Journal of Hydrology 383 (2010) 330–340

Contents lists available at ScienceDirect

Journal of Hydrology

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

# Estimation of ice thickness on lakes using artificial neural network ensembles

I. Zaier [a], C. Shu [a,*], T.B.M.J. Ouarda [a], O. Seidou [b,1], F. Chebana [a]

[a] NSERC/Hydro-Quebec Statistical Hydrology Chair, Canada Research Chair on the Estimation of Hydrometeorological Variables, INRS-ETE, University of Quebec, 490 de la Couronne, Quebec, Qc, Canada G1K 9A9

[b] University of Ottawa, 550 Cumberland St., Ottawa, Ontario, Canada K1N 6N5

ARTICLE INFO

*Article history:*
Received 26 February 2009
Received in revised form 17 November 2009
Accepted 8 January 2010

This manuscript was handled by Dr. A. Bardossy, Editor-in-Chief, with the assistance of Attilio Castellarin, Associate Editor

*Keywords:*
Ice thickness
Artificial neural networks
Lake ice growth
Neural network ensemble
Bagging
Boosting

SUMMARY

Artificial neural network ensemble (ANN ensemble) prediction is a technique in which the outputs of a set of separately trained ANNs are combined to form one unified prediction. ANN ensemble models are developed in this paper to improve the results of single artificial neural network (single ANN) for the estimation of the ice thickness in a number of selected Canadian lakes during the early winter ice growth period. An effective ensemble consists of a set of ANNs that may not be highly performing when they are used separately, but have their prediction errors greatly reduced once combined. This paper evaluates the effectiveness of a number of ensemble techniques including randomization, bagging and boosting for creating members of an ensemble, then averaging and stacking techniques for combining ensemble members. The experiments show that, in the context of estimation of lake ice thickness, boosting is much better than randomization, and sometimes better than bagging. Stacking was found to be more competitive than averaging. Overall, ANN ensemble models for the estimation of ice thickness proved to be more accurate than single ANN models, especially when boosting is used for combining ensemble members and when stacking is used to combine the outputs from individual members. ANN ensembles achieve the best generalization performance when the ensemble size is increased to around 20.



## Introduction

Ice forms in nearly every Canadian lake for a period that ranges from days to several months every year. Extreme events resulting from ice-jamming are the major causes of much economic damage to properties and infrastructures. Models of ice thickness in lakes provide useful information to deal with these problems, and serves our general aim to better understand lake ice processes. For instance, lake ice thickness, date of ice break-up and other ice characteristics are useful indices of climate change which can be modeled and forecasted.

The evolution of ice thickness in lakes is influenced by many interrelated processes. However, the site-specific nature of these complex processes makes ice thickness difficult to predict using physically-based models. The main drawback of the numerical physically-based models is that they require many physical parameters that are hard to collect. Most numerical ice growth models adopt versions of energy budget with different complexity such as the Canadian Lake Ice Model CLIMO used by Ménard et al. (2002a,b), which is a modified version of a one-dimensional sea ice model (Flato and Brown, 1996) and has been described in detail by Duguay et al. (2003). Some models are applied to a specific aspect of ice development such as, ice cover initiation (Schulyakovskii, 1966), border ice formation (Matousek, 1984; Svensson et al., 1989), frazil ice formation (Omstedt, 1985a,b; Svensson and Omstedt, 1994) and ice cover growth (e.g., Schulyakovskii, 1966; Lock, 1990). Other models are more complete and may simulate ice formation, transport, growth and decay (Shen and Chiang, 1984; Shen and Ho, 1986; Shen et al., 1990, 1995). As a result, the development of more analytical models is required. A previous study by Seidou et al. (2006) has shown that artificial neural networks (ANN) can be a valuable alternative to complex thermodynamic lake ice growth models, especially when data are not available in sufficient quality and quantity.

ANN models are generally considered as 'black box' models that are able to capture underlying relationships when presented with input and output data. They have been successfully used in hydrology for solving various problems, such as data classification (e.g., Liang and Hsu, 1994), river discharge prediction (e.g., Shamseldin, 1997), regional flood frequency analysis (Shu and Burn, 2004; Shu and Ouarda, 2007, 2008), water quality evaluation and forecasting (Zhang et al., 1994), estimating river streamflow affected by ice conditions (Chokmani et al., 2008), rainfall estimation (e.g., Xiao and Chandrasekar, 1997) and stream flow under ice estimation

* Corresponding author. Tel.: +1 418 654 3842; fax: +1 418 654 2600.
*E-mail address:* cshu@air-worldwide.com (C. Shu).
1 Tel.: +1 613 562 5800x6143.

0022-1694/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jhydrol.2010.01.006

(Ouarda et al., 2003; Chokmani et al., 2008). The suitability of ANNs for modelling complex systems has resulted in an increase in the popularity of ANN models and their use in an ever increasing number of applications.

Recent studies show that a new approach called ANN ensemble which utilizes multiple ANNs can improve the generalization ability of a single ANN. In an ANN ensemble, a number of ANNs trained for the same purpose as a single ANN are combined to generate a unique output (Shu and Burn, 2004). ANN ensemble approaches have been used successfully in several domains, such as time series modelling, chemistry, robotics, automatic control and medical diagnosis. In the area of forecasting, it has been shown that better results can be achieved by combining forecasts than by choosing the best one (Bates and Granger, 1969). For details concerning the theoretical studies of the ensemble approaches, the reader is referred to the works by Krogh and Vedelsby (1995) and Hansen and Salamon (1990). Cannon and Whitfield (2002) and Shu and Burn (2004) provide a general overview of popular ensemble methods. In this paper, six ANN ensemble models are implemented by using a combination of five ensemble modelling techniques in order to model the ice thickness in a number of selected Canadian lakes, and the results are compared with those of the single ANN models.

