Advanced Engineering Informatics 34 (2017) 80–89

Contents lists available at ScienceDirect

Advanced Engineering Informatics

journal homepage: www.elsevier.com/locate/aei

Stream flow predictions using nature-inspired Firefly Algorithms and a Multiple Model strategy – Directions of innovation towards next generation practices

R. Khatibi[a,*], M.A. Ghorbani[b], F. Akhoni Pourhosseini[c]

[a] GTEV-ReX Limited, Swindon, UK
[b] Water Engineering Department, University of Tabriz and Engineering Faculty; also Near East University, 99138 Nicosia, North Cyprus, Mersin 10, Turkey
[c] Department of Water Engineering, Faculty of Agriculture, University of Tabriz, Tabriz, Iran

ARTICLE INFO

Keywords:
Backpropagation
Firefly Algorithm (FFA)
Multi-Layer Perceptron (MLP)
Multiple Modelling
Prediction
Stream Flow

ABSTRACT

Stream flow prediction is studied by Artificial Intelligence (AI) in this paper using Artificial Neural Network (ANN) as a hybrid of Multi-Layer Perceptron (MLP) with the Levenberg–Marquardt (LM) backpropagation learning algorithm (MLP-LM) and (ii) MLP integrated with the Fire-Fly Algorithm (MLP-FFA). Monthly stream flow records used in this prediction problem comprise two stations at Bear River, the U.S.A., for the period of 1961–2012. Six different model structures are investigated for both MLP-LM and MLP-FFA models and their results were analysed using a number of performance measures including Correlation Coefficients (CC) and the Taylor diagram. The results indicate a significant improvement is likely in predicting downstream flows by MLP-FFA over that by MLP-LM, attributed to identifying the global minimum. In addition, an emerging multiple model (ensemble) strategy is employed to treat the outputs of the two MLP-LM and MLP-FFA models as inputs to an ANN model. The results show yet another further possible improvement. These two avenues for improvements identify possible directions towards next generation research activities.

1. Introduction

Stream flow prediction based on deriving correlations between modelled results and recorded time-series is often one of testing grounds for newly emerging data-driven techniques. This is evidenced indirectly by Sivakumar and Berndtsson [44] presenting the outcome of an internet search on the number of hydrological publications using Artificial Neural Networks ANN) during 1990–2010. This paper is focussed on investigating the integration of the Fire-Fly Algorithms FFA) developed by Yang [50] with the well-established feedforward Multi-Layer Perceptrons (MLP). Whilst ANN is quite well established in stream flow forecasting, MLP-FFA is yet to be applied.

The capability for predicting flows has undergone a radical development over the years since 1960, of which one class of techniques use a type of transfer function by seeking correlation and autocorrelation between flow values at one or more sections of the same river. Whilst prediction techniques based on distributed models are precluded in this paper (from hydrological routing to those based on the Saint-Venant equations), bottom-up data-driven (or data mining) techniques have emerged over the years since the 1960s. Up to 1990, the focus was on such modelling strategies as: traditional transfer functions regression analysis or statistical methods such as ARIMA models of Auto-Regressive Integrated Moving Averages, see Box and Jenkins [3] and Makridakis and Hibon [29].

Data-driven modelling techniques have undergone a radical shift since the late 1980s as further techniques emerged based on Artificial Intelligence (AI). These include: ANN models, see Thirumalaiah [47], Eğrioğlu et al. [7], Rojas [38] ASCE TF [1]; Genetic Programming, see Koza [27], Savic [42] and Kostić et al. [25]; Genetic Expression Programming (GEP), see Ferreira [11], Khatibi et al. [20]; and fuzzy logic, see Kothari et al. [26]; as well as machine learning techniques such as SVM, see Vapnik [48] and Ghorbani [13]. Applications of these techniques for single or more stations to predict hourly, daily or monthly stream flows have been investigated and successful results have been reported.

