ELSEVIER

Cold Regions Science and Technology 35 (2002) 115–122

cold regions science and technology

www.elsevier.com/locate/coldregions

Application of neural networks to predict ice jam occurrence

Darrell D. Massie [a,*], Kathleen D. White [b], Steven F. Daly [b]

[a]Department of Civil and Mechanical Engineering, United States Military Academy, West Point, NY 10996, USA
[b]US Army Engineer Research and Development Center, Cold Regions Research and Engineering Laboratory, 72 Lyme Road, Hanover, NH 03755, USA

Received 3 November 2001; received in revised form 28 May 2002; accepted 28 May 2002

Abstract

Artificial neural networks show potential for modeling the behavior of complex nonlinear processes, such as those involved in the occurrence of breakup ice jams. Because breakup ice jams and related flooding occur suddenly, ice jam prediction methods are desirable to provide early warning and to allow rapid, effective ice jam mitigation. Unlike open-water flooding, however, an analytical description of all the complex physical processes involved is not available. As a result, breakup ice jam prediction models have historically been limited to classical empirical single-variable threshold-type analyses to statistical methods such as logistic regression and discriminant function analysis. A neural network is shown to improve the error rates of ice jam prediction at Oil City, PA. The neural network input vector is determined and the methods used to appropriately account for the relatively low occurrence of jams are addressed. The neural network prediction proves to be more accurate than the current method used at this site, with a false positive error rate of 5.9% and a false negative error rate of 7.4%.
Published by Elsevier Science B.V.

Keywords: Ice jam; Ice jam prediction; Neural network

1. Introduction

Breakup ice jams occur during periods of thaw when increased discharge due to snowmelt and/or precipitation cause the forces on an ice cover to exceed its strength, resulting in the breakup of the ice cover. The broken ice is transported down the river until the river's transport capacity is exceeded. This forms an accumulation that obstructs flow, creates backwater, and can cause flooding. Other adverse impacts include bridge or levee failure, structural damage to riverine structures, and scour and erosion of riverbed and banks. Breakup ice jam stages can be much higher than open-water events with the same discharge due to the reduction in channel conveyance associated with the ice jam and due to the displacement effect of the ice itself. In addition, stage rises associated with breakup ice jams can be quite rapid, similar to flash floods. These rapid increases in stage can make it difficult to plan or execute ice jam mitigation measures such as evacuation or blasting. Depending on the jam characteristics, a prediction method might significantly increase warning time.

Ice jam prediction can be difficult because the formation and progression of breakup ice jams result from a complex interaction between hydrologic, hydraulic, and meteorological processes that are far

* Corresponding author.
E-mail address: id4747@usma.edu (D.D. Massie).

0165-232X/02/$ - see front matter. Published by Elsevier Science B.V.
PII: S0165-232X(02)00056-3

more complex and less well understood than for open-water flooding or freeze-up ice jam formation and progression. Due to the complexity of these interactions, an analytical model has not yet been formulated to describe the formation of breakup ice jams beginning with the hydrologic and meteorological processes that result in ice cover breakup and addressing ice transport, stoppage, and accumulation. The lack of such a model prevents the use of prediction methods based on dynamic analyses or other deterministic methods, which theoretically could be transferred fairly easily between locations. Thus, the small numbers of ice jam prediction models that have been developed are limited to highly site-specific classical empirical threshold-type analyses, to statistical methods such as logistic regression and discriminant function analysis (White and Daly, 2002). In addition, because of the lack of an analytical model of breakup jam processes, the selection of variables used in ice jam prediction models can be arbitrary and may be highly variable from site to site.

The difficulties in developing and utilizing a breakup ice jam prediction model can be illustrated at the confluence of the Allegheny River and Oil Creek at Oil City, PA, a location where breakup ice jams have frequently formed in the past. This area has a number of characteristics that predispose it towards serious jam events. Downstream of the confluence with Oil Creek, the Allegheny River significantly slows due to dredging that deepened the channel and decreased velocity. Because of the sudden reduction in slope caused by the deepened channel, freeze-up jams often occur in this portion of the river during early winter. Deck and Gooch (1981) reviewed historic ice jam events at Oil City and concluded that the most damaging ice jam floods occurred when transport of a broken ice cover on Oil Creek is halted by the presence of a freeze-up jam on the Allegheny River (Fig. 1). Oil Creek has a small, relatively steep drainage area and exhibits quick response to rainfall events. It is thus likely to experience an ice cover breakup and movement before the ice cover and freeze-up jam on the larger Allegheny River begins to break up and move. If the freeze-up jam on the Allegheny River is severe enough to restrict ice and water movement at the confluence of the two rivers, then movement of the broken ice emanating from Oil Creek is stopped and a breakup jam inevitably occurs.

