

Neural Network Predictive Control of a Bioprocess

Dorin Sendrescu and Emil Petre

Department of Automatic Control, University of Craiova, Craiova, Romania

e-mail: dorins@automation.ucv.ro, epetre@automation.ucv.ro

Abstract. This paper deals with the design of a nonlinear model predictive control (NMPC) scheme for the regulation of the acetate concentration in a biomethanation process – wastewater biodegradation with production of methane gas that takes place inside a Continuous Stirred Tank Bioreactor. The NMPC control structure is based on the nonlinear model of the process whose parameters are identified using the distributions method and the unknown reaction rates are estimated using a radial basis neural network. Minimization of the cost function is realized using the Levenberg–Marquardt numerical optimization method. Some simulation results are given to illustrate the efficiency of the proposed control strategy.

Keywords–neural network; biotechnological process; model predictive control

1. Introduction

In the last period, the control of biotechnological processes has been an important problem attracting wide attention. The main engineering motivation in applying control methods to such processes is to improve operational stability and production efficiency. But the use of modern control for these bioprocesses is still low. Nonlinear model predictive control (NMPC) is needed especially for nonlinear, unsteady batch processes where a trajectory needs to be followed from the prediction of a nonlinear model. It is especially useful for processes operating at or near singular points that cannot be captured by linear controllers and where higher order information is needed. NMPC uses the nonlinear dynamic model to predict the effect of sequences of control steps on the controlled variables.

This paper deals with the predictive control of a wastewater biodegradation process. The dynamics of this biotechnological process are described by a set of nonlinear differential equations obtained from the reaction scheme and the unknown reaction rates are estimated using a radial basis neural network. For the estimation of unknown parameters of the process, the distribution approach is used. The parameter identification of deterministic nonlinear continuous-time systems (NCTS), modeled by polynomial type differential equation has been considered by numerous authors [1], [2]. In this paper, we use an identification method for a class of NCTS considering that the unknown parameters can appear in rational relations with measured variables. Using techniques utilized in distribution approach, the measurable functions and their derivatives are represented by functionals on a fundamental space of testing functions. Such systems are common in biotechnology [3]. The main idea is to use a hierarchical structure of identification. First, some state equations are utilized to obtain a set of linear equations in some parameters. The results of this first stage of identification are utilized for expressing other parameters by linear equations. This process is repeated until all parameters are identified.

The paper is organized as follows. Section II is devoted to description and modeling of an anaerobic digestion bioprocess. The nonlinear model predictive control strategy is presented in Section III. Simulations results presented in Section IV illustrate the performance of the proposed control algorithms and, finally, Section V concludes the paper.

2. Process Modeling

2.1. Analytical approach of process modeling

In the conventional MPC controller, a linear predictive model is used because the theory of the identification of a linear system has well been established. The nonlinear part of the system response is treated as disturbance. But, a linear model, no matter how well has it been structured and tuned, may be acceptable only in the case where the system is working around the operating point. If the system is highly nonlinear, such as biotechnological processes, control based on the prediction from a linear model may result in unacceptable response. In some cases, remarkable static errors exist, and in other cases, oscillation or even instability may occur. Therefore, some kinds of non-linear models should be used to describe the behaviour of a highly non-linear system.

In this paper one considers a biomethanation process – wastewater biodegradation with production of methane gas that takes place inside a Continuous Stirred Tank Bioreactor whose reduced model is presented in [4]. It is a two phase process. In the first phase, the glucose from the wastewater is decomposed in fat volatile acids (acetates, propionic acid), hydrogen and inorganic carbon under action of the acidogenic bacteria. In the second phase, the ionised hydrogen decomposes the propionic acid $\text{CH}_3\text{CH}_2\text{COOH}$ in acetates, H_2 and carbon dioxide CO_2 . In the first methanogenic phase, the acetate is transformed into methane and CO_2 , and finally in the second methanogenic phase, the methane gas CH_4 is obtained from H_2 and CO_2 , [5]. The following simplified reaction scheme is considered,



