ADVANCED MPC FOR LARGE SCALE DYNAMIC SYSTEMS BASED ON MODEL REDUCTION TECHNIQUE AND FEEDFORWARD ARTIFICIAL NEURAL NETWORK

Weiguo Xie*
University of Minnesota - Duluth
Duluth, MN 55812

Abstract

Model Predictive Control (MPC) has been widely used in process industries. However, in the real world most systems exhibit nonlinear dynamics, rendering the application of linear controllers. In order to apply MPC for nonlinear distributed-parameter systems with unknown dynamics, as a “black-box” system, a data-driven model reduction-based feedforward artificial neural network (ANN) approach has been developed for MPC control. An off-line model reduction technique, the proper orthogonal decomposition (POD) method, is first applied to extract accurate non-linear low-order models from the non-linear dynamic large-scale distributed system. Then a series of successive feedforward ANNs are trained based on the time coefficients of POD basis functions to obtain the model for the system. A benchmark case study to use cooling zones to stabilize a tubular reactor with recycle will be used as an illustrative example to show this methodology.

Keywords

Model predictive control, Feedforward artificial neural network, Proper orthogonal decomposition.

Introduction

Model predictive control (MPC) has been applied efficiently for a large number of industrial processes. In general, nonlinear MPC is mostly used in batch operations, while linear MPC is more often applied in continuous operations (Heath et al., 2006). Meanwhile, a non-linear model requires extensive computation to perform nonlinear MPC control strategy due to its iterative feature, especially for the cases with a large number of inputs.

Model reduction techniques can be applied in MPC to significantly reduce the complexity of sophisticated nonlinear dynamic systems, which leads to the successful design and implementation of the MPC control strategy (Bonis et al., 2012; Bonis et al., 2014; Li and Christofides, 2008; Xie and Theodoropoulos, 2010; Xie et al., 2011; Xie et al., 2012; Xie et al., 2015). The Proper Orthogonal Decomposition (POD) method has been effectively applied in non-linear MPC frameworks (Garcia et al., 2008; Li and Christofides, 2008; Xie and Theodoropoulos, 2010; Xie et al., 2015) and for MPC application with mesoscopic simulators (Oguz and Gallivan, 2008). An off-line POD model reduction technique associated with a Trajectory Piecewise-Linear (TPWL) method has been developed using linear MPC for nonlinear large-scale systems (Xie and Theodoropoulos, 2010; Xie et al., 2011; Xie et al., 2012). This POD-FEM-based reduced model is 1-dimensional nonlinear only in the time. However, this method requires the knowing detailed governing equations and the application of Galerkin projection.

---

* To whom all correspondence should be addressed

Since 1940s (Hebb, 1949; McCulloch and Pitts, 1943), especially with the development of modern computers, artificial neural networks (ANNs) have been applied in many areas including system identification and control, process modeling, data processing, visualization, and etc. It has been stated that multilayer feed-forward networks with only one hidden layer can be universal approximators for any nonlinear functions (Hornik et al., 1989; Hornik, 1993). ANN-based dynamic modeling and control has been applied for chemical process systems (Bhat and McAvoy, 1990). More recently, ANN associated with POD model reduction has been applied for nonlinear dynamic reaction systems (Shvartsman et al., 2000; Xie et al., 2012; Xie et al., 2015). A radial basis function (RBF) neural network has been used to model distributed parameter systems (DPSs) and then implemented in MPC configurations (Aggelogiannaki and Sarimveis, 2008). However, it took a long time to implement the online RBF neural network for predicting the temporal evolution of the system. In order to control systems with unknown dynamics, a POD model reduction-based successive Elman neural network approach has been applied to simulate the nonlinear large-scale distributed system, and then nonlinear MPC control strategy has been applied on the POD/ANN reduced model (Xie et al., 2012; Xie et al., 2015). A combined methodology of POD and rANN (recurrent ANNs) has been employed to produce a reduced order system, without using the equations, which is then employed within a new nonlinear Multiparametric MPC algorithm (Petsagkourakis et al., 2018).

This paper describes the development of a POD-based artificial neural network technique on MPC application for any "black-box" system with unknown dynamics. The rest of paper is organized as follows: There is a brief introduction of POD/ANN model reduction and the feedforward neural network. Then, a case study of a tubular reactor is used to illustrate the features of POD/ANN-MPC control strategy. Finally, the conclusions of this work are discussed.

