International Conference on Water Resources, Coastal and Ocean Engineering (ICWRCOE 2015)

Soft computing approach for rainfall-runoff modelling: A review

Vinay Chandwani*, Sunil Kumar Vyas, Vinay Agrawal, Gunwant Sharma

Department of Civil Engineering, Malaviya National Institute of Technology Jaipur, JLN Marg, Jaipur, Rajasthan 302017, India

Abstract

Enormous cost and manpower utilization encountered in constructing a water resource project demands a great deal of attention in devising precise Rainfall-Runoff models for its successful performance. These models are dependent on the physiographic, climatic and biotic characteristics of the basin. These factors sometimes induce either a linear, non-linear or highly complex behaviour among the rainfall and runoff parameters. The unstructured nature of Rainfall-Runoff relations has diverted the attention of researchers towards Soft Computing tools which, harnesses reasoning, intuition, consciousness and wisdom possessed by human beings. Soft Computing being a multi-disciplinary field uses a variety of statistical, probabilistic and optimization tools which complement each other to produce its three main branches viz., Neural Networks, Genetic Algorithms and Fuzzy Logic. These techniques, whether complementing each other or working on their own, are able to model complex or unknown relationships which are either nonlinear or noisy. The review paper presents an introduction to these techniques and discusses their applications in modelling Rainfall-Runoff relations which to some extent have replaced time consuming conventional mathematical techniques with wiser and time saving computing tools.

© 2015 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of organizing committee of ICWRCOE 2015

Keywords: rainfall-runoff modelling; soft computing; artificial neural networks; genetic algorithms; fuzzy logic; hybrid soft computing.

1. Introduction

The advent of computers and development in hardware and software has led to the emergence and rapid growth in the field of computational intelligence. The field of computational intelligence has brought about a revolutionary change in the development of new non-conventional techniques of data processing and simulation. Integration of intelligence by replicating human reasoning and behaviour into the computing environment enhances its capability.
to analyze the information subjected to a dynamically changing environment. Soft computing refers to a collection of computational methodologies inspired by inherent vagueness, consciousness, intuition and wisdom of human beings and real life uncertainty. In contrast to conventional computing techniques which rely on exact solutions, soft computing aims at exploiting given tolerance of imprecision, the trivial and uncertain nature of the problem to yield an approximate solution to a problem in quick time. Soft Computing being a multi-disciplinary field uses a variety of statistical, probabilistic and optimization tools which complement each other to evolve distinct computational methodologies namely, Neural Networks, Evolutionary Computation, Fuzzy Systems, Machine Learning and Probabilistic Reasoning. Among the various sub-sets of Soft Computing, Neural Networks, Genetic Algorithms and Fuzzy Logic are the major players and are commonly used for problems related to real life applications. Artificial Neural Networks (ANN) inspired by learning ability of the human brain, are able to imbibe the subtle relationships between independent and dependent variable whose interactions are unknown, non-linear or too complex to represent. Genetic Algorithms (GA) represents a stochastic search and optimization computational tool that revolves around the evolutionary theories of natural genetics and natural selection. Fuzzy Logic (FL) helps in solving real life problems which are always in some way or the other prone to ambiguity and uncertainty.

Rainfall and corresponding runoff generated are most important hydrological processes which depend on the local physiographic, climatic and biotic factors. Rainfall over a catchment is the fall of moisture from the atmosphere to the earth's surface in liquid (rainfall) or frozen form (snow, hail, sleet, freezing rain). Runoff of a catchment area in any specified period is the total quantity of water draining into a stream or into a reservoir in that period, which can be expressed as millimeters of water over a catchment or the total volume of water in cubic meter or hectare meter. Water available in a particular time period at a particular location in a particular area in a river basin can be determined with the help of the rainfall runoff regression model i.e. hydrological model. The main components of a hydrological model are precipitation, evaporation, transpiration, interception, infiltration, stream flow, variability in time and space. When the rainfall intensity exceeds the infiltration rate than the runoff starts immediately. Physiography of the local area largely affects the rainfall and the generated runoff from the catchment. The infiltration capacity of the land depends on porosity of a soil which also measures the resistance of water flow into the deeper layers. The top soil having more than 20% clay or loam content forms a cap with lower infiltration capacities and indirectly produces more runoff as compared to the loose sandy soils. The runoff efficiency, i.e. volume of runoff per unit area increases with the decreasing size of the catchment. Vegetation cover density, thick layer of mulch of leaves or grasses, lower flow velocity in flat area provides longer time concentration/ absorption hence low runoff generation. Undulating land has more runoff than flat due to getting the additional energy received from the slope and less time for infiltration. The greater drainage density gives more runoff. The ratio of maximum rainfall at a point to the mean rainfall in the catchment called distribution coefficient, also affects the runoff generated. Direction of the prevailing wind increases or decreases the flowing velocity and indirectly affects the runoff. The average basin characteristics, i.e. temperature, wind velocity, relative humidity and annual rainfall also affect the rainfall runoff relation in the basin.

