DYNAMIC VALIDATION OF MODELS: A CASE STUDY ON PLUG-FLOW REACTOR

Aino Ropponen, Kimmo Konkarikoski, Risto Ritala

Tampere University of Technology
The Institute of Measurement and Information Technology

aino.ropponen@tut.fi (Aino Ropponen)

Abstract

Simulators are important components of operational decision support systems (ODSS). By running a simulator in parallel with the real process more accurate information of the process state can be achieved for the ODSS. However, in continuous use the properties of the system may change which may cause parameters to drift and thus cause inaccuracy to the simulator estimates and forecasts. In this paper we introduce a systematical method to update model parameters through reference measurements, such as laboratory analyses, of the output state. Updating is based on Bayesian estimation of parameters. Information about parameters is described as probability densities and thus takes into account also the uncertainty of the information. For linear models parameter updating is a well-known and rather simple operation because all the relevant probability densities are Gaussian and updating is Bayesian tracking of their means and covariance matrices. For nonlinear models the distributions are non-Gaussian which makes the task more complex. In this work we have approximated nonlinear distributions with Gaussian mixture models (GMM). GMM is a linear combination of several Gaussians and enables to describe more complex distributions keeping still the calculations manageable. We demonstrate the method in this paper with a plug-flow reactor which is an example of a dynamic and nonlinear process. We have simulated a bleaching tower process in papermaking as an example of the plug-flow reactor.

Keywords: Uncertainty, Decision support, Bayesian methods, Gaussian mixture model.

Presenting Author's biography

Aino Ropponen. Research scientist and PhD student at Tampere University of Technology, Finland. Master of Science in Technology, graduated in January 2007.

1 Introduction

An application area of simulation with particularly high economic potential is within operator decision support systems (ODSS), enabling dynamic optimization. In decision support the uncertainty of simulator estimates and predictions must be known as decisions hinge both on predicted values and their uncertainty, see e.g. [1]. The challenge in applying simulation models within ODSS is that the model parameter validity and the related uncertainty must be continuously monitored and parameters updated.

Simulation model in an ODSS is a soft sensor which estimates present process state or predicts future state evolution. As the parameters of this soft sensor are uncertain, also the uncertainty of the estimate needs to be determined. If the uncertainty is not considered important component of information in decision making is neglected.

Uncertainties of parameters are described as probability densities see e.g. [2]. By describing the parameter information as probability densities, both the estimates and the uncertainty of the estimates can be assessed.

Model parameters are initialized with history data of the process. As time evolves, the properties of the system may drift due to changes in conditions not included in the model: the model parameters are no longer valid. Knowledge about the system decreases over time and the uncertainty of the parameters increases unless reference measurements about the model output are obtained to update the parameter information. In this paper we assume parameter drifting to be unbiased and thus describe the parameter information degradation as a random walk process.

This paper is organized as follows. In Section 2 we review the general case of dynamic parameter validation method and the update mechanism for the parameters in general. We shall also discuss the degradation of parameter information over time. In Section 3 we present Gaussian mixture model (GMM) to describe nonlinear probability densities and define the method of dynamic validation with GMMs. In Section 4 we apply the method to a simple plug-flow reactor model which is used as an example of bleaching tower in mechanical pulping. The case is nonlinear and dynamic and we show through simulations how the extent, accuracy and frequency of the reference data affect the quality of the model.

2 Dynamic validation

The method of dynamic validation consists of three stages. The output state information is estimated according to the process model always when input data is available. The growth of parameter and state uncertainties between the measurement instants are described as a stochastic process. The parameter information is updated if reference measurement of the output state is available.

Let the process model to be

$$y = F(x,\theta) + \varepsilon , \tag{1}$$

where $y \in R$ is the output, $x \in R^n$ the input data vector and $\theta \in R^m$ the parameter vector. The uncertainty of the model is described by $\varepsilon \sim N\left(0, \sigma_{\text{mod}}^2\right)$. Although model appears static, it covers all finite-impulse response models with proper choice of input variables.

