World Academy of Science, Engineering and Technology
International Journal of Chemical and Molecular Engineering
Vol:1, No:12, 2007

Application of Feed Forward Neural Networks in Modeling and Control of a Fed-Batch Crystallization Process

Petia Georgieva, and Sebastião Feyo de Azevedo

***Abstract*—**This paper is focused on issues of nonlinear dynamic process modeling and model-based predictive control of a fed-batch sugar crystallization process applying the concept of artificial neural networks as computational tools. The control objective is to force the operation into following optimal supersaturation trajectory. It is achieved by manipulating the feed flow rate of sugar liquor/syrup, considered as the control input. A feed forward neural network (FFNN) model of the process is first built as part of the controller structure to predict the process response over a specified (prediction) horizon. The predictions are supplied to an optimization procedure to determine the values of the control action over a specified (control) horizon that minimizes a predefined performance index.

The control task is rather challenging due to the strong nonlinearity of the process dynamics and variations in the crystallization kinetics. However, the simulation results demonstrated smooth behavior of the control actions and satisfactory reference tracking.

***Keywords*—**Feed forward neural network, process modeling, model predictive control, crystallization process.

I. Introduction

THE phenomenon of crystallisation occurs in a large group of pharmaceutical, biotechnological, food and chemical processes. These kind of industrial productions are usually performed in a batch or fed-batch mode which is related with the formulation of a control problem in terms of economic or performance objective at the end of the process. The crystallisation quality is evaluated by the particle size distribution (PSD) at the end of the process which is quantified by two parameters - the average (in mass) particle size (MA) and the coefficient of particle variation (CV). The main challenge of the batch production is the large batch to batch variation of the final PSD. This lack of process repeatability is caused mainly by improper control policy and results in final product recycling and loss increase.

Due to the highly competitive nature of the today's crystallization industry, model-based predictive control becomes increasingly accepted as one of the approaches that can overcome the problem of repeatability and can drive the process to its optimal state of profit maximization and cost minimization [1], [2]. However, the crystallisation occurs through the complex mechanisms of particle nucleation, subsequent particle growth and agglomeration or aggregation, phenomena that are physically not well understood therefore their reliable modelling is still a challenging task [3]. For example many of the reported crystallizer models neglect the agglomeration effect but it leads in general to biased estimation of CV and MA [4].

Development of a reliable model facilitates effectively all subsequent steps in process optimization, control and operation monitoring. There are two main modelling paradigms - analytical (based on the first principles rules) which has been the traditional way of process modelling since many years and data-driven (based on the process data) which became nowadays practically meaningful due to the rapid growth of computational resources. One of the most successful data-driven modelling techniques are the artificial neural networks (ANNs). Their ability to approximate complex non-linear relationships without prior knowledge of the model structure makes them a very attractive alternative to the classical modelling techniques [5]- [7].

The purpose of this paper is twofold. On one hand we discuss and evaluate the benefits of applying hybrid strategy for dynamic behaviour modelling of crystallization processes combining analytical and ANN approaches. This mixed strategy is termed as Knowledge Based Hybrid Modelling (KBHM). On the other hand, we introduce Model Predictive Control (MPC) based on a Feed Forward Neural Network (FFNN) nonlinear model to an industrial fed-batch evaporative sugar crystallization, which belongs to the general class of crystallization processes. The aim is to regulate the process such that the supersaturation tracks a desired, process-dependent reference signal.

II. Process Operation

Crystallisation occurs through the mechanisms of nucleation, growth and agglomeration. The process is characterised by strongly non-linear and non-stationary dynamics and can be divided into several sequential phases.

Charging: During the first phase the pan is partially filled

P. Georgieva is with the Department of Electronics and Telecommunications/IEETA, University of Aveiro, 3810-193 Aveiro, Portugal (corresponding author, phone: +351-234-370531; fax: +351-234-370-545; e-mail: petia@det.ua.pt).

S. Feyo de Azevedo is with the Department of Chemical Engineering, Faculty of Engineering, University Porto, Rua Dr. Roberto Frias s/n, 4200-465 Porto, Portugal (e-mail: sfeyo@fe.up.pt).

with a juice containing dissolved sucrose (termed liquor).

