THE JOURNAL OF NAVIGATION (2002), **55**, 137–143. © The Royal Institute of Navigation
DOI: 10.1017/S0373463301001576 Printed in the United Kingdom

Predicting Sea Ice Conditions using Neural Networks

A. El-Rabbany[1], G. Auda[1] and S. Abdelazim[1]

(*Ryerson Polytechnic University*, *Toronto*)

Safe and efficient marine navigation in ice-infested waters requires comprehensive and timely information on the sea ice conditions. These include information on the ice concentration and type, ice edge location, icebergs and open leads. The Canadian Ice Service is responsible for providing ice information in Canadian waters, mainly through its daily ice charts. Unfortunately, due to the difference in time between the ice chart production and its use by mariners, the ice information is always out of date. This problem might be overcome by developing a neural network-based system for predicting the ice conditions over time. A supervised neural network is trained to predict the ice conditions at a given location and time using the current ice charts, which are provided by the Canadian Ice Service. The input ice data is mapped to an output vector that gives the predicted ice conditions. The traditional non-modular feed-forward neural network structure failed to map the required function, and hence, was modularized to give better prediction performance. Each neural module was responsible for the prediction of a 5×5 km area, while the ice characteristic of interest was the total concentration.

KEY WORDS

1. Sea Ice. 2. Modelling. 3. Safety.

1. INTRODUCTION. Safe and efficient marine navigation in ice-infested waters requires comprehensive and timely information on the sea ice conditions (El-Rabbany, 2000). These include information on the ice concentration and type, ice edge location, icebergs and open leads. Locating the ice edge is the most critical information to ensure safe passage. Validation experiments have shown that radar remote sensing, particularly RADARSAT, has the capability of providing such comprehensive information. The ScanSAR modes of RADARSAT are recommended for comprehensive monitoring of the sea ice conditions. The Canadian Ice Service (CIS) is primarily using these modes, along with other sources of information such as other remote sensing satellites, for extracting the sea ice information in the form of daily ice charts (Ramsay *et al*., 1998). The fine and the standard beam modes, with shallow incident angle, are recommended for the determination of the location of icebergs (Randell *et al*., 1998). The availability of the ice charts enables the mariners to make crucial decisions regarding the selection of the best possible navigation routes.

The ScanSAR mode of RADARSAT has revisiting capabilities varying from one day for high latitude regions to four days for mid-latitude regions. Although this might be sufficient for polar areas, it may not be so for mid-latitude areas such as the Gulf of St. Lawrence and the Great Lakes, especially if the velocity of movement of

https://doi.org/10.1017/S0373463301001576 Published online by Cambridge University Press

the ice is high. In addition, due to the difference in time between the ice chart production and its use by mariners, the ice information is always out of date by hours or even days (Canadian Ice Service, 2001). This problem might be overcome by developing a neural network-based model for predicting the ice conditions over time.

Artificial neural networks have been successfully used for a substantial number of engineering applications (Principe *et al.*, 2000). Feed-forward back-propagation-trained networks are one of the most popular and successful paradigms. The power of these neural networks lies in their ability to model highly non-linear functions and mappings through supervised learning by example. This ability motivated us to use feed-forward networks to model ice concentration variations, which we perceived as likely to have fixed patterns/trends with time and variations in the environmental conditions.

2. ICE CHART PRODUCTION AND FORMAT.

The Canadian Ice Service is responsible for providing ice information in Canadian waters, mainly through its daily ice charts (Figure 1). To do this, the CIS uses various space-

