Using Neural Networks for Identification and Control of Systems

Jhonatam Cordeiro
Department of Industrial and Systems Engineering
North Carolina A&T State University, Greensboro, NC 27411
jcrodrig@aggies.ncat.edu

Abstract

The present work addresses the utilization of Artificial Neural Networks (NN) for the identification and control of systems, in special to control nonlinear dynamic systems or systems with some degree of uncertainty. Because NNs have an inherent ability to approximate functions and to adapt to changes in input and parameters, they can be used to control systems too complex for linear controllers, such as PID controllers. In the present work a mathematical basis for NN is presented, the mathematical representation of a process unit, or neuron, and how they can be put together in order to form nets that can learn from external data. In sequence, it is presented structures of inputs that can be used along with NN to model nonlinear systems. The most common configurations of input vectors for the training of NN are highlighted. Following, a method of control is presented that take advantage of NN, where a NN is used to build a predictive nonlinear controller using a model predictive control (MPC) structure. Two nonlinear systems were used to test the identification and control of the structures proposed. The results shows the NN used were efficient in modeling and controlling the nonlinear plants.

Introduction

The modern feedback controlling systems are responsible for the success of several operational systems and are applied in the military, aerospace, manufacturing industry and other fields (Franklin, Powell, & Emami-Naeini, 2006). The function of the feedback controller is to induce an input in a system so that it would respond with a desired output. There are many methods to design a controller; the most popular and widely used in the industry are based on state space or frequency analysis. Such design techniques have yielded successful applications such as the control of pitch, yaw and roll in aircrafts, satellite positioning and air conditioning control (Franklin et al., 2006). Proportional, Integral and Derivative (PID) controller are still the most popular type of control used in the manufacturing industry, mostly given its simplicity and fact that most applications where it is used are highly linear (Visioli, 2006). However, the increasing complexity of some systems challenges the classic feedback control theory. Challenges such as nonlinearities, rapid change conditions, black-box systems or high level of uncertainty can make classic controllers, such as the PID, have poor performance.

The operation of a complex system requires the controller to be smart in a way, to adapt and learn from changes in the system dynamics, noise or external output. A solution for the control of complex system is to use control structures inspired in biological systems. Biological systems are adaptive and resilient to the environmental changes where they are inserted. Bacteria constantly change their DNA sequencing so that they remain unknown to the defense systems of other creatures. Most animals have a neural system that allows than to sense the environment and to rationalize a best course of action, such as when to run from or fight a predator, or in the case of humans, how to solve a mathematical equation. Fuzzy logic, Evolutive Algorithms and Artificial Neural Networks (NN) are among the theories developed with an inspiration in biological systems. Neural networks tries to mimic the biological neural system, it presents an inherent capacity for learning, adapt and parallel computing (S. Haykin & Network, 2004). With that NNs have being gaining exposure for its successful utilization for modeling complex non-linear systems.

Most NN applications are designed in an open loop, such as designs for pattern recognitions (Ebrahimzadeh & Ranaee, 2010), classification (Krizhevsky, Sutskever, & Hinton, 2012) and function approximation (Zainuddin & Pauline, 2011). However, the use of NN in a feedback control loop has proven to be efficient when controlling nonlinear systems. Chen, M. (Chen, Ge, & Voon Ee How, 2010) proposed a NN structure to control nonlinear systems with multiple inputs and multiple outputs [MIMO]. Dierks, T and Sarangapani, J.(Dierks & Jagannathan, 2010) used NN in a feedback loop to control a Quadrotor

UAV and Addeh, J. et al.(Addeh, Ebrahimzadeh, Azarbad, & Ranaee, 2014) used NN for statistical process control.

The present work investigates the application of NN for identification and control of systems. For the identification process, the NN is placed in parallel with the model and random step signals are generated for input. The plant's response to the signals can be used for training the NNs. The trained NN are then used in a predictive control structure. Matlab was used to implement the NNs, plants and input signals. The goal of this work is to investigate if a system can automatically experiment with a plant, learn from the experiment and control the plant, automatically and without needing a mathematical model of the plant or a fine tune the of the controller(Nørgård, Ravn, Poulsen, & Hansen, 2000).

