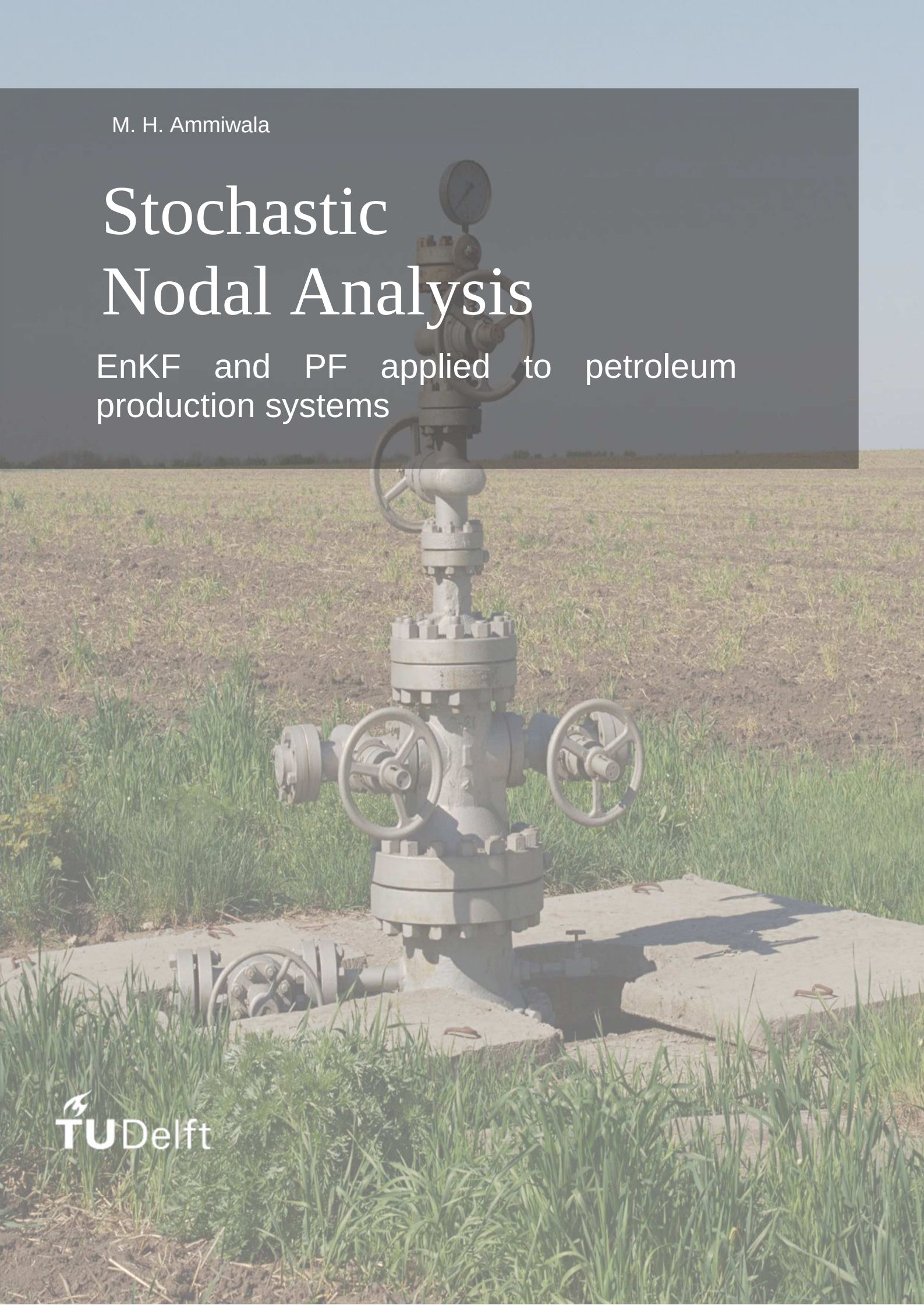
M. H. Ammiwala

Stochastic Nodal Analysis

EnKF and PF applied to petroleum production systems



TU Delft

Stochastic Nodal Analysis

EnKF and PF applied to petroleum production systems

By

M.H. Ammiwala

in partial fulfilment of the requirements for the degree of

Master of Science
in Applied Earth Sciences

at the Delft University of Technology,
to be defended publicly on Thursday August 31, 2017 at 10:30 AM.

Supervisor:	Prof. dr. ir. J.D. Jansen	TU Delft
Thesis committee:	Prof. dr. ir. A.W. Heemink,	TU Delft
	Dr. ir. F.C. Vossepoel,	TU Delft

TU Delft

Abstract

A petroleum production system is generally modelled based on the concept of nodal analysis, where the entire system is broken down into discrete elements such as near-well bore, tubing, surface choke and flow line. Operating flow rates and pressures can be estimated with a nodal analysis procedure by calculation of the intersection of performance curves. Input parameters in nodal analysis of production systems are considered deterministic, however, some of these parameters are better represented as distributions. In this report, the ensemble-based data assimilation methods "ensemble Kalman filter" (EnKF) and "particle filter" (PF) are applied to steady-state models of a production system for tuning of uncertain model parameters during the test separator phase. The performance of the EnKF and the PF is tested with the use of twin experiments. The calibrated model parameters of the choke, tubing and the near-well bore elements with EnKF and PF can be used to create an ensemble of performance curves leading to an ensemble of operating flow rates and pressures. The foreseen next step is to use the posterior distributions of model parameters as inputs for soft sensing of flow rates during semi-steady-state production for a single phase oil reservoir, where the oil rate and reservoir pressure are considered as unknown parameters. In the twin experiments as used in this thesis, a total number of three steady-state pressure drop measurements was used to estimate a total of six independent parameters which constitutes an ill-posed problem, resulting in non-unique parameter estimates. It is recommended to alleviate this issue by either reducing the number of parameters or by using multiple separator tests at different flow rates.

Acknowledgements

I owe my deepest gratitude to my supervisor Prof. J.D. Jansen for his continuous support and encouragement throughout the challenging period of this study. The weekly meetings, discussions, and feedback helped me to remain motivated and focused. Under his supervision I have acquired skills which have not only helped me in this study but will always be useful in my professional career and personal life. I am grateful to him for introducing me to this interesting subject of applying data assimilation to production systems.

