CONF-960482--4
ANL/RA/CP--88854

Paper to be presented at the 1996 Society of Computer Simulation Multiconference, New Orleans, Louisiana, April 8-11, 1996.

Recurrent Neural Networks for NO_X Prediction in Fossil Plants

by

J. Reifman, J. E. Vitela, E. E. Feldman, T.Y. C. Wei
Argonne National Laboratory
9700 South Cass Avenue
Argonne, IL 60439

*Work supported by the U.S. Department of Energy, Energy Research Programs under Contract W-31-09-ENG-38.

RECURRENT NEURAL NETWORKS FOR NO_X PREDICTION IN FOSSIL PLANTS

Jaques Reifman, Javier E. Vitela, Earl E. Feldman, and Thomas Y. C. Wei
Argonne National Laboratory
Argonne, Illinois 60439

KEY WORDS: Recurrent Artificial Neural Network, NO_X Emission

ABSTRACT

We discuss the application of recurrent (dynamic) neural networks for time-dependent modeling of NO_X emissions in coal-fired fossil plants. We use plant data from one of ComEd's plants to train and test the network model. Additional tests, parametric studies, and sensitivity analyses are performed to determine if the dynamic behavior of the model matches the expected behavior of the physical system. The results are also compared with feedforward (static) neural network models trained to represent temporal information.

INTRODUCTION

Since their reintroduction by Rumelhart et al. (1986), artificial neural networks (NNs) have been widely used as an alternative approach for modeling complex and highly nonlinear dynamic systems when the exact analytical model is unknown or unavailable. NNs are prime candidates for system modeling due to their ability to approximate large classes of nonlinear functions sufficiently accurately in cases where the system is known in terms of its inputs and outputs. This is the case for emissions generated in coal-fired power plants where input variables and parameters into the process and emissions measured by Continuous Emission Monitor systems are known.

Recently, two different approaches have been proposed (Reinschmidt and Ling 1994; Adali et al. 1995) to predict NO_X emissions in fossil plants with feedforward (static) neural networks (FNNs). In Reinschmidt and Ling's approach, FNNs were used to model static relationships between changes in the control (input) variables and the corresponding new steady-state NO_X values. That is, the dynamics of the combustion process were not represented. In Adali et al.'s approach, temporal information was represented by adding inputs to the network input layer associated with input values at previous time steps. Although this approach allows FNNs to represent the system dynamics, the number of previous time steps is problem dependent and, in general, can only be obtained after extensive testing.

In this work, we apply recurrent (dynamic) neural networks (RNNs) for time-dependent modeling of NO_X emissions in fossil plants. Unlike FNNs in which the outputs depend solely on the present inputs, RNNs contain recurrent connections that together with time delays allow for the internal representation of time-dependent information. Thus, RNNs can learn complex dynamic behavior that requires internal storage of information over time periods for later use, including the mapping of ordered sequences of inputs into ordered sequences of outputs characteristic of dynamic systems.

RECURRENT NEURAL NETWORK

The supervised RNN learning algorithm developed by Williams and Zipser (1989) is used in this project. This learning algorithm only moves forward in time, therefore, it has (1) the advantage that it is simpler and easier to implement, and (2) the ability to operate continually while the network runs with no need to precisely define training intervals. The disadvantage is that it is computationally expensive to train.

Figure 1 illustrates the architecture of the RNN. The network has n processing units, with m external input lines. The network is fully connected, with a unique weight between every pair of units (including itself), and also from each input line to each processing unit. While in this architecture there is no distinction between hidden units and output units, not all units are "visible" to the external world. In general, only a subset of the n units are used as output units for which specified target values exist.

