International Journal of Advanced Biochemistry Research

ISSN Print: 2617-4693
ISSN Online: 2617-4707
IJABR 2024; 8(4): 108-112
www.biochemjournal.com
Received: 26-01-2024
Accepted: 30-02-2024

Anusmita Bhowmik
School of Agriculture, Swami Vivekananda University, Barrackpore, West Bengal, India

Ankana Moulik
School of Agriculture, Swami Vivekananda University, Barrackpore, West Bengal, India

Tanmoy Majhi
Assistant Professor, Swami Vivekananda University, Barrackpore, West Bengal, India

Corresponding Author:
Tanmoy Majhi
Assistant Professor, Swami Vivekananda University, Barrackpore, West Bengal, India

Comparative study of artificial neural network and hydrological models towards runoff estimation

Anusmita Bhowmik, Ankana Moulik and Tanmoy Majhi

DOI: https://doi.org/10.33545/26174693.2024.v8.i4b.931

Abstract
Runoff is a vital hydrological factor, governing water flow into systems and returning excess precipitation to the seas. Water resource managers use model-derived runoff data to comprehend, regulate, and monitor water resources. The process of obtaining this data is arduous. The study provides a framework for grasping the model-specific building blocks, their influence on closure effects, and parameter calibration's simplicity and physical meaning. It surpasses mere flow rate replication, delving into deeper model capabilities in the process of integrating model outcomes, three distinct approaches are considered: the simple average technique (SAM), the weighted average method (WAM), and the neural network method (NNM). In forecasting stream flow utilizing seven parameters, the SIMHYD model examines daily rainfall and areal potential evapotranspiration data. The foremost and pivotal component of the ARNO model entails delineating the soil moisture equilibrium, while the subsequent aspect involves depicting the course of runoff transfer towards basin outflow. The degree of calibration achieved in a conceptual rainfall-runoff (CRR) model governs its potential efficacy in practical implementation. In spite of the prevalent application of conceptual rainfall-runoff (CRR) models, research findings indicate that achieving distinct optimal parameter values through automatic calibration techniques is often challenging. The adoption of Artificial Neural Networks (ANNs) has become increasingly common in the analysis of hydrological and water resource complexities. Furthermore, the Probability Distributed Model (PDM), a compendium of model functions, was developed as a lumped rainfall-runoff model capable of elucidating a spectrum of hydrological behaviours at the catchment scale. As a result, this paper offers insights into both the merits and limitations of various rainfall-runoff models for the readership's enrichment.

Keywords: Runoff, artificial neural networks, probability distributed model, ARNO model, weighted average method

Introduction
Rainfall - runoff modelling is a part of hydrological studies which is used to understand watershed yields and responses, predict water availability, and anticipate changes over time. These models are frequently used in real-time river flood forecasting systems. In such systems, a rainfall-runoff model is chosen from among a number of competing alternative models based on factors such as accuracy, user familiarity, and convenience of use, catchment type, and accessible data. In order to represent the hydrologic cycle, these models are linked together to process elements that describe physical concepts under the assumption that the model parameters would also have physical meaning and could thus be assigned values independently of observed data. The rationale behind this is that each model's output captures some key attributes of the information that is currently known about the process being simulated, offering a source of data that may be distinct from that of other models.

The simple average technique (SAM), the weighted average method (WAM), and the neural network method (NNM) are tested as ways to combine the predicted outputs of various models. Many authors have utilized the first two of these three techniques frequently. Makridakis and Winkler *et al.* 1983) [28]. The actual observed outputs of SAM & WAM are functionally related to the estimated outputs of the separate models via linear regression relationships. In this work, the application of SIMHYD lumped conceptual daily rainfall-runoff model is also discussed where daily precipitation and potential evapotranspiration are used as input data to simulate daily runoff (Chiew and Siriwardena *et al.* 2005) [5]. This study also provides a detailed description of the ARNO model, a semi-distributed conceptual rainfall-runoff model that is currently widely used in research on land-surface-atmosphere processes and as a practical flood forecasting tool on several catchments throughout the globe. The basic physical principles underlying the hydrologic cycle are approximated in the frameworks of conceptual rainfall-runoff (CRR) models. The soil moisture accounting phase of the hydrologic cycle is often modelled by CRR models as a collection of interconnected subsystems, each of which represents a specific step in the processing of a hydrologic event. The Artificial Neural Models (ANNs) models are very much effective in making forecasts of rainfall as well as for runoff parameters. The findings of ANN aids in decision-making of the planning and management of water resources, also aiding the urban designers in taking the required precautions in facing the prospect of any adverse and detrimental forecasts. The Probability Distribution Model (PDM) is also a generic conceptual rainfall runoff model that translates the rainfall & potential evaporation parameters to flow at the catchment outlet (Moore *et al*. 2007) [9]. The majority of these rainfall-runoff models can be effectively adjusted to replicate the measured rainfall - runoffs. Applications for rainfall-runoff models range from the assessment of catchment water output to the estimation of the effects of land use and climate change on runoff characteristics. The theory is that each model's output captures some key elements of the information that is currently accessible about the process being simulated, offering a source of knowledge that may be distinct from that of other models.

