Application of Varying Parameters Modelling with Gaussian Processes ⋆

Juš Kocijan * **Kristjan Ažman** **

* *Jozef Stefan Institute, Ljubljana, Slovenia and University of Nova Gorica, Nova Gorica, Slovenia (e-mail: jus.kocijan@ijs.si).*
** *Elektrina, d.o.o., Mengeš, Slovenia (e-mail: kristjan.azman@elektrina.si)*

Abstract: This paper describes an application of the method for modelling nonlinear dynamic systems from measurement data. The method merges the linear local model blending approach in the velocity-based linearisation form with Bayesian Gaussian process modelling. The new Fixed-Structure Gaussian Process model has a predetermined linear model structure with varying and probabilistic parameters represented by Gaussian process models. These models have several advantages for the modelling of local model parameters as they give us adequate results, even with small data sets. Furthermore, they provide a measure of confidence in the prediction of the varying parameters and information about the dependence of the parameters on individual inputs. The Fixed-Structure Gaussian Process model can be used for the extended local linear equivalence class of nonlinear systems. The modelling method is applied to modelling of a semi-industrial gas-liquid plant that exhibits variable dynamics depending on its operating region.

Keywords: Nonlinear models, identification, Gaussian processes, linear parameter varying model.

1. INTRODUCTION

Modelling with local model networks (LMNs) is an attractive method for modelling nonlinear dynamic systems from data for the so-called *divide and conquer* control design (Murray-Smith and Johansen (1997)). In this approach the global behaviour is represented by a network of simple local models, where each local model describes some particular operating region and the global behaviour is realised by blending the dynamics of the local models. Important issues with this approach are the realisation of the model blending (Leith and Leithead (1999); Murray-Smith *et al.* (1999); Johansen *et al.* (2000)) and the issue of scheduling-vector selection. The scheduling vector is the vector defining the current operating region and assists the blending mechanism to accurately match the nonlinear dynamics.

One way to deal with the issue of accurately modelling off-equilibrium behaviour based on local linear models is representing the nonlinear system with velocity-based linearisation (VBL), see Leith and Leithead (1999). VBL, in contrast to the conventional Taylor-series-expansion approach, enables the representation of the system at every operating point and not just in the equilibrium regions. Nevertheless, the blending and scheduling mechanisms still need to be determined.

The accurate representation of nonlinear dynamics with locally valid linear models is important from the control-design point of view, because local controllers can be designed for each of the corresponding local models and blended to local controller network or blended gain-scheduled controller (e.g., Kocijan *et al.* (2002)).

In this paper a parametric approach with a fixed linear model structure and varying parameters, i.e., a linear parameter varying model or LPV model, based on the Gaussian process (GP) models, called the Fixed-Structure Gaussian process (FSGP) model (Ažman and Kocijan (2009)) is used for modelling a nonlinear process. The FSGP model is a model with predetermined linear structure with varying and probabilistic parameters represented by Gaussian process models. As such, the FSGP model is, from the application point of view (e.g. the design), more transparent than other known nonparametric models of dynamic systems based on Gaussian processes (Ažman and Kocijan (2007); Kocijan *et al.* (2005)). It is the information about the model's structure that distinguishes the FSGP model from other GP-based models. The idea of approximating the functional dependence of varying parameters is not new. An approach using Radial Basis Function neural networks can be found in Peng *et al.* (2003). The novelty of the proposed model is that the varying parameters are represented by Gaussian process models, which brings benefits in comparison with other GP-based approaches and other LPV approaches.

There are several reasons to use GP models for modelling the varying parameters of the LPV model:

- They tend to achieve acceptable modelling results, even with relatively small training data sets.

⋆ The authors gratefully acknowledge the contribution of the Slovenian Research Agency, Grant No. P2-0001.

- Besides predicting the output the GP model also gives a measure of the prediction confidence, dependent on the density of the training data.
- When local models are blended, with GP models predicting the local models' parameters, the GP models also provide information about the dependence of the parameters on individual regressors.

The FSGP model therefore addresses problems such as unknown structure due to the non-parametricity of the model, present at regular GP models, the confidence of the varying parameters' predictions and the small number of data for the identification of these varying parameters. The paper of Gregorčič and Lightbody (2005) describes an alternative approach to the one suggested in this paper. Their method merges LMNs and conventionally blended local linear GP models, but without using VBL and its benefits, and consequently with an additional optimisation of the blending functions.

