PERGAMON

Chemical Engineering Science 58 (2003) 3699–3713

Chemical Engineering Science

www.elsevier.com/locate/ces

Knowledge-based hybrid modelling of a batch crystallisation when accounting for nucleation, growth and agglomeration phenomena

P. Georgieva[a], M. J. Meireles[b], S. Feyo de Azevedo[a,*]

[a] *Department of Chemical Engineering, Faculty of Engineering, University of Porto, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal*
[b] *Department of Chemical Engineering, Polytechnic Institute of Porto, Av. António Bernardino de Almeida, 431, 4200-072 Porto, Portugal*

Received 10 September 2002; received in revised form 7 May 2003; accepted 14 May 2003

Abstract

This paper reports on the application of knowledge-based hybrid (KBH) modelling to an industrial scale (fed-) batch evaporative crystallisation process in cane sugar refining. First, principles models of the process lead in general to good description of process state, except for the prediction of the main crystal size distribution (CSD) parameters—mean size and the coefficient of variation. This is due to difficulties in expressing accurately nucleation and crystal growth rates and especially the complex phenomena of agglomeration in the relevant population balance. A hybrid model is proposed, which combines a partial mechanistic model that reflects the general mass, energy and population balances with a neural network to express growth rate, nucleation kinetics and agglomeration phenomena. Results obtained demonstrate a better agreement between experimental data and hybrid model predictions than that observed with the complete mechanistic model.

Keywords: Sugar crystallisation; Agglomeration; Hybrid mathematical modelling; Parameter identification; Neural network

1. Introduction

The traditional way of process modelling has been for many years by mathematical equations that essentially express fundamental principles and incorporate transport kinetics and equilibrium relationships. In general terms, this type of fully mechanistic models, when feasible, have the potential to extrapolate beyond the regions (experimental conditions) within which they are constructed. Many are, however, the areas, in chemical and biochemical engineering, where rigorous scientific knowledge of fundamental mechanisms is still scarce, their mathematical representation being for this reason of limited or questionable accuracy. Empirical correlations are typically employed to characterise basic transport or kinetic phenomena. Such approach has been over the years most useful, but it has recognised limitations particularly for expressing time varying states.

Independently of the undisputable relevance of pursuing studies that may lead to this far-reaching goal of a priori design, new methods (e.g. data-based modelling techniques) are being developed that are able to extract process knowledge from existing process measured data, in this way overcoming difficulties of knowledge expression.

The rapid growth of available computational resources led in recent years to the development of a wide number of data-based modelling methods. They range from well-established statistical techniques, such as multiple linear regression (MLR) and principal component regression (PCR), to non-linear techniques such as non-linear principal component analysis (NLPCA) and non-linear autoregressive moving average with exogenous input (NARMAX). Since about a decade a new group of data-driven modelling techniques has been under scrutiny, the most popular among them being the artificial neural network (ANN) models. ANNs proved to be a powerful modelling alternative to the conventional mathematical description when there is limited understanding of the physical phenomena.

Data-based modelling alone has a number of limitations, namely related to the confidence of employing models outside identification (the so-called training) and validation regions. With such identified difficulties, it seems today only natural that the next step was the emergence of methods where different forms of available knowledge for process

* Corresponding author. Fax: +351-22-508-1632/1449.
E-mail address: sfeyo@fe.up.pt (S. Feyo de Azevedo).

0009-2509/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/S0009-2509(03)00260-4

modelling are combined, in this way trying to maximise knowledge expression. As example, ANN can be employed for modelling complex and poorly known parts of the process otherwise modelled by application of conservation laws (Roubos et al., 1999; Russell and Bakker, 1997; Schubert, Simutis, Dors, Havlik, & Lubbert, 1994a, b; Zorzetto, Maciel Filho, Wolf-Maciel, 2000). This is generally termed as knowledge-based hybrid (KBH) modelling and is today well documented in the literature (Peres, Oliveira, & Feyo de Azevedo, 2001).

In this paper, we discuss KBH modelling of industrial scale (fed-) batch evaporative crystallisation in cane sugar refining.

Sugar crystallisation is a non-linear and non-stationary process. Crystallisation occurs through mechanisms of nucleation, growth and agglomeration. The relation between operating conditions, process state and those crystallisation mechanisms is poorly known. Consequently, the search for efficient methods for process description is linked both to the scientific interest of understanding fundamental mechanisms of the crystallisation process and to the relevant practical interest of process modelling.

The available first principles models lead in general to reasonable description of process states associated with mass and energy balances. Nevertheless, exact prediction of the main crystal size distribution (CSD) parameters—the mean (MA) and coefficient of variation (CV) of the mass-size distribution function is hampered by the lack of reliable formal presentation of the crystallisation mechanism (Feyo de Azevedo, Chorão, Gonçalves, & Bento, 1994; Chorão & Feyo de Azevedo, 1996). More precisely, this is due to difficulties in expressing accurately nucleation and crystal growth rates and especially the complex phenomena of agglomeration in the relevant population balances. This type of difficulty of knowledge expression has been adequately solved in other areas through adopting knowledge-based hybrid approaches, which combine mechanistic, data-driven and fuzzy sources of knowledge (Thompson & Kramer, 1994; Feyo de Azevedo, Dahm, & Oliveira, 1997; Simutis, Oliveira, Manikowski, Feyo de Azevedo, & Lubbert, 1997). A hybrid model of sugar crystallisation was reported recently by Lauret, Boyer, and Gatina (2000), but without accounting for nucleation and agglomeration phenomena.

For the industrial set-up considered in the present work there is experimental evidence that agglomeration occurs and neglecting it would lead to underestimation of CSD characteristics. Therefore, the main objective was to develop a reliable model of the process when accounting for nucleation, crystal growth and particle agglomeration phenomena.

Serial hybrid modelling has been adopted as a complementary structure of two modules: a mechanistic and a data-driven module.

The mechanistic part consists of a set of non-linear algebraic-differential equations, resulting from relevant mass, energy and population balances. The latter is expressed in particle volume co-ordinates, including the first four moments, from where CSD parameters are inferred.

The data-driven module supplies the growth rate, nucleation and agglomeration kinetic parameters. After consideration of alternative topologies, a feedforward ANN with four inputs (supersaturation, massecuite temperature, purity of solution and volume fraction of crystals) was adopted.

Based upon a data bank of process input and output measurements, the ANN was trained and validated. The global hybrid model was then evaluated for its ability to predict the process outputs (temperature of massecuite and brix of solution) and the CSD parameters (MA and CV). The results obtained demonstrate good interpolative and extrapolative capabilities of the hybrid model. This paves the way for computing more accurate optimal operation profiles and more accurate software sensors for on-line process monitoring.

2. Vacuum pan crystalliser

2.1. Process description

The industrial unit of concern is of the fixed calandria type with specially designed tubes for the circulation of heavy viscous massecuite (i.e. a suspension of sugar crystals in heavy syrup). It is equipped with a total condensing steam valve and with a propeller stirrer. The unit has a cylindrical form and characteristics as listed in Table 1. The average operating time of a batch cycle is approximately 90 min. It is divided into several sequential phases as listed in Table 2.

Table 1
Characteristics of the crystalliser

Type	Fixed calandria, total condensation
Diameter	3.92 m
Working volume	35 m^3
Heat transfer area	194 m^2
Tube length	0.92 m
Tube diameter	0.098 m
Stirrer speed	58.6 rpm
Stirrer power	15 kW

Table 2
Batch cycle

Phase	Time/size
Charging	$\Delta t \approx 10$ min, $V_0 \approx 12$ m^3
Concentration	$\Delta t \approx$ 10–15 min
Seeding	MA ≈ 12 μm
Setting the grain	$\Delta t \approx$ 2–5 min
Crystallisation	$\Delta t \approx 60$ min
Tightening	$\Delta t \approx 5$ min
Discharge	
Cleaning	

Supersaturation is the driving force for crystallisation. In industrial conditions, it can be expressed as the ratio of concentration of sucrose in solution to its saturation concentration for the same process conditions. Crystal growth only occurs when sucrose concentration exceeds the saturation state. All along the batch supersaturation must be kept within safe limits in order to avoid the formation of new crystals (the production of false grain) or the dissolution of the existing ones. This safe zone is called the metastable zone ($1 < S < 1.2$).

The process operation objective is to maximize yield of sugar with high quality, the latter being essentially measured by the purity, shape and particle size distribution of the crystal population.

At the end of a normal industrial batch the expected main characteristics of the final product, i.e. the mass fraction of crystals, the mass averaged crystal size (MA) and the coefficient of variation (CV), should be within the following desired ranges: (i) mass percent of crystals between 56 and 58%; (ii) MA of 0.5–0.6 mm; and (iii) CV of 28–32%.

2.2. Pan controlled operation

Vacuum pan control involves the manipulation of feed flow rate of sugar liquor (a juice containing dissolved sucrose) and of vacuum pressure. Typical operation data are summarised in Table 3.