Concentration: The next phase is the concentration. The liquor is concentrated by evaporation, under vacuum, until the supersaturation reaches a predefined value. At this stage seed crystals are introduced into the pan to induce the production of crystals. This is the beginning of the third (crystallisation) phase.

Crystallisation (main phase): In this phase as evaporation takes place further liquor or water is added to the pan in order to guarantee crystal growth at a controlled supersaturation level and to increase total contents of sugar in the pan. In most cases, due to economical reasons, the liquor is replaced by other juice of lower purity (termed syrup).

Tightening: The fourth phase consists of tightening which is principally controlled by the evaporation capacity. The pan is filled with a suspension of sugar crystals in heavy syrup, which is dropped into a storage mixer. At the end of the batch, the final massecuite undergoes centrifugation, where final refined sugar is separated from the (mother) liquor.

The unit contains 15 sensors for the following properties and operating variables: *i) inside the pan* - massecuite temperatures at three locations; brix of solution; level; massecuite consistency; stirrer current; vacuum pressure and temperature. *ii) feed conditions* - temperature, brix and flow rate of feed liquor and feed syrup. *iii) steam conditions* - temperature, pressure and flow rate of steam.

Brix is the concentration of total dissolved solids (sucrose plus impurities) in the solution. Supersaturation is not a measured variable but can be determined from the available measurements. More details about the process can be found elsewhere [4], [8].

III. Crystalization Process Modelling

A. Analytical Prior Knowledge Approach (White Box Model)

The traditional way of process modelling for many years has been by mathematical equations. Since the analytical models capture physical behaviour they have the potential to extrapolate beyond the regions for which the model was constructed. The general first principles model describing a batch crystallisation process consists of three parts [4].

Mass Balance

The mass of water (M_w), impurities (M_i), dissolved sucrose (M_s) and crystals (M_c) are included in the following set of conservation mass balance equations

$$\frac{dM_w}{dt} = F_f \rho_f (1 - B_f) + F_w \rho_w - J_{vap} \quad (1)$$

$$\frac{dM_i}{dt} = F_f \rho_f B_f (1 - Pur_f) \quad (2)$$

$$\frac{dM_s}{dt} = F_f \rho_f B_f Pur_f - J_{cris} \quad (3)$$

$$\frac{dM_c}{dt} = J_{cris} \quad (4)$$

where Pur_f and ρ_f are the purity (mass fraction of sucrose in the dissolved solids) and the density of the incoming feed. F_f is the feed flowrate considered as the process input.

Energy Balance

The general energy balance model is

$$\frac{dT_m}{dt} = aJ_{cris} + bF_f + cJ_{vap} + d \quad (5)$$

where J_{vap} is the evaporation rate and *a*, *b*, *c*, *d* are parameters incorporating the enthalpy terms and specific heat capacities derived as functions of physical and thermodynamic properties [8].

Population Balance

Mathematical representation of the crystallization rate can be achieved through basic mass transfer considerations [9] or by writing a population balance represented by its moment equations [10]. Employing a population balance is generally preferred since it allows to take into account initial experimental distributions and, most significantly, to consider complex mechanisms such as those of size dispersion and/or particle agglomeration/aggregation. Hence, the population balance is expressed by the leading moments of PSD in volume coordinates ($\tilde{\mu}_i$) since agglomeration must obey mass conservation low,

$$\frac{d\tilde{\mu}_0}{dt} = \tilde{B}_0 - \frac{1}{2}\beta' \tilde{\mu}_0^2 \quad (6)$$

$$\frac{d\tilde{\mu}_1}{dt} = G_v \tilde{\mu}_0 \quad (7)$$

$$\frac{d\tilde{\mu}_2}{dt} = 2G_v + \beta' \tilde{\mu}_1^2 \quad (8)$$

and the crystallisation rate is determined as

$$J_{cris} = \rho_c \frac{d\tilde{\mu}_1}{dt}. \quad (9)$$

$\tilde{B}_0$, G_v and β' are the kinetic variables nucleation rate, volume growth rate and the agglomeration kernel, respectively. It is difficult to formulate physically based analytical models for the kinetic variables (Fig. 1*)*. Here, the empirical correlations have a long tradition and there exist in the literature a large number of empirical equations for them [4], [8], [11]. The decision which of them provides the best approximation of the crystallisation process in hand is very difficult.

