Korean J. Chem. Eng., **17**(6), 696-703 (2000)

Hybrid Neural Network Approach in Description and Prediction of Dynamic Behavior of Chaotic Chemical Reaction Systems

Hwi Jin Kim and Kun Soo Chang†

Department of Chemical Engineering and Automation Research Center,
Pohang University of Science and Technology, San 31 Hyoja Dong, Pohang 790-784, Korea
(*Received 30 June 2000 • accepted 29 August 2000*)

Abstract−A chaotic system with available prior knowledge is identified with both the sequential hybrid neural network and the standard Artificial Neural Network (ANN). The identified models are validated with phase portrait, return map, the largest Lyapunov exponent and correlation dimension instead of using Sum of Square Errors (SSE). Interpolation and Extrapolation capability of the models are compared. This is demonstrated for nonisothermal, irreversible, first-order, series reaction A→B→C in a CSTR.

Key words: Chaos, Hybrid Neural Network, Modified Error Back Propagation, System Identification, Validation

INTRODUCTION

Industrial chemical processes involving chemical reactions, heat and mass transfer, separations and fluid flow are inherently and strongly nonlinear and exhibit complicated dynamic behavior. In the past, a considerable number of studies have been carried out for the processes showing multiple steady states, oscillatory behavior and chaos [Kim, 1998]. Most of the studies were based on the mathematical models of systems derived from governing physical laws. In actual industrial processes, however, it is usually very difficult to obtain rigorous mathematical models of the systems because of both the complexity of the systems and the lack of available system parameters. An alternative method is to use the standard black-box Artificial Neural Network (ANN) based only on the input-output data of the systems. Recently, it has been widely used as a universal function approximator when there is no prior knowledge about the systems because of its ability to describe nonlinear systems. It has been proved that the standard ANN can approximate arbitrary complex functions well and describe even complex nonlinear phenomena such as steady state multiplicity and oscillatory behavior only if the internal parameters such as the number of inputs, neurons, layers and the transfer functions of neurons are properly chosen, and a sufficiently large data set with desired property is available. This inherent capability of the standard ANN is due mainly to the combination of nonlinear transfer functions used for each node. The standard ANN also has noise smoothing effect if the internal parameters are properly chosen or if the batch mode learning of back-propagation is used.

It has, however, many disadvantages. If the training data set does not have proper quality and the internal parameters are not properly chosen, the standard ANN suffers from serious malfunction. For example, if the training data set is corrupted with noise and the number of internal parameters is more than needed, the standard ANN fits the noise as well as the system dynamics. It is also impossible to realize what kind of interaction occurs between process variables since all components of the standard ANN are thought to be partially responsible for the output of the network. Although the standard ANN can be simplified by various pruning techniques, such as sensitivity analysis of weights with respect to the output error of ANN, we still cannot give physical meanings to the resulting network. Therefore, we cannot guarantee the extrapolation capability of the standard ANN beyond the limits of training data, even in case the standard ANN is trained very well.

Another recourse is to use the hybrid (structured) neural network approach. If prior knowledge about a system is available, it is smart to incorporate the prior knowledge into the black box model of the system. Recently, there have been many attempts and an excellent summary of the subject is given in the paper of Thompson and Kramer, 1994. In the hybrid neural network approach, the first principle models from physical considerations such as mass and energy balance or empirical correlation are used as prior knowledge about a system, and the ANN model complements the uncertain parts of the first principle models. It is also possible to regard the hybrid neural network model as the ANN constrained by the first principle models. There can be several approaches in the actual implementation of the hybrid neural network, but here we deal with only the sequential hybrid neural network approach, where the ANN model is combined with the first principle models in series. In the approach, the ANN serves as a nonparametric estimator of the unmeasured process parameters which are the intermediate values to be used in the first principle models, and estimates the dependence of the process parameters on the state variables of the system. In this sense, the approach provides more general parameter estimation strategy and usually gives better estimations than classical parameter estimation schemes such as nonlinear programming (NLP) optimization and kalman filter parameter estimation [Psichogios and Ungar, 1992]. The approach has many advantages. Because the ANN component approximates only the uncertain parts of the first principle models, the size of the ANN can be drastically reduced. Therefore, the training is more focused, and then potential error sources are greatly reduced. As a result, the sequential hybrid neural network model usually shows better performance and is more robust to noise than the standard ANN model. Moreover, be-

†To whom correspondence should be addressed.
E-mail: kschang@postech.ac.kr

cause the hybrid model is based mainly on the first principle models, it can also be used for extrapolation purpose as well as interpolation.

