Proceedings of the 3rd IFAC Workshop on Automatic Control in Offshore Oil and Gas Production, Esbjerg, Denmark, May 30 - June 1, 2018

Physical-stochastic (greybox) modeling of slugging

Jan Kloppenborg Møller * **Goran Goranović** ** **Niels Kjølstad Poulsen** *** **Henrik Madsen** ****

* *DTU Compute, Asmussens Allé Building 303B, Kgs. Lyngby DK-2800 (e-mail: jkmo@dtu.dk).*
** *gogo@dtu.dk*
*** *nkpo@dtu.dk*
**** *hmad@dtu.dk*

Abstract: We use state-based stochastic greybox modeling - combining physics and statistics - to model the slugging phenomenon. We extend the model of DiMeglio et al. (2010) to include random components and variable flow coefficients, providing 30 seconds prediction intervals. Altogether six models, each comprising no more than ten equations, are fitted to off-shore riser training data and then cross-validated on new data sets. We use advanced statistical methods to 1) obtain optimal parameters of a given model fitted to measurements, 2) give model predictions with uncertainty intervals, and 3) quantitatively measure the *relative* goodness of the extended models. These features of our reductive method are general and can be applied to any data sets. For the slugging data, simpler models are preferable over the more complex ones (although the differences are minute for practical purposes in oil and gas industry) and a high statistical significance obtained on the training data does not imply improved long term prediction on independent data. Better physical (mechanistic) models to capture slugging oscillations are needed, ultimately to develop effective control strategies.

Keywords: stochastic greybox modeling, slugging

1. INTRODUCTION

In this report the analysis is performed using greybox models consisting of a continuous time stochastic model (represented by a set of Stochastic Differential Equations, SDEs) and a set of discrete time observation equations.

We elaborate on our detailed methodology, Kristensen and Madsen (2003b), to demonstrate how physical models and statistical analyses are to be used to tackle and interpret the Big Data problems, such as slugging, in oil&gas industry.

Our report essentially tests how good is the (extended) model of slugging by DiMeglio et al. (2010) when applied to our data sets. As we explain below, the model extensions are guided by statistical measures on data and are not necessarily interpretable in physical terms.

1.1 General formulation

The most general set up which can be handled by the estimation procedure used here is given by the continuous time SDE

$$d\boldsymbol{x}_t = \boldsymbol{f}(\boldsymbol{x}_t, \boldsymbol{u}_t, t, \boldsymbol{\theta})dt + \boldsymbol{\sigma}(\boldsymbol{u}, t, \boldsymbol{\theta})d\boldsymbol{w}, \quad (1)$$

where x_t is the hidden state, $\boldsymbol{u}$ is input and $\boldsymbol{\theta}$ parameters; $\boldsymbol{f}$ is the drift term, $\boldsymbol{\sigma}$ the diffusion term and $\boldsymbol{w}$ the Wiener process, Øksendal (2003). The drift term describes the physical ODE part of the system, while the diffusion term describes the random perturbations of the system in continuous time including the lack of knowledge of the system. In addition to (1), the stochastic state space formulation consist of a discrete time observation equation

$$\boldsymbol{y}_k = \boldsymbol{h}(\boldsymbol{x}_k, \boldsymbol{u}_k, t_k, \boldsymbol{\theta}) + \boldsymbol{e}_k, \quad (2)$$

where subscript k refer to the observation, state or input at time t_k, and the observation error $\boldsymbol{e}_k$ is assumed to follow a Gaussian distribution with expectation $\boldsymbol{0}$ and variance $\boldsymbol{S}(\boldsymbol{u}_k, t_k, \boldsymbol{\theta})$. In the numerical scheme we estimate $\boldsymbol{\theta}$ as well as $\boldsymbol{e}_k$.

One of the strengths of our modeling method is the reduction of complexity, coupled to the quantification of uncertainties. Our three-state models have about 10 equations and less than 30 variables in total, in contrast to the more comprehensive physical models, e.g. with over 20 equations and more than 100 variables, E. Jahanshahi (2014).