From the control point of view three main phases can be identified, viz. during the first phase (about 20 min) the pan is partially filled (about 30%) with liquor. The liquor is concentrated by evaporation, under vacuum, until supersaturation reaches a predefined value (typically $S = 1.15$). At this point seed crystals are introduced into the pan to induce the production of crystals.

That is the beginning of the second (crystallisation) phase, which lasts for about 60 min. 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. Near to the end of this phase and for economical reasons, the liquor is replaced by other juice of lower purity (termed syrup).

The third phase consists of tightening, which is principally controlled by the evaporation capacity. At the end of the batch, the final massecuite undergoes centrifugation, where final refined sugar is separated from liquor (mother liquor) that is recycled to the process.

Table 3
Typical operation data

Average batch time	90 min
Final volume	32–33 m^3
Vacuum pressure	0.24–0.26 bar
Brix of feed syrup	76–80°Bx
Brix of feed liquor	71–73°Bx
Purity of feed syrup	85–95%
Purity of feed liquor	99–99.9%
Temperature of feed	65–75°C
Steam flow	7–9 ton/h
Steam pressure	1.9–2.2 bar
Water added	1–3 m^3
Mass of seed	0.16 kg
MA of seed	10–12 μm
CV of seed	80%

2.3. Available measurements

The on-line collected physical measurements related to a batch are summarised in Table 4 (sampled input measurements) and Table 5 (sampled output measurements). Additional constant operation and control variables are listed in Table 6. Data from 15 (industrial runs) batches are available, sampled during a period of 4 months. The measurements are related to:

(i) *Feed conditions*: flowrate, brix (measured by refractometry), temperature and purity of the fed liquor/syrup and the flowrate and temperature of the water;

Table 4
Measured input variables

Notation	Process variable
F_f	Liquor/syrup feed flowrate
B_f	Feed brix
T_f	Temperature of feed
F_w	Water feed flowrate
F_s	Steam flowrate
T_s	Steam temperature
P_s	Steam pressure
I_{agit}	Intensity of stirrer current
P_{vac}	Vacuum pressure
T_{vac}	Vacuum temperature

Table 5
Measured output variables

Notation	Process variable
B_{sol}	Brix of solution
T_m	Temperature of solution
L	Level in the pan

Table 6
Constant operational and control variables

Notation	Process variable
T_w	Temperature of feed water
Pur_f	Purity of liquor/syrup
t_{syr}	Initial time for introducing syrup
t_{seed}	Initial time of seeding

(ii) *Steam conditions*: flowrate, temperature and pressure of the steam;
(iii) *Variables inside the pan*: pressure and temperature of the vacuum, the brix and the temperature of the solution, the level in the pan.

Supersaturation and mass fraction of crystals are not measured variables, but they are readily determined on the basis of available measurements. The main challenge in process monitoring is the estimation of CSD properties.

3. Partial mechanistic model

The modelling of the crystalliser is conceptually obtained by writing the appropriate mass and energy balances together with a mathematical representation of the crystallisation rate. The latter can be achieved through basic mass transfer considerations (Ditl, Beranek, & Rieger, 1990) or by writing a population balance represented by its moment equations (Hulburt & Katz, 1964; Jones, 1974; Randolph & Larson, 1988; Feyo de Azevedo, Chorão, Gonçalves, & Bento, 1993; Feyo de Azevedo et al., 1994). 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.

Considering some form of particle birth and death by agglomeration or aggregation mechanisms is required though a price is paid on model complexity, which may preclude its use in model-based state estimation and control strategies. A number of publications propose the theory and alternative simplifying mathematical procedures (Hulburt & Katz, 1964; Randolph & Larson, 1988; Hartel & Randolph, 1986; Franck, David, Villermaux, & Klein, 1988; Hounslow, Ryall, & Marshall, 1988; Hounslow, 1990). None of these are applied to sugar crystals but the bases for application are by all means equivalent.

3.1. Model adopted

Process modelling involves some simplifying assumptions. In developing the model equations it was accepted that (i) the massecuite is well mixed in the whole pan volume; (ii) its temperature is uniform, given by an averaged value between two temperature measurements; (iii) thermal dynamic delays due to the unit thermal capacity are negligible; (iv) the shape and volume crystal coefficients do not change during crystallisation and (v) all steam supplied is condensed.

The state vector includes temperature, mass of water, dissolved sucrose and dissolved impurities in the liquor, mass of sucrose crystals and the population distribution moments.

Main intensive variables such as supersaturation and crystal contents are calculated from the set of state variables. An empirical model has been developed for the description of the evaporation rate.

3.2. Mass balances

The following set of conservation mass balance equations can be written for water (M_w), impurities (M_i), dissolved sucrose (M_s) and crystals (M_c):

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

$$\frac{dM_i}{dt} = F_f\rho_f B_f(1 - Pur_f), \tag{1b}$$

$$\frac{dM_s}{dt} = F_f\rho_f B_f Pur_f - J_{cris}, \tag{1c}$$

$$\frac{dM_c}{dt} = J_{cris}. \tag{1d}$$

The main process non-linearities are included in the crystallisation rate J_{cris}, which has to be computed from the population balance. The heat input and the evaporation rate are given by the following correlations:

$$Q = \alpha_s F_s \Delta H_s, \tag{2}$$

$$J_{vap} = \frac{W + Q}{\lambda_{w(vac)}} + k_{vap}(T_m - T_{w(vac)} - BPE), \tag{3}$$

where α_s and k_{vap} are equipment-dependent parameters determined experimentally with data taken from the concentration period at the beginning of the crystallisation cycle before seeding takes place. BPE is the boiling point elevation and the term $(T_m - T_{w(vac)} - BPE)$ represents the superheat in the massecuite. The procedure for obtaining empirical correlations (2) and (3) is presented elsewhere (Feyo de Azevedo et al., 1993), nevertheless it is briefly described again in Appendix B.

3.3. Energy balance

The energy balance is described by Eq. (4), where T_m represents the temperature of massecuite. The enthalpy terms and specific heat capacities included are derived as functions of physical and thermodynamical properties, as given in Appendix A.

$$\frac{dT_m}{dt} = aJ_{cris} + bF_f + cJ_{vap} + d, \tag{4}$$

$$a = \frac{1}{M_{sol}C_{Psol} + M_cC_{Pc}} \times \left[H_{sol} - H_c + (1 - B_{sol})\frac{dH_{sol}}{dB_{sol}} + \frac{1 - Pur_{sol}}{B_{sol}}\frac{dH_{sol}}{dPur_{sol}}\right], \tag{4a}$$

$$b = \frac{\rho_f}{M_{\mathrm{sol}} C_{P\mathrm{sol}} + M_c C_{Pc}} \times \left[H_f - H_{\mathrm{sol}} + (B_f - B_{\mathrm{sol}}) \frac{\mathrm{d}H_{\mathrm{sol}}}{\mathrm{d}B_{\mathrm{sol}}} + \frac{B_f}{B_{\mathrm{sol}}} (\mathrm{Pur}_f - \mathrm{Pur}_{\mathrm{sol}}) \frac{\mathrm{d}H_{\mathrm{sol}}}{\mathrm{dPur}_{\mathrm{sol}}} \right], \quad (4b)$$

$$c = \frac{1}{M_{\mathrm{sol}} C_{P\mathrm{sol}} + M_c C_{Pc}} \left[H_{\mathrm{sol}} - H_{\mathrm{vap}} - B_{\mathrm{sol}} \frac{\mathrm{d}H_{\mathrm{sol}}}{\mathrm{d}B_{\mathrm{sol}}} \right], \quad (4c)$$

$$d = \frac{1}{M_{\mathrm{sol}} C_{P\mathrm{sol}} + M_c C_{Pc}} \times \left[W + Q + F_w \rho_w (H_w - H_{\mathrm{sol}} + B_{\mathrm{sol}}) \frac{\mathrm{d}H_{\mathrm{sol}}}{\mathrm{d}B_{\mathrm{sol}}} \right]. \quad (4d)$$

3.4. Population balance (in volume co-ordinates)

The population balance represented by Eq. (5) below and the subsequent transformations leading to moment equations (10) are based on the theory published by Hulburt and Katz (1964) and Randolph and Larson (1988). This type of approach was also adopted by Tavare and Garside (1993) for a perfectly mixed crystalliser.

Hulburt and Katz (1964) assume that particle birth is only due to collisional agglomeration and that both birth and death rates are function of the mass of colliding particles. The direct consequence of this concept is that population balances to include these agglomeration mechanisms should be written in terms of crystal volume density functions (i.e. related to mass) rather than linear size density functions.

