American Journal of Civil and Environmental Engineering
2018; 3(4): 87-95
http://www.aascit.org/journal/ajcee

American Association for Science and Technology

Water Temperature Forecasting in a Small Stream Using Neural Networks with a Bayesian Regularization Technique

Mohamed Nohair, Mohssine El Marrakchi, El Mati Khoumri, Chaymae Jermouni, Sara Azmi

Laboratory of Physical Chemistry & Bioorganic Chemistry, Energy, Interfacial Electrochemistry and Chemometrics, Faculty of Sciences and Technics, Mohammedia, Morocco

Email address

nohairmohamed@yahoo.fr (M. Nohair)

Citation

Mohamed Nohair, Mohssine El Marrakchi, El Mati Khoumri, Chaimae Jermouni, Sara Azmi. Water Temperature Forecasting in a Small Stream Using Neural Networks with a Bayesian Regularization Technique. *American Journal of Civil and Environmental Engineering.* Vol. 3, No. 4, 2018, pp. 87-95.

Received: July 11, 2018 ; **Accepted**: August 2, 2018; **Published**: September 3, 2018

Abstract: Physical processes influencing water temperature in a river are highly complex and uncertain, which makes it difficult to capture them in some form of deterministic model. Accurate forecasting of water temperature in a river is important, as it has implications on the quality of water and the lives that depend on it. Here we develop a model of forecasting which allows estimation and forecasting of water temperature at short and middle term, It intends to forecast the water temperature of days (*t+i, i=1,2…*), t is the current time. Due the strong dependence between water temperature at the current time and those for the past, the projected model builds easily itself, by investigating, for each stage of forecasting, the function relating input and output relationships. For this, a multi-step-ahead forecasting model based on the neural networks with the Bayesian regularization technique, is formulated for establishing linkages between water temperature and influencing variables. The results show that the elaborated model is robust and reliable and gives good results. It allows us to forecast the water temperature with high success. To test the ability of the model for the prediction, the observed data of the average daily water temperature during a period of five years (1998-2002) is considered for analysis. The first three years serve for the training and the remaining for the test. The model produced a standard coefficient R about 98, while the standard deviation s does not exceeds 0.6°C. We noticed there are a few cases presenting an error between 1 and 1.5°C (On average three cases for all steps of forecasting).

Keywords: Water Temperature, Small Stream, Forecasting, Neural Networks Analysis, Bayesian Regularisation

1. Introduction

The forecast problem in hydrological sciences

The forecast is defined as the estimation of the future conditions of the phenomena for the given period, from the past and current observations. The general objective is to supply the best estimations of what can happen from a given point to a future precise date. For example, if *t* indicates the beginning time of the process, the calculation of the forecast of a process defined beforehand by the function *f* consists in readjusting this last one to take into account steps of forecast. These steps can be units of measure of time (Hour, day….). Due to the strong correlation between the explanatory variable of two consecutive steps, the projected model builds itself easily in the following way:

$$\begin{cases} \hat{y}(t) = f(var\ 1, var\ 2, var\ 3 \dots\dots var\ n) \\ \hat{y}(t+1) = f(\hat{y}(t), var\ 1, var\ 2, var\ 3 \dots\dots var\ n) \\ \hat{y}(t+2) = f(\hat{y}(t+1), \hat{y}(t), var\ 1, var\ 2, var\ 3 \dots\dots var\ n) \\ \vdots \\ \hat{y}(t+n) = f(\hat{y}(t+n-1) \dots\dots (\hat{y}(t+n-k) \dots\dots \hat{y}(t+1), \hat{y}(t), var\ 1, var\ 2, var\ 3 \dots\dots var\ n) \end{cases} \tag{1}$$

Where vari represents the explanatory variable at the current time t, and $\hat{y}$ represents estimated response of the hydrological process, which is the water temperature in this study.

Thus for every stage, we formulate a function f_i, by adding to every stage the previous dependent variable in the explanatory variables. We thus build simply many sets of data as steps of forecasts.

Generally, the characteristic elements of the forecast are presented as follow:

1. Variable to be planned and the explanatory variables,
2. Horizon of forecast (i = 1, 2,… k),
3. Methods of calculation or estimation (that is the nature of the function *f*),
4. Objective of the forecast (alert of floods, Rain-fall-runoff modelling, planning of the operation of reservoirs, the projects of irrigation, water quality…),
5. Type of wished results (numerical, graphic values, or distribution of probability).