where: S_1 represents the glucose substrate, S_2 the acetate substrate, X_1 is the acidogenic bacteria, X_2 the acetoclastic methanogenic bacteria and P_1 represents the product, i.e. the methane gas. The reaction rates are denoted by Φ_1, Φ_2 . The corresponding dynamical model is:

$$\frac{d}{dt} \begin{bmatrix} X_1 \\ S_1 \\ X_2 \\ S_2 \\ P_1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -k_1 & 0 \\ 0 & 1 \\ k_2 & -k_3 \\ 0 & k_4 \end{bmatrix} \begin{bmatrix} \Phi_1 \\ \Phi_2 \end{bmatrix} - D \begin{bmatrix} X_1 \\ S_1 \\ X_2 \\ S_2 \\ P_1 \end{bmatrix} + \begin{bmatrix} 0 \\ DS_{in} \\ 0 \\ 0 \\ -Q_1 \end{bmatrix} \quad (2)$$

where S_{in} is the influent substrate, Q_1 the methane gas outflow rate, D the dilution rate and the state vector of the model is:

$$\xi = [X_1 \ S_1 \ X_2 \ S_2 \ P_1]^T = [\xi_1 \ \xi_2 \ \xi_3 \ \xi_4 \ \xi_5]^T \quad (3)$$

whose components are concentrations in (g/l). The reaction rates are nonlinear functions of the state components, expressed as

$$\Phi = \Phi(\xi) = \begin{bmatrix} \Phi_1(\xi) \\ \Phi_2(\xi) \end{bmatrix} \quad (4)$$

The reaction rates for this process are given by the Monod law

$$\Phi_1(\xi) = \mu_1 \frac{S_1 \cdot X_1}{K_{M_1} + S_1} \quad (5)$$

and the Haldane kinetic model:

$$\Phi_2(\xi) = \mu_2 \frac{S_2 \cdot X_2}{K_{M_2} + S_2 + S_2^2 / K_i} \quad (6)$$

where K_{M_1}, K_{M_2} are Michaelis-Menten constants, μ_1, μ_2 represent specific growth rates coefficients and K_i is the inhibition constant.

For simplicity, shall we denote the plant parameters by the vector:

$$\theta = [\theta_1 \ \theta_2 \ \theta_3 \ \theta_4 \ \theta_5 \ \theta_6 \ \theta_7 \ \theta_8 \ \theta_9]^T \quad (7)$$

where:

$$\begin{aligned} \theta_1 = k_1; \ \theta_2 = k_2; \ \theta_3 = k_3; \ \theta_4 = k_4; \ \theta_5 = \mu_1; \\ \theta_6 = \mu_2; \ \theta_7 = K_{M_1}; \ \theta_8 = K_{M_2}; \ \theta_9 = K_i \end{aligned} \quad (8)$$

Because the dilution rate D can be externally modified, it will be considered the third component of the input vector

$$u = [u_1 \ u_2 \ u_3]^T \quad (9)$$

The other two components of u are the concentration S_m and the methane gas outflow rate Q_1 so,

$$u_1 = S_m; \ u_2 = Q_1; \ u_3 = D \quad (10)$$

Usually Q_1 depends on state variables, $Q_1 = \psi(\xi)$; determining a feedback to the input u_2 . Written explicitly by components, the state equation (2), within the above notations, takes the form:

$$\dot{\xi}_1 = \Phi_1 - u_3 \cdot \xi_1 \quad (11)$$

$$\Phi_1 = \theta_5 \cdot \frac{\xi_1 \cdot \xi_2}{\theta_7 + \xi_2} \quad (12)$$

$$\dot{\xi}_2 = -\theta_1 \cdot \Phi_1 - u_3 \cdot \xi_2 + u_1 \cdot u_3 \quad (13)$$

$$\dot{\xi}_3 = \Phi_2 - u_3 \cdot \xi_3 \quad (14)$$

$$\Phi_2 = \theta_6 \cdot \frac{\xi_3 \cdot \xi_4}{\theta_8 + \xi_4 + \theta_9 \cdot \xi_4^2}, \ \theta_9' = \frac{1}{\theta_9} \quad (15)$$