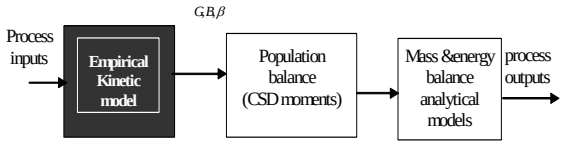


Fig. 1 Analytical model

B. ANN (Black Box Model)

An obvious advantage of the ANN modelling is its universal character in approximating different physical phenomena with similar computational structure. It saves time and efforts for identifying parameters, in contrast to the case when an analytical model is designed. Therefore ANNs are nowadays known as powerful computing structures for data processing and information storage. However, they have some remarkable disadvantages. The ANN approach suffers of the lack of transparent structure and physical understanding of the network parameters. The resulting black-box (input-output) model in general does not provide the transparency desired to enhance the process understanding. It relies only on the recorded data and does not exploit any other source of knowledge available for the process in hand.

A complete FFNN model of the sugar crystallisation was also developed. It has single input single output structure and one hidden layer with 7 sigmoid activation functions (Fig. 2). The input is related to on-line collected physical measurements of the feed flow rate. The network output is the supersaturation for which historical data are also available. The FFNN parameters were tuned applying the Levenberg-Marquart optimisation procedure [12].

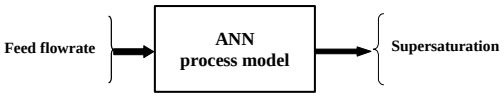


Fig. 2 FFNN model

C. Knowledge-Based Hybrid Modeling (Grey Box Model)

Knowledge-based hybrid modelling (KBHM) is a quite efficient alternative of the two modelling techniques discussed above [13]. The idea of KBHM is to complement the analytical model with the data-driven approach. In the design of such models it is possible to combine theoretical and experimental knowledge as well as process information from different sources: theoretical knowledge from physical and mass conservation laws; experimental data from laboratory plant experiments; experimental data from real plant experiments; data from regular process operation; knowledge and experience from qualified process operators. The clear advantages of KBHM compared with the data-based modelling are first with respect to more physical transparency of the model parameters and secondly less training data is required [14].

Our solution for a KBHM of the crystallization process combines a partial analytical model reflecting the mass, energy and population balances (1-9) with a feed-forward ANN for modelling the nucleation rate (B^{NN}), the growth rate (G^{NN}) and the agglomeration kernel (β^{NN}) (see Fig. 3). The ANN has 4 inputs, 3 outputs and one hidden layer with 9 sigmoid activation functions. The temperature of masecute (T_m), the supersaturation (S), the purity of the solution (Pur_{sol}) and the volume fraction of crystals (υ_c) are considered as the networks inputs because they all affect directly the kinetic parameters.

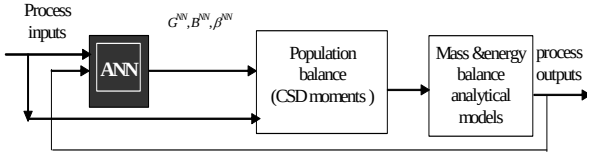


Fig. 3 KBHM

Hybrid ANN Training – Sensitivity Approach

The training of an ANN requires that the network weights are determined in such a way that the error between the network output and the corresponding target output becomes minimal. In the hybrid system, however, the target outputs are not available since the kinetic parameters are not measured. Therefore, an alternative training procedure was required. Our solution was to build a hybrid ANN training structure where the network outputs go through some fixed (known) part of the analytical model and to compare this hybrid model output with the available data (Fig. 4).

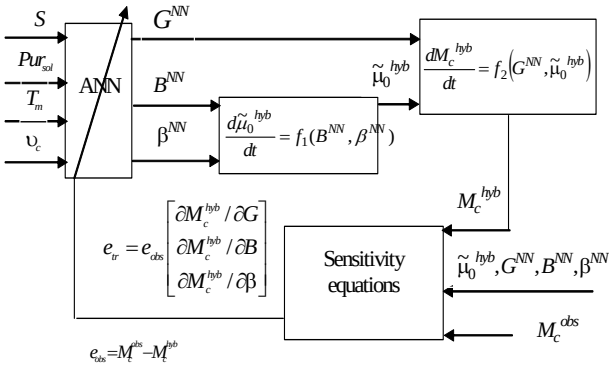


Fig. 4 Hybrid ANN training procedure

The error for updating the network weights is a function of the observed error and the gradient of the hybrid model output with respect to the ANN output. The mass of crystals is considered as most appropriate to serve as a target output in the hybrid ANN training. According to equations (4), (6) and (9), the mass balance of crystals can be rewritten as

$$\frac{dM_c^{hyb}}{dt} = 3(k_v \rho_c)^{1/3} (\tilde{\mu}_0^{hyb})^{1/3} (M_c^{hyb})^{2/3} G^{NN} \quad (10)$$

(10) is incorporated in the hybrid training structure but in order to integrate it the zero moment ($\tilde{\mu}_0$) is required.