The remainder of this paper is composed of six parts: a general introduction of the single ANN model for ice growth modelling (Section "Single artificial neural network"), an overview of general artificial neural network ensemble approaches (Section "General ensemble approaches"), a description of the data used in this study (Section "Data"), a description of the methodology adopted for this study, including single and ensemble ANN models for ice growth estimation, performance criteria and evaluation procedure (Section "Methodology"), results and discussion (Section "Results and discussion"), and finally, conclusions and future work (Section "Conclusions and future work").

## Single artificial neural network

This section presents an overview of the architecture and characteristics of the single ANN proposed for ice growth modelling. The type of ANN selected in this paper is a multilayer perceptron (MLP) feed-forward network which maps sets of input data onto a set of appropriate outputs. The architecture of a MLP with single output is shown in Fig. 1. MLP is the most popular ANN architecture in use today. Reviews of the ANN from a statistical perspective have been given by a number of authors (e.g. Ripley, 1993; Cheng and Titterington, 1994; White, 1989). The most widely used training algorithm for a MLP is the error back-propagation algorithm. This popular algorithm was described firstly by Werbos (1974). However, it was only in 1986 that it was introduced and popularized by Rumelhart and McClelland (1986). The back-propagation algorithm requires that the transfer function used by the artificial neurons be differentiable. It works by iteratively changing the network's interconnecting weights such that the overall error between observed values and network outputs is reduced. Network geometry is generally defined by the number of hidden layer nodes and the number of nodes in each of these layers. It determines the number of model parameters that need to be estimated. The single ANN used in this study is a one hidden layer MLP with sigmoid neurons in the hidden layer and a linear neuron in the output layer. It has been shown that ANNs with one hidden layer can approximate any continuous function, given sufficient degrees of freedom (Funahashi, 1989 and Hornik et al., 1989). The optimum number of neurons in the hidden layer was identified using a trial and error approach.

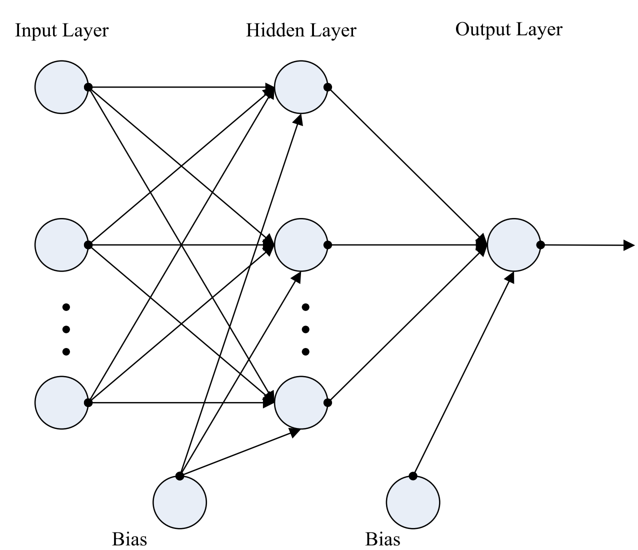


**Fig. 1.** Architecture of single ANN.

## General ensemble approaches

Recent studies have shown that the robustness and reliability of an ANN can be significantly improved by appropriately combining several ANN models into an ANN ensemble (Jacobs et al., 1991; Wolpert, 1992; Perrone and Cooper, 1993; Jordan and Jacobs, 1994; Sridhar et al., 1996; Zhang et al., 1997). The construction of an ANN ensemble requires two major steps. The first step is to create individual ensemble members and the second step is to find the appropriate combination of outputs from the ensemble members to produce the unique ensemble output (Sharkey, 1999). Various methods have been developed for creating ensembles, such as bagging and boosting. Fig. 2 illustrates the two steps for creating an ANN ensemble. For general information and the comparison of these methods, the reader is referred to Opitz and Maclin (1999), Sharkey (1999) and Shu and Burn (2004).

### *Approaches for generating individual ensemble members*

The main objective of combining ANNs in an ensemble is to improve the generalization ability over the single ANN. If the networks in an ensemble share the same characteristics, the ensemble will show a similar performance to the component single ANNs. Thus it is crucial to create individual networks with diverse characteristics while maintaining their individual generalization

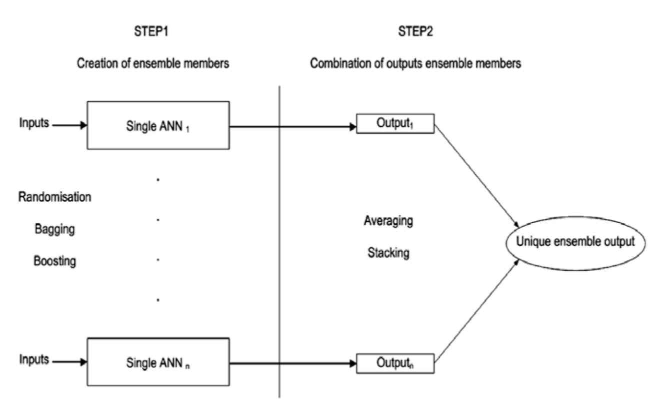


**Fig. 2.** Architecture of ANN ensemble.