5

Fuzzy Nonlinear Function Approximation (FNLLA) Model for River Flow Forecasting

P.C. Nayak[1], K.P. Sudheer[2] and S.K. Jain[3]
[1]*Deltaic Regional Centre, National Institute of Hydrology, Kakinada*
[2]*Dept of Civil Engineering, Indian Institute of Technology Madras,*
[3]*NEEPCO, Department of Water Resources Development and Management, Indian Institute of Technology, Roorkee, India*

1. Introduction

It is well understood that the limitations of hydrological measurement techniques warrants for modeling of hydrological processes in a basin. However, most hydrologic systems are extremely complex and modeling them with the available limited measurements is a difficult task. The basic purpose of a model is to simulate and predict the operation of the system that is unduly complex, and also to predict the effect of changes on this operation. It is well known, of various hydrological processes, the rainfall-runoff process is the most complex hydrologic phenomenon to comprehend due to tremendous spatial and temporal variability of basin characteristics and rainfall patterns, as well as a number of other variables associated in modeling the physical processes (Tokar and Markus, 2000). The transformation from rainfall to basin runoff involves many hydrologic components that are believed to be highly nonlinear, time varying, spatially distributed, and not easily described by simple models. The artificial neural network (ANN) and Fuzzy Inference System (FIS) approaches are becoming increasingly popular in the context rainfall-runoff modeling due to their various advantages. This Chapter discusses an effective integration of these two models in a different manner.

It is a common belief that the ANN (and FIS models to an extent) models of the rainfall runoff process are purely black models as they do not explain the process being modeled, but for a few recent studies (Wilby et al., 2003; Jain et al., 2004; Sudheer, 2005). However, it must be realized that the hydro-meteorological data that are employed in developing rainfall runoff models (ANN, FIS or conceptual) contain important information about the physical process being modeled, and this information gets embedded or captured inside the model. For instance, a flow hydrograph, which is normally used as the output variable in an ANN rainfall runoff model, consists of various components that result from different physical processes in a watershed.

For example, the rising limb of a runoff hydrograph is the result of the gradual release of water from various storage elements of a watershed due to gradual repletion of the storage due to the rainfall input. The rising limb of the hydrograph is influenced by varying infiltration capacities, watershed storage characteristics, and the nature of the input *i.e.* intensity and duration of the rainfall, and not so much by the climatic factors such as temperature and evapotranspiration etc. (Zhang and Govindaraju, 2000). On the other hand, the falling limb of a watershed is the result of the gradual release of water from various storages of the watershed after the rainfall input has stopped, and is influenced more by the storage characteristics of the watershed and climatic characteristics to some extent. Further, the falling limb of a flow hydrograph can be divided into three parts: initial portion just after the peak, middle portion, and the final portion. The initial portion of the falling limb of a flow hydrograph is influenced more by the quick-flow (or interflow), the middle portion of the falling limb is more dominated by the delayed surface flow, and the final portion of the falling limb (of smaller magnitudes) is dominated by the base flow. Hence it is apparent that a local approximation technique, which maps the changing dynamics using different functions, would be an effective way to model the rainfall runoff process. Local approximation refers to the concept of breaking up the domain into several small neighbourhood regions and analysing these separately. This argument is supported by the results of Sudheer (2005), wherein he proposes a procedure to extract knowledge from trained ANN river flow models.

Farmer and Sidorowitch (1987) have found that chaotic time series prediction using local approximation techniques is several orders of magnitude better. In their approach, the time series is first embedded in a state space using delay coordinates, and the underlying nonlinear mapping is inferred by a local approximation using only nearby states. This approach can be easily extended to higher order local polynomial approximations. Singer et al. (1992) derived the local approximation as a state dependant autoregressive modelling. However, this becomes complex with large data sets as the inefficient computation of nearby state search makes the implementation much harder. In order to overcome this limitation we proposed to simplify the signal representation (input) with a vector clustering procedure. That is, the local model fitting is based on statistically averaged prototypes instead of the original state vector samples. Also, the nearby state search can be significantly simplified with all prototypes organized according to a certain metric such as pattern similarity.

Most of the applications of local approximation technique (e.g. FIS) employ linear relationship as an effective local approximation. However, the rainfall runoff process being highly nonlinear, a nonlinear local approximation would be a better approach. The objective of this Chapter is to illustrate that a nonlinear local approximation approach in modelling the rainfall-runoff process may offer better accuracy in the context of river flow forecasting. More specifically, the chapter discusses about subdividing the data into subspaces and evaluates the limitations of linear and nonlinear local approximations. The proposed approach is illustrated through a real world case study on two river basins. Both the applications are developed for river flow forecasting, one on a daily time step and the other on an hourly time step.

2. Theoretical considerations of local approximation

Consider modeling a river flow time series such that $y_t = f(x_t, \lambda)$, where it is required to forecast the value of flow (y_t) and $\mathbf{x}_t$ is the input vector to the model at time t. Generally, the modeler uses a set of n 'candidate' examples of the form ($\mathbf{x}_i$, y_i), $i=1,2,\ldots,n$, and finds an optimal set of parameter vector (λ) by calibrating an appropriate model. The inputs to the model typically are the previous values of the time series and the output will be the forecast value. The model is normally trained and tested on training and testing sets extracted from the historical time series. In addition to previous time series values, one can utilize the values or forecasts of other time series (or external variables) that have a correlation or causal relationship with the series to be forecasted as inputs. For a river flow forecasting problem, such exogenous time series could be the rainfall or evaporation over the basin. Each additional input unit in a model adds another dimension to the space in which the data cases reside, thereby making the function to be mapped more complex. The model attempts to fit a response surface to these data.

For many applications in data driven modeling (regression, classification etc.) an estimate of expected response (output) is desired at or close to one fixed predictor (input) vector. This estimate should depend heavily on predictor vectors in the sample which are close to the given fixed predictor vector. The predictive relationship between the current state $\mathbf{x_t}$ and the next value of the time series can then be expressed as:

$$y(t+T) = f_T^0(x(t)) \tag{1}$$

where T is the prediction time horizon. The problem of one step ahead (T=1) predictive modeling is to find the mapping $f_T^0 :\ R^N\ to\ R^1$. In local approximation, a local predictor is constructed based on the nearby neighbors of $\mathbf{x}_t$, that is, fitting a polynomial to the pairs ($\mathbf{x}(t_i)$, $y(t_i+T)$) with $\mathbf{x}(t_i)$ being the nearest neighbors of $\mathbf{x}(t)$ for $t_i<t$. The original signal can also be viewed as an evolution of the state $\mathbf{x}(t)$ of a dynamical system in $\mathbf{R}^N$. The signal history (input samples) compose the map from state space of dimension N to a scalar space, the parameters of the mapping can be estimated by interpolating $f(\mathbf{x})$ from noisy signal samples. The local modeling is superior and simpler under the condition that the given dynamics is locally smooth and a long enough signal history is available (Singer et al., 1992). Under this condition, $f(\mathbf{x})$ can be approximated by the first few terms of multi dimensional Taylor series expansion, resulting in,

$$f(\mathbf{x}) = \mathbf{a}^T\mathbf{x} + b \tag{2}$$

Based on the fuzzy modeling approach originally developed by Takagi and Sugeno (1985), the global operation of a nonlinear process is divided into several local operating regions. Within each local region, R_i, a reduced order linear model is used to represent the processes behavior. Fuzzy sets are used to define the process operating conditions such that the dynamic model of a nonlinear p`rocess can be described in the following way:

R_i: IF operating condition i

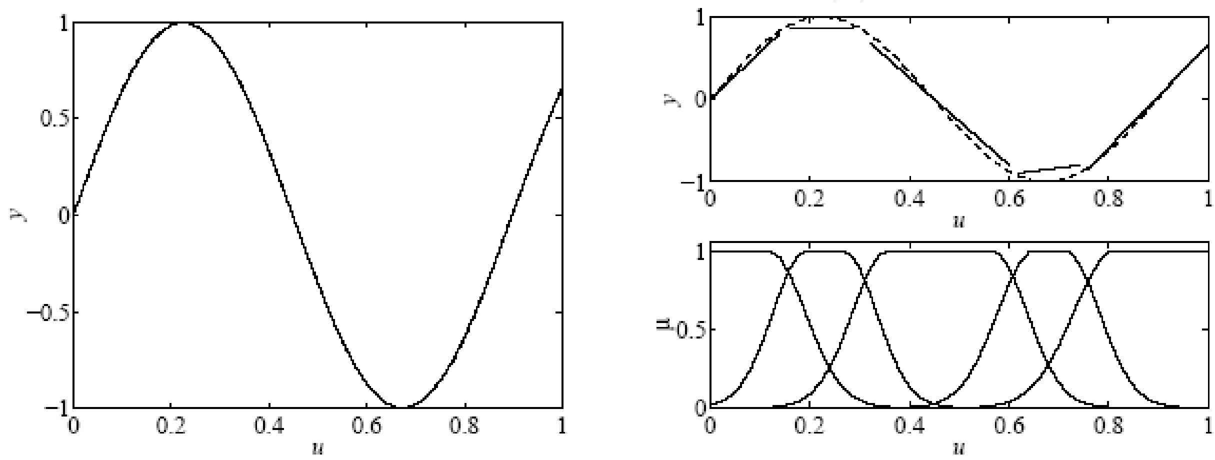
$$\text{THEN } \hat{y}_i(t) = \sum_{j=1}^{no} a_{ij} y(t-j) + \sum_{j=1}^{ni} b_{ij} u(t-j) \ \ (i=1, 2, 3, \ldots, nr) \tag{3}$$

The final model output is obtained by firing strength

$$\hat{y}(t) = [\sum_{i=1}^{nr} \mu_i \hat{y}_i(t)] / (\sum_{i=1}^{nr} \mu_i) \tag{4}$$

where y is the process output, $\hat{y}$ is the model prediction, u is the process exogenous input, $\hat{y}_i$ is the prediction of the processes output in the ith operating region, nr is the number of fuzzy operating regions, ni and no are the time lags in the input and the output respectively, μ_i is the membership function for the ith model, a_{ij} and b_{ij} are the local linear parameters, and t represents the discrete time.

Operating regions of a process can be defined by one or several process variables. A number of fuzzy sets, such as 'low, medium and high' are defined for each of these process variables. An operating condition is constructed through logical combinations of these variables which are used to define the process operating regions and assigned the fuzzy sets: 'low, medium and high'.



(a) The data (solid) and the model (dashed-dotted).

(b) Local models (top) and membership functions (bottom).

Fig. 1. Takagi-Sugeno fuzzy model computational procedure

The graphical representation for computational procedure for TS fuzzy model is presented for sine curve (Fig 1). From the figure it can be observed that if MFs can be developed for a given data set, then using the membership grade of each predictor variable, local linear models can be developed and their parameters can be estimated. The overlapping of different MFs and different local fuzzy regions are graphically presented in Fig 2. From the figure it is seen that different local linear models are developed from different fuzzy regions and fuzzy reasoning is applied to estimate the model output. Fuzzy regions are represented with different membership grade used in the fuzzy-if- then rules which is the crux of fuzzy modeling approach. Therefore, classification of different local regions in an input/output data set is very important in fuzzy modeling approach and fuzzy clustering technique is widely used for such purposes. In the clustering approach, classification is carried out using different distance measures. The degree of similarity can be calculated by using a suitable distance measure. Based on the similarity, data vectors are clustered such that the data within a cluster are as similar as possible, and data from different clusters are as dissimilar as possible.

Fig 3 gives an example of two clusters in $\mathbb{R}^2$ with prototypes $\mathbf{v}_1$ and $\mathbf{v}_2$. The partitioning of the data is expressed in the *fuzzy partition matrix* $\mathbf{U} = [\mu_{ij}]$ whose elements are the

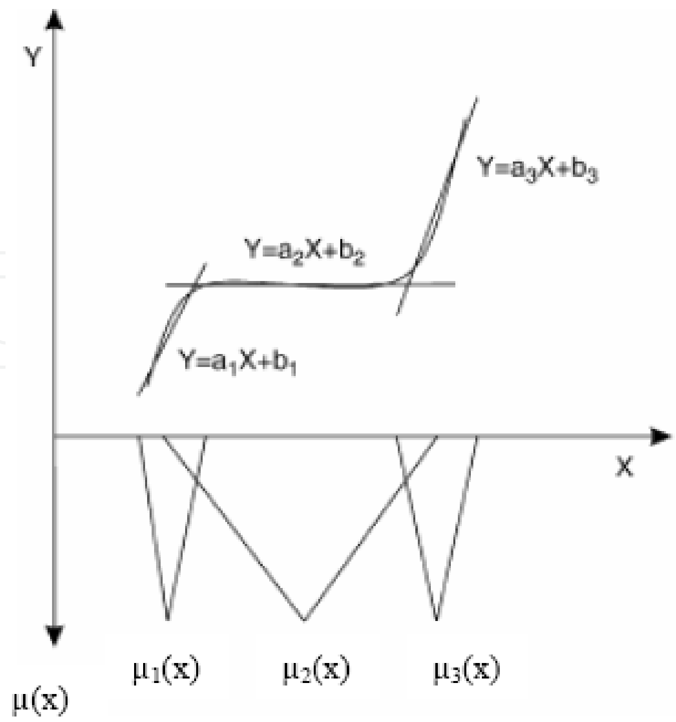


Fig. 2. Fuzzy local liner model developments

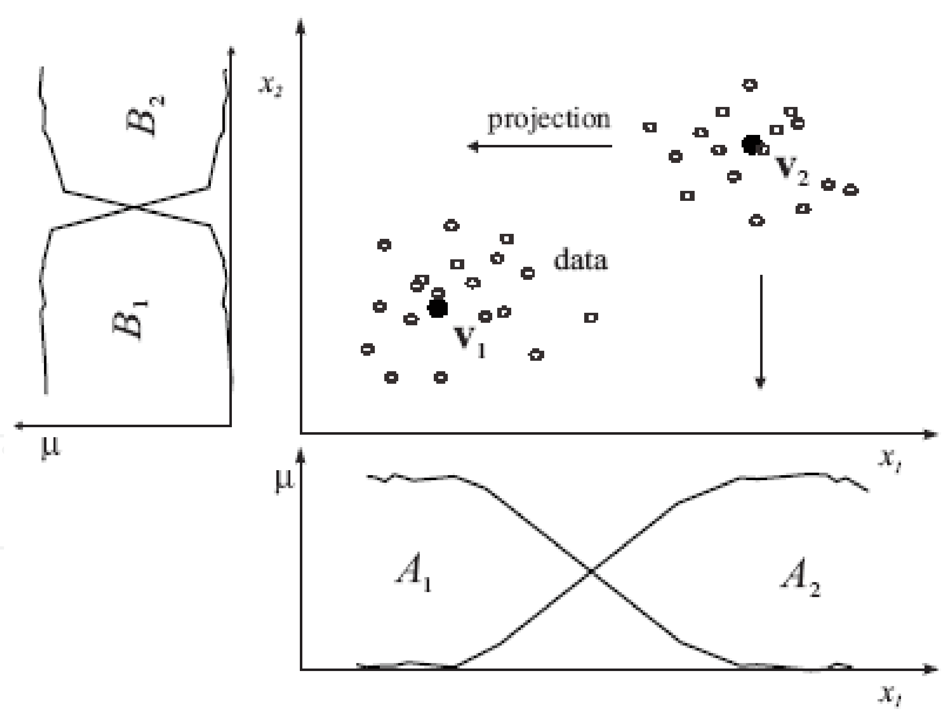


Fig. 3. Identification of membership functions through fuzzy clustering.