The population balance in volume co-ordinates is given by the following partial differential equation:

$$\frac{\partial \tilde{n}(v)}{\partial t} + \frac{\partial G_V \tilde{n}(v)}{\partial v} = \tilde{B}(v) - \tilde{D}(v), \quad (5)$$

where $\tilde{n}(v)$ is the number-volume distribution function and G_V is the overall volume growth rate. As the total working volume of the fed-batch crystalliser is time varying, the population density and other relevant quantities will be defined on the basis of the total vessel volume at any time and denoted by $\sim$ over the corresponding symbol. The agglomeration mechanism is represented by empirical birth ($\tilde{B}(v)$) and dead functions ($\tilde{D}(v)$), respectively. The aggregation of two particles of volume u and $v-u$ into a new particle of volume v is mathematically formulated by the rate of particle appearance:

$$\tilde{B}(v) = \tfrac{1}{2} \beta' \int_0^v \tilde{n}(u) \tilde{n}(v-u)\, \mathrm{d}u \quad (6)$$

and the rate of particle disappearance of volume v:

$$\tilde{D}(v) = \beta' \tilde{n}(v) \int_0^\infty \tilde{n}(v-u)\, \mathrm{d}u. \quad (7)$$

The agglomeration kernel defined as

$$\beta = \beta' V_m \quad (8)$$

is a measure of the frequency of collisions between particles of volumes u and $v-u$ that are successful in producing a particle of volume v. It is assumed to be independent of the crystal volume, at least over a small time interval for parameter identification of the dynamic process. In case the agglomeration is not considered, the right-hand side of Eq. (5) is equal to zero. Defining the jth moment of a particle number-volume distribution function by

$$\mu_j(v) = \int_0^\infty v^j n(v)\, \mathrm{d}v \quad (9)$$

and following the procedure developed by Hulburt and Katz (1964), the population balance Equation (5) is conveniently replaced for modelling purposes by the set of ordinary differential equations representing distribution moments in volume co-ordinates, as follows:

$$\frac{\mathrm{d}\tilde{\mu}_j}{\mathrm{d}t} = G_v \left[j\tilde{\mu}_{j-1} - \tilde{n}(v) v^j \Big|_0^\infty \right] + \beta' \left[\frac{1}{2} \left(\sum_{k=0}^{j} \binom{j}{k} \tilde{\mu}_k \tilde{\mu}_{j-k} \right) - \tilde{\mu}_0 \tilde{\mu}_j \right], \quad j = 0, 1, 2, \ldots \quad (10)$$

with initial and boundary conditions $\tilde{n}(0, v) = 0$, $\tilde{n}(t, \infty) = 0$, $\tilde{n}(t, 0) = \tilde{B}_0 / G_v$.

$\tilde{B}_0$ is the nucleation rate at near zero crystal volume and

$$\binom{j}{k} = \frac{j!}{k!(j-k)!}$$

is the binomial coefficient.

For population balance modelling of the process in hand, the first four moments from Eq. (10) are derived as

$$\frac{\mathrm{d}\tilde{\mu}_0}{\mathrm{d}t} = \tilde{B}_0 - \tfrac{1}{2} \beta' \tilde{\mu}_0^2, \quad (11a)$$

$$\frac{\mathrm{d}\tilde{\mu}_1}{\mathrm{d}t} = G_v \tilde{\mu}_0, \quad (11b)$$

$$\frac{\mathrm{d}\tilde{\mu}_2}{\mathrm{d}t} = 2 G_v \tilde{\mu}_1 + \beta' \tilde{\mu}_1^2, \quad (11c)$$

$$\frac{\mathrm{d}\tilde{\mu}_3}{\mathrm{d}t} = 3 G_v \tilde{\mu}_2 + 3 \beta' \tilde{\mu}_1 \tilde{\mu}_2, \quad (11d)$$

with initial conditions $\tilde{\mu}_j(0) = 0$, $j = 0, 1, 2, 3$.

The zeroth and the first moment represent, respectively, total number and total volume of particles. Finally, the crystallisation rate is related to the derivative of the first moment as

$$J_{\mathrm{cris}} = \rho_c \frac{\mathrm{d}\tilde{\mu}_1}{\mathrm{d}t}. \quad (12)$$

Note that no definition is provided for nucleation, growth rate and aggregation, hence classifying the model written as a partial mechanistic model. Also, later in the text, in Section 6.4, the reason will be given of why moments of orders 2 and 3 are included in the model.

4. KBH model

The kinetic parameters that reflect the crystallisation mechanism (Eqs. (11)) are the nucleation rate ($\tilde{B}_0$), the agglomeration kernel (β') and the linear growth rate (G), from which the overall volume growth rate is determined

$$G_v = 3k_v^{1/3}\left(\frac{v}{\tilde{\mu}_0}\right)^{2/3} G. \tag{13}$$

In the case of sugar crystallisation, the process variables that affect the kinetic parameters are rather well known. These variables are mainly temperature, supersaturation, purity of the solution and volume fraction of crystals, which is the ratio of the total volume in the pan and the volume of the crystals. Modelling of such phenomena is far from established and indeed a variety of empirical models of fuzzy accuracy are available and listed in the literature (Chorão & Feyo de Azevedo, 1996; Meireles & Feyo de Azevedo, 1998; Lauret, Boyer, & Gatina, 2000).

Hence, instead of searching for an analytical expression of the kinetic parameters what is here proposed is to follow a KBH strategy that consists of complementing the prior physical knowledge formulated by the mass, energy and population balances with a neural network to model kinetic parameters, i.e. with the latter approach extracting knowledge hidden in the experimental data available (in case it is sufficient and reliable) for building the kinetic information required by the mechanistic model.

To summarise, the KBH model of the sugar crystallisation process (Fig. 1a) consists of the set of algebraic–differential equations (1)–(4), (11)–(13), i.e. the mechanistic part, with an ANN representing the kinetic parameters (G, $\tilde{B}_0$, β') in the framework of the overall process model, coupled with the correlations for the physical properties given in Appendix A.

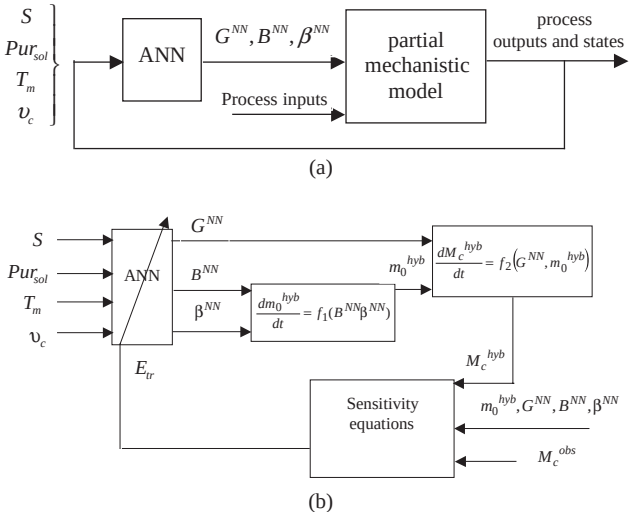

Fig. 1. (a) KBH model structure. (b) Hybrid training procedure.

5. Hybrid ANN training

5.1. Model structure and database construction

The overall KBH model structure and related training scheme are represented in Fig. 1. The first decision to be made is on the network inputs to be chosen. It is known that the kinetic parameters are affected by the temperature (T_m), supersaturation (S), purity of the solution ($\mathrm{Pur_{sol}}$) and volume fraction of crystals (υ_c). They all are considered as networks inputs. Except for the temperature, which is directly measurable, the other network inputs are function of mass balance states that are not directly measured (see the auxiliary correlations given in Appendix A). However, they can be reliably estimated based on the available measurements and simple computational algorithms (the so-called software sensors, see Feyo de Azevedo et al., 1993, 1994 for more details). Thus, supersaturation, purity of the solution and volume fraction of crystals are inferred from the estimated mass balance states and together with the physical measurements are saved in a large database for each of the available batches.

On-line measurements with sampling time of 10 s are performed. Though the batches have approximately the same duration, they are not completely equal in time and consequently the data length per batch is variable. The network training is performed with a sequence of data points corresponding to the crystallisation and tightening phases (i.e. between 380 and 480 data points per batch) of eight batches. The other seven batches are used for testing and validation purposes.

5.2. Hybrid training procedure—sensitivity approach

A typical problem arising in hybrid network modelling is that the usual training procedures for ANN are not applicable. The training of a neural network requires that the networks weights are determined in such a way that the sum of the squared error between the network output and the corresponding target output (available data) becomes minimal. In the hybrid system, however, the target outputs are not available since the kinetic parameters are not measured. Hence, a new training procedure must be developed. There are two approaches.

The first one is to try to obtain some estimation of the network target output simply by applying interpolation techniques to the experimentally obtained data to get required time derivatives (Chen, Bernard, Bastin, & Angelov, 2000; Graefe et al., 1999). Then the network part is directly trained with its input and output data in the conventional way. The procedure is quite simple but usually suffers from low accuracy, particularly when the estimated output has a vivid time trajectory (as it is the case), or if the sampled data are noisy, asynchronous and with large time increments.

A promising alternative is the sensitivity approach proposed by Psichogios and Ungar (1992) where the network

output is propagated through the analytical part of the model. The predicted hybrid model output is compared with the available data. The error signal for updating the network parameters is a function of the observed error and the gradient of the hybrid model output with respect to the ANN output. The intuition behind the scheme is that the network parameters are changed proportionally to their effect on the prediction of the process output. This approach is applied for training of the present model.

