RESEARCH PAPER | International Journal of Agricultural Engineering | Volume 10 | Issue 2 | October, 2017 | 246-252

e ISSN–0976–7223 ■ Visit us : *www.researchjournal.co.in* ■ **DOI: 10.15740/HAS/IJAE/10.2/246-252**

Suspended sediment load estimation using neuro-fuzzy and multiple linear regression: Vamsadhara River Basin, India

■ **SHREYA NIVESH AND PRAVENDRA KUMAR**

Received : 01.03.2017; **Revised :** 03.07.2017; **Accepted :** 17.07.2017

See end of the Paper for authors' affiliation

Correspondence to :

SHREYA NIVESH
Department of Soil and Water Conservation Engineering, College of Technology, Govind Ballabh Pant University of Agriculture and Technology, Pantnagar, U.S. NAGAR (UTTARAKHAND) INDIA
Email : shreyanivesh@gmail.com

■ **ABSTRACT :** Soil erosion by water is the most serious form of land degradation resulting in loss of crop productivity by 0.2-10.9 q/ha (66% total production loss) for cereals, 0.1-6.3 q/ha for oilseeds (21% total production loss) and 0.04-4.4 q/ha for pulses (13% total production loss) estimated across states, which has a direct bearing on food security of the country. Therefore, a major challenge still remaining is the accurate prediction of the catchment sediment yield responses to the rainfall-runoff events. One viable approach to this challenge is the use of suitable statistical and soft-computing techniques for the efficient management of watersheds and ecosystems. The present study deals with the development of adaptive neuro-fuzzy inference system (ANFIS) and multiple linear regression (MLR) models to estimate the suspended sediment load from Vamsadhara river catchment comprising of 7820 km^2, situated between Mahanadi and Godavari river basins in south India. Considering the active monsoon period, 70% data were used for model calibration and remaining 30% data were used for model validation. Results revealed that the Neuro-Fuzzy models are in good agreement with the observed values and present better performance in comparison to the statistical models.

■ **KEY WORDS :** Adaptive neuro-fuzzy inference system, Multiple linear regression, Calibration, Validation, Soft-computing

■ **HOW TO CITE THIS PAPER :** Nivesh, Shreya and Kumar, Pravendra (2017). Suspended sediment load estimation using neuro-fuzzy and multiple linear regression: Vamsadhara River Basin, India. *Internat. J. Agric. Engg.*, **10**(2) : 246-252, **DOI: 10.15740/HAS/IJAE/10.2/246-252.**

Correct estimation of sediment load carried by rivers is of utmost importance in the soil and water conservation practices in the watershed and also in large number of hydro-environmental issues such as planning, design and operations of reservoirs, dams and environmental impact assessment. Land which includes complex mixture of minerals, water, air, organic matter, and countless organisms serves as a natural medium for the growth of plants and capable of supporting their life and is vital to life on earth. The extent to which water is plentiful or scarce, clean or polluted, beneficial or destructive, influence the extent and quality of human life and is likely to become a critical scarce resource in the coming decades due to relentless increase in population and the resulting spurt in the demand for water required. Therefore, the conservation of land and water resources is of significant concern these days.

A number of linear and non-linear models have been developed since 1930's to simulate and forecast various hydrological processes and variables (Yang, 1996). Hydrologic simulation models are rapidly being improved with increased advances in computer techniques that facilitate their capability to interface with emerging technologies to provide more powerful tools for operational applications. Multiple linear regression (MLR) is a statistics based technique that uses several independent variables to predict the outcome of a dependent variable. In recent years, regression models have been successfully employed in modelling a wide range of hydrologic processes like soil temperature (Bilgili, 2010; Tabari *et al.*, 2010 and Marofi *et al.*, 2011); flood flows (Engeland and Hisdal, 2009 and Eslamian *et al.*, 2010); and sediment prediction (Wang and Linker, 2008 and Chang *et al.*, 2008).