Abbreviations: AI, Artificial Intelligence; ANN, Artificial Neural Networks; D, downstream station; FFA, Fire-Fly Algorithm; GA, Genetic Algorithm; GEP, gene expression programming; LM, Levenberg-Marquardt algorithm; MLP, Multi-Layer Perceptron; MLP-FFA, MLP synthesised with FFA; MLP-LM, MLP synthesised with the LM algorithm; MM, Multiple Models; MM-ANN, Multiple Models, in which lower order models are driven by ANN; MM-SA, Multiple Models, in which lower order models are driven by Simple Average; MM-SVM, Multiple Models, in which lower order models are driven by SVM; R^2, Correlation Coefficient; RMSE, Root Mean Square Error; SA, Simple Averaging; SD, Standard Deviation; MAE, Mean Absolute Error; SVM, Support Vector Machine; U, upstream station; XOR, exclusive OR gate

* Corresponding author.
E-mail addresses: gtev.rex@gmail.com (R. Khatibi), ghorbani@tabrizu.ac.ir (M.A. Ghorbani), F.Akhonipor93@ms.tabrizu.ac.ir (F.A. Pourhosseini).

http://dx.doi.org/10.1016/j.aei.2017.10.002
Received 25 March 2017; Received in revised form 29 August 2017; Accepted 3 October 2017
1474-0346/

Although first artificial neuron models go back to McCulloch and Pitts [30] as inspiration to emulate natural neural activities, developments of ANN-based working tools for research and practice emerged in the 1990 s after the following basic developments: (i) the concept of MLP emerged earlier in the 1940s and rooted in single neuron perceptrons; (ii) despite the incorporation of various Boolean logic gates to perceptrons, a certain logic gate (XOR) remained elusive but its incorporation led to the development of MLPs; (iii) MLPs driven by XOR gates opened up the door for the rediscovery of the backpropagation algorithms by Rumelhart et al. [40]; and (iv) in time, best practice procedures emerged as ANNs were transformed into tools from the 1990s onward. The mathematical basis of ANN models is described by the universal approximation theorem, according to which a single hidden layer with a finite number of neurons approximates to continuous functions, see Cybenko [6] and Hornik [17]. Hence, ANN capabilities are not confined to a particular problem but cope with diverse applications. These applications are wide and since the 1990s they include stream flows, and in this paper, it serves as the baseline for comparative studies.

ANN models are embedded with parameters that their values need to be identified using optimisation techniques. Classical optimisation capabilities cover the strategies from traditional search algorithm to gradient techniques. These were enriched by new technique by emulating deeper concepts of evolutionary processes. These nature-inspired algorithms include Genetic Algorithm (GA), which serves as an optimisation algorithm, see Goldberg [14]; as well as a host of prediction tools, such as ANN, GP and fuzzy logic. However, as remarked by Flood [12] that progress in ANN applications then had largely stagnated. One avenue for innovation since then has been the emergence of a new generation of nature-inspired optimisation algorithms are being developed in recent years, which emulate the working of species, where the first generation techniques emulate deep concepts of evolutionary processes. New algorithms include Whale Optimisation Algorithm, Moth-Flame Optimisation algorithm, Grey Wolf Optimiser; Firefly Algorithm (FFA). None of these techniques are expression of the best but an identification of a diversity of working orders and possibilities. This paper uses FFA to build on MLP to serve as a strategy for backpropagation problems. FFA is a swarm intelligence optimisation technique based on the movement of fireflies [50] and its applications are widening, see Ghorbani et al. [15,16] and Raheli et al. [37].

It is customary to investigate several models and select the best performing model but reject the others. Contrary to this wide practice, there are some arguments in favour of pluralism in modelling and the need to understand the individual models before rejecting them; see Khatibi et al. [20], Khatibi et al. [21], Khatibi et al. [24], Ghorbani et al. [13]. The question is then how can the synergy among multiple models be used? One way forward is to treat the available multiple or ensemble modelling results together to produce more representative synthetic results. A simple way is to take their simple Ensemble Averages. The idea has a long history, see Clemen [5] for a review. Simple averaging is tantamount to applying constant and equal weights to the contribution of individual models but AI techniques have been applied to multiple models by combining them through using GA to assign linearly varying weights, e.g. see Kadkhodaie-Ilkhchi et al [18]. The linearity limitation has been removed by using ANN or SVM models, as introduced by Nadiri et al. [32], Tayfur et al. [45], Nadiri et al. [33] and Nadiri et al. [34]. The paper explores the use of multiple models to identify a second direction for possible innovations in data-driven modelling practices.