The hybrid network training can be performed in different ways depending on first principles knowledge involved. Taking into account that temperature measurements are available, the most straightforward way is to use only the energy balance equation to generate the error signal for network training. However, the sensitivity equations that have to be integrated are rather complex and difficult to formulate (see Eqs. (4)). Moreover, even if the gradients of the temperature with respect to the network outputs can be reliably computed, the training time would be considerably long. To use the whole set of mass balance equations for network training is another alternative but obviously the same problems arise, even stronger, due to the larger dimension of the analytical part involved. In fact, the aim is to get a reliable ANN model of the poorly known kinetic parameters while keeping the complexity of the network training as simple as possible. Therefore, a more practical way is to choose carefully only part of the mechanistic model, which is most directly influenced by the kinetic parameters. The mass of crystals is considered as most appropriate to serve as a target output in the hybrid network training. According to Eqs. (1d), (11b), (12) and (13), the mass balance of crystals can be rewritten as

$$\frac{\mathrm{d}M_c^{\mathrm{hyb}}}{\mathrm{d}t} = 3(k_v\rho_c)^{1/3}(\tilde{\mu}_0^{\mathrm{hyb}})^{1/3}(M_c^{\mathrm{hyb}})^{2/3}G^{NN}. \tag{14}$$

Eq. (14) is incorporated in the hybrid training structure but in order to integrate it the zero leading moment is also required and therefore its balance equation is also involved in the network training stage,

$$\frac{\mathrm{d}\tilde{\mu}_0^{\mathrm{hyb}}}{\mathrm{d}t} = B^{NN} - \tfrac{1}{2}\,\beta^{NN}(\tilde{\mu}_0^{\mathrm{hyb}})^2. \tag{15}$$

Superscripts hyb and *NN* are used to point out variables obtained during the hybrid network training. The training procedure is shown in Fig. 1b.

The ANN is provided with normalised (between 0.1 and 0.9) data inputs, and the outputs of the network give estimates of the growth rate, nucleation and agglomeration kinetic parameters. After a denormalisation, these estimates are propagated through the analytical expression equations (14) and (15).

The training database consists of 3520 points (corresponding to the crystallisation and tightening phases of eight process runs) for each of the inputs and for the target. Every iteration of the network training procedure includes the integration along time of balance equations (14) and (15) for each of the eight training runs, with the required resetting of initial conditions for each run.

Neural network training is carried out in a batch mode, where the network parameters are updated only after complete processing of the whole training data set and related computing of the performance function. For each time point j of the training set, the predicted mass of crystals is compared with the target data to obtain the observed error

$$(\xi_{\mathrm{obs}})_j = (M_c^{\mathrm{obs}})_j - (M_c^{\mathrm{hyb}})_j, \quad j = 1,\ldots,n, \tag{16}$$

where n is the length of the training data. The corresponding error signal vector $(E_{tr})_j$ is obtained by multiplying the observed error by the gradient of the hybrid model output with respect to the network outputs

$$(E_{tr})_j = (\xi_{\mathrm{obs}})_j \left[(\lambda_G)_j \;\; (\lambda_B)_j \;\; (\lambda_\beta)_j\right]^{\mathrm{T}}, \quad j = 1,\ldots,n, \tag{17}$$

where

$$(\lambda_G)_j = \left(\frac{\partial M_c^{\mathrm{hyb}}}{\partial G}\right)_j, \quad (\lambda_B)_j = \left(\frac{\partial M_c^{\mathrm{hyb}}}{\partial B}\right)_j,$$

$$(\lambda_\beta)_j = \left(\frac{\partial M_c^{\mathrm{hyb}}}{\partial \beta}\right)_j. \tag{18}$$

A global matrix error signal $E_{tr} = \{(e_{tr})_{ij}\}$ is defined, where each scalar element $(e_{tr})_{ij}$ corresponds to the error associated with the i network output and the j point of the training data sequence. Hence, the performance function to be minimised at each pass of training is the classical sum of squared errors

$$J = \frac{1}{3n}\sum_{i=1}^{3}\sum_{j=1}^{n}[(e_{tr})_{ij}]^2. \tag{19}$$

A modified version of the Levenberg–Marquardt optimisation algorithm is used for training and the network parameters are updated according to the scheme

$$w_{k+1} = w_k - [Q^{\mathrm{T}}Q + \xi I]^{-1}QE_{tr}, \quad Q = \frac{\partial E_{tr}}{\partial w} \tag{20}$$

with w_k being the vector of network parameters at the kth training iteration.

At each epoch, the adaptive scalar value ξ is decreased if the performance function is reduced or it is increased when a tentative step increases the performance function (see Demuth & Beale, 1997 for more details).

5.3. Sensitivity equations

The gradients, Equations (18), can be computed through integration of the sensitivity equations with zero initial conditions:

$$\frac{\mathrm{d}\lambda_G}{\mathrm{d}t} = \frac{\partial f_2}{\partial M_c^{\mathrm{hyb}}}\lambda_B + \frac{\partial f_2}{\partial G}, \quad \lambda_G(0) = 0, \tag{21a}$$

$$\frac{d\lambda_B}{dt} = \frac{\partial f_2}{\partial M_c^{hyb}}\lambda_B + \frac{\partial f_2}{\partial \tilde{\mu}_0^{hyb}}\left(\frac{\partial \tilde{\mu}_0^{hyb}}{\partial B}\right), \quad \lambda_B(0) = 0, \quad (21b)$$

$$\frac{d\lambda_\beta}{dt} = \frac{\partial f_2}{\partial M_c^{hyb}}\lambda_\beta + \frac{\partial f_2}{\partial \tilde{\mu}_0^{hyb}}\frac{\partial \tilde{\mu}_0^{hyb}}{\partial \beta}, \quad \lambda_\beta(0) = 0, \quad (21c)$$

where f_2 is the right-hand side of Eq. (14),

$$f_2 = 3(k_v\rho_c)^{1/3}(\tilde{\mu}_0^{hyb})^{1/3}(M_c^{hyb})^{2/3}G^{NN}. \quad (21d)$$

Note that while λ_G can be obtained straightforward, λ_B and λ_β depend on the gradients of the zero moment with respect to B and β, respectively. In order to determine them the same strategy is used leading to integration of the following sensitivity equations with zero initial conditions

$$\frac{d\chi_B}{dt} = \frac{\partial f_1}{\partial \tilde{\mu}_0}\chi_B + \frac{\partial f_1}{\partial B}, \quad \chi_B(0) = 0, \quad (22a)$$

$$\frac{d\chi_\beta}{dt} = \frac{\partial f_1}{\partial \tilde{\mu}_0}\chi_\beta + \frac{\partial f_1}{\partial \beta}, \quad \chi_\beta(0) = 0, \quad (22b)$$

where f_1 is the right-hand side of Eq. (13),

$$f_1 = B^{NN} - \tfrac{1}{2}\hat{\beta}^{NN}(\tilde{\mu}_0^{hyb})^2, \quad \chi_B = \frac{\partial \tilde{\mu}_0}{\partial B}, \quad \chi_\beta = \frac{\partial \tilde{\mu}_0}{\partial \beta}. \quad (22c)$$

To summarise, in order to get the gradients, Eqs. (18), and compute the training error signals, Eqs. (17), the following system of sensitivity differential equations has to be solved:

$$\frac{d\chi_B}{dt} = 1 - \beta^{NN}\tilde{\mu}_0^{hyb}\chi_B, \quad \chi_B(0) = 0, \quad (23a)$$

$$\frac{d\chi_\beta}{dt} = -\tilde{\mu}_0^{hyb}(\beta^{NN}\chi_\beta + 0.5\tilde{\mu}_0^{hyb}), \quad \chi_\beta(0) = 0, \quad (23b)$$

$$\frac{d\lambda_G}{dt} = (k_v\rho_c\tilde{\mu}_0^{hyb})^{1/3}\left[2(M_c^{hyb})^{-1/3}G^{NN}\lambda_G + 3(M_c^{hyb})^{2/3}\right], \quad \lambda_G(0) = 0, \quad (23c)$$

$$\frac{d\lambda_B}{dt} = (k_v\rho_c)^{1/3}G^{NN}\left(\frac{M_c^{hyb}}{\tilde{\mu}_0^{hyb}}\right)^{2/3}\left[2\left(\frac{M_c^{hyb}}{\tilde{\mu}_0^{hyb}}\right)^{-1}\lambda_B + \chi_B\right], \quad \lambda_B(0) = 0, \quad (23d)$$

$$\frac{d\lambda_\beta}{dt} = (k_v\rho_c)^{1/3}G^{NN}\left(\frac{M_c^{hyb}}{\tilde{\mu}_0^{hyb}}\right)^{2/3}\left[2\left(\frac{M_c^{hyb}}{\tilde{\mu}_0^{hyb}}\right)^{-1}\lambda_\beta + \chi_\beta\right], \quad \lambda_\beta(0) = 0. \quad (23e)$$

6. Hybrid model performance and comparison with a complete mechanistic model

In this section, the hybrid model is qualitatively compared with a complete mechanistic model of sugar crystallisation that was extensively studied and reported in previous works (Meireles & Feyo de Azevedo, 1998; Najim, Ruiz, Feyo de Azevedo, & Gonçalves, 1996). The performance of the two models is examined with respect to prediction quality of process outputs (temperature of massecuite and brix of solution) for which industrial data are available. Next, the final CSD computed by the two models is compared with the experimental crystal size distribution analysis at the end of the process run.

