Modeling of River Water Temperatures using Feed-forward Artificial Neural Networks

Cindie Hébert[1], Daniel Caissie[2], Mysore G. Satish[1] and Nassir El-Jabi[3]
[1]*Faculty of Engineering, Dalhousie University, Halifax, NS, B3J 2X4, Canada*
[2]*Fisheries and Oceans, Moncton, NB, E1C 9B6, Canada*
[3]*Faculty of Engineering, Université de Moncton, Moncton, NB, Canada*

Keywords: River/Streams, Modeling, Temperature, Artificial Neural Network.

Abstract: Water temperature influences most physical, chemical and biological processes of the river environment. It plays an important role in the distribution of fishes and on the growth rates of many aquatic organisms. It is therefore important to develop water temperature models in order to effectively manage aquatic habitats, to study the thermal regime of rivers and to have effective tools for environmental impact studies. The objective of the present study was to develop a water temperature model based on artificial neural networks (ANN) for two thermally different watercourses. The ANN model performed best in summer and autumn and showed a poorer (but still good) performance in spring. The many advantages of ANN models are their simplicity, low data requirements, their capability of modelling long-term series as well as have an overall good performance.

1 INTRODUCTION

Water temperature has both economic and ecological significance when considering issues such as water quality and biotic conditions in rivers (Caissie, 2006). As such, fish habitat suitability is highly dependent on stream water temperatures. It is therefore important to use adequate water temperature modeling approaches to effectively predict water temperature variability.

Since the 1990's, artificial neural networks (ANN) have been widely used in the field of hydrology, namely in modeling of precipitation and runoff, water demand predictions, groundwater, and water quality modeling (Govindaraju, 2000). One of the main reasons was the fact that ANN has the capacity to recognize relations between input and output variables without necessarily requiring any physical explications. This approach can be very useful in hydrology because most relationships are non-linear, very complex, and sometimes unknown. Although ANNs have been applied in many hydrological studies in recent decades, very few of these studies have dealt with the modeling of river water temperatures (Risley et al., 2003); (Bélanger et al., 2005); (Sivri et al., 2007); (Chenard and Caissie, 2008), especially at the hourly time step (Risley et al., 2003).

Therefore, the objective of this component of the study was to develop an ANN model to predict hourly river water temperatures using minimal and accessible input data. This model was applied to two thermally different watercourses and its performance was compared to other water temperatures models.

2 METHODOLOGY

2.1 Study Area

The two study sites were located on the Miramichi river system (New Brunswick, Canada), which is world renowned for its population of Atlantic salmon. The first study site was located on the Little Southwest Miramichi River (LSWM) at approximately 25 km from the river mouth. The drainage area of this basin is 1190 km^2 (Johnston, 1997). The LSWM has a river width of approximately 80 m, with a depth of 0.55 m on average during mean flow conditions. No lateral variation of water temperatures were observed due to the well-mixed nature of the river (Caissie et al., 2007). The canopy closer was less than 20%.

The second study site was located on Catamaran Brook (Cat Bk) approximately 8 km upstream of the

Hébert C., Caissie D., G. Satish M. and El-Jabi N..
Modeling of River Water Temperatures using Feed-forward Artificial Neural Networks.
DOI: 10.5220/0004158005580562
In *Proceedings of the 4th International Joint Conference on Computational Intelligence* (NCTA-2012), pages 558-562
ISBN: 978-989-8565-33-4

mouth. It is the site of a 15-year multidisciplinary hydrobiological research study aimed at quantifying stream ecosystem processes and the impact of timber harvesting (Cunjak, Caissie and El-Jabi, 1990). Catamaran Brook has a drainage area of 27 km^2 at the study site, an average stream width of 9 m and a depth of 0.21 m. Catamaran Brook is well-mixed due to high turbulence, similar to LSWM, but the brook is more sheltered by streamside vegetation and upland slopes. The canopy closer for Catamaran was estimated at 55%-65%.

2.2 Water Temperature Model

Water temperature data for the ANN were collected for the period of April 15 (day 105) to October 31 (day 304) and for years between 1998 and 2007 at both CatBK and LSWM. This period corresponded approximately to the period of the year without ice cover, i.e., open water condition. Some years had missing data for a few days and these days were not included in the ANN model. Data were separated into two samples: training data (1998-2002) and validation data (2003-2007).