2.1 Diffusion process

In this paper we assume that if no new reference information about the model output y becomes available, the knowledge about the model parameters θ decreases via a random walk or diffusion process

$$\frac{\partial f_\Theta(\theta;t)}{\partial t} = \sum_{i,j}^{p} D_{\Theta,ij} \frac{\partial^2 f_\Theta(\theta;t)}{\partial \theta_i \theta_j} , \tag{2}$$

in which f_Θ is the probability density of the parameter vector, and the symmetric diffusion matrix D determines the speed for the degradation. For Gaussian distributions diffusion process keeps the mean constant but the covariance grows as $\Sigma(t_n) = \Sigma(t_{n-1}) + D(t_n - t_{n-1})$. The choice of D depends on the process. It can be understood either as a tuning parameter or identified from the data.

2.2 Validation of parameter distribution

In the beginning, parameter distribution is determined on the basis of history data and/or expert knowledge of the system. As far as no new information of the output state become available, the information degrades through the diffusion process.

If a reference measurement of the output state becomes available, the parameter distribution can be updated using Bayesian inference [3]

$$\begin{aligned} f_{\Theta|x_{meas},y_{ref}}^{(post)}(\theta \mid x_{meas}, y_{ref}; t_n) \\ = C \cdot f_\Theta^{(ap)}(\theta;t_n) L(\theta \mid x_{meas}, y_{ref}; t_n), \end{aligned} \tag{3}$$

where C is a normalization constant, $f_\Theta^{(ap)}(\theta;t_n)$ is the degraded a priori information of the parameters and

$$\begin{aligned} L(\theta \mid x_{meas}, y_{ref}; t_n) \\ = \left(2\pi\left(\sigma_{\text{mod}}^2 + \sigma_{ref}^2\right)\right)^{-1/2} \exp\left[-\frac{\left(y_{ref} - F(x_{meas},\theta)\right)^2}{2\left(\sigma_{\text{mod}}^2 + \sigma_{ref}^2\right)} \right] \end{aligned} \tag{4}$$

is the likelihood function of the parameter, given the new reference data $f_{Y_{ref}|y}(y_{ref} \mid y;t_n) = N(y_{ref}; y, \sigma_{ref}^2)$.

2.3 Validation of output distribution

For output estimate, information can be updated always when input data measurement x_{meas} is available at t_n. Then the distribution for the output estimate is

$$f_{Y|x_{meas}}^{(est)}(y \mid x_{meas};t_n) = \int_{\Theta} \left(2\pi\sigma_{\mathrm{mod}}^2\right)^{-1/2} \exp\left[-\frac{(y-F(x_{meas},\theta))^2}{2\sigma_{\mathrm{mod}}^2}\right] f_{\Theta}(\theta;t_n)d^m\theta \quad (5)$$

which can be combined with a priori information, and the distribution of the output is

$$f_{Y|x_{meas}}^{(post)}(y \mid x_{meas};t_n) = C_Y \cdot f_{Y|x_{meas}}^{(est)}(y \mid x_{meas};t_n) f_Y^{(ap)}(y;t_n). \quad (6)$$

Similarly to parameter information, we assume the a priori information to be based on earlier output estimates and it to degrade through a scalar diffusion process characterized by diffusion constant D_Y. Hence Eq. (6) introduces a filtering on output predictions – possibly nonlinear and with irregular sampling interval.

If the reference information of the output state is available, it can be combined with the Eq. (6) in Bayesian fashion.

3 Gaussian mixture model

If the model is linear, all the probability densities introduced in Section 2 are Gaussian and the parameter validation is straight forward: updating of probability densities reduces to updating of mean vectors and covariance matrices [2]. If the model is nonlinear the probability densities will not in general be Gaussian. In this paper the probability densities are approximated with Gaussian mixture models (GMM) [4] that are linear combinations of Gaussians:

$$f_X(x \mid \Psi;t_{n-1}) = \sum_{i=1}^{I} c_i N(x \mid \mu_i, \Sigma_i;t_n), \quad (7)$$

where Ψ is a collection of the distribution parameters (c_i, μ_i, Σ_i) and I is the number of Gaussians in probability density approximation. The probability density parameters are identified with Expectation Maximization (EM) algorithm [4, 5] from data or histograms.