Methodology

The processing unit of a NN is a neuron. The mathematical model of an artificial neuron tries to mimic the behavior of a biological neuron. An artificial NN is based on the approximation models of how a biological neuron processes the electric impulses it receives from other neurons or external stimuli. The model used in the present work is the perceptron of Rosenblatt (Rosenblatt, 1958). Figure 1 shows the schematic for this model of neuron and eq. 1 shows the model of a single neuron.

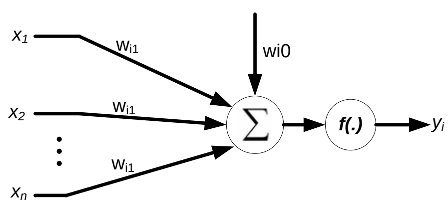

Figure 1. Schematic of artificial neuron

$$y_i = f(\sum_{j=1}^{n} w_{ij}x_j + w_{i0}) \quad (1)$$

where y_i is the output of neuron i, x_j is the j-th input, w_{ij} is the weigh given to input j when it is going to neuron i and w_{i0} is the bias of the neuron. Activation function of the neuron $f(.)$ can have several forms, such as sigmoid, linear, step or a radial basis function (S. S. Haykin, Haykin, Haykin, & Haykin, 2009).

Networks of neurons can be built by aligning neurons in single layers and by grouping the layers, forming a multilayer network. Figure 2 shows a NN with two layers of neurons.

Where x_n are the inputs of the network, L is the number of neuron in the first layer, v_{Ln} is the weigh from n-th input to L-th neuron. m is the number of neuron on the sec-

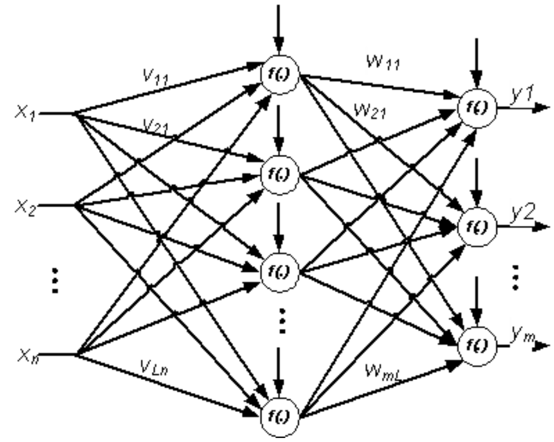

Figure 2. Schematics of a feedforward neural network

ond layer, w_{mL} is the weight from the output of L-th neuron on the first layer to the m-th neuron on the second layer. v_{0L} is the bias of the L-th neuron and w_{0m} if the bias of the m-th neuron.

Multilayer structures of NN, as the one shown in figure 2, are universal approximators, meaning that they can approximate or model any input pattern (Hornik, Stinchcombe, & White, 1989). This universal approximation feature makes NN feasible for modeling non-linear dynamics systems. By changing the NN input arrangements, it is possible to include temporal information about the system to be modeled. Among other input arrangement, the input structure used to model linear systems can be highlighted, such as the linear models of finite input response (FIR) and the autoregressive with exogenous input (ARX)(Nørgård et al., 2000). Figures 3 (a) and (b) shows the use of NN with input arrangement of a FIR and ARX model, respectively.

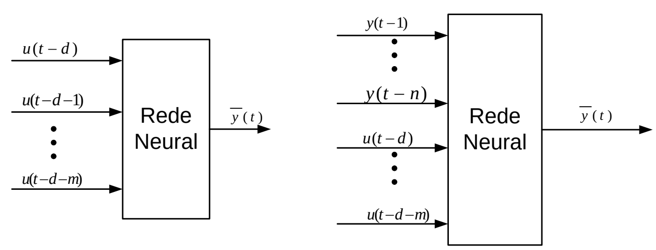

Figure 3. NN input structure: FIR (a) and ARX (b)

Figure 3 shows that the NN inputs include past values of the plant's input signal as well as the plant's response signal to the inputs. This input configuration gives the NN enough information to model the dynamics of a system.