The learning algorithm is based on gradient-descent, where at each iteration the network weights, w_{ij}, connecting units j and i, are updated such that the overall network error E(k) at time step k is minimized. The weight update Δw_{ij} is proportional to its gradient component $\partial E(k)/\partial w_{ij}$, which is recursively calculated forward in time given the initial value

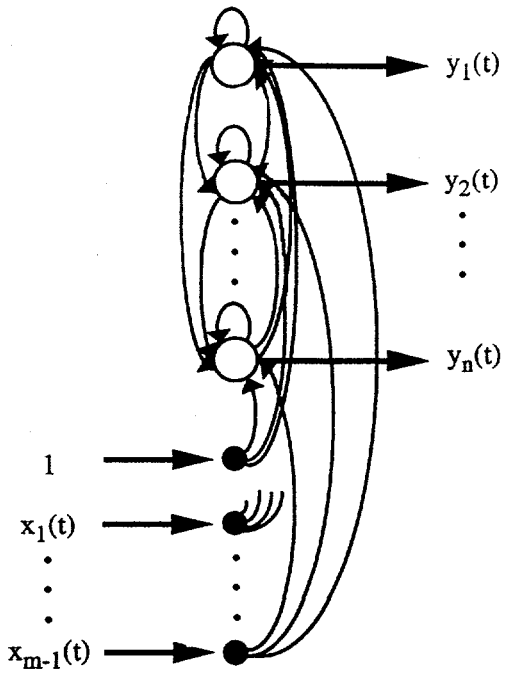

Fig. 1. Architecture of the Williams and Zipser Fully Recurrent Neural Network

$\partial E(k=0)/\partial w_{ij}$. The forward recursion in time is possible because an input at time k does not influence the output of any unit until time k+1. Therefore, the weights may be updated after each time step k, without accumulating the weight updates and making the weight changes at the end of the trajectory. For an in-depth description of the algorithm the reader should refer to Williams and Zipser (1989).

ComEd PLANT NO_X DATA

Plant data taken from tests conducted on October 10, 1993 at the ComEd Will County Unit 3 (WCU-3) coal-fired electric power plant were used in this project. WCU-3 is a 278 MWe Combustion Engineering tangentially-fired twin furnace where the coolant (water/steam) in this plant first passes through a superheat furnace, a high-pressure turbine, and then through a reheat furnace. Figure 2 shows a subset of the collected plant data used in the development of the NN NO_X models which includes: 1) NO_X in parts per million (ppm), 2) excess oxygen, i.e., O_2, in % by volume, 3) boiler master fuel flow rate (BM), which corresponds to the rate at which a mixture of coal and air are provided to both the superheat and reheat furnaces, in %, 4) the angle of the burners in the superheat furnace (ABSF), in degrees, and 5) the angle of the burners in the reheat furnace (ABRF), in degrees. The values of NO_X, O_2, and BM in the measured data are for both furnaces combined. The burner angles range from −30° to +30°, where 0° corresponds to horizontal.

As the circles on the graphs of Figure 2 indicate, 41 consecutive sets of data were measured at 2-minute intervals. In the NN models, five input quantities (NO_X, O_2, BM, ABSF, and ABRF) measured at time i are used to predict NO_X output quantities at time i+1. Thus, the 41 consecutive sets of equally-spaced data points represent 40 input/output pairs where the first point is only used as an input and the 41st is only used as an output. Prior to input into the NN, the five variables were normalized. NO_X, O_2, and BM were linearly transformed so that each normalized variable had a minimum value of 0.2 and a maximum of 0.8. ABSF and ABRF were normalized so that −30° corresponds to 0.2 and +30° corresponds to 0.8.

In the training and operation of the NN, NO_X at time i was used as an input to help predict the output NO_X at time i+1. During training, measured NO_X was used as input. After training, during operation, calculated NO_X at time i was fed back and used as input for predicting NO_X at time i+1. Using calculated NO_X as input provided essentially the same results as using measured NO_X for the existing 41 data points and allowed the trained NN to be tested under different conditions where measured data were not available.

