# Integration of GIS & Soft Computing Techniques for Enhancing Flood Forecasting and Monitoring on the River – A case Study of Wainganga Tributary of River Godavari.

**Introduction:**

Existing flood forecasting models are highly data specific, and their operational performance depends upon the science used to build and operate these models as well as their ability to respond to dynamic and rapidly changing events. Soft computing is one of the latest approaches for the development of systems, which possess computational intelligence. It attempts to integrate several different computing paradigms including artificial neural networks, fuzzy logic and genetic algorithms to create 'smart' systems. The methodology may provide a good performing, low-cost solution that may readily be integrated into existing operational flood forecasting systems to provide a performance enhancement.

Geographical Information System (GIS) is an information system that is specially designed for handling spatial (or geographical) data. It combines a set of interrelated software components that create, edit, manipulate, analyze and display data both in text and graphic forms. GIS supports spatial analysis and modeling within the discipline of geography (e.g. location, proximity and spatial distribution), so that it becomes a vital tool for modern geography.

Integration of GIS and Soft Computing techniques for flood forecasting of Wainganga, a tributary of River Godavari, would help the user community to monitor the floods effectively. The river level forecaster of soft computing can able to forecast the river levels of the given points with necessary inputs, whereas, GIS helps in delineating the extent of flood coverage of the Wainganga tributary.

**Objectives:**

To assess the potential improvements in performance that can be achieved by using soft computing technologies for real-time flood forecasting.

To develop digital dataset for Wainganga (Pranahita) tributary, which is frequently affected with floods.

To integrate GIS and flood forecaster for monitoring the river level to avert the damage due to floods.

**Datasets:**

Imageries, Topographical sheets, field Survey data for creating spatial data.

Intensity of Rainfall, Transport time, Hydrographs, Threshold values (high / intermediate / low), river level etc..

**Methodology:**

Development of digital data:

Flood prone area can be obtained from classified imageries of a particular location from pre and post flood imageries of Godavari River. The topographical sheets of corresponding location can be digitized and geocoded. Field survey using GPS shall be carried for finding suitability of rainfall gauge stations and watershed mapping. The severity of rainfall in the upstream areas i.e., Pranahita and upper Wainganga tributaries cause reaching the alarming levels of River Flow at upstream & downstream flow of Godavari at Badrachalam.

Training Neural nets:

Data files can be uploaded that define a river (sub) network. These data are then used to build a database for training neural net models. Uploading long time-series for river gauging and rainfall stations. Perhaps a better method is to upload several, shorter time-series, each defining a flood event for the river (sub) network. Sections of the hydrograph are then classified into high/low or high/intermediate/low sets and a neural net model is trained for each set. New, unseen data may then be uploaded to test the trained models.

Model building can proceed as follows:

1. Uploading principal gauging station time-series, plus up to five additional gauging station/rainfall gauge time-series defining one or more events for a river (sub) network. If necessary, upload an additional input, such as mean daily temperature, weekly soil moisture, etc.
2. Specifying the mean transport time from upstream stations to the principal station.
3. Specifying the threshold for the classification of hydrographs into high/low or high/intermediate/low sets.
4. Specifying database construction parameters (encoding, normalisation, data selection), data resolution and prediction time.
5. Uploading data for further flood events of interest, if required.
6. Specifying neural net training parameters and training time.
7. Checking progress of training and continue if necessary.
8. Saving neural net models.
9. Uploading unseen data to test the neural net models.

Three soft computing techniques can be employed in this study: artificial neural networks, fuzzy logic and genetic algorithms. A brief explanation of each technique follows.

*Artificial neural networks*

An artificial neural network is a type of biologically inspired computational model based loosely on the functioning of the human brain. The broad set of characteristics which both real and artificial neural networks share is the ability to learn and adapt, generalization to unseen data, distributed processing, robustness and fault tolerance. The mapping of inputs to outputs is performed by a series of simple processing nodes or neurons distributed in a set of interconnected layers. The connections are associated with adaptable weights that change as the network learns the correct patterns. To date, artificial neural networks have been applied successfully to a variety of application areas including pattern recognition, classification, optimization problems and dynamic modeling.