1.2 The initial system of equations

All the variables of the system are listed in Table A.1. The initial state variables (based on DiMeglio et al. (2010)) [1] are

$$dm_{GB} = ((1-\epsilon)(w_{GI} + u_{RLG,t}) - w_{G,t})dt + \sigma_1 dw_{1,t} \quad (3)$$

$$dm_{GR} = (\epsilon(w_{GI} + u_{RLG,t}) + w_{G,t} - w_{GO,t})dt + \sigma_2 dw_{3,t} \quad (4)$$

$$dm_{LR} = (w_{LI} - w_{LO,t})dt + \sigma_3 dw_{3,t}, \quad (5)$$

[1] First listed in the report Zugno et al. (2011), available on request.

where m_{GB} is the mass of gas in the gas bubble, m_{GR} is the mass of gas in the riser and m_{LR} is the mass of liquid in the riser. We further have for the individual components

$$w_{G,t} = C_g \max(0, P_{GB,t} - P_{BOT,t}) \tag{6}$$

$$w_{GO,t} = w_{O,t} \frac{m_{GR,t}}{m_{GR,t} + m_{LR,t}} \tag{7}$$

$$w_{LO,t} = w_{O,t} \frac{m_{LR,t}}{m_{GR,t} + m_{LR,t}} \tag{8}$$

$$w_{O,t} = u_{SCV,t} C_c \sqrt{P_{TOP,t} - P_{ACV,t}}, \tag{9}$$

where the pressures $P_{TOT,t}$ and P_{BOT} and are given by the algebraic equation

$$P_{GB,t} = \frac{RT}{M \cdot V_B} m_{GB} \tag{10}$$

$$P_{TOP,t} = \frac{R \cdot T}{M \left(V_r - \frac{m_{LR,t} + m_{LST}}{\rho_L} \right)} \tag{11}$$

$$P_{BOT,t} = P_{TOP,t} + \frac{g \sin(\theta)}{A} (m_{LR,t} + m_{LST}). \tag{12}$$

Temperature was not provided, and was put as constant equal to the average value of the temperature in the data set presented in Zugno et al. (2011).

The observation equations are

$$Y_{BOT,k} = P_{BOT,k} + e_{1,k} \tag{13}$$

$$Y_{TOP,k} = P_{TOP,k} + e_{2,k} \tag{14}$$

$$Y_{GO,k} = \frac{w_{GO,k}}{M} + e_{3,k} \tag{15}$$

$$Y_{LO,k} = \frac{w_{LO,k}}{\rho_L} + e_{4,k}, \tag{16}$$

where $e_{i,k}$ are mutually independent Gaussian white noise processes, with the variance of $e_{i,k}$ equal s_i^2.

1.3 The final system of equations

There is an assumed delay between the riser and the separator, initially modeled by introducing more state variables, Jónsdottir et al. (2010). The estimation though yielded very large values of the diffusion coefficient and the time delay i.e. large uncertainty intervals. We disregarded the observed liquid flow in the further analysis.

To evaluate the likelihood, the state of the system (given by equations (3)-(5)) is estimated, implying that the initial state of the system is also estimated. This was difficult to estimate in some situations. As it is anticipated that this problem is caused by the strong non-linearities (given by (10)-(12)) in the transfer between the states and the observation, we formulated the state equations in the pressure domain rather than in the mass domain. A full stochastic transformation of the state, given by Itó's formula (see e.g Øksendal, 2003), would give a complicated expression for the diffusion term in the transformed domain. As the stochastic part of formulation in (3)-(5) does not include any physical hypothesis on the diffusion, we will consider the transformed deterministic part of the equations with additive diffusion and note that the interpretation of the diffusion is different in the transformed domain. A small simplification of the transformed system equations is obtained by considering the pressure difference between the top pressure and the bottom pressure ($\Delta_{p,t} = P_{BOT,t} - P_{TOP,t}$), rather than the bottom pressure.

The resulting system of equations are given by

$$dP_{GB,t} = \frac{R \cdot T}{M \cdot V_r} [(1-\epsilon)(w_{GI} + u_{RLG,t}) - w_{G,t}] dt + \sigma_1 dw_{1,t} \tag{17}$$

$$d\Delta_{p,t} = \frac{g \sin(\theta)}{A} (w_{LI} - w_{LO,t}) dt + \sigma_2 dw_{2,t} \tag{18}$$

$$dP_{TOP,t} = \frac{1}{V_r - \frac{A\Delta_p}{\sin(\theta) g \rho_L}} \left[\frac{R \cdot T}{M} [\epsilon(w_{GI} + u_{RLG,t} M) + w_{G,t} - w_{GO,t}] + \frac{P_{TOP,t}}{\rho_L (w_{LI} - w_{LO,t})} \right] dt + \sigma_3 dw_{3,t}, \tag{19}$$

