EXPLORATION OF STATIC AND TIME DEPENDENT NEURAL NETWORK TECHNIQUES FOR THE PREDICTION OF ICE ACCRETION ON OVERHEAD LINE CONDUCTORS

E. Larouche[+], *J. Rouat**[+], *G. Bouchard*[+] and *M. Farzaneh*[+]

[+]NSERC/Hydro-Québec/UQAC Industrial Chair on Atmospheric Icing of
Power Network Equipment (CIGELE), Chicoutimi, Quebec, Canada
*ERMETIS/DSA/UQAC, Chicoutimi, Quebec, Canada

Abstract

In order to predict the ice accretion on overhead line conductors, five artificial neural network (ANN) architectures were explored and compared. Two static networks, Multilayer Perceptron and Radial Basis Functions, as well as two time dependent networks, Finite Impulse Response and Elman, were compared with multiple linear regression (ADALINE). Results indicated that the FIR network yielded the best prediction and would be a good candidate as a predictor. The static networks yielded poor results and were ill-adapted for the task.

1. INTRODUCTION

Atmospheric icing of structures is a phenomenon that may have catastrophic technological and economical consequences on a large number of human activities. A prominent example of the disruptive effect of atmospheric icing manifests itself in electrical power and distribution networks. The impact can be temporary paralysis or even complete collapse in very severe cases.

The development of communication technologies and information processing systems allowed to use the real time monitoring icing system, which in turn made possible to increase our knowledge of atmospheric icing of structures and the development of ice models.

The models for predicting the atmospheric icing of structures can be subdivided into two categories: i) “physical” models based on a mathematical description of the ice accretion process and ii) “empirical” models based on statistical relationships existing between measurable instrumental variables [1].

The present study fits in the second category and aims at adapting the most adequate neural network architectures to the prediction of ice accretion on overhead line conductors. It includes a comparison between neural networks and linear standard statistical techniques, thus allowing to assess the performance of neural models. In recent years, neural networks have indeed generated a strong interest for the forecasting of time-series and the modelling of complex non-linear dynamical systems [2]. One can cite the work of McComber et al. [1][3] who studies the prediction of icing on transmission lines by using a feedforward neural network with time delays.

2. ICING DATA

To train the neural networks we used data from the Mt-Belair icing test site, located near Quebec City. As part of its real-time monitoring system SYGIVRE, Hydro-Québec has equipped this site with instrumentation. The standard meteorological variables are measured: ambient temperature, wind velocity and direction, relative humidity and precipitation rate. Also, an icing rate meter (IRM) monitors the severity of atmospheric icing. Load cells are installed on insulator strings to measure ice loads (on the 315 kV line). Data are recorded continuously with a hourly rate.

An experimental database was created using 33 icing events that occurred in the period of February 98 to April 99 inclusively. This gives a total of about 1250 data points. Since an icing event is divided into accretion, persistence and ice shedding phases, a second database was created using only the accretion phase giving about 520 data points. Each database was further divided into two data sets for the purpose of training and testing the neural models. The 75/25 rule was used, i.e. 75% of the data for training and 25 % for validation.

The neural networks in this study make use of the following input variables: temperature, normal wind speed and IRM signals. The load cell signal constitutes the output variable. The data were pre-processed before feeding to the neural networks. For each icing event, the load signal was first rectified in order to remove the offset at the start of the event. Then, each variable data set was range-normalized to unity by an affine transformation.

3. NEURAL NETWORKS

An artificial neural network (ANN) is a distributed information system made of simple inter-connected processing elements (neurons). ANN are loosely inspired by real biological neural systems and demonstrate to some extent capabilities of their living counterparts like non-linearity, learning and adaptation [4]. In the context of prediction task, artificial neural networks are known to approximate arbitrary smooth non-linear static input/output functional relationships given enough processing elements.

Usually, networks are trained by presenting pairs of matching input/output patterns (supervised learning). The knowledge acquired by the network is stored in the neuron interconnection strengths (synaptic weights) by an adaptation process (learning algorithm). A single neuron is a non-linear output of the weighted sum of its inputs. ANN come in a variety of topological and architectural forms and many learning algorithms have been devised so far.