The FSGP approach (Ažman and Kocijan (2009)) can be used as a simple but still effective and (in terms of prediction) potentially fast engineering tool. As the model is parametric, it can be used for a wider range of control-design methods, and not only model-based predictive control, e.g., gain-scheduling control (Ažman and Kocijan (2006)).

The paper is organized as follows. In Section 2 the fundamentals of GP modelling are presented. The FSGP modelling approach is described in Section 3 and modelling case study, namely gas-liquid separator, is given in Section 4. The last section emphasises the main results and concludes the paper.

2. MODELLING WITH GAUSSIAN PROCESSES

The Gaussian process model is an example of the use of a flexible, probabilistic, non-parametric model with uncertainty predictions. It fits naturally in the Bayesian modelling framework in which instead of parameterising mapping function $f(\mathbf{x})$, a prior is placed directly on the space of possible functions $f(\mathbf{x})$ which could represent the nonlinear mapping from input vector $\mathbf{x}$ to output y. Its use and properties for modelling are reviewed in Rasmussen and Williams (2006).

A Gaussian process is a Gaussian random function, fully described by its mean and variance. Gaussian processes can be viewed as a collection of random variables $f(\mathbf{x}_i)$ with joint multivariate Gaussian distribution: $f(\mathbf{x}_1), \ldots, f(\mathbf{x}_n) \approx \mathcal{N}(0, \mathbf{\Sigma})$, where $\mathcal{N}(\cdot, \cdot)$ stands for Gaussian distribution determined with mean value and variance, where covariance matrix element $\Sigma_{pq} = \mathrm{Cov}(y_p, y_q) = C(\mathbf{x}_p, \mathbf{x}_q)$ gives the covariance between values of the functions $y_p = f(\mathbf{x}_p)$ and $y_q = f(\mathbf{x}_p)$. Thus, the mean function (usually assumed to be zero) and the covariance function fully specify the Gaussian process. Note that the covariance function $C(\cdot, \cdot)$ can be any function having the property of generating a positive definite covariance matrix.

A common choice is

$$C(\mathbf{x}_p, \mathbf{x}_q) \;=\; v_1 \exp\left[-\frac{1}{2}\sum_{d=1}^{D} w_d (x_p^d - x_q^d)^2\right] + \delta_{pq} v_0, \quad (1)$$

where $\mathbf{\Theta} = [w_1, \ldots, w_D, v_1, v_0]^T$ are the 'hyperparameters' of the covariance functions, v_0 is the estimated noise variance, v_1 is the estimate of the vertical scale of variation, D is the input dimension, x_p^d and x_q^d are d^{th} components of input vectors $\mathbf{x}_p, \mathbf{x}_q$ and $\delta_{pq} = 1$ if $p = q$ and 0 otherwise. The covariance function (1) is composed of two parts: the Gaussian covariance function for the modelling of system function and the covariance function for the modelling of noise. The noise, in our case, is presumed to be white. Other forms of covariance functions suitable for different applications can be found in Rasmussen and Williams (2006).

For a given problem, the parameters in $\mathbf{\Theta}$ are learned (identified) using the data at hand. After the learning, one can use the w parameters as importance indicators of the corresponding input components (dimensions) are: if w_d is zero or near zero it means that the input in dimension d contain little information and could possibly be removed.

Consider a set of N D-dimensional input vectors $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N]$ and a vector of output data $\mathbf{y} = [y1, y2, \ldots, y^N]^T$. Based on the data $(\mathbf{X}, \mathbf{y})$, and given a new input vector $\mathbf{x}^*$, we wish to find the predictive distribution of the corresponding output y^*. Unlike other models, there is no model parameter determination as such, within a fixed model structure. With this model, most of the effort consists of *tuning* the parameters of the covariance function. This is done by maximisation of the log-likelihood

$$\begin{aligned}\mathcal{L}(\mathbf{\Theta}) &= \log(p(\mathbf{y}|\mathbf{X})) \\ &= -\frac{1}{2}\log(|\,\mathbf{K}\,|) - \frac{1}{2}\mathbf{y}^T\mathbf{K}^{-1}\mathbf{y} - \frac{N}{2}\log(2\pi) \qquad (2)\end{aligned}$$

where $\mathbf{\Theta}$ is the vector of hyperparameters and $\mathbf{K}$ is the $N \times N$ training covariance matrix. The calculation of the log-likelihood and its derivatives due to the optimisation algorithm involves the computation of the inverse of the $N \times N$ covariance matrix $\mathbf{K}$ at every iteration, which can become computationally demanding for large N. Nevertheless, the number of parameters to be optimised is small ($D + 2$, see Equation (1)), which means that optimisation convergence might be faster and that the 'curse of dimensionality' so common to black-box identification methods is circumvented or at least decreased.