where $m_{GR,t}$, and $m_{LR,t}$ are described by the algebraic equations

$$m_{GR,t} = M \left(V_r - \frac{A\Delta_{p,t}}{\sin(\theta) g \rho_L} \right) \frac{P_{TOP,t}}{R \cdot T} \tag{20}$$

$$m_{LR,t} = \frac{A\Delta_{p,t}}{g \sin(\theta)} - m_{LST}, \tag{21}$$

the algebraic expression for m_{GB} is not needed because it enters through its effect on the pressure of the gas bubble only. The observation equation for the reduced and transformed system is given by

$$Y_{BOT,k} = \Delta_{p,k} + P_{TOP,t} + e_{1,k} \tag{22}$$

$$Y_{TOP,k} = P_{TOP,k} + e_{2,k} \tag{23}$$

$$Y_{GO,k} = \frac{w_{GO,k}}{M} + e_{3,k}, \tag{24}$$

with the total of six parameters to describe the stochastic behaviour of the model (three diffusion parameters and three observation variances). The most important state variable is the top pressure, and it was therefore decided to emphasize this state variable in the estimation. One way of doing this is by fixing the observation variance for the top pressure at a lower level than the observation variance of the bottom pressure. It was therefore decided set $s_2^2 = \frac{s_1^2}{2}$.

2. DATA

Data sets are collected in a field in August 2010 and January 2011, and fully described in Ref. Cao (2011). All data sets were given with a time resolution of 3 seconds.

The first step of the data analysis is to choose one data set which will serve as the training set for model development. The aim of the modeling exercise is to be able to predict slugging, in order for this to be realistic the chosen training set should clearly hold information about this transition. One data set that hold this kind of information is the measurement series on January 28 at 1400 hours (Fig. 1).

The data show the increased opening of the slug control valve lead to a transition into the slugging phase after approximately 2.5 hours of recording. To control the system the valve was choked and the system settle in the stationary phase after approximate 4 hours of measuring.

The likelihood estimation relies on the assumption of Gaussian observations, as stated in Section 1.1. Serious departure form this assumption will jeopardize the conclusions which can be drawn from data. Unfortunately, the provided data contained many repeated value of the measurements, as illustrated in Fig. 2. Such measurements cannot be assumed to come from a continuous-discrete time

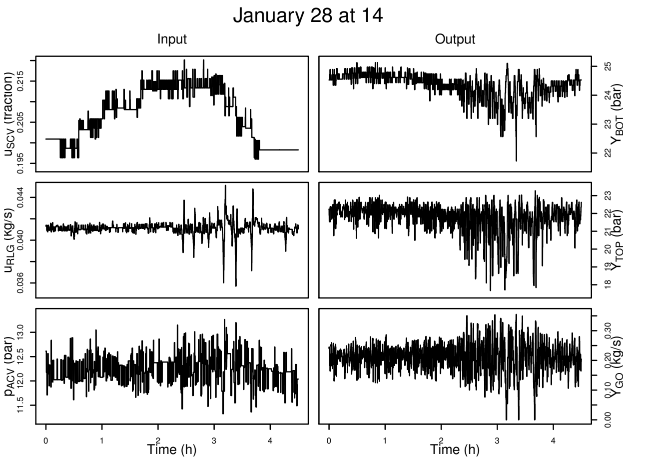

Fig. 1. Observations of the training set. Left column: The input variables. Right column: The output variables.

Table 1. Number of shifts (equal the number of datapoints used for estimation) and number of levels in the training set. Total number of observations is 5400

	no. of shifts	levels
Y_{BOT}	803	68
Y_{TOP}	1382	171
Y_{GO}	1558	1467

state space model as described in Section 1.1, and we thus removed repeated measurements, using only data where the measurements change. We did not wish to impose any models on data, such as the first order hold interpolations. Likewise, no variables, e.g. liquid production, were taken to be directly known if they were not observed - we went with the measurement we had.

As illustrated in Table 1 the situation shown in Fig. 2 is not unique, and for bottom pressure the are only 803 observations in the reduced dataset. The reduction correspond to a reduced sampling frequency, the numbers in Table 1 imply an average sampling frequency of 20, 12, and 10 seconds for Y_{BOT}, Y_{TOP}, and Y_{GO} respectively. Of courses this imply that we might not be able to capture some of the fast dynamics of the system. In addition to many repeated measurements, the number of different values of the measurements is also significantly lower than the number of measurements in the reduced observation set. This also imply that the measurement does not span the actual values.