4. NEURAL ARCHITECTURES

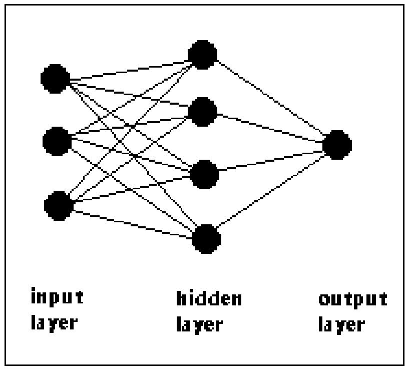
Common to all architectures, the number of input and output neurons was fixed to three and one respectively.

4.1 Static architectures

4.1.1 Multilayer perceptron (MLP)

The multilayer perceptron is the most common neural network encountered. It is comprised of multiple layers (input, hidden, output) with signal flowing from the input layer to the output layer (feedforward direction). The neuron has a sigmoidal transfer function given by $f(x)=1/(1+\exp(-x))$. To perform training, we used the backpropagation with a momentum equal to 0.9 [4]. (See fig.1)

Fig. 1: Multilayer perceptron network model



4.1.2 Radial-Basis Function network (RBF)

A Radial-Basis Function Neural Network has also a feedforward topology as the Multilayer Perceptron. But the hidden layer is comprised of Gaussian interpolating functions while the output layer is linear. It is based on a radial decomposition of the input space [4]. The spread constant for the hidden neurons was set to 0.1. The number of neurons in the hidden layer was progressively increased one step at a time by using a constructive algorithm that automatically choose the appropriate center of the radial-basis functions. The maximum number of hidden neurons was set as a design parameter (from 25 to 100).

4.2 Time dependent architectures

4.2.1 Temporal processing structure

Temporal processing involves processing of patterns through time. Application of neural nets to treat temporal problems appeared more recently in the research literature. The appropriate response at a particular point in time depends not only on the current input, but potentially on all previous inputs. Prediction can be viewed as made of two conceptually distinct components. The first is a short-term memory that retains aspects of the input sequence relevant to predictions. The second makes a prediction based on the short-term memory. In a neural net setting, the predictor will mostly be a feedforward component, while the short-term memory will often have internal recurrent connections. Many memory architectures have been proposed over the years [5] (tapped-delay line, exponential trace, gamma memory, etc.).

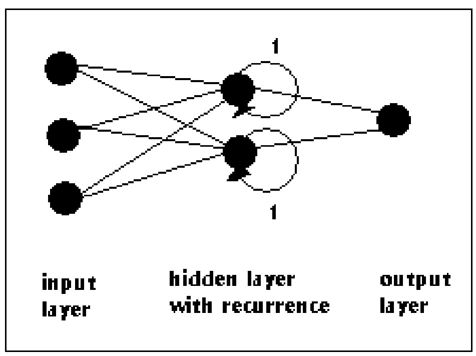
4.2.2 Finite Impulse Response network (FIR)

This network was introduced by Eric Wan in 1993 [6]. It resembles a standard feedforward network (MLP) except that each synapse is replaced with an adaptive FIR linear filter that is a weighted sum of past values of its inputs. The body of the neuron receives the filtered signals and then computes a squashing non-linear function of its inputs. Neurons are arranged in successive layers with synaptic filter connections. Training the network is accomplished through a modification of the standard backpropagation algorithm called *temporal backpropagation*. In our experiment, the network had only one hidden layer. For the first two layers (input and hidden), an hyperbolic tangent transfer function has been used while the output neuron was linear. The delay order of the synaptic filters was chosen to be 3:2.

4.2.3 Elman recurrent network

Elman proposed in 1990 a recurrent neural architecture for the task of phoneme recognition [4,7]. The Elman network comprises recurrent connections from hidden non-linear neurons (sigmoidal) to a layer of context units consisting of unit delays. The hidden neurons also feed the output neurons (linear). The role of the context units is to provide the network with a dynamic memory that is expectedly relevant for a prediction task. (See fig.2) According to [4,7], the hidden neurons may discover abstract representations of temporal information.