Therefore its balance equation is also involved in the network training stage (see also (6)) ,

$$\frac{d\tilde{\mu}_0^{\,hyb}}{dt} = B^{NN} - \frac{1}{2}\beta^{NN}\left(\tilde{\mu}_0^{\,hyb}\right)^2 . \tag{11}$$

Superscripts *hyb* and *NN* are used to point out variables obtained during the hybrid network training. The network outputs give estimates of the growth rate, nucleation and agglomeration kinetic parameters. These estimates are propagated through (10-11). The error signal for updating the network parameters is

$$e_{tr} = e_{obs}\left[\lambda_G \quad \lambda_B \quad \lambda_\beta\right]^T \tag{12}$$

It is obtained by multiplying the observed error

$$e_{obs} = M_c^{\,obs} - M_c^{\,hyb} \tag{13}$$

with the gradient of the hybrid model output with respect to the network outputs

$$\lambda_G = \frac{\partial M_c^{hyb}}{\partial G} \tag{14}$$

$$\lambda_B = \frac{\partial M_c^{hyb}}{\partial B} \tag{15}$$

$$\lambda_\beta = \frac{\partial M_c^{hyb}}{\partial \beta} . \tag{16}$$

The gradients (14-16) can be computed through integration of the sensitivity equations

$$\frac{d\lambda_G}{dt} = \frac{\partial f_2}{\partial M_c^{\,hyb}}\lambda_B + \frac{\partial f_2}{\partial G} , \ \lambda_G(0) = 0 \tag{17}$$

$$\frac{d\lambda_B}{dt} = \frac{\partial f_2}{\partial M_c^{\,hyb}}\lambda_B + \frac{\partial f_2}{\partial \tilde{\mu}_0^{\,hyb}}\left(\frac{\partial \tilde{\mu}_0^{\,hyb}}{\partial B}\right) , \ \lambda_B(0) = 0 \tag{18}$$

$$\frac{d\lambda_\beta}{dt} = \frac{\partial f_2}{\partial M_c^{\,hyb}}\lambda_\beta + \frac{\partial f_2}{\partial \tilde{\mu}_0^{\,hyb}}\frac{\partial \tilde{\mu}_0^{\,hyb}}{\partial \beta} , \ \lambda_\beta(0) = 0 , \tag{19}$$

Note, that while λ_G can be straightforward obtained, λ_B and λ_β depend on the gradients of $\tilde{\mu}_0$ with respect to $\tilde{B}_0$ and β', respectively. In order to determine them the same strategy is applied leading to integration of the following sensitivity equations with zero initial conditions

$$\frac{d\chi_B}{dt} = \frac{\partial f_1}{\partial \tilde{\mu}_0}\chi_B + \frac{\partial f_1}{\partial B} , \quad \chi_B(0) = 0 \tag{20}$$

$$\frac{d\chi_\beta}{dt} = \frac{\partial f_1}{\partial \tilde{\mu}_0}\chi_\beta + \frac{\partial f_1}{\partial \beta} , \quad \chi_\beta(0) = 0 , \tag{21}$$

where $f_1 = B^{NN} - \frac{1}{2}\hat{\beta}^{NN}\left(\tilde{\mu}_0^{\,hyb}\right)^2$, $\chi_B = \frac{\partial \tilde{\mu}_0}{\partial B}$, and $\chi_\beta = \frac{\partial \tilde{\mu}_0}{\partial \beta}$.

The network parameters were tuned applying the Levenberg-Marquart optimization procedure [12].

