
Inter-Comparison of an Evolutionary Programming Model of Suspended Sediment Time-Series with Other Local Models

M. A. Ghorbani, R. Khatibi, H. Asadi and P. Yousefi

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/47801>

1. Introduction

The experience of applying evolutionary computing to time series describing local physical problems has benefited the modelling culture by showing that many different mathematical formulae can be produced to describe the same problem. This experience brings into the focus the roles of pluralism in the modelling culture as opposed to searching for the best model, where physical problems provide relevance and context to the choice of modelling techniques. Both of these roles are often overlooked and do not directly influence research agenda. Although the focus of this paper is on evolutionary computing, it also promotes a pluralistic modelling culture by studying other modelling techniques, as well as by keeping the role of physical problems in the foreground.

Estimating suspended sediment loads is a problem of practical importance and includes such problems as changing courses in rivers, loss of fertile soil, filling reservoirs and impacts on water quality. The study of these problems in the short-run are referred to as sediment transport and erosion for those in the long-run. Past empirical capabilities remain invaluable but are not sufficient on their own as management and engineering solutions often require an insight into the problem. Empirical knowledge has been incorporated into the body of distributed modelling techniques giving rise to sophisticated modelling software tools but their applications require a great deal of resources. There remains a category of problems, often referred to as time series analysis, which uses the sequences of time variations and predicts the future values. This category of models provides useful information to management of local problems. For instance, such models may be used to schedule dredging requirements or other maintenance activities. Time series analysis is developing into local management tools and it is a focus of this chapter.

The aim of this chapter is to predict suspended sediment load of a river into the future. Besides the traditional empirical Sediment Rating Curve (SRC), there are several strategies for analysing such time series and *evolutionary computing* is one of Artificial Intelligence (AI) approaches, which broadly include capabilities for searching and recognising patterns among others. This chapter also employs Artificial Neural Network (ANN), which is another AI approach. Yet another strategy is to regard time series as outcomes of many random drivers and this assumption is supported by a whole body of probabilistic approaches, where this chapter uses Multi-Linear Regression (MLR) analysis to model the same data. Over the past few decades, research has increasingly focused on the application of deterministic chaos (or chaos theory or dynamic systems) showing that many of apparently randomly varying system behaviours can be explained by deterministic chaos. The concept behind this modelling strategy is that the particular data can largely be explained by deterministic behaviour, where in time the system evolves asymptotically towards an *attractor*. Its random-looking variations are assumed to be an internal feature of the system and depending on its initial conditions, its state under a certain range may become highly erratic but with a predictable behaviour. Evidently, none of these strategies are identical and different models rarely produce identical results. This chapter therefore compares the performance of these modelling strategies for solving an engineering problem.

The study employs 26 years of the Mississippi River data recorded at Tarbert + RR Landings and involve both flows and suspended sediment load. The river discharges about 200 million metric tons of suspended sediment per year to the Gulf of Mexico, where it ranks about sixth in the world today.

2. Literature review

Sediment Rating Curve (SRC) is an empirical approach used by practitioners in the engineering studies of sediment and erosion problems. The log linear rating curve method has been used widely and Sivakumar and Wallender (2005) outline the many flaws associated with this technique, including the lack of fit due to missing variables (e.g. Miller, 1951), retransformation bias (e.g. Ferguson, 1986), and non-normality of the error distribution (e.g. Thomas, 1988). According to Sivakumar and Wallender (2005), the technique has been modified including, among others, use of separate curves for different seasons (Miller, 1951), stratifying the data according to the magnitude of flow and applying a separate curve for each stratum (Glysson, 1987), and use of a single multivariate model instead of multiple rating curves (Cohn *et al.*, 1992). Sivakumar and Wallender (2005) argue that there is not a simple (and universal) 'water discharge-suspended sediment concentration-suspended sediment load' relationship. A brief overview of past studies is as follows.

Kisi, et al (2008) review the application of ANN and neuro-fuzzy techniques to time series analysis of sediment loads at various timescales, uncertainty in the data. Variations of these techniques have also been reported by Jain (2001), Tayfur (2002), Cigizoglu (2004), Kisi (2004), Raghuwanshi *et al.* (2006), Cigizoglu & Kisi (2006). Other studies on the application

of ANN to suspended sediment include that by Wang et al (2008), who applied ANN to derive the coefficients of regression analysis for their SRC model.

Aytek and Kishi (2008) used the GP approach to model suspended sediment for two stations on the Tongue River in Montana, USA, and indicate that the GP formulation performs quite well compared to sediment rating curves and multi linear regression models.

Chaotic signals have also been identified in time series of suspended sediment loads by Sivakumar and Jayawardena (2002, 2003), Farmer and Sidorowich, 1987). The outcomes revealed the usefulness of these methods towards an effective prediction capability.

Overall, a general understanding of the analysis of suspended sediment load is yet to emerge and one way to gain an insight into the problem is to carry out inter-comparison studies of the performance of a host of models applied to diversity of rivers of different shapes and sizes.

3. Study area and data

3.1. Understanding the problem

Sediment transport is concerned with entrained soil materials carried in water by erosion on the catchment and within channels. Sediment particles are categorised as follows (i) the saltation load (not discussed here); (ii) bedload (not discussed here) and (iii) suspended load including clay ($< 62\mu\text{m}$ in particle diameter), silt and sand. Suspended load (both as “fine-grained sediment” and “wash load”) is directly a result of the turbulence in water and forms a large proportion of the transported load, where the turbulence is a measure of the energy in the water to carry the load.

Sediment discharge is a measure of the mass rate of sediment transport at any point in space and time and determines whether the load is being transported or deposited. The whole process comprises soil erosion, sediment transport and sediment yield, where the deposited load delivered to a point in the catchment is referred to as sediment yield and is expressed as tons per unit area of the basin per year, measured at a point. Estimation of sediment yield (and soil erosion) is essential for management but these and mathematical models are used to gain an insight into the underlying processes. Sediment yield is estimated by (i) direct measurement, (ii) using local time series models to predict future states; (iii) using mathematical models to study jointly both erosion and sediment processes.

Suspended sediment forms most of the transported load and can be affected by many parameters including rainfall, land use pattern, slope, soil characteristics, e.g. soil moisture content but their considerations lead to distributed models, which are complex. Recorded suspended sediment derives distributed models by serving them as boundary conditions or input sources but their inherent information is not tapped on. There is a case for local models to study the information contained in recorded sediment loads alone in terms of flow and sediment hydrographs. This chapter is concerned with the study of the suspended load of a river, as discussed below.

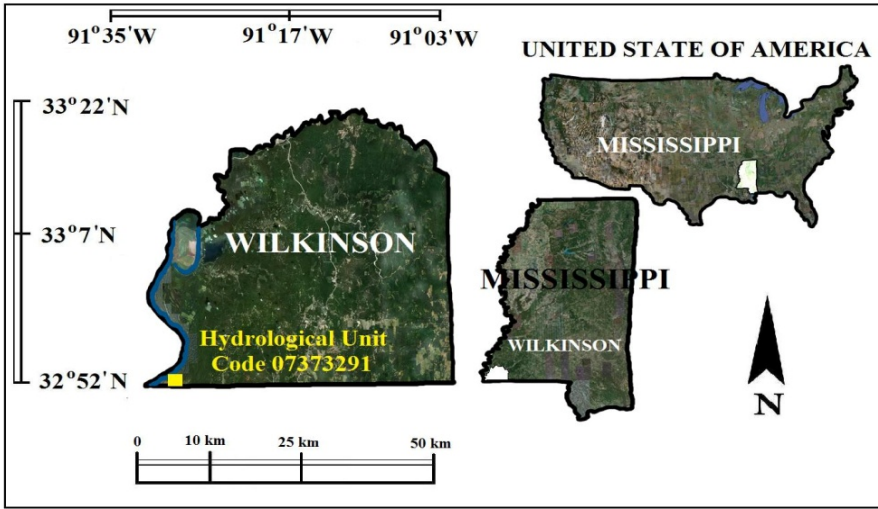


Figure 1. Mississippi River Station at Tarbert + RR Landings (<http://pubs.usgs.gov/circ/circ1133/geosetting.html>).

3.2. Study area

The flow–sediment time series data of a Mississippi river Station is used in the study, the location of which is shown in Figure 1. The gauge is situated at Tarbert + RR Landings, LA (USGS Station no. 07373291, latitude 30°57'40", longitude 91°39'52") and it is operated by the US Geological Survey (USGS) – the location map is shown in Figure 1. The Mississippi River discharges an average of about 200 million metric tons of suspended sediment per year to the Gulf of Mexico and to the ocean.

3.3. Review of data records

Daily suspended sediment measurements for the above station have been made available by the USGS from April 1949. The data used herein span over a period of about 26 years (amounting to 9496 datapoints) starting on October 1, 1949. Figures 2 show the variation of the daily suspended sediment and stream flow series observed at the above station.

Of the 26 water-years of the data sample of daily records of flow and suspended sediment (9496 datapoints), the first 25 water years of data (9131 datapoints) were used to train the models and the remaining 365 datapoints of daily records were used for testing. The statistical parameters of stream flow and sediment concentration data are shown in Table 1. These results show that the overall contribution of the datapoints in the test period is average; its individual characteristics in terms of kurtosis show the annual hydrographs to be less peaked and more flat but at the same time, the suspended sediment load during the year is significantly high. Thus, the minimum values during this year were significantly above the average but persistent and though less dynamic.