Model Reduction Methodology

POD/ANN Model Reduction

Proper orthogonal decomposition (POD) applies the spectral theory of compact and self-adjoint operators from the Karhunen-Loeve decomposition theorem (Wong, 1971). POD can capture the most "energy" in an average sense for efficient linear approximation in terms of data compression (Holmes et al., 1996). The "energy" of a given mode corresponds to the magnitude of the eigenvalue for the mode. The method of snapshots (Sirovich, 1987) is often used to obtain the reduced set of POD global basis functions.

In Figure 1, it shows main steps for POD/ANN model reduction method:

(i) An empirical collection of time evolving data points from the response of dynamic "black box" system for the chosen appropriate range of parameters;

(ii) Construction of a two-point correlation matrix from these dynamic responses;

(iii) Calculation of the empirical global basis functions with low-order set of m<<N (N being the dimension of the full model) global basis functions, in which m is determined by capturing most of the system's "energy" through eigenvalue analysis of this two-point correlation matrix;

(iv) Expression of the state variables $x(z,t)$ of the system (where z are spatial coordinates) as linear combinations of the eigenfunctions $\varpi(z)$ which are functions only of space and of some coefficients $a(t)$, which are functions only of time:

$$x(z,t) = \sum_{j=1}^{m} \alpha_j(t)\varpi_j(z) + \bar{x}(z) \qquad (1)$$

$\bar{x}(z)$ being the average snapshot.

(v) Training for neural network on time coefficients of POD method to obtain POD/ANN reduced model, which is 1-dimensional nonlinear only in the time.

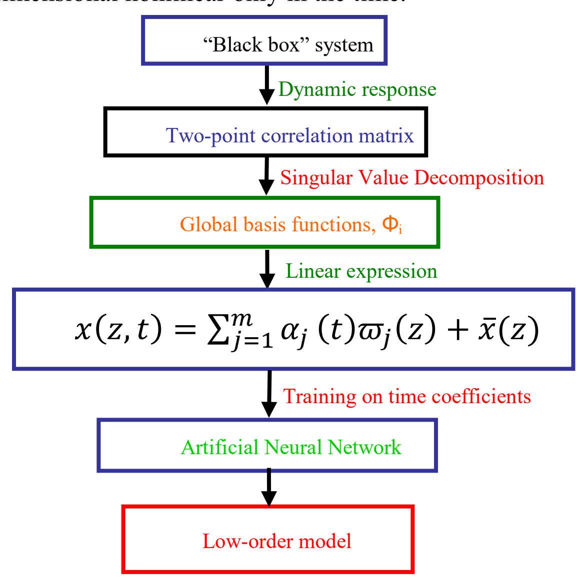


Figure 1. Schematic diagram for POD/ANN model reduction approach

Feedforward Artificial Neural Network

Because appropriate artificial neural network can be good approximators for any nonlinear functions (Hornik et al., 1989; Hornik, 1993), ANN can be used for POD time coefficients modelling. An Elman neural network (Elman, 1990) with two recurrent tansig layers and one output purelin layer has been applied for POD time coefficients modelling (Xie et al., 2012; Xie et al., 2015). A feedforward neural network is an artificial neural network wherein connections between the nodes do not form a cycle. The feedforward neural network from the Neural Network Tool-box in MATLAB is applied in this research. A feedforward neural network, as shown in Figure 2, has good features to recognize and generate temporal patterns, which are suitable for simulation of dynamic systems.

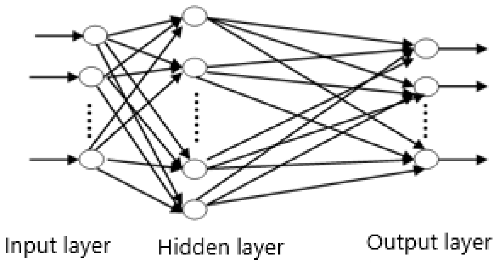


Figure 2. Neural network topology of a standard feed-forward neural network with no recursive connections.

The feedforward neural network has been chosen for POD/ANN model reduction method for the simulation of non-linear dynamic large-scale distributed system in this research.

