WATER RESOURCES RESEARCH, VOL. 42, W11407, doi:10.1029/2005WR004622, 2006

Modeling ice growth on Canadian lakes using artificial neural networks

O. Seidou,[1] T. B. M. J. Ouarda,[1] L. Bilodeau,[2] M. Hessami,[1] A. St-Hilaire,[1] and P. Bruneau[2]

Received 5 October 2005; revised 30 May 2006; accepted 7 July 2006; published 14 November 2006.

[1] This paper presents artificial neural network (ANN) models designed to predict ice in Canadian lakes and reservoirs during the early winter ice thickness growth period. The models fit ice thickness measurements at one or more monitored lakes and predict ice thickness during the growth period either at the same locations for dates without measurements (local ANN models) or at any site in the region (regional ANN model), provided that the required meteorological input variables are available. The input variables were selected after preliminary assessments and were adapted from time series of daily mean air temperature, rainfall, cloud cover, solar radiation, and average snow depth. The results of the ANN models compared well with those of the deterministic physics-driven Canadian Lake Ice Model (CLIMO) in terms of root-mean-square error and in terms of relative root-mean-square errors. The ANN models predictions were also marginally more precise than a revised version of Stefan's law (RSL), presented herein. They reproduced some intrawinter and interannual growth rate fluctuations that were not accounted for by RSL. The performance of the models results in good part from a careful choice of input variables, inspired from the work on deterministic models such as CLIMO. ANN models of ice thickness show good potential for the use in contexts where ad hoc adjustments are desirable because of the limited availability of measurements and where poor data nature, availability, and quality precludes using deterministic physics-driven models.

Citation: Seidou, O., T. B. M. J. Ouarda, L. Bilodeau, M. Hessami, A. St-Hilaire, and P. Bruneau (2006), Modeling ice growth on Canadian lakes using artificial neural networks, *Water Resour. Res.*, *42*, W11407, doi:10.1029/2005WR004622.

1. Introduction

[2] This paper investigates the capacity of artificial neural networks to simulate the growth of ice thickness on Canadian lakes. Although the growth of ice thickness can be modeled with numerical physically driven models, the lack of sufficient and precise data is often a limitation to the applicability of this kind of models. On the other hand, extensive records of meteorological variables such as temperature, rainfall and snow on the ground are widely available over the Canadian territory, along with ice thickness measurements on a number of lakes. Neural network models are often proposed for situations where the physics of the involved processes may be complex or not fully defined, and where an extensive data set is available for training a network while being incomplete from the point of view of deterministic physically driven models [e.g., *Olsson et al.*, 2001; *Cannon and Whitfield*, 2002; *Hewett*, 2003].

[3] Several artificial neural network (ANN) models of lake ice thickness growth are considered in this study. One of these ANN models is regional in scope and is applicable to any location in the country, while the others are specific to the measurement station to which they were adjusted. The results of the ANN models are compared to those of a modified version of Stefan's law and to published results of the deterministic model CLIMO (Canadian Lake Ice Model) [*Menard et al.*, 2002a, 2002b; *Duguay et al.*, 2003].

[4] The remainder of this paper is composed of six parts: an overview of ice growth models based on thermodynamic principles (section 2), an introduction to artificial neural networks (section 3), the methodology for choosing input variables and rating results (section 4), the case study of ice growth on Canadian lakes (section 5), results and discussion (section 6), and finally, conclusions (section 7).

2. Ice Growth Models Based on Thermodynamic Principles

[5] This section presents an overview of ice growth models based on thermodynamic principles. Special attention is given to the formula known as Stefan's law, and to a modified formula referred herein as the revised Stefan's law (RSL), as this formula is later used with the same data sets as the ANNs in order to compare their performances in predicting ice thickness.

[6] Formation and evolution of river and lake ice are governed by heat fluxes in the water body, heat transfers at the interfaces of air-water, air-ice, water-bed, and water-ice, and by radiation exchanges with atmosphere. The different

[1]Centre Eau, Terre et Environnement, Institut National de la Recherche Scientifique, Quebec, Quebec, Canada.
[2]Hydro-Québec, Montreal, Quebec, Canada.

0043-1397/06/2005WR004622$09.00

components of the energy budget are not easy to quantify since they are related to hydraulic conditions (turbulence and velocity distribution), turbidity (which has an effect on the absorption of radiation energy), the albedo and the depth and compactness of snow above the ice cover.

[7] All numerical ice growth models use a more or less simplified version of energy budget. Some of these models apply to a specific aspect of ice development such as the ice cover initiation [*Schulyakovskii*, 1966], border ice formation [*Matousek*, 1984; *Svensson et al.*, 1989], frazil ice formation [*Omstedt*, 1985a, 1985b; *Svensson and Omstedt*, 1994], and ice cover growth [e.g., *Beltaos*, 1995; *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].

[8] They all use energy balance to compute ice evolution. They differ from each other by the level of details used to describe the different aspects of the phenomenon: hydraulics, heat transfer and radiation. Lake ice models are simpler since water velocity is low enough to justify the use of one-dimensional energy balance models [e.g., *Stefan and Fang*, 1997; *Fang et al.*, 1996; *Duguay et al.*, 2003]. For Canadian lakes, *Ménard et al.* [2002a, 2002b] used a one-dimensional thermodynamic lake ice model called CLIMO (Canadian Lake Ice Model) which computes vertical water temperature profiles by solving the heat equation taking account of solar radiation penetrating the water body, ice cover and snow on ice. CLIMO 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]. The inputs of the model are daily mean temperatures, wind speed, relative humidity, cloud cover, snow depth on the ice and lake latitude. Outputs are dates of freeze/thaw and ice thickness. *Ménard et al.* [2002a] used CLIMO to simulate ice growth at station YZF (Back Bay) for the years 1960–1991, and it reproduced ice thickness with a mean quadratic error of 9 cm. They also simulated ice thickness for the years 1977–1990 at station YFL (Fort Reliance) and obtained a root-mean-square error of 18 cm.

[9] Another thermodynamic lake ice model was presented by *Fang et al.* [1996] to compute ice thickness and dates of freeze/thaw on lakes located in Minnesota, USA. The model solves the heat equation along a vertical axis and uses wind-triggered surface layer mixing and water temperature as criteria for predicting the ice cover formation date. The mean quadratic error for ice thickness predictions was 2 cm (for a maximal observed thickness of 55 cm) when this model was validated on lake Ryan located in Minnesota, USA.

[10] This kind of model requires types of data that are not always available, unless there is a meteorological station close to the site. Consequently, in practice, simplified formulas are used based on air temperature.

[11] The simplest and the most widely used formula is the Stefan's law (SL) which could be derived by simplifying the equations obtained using energy balance [e.g., *Lock*, 1990]:

$$H_I = k\sqrt{D_d} \tag{1}$$

where H_I is the ice thickness, D_d is the sum of degree-days below the freezing point since the onset of the ice cover in any given year, and k is a constant. In practice k is used as an adjustable parameter with a value that is lower than the theoretical value to account for varying conditions of exposure and insulation; *Michel* [1971] gives a range of values adapted for a variety of lakes and rivers.

[12] The date of the onset of the ice cover is a basic parameter for using Stefan's law as it determines the date at which the accumulation of freezing degree-days is started for any given winter. For the majority of sites of interest in this study, this date is not known so that the normal usage of Stefan's law was not feasible. It was therefore decided to use a "revised Stefan's law" (RSL), which is based on D_g, the accumulation of freezing degree-days starting with the first day of below freezing air temperature in any given season. RSL presents one more adjustable parameter, C, the effective number of degree-days to be subtracted from D_g in order to obtain D_d.

$$H_I = \begin{cases} k\sqrt{D_g - C} & if \quad D_g \geq C \\ 0 \quad if \quad D_g < C \end{cases} \tag{2}$$

The parameter C will normally take on a positive value since the date of ice onset arrives several days or weeks after the occurrence of the first day of freezing daily mean air temperature.

3. Artificial Neural Networks

[13] An artificial neural network (ANN) is a set of simple computational units or processing nodes grouped in layers and working in parallel. It is called a neural network because the processing nodes mimic the behavior of biologic neurons. These nodes are also called neurons. In the network, each layer makes an independent processing of the information and forward the results to the next layer. The information given to the network passes from the input layer to the output layer through optional intermediate layers (or hidden layers). An ANN model can have more than one hidden layer. However, research has shown that a single hidden layer is sufficient for ANNs to approximate any complex nonlinear function [*Cybenko*, 1989; *Hornik et al.*, 1989]. A larger number of hidden layers can speed up the learning process, but many experimental results seem to confirm that one hidden layer is sufficient for prediction and forecasting problems [*Zhang et al.*, 1994; *Coulibaly and Anctil*, 1999; *Coulibaly et al.*, 2000]. Artificial neural networks have the ability to memorize empirical knowledge and make it available for use. The empirical knowledge refers here to the unknown relationship between observed data series. This empirical knowledge is acquired through the learning algorithm, which essentially modifies internal neuron parameters to fit the outputs of the neural network to the observed response variable. The acquired knowledge is memorized in the synaptic weights, obtained by a training or adaptation process. Finally, this knowledge is restituted when the model is used to simulate the response variable on new input data sets. The use of ANNs is deemed to be particularly useful when the physical processes are complex and not fully defined, when the model has many uncertainties (model coefficients and/or input parameters), and when there is extensive data for training the network.

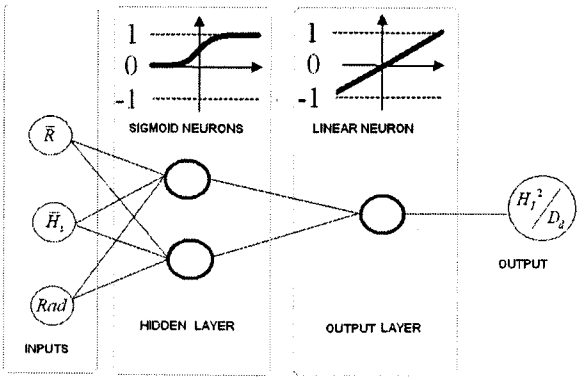

Figure 1. Architecture of a feed forward network with two neurons at the hidden layer and one neuron at the output layer.

