

A GAMMA MEMORY NEURAL NETWORK FOR SYSTEM IDENTIFICATION

Mark A. Motter
Facility Automated Controls
NASA Langley Research Center
Hampton, Virginia 23681

Jose C. Principe
Department of Electrical Engineering
University of Florida
Gainesville, Florida 32611

ABSTRACT

A gamma neural network topology is investigated for a system identification application. A discrete gamma memory structure is used in the input layer, providing delayed values of both the control inputs and the network output to the input layer. The discrete gamma memory structure implements a tapped dispersive delay line, with the amount of dispersion regulated by a single, adaptable parameter. The network is trained using static back propagation, but captures significant features of the system dynamics. The system dynamics identified with the network are the Mach number dynamics of the 16 Foot Transonic Tunnel at NASA Langley Research Center, Hampton, Virginia. The training data spans an operating range of Mach numbers from 0.4 to 1.3.

INTRODUCTION

Most neural network-based identification models for dynamic systems employ tapped delay lines to present delayed versions of the either the system input, system output, or both, to the input layer of a neural network [Narendra, 1991]. In this paper, we investigate the use of the *gamma memory* [De Vries and Principe, 1992] instead of a tapped delay line to provide the pertinent input-output history to the neural network. The output of each tap of a discrete gamma memory structure can be expressed by the following recurrence relations:

$$x_0(t) = x(t); \quad x_k(t) = (1 - \mu)x_k(t-1) + \mu x_{k-1}(t-1); \quad k = 1, \dots, K.$$

The discrete gamma memory structure, shown in Figure 1, can be seen as a tapped dispersive delay line. The gamma memory allows a tradeoff between memory resolution and memory depth via the parameter μ . For a K -th order gamma memory, the memory depth D is well approximated by $D = K / \mu$ and the resolution is $R = \mu$. Thus, for a K -th order gamma memory, the balance between depth and resolution of the memory is characterized by the relation: $K = D \times R$. This relationship can be used to significantly reduce the size of the input layer to the network while still providing the necessary memory depth.

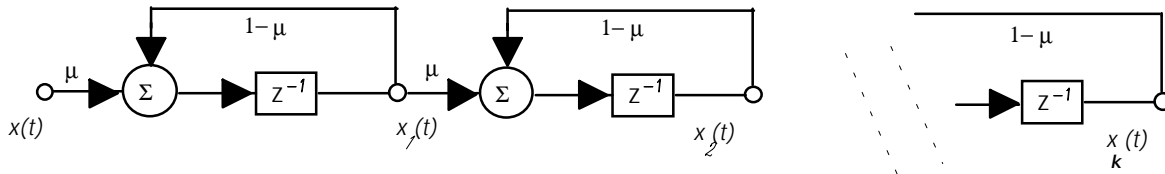


Figure 1. Discrete gamma memory structure

The 16-Foot Transonic Tunnel at the NASA Langley Research Center, Hampton, Virginia, is a closed circuit, single-return, continuous-flow, atmospheric tunnel with a Mach number capability from 0.20 to 1.30. When the tunnel began operation in November 1941, it had a circular test section that was 16 feet in diameter and maximum Mach number of 0.71. [Peddrew, 1981] Numerous upgrades to both the test section and drive system have expanded the test envelope of this facility. Currently, Mach numbers up to 1.05 are achieved using the tunnel main drive fans only. Mach numbers from 1.05 to 1.3 require the combination of test section plenum suction with the tunnel fans. The tunnel fans, 34 feet in diameter, are driven from 60 to 372 rpm by a 50 MW electric drive system. An air removal system using a 30 MW compressor and 10-Foot diameter butterfly valve

provides test section plenum suction . At Mach numbers above 1.275, the 10-Foot valve is fully open and increases in Mach number are obtained from increased power to the tunnel main drive fans.

Some of the salient features of the Mach number dynamics of the 16 Foot Tunnel are:

- The nominal dynamics vary significantly over the operational range of the tunnel;
- The control input to the tunnel fan drive system is bang-zero-bang: (+1 raise, -1 lower, 0 maintain speed);
- There is transport lag (pure delay) that varies over the operational range;
- The dynamics can change dramatically at any given operating point due to test conditions (blockage);
- The test section Mach number computed from pressure measurements is noisy;
- Power consumption is significant: 20 MW @ Mach 0.7, 80 MW @ Mach 1.3.