RECURRENT NEURAL NETWORK MODEL OF NO_X PRODUCTION

With the RNN learning algorithm described above, we trained a 10-processing unit (n=10) network with 5 external inputs (plus a bias) to predict NO_X. One of the 10 processing units served as an output unit. Out of the 40 input/output pairs, the first 20 pairs were used for training and the last 20 for testing. Figure 3 shows the measured NO_X (solid line) and the network predicted values (circles), which have a maximum difference of about 15%. The first 20 circles have plus signs inside them to indicate that these data points were used for training. This explains the excellent agreement between the measured and calculated values for the first 20 points.

Although physical principles were not applied in the development of the RNN used to generate the predicted values of NO_X in Figure 3, a proper RNN NO_X model should be able to simulate the dynamic behavior of the physical process it is intended to represent. To investigate the behavior of the model and to determine the dynamic nature

of its responses, we devised a number of simple tests. The initial steady-state values for these dynamic tests, illustrated in the first row of Table I, correspond to those at time 1:40:27 a.m. of Figure 2, i.e., the first data point. A measured data point was chosen so that the tests would be initiated from an actual operating point of the power plant being modeled.

The first test was designed to show that if all four control variables (O_2, BM, ABSF, and ABRF) were held constant, NO_X would reach and maintain the same steady-state value regardless of its initial value. This behavior was observed with the RNN model through repeated simulations where the same initial steady-state NO_X value was obtained for different initial guesses of NO_X.