In the early 1980s, the US Army Corps of Engineers Pittsburgh District, with design assistance from the Engineer Research and Development Center's Cold Regions Research and Engineering Laboratory (CRREL), constructed two ice control structures near Oil City to reduce ice jam flood occurrence. The first ice structure, a floating ice boom, was installed in 1982 on the Allegheny River just upstream from the confluence with Oil Creek. This structure is intended to decrease the frequency and severity of freeze-up jam formation on the Allegheny River near the confluence with Oil Creek. The second structure, completed in winter 1988–1989, is a low, gated, overflow weir located on Oil Creek about 5.3 miles upstream from its confluence with the Allegheny River. The purpose of this structure is to control ice movement in Oil Creek above the structure. However, during many winters, there is sufficient ice volume between the Oil Creek ice control structure and the Allegheny River confluence to form breakup ice jams. Although flooding since the construction of the two projects has not been significant, a method to predict the likelihood of jam occurrence is considered necessary in order to further reduce damages. Because of its frequent jam history and the need for a workable jam prediction model, this site was selected for development and testing of a breakup jam prediction model.

White and Daly (2002) developed the breakup ice jam prediction model currently used at Oil City. The model uses a multivariable threshold model to identify potential ice jam events, followed by a discriminant function analysis. As applied, the discriminant function analysis uses linear combinations of the variables to predict the group membership (e.g., jam or no jam) of new individuals by computing the generalized squared distance from a given observation to each group. An observation is classified into the group from which it has either the lowest generalized squared distance or the highest posterior probability. They also used a bootstrap simulation to provide confidence intervals for use by forecasters in ice jam prediction risk assessment. Applying the method to historical data, they found an 8.3% false positive error rate and a 15.6% false negative error rate.

To achieve a robust, useable prediction method, we seek to minimize both false positive and false negative errors. False positive errors predict a jam when one does not occur. Unfortunately, a high frequency of