Fig. 2: Elman network model



4.3 Error criterion

As a performance measure a well accepted error criteria was used: the Normalized Mean Square Error (NMSE). It is the ratio of the prediction variance over the natural variance of the target variable. A value of NMSE=1 corresponds to the prediction of the statistical average.

$$NMSE = \frac{1}{\sigma^2}\frac{1}{N}\sum(y_i - \hat{y}_i)^2$$

5. RESULTS

Each architecture has been uniformly trained and tested on the same data sets as described in section 2. For each architecture, the design parameters were varied to study their effect on prediction performance. We also, experimented with a varying number of neurons in the hidden layer .

Fig.3: MLP results for the variation of the number of hidden neurons

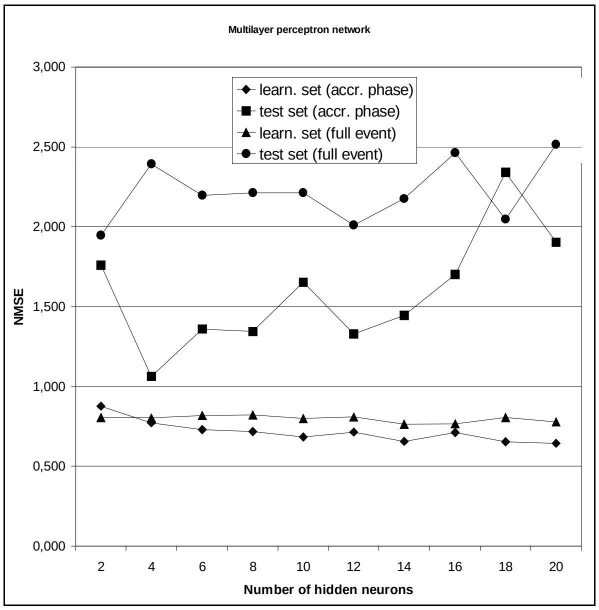


Fig.4: RBF results for the variation of the number of hidden neurons

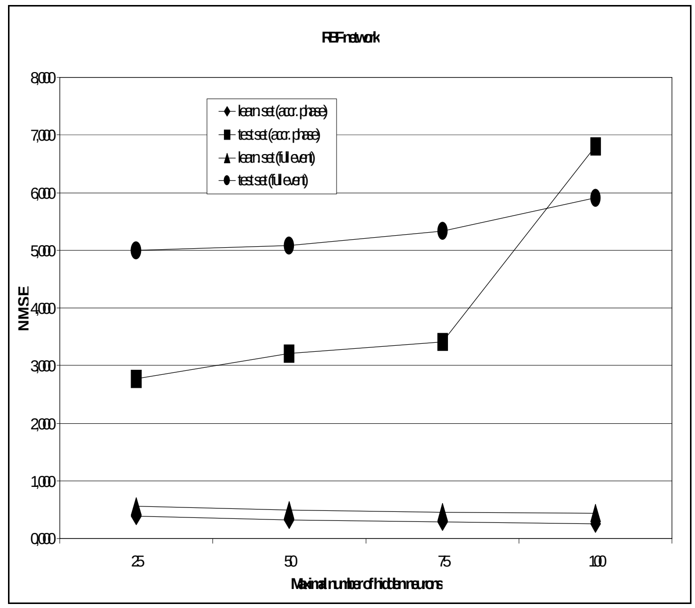