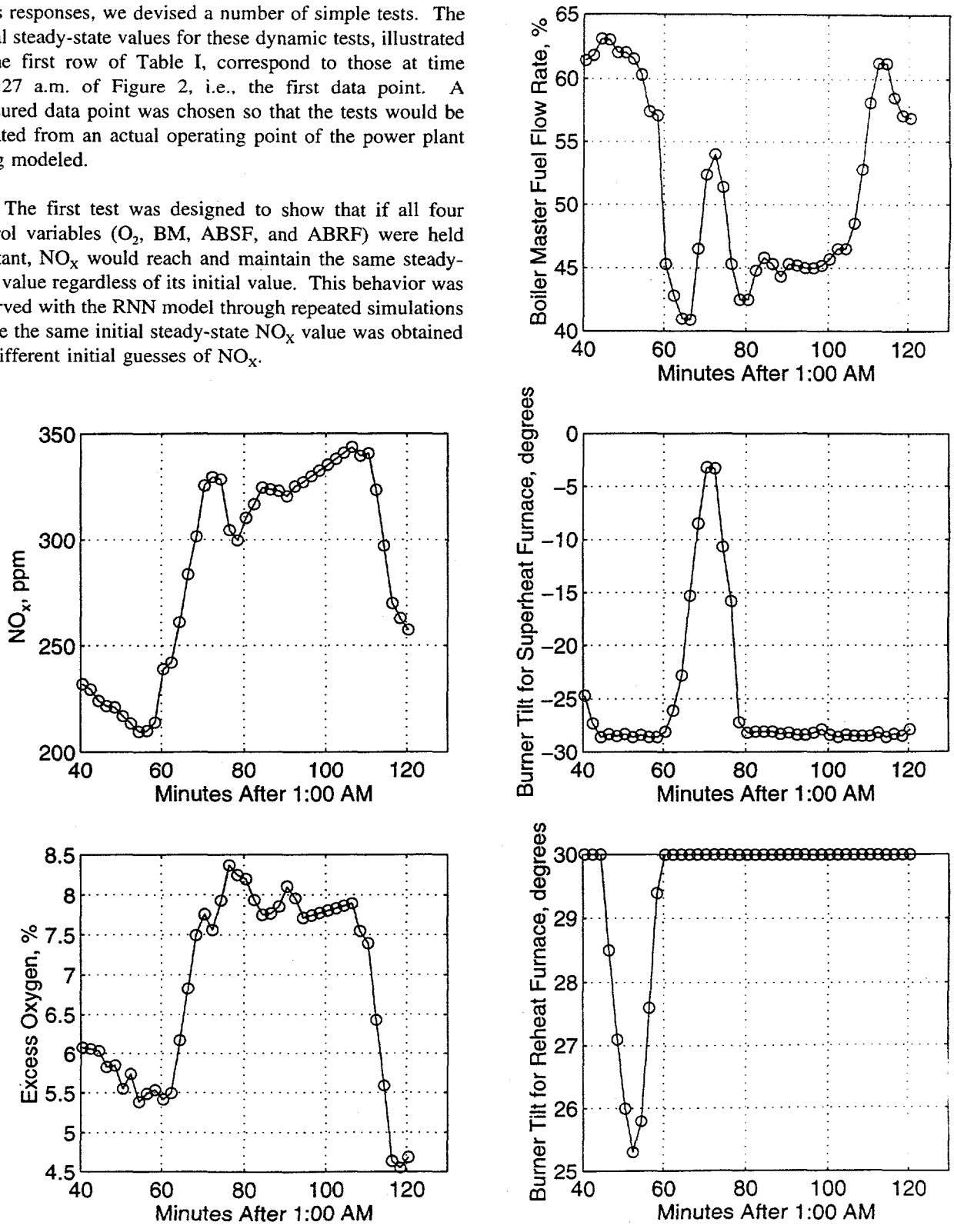

Fig. 2. Subset of Measured Plant Data Used to Develop the Neural Network NO_X Models

Table I—Local NO_X Sensitivity

Variable	NO_X	Excess Oxygen (O_2)	Boiler Master	Superheat Furnace Burner Tilt	Reheat Furnace Burner Tilt
Measured Value	232.0 ppm	6.077 %	61.4 %	-24.7 degree	30.0 degree
Sensitivity	N/A	33.22 ppm/%	-17.82 ppm/%	-2.292 ppm/deg.	3.624 ppm/deg.
Time Constant (minutes)	N/A	6.5	6.3	6.8	4.7

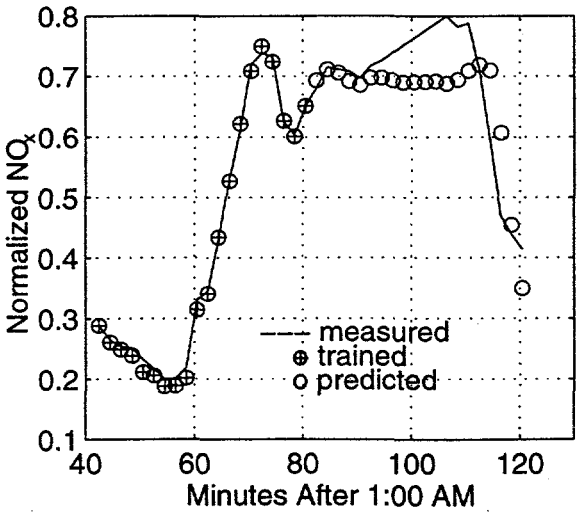

Fig. 3. NO_X Predicted by the Recurrent Neural Network

The objective of the second test was to demonstrate that if the model initially representing a steady-state NO_X condition is temporarily perturbed, by changing the value of a control variable for one time step and then immediately returned to its original value, the model returns to the same initial steady-state NO_X condition. Starting from an initial steady-state condition, each one of the four control variables was separately perturbed, one at a time, by an arbitrary amount for one time step (2 minutes) and then returned to its original value. As expected, in each case, NO_X was temporarily perturbed before returning to its initial steady-state value.

