DYNAMIC ESTIMATION AND UNCERTAINTY QUANTIFICATION FOR MODEL-BASED CONTROL OF DISCRETE SYSTEMS

João F.M. Gândara [*], **Belmiro P.M. Duarte** [**], **Nuno M.C. Oliveira** [***]

[*] *Department of Food Science and Technology, ESAC, Polytechnic Institute of Coimbra. Bencanta, 3040-316 Coimbra, Portugal. Tel. +351-239-802940.*
[**] *Department of Chemical Engineering, ISEC, Polytechnic Institute of Coimbra. R. Pedro Nunes, 3030-199 Coimbra, Portugal. Tel. +351-239-790200.*
[***] *Department of Chemical Engineering, University of Coimbra. Pólo II, Pinhal de Marrocos, 3030-290 Coimbra, Portugal. Tel. +351-239-798700.*

Abstract: This paper presents an approach to estimate the outputs and the uncertainty associated to the forecast for discrete dynamic systems represented by state-space models. The complete strategy includes three steps: 1. process identification based on a data sample; 2. estimation of the current process state based on the information available during a moving past horizon, which may contain lack of observations; 3. forecast of process states, process outputs and uncertainty along the future horizon. This procedure can be incorporated in control strategies that explicitly consider model uncertainty.

Keywords: Estimation, process monitoring, optimal sampling, quality control.

1. INTRODUCTION

Traditional discrete process control applications assume that the sampling period used for interaction with the process, either through measurements or actuations, is fixed. This parameter is often chosen during the initial design phase of the control system, and before the specification of the control law to be used. However, the recent development of sophisticated control strategies, such as model-based approaches, and the integration of process information acquired from a number of distinct sources, has placed more emphasis on the choice and on-line adjustment of sampling policies, mostly for economical reasons.

The tasks of process and quality control commonly require the use of off-line analytical equipment to measure key product characteristics, such as concentrations and properties of particle systems; this can involve scarce and expensive human and equipment resources. In certain situations, the effective allocation of analytical resources can benefit from an economic performance analysis that simultaneously considers the relative value and costs associated with new information that can be introduced in an optimization problem.

Previous work on the selection of appropriate sampling intervals for process control with basis on economic criteria has been considered by MacGregor (1976), Abraham (1979), and Kramer (1989). The approach followed by these authors assumed the availability of a linear dynamic model of the process, incorporating a stochastic component, used to predict the average performance of the controlled system when a larger sampling interval, equal to integer multiples of the basic sampling interval, is selected. This requires the use of a cost function that considers the cost of being off-specification, in terms of the variance of the observed errors, and the costs of taking new samples and making further process adjustments.

In this paper we propose a strategy for the forecast of the quality variables and their uncertainty, which are used to predict the probabilities of these variables being outside their quality specifications. Before the forecasts are made, it is necessary to estimate the current process state. For this a procedure is developed which is capable of effectively dealing with incomplete data sets. All these tasks are accomplished using a state-space model with a stochastic component. Finally, the proposed strategy is tested using a simulated continuous fermenter for ethanol production.

2. PROCESS MODEL

The approach described in this paper is applied to state-space models of the family $\mathcal{M}_1$

$$\mathcal{M}_1(A, B, C, D, K, \mathrm{Cov}(e)) = \\ = \begin{cases} x(t_{k+1}) = A\ x(t_k) + B\ u(t_k) + K\ e(t_k) \\ y(t_k) = C\ x(t_k) + D\ u(t_k) + e(t_k) \end{cases} \tag{1}$$

where $x(t_k) \in \mathbf{R}^{n_\mathrm{s}}$ is the vector of states at discrete time t_k, $u(t_k) \in \mathbf{R}^{n_\mathrm{i}}$ is the vector of inputs, $y(t_k) \in \mathbf{R}^{n_\mathrm{o}}$ is the vector of outputs and $e(t_k) \in \mathbf{R}^{n_\mathrm{o}}$ is the vector of stochastic components included in the state variables. The matrices A, B, C, D, K and $\mathrm{Cov}(e)$ are time invariant parameters.