Fig.5: FIR results for the variation of the number of hidden neurons

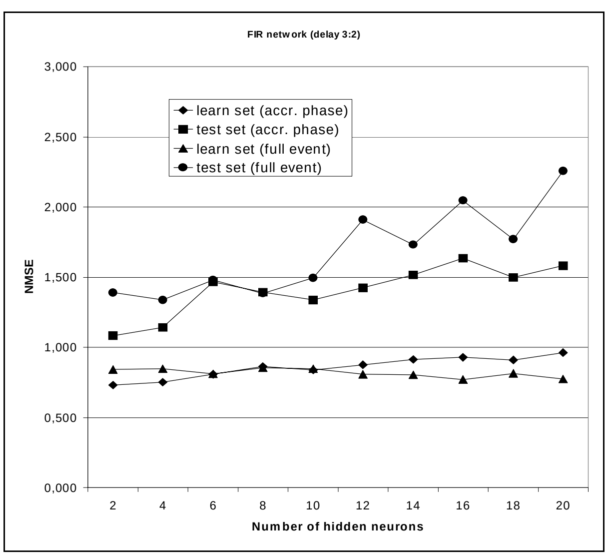
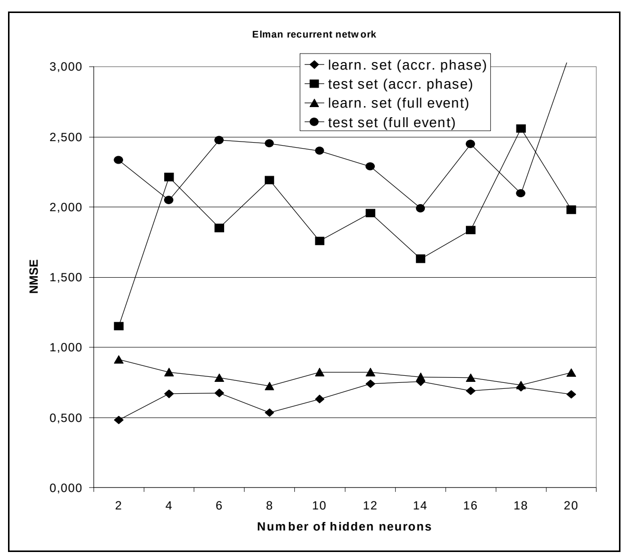
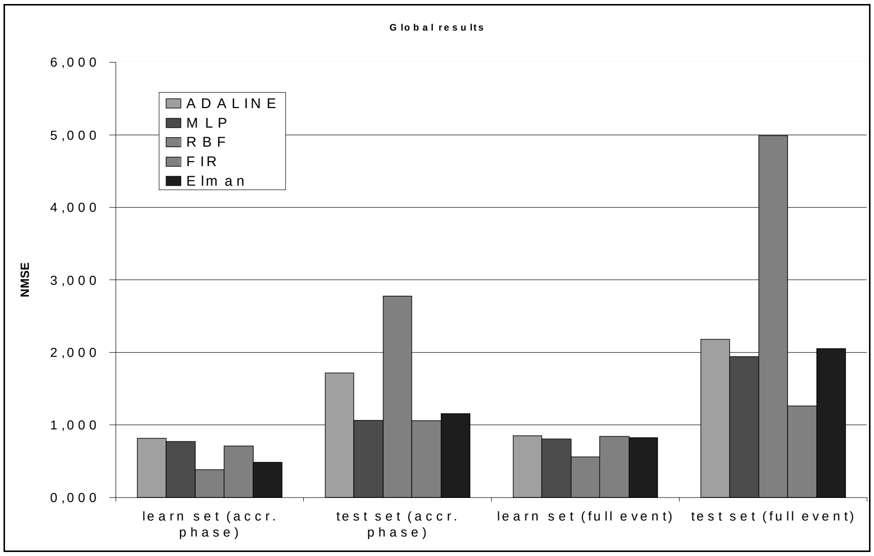


Fig. 6: Elman results for the variation of the number of hidden neurons



For each network, the architecture yielding the best performance on the testing set was retained. A global comparison of the neural networks is summarized in fig.7.. To establish a comparison baseline the ADALINE linear neural network has been also included. Although it makes use of an adaptive iterative procedure (LMS), it is entirely equivalent to a multiple linear regression [4].

Fig. 7: Summary of performance results



The FIR network shows the best overall performance when measured on the testing sets. To give a more precise idea of its behaviour, the prediction curves on the accretion phase test data is shown in fig.10. The ADALINE and MLP networks are also included for comparison (figs 8-9).