membership degrees of the data vectors $\mathbf{x}_k$ in the fuzzy clusters with prototypes $\mathbf{v}_j$. The antecedent membership functions are then extracted by projecting the clusters onto the individual variables. The performance of models of this kind depends heavily on the definition of neighboring state space. Euclidian distance is the most commonly employed measure of closeness to find out the neighboring state space. However, it is not usually an appropriate measure of closeness due to the curse of dimensionality. For instance, consider that the predictor distribution is uniform on a ball of radius unity in D dimensional space. For an expected response at or close to origin of the ball, it would be reasonable to use only sample vectors inside a ball of radius $\rho < 1$, assuming Euclidean distance as a measure of closeness. But, since the probability of a sample vector lying in the smaller ball is ρ^D, it is necessary that sample sizes are exponential in D to get enough close vectors for accurate estimation. To avoid this curse, we might assume that $f_i(\mathbf{x})$ has a ridge approximation (i corresponds to the fuzzy region),

$$f_i(x) \cong \sum_{n=1}^{l} c_n g_n(a^T x) \tag{5}$$

in which g_n is a transformation function of the linear combiner of x; a^T, c_n are the parameters of the ridge function; l is the number of sub domains. Also note that the linear local approximation is valid under the condition that the dynamics of $f(\mathbf{x})$ is locally smooth. However, equation 5 typically accounts for the nonlinear local dynamics of $f(\mathbf{x})$. A close examination of equation 5 reveals that the ridge function is an ANN with single hidden layer having l nodes, and a linear transfer function on the output layer.

Hence it is apparent that if the state space is classified into sub-domains, and each of these domains is modeled independently by a neural network approach, which when combined together, the resulting model may provide a better global modeling of the nonlinear dynamics in the state space. It appears that this heuristic has not been addressed or confirmed by empirical trials. The present study illustrates this heuristic by comparing the performance of models developed using the proposed approach, local linear approximation and a global nonlinear approach.

3. Methodology

In general, the above discussed nonlinear local model fitting is composed of two steps: a set of nearby state searches over the signal history and model parameter fitting. For a given signal, this procedure results in a set of local model parameters which when combined together provide a single function over the entire space. Since the neighborhood search is performed over the whole signal history a lot of redundant computation results which in turn hinders effective implementation of this approach. These redundant computations can be avoided by classifying the state spaces into homogenous subspaces by means of an appropriate vector clustering technique.

4. Clustering for classification

The objective of cluster analysis is the classification of objects according to similarities among them, and organizing of data into groups. Clustering techniques are among the *unsupervised* methods, they do not use prior class identifiers. The main potential of clustering is to detect the underlying structure in data, not only for classification and pattern recognition, but for model reduction and optimization. Various definitions of a cluster can be formulated, depending on the objective of clustering. Generally, one may accept the view that a cluster is a group of objects that are more similar to one another than to members of other clusters. The term "similarity" should be understood as mathematical similarity, measured in some well-defined sense. In metric spaces, similarity is often defined by means of a *distance norm*. Distance can be measured among the data vectors themselves, or as a distance from a data vector to some prototypical object of the cluster.

Since clusters can formally be seen as subsets of the data set, one possible classification of clustering methods can be according to whether the subsets are fuzzy or crisp (hard). Hard clustering methods are based on classical set theory, and require that an object either does or does not belong to a cluster. Hard clustering in a data set **X** means partitioning the data into a specified number of mutually exclusive subsets of **X**. The number of subsets (clusters) is denoted by c. Fuzzy clustering methods allow objects to belong to several clusters simultaneously, with different degrees of membership. The data set **X** is thus partitioned into c fuzzy subsets. In many real situations, fuzzy clustering is more natural than hard clustering, as objects on the boundaries between several classes are not forced to fully belong to one of the classes, but rather are assigned membership degrees between 0 and 1 indicating their partial memberships. Fuzzy clustering can be used to obtain a partitioning the data where the transitions between the subsets are gradual rather than abrupt. Most analytical fuzzy clustering algorithm is fuzzy c-means (FCM) clustering. The FCM clustering algorithm is based on the minimization of an objective function called *C-means functional.* In the current study subtractive clustering algorithm is used for classification. Subtractive clustering method (Chiu, 1994) is an extension of the FCM and mountain clustering method (Yager and Filev, 1994), where the potential is calculated for the data rather than the grid points defined on the data space. As a result, clusters are elected from the system training data according to their potential. Subtractive clustering compared to mountain clustering has an advantage that there is no need to estimate a resolution for the grid.

5. Fuzzy non linear local approximation model

As discussed earlier a novel hybrid model is proposed herein which performs independent nonlinear local approximation, and combines each of them using the fuzzy framework. The architecture of the proposed model is presented in Fig 4. The method is based on the concept that the input space is divided into sub regions of similar dynamics using an appropriate clustering algorithm (subtractive clustering algorithm in this study), and modeling of each of these regions is carried out using nonlinear local function approximation (ANN in this study). The proposed model is termed as Fuzzy Non Linear Local Approximation (FNLLA) model, which is based on fuzzy concept and neural technique is applied for nonlinear local function approximation.

In Fig 4, the gating is done to identify the membership grade associated with any given input vector for each of the clusters. This input vector will be passed to each of the developed ANN, and the output from each ANN is combined by computing the weighted mean with the membership grade in each cluster.