$$\dot{\xi}_4 = \theta_2 \cdot \Phi_1 - \theta_3 \cdot \Phi_2 - u_3 \cdot \xi_4 \quad (16)$$

$$\dot{\xi}_5 = -u_3 \cdot \xi_5 + \theta_4 \cdot \Phi_2 - u_2 \quad (17)$$

2.2. Parameters estimation

For parameters estimation the distribution based method was used. In this approach the set of nonlinear differential equations describing the state evolution is mapped into a set of linear algebraic equations respect to the model parameters. Using techniques utilized in distribution approach, the measurable functions and their derivatives are represented by functionals on a fundamental space of testing functions. The main advantages of this method are that a set of algebraic equation with real coefficients results and the formulations are free from boundary conditions.

Let Φ_n be the fundamental space from the distribution theory of the real functions $\varphi: R \rightarrow R, t \rightarrow \varphi(t)$ and $q: R \rightarrow R, t \rightarrow q(t)$ a function which admits a Riemann integral on any compact interval T from \mathbf{R} . Using this function, a unique distribution

$$F_q: \Phi_n \rightarrow R, \varphi \rightarrow F_q(\varphi) \in R \quad (18)$$

can be built by the relation:

$$F_q(\varphi) = \int_R q(t)\varphi(t)dt, \forall \varphi \in \Phi_n \quad (19)$$

In distribution theory, the notion of k-order derivative is introduced. If $F_q \in \Phi_n$, then its k-order derivative is a new distribution $F_q^{(k)} \in \Phi_n$ uniquely defined by the relations:

$$F_q^{(k)}(\varphi) = (-1)^k F_q(\varphi^{(k)}), \forall \varphi \in \Phi_n \quad (20)$$

$$\varphi \rightarrow F_q^{(k)}(\varphi) = (-1)^k \int_R q(t) \varphi^{(k)}(t) dt \in R \quad (21)$$

where

$$\varphi^{(k)} : R \rightarrow R, t \rightarrow \varphi^{(k)}(t) = \frac{d^k \varphi(t)}{dt^k} \quad (22)$$

is the k -order time derivative of the testing function.

When $q \in C^k(R)$, then

$$F_q^{(k)}(\varphi) = \int_R q^{(k)}(t) \varphi(t) dt = (-1)^k \int_R q(t) \varphi^{(k)}(t) dt, \quad (23)$$

that means the k -order derivative of a distribution generated by a function $q \in C^k(R)$ equals to the distribution generated by the k -order time derivative of the function q .

So, in place of the states and their time derivatives of a system one utilizes the corresponding distributions and, in some particular cases, it is possible to obtain a system of equations linear in parameters. If the system is compatible then all the model parameters are structurally identifiable.

Consider all state variables accessible for measurements. The dynamical system (11) - (17) contains rational dependencies between parameters and measured variables. To obtain linear equations in unknown parameters, the identification problem is split in several simpler problems. Based on the specific structure of this system, it is possible to group the state equations, in such way to determine five interconnected relations. They are organized in a hierarchical structure. In the first stage some state equations are utilized to obtain a set of linear equations in some parameters. If these parameters are identified then they can be used as known parameters in the following stages. This process is repeated in the other stages until all the parameters are identified.

3. Nonlinear Model Predictive Control

3.1. Problem formulation

Consider the following discrete-time, time-invariant nonlinear system:

$$\xi_{k+1} = f(\xi_k, u_k) \quad (24)$$

with ξ_k the state vector, u_k the control signal, corresponding in our case with the discretization of system (2). The objective is to regulate the state vector (or a particular output signal) to a specified setpoint value while guaranteeing that certain input and state constraints:

$$\begin{aligned} \xi_{\min} &\leq \xi_k \leq \xi_{\max} \\ u_{\min} &\leq u_k \leq u_{\max} \end{aligned} \quad (25)$$

Nonlinear model predictive control treats such a constrained control problem by repeatedly solving the following optimization problem:

$$\min \sum_{i=0}^{N-1} (\xi_{ref} - \xi_{k+i})^T Q (\xi_{ref} - \xi_{k+i}) + u_{k+i}^T R u_{k+i} \quad (26)$$

$$s.t. \begin{cases} \xi_{k+1} = f(\xi_k, u_k) \\ \xi_{\min} \leq \xi_k \leq \xi_{\max} \\ u_{\min} \leq u_k \leq u_{\max} \end{cases} \quad (27)$$

Among the whole sequence resulting from the on-line optimization, only the first optimal control is applied as input to the system [6]. At the next sampling time, the current state is obtained (measured or

estimated) and the optimization problem (26), (27) is solved again with this new initial state value, according to the well-known receding horizon principle.