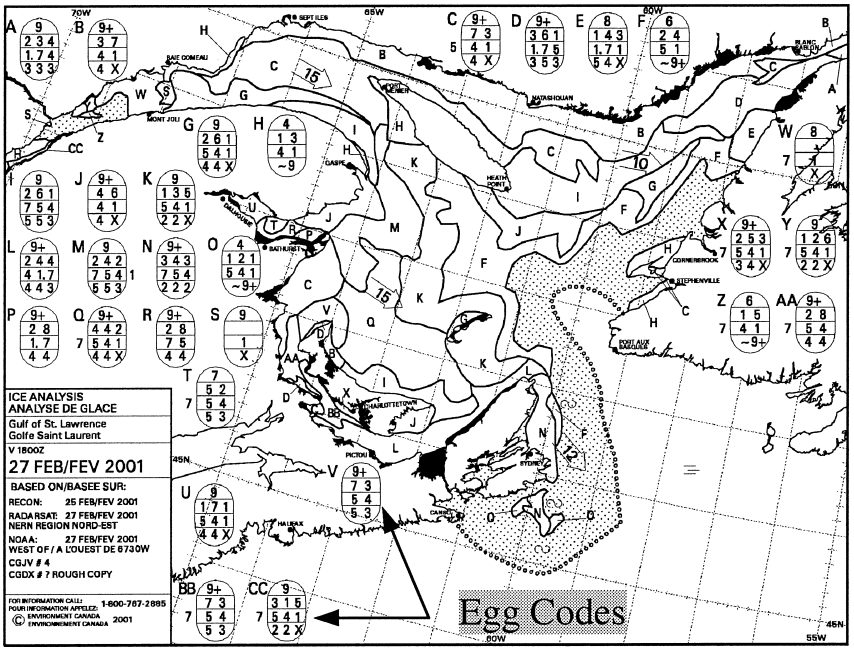


Figure 1. An example of the daily CIS Ice Charts (courtesy of the Canadian Ice Service).

borne and airborne remote sensing sensors, shore station observations and shipboard ice observations (Canadian Ice Service, 1994). The charts use the North American Datum 1927 (NAD 27) and the Lambert conical projection. The World Meteorological Organization (WMO) symbology for ice information, frequently referred to as the 'Egg Code', is used to describe the ice conditions (Figure 1). Boundaries are drawn around the ice areas with different concentrations; each is represented by an egg code (Raney and Falkingham, 1994).

https://doi.org/10.1017/S0373463301001576 Published online by Cambridge University Press

An egg code is an oval-shaped symbol, which contains three parts that describe the concentration of the ice, the stage of development (age) of the ice concentration and the predominant form of ice (floe size). These are expressed by up to 12 numerical values. The concentration of the ice represents the ratio between the area of the water surface covered by ice and the total area, and is expressed in tenths. The value of the ice concentration varies from 10/10 for consolidated ice to 1/10 for open water. The single uppermost parameter in the egg code represents the total concentration, which includes all stages of development. The second row in the egg code matrix contains the partial concentration for the thickest (left), the second thickest (middle) and the third thickest (right) ice types. The partial concentration field may contain two numbers, if only two ice types are present in the area (see Figure 1). If there is one ice type only, the partial concentration field will be left blank, as the concentration will be presented by the total concentration (Figure 1). The third field in the ice code contains the stages of development (age) for the ice types reported in the partial concentration field. Thicker ice refers to older ice, and vice versa. Various codes are used, depending on the stage of development. For example, a code of '1' is a assigned to the *new ice* (less than 10 cm in thickness), while a code of '9' is assigned to the *second stage thin first-year ice* (50–70 cm in thickness). Medium/thick first-year ice as well as old ice are assigned a dot (·) as part of their code. The last field in the code represents the predominant forms of the sea ice (floe sizes) corresponding to the stages of development identified in the previous field. Various codes are given to various floe sizes, which vary from '0' for the *pancake ice* to '7' for the *giant floe* (width greater than 10 km). Fast ice and icebergs are given the codes of '8' and '9', respectively. Undetermined ice form, unknown or no form is assigned the code of 'X' (see Figure 1).

3. MODELLING ICE CHARTS USING NEURAL NETWORKS.

The data set we used for our preliminary experiments was the daily ice charts of the Gulf of St. Lawrence for the year 2001. Several steps were needed to process those charts in order to obtain appropriate (numerical) values for the neural network simulator. First, the contours of the different regions were traced (Figure 2); hence,

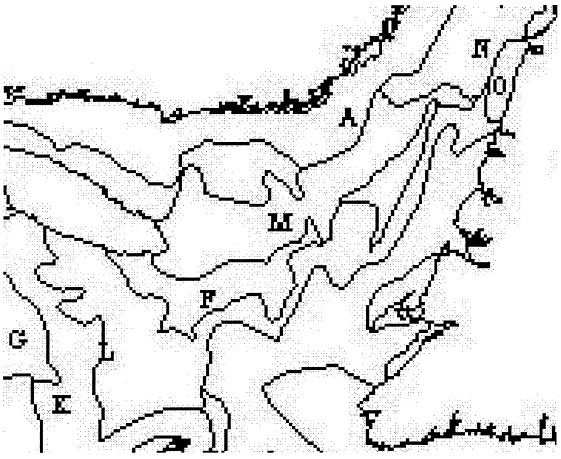


Figure 2. A partial zoom on one of the maps to show the different contours and regions.