System Identification

In this work the ARX structure were used for the identification of systems. The NN are placed in parallel with the plant and a series step signals, with normally distributed amplitudes, are generated as input to the plant. The input and output response of the plant are then arranged as an input to the NN. With the appropriate inputs, the NN are trained to best mimic the output of the plant. A validation set of input signals and plant output signals are used to test the NN on the ability to represent the plant for signals not previously used in the training. Matlab was used to implement the training inputs. Figure 4 illustrates an input signal used for identification, where the lower and upper bound of the signal is 0 and 1, and the steps last for 20 samples.

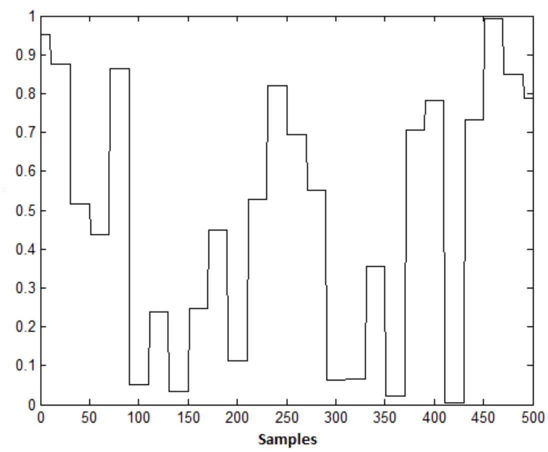

Figure 4. Example of a control signal used for NN training.

Neural Networks approximate functions by changing the weights for the neuron inputs. Many optimization methods can be used to optimize the matrix of NN weights, but in this work the Levenberg-Marquardt is used given its better performance in comparison with other methods (Moré, 1978).

Control Structure

Two strategies can be used for controlling dynamic systems: feedback control and optimization. The feedback control loop strategy includes using the controller and the plant in the same control loop, in a way that the entire system can described as a transfer function. Using optimization to control systems includes defining the system as an objective problem where the independent variables are the inputs to the plant. For this work the optimization strategy is used and the control structure follows that of a model prediction control (MPC), as initially presented by Clarke et. al. (Clarke, Mohtadi, & Tuffs, 1987). Figure 5 show a simplified structure of the predictive control structure.

In predictive control, the control problem is transformed into an optimization problem where the goal is to minimize the error between reference and output as well as the variability of control input at each interaction. The objective function presented in eq. 2 is minimized

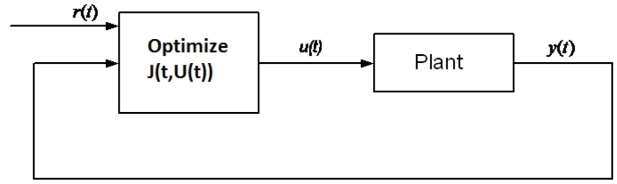

Figure 5. Simplified predictive control structure

$$J\big(t,U(t)\big) = \sum_{i=N_1}^{N_2}[r(t+i) - \hat{y}(t+i)]^2 + \rho\sum_{i=1}^{N_u}[u(t+i) - u(t+i-1)]^2 \quad (2)$$

where, N_1, N_2 and N_u are the minimum, maximum and control prediction horizon, respectively. $r(t)$ and $\hat{y}(t)$ is the reference and output prediction, respectively, from time t. $u(t)$ is the control signal at time t and ρ is a penalty weight given to the variation in control signal. The vector of control signals $U(t)$ are optimized at every time step t, where $U(t) = [u(t), u(t+1), \dots, u(t+N_u-1)]$.

The NN model of the plant is used to predict the plant's output $\hat{y}(t+i)$ at i steps ahead of t, then the prediction is used to optimize the future control signals. Figure 6 shows the structure of a MPC where a NN is used as the plant's model, where $\hat{Y}(t) = [\hat{y}(t+N_1), \hat{y}(t+N_1+1), \dots, \hat{y}(t+N_2)]$.