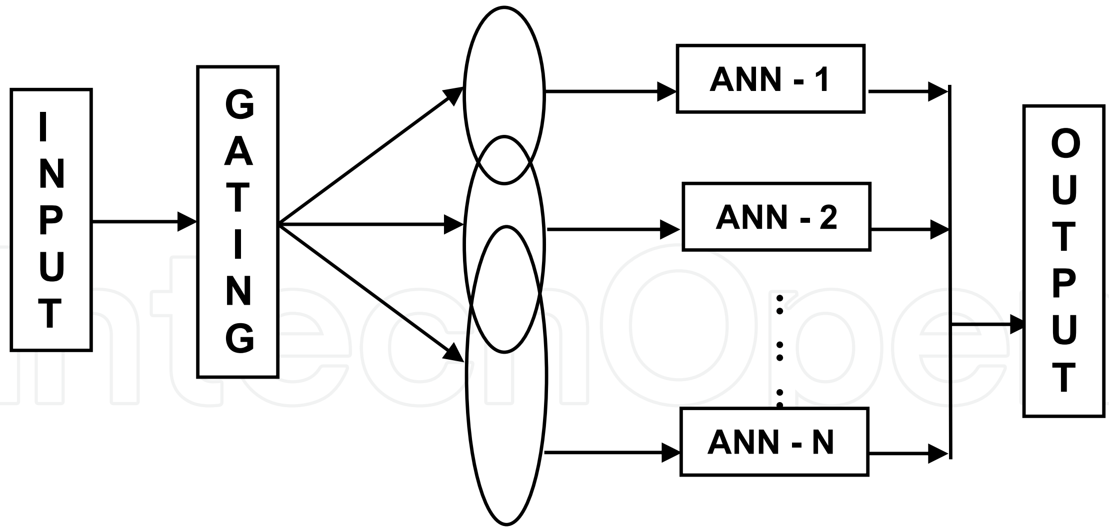


Fig. 4. Schematic representation of the proposed FNLLA model

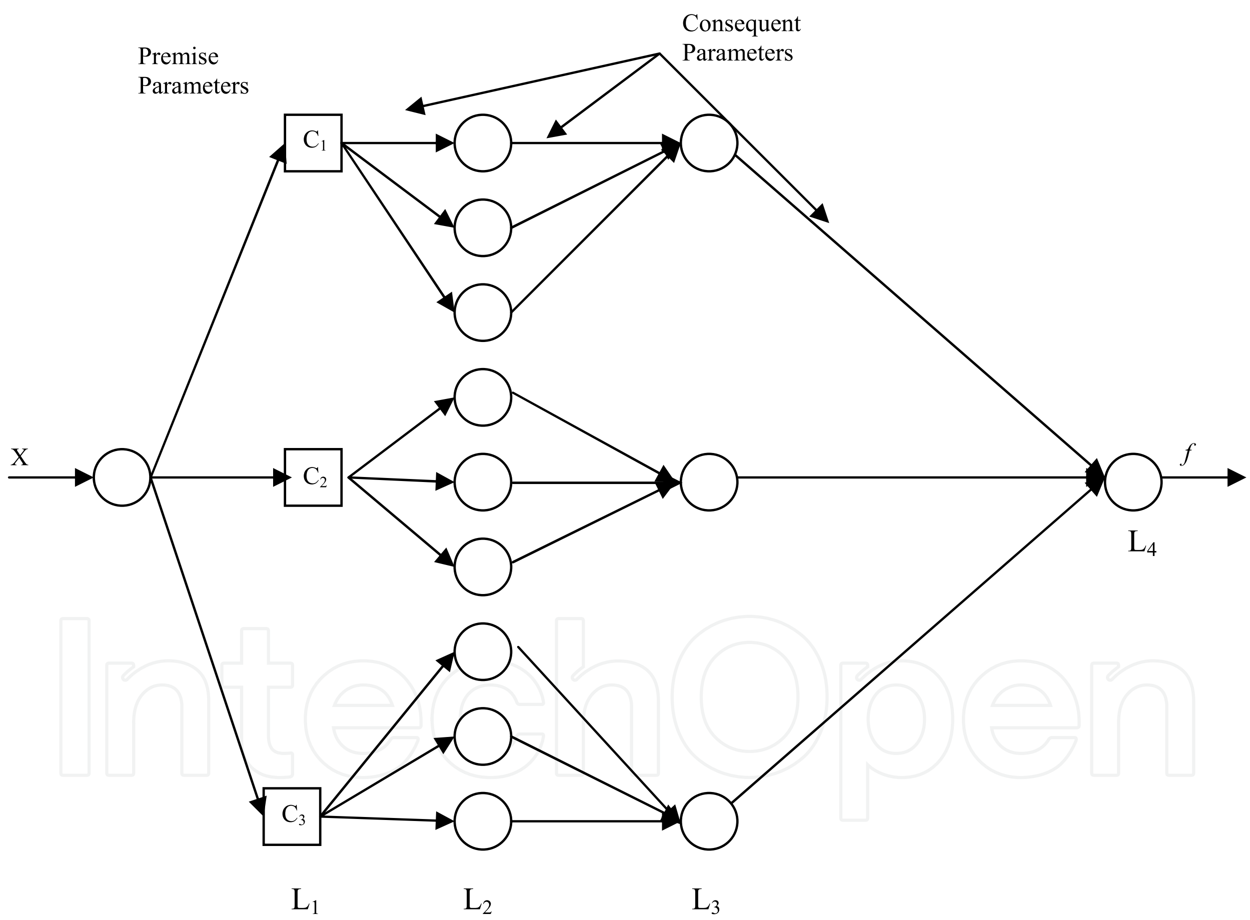


Fig. 5. Computational architecture of the proposed FNLLA model

The architecture of FNLLA is depicted in Fig 5, in which the fuzzy computing scheme is represented in an adaptive neural network structure. The consequent part of each of the fuzzy rule is a nonlinear function (ANN in this case). The computations are performed in 4 layers in FNLLA. In the layer 1, (L_1), the incoming input vector (**x**) is passed to different sub regions (C_1, C_2 and C_3) and the associated MFs are computed as:

$$\mu_j = e^{-\alpha \left\| x_j - x_i \right\|^2} \quad (6)$$

in which x_i is the i^{th} input vector, x is the cluster centre and j the sub-region number, and α is a function of the cluster radius. In layer 2 (L_2), the weights of the hidden nodes of the ANN consequent models are estimated, and the output from each ANN consequent model is arrived at layer 3 (L_3). At layer 4 (L_4), each of these consequent outputs are combined to arrive at the final output.

6. Demonstrative case examples

The proposed model is illustrated through two examples by developing rainfall-runoff models: (a) for Kolar basin up to the Satrana gauging site in India, (b) the Kentucky basin, USA. The Kolar River is a tributary of the river Narmada that drains an area of about 1350 sq km before its confluence with Narmada near Neelkanth. In the present study the catchment area up to the Satrana gauging site is considered, which constitutes an area of 903.87 sq km (Fig 6). The 75.3 km long river course lies between north latitude $21^0 09'$ to $23^0 17'$ and east longitude $77^0 01'$ to $77^0 29'$. For the study, rainfall and runoff data on an hourly interval during the monsoon season (July, August, and September) for three years (1987–1989) are used. The rainfall data available were in the form of areal average values in the basin. The total available data has been divided into two sets, calibration set (data during the years 1987–1988) and validation set (data during the year 1989). Different models for lead times of up to 6 hours have been developed in the study.