Soft computing based techniques are becoming a strong tool for providing environmental, irrigation and drainage, soil and water conservation and civil engineers with sufficient details for design purposes and management works. Specific applications of Soft computing have been applied in hydrology and hydraulics including real-time flood forecasting and rainfall-runoff modelling (Zhu *et al.*, 1994; See and Openshaw, 2000; Stuber *et al.*, 2000; Hundecha *et al.*, 2001; Xiong *et al.*, 2001; Giustolisi and Lauchelli, 2005 and Nayak *et al.*, 2004 and 2005); stage–discharge relationship modelling (Lohani *et al.*, 2006 and Kisi and Cobaner, 2009); reservoir inflow forecasting (Bae *et al.*, 2007); river flow modelling (Zounemat-Kermani and Teshnehlab, 2008); estimation of suspended sediment and scour depth near pile groups (Tayfur, 2002; White, 2005; Cigizoglu and Kisi, 2006; Tayfur and Guldal, 2006; Ardiclioglu *et al.*, 2007 and Sadeghi *et al.*, 2013); Fuzzy rule base approach for developing soil a protection index map: a case study in the upper awash basin, Ethopian highlands (Oinam *et al.*, 2014); Fuzzy intelligence system for land consolidation-a case study for Shunde, China (Wang *et al.*, 2015) and a new approach for modelling suspended sediment using evolutionary fuzzy approach (Kisi, 2016); Suspended sediment transport dynamics in rivers : Multi-scale driver of temporal variation (Vercruysse *et al.*, 2017). The present study deals with the development, performance evaluation and validation of Neuro-fuzzy and regression models for predicting sediment load from the Vamsadhara river basin situated between Mahanadi and Godavari river basins in south India.

■ METHODOLOGY

Study area :

The present study was undertaken in Vamsadhara river basin comprising of 7820 km^2, situated within the geographical coordinates of 18^{0} $15^{|}$ to 19^{0} $55^{|}$ N latitudes and 83^{0} $15^{|}$ to 84^{0} $20^{|}$ E longitudes in between Mahanadi and Godavari river basins falls in the state of Orissa and the rest 26% in Andhra Pradesh. Hydrological data were collected by India Meteorological Department (IMD) and Central Water Commission (CWC), Godavari Mahanadi Circle Division, South Eastern Region, Bhubaneswar, Orissa at six sites: Kutraguda, Mohana, Gudari, Mohandragarh, Gunpur, and Kashinagar. The measurements include rainfall in the units of millimetres, discharge in the units of m^3/sec and sediment concentration in the units of kg/m^3. The daily weighted rainfall for the study area was found by considering the Theissen polygons. The location of the study area is shown in Fig. A.

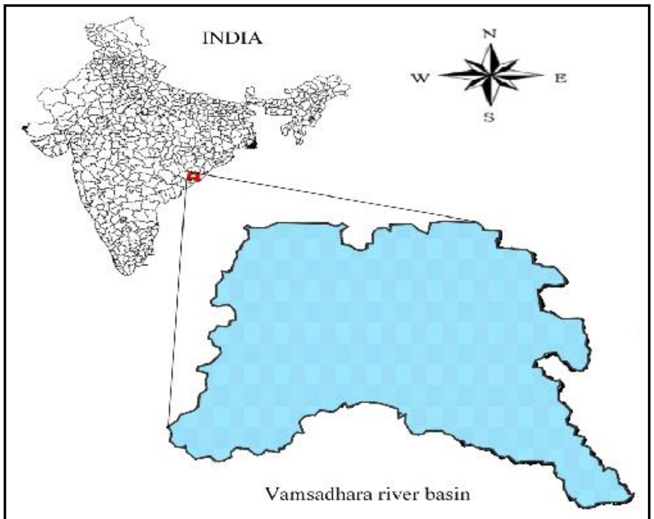

Fig. A : Location map of the study area

Methodologies :

Adaptive neuro-fuzzy inference system :

An adaptive neuro-fuzzy inference system or adaptive network-based fuzzy inference system (ANFIS) is a kind of artificial neural network that is based on Takagi–Sugeno fuzzy inference system. The technique was developed in the early 1990s. It integrates both neural networks and fuzzy logic principles. It has potential to capture the benefits of both techniques in a single framework. Its inference system corresponds to a set of fuzzy IF–THEN rules that have capability to learn to approximate nonlinear functions. The architecture of Adaptive Neuro-Fuzzy Inference System (ANFIS) consists of a five layers feed forward neural network. Description of each layer is given as follows:

Layer 1: Fuzzification layer:

Each node in this layer produces membership grades of an input variable. The output of the i^th^ node in layer l is denoted as Oi^1. Assuming a generalized bell function as the membership function, the output Oi^1 can be computed as:

$$O_i^1 \, Z \, \pi_{A_i}(x) \, Z \, \frac{1}{1 \, H \, [(x - c_i)/ai]^{2b_i}} \qquad (1)$$

where $\{a_i, b_i, c_i\}$ are adaptable variables known as premise parameters.