https://doi.org/10.1017/S0373463301001576 Published online by Cambridge University Press

regions were defined and linked to their ice concentration values, according to the values in the egg code database. In order to digitize the data, the whole Gulf area had to be divided into smaller regions of (assumed) homogeneous ice conditions. We decided to use the resolution of the maps as a guide to define the dimensions of those smaller regions. In order to calculate the resolution of the maps, we repeatedly zoomed into the contours provided in the ice charts until we were able to view the shortest possible straight line that connects two points on the map. Since this line was about 500 metres, we decided to consider our smallest (homogenous) area/unit to be 5 by 5 kilometres. In addition, since the whole geographic area under consideration is about 13292 kilometres square, the map was divided into 529 homogeneous areas, 5 by 5 kilometres each. A rasterization process was carried out in order to convert the maps and the concentration values database into 18 grids of numbers, each grid representing one reading of the egg index.

Building the neural network model for the charts went through several phases. Some details are given below for the sake of providing the reader some practical experience. The design parameters of this model are the number of neural network inputs, outputs, and hidden nodes, as well as the structure of the feed-forward network itself. The different designs will be outlined in this section and the corresponding experimental results will be presented in the following section.

The inputs to the neural network should represent the date information as well as the environmental/weather conditions. However, since the environmental/weather conditions are not available to us so far, we can merely use the date of the chart as the input to the network. Two formats for representing this input were considered, namely, binary and continuous. The binary inputs were a total of 15 bits (Figure 3a):

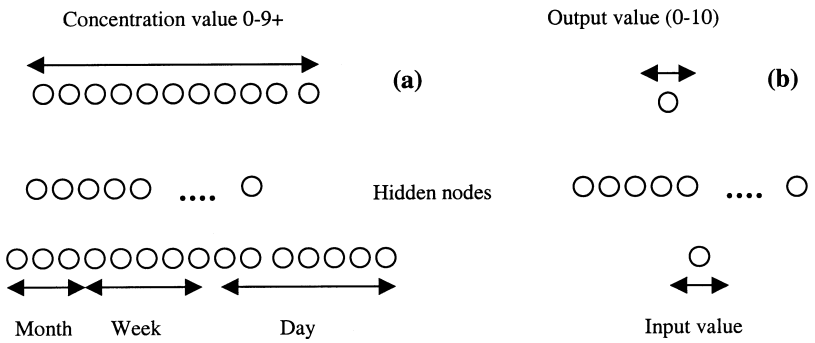


Figure 3. Binary versus continuous formats for the networks' inputs and outputs.

3 bits representing the month, 5 bits for the week (each representing one week of that month, ignoring the fact that the 5th week is always shorter), and 7 bits for the day (each representing one day of the week). The continuous representation consisted of one bit that simply counts the day in the year, regardless of the month or week.

The outputs of the network (which represent the ice concentration) were also tested in both binary and continuous formats. The binary format was 11 bits (one bit for zero to 9 concentration levels, and one for the 9+ level), while the continuous format was just one bit that has a value from zero to ten (considering the 9+ level as a 10 value, Figure 3b).

The choice of the number of hidden nodes was, as usual, subject to some trial and

https://doi.org/10.1017/S0373463301001576 Published online by Cambridge University Press

error experiments. The structures of neural networks that were considered for these experiments were: a non-modular massive network, a modular decoupled network, and a modular cooperative network. The non-modular network was a massive one that receives the input date and produces 529 output vectors, one vector for each 5×5 km square area. The dimension of the vector is 11 for the binary output and 1 for the continuous output (Figure 4).