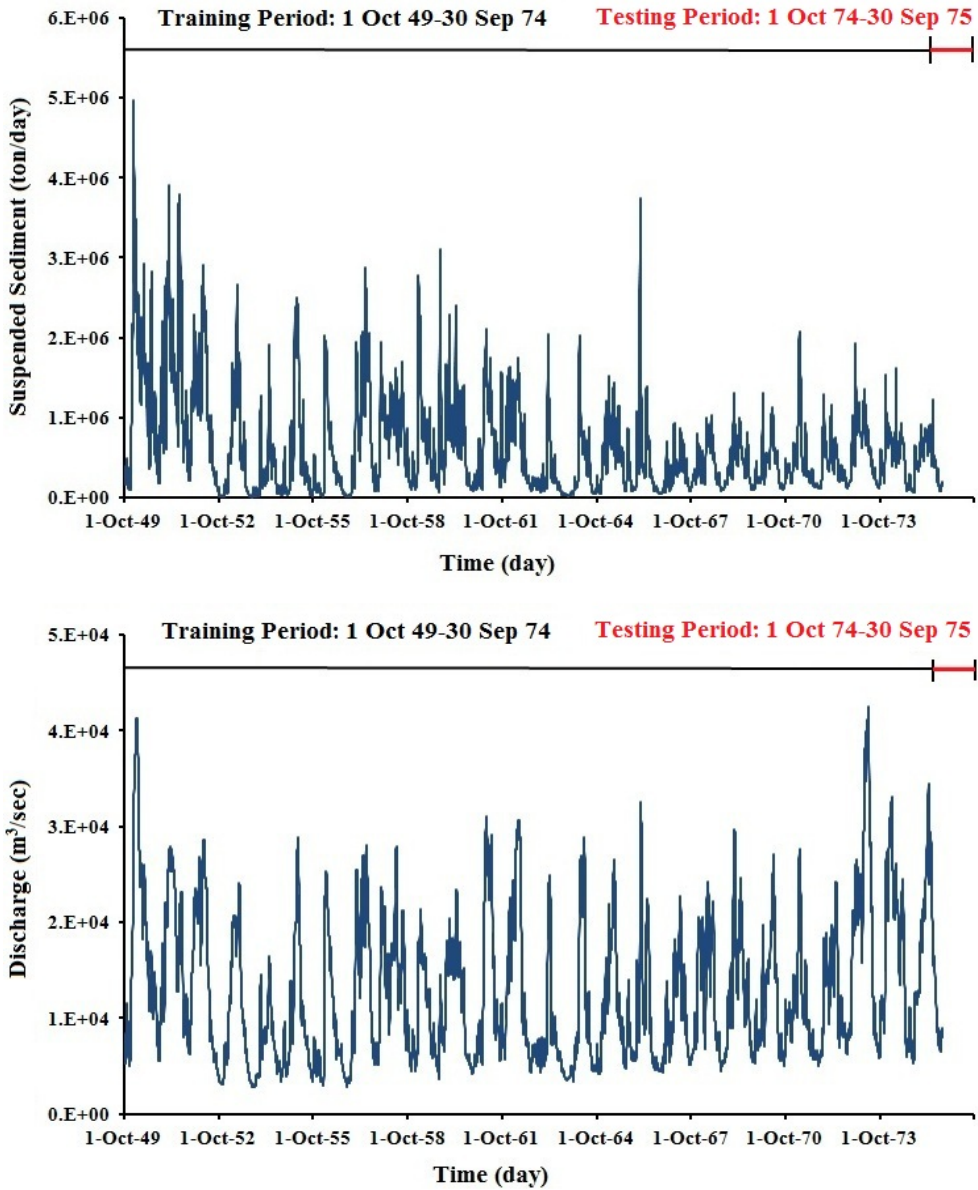


Figure 2. Variation of Daily Suspended Sediment and Flow Data in the Mississippi River Basin

Data Type	Training set		Testing set		All Dataset	
	Suspended Sediment (ton/day)	Discharge (m ³ /sec)	Suspended Sediment (ton/day)	Discharge (m ³ /sec)	Suspended Sediment (ton/day)	Discharge (m ³ /sec)
Datapoints	9131	9131	365	365	9496	9496
Mean	6.37E5	1.27E4	4.52E5	1.58E4	6.30E5	1.28E4
St dev	6.30E5	7.60E3	2.53E5	7.80E3	6.21E5	7.60E3
Max	4.97E6	4.25E4	1.22E6	3.45E4	4.97E6	4.25E4
Min.	4.00E3	2.80E3	6.70E4	5.40E3	4.00E3	2.80E3
CV	1.0	0.60	0.56	0.49	0.99	0.59
Skew	1.78	0.95	0.10	0.51	1.82	0.93
Kurt	3.98	0.34	-0.98	-0.85	4.20	0.27

*Data = Number of Data; Std = Standard Deviation; Max = Maximum Value; Min = Minimum Value; CV = Coefficient of Variation; Skew = Skewness; Kurt = Kurtosis

Table 1. Statistical Parameters for Dataset from the Mississippi River Basin

3.4. Overview of the models

The sediment rating curve method is the traditional method for converting measured flows to predict suspended sediment load and this paper aims to test the performance of evolutionary computing models but uses a host of other techniques for the inter-comparison purpose. These models are outlined in this section but evolutionary computing is explained in more detail. Their underlying notion is that past values contain a sufficient amount of information to predict the future values and a systematic way of representing this notion is purported in Table 2 in terms of a selection of models. These models, in essence, are reminiscent of regression analysis but GEP, ANN and MLR models approach the problem in their own individual ways to unearth the structure of the information inherent in time series. Notably, the SRC model is expressed by Model 1 and the deterministic chaos model is expressed by Model 0. These models will all be evaluated by using coefficient of Correlation (CC), Relative Absolute Errors (RAE) and Root Mean Square Errors (RMSE).

Model	Input variables	Output variables	The Structure
Model 0	$S_{t-1}, S_{t-2}...$	S_t	Chaos
Model 1	Q_t	S_t	ANN, SRC
Model 2	Q_t, S_{t-1}	S_t	GEP, ANN, MLR
Model 3	Q_t, Q_{t-1}	S_t	GEP, ANN, MLR
Model 4	Q_t, Q_{t-1}, S_{t-1}	S_t	GEP, ANN, MLR
Model 5	Q_t, Q_{t-1}, Q_{t-2}	S_t	GEP, ANN, MLR
Model 6	$Q_t, Q_{t-1}, Q_{t-2}, S_{t-1}$	S_t	GEP, ANN, MLR

Where Q_t and S_t represent respectively flow and suspended sediment load at day t .

Table 2. Modelling Structures of the Selected Modelling Techniques

3.4.1. *Sediment rating curve*

Sediment rating-curve is a flux-averaged relationship between suspended sediment, S , and water discharge, Q , expressed as a power law in the form of: $S = aQ^b$, where a and b are coefficients. Values of a and b for a particular stream are determined from data via a linear regression between $(\log S)$ and $(\log Q)$. The SRC model is represented in terms of Model 1 in Table 2. For more critical views on this model, references may be made Kisi (2005) and Walling (1977), among others.

3.4.2. *Evolutionary computing*

Evolutionary computing techniques apply optimisation algorithms as a tool to facilitate the mimicking of natural selection. A building block approach to generalised evolution driven by natural selection is yet to be presented, although Khatibi (2011) has outlined a rationale for it. Traditional understanding of natural selection for biological species is well developed, according to which the process takes place at the gene level of all individuals of all species carrying hereditary material for reproduction by inheriting from their parents and by passing on a range of their characteristics to their offspring. The process of reproduction is never a perfect copying process, as mutation may occur from time to time in biological reproductions involving the random process of reshuffling the genes during sexual reproduction. The paper assumes preliminary knowledge on genes, chromosomes, gene pool, DNA and RNA, where the environment also has a role to play. The environment for the production of proteins and sexual reproduction is different than the outer environment for the performance of the individual entities supported by the proteins or produced by sexual reproduction. The outer environment is characterised by (i) being limited in resources, (ii) having no foresight, (iii) organisms tend to produce more offspring than can be supported, a process that is driven by positive feedback loops, and (iv) there is a process of competition and selection. Some of these details are normally overlooked or simplified in evolutionary computing and therefore the paper stresses the point that natural selection takes place at the gene level and this is not directly applicable to that at the social level.

Facts on natural selection are overwhelming but there are myths as well, e.g. the myth of "the survival of the fittest" and this term is widely used in evolutionary computing. Although the fittest has a selective advantage to survive, this is not a guarantee for the survival in the natural world. An overview of the dynamics of natural selection in an environment is that (i) the environment can only support a maximum population of certain size, but there is also a lower size at the critical mass below which a population is at risk of losing its viability; (ii) there is a process of reproduction, during which natural selection operates at the gene level, although there are further processes operating at the individual levels beyond the direct reach of natural selection (e.g. interactions among the individuals catered for by other mechanisms or each individual is under selection pressure by the environment); (iii) the process of reproduction is associated with mutation, which gives rise to the production of gene pools.

A great deal of the above overview has been adopted in evolutionary computing, the history of which goes back to the 1960s when Rechenberg (1973) introduced evolution strategies. The variants of this approach include genetic algorithm (Holland, 1975), evolutionary programming (Fogel et al, 1966), genetic programming (Koza, 1992) and Gene Expression Programming (GEP), Ferreira (2001a). This paper uses the latter approach, which in a simple term is a variation of GP but each of these techniques have differences with one another. These techniques have the capability for deriving a set of mathematical expressions to describe the relationship between the independent and dependent variables using such functions as mutation, recombination (or crossover) and evolution.

This chapter is concerned with GEP and one of the important preliminary decisions in its implementations is to establish the models represented in Table 2 (Models 2-6). There is no prior knowledge of the appropriateness of any of these models and therefore this is normally fixed in a preliminary modelling task through a trial-and-error procedure. Whichever the model choice (Model 2 – Model 6 or similar other ones), each implementation of GEP builds up the model in terms of the values of the coefficients (referred to as terminals) and the operations (functions) through the procedure broadly outlined in Figure 3.