I also express my gratitude to Dr. F.C. Vossepoel for her input in this study, specially about the application of the particle filter. Her support helped me a lot to understand the basic concepts of data assimilation. I would take this opportunity to also thank Dr. P.J van den Hoek, whose field experience helped me to formulate the problem with better understanding.

This academic achievement would not have been possible without the unconditional support of my parents and my family. I would like to thank them for their love and believe in me. I would also like to thank Delft University of Technology (TU Delft) for giving me an opportunity to be mentored by learned professors and meeting wonderful students.

Finally, I would like to thank all the friends I have made during my time at the TU Delft with whom I shared many memorable experiences. I will always cherish the good memories in times to come.

M.H. Ammiwala
Amsterdam, Augustus 2017

TABLE OF CONTENTS

1 INTRODUCTION

A petroleum production system is generally modelled based on the concept of nodal analysis, where the entire system is broken down into discrete elements such as near-well bore, tubing, surface choke and flow line [1]. Nodal analysis enables one to optimize the design of production facilities in field development by calculation of flow rates and pressures for different system configurations. Nodal analysis could also be used during operation of the field to interpret production measurements. The involved models give a steady-state description of the pressure drop across each element for given flow rates of oil, gas and water while assuming all other parameters as known values. Some parameters such as well length, choke diameter of recently placed choke and initial reservoir pressure are known values or could be measured with good accuracy. However, other parameters such as the reservoir pressure during production, permeability and model parameters are much more uncertain.

Data assimilation, also known as “computer-assisted history matching” is a frequently used technique in weather forecasting, oceanography [2] and reservoir engineering [3] to update uncertain model parameters with incoming measurements. Data assimilation consist of three components: (1) measurements, (2) a model and (3) a data assimilation method.

Data assimilation has already been applied to discrete elements of a production system (e.g. multi-phase flow in wells [4]) but has not been applied to an entire production system. Ensemble-based data assimilation methods, such as the Ensemble Kalman filter (EnKF) and the particle filter (PF), create multiple realisations of the output of the model (i.e. of the predicted measurements) by sampling uncertain model parameters and measurement errors based on prior distributions of those uncertain parameters and errors. Subsequently, the mismatch between the predicted measurements, based on the model, and the actual measured values is minimized in some averaged sense and used to update the distribution of the uncertain parameters with EnKF and PF. Assessment of the uncertain parameters is done by comparing the prior and posterior distributions of the uncertain parameter, aiming for a posterior distribution that has a smaller spread, i.e. smaller variance of the uncertain parameters.

The scope of this research is to apply the stochastic data assimilation methods EnKF and PF to a petroleum production system based on production data such as pressures and flow rates of oil, gas and water. We envisage that the data assimilation methods EnKF and PF could be applied for two production phases in a production system consisting of, e.g., the elements choke, tubing and near-well bore.

In the first production phase, called the "test separator phase", flow rates of gas, oil and water and pressures at different locations like at the bottom of the well, at the tubing head and the downstream part of the surface choke are measured. The goal of the "test separator phase" is to tune the uncertain model parameters. Earlier examples of model parameter tuning for production system models include the updating of a drift flux

multi-phase flow model with EnKF [5], and the updating of drift-flux model parameters on real data with the extended Kalman filter [6].

In the second production phase, called the "flow rate estimation phase", the measurements of the pressure are available, while the flow rates are estimated with EnKF and PF. This type of flow rate estimation, also called “soft sensing”, has already been applied with the auxiliary particle filter for transient well flow models [4, 7] and for steady-state models in wells by [8-10] based on pressure and/or temperature measurements. The focus of this thesis is to investigate the application of the data assimilation methods EnKF and PF on a production system with twin experiments, i.e. test the performance of a data assimilation method in the absence of real data. The two goals for applying data assimilation for the steady-state models of the production system are 1) tuning of uncertain model parameters (during the test separator phase), and 2) soft sensing of flow rates (during the flow rate estimation phase). In this thesis, only the tuning of uncertain model parameters during the test separator phase will be discussed. Use of the results for flow rate estimation is recommend for future research.

In Chapter 2 details of the data assimilation methods EnKF and PF are described for the application to a petroleum production system. The production system used throughout this report consists of the elements: a surface choke, a vertical well and near-well bore. In Chapters 3, 4, and 5 data assimilation is applied to the discrete elements of the production system, i.e. the choke, tubing and near-well bore for the purpose of model parameter estimation during the test separator phase. Chapter 3 focuses on tuning of model parameters of the extended Gilbert equation [1, 11] which describes the pressure drop over the choke. In Chapter 4 data assimilation is applied for multi-phase flow through wells with the Mukherjee and Brill correlation [12], by tuning two introduced model parameters: correction factors for the gravity component and friction component of the pressure loss in wells. Chapter 5 explores data assimilation applied to the near-well bore element of the production system, taking into consideration the uncertain model parameter, which is chosen as a correction factor for the productivity index. In Chapter 6 data assimilation is applied, in the test separator phase, to the entire production system, i.e. the tubing, choke and near-well bore elements simultaneously. Chapter 7 concludes this thesis with a discussion and directions for future research.

2 DATA ASSIMILATION APPLIED TO PRODUCTION SYSTEMS

This chapter presents how two stochastic data assimilation methods known as the "ensemble Kalman filter" (EnKF) and the "particle filter" (PF) can be used for model parameter and flow rate estimation in a production system.

First, the classical use of nodal analysis to predict flow rates based on two fixed pressures in a production system is explained. Current deterministic nodal analysis models take all parameters and inputs as certain (i.e. deterministic) values. However, in reality, many of the parameters are uncertain (i.e. stochastic). Uncertain model parameters and measurement errors (input errors) can be described with probability distribution functions (pdfs) instead of with single values. Samples of these pdfs (also known as 'realizations') can be used to generate multiple model predictions of pressures and flow rates, i.e. ensembles of predicted values.