The consideration of all these elements constitutes the problem of the forecast in the environment [1-6]. Two types of approaches can be used in the forecast: the determinist approach and the empirical approach (or stochastic). The determinist approach is based on the physical simulation of the system. The determinist model supposes that an exact calculation of the parameters is possible. But often, this hypothesis is not realistic, because the natural phenomena are very unpredictable. The determinist models are ultimately limited by the large number of parameters to be measured and by the limits of the current knowledge of the complex natural systems. On the other hand, the stochastic approach allows ignoring the limits of the physical knowledge of the system. Contrary to the determinist models, they work as "black boxes", that is without any consideration of the internal structure of the system. Nevertheless, there are techniques which allow limiting the defects of this problem. Among them, neural network technique [7] is known to be very effective in representing the relationships, which could exist between variables in complex systems. For this, it has been largely employed to hydrology application [8]. Its applications concern: the classification of the hydrological data [9], flow forecasting of rivers and alert of floods [10-11], the quality of the water [12-13], forecasting of water consumption in urban domain [14], the estimation of the precipitation-runoff [15], forecasting the natural contributions in the reservoirs for irrigation or hydroelectric production [16], and forecasting the water temperature in a river [17].

The focus of the present work is to define a model that allows forecasting water temperature in a small stream, based on the neural networks with the Bayesian regularization technique [18-19].

2. Neural Networks Analysis and Bayesian Regularization Technique for Analysis

Regression method based on neural network [7] is widely used for its ability to model complex non-linear relationships without any prior assumptions about the nature of the relationships, and this represents its greatest advantage. The solution of a problem is not explicitly encoded in the network, but is learned by supplying examples of previously solved problems to the network. After the network has learned to solve the example problems, it is said to be trained. New data from the same knowledge domain then can be input to the trained neural network which then outputs a solution.

From a practical point of view, an artificial neural network is simply a computer program that transforms an *m*-variable input into an *n*-variable output. Artificial neural networks (ANNs) appear to be very promising in obtaining models that convert structural features into different properties of process. The computational neural network used in this study was a three-layer (input-hidden-output), fully connected, feed-forward network [21]. The input layer contains one node for each variable. The output layer has one node generating the scaled estimated value of the water temperature. Although there are neither theoretical nor empirical rules to determine the number of hidden layers or the number of neurons in this layer, one layer seems to be sufficient in most applications of ANNs. Networks with biases, a sigmoid layer, and a linear output layer are capable of approximating any function with a finite number of discontinuities [7]. To do this, input vectors and the corresponding target vectors are used to train a network until it can approximate a function relating between them in a training phase.

2.1. Training Phase

The training phase was realized by using the standard back-propagation method [22]. It is a generalizion of the Widrow-Hoff learning rule [23] for multiple-layer networks and nonlinear differentiable transfer functions. Standard back-propagation is a gradient descent algorithm, in which the network weights are moved along the negative of the gradient of the performance function *J* (equation (2)). The term back-propagation refers to the manner in which the gradient (equation (3) & equation (4)) is computed for nonlinear multilayer networks.

$$J(w) = \sum_{k=1}^{N} (y^k - g(x^k, w))^2 = \sum_{k=1}^{N} J^k(w) \qquad (2)$$

The function *J*

y: observed target, *g(x,w)*= calculated target

$$\left(\frac{\partial J^k}{\partial w_{ij}}\right)=\left(\frac{\partial J^k}{\partial v_i}\right)\left(\frac{\partial vi}{\partial w_{ij}}\right)_k=\delta_i^k x_j^k \tag{3}$$

The Gradient of the function J

And

$$\delta_j^k=\left(\frac{\partial J^k}{\partial v_i}\right)=-2g(x^k,w)\left(\frac{\partial g(x,w)}{\partial v_i}\right)_k \tag{4}$$

The partial gradient related to neuron i

The same approach is applied to hidden neurons; the parameters are modified by following the learning rule based on the back-propagation method (equation (5))

$$w(i)=w(i-1)-\mu_i \nabla J(w(i-1)) \tag{5}$$

μ_i and i represent respectively the pass and iteration training.

The weights of connections between the neurons were initially assigned random values uniformly. During training, the weights and biases of the network are iteratively adjusted to minimize the network performance function *J*. The training was followed by examining the RMS error (RMS stands for root mean square, that is the square root of the average residual) for the total set. Training was stopped when there was no further improvement in the test set RMS error. We also computed both the correlation coefficient and the standard deviation between the observed and predicted values (equation (6)):

$$R^2=1-\frac{\sum(y_{obs}-y_{calc})}{\sum(y_{obs}-y_{mean})}$$

$$s^2=\frac{\sum(y_{obs}-y_{calc})^2}{n} \tag{6}$$

Both y and n represents water temperature and the size of data.

The algorithm explained previously represents the simplest implementation of back-propagation learning. In several cases, this algorithm has been widely criticized for slow convergence, inefficiency, and lack robustness, as the network could get stuck in a shallow local minimum. To avoid the problem of the convergence and to reduce the training time, we can use algorithms based on steepest descent gradient with momentum [7]. Momentum allows a network to respond not only to the local gradient, but also to recent trends in the error surface. Acting like a lowpass filter, momentum allows the network to ignore small features in the error surface. The momentum is added to back-propagation learning by making weight changes equal to the sum of a fraction of the last weight change and the new change suggested by the back-propagation rule. The performance of the algorithm is very sensitive to the proper setting of the learning rate. The optimal learning rate changes during the training process, as the algorithm moves across the performance surface. We use a steepest descent algorithm by allowing the learning rate to change during the process with momentum training. There are also other algorithms based on the conjugate gradient; a search is performed along conjugate directions, which produces generally faster convergence than steepest descent directions. Among them, the Levenberg-Marquardt algorithm [24] was designed to approach second-order training speed.