The optimum design, operation, maintenance and use of existing or proposed water resources projects in a particular River Basin requires detailed knowledge of the rainfall and the corresponding runoff generated in a particular time interval/period. Besides planning and development of water resources projects, the rainfall-runoff models helps in formulating flood control measures, drought management, optimization of reservoir operation, water supply, etc. In addition to its multifaceted and wide applicability in judging the overall water balance scenario of an area under study, the functional relationships between rainfall and runoff is extremely complex due its dependence on the factors discussed in the preceding paragraph. In such cases wherein the interactions among the variables are complex, the conventional mathematical techniques in the form of regression equations do not provide a perfect representation of the rainfall-runoff phenomenon. Soft computing tools offer a simplified approach over conventional hard computing in dealing with the real life phenomenon associated with imprecise, noisy, complex and ambiguous nature of information. The review paper is an attempt to provide a comprehensive introduction to three major fields of Soft Computing viz., Artificial Neural Networks (ANN), Genetic Algorithms (GA) and Fuzzy Logic (FL) and their applications in modelling complex rainfall-runoff relationships.

The review paper has been divided into sections. A brief introduction to three major soft computing techniques viz., Artificial Neural Networks (ANN), Genetic Algorithms (GA) and Fuzzy Logic (FL) is given in Section 2. Section 3 deals with various applications of soft computing applied to rainfall-runoff modelling. This section has been sub-divided to give a detailed review about the application of individual and hybrid soft computing techniques
in rainfall-runoff modelling. The conclusions of the literature review have been briefly dealt in Section 4.

2. Soft computing techniques

The umbrella of Soft Computing techniques is widening its scope with every passing year due to increasing demand for time saving and fault tolerant computational tools. In contrast to analytical methods, Soft Computing methodologies mimic consciousness and cognition in several important aspects: they learn from experience; they can universalize into domains where direct experience is absent; and, through parallel computer architectures that simulate biological processes, they can perform the mapping from inputs to the outputs faster than inherently serial analytical representations (Chaturvedi, 2008). The collection of techniques under Soft Computing render low cost solutions to imprecisely formulated problems and attempt to inculcate the behaviour and learning ability of human beings into computers. The following sub-sections deal with an introduction to the forerunner techniques of Soft Computing viz., Artificial Neural Networks, Genetic Algorithms and Fuzzy Logic.

2.1. Artificial neural networks

Artificial Neural Networks (ANN) represents the learning algorithms and architectures inspired by the working and structure of the human brain. Although they represent a much simplified version of the human brain, yet these computational models inspired by biological neural network has provided new directions to solve problems arising in natural tasks. Haykin (2009) has described a neural network as a massively parallel distributed processor made up of simple processing units, which has a natural propensity for storing experiential knowledge and making it available for use. They present an information learning paradigm comprising of processing elements called “artificial neurons” or “neurons”, which are arranged in layers. Individually, the neurons perform trivial functions, but collectively, in the form of a network, they are capable of solving complicated problems (Flood and Kartam, 1994). The architecture of ANN comprises of three basic identities viz., the weighted connections between the neurons, the learning algorithm for updating the weights and the activation function acting on the weighted sum of input signal fed into the neuron. ANNs are well suited for problems whose solutions require knowledge that is difficult to specify but for which there are enough data or observations (Zhang et al., 1998). ANN’s ability to offer model free, parallel processing of noisy data and adaptability to changing instances of the problem, give it an edge over conventional data processing techniques. The multi-disciplinary applications of ANN is attributed to its ability of deriving complex, non-linear and unknown relationships among independent and dependent variables through a learning process, thereby working as a universal function approximator and therefore has been a field of interest for predicting behaviour of engineering and natural systems.