The described approach can be easily utilised for regression calculation. Based on training set $\mathbf{X}$ a covariance matrix $\mathbf{K}$ of size $N \times N$ is determined. As already mentioned, the aim is to find the distribution of the corresponding output y^* at some new input vector $\mathbf{x}^* = [x_1(N+1), x_2(N+1), \ldots, x_D(N+1)]^T$.

For a new test input $\mathbf{x}^*$, the predictive distribution of the corresponding output y^*, over cases in the training set $(\mathbf{X}, \mathbf{y})$ is $y^*|(\mathbf{X}, \mathbf{y}), \mathbf{x}^*$ and is Gaussian, with mean and variance

$$\begin{aligned}\mu(\mathbf{x}^*) \;&=\; \mathbf{k}(\mathbf{x}^*)^T\ \mathbf{K}^{-1}\ \mathbf{y}, \qquad (3)\\ \sigma^2(\mathbf{x}^*) \;&=\; \kappa(\mathbf{x}^*) \;-\; \mathbf{k}(\mathbf{x}^*)^T\ \mathbf{K}^{-1}\ \mathbf{k}(\mathbf{x}^*),\end{aligned}$$

where $\mathbf{k}(\mathbf{x}^*) = [C(\mathbf{x}^1, \mathbf{x}^*), \ldots, C(\mathbf{x}^N, \mathbf{x}^*)]^T$ is the $N \times 1$ vector of covariances between the test and training cases, and $\kappa(\mathbf{x}^*) = C(\mathbf{x}^*, \mathbf{x}^*)$ is the covariance between the test input and itself. Vector $\mathbf{k}(\mathbf{x}^*)^T\ \mathbf{K}^{-1}$ in (3) can be interpreted as a vector of *smoothing* terms which weights

the training outputs $\mathbf{y}$ to make a prediction at the test point $\mathbf{x}^*$.

Reasons to select modelling with Gaussian process models are small amounts of data relative to the number of selected regressors, data corrupted with noise and measurement errors and the need for the measure of model prediction confidence. If there is not enough data or it is heavily corrupted with noise, even the Gaussian process model cannot perform well, but in that case the inadequacy of the model and the identification data is indicated through higher variance of the predictions.

3. MODELLING THE NONLINEAR SYSTEM

FSGP approach to modelling nonlinear system that is a combination of velocity-based linearisation and modelling with Gaussian processes is briefly introduced in this section. More details and description of FSGP for discrete time systems is given in Ažman and Kocijan (2009).

Consider the delayed nonlinear system

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \mathbf{F}(\mathbf{x}(t), \mathbf{u}(t)),\\ \mathbf{y}(t) &= \mathbf{G}(\mathbf{x}(t), \mathbf{u}(t)),\end{aligned} \tag{4}$$

which may be reformulated, without loss of generality, in form denoted as extended local linear equivalence (ELLE)(Leith and Leithead (1999))

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) + \mathbf{f}(\boldsymbol{\rho}),\\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) + \mathbf{g}(\boldsymbol{\rho}),\end{aligned} \tag{5}$$

where $\mathbf{x}(t) \in \mathcal{R}^n$, $\mathbf{u}(t) \in \mathcal{R}^m$, $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ are appropriately dimensioned constant matrices, $\mathbf{f}(\cdot)$ and $\mathbf{g}(\cdot)$ are nonlinear functions and $\boldsymbol{\rho} = \boldsymbol{\rho}(\mathbf{x}(t), \mathbf{u}(t)) \in \mathcal{R}^q, q \leq m+n$, embodies the nonlinear dependence of the dynamics on the state and input with $\nabla_x \boldsymbol{\rho}$, $\nabla_u \boldsymbol{\rho}$ constant (Leith and Leithead (1999)). Differentiating (5) an alternative representation of the nonlinear system is