Case Study

Tubular Reactor Case

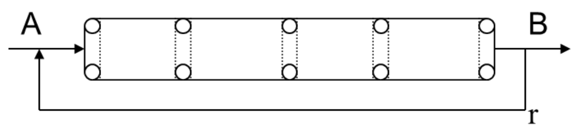


Figure 3. Tubular reactor with recycling

The tubular reactor with recycling in Figure 3 can be modeled by two set of partial differential equations (Jensen and Ray, 1982) in a spatial domain $z \in [0,1]$:

$$C_t = -\frac{\partial C}{\partial z} + \frac{1}{Pe_C}\frac{\partial^2 C}{\partial z^2} - f(C,T) \quad (2)$$

$$T_t = -\frac{\partial T}{\partial z} + \frac{1}{Pe_T}\frac{\partial^2 T}{\partial z^2} + B_T f(C,T) + \beta_T(T_c - T) \quad (3)$$

where, C and T are dimensionless concentration and dimensionless temperature, respectively. T_C corresponds to the dimensionless temperature of the cooling medium and $f(C,T) = B_c C exp(\frac{\gamma T}{1+T})$ is the reaction term. The values of parameters used are: $Pe_c = 7.0$, $Pe_T = 7.0$, $B_c = 0.1$, $B_T = 2.5$, γ=10.0 and β_T=2.0, with Pe_c, Pe_T being the Peclet numbers for mass and heat transport respectively, B_C being the (dimensionless) heat transfer coefficient, B_T being the (dimensionless) temperature rise and γ being the activation energy. For a given recycling ratio r, the boundary conditions for concentration and temperature at $z = 0$ become (Antoniades and Christofides, 2001):

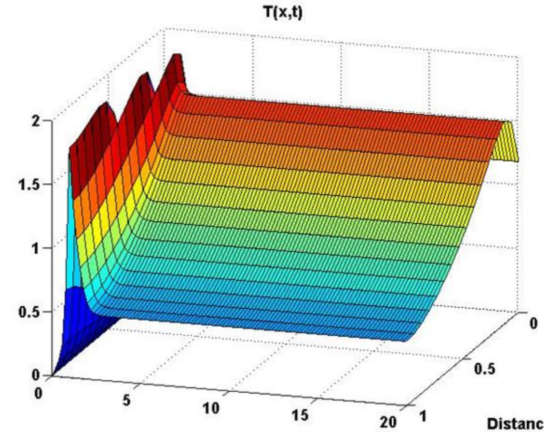
$$\frac{\partial C}{\partial z} = -Pe_C[(1-r)(1+C_0) + rC(t,1) - C(t,0)] \quad (4)$$

$$\frac{\partial T}{\partial z} = -Pe_T[(1-r)(1+T_0) + rT(t,1) - T(t,0)] \quad (5)$$

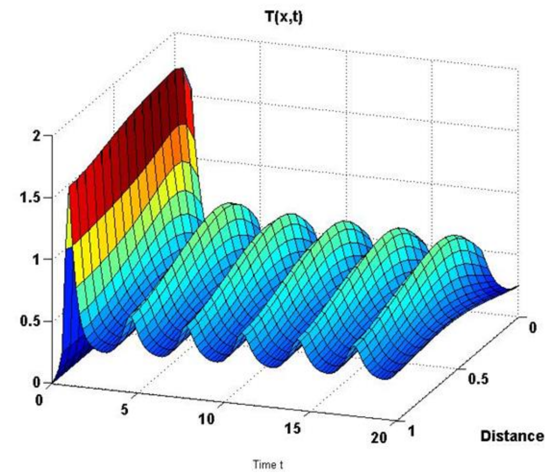
The boundary conditions at $z = 1$ are $dC/dz = 0$ and $dT/dz = 0$. The reactor exhibits oscillations at C_0=T_0=T_c=0 for r=0.5 (Alonso et al., 2004). The model was discretized into 16 nodes for the spatial domain, and the full-scale FEM method has been applied to solve this model. The data from FEM method are used as the full model for further process in training or testing neural network.

Control Objective

It can be seen that the outlet temperature of a tubular reactor shows stable behavior for r=0 (Figure 4a) while it undergoes sustained oscillations for r=0.5 (Figure 4b).