6.1. Empirical model for the kinetic parameters

The mechanistic model consists of the same mass, energy and population balances, Eqs. (1)–(4), (11)–(13), with kinetic parameters given by the following empirical correlations:

$$G = K_g \exp\left[-\frac{57000}{R(T_m + 273)}\right](S-1) \times \exp[-13.863(1 - P_{sol})]\left(1 + 2\frac{v}{V_m}\right), \quad (24)$$

$$\tilde{B}_0 = K_n \times 2.894 \times 10^{12} G^{0.51}\left(\frac{\tilde{\mu}_1}{k_v V_m}\right)^{0.53} V_m, \quad (25)$$

$$\beta' = K_{ag} G\left(\frac{\tilde{\mu}_1}{V_m^2}\right). \quad (26)$$

K_g, K_n and K_{ag} are kinetic constants, optimised on the basis of data from eight industrial runs following a classical non-linear least-squares regression. Their final values are given in Table 7. It has been proved that this model exhibits a superior performance when compared with a fully mechanistic model including nucleation and growth with dispersion but no agglomeration effects. Values of constant parameters required in both models are given in Table 8.

6.2. Results of the training phase

Time trajectories of the process outputs T_m and Bx_{sol} are depicted on Fig. 2. They serve as direct indicators of the model quality since industrial data for them are present. Data are denoted by bold dash-dotted lines, the dashed lines correspond to the predicted values by the mechanistic model and by solid lines are depicted the same outputs computed by the hybrid model. The predictions of both models are similar and closely match the observed data. This is not surprising since data from the same batches was used for hybrid network training and for tuning the optimisation parameters in the empirical kinetic correlations (Eqs. (24)–(26)).

Table 7
Optimised kinetic constants

K_g	K_n	K_{ag}
265	0.721	1.36×10^{-5}

Table 8
Values of constant model parameters

α_s	k_{vap} (kg/s°C)	W (J/s) (stirring power)	ρ_c (kg/m^3) (crystal density)	k_v (volume shape factor)	R (J/molK) (gas constant)
1.02	0.03	15000	1580	0.65	8.314

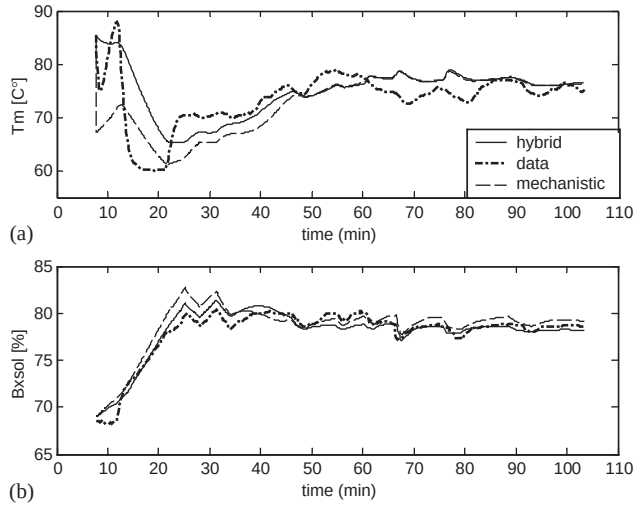

Fig. 2. Batch 1 (one of the batches used for hybrid network training and empirical kinetic model optimisation): process output predictions by the hybrid model (solid line) and the mechanistic model (dashed line) compared with industrial process data (bold dash-dot line). Phases of operation in min (start–end): concentration (7–22); crystallisation (22–95); tightening (95–103).

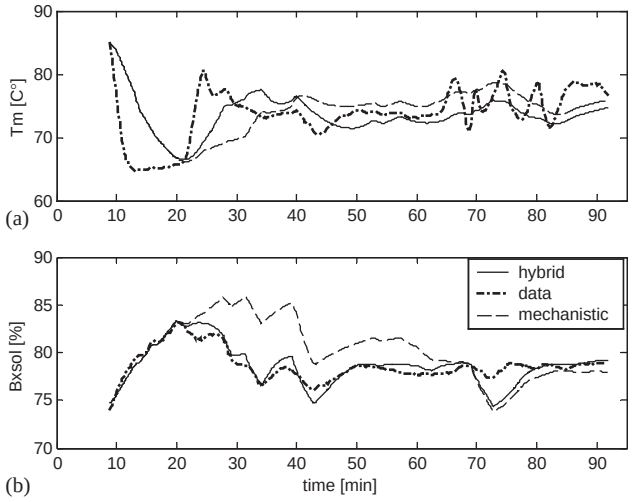

Fig. 3. Batch 2 (one of the batches used for hybrid network training): process output predictions by the hybrid model (solid line) and the mechanistic model (dashed line) compared with industrial process data (bold dash-dot line). Phases of operation in min (start–end): concentration (8–20); crystallisation (20–85); tightening (85–92).

The actual network in the framework of the hybrid model has one hidden layer with 10 activation functions of *tansig* type and the output layer has three activation functions of *logsig* type. This particular choice was achieved by starting with a large dimension hidden layer (25 nodes in this case) and reducing the number of the nodes gradually, down to the point where the performance function degradation is kept below 5%.

Fig. 3 shows results of a new batch that is also used for hybrid network training, but now the optimisation constants of the mechanistic model are the same as those obtained from the previous batch (Batch 1). The hybrid model gives better predictions of Bx_{sol} particularly in the rapid crystal growth phase (the period 20–70 min). For this period, the kinetic parameters exhibit quite vivid non-linear dynamics and the ANN model (with 38 tuning parameters) succeeds in capturing such dynamics better than the empirical kinetic model (with three tuning parameters). In contrast to brix, the predictions of the two models with respect to T_m are comparatively the same. The physical interpretation is that while the kinetic parameters are strongly influenced by temperature variations, the temperature itself is not substantially affected by the kinetic parameters. The same results are also observed at the validation stage. During the first phase of the batch cycle kinetic parameters can be taken as equal to zero since no crystals are introduced. For this reason, the hybrid network is trained only with data corresponding to the second and third phases and the plots of the figures start not at the beginning of the batch cycle but at the moment of concentration, which is different for each batch.

6.3. Model validation—analysis of process outputs along operation

More important are results related to the generalisation properties of the two models. Once the training was over the network weights were frozen and tested on new (validation) batches. The kinetic constants of the mechanistic model were also kept as they were estimated by the previous batch (Batch 1). As seen from Figs. 4 and 5, the hybrid model clearly exhibits superior performance with respect to predictions of brix. Accurate prediction of this property is directly related to reliable computation of supersaturation (see Appendix A), which is the driving force for crystallisation. As for the temperature, both models present good extrapolation properties apart from the initial transient period (10–15 min) before the temperature settles down.

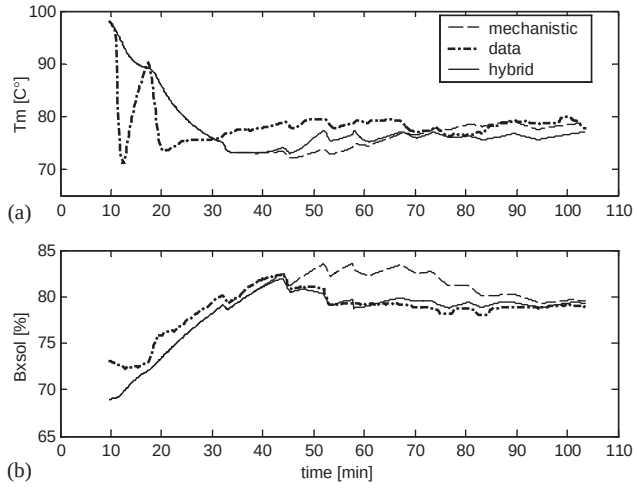

Fig. 4. Batch 3 (validation stage): process output predictions by the hybrid model (solid line) and the mechanistic model (dashed line) compared with industrial process data (bold dash-dot line). Phases of operation in min (start–end): concentration (10–26); crystallisation (26–90); tightening (90–104).

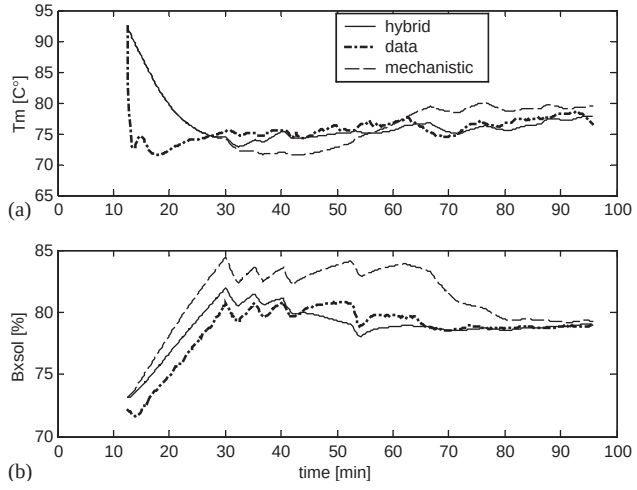

Fig. 5. Batch 4 (validation stage): process output predictions by the hybrid model (solid line) and the mechanistic model (dashed line) compared with industrial process data (bold dash-dot line). Phases of operation in min (start–end): concentration (12–28); crystallisation (28–87); tightening (87–95).

