Predictive control of a gas-liquid separation plant based on a Gaussian Process model

Bojan Likar[1], Juš Kocijan[1,2]

[1]Jozef Stefan Institute, Jamova 39

SI-1000 Ljubljana, Slovenia

tel: +386 1 4773 661, fax: +386 1 4773 994

E-mail: jus.kocijan@ijs.si

[2]University of Nova Gorica, Vipavska 13

SI-5000 Nova Gorica, Slovenia

Abstract

Gaussian process models provide a probabilistic non-parametric modelling approach for black-box identification of nonlinear dynamic systems. The Gaussian processes can highlight areas of the input space where prediction quality is poor, due to the lack of data or to its complexity, by indicating the higher variance around the predicted mean. Gaussian process models contain noticeably less coefficients to be optimised. This paper demonstrates feasibility of application and realisation of a control algorithm based on a Gaussian process model. The extra information provided by the Gaussian process model

is used in predictive control, where optimisation of the control signal takes the variance information into account. The feasibility of Gaussian process model usage for predictive control in industrial practice is demonstrated via the control of a gas-liquid separation plant.

1 Introduction

The popularity of MPC can be attributed largely to the ability of MPC algorithms to deal with constraints that are frequently met in control practice and are often not well addressed by other approaches. MPC algorithms can handle hard state and rate constraints on inputs and states that are usually, but not always, incorporated in the algorithms via an optimisation method. Linear model predictive control approaches [Maciejowski, 2002] started appearing in the early eighties and are well-established in control practice (e.g. [Qin and Badgwell, 1997] for an overview). Nonlinear model predictive control (NMPC) approaches [Allgöwer et al., 1999] started to appear about ten years later and have also found their way into control practice (e.g. [Qin and Badgwell, 2000],[Young et al., 2001]) though their popularity can not be compared to linear model predictive control. This is due to the difficulties associated with nonlinear model construction and with the lack of the necessary confidence in the model. There have been a number of contributions in the field of nonlinear model predictive control dealing with issues such as stability, efficient computation, optimisation, constraints and others. Some recent work in this field can be found in [Allgöwer and Zheng, 2000], [Kouvaritakis and Cannon, 2001]. NMPC algorithms are based on various nonlinear models. Often these models are developed as first principles models, but other approaches - like black-box identification approaches -

are also popular. Various predictive control algorithms are based on neural networks models e.g. [Nørgaard et al., 2000], fuzzy models e.g. [Kavšek-Biasizzo et al., 1997] or local model networks e.g. [Johansen et al., 1995]. Nonlinear model-based predictive control, as the name implies, critically depends on the nonlinear plant model. The better the model, the better the control. This is where Gaussian process models can offer significant advantages. Gaussian process models provide a measure of confidence, which would be of help in NMPC design as noted in [Tsai et al., 2002], where a different approach to the same problem was described. The Gaussian process model is an example of a probabilistic non-parametric black-box model that also provides information about prediction uncertainties which are difficult to evaluate appropriately in nonlinear parametric models. The majority of work on Gaussian processes presented up to now considers the modelling of static nonlinearities. The use of Gaussian processes in modelling dynamic systems is a recent development e.g. [Murray-Smith and Girard, 2001], [Girard et al., 2003], [Kocijan et al., 2003a], [Kocijan et al., 2003b],
[Girard and Murray-Smith, 2005] and some control algorithms based on such an approach are described in [Murray-Smith and Sbarbaro, 2002], [Gregorčič and Lightbody, 2003]. This approach to modelling is not considered as a replacement of any existing method, but rather as a complementary approach to modelling. The drawback of Gaussian process models is the considerable computational burden. This burden may be perceived as an obstacle for Gaussian process model usage in industrial control applications. The purpose of this paper is to demonstrate the feasibility of application and realisation of a control algorithm based on a Gaussian process model on a process plant and to highlight some of the potentials. More about the benefits of dynamic systems modelling with Gaussian processes can be found in e.g. [Kocijan et al., 2003a], [Girard and Murray-Smith, 2005].

The paper is organized as follows. Dynamic Gaussian process models are briefly introduced in the following section. The control algorithm principle is described in Section 3. The example in Section 4 illustrates the operation of NMPC on a gas-liquid separator plant. Conclusions are stated at the end of the paper.

2 Modelling of Dynamic Systems with Gaussian Processes

A Gaussian process is an example of the use of a flexible, probabilistic, non-parametric model with uncertainty predictions. Its use and properties for modelling are reviewed in [Williams, 1998].

A Gaussian process is a collection of random variables which have a joint multivariate Gaussian distribution. Assuming a relationship of the form $y = f(\mathbf{x})$ between an input $\mathbf{x}$ and output y, we have $y^1, \ldots, y^n \sim \mathcal{N}(0, \Sigma)$, where $\Sigma_{pq} = \mathrm{Cov}(y_p, y_q) = C(\mathbf{x}_p, \mathbf{x}_q)$ determines the covariance between output points corresponding to input points $\mathbf{x}_p$ and $\mathbf{x}_q$. Thus, the mean $\mu(\mathbf{x})$ (usually assumed to be zero) and the covariance function $C(\mathbf{x}_p, \mathbf{x}_q)$ fully specify the Gaussian process. Note that the covariance function $C(.,.)$ can be any function having the property of generating a positive definite covariance matrix.

A common choice is

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

where $\boldsymbol{\Theta} = [w_1 \ldots w_D \; v_0 \; v_1]^T$ are the 'hyperparameters' of the covariance functions, δ_{pq} is the Kronecker operator, and D is the input dimension. Other forms of covariance functions suitable

for different applications can be found in [Rasmussen, 1996]. For a given problem, the parameters are learned (identified) using the data at hand. After the learning the w parameters can be used as indicators of 'how important' the corresponding input components (dimensions) are: if w_d is zero or near zero it means that the inputs 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 $\mathbf{y} = [y^1, y^2, \ldots, y^N]^T$. 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 maximizing the log-likelihood of the parameters, which is computationally relatively demanding since the inverse of the data covariance matrix ($N \times N$) has to be calculated at every iteration. 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 utilized for regression calculation. Based on a 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 is $y^*|(\mathbf{X}, \mathbf{y}), \mathbf{x}^*$

and is Gaussian, with mean and variance

$$\mu(\mathbf{x}^*) = \mathbf{k}(\mathbf{x}^*)^T \, K^{-1} \, \mathbf{y}, \tag{2}$$

$$\sigma^2(\mathbf{x}^*) = k(\mathbf{x}^*) - \mathbf{k}(\mathbf{x}^*)^T \, \mathbf{K}^{-1} \, \mathbf{k}(\mathbf{x}^*), \tag{3}$$

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 $k(\mathbf{x}^*) = C(\mathbf{x}^*, \mathbf{x}^*)$ is the covariance between the test input and itself.

Gaussian processes can, like neural networks, be used to model static nonlinearities and can therefore be used for modelling dynamic systems [Girard et al., 2003], [Kocijan et al., 2003a], [Kocijan et al., 2003b], [Girard and Murray-Smith, 2005] if delayed input and output signals are fed back and used as regressors. In such cases an autoregressive model is considered, such that the current output depends on previous outputs, as well as on previous control inputs.

$$\begin{aligned} \mathbf{x}(k) &= [\hat{y}(k-1), \hat{y}(k-2), \ldots, \hat{y}(k-L), u(k-1), \\ &\quad u(k-2), \ldots, u(k-L)]^T, \\ \hat{y}(k) &= f(\mathbf{x}(k)) + \epsilon, \end{aligned} \tag{4}$$

where k denotes the consecutive number of the data sample. Let $\mathbf{x}$ denote the state vector composed of the previous outputs y and inputs u up to a given lag L, and ϵ is white noise.

For multi-step ahead prediction the uncertainty of future predictions which provides the 'inputs' for estimating further means and uncertainties must be taken into account. The partial overview of results expounded in [Girard et al., 2003] is presented as follows.