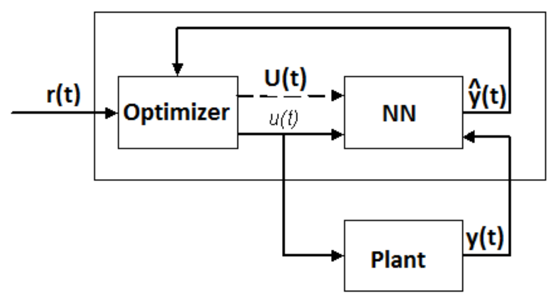

Figure 6. Predictive Control with NN model

Plants

Two models of non-linear plants were used to test the controller structure proposed. One is the non-linear model of a valve (Nørgård et al., 2000). The plant's mathematical model is shown on eq. 3. In order to illustrate the plant's non-linearity, figure 7 shows the output of the plant for a slope input.

$$x(t) = 1{,}4138x(t-1) - 0{,}6065x(t-2) + 0{,}1044u(t-1) + 0{,}0883u(t-2)$$

$$y(t) = \frac{x(t)}{\sqrt{0{,}1 + 0{,}9x(t)^2}} \quad (3)$$

A nonlinear model of a tank used for chemical reaction was also used to test the NN MPC controller. Figure 8

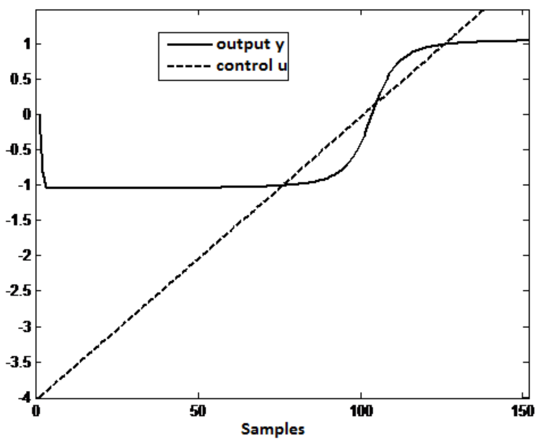

Figure 7. Control and output signal of the valve.

shows the schematics of the tank and eq. 4 illustrates the equation that models the tank

$$\frac{dh(t)}{dt} = w_1(t) + w_2(t) - 0.2\sqrt{h(t)}$$

$$\frac{dC_b(t)}{dt} = \left(C_{b1} - C_b(t)\right)\frac{w_1(t)}{h(t)} + \left(C_{b2} - C_b(t)\right)\frac{w_2(t)}{h(t)} - \frac{k_1 C_b(t)}{\left(1 + k_2 C_b(t)\right)^2} \quad (4)$$

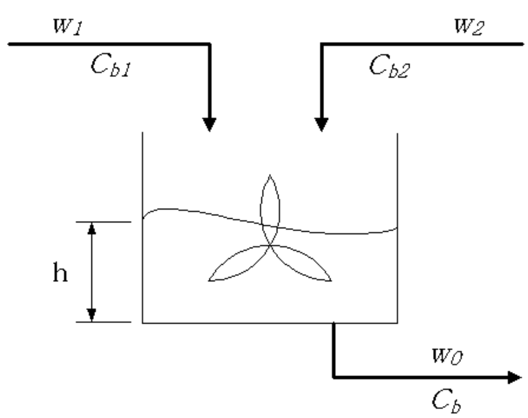

Figure 8. Schematic of chemical reaction tank

where $h(t)$ is the level of liquid in the tank, $C_b(t)$ is the concentration of the output product, $w_1(t)$ is the flow of concentrated C_{b1}, and w_2 is the flow of solvent C_{b2}. In this work, the concentration C_{b1} and C_{b2} is 24.9 and 0.1, respectively, following the work of Nørgård et. al. (Nørgård et al., 2000) for the same plant. The goal is to control the output concentration $C_b(t)$ by varying the flow $w_1(t)$. The flow $w_2(t)$ is left at a constant rate of 0.1.

Results and Discussion

Both plants are single input, single output models. In order to control the plants the first step was to model the plant using NN.

Valve

To identify the dynamics of the nonlinear valve, an input signal was generate consisting of 6000 samples, with amplitude varying from 0 to 1 at every 20 samples.

It was assumed that a NN with two layers would be sufficient to model the nonlinear dynamic of the valve. The first layer had 15 neurons and the second layer has one neuron. The second layer serves as a summation of outputs from the first layer and due that it has a linear activation function, while the first layer has a hyperbolic tangent activation function. The input structure $X(t)$ used for the NN followed an ARX structure with delayed inputs and delayed outputs as illustrated in eq. 5

$$X(t) = \begin{bmatrix} y(t-1) \\ \vdots \\ y(t-n_a) \\ u(t-d) \\ u(t-d-1) \\ \vdots \\ u(t-d-n_b+1) \end{bmatrix} \quad (5)$$

where n_a=3, n_b=8 and d = 2.