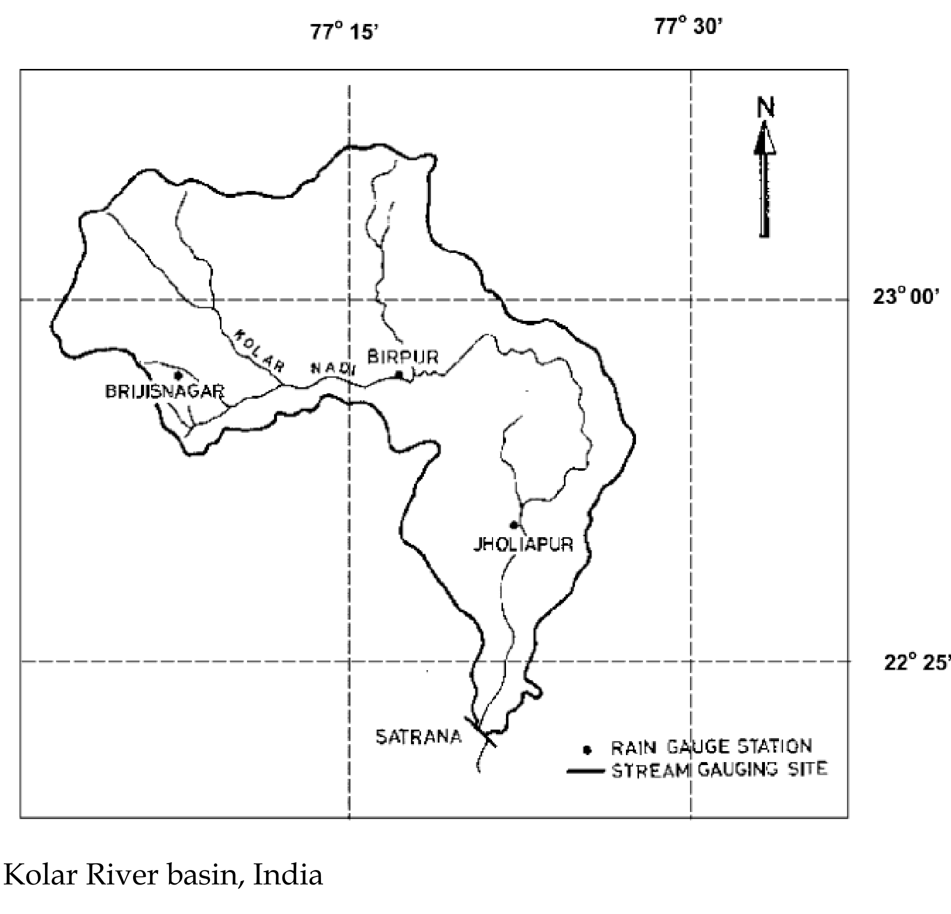


Fig. 6. Map of Kolar River basin, India

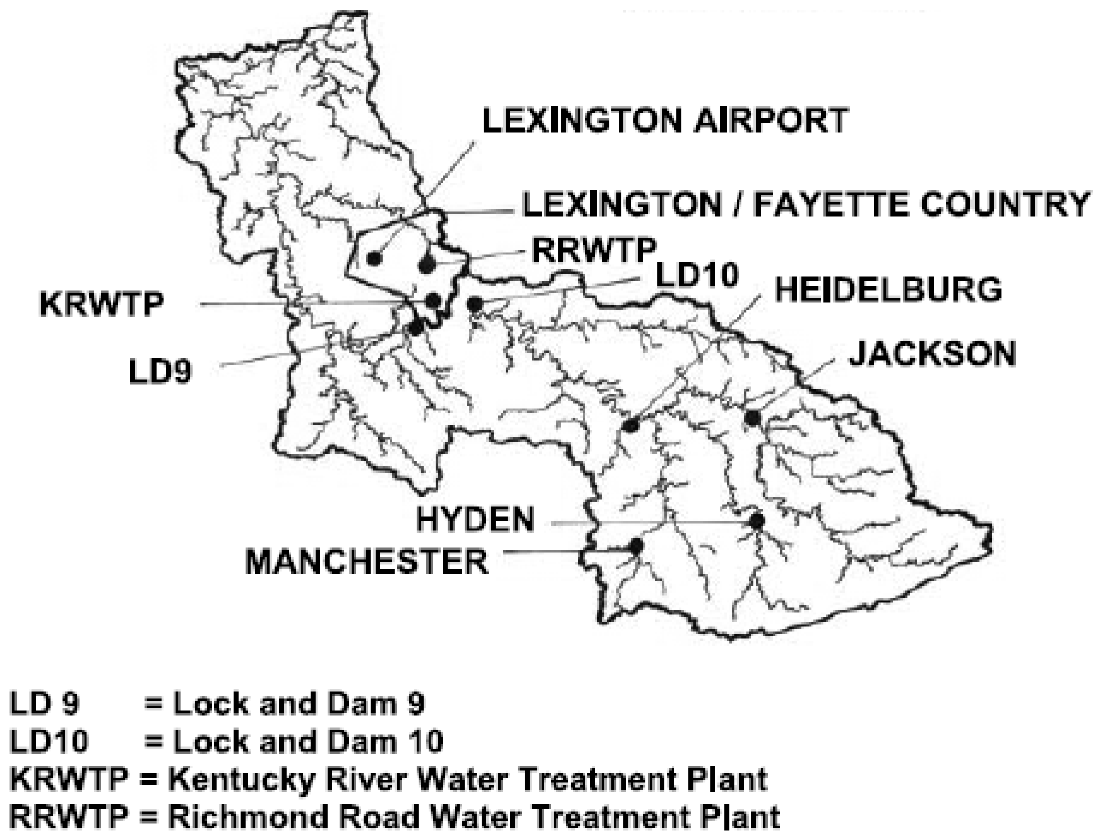


Fig. 7. Study area map of Kentucky River basin in USA

The Kentucky River basin covers over 4.4 million acres of the state of Kentucky. Forty separate counties lie either completely or partially within the boundaries of the river basin. The Kentucky River is the sole water source for several water supply companies of the state. There is a series of fourteen locks and dams on the Kentucky River, which are owned and operated by the US Army Corps of Engineers. The drainage area of the Kentucky River at Lock and Dam 10 (LD10) near Winchester, Kentucky is approximately 6,300 km^2 (Fig 7).

The data used in the study presented in this paper include average daily streamflow (m^3/s) from the Kentucky River at LD10, and daily average rainfall (mm) from five rain gauges (Manchester, Hyden, Jackson, Heidelberg, and Lexington Airport) scattered throughout the Kentucky River Basin. The total length of the available rainfall runoff data was 26 years (1960-1989 with data in some years missing).

The input vector identified, according to Sudheer et al. (2002) for modeling the river flow in Kolar, included a total number of 4 variables. Accordingly, the functional form of the model, in the case of Kolar, for rainfall runoff modeling is given by:

$$Q(t) = f[R(t-9), R(t-8), R(t-7), Q(t-1)] \quad (7)$$

where Q(t) and R(t) are river flow and rainfall respectively at any time t in hour.

The functional form for rainfall-runoff dynamic for Kentucky River basin is given by:

$$Q(t) = f[R(t), R(t-1), R(t-2), Q(t-2), Q(t-1)] \quad (8)$$

Different statistical indices that are employed to estimate the model performance include coefficient of correlation (CORR), efficiency (EFF), root mean square error (RMSE) and noise to signal ration (NS).

7. Results and discussions

7.1 Parameter Estimation in FNLLA

The optimal number of clusters in FNLLA has been obtained by varying cluster radius in the subtractive clustering algorithm, combined with the ANN model development. Different ANN models are developed for different clusters of input space changing hidden neuron from 2 to 10. Single hidden layer with sigmoid function nodes is used in the ANN. The sigmoid activation function is considered in the output layer also. A standard back propagation algorithm with adaptive learning rate and momentum factor has been employed to estimate the network parameters for different clusters. In order to have a true evaluation of the proposed nonlinear local approximation in fuzzy models, the result obtained for both the basins from FNLLA was compared with FIS, which performs a linear local approximation. In FIS model, subtractive clustering has been used for fuzzy model identification which includes optimal number of if-then-rule generation and consequent parameters are optimized using least square error (LSE) technique.

Kolar River basin		Kentucky River basin	
Cluster Radius	Number of clusters	Cluster Radius	Number of clusters
0.050 - 0.060	7	0.0010 - 0.0080	4
0.061 - 0.074	6	0.0081 - 0.0100	3
0.075 - 0.079	5	0.0110 - 0.2000	2
0.080 - 0.100	4	0.2100 - 1.0000	1
0.110 - 0.200	3		
0.210 - 0.300	2		
0.030 - 1.000	1		

Table 1. Partition of input space for Kolar and Kentucky basin

In the analysis, the radius of influence (r_a) of the cluster centre is fixed by various trials, which is the foremost interest for the current study. The value of r_a is varied from 0.001 to 1.0 with a step size of 0.01; at each stage number of clusters is estimated. From the current data set 1 to 7 clusters are found while changing the cluster radius from 0.05 to 1 for Kolar river basin and maximum 4 clusters are observed by changing radius from 0.001 to 1 for Kentucky basin. The number of clusters identified corresponding to various radiuses for both the basins are presented in Table 1. It is evident from Table 1 that as the cluster radius increases, the number of cluster decreases. Different ANN models are developed for data belonging to different clusters (effectively representing different ranges of flow), and as discussed earlier hidden neurons are by trial and error procedure. The stopping criteria for ANN model building was maximum efficiency.