The developed hourly ANN model of this study used six input nodes: air temperature (°C) of the present and previous hour, time of day (hour), the time of year (day), daily mean water temperature (simulated) (°C), and the mean daily water level (m). The selection of air temperature, as input data, was based on the availability of data and their strong correlation to water temperatures (Cluis, 1972); (Song and Chien, 1977); (Stefan and Preud'Homme, 1993); (Mohseni and Stefan, 1999); (Bélanger et al., 2005); (Chenard and Caissie, 2008). Daily mean water temperatures were first predicted from a daily ANN model (using air temperature (°C) of the present and previous day (°C), daily water level (m), and time of year (day)). The air temperature of the previous day was used as input because air and water temperature are strongly correlated (Kothandaraman, 1971); (Cluis, 1972). These daily mean water temperatures were then used as input data into the hourly ANN water temperature model. During the training, the observed daily mean water temperatures were used; however during the validation the simulated daily mean water temperatures were used to simulate the hourly temperatures. The output of the developed ANN model was hourly water temperature at both Cat Bk and LSWM.

The feed-forward backpropagation ANN model was created using Matlab Student 7.1. For the application within the present study, the supervised learning process was used. The ANN model was adjusted for the minimum difference between predicted and observed water temperatures. The ANN model achieved optimal six input nodes, five hidden nodes in one hidden layer and only one output node. The transfer function (f(n)) used between each node was the hyperbolic tangent sigmoid transfer function, described as follows:

$$f(n) = \frac{2}{\left(1 + e^{(-2n)}\right)} - 1 \tag{1}$$

This function also represents well the non-linear processes usually found in hydrology (Smith, 1993); (Jain et al., 1996).

2.3 Modeling Performance Criteria

Three criteria were used to compare modeling performances: the root-mean-square error (RMSE), the coefficient of determination (R^2), and the bias (Bias). They were selected because they are often used in modeling studies and results from these performance criteria were also available for other water temperature models at Cat Bk and LSWM. The root-mean-square error (RMSE) represents the mean errors associated to the model. The coefficient of determination (R^2) represents the percentage of variability that can be explained by the model. The bias is an indication of the overestimation or underestimation of the water temperature model and represents the mean of errors.

3 RESULTS

Results of the ANN models (RMSE, R^2, and bias) are represented in Table 1. The ANN model generally provided the best results at Cat Bk with a root-mean-square error (RMSE) of 0.63°C for the training and 1.19°C for the validation period. At Cat Bk, the coefficient of determination (R^2) was 0.986 (training) and 0.948 (validation). The bias was at 0.00°C for the training period and -0.28°C for the validation period. For the LSWM, the ANN model performed comparably well, especially during the training (RMSE = 0.69°C and R^2 = 0.989). However, during the validation period, the RMSE was higher at 1.62°C and a correspondingly lower R^2 at 0.930. The bias for LSWM was 0.00°C (training) and 0.05°C (validation). Overall (all years), the ANN model performed well for both watercourses with RMSE of 0.94°C (Cat Bk) and 1.23°C (LSWM) and with R^2 of 0.967 (Cat Bk) and

0.962 (LSWM). Water temperatures were slightly underestimated at Cat Bk with bias of –0.13°C and the overall bias for LSWM was very low (0.02°C).

Table 1: Results of the ANN water temperature models.

	Training (1998- 2002)	Validation (2003-2007)	All years (1998-2002)
Cat. Bk			
RMSE	0.63	1.19	0.94
R^2	0.986	0.948	0.967
Bias	0.00	-0.28	-0.13
LSWM			
RMSE	0.69	1.62	1.23
R^2	0.989	0.930	0.962
Bias	0.00	0.05	0.02

Table 2: Results of the seasonal analysis.

	Training (1998-2002)	Validation (2003-2007)	All years (1998-2002)
Spring			
Cat. Bk			
RMSE	0.70	1.38	1.06
R^2	0.979	0.920	0.951
Bias	0.01	-0.02	-0.01
LSWM			
RMSE	0.85	1.76	1.38
R^2	0.979	0.922	0.947
Bias	0.02	0.78	0.39
Summer			
Cat. Bk			
RMSE	0.64	1.02	0.85
R^2	0.942	0.865	0.901
Bias	0.00	-0.32	-0.16
LSWM			
RMSE	0.67	1.61	1.23
R^2	0.961	0.776	0.868
Bias	-0.02	-0.20	-0.11
Autumn			
Cat. Bk			
RMSE	0.47	1.25	0.94
R^2	0.979	0.856	0.915
Bias	0.00	-0.53	-0.27
LSWM			
RMSE	0.52	1.39	1.00
R^2	0.985	0.890	0.943
Bias	0.02	-0.47	-0.19