3. MODEL DEVELOPMENT

Parameter estimation in this work is based on a well-developed statistical theory for testing, namely the likelihood estimation by the Extended Kalman Filter (EKF) Jazwinsky (1970). Note that in this approach we do not need a state for each parameter. The method is described in Kristensen et al. (2004b) and has been implemented in our software on continuous-time stochastic modeling the CTSM-R, Kristensen and Madsen (2003a,b), used to obtain the filter predictions for our models.

The basic assumption in the estimation procedure is that the one-step predictions of the observations can be

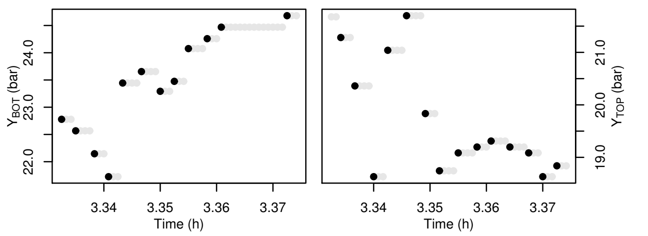

Fig. 2. Observations of the training set, left column show the bottom pressure while right column show the top pressure. grey dots indicate the provided data set while black dots indicate the data used for estimation.

described by a Gaussian random variable i.e.

$$\hat{\boldsymbol{Y}}_{k+1|k} = \boldsymbol{Y}_{k+1}|\boldsymbol{Y}_k \sim N(\boldsymbol{h}(\hat{\boldsymbol{X}}_{k+1|k}, \boldsymbol{u}_k, t_k, \boldsymbol{\theta}), \Sigma^{yy}_{k+|k}), \tag{25}$$

where $\hat{\boldsymbol{X}}_{k+1|k}$, and $\Sigma^{yy}_{k+|k}$ are obtained by the filtering equations. In the case of the EKF this amounts to a set of ordinary differential equation which are solved numerically. The likelihood is simply given as the product of all probability density functions (defined by (25)) taken at the observations

$$L(\boldsymbol{\theta}; \mathcal{Y}_N) = p(\boldsymbol{Y}_1|\boldsymbol{X}_0, \boldsymbol{\theta}) \prod_{i=1}^{N-1} p(\boldsymbol{Y}_{i+1}|\boldsymbol{Y}_i, \boldsymbol{\theta}), \tag{26}$$

where $p(\boldsymbol{Y}_1|\boldsymbol{X}_0, \boldsymbol{\theta})$ is the density for given $\boldsymbol{X}_0$, and $p(\boldsymbol{Y}_{i+1}|\boldsymbol{Y}_i, \boldsymbol{\theta})$ is defined by (25). All parameter estimation and the initial model development is based on maximum (log-)likelihood estimation.

The one-step state prediction ($\hat{\boldsymbol{X}}_{k+1|k}$) gives the expected path of the state variable given the model, the observations, and the filter assumptions. By formulating an extended state space model given by

$$d\boldsymbol{x}_t = \boldsymbol{f}(\boldsymbol{x}_t, \boldsymbol{u}_t, t, \tilde{\boldsymbol{\theta}}, \theta_{i,t})dt + \boldsymbol{\sigma}(\boldsymbol{u}, t, \tilde{\boldsymbol{\theta}})d\boldsymbol{w} \tag{27}$$

$$d\theta_{i,t} = \sigma_\theta dw_\theta, \tag{28}$$

where $\tilde{\boldsymbol{\theta}} = \boldsymbol{\theta} \setminus \theta_i = \{\theta_1, .., \theta_{i-1}, \theta_{i+1}, .., \theta_p\}$ and p is the dimension of the parameter space, it is possible to find the expected path of the random walk parameter θ_i. The methodology is presented in Kristensen et al. (2004a), and in the reference the selection of θ_i is based on values in the diffusion matrix $\boldsymbol{\sigma}$. Such a strategy is however complicated when the parameters cannot be attributed to a specific state (as is the case in the example considered here). We therefore estimated the extended state space model with each of the parameters treated as random walk parameters (one at the time).

The random walk parameter is then plotted as a function of the state predictions and the slug control valve opening (not shown); based on these plots model extensions were formulated and tested in a likelihood framework, see Table A.2.