The test section Mach number, referred to in the rest of this paper simply as *the* Mach number, is computed from a calibrated ratio of two measured quantities, the airstream stagnation pressure and the plenum or tank static pressure. The ratio of these two measured quantities is referred to as the plenum Mach number. A tabulated wind-tunnel calibration provides the correlation between the test section airstream Mach number and the plenum Mach number. The relationship between the calibrated pressure ratio and test section Mach number, M , is [John, 1984; Mercer *et al.*, 1981]:

$$\frac{P_{stagnation}}{P_{static}} = \left(1 + \frac{\gamma - 1}{2} M^2 \right)^{\frac{\gamma}{\gamma - 1}} ; \quad M = \sqrt{5 \left[\left(\frac{P_{stagnation}}{P_{static}} \right)^{\frac{-2}{\gamma}} - 1 \right]} ; \quad \gamma_{air} = 1.4 .$$

A large volume of test data relating the tunnel fan drive system control input (+1, 0, -1) , and the Mach number, is available for nominal operating conditions over most of the operating range. The objective of this investigation was to determine if a single network, employing gamma memory to provide the pertinent input-output history , could capture some significant features of the dynamics of this process over its entire operating range.

NETWORK ARCHITECTURE

The static feed forward network consisted of fifty input nodes, a hidden layer with eight neurons, and a single linear output node. The neurons in the hidden layer used the logistic function to implement the sigmoid non linearity. The discrete gamma memory structure contained nineteen gamma memory units for the control input ($m=19, \mu=0.8$), nineteen gamma units for the integral-of-the-input ($m=19, \mu=0.8$), and ten gamma memory units for providing past values of the network output to the input layer ($n=10, \mu=1$), i.e. a tapped delay line. Thus, the network architecture provided the most general form of a system identification model, where the output is expressed as a non linear function of both the input and output histories. The integral-of-the-input was included based on a priori knowledge of the plant characteristics. A diagram of the network architecture is shown in Figure 2.

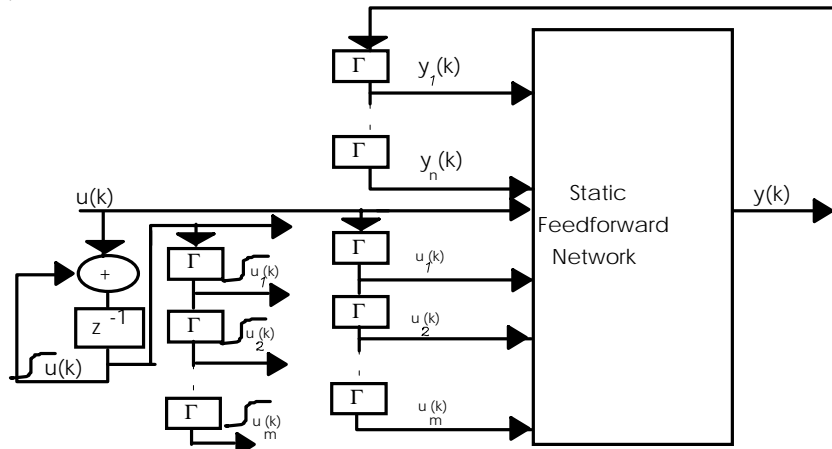


Figure 2. Network Architecture

TRAINING

The network was trained using input-output data sequences obtained from measurements of the 16 Foot Transonic Tunnel. The input data for training the network was the control input sequence, consisting of either a raise (+1), lower (-1), or no change (0) command to the tunnel drive system. The output data for training the network was the Mach number resulting from the control input sequence. Each training sequence was composed of fifty points, each point separated in time by approximately 0.25 seconds. Four typical training sequences are shown in Figure 3.