2.2. Advantages of the Bayesian Regularisation Technique

Methods, including back-propagation neural nets, still present some problems, and principal among these are overtraining, overfitting, and selection of the best model from a number of models obtained in the validation process. Overfitting results from the use of too many adjustable parameters in modelling the training data. It can be avoided by the use of a validation set. However, validation procedures produce a family of similar models, and it is not clear which of these models is preferred. The purpose is to choose a model capable of not only learning the data, but also capable of offering a good generalization for prediction. To do this, one of the used techniques consists of dividing the data set into three sets, the first one is used for the learning; the second for adjusting the model and avoiding the overfitting by making an early stopping technique of the training when its error increases during the test; and the third set serves for testing the performance of the model. The choice of the second set for the validation of the model presents a major problem to get a model with good generalization for prediction. To ovoid these deficiencies, we use a method based on the technique of Bayesian regularization. It consists of penalizing the high values of the weights by modifying the cost function *J*. This technique forces the parameters (for example, the weights) not to take high values, and consequently to avoid overfitting. The advantages of Bayesian methods are that they produce models that are robust and well matched to the data, and which make optimal predictions. At the end of training, a Bayesian-regularized neural network has optimal generalization qualities. There is no need for a validation set, since the application of the Bayesian statistics provides a network that has maximum generalization. We have used the Bayesian-regularized neural network package included in the MATLAB Toolbox [25].

The Bayesian neural networks are classical back-propagation neural networks that incorporate the Bayesian regularization algorithm for finding the optimum weights; the basic method used in the network training is derived from the Levenberg-Marquardt algorithm and the MATLAB implementation of the algorithm uses an automatic adjustment of the learning rate and the Marquardt parameters. The Bayesian regularization takes place within the Levenberg-Marquardt algorithm and uses back-propagation to minimize the linear combination of squared errors and weights.

To predict the ability of the model in forecasting, we divided the initial data into two sets. The first one serves for the training; the second is used to test the validity of the model.

3. Forecasting Water Temperature

The temperature is an essential parameter in the quality control of water [26-28]. Other physical, chemical or biological parameters would be considered, but the temperature is the factor of the prime importance [29-30] because the effects which result from it can influence the other factors. The water quality can be affected when its temperature increases, by modifying, for example, the dissolved quantity of oxygen; consequently, the aquatic living conditions of the fauna and the flora will be modified forever [31-32]. When the temperature increases, the quantity of oxygen decreases, this leads to the development of micro-organisms capable to live in absence of oxygen with following consequences:

1. malodorous putrid gases
2. Reduction in the ability of the water purge

It is very difficult to define a determinist model to predict or forecast river water temperature, because there are many climatic factors having an impact on it. The physical knowledge of each parameter and its contribution to water temperature appears to be no realistic. For this, we develop a model of forecasting, based on back-propagation neural network networks and Bayesian regularization technique, which allows estimation and forecasting of water temperature at short and middle term, It intends to forecast water temperature of days (*t+i, i=1,2…*).

The data set served for this study contains only three climatic parameters to establish relationship with water temperature in a river:

1. Air temperature
2. Water flow
3. Quantity of rainfall on the river.

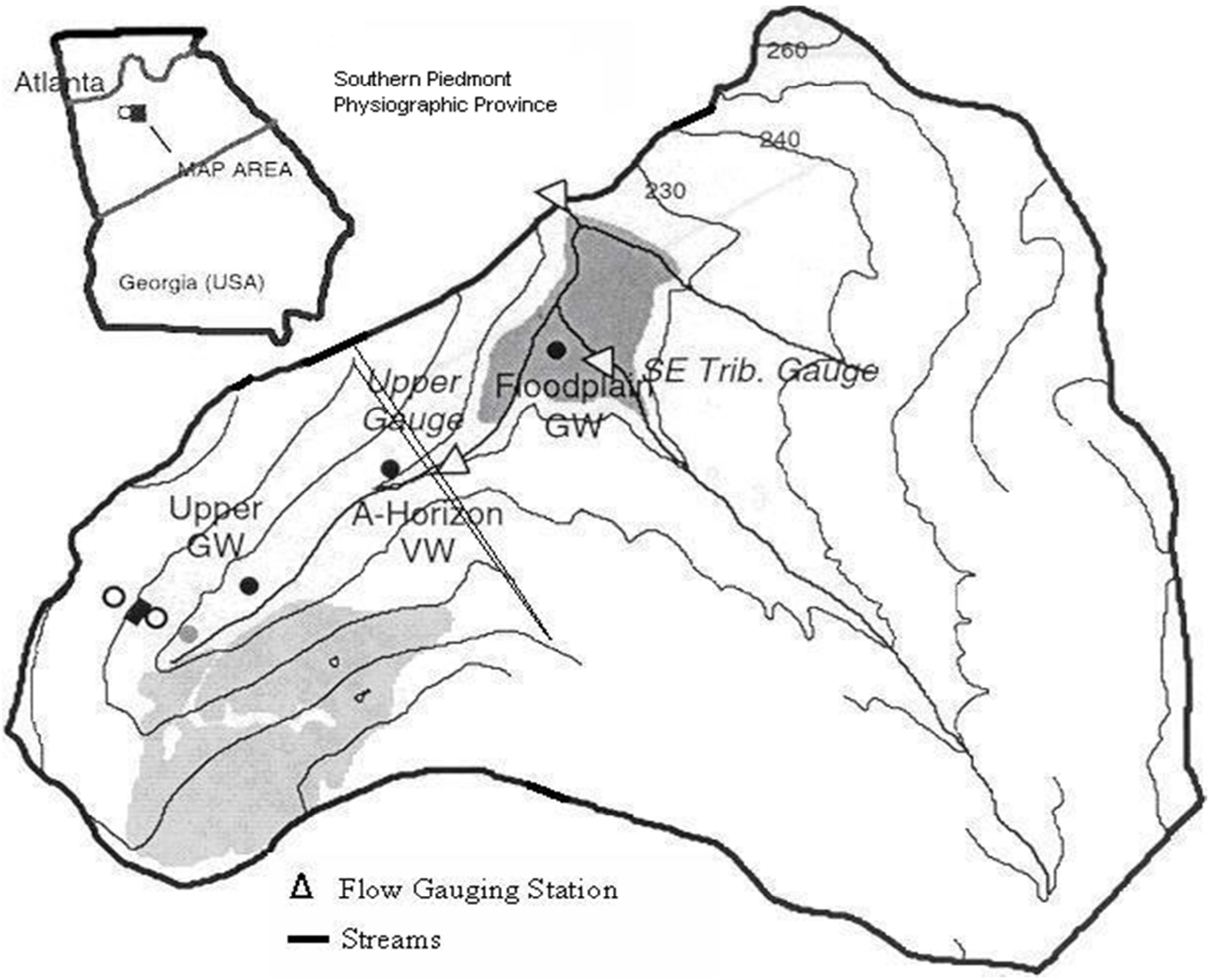

Figure 1. *Map of PMRW.*

4. The Study Region

The data set is taken from a small stream in Panola Mountain [33] (figure 1). It is located in the Piedmont physiographic province, approximately 25 km southeast of Atlanta, Georgia, USA. The climate in the area is classified as warm temperate subtropical, with an annual average temperature of 16°C. The long-term annual average precipitation in the area is 1.24 m; annual average runoff is

30% of precipitation, but this percentage varies widely from year to year as a function of both rainfall amount and timing. The water and air temperature, water flow and the quantity of rainfall on the river are taken from 1998 to 2002. The hourly values taken for the first three variables during the day are converted to daily average; the quantity of the precipitation was transformed into a cumulative variable of the day in question. The day begins at the 1 am in the morning, and ends at midnight.

5. Results

As indicated above, the forecasting of water temperature can be easily realized according to the equation 1, but this method has been found to have some limitations. This is essentially due to:

The error propagates by adding to every stage the previous dependent variable in the explanatory variables

The explanatory variables, such as air temperature, water flow and the quantity of rainfall on the river at the current time t, are constantly used in all the functions f_i.

A simple statistical analysis between climatic parameters considered above shows that air temperature is the explanatory parameter of the prime importance; the remainder factors have minor influencing water temperature at the current time, and at the days ($t+i$, $i=1,..$) (table 1). According to this, and due to the strong correlation between water temperature at the current time and those for the previous days, we firstly built model relating them by means the relation in the equation 7 formulated as below:

$$\hat{y}_{eau}(t) = f(y_{eau}(t-1), y_{eau}(t-2), y_{eau}(t-3), x_{air}(t)) \tag{7}$$

$\hat{y}_{eau}$ and y_{eau} corresponds respectively to both the estimated and observed values of water temperature at the current time; x_{air} to the air temperature.

The function f is derived from three-layer (input-hidden-output) feed-forward networks. A sigmoid transformation of the parameters factors values is performed in the first layer, the input layer. The output layer consisted of one neurone and represents water temperature. For a better flexibility of the neural networks model, only the previous three days are retained. For days (i>3), there is not much difference; so we selected a model with least number of parameters, that is, the four-parameters model ($y_{eau}(t-1)$, $y_{eau}(t-2)$, $y_{eau}(t-3)$, $x_{air}(t)$) for computing water temperature in the river considered in this study.