In any ANN, important steps are in the selection of appropriate number of layers and of neurons in each layer, the choice of the transfer function used for each neuron and the training algorithm in order to obtain a good identified model, and also the validation of the model. Usually, a trial and error procedure based on criterion of minimization of sum of squares of ANN training errors and comparison of the time series of the original system with the model generated time series by calculating the mean square error between them are used for this purpose. For chaotic systems, however, this criterion may not provide useful information. Identified models can show different dynamical behaviors even though the training errors are roughly the same, and the criterion of minimization of the mean square error between the time series is just a necessary condition for an identified model to capture the dynamical properties of the system; it is definitely not sufficient. In the case of a chaotic system, although the initial prediction of an identified model can be very accurate, predicted values diverge from the original time series at much later prediction times no matter how good the model is. This is due to the inaccuracies in the model and the existence of positive Lyapunov exponents. Because nearby trajectories diverge locally in state-space for a chaotic system, the initial error due to the modeling error, however small, is magnified. The model generated time series thus becomes completely different from the original time series in the long run. Therefore, more sophisticated criteria are required. One of the criteria is to compare attractors (phase portraits), and reconstructed attractors (or return maps, the 2-dimensional projection of reconstructed attractors). Because there exists a smooth invertible transformation between the original states and the reconstructed states with appropriately chosen delay time and embedding dimension, we can check if an identified model captures the original dynamic behavior of the system by comparing the reconstructed attractors. In many cases, however, although the location and the overall shape of the attractors look similar and thus the dynamic behavior of the system seems to have been reasonably captured, detailed characteristics such as the density of trajectories in some region of the attractor and the local divergence rate of nearby trajectories are somewhat different. Therefore, other criteria like Lyapunov exponent and correlation dimension that quantify numerically the matching between the dynamic behaviors are also required. From the criteria, we can determine and validate the optimal ANN model describing the system's chaotic dynamical behavior.

In this paper, we identify a chaotic chemical reaction system with both the sequential hybrid neural network and the standard ANN, and validate the identified models with the criteria used for nonlinear dynamics instead of sum of square error (SSE). Then we compare the interpolation and the extrapolation capability of the optimal sequential hybrid neural network model with those of the optimal standard ANN model.

THEORETICAL BACKGROUNDS

In the sequential hybrid neural network, the ANN component estimates the unmeasured process parameters which are intermediate values to be used in the first principle model. The inputs to the ANN component are current state variables and current manipulated variables, and then the ANN estimates current process parameters. The obtained parameter values are considered as constants between sampling instants. Then the parameter values with the inputs to the ANN component are propagated through the first principle model. The outputs of the first principle model are the estimates of the process state variables for the next sampling time. The schematic representation of the sequential hybrid neural network is given in Fig. 1.

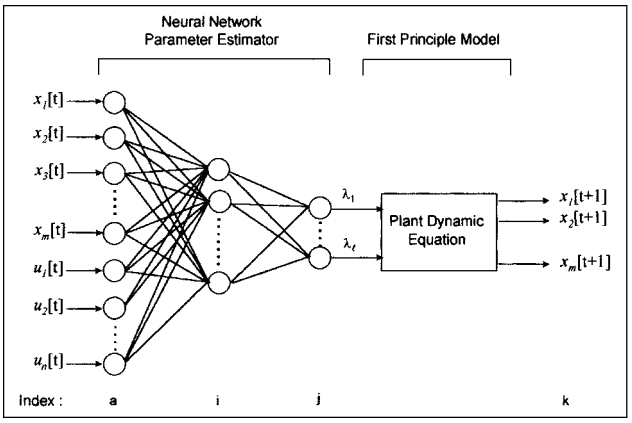

Fig. 1. Schematic representation of the sequential hybrid neural network.

1. Modified Error Back Propagation

In the standard ANN, weights are updated by using the error signals between the outputs of the ANN and the target values as driving force. One of the most famous methods is the error back-propagation algorithm [Runmelhart et al., 1986], where the output errors of the ANN are back-propagated through the network so that weights are updated in the local direction of steepest descent of the error signals. In the sequential hybrid neural network, however, the standard error back-propagation algorithm cannot be applied directly, because the target values of the outputs of the ANN component are unmeasured process parameters and therefore the output errors of the ANN component are not directly available. Therefore, the modified error-back propagation algorithm [Psaltis et al., 1988] is introduced, where the errors between the outputs of the first principle model part (the plant) and the target values of the process state variables are translated into the error signals for the outputs of the ANN component by Jacobian (differential gain) of the plant. In the algorithm, the plant can be thought of as an additional, but unmodifiable, layer since the output errors of the plant are propagated back through the plant without modifying anything. The translated error signals are then used as the driving force to update the weights of the ANN component. More details about the modified back propagation algorithm are given as follows. The objective function to be minimized can be expressed as

$$J = \frac{1}{2}\sum_k (d_k - y_k)^2 \tag{1}$$

where y_k and d_k denote the k th plant output and the k th target value, respectively. If we differentiate the objective function (J) with respect to the ANN's weight between the i th neuron of the m−1 th

layer and the j th neuron of the m th layer, w_{ij}^m, we obtain

$$\frac{\partial J}{\partial w_{ij}^m} = -\sum_k \varepsilon_k \frac{\partial y_k}{\partial w_{ij}^m} \tag{2}$$

where $\varepsilon_k = d_k - y_k$ is the k th plant output error and m denotes the final layer of the ANN. From the chain rule, the gradient becomes

$$\frac{\partial J}{\partial w_{ij}^m} = -\sum_k \varepsilon_k \frac{\partial y_k}{\partial w_{ij}^m} = -\sum_k \varepsilon_k \frac{\partial y_k}{\partial \lambda_j}\frac{\partial \lambda_j}{\partial w_{ij}^m} = -\frac{\partial \lambda_j}{\partial w_{ij}^m}\sum_k \varepsilon_k \frac{\partial y_k}{\partial \lambda_j} \tag{3}$$