(a)



(b)

Figure 4. Temperature profiles of a tubular reactor under the condition (a) $r = 0$, (b) $r = 0.5$

The control objective was to stabilize the reactor with r=0.5 to behave like the system with r=0 by introducing a number of jacket temperature zones (actuators). The objective function is as follows:

$$J = \min_{du}\left(T(t) - T_{ref}\right)^T Q\left(T(t) - T_{ref}\right) + DU^T RDU \quad (6)$$

where, $T_{ref}(t)$ is the reference state (r=0) and DU is the control on the actuators.

Then, the following objective function can be obtained by applying Eq. (1) from POD method to replace temperature term.

$$J = \min_{du}\left(\left(\sum_{k=1}^{m}\alpha_{k\_T}(t)\varpi_{k\_T}(x) + \overline{T_{16}}\right) - T_{ref}\right)^T Q\left(\left(\sum_{k=1}^{m}\alpha_{k\_T}(t)\varpi_{k\_T}(x) + \overline{T_{16}}\right) - T_{ref}\right) + DU^T RDU \quad (7)$$

which is a quadratic function due to the linear POD representation of the state variables as seen in Eq. (1). Q and R are non-negative definite matrices. Applying POD on the nonlinear Eqs. (2) - (5) resulted in a reduced set of nonlinear characteristic equations of the system which are functions of the time coefficients *a(t)*.

Data Sampling

A method based on orthogonal experimental design methodology has been applied on data sampling. There were 8 jacket temperature zones (actuators) for the implementation of the computed control actions. In Taguchi's orthogonal experimental design, the use of the L_{12} orthogonal array has been highly recommended and many successful cases have been reported (Tsai, 1995). The first 8 columns of L_{12} (2^{11}) orthogonal array, as listed in Table 1, has been used in this work.

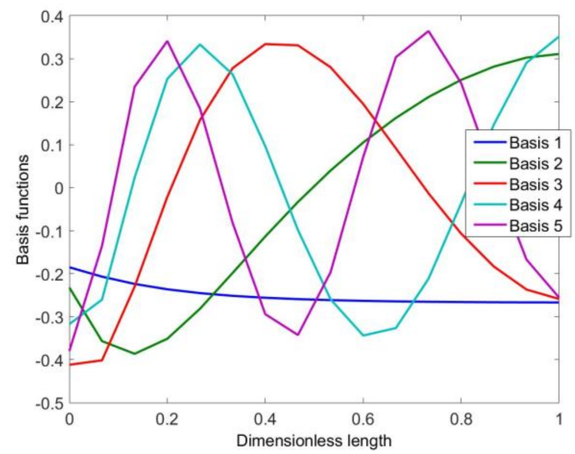
Table 1: $L_{12}(2^{11})$ orthogonal array from Taguchi's orthogonal experimental design [25], with symbol 1 being replaced by 0 and -1 being replaced by 1.

Run	A	B	C	D	E	F	G	H	I	J	K
1	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	1	1	1	1	1	1
3	0	0	1	1	1	0	0	0	1	1	1
4	0	1	0	1	1	0	1	1	0	0	1
5	0	1	1	0	1	1	0	1	0	1	0
6	0	1	1	1	0	1	1	0	1	0	0
7	1	0	1	1	0	0	1	1	0	1	0
8	1	0	1	0	1	1	1	0	0	0	1
9	1	0	0	1	1	1	0	1	1	0	0
10	1	1	1	0	0	0	0	1	1	0	1
11	1	1	0	1	0	1	0	0	0	1	1
12	1	1	0	0	1	0	1	0	1	1	0

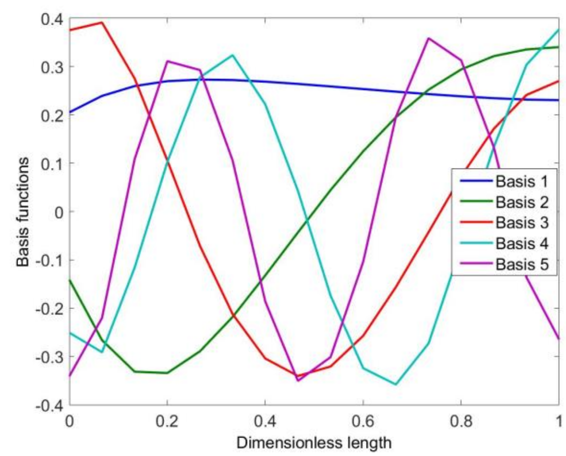
Taking 11 samples over the range of cooling temperature [-1, 1], we have $12 * 11 = 132$ runs. The full-scale FEM model was used for sampling and the sampling time was 15s.