Layer 2: Rule layer:

Every node in this layer multiplies with the incoming signals

$$O_i^2 = w_i = \mu_{Ai}(x)\,\mu_{Bi}(y),\ i = 1, 2 \qquad (2)$$

Layer 3: Normalization layer :

The i^{th} node of this layer calculates the normalized firing strengths as

$$O_i^3 = \overline{w_i} = \frac{w_1}{w_1 + w_2}\ ;\ i=1, 2 \qquad (3)$$

Layer 4: Defuzzification layer :

Node i in this layer calculates the contribution of the i^{th} rule towards the model output, with the following node function:

$$O_i^4 = \overline{w_i} f_i = \overline{w_i}(p_i x + q_i y + r_i) \qquad (4)$$

where w_i is the output of layer 3 and $\{p_i, q_i, r_i\}$ is the parameter set.

Layer 5: Single summation neuron :

The single node in this layer calculates the overall output of the ANFIS as reported by Jang and Sun (1993)

$$O_i^5 = \sum \overline{w}_i f_i = \frac{\sum w_i f_i}{\sum_i w_i} \qquad (5)$$

Multiple linear regression :

Regression analysis is used when two or more variables are thought to be well connected by a linear relationship systematically. MLR applies to problems in which records have been kept of one variable, *y*, the dependent variable, and several other variables $x_1,..., x_k$, the independent variables, and in which the objective requires the relationship between the variable *y* and the variables $x_1, ..., x_k$ to be investigated. In the present study the multiple linear regressions analysis was performed on the same data set to estimate sediment concentration and the regression equation used is defined as

$$\mathbf{S_t = a + bP_t + cQ_t + dQ_{t-1} + eS_{t-1}}$$

where a, b, c, d and e are constants and P_t, Q_t, Q_{t-1}, and S_{t-1} are the variables.

Model architecture :

For the present study MATLAB R2009a software was used to model suspended sediment load. Four years daily data of rainfall, stream flow and suspended sediment concentration of monsoon season from June 1, 1997 to October 31, 2000 was used. 70% data (428 data sets) were used for training and 30% data (184 data sets) were used for testing. Two daily input data groups were employed in this study. Input 1 consists of P_t, Q_t, Q_{t-1}, S_{t-1} as inputs to the model to predict S_t. Input 2 consist of P_{t-1}, Q_t, Q_{t-1}, S_{t-1}. The architecture of Adaptive Neuro-Fuzzy Inference System (ANFIS) networks was developed using four types: Triangular (trimf), Trapezoidal (trapmf), Gaussian (Gaussmf) and generalized-bell (gbellmf) membership functions with number of membership functions per input varying from 3 to 5. Fuzzy model used was Takagi-Sugeno-Kang type with maximum number of epochs 30 considering back propagation learning algorithm.

Model performance :

Three performance indicators were used to examine the goodness to fit of the ANFIS and MLR models to the testing data. These measures include the root mean square error (RMSE), correlation co-efficient (r) and co-efficient of efficiency (CE).