where λ_j is the j th output of the ANN, that is the j th unmeasured process parameter, and $\partial y_k/\partial \lambda_j$ is the differential gain of the plant. If we define p and q as the input to the neuron and the output from the neuron, respectively, the weighted sums of the outputs of the m−1 th layer, p_j^m, and the output of the i th neuron of the m th layer, q_i^m can be written as

$$p_j^m = \sum_i w_{ij}^m q_i^{m-1}, \quad q_j^m = f_j^m(p_j^m) \tag{4}$$

where f_j^m is the transfer function of the j th neuron of the m th layer. If we incorporate the above notation into the gradient expression,

$$\frac{\partial J}{\partial w_{ij}^m} = -\frac{\partial \lambda_j}{\partial w_{ij}^m}\sum_k \varepsilon_k \frac{\partial y_k}{\partial \lambda_j} = -f_j^{m'}(p_j^m) q_i^{m-1} \sum_k \varepsilon_k \frac{\partial y_k}{\partial \lambda_j} = -q_i^{m-1}\delta_j^m \tag{5}$$

where, $\delta_j^m = f_j^{m'}(p_j^m)\sum_k \varepsilon_k \frac{\partial y_k}{\partial \lambda_j}$

And therefore the amount of weight updated is

$$\Delta w_{ij}^m = -\alpha \frac{\partial J}{\partial w_{ij}^m} = \alpha q_i^{m-1}\delta_j^m \tag{6}$$

where α is the learning rate. For all other layers, the gradient of the objective function with respect to the weight and the amount of weight updated can be derived from the similar procedure as above, and the results are given as follows.

$$\frac{\partial J}{\partial w_{ai}^{m-1}} = -f_i^{(m-1)'}(p_i^{m-1}) q_a^{m-2} \sum_j \delta_j^m w_{ij}^m = -q_a^{m-2}\delta_i^{m-1} \tag{7}$$

where, $\delta_i^{m-1} = -f_i^{(m-1)'}(p_i^{m-1}) q_a^{m-2} \sum_j \delta_j^m w_{ij}^m$

$$\Delta w_{ai}^{m-1} = -\alpha \frac{\partial J}{\partial w_{ai}^{m-1}} = \alpha q_a^{m-2}\delta_i^{m-1} \tag{8}$$

In numerical calculation, the differential gain, $\partial y_k/\partial \lambda_j$, is usually approximated by determining how the plant outputs change as the unmeasured parameters change at the operating point, that is, the numerical derivative.

$$\frac{\partial y_k}{\partial \lambda_j} = \frac{y_k(\lambda_j + \Delta\lambda_j) - y_k(\lambda_j)}{\Delta \lambda_j} \tag{9}$$

PROCESS MODEL

We consider the dynamic behavior occurring in a nonisothermal CSTR with two irreversible consecutive first-order reactions, A→B→C; the first exothermic, the second endothermic. The system can be described by the following dimensionless differential equations [Kahlert et al., 1981]:

$$\frac{dx_1}{dt} = 1 - x_1 - Dax_1 \exp\left[\frac{x_3}{1+\varepsilon x_3}\right] \tag{10}$$

$$\frac{dx_2}{dt} = -x_2 + Dax_1 \exp\left[\frac{x_3}{1+\varepsilon x_3}\right] - DaSx_2 \exp\left[\frac{\kappa x_3}{1+\varepsilon x_3}\right] \tag{11}$$

$$\frac{dx_3}{dt} = -x_3 + DaBx_1 \exp\left[\frac{x_3}{1+\varepsilon x_3}\right] - DaB\alpha Sx_2 \exp\left[\frac{\kappa x_3}{1+\varepsilon x_3}\right] - \beta(x_3 - u) \tag{12}$$

where the variables x_1, x_2 denote the dimensionless concentrations of species A, B, x_3 is the dimensionless temperature in the reactor, Da is the Damköhler number, ε is the dimensionless activation energy, S is the ratio of the two rate constants, κ is the ratio of activation energies, B is the dimensionless adiabatic temperature rise, α is the ratio of heat effects, β is the dimensionless heat transfer coefficient, and u is the dimensionless coolant bath temperature and can be viewed as an externally manipulable variable. The system is known to show deterministic chaos when the system parameter values are Da=0.26, ε=0.0, S=0.5, κ=1.0, B=57.77, α=0.42, β=7.9999, and u=0.0, that is, when there is no control action. Fig. 2-3 show the 3-D phase portrait of the system and the 2-D projection of the 3-D phase portrait, respectively. The simulation was carried out on IBM RS6000/370 using the IMSL subroutine ode_adams_gear. Fig. 4 shows the second return map of the state variable x_3.

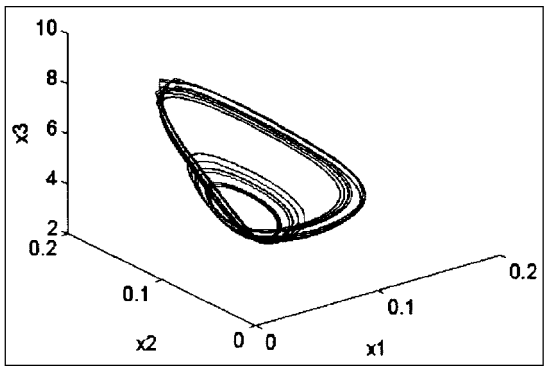

Fig. 2. 3-D phase portrait of the system.