Reduced Model vs. Full Model

It is shown in Figure 5a and 5b that five global basis functions for concentration and temperature were computed based on the 132 samples collected. *l*=5, eigenfunctions for concentration (or temperature) capture 99.7% (or 98.3%) of the system's energy.



(a)



(b)

Figure 5. Global basis functions for (a) concentration (b) temperature from the sampling data of tubular reactor with $r = 0.5$.

The feedforwardnet function in MATLAB with a two-layer feedforward network has been applied in this research. The network has one hidden layer with 20 neurons and an output layer. All data sampling collected with Taguchi's orthogonal experimental design method has been used to train the feedforward neural network. In general, neural network training is very important for the quality of the approximation to the original model. Successive feedforward neural networks with short time slots have been used to simulate the original model to avoid bad approximation using neural networks for long time period. In addition, a least square optimization has been applied to the sampling cases of feedforward neural network to catch the dynamics on time coefficients of POD/ANN reduced model. After simulation of the sampling cases, this optimization step will not be used for further MPC control.

A training result of the first feedforward ANN from MATLAB is shown in Figure 6. There are 8 inputs from the 8 actuators (cooling zones). The training algorithm is Levenberg-Marquardt method. It took 32 iterations to reach minimum gradient and the performance (Mean-Squared Error) is 4.34e-5, which is also acceptable. Then, the successive feedforward ANNs can be obtained by using the similar training method to simulate the tubular reactor system.

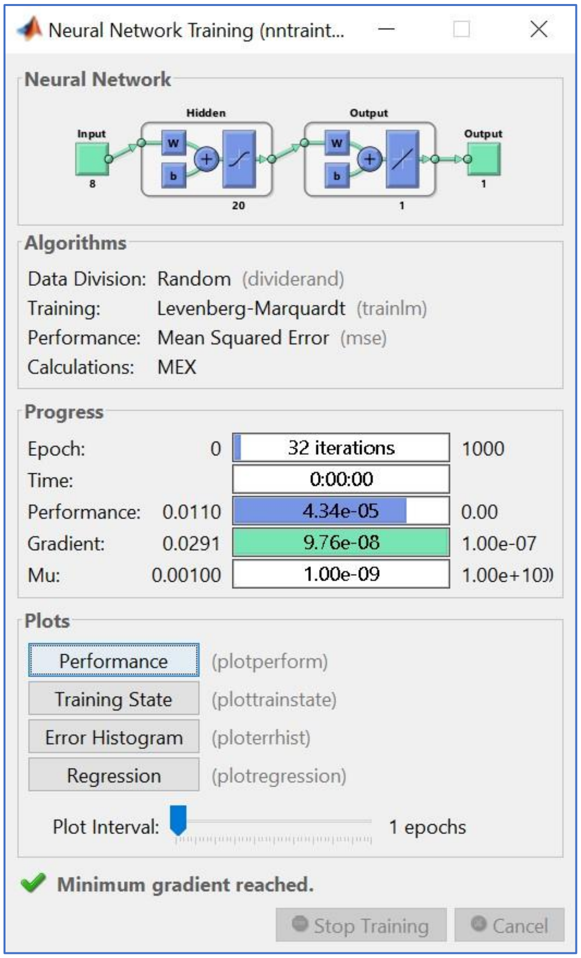


Figure 6. A training result of feedforwardnet in MATLAB.

The comparison of temperature between full and reduced model for the dynamics of the middle and output points for $r = 0.5$ is shown in Figure 7. It can be seen that the reduced model has a good prediction on the complex reactor dynamics.