6.4. Model validation—CSD predicted by the KBH model

The quality of crystal size distribution is expressed at industrial level in terms of the norms approved by (International Commission for Uniform Methods of Sugar Analysis) (ICUMSA), i.e. by the MA and CV of the mass-size distribution function $m(L)$.

There are no techniques for (on-line) direct measurement of such parameters and generally data are limited to measurements made at the end of each batch by laboratory (sieve) analysis.

Mathematically, the moments of the mass-size distribution are defined as

$$\eta_j(L) = \int_0^\infty L^j m(L)\,\mathrm{d}L, \quad j = 0, 1, 2, 3, \ldots, \tag{27}$$

and the following formulas apply:

$$\mathrm{MA} = \bar{L} = \int_0^\infty Lm(L)\,\mathrm{d}L \Big/ \int_0^\infty m(L)\,\mathrm{d}L, \tag{28}$$

or equivalently

$$\mathrm{MA} = \eta_1/\eta_0 \tag{29}$$

and

$$\mathrm{CV} = \sigma/\bar{L}, \tag{30}$$

where

$$\sigma^2 = \frac{1}{\eta_0} \int_0^\infty (L - \bar{L})^2 m(L)\,\mathrm{d}L. \tag{31}$$

Hence, in terms of moments

$$\mathrm{CV} = (\eta_0\eta_2/\eta_1^2 - 1)^{1/2}. \tag{32}$$

Since in the model adopted (Eqs. (11)) the population balance is expressed as number-volume distribution moments, in order to account for the agglomeration mechanism, a relation between mass-size ($\eta_j(L)$) and number-volume ($\tilde{\mu}_j(v)$) distribution moments is required for the validation procedure.

At this point it is relevant to recall the definitions and relationships between crystal mass-size distribution function ($m(L)$), number-volume distribution function ($n(v)$) and number-size distribution function ($n(L)$):

$$v = k_v L^3, \tag{33}$$

$$n(L)\,\mathrm{d}L = n(v)\,\mathrm{d}v, \tag{34}$$

$$m(L) = \rho_c k_v L^3 n(L)/M_c, \tag{35}$$

where $M_c = \rho_c \mu_1(v)$ is the total crystal mass, with $\mu_1(v)$ representing the total volume of particles (Eq. (9)).

Now, substituting Eqs. (33)–(35) into Eq. (27), a relation between $\eta_j(L)$ and $\mu_j(v)$ is obtained as:

$$\eta_j(L) = \frac{1}{M_c}\left(\frac{\rho_c}{k_v^{k-1}}\mu_k(v)\right),$$

$$k = \frac{j}{3} + 1, \quad j = 0, 3, 6, \ldots\,. \tag{36}$$

The recovering method adopted in this work is based on the assumption that the final sugar mass-size distribution can be reasonably well represented by a normal distribution function (Meireles & Feyo de Azevedo, 1998).

According to Randolph and Larson (1988), the jth moment of a normal distribution (for any type of particle distribution function, but here applied to the mass-size distribution) can be defined as a function of the mean ($\bar{L}$) and

Table 9
Final CSD—experimental data versus hybrid and mechanistic model predictions

Batch no.	Experimental data		Hybrid model		Mechanistic model	
	MA (mm)	CV (%)	MA (mm)	CV (%)	MA (mm)	CV (%)
1	0.479	32.6	0.518	28.12	0.583	21.26
2	0.559	33.7	0.491	29.34	0.542	18.43
3	0.680	43.6	0.575	32.65	0.547	18.69
4	0.494	33.7	0.514	35.16	0.481	14.16
5	0.537	32.5	0.599	29.52	0.623	24.36
6	0.556	35.5	0.496	28.49	0.471	13.642
7	0.560	31.6	0.658	35.67	0.755	34.9
8	0.530	31.2	0.609	36.87	0.681	27.39
av. error			7.8%	13.4%	13.7%	36.1%

the variance (σ) of the distribution as follows:

$$\eta_j(L) = \sum_r \left[(2^{1/2}\sigma)^{j-r} (\bar{L})^r \frac{j!.1.3\ldots(j-r-1)}{(j-r)!r!2^{(j-r)/2}} \right], \quad (37)$$

where $r = 0, 2, 4, \ldots, j$ for j even and $r = 1, 3, 5, \ldots, j$ for j odd.

Now, to compute $\bar{L}$ and σ we make use of the third and the sixth mass-size distribution moments of Eq. (37), i.e.

$$\eta_3/\bar{L}^3 = 1 + 3\left(\frac{\sigma}{\bar{L}}\right)^2, \quad (37a)$$

$$\eta_6/\bar{L}^6 = 1 + 15\left(\frac{\sigma}{\bar{L}}\right)^2 + 45\left(\frac{\sigma}{\bar{L}}\right)^4 + 15\left(\frac{\sigma}{\bar{L}}\right)^6. \quad (37b)$$

Let

$$X = \left(\frac{\sigma}{\bar{L}}\right)^2. \quad (38)$$

Solving together Eqs. (37a), (37b) and (38) leads to

$$\bar{L} = \left(\frac{\eta_3}{1+3X}\right)^{1/3} \quad (39)$$

and

$$15\eta_3^2 X^3 + (45\eta_3^2 - 9\eta_6)X^2 + (15\eta_3^2 - 6\eta_6)X + \eta_3^2 - \eta_6 = 0. \quad (40)$$

Numerical computation will then go through the following steps:

(a) For a given set of number-volume moments $\tilde{\mu}_k(v)$, $\eta_3(L)$ and $\eta_6(L)$ are obtained from Eq. (36).
(b) The roots (only positive X) of polynomial equation (40) are computed.
(c) $\bar{L}$ follows from Eq. (39).
(d) Finally, a model-based estimation of MA and CV from Eqs. (28) and (30) is computed.

MA and CV predicted by the mechanistic and the hybrid models following the above procedure are compared with the experimental data available at the end of eight (validation) batches. The results, summarised in Table 9, confirm the practical advantage of the hybrid model.

Table 9 and the process output predictions indicate that introduction of knowledge from different sources (first principle model and available data record for the process investigated here) can be effectively combined to obtain an improved sugar crystallisation model. The hybrid model obtained seems to posses the advantages of both fully analytical and fully data-driven modelling approaches without being too time-consuming and expensive procedure.

7. Conclusions

The difficulty in crystallisation modelling is essentially on the accurate description of the CSD with the respective quantities—MA and CV. The experience with models neglecting agglomeration and/or nucleation mechanisms shows that the CSD predictions do not correspond to the experimentally obtained MA and CV at the end of the process run. Therefore, accurate modelling can only be achieved by incorporating agglomeration and nucleation mechanisms.

In this paper, a KBH model of a sugar crystallisation process when accounting for nucleation, growth and agglomeration phenomena is developed and examined. The concept of serial hybrid modelling is adopted where only the poorly known kinetic parameters (nucleation rate, growth rate, agglomeration kernel) are replaced by a feedforward ANN. This black-box approximation is incorporated in the structure of the classical mechanistic model reflecting the mass, energy and population balances. The KBH model demonstrates good agreement with the experimental data available.

The hybrid modelling offers a reasonable compromise between the extensive efforts to get a fully parameterised structure, as are the mechanistic models and the poor generalisation of the complete data-based modelling approaches.

A reliable description of the kinetic parameters is of special importance not only for the academic understanding of the crystallisation phenomena, but also for the purposes of

optimising the manipulated input time profiles, with the objective to obtain sugar crystals with desired quality characteristics.