Process identification is performed based on a complete data sample (including all process dynamic features) by employing a subspace projection algorithm (*N4SID*), an approach devoted to discrete systems identification (Van Overschee, 1994). The data represents the open–loop process behavior along the time horizon $N \times \Delta t$, where N is the number of records used for identification and Δt is the sampling interval. The *N4SID* algorithm only requires the knowledge of the system order, thus avoiding the need of *a priori* parametrization, and is non–iterative, avoiding the need of optimization schemes with corresponding problems, such as the convergence rate and the existence of local minima (Ljung, 1999).

The order of the system, n_s, is determined applying an information–based criterion, the Akaike Information Criterion (AIC), to measure the model fitness to process data (Akaike, 1972). The *AIC* metric of the models of the family $\mathcal{M}_1$ with order $n \in \{1, \cdots, n_\mathrm{s}{}^{\max}\}$, $\mathcal{M}_1^n$, is represented as:

$$AIC(\hat{\theta}^n) = \\ \log\left\{\frac{1}{N}\sum_{i=1}^{N}[\epsilon(i,\hat{\theta}^n)]^2\right\} + \frac{\dim(\hat{\theta}^n)}{N} \tag{2}$$

where $\hat{\theta}^n$ is the vector of parameter estimates included in the model $\mathcal{M}_1^n$, $\hat{\theta}^n = \{\hat{A}, \hat{B}, \hat{C}, \hat{D}, \hat{K}, \mathrm{Cov}(e)\}^n$, and $\epsilon(i, \hat{\theta}^n)$ is the error of estimates of outputs $i \in \{1, \cdots, N\}$. The order of the system is determined as:

$$n_\mathrm{s} = \underset{n \in \{1, n_\mathrm{s}{}^{\max}\}}{\arg\min}\ AIC(\hat{\theta}^n) \tag{3}$$

where $n_\mathrm{s}{}^{\max}$ is the maximum order iterated. Process identification is performed off–line and the model is to be updated whenever process modifications are detected.

3. STATE ESTIMATION

The estimate of the current state process, $\hat{x}(t_0)$, can be obtained from the information available in the form of the inputs and the measurements obtained from sampling the process in the current, t_0, and past sampling times. We consider the receding horizon $\mathcal{H}_r$ comprising the last r discrete sampling times, $\mathcal{H}_r = \{t_{-r+1}, t_{-r+2}, \cdots, t_{-1}, t_0\}$. A possible approach to this problem (Brookner, 1998) consists on obtaining a set of equations in order to $\hat{x}(t_0)$. This is achieved by recursive substitution of all the state variables in the model equations at every sampling time in $\mathcal{H}_r$. This approach involves the use of negative powers of the transition matrix, A, and may lead to ill-conditioned problems, especially for stable systems and large horizons. The approach used in this paper consists on the simultaneous solution of all the model equations in the horizon $\mathcal{H}_r$. Although this leads to larger problems, it is a numerically stable procedure, avoiding ill-conditioning.

In the proposed approach, the problem of estimating the current state process is dealt with by solving the state equations at every sampling time in the horizon $\mathcal{H}_r$, together with the output equations referring to the available measurements

$$\begin{cases} \hat{x}(t_k) = A\ \hat{x}(t_{k-1}) + B\ u(t_{k-1}),\ t_k \in \mathcal{H}_r \\ y(t_k) = C_k\ \hat{x}(t_k) + D_k\ u(t_k),\ t_k \in \mathcal{H}_r, \end{cases} \tag{4}$$

where $y(t_k) \in \mathbf{R}^{n_{\mathrm{o},k}}$, $0 \le n_{\mathrm{o},k} \le n_\mathrm{o}$ is the vector containing the variables measured at sampling time t_k. When no output is measured we have $n_{\mathrm{o},k} = 0$, and when all outputs are measured $n_{\mathrm{o},k} = n_\mathrm{o}$. Matrices $C_k \in \mathbf{R}^{n_{\mathrm{o},k} \times n_\mathrm{s}}$ and $D_k \in \mathbf{R}^{n_{\mathrm{o},k} \times n_\mathrm{i}}$ contain the rows of C and D

corresponding to the measured variables at t_k. We assume that all inputs in the horizon are perfectly known.