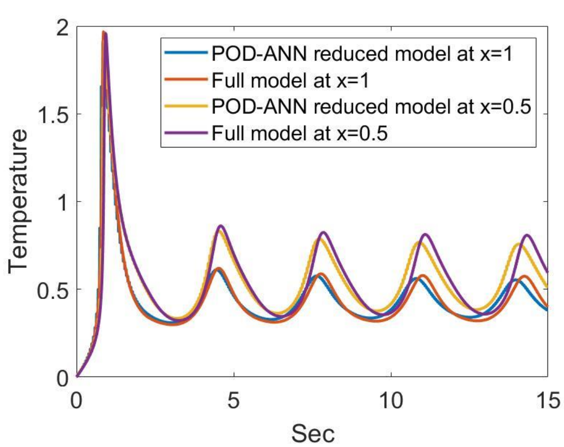
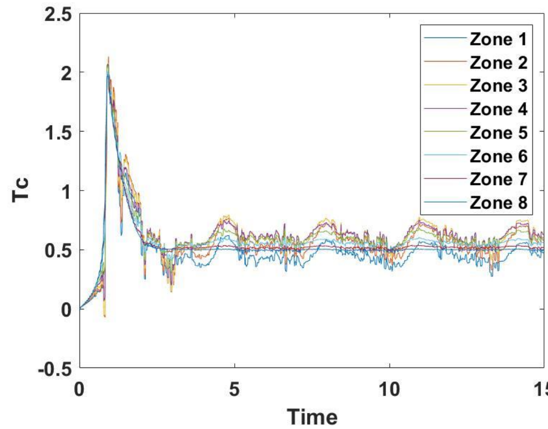


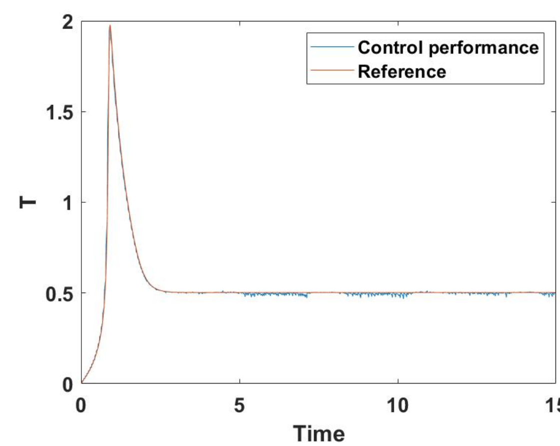
Figure 7. Comparison between temperature predictions of full model and POD-ANN reduced model at the middle and output points for tubular reactor with $r = 0.5$.

POD/ANN-MPC Results

In Figure 8, results of the POD/ANN-MPC for 8 actuators are shown. Figure 8a shows the control law for the 8 actuators (zones). The control output and the reference profile are shown in Figure 8b. It can be seen that the tubular reactor outlet temperature is efficiently stabilized by the 8 actuators.



(a)



(b)

Figure 8. POD/ANN-MPC results (a) time evolution of control profile for the 8 actuators (b) control and reference profile of the tubular reactor outlet temperature.

Conclusion

A POD model reduction-based artificial neural network (ANN) approach has been developed for the nonlinear MPC application for nonlinear large-scale distributed systems. This efficient model reduction-based technique combines the proper orthogonal decomposition (POD) with the feedforward artificial neural network (ANN). The POD-ANN methodology enables the use of nonlinear MPC for large scale non-linear "black-box" systems. This method can effectively facilitate the use of nonlinear MPC for large scale distributed systems, as it was demonstrated in the case study, where stabilization of a tubular reactor undergoing sustained oscillations was performed. For the future work, it has been proposed to implement a piecewise linear MPC associated with POD-ANN for control of large scale non-linear "black-box" systems, by which high computational efficiency can be achieved.

Acknowledgments

The author would like to acknowledge Professor Constantinos Theodoropoulos and Dr. Ioannis Bonis (former PhD student and Postdoctoral Research Associate) at the University of Manchester, UK for the previous background work and the case study in the optimization and control for large scale distributed systems during the time period of the EC FP6 Project: CONNECT [COOP-2006-31638] and the EC FP7 project CAFÉ [KBBE-212754].

References

Aggelogiannaki, E., Sarimveis, H. (2008). Nonlinear model predictive control for distributed parameter systems using data driven artificial neural network models. *Computers & Chemical Engineering,* 32(6), 1225-1237.

