Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Agricultural Water Management 78 (2005) 195–208
Agricultural water management
www.elsevier.com/locate/agwat

Comparison of Artificial Neural Network and regression models for sediment loss prediction from Banha watershed in India

A. Sarangi *, A.K. Bhattacharya

Water Technology Centre, IARI, Pusa Campus, New Delhi 110012, India

Accepted 2 February 2005
Available online 24 February 2005

Abstract

Two Artificial Neural Network (ANN) models, one geomorphology-based (GANN) and another non-geomorphology-based (NGANN) for the prediction of sediment yield were developed and validated using the hydrographs and silt load data of 1995–1998 for the Banha watershed in the Upper Damodar Valley in Jharkhand state in India. The sediment loads predicted by these models were compared with those predicted by an earlier developed regression model for the same watershed. It was revealed that the feed-forward ANN model with back propagation algorithm performed well for both the GANN and NGANN models. However, the GANN predicted better with highest coefficient of determination (R^2) of 0.98, model efficiency (E) of 0.96 and absolute average deviation (AAD) of 0.0017 in comparison to NGANN ($R^2 = 0.94$, $E = 0.81$, AAD = 0.006). The regression model performance was inferior ($R^2 = 0.940.78$, $E = 0.72$, AAD = 0.023) to the ANN models. The Neural-work-ProII-plus and MATLAB software were used for development of the ANN models. It was also revealed that association of geomorphological parameters viz. relief factor, form factor and drainage factor with runoff rate resulted in a better prediction of sediment loss.

Keywords: ANN; Hydrology; Runoff rate; Sediment Loss; Geomorphology; Regression model; Model efficiency

* Corresponding author. Tel.: +91 9811400885.
E-mail address: arjamadutta.sarangi@elf.mcgill.ca (A. Sarangi).

0378-3774/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.agwat.2005.02.001

1. Introduction

In the past decades, great strides have been made in conceptualizing the runoff and sediment yield processes from watersheds through modeling. Models are classified based on their comprehensiveness in representing the physical processes involved. With increasing comprehensiveness, models are classified as black-box models, conceptual models and physically based distributed models. The last of the three can be considered the better choice in a rigorous theoretical sense. However, the significant data need of such models and their marginally superior results compared to the others make them an unfavorable choice in operational hydrology (Gautam et al., 2000). Lumped conceptual models are favoured, as they can be based on a sound conceptual framework due to their limited data need. But they require lengthy calibration and parameterization processes. Amongst the soft computing tools viz. Genetic Algorithm (GA), Simulated Annealing (SA), Multivariate Adaptive Regression Splines (MARS) and Artificial Neural Networks (ANNs), the ANNs are most frequently used for hydrological modeling.

The first fundamental concepts related to neural computing were developed by McCulloch and Pitts (1943), and much of the ANN activities have been centered on back-propagation and its extensions (Salas et al., 2000). The ANN technique mimics the cognitive response of the human brain. The ANN functions as a data-mining tool, in which the input and output data set has to be fed to the software and trained before validating the model. The network function is determined by the connections between elements. The neural networks need to be trained to perform a particular function by adjusting the values of the connections (weights) between elements. The weights are adjusted based on a comparison of ANN output and the target, until they match. ANNs have an advantage over deterministic models in that the data needs are usually less and they are well suited for long-term forecasting. The disadvantage of the ANN is that it is based on a 'black box' approach since the internal structure of the model is generally not known and must be developed by a trial and error process. There has been a growing trend for the use of ANNs in the areas of hydrologic and water quality modeling (Sharma et al., 2003; Yu et al., 2004; ASCE, 2000a,b; Maier and Dandy, 2000), and land drainage engineering (Yang et al., 1996; Shukla et al., 1996). Despite the black-box nature of the ANN, it has the flexibility in inclusion of parameters and in capturing the non-linearity of rainfall-runoff-sediment yield processes, making it more attractive for modeling hydrological processes (Hsu et al., 1995). The main advantage of the ANN approach over traditional methods is that it does not require an explicit description of the complex nature of the underlying process in a mathematical form (Sudheer et al., 2002).

Cannon and Whitfield (2002) suggested ANNs to be superior to stepwise linear regression procedures while conducting a study on predicting runoff from 5-day mean stream flow atmospheric data from 21 watersheds of British Columbia, Canada. Nagy et al. (2002) used a feed-forward three-layer back propagation (BP) ANN model to predict the sediment concentration in rivers using eight input parameters reflecting sediment and riverbed information. The ANN approach provided better results than other formulas used for estimation of sediment concentration. Sudheer et al. (2002)