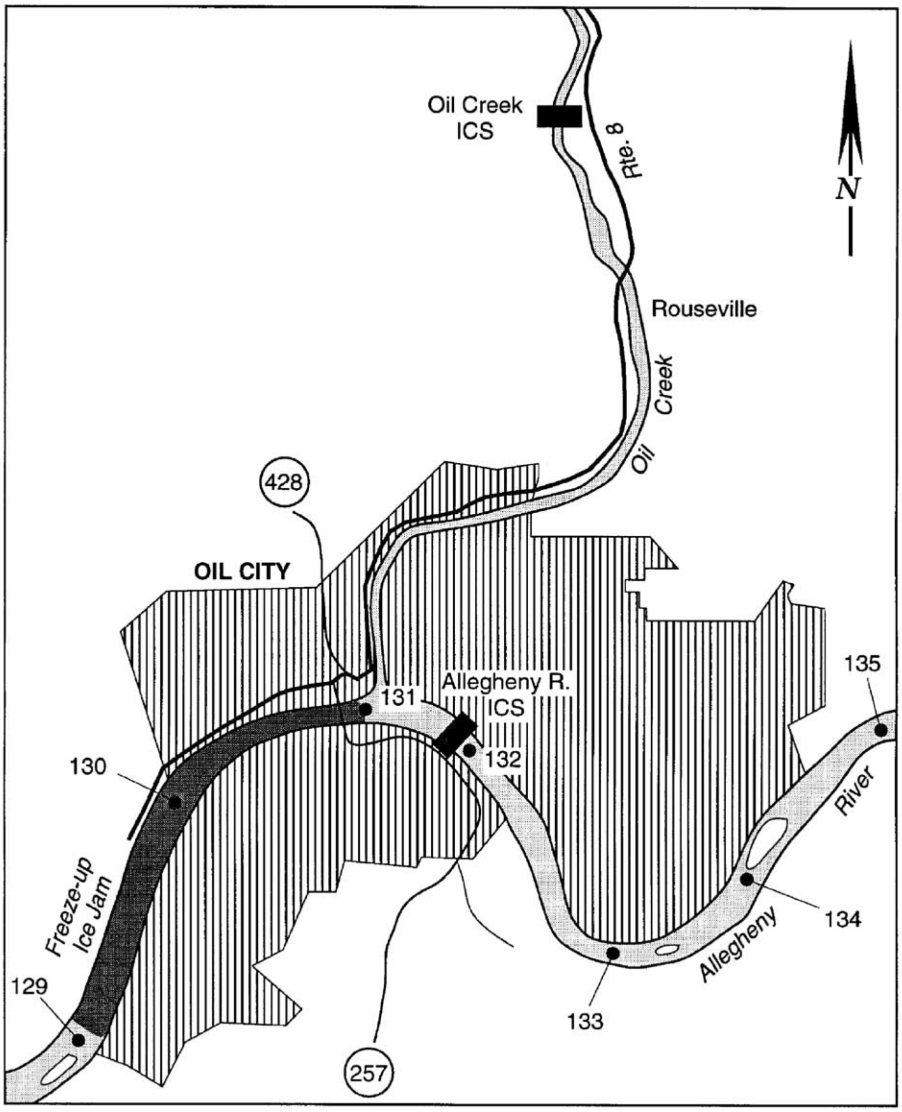

Fig. 1. Location map of Oil City, PA.

these false positive errors can lead to the "cry wolf" syndrome, in which warnings for actual jams are ignored and a generally low confidence in the model prediction results. False negative errors predict that a jam will not occur, but in actuality, one does. False negative errors are highly undesirable since a jam could lead to significant damage with no warning or preparation for the event, again leading to low confidence in the model prediction results. The false positive error rate of the method described herein appears reasonable given the complex processes involved in ice jams. However, the false negative error rate is too high for comfort, and increased prediction accuracy is desired.

Nomenclature

AFDD above freezing degree days (°C day)
Δ_{15}AFDD change in above freezing degree days for specified period, in this case, 15 (°C day)

$P(x\,|\,C_J)$ probability of a jam occurrence given input vector x
$P(x\,|\,C_N)$ probability of a non-jam occurrence given input vector x
FDD freezing degree days (°C day)
T_a average daily air temperature (°C)

2. Neural networks

2.1. Approach

Neural networks are a form of artificial intelligence that consist of nonlinear computer algorithms that "learn" with feedback to reproduce the existing relationship between input and output variables of complex nonlinear systems (Rumelhart and McClelland, 1986; Cowan and Sharp, 1988; Wasserman, 1989; Bishop, 1995). NN are particularly well suited for the type of problems posed by breakup jam prediction because they are easily configured to map several input variables to multiple output variables. Several types of NN structures are available; this study used a cascaded, feed-forward network without recursion. This type of NN has a structure similar to that of Fig. 2, where nodes (shown as circles), also known as neurons, within each layer are connected by weighting factors (shown as lines). The goal of the network of layers is to map the relationship between the input vector and output vector. The nodes collect information from weighted upstream node output, process the information using an activation function, and pass the information to the next layer using more weights. Use of a nonlinear activation function, such as the sigmoid function (Massie, 2001), results in nonlinear mapping. The size of input and output vectors can vary, as can the number of nodes and layers. More nodes and weights in the network architecture allows for more complex modeling of nonlinear systems, but is computationally intensive and may result in a network that does not obtain a generalized solution.

The development of a NN requires selection of the number of layers, the number of nodes in each layer, the activation function of each layer, and the training algorithm, which is used to minimize the error between the input and output vectors (Wasserman, 1989). Once the architecture has been determined, the network is trained and then tested. In the training (or learning) phase, the NN is taught to match a known set of corresponding input and output values in order to "learn" the relationship existing between them, and at the same time, modification of the weights associated with each neural connection by the training algorithm. Training is the most time-consuming phase of NN development and it is critical for the success of the neural network as a predictive model. In the testing phase, also known as generalization, the NN is tested using another known set of corresponding