Root mean square error (RMSE) :

It yields the residual error in terms of the mean square error expressed as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(S_{o,i} - S_{e,i})^2}{N}}$$

Correlation co-efficient (r) :

It is a measure of how well the estimated values from an estimated model fit with the real-life data. It is expressed as:

$$r = \frac{\sum_{i}^{N} [(S_{o,i} - \overline{S}_{o,i})(S_{e,i} - \overline{S}_{e,i})]}{\sqrt{\sum_{i}^{N} (S_{o,i} - \overline{S}_{o,i})^2 \sum_{i}^{N} (S_{e,i} - \overline{S}_{e,i})^2}}$$

Co-efficient of efficiency (CE) :

The Nash–Sutcliffe model efficiency co-efficient is used to assess the predictive power of hydrological models and is expressed as:

$$CE = \left\{ 1 - \frac{\sum_{i}^{N} (S_{o,i} - S_{e,i})^2}{\sum_{i}^{N} (S_{o,i} - \overline{S}_{o,i})^2} \right\} *100$$

where, $S_{o,i}$ and $S_{e,i}$ are the observed and estimated suspended sediment concentration; $\overline{S}_{o,i}$ and $\overline{S}_{e,i}$ are the average observed and estimated suspended sediment concentration, respectively for the i^{th} data set and N is the total number of observations.

RESULTS AND DISCUSSION

Various graphical and statistical indicators were used to evaluate the performance of the sediment ANFIS and regression models (Fig. 1). These performance evaluation indicators of the models are given in the Tables 1, 2 and 3.

Table 1 reveals that results produced by ANFIS case-1 models which take concurrent rainfall and runoff; and antecedent runoff and sediment load of time step t-1 considering generalised bell membership function perform better than the other models. The RMSE, which was 249.92 kg/sec in case of ANFIS-4 model having trapezoidal membership function with three membership functions reduced to 44.02 kg/sec in case of ANFIS-12 model having generalized membership function with number of membership functions five period. There is an improvement in the value of correlation co-efficient (r) from 0.89 to 0.99 and co-efficient of efficiency from 69.76% to 99.06%. This indicates that previous day runoff and sediment load; concurrent day rainfall and runoff have significant influence on the sediment yield.

The ANFIS case-2 models were developed to see the effect of antecedent rainfall with time step t-1 in addition to Q_t, Q_{t-1}, S_{t-1} using triangular, trapezoidal, Gaussian and generalized-bell membership functions with different number of membership functions. From Table 2, it can be seen that this scenario is inferior in all aspect of statistical indicators *i.e.*, root mean square error, correlation co-efficient and co-efficient of efficiency. For the ANFIS-23 model the RMSE, r and CE values are 52.99 kg/sec, 0.99 and 98.64 %, respectively. Results reveal that previous day rainfall does not have significant influence to sediment yield from the river basin.

Various graphical and statistical indicators were used to evaluate the performance of the sediment yield regression models as shown in Table 3. The Table 3 reveals that regression model MLR-1 is better than the MLR-2.

Table 1 : Performance indicators of various ANFIS models for case-1

Model	Network	RMSE	r	CE
ANFIS-1	trimf-3	236.18	0.90	72.99
ANFIS-2	trimf-4	220.77	0.92	76.41
ANFIS-3	trimf-5	200.96	0.93	80.45
ANFIS-4	trapmf-3	249.92	0.89	69.76
ANFIS-5	trapmf-4	239.09	0.90	72.33
ANFIS-6	trapmf-5	205.17	0.92	79.62
ANFIS-7	Gaussmf-3	181.19	0.94	84.11
ANFIS-8	Gaussmf-4	110.37	0.97	91.68
ANFIS-9	Gaussmf-5	47.79	0.95	96.03
ANFIS-10	gbellmf-3	126.00	0.97	92.31
ANFIS-11	gbellmf-4	114.83	0.98	93.62
ANFIS-12	gbellmf-5	44.02	0.99	99.06

Table 2 : Performance indicators of various ANFIS models for case-2

Model	Network	RMSE	r	CE
ANFIS-13	trimf-3	224.40	0.91	79.16
ANFIS-14	trimf-4	207.00	0.92	75.63
ANFIS-15	trimf-5	162.49	0.95	87.22
ANFIS-16	trapmf-3	216.93	0.91	77.22
ANFIS-17	trapmf-4	184.94	0.93	83.44
ANFIS-18	trapmf-5	168.92	0.95	86.19
ANFIS-19	Gaussmf-3	138.75	0.93	90.68
ANFIS-20	Gaussmf-4	136.74	0.96	90.95
ANFIS-21	Gaussmf-5	84.62	0.98	96.54
ANFIS-22	gbellmf-3	129.17	0.96	91.86
ANFIS-23	gbellmf-4	52.99	0.99	98.64
ANFIS-24	gbellmf-5	91.46	0.95	97.51