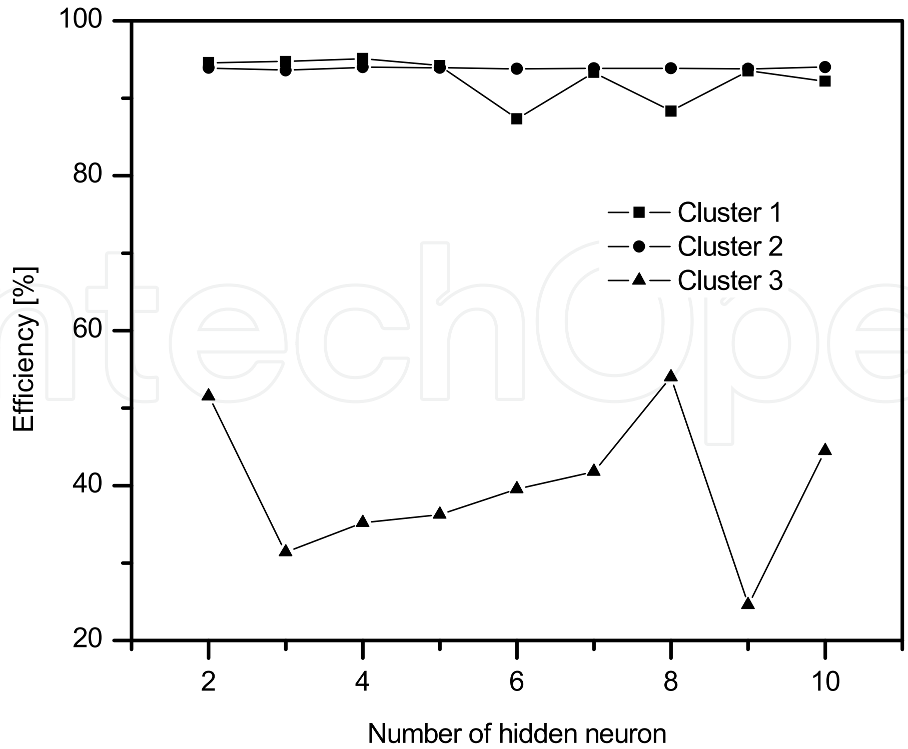


Fig. 8. Variation of Efficiency Plot for three clusters with respect to hidden neurons

It was observed that model performance was good when input space is classified into 3 clusters for both the basins. The efficiency plots with respect to different hidden neurons for 3 ranges of flows are presented in Fig.8 for Kolar basin. From the Fig 8 it is observed that the maximum efficiency is for the models having hidden neurons are 4, 4 and 8 for individual clusters for Kolar River basin. Similar procedure has been followed for optimization of model parameter for Kentucky River basin. The optimum hidden neuron obtained 3, 3 and 3 for Kentucky basin. The identified optimal clustering radius with number of if-then rules for Kolar and Kentucky basin are furnished in Table 2.

Model	Kolar River basin	Kentucky River basin
FIS	0.20 [3 rules]	0.008 [4 rules]
FNLLA	0.12 [3 ranges/rules]	0.010 [3 ranges/rules]

Table 2. Optimal model structure for three sub-domains for Kolar and Kentucky basin

The summary statistics of the flow data belonging to each of the identified clusters are presented in Table 3. It is clear from Table 3 that the clustering based on the input vector clearly identifies distinct clusters that have different nonlinear dynamics. This is evident from overlapping clusters; the cluster C1 contain flow range from 4.67 m^3/s to 240.55 m^3/s, and the cluster C2 contain flow range from 3.63 m^3/s to 240.24 m^3/s, in the case of Kentucky river basin data. A similar observation is found in the case of Kolar basin data too. It is worth mentioning that the consequent ANN models for each of these clusters preserve the summary statistics much effectively in both basins (see Table 3). Also, it is evident that the classification of the data into different clusters is according to the range of flow, though not exclusively forced by the FNLLA.

Statistic	C1		C2		C3	
	Observed	Computed	Observed	Computed	Observed	Computed
Kentucky River Basin						
Mean	13.45	13.19	32.25	32.13	239.16	239.08
Standard Deviation	13.08	12.05	25.83	20.79	281.21	272.51
Minimum	3.63	4.67	9.61	15.09	15.99	31.55
Maximum	240.24	240.55	297.43	266.96	2553.90	2293.20
Kolar River Basin						
Mean	4.09	4.07	12.47	12.43	142.73	138.69
Standard Deviation	1.94	1.74	4.55	4.27	308.65	290.72
Minimum	1.62	1.58	6.83	7.04	0.88	19.26
Maximum	29.03	28.86	125.57	124.84	2427.70	2062.40

Table 3. Summary statistics of the river flow in the identified sub-domain by FNLLA Model

Lead time	1-hour			3-hour			6-hour		
Cluster	C1	C2	C3	C1	C2	C3	C1	C2	C3
Calibration									
Correlation	0.82	0.98	0.91	0.64	0.90	0.82	0.94	0.77	0.91
Efficiency (%)	66.91	95.40	83.33	39.21	79.58	67.56	24.49	53.18	53.35
RMSE	2.62	66.18	0.79	5.88	139.41	1.92	11.03	210.71	3.43
Noise to Signal Ratio	0.58	0.21	0.41	0.78	0.45	0.57	0.39	0.69	0.41
Validation									
Correlation	0.93	0.98	0.45	0.63	0.87	0.48	0.91	0.75	0.74
Efficiency (%)	87.08	95.28	81.82	37.41	74.55	64.24	30.15	52.98	55.06
RMSE (m^3/s)	1.43	53.97	1.50	4.03	125.44	2.43	9.45	170.65	3.15
Noise to Signal Ratio	0.36	0.22	0.96	0.79	0.51	0.93	0.49	0.69	0.43

Table 4. Cluster wise FNLLA performance for Kolar Basin

The results of the FNLLA were first analyzed for its effectiveness in capturing the nonlinear dynamics at local level. In order to achieve this, the performance indices were computed for each sub-domain for Kolar basin for the calibration and the validation period, and are presented in the Table 4. Note that in the study the FNLLA model classified 1975, 1782 and 593 patterns as low, medium and high flow respectively using the subtractive clustering algorithm. It is evident from the Table 4 that the nonlinear consequent models of the FNLLA are effective in capturing the nonlinear dynamics in each sub regions.

7.2 Performance of FNLLA at 1 step-ahead forecast

The values of various evaluation measures during calibration and validation period for FNLLA and FIS for both the basins for 1-hour lead forecast are summarized in the Table 5, from which it can be observed that both the models possess high value of correlation (0.96 and more) between the forecasted and the observed river flow at 1-hour ahead. The high value for the efficiency index indicates a very satisfactory model performance in capturing the nonlinear dynamics involved in the rainfall-runoff processes. Note that the FNLLA performs with higher efficiency in the case of Kolar basin compared to the FIS model. The value of RMSE varies from 24 m^3/sec to 42 m^3/sec for Kolar basin indicating a very good performance by both the model; the RMSE values are relatively less in Kentucky basin for FNLLA model. It is also noted that the NS ratio for the FNLLA model is less than that corresponding to the FIS model.

Statistical Indices	FNLLA		FIS	
	Calibration	Validation	Calibration	Validation
Kolar River Basin				
Correlation	0.98	0.99	0.98	0.98
Efficiency	95.84	96.72	95.91	96.61
RMSE (m^3/s)	42.39	23.29	42.06	24.31
Noise to Signal ratio	0.20	0.11	0.34	0.12
Kentucky River Basin				
Correlation	0.98	0.97	0.96	0.96
Efficiency	95.68	93.87	92.17	91.68
RMSE (m^3/s)	50.12	51.87	67.45	60.41
Noise to Signal ratio	0.21	0.25	0.28	0.29

Table 5. Statistical Indices for 1-hour and 1- day lead forecast for Kolar and Kentucky basin

The scatter plots of flows for 1-hour lead forecast validation period for both the models are presented in Fig 9 for Kolar basin (Fig 10 for Kentucky basin for 1 day lead). These plots give clear indication of the simulation ability of the developed model across the full range of flows. It is noted that most of the flows tend to fall close to the 45^0 line (reduced scattering), showing a good agreement between observed and forecasted flows. From the plot it is observed that both the models are quite competent in forecasting river flow at 1 hour/day lead time. In general it is noted that the mapping of the low flow region is relatively better compared to high flow region.