Defining the enlarged vectors of states and inputs

$$\hat{\mathcal{X}}_r = \left[\hat{x}(t_0)^{\mathrm{T}}\ \hat{x}(t_{-1})^{\mathrm{T}}\ \cdots\ \hat{x}(t_{-r+1})^{\mathrm{T}}\right]^{\mathrm{T}},$$
$$\mathcal{U}_r = \left[u(t_0)^{\mathrm{T}}\ u(t_{-1})^{\mathrm{T}}\ \cdots\ u(t_{-r+1})^{\mathrm{T}}\right]^{\mathrm{T}},$$

respectively, the system of equations (4) can be formulated as

$$\mathcal{J}_r\ \hat{\mathcal{X}}_r = \mathcal{L}_r, \qquad (5)$$

where $\mathcal{L}_r = \mathcal{I}_r + \mathcal{N}_r\mathcal{U}_r$. Vector $\mathcal{I}_r$ contains the output information in the considered horizon:

$$\mathcal{I}_r = \left[y(t_0)^{\mathrm{T}}\ \mathbf{0}_{n_s}^{\mathrm{T}}\ y(t_{-1})^{\mathrm{T}}\ \mathbf{0}_{n_s}^{\mathrm{T}}\ \cdots\ y(t_{-r+1})\right]^{\mathrm{T}},$$

where $\mathbf{0}_{n_s}$ is a vector containing n_s zeroes. Matrix $\mathcal{N}_r$ is formed by the model matrices B and D_k, and $\mathcal{J}_r$ is a sparse and structured matrix formed by matrices C_k, A and I_{n_s} (n_s–dimension identity matrix). For instance, using a generic receding horizon with dimension $r = 4$, we have:

$$\mathcal{J}_r = \begin{pmatrix} C_0 & & & \\ I_{n_s} & -A & & \\ & C_{-1} & & \\ & I_{n_s} & -A & \\ & & C_{-2} & \\ & & I_{n_s} & A \\ & & & C_{-3} \end{pmatrix};$$

$$\mathcal{N}_r = \begin{pmatrix} -D_0 & & & & & & \\ & B & & & & & \\ & -D_{-1} & & & & & \\ & & & B & & & \\ & & & -D_{-2} & & & \\ & & & & & B & \\ & & & & & -D_{-3} & \end{pmatrix}.$$

The process state estimate along $\mathcal{H}_r$, $\hat{\mathcal{X}}_r$, which includes the current process state, designated as $\hat{x}(t_0)$, is determined by solving the least-squares problem

$$\min_{\hat{\mathcal{X}}_r} \parallel \mathcal{L}_r - \mathcal{J}_r\ \hat{\mathcal{X}}_r \parallel^2 \qquad (6)$$

The optimality of $\hat{\mathcal{X}}_r$ is achieved whenever problem (6) is determined or over-determined. Taking into account that $\mathcal{J}_r \in \mathbf{R}^{m\times n}$, with

$$m = r\ n_s + \sum_{i=0}^{r-1} n_{\mathrm{o},-i}, \quad n = (r+1)\ n_s,$$

the necessary condition that leads to the determination or over-determination of problem (6) is $m \geq n$, thus requiring

$$\sum_{i=0}^{r-1} n_{\mathrm{o},-i} \geq n_s.$$

Therefore, the problem is determined if at least n_s observations are available in $\mathcal{H}_r$.

Problem (6) is solved by performing a QR decomposition of matrix $\mathcal{J}_r$,

$$\mathcal{J}_r = \left[Q_1\ Q_2\right]\begin{bmatrix} R_1 \\ 0 \end{bmatrix} \qquad (7)$$

and subsequently applying a back-substitution procedure to determine $\hat{\mathcal{X}}_r$. In closed form, the problem solution is represented as

$$\hat{\mathcal{X}}_r = R_1^{-1}Q_1^{\mathrm{T}}\mathcal{L}_r. \qquad (8)$$

The outputs in the horizon $\mathcal{H}_r$ are affected by errors, mainly measurement ones. Since the state estimates are obtained from the available measurements, they also include an error component,

$$\hat{\mathcal{X}}_r = \mathcal{X}_r + e_{\mathcal{X}}.$$

Assuming that $E\{e(t_k)\} = 0, \forall t_k \in \mathcal{H}_r$, then $\hat{\mathcal{X}}_r$ is an unbiased estimator of $\mathcal{X}_r$ (Ikonen, 2002), leading to $E\{e_{\mathcal{X}}\} = 0$. Using (8), the covariance of this error can be computed as

$$\mathrm{Cov}(e_{\mathcal{X}}) = R_1^{-1}Q_1^{\mathrm{T}}\mathrm{Cov}(e_r)Q_1(R_1^{-1})^{\mathrm{T}},$$

where e_r is a vector containing all the errors in the horizon $\mathcal{H}_r$ (Brookner, 1998), which are due both to modelling and to measurement components.