Table 1. *Correlation analysis of input-output relationships.*

	Flow(t)	Rainfall(t)	Air T(t)	Water T(t-1)	Water T(t-2)	Water T(t-3)	Water T(t)
Water flow(t)	1.00						
Rainfall(t)	0.50	1.00					
Air T(t)	-0.16	0.01	1.00				
Water T(t-1)	-0.22	-0.01	0.91	1.00			
Water T(t-2)	-0.22	-0.02	0.87	0.98	1.00		
Water T(t-3)	-0.22	-0.03	0.85	0.95	0.98	1.00	
Water T(t)	-0.21	0.01	0.94	0.96	0.95	0.93	1.00

T represents the temperature and t the current time

To ensure that neural networks can give reliable prediction, it has been proposed, based on empirical observations [34]; to use only network architecture with ρ parameter greater than 2 (ρ is the ration of the number of patterns in the training to the number of connections). The data set contains more than 1400 values of water temperature over a period of five years; in addition to this, the number of input nodes is represented by four parameters, which are the influencing factors to river water temperature. Based on these observations, an optimal architecture network can be easily obtained by varying the number of hidden neurons without limits. This was carried out iteratively. The size of the hidden layer was varied from 1 to 10 hidden units, and 20 networks were trained for each architecture. Plotting the error for the validation set against the number of hidden units allowed the optimal architecture to be determined. We started from one neuron in the hidden layer, the statistical indices of the correlation between experimental and predicted water temperature for the whole data set improves with the increase of their number; the optimal number retained of neurons in the hidden layer was found six. Any improvement was observed when this number increases over six.

Figure 2 shows the perfect relation, derived by the neural networks model, between water temperature and new influencing factors. Statistical criteria of the model are fairly good. Indeed, we have a model with about 98% of the total variance and a standard deviation equal to 0.68°C. The plot in the figure 3 shows the residuals plotted for the all data (1998-02). The 95% confidence intervals about these residuals are plotted as error bars. The first observation is that only a few cases present an error superior than 2°C. The residuals are the difference between the observed and predicted values. The residuals are useful for detecting failures in the model assumptions, since they correspond to the errors in the model equation. By assumption, these errors have independent normal distributions with mean zero and a constant variance.