Table 3 : Comparison of selected ANFIS models with MLR models

Model	RMSE	r	CE (%)
ANFIS-12	44.02	0.99	99.06
ANFIS-23	52.99	0.99	98.64
MLR-1	188.28	0.91	82.82
MLR-2	194.65	0.90	81.64

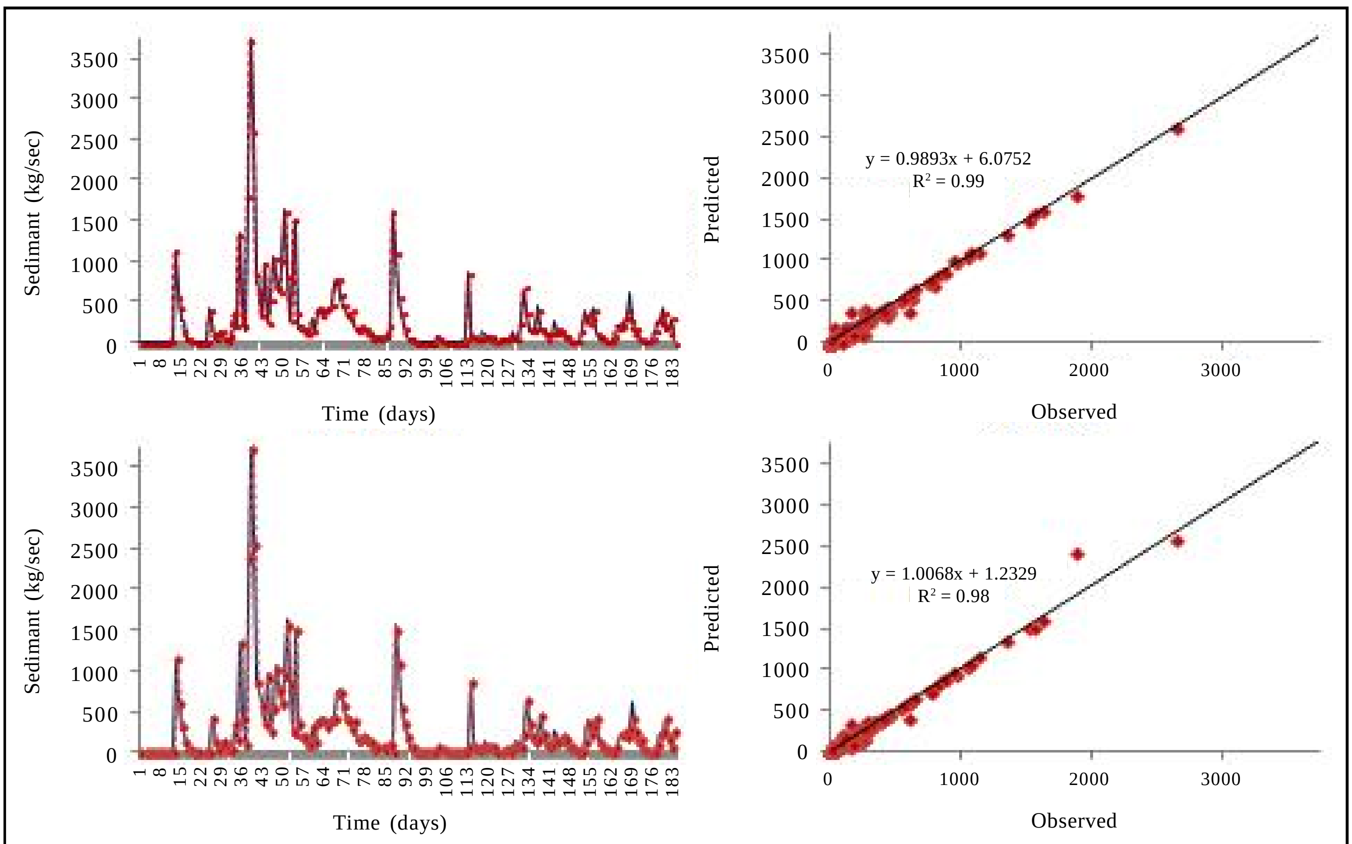

Fig. 1 : Time series and Scatter plots of observed and estimated suspended sediment load for ANFIS-12 and ANFIS-23 models