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \mathbf{w}(t),\\ \dot{\mathbf{w}}(t) &= \mathbf{A}(\boldsymbol{\rho})\mathbf{w}(t) + \mathbf{B}(\boldsymbol{\rho})\dot{\mathbf{u}}(t),\\ \dot{\mathbf{y}}(t) &= \mathbf{C}(\boldsymbol{\rho})\mathbf{w}(t) + \mathbf{D}(\boldsymbol{\rho})\dot{\mathbf{u}}(t).\end{aligned} \tag{6}$$

where

$$\begin{aligned}\mathbf{A}(\boldsymbol{\rho}) &= \mathbf{A} + \nabla\mathbf{f}(\boldsymbol{\rho})\nabla_{\mathbf{x}}\boldsymbol{\rho}, \quad \mathbf{B}(\boldsymbol{\rho}) = \mathbf{B} + \nabla\mathbf{f}(\boldsymbol{\rho})\nabla_{\mathbf{u}}\boldsymbol{\rho},\\ \mathbf{C}(\boldsymbol{\rho}) &= \mathbf{C} + \nabla\mathbf{g}(\boldsymbol{\rho})\nabla_{\mathbf{x}}\boldsymbol{\rho}, \quad \mathbf{D}(\boldsymbol{\rho}) = \mathbf{D} + \nabla\mathbf{g}(\boldsymbol{\rho})\nabla_{\mathbf{u}}\boldsymbol{\rho},\end{aligned} \tag{7}$$

where

$$\nabla_{\mathbf{x}}\mathbf{F}(\mathbf{x}(t), \mathbf{u}(t)) = \mathbf{A}(\boldsymbol{\rho}) \approx \boldsymbol{\alpha} \tag{8}$$

$$\nabla_{\mathbf{u}}\mathbf{F}(\mathbf{x}(t), \mathbf{u}(t)) = \mathbf{B}(\boldsymbol{\rho}) \approx \boldsymbol{\beta} \tag{9}$$

$$\nabla_{\mathbf{x}}\mathbf{G}(\mathbf{x}(t), \mathbf{u}(t)) = \mathbf{C}(\boldsymbol{\rho}) \approx \boldsymbol{\gamma} \tag{10}$$

$$\nabla_{\mathbf{u}}\mathbf{G}(\mathbf{x}(t), \mathbf{u}(t)) = \mathbf{D}(\boldsymbol{\rho}) \approx \boldsymbol{\delta} \tag{11}$$

Elements of matrices $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\delta}$ are modelled with Gaussian process models, which are learned from data. The learning data are coefficients of individual linearisations obtained from local linear models identified in various equilibrium as well as off-equilibrium regions.

This paper will not focus on details of how the local models are obtained. More on the identification of the the local models and the associated issues can be found e.g. in Murray-Smith and Johansen (1997), Murray-Smith *et al.* (1999) and Johansen *et al.* (2000) and references in there. Nevertheless, the identified linear local models need to be of the same order, must describe corresponding region satisfactorily well and must be located in the region of interest.

It is important to note that a local linear input-output model only specifies parameters up to a co-ordinate transformation (Leith and Leithead (2002)).

The FSGP system modelling consist roughly of two stages. The first stage is the identification of local linear models in equilibrium and off-equilibrium points. The results of the first stage are sets of derivatives of nonlinear function which are at the same time coefficients of linear local models.

In the second stage GP models are learned for each set of identified coefficients/parameters. Each set has a functional dependency on system states and inputs. GP model can be used as a relevance detector, i.e., via GP model training a necessary regressors, namely states and inputs to which parameters are functionally linked, are revealed.

As the local models usually can not be identified for every point in operating region the smoothing property of GP models can be used to acquire the values of the local model parameters for the operating points lying between the points where linear local models were identified. This blending of the parameters is in our case realized with modelling each local model parameter with own GP model. The input into GP models is the vector of regressors, which can be interpreted as vector of scheduling variables $\boldsymbol{\rho}(t)$. The output of the particular GP models is corresponding most likely value of the parameter θ_i and associated variance var θ_i, where index i denotes i-th parameter from vector of parameters $\boldsymbol{\theta} = \left[[\alpha_{ij}]^T [\beta_{ij}]^T [\gamma_{ij}]^T [\delta_{ij}]^T \right]^T$. Variances var θ_i are expressing the confidence of the predicted values of the parameters. The confidence depends on the amount of information available for modelling. Better the region is modelled – with more local models – lower is the variance.

The nonlinear system model is realised using velocity based linearisation (Leith and Leithead (1999)). The realisation in Leith and Leithead (1999) has been originally developed as a framework for continuous systems, but it can be extended to discrete systems (Nyström *et al.* (2002)).

The obtained nonlinear model can be viewed as a parametric model with probabilistic and variable parameters $\boldsymbol{\theta}(\boldsymbol{\rho}(t))$ – FSGP model. Each parameter depends on the vector of scheduling variables $\boldsymbol{\rho}(t)$, which can consists of all states and inputs or a subset of them.