3.1 Model 0: the benchmark model

The first step of the model development procedure is to estimate the parameters of the benchmark model presented in Section 1.1 and 1.3.

All parameters are restricted to positive real numbers and good practice for estimation of such parameters is to estimate them in a transformed domain. Most of the parameters are however expected to be well determined and transformation in such cases is not necessary; diffusion parameters and observation variances are on the other hand often difficult to estimate. To ensure stable estimates these are therefore estimated in the log-domain. Also, the parameter ϵ, the opening of the virtual valve, is restricted to the interval (0,1) and we estimate the logit transform of ϵ rather than ϵ itself, i.e. we estimate $\tilde{\epsilon}$ and use the back-transform

$$\epsilon = \frac{e^{\tilde{\epsilon}}}{1 + e^{\tilde{\epsilon}}} \tag{29}$$

such that $\tilde{\epsilon} \in \mathbb{R}$.

The estimation results are given in the first column of table A.3. Most of the parameters are well determined, except for m_{LST}, $\tilde{\epsilon}$, and $\tilde{\sigma}_3$. This imply that these parameters could be removed from the model, but we keep the non-significant ones and focus on model extensions rather than reductions. For $\tilde{\epsilon}$ the conclusion would be $\tilde{\epsilon} = 0$, which actually means that $\epsilon = \frac{1}{2}$ which is not a meaningful alternative. The large standard deviation does however imply that the 95% confidence interval for ϵ equals [0.03,0.65].

Table A.3 is one of the highlights of our method, listing (for all models) the optimal parameter values and their 95% confidence intervals, obtained via the likelihood estimation of Section 3.

3.2 Model extension

We have extended the benchmark model five times (Models 1-6), essentially varying the constant coefficients guided by the trends in data, see the text below equations (27) and (28). The extensions are shown in Table A.2: Diffusive coefficients σ_1 and σ_2 are made functions of the control valve opening u_{SCV} (Models 1 and 2), w_{LI} and V_B have linear dependence on u_{SCV} and pressure Δ_p, respectively (Models 3 and 4) and C_g is linear or quadratic function of P_{GB} (Models 5 and 6). All model extensions are nested and likelihood ratio tests can therefore be applied - the improvements are significant. We emphasize that goodness of models can be judged only relatively, see the next Section. Also, even though statistical extensions might not give any meaningful interpretation, they should be judged against a required purpose, e.g. prescribed time-step ahead predictions.

In Fig. 3 shown are the 30 second prediction (red line) along with the confidence limits (grey area) and the simulation i.e. the pure prediction of model without data update (blue line), for the final Model 6. The amplitude of the simulation is still small compared to what is seen from data - in other words, oscillatory behavior of the slugging is not captured by the model (even though the p-values are extremely small, indicating high significance, Table A.2). The simulation is however the unconditional expectation not a real i.e. single-run simulation, and some averaging should be expected.

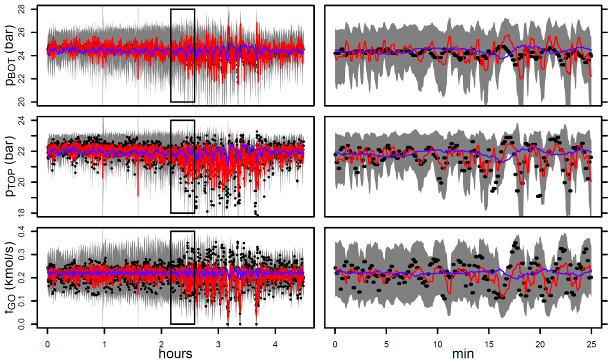

Fig. 3. Modeling results for Model 6. Left column: Observed values of pressure and flow (black dots), 30 second predictions of the observations (red lines), 95% confidence limits (grey area), and simulated values (blue lines). Right column show a zoom on the region where the system enters the slugging regime.

4. CROSS VALIDATION

The next step is to test model performance on independent datasets. To do this the remaining datasets from the January trial is divided into two groups, one with similar operating conditions as the training-set. Here we define similar operating conditions as dataset where the slug control valve opening is in the same range as in the test-set. The final goal of the model is to make good 30 second predictions, but the simplest evaluation is obtained by comparing the likelihood (i.e. 3 seconds predictions) for the different test sets across the models. Even though this is not the final goal this simple evaluation will give an indication of the performance of the 30 second predictions.