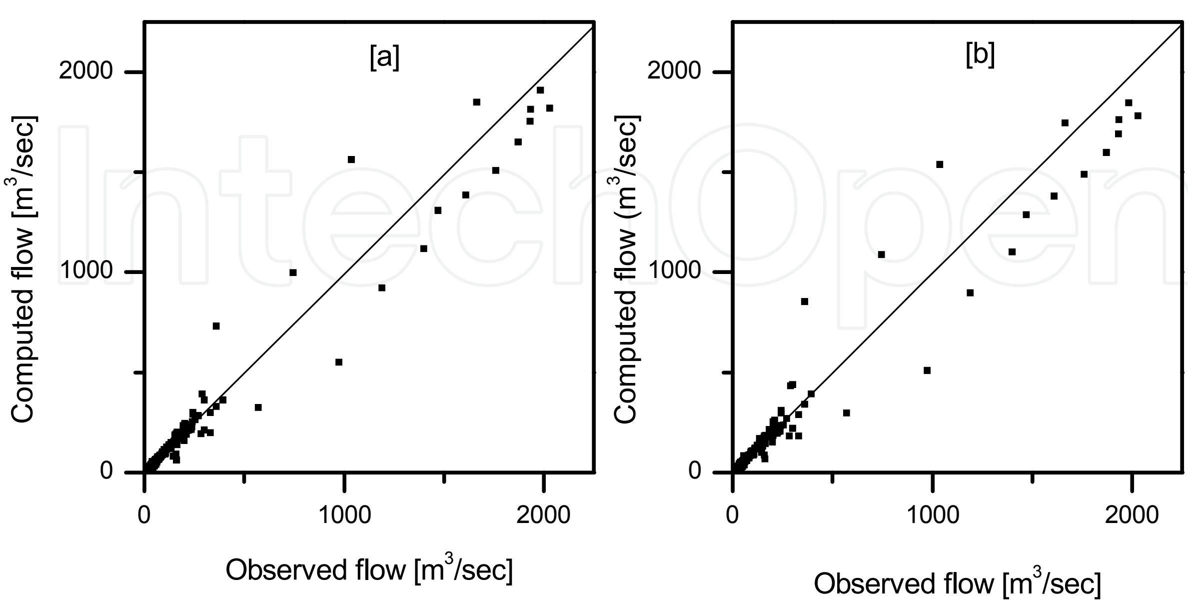


Fig. 9. Scatter plots for observed and computed flows by both models at 1 hour lead time for Kolar basin (a) FIS (b) FNLLA

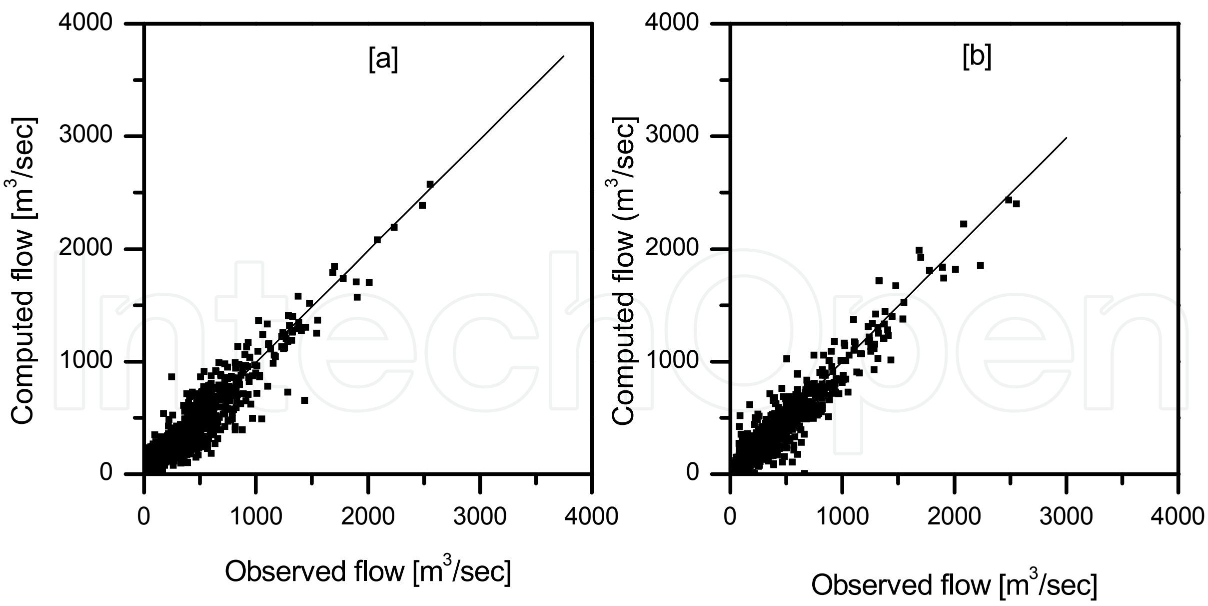


Fig. 10. Scatter plots for observed and computed flows (a) FIS (b) FNLLA model for 1 day lead forecast for Kentucky basin

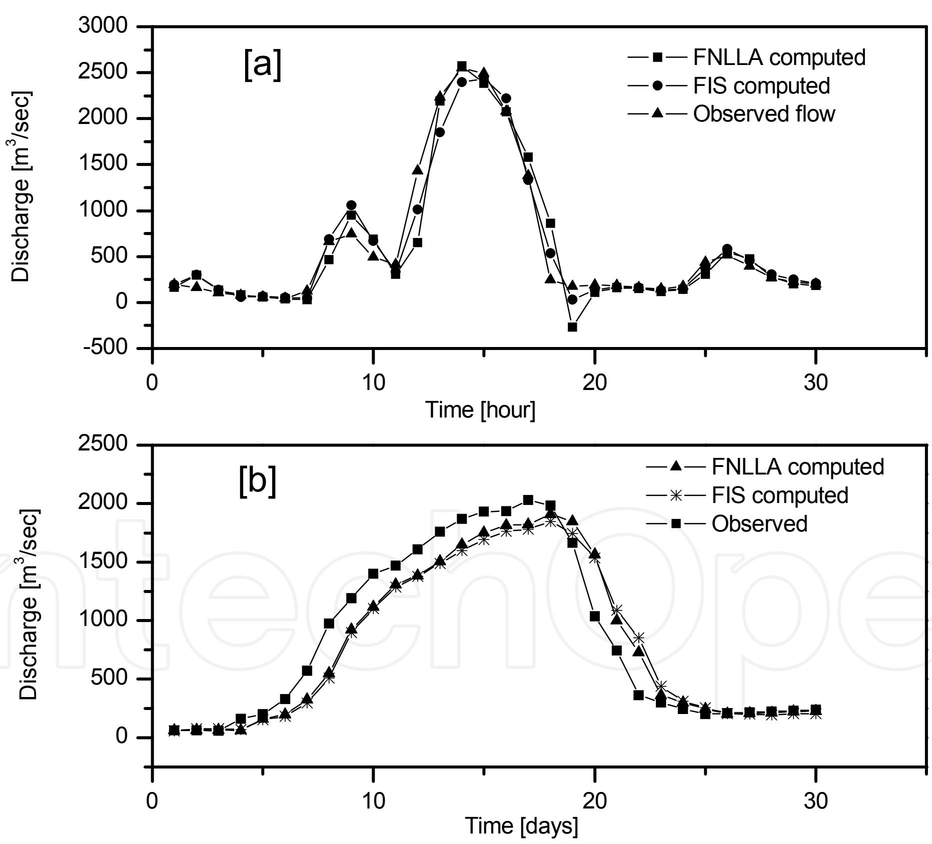


Fig. 11. Comparison plot between FIS and FNLLA model for a typical storm event for (a) Kolar basin (b) Kentucky basin

The forecasted hydrograph for a typical flood event (during validation period) for both the basins by both the models are presented along with its observed counterpart in Fig 11. It can be observed from Fig 11 that FNLLA is preserving the peak flows effectively than the FIS for

1 hour/day lead forecast for Kolar and Kentucky basin, while in low and medium ranges of flow the performance of both models is similar. This observation effectively brings out the capturing of nonlinear dynamics at the local regions of the input space.