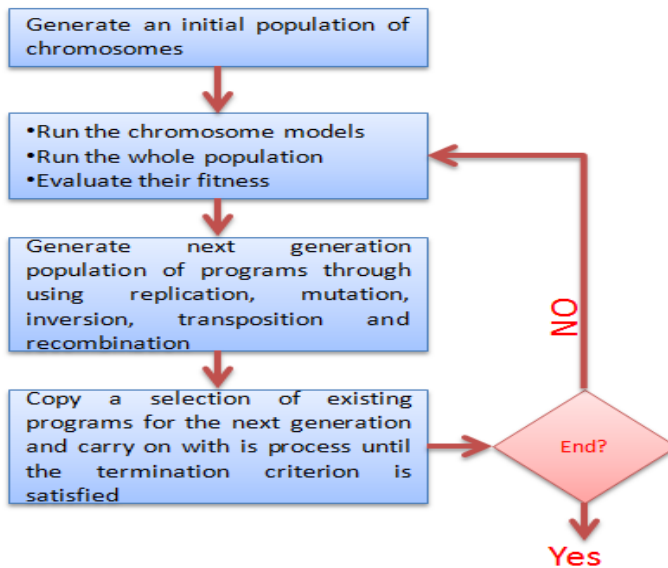


Figure 3. Simplified Outline of Implementation of Evolutionary Programming Models

The working of a gene expression program depicted in Figure 3 is outlined as follows. A chromosome in GEP is composed of genes and each gene is composed of (i) terminals and (ii) functions. The gene structures and chromosomes in GEP are illustrated for the solution that is obtained for the dataset used in this study (see Section 4.2). The terminals as their names suggest are composed of constants and variables and the functions comprise mathematical operations, as shown by (4.a)-(4.f).

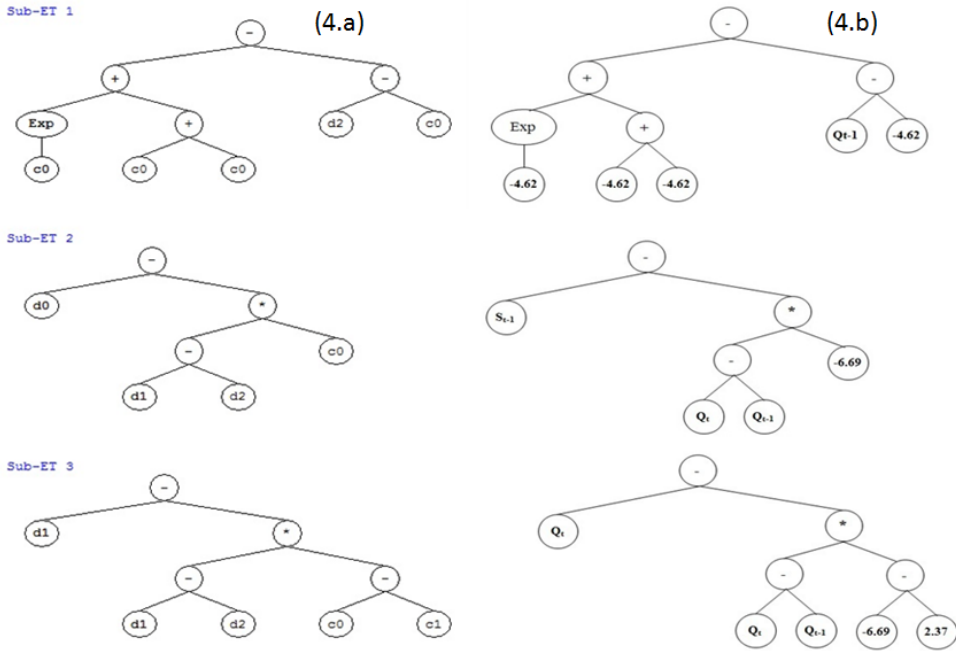


Figure 4. Expression Trees – (a) typical expression tree; (b) the selected GEP model in this study

As the term terminal suggests, it comprises a set of values at the tail-ends of the genes of the chromosomes and these are made meaningful by the functions making up the other component of the genes of the chromosomes. In GEP, these are represented by a bilingual notation called Karva language of (i) genetic codes, which are not deemed necessary for a description here and (ii) expression trees (or parse trees), as illustrated by Figure 4.a and the recommended solution is shown in Figure 4.b, which is transcribed by Equation (4) in Section 4.2. The initial chromosomes of the initial population are no different than the solution shown in Figure 4.b but their difference is that the composition of each of the initial chromosomes is selected often in random and then GEP is expected to improve them through evolution by the strategy of selections, replication and mutation but there are other facilities that not mentioned facilitating a more robust solution and these include inversion, transposition and recombination. The improvements are carried out through selection from one generation to another and this is why this modelling strategy is called evolutionary computation. The main strength of this approach is that it does not set up any system of equation to predict the future but it evaluates the fitness of each chromosome and selects from those a new population with better performance traits.

The GEP employed in this study is based on evolving computer programs of different sizes and shapes encoded in linear chromosomes of fixed lengths, Ferreira, 2001a; Ferreira, (2001b). The chromosomes are composed of multiple genes, each gene encoding a smaller subprogram. Furthermore, the structural and functional organisation of the linear

chromosomes allows the unconstrained operation of important genetic functions, such as mutation, transposition and recombination. It has been reported that GEP is 100-10,000 times more efficient than GP systems (Ferreira, 2001a; Ferreira, 2001b) for a number of reasons, including: (i) the chromosomes are simple entities: linear, compact, relatively small, easy to manipulate genetically (replicate, mutate, recombine, etc); (ii) the parse trees or expression trees are exclusively the expression of their respective chromosomes; they are entities upon which selection acts, and according to fitness, they are selected to reproduce with modification.

3.4.3. *Artificial Neural Networks (ANNs)*

Whilst evolutionary programming emulates the working of Nature, ANNs emulate the workings of neurons in the brain. Both the brain and ANNs are parallel information processing systems consisting of a set of neurons or nodes arranged in layers but this is where the parallel ends. The actual process of information processing in the brain is a topical research issue but the drivers of ANNs are polynomial algebra and there is no evidence that the brains of humans, monkeys or any other animals employ algebraic computations such as optimisation methods. Although there is a great incentive to understand the working of the brain, it is not imperative to be constrained by it and the use of algebra in ANNs is not criticised here but awareness is raised as these two processes are not identical.

The ANN theory has been described in many books, including the text by Rumelhart et al. (1986). The application of ANNs has been the subject of a large number of papers that have appeared in the recent literature. There are various implementations of ANNs but the type used in this study is a Multi-Layer feedforward Perceptron (MLP) trained with the use of back propagation learning algorithm with the following functions: (i) the input layer accepts the data, (ii) intermediate layer processes them, and (iii) the output layer displays the resultant outputs. The number of hidden layers is decided in a preliminary modelling process by finding the most efficient number of layers through a trial-and-error procedure. Each neuron in a layer is connected to all the neurons of the next layer, and the neurons in one layer are not connected among themselves. All the nodes within a layer act synchronously.

This study implements the ANN models in terms of Models 1-6 of Table 2 and Figure 5 shows one of the implementation selected. For each of these models, the data passing through the connections from one neuron to another are multiplied by weights that control the strength of a passing signal. When these weights are modified, the data transferred through the network changes; consequently, the network output also changes. The signal emanating from the output node(s) is the network's solution to the input problem.

In the back-propagation algorithm, a set of inputs and outputs is selected from the training set and the network calculates the output based on the inputs. This output is subtracted from the actual output to find the output-layer error. The error is back propagated through the network, and the weights are suitably adjusted. This process continues for the number of

prescribed sweeps or until a prescribed error tolerance is reached. The mean square error over the training samples is the typical objective function to be minimized. After training is complete, the ANN performance is validated. Depending on the outcome, either the ANN has to be retrained or it can be implemented for its intended use.

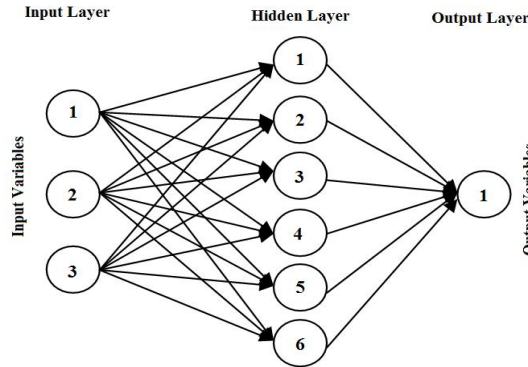


Figure 5. Implementation of the ANN Models and its Various Layers

3.4.4. Multi Linear Regression (MLR)

An overview of the data presented in Figure 2 invokes the thought that other than the annual trend within the data, the underlying process is probably random and a more rational way of explaining the data would be through probabilistic approaches. One such method applied to the selected data is the Multi-Linear Regression (MLR) model. It fits a linear combination of the components of a multiple signals x (e.g. recorded flows and suspended sediment timeseries as defined by the Models 2-6 in Table 2) to a single output signal y , as defined by (1.a) (e.g. predicted suspended sediment load) and returns the residual, r , i.e. the difference signal, as defined by (1.b):

$$y = b + \sum_{i=0}^N a_i x_i \tag{1a}$$

$$r = y - a_1 x_1 - a_2 x_2 - \dots - b \tag{1b}$$

Where x_i is defined in Table 2 in terms of various models and a_i values are called regression coefficients, which are estimated by using the least square or any other similar method. In this study, the coefficients of the regressions were determined using the least square method.

3.4.5. Chaos theory

A cursory view of the suspended sediment record of the Mississippi River in Figure 2 provides no clue to a strategy for its underlying patterns, if any, although annual trend

superimposed on random variation may be an immediate reaction of a hydrologist. Another strategy to explore such possible patterns is through the application of chaos theory, more specifically through the “phase-space diagram” as shown in Figure 6 for this river data. A point in the phase-space represents the state of the system at a given time. The narrow dark band in the figure signifies strong determinism but its scattered band signifies the presence of noise and therefore there is a possibility to explain this set of data by chaos theory. The dark band signifies convergence of the trajectories of the phase-space with a fractal dimension towards the attractor of the data, where the dynamics of the system can be reduced to a set of deterministic laws to enable the prediction of its future states.