Notation

$B(v)$	birth rate function (based on total vessel volume), $1/m^6$ s
B_0	nucleation rate, $1/m^3$ s
BPE	boiling point elevation, °C
Bx	brix (mass fraction of dissolved solids)
CV	coefficient of variation, %
$D(v)$	death rate function (based on total vessel volume), $1/m^6$ s
F	feed flowrate, m^3/s
F_s	steam flowrate, kg/s
G	linear growth rate, m/s
G_v	overall crystal volume growth rate, m^3/s
H	specific enthalpy, J/kg
k_v	volume shape factor
k_{vap}	parameter of the evaporation rate [Eq. (3)]
K_{ag}	crystal agglomeration kinetic constant
K_g	crystal growth kinetic constant
K_n	crystal nucleation kinetic constant
L; $\bar{L}$	particle size; mean size of a population distribution
$m(L)$	crystal mass-size distribution function, 1/m
M	mass, kg
MA	mass averaged crystal size, m
$n(v)$	crystal number-volume distribution function, $1/m^6$
$n(L)$	crystal number-size distribution function, $1/m^4$
Pur	purity (mass fraction of sucrose in the dissolved solids)
Q	heat input, J/s
R	gas constant, J/molK
S	supersaturation
t	time, s
T	temperature, °C
v	crystal volume, m^3
V	volume, m^3
W	stirring power, J/s

Greek letters

α_s	parameter of the heat input [Eq. (2)]
β'	agglomeration kernel at any time for vessel volume, 1/s
β	agglomeration kernel, m^3/s
$\eta_j(L)$	j-moment of mass-size distribution function $m(L)$, m^j
$\lambda_{w(vac)}$	latent heat of water evaporation, J/kg
$\tilde{\mu}_j(v)$	j-moment of number-volume distribution function $n(v)$, $m^{3j}/(m^3)$
ρ	density, kg/m^3
σ	variance of a population distribution
υ_c	volume fraction of crystals

Subscripts

c	crystals
cris	crystallisation
f	feed
i	impurities
m	massecuite
s	dissolved sucrose
sat	saturation
seed	seeding
sol	solution
syr	syrup
vac	vacuum
vap	evaporation
w	water

Superscripts

hyb	hybrid
NN	neural network
obs	observed (measured) data
$\sim$	based on total vessel volume, ($\tilde{x} = xV_m$)

Acknowledgements

Dr. Petia Georgieva is on leave from the *Institute of Control and System Research, Bulgarian Academy of Sciences, Sofia, Bulgaria*, supported by EU Research Project HPRN-CT-2000-00039. This work was further financed by the Portuguese Foundation for Science and Technology within the activity of the Research Unit *Institute for Systems and Robotics-Porto.*

Appendix A. Correlations for physical properties

Feed

Density of the feed liquor/syrup (kg/m^3):

$$\rho_f = \left(1000 + \frac{Bx_f(200 + Bx_f)}{54}\right)\left(1 - 0.036\,\frac{T_f - 20}{160 - T_f}\right).$$

Specific heat capacity of the feed liquor/syrup (J/(kg °C)):

$$Cp_f = 4186.8 - 29.7Bx_f + 4.61Bx_f\mathrm{Pur}_f + 0.075Bx_fT_f.$$

Specific enthalpy of the feed liquor/syrup (J/kg):

$$H_f = Cp_fT_f.$$

Solution

Density of pure solution (kg/m^3):

$$\rho^*_{\mathrm{sol}} = \left(1000 + \frac{Bx_{\mathrm{sol}}(200 + Bx_{\mathrm{sol}})}{54}\right)\left(1 - 0.036\,\frac{T_m - 20}{160 - T_m}\right).$$

Density of impure solution (kg/m^3):

$$\rho_{\mathrm{sol}} = \rho^*_{\mathrm{sol}} + 1000(-1 + \exp[(-6.927 \times 10^{-6} Bx^2_{\mathrm{sol}} - 1.164 \times 10^{-4} Bx_{\mathrm{sol}})(\mathrm{Pur}_{\mathrm{sol}} - 1)]).$$

Specific heat capacity of solution (J/(kg °C)):

$$Cp_{\mathrm{sol}} = 4186.8 - 29.7Bx_{\mathrm{sol}} + 4.61Bx_{\mathrm{sol}}\mathrm{Pur}_{\mathrm{sol}} + 0.075Bx_{\mathrm{sol}}T_m.$$

Specific enthalpy of solution (J/kg):

$$H_{\mathrm{sol}} = Cp_{\mathrm{sol}}T_m.$$

Density of massecuite (kg/m^3):

$$\rho_m = \frac{\rho_{\mathrm{sol}}\rho_c}{\rho_c - \omega_c(\rho_c - \rho_{\mathrm{sol}})}.$$

Volume of the massecuite:

$$V_m = (M_c + M_{\mathrm{sol}})/\rho_m.$$

Crystals

Crystal volume (m^3):

$$v = M_c/\rho_c.$$

Mass fraction of crystals:

$$\omega_c = \frac{M_c}{M_c + M_{\mathrm{sol}}}.$$

Specific heat capacity of crystals (J/(kg °C)):

$$Cp_c = 1163.2 + 3.488T_m.$$

Specific enthalpy of crystals (J/kg):

$$H_c = Cp_cT_m.$$

Water

Density of water (kg/m^3):

$$\rho_w = 1016.7 - 0.57T_w.$$

Saturation temperature of water (°C):

$0.1 < P_{\mathrm{vac}} < 1$ bar,

$$T_{w(\mathrm{vac})} = 122.551\exp(-0.246P_{\mathrm{vac}})(P_{\mathrm{vac}})^{0.413},$$

$1 < P_s < 3$ bar,

$$T_{w(s)} = 100.884\exp(-1.203 \times 10^{-2}P_s)(P_s)^{0.288}.$$

Latent heat of water evaporization (kJ/kg):

$0.1 < P_{\mathrm{vac}} < 1$ bar, $\lambda_{w(\mathrm{vac})} = 2263.28 - 58.21\ln(P_{\mathrm{vac}})$,

$1 < P_s < 3$ bar, $\lambda_s = 2257.51 - 85.95\ln(P_s)$.

Specific enthalpy of water (J/kg):

$$H_w = 2323.3 + 4106.7T_w + 0.6T^2_w,$$

$$H_{w(s)} = 2323.3 + 4106.7T_{w(s)} + 0.6T^2_{w(s)},$$

$$H_s = 2\,491\,860 - 13\,270P_s + (1947.5 + 37.9P_s)T_s,$$

$$H_{\mathrm{vac}} = 2\,499\,980 - 24\,186P_{\mathrm{vac}} + (1891.1 + 106.1P_{\mathrm{vac}})T_m,$$

$$\Delta H_s = H_s - H_{w(s)}.$$

Boiling point elevation [°C]:

$$\mathrm{BPE} = (0.03 - 0.018\ \mathrm{Pur}_{\mathrm{sol}})(T_{w(\mathrm{vac})} + 84)\,\frac{Bx_{\mathrm{sol}}}{100 - Bx_{\mathrm{sol}}}.$$

Supersaturation nucleation bound:

$$S^* = 1.129 - 0.284(1 - \mathrm{Pur}_{\mathrm{sol}}) + [2.333 - 0.0709(T_m - 60)](1 - \mathrm{Pur}_{\mathrm{sol}})^2.$$

Auxiliary correlations

Mass of solution:

$$M_{\mathrm{sol}} = M_s + M_i + M_w.$$

Purity of the solution (the mass fraction of sucrose in the dissolved solids):

$$\mathrm{Pur}_{\mathrm{sol}} = \frac{M_s}{M_s + M_i}.$$

Brix of solution (mass of total dissolved solids per 100 mass units of solution) (%):

$$B_{\mathrm{sol}} = \frac{M_s + M_i}{M_{\mathrm{sol}}}, \quad Bx_{\mathrm{sol}} = 100B_{\mathrm{sol}}.$$

Volume fraction of crystals:

$$\upsilon_c = \frac{v}{V_m}.$$

Saturation brix of solution:

$$Bx_{\mathrm{sat}} = 64.447 + 8.222 \times 10^{-2}T_m + 1.66169 \times 10^{-3}T^2_m - 1.558 \times 10^{-6}T^3_m - 4.63 \times 10^{-8}T^4_m.$$

Supersaturation:

$$S = \frac{Bx_{\mathrm{sol}}/(100 - Bx_{\mathrm{sol}})}{Bx_{\mathrm{sat}}/(100 - Bx_{\mathrm{sat}})C_{\mathrm{sat}}},$$

$$C_{\mathrm{sat}} = 0.1\,\frac{Bx_{\mathrm{sol}}}{100 - Bx_{\mathrm{sol}}}(1 - \mathrm{Pur}_{\mathrm{sol}}) + 0.4 + 0.6\exp\left(-0.24\,\frac{Bx_{\mathrm{sol}}}{100 - Bx_{\mathrm{sol}}}(1 - \mathrm{Pur}_{\mathrm{sol}})\right).$$

Appendix B. Empirical model for the evaporation rate

In industrial applications there are a number of unknowns, which hamper the use of pure theoretical considerations in

the modelling of heat fluxes. It is the case of unquantifiable heat losses all over the unit, of the unreliable measurements of steam heat contents and of the inaccurate expressions available to represent evaporation rates. A possible way to cope with such type of problems is to consider empirical correction factors on the theoretical relationships, with parameters tuned by experiment.