Fig. 8: ADALINE prediction curves (solid: actual data; dashed: forecast)

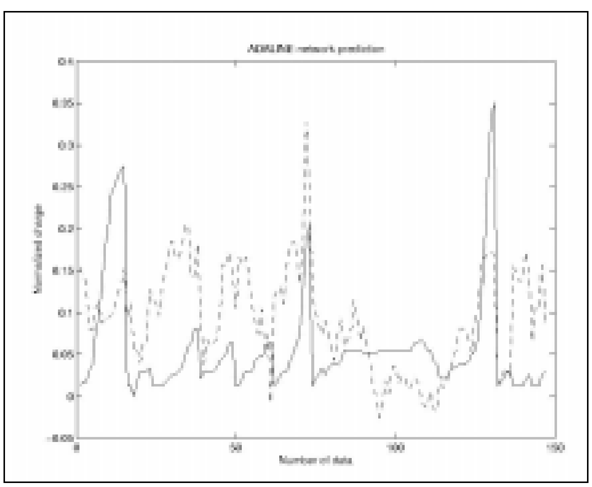


Fig. 9: MLP prediction curves (solid: actual data; dashed: forecast)

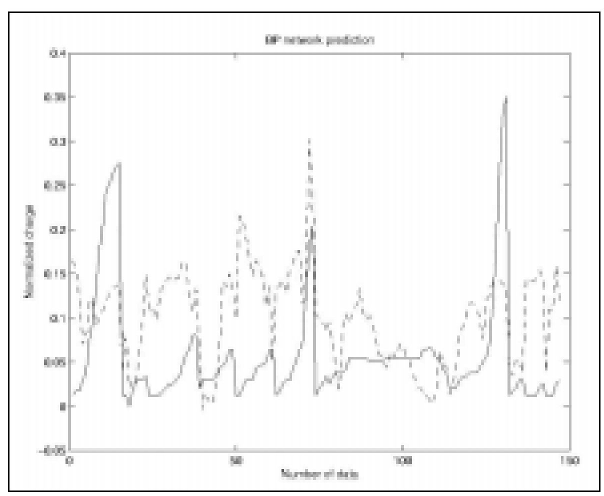


Fig. 10: FIR network prediction curves (solid: actual data; dashed: forecast)

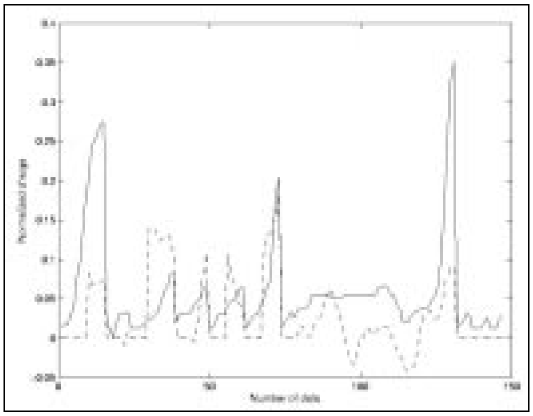
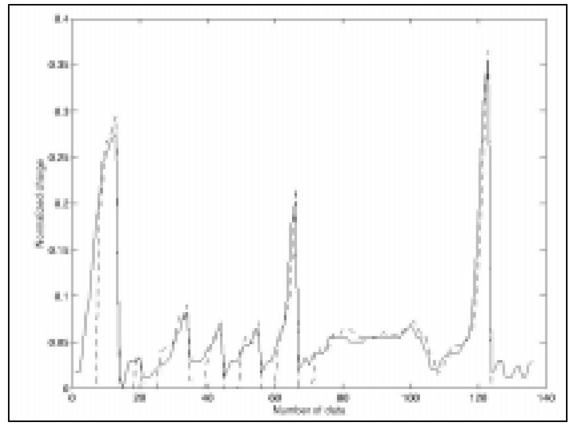


Fig. 11: FIR one-step ahead prediction (solid: actual data; dashed: forecast)



6. DISCUSSION

The global performance summary of fig.7 allows us to make some interesting observations. The data for the accretion phase demonstrate less error in the test set than for the full events data. This corroborates the expected fact that the accretion phase data are more homogeneous.