It is a "black box" type of model for which the user has no control on the internal behavior, and it can poorly perform when used for generalization especially when the number of neurons is too high for the complexity of the problem (overfitting). In hydrology, ANNs have been used for hydrologic data classification [e.g., *Liang and Hsu*, 1994], river discharge prediction [e.g., *Shamseldin*, 1997], evaluation and forecasting of water quality [e.g., *Zhang et al.*, 1994], inflow forecasting for irrigation and hydroelectric dams [e.g., *Coulibaly et al.*, 2000], rainfall estimation [e.g., *Xiao and Chandrasekar*, 1997], and correction of streamflow under ice [*Ouarda et al.*, 2003].

4. Methodology

[14] The capabilities of RSL and ANNs to model ice growth on individual Canadian lakes are investigated. The performances of the ANN models are investigated as function of its input variables and their internal structure, and a procedure is set up to select the best performing models.

[15] Two kinds of ANN models were considered in this study: (1) site specific ANN models were trained and validated with ice thickness data from a single site and (2) a regional ANN model was also constructed in order to simulate ice growth for sites without field measurements of ice thickness. Each model is characterized by the set of its input variables (referred to as a combination of input variables), and the ANN architecture (defined by the number of neurons on the hidden layer).

[16] Choosing the best set of input variables and assessing the performance of a model requires some performance criteria, some model validation procedures, and a method of picking the best model among all possible configurations. Since our models are data-driven, two validation procedures were used to ensure that the final ANN models are properly trained and can be used with confidence for generalization purposes. These validation procedures are described at section 4.4.

4.1. Architecture of the Artificial Neural Networks

[17] The artificial neural network used in this research was a one-hidden-layer neural network with sigmoid neurons in the hidden layer and a linear neuron in the output layer [*Hagan*, 1996]:

$$a^{m+1} = f^{m+1}\left(W^{m+1}a^m + B^{m+1}\right) \qquad m = 0, 1 \tag{3}$$

where a^0 is the network input, a^1 is the output of the hidden layer, a^2 is the network output, f^1 and f^2 are respectively sigmoid and linear transfer functions:

$$f^1(n) = \frac{1}{1+e^{-n}} \tag{4}$$

$$f^2(n) = n \tag{5}$$

W^{m+1} and B^{m+1} are network weights and biases which are defined with the following formulas:

$$W^{m+1} = \begin{bmatrix} w_{1,1}^{m+1} & \cdots & w_{1,S_m}^{m+1} \\ \vdots & \ddots & \vdots \\ w_{S_{m+1},1}^{m+1} & \cdots & w_{S_{m+1},S_m}^{m+1} \end{bmatrix}, \qquad B^{m+1} = \begin{bmatrix} b_1^{m+1} \\ \vdots \\ b_{S_{m+1}}^{m+1} \end{bmatrix} \tag{6}$$
$$m = 0, 1$$

where S_0 is the number of input variables, S_1 is the number of neurons at the first layer and S_2 is the number of neurons at the second layer.

[18] One-hidden-layer neural networks with sigmoid activation function on the hidden layer and a linear activation neuron were shown to be able to approximate any bounded continuous function with arbitrarily small error [*Cybenko*, 1989], provided the number of neurons in the hidden layer is sufficient. Because of the neurons with sigmoid activation functions in the hidden layer, this ANN model is not a linear model. Other architectures may have been chosen, but there are no rules for choosing the number and the sizes of an ANN model. For instance, it is believed that increasing the number of layers can increase the learning capabilities of the model, but it also increases the number of parameters and thus the length of the series required to properly train it. For the onset of this study, the simpler and most popular architecture was chosen.

[19] Figure 1 shows the architecture of a one-hidden-layer neural network with three input variables, two neurons in the hidden layer and one neuron in the output layer. The synaptic weights are obtained using a supervised training algorithm using Bayesian regulation [*Mackay*, 1992, 1995]. In supervised training, both the inputs and the outputs are provided. The network then processes the inputs and compares its resulting outputs with the desired outputs. Errors are then propagated back through the system, causing the system to adjust the weights which control the network.

[20] An iterative trial and error process, described in more details in the rest of the paper, will be used to set the optimal number and nature of input variables as well as the number of neurons in the hidden layer.

4.2. Selection of Input and Output Variables

[21] ANNs are data-driven models, so their performance for a given problem relies on the relevance of the input/output variables that are considered in the training process, and on the complexity of the relationship between inputs

and outputs. As it is the case for physically driven models, problem parameterization can dramatically change algorithm efficiency. The purpose of the ANN being the prediction of ice thickness, the formulation of its input and output variables was guided by these practical considerations: (1) inputs should be meaningful for the ice growth process and available at all ice measurement stations, (2) the sensitivity of the output variable to input variations should be as high as possible to enhance training efficiency, and (3) the sensitivity of the output variable to input variations should have the same order of magnitude for all input variables.

4.2.1. Output Variable

[22] At first, it seemed obvious that H_I, the ice thickness, should serve as the output of the ANN. Early runs of the ANN showed however that the ANN relied almost exclusively on the sum of freezing degree-days and made little use of the other variables (presented below). The reason for this is that the influence of the sum of freezing degree-days on the output function was so strong that the training algorithm (which is basically a numerical optimization algorithm) was unable to account for the effect of the other variables. Other variables closely related to the ice thickness were then sought with the intent of improving the fit between observations and predictions of ice thickness.

[23] After some exploration, the output variable that was retained for this study is the parameter H_I^2/D_d, which corresponds to the Stefan coefficient. Once a time series of the growth rate of the square of the ice thickness has been computed, a time series of ice thickness is easily reconstructed. The parameter H_I^2/D_d is drawn directly from the classic Stefan's law and represents an instantaneous evaluation of the square of its constant k. Using this output variable led to a reduction of the weight of the degree-days input variable in the ANNs, compared to that of other input variables, and presumably allowed the ANN to extract additional information from the other input variables. It was found that this choice of an output variable increased appreciably the precision of the ice thickness prediction and led to a better comprehension of the effect of the other variables in modeling the ice thickness.

4.2.2. Input Variables

[24] In the present study, selecting the input variables of the ANNs was a two step process. First, candidate physically observed variables were identified. Then, ANN input variables were formulated using the physically observed variables. These physically observed variables were retained according to their availability and recognized significance for the heat budget involved in ice growth: (1) daily mean air temperature, (2) daily total solar radiation, (3) daily rainfall, and (4) daily snow depth on the ground (as measured at weather stations). Longitude and latitude were also tested as additional parameters for the regional neural network model.

[25] The preceding variables were used to construct a variety of variables, to be used singly or in linear combinations as input variables for the ANNs. When these variables are sums, the summation starts on the first day of frost based on the daily mean air temperature and ends on the day assigned to the variable. The following variables were considered and computed for each winter: (1) the sum of freezing degree-days derived from the air temperature

Table 1. Tested Combinations of Meteorological Variables

Combinations of Meteorological Variables	Variables[a]
1	D_d ; Rad_{nc}
2	D_d ; Rad_{nc} + 0.25 Rad_c
3	D_d ; Rad_{nc} + 0.50 Rad_c
4	D_d ; Rad_{nc} + 0.75 Rad_c
5	D_d ; Rad_{nc} + Rad_c
6	D_d ; Rad_{nc}; $\bar{H}_S$
7	D_d ; Rad_{nc} + 0.25 Rad_c; $\bar{H}_S$
8	D_d ; Rad_{nc} + 0.50 Rad_c;$\bar{H}_S$
9	D_d ; Rad_{nc} + 0.75 Rad_c; $\bar{H}_S$
10	D_d; Rad_{nc} + Rad_c; $\bar{H}_S$
11	D_d ; Rad_{nc};$\bar{R}$
12	D_d ; Rad_{nc} + 0.25 Rad_c;$\bar{R}$
13	D_d; Rad_{nc} + 0.50 Rad_c;$\bar{R}$
14	D_d; Rad_{nc} + 0.75 Rad_c;$\bar{R}$
15	D_d; Rad_{nc} + Rad_c;$\bar{R}$
16	D_d; Rad_{nc};$\bar{H}_s$; $\bar{R}$
17	D_d ; Rad_{nc} + 0.25 Rad_c; $\bar{H}_S$; $\bar{R}$
18	D_d ; Rad_{nc} + 0.50 Rad_c; $\bar{H}_S$; $\bar{R}$
19	D_d ; Rad_{nc} + 0.75 Rad_c; $\bar{H}_S$; $\bar{R}$
20	D_d; Rad_{nc} + Rad_c;$\bar{H}_S$; $\bar{R}$
21	D_d ; $\bar{H}_S$
22	D_d;$\bar{R}$
23	D_d ;$\bar{H}_S$; $\bar{R}$

[a]Note is used only for data preprocessing and postprocessing.

D_g(°Cday), as was already described earlier in this paper, (2) the sum of solar radiation during the period of ice growth for days with precipitation (W day/m^2), divided by the sum of degree-days Rad_c (this quantity is a proxy to the quantity of solar radiation attenuated by cloud cover), (3) the sum of solar radiation during the period of ice growth for days without precipitation (W day/m^2), divided by the sum of degree-days Rad_{nc}, (4) the average daily rainfall (over time) during the ice growth period $\bar{R}$ (mm), and (5) the average on-ground snow depth (over time) during the ice growth period $\bar{H}_S$(cm).

[26] The sum of degree-days D_g was used to compute H_I^2/D_g on the calibration data set (preprocessing), and to reconstruct ice thickness time series from simulated series (postprocessing). It will be referred to as an input variable in the remainder of the text.

[27] The sum of solar radiation was split into two parts because only the total amount that reaches the ground should be considered. Consequently, attenuation due to cloud cover has to be accounted for by multiplying Rad_c by a positive factor α smaller than or equal to 1. Five sets of combinations of Rad_{nc} and Rad$_c$ were considered: Rad_{nc}, Rad_{nc}+0.25 Rad_c, Rad_{nc} Rad_c, Rad_{nc}+0.75 Rad_c, Rad_{nc}+ Rad_c, corresponding respectively to α = 0, 0.25, 0.50, 0.75 and 1. A total of 23 sets of combinations of meteorological variables (listed in Table 1) were tested as inputs for the artificial neural network models.