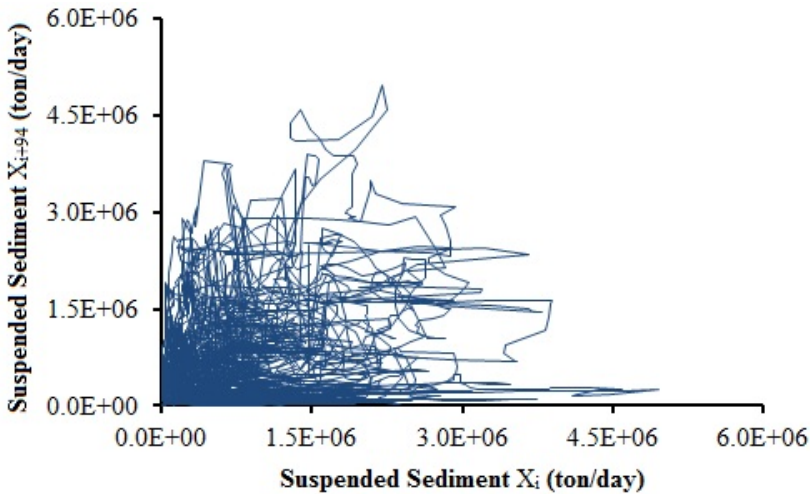


Figure 6. Phase-space Diagram of Daily Suspended Sediment Data in the Mississippi River Basin

Chaos theory is a method of nonlinear time series analysis and involves a host of methods, essentially based on the phase-space reconstruction of a process, from scalar or multivariate measurements of physical observables. This method is implemented in terms of Model 0 of Table 2. It is largely based on the representation of the underlying dynamics through reconstruction of phase-space, originally given by Takens, 1981. It is implemented in terms of two parameters of delay time and embedding dimension, according to which given a set of physical variables and an analytical model describing their interactions, the dynamics of the system can be represented geometrically by a single point moving on a trajectory, where each of its points corresponds to a state of the system. The phase-space diagram is essentially a co-ordinate system, whose coordinates represent the variables necessary to completely describe the state of the system at any moment.

One difficulty in its construction is that in most practical situations, information on every variable influencing the system may not be available. However, a time series of a single variable may be available, which may allow the construction of a (pseudo) phase-space. The idea behind such a reconstruction is that a non-linear system is characterized by self-

interaction, and a time series of a single variable can carry the information about the dynamics of the entire multi-variable system. The trajectories of the phase-space diagram describe the evolution of the system from some initial state, and hence represent the history of the system.

This paper applies chaos theory to analyse the suspended sediment load of the Mississippi River data in a similar fashion to the other modelling strategies described above. It uses the local prediction method for training and testing, as outlined below, but it is a traditional practice to apply several methods to build evidence for the existence of chaotic signals in a particular data. These techniques employ the delay-embedding parameters of τ and m , which are unknown a-priori. The following methods are used in this chapter:

1. Average Mutual Information (AMI) is used to estimate τ ; and the minimization of the False Nearest Neighbours to do that of the optimal values for the embedding dimension, m .

AMI (Fraser and Swinney, 1986) defines how the measurements $X(t)$ at time t are related, from an information theoretic point of view, to measurements $X(t + \tau)$ at time $t + \tau$. The average mutual information is defined as:

$$I(\tau) = \sum_{X(i), X(i+\tau)} P(X(i), X(i + \tau)) \log \left[\frac{P(X(i), X(i + \tau))}{P(X(i))P(X(i + \tau))} \right] \quad (2a)$$

where i is total number of samples. $P(X(i))$ and $P(X(i + \tau))$ are marginal probabilities for measurements $X(i)$ and $X(i + \tau)$, respectively, whereas $P(X(i), X(i + \tau))$ is their joint probability. The optimal delay time τ minimises the function $I(\tau)$ for $t = \tau$, $X(i + \tau)$ adds the maximum information on $X(i)$.

The False Nearest Neighbours procedure (Kennel et al., 1992) is a method to obtain the optimum embedding dimension for phase-space reconstruction. By checking the neighbourhood of points embedded in projection manifolds of increasing dimension, the algorithm eliminates 'false Neighbours': This means that points apparently lying close together due to projection are separated in higher embedding dimensions. when the ratio between the number of false neighbours at the dimension $m + 1$ and m is below a given threshold, generally smaller than 5%, each $m' > m + 1$ is an optimal embedding. A poor reconstruction of few embedding states with several components is obtained if m' is too large and the next analyses should not be performed.

2. **Correlation Dimension (CD) method:** is a nonlinear measure of the correlation between pairs lying on the attractor. For time series whose underlying dynamics is chaotic, the correlation dimension gets a finite fractional value, whereas for stochastic systems it is infinite. For an m -dimensional phase-space, the correlation function $C_m(r)$ is defined as the fraction of states closer than r (Grassberger and Procaccia, 1983; Theiler, 1986):

$$C(r) = \lim_{N \rightarrow \infty} \frac{2}{N(N-1)} \sum_{i,j=1}^N H(r - |Y_i - Y_j|) \quad (2b)$$

where H is the Heaviside step function, \tilde{Y}_i is the i^{th} state vector, and N is the number of points on the reconstructed attractor. The number w is called Theiler window and it is the correction needed to avoid spurious results due to temporal correlations instead of dynamical ones. For stochastic time series $C_m(r) \propto r^m$ holds, whereas for chaotic time series the correlation function scales with r as:

$$C_m(r) \propto r^{D_2} \tag{2c}$$

where D_2 , correlation exponent, quantifies the degrees of freedom of the process, and defined by:

$$D_2 = \lim_{r \rightarrow 0} \frac{\ln C(r)}{\ln r} \tag{2d}$$

and can be reliably estimated as the slope in the $\ln C_m(r)$ vs. $\ln(r)$ plot.

3. **Local Prediction Model:** The author’s implementation of the local prediction method for deterministic chaos is details in Khatibi et al (2011) but the overview is that a correct phase-space reconstruction in a dimension m facilitates an interpretation of the underlying dynamics in the form of an m -dimensional map, f_T , according to

$$Y_{j+T} = f_T(Y_j) \tag{2e}$$

where Y_j and Y_{j+T} are vectors of dimension m , describing the state of the system at times j (i.e. the current state) and $j+T$ (i.e. the future state), respectively. The problem then is to find an appropriate expression for f_T (i.e. F_T). Local approximation entails the subdivision of the f_T domain into many subsets (neighbourhoods), each of which identifies some approximations F_T , valid only in that same subset. In other words, the dynamics of the system is described step-by-step locally in the phase-space. In this m -dimensional space, prediction is performed by estimating the change of X_i with time, which are observed values of discrete scalar timeseries, with delay coordinates in the m -dimensional phase space. The relation between the points X_t and X_{t+p} (the behaviour at a future time p on the attractor) is approximated by function F as:

$$X_{t+p} \cong F(X_t) \tag{2f}$$

In this prediction method, the change of X_t with time on the attractor is assumed the same as those of nearby points, $(X_{T_h}, h = 1, 2, \dots, n)$. Herein, X_{t+p} is determined by the d^{th} order polynomial $F(X_t)$ as follows (Itoh, 1995):

$$x_{t+p} \cong f_0 + \sum_{k_1=0}^{m-1} f_{1k_1} X_{t-k_1\tau} + \sum_{\substack{k_2=k_1 \\ k_1=0}}^{m-1} f_{2k_1k_2} X_{t-k_1\tau} X_{t-k_2\tau} + \dots + \sum_{\substack{k_d=k_{d-1} \\ \vdots \\ k_2=k_1 \\ k_1=0}}^{m-1} f_{dk_1k_2\dots k_d} X_{t-k_1\tau} X_{t-k_2\tau} \dots X_{t-k_d\tau} \tag{2g}$$

Using n of X_{T_h} and $X_{T_{h+p}}$ for which the values are already known, the coefficients, f_i , are determined by solution of the following equation:

$$X \equiv Af \tag{2h}$$

where

$$X = (X_{T_{1+p}}, X_{T_{2+p}}, \dots, X_{T_{n+p}}) \tag{2i}$$

and $f = (f_0, f_{10}, f_{11}, \dots, f_{1(m-1)}, f_{200}, \dots, f_{d(m-1)(m-1)\dots(m-1)}) \tag{2j}$

and A is the $n \times (m+d)!/m!d!$ Jacobian matrix which in its explicit form is:

$$A = \begin{bmatrix} 1 & X_{T_1} & X_{T_{1-\tau}} & \dots & X_{T_{1-(m-1)\tau}} & X_{T_1}^2 & \dots & X_{T_{1-(m-1)\tau}}^d \\ 1 & X_{T_2} & X_{T_{2-\tau}} & \dots & X_{T_{2-(m-1)\tau}} & X_{T_2}^2 & \dots & X_{T_{2-(m-1)\tau}}^d \\ \vdots & & & & & \vdots & & \vdots \\ 1 & X_{T_n} & X_{T_{n-\tau}} & \dots & X_{T_{n-(m-1)\tau}} & X_{T_n}^2 & \dots & X_{T_{n-(m-1)\tau}}^d \end{bmatrix} \tag{2k}$$

In order to obtain a stable solution, the number of rows in the Jacobian matrix A must satisfy:

$$n \geq \frac{(m+d)!}{m!d!} \tag{2l}$$

As stated by Porporato and Ridolfi (1997), even though F -values are first degree polynomials, the prediction is nonlinear, because during the prediction procedure every point $x(t)$ belongs to a different neighbourhood and is therefore defined by different expressions for f (Koçak,1997).

4. Setting up models and preliminary results

4.1. Performance of sediment rating curve

The SRC model was implemented by using a simple least squares method leading to

$$S=13.2Q^{1.14} \tag{3}$$

The performance of this model is summarised in Table 3 and shown in Figure 7. Evidently, its performance is poor and the concern raised in the literature on this model is confirmed. This is a sufficient justification to search for reliable models.