SAM (Simple Average Method): The easiest way to combine the results of many independent models is to use the simple approach. A combined estimate of the discharge, Qci of the ith time period using the Simple Average Method predicts the discharges of N rainfall-runoff models. This method can yield forecasts that are superior to those of the individual models (Makridakis *et al*. 1982) [30] and its accuracy mostly relates on the number of models related to the real forecasting capability of the particular models that are included in the basic average system (Makridakis and Winkler *et al*. 1983) [28]. Using simple average method, this strategy has demonstrated that combining predictions also lowers the variability of forecasting mistakes and, thus, the risk related to the selection of forecasting methods. (Makridakis and Winkler *et al*. 1983) [28]. The theory is that each model's output captures some key elements of the information that is currently accessible about the process being simulated, offering a source of knowledge that may be distinct from that of other models.

WAM (Weighted Average Method): The use of a weighted average would be taken into consideration when some of the individual models chosen for combination appears to be consistently more accurate than others, in which case the basic average strategy for combination can be highly ineffective (Armstrong *et al*. 1989) [1]. The fundamental downside of WAM consists of the fact that it can be susceptible from multicollinearity, which may contribute to unstable weight estimations (Winkler *et al*. 1989) [27] lowering the significant benefits achieved by merging the numerous model outputs.
This method is particularly relevant in the current hydrological setting, when the amount of multicollinearity grows with the forecasting capacity of the different models, or even when the outputs of the numerous models utilized are highly similar but are not necessarily excellent. The SAM and WAM makes the assumption that the true measured outputs are functionally connected to the estimated outputs of the various models through linear regression relationships.

NNM (Neural Network Method): The neural network approach is an alternate to the simple average (SAM) and weighted average methods (WAM) for mixing the outputs of multiple models, and it may be used to determine whether a more sophisticated connection is required for such combinations. It is a non-parametric strategy in the sense that the relationship's precise mathematical form is unknown. This third technique (NNM) is suggested to assess if a more complicated connection is required to combine the predicted outputs of the various models. The employed multi-layer feed forward neural network used in this study has an input layer, an output layer, and just one 'hidden' layer between the input and output layers. A layer is often said to be a collection of neurons that have the same pattern of connection paths with neurons from neighbouring layers. As its name suggests, this approach makes use of the neural network model structure, a very potent computational technique for simulating complex non-linear relationships, especially when the explicit form of the relationship between the variables is unknown. The NNM has the capacity to combine data from disparate sources in terms of their geographical location. The Weighted Average Method (WAM) & the Neural Network Method (NNM) are considered to be more resilient and efficient methods of combination used in rainfall-runoff models.

SIMHYD Model: A conceptually aggregated daily rainfall runoff model is SIMHYD. Daily precipitation and potential evapotranspiration are used as input data to model daily surface runoff and base flow. There are 7 parameters of this model named as (Chiew and Siriwardena *et al*. 2005) [5].

- **INSC:** Interception Store Capacity (mm).
- **COEFF:** Maximum Infiltration Loss (mm).
- **SQ:** Infiltration Loss Exponent.
- **SMSC:** Soil moisture store capacity (mm).
- **SUB:** Constant of proportionality in interflow equation
- **CRAK:** Constant of proportionality in groundwater recharge equation.
- **K:** Base flow linear recession parameter.

The majority of catchments do not require COEFF and SQ optimization since SIMHYD generates little to no infiltration surplus runoff apart from tropical catchments (Chiew and Siriwardena *et al*. 2005) [5]. The outcomes suggest that the SIMHYD model can be trained sufficiently to replicate the runoffs that were observed from rainfall-runoff data (Chiew and Siriwardena *et al*. 2005) [5]. SIMHYD operates on an everyday time phase, but is adjusted against monthly runoff and thus eliminates the requirement for routing and also the accompanying mistakes (Chiew and Siriwardena *et al*. 2005) [5].