Alonso, A. A., Frouzakis, C. E., Kevrekidis, I. G. (2004). Optimal sensor placement for state re-construction of distributed process systems. *AIChE Journal*, 50, 1438-1452.

Antoniades C., Christofides, P. D. (2001). Integrated optimal actuator/sensor placement and robust control of uncertain transport-reaction processes. *Chem Eng Sci.*, 56, 4517–4535.

Bhat, N., McAvoy, T.J. (1990). Use of neural nets for dynamic modeling and control of chemical process systems, *comp. chem. engg.*, 14,(4/5) 573-582.

Bonis, I., Xie, W., Theodoropoulos, C. (2012). A linear model predictive control algorithm for nonlinear large-scale distributed parameter systems. *AICHE Journal*, 58(3), 801-811.

Bonis, I., Xie, W., Theodoropoulos, C. (2014). Multiple model predictive control of dissipative PDE systems. *IEEE Transactions on Control Systems Technology*, 22(3), 1206-1214.

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, 14, 179-211.

Garcia, M. R., Vilas, C., Banga, J. R., Alonso, A. A. (2008). Exponential Observers for Distributed Tubular (Bio) Reactors. *AIChE Journal,* 54, 2943–2956.

Heath, W. P., Li, G., Wills, A. G., Lennox, B. (2006). The robustness of input constrained model predictive control to infinity-norm bound model uncertainty," *ROCOND06, 5th IFAC Symposium on Robust Control Design*, Toulouse, France, 5th-7th July, 2006.

Hebb, D. O. (1949). The organization of behavior. New York: Wiley & Sons.

Holmes, P., Lumley, J. L., Berkooz, G. (1996). Turbulence, coherent structures, dynamical systems and symmetry. Cambridge University Press.

Hornik, K., Stinchcombe, M., White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, 2, 359–366.

Hornik, K. (1993). Some new results on neural network approximation. *Neural Networks*, 6, 1069–1072.

Jensen K. F., Ray, W. H. (1982). The bifurcation behaviour of tubular reactors. *Chemical Engineering Science,* 37, 199–222.

Li, M. H., Christofides, P. D. (2008). Optimal control of diffusion-convection-reaction processes using reduced-order models. *Computers & Chemical Engineering*, 32, 2123-2135.

McCulloch, W., Pitts, W. (1943). A logical calculus of ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5, 115-133.

Oguz, C., Gallivan, M. A. (2008). Optimization of a thin film deposition process using a dynamic model extracted from molecular simulations. *Automatica,* 44, 1958–1969.

Petsagkourakis, P., Theodoropoulos. C. (2018). Data driven reduced order nonlinear multiparametric mpc for large scale systems. *Computer Aided Chemical Engineering*, 43, 1249-1254.

Shvartsman, S. Y., Theodoropoulos, C., Rico-Martinez, R., Kevrekidis, I. G., Titi, E. S., Mountziaris, T. J. (2000). Order reduction for nonlinear dynamic models of distributed reacting systems. *Journal of Process Control*, 10, 177-184.

Sirovich, L. (1987). Turbulence and the dynamics of coherent structures, parts i-iii. *Quart. Appl. Math*., 45, 561-590.

Tsai, H. T. (1995). The partial aliasing patterns for Taguchi's L12 orthogonal array. *International Journal of Quality & Reliability Management,* 12 (5), 54-60.

Wong, E. (1971). Stochastic Process in Information and Dynamical Systems. McGraw-Hill.

Xie, W., Theodoropoulos, C. (2010). An off-line model reduction-based technique for on-line linear MPC applications for nonlinear large- scale distributed systems. *Computer Aided Chemical Engineering*, 28, 409-414.

Xie, W. Bonis, I., Theodoropoulos, C. (2011). Off-line model reduction for on-line linear MPC of nonlinear large-scale distributed systems. *Computers & Chemical Engineering*, 35, 750-757.

Xie, W., Bonis, I., Theodoropoulos, C. (2012). Linear MPC based on data-driven Artificial Neural Networks for large-scale nonlinear distributed parameter systems. *Computer Aided Chemical Engineering,* 30, 1212-1216.

Xie, W., Bonis, I., Theodoropoulos, C. (2015). Data-driven model reduction-based nonlinear MPC for large-scale distributed parameter systems. *Journal of Process Control*, 35, 50-58.