The purpose of the third test was to demonstrate that if the model is perturbed by a step change in a control variable, the model reaches a new steady-state condition. Figure 4 shows the results of such a test where the value of O_2 undergoes a step-change increase in value of 1% from its initial value (6.077%) at 0 minutes. The RNN model results show an initial reduction in NO_X, followed by an increase and an overshoot before a final steady-state value is achieved. Although it is very encouraging that a step-change in O_2 resulted in a step-change in NO_X, the initial decrease in NO_X is not expected to be observed in the physical system. The 1% step-change test was also performed for the other three control variables, producing similar results. These tests allowed us to calculate the sensitivities to a change in control variable and the associated time constants, taken to be the time required to achieve the first 63.2% (i.e., $1 - e^{-1}$) of the total change in NO_X, as illustrated in the last two rows of Table I. The time-constant results are reasonable in that experience indicates that NO_X is fully stabilized in about 20 minutes after a change in setpoint of one of the four control variables.

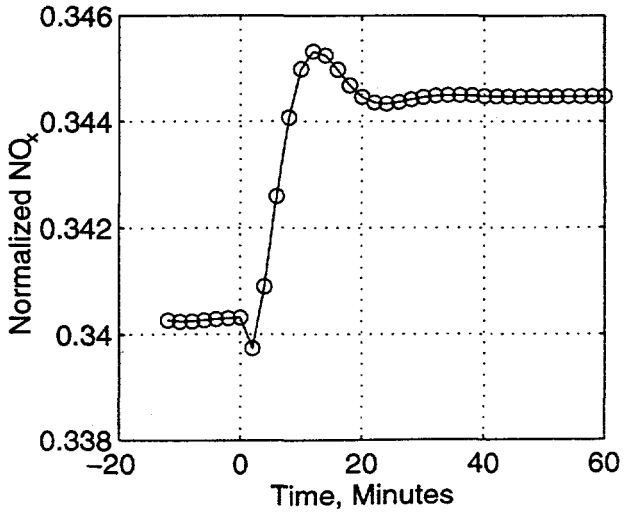

Fig. 4. NO_X Response to a 1% Step-Change in O_2 for the Recurrent Neural Network Model

In the fourth test, the O_2 input is perturbed with the same amplitude and from the same initial conditions as in the step-change test but a sine wave, whose period is 32 minutes, was used. As expected, the response of the model to a small-amplitude sine-wave disturbance is a sine wave of the same frequency, but delayed in time (see Figure 5). The dashed curve in the figure was added to indicate the shape and phase difference of the O_2 disturbance. The solid curve

indicates the network predicted NO_X. The amplitude of the NO_X output sine wave is somewhat greater than the amplitude of the step-change in NO_X from the previous test. This is physically unrealistic since gains greater than 1.0 are not expected here. Sine waves of period 8, 16, and 64 minutes were also analyzed and all produced smaller gains than that for 32 minutes. As expected, the phase (lag) for the four frequencies showed increasingly negative values with increasing frequency.

From the results of the tests, we may conclude that the RNN NO_X model behaves much like a physical model in spite of the limited amount of data available for training and testing. Since the RNN NO_X model represents a nonlinear system, the results apply to only a single operating point and the sensitivities provided in Table I are also dependent on input amplitude.

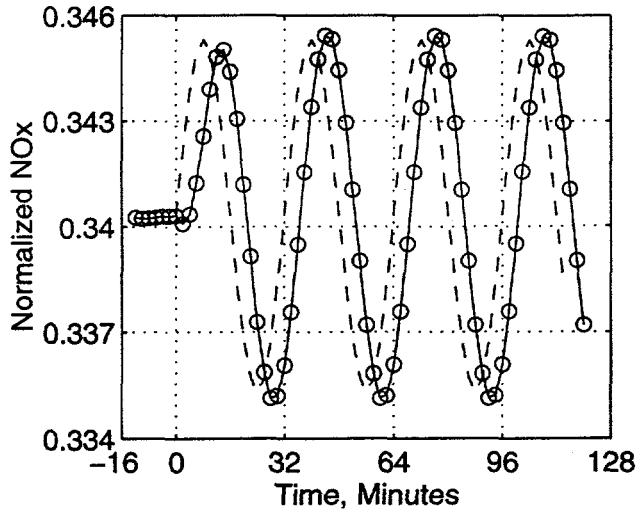

Fig. 5. NO_X Response to a Sinusoidal Perturbation in O_2 for the Recurrent Neural Network Model