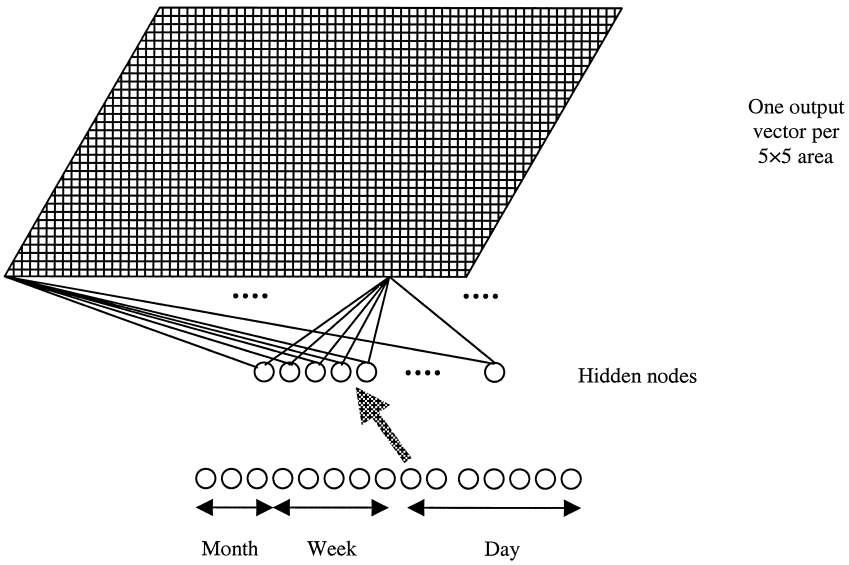


Figure 4. The massive network's hidden nodes are fully connected to the input vector and all of the 529 output vectors.

Given the size of the above network, preliminary training attempts showed serious performance problems. The network's training error was high and did not show a tendency to decrease, even with a very high number of training epochs. We think that the structure of the network is the reason behind this problem. Calculating the set of the large number of weights in the neural network that performs the desired mapping function is hindered by the high coupling (interdependency) the back-propagation training algorithm is introducing among hidden nodes (Auda *et al*., 1994). This high coupling does not cope with the variation in complexity of sub-parts of the task under consideration.

So, for the purpose of our experiments, we used a modular network with decoupled outputs (Figure 5). In other words, every output vector was connected to one dedicated hidden layer and not coupled with other output vectors in the learning equation.

Finally, the cooperative modular network is a modular network where modules 'cooperate' by exchanging information. In terms of the neural network design, this cooperation takes place through a set of additional connections between the output layer of each module and the input of four 'neighbouring' modules.

4. EXPERIMENTAL PROCEDURES AND RESULTS. We have used the NeuralWorks commercial software (NeuralWare Inc., 2000) for building the

https://doi.org/10.1017/S0373463301001576 Published online by Cambridge University Press

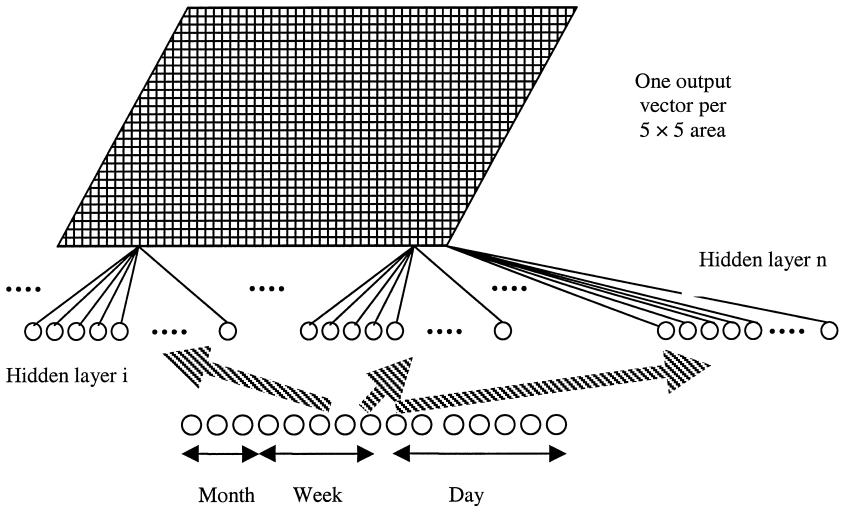


Figure 5. The decoupled modular network has a different set of hidden nodes for every output vector.

Table 1. A summary of the results of the experiments that shows the accuracy of the predictions.