The second part of this chapter describes the stochastic data assimilation methods EnKF and PF. These methods update the pdfs of model parameters based on the mismatch between predicted measurements and actual measurements. The chapter concludes with a generic description of the application of EnKF and PF to a production system for two production phases: (1) the "test separator phase" and (2) the "flow rate estimation phase". Details of the application are then discussed in the following chapters.

2.1 Stochastic nodal analysis

Nodal analysis is a commonly used technique to compute the pressures and flow rates in an oil and gas production system [1]. It represents the production system as a group of elements in between nodes. A simple production system consisting of a choke, a tubing and a near-well bore element are used in this report; see Figure 2.1. The pressure drop over an element is described by a mathematical function g which calculates the output pressure p_{out},

$$p_{out} = g(p_{in}, q_{g,sc}, q_{o,sc}, q_{w,sc}), \tag{2.1}$$

starting from an input pressure p_{in} and gas, oil and water flow rates at standard conditions $q_{g,sc}$, $q_{o,sc}$, and $q_{w,sc}$. The pressure drop over an element is typically calculated for known values of the three flow rates (or, alternatively, the oil flow rate, gas/oil ratio (GOR) and water cut). However, often the pressures at two points in the system are known (e.g. from measurements), whereas the oil flow rate is unknown. In that case the nodal analysis procedure has to be repeated in an iterative fashion (changing the oil flow rate, for given values of the GOR and water cut until the pressure drop matches the difference between the known pressures.) Alternatively, the pressure drop over two system elements, one upstream and one downstream of a connecting "analysis" node is computed for a large number of oil flow rates, and the flow rate at which the pressures at the analysis node match is then the output of the nodal analysis procedure [1].

With the nodal analysis procedure the oil rate $q_{o,sc}$ in the simple production system displayed in Figure 2.1 can be calculated for given values of reservoir pressure p_R and flow line pressure p_{fl}.

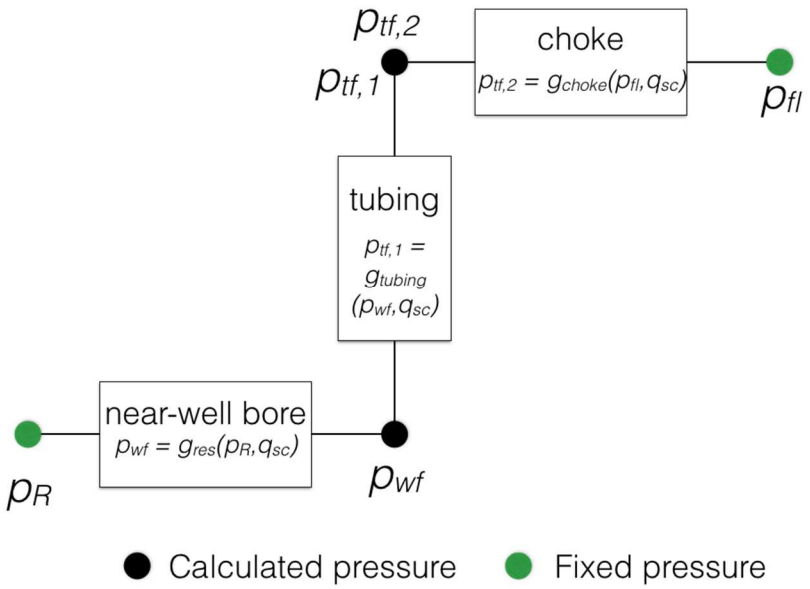


Figure 2.1: Nodal analysis configuration of a production system with fixed reservoir pressure p_R and flow line pressure p_{fl} with analysis node at the tubing head. The bottom hole pressure p_{wf} and two values of the tubing head pressure, $p_{tf,1}$ and $p_{tf,2}$ are calculated for different flow rates q_{sc}: one for the system elements upstream of the analysis node, and one for the downstream part of the system. The output flow rate corresponds to the value for which the two tubing head pressures become equal.

Suppose the analysis node is the tubing head node, see Figure 2.1. Two values of the tubing head pressure, $p_{tf,1}$ and $p_{tf,2}$, are calculated, each for a range of different oil rates, resulting in a "tubing performance curve" and a choke performance curve"; see Figure 2.2. The calculated tubing head pressure $p_{tf,1}$ is a function of the pressure loss over the near-well bore area and the tubing, starting from a fixed reservoir pressure p_R. The calculated tubing head pressure $p_{tf,2}$ is a function of the pressure loss over the choke, starting from a fixed flow line pressure p_{fl}. With the choke performance curve and the tubing performance curve plotted in Figure 2.2, the operating point at the tubing head can be determined as the intersection of the two curves.

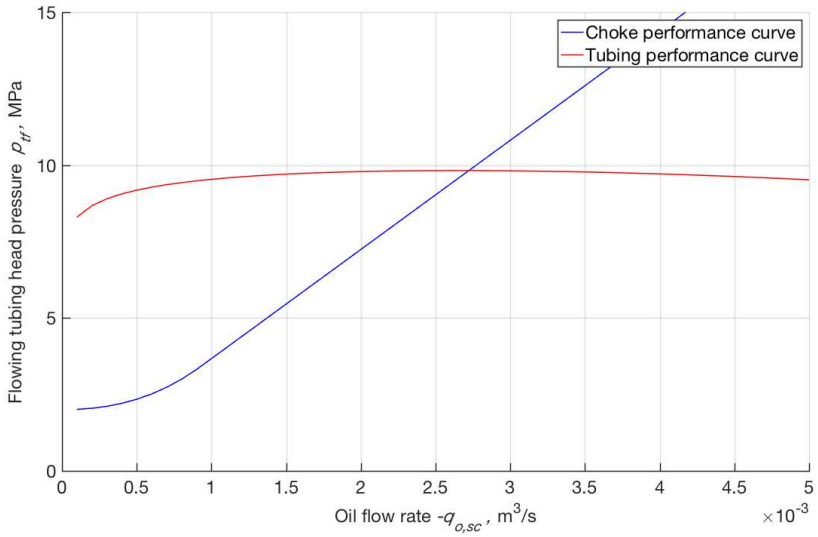


Figure 2.2: Tubing performance curve and the choke performance curve. The intersection between performance curves gives operating oil rate and tubing head pressure.