A summary of likelihoods for each of the models is given in Table A.4, when operating conditions are similar to the test set. Model 5 is the best performing model while Model 6 has a very poor performance (probably due to the behavior of the parameter C_g, when the second order polynomial is introduced). It was not possible to calculate the likelihood for Model 5 on part of the test set where operating conditions was not similar to the training-set (Test 2). This is probably due to the fact that C_g can obtain negative values.

In general, complex models are preferred on Test 1, while simple models are preferred on Test 2. This implies that extrapolation of the results obtained on the training-set is not possible, and further, when the best performing model on Test 1 cannot be evaluated on Test 2, some more robust models should be in place. Note, however, that our precision is very high and that the obtained parameter differences seen in table A.3 may not be practically important in industry. But the method will work, of course, also when model differences turn out to be large, or decisive.

5. CONCLUSION

We have applied detailed physical-stochastic, greybox, modeling to the problem of severe slugging.

The data used for this work have many repeated measurements, which imply that the time resolution in the measurements is not 3 seconds, as would be indicated from the time distance between observations. The number of observations removed due to repeated measures vary between the observations, and for bottom pressure almost

7 in 8 measurements were removed, which give an average time resolution of about 20 seconds. For top pressure the time resolution is about 12 seconds and for gas flow it is about 10 seconds. Due to this it might be difficult to detect the fast dynamics in the system.

The SDE models discussed in this report are all quite complex and even the simplest benchmark model is a highly non-linear SDE-model. The model development applied in Section 3 demonstrate that it is possible to extract information embedded in the observation by random walk parameter identification. The identified model extensions were all highly significant on the training set. Cross validation did however show that the most complex models were not robust and the performance on other data sets were generally poor. In other words, when operating conditions are not similar to those in the training-set, cross validation favoured simpler models. The most likely cause for this break down is that extrapolation led to negative values of the parameter C_g (in the case of Model 5) or very extreme values of C_g (in the case of Model 6).

It is clear that better overall performance can be obtained by including more data sets in the training-set. These should include a larger range of operating conditions to give a more globally applicable model. The benchmark model (Model 0) is based on a mechanistic understanding of the system, while the model extensions are based purely on statistical reasoning. They, too, could include mechanistic reasoning.

Stochastic greybox models can be successfully employed for control purposes, in, for example, model predictive controllers, Halvgaard et al. (2012). For the control of slugging though, a better mechanistic base model is needed to capture the behavior near the equilibrium point i.e., the occurrence of (the onset of) oscillations.

ACKNOWLEDGEMENTS

G.G. acknowledges the support of the Danish Hydrocarbon Research and Technology Center (DHRTC).

REFERENCES

Cao, Y. (2011). Slug control field trial report. Technical report, Report no. 11/YC/563, Dong Energy E& P A/S.

DiMeglio, F., Kaasa, G.O., Petit, N., and Alstad, V. (2010). Reproducing slugging oscillations of a real oil well. In *49th IEEE Conference on Decision and Control*, 4473–4479. Atlanta GA, USA.

E. Jahanshahi, S.S. (2014). Simplified dynamic models for control of riser slugging in offshore oil production. *Oil and Gas Facilities*, 64–79.

Halvgaard, R., Bacher, P., Perers, B., Andersen, E., Furbo, S., Jørgensen, J.B., K.Poulsen, N., and Madsen, H. (2012). Model predictive control for a smart solar tank based on weather and consumption forecasts. *Energy Procedia*, 30, 270–278.

Jazwinsky, A.H. (1970). *Stochastic Processes and Filtering Theory.* Academic Press, New York, ASU.

Jónsdottir, H., Jacobsen, J.L., and Madsen, H. (2010). A grey-box model describing the hydraulics in a creek. *Environmetrics*, 12, 347–356.

Kristensen, N.R. and Madsen, H. (2003a). Continuous time stochastic modelling ctsm 2.3 - mathematics guide. Technical report, Technical University of Denmark.

Kristensen, N.R. and Madsen, H. (2003b). Continuous time stochastic modelling ctsm 2.3 - user's guide. Technical report, Technical University of Denmark.

Kristensen, N., Madsen, H., and Jørgensen, S. (2004a). A method for systematic improvement of stochastic grey-box models. *Computers & Chemical Engineering*, 28, 1431–1449.