FEEDFORWARD NEURAL NETWORK MODEL OF NO_X PRODUCTION

To apply a feedforward (static) network to represent the dynamics of the combustion process in the furnace, we trained a three-layer FNN in which the measured input values of NO_X and the four other variables at time steps i-2, i-1, and i were used to predict the output value of NO_X at time step i+1. Thus, the network had 15 (3 × 5) input units and 1 output unit. There were 5 units in the hidden layer. This arrangement reduces the total number of available input/output pairs from 40 to 38. The first 18 were used for training the network and the remaining 20 for testing.

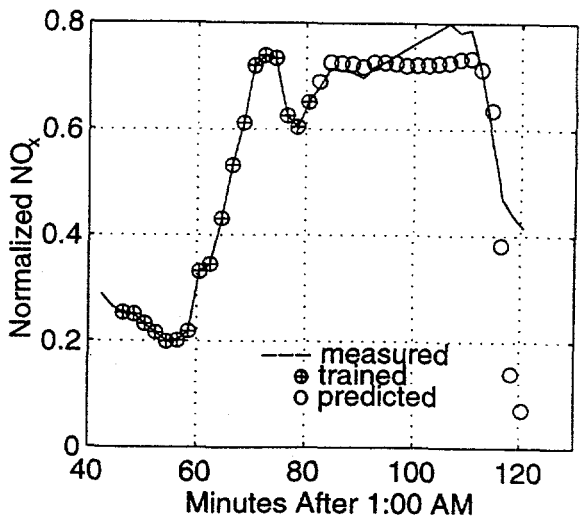

Fig. 6. NO_X Predicted by the Feedforward Neural Network

The results of the FNN NO_X model are illustrated in Figure 6. The network was trained with the conjugate gradient method (Reifman and Vitela 1994) which expedites training and automatically selects the momentum parameter and learning parameter. The predicted results for the test data are comparable to those for the RNN model in Figure 3, except for the last two time steps which have much larger errors. In spite of the reasonable results for the test data, the frequency-response tests, described below, produced unreasonable results, indicating the importance of model testing.

For testing the FNN model, the same initial steady-state values of O_2, BM, ABSF, and ABRF used for frequency response testing of the RNN are used here. The same step-change and sinusoidal tests were also repeated for the FNN model. The results of the O_2 1% step-change test, Figure 7, show a considerable amount of oscillation, which eventually dies out, that was not observed in the RNN model in Figure 4. The period of this oscillation appears to be about 16 minutes. Similar oscillations were observed for step changes in the other three control variables. Also, the time to achieve a new steady-state NO_X condition (approximately 150 minutes) is too long. The sinusoidal tests for O_2 produced a maximum gain greater than 5, which is totally unrealistic.

Although the tested FNN model produced unrealistic responses for the frequency-response tests, we should not conclude that the results can not be improved. By varying the number of time delays, network topology, and initial weights, we should be able to obtain a more realistic model.

The focus of this work was to apply RNNs to model the dynamic behavior of time-dependent processes. The FNN model was introduced only for comparison purposes.

SUMMARY AND CONCLUSIONS

This paper presents the results of our initial studies in the application of recurrent neural networks for modeling NO_X emissions in coal-fired fossil plants. Recurrent networks internally account for the representation of time-dependent information and are able to learn complex tasks that require the retention of information for later use. This is the key advantage of this type of network over the feedforward type, where temporal information can only be represented by adding inputs to the network input layer associated with an arbitrary number of previous input time steps. The precise number of time steps is problem dependent and can only be determined through extensive testing. The disadvantage of recurrent networks is that they are computationally expensive to train due to their additional recurrent connections.