[28] Other important variables for ice growth would have been variables based on lake morphology such as maximum depth, mean depth, surface area, perimeter length and other variables that could be derived from these. For most of the lakes considered, depth information could not be found within the scope of this study. Other parameters based on the surface area and lake shape on maps were not used in the end. It should be noted that this information is never-

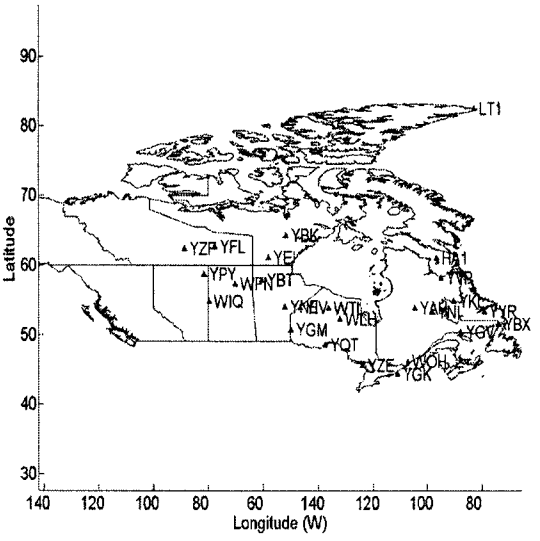

Figure 2. Location of the 26 lake ice measurement stations.

theless considered to be relevant, particularly for the evaluation of the date of freezeup. It is felt by the investigators that the use of morphological lake data could provide interesting avenues for the use of the ANNs for the modeling of lake ice in future studies. They would require a specialized focus on sources of data that could not be explored within the framework of the present study.

4.3. Performance Criteria

[29] The performance of a given ANN model is evaluated using five performance criteria: root-mean-square error (RMSE), relative root-mean-square error (RRMSE), model explained variance (r^2), Nash criterion (NASH), and bias (BIAS). The criteria are defined as follows:

$$\mathrm{RMSE} = \left(\frac{1}{n}\sum_{k=1}^{n}\left(H_I^k - \hat{H}_I^k\right)^2\right)^{\frac{1}{2}} \tag{7}$$

$$\mathrm{RRMSE} = \left(\frac{1}{n}\sum_{k=1}^{n}\left(\frac{H_I^k - \hat{H}_I^k}{\bar{H}_I}\right)^2\right)^{\frac{1}{2}} \tag{8}$$

$$r^2 = \frac{\mathrm{cov}\left(\left[H_I^1,\ldots,H_I^n\right],\left[\hat{H}_I^1,\ldots,\hat{H}_I^n\right]\right)^2}{\mathrm{var}\left(\left[H_I^1,\ldots,H_I^n\right]\right)\mathrm{var}\left(\left[\hat{H}_I^1,\ldots,\hat{H}_I^n\right]\right)}\mathrm{S} \tag{9}$$

$$\mathrm{BIAS} = \frac{1}{n}\sum_{k=1}^{n}\left(H_I^k - \hat{H}_I^k\right) \tag{10}$$

$$\mathrm{NASH} = 1 - \frac{\sum_{k=1}^{n}\left(H_I^k - \hat{H}_I^k\right)^2}{\sum_{i=1}^{2}\left(H_I^k - \bar{H}_I\right)^2} \tag{11}$$

where $H_I^k, i = 1,..,n$ and $\hat{H}_I^k, \mathrm{i} = 1,\ldots,n$ are observed and simulated ice thicknesses.

4.4. Validation Procedures

[30] Two validation procedures were used: the leave-one-out cross-validation procedure to find the best combi-

Table 2. Ice Thickness Measurement Stations

Station Code	Station Name	Water Body	Longitude, °W	Latitude, °N
HA1[a]	Quaqtaq	Unnamed Lake	68.36	61.03
LT1	Alert	Upper Dumbell Lake	61.5	82.46
WFN[a]	Cree Lake	Cree Lake	105.15	57.33
WIQ	Primrose Lake	Primrose Lake	109.93	54.76
WLH[a]	Lansdowns house	Attawapiskat Lake	86.09	52.21
WHO	Sainte agathe des monts	Lac des sables	73.69	46.03
WTL[a]	Big trout Lake	Big trout Lake	88.11	53.81
YAH[a]	La grande IV	Lac la Tarrière	72.36	53.76
YBK[a]	Baker Lake	Baker Lake	95.96	64.3
YBT[a]	Brochet	Brochet bay of Reindeer Lake	100.31	57.86
YBX[a]	BlancMsablon	Lac à la Truite	56.81	51.45
YEI[a]	Ennadai Lake	Ennadai Lake	99.09	61.11
YGK[a]	Kingston	Lake Ontario-Horsey Bay	75.5	44.2
YGM[a]	Gimli	Lake Winnipeg	95.01	50.61
YGV	Havre Saint Pierre	Patterson Lake	63.88	50.29
YIV[a]	Island Lake	Island Lake	93.31	53.84
YKL[a]	Schefferville	Knob Lake	65.19	54.78
YNE	Norway house	Little Playgreen Lake	96.16	53.98
YNI	Nitchequon	Nitechequon Lake	69.08	53.18
YPY[a]	Fort chipewyan	Lake Athabasca	110.83	58.7
YQT	Thunder bay	Thunder Bay	88.78	48.43
YVP[a]	Kuujjuaq	Stewart Lake	67.53	58.11
YYR[a]	Goose bay	Terrington Basin	59.58	53.33
YZE	South baymouth	Huron Lake (South Bay)	81.98	45.54
YZF[a]	Yellowknife	Great Slave Lake (Back Bay)	114.34	62.45
YFL[a]	Quaqtaq	Great Slave Lake (Fort Reliance)	108.86	62.70

[a]Stations with enough data to calibrate local ANN models.

nation of input variables and ANN structure, and the standard split sample procedure in which the data is split in two subsets, the first being used to train the neural network and the latter being used to check the model performance.

[31] The leave-one-out procedure aims to assess the generalization capability of a given ANN model. In the leave-one-out procedure, the data is divided in subsets. A subset is represented by all measurements during a year (for local ANN models and RSL) or all measurements at a given site (for the regional ANN model). All subsets but one are used to train the neural network. Then the trained ANN model is used to simulate the subset that is left out. The same process is repeated so that every subset of the data is left out once. The performance indices are then computed using the observed and simulated values over the whole data set. The pair of input variables and ANN structure which best explain the data is then selected for the rest of the procedure. Unfortunately, the leave-one-out procedure does not give a final model since the neural network is trained several times (one time for each subset) to compute the performance indices.

[32] The objective of the split sample validation procedure is to have a trained neural network to use for the rest of the study. In this procedure, the data is repeatedly and randomly divided into two parts containing 80% and 20% of the observations. The first part is used to train the ANN and the other is used to rate it. The operation is performed twenty times and the parameter set of the run which gives the smallest RMSE is retained and completes the assembly of the model.

4.5. Selection of the Best Combination of Meteorological Variables and ANN Structure

[33] First, all possible pairs composed of a set of input meteorological variables (listed in Table 1) and a network structure (defined by the number of neurons in the hidden layer) are formed, given that the number of neurons on the hidden layer is constrained to be less than 10 for computational purposes. For example, the pair (3, 7) represents an ANN model with seven neurons on the hidden layer and for which the inputs variables are D_d, Rad_{nc} + 0.50 Rad_c. The performance criteria of each pair are then rated using the leave-one-out procedure.

[34] When some values of an input variable (such as sum of rain or snow on ground) are missing at a given site, the combinations of meteorological inputs containing that variable cannot be used. In this case, only a part of the 23 possible combinations of inputs are tested when searching the best local ANN model. To avoid this situation with the regional ANN model, it was trained and validated with the data of the measurement stations where the entire input variable were available.

[35] The best pair is then chosen, and the combination of meteorological variables in this pair is considered to be the one which best explains the data, while the number of neurons in the pair defines the best ANN structure.

5. Case Study

5.1. Data

[36] The data used in this study is of three types: (1) ice thickness data from the *Canadian Ice Service* [2005],

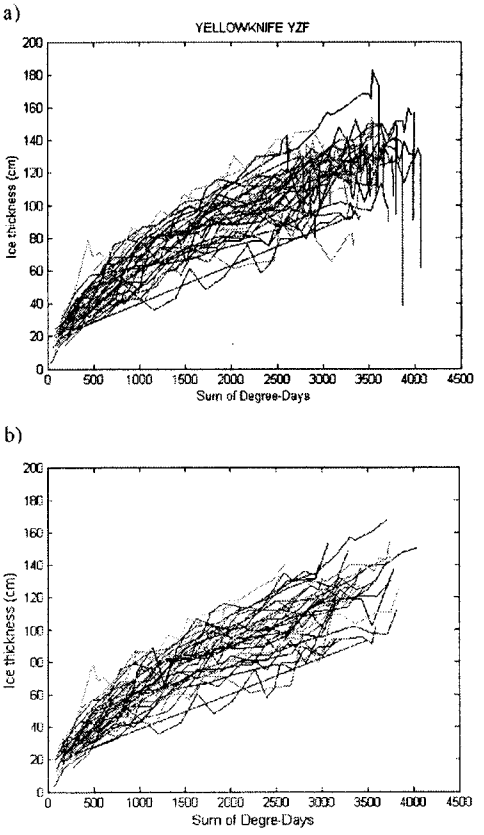

Figure 3. Identification of ice growth phase in the data sets at station YZF: (a) all observations and (b) ice growth phase.

(2) daily meteorological data obtained from Environment Canada, and (3) incident solar radiation at the top of atmosphere which have been computed as function of time and geographic location of the studied sites using the formulas presented in Appendix A.

[37] The data set contains ice thickness and snow depth measurements on 53 sites located on Canadian lakes [*Lenormand et al.*, 2002]. Only 26 of these sites have years with enough measurements to calibrate Stefan's law and the local ANN. Eight of these stations were not used for calibration and validation of the regional ANN because of data availability issues that will be further explained. The 26 ice measurement stations are listed in Table 2, and their geographic locations are presented in Figure 2.

[38] The weather data was provided by the national climatic archives of Environment Canada and contains the daily data of temperature, precipitation (snow and rain) and snow on the ground for more than 10,000 stations distributed all over the Canadian territory.