The performance curves in Figure 2.2 are based on two fixed pressures and consider all other parameters as deterministic parameters. However, many of these parameters are stochastic and could better be described with a pdf. This implies that the output pressure will also have a pdf which is a function of the pdfs of the uncertain parameters. Because there are multiple uncertain parameters in a production system, and because the functions g describing the pressure drops are nonlinear, closed-form calculation of the pdf of the output pressure is normally not possible. A different approach to represent the output pdf is with a discrete distribution rather than a continuous probability distribution. This is done by random sampling of the uncertain parameters from the pdf, to obtain "realizations" of outputs. In Figure 2.3, realizations of the choke performance curve and tubing performance curve are displayed, which are based on samples from the pdfs of uncertain parameters. This leads to an ensemble of possible operating pressures and flow rates as output.

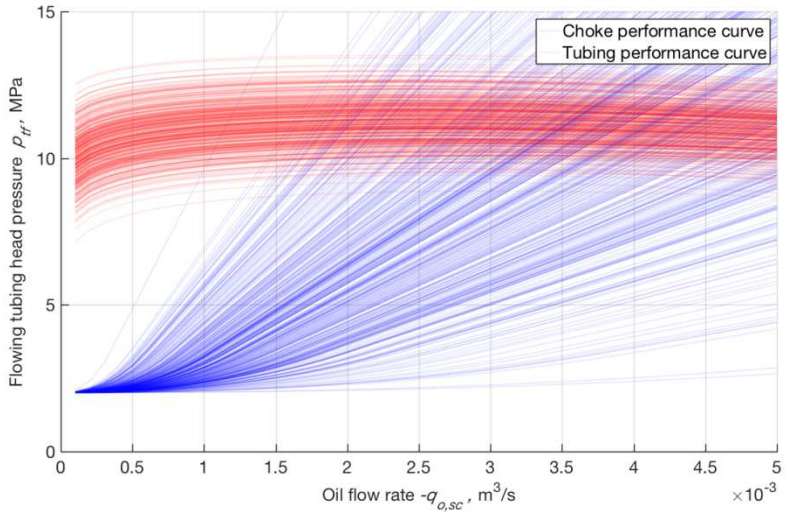


Figure 2.3: Realizations of choke performance curves and tubing performance curves leading to multiple operating flow rates and pressures.

2.2 Stochastic data assimilation methods EnKF and PF

The distribution of the uncertain parameters can be updated with one of the ensemble-based data assimilation methods "ensemble Kalman filter" (EnKF) and "particle filter" (PF). Ensemble-based data assimilation methods are typically used because of their easy implementation [13].

2.2.1 Definitions

Ensemble-based sequential data assimilation methods are typically used for dynamic models in which the model parameters are defined as static (i.e. fixed in time), and the input and output variables as dynamic (i.e. varying over time) [14]. In our case (semi-)steady-state models are used to describe the pressure drop, which implies that the input and output variables either do not change over time or, in the semi-steady-state case, decline very slowly at a constant rate. The following concepts are introduced for our (semi-)steady-state data assimilation problem:

1. Model parameters – each mathematical function in an element requires the definition of model parameters. Some of these model parameters could be seen as deterministic values, e.g. fluid properties, well length, or well radius. However, other model parameters are subject to more uncertainty, and could therefore be

better described as stochastic. Stochastic model parameters are updated with data assimilation and are stored in the uncertain model parameter vector **m**. Their initial distribution is referred to as the "prior distribution"; their updated distribution is referred to as the "posterior distribution".

2. Model inputs – together with the model parameters, these variables are required to calculate the modelled output. The model inputs are often measurements containing measurement errors with an (assumed known) pdfs. Model inputs are stored in a model input vector **u** and are not updated during the data assimilation procedure. (Sometimes inputs may therefore also be stochastic model parameters that have been updated in a previous data assimilation step, but are now left unchanged.)
3. Modelled outputs – depend on the model inputs and the model parameters. The uncertain inputs **u** and uncertain parameters **m** lead to multiple realizations of the modelled output stored in a model output vector **y,** see Figure 3. Inputs in a simple 'nodal' production system model are typically the flow rates $q_{o,sc}$, $q_{g,sc}$ and $q_{w,sc}$ and the input pressure p_{in}, to calculate the output p_{out}.
4. Measured outputs – the mismatch between actual measured values and predicted values is used to update the uncertain model parameters. The actual measured values are subject to uncertainty due to measurement errors. They are stored in a measured output vector **d**. The mismatch is therefore defined as **(y – d)**.

The PF and the EnKF make use of Monte Carlo sampling of the distributions of uncertain inputs, model parameters and measurement errors. The samples from the distributions are often called "realizations", "ensemble members" or "particles" depending on the data assimilation method used. Equation 2.1 can be rewritten for data assimilation purposes as

$$\mathbf{y} = \mathbf{g}(\mathbf{u}, \mathbf{m}), \tag{2.2}$$

where the vector-valued forward model **g** calculates the outputs **y** with a realization of input parameters **u** and model parameters **m.** (We use bold lower-case letters to indicate vectors and bold capitals to indicate matrices). Suppose that there are N_m uncertain model parameters defined and that each uncertain model parameter is initially randomly sampled from independent Gaussian distributions:

$$\mathbf{m}^i = \left[m_1^i \sim N\left(\mu_{m,1}, \sigma_{m,1}^2\right) \cdots\ m_j^i \sim N\left(\mu_{m,j}, \sigma_{m,j}^2\right) \cdots\ m_{N_m}^i \sim N\left(\mu_{m,N_m}, \sigma_{m,N_m}^2\right)\right]^T, \tag{2.3}$$

where $\mathbf{m}^i$ is the i^{th} realization of model parameter vector **m**, and $\mu_{m,j}$ and $\sigma_{m,j}^2$ are the mean and variance of model parameter m_j. The prior model parameter matrix,