The recurrent network model provided better results than those obtained with the feedforward network trained to represent time dependencies. Although the predicted results for the test data with the two types of networks were comparable, the recurrent network model provided a better match with the expected physical behavior of the dynamic process during simulated tests with step and sinusoidal changes in the control variables. However, due to the small number of feedforward network topologies tested, limited amount of data available, and key control variables and state variables possibly not being represented (reflected in the less than perfect match between measured and network predicted NO_X), these initial results need to be further investigated. A more comprehensive data set is needed to better compare the dynamic behavior of models based on the two types of networks.

ACKNOWLEDGMENTS

This project was supported by the Laboratory Directed Research and Development program at Argonne National Laboratory. The authors would like to thank J. S. Herzau, D. G. Kubik, and K. E. Wanninger from ComEd for providing plant data and technical support.

REFERENCES

Adali, T., B. Bakal, and R. Fakory. 1995. "NO_X and CO Prediction in Fossil Fuel Plants Using Artificial Neural Networks." In *Proceedings of the 1995 Society for Computing Simulation Multiconference* (Phoenix, Arizona, Apr.), 1-10.

Reifman, J. and J.E. Vitela. 1994. "Accelerating Learning of Neural Networks with Conjugate Gradients for Nuclear Power Plant Applications." *Nuclear Technology*, vol. 106, no. 2, 225-241.

Reinschmidt, K.F. and B. Ling. 1994. "Neural Networks for NO_X Control." In *Proceedings of the 1994 American Power Conference* (Chicago, Illinois, Apr. 25-27), vol. 56-I, 354-359.

Rumelhart, D.E., G.E. Hinton, and R.J. Williams. 1986. "Learning Internal Representations by Error Propagation." *Parallel Distributed Processing: Explorations in the Microstructure of Cognition.* D. E. Rumelhart and J. C. McClelland, eds., The MIT Press, Cambridge, Massachusetts, vol. I, 319-362.

Williams, R.J. and D. Zipser. 1989. "A Learning Algorithm for Continually Running Fully Recurrent Neural Networks." *Neural Computation*, vol. 1, 270-280.

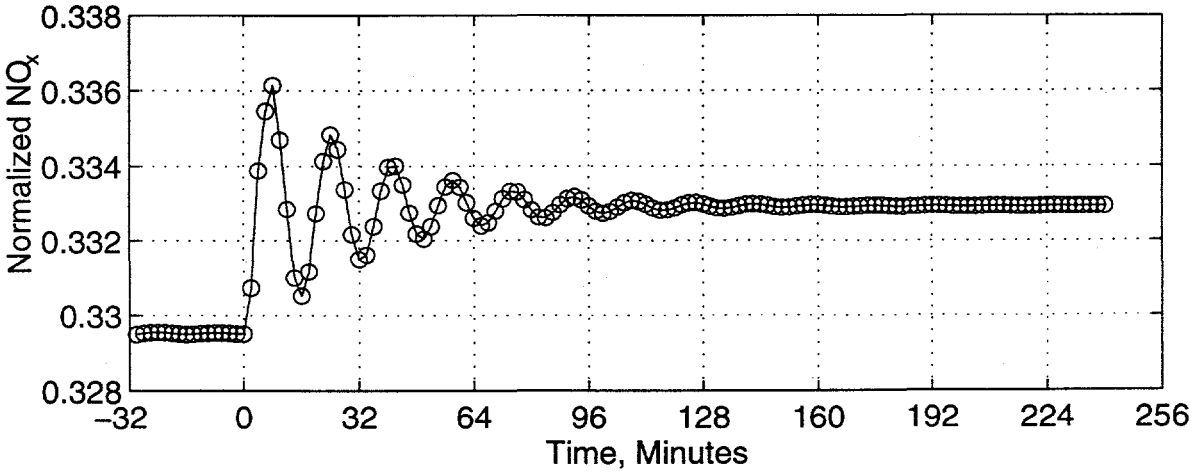

Fig. 7. NO_X Response to a 1% Step-Change in O_2 for the Feedforward Neural Network Model

DISCLAIMER