IV. ANN-Based Model Predictive Control

A. Problem Formulation

Nonlinear model predictive control (NMPC) is an optimisation-based multivariable constrained control technique that uses a nonlinear dynamic model for the prediction of the process outputs. At each sampling time the model is updated on the basis of new measurements and state variables estimates. Then the open-loop optimal manipulated variable moves are computed over a finite (predefined) prediction horizon with respect to some performance index, and the manipulated variables for the subsequent prediction horizon are implemented. Then the prediction horizon is shifted or shrunk by usually one sampling time into the future, and the previous steps are repeated. The optimal control problem in the NMPC framework can be mathematically formulated as:

$$\min_{u_{\min} \le u(t) \le u_{\max}} J = \varphi(x(t), u(t), P), \tag{22}$$

subject to:

$$\dot{x} = f(x(t), u(t), P), \quad 0 \le t \le t_f , \quad x(0) = x_0 \tag{23.1}$$

$$y(t) = h(x(t), P) \tag{23.2}$$

$$g_j(x) = 0, \quad j = 1,2,.....p \tag{24.1}$$

$$v_j(x) \le 0, \quad j = 1,2,.....l \tag{24.2}$$

where (22) is the performance index, (23) is the process model, function *f* is the state-space description, function *h* is the relationship between the output and the state, *P* is the vector of possibly uncertain parameters and t_f is the final batch time. $x(t) \in R^n, u(t) \in R^m$ and $y(t) \in R^p$ are the state, the manipulated input and the control output vectors, respectively. The manipulated inputs, the state and the control outputs are subject to the following constraints,

$x(t) \in X, u(t) \in Z, y(t) \in Y$ in which X , Z and Y are convex and closed subsets of R^n, R^m and R^p. g_j and v_j are the equality and inequality constrains with p and l dimensions respectively.

B. Closed Loop ANN-MPC Structure

The particular closed loop MPC structure considered in this work is illustrated in Fig. 5.

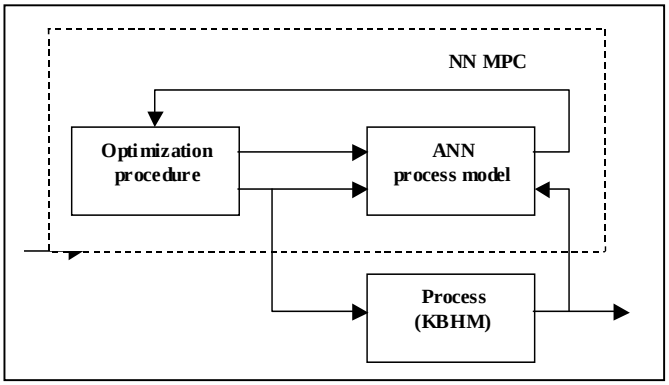


Fig. 5 ANN-based model predictive control (ANN-MPC)

For the simulation purposes the KBHM process model introduced in section II was implemented as the simulation model (see Fig. 3). The controller consists of the feed forward ANN process model discussed also in section II (see Fig. 2) and an optimization block. The ANN model predicts future process responses to potential control signals over the prediction horizon. The predictions are supplied to the optimization block to determine the values of the control action over a specified (control) horizon that minimize the following performance index

$$\min_{u_{\min} \le [u(t+k), u(t+k+1), \ldots u(t+c))] \le u_{\max}} P = \lambda_1 \sum_{k=1}^{p} \left(y_r(t+k) - y_m(t+k)\right)^2 - \lambda_2 \sum_{k=1}^{c} \left(u(t+k-1) - u(t+k-2)\right)^2 \quad (25)$$

The prediction horizon p is the number of time steps over which the prediction errors are minimized and the control horizon c is the number of time steps over which the control increments are minimized, y_r is the desired response and y_m is the network model response.

$u(t+k), u(t+k+1), \ldots . u(t+c)$ are tentative values of the future control signal, which are limited by $u_{\min}$ and $u_{\max}$ and parameterized as peace wise constant. λ_1 and λ_2 determine the contribution of the sum of the squares of the output error and the control increments over the performance index. The length of the prediction horizon is crucial for achieving tracking and stability. For small values of p the tracking deteriorates but for high p values the bang-bang behaviour of the process input might be a real problem. The MPC controller requires a significant amount of on-line computation, since the optimization (25) is performed at each sample time to compute the optimal control input. At each step only the first control action is implemented to the process (in this case to the simulation KBHM).