The structure and sparsity of matrices $\mathcal{J}_r$ and R_1 is exploited by the use of a tailored algorithm for the QR decomposition. This algorithm avoids the decomposition of the full matrix, and performs successive factorizations of smaller blocks, reducing the number of necessary operations. From this point of view, an important feature of this problem is that the the structure of R_1 is not dependent of the number of variables measured in the horizon $\mathcal{H}_r$.

The receding horizon $\mathcal{H}_r$ is updated every time the procedure is used, in order to incorporate the new information, as it becomes available. If the information in $\mathcal{H}_r$ is not sufficient to ensure the optimality of $\hat{x}(t_0)$, then its size can be increased.

4. PREDICTION

The current process estimate, together with the process model, can be used to forecast the quality variables in a future horizon $\mathcal{H}_f$, comprising f discrete sampling times ahead of t_0, $\mathcal{H}_f = \{t_1, \cdots, t_f\}$. This forecast can be split in two components, the expected values of the quality variables, $\hat{y}$ and the uncertainty associated with the forecast, ε:

$$y(t_k) = \hat{y}(t_k) + \varepsilon(t_k),\ t_k \in \mathcal{H}_f. \qquad (9)$$

Both these components are used to forecast the probability of a given quality variable being out of a set of predetermined specifications.

The expected values of the quality variables are predicted using the deterministic part of the model (1). Using the the previously obtained state estimate as the initial condition:

$$\mathcal{M}_2(A,B,C,D,K,\text{Cov}(e)) = \begin{cases} x(t_k) = A\ x(t_{k-1}) + B\ u(t_{k-1}), & t_k \in \mathcal{H}_f \\ \hat{y}(t_k) = C\ x(t_k) + D\ u(t_k), & t_k \in \mathcal{H}_f \\ x(t_0) = \hat{x}(t_0) \end{cases} \tag{10}$$

Again, we assume that the future profile of the input variables has been determined (using, for example, a MPC-type strategy) and is known.

The uncertainty in the obtained forecast of the quality variables can also be predicted, using the stochastic part of the process model (1). This uncertainty is not only due to the error terms, but also to the uncertainty in the value of $\hat{x}(t_0)$. Propagating (1) in the prediction horizon, and including a term due to the error in the initial condition, we obtain:

$$\begin{aligned} \varepsilon(t_k) &= y(t_k) - \hat{y}(t_k) \\ &= C \sum_{j=0}^{k} A^j K e(t_j) + e(t_k) + C A^k e_{x_0},\ t_k \in \mathcal{H}_f. \end{aligned} \tag{11}$$

From the above equation we can conclude that $E\{\varepsilon(t_k)\} = 0$, since $E\{e(t_k)\} = 0, t_k \in \mathcal{H}_f$, and $E\{e_{x_0}\} = 0$.

The result obtained in (11) is useful, not to predict the actual value of the error, but because it allows us to obtain a measure of the uncertainty in the predictions $\hat{y}(t_k)$. Based on the work of Seppala (1998), the variance of the ith element of $\varepsilon(t_k)$ is computed as

$$\begin{aligned} \hat{\sigma}_i(t_k) = \text{var}(\varepsilon_i(t_k)) = \langle \Psi_i^{\mathrm{T}}\ \Psi_i,\ \text{Cov}(e)\rangle + \\ \langle (\Omega_i)^{\mathrm{T}} \Omega_i, \text{Cov}(e_{x_0})\rangle,\ t_k \in \mathcal{H}_f, \end{aligned} \tag{12}$$

where $\langle \bullet, \bullet \rangle$ is the internal product operator, and Ψ_i and Ω_i are the ith rows of matrices Ψ and Ω:

$$\Psi = C \left(\sum_{j=0}^{k} A^j \right) K + I \tag{13}$$

$$\Omega = C A^k. \tag{14}$$

Note that, since $E\{e(t_k)\} = 0$, the variance of $\hat{y}_i(t_k)$ is the same as that of $\varepsilon_i(t_k)$.