Model Input	Training			Testing		
	CC	MAE	RMSE	CC	MAE	RMSE
Model 1: Q_t	0.76	2.62E5	4.11E5	0.82	3.89E5	4.86E5

Table 3. Statistical Performance of the Sediment Rating Curve for the Training and Test Periods

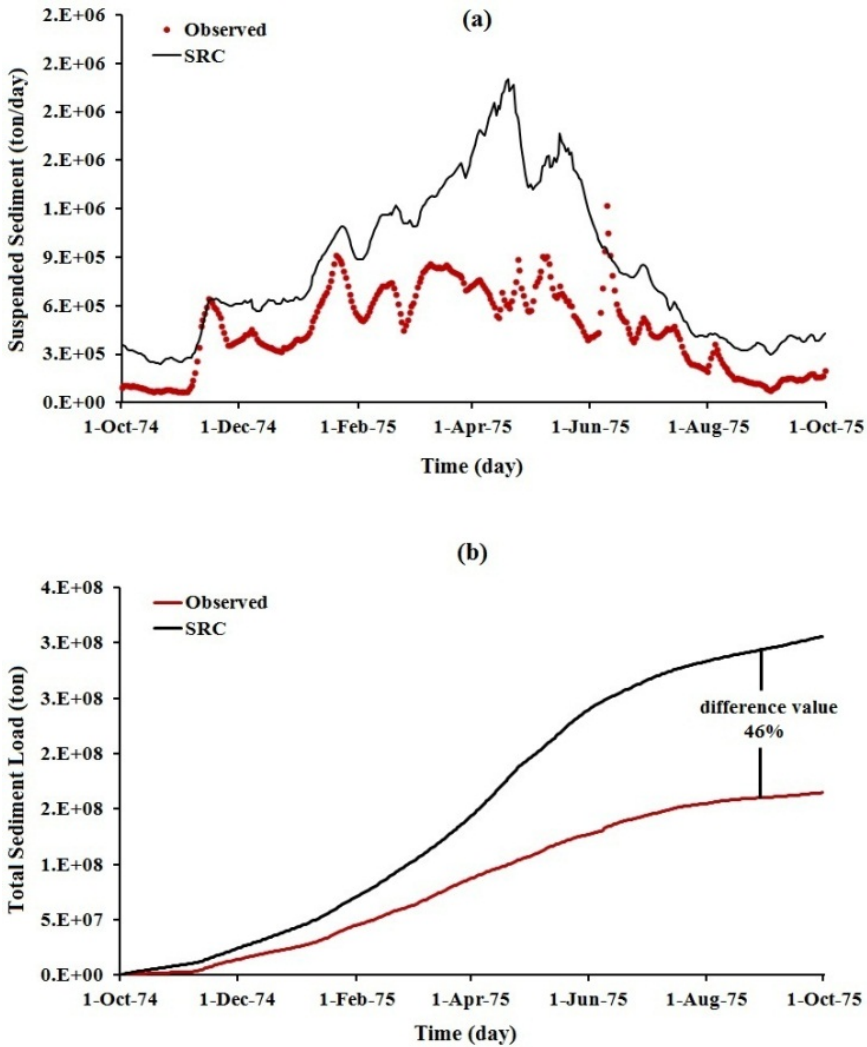


Figure 7. Comparison of Observed Suspended Sediment with that Modelled by SRC; (a) hydrograph, (b) cumulative values

4.2. Implementation of GEP

The preliminary investigation for the construction of a relationship between flows and suspended sediment in GEP requires: (i) the setting of the functions, as discussed below; (ii) the fitness function; and (iii) a range of other parameters, but the default values, given in Table 11, were sufficient in this study. The following functions were investigated:

$$\{+, -, \times\} \tag{4a}$$

$$\{+,-, \times, x\} \tag{4b}$$

$$\{+,-, \times, x^2\} \tag{4c}$$

$$\{+,-, \times, x^3\} \tag{4d}$$

$$\{+,-, \times, e^x\} \tag{4e}$$

$$\{+,-, \times, \ln(x)\} \tag{4f}$$

The performance of each function was investigated in terms of CC, MAE, and RMSE and the results are shown in Table 4.a for the training periods. The results show that (i) the model performances are more sensitive to the choice of independent variables than the function choices; (ii) the models not including suspended sediment time series perform poorly; and (iii) the model performance is not overly sensitive to the choice of the function. Appendix I, Table 11 specifies the fitness function to be Root Relative Squared Errors (RRSE).

Model		Q_t	Q_t, S_{t-1}	Q_t, Q_{t-1}	Q_t, Q_{t-1}, S_{t-1}	Q_t, Q_{t-1}, Q_{t-2}	$Q_t, Q_{t-1}, Q_{t-2}, S_{t-1}$
(4.a): {+,-,×}	CC	0.77	0.99	0.78	0.99	0.78	0.99
	MAE	2.79E5	3.88E4	2.78E5	3.25E4	2.79E5	3.37E4
	RMSE	4.13E5	8.43E4	4.07E5	7.74E4	4.05E5	7.98E4
(4.b): {+,-,×,x}	CC	0.77	0.99	0.77	0.99	0.77	0.99
	MAE	2.82E05	3.87E04	2.81E05	3.80E04	2.78E05	3.27E04
	RMSE	4.15E05	8.42E04	4.14E05	8.36E04	4.10E05	7.75E04
(4.c): {+,-,×,x ² }	CC	0.77	0.99	0.78	0.99	0.78	0.99
	MAE	2.82E05	3.89E04	2.78E05	3.25E04	2.76E05	3.45E04
	RMSE	4.15E05	8.43E04	4.08E05	7.74E04	4.05E05	8.02E04
(4.d): {+,-,×,x ³ }	CC	0.77	0.99	0.77	0.99	0.77	0.99
	MAE	2.43E5	3.89E4	2.76E5	3.21E4	2.76E5	3.41E4
	RMSE	4.05E5	8.43E4	4.12E5	7.76E4	4.13E5	8.03E4
(4.e): {+,-,×,e ^x }	CC	0.77	0.99	0.77	0.99	0.77	0.99
	MAE	2.81E5	3.88E4	2.81E5	3.56E4	2.42E5	3.64E4
	RMSE	4.15E5	8.43E4	4.14E5	8.16E4	4.00E5	8.23E4
(4.f): {+,-,×,ln(x)}	CC	0.76	0.99	0.76	0.99	0.78	0.99
	MAE	2.56E5	3.89E4	2.64E5	3.25E4	2.60E5	3.21E4
	RMSE	4.09E5	8.42E4	4.10E5	7.72E4	4.02E5	7.72E4

Table 4. a. Statistical Performance of a Selection of Functions for the Training Period

The performance of the GEP model is presented in Table 4.b, according to which there is not much to differences between performances of a number of the alternative models but (4.e) is selected in this study for the prediction purposes (its expression tree is given in Figure 4) and given below.

$$S_t = S_{t-1} + 16.77Q_t - 16.77Q_{t-1} - 13.87 \tag{5}$$

Model		Q_t	Q_t, S_{t-1}	Q_t, Q_{t-1}	Q_t, Q_{t-1}, S_{t-1}	Q_t, Q_{t-1}, Q_{t-2}	$Q_t, Q_{t-1}, Q_{t-2}, S_{t-1}$
(4.a): {+, -, ×}	CC	0.83	0.99	0.84	0.99	0.84	0.99
	MAE	3.99E5	2.34E4	4.01E5	2.32E4	4.08E5	2.08E4
	RMSE	4.72E5	3.87E4	4.73E5	4.06E4	4.79E5	3.75E4
(4.b) {+, -, ×, x}	CC	0.83	0.99	0.84	0.99	0.84	0.99
	MAE	3.99E05	2.33E04	4.01E05	2.28E04	3.96E05	2.34E04
	RMSE	4.69E05	3.87E04	4.70E05	3.08E04	4.65E05	4.04E04
(4.c) {+, -, ×, x ² }	CC	0.83	0.99	0.84	0.99	0.84	0.99
	MAE	4.00E05	2.35E04	4.01E05	2.32E04	3.79E05	2.31E04
	RMSE	4.69E05	3.86E04	4.71E05	4.06E04	4.67E05	3.96E04
(4.d): {+, -, ×, x ³ }	CC	0.83	0.99	0.83	0.99	0.83	0.99
	MAE	3.87E5	2.35E4	4.00E5	2.10E4	3.96E5	2.06E4
	RMSE	4.94E5	3.87E4	4.75E5	3.81E4	4.69E5	3.69E4
(4.e): {+, -, ×, e ^x }	CC	0.83	0.99	0.84	0.99	0.84	0.99
	MAE	3.98E5	2.35E4	3.97E5	2.06E4	3.80E5	2.17E4
	RMSE	4.67E5	3.86E4	4.66E5	3.64E4	4.84E5	3.73E4
(4.e): {+, -, ×, ln(x)}	CC	0.83	0.99	0.83	0.99	0.83	0.99
	MAE	3.90E5	2.35E4	3.90E5	2.40E4	3.81E5	2.29E4
	RMSE	4.93E5	3.86E4	4.81E5	4.18E4	4.69E5	4.05E4

Table 4 b. Statistical analysis of the estimated values for the test period

Figure 8 compares modelled suspended sediment against their observed values, according to which the improvement by GEP is remarkable compared with SRC. Overall, the GEP modelling results follow observed values rather faithfully both in values and patterns, although there are still discrepancies in predicted values.