The training performance of the NN is show in figure 9, where it can be seen that the validation performance based on the Mean Squared Error (MSE) is negligible. Figure 10 illustrates the plant output and the NN output for the same set of inputs. Notice that the NN is capable of closely represent the valves dynamic.

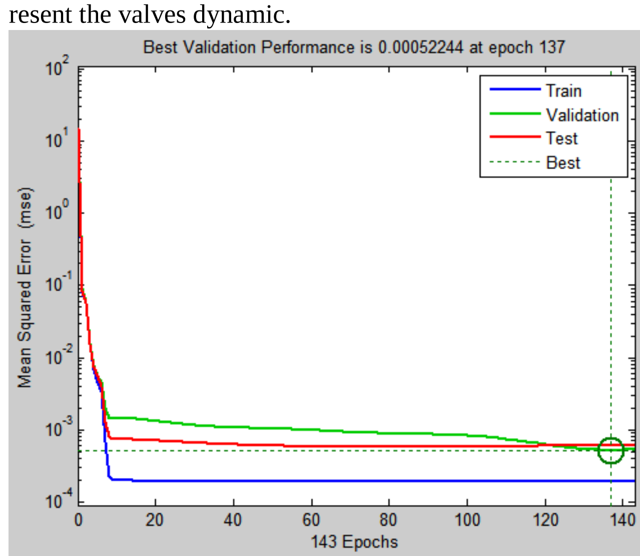

Figure 9. Training performance of NN to model nonlinear valve

Using the NN model of the plant in the MPC control loop it was possible to control the output of the plant. The optimization method used is a classic levenberg-marquardt. The minimum, maximum and control prediction horizon are: $N_1 = 1, N_2 = 7$ and $N_u = 1$. The penalty for signal control variation is $\rho = 10$. Figure 11 shows the inputs and outputs of the system, using a MPC controller and the NN as a model for prediction. The red dotted line is the reference of the system (r), the light line is the control signal of the plant (u) and the bold line is the output of the plant.

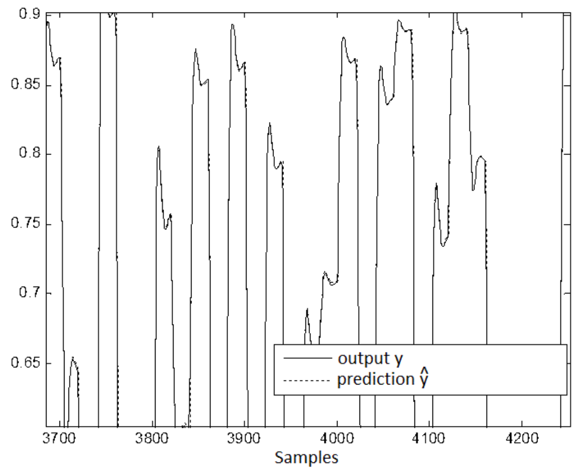

Figure 10. Output of valve and NN model for the same set of inputs

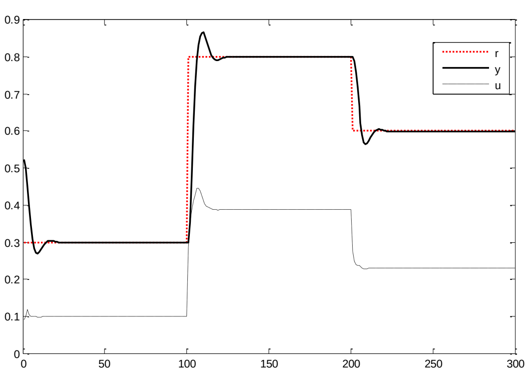

Figure 11. Signals of reference, control and plant output for a MPC controller using a NN model for the control of a valve

It can be seen that despite the non-linearity of the plant, the controller was able to efficiently control the plant, with a rapid response time.

Reaction Tank

To identify the dynamics of the nonlinear reaction tank, an input signal was generate consisting of 4000 samples, with amplitude varying from 0 to 5 at every 20 samples.