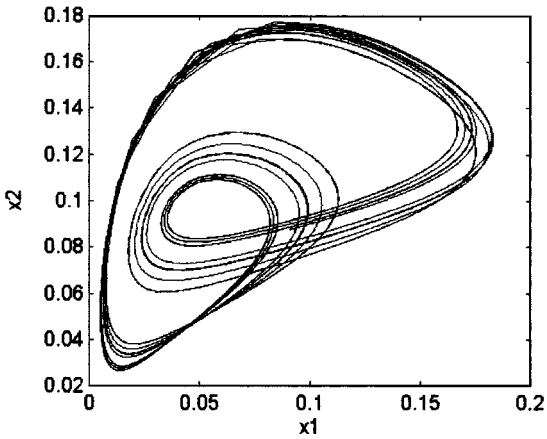

Fig. 3. x_1−x_2 plane projection of 3-D phase portrait of the system.

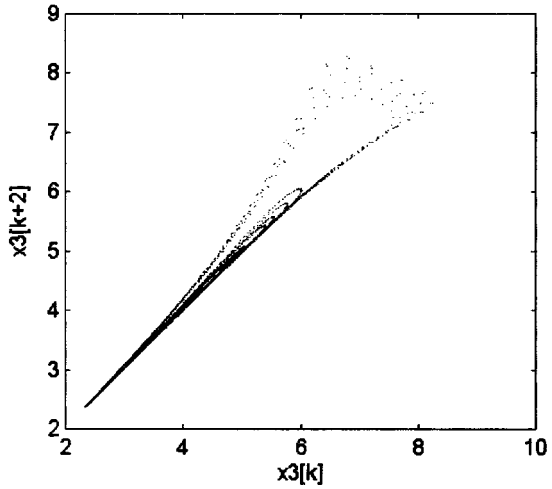

Fig. 4. Second return map of the system.

The largest Lyapunov exponent and correlation dimension are calculated for the time series data of the system using in-house implementations of the Wolf's algorithm [1985] for the largest Lyapunov exponent and the Grassberger and Procaccia algorithm [1983] for correlation dimension. The obtained values are 0.00446 for the largest Lyapunov exponent and 1.535 for correlation dimension as summarized in Table 1. The bifurcation analysis of the system

Table 1. Summary of the largest Lyapunov exponent and correlation dimension

	The largest Lyapunov exponent	Correlation dimension
Original system	0.00446	1.535
Hybrid model with h=4, (trained with full data)	0.00446	1.541
Hybrid model with h=7, (trained with full data)	0.00436	1.854
Standard ANN with h=8, (trained with full data)	0.003942	1.402
Hybrid model with h=4, (trained with partial data)	0.005129	1.301

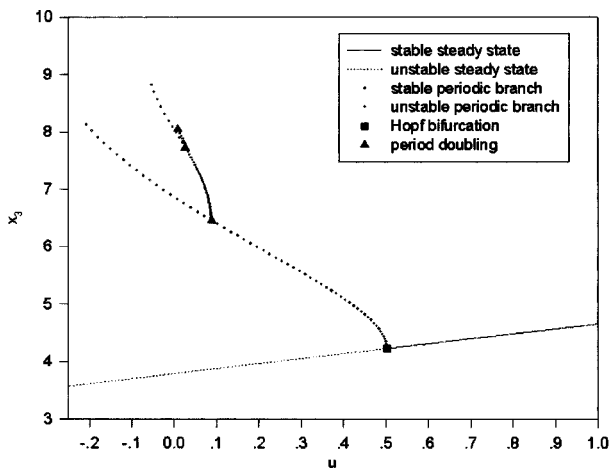

Fig. 5. Bifurcation diagram of the system.

equation is also carried out by using numerical continuation techniques which are implemented in the software package AUTO [Doedel, 1986]. From the analysis, we obtain the bifurcation diagram as shown in Fig. 5. All the detailed analysis is given in the paper of Kim, 1998.

SYSTEM IDENTIFICATION

We assume that the first principle model of the system in the same form as above is available from the mass and energy balance. All of the parameters are assumed to be available from the individual experiments or the literature except the ratio of the two rate constants, S. The parameter values are Da=0.26, ε=0.0, κ=1.0, B=57.77, α=0.42, β=7.9999, which are the same values as in section 3, and S is assumed to vary in complex ways with chemical composition and temperature of the system, that is, the state variables. Here we consider only the case with no control action (u=0).