Hence, the first empirical parameter is introduced in the heat input, Q, given by

$$Q = \alpha_s F_s \Delta H_s. \tag{A.1}$$

α_s may be interpreted as a corrective term of the heating steam enthalpy, which includes heat losses and errors in the steam flow and pressure sensors. Experimental data for the estimation of α_s is collected from the concentration period, when no crystals are present in the pan and the input fluxes are closed ($F_f = 0,\ F_w = 0$). During this phase the energy balance equations (4) simplify to

$$\frac{\mathrm{d}T_m}{\mathrm{d}t} = cJ_{\mathrm{vap}} + d \tag{A.2}$$

with

$$c = \frac{1}{M_{\mathrm{sol}} C_{P\mathrm{sol}}} \left[H_{\mathrm{sol}} - H_{\mathrm{vap}} - B_{\mathrm{sol}} \frac{\mathrm{d}H_{\mathrm{sol}}}{\mathrm{d}B_{\mathrm{sol}}} \right],$$

$$d = \frac{1}{M_{\mathrm{sol}} C_{P\mathrm{sol}}} [W + Q]. \tag{A.2a}$$

During this period, it is possible to calculate the mass of evaporated water, solving (A.2) with respect to J_{vap}:

$$J_{\mathrm{vap}} = \frac{W + Q - M_{\mathrm{sol}} C_{P\mathrm{sol}} (\mathrm{d}T_m/\mathrm{d}t)}{H_{\mathrm{vap}} - (H_{\mathrm{sol}} - B_{\mathrm{sol}} (\mathrm{d}H_{\mathrm{sol}}/\mathrm{d}B_{\mathrm{sol}}))}. \tag{A.3}$$

The integral of J_{vap} over the concentration period is obtained by material balance, employing the experimental brix of solution. Consequently, the integration of Eq. (A.3) over the same period leads to the analytical solution for α_s since all other terms and parameters are known either from theoretical or experimental data.

In a second stage, an alternative expression of J_{vap} is obtained by taking the practical view that the evaporation rate should include the effect of an existing superheat in the massecuite (Batterham & Norgate, 1975), expressed in terms of the temperature difference ($T_m - T_{w(\mathrm{vac})} - \mathrm{BPE}$). The simplified form for J_{vap} can be written as

$$J_{\mathrm{vap}} = \frac{W + Q}{\lambda_{w(\mathrm{vac})}} + k_{\mathrm{vap}} (T_m - T_{w(\mathrm{vac})} - \mathrm{BPE}). \tag{A.4}$$

Identically to the calculation of α_s, the empirical parameter k_{vap} is obtained from the analytical solution of the integral of Eq. (A.4) over the concentration period. The optimised values for α_s and k_{vap} obtained by the analysis of experimental data taken from nine batches in different days are reported in Table 8. Fig. 6 shows the evolution of mass of water in solution, comparing the experimental profile with those obtained by employing Eq. (A.3), considered as a theoretical model, and Eq. (A.4), adopted as the empirical correlation, for the evaporation rate in the overall crystallisation model (see Eq. (3)). The values for the empirical parameters are taken from Table 8 and $\mathrm{d}T_m/\mathrm{d}t$ in Eq. (A.3) is approximated by $\Delta T_m/\Delta t$, with $\Delta t = 10$ s.

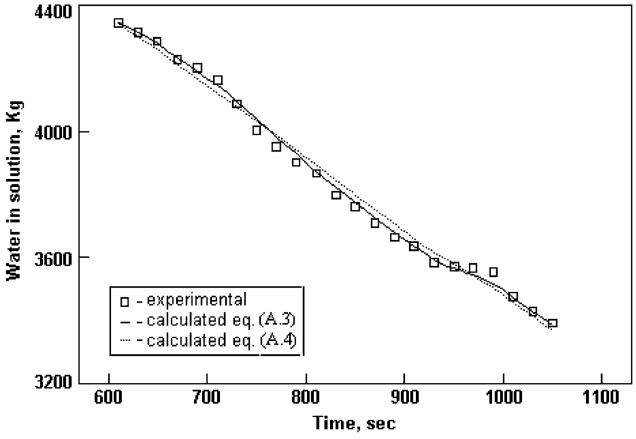

Fig. 6. Experimental and calculated mass of water in solution during the concentration period.

References

Batterham, R. J., & Norgate, T. E. (1975). Boiling point elevation and superheat in impure cane sugar solutions. *International Sugar Journal*, *77*, 359–364.

Chen, L., Bernard, O., Bastin, G., & Angelov, P. (2000). Hybrid modelling of biotechnological processes using neural networks. *Control Engineering Practice*, *8*, 821–827.

Chorão, S., & Feyo de Azevedo, S. (1996). A discretized population balance approach for the modelling of industrial sucrose crystallisation. *13th symposium on industrial crystallisation*, Toulouse, France, 16–19 September, 1996 (pp. 719–725).

Demuth, H., Beale, M. (1997). *Neural Network Toolbox User's Guide*. Version IV, Math Works, Inc.

Ditl, P., Beranek, L., & Rieger, Z. (1990). Simulation of a stirred sugar boiling pan. *Zuckerindustric*, *115*, 667–676.

Feyo de Azevedo, S., Chorão, J., Gonçalves, M. J., & Bento, L. (1993). On-line monitoring of white sugar crystallisation through software sensors—Part I. *International Sugar Journal*, *95*, 483–488.

Feyo de Azevedo, S., Chorão, J., Gonçalves, M. J., & Bento, L. (1994). On-line monitoring of white sugar crystallisation through software sensors—Part II. *International Sugar Journal*, *96*, 18–26.

Feyo de Azevedo, S., Dahm, B., & Oliveira, F. R. (1997). Hybrid modelling of biochemical processes: A comparison with the conventional approach. *Computers and Chemical Engineering*, *21*, S751–S756.

Franck, R., David, R., Villermaux, J., & Klein, J. P. (1988). Crystallisation and precipitation engineering—II. A chemical reaction engineering approach to salicylic acid precipitation: Modelling of batch kinetics and application to continuous operation. *Chemical Engineering Science*, *43*, 69–81.

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.

Hartel, R. W., & Randolph, A. D. (1986). Mechanisms and kinetic modelling of calcium oxalate crystal aggregation in a urinelike liquor, part II. Kinetic modelling. *A.I.Ch.E. Journal*, *32*, 1186–1199.

Hounslow, M. J. (1990). Nucleation, growth and aggregation rates from steady state experimental data. *A.I.Ch.E. Journal*, *36*, 1748–1759.

Hounslow, M. J., Ryall, R. L., & Marshall, V. R. (1988). A discretised population balance for nucleation, growth and aggregation. *A.I.Ch.E. Journal*, *34*, 1821–1831.

Hulburt, H. M., & Katz, S. (1964). Some problems in particle technology: A statistical mechanical formulation. *Chemical Engineering Science*, *18*, 555–564.

Jones, A. (1974). Optimal operation of a batch cooling crystalliser. *Chemical Engineering Science*, *29*, 1075–1087.

Lauret, P., Boyer, H., & Gatina, J. C. (2000). Hybrid modelling of a sugar boiling process. *Control Engineering Practice*, *8*, 299–310.

Meireles, M. J., & Feyo de Azevedo, S. (1998). Modelling the operation of a sugar industrial evaporative crystalliser. *CHEMPOR'98*, Lisbon, Portugal, 26–28 September, 1998 (pp. 807–814).

Najim, K., Ruiz, V., Feyo de Azevedo, S., & Goncalves, M. J. (1996). Modelling and adaptive control of a batch evaporative crystallise. *Journal of Systems Engineering*, *6*, 233–241.

Peres, J., Oliveira, R., & Feyo de Azevedo, S. (2001). Knowledge based modular networks for process modelling and control. *Computers and Chemical Engineering*, *25*, 783–791.

Psichogios, D. C., & Ungar, L. H. (1992). A hybrid neural network—first principles approach to process modelling. *A.I.Ch.E. Journal*, *38*(10), 1499–1511.

Randolph, A. D., & Larson, M. A. (1988). *Theory of particulate processes—analyses and techniques of continuous crystallisation*. New York: Academic Press.

Roubos, J. A., Krabben, P., Setnes, M., Babuska, R., Heijnen, J. J., & Verbrugen, H. B. (1999). Hybrid model development for fed-batch bioprocesses; combining physical equations with the metabolic network and black-box kinetics. *6th Workshop on fuzzy systems*, September 8–9, Brunel University, Uxbridge (pp. 231–239).

Russell, N. T., & Bakker, H. H. C. (1997). Modular modelling of an evaporator for long-range prediction. *Artificial Intelligence in Engineering*, *11*, 347–355.

Schubert, J., Simutis, R., Dors, M., Havlik, I., & Lubbert, A. (1994a). Hybrid modelling of yeast production processes—combination of a priori knowledge on different levels of sophistication. *Chemical Engineering Technology*, *17*, 10–20.

Schubert, J., Simutis, R., Dors, M., Havlik, I., & Lubbert, A. (1994b). Bioprocess optimisation and control: Application of hybrid modelling. *Journal of Biotechnology*, *35*, 51–68.

Simutis, R., Oliveira, R., Manikowski, M., Feyo de Azevedo, S., & Lubbert, A. (1997). How to increase the performance of models for process optimisation and control. *Journal of Biotechnology*, *59*, 73–89.

Tavare, N., & Garside, J. (1993). Silica precipitation in a semi-batch crystalliser. *Chemical Engineering Science*, *48*, 475–488.

Thompson, M. L., & Kramer, M. A. (1994). Modelling chemical processes using prior knowledge and neural networks. *A.I.Ch.E. Journal*, *40*(8), 1328–1340.

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