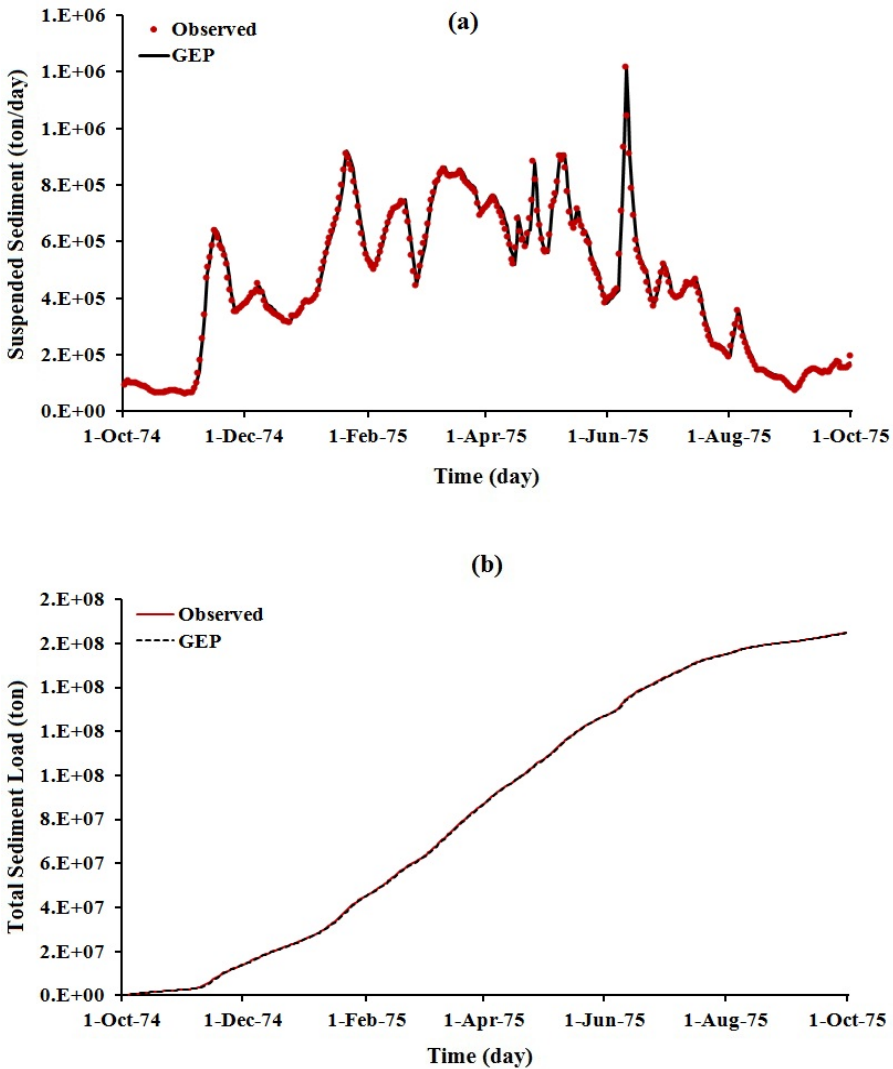


Figure 8. Comparison of Observed Suspended Sediment with that Modelled by GEP; (a) hydrograph, (b) cumulative values

4.3. Implementation of ANN

ANN implements another AI approach to the data represented in Figure 2 by another strategy, as described in Section 3.4.3. A preliminary investigation was carried out to make decisions on the choice of the models given in Table 2 (Models 1-6) and the ANN structure in terms of the neuron structure of the various layers. Table 5 presents model structures investigated. The preliminary modelling task also included a normalisation function for the

data. In this study, MATLAB was employed to develop the ANN model and its *mapstd* function was selected for the normalisation (further defaults values are given in Table 11). The investigated ANN model structures are defined in Table 5 and their results for both the training and testing periods are presented in Table 6.

Model Identifier	Model Inputs	Training	Testing
Model 1	Q_t	2-5-1	2-5-1
Model 2	Q_t, S_{t-1}	3-5-1	3-5-1
Model 3	Q_t, Q_{t-1}	3-7-1	3-7-1
Model 4	Q_t, Q_{t-1}, S_{t-1}	4-6-1	4-6-1
Model 5	Q_t, Q_{t-1}, Q_{t-2}	4-9-1	4-9-1
Model6	$Q_t, Q_{t-1}, Q_{t-2}, S_{t-1}$	5-12-1	5-12-1

Table 5. ANN Structure (number of nodes in layers)

The performances of Models 1-6 are shown in Table 6 in terms of the values of three statistical indices of CC, MAE and RMSE. The performance of different models in terms of CC is remarkably high but Model 4 (Q_t, Q_{t-1}, S_{t-1}) produce less deviations, which is selected for the final run.

Model Inputs	Model Training			Model Testing		
	CC	MAE	RMSE	CC	MAE	RMSE
Q_t	0.999	2.32E4	2.70E4	0.999	2.16E4	2.30E4
Q_t, S_{t-1}	0.999	2.59E4	3.12E4	0.996	2.17E4	2.64E4
Q_t, Q_{t-1}	0.999	2.00E4	2.79E4	0.981	4.19E4	4.84E4
Q_t, Q_{t-1}, S_{t-1}	0.999	2.01E4	2.51E4	0.998	1.18E4	1.37E4
Q_t, Q_{t-1}, Q_{t-2}	0.991	7.57E4	8.42E4	0.942	8.47E4	8.41E4
$Q_t, Q_{t-1}, Q_{t-2}, S_{t-1}$	0.995	5.66E4	6.42E4	0.976	4.63E4	5.47E4

Table 6. Statistical Performance of the Selected Model Structure for the Training and Testing periods

4.4. Implementation of the MLR model

The MLR modelling strategy was implemented using Mathematica to derive regression coefficients for both periods of model fitting (training in the AI terminology) and testing using different statistical indices (CC, MAE and RMSE) given in Table 7, which shows that Model 2 (Q_t, S_{t-1}) performs relatively better than the others. The regression equation suggested by this technique is given by:

$$S_t = 0.24Q_t + 30.99S_{t-1} \tag{6}$$

Model Inputs	Model Training			Model Testing		
	CC	MAE	RMSE	CC	MAE	RMSE
Q_t	0.77	2.41E5	4.05E5	0.833	3.85E5	4.96E5
Q_t, S_{t-1}	0.994	3.90E4	8.40E4	0.988	2.40E4	3.90E4
Q_t, Q_{t-1}	0.78	2.40E5	3.97E5	0.837	3.85E5	4.98E5
Q_t, Q_{t-1}, S_{t-1}	0.993	3.30E4	7.60E4	0.987	2.50E4	4.10E4
Q_t, Q_{t-1}, Q_{t-2}	0.779	2.40E5	3.95E5	0.840	3.84E5	4.96E5
$Q_t, Q_{t-1}, Q_{t-2}, S_{t-1}$	0.993	3.30E4	7.70E4	0.987	2.50E4	4.10E4

Table 7. Statistical analysis of the estimated values for the train and test period

4.5. Implementation of the deterministic chaos model

A visual assessment for the existence of chaotic behaviour in the suspended sediment time series was presented in Figure 9, although it was not conclusive evidence but just invoked the possibility of the existence of a low-dimensional chaos. Traditionally, several techniques are employed to show the existence of low-dimensional chaos and below the results of the determination of the dimensions of the phase-state diagram are given:

1. Using the AMI method, the delay time, is estimated for the data as the intercept with the x-axis of the curves by plotting the values of the AMI evaluated by the TISEAN package (Hegger *et al.*, 1999) against delay times progressively increased from 1 to 100. The value of delay time is calculated as the first (local) minimum in the variation of AMI against varying delay time from 1 to 100 day. The results are shown in Figure (9.a), signifying a well-defined first minimum at delay time of 94 day. The delay time is then used in the determination of the sufficient embedding dimension using the percentage of false nearest neighbours for the time series. Figure (9.b) shows the results of the false nearest neighbours method for embedding dimension m , by allowing it to vary from 1 to 40 and hence its value is 28.
2. The presence of chaotic signals in the data is further confirmed by the correlation dimension method. Figure (10.a) shows the relationship between correlation function $C(r)$ and radius r (i.e. $\ln C(r)$ versus $\ln(r)$) for increasing m , whereas Figure (10.b) shows the relationship between the correlation dimension values $D_2(m)$ and the embedding dimension values m . It can be seen from Figure (10.b) that the value of correlation exponent increases with the embedding dimension up to a certain value and then saturates beyond it. The saturation of the correlation exponent is an indication of the existence of deterministic dynamics. The saturated correlation dimension is 3.5, ($D_2=3.5$). The value of correlation dimension also suggests the possible presence of chaotic behaviour in the dataset. The nearest integer above the correlation dimension value ($D_2=4$) is taken as the minimum dimension of the phase space.

- Local prediction algorithm is used to predict suspended sediment time series. The procedure involves varying the value of the embedding dimension in a range, say 3-8, and estimating the CC and RMSE. The embedding function with the highest coefficient of correlation is selected as the solution. These are given in Table 8 for Mississippi River basin for the dataset with daily time interval, as well as a selection of other time steps. It shows that the best predictions are achieved when the embedding dimension is $m=3$ produce the best results.

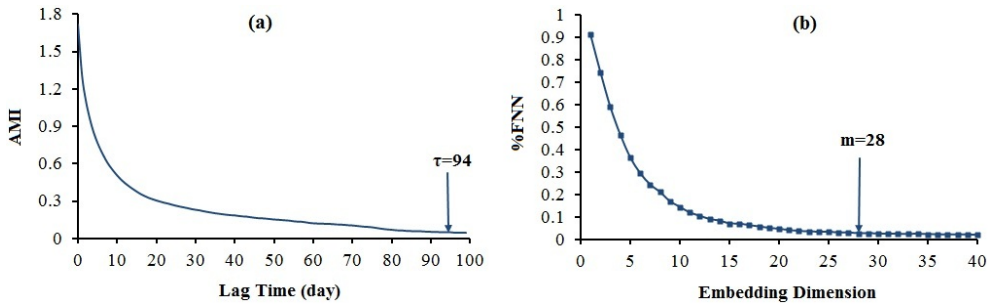


Figure 9. Analysis of the Phase-Space Diagram of Suspended Sediment Data in the Mississippi River basin; (9.a): Average Mutual Information; (9.b) Percentage of false nearest neighbours

m	CC	RMSE
3	0.988	4.00E4
4	0.988	4.10E4
5	0.986	4.30E4
6	0.985	4.60E4
7	0.986	4.40E4
8	0.987	4.20E4

Table 8. Local Prediction Using Different Embedding Dimension for the Mississippi River Dataset

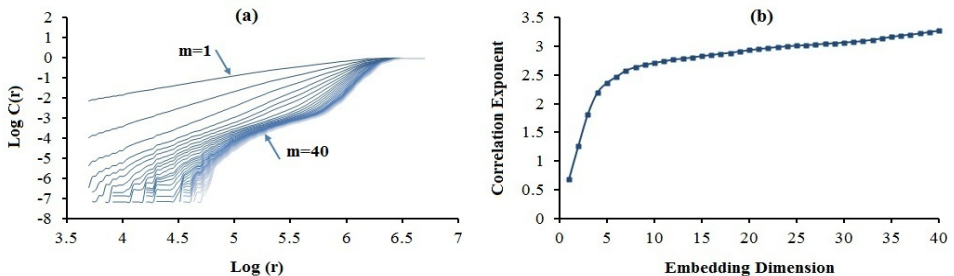


Figure 10. Correlation Dimension Method to Identify the Presence of Chaos Signal in the Dataset; (10.a): Convergence of $\log C(r)$ versus $\log(r)$; (10.b): saturation of correlation dimension $D_2(m)$ with embedding dimension m – this signifies chaotic signals in the Dataset

5. Inter-comparison of the models and discussion of results

Table 9 summarises the performance and main features of each and all of the modelling strategies. The results presented so far confirms the experience that the traditional SRC model performs poorly and may only be used for rough-and-ready assessments. However, the results by the GEP model show that considerable improvements are likely by using it. This section also analyses the relative performance of the various modelling strategies. An overall visual comparison of all the results is presented in Figure 11, according to which GEP, ANN, MLR and local prediction models perform remarkably well and similar to one another.