As the prediction instant moves further away from the current sampling time, the second term in (12) continuously increases, due to the accumulation of the model uncertainty. The behavior of the second term, depends on the process stability. If the process is unstable, this term will also increase. If the process is stable, the contribution of the initial error for the forecast variance continuously decreases as t_k mover further away from the current sampling time, since $A^k e_{x_0}$ goes to zero as k increases.

With the obtained information it is possible to predict the probability of a given quality variable being outside it specifications, LS_i and US_i, lower and upper specification limits, respectively. If the random noise, $e(t)$, is well described by a stationary Gaussian distribution, then this probability can be predicted by:

$$\begin{aligned} pf_i(t) = \\ \int_{-\infty}^{LS_i} \frac{1}{\sqrt{2\ \pi}\ \hat{\sigma}_i(t_k)}\ \exp\left[-\frac{(z - \hat{y}_i(t_k))^2}{2\ (\hat{\sigma}_i(t_k))^2} \right] \mathrm{d}z + \\ + \int_{US_i}^{+\infty} \frac{1}{\sqrt{2\ \pi}\ \hat{\sigma}_i(t_k)}\ \exp\left[-\frac{(z - \hat{y}_i(t_k))^2}{2\ (\hat{\sigma}_i(t_k))^2} \right] \mathrm{d}z, \\ t_k \in \mathcal{H}_f \end{aligned} \tag{15}$$

A large value of this probability can be due either to a shift in the process or to the increase of the uncertainty in the forecasts. The first can be solved by taking appropriate control measures, in order to drive the process back to the central value of the specifications. If the uncertainty in the forecasts is too large, then a a new measurement of the variable for which pf is too large should be made, before or at the sampling time where this occurs. With this new information, the estimation procedure described earlier is repeated, in order to get a new estimate of the process state, with reduced uncertainty.

5. APPLICATION EXAMPLE

The proposed strategy was tested using a nonlinear dynamical model of a continuous fermenter for ethanol production using glucose (Chmúrny, 2000). The measured variables considered in this model are:

(1) biomass concentration, x;
(2) substrate concentration, s;
(3) product concentration, P;
(4) biomass concentration in the output solution, x_v;
(5) carbon dioxide production rate, r_{CO_2};
(6) rate of base consumption, r_z.

The carbon dioxide production and the base consumption rates can be easily measured by online sensors. The concentrations require more complex analytic methods, and thus, are more expensive to obtain. The input variables are D, the ratio between the fermenter feed rate and its volume, and s_0, the substrate concentration in the feed. In this particular model, the temperature and the volume are assumed to be constant.

5.1 Model identification

The fermenter dynamical model was used to generate a complete data set, with random errors added to the output variables. The *N4SID* algorithm was applied to this data in order to obtain a state-space model, in the form $\mathcal{M}_1(\bullet)$. The substrate concentration variable was not used for identification purposes since its behavior gave rise to much worse models, from the point of view of stability and data fitting. The dimensions of the obtained linear model are $n_s = 5$, $n_i = 2$ and $n_o = 5$ (all of the mentioned above, except the substrate concentration). The sampling interval used is $\Delta t = 0.05\,\text{h}$.

5.2 State estimation

For the current state estimation we have considered an horizon $\mathcal{H}_r$ with dimension $r = 50$. The data for this horizon was generated using the original dynamical model with random error added to the measured variables. This data is different from the one used for identification purposes, but the error has the same characteristics. Not all of the measurements available in $\mathcal{H}_r$ were used. The decision regarding the availability of the measurements was modelled by an independent random binary signal. For the presented results, the number of measurements considered is of 121 out of 250 possible.

The estimated value of $\hat{y}(t_0)$ obtained is:

$$\hat{y}(t_0) = \begin{bmatrix}1.03 & 1.03 & 1.01 & 1.01 & 1.03\end{bmatrix}^{\mathrm{T}},$$

normalised, and the deviation from the value obtained by simulation, in percentage, is

$$\begin{bmatrix}-1.39 & -1.51 & -1.44 & -3.23 & -0.91\end{bmatrix}^{\mathrm{T}}.$$

The profiles of the estimated and measured values of outputs variables x and P, in the horizon $\mathcal{H}_r$, are presented in Figure 1. The difference between all the estimated outputs and all the measurements (both the used and the deleted in the estimation procedure) is presented in Figure 2.