4. CASE STUDY

4.1 Process Plant Description

The semi-industrial process plant used for the case study in the paper is the unit for separating gas from liquid that forms the part of a larger pilot plant positioned at Department of systems and control, Jozef Stefan Institute. The screen shot from SCADA presenting the scheme of this plant is given in Figure 1.

The role of the separation unit is to capture flue gases under low pressure from the effluent channels by means of water flow, to cool them down and then supply them under high-enough pressure to other parts of the pilot plant.

The flue gases coming from the effluent channels are absorbed by the water flow into the water circulation pipe through injector.

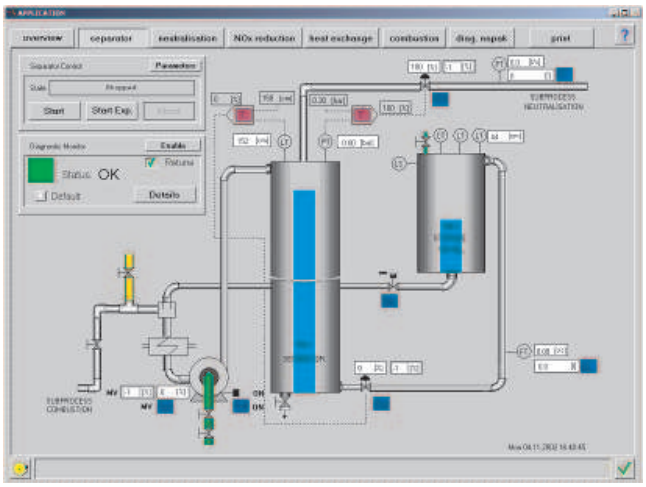

Fig. 1. SCADA scheme of the gas-liquid separator

The water flow is generated by the water ring pump. The speed of the pump is kept constant. The pump feeds the mixture of water and gas into the tank, where gas is separated from water. Hence the accumulated gas in tank forms a sort of 'gas cushion' with increased internal pressure. Owing to this pressure, the flue gas is blown out from tank into the neutralization unit. On the other side, the 'cushion' forces water to circulate back to the reservoir. The quantity of water in the circuit is constant.

In order to understand the basic relations among variables and to illustrate the nonlinearity of the process a mathematical model is introduced. The gas-liquid separation pressure sub-system of interest can be described by a set of two equations.

$$\begin{aligned}\frac{dp_1}{dt} &= \frac{1}{S_1(h_{T_1}-h_1)}\Big(p_0\left(\alpha_0+\alpha_1 p_1+\alpha_2 p_1^2-k_1 R_1^{u_1-1}\sqrt{p_1}\right)\\ &\quad+(p_0+p_1)\left(\Phi_w-k_2 R_2^{u_2-1}\sqrt{p_1+k_w(h_1-h_{T_2})}\right)\Big),\\ \frac{dh_1}{dt} &= \frac{1}{S_1}\left(\Phi_w-k_2 R_2^{u_2-1}\sqrt{p_1+k_w(h_1-h_{T_2})}\right),\end{aligned} \tag{12}$$

where u_i is the command signal of valve Vi, $i=1,2$, where V1 is the valve on output from tank R4.1 to other subprocess and V2 is the valve between tanks R4.2 and R4.1, h_i is the level in tank R4.i, $i=1,2$, p_1 is the relative air pressure in tank R4.1, S_i is the section area of tank R4.i, p_0 is atmospheric pressure, h_{Ti} is height of tank R4.i, $i=1,2$, R_i is the open-close flow ratio of valve Vi, $i=1,2$, k_i is the flow coefficient of valve Vi, $i=1,2$, Φ_w is the known constant water flow through pump, $\alpha_i; i=1,2,3$ are constant parameters.

From the model presented, it can be deduced that the nonlinear process is of a multivariable nature (two inputs and two outputs with dynamic interactions between the channels). In our case a level feedback control was already implemented. Consequently the dynamic system could be approached as a single-input single-output dynamic system with the command signal of valve V_1 as the input and the pressure in tank R4.1 as the output. As can be seen from Eqs. (12) pressure is nonlinearly related to level h_1 and valve command signal u_1 which results in different dynamic behaviour depending on the operating region.

User-friendly experimentation with the process plant is enabled through interface with the Matlab/Simulink environment. This interface enables PLC access with the Matlab/Simulink using OPC protocol via TCP/IPv4 over Ethernet IEEE802.3. Control algorithms for experimentation can be prepared in Matlab code or as Simulink blocks and extended with functions/blocks, which access PLC. In our case all schemes for data acquisition were put together as Simulink blocks.