First, we deal with the sequential hybrid neural network which utilizes the above first principle model as the prior knowledge. In this method, the dependence of the ratio of the two rate constants, S, on the state variables is described by the ANN component, and then the ANN component is combined with the first principle model to compose the sequential hybrid neural network. We use the three layer feed forward neural network as the ANN component. The inputs to the ANN component are the state variables x_1, x_2 and x_3. Each neuron in the hidden layer has the sigmoidal activation function, while the linear activation function is used for the output layer. The biases of the neurons in the input layer are assumed to be zero. We train the sequential hybrid neural network by the modified error-back propagation algorithm proposed by Psaltis et al.: however, we improve the algorithm with momentum and an adaptive learning rate to increase the speed and the performance. Momentum helps the network avoid being trapped into local minimum, and the adaptive learning rate accelerates the training speed by keeping the learning step size as large as possible while keeping learning stable. The training data are obtained by integrating the system equations in section 3 and by sampling at every 0.001 dimensionless time. The training is carried out on DEC Alpha Server 2100 using MATLAB. We adapt only the number of hidden nodes and determine the optimal model which best describes the chaotic trajectory of the system according to the criteria of phase portrait, return map, the largest Lyapunov exponent and correlation dimen-

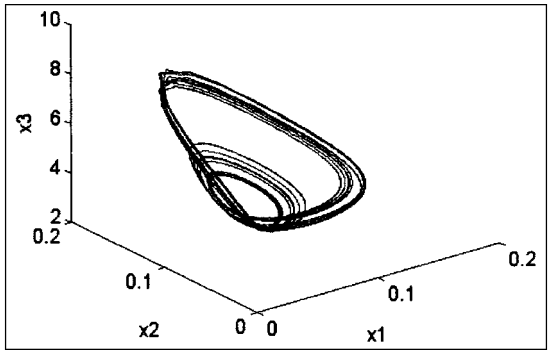

Fig. 6. 3-D phase portrait of the hybrid model with h=4 (trained with full data).

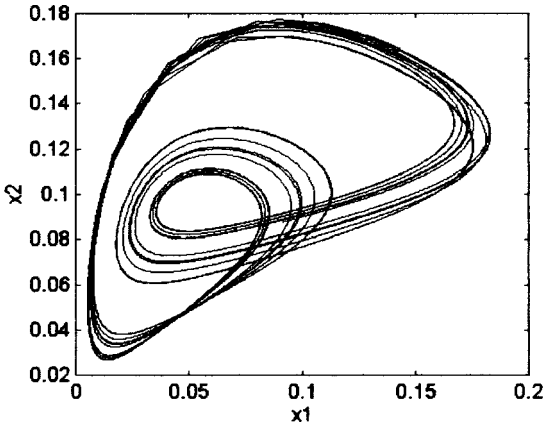

Fig. 7. x_1–x_2 plane projection of the 3-D phase portrait (The hybrid model with h=4, trained with full data).

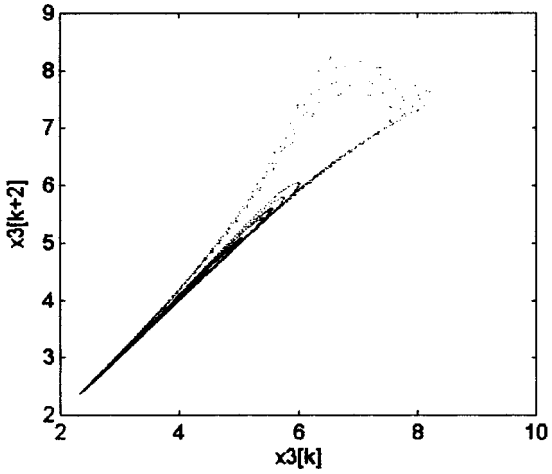

Fig. 8. Second return map of the hybrid model with h=4 (trained with full data).

sion.

Among an enormous number of candidates, we find two candidates by trial and error which seem to describe the chaotic behavior of the system closely. The numbers of hidden nodes (h) are 4 and 7, respectively. Fig. 6 denotes a 3-D phase portrait of the hybrid model with 4 hidden nodes, and Fig. 7 is the x_1–x_2 plane projection of the 3-D phase portrait. Fig. 8 shows the second return map reconstructed from the time series data of the state variable x_3. Figs. 9-11 denote the corresponding results when the number of hidden nodes is 7. The figures say that the hybrid model with 4 hidden nodes describes the chaotic dynamics of the system better than the hybrid model with 7 hidden nodes. In addition, we also calculate the largest Lyapunov exponent and correlation dimension to check the matching between the dynamic behaviors quantitatively. The calculations are carried out by using the same method as before. The obtained values are summarized in Table 1. From the results, we conclude that the hybrid model with 4 hidden nodes is the optimal model and, moreover, the model describes the chaotic dynamics of the system almost perfectly.

Next we compare the performance of the obtained optimal hybrid model with that of the optimal standard ANN model. We use the three layer feed forward neural network as shown in Fig. 12.

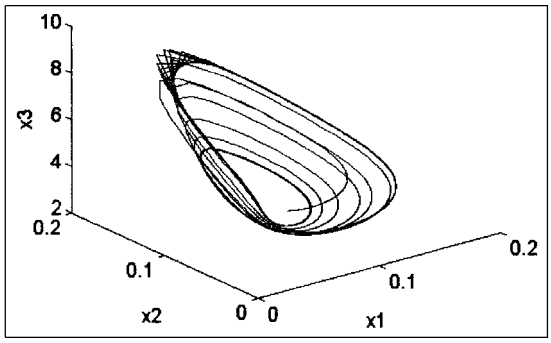

Fig. 9. 3-D phase portrait of the hybrid model with h=7 (trained with full data).