GMMs are particularly attractive for parameter validation as their diffusion is solved analytically as

$$f_X(x \mid \psi;t_n) = \sum_{i=1}^{I} c_i N(x \mid \mu_i, \Sigma_i + D(t_n - t_{n-1})). \quad (8)$$

This is shown in the Fig. (1).

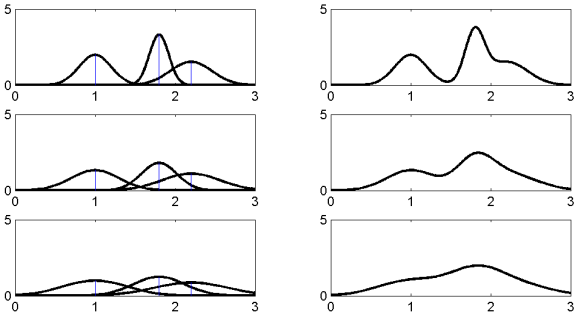

Fig. 1 An example of diffusion for GMM. The original GMM consist of three Gaussians which are shown in the upper figure on the left. The combination of those is shown on the right. The lower figures shows the distribution after time steps t = 1 and t = 2.

Furthermore, if the likelihood and a priori density in Eq. (3) are GMMs, then the posteriori density is also, although of higher order. The growth of the order can be prevented by estimating posterior density with a GMM of the desired order.

3.1 Identification of distribution parameters

Let us assume that the histogram distribution of the variable x is known as

$$f_X^{hist}(x_h) = f_X(x_h)V(x_h), \quad (9)$$

where x_h is the center of histogram bar and $V(x_h)$ the volume of it. This distribution is approximated using Gaussian mixture model. Kullback-Leibler (KL) divergence [4] is used as a measure of the difference between these distributions. By minimizing the KL-divergence, the distance of the real and the approximated distribution is minimized

$$\min_{\Psi} D_{KL}\left(f_X^{hist}(x_h) \,\|\, f_X(x,\Psi)\right) = \min_{\Psi} \sum_h f_X^{hist}(x_h) \log\left[\frac{f_X^{hist}(x_h)}{f_X(x,\Psi)}\right] \quad (10)$$

$$\Leftrightarrow \max_{\Psi} \sum_h f_X^{hist}(x_h) \log f_X(x,\Psi).$$

This can be solved using the well-known EM-algorithm [4, 5]. The target of the EM-algorithm is to find optimal values for the parameter vector Ψ of GMM, Eq. (7).

In the beginning of the EM-algorithm the initial values of parameters of the Gaussians are chosen, e.g. randomly. Then in E-step the probability p_{kj} that data point x_j belongs to the Gaussian k is

$$p_{kj}^{(+)} = \frac{c_k^{(-)} N\left(x_j \mid \mu_k^{(-)}, \Sigma_k^{(-)}\right)}{\sum_{i=0}^{I} c_i^{(-)} N\left(x_j \mid \mu_i^{(-)}, \Sigma_i^{(-)}\right)}, \quad (11)$$

where $k \in I$. The superscript (-) indicate values calculated in the previous iteration step and (+) the new values.

In M-step the parameters are updated using the probabilities calculated in E-step

$$
\begin{aligned}
c_k^{(+)} &= \frac{1}{N}\sum_{j=1}^{N} p_{kj}^{(+)} \\
\mu_k^{(+)} &= \frac{\sum_{j=1}^{N} p_{kj}^{(+)} x_j}{\sum_{j=1}^{N} p_{kj}^{(+)}} \\
\Sigma_k^{(+)} &= \frac{\sum_{j=1}^{N} p_{kj}^{(+)} \left(x_j - \mu_k^{(+)}\right)\left(x_j - \mu_k^{(+)}\right)^T}{\sum_{j=1}^{N} p_{kj}^{(+)}}.
\end{aligned}
\tag{12}
$$

These steps alternate until convergence to the optimal values.