It was assumed that a NN with two layers would be sufficient to model the nonlinear dynamic of the tank. The first layer had 12 neurons and the second layer had just one neuron. The input structure $X(t)$ used for the NN followed an ARX structure with delayed inputs, delayed outputs and following the format in eq. 5, with n_a= 4, n_b=5 and d = 1.

The training performance of the NN in modeling the reaction tank is shown in figure 12, where it can be seen that MSE is negligible. Figure 13 illustrates the plant's output and NN prediction for the same set of inputs, it can be seen that both signals overlap, suggesting that the plant model the plant's dynamic efficiently.

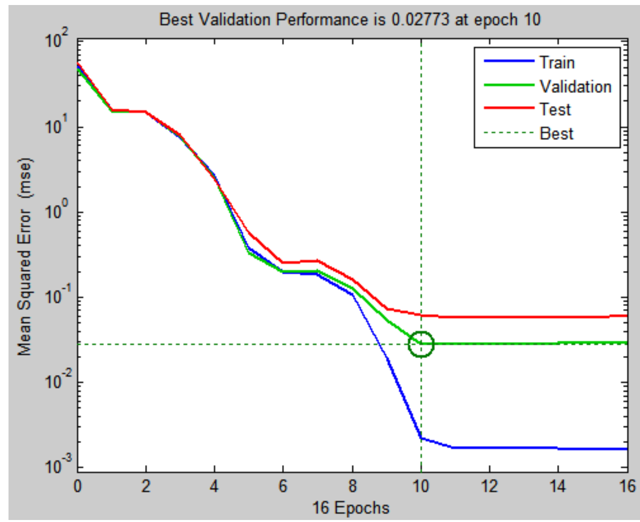

Figure 12. Training performance of NN to model nonlinear reaction tank

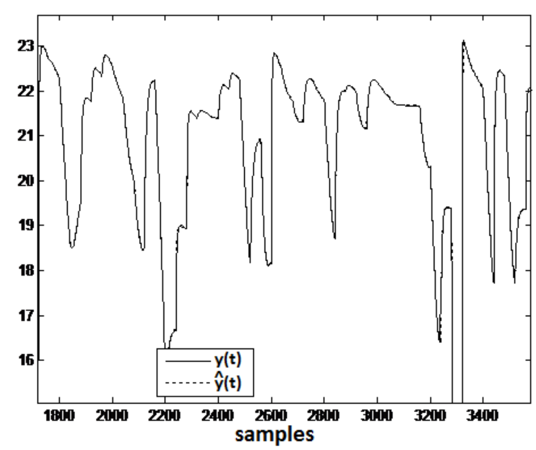

Figure 13. Concentration outputs from reaction tank and NN model for the same set of inputs

The NN model was used in a MPC structure to control the reaction tank plant. The method of levenberg-marquardt was used for optimization in the MPC structure. The minimum, maximum and control prediction horizon are: $N_1 = 1, N_2 = 7$ and $N_u = 2$. The penalty for signal control variation is $\rho = 0.05$. Figure 14 illustrates the inputs and outputs of the system, where the red dotted line is the reference of the system (r), the light line is the control signal of the plant (u) and the bold line is the output of the plant (y).

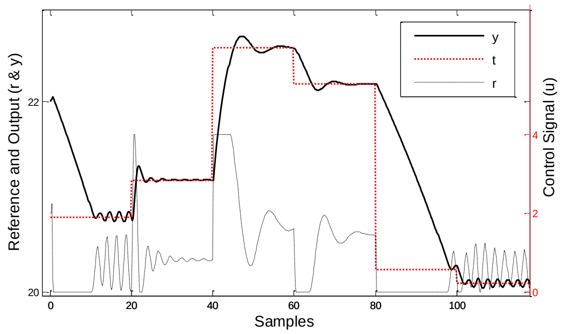

Figure 14. Signals of reference, control and plant output for a MPC structure using a NN model for the control of a chemical reaction tank

Despite of the non-linearity of the plant, the controller was able to execute control of the plant. However, it can be seen that, for lower reference levels, the controller is not able to stabilize the output of the plant around that reference value. This can be due the fact that the NN did not model the dynamics of the plant for such lower reference levels. The training set used for identification of the plant should have included more data in the lower reference levels. The NN do not correctly represent the plant for such low levels of reference, therefore it predicts erroneous plant's output. That makes the optimizer to optimize an objective function that is not representative of the plant, causing the marginal control performance in the lower reference levels.