5.2. Proxy Variable for Solar Radiation at Ground Level

[39] Solar radiation was considered as a relevant variable at the onset of this study because radiative fluxes were recognized to be a significant part of the lake energy by most authors [e.g., *Lock*, 1990; *Fang et al.*, 1996; *Menard et al.*, 2002a, 2002b; *Duguay et al.*, 2003]. However, solar

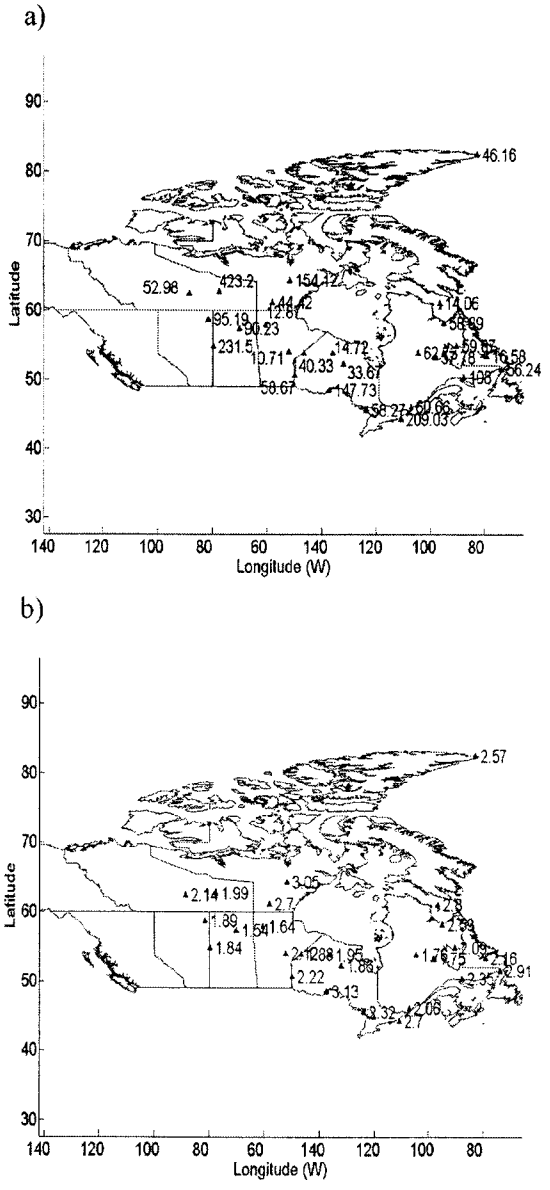

Figure 4. Spatial variability of RSL parameters: a) parameter C; b) parameter k.

radiation measurements at the ground level were not available in sufficient quantity since they are not measured at all weather stations. As an alternative, assumed ground level solar radiation was calculated using solar radiation on top of the atmosphere multiplied by a factor α; α is equal to the unity on days without precipitation, and lower than the unity on days when rain or snow was observed. The values of solar radiation on top of the atmosphere were calculated as functions of the date and the geographic location according to an algorithm given in Appendix A.

5.3. Spatial Interpolation of Weather Data

[40] As geographic positions of the ice measuring sites did not coincide with that of the weather stations, the air temperature on a given date at a given point is interpolated using the observed values at the nearby stations: a data processing program seeks the nearest weather station successively in the northeast, northwest, southwest, and southeast quadrants. the temperature of the ice measurement site is computed as the weighted average of the temperatures at the weather stations. The weights are inversely proportional to the distances from the site of measurement to the weather stations. If historical measurements are not available at the ten closest stations in a quadrant on a given date, the data is considered missing. When a variable (such as rain, degree-days or snow) has missing values, the combinations of inputs containing this variable cannot be used. This may happen when weather stations are too far from the ice measurement station. In this case only part of the 23 possible combinations of input listed in Table 1 can be considered. At eight of the 26 lake ice thickness measurement stations, some irregularities were observed in the data (such as a long delay between first freezing air temper-

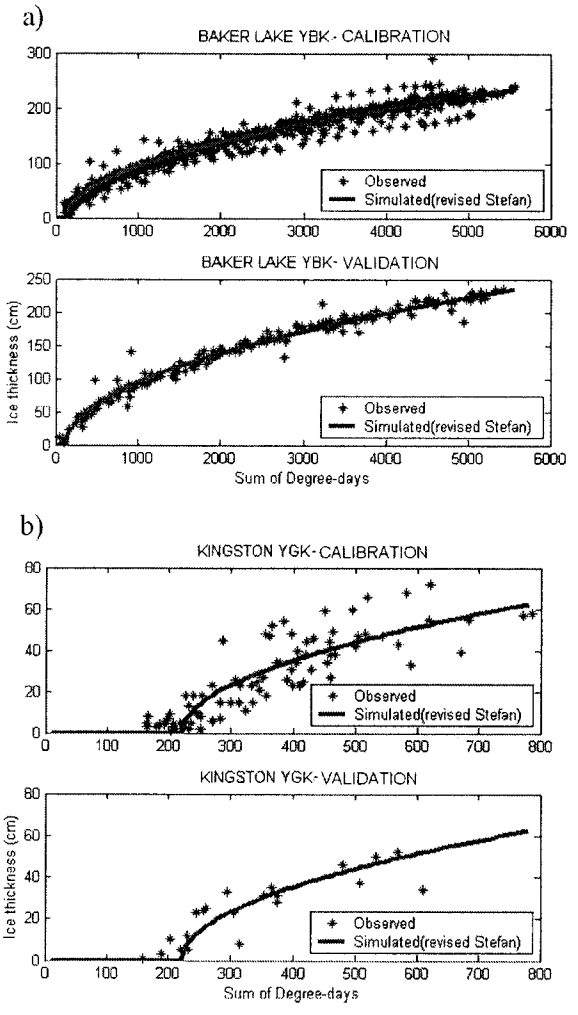

Figure 5. Observed and simulated (RSL) ice thicknesses versus sum of degree-days at some ice measurement stations: (a) YBK (Baker Lake) and (b) YGK (Kingston).

Table 3. Performance Criteria for the Revised Stefan's Law

Station	k, cm $d^{-0.5}$ $°C^{-0.5}$	C, d °C	Calibration RMSE, cm	Calibration r^2	Calibration RRMSE	Calibration BIAS, cm	Validation RMSE, cm	Validation r^2	Validation RRMSE	Validation BIAS, cm
HA1	2.80	14.06	4.97	0.98	0.05	1.02	9.81	0.96	0.09	−3.56
LT1	2.57	46.16	11.94	0.90	0.12	3.16	30.29	0.72	0.19	47.30
WFN	1.54	90.23	4.17	0.86	0.07	0.10	14.84	0.68	0.16	−6.68
WIQ	1.84	231.50	6.21	0.84	0.17	−2.94	8.49	0.74	0.28	2.90
WLH	1.86	33.67	5.87	0.70	0.10	−1.76	9.34	0.83	0.14	3.61
WOH	2.06	60.66	4.94	0.69	0.08	0.73	8.41	0.82	0.18	1.20
WTL	1.95	14.72	5.61	0.91	0.08	0.49	13.82	0.87	0.15	−8.09
YAH	1.76	62.73	5.53	0.93	0.10	−1.99	10.65	0.89	0.16	7.20
YBK	3.05	154.12	6.07	0.97	0.06	4.27	11.69	0.96	0.10	−4.22
YBT	1.64	12.87	6.50	0.86	0.11	−0.28	7.32	0.92	0.14	−1.35
YBX	2.91	56.24	7.92	0.72	0.23	1.08	21.86	0.83	0.50	19.12
YEI	2.70	44.42	7.29	0.81	0.08	4.06	13.64	0.92	0.13	2.94
YGK	2.70	209.03	4.83	0.71	0.40	3.17	10.51	0.77	0.91	8.31
YGM	2.22	58.67	5.90	0.60	0.07	1.63	14.20	0.85	0.15	−9.81
YGV	2.35	108.00	5.15	0.87	0.10	1.24	32.87	0.74	1.13	28.97
YIV	1.88	140.33	5.14	0.69	0.08	0.15	11.22	0.76	0.19	1.78
YKL	2.09	59.87	6.29	0.92	0.07	0.83	11.38	0.91	0.15	−6.51
YNE	2.12	10.71	8.05	0.80	0.21	2.88	9.26	0.88	0.17	10.46
YNI	1.75	57.78	5.81	0.92	0.11	2.17	9.41	0.89	0.13	−5.44
YPY	1.89	95.19	10.30	0.55	0.26	1.25	13.83	0.70	0.29	8.65
YQT	3.13	147.73	4.19	0.74	0.08	−1.06	11.24	0.80	0.92	51.10
YVP	2.53	58.89	6.24	0.86	0.08	2.78	11.44	0.95	0.12	6.28
YYR	2.16	16.58	7.80	0.73	0.16	1.62	13.68	0.67	0.19	3.05
YZE	2.32	58.27	5.75	0.70	0.13	−0.92	6.95	0.74	0.16	2.30
YZF	2.14	52.98	6.57	0.85	0.09	2.21	13.80	0.84	0.16	0.27
YFL	1.99	423.2	15.24	0.80	0.09	−0.30	2.21	13.80	0.16	0.27
Maximum	3.13	423.20	15.24	0.98	0.40	4.27	32.87	13.80	1.13	51.10
Minimum	1.54	10.71	4.17	0.55	0.05	−2.94	2.21	0.67	0.09	−9.81
Mean	2.23	89.18	6.70	0.80	0.12	0.98	12.78	1.32	0.27	6.16

atures and the ice growth starting period, or missing values in some of the interpolated weather variables) so only the 18 remaining stations were used for the construction of the ANN models. These stations are indicated in Table 2.

5.4. Elimination of the Period of Thinning at the End of the Winter in the Data Set

[41] The model developed in the present study is applicable to the growth phase of the ice thickness. It is therefore important to eliminate the period of ice thinning which happens at the end of winter. For this purpose, we use an empirical iterative procedure: starting from the end of the data series and moving toward its beginning, all measurements are successively considered for eventual removal. All the dates when the measured ice thickness happened to be higher than the average ice thickness of the remaining subsequent measurements. Previously eliminated values are not accounted for when computing the average ice thickness of the remaining subsequent measurements. When a given measurement is removed, all subsequent measurements are also removed. This practical procedure insures that the thinning period, and part of the period when the thickness stagnates are eliminated. The results of the application of such a procedure to station YZF are illustrated in Figure 3. Figure 3a presents the whole set of measurements including the thinning period and displays several step drops on the ice thickness which would have affected the training of the ANN models. The retained measurements after application of the proposed procedure are illustrated in Figure 3b. These results correspond to what was expected when designing the procedure and were used for subsequent analyses.