The growth of the GMM order when combining information in Eq. (3) and (6) can be prevented by using direct histogram-based algorithm instead of the regular EM-algorithm. In direct histogram-based algorithm the center of the each histogram bins are the means for the Gaussians in GMM (Fig. 2). The height of the histogram defines the height of the Gaussian. The covariance matrix is chosen by a user between the means of the Gaussians. This method is not as accurate as the EM-algorithm, and the number of Gaussians is larger, but it requires less computation and the order is not growing.

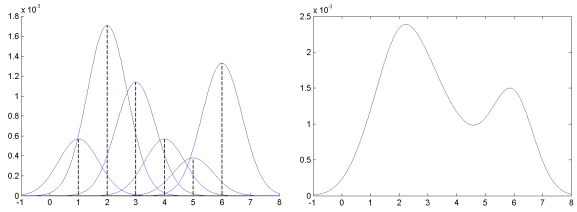

Fig. 2 An example of histogram based method for GMM. The dotted line represents the center of the histogram bar. The GMM of the histogram is shown in the right.

3.2 Validation of parameter distribution

If the reference measurement of the output state is available, the likelihood function of the parameters (Eq. 4) can be identified and approximate with the GMM as

$$
f_\Theta^{(ref)}\left(\theta \mid \Psi_{ref}; t_n\right) = \sum_{j=1}^{J} d_j N\left(\theta \mid \nu_j, \Gamma_j\right). \tag{13}
$$

By combining that with a priori information

$$
f_\Theta^{(ap)}\left(\theta \mid \Psi_{ap}; t_n\right) = \sum_{i=1}^{I} c_i N\left(\theta \mid \mu_i, \Sigma_i\right) \tag{14}
$$

we get a posteriori distribution as

$$
\begin{aligned}
&f_\Theta^{(post)}\left(\theta \mid \Psi_{post}; t_n\right) \\
&\quad = C \cdot f_\Theta^{(ap)}\left(\theta \mid \Psi_{ap}; t_n\right) f_\Theta^{(ref)}\left(\theta \mid \Psi_{ref}; t_n\right) \\
&\quad = C \cdot \sum_{i=1}^{I} c_i N\left(\theta \mid \mu_i, \Sigma_i\right) \sum_{j=1}^{J} d_j N\left(\theta \mid \nu_j, \Gamma_j\right).
\end{aligned}
\tag{15}
$$

As a product of two Gaussians is a Gaussian form the a posteriori distribution is also a Gaussian mixture and can thus be written as

$$
f_\Theta^{(post)}\left(\theta \mid \Psi_{post}; t_n\right) = \sum_{i=1}^{I}\sum_{j=1}^{J} \tilde{\tilde{c}}_{ij} N\left(\theta \mid \tilde{\mu}_{ij}, \tilde{\Sigma}_{ij}\right), \tag{16}
$$

where the parameters are related through

$$
\begin{aligned}
\tilde{c}_{ij} &= c_i d_j (2\pi)^{-d/2} \det\left(\Sigma_i + \Gamma_j\right)^{-1/2} \\
&\quad \cdot \exp\left(-\frac{1}{2}\left(\mu_i - \nu_j\right)^T \left(\Sigma_i + \Gamma_j\right)^{-1}\left(\mu_i - \nu_j\right)\right) \\
\tilde{\mu}_{ij} &= \tilde{\Sigma}_{i,j}\left(\Sigma_i^{-1}\mu_i + \Gamma_j^{-1}\nu_j\right) \\
\tilde{\Sigma}_{ij}^{-1} &= \Sigma_i^{-1} + \Gamma_j^{-1}
\end{aligned}
\tag{17}
$$

and the probability density is properly normalized by choosing

$$
\tilde{\tilde{c}}_{ij} = \frac{\tilde{c}_{ij}}{\sum_{i=1}^{I}\sum_{j=1}^{J} \tilde{c}_{ij}}. \tag{18}
$$

If the means are fixed, i.e. the distribution parameters are identified using histogram-based algorithm, only the heights and covariance matrices need to be calculated. When regular GMM is applied, the expression (Eq. 16) is truncated to prevent increase of GMM order. The truncation may be based on full reduced-order GMM identification or elimination of terms by minimizing the KL-distance between the full and truncated probability densities.