In the case of multi-step ahead prediction a prediction at $\mathbf{x}^*$ is sought, where input vector $\mathbf{x}^*$ contains also uncertain inputs fed back from outputs. Within a Gaussian approximation, input values can be described by normal distribution $\mathcal{N}(\boldsymbol{\mu}_{\mathbf{x}^*}, \boldsymbol{\Sigma}_{\mathbf{x}^*})$, where $\boldsymbol{\mu}_{\mathbf{x}^*}$ and $\boldsymbol{\Sigma}_{\mathbf{x}^*}$ are the vector

and the matrix of the input mean values and variances respectively. To obtain a prediction the predictive distribution $p(y^*|(\mathbf{X},\mathbf{y}),\mathbf{x}^*)$ needs to be integrated over the input distribution, that is

$$p(y^*|(\mathbf{X},\mathbf{y}),\boldsymbol{\mu}_{\mathbf{x}^*},\boldsymbol{\Sigma}_{\mathbf{x}^*}) = \int_{-\infty}^{+\infty} p(y^*|(\mathbf{X},\mathbf{y}),\boldsymbol{\mu}_{\mathbf{x}^*})p(\mathbf{x}^*|\boldsymbol{\mu}_{\mathbf{x}^*},\boldsymbol{\Sigma}_{\mathbf{x}^*})d\mathbf{x}^*, \tag{5}$$

where $p(y^*|(\mathbf{X},\mathbf{y}),\mathbf{x}^*) = \frac{1}{\sqrt{2\pi\sigma^2(\mathbf{x}^*)}}\exp\left[-\frac{(y^*)-\mu(\mathbf{x}^*))^2}{\sigma^2(\mathbf{x}^*)}\right]$. Since $p(y^*|(\mathbf{X},\mathbf{y}),\mathbf{x}^*)$ is a nonlinear function of $\mathbf{x}^*$, the new predictive distribution $p(y^*|(\mathbf{X},\mathbf{y}),\boldsymbol{\mu}_{\mathbf{x}^*},\boldsymbol{\Sigma}_{\mathbf{x}^*})$ is not Gaussian and this integral cannot be solved without using approximations.

Approximations can be roughly divided into numerical, for example Monte-Carlo numerical methods, and analytical approximations. In order to achieve computational simplicity the analytical approximation was used, which consists of computing only the first two moments, namely the mean and variance of $p(f(\mathbf{x}^*)|(\mathbf{X},\mathbf{y}),\mathbf{x}^*)$.

To distinguish between $\mu(\mathbf{x}^*)$ and $\sigma^2(\mathbf{x}^*)$, the mean and variance of the Gaussian predictive distribution in the case when there are no uncertain inputs, $m(\boldsymbol{\mu}_{\mathbf{x}^*},\boldsymbol{\Sigma}_{\mathbf{x}^*})$ denotes the mean and $v(\boldsymbol{\mu}_{\mathbf{x}^*},\boldsymbol{\Sigma}_x)$ the variance of the non-Gaussian predictive distribution $p(y^*)|(\mathbf{X},\mathbf{y}),\boldsymbol{\mu}_{\mathbf{x}^*},\boldsymbol{\Sigma}_{\mathbf{x}^*})$, corresponding to $\mathbf{x}^* \sim \mathcal{N}(\boldsymbol{\mu}_{\mathbf{x}^*},\boldsymbol{\Sigma}_{\mathbf{x}^*})$. This can be interpreted as a Gaussian approximation, such that

$$p(y^*|(\mathbf{X},\mathbf{y}),\boldsymbol{\mu}_{\mathbf{x}^*},\boldsymbol{\Sigma}_{\mathbf{x}^*}) \approx \mathcal{N}(m(\boldsymbol{\mu}_{\mathbf{x}^*},\boldsymbol{\Sigma}_{\mathbf{x}^*}), v(\boldsymbol{\mu}_{\mathbf{x}^*},\boldsymbol{\Sigma}_{\mathbf{x}^*})). \tag{6}$$

The predictive mean and variance of the output corresponding to a noisy input $\mathbf{x}^*$ are obtained by solving [Girard et al., 2003]

$$m(\boldsymbol{\mu}_{\mathbf{x}^*},\boldsymbol{\Sigma}_{\mathbf{x}^*}) = E_{\mathbf{x}^*}[\mu(\mathbf{x}^*)], \tag{7}$$

$$v(\boldsymbol{\mu}_{\mathbf{x}^*},\boldsymbol{\Sigma}_{\mathbf{x}^*}) = E_{\mathbf{x}^*}[\sigma^2(\mathbf{x}^*)] + \mathrm{var}_{\mathbf{x}^*}[\mu(\mathbf{x}^*)]. \tag{8}$$

Instead of working with the expressions of $\mu(\mathbf{x}^*)$ and $\sigma^2(\mathbf{x}^*)$, equations (7) and (8) are solved by approximating directly $\mu(\mathbf{x}^*)$ and $\sigma^2(\mathbf{x}^*)$ by their first and second order Taylor expansions respectively around $\boldsymbol{\mu}_{\mathbf{x}^*}$. The second order expansion is required in order to obtain a correction term for the new variance. This is a relatively rough approximation. Other Gaussian approximations can be found in [Girard et al., 2003] and [Girard and Murray-Smith, 2005].

The first order Taylor expansion of $\mu(\mathbf{x}^*)$ around $\mathbf{x}^* = \boldsymbol{\mu}_{\mathbf{x}^*}$ gives

$$\mu(\mathbf{x}^*) = \mu(\boldsymbol{\mu}_{\mathbf{x}^*}) + (\mathbf{x}^* - \boldsymbol{\mu}_{\mathbf{x}^*})^T \boldsymbol{\mu}'(\boldsymbol{\mu}_{\mathbf{x}^*}) + O(||\mathbf{x}^* - \boldsymbol{\mu}_{\mathbf{x}^*}||^2), \tag{9}$$

where $\boldsymbol{\mu}'(\boldsymbol{\mu}_{\mathbf{x}^*}) = \left.\frac{\partial \mu(\mathbf{x}^*)}{\partial \mathbf{x}^*}\right|_{\mathbf{x}^*=\boldsymbol{\mu}_{\mathbf{x}^*}}$ and $O(\cdot)$ are Taylor expansion higher order terms.

It follows that

$$E_{\mathbf{x}^*}[\mu(\mathbf{x}^*)] \simeq \mu(\mathbf{x}^*) \tag{10}$$

and furthermore

$$\mathrm{var}_{\mathbf{x}^*}[\mu(\mathbf{x}^*)] \simeq \boldsymbol{\mu}'(\boldsymbol{\mu}_{\mathbf{x}^*})^T \boldsymbol{\Sigma}_{\mathbf{x}^*} \boldsymbol{\mu}'(\boldsymbol{\mu}_{\mathbf{x}^*}) = \mathrm{Tr}[\boldsymbol{\mu}'(\boldsymbol{\mu}_{\mathbf{x}^*}) \boldsymbol{\mu}'(\boldsymbol{\mu}_{\mathbf{x}^*})^T \boldsymbol{\Sigma}_{\mathbf{x}^*}]. \tag{11}$$

For the computation of the new variance, $\sigma^2(\mathbf{x}^*)$ was approximated by its second order Taylor expansion around $\boldsymbol{\mu}_{\mathbf{x}^*}$:

$$\sigma^2(\mathbf{x}^*) = \sigma^2(\boldsymbol{\mu}_{\mathbf{x}^*}) + (\mathbf{x}^* - \boldsymbol{\mu}_{\mathbf{x}^*})^T \boldsymbol{\sigma}^{2\prime}(\boldsymbol{\mu}_{\mathbf{x}^*}) + \frac{1}{2}(\mathbf{x}^* - \boldsymbol{\mu}_{\mathbf{x}^*})^T \boldsymbol{\sigma}^{2\prime\prime}(\boldsymbol{\mu}_{\mathbf{x}^*})(\mathbf{x}^* - \boldsymbol{\mu}_{\mathbf{x}^*}) + O(||\mathbf{x}^* - \boldsymbol{\mu}_{\mathbf{x}^*}||^3), \tag{12}$$

where $\boldsymbol{\sigma}^{2\prime}(\boldsymbol{\mu}_{\mathbf{x}^*}) = \left.\frac{\partial \sigma^2(\mathbf{x}^*)}{\partial \mathbf{x}^*}\right|_{\mathbf{x}^*=\boldsymbol{\mu}_{\mathbf{x}^*}}$ and $\boldsymbol{\sigma}^{2\prime\prime}(\boldsymbol{\mu}_{\mathbf{x}^*}) = \left.\frac{\partial^2 \sigma^2(\mathbf{x}^*)}{\partial \mathbf{x}^* \partial \mathbf{x}^{*T}}\right|_{\mathbf{x}^*=\boldsymbol{\mu}_{\mathbf{x}^*}}$, leading to

$$E_{\mathbf{x}^*}[\sigma^2(\mathbf{x}^*)] \simeq \sigma^2(\boldsymbol{\mu}_{\mathbf{x}^*}) + \frac{1}{2}\mathrm{Tr}[\boldsymbol{\sigma}^{2\prime\prime}(\boldsymbol{\mu}_{\mathbf{x}^*}) \boldsymbol{\Sigma}_{\mathbf{x}^*}] \tag{13}$$

The final expression for variance is obtained by replacing (11) and (13) in (8).