V. Numerical Implementation of ANN-MPC

The numerical implementation of ANN-MPC control is schematically presented in Fig. 6. The control problem is simulated in Matlab/Simulink framework as a set of modules. The Matlab NN Toolbox is also required. The controller is designed as an independent block and the process is simulated as a KBHM model. The KBHM model is coded as an S-function required by Simulink. The architecture of the ANN process model (which is part of the MPC structure) consists of one hidden layer with 7 sigmoid squashing activation functions and one output linear layer. The ANN is trained offline, in batch mode with input-output data generated by the KBHM. Before introducing to the ANN, data is normalized in the range (-1,1) and after processing over the network the network outputs are denormalized.

Simulation results are summarized in Fig. 7. The process manipulated input (the feed flow rate) is depicted in subplot (a) and the process controlled output (the supersaturation) is depicted in subplot (b) for prediction horizon p=10 and control horizon c=4. The peace-wise constant output reference was determined based on the optimal profile derived by an off-line dynamic optimization [15]. During the stages of concentration and feeding with liquor, the reference was set at S_{ref}=1.15 and afterwards was reduced to S_{ref}=1.05. A smooth transition between the two levels was determined to overcome possible overreaction of the tracking controller. The graphics shows satisfactory reference tracking with an acceptable smooth behaviour of the control input which stays within the technological constrains defined with u_{max}=0.015 [m3/s]. Higher the prediction horizon better would be the tracking but to the expense of more vivid manipulated input. For higher values of p (>10), the control action faced saturation problems.

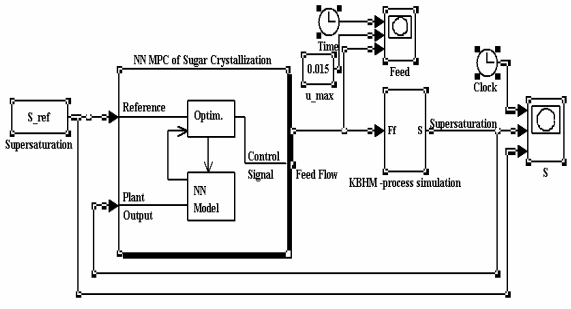
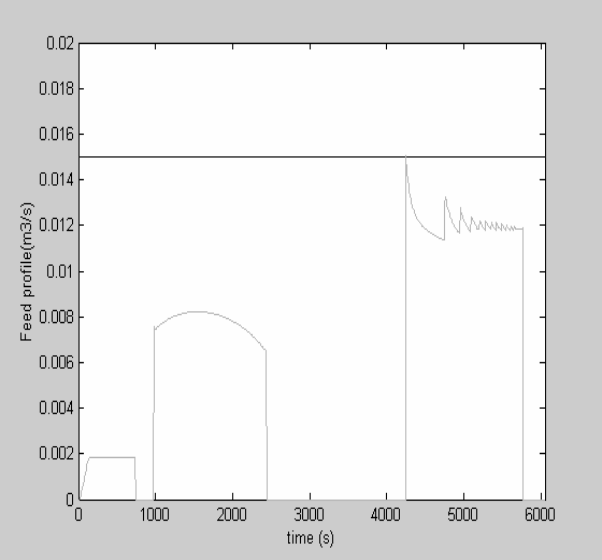
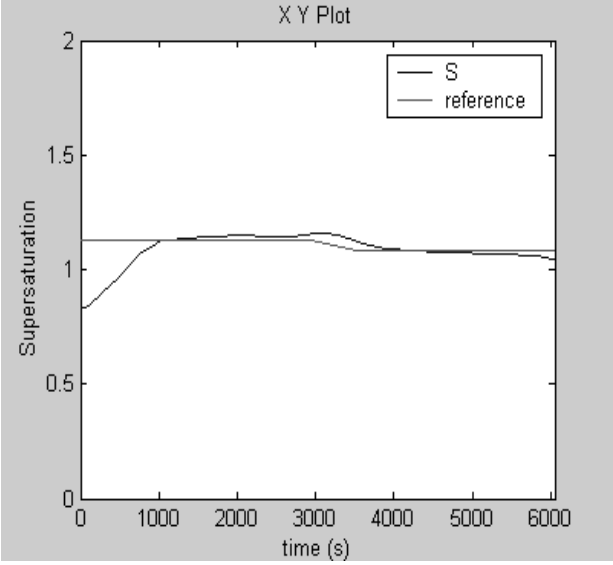


Fig. 6 ANN MPC – simulation scheme



a) Feed rate profile and its max allowed value



b) Supersaturation profile and its ref. trajectory

Fig. 7 ANN-MPC simulations p=10, c=4

VI. Conclusion

The application of FFNN at two stages of sugar crystallization process automation, namely modelling and control is presented in this paper.