6. Results and Discussion

[42] The data sets described in section 5 were used to calibrate the RSL, and to train the local and regional ANN models. ANN training and simulations were performed using the Neural Networks toolbox of Matlab [*The Mathworks*, 2005]. The training function (trainbr function in the Matlab environment) uses Bayesian regulation [*Mackay*, 1992, 1995] to enhance generalization capabilities. The maximum number of epochs is set to 1000, and the minimum gradient (for minimization of the objective function) to 1×10^{-10}. The training algorithm also uses an adjustable parameter μ_{max} which controls how far the next values of the parameters will be searched during the optimization process. μ_{max} was set to its default value in Matlab, i.e., 1×10^{10}.

6.1. Comparison of RSL and Local ANN Models

[43] The spatial variations of RSL parameters are illustrated in Figure 4, and no trend could be found with respect to geographical location. It was found that RSL represents an excellent ice growth model when it is calibrated for a given site. Observed and simulated data are represented for some of the ice thickness measurement stations in Figure 5. The parameters of Stefan's law are obtained by the least squares method using 80% of the data. There is a very good agreement between simulated data and observations, with a

Table 4. Performance Criteria for Revised Stefan's Law and Local ANN Models[a]

Station	Best Combination	Number of Neurons on the Hidden Layer	RMSE (Stefan), cm	R^2 (Stefan)	RRMSE (ANN), cm	r^2 (ANN)
HA1	10	1	8.94	0.95	8.02	0.96
WFN	5	9	11.59	0.71	11.07	0.74
WLH	23	4	14.54	0.59	11.28	0.76
WTL	9	1	10.54	0.85	10.50	0.85
YAH	22	3	8.96	0.81	7.97	0.85
YBK	15	1	11.59	0.96	11.72	0.96
YBT	9	1	8.53	0.86	8.52	0.86
YBX	4	10	13.63	0.66	13.03	0.69
YEI	10	1	15.55	0.86	14.27	0.88
YGK	7	9	9.16	0.74	13.43	0.47
YGM	9	3	14.70	0.59	10.99	0.77
YIV	10	2	12.03	0.69	11.47	0.72
YKL	1	9	9.72	0.89	9.36	0.90
YPY	11	9	19.98	0.54	19.88	0.55
YVP	20	2	13.22	0.88	12.25	0.90
YYR	14	9	12.86	0.69	11.91	0.74
YZF	10	2	13.73	0.83	12.48	0.86
YFL	5	10	17.10	0.74	16.78	0.75

[a]Best combination and optimal number of neurons on the hidden layer and performance indices obtained during the leave-one-out cross-validation procedure.

mean RRMSE of 0.12 for the calibration set and 0.27 for the validation set (Table 3). Hence RSL may be an excellent model for engineering purposes when errors of a few centimeters do not matter. However, its major drawback is that parameter values are variable from lake to lake. It is thus difficult to determine which values to apply for a lake without observations, unlike the regional ANN model developed in this paper which can readily be used for such lakes.

[44] Table 4 illustrates the performance of the local ANN models and the revised Stefan's law when using the leave-one-out cross-validation procedure. It turned out that it is always possible to find a combination of meteorological variables such that the neural networks perform better than the Stefan's law except at two stations: the station YBK (Baker Lake: Figure 5a) and station YGK (Kingston: Figure 5b). The data of YBK (Baker Lake) present very little variability and the two models had quasi equal performances. The data of the second station come from Lake Ontario, one of the Great Lakes with an average depth of 91 m and it comes as no surprise that the growth of the ice starts rather late comparatively to all the others (degree-days of about 200).

[45] The number of combinations for which ANN models perform better than the Stefan's law vary from two at station YBT (Brochet lake) to more than 120 at station YAH (La Terrière Lake). The best combinations (in terms of RMSE) for each station are given in Table 4. The best combination is different for each station, but 86% contain the proxy for radiation, 55% contain snow on the ground and 33% contain rainfall. The dominating factor apart from the degree-days in the ice growth process is then solar radiation, followed by snow and rainfall.

[46] Once the best combination of variables for a given station selected, the local neural network model for this station is obtained following the procedure described in section 4.4. The performance indices of the retained local ANN models are listed in Table 5.

[47] The variation of the ice thicknesses simulated with both the local ANN and RSL at station YGM when performing the leave-one-out procedure are presented in Figure 6. Figures 6a and 6b both correspond to a case where a year of data has been excluded. It can be seen that the ANN model follows data variations more closely than RSL. The ANN model adapts well to the intrawinter and interannual variations of the ice growth regime while RSL is always represented by a single curve that tends to follow an average interannual regime.

6.2. Comparison of the Retained Local ANN Models With ANN Models With Sigmoid Output Function and Multiple Linear Regression

[48] In this study, the choice of one-hidden-layer neural networks with linear output functions and sigmoid neurons

Table 5. Performance Criteria of RSL and the Retained Local ANN Models[a]

Station	Combination of Meteorological Variables	RMSE (RSL), cm	RMSE (RNA), cm	r^2 (RSL)	r^2 (ANN)	RRMSE (RSL)	RRMSE (ANN)	NASH (RSL)	NASH (ANN)	BIAS (RSL), cm	BIAS (ANN), cm
HA1	10	10.11	5.00	0.98	0.99	0.11	0.05	0.91	0.98	−8.45	−3.65
WFN	5	7.84	7.76	0.85	0.85	0.13	0.13	0.83	0.84	1.55	2.10
WLH	23	5.30	4.71	0.93	0.96	0.11	0.10	0.90	0.92	3.03	−2.82
WTL	9	8.85	7.36	0.93	0.93	0.13	0.11	0.90	0.93	4.54	1.33
YAH	22	7.35	6.99	0.93	0.93	0.14	0.13	0.87	0.88	4.63	4.12
YBK	15	7.59	7.28	0.98	0.99	0.06	0.06	0.98	0.98	0.28	0.21
YBT	9	8.31	7.38	0.88	0.89	0.14	0.12	0.85	0.88	−1.93	0.26
YBX	4	11.82	9.63	0.76	0.83	0.24	0.20	0.71	0.81	4.20	3.32
YEI	10	14.51	11.77	0.88	0.93	0.16	0.13	0.87	0.91	3.72	3.21
YGK	7	7.36	10.27	0.92	0.91	0.27	0.37	0.86	0.72	3.76	−3.35
YGM	9	14.42	6.14	0.64	0.93	0.18	0.08	0.63	0.93	0.83	0.18
YIV	10	8.33	7.34	0.88	0.90	0.15	0.13	0.84	0.87	3.79	1.80
YKL	1	10.18	9.66	0.92	0.92	0.14	0.13	0.90	0.91	3.83	2.47
YPY	11	18.11	17.57	0.69	0.70	0.31	0.30	0.67	0.69	0.32	−2.04
YVP	20	10.80	7.40	0.94	0.96	0.13	0.09	0.90	0.95	−5.45	1.30
YYR	14	9.64	9.15	0.77	0.81	0.14	0.14	0.77	0.79	−0.58	2.61
YZF	10	9.50	8.00	0.93	0.94	0.13	0.11	0.92	0.94	−4.25	−2.19
YFL	5	13.36	7.45	0.93	0.95	0.18	0.10	0.83	0.94	9.45	3.15

[a]Data sets randomly split in 80-20% subsets.

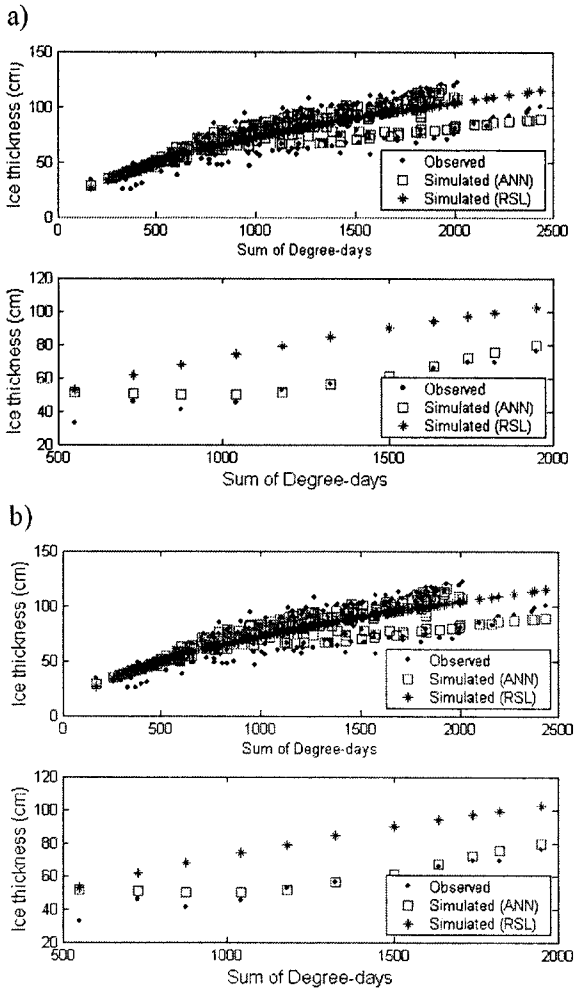

Figure 6. Comparison of local ANN models and RSL at station YGM (leave-one-out procedure) for (a) 1977 and (b) 1980.

on the hidden layer was due to the demonstrated value of this kind of architecture. In theory, the capability of an ANN model to learn the hidden logic in a data set is not supposed to heavily rely on the shape of the activation function unless there is a major incompatibility between the range of the output function and the range of the training data sets (e.g., binary data and continuous activation function for the output neuron). To investigate this point, previously described local ANN models (with sinusoidal activation function for hidden neurons and linear activation function for output neuron) are compared to local ANN models with a sigmoid output activation function for all neurons. The comparison exercise was also performed for multiple linear regression (with an additional intercept parameter). The same procedures for the selection of explanatory variables were applied and the best models were selected for each site. The change of RMSE (negative when the new model is better; positive when the new model is worse) due to the use of these two new models at the ice measurement sites is plotted in Figure 7. The use of a sigmoid activation function resulted in RMSE reductions of up to 10% at some sites, and an increase of RMSE of up to 15% at others (Figure 7, top). In general, the use of a sigmoid activation function for the output neuron gives slightly better results. On the other hand, the use of multiple linear regression always resulted in an increased RMSE (Figure 7, bottom), confirming the non linear nature of the ice growth process.