$$\mathbf{M}_{prior} = \left[\mathbf{m}^1\ \mathbf{m}^2 \cdots\ \mathbf{m}^i \cdots \mathbf{m}^{N_r}\right], \tag{2.4}$$

contains N_r realizations of the model parameters. Realizations of N_u measured inputs are defined as

$$\mathbf{u}^i = \left[u_1^i\ \ u_2^i \cdots\ u_j^i \cdots\ u_{N_u}^i\right]^T, \tag{2.5}$$

where a realization for each input u_j is defined as,

$$u_j = u_{j,obs} + \epsilon_{u,j}\,, \tag{2.6}$$

where $u_{j,obs}$ is the measurement of the input j, and $\epsilon_{u,j}$ is a randomly sampled measurement error from the Gaussian distribution,

$$\epsilon_{u,j} \sim N\left(0, \epsilon_{u,j}^2\right), \tag{2.7}$$

with a zero mean and variance $\epsilon_{u,j}^2$, which is assumed to be known. The input matrix,

$$\mathbf{U} = [\,\mathbf{u}^1\ \mathbf{u}^2\ \cdots\ \mathbf{u}^i\ \cdots\ \mathbf{u}^{N_r}], \tag{2.8}$$

contains N_r realizations of the uncertain input parameters. Similarly, the modelled output matrix is defined as

$$\mathbf{Y} = [\,\mathbf{y}^1\ \mathbf{y}^2\ \cdots\ \mathbf{y}^i\ \cdots\ \mathbf{y}^{N_r}], \tag{2.9}$$

for N_r realizations of the output vector,

$$\mathbf{y}^i = \left[y_1^i\ y_2^i \cdots y_j^i \cdots y_{N_y}^i\right]^T, \tag{2.10}$$

where each realization,

$$\mathbf{y}^i = \mathbf{g}\left(\mathbf{u}^i, \mathbf{m}^i\right), \tag{2.11}$$

is calculated with forward model $\mathbf{g}$ with a sample of input parameters $\mathbf{u}$ and model parameter $\mathbf{m}$. The resulting parameter-output matrix is defined as,

$$\mathbf{Z}_{prior} = \left[\mathbf{z}^1\,\mathbf{z}^2 \cdots \mathbf{z}^i \cdots \mathbf{z}^{N_p}\right] = \left[\mathbf{M}_{prior}^T\,\mathbf{Y}^T\right]^T\,, \tag{2.12}$$

with parameter-output vectors defined as

$$\mathbf{z} = [\mathbf{m}^T\ \mathbf{y}^T]^T, \tag{2.13}$$

containing realizations of model parameters and modelled outputs. The measurement matrix is defined as,

$$\mathbf{D} = \left[\mathbf{d}^1\ \mathbf{d}^2 \cdots\ \mathbf{d}^i \cdots \mathbf{d}^{N_r}\ \right], \tag{2.14}$$

where

$$\mathbf{d}^i = \left[d_1 \cdots d_j \cdots d_{N_d}\right]^T, \tag{2.15}$$

contains N_r realizations of measurements for N_d pressure measurements. Each realization of a measurement is defined as,

$$d_j = d_{j,obs} + \epsilon_{d\,j}, \tag{2.16}$$

where $d_{j,obs}$ is the observed value and the measurement error ϵ_{d_j} is randomly sampled from a Gaussian distribution,

$$\epsilon_{d_j} \sim N\left(0, \epsilon_{d_j}^2\right), \tag{2.17}$$

with a zero mean and variance of ϵ_d^2, the measurements errors which are stored in

$$\boldsymbol{\epsilon}_d = \left[\epsilon_{d,1}\epsilon_{d,2} \cdots \epsilon_{d,j} \cdots \epsilon_{d,N_d}\right]^T. \tag{2.18}$$

The measurement errors are assumed to be known and uncorrelated leading to a diagonal covariance matrix of the measurements:

$$\mathbf{C}_d = diag\left(\left[\epsilon_{d,1}^2\epsilon_{d,2}^2 \cdots \epsilon_{d,j}^2 \cdots \epsilon_{d,N_d}^2\right]\right). \tag{2.19}$$

Sequential data assimilation methods require an initial distribution of the uncertain model parameters $\mathbf{m}_0$ which is used as the prior distribution $\mathbf{m}_{prior}$ for the first data assimilation step. During each data assimilation step, the distribution of the model parameters is updated with incoming data.

2.2.2 Particle filter (PF)

The particle filter is preferred because of the easy implementation and because it can cope with nonlinear systems and non-Gaussian pdfs [15]. Like all ensemble-based methods, the prior distribution of the modelled output is based on sampling of the uncertain model parameters. To obtain the exact posterior distributions of the model parameters, large amounts of particles (i.e. realizations) are needed to create every possible combination of model parameters and inputs to calculate the modelled output $\mathbf{y}$. This is computationally challenging and becomes difficult when the number of the uncertain model parameters N_m increases. In this thesis the number of uncertain model parameters is low compared to geophysical systems [16]. However, the uncertainty variance in model parameters leads to a large spread in predicted outputs, see, e.g,. Figure 2.3. Each data assimilation problem needs as specific implementation of the particle filter ranging from basic importance sampling (SIS), to the auxiliary particle filter (ASIR). The most basic implementation of the particle filter is SIS, where only the relative weight of each particle is changed.