3.2. The Neural Network Model

The predictive model for a conventional MPC controller is usually a linear model which is preferred as being more intuitive and requiring less a priori information for its identification. MPC based on linear models is acceptable if the process operates at a single setpoint and the primary use of the controller is the rejection of disturbances. Many chemical processes, including polymer reactors, do not operate at a single setpoint. However, these models are not suitable for a nonlinear system such as biotechnological processes. To solve this problem neural networks are proposed to obtain the estimated output used by the MPC controller, because the neural networks have the ability to map any nonlinear relationships between an input and output set. There have also been many reports on the application of neural network to bioprocesses modelling and identification [6], [8].

In this paper the process model is obtained using a radial basis neural network (RBNN) with adjustable parameters to approximate the reaction rates Φ_1 and Φ_2 from model (2). A RBNN is made up of a collection of $p > 0$ parallel processing units called nodes. The output of the i th node is defined by a Gaussian function $\gamma_i(x) = \exp(-|x - c_i|^2 / \sigma_i^2)$, where $x \in \mathfrak{R}^n$ is the input to the NN, c_i is the centre of the i -th node, and σ_i is its size of influence. The output of a RBNN, $y_{NN} = F(x, W)$, may be calculated as [10]

$$F(x, W) = \sum_{i=1}^p w_i \gamma_i(x) = W^T(t) \Gamma(x) \quad (28)$$

where $W(t) = [w_1(t) \ w_2(t) \ \dots \ w_p(t)]^T$ is the vector of network weights and $\Gamma(x)$ is a set of radial basis functions defined by $\Gamma(x) = [\gamma_1(x) \ \gamma_2(x) \ \dots \ \gamma_p(x)]^T$.

Given a RBNN, it is possible to approximate a wide variety of functions $f(x)$ by making different choices for W . In particular, if there is a sufficient number of nodes within the NN, then there is some W^* such as $\sup_{x \in S_x} |F(x, W^*) - f(x)| < \varepsilon$, where S_x is a compact set, $\varepsilon > 0$ is a finite constant, provided that $f(x)$ is continuous [10]. The RBNN is used to estimate the reaction rates Φ_1 and Φ_2 (that are considered unknown) using some state measurements.

3.3. The Neural Network Model Predictive Control Algorithm

The model predictive control is a strategy based on the explicit use of system model to predict the controlled variables over a certain time horizon, called the prediction horizon. The control strategy can be described as follows:

- 1) Using the on-line measurements the unknown dynamics of the system are estimated using an ANN.
- 2) At each sampling time, the value of the controlled variable y_{t+k} is predicted over the prediction horizon $k=1, \dots, N_y$. This prediction depends on the future values of the control variable u_{t+k} within a control horizon $k=1, \dots, N_u$.
- 3) A reference trajectory y_{t+k}^{ref} , $k=1, \dots, N$ is defined which describes the desired system trajectory over the prediction horizon.
- 4) The vector of future controls u_{t+k} is computed such that an objective function (a function of the errors between the reference trajectory and the predicted output of the model) is minimised.
- 5) Once the minimisation is achieved, the first optimised control actions are applied to the plant and the plant outputs are measured. These measurements are used as the initial states of the model to perform the next iteration. Steps 1 to 5 are repeated at each sampling instant; this is called a receding horizon strategy.

The strategy of the NMPC based control is characterized by scheme represented in “Fig. 1”.