6.3. Influence of the Data Sets Characteristics on the Best Performing Solutions

[49] The retained combinations of explanatory variables for the local ANN models are very different from site to site. Consequently, it is important to investigate whether the characteristics of the input variables have an effect on the best performing variables. Figure 8a presents the mean snow depth on the ground for the ice measurement stations listed in Table 4. Figure 8b presents the list of sites where the best combination of explanatory variables includes snow. Similar results are presented for rainfall (Figure 9a and 9b) and radiations (Figure 10a and 10b). As for the parameters of the revised Stefan's law, there seems to be no clear relationship between the magnitude of a variable, and whether or not it is chosen as an explanatory variable. Figure 11 presents the maximum ice measurements (Figure 11a) as well as the RMSE of the regional model at these sites (Figure 11b). The RMSE of the regional model seems to be constant all over the territory. The bias of the regional model is presented on Figure 12. It can be seen on Figure 12 that the regional model slightly underestimates ice thicknesses in the northwestern part of the country where ice thickness is generally large. Conversely, it slightly overestimates ice thickness in the southwestern part of the country, where ice thickness is smaller.

6.4. Limited Comparison of Local ANN Models and Deterministic Model CLIMO

[50] The deterministic model CLIMO was used by *Ménard et al.* [2002b] to simulate ice growth at stations

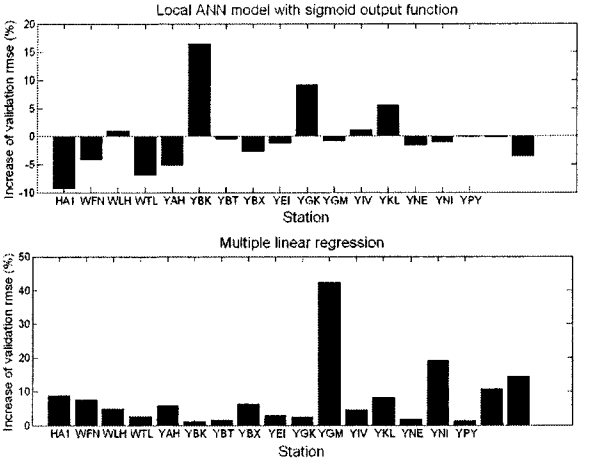

Figure 7. Increase in RMSE when switching from local ANN with sigmoid activation function for hidden neurons and linear activation function for the output neuron to (top) ANN model with sigmoid activation function for all neurons and (bottom) multiple linear regression model.

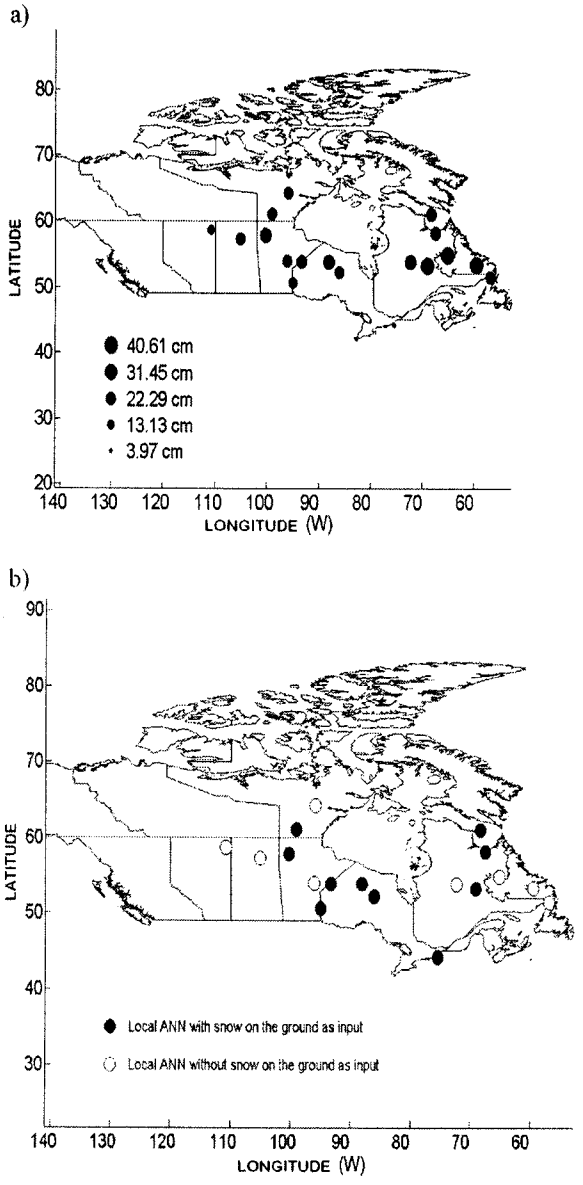

Figure 8. Effect of the statistical characteristics of input variables on the best performing solution: (a) mean snow depth on the ground during the ice growth period and (b) usage of snow depth on the ground as explanatory variable by local ANN models.

YZF (Great Slave Lake–Back Bay) for the years 1960–1991, and at station YFL (Great Slave Lake–Fort Reliance) for years 1977–1990. They obtained an RMSE of 9 cm at YZF and 18 cm at YFL. Since those stations are included in our database, the results of the two studies were compared. It can be seen in Table 4 that RSL gives a RMSE of 13.73 cm (12.48 cm for the local ANN model) using the "leave-one-out" method on the data of the years 1958–1996 and 2002–2003 at station YZF. When simulated on the same period than CLIMO (years 1960–1991) using the leave-one-out procedure, RSL and the local ANN model respectively gave an RMSE of 14.10 and 14.31 cm, which is 13% and 15% higher than CLIMO. Both models performed better than CLIMO when simulating ice growth at station YFL (Fort Reliance) on the data of years 1977–1990. The RMSE was 17.10 cm for revised Stefan's law and 16.78 cm for the local ANN model, slightly lower than the 18 cm obtained with CLIMO. Despite the much higher complexity of the deterministic model, its performance is comparable to those of the local ANNs and RSL.

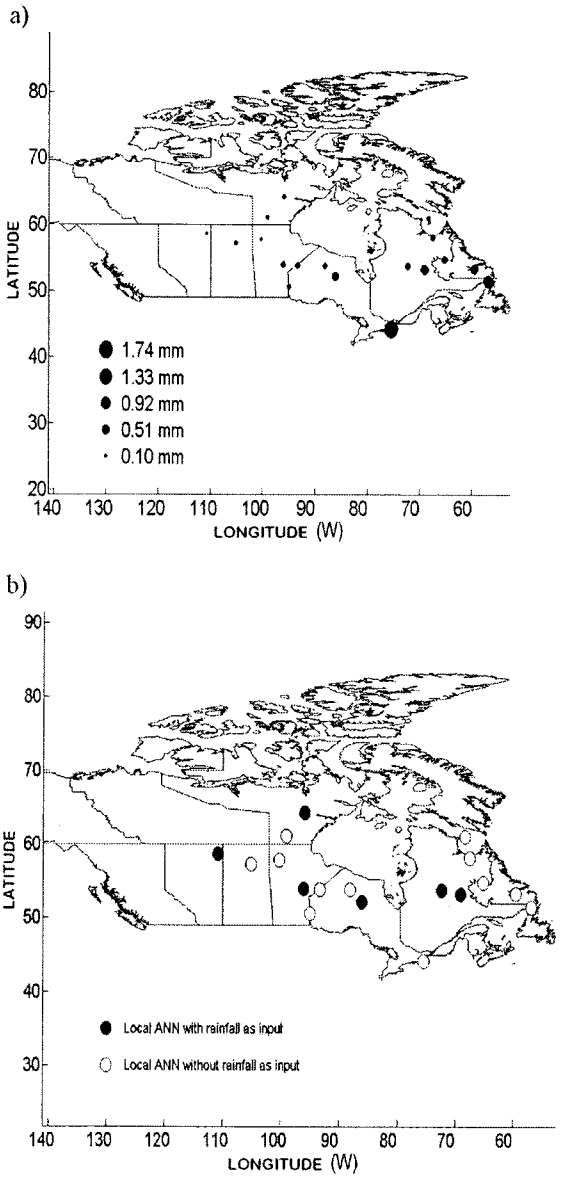

Figure 9. Effect of the statistical characteristics of input variables on the best performing solution: (a) mean daily rainfall during the ice growth period and (b) usage of rainfall as explanatory variable by local ANN models.

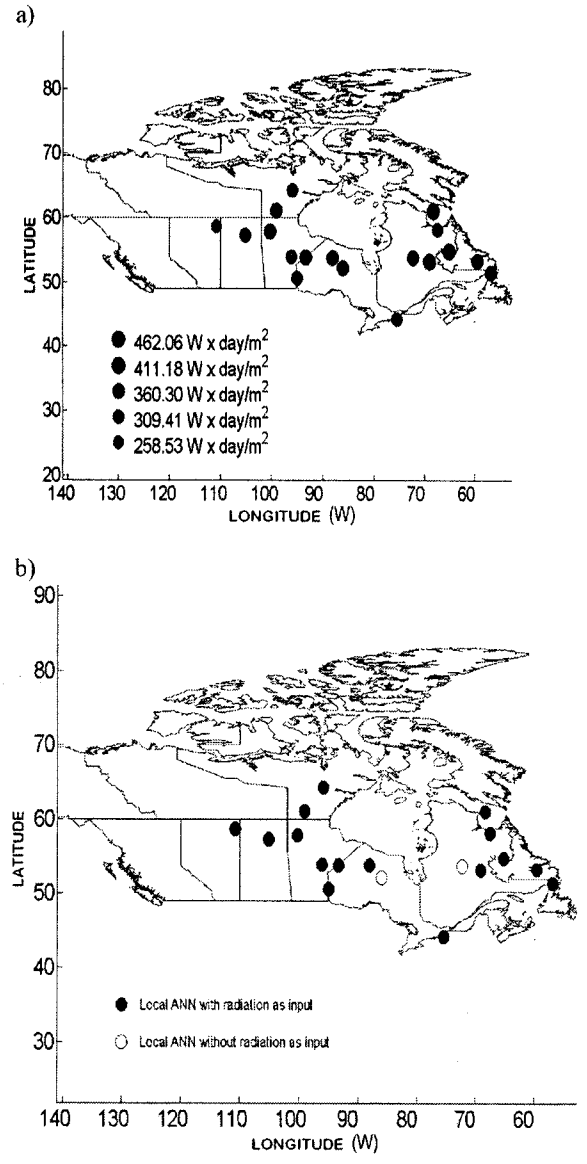

Figure 10. Effect of the statistical characteristics of input variables on the best performing solution: (a) mean sum of solar radiation during the ice growth period and (b) usage of radiation as explanatory variable by local ANN models.