	Prediction accuracy	
Network structure	Binary inputs and outputs	Continuous inputs and outputs
Massive single network	No convergence	No convergence
Decoupled modular net	63·3%	85·9%
Cooperative (given accurate neighbours' predictions)	96·1%	100%

neural networks, defining their parameters, and testing their performance, along with some simple software algorithms for pre- and post-processing the data. Three data sets were used; learning, validation, and testing sets. The learning stopping criteria was to save the best network, with respect to the validation set, over 100 000 iterations over the data records. A test for the best mapping performance was carried out every 2000 iterations. The initialization scheme for the weights was a uniform randomization scheme, with high and low limits of $+1$ and -1, respectively. The input samples were introduced to the neural network via a Shuffle and Deal randomization scheme (NeuralWare Inc., 2000). The default back-propagation momentum was 0·4 and the default-learning coefficient was 0·5. The Extended Delta Bar Delta (NeuralWare Inc., 2000) was used to further adjust the parameters during learning. Runs were done several times, and the best performing network was chosen. Table 1 summarizes the average results obtained over the grid of neural networks for the binary and continuous input/output formats, and the different structures used in the experiments.

As mentioned above, the massive neural network was of no practical use. On the other hand, the decoupled and cooperative modular networks proved themselves capable of learning the required functions. Using continuous values proved to be

https://doi.org/10.1017/S0373463301001576 Published online by Cambridge University Press

much better for the performance of the network modules. This is due to the fact that there is usually a small/incremental difference in the ice concentration between consecutive days. However, using binary inputs, the start of every week and/or every month introduces a significant Cartesian change in the corresponding input vector, which 'confused' the neural networks during learning.

Notice that a significant performance enhancement was achieved when decoupled modular networks were used, and yet another significant enhancement is achieved when cooperative networks were used. The drawback of the cooperative networks, however, is their sensitivity to the accuracy of the different neighbouring inputs. For the purpose of our experiments, we only introduced accurate predictions given by the neighbouring networks. This explains the very high prediction accuracy reached, which we do not think is achievable in real situations, given the unavoidable inaccuracy in the predictions made by neural networks. Therefore, to use the power of this structure, we plan to include environmental conditions (e.g. the water temperature and salinity) in the input vector in order to improve the predictions.

5. CONCLUSIONS.

This paper showed how neural networks could be used to predict the total ice concentration for the area of the Gulf of St. Lawrence. Several structures and input/output formats were used. The non-modular massive structure failed to learn the required mapping; hence, a modularization design was proposed. In addition, due to the incremental nature of the changes in the ice concentration from day to day, the binary input/output format was not as efficient as the continuous format. An average prediction performance of 85·9 % was achieved using the available data.

ACKNOWLEDGEMENTS

This research is supported by the Centre for Research in Earth and Space Technology (CRESTech), an Ontario Centre of Excellence. The ice charts were provided by the Canadian Ice Service (CIS) in Ottawa.

REFERENCES

Auda, G., Kamel, M. and Raafat, H. (1994). A new neural network structure with cooperative modules. *Proceedings of the IEEE*, Vol. 3, 1301–1306, Florida, USA.

Canadian Ice Service (1994). *MANICE – Manual of standard procedures for observing and reporting ice conditions*, *8th Edition*. Environment Canada, Ottawa, Canada.

Canadian Ice Service (2001). *ECDIS Ice Objects*, *Version 3.0*. Environment Canada, Ottawa, Canada.

El-Rabbany, A. (2000). Integrated Navigation Chart System. Invited presentation at the *Ice in ECDIS Workshop*, St. John's, Newfoundland, 5–6 June.

NeuralWare Inc. (2000). *NeuralWorks II/Plus*: *Software Manual*, PA, USA.

Principe, J., Euliano, N. and Lefebvre, W. (2000). *Neural and adaptive systems*. John Wiley & Sons, Inc.

Ramsay, B., Manore, M., Weir, L., Wilson, K. and Bradley D. (1998). Utilization of RADARSAT data in the Canadian Ice Service. *Application Development and Research Opportunity* (*ADRO*) *Symposium*, Montreal, October.

Randell, C., Rokonuzzaman, M., Youden, J. and Khan R. (1998). Evaluation of RADARSAT for detection, classification and discrimination of icebergs. *Application Development and Research Opportunity* (*ADRO*) *Symposium*, Montreal, October.

Raney, R. K. and Falkingham, J. C. (1994). RADARSAT and operational ice information. In: Haykin, S, Lewis, E. O, Raney, R. K. and Rossiter, J. R. (eds), *Remote Sensing of Sea Ice and Icebergs*. John Wiley & Sons, Inc.

https://doi.org/10.1017/S0373463301001576 Published online by Cambridge University Press