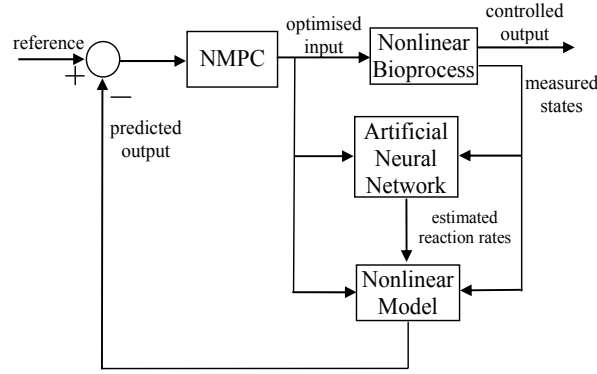


Figure 1. NMPC control scheme.

When a solution of the nonlinear least squares (NLS) minimization problem cannot be obtained analytically, the NLS estimates must be computed using numerical methods. To optimize a nonlinear function, an iterative algorithm starts from some initial value of the argument in that function and then repeatedly calculates next available value according to a particular rule until an optimum is reached approximately. Between many different methods of numerical optimization the Levenberg-Marquardt (LM) algorithm was chosen to solve the optimisation problem.

The LM algorithm is an iterative technique that locates the minimum of a multivariate function that is expressed as the sum of squares of non-linear real-valued functions. It has become a standard technique for non-linear least-squares problems, widely adopted in a broad spectrum of disciplines. LM can be thought of as a combination of steepest descent and the Gauss-Newton method. When the current solution is far from the correct one, the algorithm behaves like a steepest descent method. When the current solution is close to the correct solution, it becomes a Gauss-Newton method.

4. Simulation Results

In this Section we will apply the designed nonlinear model predictive control in the case of the anaerobic digestion bioprocess presented in Section 2. In order to control the output pollution level y , as input control we chose the dilution rate, $u = D$. The main control objective is to maintain the output y at a specified low level pollution $y_d \in \mathfrak{R}$. We will analyze the realistic case where the structure of the system of differential equation (2) is known and specific reaction rates Φ_1 and Φ_2 (described by “(6)” and “(7)”) are completely unknown and must be estimated. Using a RBNN from subsection III.B, one constructs an on-line estimate of Φ_1 respectively of Φ_2 .

The performance of the nonlinear predictive controller presented in subsection III.C has been tested through extensive simulations by using the process model “(2)”. The values used for yield and kinetic coefficients are: $k_1 = 3.2$, $k_2 = 16.7$, $k_3 = 1.035$, $k_4 = 1.1935$, $k_5 = 1.5$, $k_6 = 3$, $k_7 = 0.113$, $\mu_1^* = 0.2 \text{ h}^{-1}$, $K_{M_1} = 0.5 \text{ g/l}$, $\mu_2^* = 0.35 \text{ h}^{-1}$, $K_{M_2} = 4 \text{ g/l}$, $K_{I_2} = 21 \text{ g/l}$, and the values $\alpha_1 = 1.2$, $\alpha_2 = 0.75$. It must be noted that for the reaction rates estimation a full RBNN with deviation $\sigma_i = 0.05$ was used. The centres c_i of the radial basis functions are placed in the nodes of a mesh obtained by discretization of states $X_1 \in [1, 12] \text{ g/l}$, $X_2 \in [0.4, 0.65] \text{ g/l}$, $S_1 \in [0.1, 1.4] \text{ g/l}$ and $S_2 \in [0.3, 1.6] \text{ g/l}$ with $dX_i = dS_i = 0.2 \text{ g/l}$, $i = 1, 2$.

The simulation results, obtained with a sample period $T_s = 6 \text{ min}$ are presented in Figs. 2 – 5. In “Fig. 2” the controlled output trajectory is presented and in “Fig. 3” the nonlinear model predictive control action (dilution rate D evolution) is depicted. The functions Φ_1 and Φ_2 provided by the RBNN are depicted versus the “real” functions in “Fig. 4” and “Fig. 5”. From these figures it can be seen that the behaviour of the control system with NMPC controller is very good, although the process dynamics are incompletely known. The control action has an oscillatory behavior, but these oscillations are relatively slow and with small magnitude.