The application for updating uncertain model parameters is different from the approach for a dynamic model for state estimation described in [15, 16]. In dynamic models states are estimated sequentially in time, i.e. the states of the current time step are predicted by a dynamic model based on the state of the previous time step and parameters. With the particle filter, the predicted states are updated based on measurements of the current time step. In our case, however, the particle filter is applied to update uncertain model parameters of a steady-state model. The uncertain model parameters $\mathbf{m}$ and uncertain inputs $\mathbf{u}$ are used in a forward model to calculate output $\mathbf{y}$ and stored in the parameter-output vector $\mathbf{z}$. Thus, the forward model calculates outputs at the same time step, instead of first performing a forward integration in time as necessary in state estimation. The posterior distribution of the parameter-output vector is approximated with the particle filter, which is based on Bayes' theorem,

$$P(\mathbf{z}|\mathbf{d}) = \frac{P(\mathbf{d}|\mathbf{z})P(\mathbf{z})}{P(\mathbf{d})}, \tag{2.20}$$

where the probability of the observations given a particle of parameter-output vector $P(\mathbf{d}|\mathbf{z})$ is multiplied with the prior distribution of the model $P(\mathbf{z})$ divided by the probability of the data $P(\mathbf{d})$ [16]. The probability of the samples is represented as,

$$P(\mathbf{z}) = \frac{1}{N_r}\sum_{i=1}^{N_r} \delta(\mathbf{z} - \mathbf{z_i}), \tag{2.21}$$

where N_r is the number of particles and $\mathbf{z}$ is a particle containing uncertain model parameters $\mathbf{m}$ and modelled outputs $\mathbf{y}$. Because of random sampling, the initial weight of each particle will be equal to $1/N_r$. The probability of the measurements can be calculated with

$$P(\mathbf{d}) = \int P(\mathbf{d}|\mathbf{z})P(\mathbf{z})d\mathbf{z}, \tag{2.22}$$

and is referred to as the normalization constant [17] used to normalize the probabilities. The posterior distribution in the particle filter is calculated with Equation 2.20 and 2.21,

$$P(\mathbf{z}|\mathbf{d}) = \sum_{i=1}^{N} w_i\, \delta(\mathbf{z} - \mathbf{z}_i), \tag{2.23}$$

with the weight of a particle defined as,

$$w_\mathrm{i} = \frac{P(\mathbf{d}|\mathbf{z}_i)}{\sum_{j=1}^{N} P(\mathbf{d}|\mathbf{z}_j)}, \tag{2.24}$$

and the probability of the observations given a set of the uncertain parameters $P(\mathbf{d}|\mathbf{z}_i)$ is usually calculated with a Gaussian distribution [16],

$$P(\mathbf{d}|\mathbf{z}_i) = \frac{1}{\sqrt{2\pi^{N_d}\,|\mathbf{C}_d|}}\; e^{-\frac{1}{2}(\mathbf{d}-\mathbf{y}_i)^T\mathbf{C}_d^{-1}(\boldsymbol{d}-\boldsymbol{y_i})}, \tag{2.25}$$

where $\mathbf{C}_d$ is the covariance matrix of the measurement errors (Equation 18), and $|\mathbf{C}_d|$ is the determinant of $\mathbf{C}_d$. In basic importance sampling (SIS), weights are assigned to each particle $\mathbf{z}$.

An example of the application of the particle filter for model calibration with basic importance sampling (SIS) is shown in Figure 2.4. The first step is to forecast predictions of $\mathbf{y}$ with Equation 2.11, i.e. particles from prior distributions of $\mathbf{m}$ and $\mathbf{u}$ give different outcomes of $\mathbf{y}$. The relative weight of each particle is calculated with Equation 2.23. The posterior distribution of the model parameter and input can be represented by plotting the values of the sampled particles and the weights of those particle, see Figure 2.4. Weights of the particle is based on calculation of the conditional probability of observation $\mathbf{d}$ given for each realization (particle) of parameter-output vector $\mathbf{z}$, Equation 2.25, which are normalized with Equation 2.24.

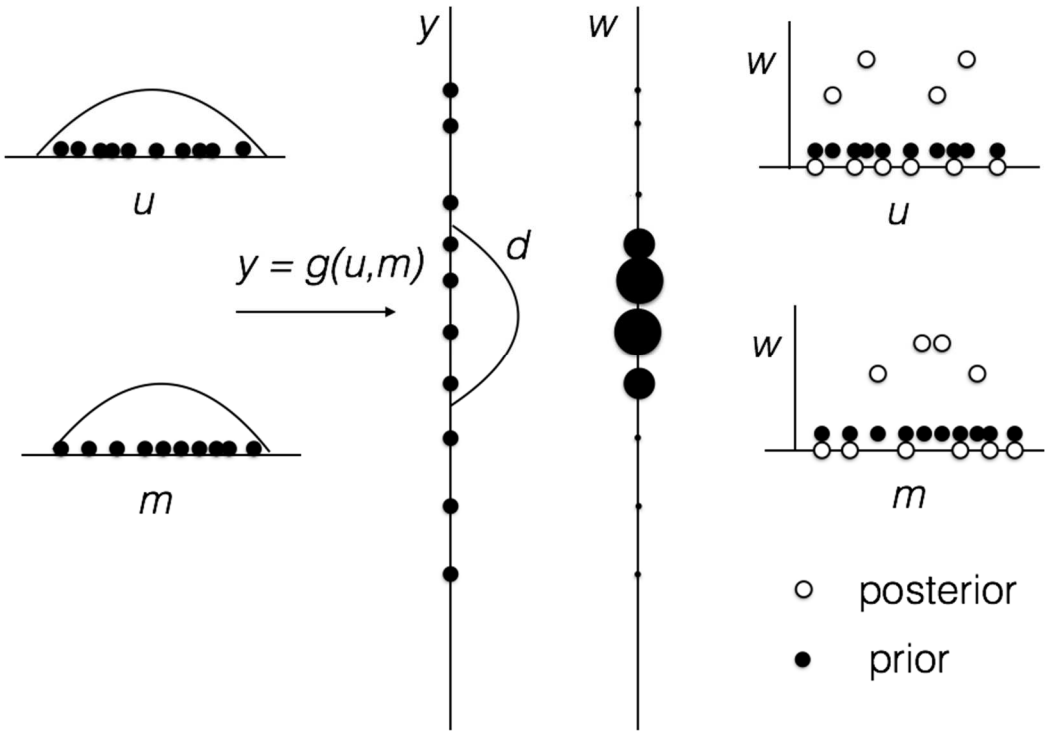


Figure 2.4: Sequential importance sampling (SIS) applied for model parameter tuning. Particles are generated based on prior distributions of u and m. Output y is calculated for each particle with g (Equation 2.11) and gives a prior distribution of the model parameters. Weight to each particle are assigned by Equation 2.24 The posterior distribution of m and u is represented by assigning the calculated weights to the particles.