4 Case: plug-flow reactor

We demonstrate the method of dynamic validation with a simulated plug-flow reactor model. The plug-flow reactor is a pipe where the flow is through the reactor. It assumed that there is no mixing in the axial direction or in the radial direction. The chemical reaction in the pipe is of type A+B → C and that continues throughout the reactor.

Let us denote the position in the flow direction in the reactor by x. The chemical reactions at each position are described with simple kinetic equations

$$
\begin{aligned}
\frac{dc_A(x)}{dt} &= -k_1 c_A(x) c_B(x) \\
\frac{dc_B(x)}{dt} &= -k_2 c_A(x) c_B(x) \\
\frac{dc_C(x)}{dt} &= k_1 c_A(x) c_B(x)
\end{aligned}
\tag{19}
$$

where $c_i(x)$ are the amounts of the components at position x. The model parameters of interest are the kinetic coefficients k_1 and k_2.

The movement of material is described by writing the mass balance in the form

$$\int_{t-\tau(t)}^{t} f_{tot}(t')dt' = V\,, \qquad (20)$$

where f_{tot} is the total volumetric flow and $\tau(t)$ is the residence time of material coming out of the reactor.

4.1 Bleaching tower in mechanical pulping

The plug-flow reactor model can be applied e.g. to describe the pulp bleaching tower in the pressurized ground wood (PGW) process in papermaking In the PGW bleaching water, fiber (component A) and bleaching chemical (B) are fed into the reactor. In the tower the fiber and the chemical react increasing pulp brightness, but at the same time produce organic substances (C). Typically, such organic substances exist also in the inflow. In the ODSS the target is to estimate the amount of the total organic carbon (TOC) after the bleaching. The TOC is of interest because high levels cause disturbances to the process.

TOC-concentration cannot be measured online. Laboratory measurements of TOC are expensive and available infrequently. By estimating TOC using online measurements of other variables the expensive laboratory measurements are needed rarely.

4.2 Behavior of the reactor

The behavior of the reactor is demonstrated by a simulation with constant parameters. In this simulation the input flow measurements are assumed to be exact.

Before the simulation starts, the reactor inflow has been 0.97 volume units/time unit (VU/TU) of water, 0.03 VU/TU fiber and 0.001 VU/TU total organic carbon (TOC). No chemical has been fed to the reactor prior to the zero time of the simulation. The volume of the reactor is 30 VU. The model parameters are $k_1 = 0.01$ 1/(TU·CVU) and $k_2 = 2$ 1/(TU·VU). The simulation runs 500 TU.

Three actions are made on the bleaching tower during the simulation period. At $t = 0$ chemical flow of 1 chemical volume units/time unit (CVU/TU) is introduced into the inflow while all the other conditions the same. At $t = 250$ chemical flow is increased to 2 CVU/TU, and at $t = 450$ inflow is increased by 2/0.97. The amount of TOC out of the reactor is shown in the Fig. 3.

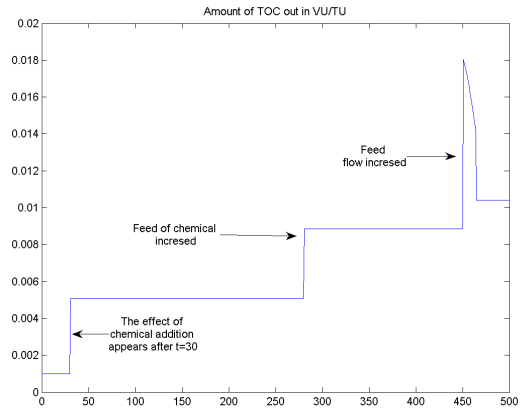

Fig. 3 Example of the PGW-simulator. The amount of TOC out of the reactor in VU/TU when parameters are kept constant $k_1 = 0.01$ 1/(TU·CVU) and $k_2 = 2$ 1/(TU·VU).

4.3 Simulation with drifting parameters

Fig. 4 represents the validation method for PGW-simulation when the parameters are not constant.