Two types of artificial neural network are used in this study: a **m**ulti-**l**ayer **p**erceptron (MLP) trained with the back propagation algorithm (Openshaw & Openshaw 1997) and a **s**elf-**o**rganising **m**ap (SOM) developed by Kohonen (1995). A MLP uses a variant of gradient descent to propagate the errors backwards and adjust the neuron weights over many training cycles until a stopping criteria is satisfied. It is essential that the training dataset contain a representative sample of all possible situations that the MLP is likely to encounter so that the network can generalise well to unseen data and therefore be used in a predictive capacity. The SOM is a type of competitive neural network in which the neurons are usually arranged in a two-dimensional matrix. By allowing the neurons to compete against each other and discover the similarities and dissimilarities within the data, the network is able to find the organisation in the input data for itself, through unsupervised learning. In the methodology described below, a SOM is used to classify the hydrological time series data prior to training with a series of MLPs.

*Fuzzy logic*

Fuzzy logic is based on the mathematics of fuzzy set theory where the classical notion of binary membership in a set has been modified to include partial membership ranging between 0 and 1 (Zadeh 1965). Fuzzy sets, in contrast to their crisp counterparts, have ambiguous boundaries and therefore gradual transitions between defined sets. These transitions allow the uncertainty and ambiguity associated with the concepts, which they represent to be modelled directly. In fuzzy modelling, a series of overlapping fuzzy sets - also referred to as **m**em**b**ership **f**unctions (MBFs) - are associated with natural language labels such as 'hot' or 'warm', which might for example correspond to the fuzzy variable 'temperature'.

The fuzzy variables are then used in a series of IF-THEN statements. The rules together with the defining membership functions make up the knowledgebase of a fuzzy model. An inference engine, which uses an extension of multivalued logic or fuzzy logic, drives the model by execution of the rules to produce the system outputs. At present, fuzzy logic models are successfully being used in many engineering applications involving process control (Jang et al. 1994). In this methodology, fuzzy logic will be used to link the individual MLPs into a single forecasting model that also incorporates expert knowledge.

*Genetic algorithms*

A **g**enetic **a**lgorithm (GA) is a non-linear search and optimisation method inspired by the biological processes of natural selection (Holland 1975). By coding a given problem into a GA framework, these smart algorithms are able to "evolve" solutions to real world problems. Unlike other search and optimisation methods such as hill-climbing and simulated annealing, the GA searches in parallel, considering many points at once and thereby reducing the chance of converging to a local optima. A GA is used to optimise the parameters of the fuzzy logic model that combines the individual neural network solutions as a fine tuning device designed to further improve performance.

*Integrating the technologies*

In the first step of the methodology, level data at Skelton were classified into different event types or hydrograph behaviours with the use of a SOM. The reason for preclassifying the data prior to training with a MLP was based in part on the results of initial experiments in which a global MLP was used to forecast level data at Skelton (See et al. 1997).

A SOM classification also has the advantage of being more generally applicable to other stations in the catchment for the building of larger spatially integrated forecasting systems. Once created it is very fast when classifying unseen data. Experiments can be undertaken with different sizes of Kohonen map: 2x2, 3x3, 4x4 and 5x5 neurons, as well as differing lengths of previous level measurements ranging from 6 to 24 hours. The best results appeared to be produced by the 4x4 SOM using the current plus last 11 hours of level data in the event profiles. The 16 clusters produced by this map gave a good differentiation between event behaviours, which were then reclassified into 5 final event types: falling, rising, high peaks, low level flat and medium level (Linda See & Stan Openshaw 1999).