6.5. Regional ANN Model

[51] The performance criteria for the five best combinations for the regional ANN model are given in Table 6. The best combination (in terms of RMSE) to explain ice growth at the regional level is combination 15 (D_g; Rad_{nc}+ Rad_c;$\bar{R}$) followed closely by combination 20 (D_g; Rad_{nc} + Rad_c; $\bar{H}_s$ $\bar{R}$). The optimal number of neurons on the hidden layer was found to be one (as indicated earlier, ANN structures with 1 to 10 neurons on the hidden layer were tested, but only one is retained corresponding to the best performance). The regional model is thus relatively parsimonious since it uses only six parameters: one bias and one weight parameter for each of the connections of the inputs (Rad_{nc}+ Rad_c and $\bar{R}$) to the hidden neuron, and one bias and one weight parameter for the connection of the hidden neuron to the output neuron.

[52] The radiation parameter for the regional model is Rad_{nc} + Rad_c (α = 1), but the choice to split the radiation value into two parts is justified by the fact that for some of the local ANN models, the optimal value of α is lower than 1. For example, for station WTL, the optimal value of α was found to be 0.75 (see Tables 4 and 1).

[53] It was also noticed when rating the different combinations of input variables for the regional ANN model that latitude and longitude do not have an influence on the result. However, geographic location has a direct influence on all input variables. In other terms, geographic location has an

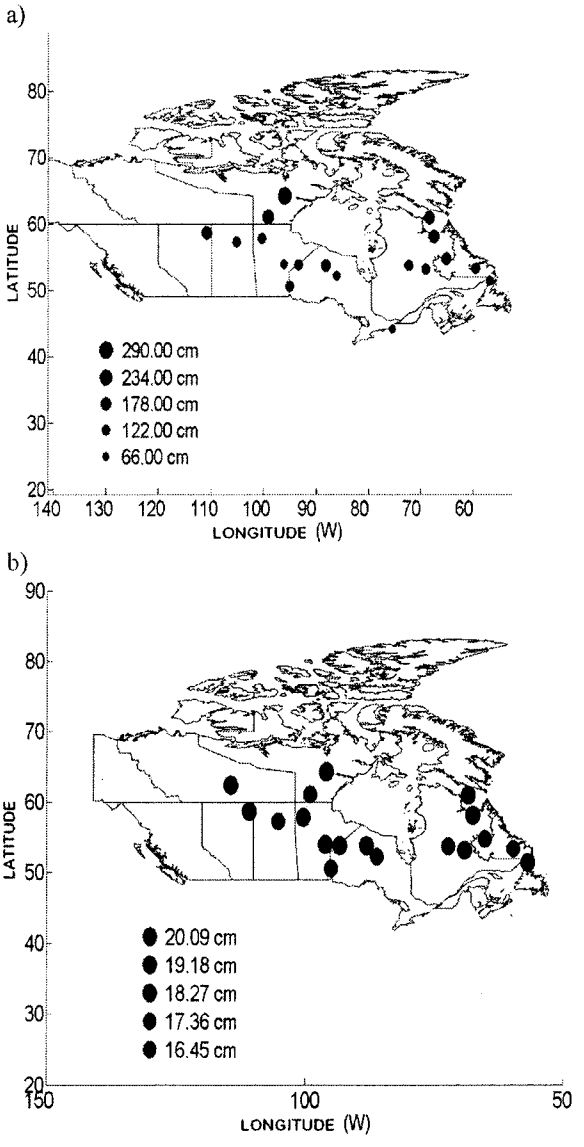

Figure 11. (a) Maximum ice thickness and (b) RMSE of the regional ANN model.

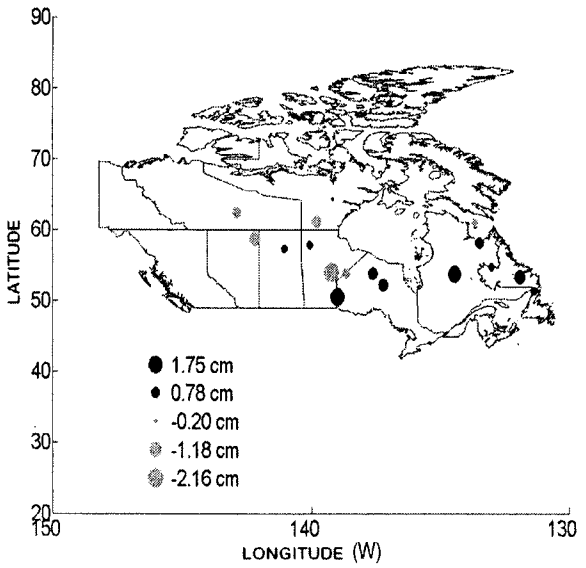

Figure 12. Bias of the regional ANN model.

influence on the range of input variables, but not directly on ice growth relationship to input variables.

[54] An unexpected result in this study is that snow depth on the ground was not identified as an important parameter at the regional level for lake ice growth. This is probably due to the great variability of this parameter in space, which was estimated by interpolation from the weather stations located at several kilometers.

[55] To illustrate the ability of the regional ANN model to reproduce ice growth dynamics simulated ice thickness (using the regional ANN) at station YZF (Great Slave Lake-Back Bay) is presented in Figure 13. It can be seen that the agreement is excellent.

7. Conclusions

[56] It is shown in this paper that artificial neural networks can be valuable alternatives to complex thermodynamic lake ice growth models, especially when data is not available in sufficient quantity and quality. The performances of the proposed ANN models were fairly good when compared to that of the deterministic Canadian Lake Ice CLIMO. They also offered much more flexibility than the Stefan's law and were able to transpose information on ice growth dynamics from monitored sites to unmonitored sites or periods using only six weight and bias parameters. They also reproduced various ice growth patterns with a few easily obtained meteorological data.

[57] The methodology presented in this paper is of the greatest importance for Nordic hydrologists since no operational model of lake ice growth exists for ungauged sites. This gap needed to be addressed for several reasons: a practical model of lake ice growth can help assess the impact of climate change on lake ecosystems, which is a topic of great concern nowadays. It can also be used to estimate the amount of water that is lost as ice left on the banks of hydroelectric reservoirs during winter operation, which may be important for large annual reservoirs in northern countries.

[58] The main limitation of the work presented in this paper is that it applies only to the ice growth period. In other empirical models, a combination of freezing degree-days and melting degree-days have provided good results in modeling the whole duration of the ice cover [e.g., *Thompson et al.*, 2005]. The extension of the developed method to the whole duration of the ice cover using a single ANN model is a challenging problem because of the differences in the two processes to be modeled. The learning performance of the ANN may hence be reduced. A possible alternative solution is to train an ANN model for the growth phase and another one for the decay phase.

Appendix A: Solar Radiation at the Top of Atmosphere

[59] These relations are obtained from the *Solar Radiation Monitoring Laboratory* [2004]. On average extraterrestrial irradiance is 1367 W/m^2. This value varies by ±3% as the earth orbits the sun.

$$I = 1.367\left(\frac{R_{av}}{R_v}\right)^2 * \cos(Z)\ \mathrm{KW/m^2} \qquad \text{(A1)}$$

where R_{av} is the average Sun-Earth distance, R is the actual Sun-Earth distance depending on the day of the year, and Z the zenith (the angle of the sun relative to a line perpendicular to the Earth's surface) which depends on the declination of the Earth, latitude and solar time. If one notes n the Julian day, E_{qt} (equation of time) the variation in minutes of the solar time (T_{solar}) compared to the standard time (T_{local}) during the year, w the solar hour angle (in radians), $Long_{local}$ the longitude of the central meridian line

Table 6. Performance Criteria of the Regional Model for the Five Best Combinations[a]

Combination of Meteorological Variables	RMSE (RSL), cm	r^2 (RSL)	RMSE (ANN), cm	r^2 (ANN)	RRMSE (RSL)	RRMSE (ANN)	NASH (RSL)	NASH (ANN)	BIAS (RSL), cm	BIAS (ANN), cm
15	25.08	0.65	18.15	0.82	0.42	0.43	0.65	0.82	−0.15	0.65
20	25.08	0.65	18.17	0.82	0.42	0.43	0.65	0.82	−0.15	0.56
19	25.08	0.65	18.17	0.82	0.42	0.43	0.65	0.82	−0.15	0.12
14	25.08	0.65	18.19	0.82	0.42	0.43	0.65	0.82	−0.15	0.23
18	25.08	0.65	18.24	0.82	0.42	0.43	0.65	0.81	−0.15	−0.48

[a]One neuron on the hidden layer and performance indices obtained during the leave one out cross-validation procedure. Note longitude and latitude are not listed as explanatory variables because it was found that they have no influence on the output of the regional model.