Consequently, within a Gaussian approximation and a Taylor expansion of $\mu(\mathbf{x}^*)$ and $\sigma^2(\mathbf{x}^*)$ around $\mathbf{x}^* = \boldsymbol{\mu}_{\mathbf{x}^*}$, the predictive distribution is again Gaussian with the following mean and variance

[Girard et al., 2003]

$$\begin{aligned} m(\boldsymbol{\mu}_{\mathbf{x}^*}, \boldsymbol{\Sigma}_{\mathbf{x}^*}) &= E_{\mathbf{x}^*}[\mu(\mathbf{x}^*)] \\ &\approx \mathbf{k}(\mu(\mathbf{x}^*)^T \mathbf{K}^{-1} \mathbf{y}, \end{aligned} \tag{14}$$

$$\begin{aligned} v(\boldsymbol{\mu}_{\mathbf{x}^*}, \boldsymbol{\Sigma}_{\mathbf{x}^*}) &= E_{\mathbf{x}^*}[\sigma^2(\mathbf{x}^*)] + \mathrm{var}_{\mathbf{x}^*}(\mu(\mathbf{x}^*)) \\ &\approx \sigma^2(\mu(\mathbf{x}^*)) + \frac{1}{2}\mathrm{Tr}\left\{ \left. \frac{\partial^2 \sigma^2(\mathbf{x}^*)}{\partial \mathbf{x}^* \partial \mathbf{x}^{*T}} \right|_{\mathbf{x}^* = \boldsymbol{\mu}_{\mathbf{x}^*}} \boldsymbol{\Sigma}_{\mathbf{x}^*} \right\} \\ &+ \left. \frac{\partial \mu(\mathbf{x}^*)}{\partial \mathbf{x}^*} \right|^T_{\mathbf{x}^* = \boldsymbol{\mu}_{\mathbf{x}^*}} \boldsymbol{\Sigma}_{\mathbf{x}^*} \left. \frac{\partial \mu(\mathbf{x}^*)}{\partial \mathbf{x}^*} \right|_{\mathbf{x}^* = \boldsymbol{\mu}_{\mathbf{x}^*}}. \end{aligned} \tag{15}$$

For a more detailed derivation see [Girard et al., 2003], where also the comparison of the approximations, here mentioned but not used, is presented. Equations (14) and (15) can be applied to the calculation of multi-step ahead prediction with propagation of uncertainty.

As can be seen from the presented relations, the obtained model not only describes the dynamic characteristics of the nonlinear system, but also provides information about the confidence in these predictions by means of prediction variance. The Gaussian process can highlight areas of the input space where prediction quality is poor, due to lack of data, by indicating higher variance around the predicted mean.

3 Model Predictive Control

It is not the purpose of this section to explain the fundamentals of model predictive control, which can be found in literature (e.g. [Maciejowski, 2002], [Allgöwer and Zheng, 2000], [Kouvaritakis and Cannon, 2001]), but rather to highlight some issues of MPC, which are related to the case when the Gaussian process model is used.

3.1 Some general issues

Model Predictive Control (MPC) is a common name for computer control algorithms that use an explicit process model to predict the future plant response. According to this prediction within the selected period, also known as the prediction horizon, the MPC algorithm optimises the manipulated variable to obtain an optimal future plant response. The input of chosen length, also known as the control horizon, is sent into the plant and then the entire sequence is repeated again in the next time period. The most common practice is to only send the first sample of the optimised input to the plant. MPC is one of the rare advanced control principles that has found its place in industrial practice, especially in process control. This is most likely due to its intuitive principle and attractive properties which enable control of complex, nonlinear, dead-time, non-minimum phase and constraint systems. The disadvantages of MPC are the need for a good process model and often the computational burden, which is an obstacle when faster systems are controlled.

Nonlinear MPC contains a nonlinear process model, in our case, a Gaussian process model. The ability of Gaussian process models to provide information about confidence in the model

is the major reason for their use in MPC.

The availability of prediction confidence information presents a possibility to constrain the action of the closed-loop system within the trusted region of the model. This is not a concept from literature, but it is not unfamiliar in control engineering practice where safety of operations is of crucial importance. The 'soft' constraints on the confidence region is, on the other hand, a well known concept. It can be found under the name 'cautious' control (e.g. [Apley, 2004], [Murray-Smith and Sbarbaro, 2002]) and it means that the cost function contains a confidence region dependant weight or penalty that forces the closed-loop system to avoid the not trusted region. The 'hard' constraints that are proposed for increased safety do not just penalise the action, but do not allow any action outside of the region of model validity.

Determination of the control signal involves optimisation of the chosen cost function. When the process model is a nonlinear one, the optimisation problem is in general non-convex. Convergence of optimisation is not assured in the case of time and computationally demanding iterative algorithms. The cost function minimum might even be nondeterminate in particular cases of certain cost functions when constraints are included into the optimisation.

The optimisation problem is a serious one and there are many approaches suggested in literature to cope with it. These suggest how to reduce the computational burden and transform the optimisation into a convex problem, but in general their applicability is limited to parametric models, while the Gaussian process model is a probabilistic nonparametric model. Using Gaussian processes for the process model means computationally increasing the burden, and therefore it is important to pay attention to this issue. One way to cope with the dimensions of the computation problem is to use so called 'soft' constraints instead of 'hard' ones and

convert the constrained optimisation problem to an unconstrained one. How soft constraints can be used in connection with Gaussian process based model predictive control is described in [Kocijan and Murray-Smith, 2004].

The alternative solution is to keep the algorithm and cost function as simple as possible. This is the reason why the Predictive Function Control algorithm has been selected to demonstrate predictive control based on the Gaussian process model.

3.2 Predictive Functional Control

Predictive Functional Control [Richalet, 1993] is a predictive control algorithm which, due to its simplicity and effectiveness, has found its way into numerous industrial applications. The method has many versions, depending on the process model used, or structure of the control signal, etc. Common to all versions is that the prediction horizon is minimised to a very small number of points. Often this is just one point, which is called a coincidence point.

Another common point is the structuring of the control signal, a simple form of which is a constant signal in the entire control horizon.

In the case of a single coincidence point the cost function is

$$\min_{\mathbf{U}(k)} \; [r(k+P) - \hat{y}(k+P)]^2, \tag{16}$$

where $\mathbf{U}(k) = [u(k) \dots u(k+P)]$ is the input signal and P is the coincidence point.

The reasons why closed-loop control with use of such a simple cost function with only one coincidence point and no penalty on control action is often satisfactory, are [Maciejowski, 2002]:

- the required performance is largely determined by the reference trajectory and there is no need for further tuning of the cost function weights to obtain the required performance,
- the polynomial structure of the future control signal (in our case constant) results in relatively few degrees of freedom being available to the optimiser, which usually results in relatively smooth input trajectories.

Only one sample of the future control signal is applied and then the entire control signal optimisation procedure is repeated in accordance with the receding horizon strategy.

4 Example

4.1 Gas-liquid separation plant

The semi-industrial process plant used for the case study in the paper is a unit for separating gas from liquid that forms part of a larger pilot plant.

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. It was designed for flue gases containing CO_2 that are produced from natural gas combustion, but in our case regular air was used. The separation unit is 2.25 m high and has 0.312 m^2 cross-section area on inside. It normally operates with gas pressure between about 0.2 bar, depending on the amount of water inflow and maximally 1 bar and liquid height of maximum 1.9 m. More details of the plant can be found in [Vrančić et al., 1995].

The flow-sheet of the separator is shown in figure 1. The flue gases coming from the effluent channels are absorbed by the water flow into the water circulation pipe through injector I4.1.