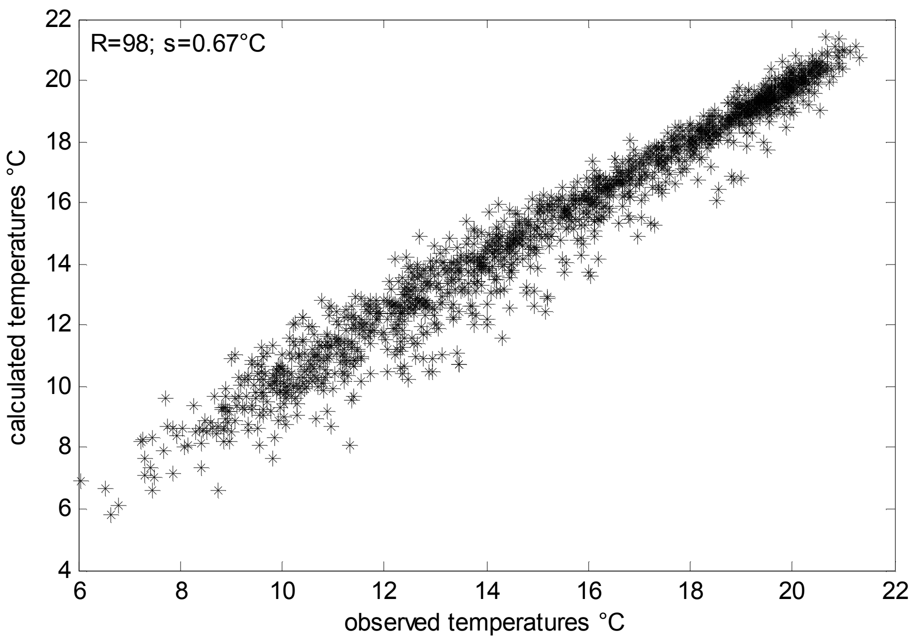

Figure 2. *Plot of predicted against experimental temperatures.*

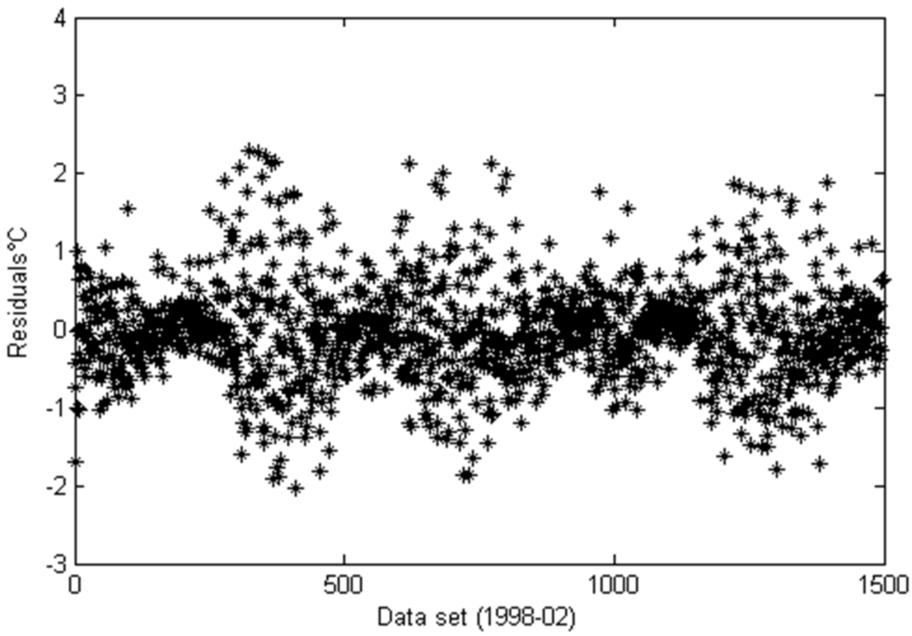

Figure 3. *The error distribution.*

The relationships formulated by the equation 7 is very interesting, because only air temperature at the current time and past water temperature are needed for estimating water temperature at the current time. Its utility can be useful to build accurate forecasting model of water temperature. In addition to estimated water temperature given by the neural networks model, air temperature data, with a projection in the future, could be given from the Meteorological Office.

The problem of forecast is then reduced to finding in every step a function *f* by means of the neural networks. In every step we have a table of data which leads to computing a function f_i relating water temperature and new influencing variables. The availability of air temperature data allows us using those taken from the data set for the study site. For this, we realized estimation by iteration (stepwise estimation). The first function f_1 calculate the water temperature at $(t+1)$ (similar to the equation 7). Another table of data is then built by adding $\hat{y}_{eau}(t+1)$ to the explanatory variables, and we compute $\hat{y}_{eau}(t+2)$. We continue the process above until the estimation of water temperature for $(t+k)$ days. The multitudes of estimations are realized according to the following relations in the equation (8):

$$\begin{aligned}
&\hat{y}_{eau}(t) = f(y_{eau}(t-1), y_{eau}(t-2), y_{eau}(t-3), x_{air}(t)) \\
&\hat{y}_{eau}(t+1) = f_1(y_{eau}(t), y_{eau}(t-1), y_{eau}(t-2), x_{air}(t+1)) \\
&\hat{y}_{eau}(t+2) = f_2(\hat{y}_{eau}(t+1), y_{eau}(t), y_{eau}(t-1), x_{air}(t+2)) \\
&\hat{y}_{eau}(t+3) = f_2(\hat{y}_{eau}(t+2), \hat{y}_{eau}(t+1), y_{eau}(t), x_{air}(t+3)) \\
&\bullet \\
&\bullet \\
&\bullet \\
&\hat{y}_{eau}(t+7) = f_2(\hat{y}_{eau}(t+6), \hat{y}_{eau}(t+5), \hat{y}_{eau}(t+4), x_{air}(t+7))
\end{aligned} \tag{8}$$

$f, f_1, f_2 \ldots f_7$ were derived from the neural networks analysis for each data set. For this, the same network consisting on 4-6-1 architecture was employed.

For the validity of the model, the data from 1998-2000 is used for the training, and the data form 2001-02 is used for the test. The Bayesian regularization is introduced to improve the generalization ability of the trained networks; it produces an effect similar to that obtained by the early-stopping technique commonly used to avoid overtraining in applications of neural networks [35].

Results of statistical indices between observed and estimated values of water temperature for the test data are reported in the table 2, and we present in the figure 4 as an example the plot of the perfect relation between them for the seventh day.

Table 2. *Statistical analysis relating to forecasting of water temperature from the current time to the seventh day.*

the day j	j=0	1	2	3	4	5	6	7
R, s	0.99	0.99	0.99	0.98	0.98	0.98	0.98	0.98
s	0.36	0.38	0.47	0.51	0.53	0.55	0.55	0.58

Both R and s represent the correlation coefficient and the standard deviation between estimated and observed water temperature for the data test.