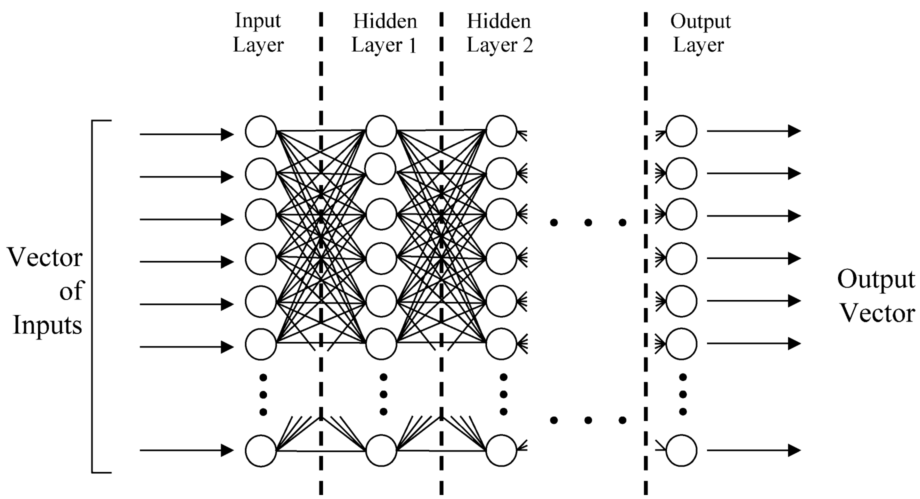

Fig. 2. Generic view of NN layered architecture.

input and output values (none of which belong to the training set) and its performance is evaluated.

When a NN is trained for classification, each output is assigned a class label and the probability of an output result into that class can be determined. For the purposes of breakup ice jam prediction, a NN could be described as a classification system that maps the meteorological, hydrologic, hydraulic, or ice values in the input vector to a class label category of either jam or no-jam. Because NN can establish nonlinear decision boundaries, they are capable of outperforming standard statistical classification methods, such as discriminant function analysis. Fig. 3 graphically depicts how a nonlinear boundary can better determine class labels when classes are tightly intermixed than a linear boundary.

The various steps taken to train and test the use of a neural network for ice jam prediction, based on data from Oil City, PA, are described below.

2.2. Variable selection

Selection of input variables is perhaps the most critical decision impacting on mapping accuracy in a neural network. The data set for Oil City contained a list of variables measured daily from the months of December through March for the period 1933 to 2000. The measured values included average air temperature and discharge for the Allegheny River and Oil Creek. The data set also indicated whether or not an ice jam had occurred on a given day. In the 67 years of observed data, 17 ice jams were recorded. Based on a statistical analysis, any data that were in obvious error and could cause skewed results were removed.

As is often the case for ice jams, an investigation of the ice, hydraulic, and meteorological conditions that lead to the formation of ice jams in Oil City (Daly et al., 1996) did not reveal a clear relationship between any of the available hydraulic and meteorological variables and the severity of ice jams at Oil City. However, inspection of the historical record revealed that ice jams could be expected when periods of intense cold are followed by sudden increases in flow in Oil Creek. After numerical experiment, Daly et al. (1996) found that examining the change in accumulated freezing degree days (AFDD) throughout the winter could be used to identify periods of intense cold. Freezing degree days (FDD) are calculated for each day of the winter season using the relationship FDD=$(0-T_a)$, where T_a is the average daily air temperature in degrees Celsius. A negative FDD value represents a temperature warmer than freezing, while a positive freezing-degree day represents temperatures below freezing. The FDD values for each day of the winter are summed to determine the AFDD each day. Daly et al. (1996) selected Δ_{15}AFDD, or difference between the AFDD on each day of the winter season and the AFDD accumulated during the previous 15 days. Graphical analysis indicated that this value was found to increase continuously during intense cold periods, reaching a peak when the cold period ended. The magnitude of the peak varied with the intensity and duration of the cold. Sudden increases in the discharge in Oil Creek were identified by determining, on each day of the winter season, the difference between the daily average discharge in Oil Creek and the daily average discharge on the previous day.

An independent statistical analysis performed in the current study confirmed that the variable selected in the earlier statistical prediction method make the most significant contributions to ice jams: a change in accumulated freezing-degree day (AFDD) and increased discharge in both Oil Creek and the Allegheny River. Final variable selection included average daily air temperature, 1- to 15-day change in AFDD, flow rates from both the Allegheny River and Oil Creek, as well as a 1-day change in the volumetric flow rates for both rivers. Table 1 summarizes input and output variables.

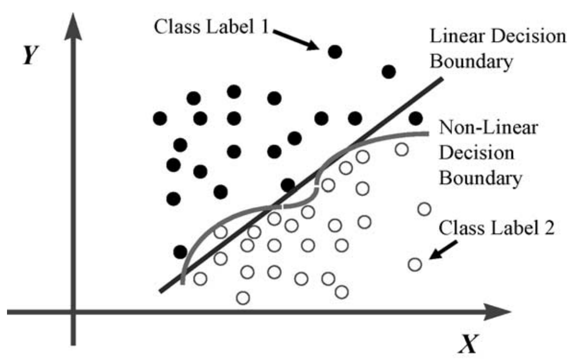

Fig. 3. Linear and nonlinear decision boundaries for establishing classes where X and Y are input variables.