2.2. Genetic algorithms

Genetic Algorithms (GA) is a computerized search and optimization tool whose methodology is inspired by the “Survival of Fittest” heuristic. One important feature in this heuristic is its durability and adaptability because it provides a flexible balance between effectiveness and necessary characteristics for survival in different environments and conditions (Johari et al., 2011). GA’s outperform the efficiency of conventional optimization techniques in searching non-linear and non-continuous spaces, which are characterized by abstract or poorly understood expert knowledge (Mellit et al., 2010). In contrast to traditional optimization techniques which start from a single point in the solution search space, GA initializes its search process using a group of solutions represented as the population of chromosomes at a time and navigates the search space using its three evolutionary operator’s viz., selection, crossover and mutation to reach the global optimum. Based on the fitness of each chromosome, the selection operator filters the fitter chromosomes from a pool of chromosome population and allows them to represent themselves in the next generation. The crossover works on the pair of chromosomes and selects a crossover point along the length of the chromosome. The genetic characteristics are then swapped across the pair of chromosomes following the cross site, in the hope of producing better off springs. In contrast to crossover that performs the exploitation of the current population, the mutation operator explores the entire population and tries to bring forth any possibility of improvement in the search that has largely converged. With every successive
generation, the GA operators help in improving the quality of the solution, thereby increasing the probability of finding global optima.

2.3. Fuzzy logic

Conceptualized by Lotif A. Zadeh in 1965, Fuzzy Logic (FL) reflects the computational methodology of thinking and solving problems inherent in human beings. FL approximates the reasoning and decision making ability of human beings in areas which are prone to imprecise, inexact and vague knowledge and therefore is a quantitative method of describing observations. FL has an ability to link multiple inputs to one output and does not require the normality, linearity and homoscedasticity as needed by traditional methods such as regression and principal component analysis (Ozger, 2011). This approach provides a simple method to draw definite conclusions from vague, ambiguous, or imprecise information (Klir and Foger, 1988). Built on linguistic principles, modelling using fuzzy logic involves fuzzification of variables, defining rule base, selecting inference method and finally applying defuzzification method for predicting responses. The fuzzification of variables is accomplished by defining the membership function, which represents the degree of belongingness of the element to the set. Its mathematical formulation is synonymous to fuzzy set theory dealing with classes without crisp boundaries. In contrast to traditional logic that deals with either “True” or “False” logic or “0” or “1” logic, the FL using the fuzzy set theory is able to deal with many valued logic that has prevailed due to uncertainty and vagueness inherent in real life phenomenon, allowing generalization of a characteristic function to a membership function. Fuzzy Logic theory and fuzzy set theory provide an excellent means for representing imprecision and uncertainty in the decision-making process and for defining the reasoning in such processes (Zadeh, 1983).

3. Applications of soft computing for rainfall-runoff modelling

3.1. Artificial neural network applications

ANN has been a preferred choice among the various soft computing techniques for modelling complex rainfall-runoff phenomenon. Tokar and Johnson (1999) employed neural networks for modelling daily runoff as function of daily precipitation, temperature and snowmelt. The study showed that ANN model provided better prediction accuracy than compared with regression and conceptual models. The study also proved that for obtaining sufficiently reliable predictions, the data should representative of the characteristics of the watershed and the climatic conditions. Rajurkar et al. (2002) used the daily rainfall and runoff data for Narmada catchment area for development of a model for transforming rainfall into runoff for a large catchment area. A linear multi-point single output model (MISO) has been used for deriving the response function relating rainfall to runoff. The predicted runoff from the MISO model is then refined by a multi-layer feed forward neural network. The study showed that integration of MISO model with ANN led to higher prediction accuracy than linear and non-linear MISO models. Wilby et al. (2003) used precipitation, evaporation and discharge data for developing conceptual and neural network rainfall-runoff model. Three different experiments were conducted with progressively decreasing amount of information, to determine the extent to which neural networks can safely imbibe the hydrological processes. The study showed that, a neural network with seven inputs and three hidden nodes were able to grasp the behaviour of the conceptual model. The correlation analysis showed that the two hidden nodes corresponded to the base flow and quick flow components whereas, the third hidden unit represented the seasonal variations in the soil moisture deficit.