At the modelling stage, a knowledge based hybrid model (KBHM) of the process was designed that possesses the advantages of both analytical and pure data based process models. The KBHM offers a reasonable compromise between the extensive efforts to get a fully parameterised structure, as are the analytical models and the poor generalisation of the complete data-based modelling approaches.

At the control stage, a model predictive control based on a FFNN model of the process was designed. Since the FFNN model captures the nonlinear nature of the process it has the potential advantage over the linear models widely used in the MPC framework. The proposed scheme guaranteed feasibility of the reference tracking in the presence of input constrains. It should, however, be pointed out that variations of initial conditions and disturbances during the batch are not treated in the paper but are most probably to appear in the reality. Future work will focus on these issues.

Acknowledgment

This work was pursued in the framework of EC RTN Project BatchPro HPRN-CT-2000-0039. It was further financed by the Portuguese Foundation for Science and Technology within the activity of the Research Unit IEETA-Aveiro, which is gratefully acknowledged.

References

[1] J. B. Rawlings, S. M. Miller, and W. R. Witkowski., "Model identification and control of solution crystallization processes: a review", Ind. Eng. Chem. Res., 32(7):1275-1296, July 1993.

[2] Nagy Z. K. and R. D. Braatz,. Robust nonlinear model predictive control of batch processes. AIChE J., 49:1776-1786, 2003.

[3] Fujiwara, M. Z. K. Nagy, J. W. Chew, and R. D. Braatz First-principles and direct design approaches for the control of pharmaceutical crystallization. J. of Process Control, 15:493-504, 2005 (invited).

[4] Feyo de Azevedo, S., Chorão, J., Gonçalves, M.J. and Bento, L., 1994, On-line Monitoring of White Sugar Crystallization through Software Sensors. Int. Sugar JNL., 96:18-26.

[5] S. Haykin S. (1994), Neural Networks: A comprehensive foundation, Prentice Hall, UK.

[6] Russell N.T. and H.H.C. Bakker (1997), Modural modelling of an evaporator for long-range prediction, Artificial Intelligence in Engineering, 11, 347-355.

[7] Zorzetto L.F.M., R. Maciel Filho and M.R. Wolf-Maciel (2000), Process modelling development through artificial neural networks and hybrid models, Computers and Chemical Engineering, 24, 1355-1360.

[8] Georgieva P., Meireles, M. J. and Feyo de Azevedo, S., 2003, KBHM of fed-batch sugar crystallization when accounting for nucleation, growth and agglomeration phenomena. Chem Eng Sci, 58:3699-3711.

[9] Ditl, P., Beranek L., & Rieger, Z. (1990). Simulation of a stirred sugar boiling pan. *Zuckerind*., 115, 667-676.

[10] A.D. Randolph, M.A. Larson, Theory of Particulate Processes – Analyses and Techniques of Continuous Crystallization, Academic Pres, 1988.

[11] P. Lauret, H. Boyer, J.C. Gatina, Hybrid modelling of a sugar boiling process, Control Engineering Practice, 8, 2000, 299-310.

[12] Hagan M. T., M. Menhaj (1994), "Training feedforward networks with the Marquardt algorithm," IEEE *Transactions on Neural Networks*, vol. 5, no.6, pp. 989–993.

[13] D.C. Psichogios, L. H. Ungar, A hybrid neural network - first principles approach to process modelling, AIChE J., 38(10), 1992, 1499-1511.

[14] Graefe, J., Bogaerts, P., Castillo, J., Cherlet, M., Werenne, J., Marenbach, P., & Hanus, R. (1999). A new training method for hybrid models of bioprocesses. *Bioprocess Engineering*, 21, 423-429.

[15] V. Galvanauskas, P. Geogieva, S. Feyo de Azevedo, Dynamic optimization of industrial sugar crystallization process based on a hybrid (mechanistic +ANN) model, submitted for publication to IJCNN, Vancouver, Canada, June 2006.