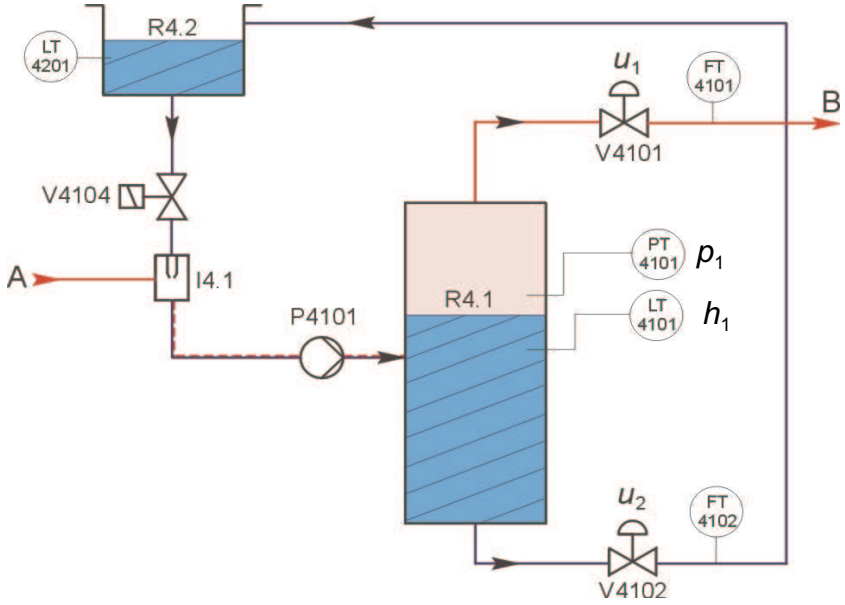

Figure 1: The flow-sheet of the gas-liquid separator

The water flow is generated by the pump P4101 (water ring). The speed of the pump is kept constant. The pump feeds the mixture of water and gas into the tank R4.1, where gas is separated from water. Hence the accumulated gas in R4.1 forms a sort of 'gas cushion' with increased internal pressure. Owing to this pressure, the flue gas is blown out from tank R4.1 into the neutralisation unit. On the other side, the 'cushion' forces water to circulate back to the reservoir R4.2. The quantity of water in the circuit is constant.

Both the gas pressure and the water level in the separator have to be controlled in order to maintain security and keep the process within an adequate working range. This might be difficult in the case when the model is uncertain and might result in potential safety risk. Two controllers are needed for this control. The first controls pressure in the unit by manipulating valve V4101. The second controller keeps the liquid level around its reference value by

manipulating valve V4102.

In order to understand the basic relations among variables and to illustrate the nonlinearity of the process a mathematical model is introduced (the derivation of the mathematical model can be found in [Vrančić et al., 1995]). The gas-liquid separation pressure sub-system of interest can be described with a set of two equations.

$$\begin{aligned}
\frac{dp_1}{dt} &= \frac{1}{S_1(h_{T_1}-h_1)}(p_0(\alpha_0+\alpha_1 p_1+\alpha_2 p_1^2-k_1 R_1^{u_1-1}\sqrt{p_1}) \\
&+ (p_0+p_1)(\Phi_w-k_2 R_2^{u_2-1}\sqrt{p_1+k_w(h_1-h_{T_2})})), \\
\frac{dh_1}{dt} &= \frac{1}{S_1}(\Phi_w-k_2 R_2^{u_2-1}\sqrt{p_1+k_w(h_1-h_{T_2})}),
\end{aligned} \tag{17}$$

where u_i is the command signal of valve V410i, i=1,2, 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 the atmospheric pressure, h_{Ti} is the height of tank R4.i, i=1,2, (2.25 m and 2 m respectively), R_i is the open-close flow ratio of valve V410i, i=1,2, (46 and 75.66 respectively), k_i is the flow coefficient of valve V410i, i=1,2, (75.1 l/(s·bar$^{1/2}$) and 0.742 l/(s·bar$^{1/2}$) respectively), Φ_w is the known constant water flow through pump P4101 $i = 1, 2$, (0.1644 l/s), $\alpha_i; i = 1, 2, 3$, are constant parameters (5.1475, 3.5634 and -5.0316 respectively).

From the model presented, it can be seen that the nonlinear process is of a multivariable nature (two inputs and two outputs with dynamic interactions between the channels). This problem can be overcome by separating closed-loop control bandwidths. The level control speed is much slower than the pressure control speed. Consequently the pressure control can be approached as a single-input single-output control problem, with the focus on solving the nonlinear control problem. As can be seen from Eqs. (17) pressure is nonlinearly related to level and input flow which results in different dynamic behaviour depending on the operating

region.

The model is presented exclusively to explain the basic dynamic characteristics of the plant. All following experiments were conducted on the actual plant.

The main control objective was to achieve uniform closed-loop performance for pressure control in the entire liquid level operating region. This means that the pressure dynamic response should be approximately equal, regardless of the liquid level. How this control objective was fulfilled with a technique different from the one proposed in this paper can be found in [Kocijan et al., 2002].

The rationale for the selected case study is as follows.

- The used plant contains features of industrial processes so control implementation and commissioning would be a good test for other real-life applications.
- The dynamics of the plant appear relatively simple, but not all features of the semi-industrial plant can be seen from the presented first-principle model. However, it is known that MPC algorithms can deal with complex dynamics regardless of the model used under the condition that the model describes the process well and reliably enough, which is the property of GP models. The stress in our example is on demonstrating the feasibility of implementation in real-time regardless of issues such as computational burden.
- The modelling approach may show its advantages on processes where first-principle modelling is more difficult (e.g. biological and medical systems), but it is not the main purpose of this paper to demonstrate benefits of modelling approach. Instead, the emphasis is on

the application and implementation of a Gaussian process model, which, we believe can be beneficial in process control and elsewhere.

- It can be argued that the plant can be modelled and controlled in other ways, but the choice to make the feasibility study on a familiar process was made purposefully so that experience from other applied and differently obtained control algorithms could be utilized and comparisons could be made.

4.2 Process identification

Since the process to be identified is characterised by equation (17) as the first order system, a model of the form (18) was identified

$$p_1(k+1) = f(p_1(k), u_1(k), h_1(k)), \tag{18}$$

consequently pressure $p_1(k)$, valve signal $u_1(k)$ and liquid level $h_1(k)$ were selected for regressors. Signals $u_1(k)$ and $h_1(k)$ were measured on-line and are deterministic. Pressure $p_1(k)$ was fed back as distribution, and the predicted mean and variance were calculated by Taylor expansion as described in the previous section. Attempts have been made to identify the system with a higher order model, but the results were not improved upon.

Based on the response and iterative cut-and-try procedure, a sampling time of 15 seconds was selected. Identification data consisted of pseudo random changes of valve signal u_1 in regions with different liquid level h_1, so that as wide a region as possible was encompassed in 967 samples for each signal.

Obtained hyperparameters of the first order Gaussian process were:

$$\begin{aligned} \boldsymbol{\Theta} &= [w_1, w_2, w_3, v_0, v_1] \\ &= [20.2759, 78.0774, 0.1517, 2.9145 \cdot 10^{-5}, 0.1162], \end{aligned} \tag{19}$$

where hyperparmeter w_1 corresponds to the pressure signal p_1, w_2 corresponds to the valve signal u_1, w_3 corresponds to the level signal h_1, v_0 is the estimated noise variance, and v_1 is the estimate of the vertical scale of variation.

The validation signals that are shown in figure 2 are different from the identification signals, though of the same kind. The response of the model to a validation signal as predicted mean value and two standard deviations or 95% confidence band and comparison with the process response are shown in figure 3. Two standard deviations or 95% confidence band are selected because this is a frequent Gaussian prediction representation, but any other value (e.g. three standard deviations or 99% confidence band) could be selected.

Fitting of the response for validation signal:

-average squared test error

$$SE = \frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2 = 3.2744 \cdot 10^{-4}, \tag{20}$$

-log density error

$$LD = \frac{1}{2N}\sum_{i=1}^{N}\left(\log(2\pi) + \log(\sigma_i^2) + \frac{(\hat{y}_i - y_i)^2}{\sigma_i^2}\right) = 149.07. \tag{21}$$

where y_i, $e_i = \hat{y}_i - y_i$ and σ_i^2 are the system's output, prediction error and prediction variance in i-th sample and N is the number of training points. The graph of standard deviation changes is shown in figure 4.