Table 2 shows the performance of the model on a seasonal basis. Spring was between April 15 and June 21 (day 105-171), summer between June 22 and September 20 (day 172-263) and autumn between September 22 and October 31 (day 264-305). For the training period, autumn showed the best performance with a RMSE of 0.47°C (Cat Bk) and 0.52°C (LSWM). Spring (training period) showed a poorer performance with RMSE of 0.70°C (Cat Bk) and 0.85°C (LSWM). RMSEs during the summer were similar at Cat Bk and LSWM with values of 0.64°C and 0.67°C. Coefficients of determination (R^2) were similar in autumn and spring with values over 0.979; however, lower values were observed in summer (0.942-0.961). The biases were generally small for both watercourses for the training period with seasonal values less than ±0.02°C.

Seasonal results were similar during the validation period, although RMSEs and biases were generally higher with lower R^2. Highest RMSEs were observed during the spring (1.38°C Cat Bk and 1.76°C LSWM) and best performances were in summer in Cat Bk (1.02°C) and autumn in LSWM (1.39°C). Summer had the lowest R^2 (0.776), whereas spring had the highest R^2 (0.922). Spring showed a general overestimation of predicted water temperature in LSWM with a bias of 0.78°C. In general (all years), the ANN model showed similar seasonal performances in Cat Bk and a better performance in summer and autumn for LSWM.

4 DISCUSSION

Most ANN models have estimated daily mean water temperatures. The modeling of hourly stream water temperature in this study was found to be as good as the modeling of daily mean stream water temperatures. For example, Chenard and Caissie (2008), who modeled daily mean stream temperatures in Catamaran Brook using an ANN, achieved similar results with overall RMSE of 0.96°C and R^2 of 0.971. Bélanger et al. (2005) calculated an overall RMSE of 1.06°C when applying an ANN model at Catamaran Brook (daily mean temperatures). The study by Bélanger et al. (2005) used only air temperature and water level as input parameters. Risley et al. (2003) have developed a more complex ANN model for 148 sites in western Oregon on a short-term period (June 21 to September 20, 1999). Three different ANN models were developed to estimate hourly water temperatures along 1st, 2nd, and 3rd order streams using meteorological data (air temperature, dewpoint temperature, short-wave solar radiation, air pressure, and precipitation), riparian habitat characteristics (stream bearing, gradients, depth, substrate, wetted widths, and canopy cover), and basins landscape characteristics (topographic and vegetative), acquired by using a geographic information system (GIS). Their results showed RMSEs ranging between 0.05°C and 0.59°C and with R^2 ranging from 0.88 to 0.99.

Comparison of seasonal performance showed that the ANN model performed best in summer or autumn, which is consistent with other temperature models (Caissie et al., 1998); (Caissie et al., 2005); (Chenard and Caissie, 2008). It could suggest the potential role of discharge in the modeling performance, as low water levels are usually observed in autumn and mid-summer, resulting in more effective thermal exchange and giving better performances. The poorer but still good performance in spring could be explained by the higher discharge caused by snowmelt, resulting in a poorer air to water temperature relationship (Caissie et al., 1998). The performance of the model was closely linked to water levels, meaning that the performance was better when water levels were low. At LSWM, the ANN model performed best in autumn for all the years, whereas at CatBk, some years had their best performance during summer. These results suggest that the thermal exchange is more efficient for less sheltered river under low flow (autumn at LSWM). CatBk is more sheltered and could potentially be influenced by other factors (ex., groundwater) reducing the efficiency of the thermal exchange. For example, Hébert, Caissie, Satish, and El-Jabi (2011) showed that the impact of groundwater on hourly water temperatures was more significant on smaller streams, like CatBk.

The training period showed better results than the validation period, which is consistent in modeling. Daily water levels used in the modeling were estimated using power functions (Caissie, 2004). Using hourly water levels instead of daily water levels could potentially improve the modeling, especially during days that discharge varied significantly. However, hourly water levels were not available for the present study.

ANN models have major advantages over more commonly used water temperature models, as they do not need many input data. In this case, only air temperature and water level time series were used to achieve good predictions. For instance, deterministic model needs many hydrological and meteorological parameters that are not always readily available (e.g., solar radiation). Another major advantage of ANN is that they are easy to use and very simple in their application. However, ANN models cannot give any physical explanation of the relationship between the input and output data. These models should therefore be used with caution, especially when using input data that are outside the range of the calibration period (Risley et al., 2003).