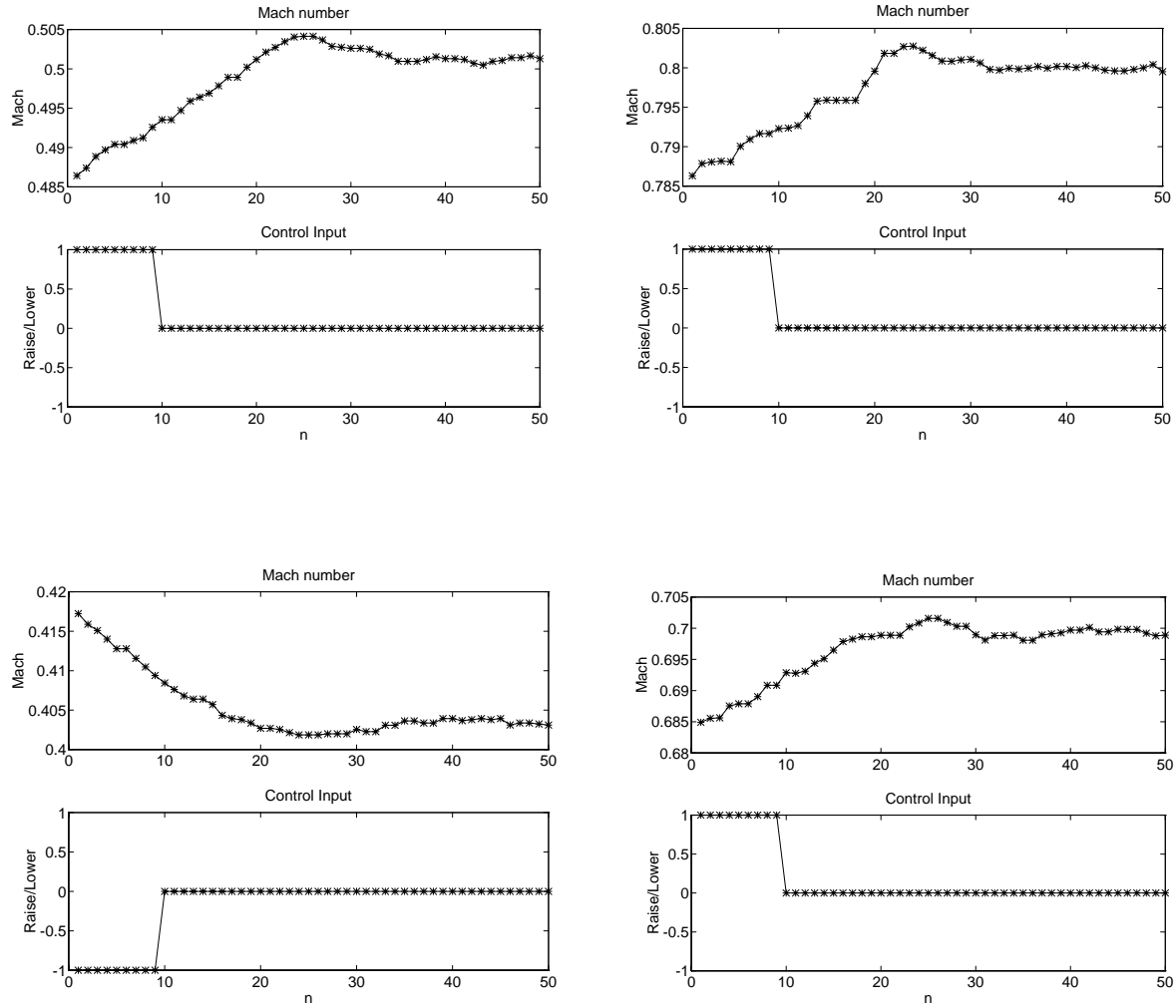


Figure 3. Four typical training sequences

Twenty input-output sequences were identified to represent the operating characteristics of the tunnel over a range of Mach numbers from 0.4 to 1.3. The first ten points of the input-output sequence were clocked in to the both the input and the output gamma memory units of the network. This provided the network with the immediate dynamic history of the tunnel. The next forty points of the training sequence were used to train the network, point by point, i.e. back propagation [Rumelhart, 1986] was employed to adjust the network weights after the presentation of each remaining point in the sequence. Thus, the network was trained to identify the response of the tunnel to control input sequences ranging in duration from zero to ten, where ten was experimentally determined to be sufficiently long to characterize the pertinent input-output history. The input is non-zero only

during the first ten points of the training sequence. Figure 4 shows the system identification model for training the network, where the network weights were adjusted for the training sequence index , n , equal to 11, 12, ... , 50.