Riad et al. (2004) used rainfall-runoff data at the Aghbalou station for developing rainfall-runoff model using ANN. The input vector of the ANN comprised of rainfall and runoff for the preceding seven days as well as the rainfall expected for the day. The expected runoff for the day formed the output vector for ANN. The overall rainfall-runoff data of last seven years were split into two parts. The rainfall-runoff data for six years comprised the training data and the data for one year was used for testing the trained ANN. The performance of ANN was compared with multiple linear regression (MLR) model. The values predicted by ANN were found to be in close agreement with the observed data. The study showed that ANN has a good ability of modelling complex hydrological processes. Jeong and Kim (2005) used two neural network models viz., single neural network (SNN) and ensemble neural network (ENN) for providing better rainfall-runoff simulation. The study showed that ENN
provided the least root mean square error and therefore performed efficiently than SNN. By comparing, ENN with existing rainfall-runoff simulation model TANK based on some probabilistic accuracy measures, the ENN outperformed the TANK model in most of test cases. The applicability of neural networks for modelling non-linear rainfall-runoff relationships was explored by Abrahart and See (2007). The neural network modelling provided promising results for producing non-linear transformations and can be helpful in case of scarce or difficult to obtain data sets. Kalteh (2008) developed rainfall-runoff model using ANN and compared that with the Neural Interpretation Diagram, Garson’s algorithm and randomization approach. The study showed that ANN not only efficiently learns the complex processes but also, provides an understanding about the complex relationships within the processes.

Solaimani (2009) used three different training algorithms for developing ANN models for forecasting rainfall. Efficiencies of gradient descent, conjugate gradient and Lavenberg-Marquardt training algorithms were compared. Monthly hydrometric and climatic data were used for development of the ANN model. The study showed that by amalgamating computational efficiency with input parameters that describe the hydro-climatology variables, an improvement in ANN prediction can be achieved. Machado et al. (2011) prepared three ANN models prepared on the basis of monthly rainfall runoff and compared them with the conceptual model IPHMEN at the monthly time scale. Back-propagation algorithm was used for training the neural network. The trained ANN was compared with IPHMEN conceptual model during calibration and validation phases. IPHMEN model provided inconsistent results with large dispersion of computed flows. The study showed that the trained ANN has good potential for accurately predicting the observed flows. Chen et al. (2013) developed the rainfall-runoff model for typhoon using artificial neural networks. A three-layered neural network was constructed having hourly rainfall data of three rainfall stations as inputs for modelling hourly flows of a station. The study showed that regression analysis is suitable for small variation in data. But for large variation in data, ANN provided a promising methodology. The ANN proved to be flexible and easy to implement computational tool for modelling the complex hydrological processes. Phuketphim et al. (2014) used three different approaches viz., Garson’s algorithm, neural interpretation diagram and sensitivity analysis, to understand the contribution of input variables and rule extraction approaches to neural network modelling of rainfall-runoff relationships. The study showed that the complexity prevalent in the rainfall-runoff models can be significantly reduced by eliminating the least significant input variables.

3.2. Genetic algorithms applications

Genetic algorithms stochastic search and optimization has been harnessed for calibrating the rainfall-runoff models. Franchini and Galeati (1997) used GA for calibration of conceptual rainfall-runoff models. The parameters which cause minimum deviation were deduced using GA. Ndiritu and Daniell (2001) presented an improved GA based on automatic search space shifting to achieve hill-climbing, automatic search space reduction to effect time-tuning, and the use of independent subpopulation search coupled with shuffling to deal with the occurrence of multiple regions of attraction for rainfall-runoff model calibration. Cheng et al. (2005) presented a parallel GA methodology for calibrating rainfall-runoff models influenced by a large number of parameters. The problem was partitioned into number of problems and run on number of computers through parallel GA (PGA). The study showed that PGA methodology significantly improves the time taken for optimization of solution and produces a stable solution. An improved GA to tackle the problem of inaccurate coding and low precision was presented by Li et al. (2009). An amendatory objective function was used in the study to balance the hydrological elements present in the constraint relations, for accurate prediction of hydrological flow process. Khazaei et al. (2014) presented an automatic calibration tool to calibrate the ARNO conceptual rainfall-runoff model. For this a simple genetic algorithm (SGA) was employed. The methodology provided sufficiently accurate predictions during calibration and validation and therefore can be applied for continuous rainfall-runoff simulation.