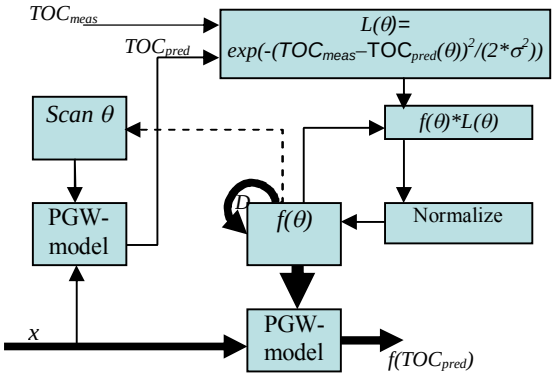

Fig. 4 Scheme of parameter validation method. Thick lines indicate operations when input measurements become available. Thin lines indicate operations when reference output measurements become available.

When online input measurements of the system become available (thick lines), output estimate is updated using the input measurements and process model, and the uncertainty of the parameter probability density is increased through the diffusion process. When the output reference measurement becomes available (thin lines), the likelihood function is calculated and combined with the a priori probability density.

It turns out that the parameter probability density cannot be properly updated using only the reference measurement of the output TOC, but also the reference measurement of the outcoming chemical concentration is needed. If the chemical reference measurement is not used, the parameter distribution spreads which make the parameter and thus the output

estimates useless. Fig. 5 and Fig. 6 show the parameter distribution with and without the reference measurement of the outcoming chemical concentration.

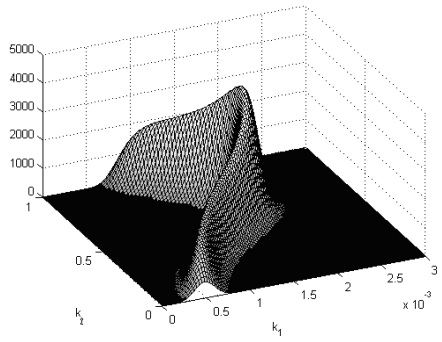

Fig. 5 The parameter distribution without the reference chemical concentration measurement.

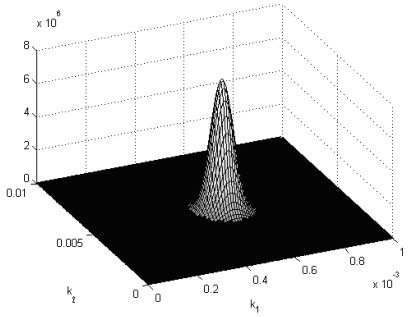

Fig. 6 The parameter distribution with the both chemical and TOC reference.

5 Results

We studied the general method with the plug-flow reactor by having one simulation model to represent the true behavior with known parameter values, drifting over time. Another simulator was acting as that in ODSS. It had originally correct parameters and a given parameter uncertainty, but the ODSS simulator was able to track the parameter drift only through reference measurements of TOC and chemical concentrations, and parameter updating of Section 2.

The parameter uncertainties for the simulation model were set to $\sigma_{k1} = 2{,}1 \cdot 10^{-5}$ and $\sigma_{k2} = 2{,}1 \cdot 10^{-4}$. The diffusion matrix for the GMM in ODSS was $D_\Theta = \begin{bmatrix} 6 \cdot 10^{-10} & 0 \\ 0 & 4{,}16 \cdot 10^{-8} \end{bmatrix}$.

The model uncertainty was set to $\sigma_{\mathrm{mod}} = 0.0005$ kg/s. The uncertainty for the reference measurement of TOC was $\sigma_{TOC} = 50$ mg/l and for chemical reference $\sigma_{chem} = 0.02$ kg/s. The diffusion coefficient used for the output estimate in ODSS was $D_Y = 10^{-7}$.

Fig. 7 shows the effect of the parameter uncertainty on estimated value of the TOC-concentration, predicted at each point with earlier model input history. The upper figure illustrates the true (black) and the estimated (grey) TOC-concentration at the output and the lower figure the uncertainty of the prediction due to parameter uncertainty described as standard deviation of the output GMM. Black circles represent the instants when the output reference values are available. Uncertainty increases between the reference measurements, but decreases discontinuously with a new reference measurement.