Table 1
Neural network input and output variables

Input variables	Transformation	Output variables
Allegheny River flow rate (cfs)	linear	ice jam
$\log_{10}$ Allegheny River flow rate (cfs)	linear	no jam
Oil Creek flow rate (cfs)	linear	
$\log_{10}$ Oil Creek flow rate (cfs)	square root	
1-day delta Oil Creek Flow rate (cfs)	loglog	
$\log_{10}$ 1-day delta Oil Creek Flow rate (cfs)	linear	
Average daily ambient temperature (0 °C)	tanh	
Above freezing degree days (0 °C)	linear	
Δ_1AFDD through Δ_{15}AFDD (0 °C) [15 fields]	linear, power, or tanh	

2.3. Data preprocessing

Data preprocessing for a neural network involves normalizing the input data and reducing its dimensionality. Each value of input and output is normalized so that no one set of values dominates the solution. Without normalization, input variables with something as simple as a change in units could produce significantly different results. One of the simplest and most common methods to normalize data is through a simple linear rescaling. This normalizing process involves subtracting the mean, $\bar{x}_i$, of each feature, x_i, and dividing by the standard deviation σ_i as shown:

$$\tilde{x}_i = \frac{x_i - \overline{x}_i}{\sigma_i} \tag{1}$$

This simple linear normalization technique results in an input space with zero mean and unit standard deviation. Other data transformations, such as power functions, tangent or logarithm, as shown in Table 1, were applied to input variables. The transformation that gave the smallest misclassification rate was used in the final network.

2.4. Network architecture and development

The network consisted of a feed-forward architecture without the use of recursion. A total of 23 fields were used as input, and the output value was made into categorical variables, which do not presuppose natural ordering (Table 1). Also, 23 input values were used because the elimination of an input variable would presuppose an understanding of input variables to jam occurrences. Since there are only two categories of output, jam or no-jam, output values were represented by values of (1,0) and (0,1), respectively.

As in many statistical methods, imbalance in the size of the data sets representing the populations to be classified results in difficulty in application. In the 67 years of observed data at Oil City, only 17 ice jams were observed and recorded out of over 7700 wintertime records. The low probability of occurrence of can cause a neural network to draw a hyperplane that does not represent a general solution. An example of this would be to classify every occurrence as a no jam $P(xC_{\mathrm{N}})$, and in this case will only misclassify 17 dates. While the misclassification of no jams $P(xC_{\mathrm{N}})$ is small, the error rate of jam predictions is unacceptably high.

To overcome this difficulty, the number of no-jam occurrences used in the network training and testing data sets must be relatively close to the number of jam occurrences. With such a small number of no-jam days available for use in training and testing, it is important to ensure that the training set includes the features needed to predict a jam. The no-jam data set was first reduced to a manageable size after recognizing that a jam always occurs when there is an increase in average daily air temperature and discharge in Oil Creek. All data that did not fit these criteria were removed, thus reducing the number of no-jam events to 2700 occurrences. After investigating a number of clustering techniques, none of which showed definitive trends, 50 no-jam days were selected at random from this data set for training and testing. The no-jam events were divided equally between training and testing sets. Since testing of the prediction method was deemed important, 6 randomly selected jam days

Table 2
Comparisons of ice jam prediction at Oil City using various techniques

Error type	Empirical	Statistical	Combined empirical and statistical	Neural network
% False positive errors	11.8% (2/17)	8.3% (1/11)	35.3% (6/17)	5.9% (1/17)
% False negative errors	40.0%	15.6%	18.0%	7.4%

were included in the training set and the remaining 11 jam days were used for testing.

During the training phase of the neural network development, five of the six jams were classified correctly and all of the no-jam occurrences were correctly classified. During the testing phase of the neural network development, all 11 jams in the testing set were correctly classified, and 28 of the no-jam occurrences were correctly classified. After testing, the entire database of 7700 records was then fed to the trained neural network to check for a generalized solution. For this test, 93% of all no-jam events were correctly classified, and 92.6% jam events were correctly classified.

3. Results and discussion

Overall, the neural network classifier functioned with an accuracy of 93%. With all the original data, the network classified ice jam events with 94.1% jam accuracy and no-jam events with 92.6%. These results can be compared with previous studies conducted at Oil City. The results of the empirical prediction method alone are described in Daly et al. (1996) and the results of a combined empirical and statistical (discriminant functional analysis) method are described in White and Daly (2002). As Table 2 illustrates, the neural network solution provides an improvement over alternative methods. For a summary of results, see Table 3.