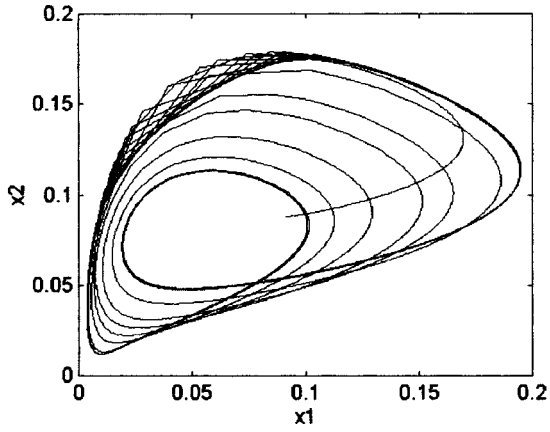

Fig. 10. x_1–x_2 plane projection of the 3-D phase portrait (The hybrid model with h=7, trained with full data).

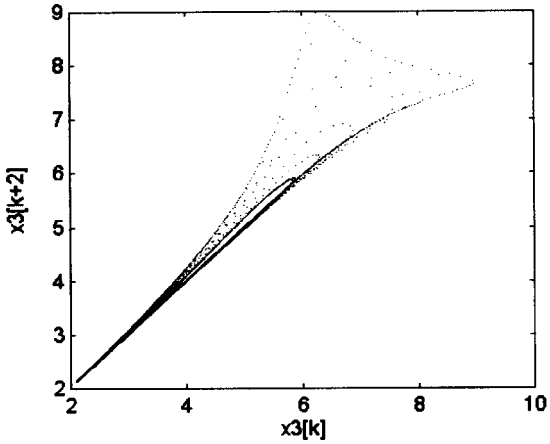

Fig. 11. Second return map of the hybrid model with h=7 (trained with full data).

The inputs to the standard ANN are the state variables x_1, x_2 and x_3. All activation functions, biases and the training data set used are same as before. Levenberg-Marquardt optimization algorithm was used to train the standard ANN, and the training was carried out on a DEC Alpha Server 2100 using MATLAB. We adapt only the number of hidden nodes and determine the optimal model which best describes the chaotic trajectory of the system according to the criteria of phase portrait, return map, the largest Lyapunov expo-

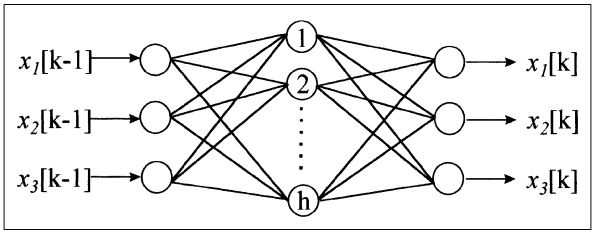

Fig. 12. Schematic representation of the standard ANN.

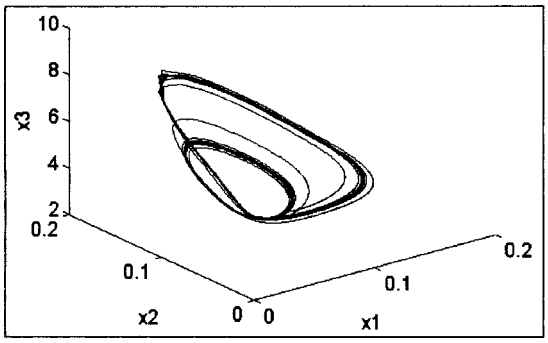

Fig. 13. 3-D phase portrait of the standard ANN with h=8 (trained with full data).

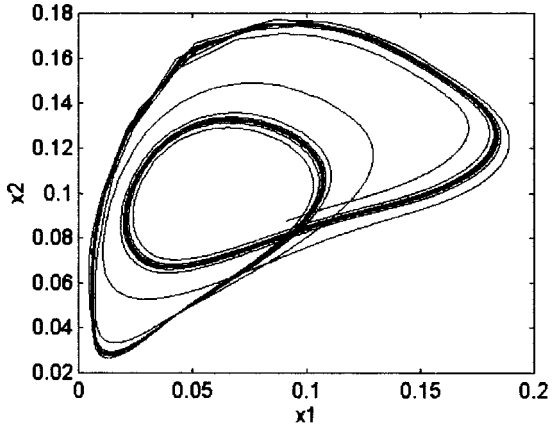

Fig. 14. x_1-x_2 plane projection of the 3-D phase portrait (The standard ANN with h=8, trained with full data).

nent and correlation dimension as before. From the extensive trial and error procedure, the optimal model is obtained when the number of hidden nodes (h) is 8 among several candidate models having roughly the same SSE. The 3-D phase portrait and its x_1-x_2 plane projection are shown in Figs. 13-14, respectively. The second return map is shown in Fig. 15. We also calculate the largest Lyapunov exponent and correlation dimension to check the matching between the dynamic behaviors quantitatively. The calculations are carried out by using the same method as before. The obtained values are summarized in Table 1. The results say that the sequential hybrid neural network shows much better interpolation capability than the standard ANN although the standard ANN also describes the chaotic dynamics of the system quite well.