ARNO Model: The soil moisture module of the ARNO model has been widely used in hydrological practice; in particular, it has become the core of a real-time operational flood forecasting system created on behalf of the Commission of the European Communities (the European Flood Forecasting Operational Real-Time System

(EFFORTS), which is already operational on several rivers in many countries, including the PO, the Arno, the Tiber, and the Fuchun in China. A number of meteorologists (Rowntree and Lean *et al*. 1994, Polcher *et al*. 1996) [21, 20] and others have recently tested it in collaborative meteorological hydrological experiments, such as the Spatial Variability of Land Surface Processes (SLAPS) project.It was effectively incorporated into the Hamburg climate model due to the formulation's simplicity (Dtimenil and Todini *et al*. 1992) [8]. The ARNO model incorporates groundwater, evapotranspiration, and snowmelt components in addition to the essential soil moisture component, all of which are sparse descriptions consistent with the general philosophy. Furthermore, the runoff contributions of the area units under consideration must be moved downstream and aggregated as they move along the slopes and the drainage network until an acceptable description of the so-called "runoff production function" has been attained. (Todini *et al*. 1996) [25].

CRR (Conceptual Rainfall - Runoff) Model: Many different conceptual rainfall-runoff (CRR) models have been created over the last few decades. Other kinds of watershed hydrological models, like physically based models, are frequently preferred to CRR models. This is due to the fact that, in many real-world situations, we may just need runoff process estimates from rainfall at the watershed outflow or a specific point, and they are also simpler to compute. The model structure and parameters of CRR models, which simulate watershed hydrological processes using techniques from mathematical physics, can be represented. In other words, the parameters directly affect the model's and its forecasts' accuracy in addition to the structure's logic. Since they have a physical meaning, the majority of CRR model parameters may theoretically be calculated either directly or indirectly using measurements or a physical approach. Calibration is the process of subjecting model parameters to past data on system responses. (Zhang *et al*. 2015) [29].

In this study, we concentrated on three frequently used daily time-step CRR models that estimate catchment-outlet runoff from catchment-averaged rainfall (P) and potential evapotranspiration (PET) data. These models all describe the catchments as being spatially lumped. The models' structural representation varies, which has an impact on their input-state-output behaviour and may result in varying degrees of performance robustness. (Guo and Zheng *et al*. 2020) [10].

ANN (Artificial Neural Network) Model: A flexible mathematical structure called an artificial neural network (ANN) is able to recognize intricate nonlinear correlations between input and output data sets. Particularly in situations when it is challenging to define the properties of the processes using physical equations, ANN models have been found to be effective and beneficial. This paper highlights the potential of such models for simulating the nonlinear hydrologic behaviour of watersheds and introduces a new method (Referred to as linear least squares simplex, or LLSSIM) for finding the structure and parameters of three-layer feed forward ANN models (Hsu *et al*. 1995) [11]. It is demonstrated that the nonlinear ANN model approach represents the rainfall-runoff connection of the medium-sized Leaf River basin more accurately than the linear ARMAX i.e. autoregressive moving average with exogenous inputs time series (Hsu *et al*. 1995) [11]. The ANN technique described here is by no means a replacement for conceptual watershed modeling because it does not offer models with physically realistic components and characteristics.

In circumstances when modelling of the internal structure of the watershed is not necessary, the ANN approach does offer a viable and efficient substitute to the ARMAX time series approach for constructing input-output simulation and forecasting models (Hsu *et al*. 1995) [11]. In hydrological applications, the feed-forward multilayer perceptron (MLP) is the most widely employed ANN. The format of a MLP with three layers is taken. It includes three layers: an input layer, a concealed layer, and the top layer and a layer for output; amount of neurons defined in the input and output layers according to the quantity of input and output variables relating to the system under study respectively. In this study, a neuron in the hidden layer or layers six neurons in a single buried layer, is usually determined through trial and error procedure. According to the theory, each layer's neurons are linked to the weighted neurons of the next layer. To get the best values for these ANNs need to be trained to handle connection weights. The outputs produced by the network are compared to the intended output values of the system that is being examined to calculate error, first the propagation of computed error, the weights of the connections and the network updated. This process is known as training. Technique is continued until a satisfactory convergence level is attained. In this research, the neural network must prevent instability in order to, twenty times to train the network, and by Average the results of all final results was acquired. Information about ANN structures, applications for training algorithms in ASCE provides a full discussion on hydrology. (Maier and Dandy *et al*. 2000, ASCE task committee *et al*. 2000a, ASCE task committee *et al*. 2000b, Dawson and Wilby *et al*. 2001, Kalteh and Mohammad *et al*. 2008) [15, 2, 3, 6, 12].