The neural network classifier only predicts whether a jam will or will not occur on a given day based on values forecast for that day. To obtain reliable future predictions, however, air temperatures and discharges must be forecast with reasonable certainty. Temperature predictions are readily available from the National Weather Service. Discharge, however, can-

Table 3
Summary of breakup jam prediction results using various methods

Water year	Actual event date	Actual event classification	Empirical model classification	Statistical model classification	Neural network classification
1936	27 February 1936	jam	jam	jam	jam
1937	25 January 1937	jam	no jam	jam	jam
1952	01 January 1952	jam	jam	jam	jam
1956	08 March 1956	jam	jam	jam	jam
1957	23 January 1957	jam	jam	jam	jam
1959	22 January 1959	jam	jam	jam	jam
1965	08 February 1965	jam	jam	no jam	jam
1966	11 February 1966	jam	jam	no jam	jam
1969	17 January 1969	jam	jam	jam	jam
1971	21 February 1971	jam	no jam	no jam	jam
1976	27 January 1976	jam	jam	no jam	jam
1977	05 March 1977	jam	jam	no jam	no jam
1979	25 February 1979	jam	jam	no jam	jam
1981	17 February 1981	jam	jam	jam	jam
1982	01 February 1982	jam	jam	jam	jam
1995	13 January 1995	jam	jam	jam	jam
1996	18 January 1996	jam	jam	jam	jam

not be predicted so readily, particularly with ice-covered conditions and snow covers with varying properties. Forecast discharge must be estimated based on anticipated precipitation amounts, snow conditions, and ground conditions. Current work includes the refinement and calibration of a watershed model for Oil Creek and use of an existing watershed model for the Allegany River. With this information, a future discharge prediction can be made, along with an associated error rate. The final product will consist of an easy-to-use internet-based program that allows for air temperature and discharge predictions up to 5 days in the future.

4. Conclusion

This study demonstrates, at least for one location, that a neural network method for predicting ice jams is capable of providing improved accuracy over statistical and classical empirical threshold-type analysis methods in obtaining a solution to a complex and elusive problem. The complex physical processes involved in the formation of breakup ice jams makes them difficult to predict. However, the sudden nature of occurrence and high stages makes jam prediction desirable. As is the case with empirical methods, neural network classifiers most likely will be site specific. However, since neural networks learn patterns with no modification of the algorithm, it is likely that they can be transported to other locations with minimal modifications. Further investigation must be made into this area.

Acknowledgements

The authors would like to acknowledge the support of the US Army Corps of Engineers' Pittsburgh District. The Civil Works Cold Regions Engineering Research Program Work Units "River Ice Forecasting" and "Statistical Methods in Ice Engineering" also supplied funding.

References

Bishop, C.M., 1995. Neural Networks for Pattern Recognition. Clarendon Press, Oxford.

Cowan, J.D., Sharp, D.H., 1988. Neural nets and artificial intelligence. Proceedings of the American Academy of Arts and Sciences 117, pp. 85–121.

Daly, S.F., Tuthill, A.M., White, K.D., 1996. Investigation of Ice Jam Formation at Oil City, PA. Contract report prepared by the US Army Cold Regions Research and Engineering Laboratory for the US Army Engineer District, Pittsburgh.

Deck, D., Gooch, G.E., 1981. Ice jam problems at Oil City, PA, Special Report 81-9, U.S. Army Engineer Research and Development Center Cold Regions Research and Engineering Laboratory, Hanover, NH.

Massie, D.D., 2001. Neural network fundamentals for scientists and engineers. Proceedings, ECOS 2001, International Conference on Efficiency, Cost, Optimisation, Simulation and Environmental Aspects of Energy and Process Systems, ASME, Istanbul, Turkey, July 4–6, 2001.

Rumelhart, D.E., McClelland, J.L., 1986. Parallel Distributed Processing: Explorations in the Microstructure of Cognition. MIT Press, Cambridge, MA.

Wasserman, P.D., 1989. Neural Computing: Theory and Practice. Van Nostrand Reinhold, New York.

White, K.D., Daly, S.F., 2002. Predicting ice jams with discriminant function analysis. Proceedings, OMAE 2002, The 21st International Conference on Offshore Mechanics and Arctic Engineering, ASME, Oslo, Norway, June 23–28, 2002.