When we compare the extrapolation capability of the models, the advantage of the sequential hybrid neural network becomes more obvious. We train the hybrid model with 4 hidden nodes and the

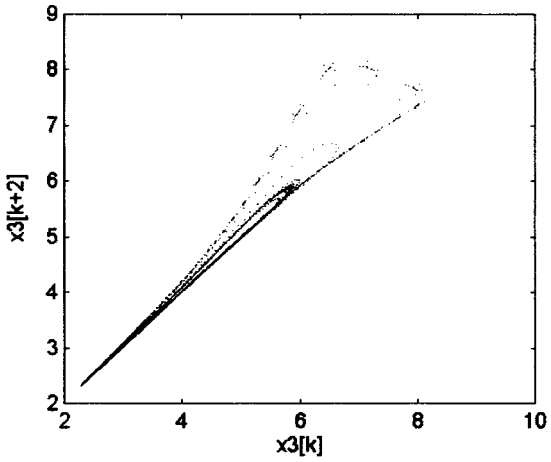

Fig. 15. Second return map of the standard ANN with h=8.

standard ANN with 8 hidden nodes using only the first one-fifth of the original training data set, and then simulate the identified models up to the same final time as the case of the original training data. The hybrid model shows quite good extrapolation capability as shown in Figs. 16-18 and Table 1; however, the standard ANN cannot extrapolate at all. Therefore, we can conclude that the sequential hybrid neural network shows better interpolation and also far better extrapolation capability than the standard ANN.

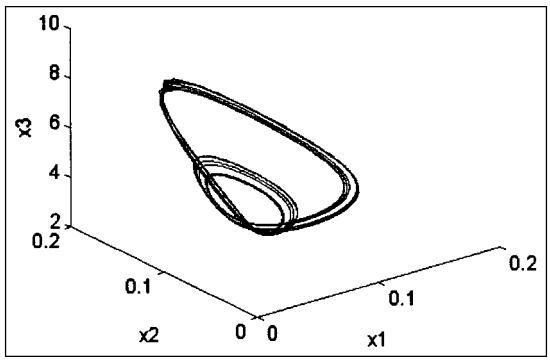

Fig. 16. 3-D phase portrait of the hybrid model with h=4 (trained with partial data).

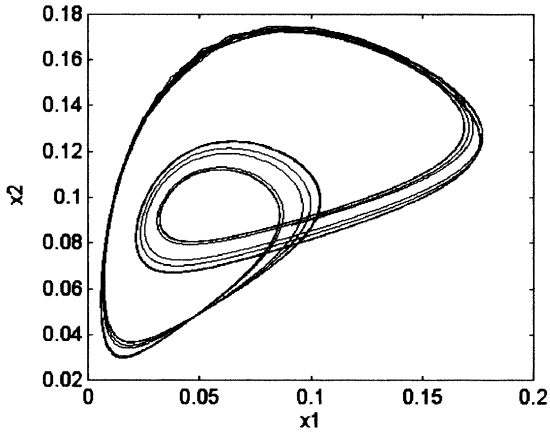

Fig. 17. x_1-x_2 plane projection of the 3-D phase portrait (The hybrid model with h=4, trained with partial data).

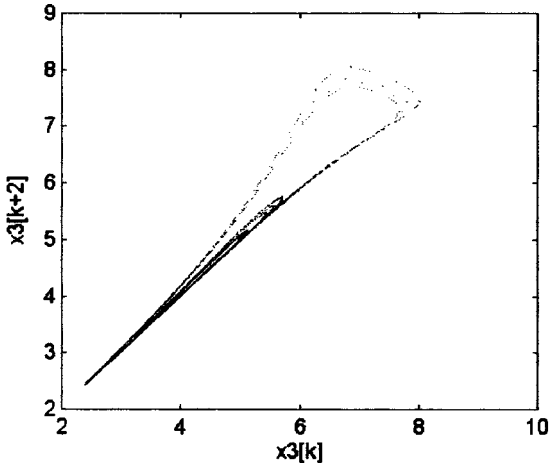

Fig. 18. Second return map of the hybrid model with h=4 (trained with partial data).

CONCLUSION

If prior knowledge about a system is available, it is smart to incorporate the prior knowledge into the black box model of the system. In this paper, we identify a chaotic chemical reaction system with both the sequential hybrid neural network and the standard ANN, and compare interpolation and extrapolation capability of the models. The identified models are validated with phase portrait, return map, the largest Lyapunov exponent and correlation dimension instead of Sum of Square Errors (SSE). This is demonstrated for a nonisothermal, irreversible, first-order, series reaction A→B→C in a CSTR. The results say that the sequential hybrid neural network shows good interpolation and extrapolation capability. When compared with the standard ANN, it shows better interpolation and also far better extrapolation capability.

ACKNOWLEDGMENTS

The authors would like to thank Korea Science and Engineering Foundation (KOSEF) for financial support through Automation Research Center at Pohang University of Science and Technology (POSTECH).

NOMENCLATURE

In Section 2

J : objective function
y_k : k th plant output
d_k : k th target value
w_{ij}^m : weight between the i th neuron of the m−1 th layer and the j th neuron of the m th layer
$\varepsilon_k = d_k - y_k$: k th plant output error
λ_j : j th unmeasured process parameter (j th output of the ANN component)
$\frac{\partial y_k}{\partial \lambda_j}$: differential gain of plant
p : input to neuron
q : output from neuron
p_j^m : weighted sums of the outputs of the m−1 th layer
q_i^m : output of the i th neuron of the m th layer
f_j^m : transfer function of the j th neuron of the m th layer
α : learning rate

In Section 3

x_1 : dimensionless concentrations of species A
x_2 : dimensionless concentrations of species B
x_3 : dimensionless temperature in the reactor
Da : Damköhler number
ε : dimensionless activation energy
S : ratio of the two rate constants
κ : ratio of activation energies
B : dimensionless adiabatic temperature rise
α : ratio of heat effects
β : dimensionless heat transfer coefficient
u : dimensionless coolant bath temperature

REFERENCES

Abarbanel, H. D. I., Brown, R., Sidorowich, J. J. and Tsimring, L. S., "The Analysis of Observed Chaotic Data in Physical Systems," *Rev. Mod. Phys.*, **65**(4), 1331 (1993).