4.2 Modelling and validation

The modelling was done in two stages as described in Section 3. In the first stage eleven linear local models were identified around equilibrium points. These were levels of liquid h_1 ranging from 0.4 meter to 1.4 meter. Pseudo-random binary signal was used as the input signal for identification. The local models were of first order and obtained with ARX method.

Result of these identifications were parameters of discrete local models. These were, together with associated parameter variances, transformed to parameters of continuous first order models with step variance transformation method.

In the second stage two Gaussian process models were trained, each on a set of eleven input data for parameters $a(\rho) = -\frac{1}{T}$ and $b(\rho) = \frac{K}{T}$ with associated variances, where a, b are state and input scalars of state-space system description, T stands for continuous system time constant and K stands for continuous system gain. Liquid level was used as a scheduling variable. It would be more complete if also input and output variables would be included in scheduling vector. However, the preliminary analysis shown in Kocijan *et al.* (2002) shows that liquid level is far the most dominant variable. The variables and associated variances can be easily extracted from obtained Gaussian process models. Predictions of GP models for both varying parameters with associated variances as well as with training data are depicted in Fig. 2 and Fig. 3.

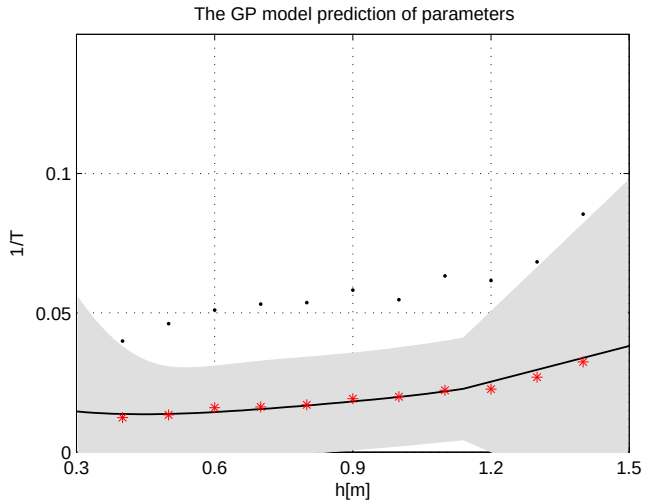

Fig. 2. Identified parameter $(1/T) = -a(\rho)$

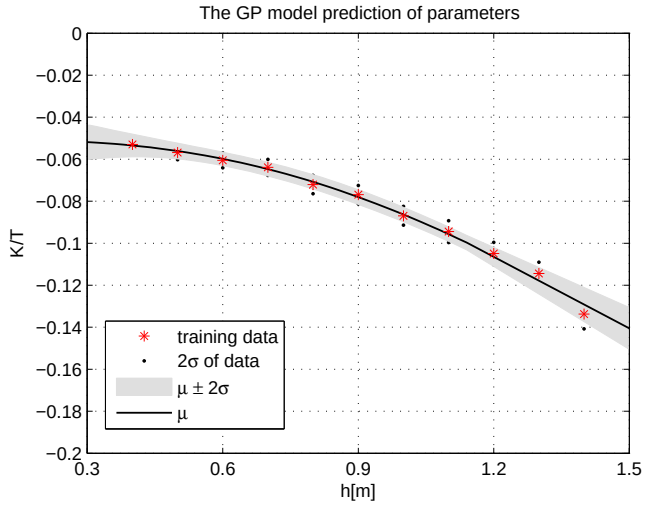

Fig. 3. Identified parameter $(K/T) = b(\rho)$

It can be seen from Fig. 2 that the cumulative variance of parameter prediction is lower than the variance of training data. This is due to the property of Gaussian process model that assigns lower variance in places where there is more data and because of that prediction is more confident.

The obtained Gaussian process models which form a complete first order model with varying parameters were then used for simulation. The responses are in Fig. 4 and input and liquid level as scheduling variable are given in Fig. 5.

The plots of varying time constant and gain obtained from Gaussian process models with associated confidence bands as well as with extra depicted variances are given in Fig. 6 and Fig. 7.

The movement of variable confidence is shown in a phase plot in Fig. 8 in which the variance of time constant is depicted in relation to the variance of gain. It can be seen from Fig. 8 that variances are simultaneously decreasing towards low value and then again increasing as the model is approaching limits of model validity.