Model	Performance	Model Structure	Outcome	Comments
SRC	Poor	Model 1	Eq. (3)	For rough-and-ready estimates
GEP	Good	Model 4	Eq. (4.e)	
ANN	Good	Model 4	→	The model is bounded to software
MLR	Good	Model 2	Eq. (6)	
Chaos	Good	Model 0	→	Needs expertise to implement

Table 9. Qualitative Overview of the Performances of Various Modelling Strategies

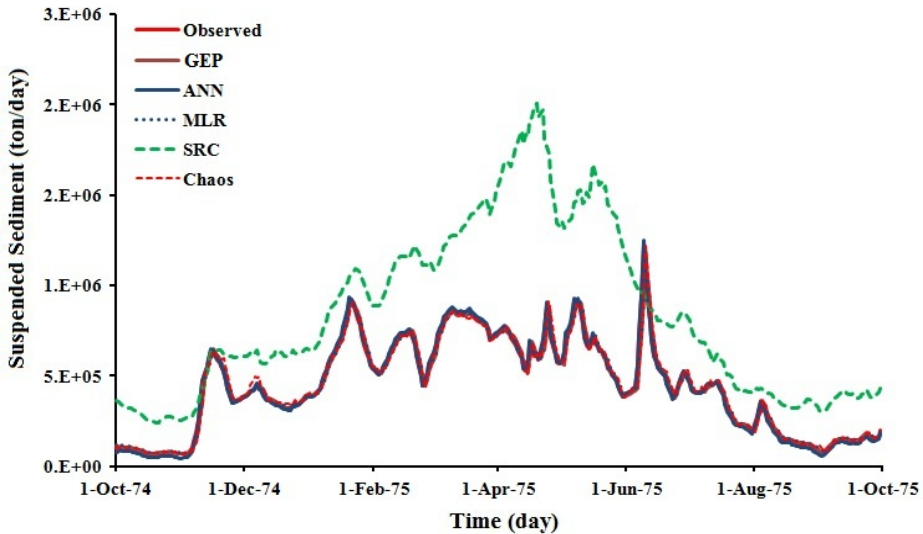


Figure 11. Model Predictions for Suspended Sediment – Performances of GP, ANN, MLR, Chaos (closest to observed), and SRC (poor)

Scatter diagrams are also a measure of performance. These are presented in Figures 12, which provides a quantitative basis that (i) SRC performs poorly and (ii) there is little to choose between the other models, although the performance of ANN stands out.

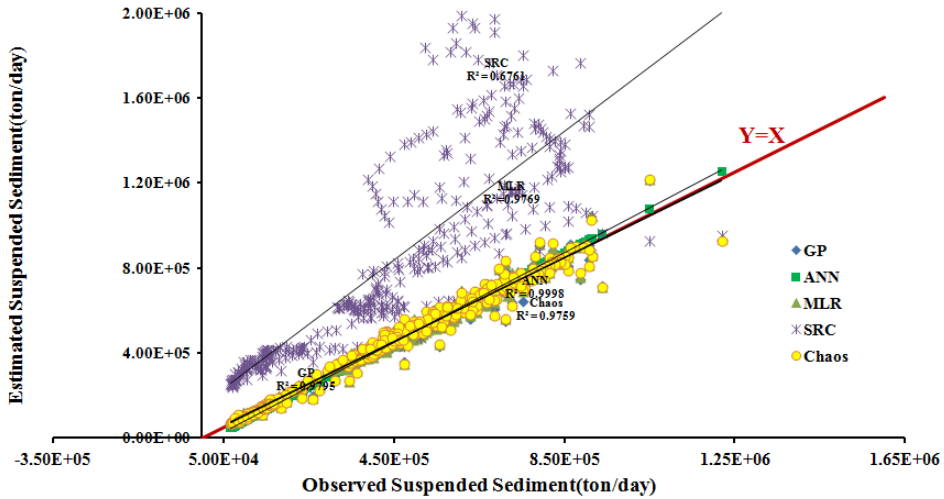


Figure 12. Scatter between Modelled and Observed Suspended Sediment Load

The relative performances of GEP, ANN, MLR and local prediction models are not still visible from Figure 12 and therefore attention is focused on the differences between the GEP and ANN models with respect to their corresponding observed values. Figures 13 shows the respective results for both the GEP and ANN models and that of ANN is remarkable, as the differences are nearly zero. It may be reported that those of local prediction model and MLR are very close to that of GEP.

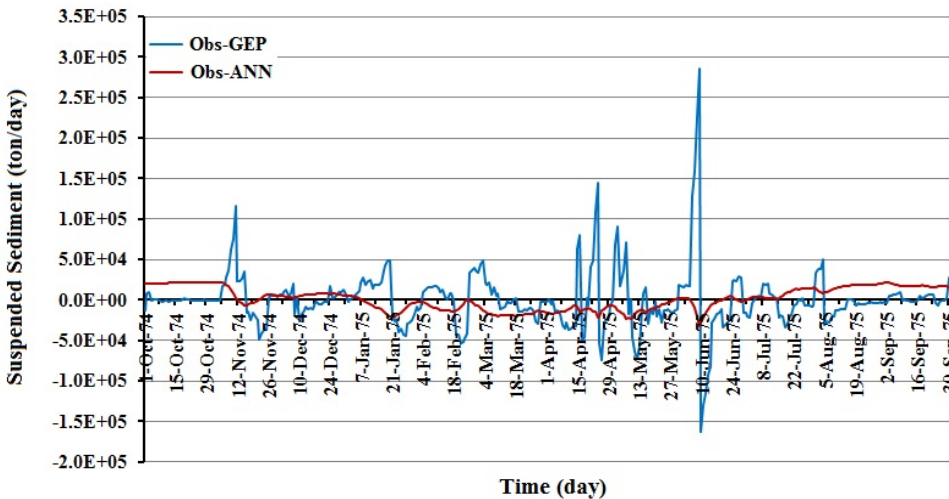


Figure 13. Performances of the ANN and GEP Models – y-ordinates: observed – modelled values

Due to the importance of the volume of transported sediment, the total predicted values are also compared with that of the observed values for the testing period and the results are

presented in Table 10. The table show that the traditional SRC model is in error by as much as nearly 50% but the other models perform well, among which the error in the performance of ANN is the lowest. It is also noted that, despite the good performance of ANN models, it is not transferrable, like the GEP models. The implementation of both ANN and deterministic chaos models require considerable expertise.

Model	Actual Val. (ton/year)	Estimation Val. (ton/year)	Dif. Val. (%)
SRC	1.65E8	3.06E8	+46 %
GEP	1.65E8	1.65E8	- 0.4 %
ANN	1.65E8	1.64E8	- 0.3 %
MLR	1.65E8	1.66E8	+0.6 %
Chaos	1.65E8	1.66E8	+0.7 %

Table 10. Total Volume of Suspended Sediment Predicted by each of the Models at Gauging Station for the Mississippi River basin

The chapter presents the performance of the GEP model, as a variation of evolutionary programming, to forecast suspended sediment load of the Mississippi River, the USA. GEP is just a modelling strategy, where any other relevant strategy is just as valid if its performance is satisfactory. The overall results show that the information contained in the observed data can be treated by the following modelling strategies:

1. **Evolutionary computing:** this produced a formula to forecast the future values in terms of recorded values of flows and suspended sediment. The results show that the strategy can be successful in identifying a number of different formulae.
2. **Emulation of the working of the brain:** this successfully fitted an inbuilt polynomial to the data. It performs better than the other tested models but is not readily transferrable as it resides in particular software applications.
3. **Regression analysis:** this produced a regression equation, according to which the future values would regress towards average recorded values, in spite of the presence of noise.
4. **Deterministic chaos:** this produced future values of suspended sediment load by identifying an attractor towards which the system performance would converge even when the internal system behaves erratically.

The only common feature in the above modelling strategies is their use of optimisation techniques. Otherwise, they are greatly different from one another but remarkably, they produce models fit for purpose and can explain the data. Undoubtedly, the data can be explained by many more sets of equations or by other possible strategies. This emphasises that models are just tools and the modelling task is to test the performance of the various models to add confidence to the results. Yet the poor performance of the traditional SRC underlines the fact that a good performance cannot be taken for granted.

A review of the data (in Section 3) shows that the overall contribution of the datapoints in the test period is average; its individual characteristics in terms of kurtosis shows that the annual hydrographs are less peaked and more flat but at the same time the suspended sediment load during the year was significantly high. Thus, the minimum values during this

year were significantly above the average but persistent and though less dynamic. However, all the four modelling strategies coped well with these data peculiarities. If the data during the test period have a more pronounced feature not very common during the training period, the various local modelling strategies are likely to perform poorly in their own unique way and one of the greatest tasks of research in modelling should be investigations to understand these unique features and not to sweep them under the carpet.