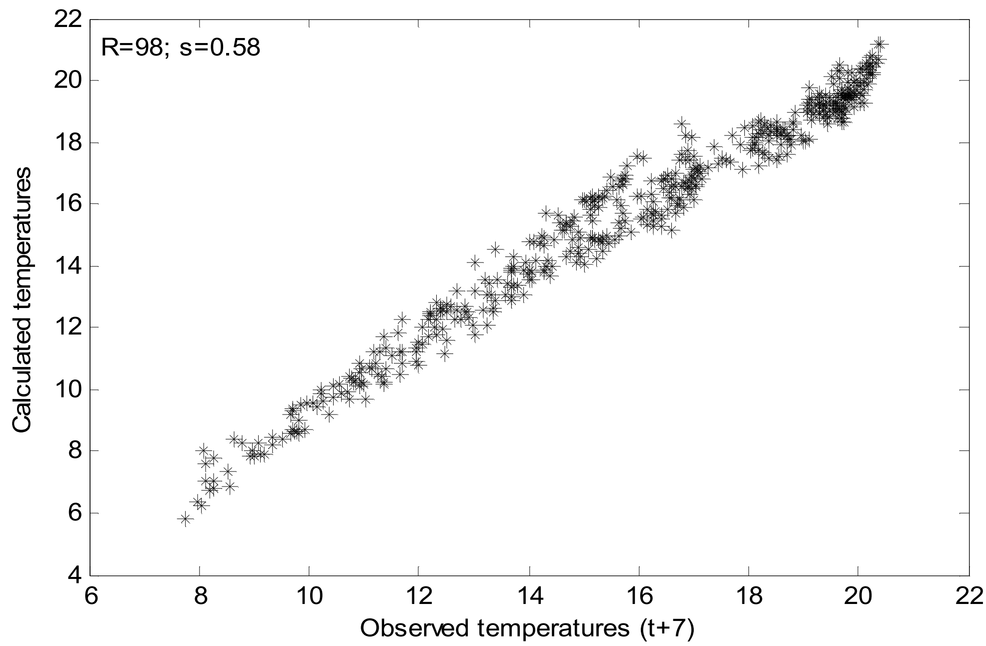

Figure 4. *Plots of predicted against experimental temperatures for the test data (2001-02).*

The statistical parameters of the models are very good. Indeed we have a model superior than 96% of the total variance and a standard deviation inferior to 0.6°C for all days. For the current time and the first day, results (R, s) are similar and are sharply better in comparison to the remainder days; the explanatory variables used in both functions f and f_1 are those taken from the data set. For the remainder of days, we use the calculated water temperature for the past.

6. Conclusion

This paper serves as a contribution to forecasting problem in hydrological sciences. The merits of this methodology are discussed through an example concerning the forecast of water temperature in a small stream. Artificial neural networks (ANNS) are employed here to investigate relationships between water temperature and influencing factors, with emphasis on Bayesian Regularization technique.

It seems firstly from this study that air temperature at the current time and past water temperature are the most dominant influencing variables, and are efficient to estimate water temperature at the current time. Efficiency of the neural networks model investigated to represent input-output relationships is not demonstrated here in comparison with other models. However, this present study has concentrated on developing a computational method as to how we have adapted it to generate multi-step-ahead neural networks for forecasting water temperature.

The size of the data set is enough of a correct investigation, as it overcomes the tendency of an overflexible network to discover nonexistent, or overly complex, data models. In addition to that, daily water temperature does not vary significantly from day to day. Bayesian regularization technique is only used here to obtain a neural networks model with high generalization ability, without needed a validation set during the training phase.

An accurate forecasting of water temperature is made possible with high success, because daily air temperature is involved in every stage of the forecast.

Acknowledgements

This research was made possible by a financial support of the AUF (www.auf.org) for a stay in the laboratory of the statistical hydrology (www.ete.inrs.ca).

References

[1] Coulibaly, P., Anctil, F. & Bobeé, B. (1999) Prévision hydrologique par réseaux de neurones artificiels: Etat de l'art. Can. J. Civ. Engrg., 26 (3), 293–304.

[2] Hipel, K. W. & McLeod, A. I. (1994) Time series modelling of water resources and environmental systems. Elsevier, Hollande.

[3] Thirumalaiah, K. & Deo, M. C. (2000) Hydrological forecasting using neural networks. J. Hydrologic Eng., 5 (2), 180-189

[4] WMO, (1994) Guide to hydrological practices. WMO-No. 168, Genève, Suisse.

[5] Garner G, Hannah DM, Sadler JP & Orr HG (2014) River temperature regimes of England and Wales: spatial patterns, inter-annual variability and climatic sensitivity. Hydrological Processes 28 (22): 5583-5598

[6] Cole JC, Maloney KO, Schmid M & McKenna JE (2014) Developing and testing temperature models for regulated systems: A case study on the Upper Delaware River. Journal of Hydrology 519: 588-598.

[7] Dreyfus, G., Martinez, J., Samuelides, M., Gordon, M., Badran, F., Thiria, S. & Hérault, L. (2002) Réseaux de neurones - Méthodologie et applications. Edition Eyrolles.

[8] Govindaraju, R. S. & Rao, A. R. (2000) Artificial Neural Networks in Hydrology. Springer; 1 edition.