PDM (Probability Distribution Model): The hydrological probability-distributed model (PDM), which is used extensively throughout the world, has also its applicability examined through a lengthy period of monitoring years, rainfall-discharge interactions for all gauging stations erected on impassable watercourses were modelled. 1456 years in total were modelled. Peak flow and volume characteristics of modelled series are contrasted with observations. With the PDM, accurate discharge values can be calculated based on the rather long time series. Water volumes and peak characteristics closely match the values that have been measured. A single-parameter approach, a parameter set approach, and a technique with predetermined cluster zones based on hydrological flow characteristics were all taken into consideration. The parameter of the single-parameter approach with less specific geographical information was provided by the set approach and their combination than by clustering on hydrological variables (Cabus *et al*. 2008) [4]. The Probability Distributed Moisture (PDM) model is a conceptual model that uses two parallel linear reservoirs in this application and most others for the routing component and a distribution of soil moisture storage capacities for soil moisture accounting (Moore *et al*. 1985, Moore *et al*. 1999, Moore *et al*. 2007) [17-19]. The simplicity of this model structure should increase the likelihood of successful regionalization via parameter regression against CDs (Lamb *et al*. 2000, Wheater 2006, Lee *et al*. 2005) [13, 16, 14]. It was chosen because it and variations of it are typically found to work at least as well as other simple models in the UK. The Pareto distribution with the following function is believed to best characterise the soil moisture storage capacity, C (mm).

$$F(C) = 1 - (1 - C/Cma).^{b}$$

Where C is the catchment's total storage capacity, Cmax is the highest capacity possible at any given site, and the parameter b regulates the catchment's spatial variability of storage capacity. The extra soil moisture calculated at each time step is equal to the effective rainfall. The potential evaporation is multiplied by the relative saturation of the catchment to determine the evaporation rate.

Declaration of competing interest

The authors declare no conflict of interest.

Credit authorship contribution statement

- **Anusmita Bhowmik:** Data curation, Writing - review & editing, Writing - original draft.
- **Ankana Moulik:** Data curation, Writing - review & editing, Writing - original draft.
- **Tanmoy Majhi:** Conceptualization, Methodology, review & editing.

Conclusion

SAM or simple average method is used in the simplest way to aggregate the output of numerous distinct models. The WAM is a deterministic hydrologic watershed model that uses a geographic information system (GIS) to depict the intricate interactions between water quantity and quality in the terrestrial part of the hydrologic cycle. As the name suggests, the NNM method makes use of the neural network model structure, a very potent computational methodology for modelling complex non-linear interactions, especially when the explicit form of the relationship between the variables involved is unclear. A conceptual rainfall-runoff model called SIMHYD uses data on daily rainfall and areal potential evapotranspiration to estimate daily streamflow. Three stores are included in the model; groundwater, soil moisture, and intercept loss. Seven variables are used in the model. The soil moisture balance is represented by the first and most significant component of the ARNO model, while the transfer of runoff to the basin outflow is represented by the second. CRR models for conceptual rainfall-runoff are frequently used to simulate and predict historical streamflow. These model structures are frequently established through calibration to a subset of the available data, with the reliability of the model being assessed using a separate subset of data. The basis of ANN is a self-adaptive mechanism that allows the model to learn from past data, record functional relationships between data, and make predictions based on current data. One of the most important criteria for managing the water resources is the accuracy of rainfall forecasts. The PDM, or Probability-Distributed Model, translates the time-series of rainfall and evaporation to river flow at the catchment outlet using a generic conceptual rainfall-runoff model. Studying all the models, we come to the conclusion that the ANN Model is the most efficient method of determining the rainfall-runoff estimation because of its self-adaptive ability. The developed ANN models could determine the relationship between the input and output data sets using its efficient neural networks.