Although FFA has been used successfully in different fields, to the best knowledge of the authors, MLP-FFA is yet to be applied to stream flow problems. This paper investigates the suitability of the hybrid MLP-FFA approach to predicting stream flows; compares its performance with its corresponding ANN model for predicting stream flows at Bear River, USA; discusses the possibility of multiple models, such as using simple average of MLP-LM and MLP-FFA, as an application of multiple models approach; and discusses possible trends for next generation modelling research and practices.

2. Methodology

The paper investigates predicting flows at a downstream station in terms of recorded flow values at the same station and/or an upstream station with various time lags using the following methodologies.

2.1. Artificial Neural Networks (ANN)

The emergence of proof-of-concept for ANNs in 1985 was the outcome of the integration of MLP with backpropagation algorithms. Perceptrons developed in 1957 as an electronic device by Frank Rosenblatt (1928–71), see Rosenblatt [39], was designed to emulate biological processes and to be capable of learning, although using initial single artificial neurons (no hidden layers) was seen unlikely to deliver the capability for learning complex operations. After a series of inventions and reinventions, e.g. XOR logic gates, a deeper insights emerged with MLPs driven by their hidden layers and thereby this led to the integration of MLPs with backpropagation algorithms, often based on least squares techniques.

Since 1985, ANNs have been transformed into flexible working tools, which serve as a modelling strategy to identify possible correlations within a dataset and incorporate an optimisation strategy. Their applications to forecasting flow/stage values within open channels have been an active area of research in the past two or three decades. Typical neural networks consist of three layers of neurons: (i) the input layer, (ii) the hidden layer, and (iii) the output layer; commonly known as feedforward Multi-Layer Perceptron (MLPs). The topology of typical neurons of an ANN is shown in Fig. 1.

This study uses the MATLAB platform and, as shown in Fig. 1, weights are attached between neurons of the input layer connected to those of the hidden layer and from neurons of the output layer to those of the hidden layer through appropriate activation functions. A summing junction adds weighted input signals through activation functions and that between the input layer and the hidden layer is the sigmoid activation function to limit the amplitude of the input data to the range of {0.0–1.0} and that between the hidden layer and the output layer is Purelin. These are expressed mathematically as follows:

$$\text{Input data at the downstream station} = \boldsymbol{Q} = (U_t, U_{t-1}, U_{t-2}, D_{t-1}, D_{t-2}) \tag{1}$$

$$\boldsymbol{O}_j = f_1\left(b_j + \sum_{i}^{I} \boldsymbol{W}_{j,i} Q_i\right) \tag{2}$$

$$\boldsymbol{O}_k = f_2\left(b_k + \sum_{i}^{I} \boldsymbol{W}_{k,j} Q_j\right) \tag{3}$$

where $\boldsymbol{Q}$ is the array of input parameters in terms of flows at the Downstream (D) and/or Upstream (U) observation stations; f_1 and f_2 are actuation functions, b_j and b_k are bias values corresponding to f_1 and f_2 and $\boldsymbol{W}$ is weight with suffices if i, j and k determining their correspondence to f_1 and f_2; in this study i takes values from 1 to 5, j takes values from 1 to 20 and k is set to 1.

The backpropagation algorithm refers to the computational phase for identifying the values of the weights. It was discovered and rediscovered, until in 1985 before its integration with MLP. The weights can be identified by minimising an error function, e.g. the Least Square Method but the identification of the minimum is an optimisation problem and often solved by the steepest gradient technique. In this study, the backpropagation phase employs the widely used Levenberg-Marquardt (LM) algorithm, which interpolates between the Gauss–Newton algorithm and the gradient descent method, for a further detail, see [41].