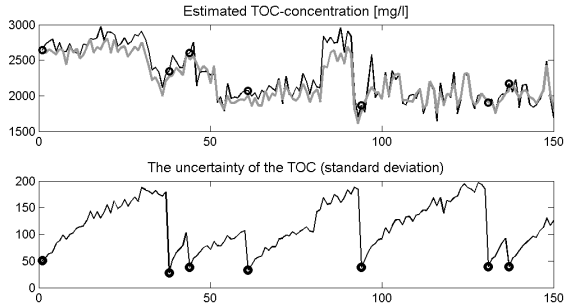

Fig. 7 Top: the true (black) and the estimated (grey) concentration of component C in output. Bottom: the uncertainty of the estimate as the standard deviation. Black circles show points of the reference measurement.

Fig. 8 shows the expectation value and the GMM variance of the parameter k_1. The true value is showed by black and the estimated value by grey. Parameter estimates are constant between the reference measurements (circles). The parameter k_2 is updated similarly (Fig. 9).

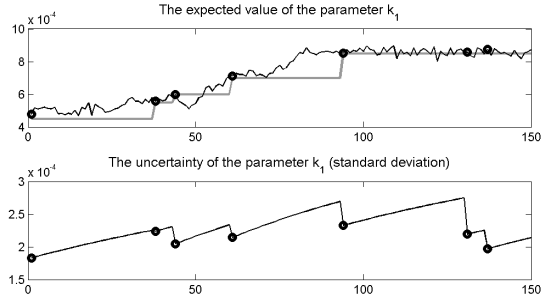

Fig.8 The expected value (top) and the uncertainty (bottom) of the parameter k_1.

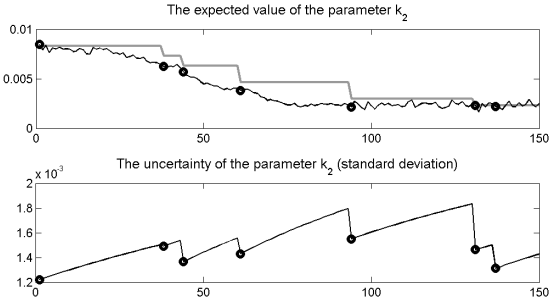

Fig. 9 The expected value (top) and the uncertainty (bottom) of the parameter k_2.

6 Conclusions

This paper reports a practical dynamic parameter validation method for nonlinear dynamic models. The validation is based on Bayesian estimation and information is described as probability densities. In this paper the parameter probability densities have been approximated with GMMs. The method is demonstrated in a case of simulated plug-flow reactor.

In this study the uncertainty of the process model as well as the uncertainty of the reference measurements has been taking into account. As a consequence the parameter information and output state estimate became uncertain. However the uncertainty of the input measurements has not been taking into account. This does not change the basic idea of dynamic validation, but keep the method simpler. If both the uncertainty of the input and the parameters are taken into account the output state estimate would not be Gaussian any longer.

References

[1] H. Jokinen, K. Latva-Käyrä, P. Pulkkinen and R. Ritala. Modelling of uncertainty: case studies on operation of papermaking process. In I. Troch and F. Breitenecker, editors, Proceedings 5th MATHMOD Conference, Symposium on Mathematical Modelling, pp. 4-1..4-10, Vienna, Austria, February 8.-10 2006.

[2] K. Latva-Käyrä and R. Ritala. Dynamic validation of multivariate linear soft sensors with reference laboratory measurements. In: J. Kappen, J. Manninen and R. Ritala, editors, Proceedings of Model Validation Workshop, COST ACTION E36, Pages 57-64, Espoo, Finland, October 6 2005. VTT Symposium 238.

[3] e.g. A. O'Hagan and J. Forster. Bayesian Inference, second ed., Kendall's Advanced Theory of Statistics. Arnold. 2004.

[4] C.M. Bishop. Pattern Recognition and Machine Learning p. 430-455. Springer. 2006.

[5] A.P. Dempster, N.M. Laird and D.B. Rubin. Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society, Series B 39:1-38,* 1977.

ISBN 978-3-901608-32-2