So, based on the above discussion, it can be concluded that ANFIS models with input variables as P_t, Q_t, Q_{t-1} and S_{t-1} and membership function generalized bell (gbellmf) with number of membership functions per input 4 and 5 can best simulate the sediment load in Vamsadhara river basin. It can also be concluded that statistical or traditional models are not capable of simulating complex and non-linear sediment yield processes whereas performance of the ANFIS models is quite satisfactory in this regard.

Conclusion :

In the present study, ANFIS and MLR models were developed for simulation of sediment yield in Vamsadhara River basin. Based on the performance evaluation indices the following conclusions were drawn from this study.

– The ANFIS-12 and ANFIS-23 outperformed the ANFIS and MLR models for estimating sediment yield for present the study area.

– The ANFIS model with membership function generalized-bell and inputs as concurrent rainfall and runoff, antecedent runoff and sediment load was found to be the best among the selected models for predicting sediment yield for the Vamsadhara river basin.

– The MLR model fits poorly for the data set under study.

– It can be concluded that neuro-fuzzy models are superior to regression models in predicting sediment load in all respects.

Acknowledgement :

The authors thank to India Meteorological Department (IMD) and Central Water Commission (CWC), Godavari Mahanadi Circle Division, South Eastern Region, Bhubaneswar, Orissa for providing the data used in the study.

Authors' affiliations:

PRAVENDRA KUMAR, Department of Soil and Water Conservation Engineering, College of Technology, Govind Ballabh Pant University of Agriculture and Technology, Pantnagar, U.S. NAGAR (UTTARAKHAND) INDIA

■ REFERENCES

Ardiclioglu, M., Kisi, O. and Haktanir, T. (2007). Suspended sediment prediction by using two different feed-forward back propagation algorithms. *Canadian J. Civil Engg.*, **34**(1): 120-125.

Bae, D.H., Jeong, D.M. and Kim, G. (2007). Monthly dam inflow forecasts using weather forecasting information and neuro-fuzzy technique. *Hydrological Sci. J.*, **52**(1): 99-113.

Bilgili, M. (2010). Prediction of soil temperature using regression and artificial neural network models. *Meteorol. & Atmospheric Phys.*, **110** : 59-70.

Chang, C.K, Ghani, A.A., Abdullah, R. and Zakaria, N.A. (2008). Sediment transport modelling for Kulim river case study. *J. Hydro-environ. Res.*, **2**(1): 47-59.

Cigizoglu, H.K. and Kisi, O. (2006). Methods to improve the neural-network performance in suspended sediment estimation. *J. Hydrol.*, **317** : 221-238.

Engeland, K. and Hisdal, H. (2009). A comparison of low flow estimates in ungauged catchments using regional regression and the HBV-model. *Water Resour. Manage.*, **23**(12): 2567-2586.

Eslamian, S.S., Ghasemizadeh, M., Biabanaki, M. and Talebizadeh, M. (2010). A principal component regression method for estimating low flow index. *Water Resour. Manage.*, **24**(11) : 2553-2566.

Giustolisi, O. and Lauchelli, D. (2005). Improving generalization of artificial neural networks in rainfall-runoff modelling. *Hydrological Sci. J.*, **50**(3): 515-528.

Hundecha, Y., Bardossy, A. and Theissen, H.W. (2001). Development of a fuzzy logic-based rainfall-runoff model. *Hydrol. Sci. J.*, **46**(3): 363-376.

Kisi, O. and Cobaner, M. (2009). Modelling river-stage discharge relationship using different neural networks. *Clean-Soil, Air, Water,* **37**(2): 160-169.

Kisi, O. (2016). A new approach for modeling suspended sediment: Evolutionary fuzzy approach. *Hydrol. & Earth System Sci.*, **58**(3) : 587-599.

Lohani, A.K, Goel, N.K. and Bhatia, K.K.S. (2006). Takagi-Sugeno fuzzy inference system for modelling stage-discharge relationship. *J. Hydrol.*, **331**(1-2): 146-160.