3.3. Fuzzy logic applications

Based on the linguistic variables, FL presents an easier way of developing rainfall-runoff model governed by highly complex hydrological processes. Hundecha et al. (2001) employed fuzzy rule based routines for generating runoff from precipitation. The methodology was applied to conceptual, modular, and semi-distributed model to
prove the effectiveness of FL. Nayak et al. (2005) employed fuzzy computing for real time flood forecasting. The study showed that fuzzy models have the ability to simulate the unknown relationships between the hydrological data. Combinations of variables were presented to the fuzzy model to study the sensitivity of prediction and assess the agreement between the precipitation, upstream runoff and total watershed runoff. Casper et al. (2007) used FL for rainfall-runoff modelling using soil moisture measurements. In the study fuzzy rule-based system (FRBS) using the Takagi-Sugeno-Kang approach has been developed using soil moisture and rainfall as input variables to predict the actual discharge at the catchment outlet. The study showed that the measurements of soil moisture at representative locations could be used as representation for the actual system state, allowing for an entirely data-driven prediction of the runoff response using rainfall. Wang and Altunkaynak (2012) compared the storm water management model (SWMM) commonly used for rainfall-runoff simulation with FL. The study showed that FL model outperformed the SWMM for large rainfall. However, the SWMM can produce the time varying hydrograph whereas fuzzy logic is subject to limitation of the methodology and is unable to generate such an output.

3.4. Hybrid soft computing applications

In the past few years there has been a tremendous inclination towards hybrid soft computing techniques for dealing with problems encountered in real life. The reason for this growth is attributed to the complementing nature of the distinct soft computing approaches. The hybridization of the techniques covers up the limitations of the individual ones and leads to development of robust computational methodologies. Srinivasulu and Jain (2006) compared the back-propagation neural network ANN trained using real coded genetic algorithm (RGA) and self-organizing map (SOM) classified input-output rainfall-runoff data. The study showed that ANN trained using RGA provided better generalization of the complex and non-linear rainfall-runoff process. Nasseri et al. (2008) presented hybridization of ANN with GA for short term rainfall forecasting, by harnessing the global search ability of GA for selection of suitable input parameters and optimal neural network architecture. The study showed that hybrid ANN outperformed the prediction ability of multi-layer perceptron (MLP) neural network.

Talei et al. (2010) used Adaptive Network-based Fuzzy Inference System (ANFIS) for event-based rainfall-runoff modelling. The results of the ANFIS were compared with an established physical-based model. The study showed that ANFIS is comparable to the physical model and is found to give a better peak flow estimation compared to the physical model. Dorum et al. (2010) studied the rainfall-runoff data using ANN and ANIFS methods. A multi regression model was also used to compare the results obtained from ANN and ANIFS models with traditional methods. The study showed that ANN and ANIFS models can be used in determination of rainfall-runoff relationships of Susurluk Basin except peak situations. Asadi et al. (2013) used GA for evolving the weights of the neural network used for rainfall-runoff process modelling. The data were preprocessed by data transformation, input variable selection and data clustering for improving the prediction accuracy of the model. The study showed that by adopting this methodology, faster training, high degree of accuracy and good adaptation of nonlinear functional relationships between rainfall and runoff is achieved.

4. Conclusions

Rainfall-runoff interactions and their correct assessment form an integral part of any hydrological study. The complex, ambiguous and non-linear factors affecting the rainfall-runoff relations, makes its mathematical modelling a difficult task. Such situations demand for nature inspired computational tools to deal with the real life phenomenon which are always subjected to imprecise, ambiguous and noisy information. Soft computing through amalgamation of probabilistic, statistical and optimization techniques into the computing environment, has presented a suitable replacement substitute to traditional mathematical techniques. Fusion of these methodologies has given a new dimension to computing, in which human behaviour resembling capability of reasoning, intuition, consciousness and wisdom can be combined through software programming. Though these computational tools do not offer exact answers, yet they provide a sound decision oriented solutions to the problems influenced by vague and noisy information.

The review paper has presented a brief introduction to three major soft computing techniques viz., Artificial Neural Networks, Genetic Algorithms and Fuzzy Logic and their application in rainfall-runoff modelling. The
review has shown that by applying these techniques, the complex analytical method can be avoided to some extent. ANN learning from experiential or historical data can form a backbone for complex, unknown and difficult to describe functional relationships between rainfall and runoff. GA through its stochastic search ability can be harnessed for calibration of these rainfall-runoff models. FL emulation of human reasoning and decision making ability can be exploited for modelling problems governed by inexact, vague and imprecise information. The paper has also significantly attended the hybrid soft computing techniques. The complementing nature of the distinct soft computing approaches has given the benefit of deriving the best from these techniques to the user. Hybridization is a technique of enriching the original procedures and covering up the limitations of the individual methodologies and thus opens up the avenues for solving new problems. It is hoped that the conjunctive use of soft computing will surely be step beyond the cognitive skills inherent in human beings.

References