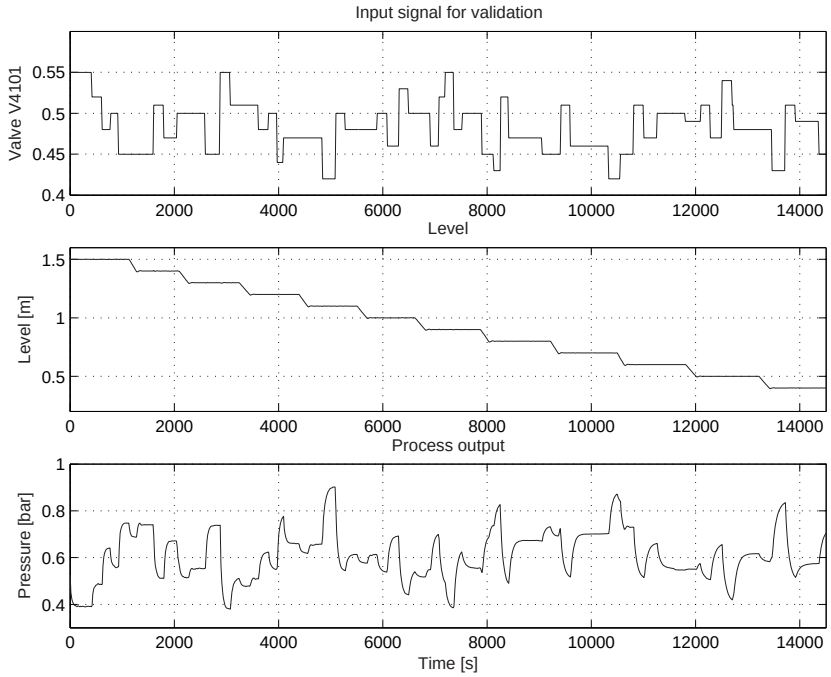

Figure 2: Model input validation signals

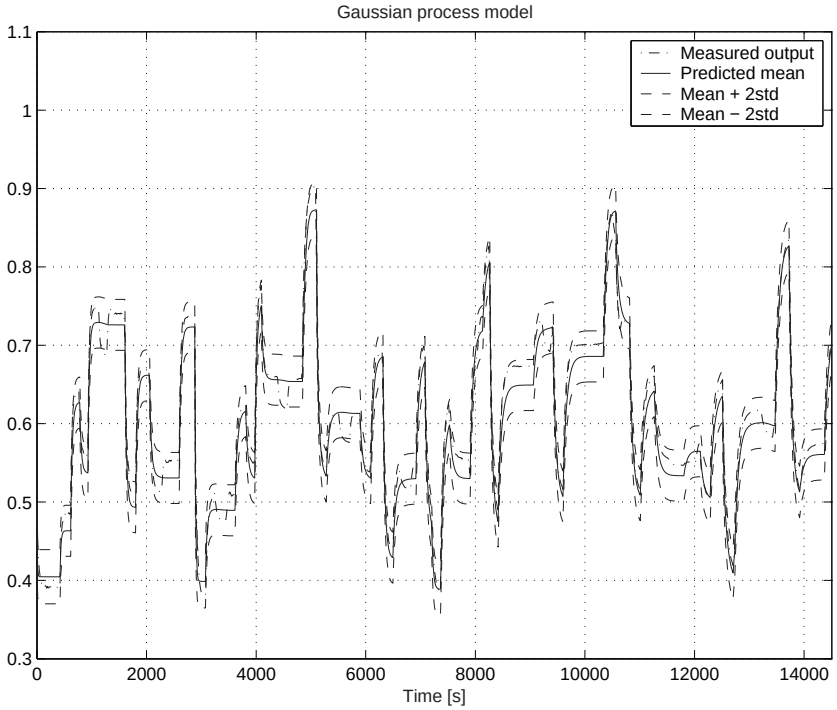

Figure 3: Response of the Gaussian process model to the excitation signal used for validation

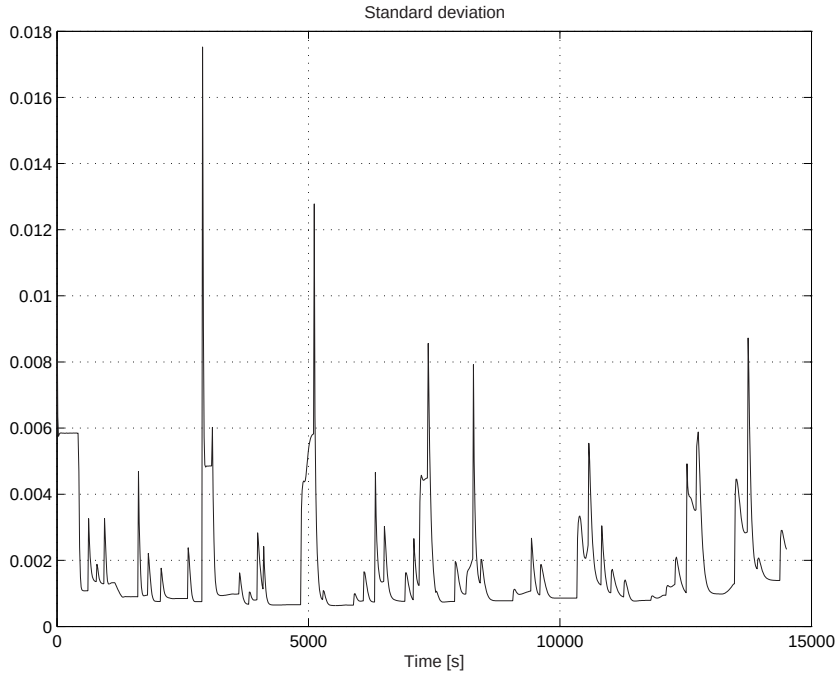

Figure 4: Standard deviation of the GP model

The relatively low value of the average squared test error in comparison with the relatively high value of the log density error shows that while the model mean values follow the process response well, the variance may not be large enough. Nevertheless, the standard deviation still clearly indicates the regions where identification data are sparse, as can be seen from figure 4, which is good enough for control purposes.

4.3 Control design and assessment

A moving-horizon minimization problem of the special form [Maciejowski, 2002]

$$\min_{\mathbf{U}(k)} \; [r(k+P) - \hat{y}(k+P)]^2, \tag{22}$$

subject to:

$$\text{var } \hat{y}(k+P) \leq \text{var}_{max}, \tag{23}$$

is used in our case, where $\mathbf{U}(k) = [u(k) \dots u(k+P)]$ is the input signal, P is the coincidence point (the point where a match between output and reference value is expected) and inequality (23) represents an arbitrarily selected constraint on output variance var_{max}. It is possible to add hard constraints on other variables, but in our case they were not taken into account. The process model is a Gaussian process.

As already mentioned, the predictive control algorithm uses the information about model prediction variances. The variances are the sum of variances that correspond to information about regions where there are varying degrees of confidence in the model accuracy, depending upon the local density of available identification data and of output response variances. When variances increase too much, a possible design option is that the response can be optimised with constrained control. The expected consequence is that the control algorithm does not allow any excursion in the region where the accuracy of the model is below the prescribed value. This is a possible way to guarantee safe operation based on known accuracy of the model.

The optimisation algorithm, which is constrained nonlinear programming, is solved at each sample time over a prediction horizon of length P, for a series of moves which is equal to the control horizon. The optimisation problem is solved with the Matlab Optimization Toolbox routine for constrained nonlinear minimisation. All these modifications, with the purpose of making the predictive algorithm as simple as possible, do not change the generality of the solution, but they do affect the numerical solution itself.

The reference trajectory, which defines the trajectory along which the plant should return to the setpoint trajectory, is very important in defining the closed-loop behaviour of the controlled plant [Maciejowski, 2002]. Often - and so it is in our case - the reference trajectory approaches

the setpoint exponentially from the current output value, with the time constant (T_{ref}) of the exponential defining the speed of response. If the current error is

$$\epsilon(k) = w(k) - y(k), \tag{24}$$

where $w(k)$ is the setpoint trajectory and $y(k)$ is the current output. The reference trajectory is chosen such that the error i steps later, if the output followed it exactly, would be

$$\epsilon(k+i) = e^{\frac{-iT_s}{T_{ref}}}\epsilon(k), \tag{25}$$

where T_s is the sampling interval. The reference trajectory is defined to be

$$r(k+i|k) = s(k+i) - \epsilon(k+i) = w(k+i) - e^{\frac{-iT_s}{T_{ref}}}\epsilon(k), \tag{26}$$

In our case the coincidence point was chosen to be 8 samples and the control horizon was one sample. The value of the coincidence point is determined iteratively as a compromise between closed-loop performance and real-time computation feasibility. This control algorithm is used for experimental assessment of the Gaussian process model based predictive control on a gas-liquid separation plant.