7.3 Performance of FNLLA at higher lead time forecasts

Both the models are further evaluated for their effectiveness to forecast flows at higher lead times. The statistical indices for higher lead forecasting are presented in Table 6 for both the basins. The superior performance of the FNLLA compared to FIS is clearly visible at higher lead times from the results presented in Table 6. It is observed that even though the performance of FNLLA deteriorates as the lead time increases, it falls down to only 79.15% efficiency at 6 hours ahead, while FIS shows only 51.00% efficiency at the same lead time. A similar behavior is exhibited by FNLLA in the case of Kentucky basin also, as the efficiency is 77.23% at 3 days ahead compared to 46.92% in the case of FIS. A similar argument holds well in the case of other performance indices too.

	FNLLA			FIS		
Kolar River Basin						
Forecast Lead time	1 hour	3 hour	6 hour	1 hour	3 hour	6 hour
Calibration						
Efficiency (%)	95.84	81.52	79.15	95.91	57.79	51.00
RMSE (m^3/s)	42.39	89.40	94.95	42.06	135.11	145.70
Noise to Signal ratio	0.20	0.43	0.923	0.34	0.65	1.95
Validation						
Efficiency (%)	96.72	79.95	77.73	96.61	50.54	46.92
RMSE (m^3/s)	23.29	60.45	62.37	24.31	92.12	96.48
Noise to Signal ratio	0.11	0.35	0.39	0.12	0.77	0.83
Kentucky River Basin						
Forecast Lead time	1 day	2 day	3 day	1 day	2 day	3 day
Calibration						
Efficiency (%)	95.68	84.21	64.88	92.17	73.58	52.50
RMSE (m^3/s)	50.12	95.82	142.9	67.45	123.93	166.20
Noise to Signal ratio	0.21	0.40	0.59	0.28	0.51	0.69
Validation						
Efficiency (%)	93.87	78.68	58.66	91.68	73.14	50.56
RMSE (m^3/s)	51.87	96.68	134.63	60.41	108.53	147.24
Noise to Signal ratio	0.25	0.46	0.64	0.29	0.52	0.70

Table 6. Performance indices for FIS and FNLLA model at higher forecast lead times

8. Summary and conclusions

The objective of this chapter was to present and illustrate a nonlinear local approximation approach in modelling the rainfall-runoff process which offers better accuracy in the context of river flow forecasting. Based on the theoretical considerations of the fuzzy modelling in the state space (input-output), it is clear that if the state space is classified into sub-domains and each of these domains is modeled independently by a neural network approach which are combined together, it may provide a better global modeling of the nonlinear dynamics in the state space. In general, the proposed nonlinear local model fitting is composed of two steps: a set of nearby state searches over the signal history and model parameter fitting. For a given signal, this procedure results in a set of local model parameters which when combined together provide a single function over the entire space. Since the neighborhood search is performed over the whole signal history, a lot of redundant computations result which in turn hinders effective implementation of this approach. These redundant computations can be avoided by classifying the state spaces into homogenous subspaces by means of an appropriate vector clustering technique. The proposed model is termed as Fuzzy Non Linear Local Approximation (FNLLA) model, which is based on fuzzy concept and neural technique is applied for nonlinear local function approximation. The partition of the state space is achieved by subtractive clustering algorithm and nonlinear local function approximation is by ANN in the proposed method. The antecedent parameters of the model are simultaneously estimated during clustering, and the standard back propagation algorithm is employed for ANN parameter estimation.

The potential of FNLLA is illustrated using two case examples: (i) data pertaining to Kolar River basin, and (ii) data corresponding to Kentucky River basin. The optimal architecture of the FNLLA model, which is defined by the number of sub-regions in the data and the structure of each ANN, is arrived after a trial and error procedure. When the performance of the FNLLA is compared with that of a pure FIS, it is observed that the FNLLA certainly possesses the advantages of nonlinear mapping. Though both the models perform similar at 1 step-ahead forecasts, the FNLLA performs much better than FIS at higher lead times. Overall, the results of the study confirm the heuristic that a nonlinear local approximation is a better approach in fuzzy modeling especially when complex nonlinear functions are being mapped.

9. References

Chiu, S. (1994) Fuzzy Model Identification Based on Cluster Estimation. *Journal of Intelligent and Fuzzy Systems*, 2(3), 267-278,

Farmer, J.D. and J.J. Sidorowitch (1987) Predicting chaotic time series. *Physical. Review Letter*, 5(59), 845-848.

Jain, A., K.P. Sudheer, and S. Srinivasulu (2004) Identification of physical processes inherent in artificial neural rainfall-runoff models. *Hydrological Processes*, 18(3), 571-581.

Singer, A.C., G. Wornell, and A. Oppenheim (1992) Codebook prediction: a nonlinear signal modeling paradigm. *IEEE, In proc. Int. Conf. Acoustics, Speech & signal processing*, San Francisco, 5, 325-328.

Sudheer, K.P. (2005) Knowledge extraction from trained neural network river flow models. *Journal of Hydrologic Engineering*, ASCE, 10(4), 264-269.

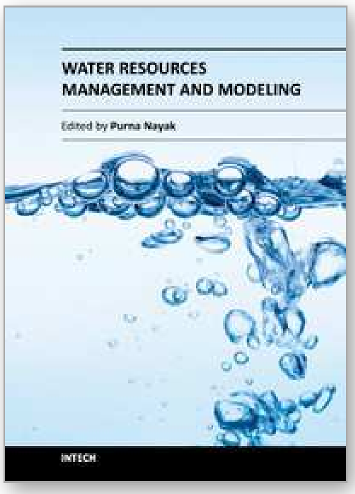
Sudheer, K.P., A.K. Gosain, and K.S. Ramasastri (2002) A data-driven algorithm for constructing artificial neural network rainfall-runoff models. *Hydrological Processes*, 16, 1325-1330.

Tokar, S., and Markus, M. (2000). *Precipitation-runoff modeling using artificial neural networks and conceptual models.* Journal of hydrologic engineering, ASCE, 5(2): 156-161.

Wilby, R.L., R.J. Abrahart, and C.W. Dawson (2003) Detection of conceptual model rainfall–runoff processes inside an artificial neural network. *Hydrological Sciences Journal*, 48(2), 163–181.

Yager, R. and D Filev (1994) Generation of fuzzy rules by Mountain clustering. *Journal of Intelligent &Fuzzy Systems*, 2(3), 209-219.

Zhang, B. and R.S. Govindaraju (2000) Prediction of watershed runoff using Bayesian concepts and modular neural networks. *Water Resources Research,* 36(3), 753-762.



Water Resources Management and Modeling

Edited by Dr. Purna Nayak

ISBN 978-953-51-0246-5
Hard cover, 310 pages
Publisher InTech
Published online 21, March, 2012
Published in print edition March, 2012

Hydrology is the science that deals with the processes governing the depletion and replenishment of water resources of the earth's land areas. The purpose of this book is to put together recent developments on hydrology and water resources engineering. First section covers surface water modeling and second section deals with groundwater modeling. The aim of this book is to focus attention on the management of surface water and groundwater resources. Meeting the challenges and the impact of climate change on water resources is also discussed in the book. Most chapters give insights into the interpretation of field information, development of models, the use of computational models based on analytical and numerical techniques, assessment of model performance and the use of these models for predictive purposes. It is written for the practicing professionals and students, mathematical modelers, hydrogeologists and water resources specialists.

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821