A general view projected by the investigation in this chapter is that the performance of modelling techniques must not be the only basis of practical applications. Equal attention must also be paid to the quality of the data used. If the data suffers from inherent uncertainties, no good model will compensate for the inherent shortfalls.

6. Conclusion

This chapter presents an investigation of the performance of the Gene Expression Programming (GEP) models of suspended sediment load of the Mississippi River, the USA. The study employs the Mississippi River data spanning 26 years involving both flows and suspended sediment load, of which the first 25 years of the data is used for training and the remaining for the prediction of one year into the future. This investigation concurs with the past findings that the performance of sediment rating curve, an empirical technique used widely in practice, can lead to gross errors. This alone underlines the value of other modelling techniques capable of producing reliable results with less than 1% of errors.

The chapter promotes a pluralist culture of modelling and although presents the GEP model as the focus, it also presents the application of other techniques to model the same data. The other models comprise: artificial neural networks, multi-linear regression analysis and deterministic chaos. The chapter outlines the modelling strategy underlying each of these techniques and the results show in spite of their differences they produce similar results inflicting less 1% of errors. The lowest errors are associated with the artificial neural networks for this set of data but each of these techniques should be considered as reliable. The volume of sediment load is an important management parameter and the error associated with each model was estimated for each model. The results show that the traditional SRC model suffers from gross errors by as much as 50% but the other tested models perform well, among which the error in the performance of ANN is the lowest. ANN is noted for its good performance but with some drawback that these models are not transferrable, like the GEP models. It is noted that the implementation of ANN requires an ANN-platform for further modelling and deterministic chaos models require considerable expertise.

Author details

M. A. Ghorbani, H. Asadi and P. Yousefi
University of Tabriz, East Azerbaijan, Iran

R. Khatibi
Consultant Mathematical Modeller, Swindon, Wilts., UK

7. Appendix

Symbols

SRC	Sediment Rating Curve	
MLR	Multi Linear Regression	
Q_t	Discharge Series	
S_t	Sediment Series	
MLR	X_i	Term of Various Model
	a_i	Values Called Regression
Chaos	T	Delay Time
	$C_m(r)$	Fraction of states
	H	Heaviside Step
	N	Number of Points
	D_2	Correlation Exponent
	Y_j	Vectors of Dimension
	M	Dimensional phase Step
	A	Jacobian Matrix
	$x(t)$	different neighbors
	R	Radius Spherical
	$C(r)$	Correlation Function

Appendix I

Table 11 Defaults Values Employed in Implementing GEP and ANN Models

GP		ANN	
Training parameters	Values	Training parameters	Values
Crossover rate	0.1	Goal	Mean Square Error
Mutation rate	0.044	Epochs	10 - 100
Inversion	0.1	Training algorithm	Trainlm
IS Transposition	0.1		
RIS Transposition	0.1		
1-point Recombination	0.3		
2-point Recombination	0.3		
Gene Recombination	0.1		
Gene Transposition	0.1		
Population (Chromosome) size	30		
Head Size	7		
Number of Genes	3		
Linking Function	Addition		
Random Numerical Constants	Yes		
Number of generation	1000		
Arithmetic functions	(4.a)-(4.f)		
Fitness Function	RRSE: RRSE: Root Relative Squared Errors		

Table 11. Default Parameter Values Used by the Model

8. References

- Aytek, A., and Kisi, O. (2008). A genetic programming approach to suspended sediment modeling. *J. Hydrol.*, 351, 288-298
- Cigizoglu, H.K. & Kisi, O. (2006) Methods to improve the neural network performance in suspended sediment estimation, *J. of Hydrology*, 317, 221-238.
- Cigizoglu, H.K., (2004) Estimation and forecasting of daily suspended sediment data by multi layer perceptrons. *Advances in Water Resources* 27, 185–195.
- Cohn TA, Caulder DL, Gilroy EJ, Zynjuk LD, Summers RM. (1992). The validity of a simple statistical model for estimating fluvial constituent loads: An empirical study involving nutrient loads entering Chesapeake Bay. *Water Resources Research* 28(9): 2353–2363.
- Farmer, D.J., Sidorowich, J.J., (1987a). Predicting chaotic time series. *Phys. Rev. Lett.* 59, 845–848.
- Farmer, J.D., Sidorowich, J.J., (1987b). Exploiting chaos to predict the future and reduce noise. In: Lee, Y.C. (Ed.), *Evolution, Learning and Cognition*. World Scientific, River Edge, NJ, pp. 277–330.
- Ferguson RI. (1986). River loads underestimated by rating curves. *Water Resources Research* 22(1): 74–76.
- Ferreira C. (2001a). Gene expression programming in problem solving. In: 6th Online World Conference on Soft computing in Industrial Applications (invited tutorial)
- Ferreira C. (2001b). Gene expression programming: a new adaptive algorithm for solving problems. *Complex Syst* 13(2):87–129
- Fraser, A.M., Swinney, H.L., (1986). Independent coordinates for strange attractors from mutual information. *Physical Review A.* 33(2), 1134-1140.
- Ghorbani MA, Khatibi R, Aytek A, Makarynskyy O, Shiri J (2010) Sea water level forecasting using genetic programming and comparing the performance with artificial neural networks. *J Comput Geosci* 36(5):620–627
- Glysson GD. (1987). *Sediment transport curves*. US Geological Survey Open File Report 87–218.
- Grassberger, P. and Procaccia, I., (1983). Characterization of strange attractors. *Physical review letters*, Vol. 50, No. 5, 346-349.
- Hegger, R., Kantz, H., Schreiber, T., (1999). Practical implementation of nonlinear time series methods: The TISEAN package. *Chaos.* 9, 413-435.
- Itoh, K., (1995). A method for predicting chaotic time-series with outliers. *Electron. Commun. Jpn.* 78 (5), 44–53.
- Jain, S.K. (2001) Development of integrated sediment rating curves using ANNs. *J. Hydraul. Eng ASCE* 127(1), 30–37.
- Kennel, M., Brown, R., Abarbanel, H.D.I., (1992). Determining embedding dimension for phase- space reconstruction using a geometrical construction. *Phys Rev A* .45, 3403–11.
- Khatibi, R., (2011), “Evolutionary Systemic Modelling for Flood Risk Management Practices,” *Journal of Hydrology* Vol. 401 Issue 1-2, Pp 36–52 (<http://dx.doi.org/10.1016/j.jhydrol.2011.02.006>)

- Kisi, O. (2004b) Multi-layer perceptrons with Levenberg-Marquardt optimization algorithm for suspended sediment concentration prediction and estimation. *Hydrol. Sci. J.* 49(6), 1025–1040.
- Kisi, O. (2005). "Daily river flow forecasting using artificial neural networks and autoregressive models", *Turkish J. Eng. Env. Sci.* vol 29, 9-20.
- Kisi, O., (2005b) "Suspended sediment estimation using neuro-fuzzy and neural network approaches", *Hydrol. Sci. J.*, 50(4), 683-696.
- Koza JR (1992) Genetic programming: on the programming of computers by means of natural selection. MIT, Cambridge
- Miller CR. (1951). *Analysis of flow-duration, sediment-rating curve method of computing sediment yield*. US Bureau of Reclamation Report: Denver, Colorado.
- Porporato, A, Ridolfi, L., (1997). Nonlinear analysis of river flow time sequences, *Water Resources Research*.33(6), 1353-1367.
- Rumelhart, D.E. and McClelland, J.L. (Eds.), (1986)."Parallel Distributed Processing". Explorations in the Microstructure of Cognition, 1. MIT Press, Cambridge
- Sivakumar, B., (2002). A phase-space reconstruction approach to prediction of suspended sediment concentration in rivers, *Journal of Hydrology*, 258(1-4), 149-162,
- Sivakumar, B., and A.W. Jayawardena, (2002). An investigation of the presence of lowdimensional chaotic behavior in the sediment transport phenomenon, *Hydrological Sciences Journal*, 47(3), 405- 416.
- Sivakumar, B., and A.W. Jayawardena, (2003). Sediment transport phenomenon in rivers: An alternative perspective, *Environmental Modeling and Software*, in press.
- Sivakumar, B., Wallender, W.,(2005). Predictability of river flow and suspended sediment transport in the Mississippi River basin: a non-linear deterministic approach. *J. Earth Surf. Process. Landforms* 30, 665–677
- Takens, F.,(1981). Detecting strange attractors in turbulence, in *Dynamical Systems and Turbulence, Lecture Notes in Mathematics 898*, D. A. Rand and L. S. Young (eds.), 366-381, Springer-Verlag, Berlin.
- Tayfur, G. & Guldal, V. (2006) Artificial neural networks for estimating daily total suspended sediment in natural streams, *Nordic Hydrology*, 37, 69-79.
- Tayfur, G. (2002) Artificial neural networks for sheet sediment transport. *Hydrol. Sci. J.* 47(6), 879–892.
- Theiler, J., (1986). Spurious dimension from correlation algorithms applied to limited time series data. *Phys Rev A.* 34. 2427-2432.
- Thomas RB. (1988). Monitoring baseline suspended sediment in forested basins: The effects of sampling of suspended sediment rating curves. *Hydrological Sciences Journal* 33(5): 499–514.
- Ustoorikar K, Deo MC (2008) Filling up gaps in wave data with genetic programming. *Marine Structures* 21:177–195
- Wilks DS. (1991). Representing serial correlation of meteorological events and forecasts in dynamic decision-analytic models. *Monthly Weather Rev* 119:1640–1662

Yu-Min Wang, Seydou Traore and Tienfuan Kerh (2008), "Monitoring Event-Based Suspended Sediment Concentration by Artificial Neural Network Models" WSEAS TRANSACTIONS on COMPUTERS, Issue 5, Volume 7, May 2008

Walling, D. E., 1977: Limitations of the rating curve technique for estimating suspended sediment loads, with particular reference to British rivers. In: Erosion and solid matter transport in inland waters (Proceedings of the Paris symposium, (July, 1977), IAHS Publ., 122, 34-38.