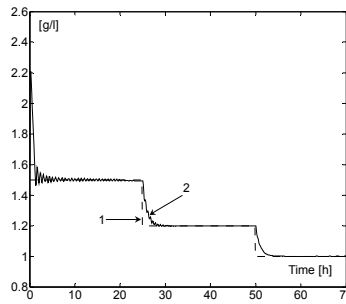


Figure 2. The controlled output evolution (reference (1) and controlled output (2)).

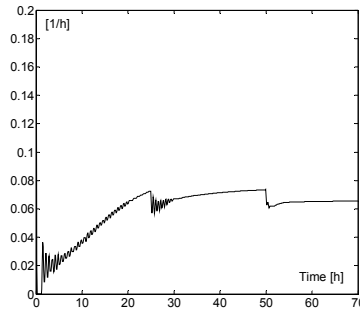


Figure 3. The nonlinear model predictive control action (dilution rate D).

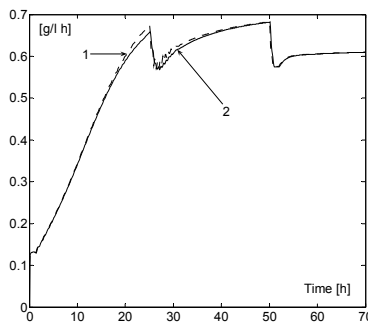


Figure 4. The real reaction rate Φ_1 (1) versus the function provided by the RBNN (2)

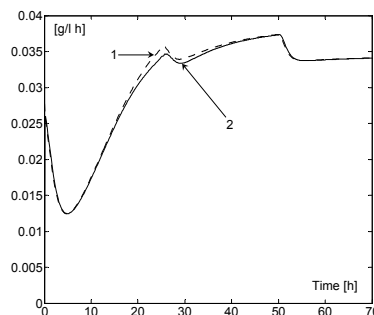


Figure 5. The real reaction rate Φ_2 (1) versus the function provided by the RBNN (2)

5. Conclusions

In this paper, a nonlinear model predictive control strategy was developed for a wastewater treatment bioprocess. The nonlinear model used by the control algorithm was obtained using the analytical description of the biochemical reactions. The unknown reaction rates are estimated using radial basis neural networks.

The nonlinear model states are used to calculate the optimal control signal applied to the system. The optimization problem was solved using the iterative Levenberg-Marquardt algorithm.

6. Acknowledgment

This work was supported by the National University Research Council - CNCSIS, Romania, under the research project TE 106, no. 9/04.08.2010.

7. References

- [1] A. Patra, H. Unbehauen, 1995. Identification of a class of nonlinear continuous time systems using Hartley modulating functions. *Int.J.Contr.* Vol.62, No.6,1431-1452.
- [2] A. Pearson, F. Lee, 1985. On the identification of Polynomial input-output differential systems. *IEEE Trans. Aut. Contr.*,vol. AC30, No8,778-782.
- [3] G. Bastin, D. Dochain, *On-line Estimation and Adaptive Control of Bioreactors*, Elsevier, Amsterdam, 1990.
- [4] D. Dochain, P. Vanrolleghem, *Dynamical Modelling and Estimation in Wastewater Treatment Processes*. IWA Publishing, 2001.
- [5] E. Petre, D. Selisteanu, *Modelling and Identification of Depollution Bioprocesses*, Universitaria, Craiova, 2005.
- [6] B. Kouvaritakis and M. Cannon, *Nonlinear predictive control: Theory and practice*. IEE, 2001.
- [7] E.F Camacho and C. Bordons, *Model Predictive Control*, second edition, Springer, 2004.
- [8] J.W. Eaton and J.R. Rawlings, "Feedback Control of Nonlinear Processes Using Online Optimization Techniques," *Computers and Chemical Engineering*, vol. 14, pp. 469-479, 1990.
- [9] J. Nocedal and S.J. Wright. *Numerical Optimization*. Springer, New York, 1999.
- [10] Spooner, J.T., Passino, K.M.: Decentralized adaptive control of nonlinear systems using radial basis neural networks. *IEEE Trans. on Autom. Control*, 44, 11, 2050-2057, 1999