatterjee, S.; Bandopadhyay, S. Global neural network learning using genetic algorithm for ore grade prediction of iron ore deposit. *Miner. Resour. Eng* **2007**, *12*, 258–269.
28. Yadav, A.; Joshi, D.; Kumar, V.; Mohapatra, H.; Iwendi, C.; Gadekallu, T.R. Capability and Robustness of Novel Hybridized Artificial Intelligence Technique for Sediment Yield Modeling in Godavari River, India. *Water* **2022**, *14*, 1917. [[CrossRef](#)]
29. Jain, A.; Srinivasulu, S. Development of effective and efficient rainfall-runoff models using integration of deterministic, real-coded genetic algorithms and artificial neural network techniques. *Water Resour. Res.* **2004**, *40*, 1–12. [[CrossRef](#)]
30. Asadi, S.; Shahrabi, J.; Abbaszadeh, P.; Tabanmehr, S. A New Hybrid Artificial Neural Networks for Rainfall-Runoff Process Modeling. *Neurocomputing* **2013**, *121*, 470–480. [[CrossRef](#)]
31. Adib, A.; Mahmoodi, A. Prediction of Suspended Sediment Load using ANN GA Conjunction Model with Markov Chain Approach at Flood Conditions. *KSCE J. Civ. Eng.* **2016**, *21*, 447–457. [[CrossRef](#)]
32. Nasser, M.; Asghari, K.; Abedini, M.J. Optimized scenario for rainfall forecasting using genetic algorithm coupled with artificial neural network. *Expert Syst. Appl.* **2008**, *35*, 1415–1421. [[CrossRef](#)]
33. Ding, Y.R.; Cai, Y.J.; Sun, P.D.; Chen, B. The use of combined neural networks and genetic algorithms for prediction of river water quality. *J. Appl. Res. Technol.* **2014**, *12*, 493–499. [[CrossRef](#)]
34. Tahmasebi, P.; Hezarkhani, A. Application of optimized neural network by genetic algorithm. In Proceedings of the International Association for Mathematical Geosciences Annual Conference, Stanford, CA, USA, 23–27 August 2009; 2, pp. 15–23.

35. Samarasinghe, S. *Neural Networks for Applied Sciences and Engineering: From Fundamentals to Complex Pattern Recognition*; CRC Press: Boca Raton, FL, USA, 2016; pp. 555–2016.
36. Yadav, A.; Chatterjee, S.; Equeenuddin, S.M. Suspended sediment yield estimation using genetic algorithm-based artificial intelligence models: Case study of Mahanadi River. *India. Hydrol Sci. J.* **2018**, *63*, 1162–1182. [[CrossRef](#)]
37. Chakrapani, G.J.; Subramanian, V. Factors controlling sediment discharge in the Mahanadi River Basin. *India. J. Hydrol.* **1990**, *117*, 169–185. [[CrossRef](#)]
38. Kant, A.; Suman, P.K.; Giri, B.K.; Tiwari, M.K.; Chatterjee, C.; Nayak, P.C.; Kumar, S. Comparison of multi-objective evolutionary neural network, adaptive neuro-inference system and bootstrap-based neural network for flood forecasting. *Neural Comput. Appl.* **2013**, *23*, 231–246. [[CrossRef](#)]
39. India-WRIS. Water Resources Information System of India. Available online: <http://india-wris.nrsc.gov.in/wrpinfo/index.php?title=Mahanadi> (accessed on 8 August 2015).
40. Chakrapani, G.J.; Subramanian, V. Rates of erosion and sedimentation in the Mahanadi River basin, India. *J. Hydrol.* **1993**, *149*, 39–48. [[CrossRef](#)]
41. Hagan, M.T.; Menhaj, M.B. Training feedforward networks with the Marquardt algorithm. *IEEE Trans. Neural Netw.* **1994**, *5*, 989–993. [[CrossRef](#)] [[PubMed](#)]
42. Goldberg, D. *Genetic Algorithms in Search, Optimization, and Machine Learning*; Addison-Wesley: Boston, MA, USA, 1989; p. 412.
43. Deb, K.; Agrawal, S.; Pratap, A.; Meyarivan, T. A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II. In *Parallel Problem Solving from Nature*; Springer: Berlin/Heidelberg, Germany, 2000; Volume 6, pp. 849–858.
44. Reddy, P.V.B. Modelling and Optimization of Wire Electrical Discharge Machining of Cr-Mo-V Special Alloy Steel Using Neuro Genetic Approach. Ph.D. Thesis, Jawaharlal Nehru Technological University, Anantapur, India, 2014.
45. Yadav, A.; Chatterjee, S.; Equeenuddin, S.M. Suspended sediment yield modeling in Mahanadi River, India by multi-objective optimization hybridizing artificial intelligence algorithms. *Int. J. Sediment Res.* **2021**, *36*, 76–91. [[CrossRef](#)]
46. Rosales-Perez, A.; Escalante, H.J.; Gonzalez, J.A.; Reyes-Garcia, C.A. Bias and variance optimization for SVMs model selection. In Proceedings of the Twenty-Sixth International FLAIRS Conference, St. Pete Beach, FL, USA, 22–24 May 2013.
47. Reinsel, G.C. *Elements of Multivariate Time Series Analysis*; Springer: New York, NY, USA, 1997.
48. Hyndman, R.J.; Athanasopoulos, G. *Forecasting: Principles and Practice*; OTexts: Melbourne, Australia, 2013; Available online: [www.OTexts.com/fpp](http://www.OTexts.com/fpp) (accessed on 6 October 2022).
49. Hejase, H.A.; Assi, A.H. Time-series regression model for prediction of mean daily global solar radiation in Al-Ain, UAE. *ISRN Renew. Energy* **2012**, *2012*, 412471. [[CrossRef](#)]
50. Ajewole, K.P.; Adejuwon, S.O.; Jemilohun, V.G. Test for stationarity on inflation rates in Nigeria using augmented dickey fuller test and Phillips-persons test. *J. Math* **2020**, *16*, 11–14.