In the estimation step, to obtain $\mathrm{Cov}(e_{\mathcal{X}})$, we assume that the outputs errors are not correlated with each other and that these errors are mainly due to errors in the measurement methods. Under these assumption, the covariance matrix of the error in $\mathcal{H}_r$, $\mathrm{Cov}(e_r)$, is diagonal with all its elements equal to

$$\sigma^2 = 8.33 \times 10^{-4},$$

the variance of the added random error.

5.3 Prediction

For the prediction phase, we have arbitrarily set the specification limits at $LS = 0.85$ and $US = 1.15$, for all output variables. We have considered a prediction horizon, $\mathcal{H}_f$, with dimension $f = 50$.

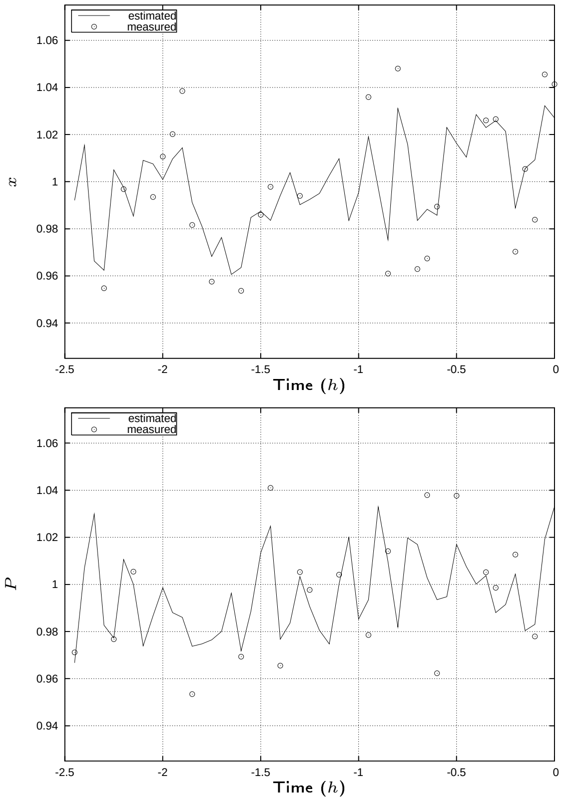

Fig. 1. Real and estimated values of variables x and P in estimation horizon.

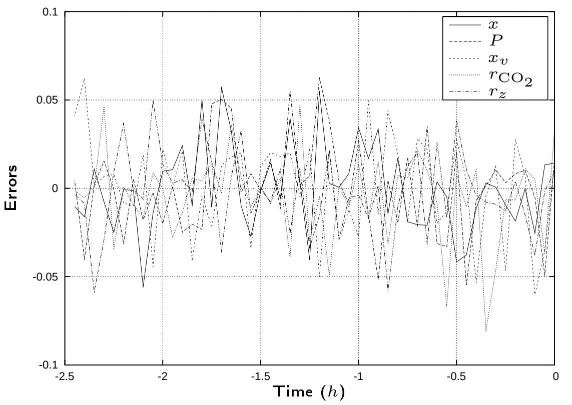

Fig. 2. Estimation errors in the horizon $\mathcal{H}_r$ for all the measured variables.

All the input variables in this horizon are kept at their reference values. The forecasts, for $\mathcal{H}_f$, of all the outputs, their variances and the probability of being out of specifications are presented in Figure 3.

We can see, in Figure 3, than, initially, the variance of the forecasts decreases, due to the the expected decrease of the second term in (12), since the system is stable. As t_k increases, the contribution from the second term becomes dominant, and the variance increases.

In Figure 3 we can see than, without further measurements being made, the probability of all the variables being within their specifications is