The data output was the difference between the current level reading and the value expected in 6 hours time, thereby producing a consistent 6 hour ahead forecast. The final MLPs had one hidden layer with 12 neurons. Conjugate gradient descent with a learning coefficient of 0.5 and momentum of 0.2 was used to train the MLPs for approximately 10,000 iterations, which appeared to ensure good model performance whilst avoiding overfitting. Although some experimentation with MLP architecture was undertaken, a recent river level prediction benchmarking exercise, which investigated the effect of neural network architecture on performance, indicated that there was little significant difference between one and two hidden layers or over different numbers of neurons (Abrahart & See, forthcoming).

<u>Integration of GIS – Soft Computation model:</u>

The information pertaining to tributaries includes rainfall – runoff models can be obtained at regular intervals. Based on historical time-series data the threshold levels of low, intermediate and high River levels are obtained. The neural nets are trained using the advantages derived from ANN / Fuzzy logic and Genetic algorithms. Alarming system may be adapted to avert the floods by comparing the trained neural nets and soft computations of real time rainfall – runoff models.

The rise / fall of river levels are indicators for whether the neighboring area of catchments are submerged / receded. The extent of spatial data of catchments is correlated with corresponding levels of river. The integration of spatial and soft computational data can able to forecast the area of submergence for a given river level.

**Results Expected**:

An integrated soft computing methodology for forecasting water level is helpful in predicting key alarm triggering levels used in operational forecasting. This could provide an operationally useful tool for improving the performance of flood forecasting systems. GIS can be used as a tool to draw the flood plain maps for a given river level, suitable measures can be effectively handled in the case of emergencies.

The large-scale application for the entire river stream can be run over Internet by uploading the data of river levels and rain gauge stations located at the strategic locations shall be useful for the administrators to take preventive measures to minimize the loss of life due to sudden flood situation.

**Bibliography:**

Linda See & Stan Openshaw, 1999. Using soft computing techniques to enhance flood forecasting on the River Ouse, School of Geography, University of Leeds, UK.

Abrahart, R.J. & P.E. Kneale. 1997. Exploring neural network rainfall-runoff modelling. *Sixth National Hydrological Symposium* , University of Salford, Manchester, 15-18 September 1997, 9.35-9.44.

Abraahart, R.J. & L. See. forthcoming. Neural network vs. ARMA modelling: constructing benchmark case studies of river flow prediction. Paper in preparation for special edition of *Hydrological Processes*.

Box, G.E.P & G.M. Jenkins. 1976. *Time series analysis: forecasting and control*. Oakland: Holden-Day.

Holland, J. 1975. *Adaptation in natural and artificial systems* . Cambridge: MIT Press.

Jang, J.-S.R., C.-T. Sun & E. Mizutani. 1997. *Neuro-fuzzy and soft computing*. New Jersey: Prentice Hall.

Kohonen, T. 1995. *Self-Organizing Maps*. Heidelberg: Springer-Verlag.

Minns, A.W. & M.J. Hall. 1996. Artificial neutral networks as rainfall-runoff models. *Hydrological Sciences Journal*. 41: 399-417.

Moore, R.J., D.A. Jones, K.B. Black, R.M. Austin, D.S. Carrington, M. Tinnion & A. Akhondi. 1994. RFFS and HYRAD: Integrated systems for rainfall and river flow forecasting in real-time and their application in Yorkshire. *BHS Occasional Paper*. 4.

Openshaw, S. & C. Openshaw. 1997. *Artificialintelligence in geography*. London: John Wiley & Sons.

See, L., S. Corne, M. Dougherty & S. Openshaw. 1997. Some initial experiments with neural network models of flood forecasting on the River Ouse. *Proceedings of the 2nd International Conference on GeoComputation*, University of Otago, Dunedin, New Zealand, 26-29 August 1997.

Smith, J. & R.N. Eli. 1995. Neural-network models of rainfall-runoff process. *Journal of Water Resources Planning and Management* . 121: 499-509.

Zadeh, L. 1965. Fuzzy sets. *Information and Control*. 8: 338-353.

Zadeh, L. 1994. Soft computing and fuzzy logic. *IEEE Software* . 11:48-56.