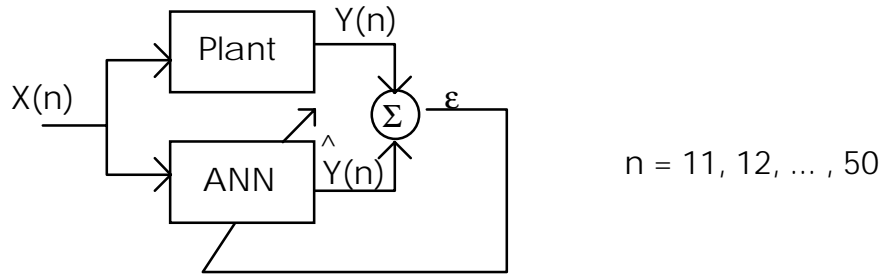


Figure 4. System Identification Model for Training the Network

The network weights were initialized to small random values . The learning rate was fixed to 0.01 for the first 100 passes through the 20 input-output training sequences. An adaptive learning rate was used after the first 100 training epochs, where the learning rate was increased by one percent if the normalized mean-squared error decreased after a training epoch and, conversely, the learning rate was decreased by half when the normalized MSE increased by more than one percent. The network was trained until a normalized mean-squared error of $1e-05$ was achieved. This level of performance was achieved after 2500 epochs, or equivalently, the presentation of 50 000 training sequences. Figure 5 shows the MSE and the learning rate during the last 300 epochs, illustrating the effect of the variable learning rate.

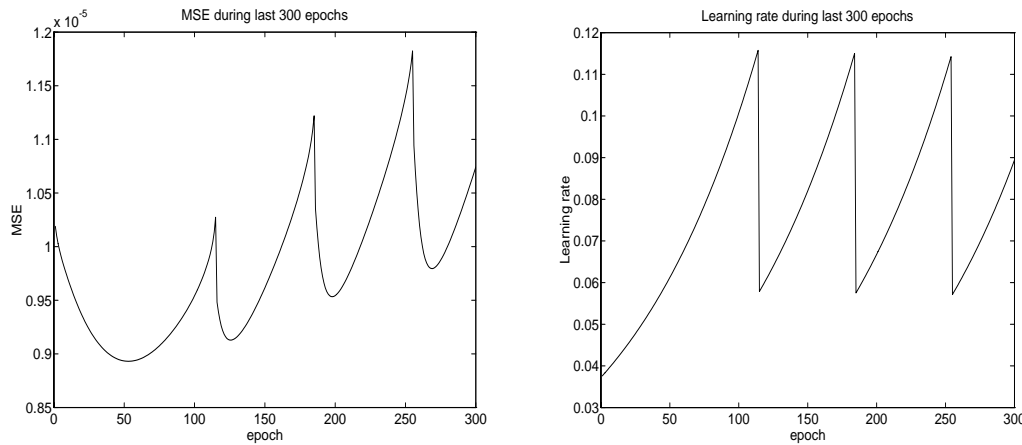


Figure 5. Normalized Mean-Squared Error and Learning Rate during the last 300 epochs

The network responses corresponding to the training sequences shown earlier are shown with their respective tunnel Mach number responses in Figure 6. Two network responses from sequences not in the training set are also shown. Typical of all responses, over the entire operating range, is the network capability to estimate the final value of the Mach number response associated with a given control input, along with some capacity to model the transient. The first ten points of each the network outputs shown is identical to the corresponding tunnel Mach number, illustrating the effect of initializing the output memory with the past ten process measurements. The network then predicts the output trajectory for the next 40 points.