Adomaitis, R. A., Farber, R. M., Hudson, J. L., Kevrekidis, I. G., Kube, M. and Lapedes, A. S., "Application of Neural Nets to System Identification and Bifurcation Analysis of Real World Experimental Data," In: Neural Networks: Biological Computers or Electronic Brains, Springer Verlag, Paris, 87 (1990).

Aguirre, L. A. and Billings. S. A., "Validating Identified Nonlinear Models with Chaotic Dynamics," *Int. J. Bifurcation and Chaos*, **4**(1), 109 (1994).

Casdagli, M., "Nonlinear Prediction of Chaotic Time Series," *Physica D*, **35**, 335 (1989).

Doedel, E. J., "AUTO: Software for Continuation and Bifurcation Problems in Ordinary Differential Equations," AUTO 86 User Manual, CALTECH (1986).

Farmer, J. D., Ott, E. and Yorke, J. A., "The Dimensions of Chaotic Attractors," *Physica 7D*, 153 (1983).

Giona, M., Lentini, F. and Cimagalli, V., "Functional Reconstruction and Local Prediction of Chaotic Time Series," *Phys. Rev. A*, **44**, 3496 (1991).

Grassberger, P. and Procaccia, I., "Measuring the Strangeness of Strange Attractors," *Physica D*, **9**, 189 (1983).

Grassberger, P., Schreiber, T. and Schaffrath, C., "Nonlinear Time Sequence Analysis," *Int. J. Bifurcation and Chaos*, **1**(3), 521 (1991).

Jordan, M. I. and Rumelhart, D. E., "Forward Models: Supervised Learning with a Distal Teacher," *Cognitive Science*, **16**, 307 (1992).

Kahlert, C., Rossler, O. E. and Varma, A., "Chaos in a Continuous Stirred Tank Reactor with Two Consecutive First-Order Reactions, One Exo-, One Endothermic," *Springer Ser. Chem. Phys.*, **18**, 355 (1981).

Kim, H. J., "Identification and Validation of Neural Network Models for Chaotic Chemical Reaction Systems," M. S. Thesis, Pohang University of Science and Technology, Korea (1998).

Kim, H. J. and Chang, K. S., "Applications of Neural Network Method to Chaotic Chemical Reaction Systems," *Comp. Chem. Eng.*, 1998 (submitted).

Kim, H. J., Lee, J. S., Han, C. and Chang, K. S., "Chaos and Fractals

in Process Systems Engineering," Proceedings of the Forty-first Annual Meeting of the ISSS, Seoul (1997).
Packard, N. H., Crutchfield, J. P., Farmer, J. D. and Shaw, R. S., "Geometry from a Time Series," *Phys. Rev. Lett.*, **45**(9), 712 (1980).
Parker, T. S. and Chua, L. O., "Practical Numerical Algorithms for Chaotic Systems," Springer-Verlag (1989).
Principe, J. C., Rathie, A. and Kuo, J. M., "Prediction of Chaotic Time Series With Neural Networks and the issue of Dynamic Modeling," *Int. J. Bifurcation and Chaos*, **2**(4), 989 (1992).
Psaltis, D., Sideris, A. and Yamamura, A. A., "A Multilayered Neural Network Controller," *IEEE Cont. Syst. Mag.*, **4**(17) (1988).
Psichogios, D. C. and Ungar, L. H., "A Hybrid Neural Network-First Principles Approach to Process Modeling," *AIChE J.*, **38**(10), 1499 (1992).
Rand, D. A. and Young, L. S., "Dynamical Systems and Turbulence," Springer Verlag (Berlin), 366 (1981).
Roux, J. C., Simoyi, R. H. and Swinney, H. L., "Observation of a Strange Attractor," *Physica 8D*, 257 (1983).
Rumelhart, D., Hinton, G. and Williams, R., "Learning Internal Representations by Error Propagation," Parallel Distributed Processing: Explorations in the Microstructures of Cognition: 1. Foundations, MIT Press (1986).
Sauer, T., Yorke, J. A. and Casdagli, M., "Embedology," *J. Stat. Phys.*, **65**(3/4), 579 (1991).
Seydel, R., "From Equilibrium to Chaos: Practical Bifurcation and Stability Analysis," Elsevier, New York (1988).
Takens, F., "Detecting Strange Attractors in Turbulence," in Lecture Notes in Mathematics, vol. 898 eds.
Thompson, M. L. and Kramer, M. A., "Modeling Chemical Processes Using Prior Knowledge and Neural Networks," *AIChE J.*, **40**(8), 1328 (1994).
Wolf, A., Swift, J. B., Swinney, H. L. and Vastano, J. A., "Determining Lyapunov Exponents from a Time Series," *Physica 16D*, 285 (1985).