For higher reference levels, the controller was able to control the concentration in the reaction tank. To maintain the steady levels of the plant's output, the control signal constantly changes in the time interval from 40s to 80s. This illustrates the optimizer trying to compensate for future changes in the output and adjusting the control signal ahead of time, so that the future output would follow the reference input. In this work, the performances of the NN in a predictive control structure are in close agreement with the work of Nørgård et. al. (Nørgård et al., 2000).

Conclusions

In this paper, multilayer, feedforward neural networks were used to identify the dynamics of two nonlinear plants, a valve and a reaction tank. Random step signals were used as input in the identification of the plants, while their responses were recorded. The set of inputs and outputs were used to train the NNs. Despite the simplicity of the NNs used, the models proved satisfactory to represent the plants for the range of inputs used in training.

The NNs of the plants were used in a control loop with a MPC structure. Given a set of control inputs, the NN were used to provide predictions of plants outputs. The output predictions are used to calculate the error from a desired reference signal. A levenberg-marquardt optimization method was used to optimize the control inputs in order to minimize the plant's output error.

For both plants in this work, the NN proved to be efficient in modeling the non-linearity of the plant. Additionally, the use of NN models in a MPC structure made possible the control of the nonlinear plants, where the controller would compensate for the plant's nonlinearities. The controller had a marginal performance in controlling the reaction tank for low reference levels. What is explained by the fact that the NN did not capture the dynamics of the plant for those levels. During the training of the NN for modeling the reaction tank, more low levels of reference should be the used so that the NN could have more information about the dynamics of the plant in those levels and therefore build a more accurate model of the plant.

The use of NN in the control of systems makes it possible for the control of nonlinear systems, black box systems and system with changing dynamics. The same methods used in this work for identification and control of systems can be extended to the identification and control of other complex systems.

References

Addeh, J., Ebrahimzadeh, A., Azarbad, M., & Ranaee, V. (2014). Statistical process control using optimized neural networks: A case study. *ISA transactions, 53*(5), 1489-1499.

Chen, M., Ge, S. S., & Voon Ee How, B. (2010). Robust adaptive neural network control for a class of uncertain MIMO nonlinear systems with input nonlinearities. *Neural Networks, IEEE Transactions on, 21*(5), 796-812.

Clarke, D. W., Mohtadi, C., & Tuffs, P. (1987). Generalized predictive control—Part I. The basic algorithm. *Automatica, 23*(2), 137-148.

Dierks, T., & Jagannathan, S. (2010). Output feedback control of a quadrotor UAV using neural networks. *Neural Networks, IEEE Transactions on, 21*(1), 50-66.

Ebrahimzadeh, A., & Ranaee, V. (2010). Control chart pattern recognition using an optimized neural network and efficient features. *ISA transactions, 49*(3), 387-393.

Franklin, G. F., Powell, J. D., & Emami-Naeini, A. (2006). Feedback control of dynamics systems. *Pretince Hall Inc*.

Haykin, S., & Network, N. (2004). A comprehensive foundation. *Neural Networks, 2*(2004).

Haykin, S. S., Haykin, S. S., Haykin, S. S., & Haykin, S. S. (2009). *Neural networks and learning machines* (Vol. 3): Pearson Education Upper Saddle River.

Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks, 2*(5), 359-366.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). *Imagenet classification with deep convolutional neural networks.* Paper presented at the Advances in neural information processing systems.

Moré, J. J. (1978). The Levenberg-Marquardt algorithm: implementation and theory *Numerical analysis* (pp. 105-116): Springer.

Nørgård, P. M., Ravn, O., Poulsen, N. K., & Hansen, L. K. (2000). *Neural Networks for Modelling and Control of Dynamic Systems-A Practitioner's Handbook*: Springer-London.

Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review, 65*(6), 386.

Visioli, A. (2006). *Practical PID control*: Springer Science & Business Media.

Zainuddin, Z., & Pauline, O. (2011). Modified wavelet neural network in function approximation and its application in prediction of time-series pollution data. *Applied Soft Computing, 11*(8), 4866-4874.