The weights assigned to the particles can be used to obtain updated statistical information, like the mean with

$$\overline{\mathbf{z}_{\boldsymbol{PF}}} = \sum_{i=1}^{N} w_i \mathbf{z}_{\mathrm{i}} = \mathbf{Z}_{prior}\, \mathbf{w}, \tag{2.26}$$

and the covariance of the posterior distribution with

$$\mathbf{C}_{z,PF} = \frac{\Delta \mathbf{Z}_{PF}\, \Delta \mathbf{Z}_{PF}{}^{T}}{N_p - 1}, \tag{2.27}$$

where $\Delta \mathbf{Z}_{\mathrm{PF}}$ is the perturbation matrix defined as

$$\Delta \mathbf{Z}_{PF} = \mathbf{Z}_{prior} - \overline{\mathbf{Z}_{PF}} = [(\mathbf{z_1} - \overline{\mathbf{z}_{\boldsymbol{PF}}}) \cdots (\mathbf{z_i} - \overline{\mathbf{z}_{\boldsymbol{PF}}}) \cdots (\mathbf{z}_{\mathbf{N_p}} - \overline{\mathbf{z}_{\boldsymbol{PF}}})]. \tag{2.28}$$

A common problem of the particle filter occurs when only a limited number of particles have a non-zero weight and the ensemble effectively becomes very small. This is referred as filter degeneracy [16]. When the spread of ensemble of realizations of output **y** is large compared to the error variance of the measurement **d,** this leads to only a limited number of particles with weights assigned. This is illustrated in the example from Figure 2.4. In SIS the particles sampled from the prior distribution are not changed during the data assimilation, only a relative weight w_i is assigned to them. When filter degeneracy occurs, a large number of particles are assigned a zero weight and are not taken into account for the estimation. A possible approach to prevent this is to represent the non-zero weighted particles as a distribution where multiple copies of particles with non-zero weights are resampled, this leads to a posterior distribution where the probability of each particles is

$1/N_r$. In this thesis resampling techniques residual resampling and regularization are applied to obtain a posterior distribution of the model parameters. Including the resampling step is referred as sequential importance resampling (SIR) [16]. An implementation of the auxiliary particle filter (ASIR) is discussed alongside with the particle filter with the two SIR implementations.

Sequential importance resampling (SIR)

A workflow to obtain the posterior distribution of the model parameters could be described by:

1. Make N_r independent samples of inputs $\mathbf{u}$ and model parameters $\mathbf{m}$, see Section 2.2.1.
2. Use realizations of $\mathbf{u}$ and $\mathbf{m}$ to calculate predicted measurements $\mathbf{y}$, see Equation 2.11.
3. Calculate weights of each realization of parameter-output vector $\mathbf{z}$ Equation 2.24.
4. The obtained distribution is the outcome of SIS, when the variance of output $\mathbf{y}$ is much larger than the error variance of $\mathbf{d}$, filter degeneracy will occur. With step 5, a posterior distribution of model parameters can be represented without any particles with a zero probability.
5. An posterior distribution of model parameters $\mathbf{m}_{posterior}$ can obtained with resampling methods. In this thesis residual resampling and regularization are applied, this additional resampling step differentiates SIS from SIR.

In residual resampling the particles with non-zero weights are copied, i.e. sampling noise is reduced [18]. The amount of copies of a particle depends on the weight, i.e. higher weight leads to more copies of particles in posterior distribution. The advantage of using residual resampling that the posterior (resampled) distribution will closely resemble the posterior distribution of SIS. The disadvantage of using residual resampling is when filter degeneracy occurs. Due the fact of limited particles with weights, it is highly likely that a misrepresented posterior distribution will be obtained with SIS. Which potentially leads to a overestimation of some model parameters in the posterior distribution when resampled with residual resampling, see Figure 2.5. In residual resampling, the weights $\mathbf{w}$ are multiplied by the number of realizations N_r [16]. Particles of the uncertain model parameters are copied n times, where n is the integer part of $N_r\mathbf{w}$. The sum of the integer part of $N_r\mathbf{w}$ will not add up to N_r, i.e. remaining particles need to be defined to get N_r particles. The remaining particles are drawn randomly from the distribution obtained by subtracting the integer parts of $N_r\mathbf{w}$ from $N_r\mathbf{w}$.

In addition to the misrepresentation of the posterior distribution, residual resampling could also potentially lead to filter divergence, i.e. when particles in the forward step do not result in different realizations of the output. This is typically the case for most deterministic models like the one under consideration here. Having multiple copies of particles with the same values of $\mathbf{m}$ could lead to filter divergence when the distribution of $\mathbf{u}$ does not have a large impact on the calculated output. Instead of discrete resampling from the weights as in residual resampling, regularization could be applied by resampling from a continuous distribution around each particle [16]. In this thesis, regularization is applied by sampling from a Gaussian distribution, with as mean the particle obtained from residual resampling and a decreased standard deviation based on Equation 2.27. This decreased standard deviation is to ensure that not too much sampling noise is added to the posterior distribution of model parameters. The advantage of using regularization is

that the added noise will increase the model parameter space, i.e. more variance in the posterior distribution, see Figure 2.5. The added noise decrease the chances of filter divergence occurring. The disadvantage is that with the added noise the posterior distribution is changed (compared to SIS), and potentially too much noise is added.
SIR can suffer from two shortcomings depending on the system under consideration: (1) if the number of particles is small, outliers could have an uneven distribution in the posterior distribution, i.e. filter divergence and (2) the posterior distribution is not guided by future observations [19].