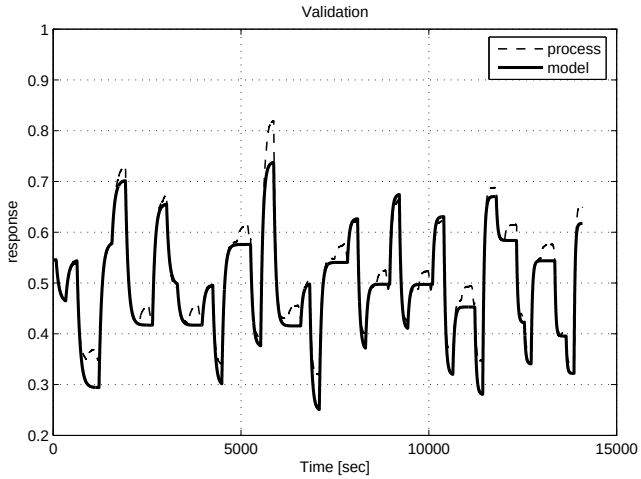

Fig. 4. Process and model response

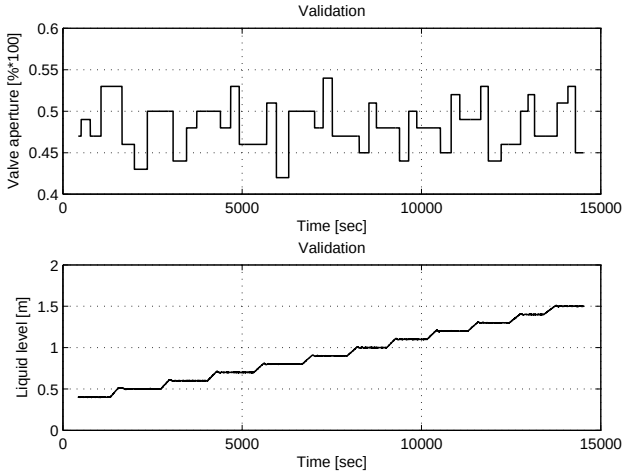

Fig. 5. Valve aperture and liquid level

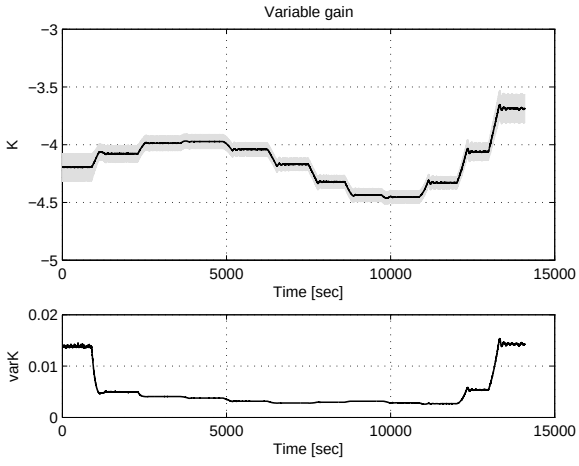

Fig. 6. Variable gain

5. CONCLUSION

Most control design methods are based on parametric models. As Gaussian process models with the advantage to give information about confidence in its predictions are non-parametric probabilistic models they are not suited

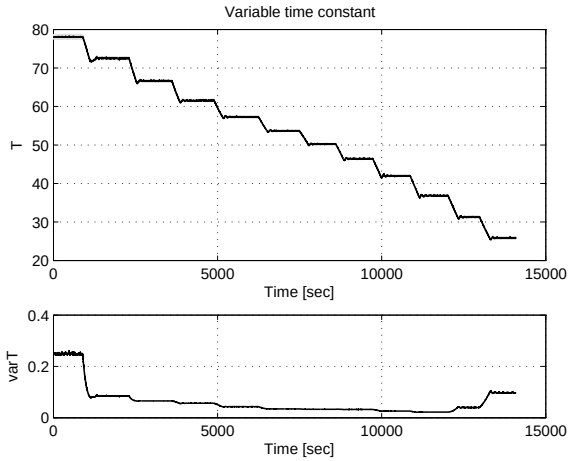

Fig. 7. Variable time constant

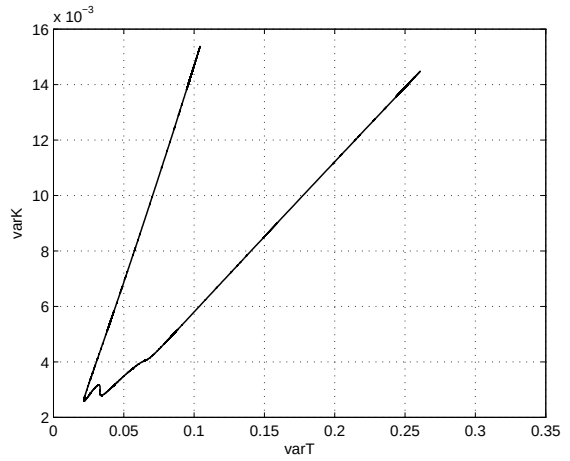

Fig. 8. Diagram with phase plot of parameter variances

for lots of design methods that are based on parametric models.