[9] Liang, R. H. & Hsu, Y. Y. (1994) Scheduling of hydroelectric generations using artificial neural networks. IEEE Proceedings, Generation Transmission and Distribution, 141 (5), 452–458.

[10] Chang, F. J., Chiang, Y. M. & Chang, L. C. (2007) Multi-step-ahead neural networks for flood forecasting / Réseaux de neurones à échéances multiples pour la prévision de crue. Hydrol. Sci. J., 52, 114-130.

[11] Birikundavyi, S., Labib, R., Trung, H. T. & Rousselle, J. (2002) Performance of neural networks in daily streamflow forecasting. J. Hydrologic Eng., 7 (5), 392-398.

[12] Maier, H. R. & Dandy, G. C. (1996) The use of artificial neural networks for the prediction of water quality parameters. Water Resour. Res., 32 (4), 1013–1022.

[13] Zhang, Q. & Stanley, S. J. (1997) Forecasting raw-water quality parameters for the north saskatchewan river by neural network modelling. Water Resour. Res., 31 (9), 2340–2350.

[14] Zhang, S. P., Watanabe, H. & Yamada, R. (1994) Prediction of daily water demands by neural networks In: Stochastic and statistical methods in hydrology and environmental engineering. 3. (ed. by K. W. Hipel, McLeod, A. I. Panu, U.S. & Singh, V. P.) 217–227, Kluwer Academic Publishers, Dortrecht, Hollande.

[15] Sezin Tokar, A. & Momcilo, M. (2000) Precipitation-runoff modelling using artificial neural networks and conceptual models. J. Hydrologic Eng., 5 (2), 156-161.

[16] Coulibaly, P., Anctil, F. & Bobeé, B. (2001) Multivariate reservoir inflow forecasting using temporal neural networks. J. Hydrologic Eng., 6 (5), 367-376.

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

[18] Menchero, A., Diez, R. M., Insua, D. R. & Müller, P. (2005) Bayesian analysis of nonlinear autoregression models based on neural networks. Neural Computation, 17, 453-485.

[19] Titterington, D. M. (2004) Bayesian methods for neural networks and related models. Statist. Sci., 19 (1), 128-139.

[20] MacKay, D. J. C. (1992) Bayesian interpolation. Neural Computation, 4, 415-447.

[21] Rudolph, M. (1997) On topology, size and generalization of non-linear feed-forward neural networks" Neurocomputing, 16, 1–22.

[22] Haykin, S. (1994) Neural networks: a comprehensive foundation, Mac-Millan, New York.

[23] Hassoum, M. H. (1995). Fundamentals of artificial neural networks. Eds., MIT Press, Cambridge.

[24] Hagan, M. T. & Menhaj, M. B. (1994) Training feedforward networks with the Marquardt algorithm. IEEE Trans. Neural Networks, 5 (6): 989-993.

[25] MATLAB 7.0, The MathWorks, Inc., Natick, MA. 2002.

[26] Cluis, D. A. (1972) Relationship between stream water temperature and ambient air temperature – A simple autoregressive model for mean daily stream water temperature fluctuations. Nordic Hydrology, 3 (2), 65-71.

[27] Caissie, D., El-Jabi, N. & Satish, M. G. (2001) Modelling of maximum daily water temperatures in small stream using air temperature. J. hydrology, 251, 14-28.

[28] Morrill, J. C., Bales, R. C. et Conklin, M. H. (2005). Estimating stream temperature from air temperature: Implications for future water quality. Journal of Environmental Engineering-Asce, 131 (1): 139-146.

[29] Mohseni, O. & Stefan, H. G. (1999) Stream temperature/air temperature relationship: a physical interpretation. J. Hydrology, 218, 128-141.

[30] Moatar, F., Bustillo, V., Ducharne, A., Thierry, D., Sauquet, E. et Vidal, J. F. (2010). ICC-HYDROQUAL– Impact du Changement Climatique sur l'hydrosystème Loire Hydrologie, Régime thermique, qualité des eaux. Plan Loire Grandeur Nature.

[31] Ducharne, A. (2008) Importance of stream temperature to climate change impact on water quality. Hydrology and Earth System Sciences, 12 (3): 797-810.

[32] Ficklin DL, Stewart IT & Maurer EP (2013) Effects of climate change on stream temperature, dissolved oxygen, and sediment concentration in the Sierra Nevada in California. Water Resources Research 49 (5): 2765-2782.

[33] Hopper, P. H. (2001) Applying the scientific method to small catchement studies: a review of the Panola Mountain experience. Hydrol. Processes, 15, 2019-2050.

[34] Burns, J. A. & Whitesides, G. M. (1993) Feed-Forwards Neural Networks in chemistry: Mathematical Systems For classification and Pattern Recognition. Chem. Rev., 93, 8, 2591-2601.

[35] Bishop, C. M. (1995) Neural Networks for Pattern Recognition; Oxford, University Press.