The real-time experiments were pursued in the environment schematically shown in Fig. 5.

This environment encompasses supervisory control on two levels: upper level with Factory Link SCADA system, and lower procedural and basic control levels implemented in two PLCs. This is one of the possible configurations of control, which can be found in industry. User-friendly experimentation with the process plant is enabled through the interface with the Matlab/Simulink environment, which was implemented on a PC computer (PII 400 MHz, 128 Mb RAM). This interface enables PLC access with Matlab/Simulink using DDE protocol via Serial Communication Link RS232 or TCP/IPv4 over Ethernet IEEE802.3. Control algorithms

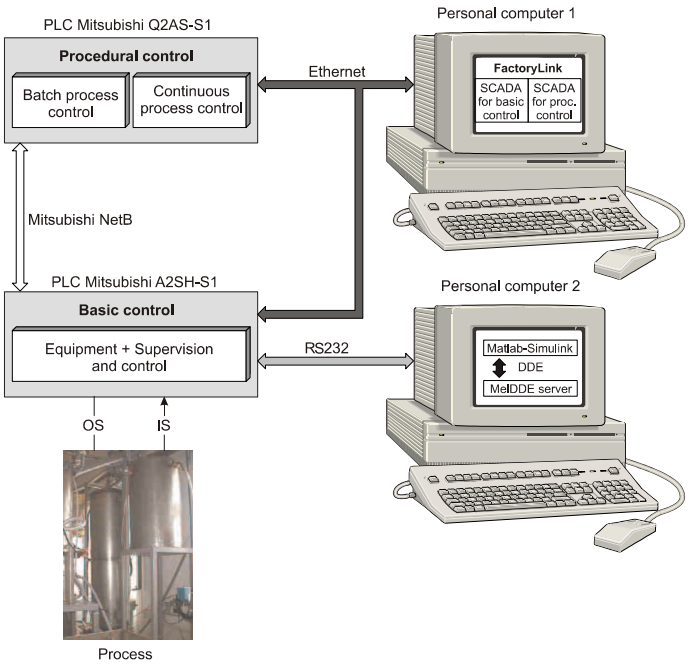

Figure 5: Experimental set-up for data acquisition and control algorithm testing

for experimentation can be prepared in Matlab code or as Simulink blocks and extended with functions/blocks, which access the PLC. This interface also enables user-friendly data acquisition for Matlab users. In our case all schemes for data acquisition as well as for control were composed as Simulink blocks.

Three experiments are presented. The first is tracking control for a square wave setpoint pressure signal with the level changing in the entire operating region. Nevertheless, the system operates within the region where the model is giving a good description of system dynamics, because this is the region where the model was identified. The time constant of the reference trajectory is $T_{ref} = 150s$. The closed-loop response is given in figure 6. The changing level and manipulative signal are shown together with the model prediction variance in figure 7.

The closed-loop performances with linear PID and gain-scheduling controller for the same process in similar operation are described in [Kocijan et al., 2002] and could be used for com-

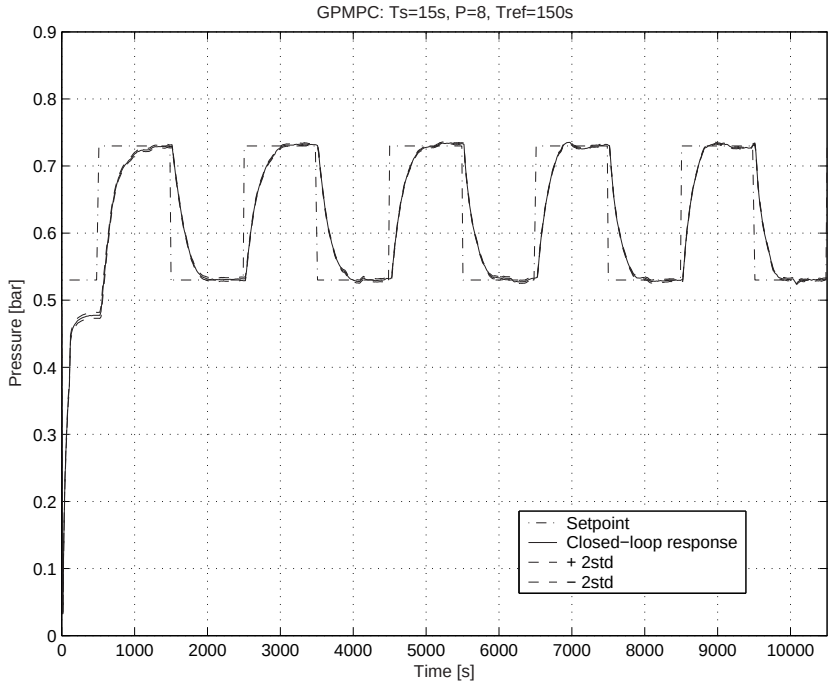

Figure 6: Closed-loop pressure response at changing liquid level

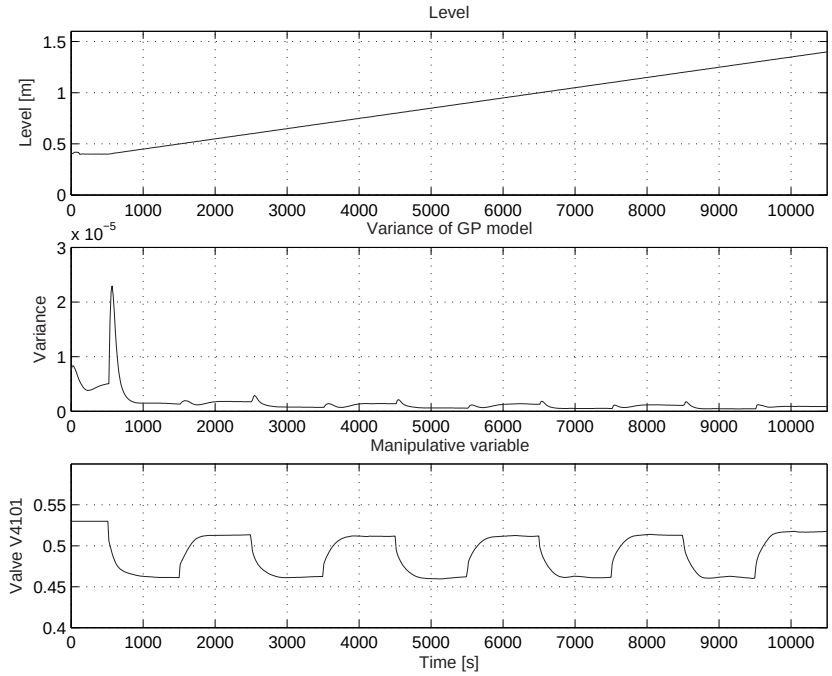

Figure 7: Liquid level, variance of GP model and manipulative variable

parison.

It can be seen from figure 6 that closed-loop performance is uniform regardless of level changes, and the predicted standard deviation from the model is consequently low and difficult to distinguish from the response in figure 6. It appears as if the hard constraint on variance ($\text{var}_{max} = 0.01$) is not coming into effect at all. However, the situation is such that if the hard constraint is not taken into account, the optimisation algorithm gets stuck in a local minimum at a distance from the global one and finds the input signal which drives the closed-loop system in the region where the model is very uncertain and plant operation becomes hazardous. The hard constraint on variance is actually used as an instrument to keep the closed-loop system within the operating region close to the global minimum. The global minimum or at least a minimum very close to the global one is then achieved by optimisation.

The further two experiments are conducted on the very edge of the region where the model was identified. This is the region where the model accuracy rapidly decreases. Both experiments are pursued at constant low liquid level, and pressure setpoint changing stepwise from values where the model is accurate toward the values where the model is not accurate any more. The time constant of the reference trajectory is $T_{ref} = 400s$.

In the second experiment, again the constraint of $\text{var}_{max} = 0.01$ is imposed on the control variable optimisation algorithm. The value of the constraint is high enough to ensure that the predicted model variance does not reach that limit value, but, on the other hand, is tight enough to prevent the optimisation algorithm from getting stuck in a local minimum at a distance from the global one, and keeps the closed-loop system within safe operation, as was the case in the previous experiment. The closed-loop response is shown in figure 8. The level and manipulative signal are shown together with the model prediction variance in figure 9.