This was the rationale for developing a parameter varying model that retains some of the advantages of Gaussian process model, but also includes some characteristics of parametric model. Fixed-structure Gaussian process (FSGP) model is the model that is suited for probabilistic description of parameters of linear parameter varying model.

It was shown in the paper that the method was successfully applied on modelling of a gas-liquid separator case study. The resulting model may not be used only to simulate the process behaviour but also to give information about confidence in these parameters regarding the region in which the system is operating.

Local models acquisition needs to pursued with a great caution to obtain credible results that can be used for FSGP model integration. The complexity of this task increases with the order of local models. More on this topic can be found in Leith and Leithead (2003).

The obtained FSGP model with information about confidence can be used for control design in at least two ways. The first approach is to avoid regions in which the model is uncertain and the second approach is to derive optimal controller based on information about uncertainties. More research is planned to investigate the utility of FSGP model for control design in the future.

REFERENCES

Ažman K. and Kocijan J. (2006). An Application of Gaussian Process Models for Control Design, In *International Control Conference ICC 2006*, Glasgow, 4 pages.

Ažman K. and Kocijan J. (2007). Application of Gaussian processes for black-box modelling of biosystems, *ISA Transactions*, 46, 443–457.

Ažman K. and Kocijan J. (2009). Fixed-Structure Gaussian Process Model, *International Journal of Systems Science*, accepted for publication.

Gregorčič G. and Lightbody G. (2005). Gaussian Process Approaches to Nonlinear Modelling for Control, in *IEE Control Series 70 – Intelligent Control Systems using Computational Intelligence Techniques*, IEE, London, 177–217.

Johansen T. A., Shorten R. and Murray-Smith R. (2000). On the Interpretation and Identification of Dynamic Takagi-Sugeno Fuzzy Models, *IEEE Transactions on Fuzzy Systems*, 8, 297–313.

Kocijan J., Žunič G., Strmčnik S., Vrančić D. (2002). Fuzzy Gain-scheduling Control of a Gas-liquid Separation Plant Implemented on a PLC, *Int. J. Control*, 75, 1082–1091.

Kocijan J., Girard A., Banko B. and Murray-Smith R. (2005). Dynamic Systems Identification with Gaussian Processes, *Mathematical and Computer Modelling of Dynamic Systems*, 11, 411–424.

Leith D. J. and Leithead W. E. (1999). Analytic Framework for Blended Multiple Model Systems Using Linear Local Models, *Int. J. Control*, 72, 605–619.

Leith D. J. and Leithead W. E. (2002). Global Reconstruction of Nonlinear Systems From Families of Linear Systems, In *IFAC World Congress 2002*, Barcelona, 6 pages.

Leith D. J. and Leithead W. E. (2003). Necessary and Sufficient Conditions for the Minimal State-Space Realisation of Nonlinear Systems from Input-Output Information, In *Proceedings of the 2003 American Control Conference*, 4, 2937–2942.

Murray-Smith R. and Johansen T. A. (Eds.)(1997). *Multiple Model Approaches to Modelling and Control*, London: Taylor and Francis.

Murray-Smith R., Johansen T. A. and Shorten R. (1999). On Transient Dynamics, Off-equlibrium Behaviour and Identification in Blended Multiple Model Structures. In *ECC European Control Conference 1999*, BA-14, Karlsruhe, 6 pages.

Nyström R. H., Aakesson B. M. and Toivonen H. T. (2002). Gain-Scheduling Controllers Based on Velocity-Form Linear Parameter-Varying Models Applied to an Example Process, *Ind. Eng. Chem. Res.*, 41, 220–229.

Peng H., Ozaki T., Haggan-Ozaki V., and Toyoda Y. (2003). A Parameter Optimization Method for the Radial Basis Function Type Models, *IEEE Transactions on Neural Networks*, 14, 432–438.

Rasmusen C. E. and Williams C. K. I. (2006). *Gaussian Processes for Machine Learning*, Cambridge: MIT Press.