Marofi, S., Tabari, H. and Zare, A.H. (2011). Predicting spatial distribution of snow water equivalent using multivariate non-linear regression and computational intelligence methods. *Water Resour. Manage.*, **25**: 1417-1435.

Nayak, P.C., Sudheer, K.P. and Ramasastri, K.S. (2004). A neuro-fuzzy computing technique for modelling hydrological time series. *J. Hydrol.*, **291** : 52-66.

Nayak, P.C., Sudheer, K.P. and Ramasastri, K.S. (2005). Fuzzy computing based rainfall-runoff model for real time flood forecasting. *Hydrol. Processes*, **19**: 955-968.

Oinam, B.C., Marx, W., Scholten, T. and Wieprecht, S. (2014). Fuzzy rule base approach for developing soil a protection index map: a case study in the upper Awash basin, Ethopian highlands. *Land Degrad. & Dev.*, **25**(5): 483-500.

Sadeghi, S.H.R., Seghaleh, M.B. and Rangaver, A.S. (2013). Plot sizes dependency of runoff and sediment yield estimates from small watersheds. *Catena*, **102**: 55-61.

See, L. and Openshaw, S. (2000). Applying soft computing approaches to river level forecasting. *Hydrol. Sci. J.*, **44**(5): 763-779.

Stuber, M., Gemmar, P. and Greving, M. (2000). Machine supported development of fuzzy-flood forecast systems. In: *European Conf. on Adv. In Flood Res. (Potsdam, Germany)* (ed. By A. Bronstert, C. Bismuth and L. Menzel), pp. 504-515.

Tabari, H., Marofi, S., Zare, A.H. and Sharifi, M.R. (2010). Comparison of artificial neural network and combined models in estimating spatial distribution of snow depth and snow water equivalent in Samsami basin of iran. *Neural Comput. Appli.*, **19** : 625-635.

Tallis, J.H. (1998). Growth and Degradation of British and Irish Blanket mires. *Environ. Reviews*, **6**(2): 81-122.

Tayfur, G. (2002). Artificial neural networks for sheet sediment transport. *Hydrol. Sci. J.*, **47**(6) : 879-892

Tayfur, G. and Guldal, V. (2006). Artificial neural networks for estimating daily total suspended sediment in natural streams. *Nordic Hydrol.*, **37** : 69–79

Vercruysse, K., Robert, C.G. and Rickson, R.J. (2017). Suspended sediment transport dynamics in rivers: Multi-scale driver of temporal variation. *Earth-Science Reviews*, **166** : 38-52.

Verstraeten, G. and Poesen, J. (2001). Factors controlling sediment transport from small intensively cultivated catchment in a temperate humid climate. *Geomorphol.*, **40** (1-2) : 123-144.

Wang, P. and Linker, L.C. (2008). Improvement of regression simulation in fluvial sediment loads. *J. Hydra. Engg.*, **134**: 1527-1531.

Wang, J., Ge, J., Hu, Y., Li, C. and Wang, L. (2015). Fuzzy intelligence system for land consolidation-a case study for Shunde, China. *Soil Earth*, **6** : 997-1006.

White, S. (2005). Sediment yield prediction and modelling. *Hydrol. Process*es, **19** : 3053-3057.

Xiong, L.H., Shamsaldein, A.Y. and O'Connor, K.M. (2001). A nonlinear combination of the forecasts of rainfall-runoff models by the first order Takagi-Sugeno fuzzy system. *J. Hydrol.*, **245**: 196-217.

Yang, C.T. (1996). *Sediment transport theory and practice*, McGraw-Hill, New York.

Zhu, M.L, Fujita, M., Hashimoto, N. and Kudo, M. (1994). Long lead time forecast of runoff using fuzzy reasoning method. *J. Japan. Soc. Hydrol. & Water Resour.*, **7**(2): 83-89.

Zounemat-Kermani, M. and Teshnehlab, M. (2008). Using adaptive neuro-fuzzy inference system for hydrological time series prediction. *Appl. Soft Comput.*, **8**: 928–936.

★★★★★ 10th Year of Excellence ★★★★★