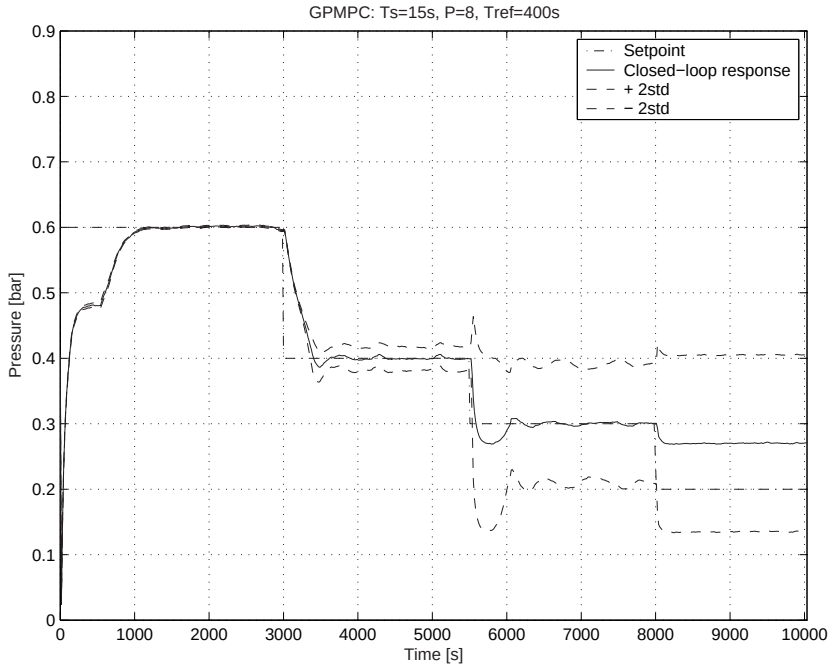

Figure 8: Closed-loop pressure response on the border of the operating region

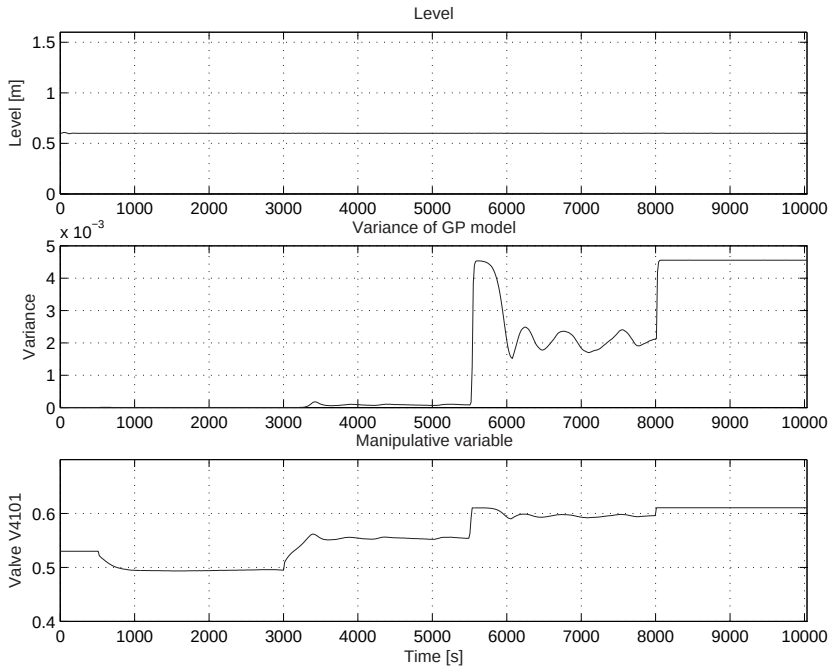

Figure 9: Liquid level, variance of GP model and manipulative variable on the border of the operating region

It can be seen from figures 8 and 9 that the standard deviation and variance of model prediction are increasing rapidly when the Gaussian process model leaves the region where the density of identification data decreased. The mismatch between the model and the real process becomes so large that firstly the performance, determined by the reference trajectory, is deteriorated (between 5500 and 8000 seconds) and further in the operation a large steady-state error occurs (beyond 8000 seconds). The latter means that the control algorithm not only does not fulfill its purpose anymore but the uncontrollable plant operation might not be within safety parameters. The value of steady-state error in the region of mismatch between the model and process differs depending on the values of process and model input variables and can be obtained only by experiment on the plant itself.

In the third experiment the constraint of $\text{var}_{max} = 0.00012$ is imposed to the control variable optimisation algorithm. This constraint value allows the algorithm to operate only in the region where the process model is good enough to guarantee performance of model predictive control with specified accuracy. The closed-loop response is shown in figure 10. The level and manipulative signal are shown together with the model prediction variance in figure 11.

The value of var_{max} is very important in the performance of the controller. The selection of var_{max} depends on the compromise that the designer is ready to make between performance and dimensions of the operation region and is constrained with available model, safety and stability of the closed-loop system. As constraints on variance are tightened, performance more closely matches the specified one, and more contracted is the region of operation. However, it is important to stress that it is the model that defines the region of operation. var_{max} only relaxes the borders of the region.

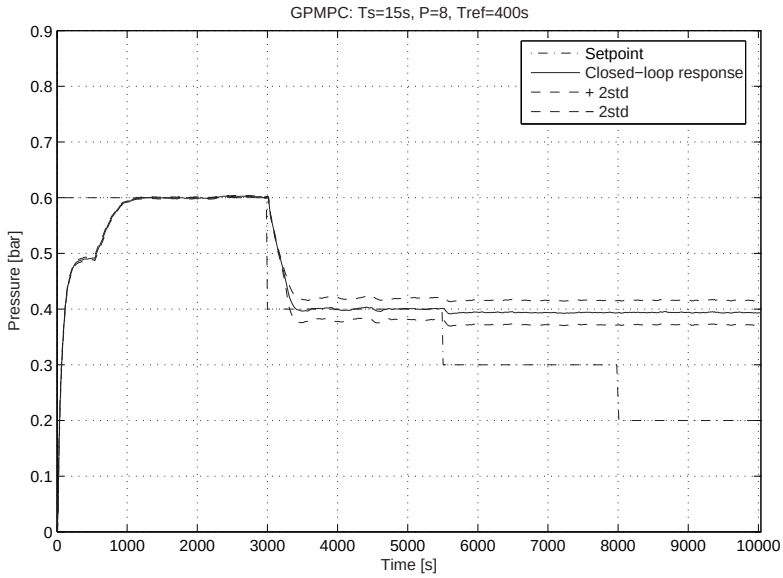

Figure 10: Closed-loop pressure response on the border of the operating region with variance constraint at 0.00012

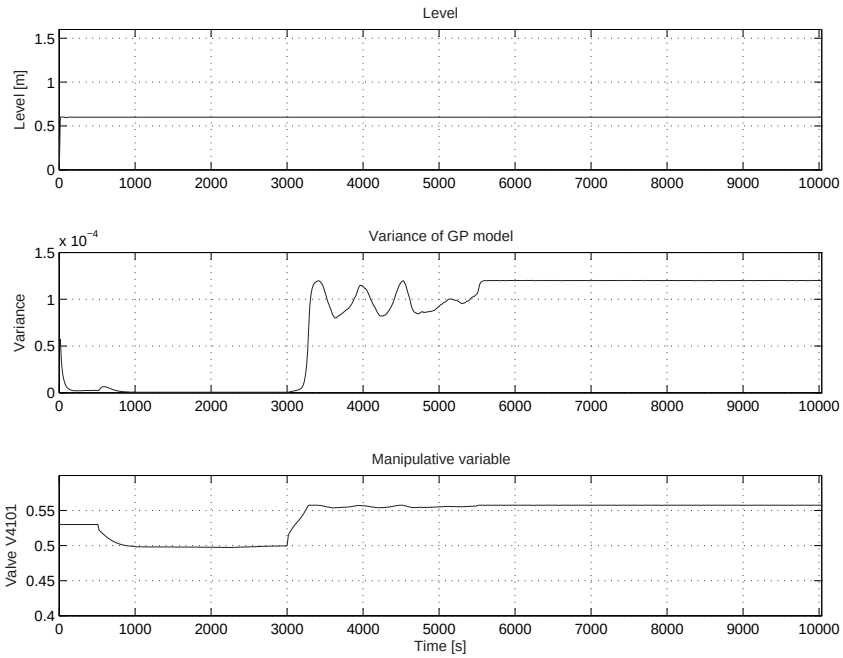

Figure 11: Liquid level, variance of GP model and manipulative variable on the border of the operating region with variance constraint at 0.00012

It can be seen from figures 10 and 11 that standard deviation and variance of model prediction are again increasing when the Gaussian process model leaves the region where the density of identification data decreased. In this region the constraint comes into effect and the closed-loop system response now avoids the region with large variance, at the cost of not tracking the setpoint and therefore an increase of steady-state error. However, the safety of operations is assured since the plant operates only within the known region.