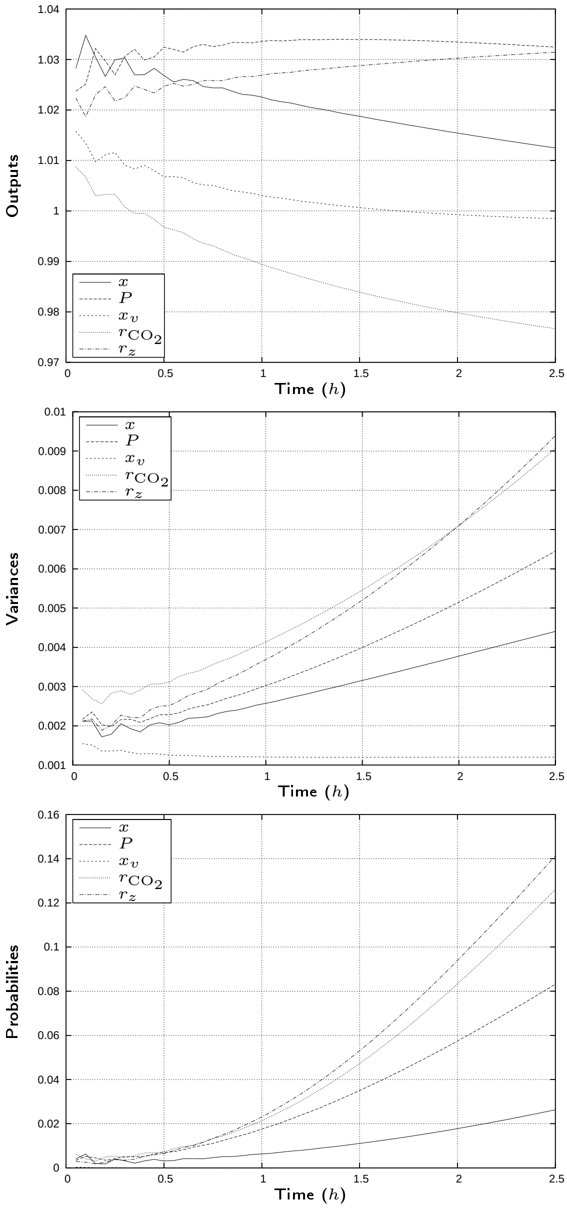

Fig. 3. Forecast of the measured variables, their variance and their probability of being out of specifications for the prediction horizon $\mathcal{H}_f$.

about 98%. From this value it would be reasonable to suppose that no further measurements would be needed up to this time. However, the decision to sample or not would require a greater insight into the system, such as sampling and quality costs.

6. CONCLUSIONS

In this paper we have presented a procedure for the forecast of process measurements and the quantification of their uncertainty. This can be used to predict the probability of a given quality variable being out of its specifications in a future horizon, such as used in model predictive control. The probability can be used to decide if it is possible to take control measures (within the available degrees of freedom) to correct the predicted trajectories, or if instead it is preferable to obtain new information from the process, by performing new types of measurements, in order to decrease the expected operating costs.

This procedure relies on the use of a state-space model which is obtained using process identification techniques. It also includes the estimation of the current process state from the information available in a receding horizon, again using the process model. The problem of missing measurements in the receding horizon is dealt with by considering an outputs vector with variable dimension. Some of the advantages of this procedure are its numerical stability and the capability of dealing with growing or shrinking receding horizons.

The overall procedure can be easily integrated in a Model Predictive Control type strategy, using an objective function that explicitly includes quality and sampling costs.

REFERENCES

Abraham, B., Box G.E.P. (1979). Sampling interval and feedback control. *Technometrics* **21**, 1–8.

Akaike, H. (1972). Information theory and an extension of the maximum likelihood principle. In: *Proc. 2nd. Int. Symp. Information Theory, Supp. to Problems of Control and Information Theory.* pp. 267–281.

Brookner, E. (1998). *Tracking And Kalman Filtering Made Easy.* John Wiley & Sons. New York.

Chmúrny, D., Chmúrny R. (2000). Simulation and control of fermentation complex systems. *Bioprocess Engineering* **23**, 221–227.

Ikonen, E., Najim K. (2002). *Advanced Process Identification and Control.* Marcel Dekker. New York.

Kramer, T. (1989). Process Control from an Economic Point of View. PhD thesis. Univ. Wisconsin–Madison.

Ljung, L. (1999). *System Identification – Theory for the user.* 2nd. ed.. Prentice Hall PTR. New Jersey.

MacGregor, J.F. (1976). Optimal choice of the sampling interval for discrete process control. *Technometrics* **18**, 151–160.

Seppala, C.T. (1998). Dynamic Analysis of Variance Methods for Monitoring Control System Performance. PhD thesis. Queen's University.

Van Overschee, P., DeMoor B. (1994). N4sid: Subspace algorithms for the identification of combined deterministic–stochastic systems. *Automatica* **30**, 75–93.