We also note more differences between the neural models in the full events test data than for the accretion phase. Moreover, the neural models make generally less error than the linear regression; maybe marginally for the learning sets but the difference is pronounced for the test sets. This gives good indication of the usefulness of neural models over simple linear techniques. Surprisingly, the RBF network shows the best performance during training, but the worst during validation.

As stated above, the best candidate for the prediction task is the FIR network. Based on the NMSE criteria, it is closely followed by the MLP. A comparison can be made in figs 8-10. It is shown that the MLP has a behaviour quite similar to that of the ADALINE. The FIR network has a markedly different behaviour. The number of false alarms is very low in comparison to that of the other networks. In fact, ADALINE and MLP have a somewhat “noisy” output that seems poorly correlated with the data they were supposed to predict. However, the FIR has some tendency to overestimate small icing events and to underestimate large ones. This suggests that the size of the training data is too small. The introduction of discontinuities at the delimitation of two successive events had to be taken into account during learning since time-dependent neural networks are generally adapted to handle long continuous time sequences. The null output observable for each event (Fig. 10 and 11) is an artefact caused by the testing methodology. This would not appear normally for a continuous application of the network.

We could give the neural models a better sensitivity to major icing events by carefully choosing the data going into training. By selectively filtering the data, a better response could possibly be achieved, with yet the drawback of further reducing the quantity of available useful data.

As it has been observed that the coupling of the real load cell signal as an additional input gives a much more precise prediction in forecast mode (one-step ahead prediction) than IRM input signal alone, another interesting approach coming from experimentation done in the course of this investigation can be suggested. It seems that neural networks could be useful in a forecast setting where the ice load could be predicted a few hours in advance. (See fig.11)

7. CONCLUSION

In the current framework of a forecasting approach based on the exclusive use of IRM signals as input, the FIR neural architecture seems to be the most promising for the task. However, the task is rather difficult and precise point-by-point forecasting seems unreasonable. It is complicated by the fact that we used noisy environmental data. The absence of a pure determinism between the input and output variables can also be easily postulated. In fact, any empirical model cannot do better than the existing correlation between the variables themselves. ANN does not constitute an exception. Neural networks need a lot of data to be trained properly and reach statistical convergence. However, in the present study, the available data points were unfortunately limited, and this indeed limited us in our conclusions. In spite of this limitation, the results showed that time dependent ANN (such as the FIR network) might assist us in the prediction of ice accretion. Other experiments with much larger databases still have to be made.

ACKNOWLEDGEMENT

This work was carried out within the framework of the NSERC/Hydro-Québec/UQAC Industrial Chair on Atmospheric Icing of Power Network Equipment (CIGELE) at the University of Quebec at Chicoutimi. The authors would like to thank all the sponsors of the project.

REFERENCES

[1] McComber P., de Lafontaine J., Laflamme J., 1998, “A Neural System to Estimate Transmission Line Icing” Proc. of 8th IWAIS, Iceland, pp. 101-106

[2] Weigend A.S. and Gershenfeld N., Proceedings of the NATO Advanced Research Workshop on Time Series Prediction and Analysis, ed. Addison-Wesley, Santa Fe, New Mexico, 1994.

[3] McComber P. et al., 1999, “Estimation of Transmission Line Icing at Different Sites Using a Neural Network”, Proc. of 9th ISOPE, Brest, France, pp. 599-606

[4] Haykin S., “Neural Networks – A Comprehensive Foundation”, Prentice-Hall, New York, 1999.

[5] Mozer M. C., “Neural Net Architectures for Temporal Sequence Processing”, In Proceedings of the NATO Advanced Research Workshop on Time Series Prediction and Analysis, ed. Addison-Wesley, Santa Fe, New Mexico, 1994.

[6] Wan E. A., “Time Series Prediction by Using a Connectionnist Network with Internal Delay Lines”, In Proceedings of the NATO Advanced Research Workshop on Time Series Prediction and Analysis, ed. Addison-Wesley, Santa Fe, New Mexico, 1994.

[7] Hertz J., Krogh A. and Palmer R. G., “Introduction to the Theory of Neural Computation”, Addison-Wesley, Redwood City, 1991.