Kristensen, N., Madsen, H., and Jørgensen, S. (2004b). Parameter estimation in stochastic grey-box models. *Automatica*, 40, 225–237.

Øksendal, B. (2003). *Stochastic Differential Equations - An Introduction with Applications, Sixth edition.* Springer.

Zugno, M., Møller, J.K., Madsen, H., and Poulsen, N. (2011). Slug control valve project. Technical report, Technical University of Denmark.

Appendix A. TABLES

Table A.1. Variables in the deterministic part of the system equation, abbreviations of type refer to; SV= state variable, AE= algebraic equation, Parameter= Parameters to be estimated, Constant= Parameters held constant throughout the analysis and Input= Observed time-varying inputs.

Variable	Meaning	Type	Value	Unit
m_{GB}	Mass of gas in the bubble	SV/AE		kg
m_{GR}	Mass of gas in riser	SV/AE		kg
m_{LR}	Mass of liquid in riser	SV/AE		kg
m_{LST}	Constant mass of liquid in riser	Parameter		kg
w_{GI}	Gas production	AE		m^3/s
w_G	Gas flow through virtual valve	AE		m^3/s
w_{GO}	Gas flow through the SCV	AE		m^3/s
w_{LI}	Liquid production	Parameter		m^3/s
w_{LO}	Liquid flow through SCV	AE		m^3/s
$u_{RLG,t}$	Riser lift gas	Input		m^3/s
$u_{SCV,t}$	SCV opening	Input		fraction
C_G	Flow coefficient	Parameter		$\frac{kg}{s\sqrt{bar}}$
C_c	Flow coefficient	Parameter		$\frac{kg}{s \cdot bar}$
$P_{GB,t}$	Pressure in the gas bubble	AE/SV		bar
$P_{BOT,t}$	Pressure at the bottom of the riser	AE/SV		bar
$P_{TOP,t}$	Pressure at the top of the riser	AE/SV		bar
$P_{ACV,t}$	Production pressure	Input		bar
ϵ	Opening of virtual valve	Parameter		fraction
V_r	Volume of the riser	Constant		m^3
ρ_L	Density of liquid	Constant	832.5	$\frac{bar}{kg}$
R	Ideal gas constant	Constant	0.08314	$\frac{m^3 \cdot bar}{K \cdot kmol}$
T	Temperature	Constant	297	K
M	Molar weight of gas	Constant	23	$\frac{kg}{kmol}$
V_B	Volume of gas bubble	Parameter		m^3
g	Gravity	Constant	$9.81 \cdot 10^{-5}$	$\frac{bar \cdot m^2}{kg}$
θ	Angle of riser	Constant	1.57	rad
A	Cross section area of riser	Constant	0.0856	m^2

Table A.2. Summary of the model development presented in Section 3. Columns 1-6: model number, model extension, log-likelihood, number of degrees of freedom, -2 times the log-likelihood ratio and p-value, respectively.

Model	Extension	log(L)	DF	$-2\log\Lambda$	p-value
0		730	12		
1	$\sigma_1 = \exp(\tilde{\sigma}_1 + \tilde{\sigma}_{11} u_{SCV})$	764	13	69.1	1.1e-16
2	$\sigma_2 = \exp(\tilde{\sigma}_2 + \tilde{\sigma}_{21} u_{SCV})$	789	14	48.8	2.9e-12
3	$w_{LI} = w_{LI,0} + w_{LI,1} u_{SCV}$	805	15	32.7	1.1e-08
4	$V_B = V_{B,0} + V_{B,1} \Delta_p$	826	16	41.8	1.0e-10
5	$C_g = C_{g,0} + C_{g,1} P_{GB}$	865	17	77.1	<1.0e-16
6	$C_g = C_{g,0} + C_{g,1} P_{GB} + C_{g,1} P_{GB}^2$	899	18	68.6	1.1e-16

Table A.3. Estimation of parameters and their confidence intervals (in parentheses) for model 1-6. Parameter that are significantly different from 0 on a 95% level are marked in bold face.