References

1. Armstrong JS. Combining forecasts: The end of the beginning or the beginning of the end? Int. J Forecast. 1989;5(4):585-588.
2. ASCE Task Committee on Application of Artificial Neural Networks in Hydrology. Artificial neural networks in hydrology. I: preliminary concepts. J Hydrol. Eng. ASCE. 2000;5(2):115-123.
3. ASCE Task Committee on Application of Artificial Neural Networks in Hydrology. Artificial neural networks in hydrology. II: hydrologic applications. J Hydrol Eng ASCE. 2000;5(2):124-137.
4. Cabus P. River flow prediction through rainfall-runoff modelling with a probability-distributed model (PDM) in Flanders, Belgium. Agric. Water Manag. 2008;95(7):859-868.
5. Chiew FHS, Siriwardena L. Estimation of SIMHYD parameter values for application in ungauged catchments. In: MODSIM 2005 International Congress on Modelling and Simulation, Modelling and Simulation Society of Australia and New Zealand, Melbourne; c2005. p. 2883-2889.
6. Dawson CW, Wilby RL. Hydrological modelling using artificial neural networks. Prog. Phys. Geog. 2001;25(1):80-108.
7. Dooge JCI, Bruen M, Dowley A. Spatial variability of land surface processes. Final Report of CEC Research Contract PL 890016 EPOCH; c1994.
8. Dümenil L, Todini E. A rainfall-runoff scheme for use in the Hamburg climate model. In: Advances in theoretical hydrology. Elsevier; c1992. p. 129-157.
9. Franchini M, Pacciani M. Comparative analysis of several conceptual rainfall-runoff models. J Hydrol. 1991;122(1-4):161-219.
10. Guo D, Zheng F, Gupta H, Maier HR. On the robustness of conceptual rainfall-runoff models to calibration and evaluation data set splits selection: A large sample investigation. Water Resour. Res. 2020;56(3):e2019WR026752.
11. Hsu KL, Gupta HV, Sorooshian S. Artificial neural network modeling of the rainfall-runoff process. Water Resour Res. 1995;31(10):2517-2530.
12. Kalteh AM. Rainfall-runoff modelling using artificial neural networks (ANNs): modelling and understanding. Casp. J Environ Sci. 2008;6:53-58.
13. Lamb R, Crewett J, Calver A. Relating hydrological model parameters and catchment properties to estimate flood frequencies from simulated river flows. In: BHS 7th National Hydrology Symposium, Newcastle-upon-Tyne. 2000;7:357-364.
14. Lee H, McIntyre N, Wheater H, Young A. Selection of conceptual models for regionalisation of the rainfall-runoff relationship. J Hydrol. 2005;312(1-4):125-147.
15. Maier HR, Dandy GC. Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. Environ Model Softw. 2000;15:101-124.
16. Mohseni U, Muskula SB. Rainfall-runoff modeling using artificial neural network-a case study of purna

sub-catchment of Upper Tapi. Basin, India. Environ Sci. Prec. 2023;25(1):1.
17. Moore RJ. The probability-distributed principle and runoff production at point and basin scales. Hydrol. Sci. J. 1985;30(2):273-297.
18. Moore RJ. Real-time flood forecasting systems: Perspectives and prospects. In: Floods and landslides: Integrated risk assessment; c1999. p. 147-189.
19. Moore RJ. The PDM rainfall-runoff model. Hydrol Earth Syst. Sci. 2007;11(1):483-499.
20. Polcher J, Laval K, Dümenil L, Lean J, Rowntree PR. Comparing three land surface schemes used in general circulation models. J Hydrol. 1996;180(1-4):373-394.
21. Rowntree PR, Lean J. Validation of hydrological schemes for climate models against catchment data. J Hydrol. 1994;155(3-4):301-323.
22. Sandoval S, Bertrand-Krajewski JL. From marginal to conditional probability functions of parameters in a conceptual rainfall-runoff model: An event-based approach. Hydrol Sci J. 2019;64(11):1340-1350.
23. Shamseldin AY, O'Connor KM, Liang GC. Methods for combining the outputs of different rainfall-runoff models. J Hydrol. 1997;197(1-4):203-229.
24. Sun S, Bertrand-Krajewski JL. Separately accounting for uncertainties in rainfall and runoff: Calibration of event-based conceptual hydrological models in small urban catchments using Bayesian method. Water Resour Res. 2013;49(9):5381-5394.
25. Todini E. The ARNO rainfall-runoff model. J hydrol. 1996;175(1-4):339-382.
26. Wheater HS. Progress in and prospects for fluvial flood modelling. Philos. Trans R Soc. Lond. A. 2002;360(1796):409-1431.
27. Winkler RL. Combining forecasts: A philosophical basis and some current issues. Int. J Forecast. 1989;5(4):605-609.
28. Winkler RL, Makridakis S. The combination of forecasts. J R Stat Soc. Ser. A. 1983;146(2):150-157.
29. Zhang C, Wang RB, Meng QX. Calibration of conceptual rainfall-runoff models using global optimization. Adv. Meteorol; c2015.
30. Makridakis S, Andersen A, Carbone R, Fildes R, Hibon M, Lewandowski R, *et al*. The accuracy of extrapolation (time series) methods: Results of a forecasting competition. Journal of forecasting. 1982 Apr;1(2):111-153.