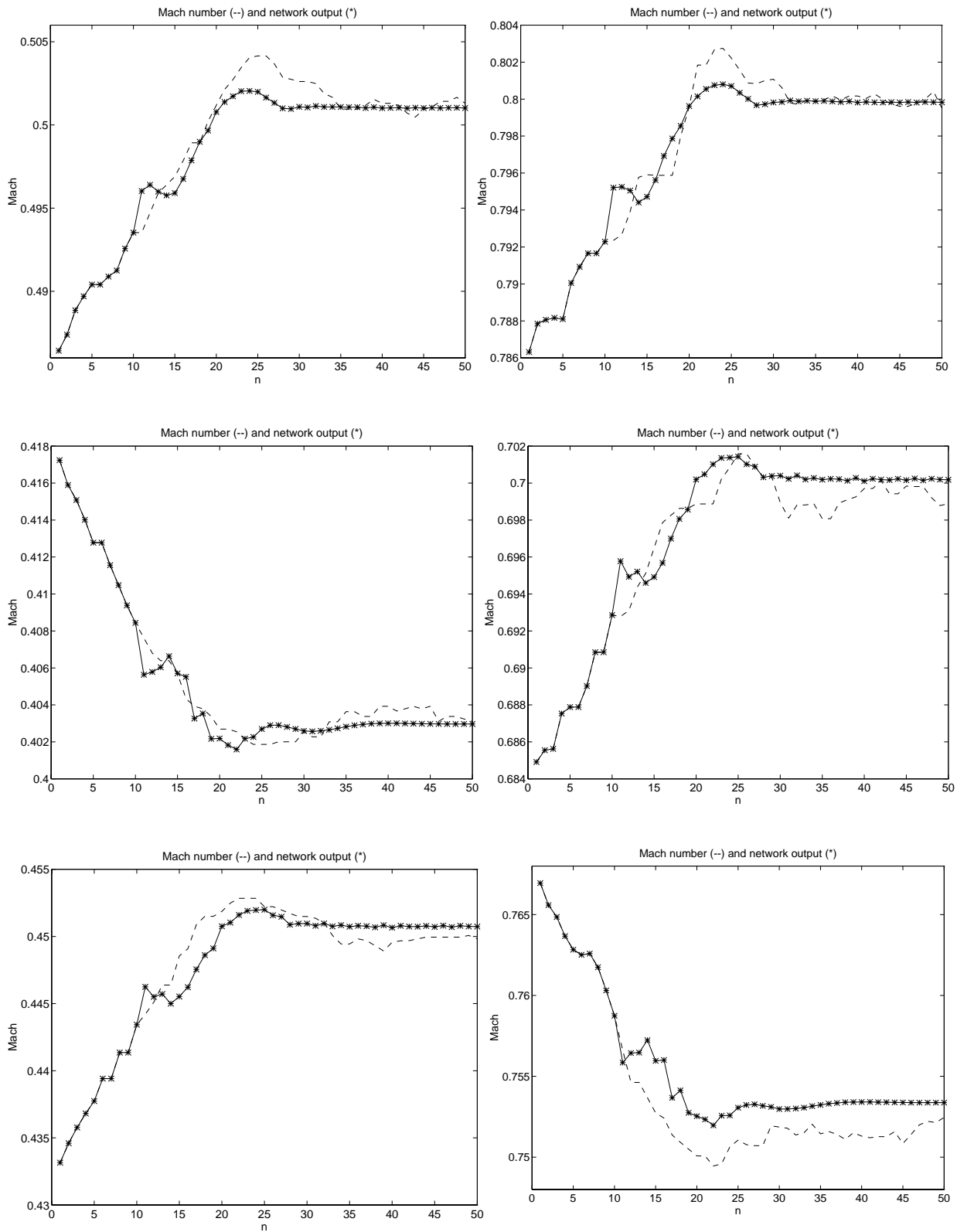


Figure 6. Typical Network Responses

CONCLUSION

The neural network topology described in this paper shows promise for identification of system dynamics over a wide operating range. A *gamma memory* was employed to vary memory depth versus resolution without changing the number of input nodes of the network. The network captured some features of the tunnel dynamics and predicted the steady-state changes in Mach number resulting from control inputs of varying duration to within 0.001, over a wide portion of the operating range of the tunnel. The current tunnel control algorithm controls the Mach number to within 0.003, indicating that this identification model could be used to enhance the current controls.

Fixed-point learning was used to train the network, as opposed to back propagation through time (BPTT) [Werbos, 1990; Hertz, 1991]. Further work will investigate the effectiveness of employing BPTT to enhance the network capability to capture transient features over the entire operating range.

ACKNOWLEDGEMENT

This work was partially supported by NSF grant 920878.

REFERENCES

- De Vries, B. and Principe, J. C. (1992). "The Gamma Model - A New Neural Model for Temporal Processing," *Neural Networks*, Vol. 5, pp.565-576
- Hertz, J., Krough, A., and Palmer, R. G. (1991). *Introduction to the Theory of Neural Computation*. Reading, MA: Addison-Wesley
- John, James E.A. (1984). *Gas Dynamics*. Newton, MA : Allyn and Bacon, Inc.
- Mercer, Charles E. , *et al.* (1984) . *Computations for the 16-Foot Transonic Tunnel - NASA Langley Research Center*. NASA Technical Memorandum 86319
- Narendra, K.S. (1991). "Adaptive Control Using Neural Networks," in *Neural Networks for Control*, W. Thomas Miller, Richard S. Sutton, and Paul J. Werbos, editors. Cambridge, MA: The MIT Press, 1991
- Peddrew, Kathryn H. (1981). *A User's Guide to the Langley 16-Foot Transonic Tunnel*. NASA Technical Memorandum 83186
- Rumelhart, D.E. , Hinton, G.E., and Williams, R.J. (1986). " Learning internal representations by error back-propagation," in *Parallel Distributed Processing*, D.E. Rumelhart and J.L. McClelland, editors. Cambridge, MA: The MIT Press
- Werbos, P.J. (1990). "Back Propagation Through Time: What it does and how to do it," in *Proceedings of the IEEE*, 78(10), 1550-1560.