5 CONCLUSIONS

This study showed that artificial neural network (ANN) could be an effective tool for the prediction of hourly stream temperatures. ANN models achieved comparable or better performances to other water temperature models reported in the literature, with RMSE of 0.94°C at Cat Bk and 1.23°C at LSWM. ANN models showed a good generalization capability by modeling well water temperature time-series. ANN was effectively applied on two thermally different streams and provides similar results and performances. As such, ANN models can be considered as effective modeling tool in water resources and fisheries management.

REFERENCES

Bélanger, M., El-Jabi, N., Caissie, D., Ashkar, F., Ribi, J.M. (2005). Estimation de la température de l'eau en rivière en utilisant les réseaux de neurones et la régression multiple. *Revue des Sciences de l'Eau,* 18(3), 403-421.

Caissie, D. (2004). *Stream temperature modeling in forest catchments*. PhD Thesis, Dalhousie University, Halifax, NS, 207p.

Caissie, D. (2006). The thermal regime of rivers: a review. *Freshwater Biology*, 51, 1389-1406.

Caissie, D., El-Jabi, N., St-Hilaire, A. (1998). Stochastic modelling of water temperatures in a small stream using air to water relations. *Canadian Journal of Civil Engineering,* 25, 250-260.

Caissie, D., Satish, M. G., El-Jabi, N. (2005). Predicting river water temperatures using the equilibrium temperature concept method with applications on Miramichi River Catchments (New Brunswick, Canada). *Hydrological Processes,* 19, 2137-2159.

Caissie, D., Satish, M. G., El-Jabi, N. (2007). Predicting water temperatures using a deterministic model: Application on Miramichi River catchments (New Brunswick, Canada). *Journal of Hydrology,* 336(3-4), 303-315.

Chenard, J., Caissie, D. (2008). Stream temperature modelling using artificial neural networks: application on Catamaran Brook, New Brunswick, Canada. *Hydrological Processes*, 22(17), 3361-3372.

Cluis, D. A. (1972). Relationship between stream water temperature and ambient air temperature - a simple autoregressive model for mean daily stream temperature fluctuations. *Nordic Hydrology,* 3, 1025-1031.

Cunjak, R. A., Caissie, D., El-Jabi, N. (1990). The Catamaran Brook Habitat Research Project: description and general design of study. *Canadian Technical Report of Fisheries and Aquatic Sciences,* 1751, 14.

Govindaraju, R. S. (2000). Artificial neural networks in hydrology. I: Preliminary concepts. *Journal of hydrologic Engineering,* 5(2), 115-123.

Hebert, C., Caissie, D., Satish, M. G., El-Jabi, N. (2011). Study of stream temperature dynamics and corresponding heat fluxes within Miramichi River catchments (New Brunswick, Canada). *Hydrological Processes*, 25, 2439-2455.

Jain, A. K., Mao, J., Mohiuddin, K. M. 1996. Artificial neural network: A tutorial. *Computer*, 31-44.

Johnston, T. A. (1997). Downstream movement of young-of-the-year fishes in Catamaran Brook and the Little Southwest Miramichi River, New Brunswick. *Journal of Fish Biology,* 51, 1047-1062.

Kothandaraman, V. (1971). Analysis of water temperature variations in large rivers. *Journal of the Sanitary Engineering Division,* 97, 19-31.

Mohseni, O., Stefan, H. G. (1999). Stream temperature-air temperature relationships: a physical interpretation. *Journal of Hydrology,* 218, 128-141.

Risley, J. C., Roehl, E. A. Jr., Conrads, P. A. (2003). Estimating water temperatures in small stream in Western Oregon using neural network models*. U.S. Geological Survey (USGS) Water-Resources Investigations Report* 02-4218, 59p.

Sivri, N., Kilic, N., Ucan, O. (2007). Estimation of stream temperature in Firtinia Creek (Rize-Turkiye) using artificial neural network model. *Journal of Environmental Biology*, 28(1), 67-72.

Smith, M. 1993. *Neural networks for statistical analysis*. Van Nostrand Reinhold, New York, 235p.

Song, C. C. S., Chen, C. Y. (1977). Stochastic properties of daily temperature in rivers. *Journal of the Environmental Engineering Division,* 103, 217-231.

Stefan, H. G., Preud'homme, E. B. 1993. Stream temperature estimation from air temperature. *Water Resources Bulletin,* 29, 27-45.