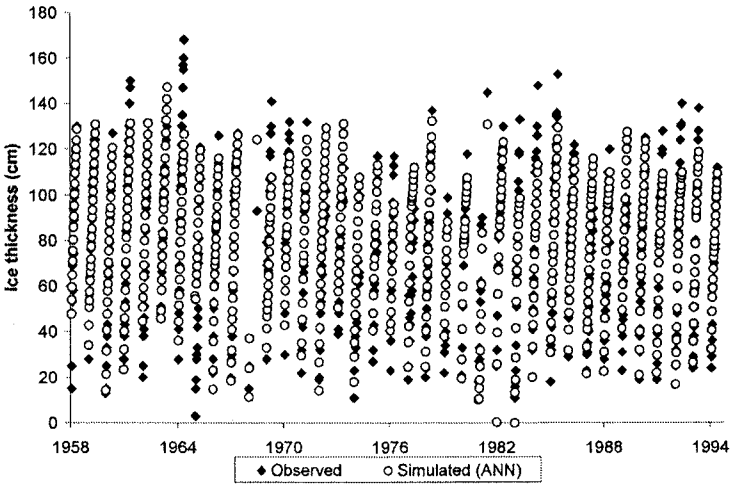

Figure 13. Simulated and observed ice thickness at station YZF.

of the time zone and $Long_{sm}$ the longitude of the point of calculation, we have the following approximations:

$$\left(\frac{R_a}{R_v}\right)^2 = 1.00011 + 0.034221 * \cos\left(2\pi\frac{n}{365}\right) + 0.001280 * \sin\left(2\pi\frac{n}{365}\right) + 0.000719 * \cos\left(4\pi\frac{n}{365}\right) + 0.000077 * \sin\left(4\pi\frac{n}{365}\right) \quad \text{(A2)}$$

$$\cos(Z) = \sin(l)\sin(d) + \cos(l)\cos(d)\cos(w) \quad \text{(A3)}$$

$$d = 23.45 * \sin\left(2\pi\frac{284+n}{365}\right) \quad \text{(A4)}$$

$$w = \pi * \frac{(12 - T_{solar})}{12} \quad \text{(A5)}$$

$$T_{solar} = Tlocal + \frac{Eqt}{60} + (Long_{sm} - Long_{local})/15 \quad \text{(A6)}$$

$$Eqt = \begin{cases} -14.2\sin\left(\pi\frac{n+7}{111}\right) & 1 \le n < 106 \\ 4.0\sin\left(\pi\frac{n-106}{59}\right) & 107 \le n < 166 \\ -6.5\sin\left(\pi\frac{n-166}{80}\right) & 167 \le n < 246 \\ 16.4\sin\left(\pi\frac{n-247}{113}\right) & 247 \le n \le 365 \end{cases} \quad \text{(A7)}$$

Notation

a^m ANN output vector of the mth layer.
B^m ANN bias vector of the connections form the (m-1)th layer to the (m)th layer.
C bias parameter of the modified Stefan's law.
C_a heat transfer coefficient from ice to air.
D_d sum of degree-day from the date of ice cover formation.
D_g sum of degree-day from the first below zero temperature.
f^m ANN transfer function of the mth layer.
H_I ice thickness.
$\bar{H}_I$ mean ice thickness.
H_I^k kth ice thickness in the data set.
$\hat{H}_I^k$ estimate of the kth ice thickness in the data set.
H_S on-ground snow depth.
$\bar{H}_S$ mean on-ground snow depth.
k coefficient of Stefan's law.
$\bar{R}$ average rainfall during the ice growth period.
Rad_c sum of solar radiation of days with precipitation (snow or rainfall).
Rad_{nc} sum of solar radiation of days without precipitation (snow or rainfall).
W^m ANN weight vector of the connections form the (m−1)th layer to the (m)th layer.

[60] **Acknowledgments.** The financial support provided by the Natural Sciences and Engineering Research Council of Canada (NSERC), Hydro-Quebec, the GEOIDE network, and the Northern Study Center (CEN) at Laval University is gratefully acknowledged. This research was partly carried out while the second author was on sabbatical leave at the National Engineering School of Tunis (Tunisia). The authors would like to thank the associate editor and three anonymous reviewers whose comments helped improve the quality of the original manuscript.

References

Beltaos, S. (1995), *River Ice Jams*, Water Resour. Publ., Highlands Ranch, Colo.

Canadian Ice Service (2005), Ice thickness data sets, http://ice-glaces.ec.gc.ca, Ottowa.

Cannon, A. J., and P. H. Whitfield (2002), Downscaling recent streamflow conditions in British Columbia, Canada using ensemble neural network models, *J. Hydrol.*, *259*, 136–151.

Coulibaly, P., and F. Anctil (1999), Real time short term natural waters inflow forecasting using recurrent neural networks, in *Proceedings of International Joint Conference on Neural Networks, 1999. IJCNN '99*, vol. 6, 3802–3805, IEEE Press, Piscataway, N. J.

Coulibaly, P., F. Anctil, and B. Bobée (2000), Daily reservoir inflow forecasting using artificial neural networks with stopped training approach, *J. Hydrol.*, *230*, 244–257.

Cybenko, G. (1989), Approximation by superposition of as sigmoidal function, *Math. Control Signals Syst.*, *2*, 303–314.

Duguay, C. R., G. M. Flato, M. O. Heffries, P. Ménard, K. Morris, and W. R. Rouse (2003), Ice covers variability on shallow lakes at high altitudes: Model simulation and observations, *Hydrol. Processes*, *17*, 3465–3483.

Fang, X., C. R. Ellis, and H. G. Stefan (1996), Simulation and observation of ice formation (freeze-over) in a lake, *Cold Reg. Sci. Technol.*, *24*, 129–145.

Flato, G. M., and R. Brown (1996), Variability and climate sensitivity of land fast arctic sea ice, *J. Geophys. Res.*, *101*, 25,767–25,777.

Hagan, M. T. (1996), *Neural Network Design*, PWS, Boston.

Hewett, R. (2003), Data mining for generating predictive models of local hydrology, *Appl. Intell.*, *19*, 157–170.

Hornik, K., M. Stinchcombe, and H. White (1989), Multilayer feedforward neural networks are universal approximators, *Neural Networks*, *2*, 359–366.

Lenormand, F., C. R. Duguay, and R. Gauthier (2002), Development of a historical ice database for the study of climate change in Canada, *Hydrol. Processes*, *16*, 3707–3722.

Liang, R. H., and Y. Y. Hsu (1994), Scheduling of hydroelectric generation using artificial neural networks, *IEE Proc., Part C*, *145*, 452–458.

Lock, G. S. H. (1990), *The Growth and Decay of Ice*, Cambridge Univ. Press, New York.

MacKay, D. J. C. (1992), Bayesian interpolation, *Neural Comput.*, *4*, 415–447.

MacKay, D. J. C. (1995), Bayesian non-linear modeling with neural networks, report, Cambridge Programme for Ind., Cambridge, U. K.

Matousek, V. (1984), Regularity of the freezing up of the water surface and heat exchange between water body and water surface, paper presented at IAHR International Symposium on Ice, Int. Assoc. of Hydraul. Eng. and Res. Hamburg, Germany.

Ménard, P., C. R. Duguay, G. M. Flato, and W. R. Rouse (2002a), Simulation of ice phenology on Great Slave Lake, Northwest Territories, Canada, *Hydrol. Processes*, *16*, 3691–3706.

Ménard, P., C. R. Duguay, G. M. Flato, and W. R. Rouse (2002b), Simulation of ice phenology on a large lake in the Mackenzie River basin, *Proc. Annu. Meet. Eastern Snow Conf.*, *59th*(3), 12.

Michel, B. (1971), *Winter Regime of Rivers and Lakes*, Cold Reg. Res. and Eng. Lab., Hanover, N. H.

Olsson, J., C. B. Uvo, and K. Jinno (2001), Statistical atmospheric downscaling of short-term extreme rainfall by neural networks, *Phys. Chem. Earth, Part B*, *26*, 695–700.

Omstedt, A. (1985a), On supercooling and ice formation in turbulent seawater, *J. Glaciol.*, *31*, 272–280.

Omstedt, A. (1985b), Modeling frazil ice and grease ice formation in the upper layers of the ocean, *Cold Reg. Sci. Technol.*, *11*, 87–98.

Ouarda, T. B. M. J., H. Gingras, S. Hamilton, H. Ghedira. and B. Bobée (2003), Estimation of streamflow under ice, paper presented at the 12th Workshop on the Hydraulics of Ice Covered Rivers, Committee on River Ice Processes and the Environment, Canadian Geophysical Union (CGU-HS), Edmonton, 19–21 June.

Schulyakovskii, L. G. (Ed.) (1966), *Manual of Ice-Formation Forecasting for Rivers and Inland Lakes*, Isr. Prog. for Sci. Transl., Jerusalem.

Shamseldin, A. Y. (1997), Application of neural network technique to rainfall-runoff modeling, *J. Hydrol.*, *199*, 272–294.

Shen, H. T., and L. A. Chiang (1984), Simulation of growth and decay of river ice cover, *J. Hydraul. Eng.*, *110*, 958–971.

Shen, H. T., and C. F. Ho (1986), Two-dimensional simulation of ice cover formation in a large river, paper presented at IAHR Ice Symposium, Int. Assoc. of Hydraul. Eng. and Res., Iowa City, Iowa.

Shen, H. T., H. H. Shen, and S. M. Tsai (1990), Dynamic transport of river ice, *J. Hydraul. Res.*, *28*, 659–671.

Shen, H. T., D. S. Wang, and A. M. W. Lal (1995), Numerical simulation of river ice processes, *J. Cold Reg. Eng.*, *107*, 107–118.

Solar Radiation Monitoring Laboratory (2004), Solar radiation basics, report, Univ. of Oreg., Corvallis. (Available at http://solardat.uoregon.edu/SolarRadiationBasics.html).

Stefan, H. G., and X. Fang (1997), Simulated climate change effects on ice and snow covers on lakes in a temperate region, *Cold Reg. Sci. Technol.*, *25*, 137–152.

Svensson, U., and A. Omstedt (1994), Simulation of supercooling and size distribution of frazil ice dynamics, *Cold Reg. Sci. Technol.*, *22*, 221–233.

Svensson, U., L. Billfalk, and L. Hammar (1989), A mathematical model of border-ice formation in rivers, *Cold Reg. Sci. Technol.*, *16*, 179–189.

The Mathworks (2005), Neural Network Toolbox, user's guide, version 5, Natick, Mass. (Available at http://www.mathworks.com/access/helpdesk/help/pdf_doc/nnet/nnet.pdf)

Thompson, R., et al. (2005), Quantitative calibration of remote mountain-lake sediments as climatic recorders of air temperature and ice-cover duration, *Arct. Antarct. Alp. Res.*, *37*(4), 626–635.

Xiao, R., and V. Chandrasekar (1997), Development of a neural network based algorithm for rainfall estimation from radar observation, *IEEE Trans. Geosci. Remote Sens.*, *35*, 160–171.

Zhang, S. P., H. Watanabe, and R. Yamada (1994), Prediction of daily water demand by neural networks, in *Stochastic and Statistical Methods in Hydrology and Environmental Engineering*, vol. 3, edited by K. W. Hipel et al., pp. 217–227, Springer, New York.

L. Bilodeau and P. Bruneau, Hydro-Québec, 855 Ste-Catherine Street East, 12th Floor, Montreal, QC, Canada H2L4P5.

M. Hessami, T. B. M. J. Ouarda, O. Seidou, and A. St-Hilaire, INRS-ETE, 490 rue de la Couronne, Quebec, QC, Canada G1K 9A9. (ousman_seidou@ete.inrs.ca)