Thus the goal, to design and implement a control algorithm which will guarantee operation only in the region of guaranteed performance, is fulfilled.

5 Conclusions

The application of Gaussian process model based predictive control is presented in the paper. In addition to other advantages, such as simple model structure and reduced sensitivity to the choice of model structure, Gaussian process models provide a measure of confidence in the form of prediction variance for the identified process model, and this measure is attractive for control design. Incorporation of confidence in the control algorithm enables design of control which avoids regions where the density of identification data is not large enough to provide a model of the required accuracy. This is a way to ensure desired closed-loop performance and circumvent some stability or robustness issues.

The Gaussian process model is a probabilistic non-parametric model. Consequently, model based predictive control is one of the suitable algorithms that can be used with such a type of process model. The drawback of Gaussian process dynamic systems models is the computational

burden necessary for the model simulation. In our case, a model containing about one thousand points has been identified. This means that calculation of the inverse covariance matrix is a relatively time consuming operation.

The main conclusion is that real-time model predictive control with such a model is demonstrated as viable on a process plant, regardless of the sizable model. It has to be noted that the choice of predictive algorithm and uncertainty propagation was made as simple as possible so that the approach would be applicable. The simplified implementation within used hardware (subsection 4.3) enabled that all computations necessary for closed-loop control, including Gaussian process model prediction, uncertainty propagation and constraint optimisation, were performed just within the limits of sampling time. Nevertheless, this fact does not reduce the general applicability of the Gaussian process model for predictive control.

The application demonstrates that using Gaussian process models offers an attractive possibility for control design. The controller has a higher level of robustness due to information contained in the model and provides a certain level of operation safety.

Activities are proceeding in the direction of reducing the computational burden and making the approach attractive also for the control of processes with faster dynamics.

Acknowledgements

This work was supported by the European Commission-funded Multi-Agent Control Research Training Network, HPRN-CT-1999-00107 and the Ministry of Science and Technology, Slovenia.

References

[Allgöwer et al., 1999] Allgöwer F., Badgwell T.A., Qin S.J., Rawlings J.B., Wright S.J.,1999, Nonlinear predictive control and moving horizon estimation - an introductory overview, In: Frank, P.M (edt.), *Advances in control: highlights of ECC'99*, (Springer), 391-449.

[Allgöwer and Zheng, 2000] Allgöwer F., Zheng A. (eds.), 2000, *Nonlinear Model Predictive Control*, Progress in system and control theory, Vol. 26, (Basel: Birkhäuser Verlag).

[Apley, 2004] Appley D.W., 2004, A cautious minimum variance controller with ARIMA disturbances, *IIE Transactions* 36 (5), 417-432.

[Girard et al., 2003] Girard A., Rasmussen C.E., Quinonero Candela, J. and Murray-Smith R., 2003, Gaussian Process Priors With Uncertain Inputs - Application to Multiple-Step Ahead Time Series Forecasting, In: *Advances in Neural Information Processing Systems*, Vol. 15, Eds. S. Becker, S. Thrun, K. Obermayer, (MIT Press), 542-552.

[Girard and Murray-Smith, 2005] Girard A. and Murray-Smith R., 2005, Gaussian Process: Prediction at a noisy input and application to iterative multiple-step ahead forecasting of time-series, In: *Switching and Learning in Feedback Systems*, Eds. R. Murray-Smith, R. Shorten, (Springer).

[Gregorčič and Lightbody, 2003] Gregorčič G., Lightbody G., 2003, Internal model control based on a Gaussian process prior model, *Proceedings of ACC'2003*, Denver, CO, (2003), 4981-4986.

[Johansen et al., 1995] Johansen T.A., Foss B.A., Sorensen A.V., 1995, Non-linear predictive control using local models - applied to a batch fermentation process, *Control Eng. Practice*, 3(3), 389-396.

[Kavšek-Biasizzo et al., 1997] Kavšek-Biasizzo K., Škrjanc I., Matko D., 1997, Fuzzy predictive control of highly nonlinear pH process, *Computers & chemical engineering*, Vol. 21, Supp. 1997, S613-S618.

[Kocijan et al., 2002] 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*, Vol. 75, No. 14, 1082-1091.

[Kocijan et al., 2003a] Kocijan J., Girard A., Banko B., Murray-Smith R., 2003, Dynamic Systems Identification with Gaussian Processes, *Mathematical and Computer Modelling of Dynamical Systems*, Vol. 11, No. 4, 411–424.

[Kocijan et al., 2003b] Kocijan J., Likar B., Banko B., Girard A., Murray-Smith R., Rasmussen C.E., 2003, A case based comparison of identification with neural network and Gaussian process models, *Preprints of IFAC ICONS Conference*, Faro, 137-142.

[Kocijan and Murray-Smith, 2004] Kocijan J., Murray-Smith, R., 2004, Nonlinear Predictive Control with a Gaussian Process Model, In: *Switching and Learning in Feedback Systems*, Eds. R. Murray-Smith, R. Shorten, (Springer).

[Kouvaritakis and Cannon, 2001] Kouvaritakis B., Cannon M. (eds.), 2001, *Nonlinear predictive control, Theory and practice*, IEE Control Engineering Series 61, (IEE).

[Maciejowski, 2002] Maciejowski J.M., 2002, *Predictive control with constraints*, (Harlow: Pearson Education Limited).

[Murray-Smith and Girard, 2001] Murray-Smith, R., Girard, A., 2001, Gaussian Process priors with ARMA noise models, *Irish Signals and Systems Conference*, Maynooth, 147-152.

[Murray-Smith and Sbarbaro, 2002] Murray-Smith R. and Sbarbaro D., 2002, Nonlinear adaptive control using nonparametric Gaussian process prior models, In: *Proc. IFAC Congress*, Barcelona.

[Nørgaard et al., 2000] Nørgaard M., Ravn O., Poulsen N.K., Hansen L.K., 2000, *Neural networks for modelling and control of dynamic systems*,(London: Springer).

[Qin and Badgwell, 1997] Qin S.J., Badgwell T.A., 1997, An overview of industrial model predictive control technology, In: Kantor J.C., Garcia C.E., Carnahan B. (eds.) *Fifth international conference on Chemical process control, AChE and CACHE*, 232-56.

[Qin and Badgwell, 2000] Qin S.J., Badgwell T.A., 2000, An overview of nonlinear model predictive control applications, In: Allgöwer F., Zheng A. (eds.), *Nonlinear model predictive control*, (Basel: Birkhauser Verlag), 369-392.

[Rasmussen, 1996] Rasmussen C.E., 1996, Evaluation of Gaussian Processes and other Methods for Non-Linear Regression, Ph.D. Dissertation, Graduate department of Computer Science, University of Toronto, Toronto.

[Richalet, 1993] Richalet J., 1993, Industrial applications of model based predictive control, *Automatica*, Vol. 19., No. 5, 1251-1274.

[Tsai et al., 2002] Tsai P.-F., Chu J.-Z., Jang S.-S., Shieh S.-S., 2002, Developing a robust model predictive control architecture through regional knowledge analysis of artificial neural networks, *Journal of Process Control*, **13**, 423-435.

[Williams, 1998] Williams C.K.I., 1998, Prediction with Gaussian processes: From linear regression to linear prediction and beyond, In: *Learning in Graphical Models*, Jordan, M.I., (edt.), (Dordrecht: Kluwer Academic), 599-621.

[Young et al., 2001] Young R.E., Bartusiak R.D., Fontaine, R.W., 2001, Evolution of an industrial nonlinear model predictive controller, *Preprints on Chemical Process Control - CPC VI, CACHE*, Tucson, AZ, 399-401.

[Vrančić et al., 1995] Vrančić, D., Juričić, D., Petrovčič, J., 1995, Measurements and mathematical modelling of a semi-industrial liquid-gas separator for the purpose of fault diagnosis. Technical Report DP-7260, Institut Jozef Stefan, Ljubljana, Slovenia.
`http://www-e2.ijs.si/Damir.Vrancic/Files/dp7260.pdf`