	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
				Drift parameters			
$w_{LI,0}$	**106.30**	**106.24**	**105.00**	**14.40**	**11.30**	10.05	**102.78**
	(1.152)	(1.385)	(1.485)	(5.333)	(1.489)	(7.61)	(1.021)
$w_{LI,1}$				**444.49**	**458.02**	**469.15**	**440.75**
				(29.854)	(10.105)	(34.208)	(69.123)
V_B	**45.54**	**46.53**	**45.84**	**45.09**	0.07	$5.5e-6$	**34.73**
	(1.94)	(2.066)	(2.304)	(2.256)	(0.546)	$(1.4e-5)$	(1.383)
$V_{B,1}$					**16.70**	**9.85**	**27.33**
					(0.943)	(0.393)	(2.960)
C_g	**1.83**	**1.87**	**1.82**	**1.87**	**2.00**	**2.66**	**2.29**
	(0.096)	(0.090)	(0.104)	(0.100)	(0.101)	(0.079)	(0.021)
$C_{g,1}$						**−0.51**	0.02
						(0.037)	(0.016)
$C_{g,2}$							**−0.29**
							(0.007)
C_c	**173.13**	**173.06**	**172.64**	**173.16**	**172.65**	**174.26**	**172.13**
	(1.773)	(2.015)	(2.199)	(2.124)	(2.113)	(1.839)	(1.63)
$\tilde{\epsilon}$ [a)]	−1.40	−1.40	−1.46	**−1.95**	**−1.62**	**−4.91**	**−1.81**
	(1.057)	(2.010)	(1.226)	(0.900)	(0.110)	(0.533)	(0.119)
w_{GI}	**4.00**	**3.94**	**3.96**	**3.98**	**4.04**	**4.07**	**4.12**
	(0.035)	(0.035)	(0.034)	(0.030)	(0.029)	(0.027)	(0.028)
m_{LST}	$1.9e-4$	$7.0e-32$	0.96	0.06	0.05	0.06	0.03
	(0.113)	$(1.2e-30)$	(0.65)	(0.149)	(0.062)	(0.062)	(0.024)
				Diffusion parameters			
$\tilde{\sigma}_1$ [b)]	**−2.39**	**−12.73**	**−15.36**	**−15.8**	**−17.02**	**−16.07**	**−13.45**
	(0.053)	(0.055)	(0.479)	(0.667)	(0.121)	(0.060)	(0.010)
$\tilde{\sigma}_{11}$ [b)]		**49.81**	**62.46**	**64.41**	**69.81**	**67.8**	**53.58**
		(0.242)	(2.334)	(3.209)	(0.476)	(0.381)	(0.053)
$\tilde{\sigma}_2$ [b)]	**−2.47**	**−2.49**	**−8.72**	**−9.44**	**−9.31**	**−8.74**	**−9.15**
	(0.029)	(0.030)	(0.878)	(0.195)	(0.678)	(0.340)	(0.015)
$\tilde{\sigma}_{21}$ [b)]			**30.13**	**33.5**	**32.83**	**29.71**	**31.90**
			(4.280)	(0.973)	(3.332)	(1.668)	(0.083)
$\tilde{\sigma}_3$	−11.12	−13.12	**−11.12**	−10.59	−10.30	−9.97	−9.76
	(160.850)	(13.938)	(4.141)	(11.600)	(16.53)	(16.933)	(8.021)
				Variance of observation noise			
$\log(s_1^2)$	**−1.49**	**−1.54**	**−1.56**	**−1.55**	**−1.54**	**−1.57**	**−1.67**
	(0.043)	(0.040)	(0.046)	(0.045)	(0.043)	(0.037)	(0.040)
$\log(s_3^2)$	**−7.97**	**−7.87**	**−7.94**	**−7.91**	**−7.86**	**−7.64**	**−7.64**
	(0.084)	(0.030)	(0.095)	(0.091)	(0.094)	(0.084)	(0.052)

[a)] $\epsilon = \exp(\tilde{\epsilon})/[1+\exp(\tilde{\epsilon})]$

[b)] $\sigma_i(\tilde{\sigma}_i, \tilde{\sigma}_{i1}, u_{SCV}) = \exp(\tilde{\sigma}_i + \tilde{\sigma}_{i1} u_{SCV})$

Table A.4. Log-score for each of the estimated models with linear hypothesis on the parameters; **L28at14** is the training set, **L test 1** is test sets with similar (to L28at14) operating conditions and **L test 2** are data sets with different operation conditions. The best performing model in each row is marked in bold face.

Model	0	1	2	3	4	5	6
L28at14	730	764	789	805	826	865	**899**
L test 1	16685	17049	17352	17506	17596	**17614**	11900
L test 2	14461	**14586**	14120	14309	14470	NA	9866
L test all	31146	31635	31472	31815	**32066**	NA	21765