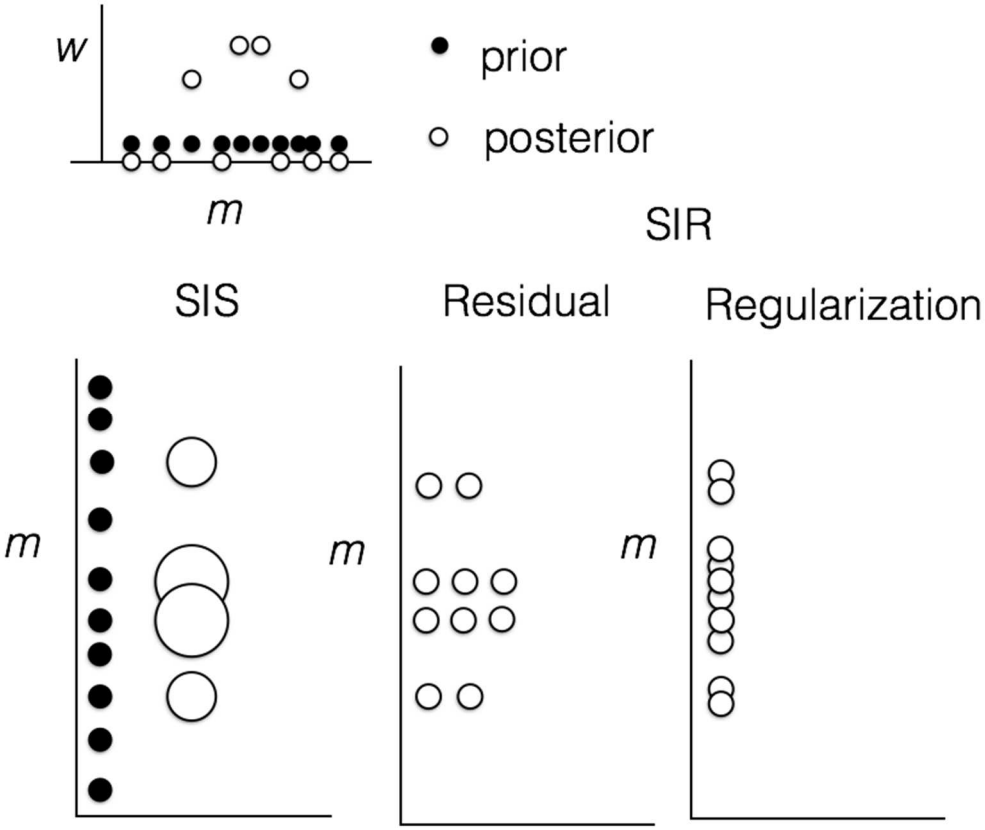


Figure 2.5: Resampling of the obtained weights of the particles with sequential importance sampling (SIS) (left figure). Sequential importance resampling (SIR) is applied with residual resampling (middle figure) and regularization (right figure).

Auxiliary particle filter (ASIR)

The auxiliary particle filter (ASIR) was introduced to take the measurements into account to estimate the posterior distribution [19]. In this thesis, a slightly different approach of ASIR is implemented which tries to tackle possible shortcomings of the SIR. The first shortcoming is the problem with outliers in the posterior distribution when filter divergence occurs, i.e. more likely to happen with residual resampling. This motivated the choice for applying regularization as resampling method. However, sampling noise is introduced with regularization, which could lead to a posterior distribution with too much variance. In the ASIR, the posterior distribution of the PF generated with regularization will be used to make predictions again at the same data assimilation step. The predicted measurements based on the posterior distribution will have a much smaller variance compared the predicted measurements obtained with the original prior distribution, this leads to less filter degeneracy when applying Equation 2.24 to the predicted measurement based on the posterior distribution of regularization. The ASIR used can be described by:

1. Use predictions $\mathbf{y}$ (based on $\mathbf{m}_{prior}$) and measurements $\mathbf{d}$ and SIR with regularization to obtain the posterior distribution based on SIR, $\mathbf{m}_{posterior,SIR}$, i.e. steps 1 – 5 of SIR.
2. Based on $\mathbf{m}_{posterior,SIR}$ and $\mathbf{u}$ predictions of $\mathbf{y}_{SIR}$ are made (note the predictions are made for the same data assimilation step), i.e. Equation 2.11 (note, the application of ASIR requires the forward model g to be run two times for the number particles N_r for each data assimilation step.).
3. The prior distribution $\mathbf{m}_{posterior,SIR}$ is resampled with regularization to generate the posterior distribution $\mathbf{m}_{posterior,ASIR}$.

The advantage of using $\mathbf{m}_{posterior,SIR}$ to make predicted measurements is that the variance will be much smaller compared to making predicted measurements based on $\mathbf{m}_{prior}$. Which will result in much more particles with an assigned weight leading to a better estimation of the posterior distribution of model parameters. However, the prominent disadvantage is that the model needs to be run twice. With ASIR a more accurate representation of the model parameters is given, in the first data assimilation step model parameters are filtered out, where after sampling noise is introduced to expand the search area near the combinations of model parameters which caused high weights with regularization, i.e. add Gaussian noise to the posterior distribution obtained with residual sampling. By using this posterior distribution 'again' as prior distribution will lead to predicted measurements, i.e. prior $\mathbf{y}_{SIR}$, which will have less variance compared to the predicted measurements $\mathbf{y}$. Because the variance is narrowed down, more particles will be assigned a weight leading to a more accurate representation of the posterior distribution with ASIR.

2.2.3 Ensemble Kalman filter (EnKF)

The ensemble Kalman filter (EnKF) is a sequential data assimilation method widely used in oceanography, reservoir history matching and weather forecasting. EnKF is a Kalman filtering technique and was put forward as an improvement of the extended Kalman filter (EKF) [13]. For more information about EnKF, see e.g. [20, 21]. EnKF is preferred as data assimilation algorithm because it is a sequential data assimilation method, which is easy to implement and is suitable for large number of variables [3].

Kalman filtering techniques are applied in two steps, the forecast step and the analysis step. In the forecast step, the prior ensemble members of model parameters and inputs are used to predict the output $\mathbf{y}$, see Equation 11. The prior parameter-output matrix $\mathbf{Z}_{prior}$ contains ensemble members, see Equation 12. In the analysis step the posterior distribution is obtained with a linear update step,

$$\mathbf{Z}_{posterior} = \mathbf{Z}_{prior} + \mathbf{K}\,(\mathbf{D} - \mathbf{Y}), \tag{2.29}$$

where $\mathbf{K}$ is the Kalman gain matrix, $\mathbf{D}$ is the matrix of observations (Equation 2.17), and $\mathbf{Y}$ contains modelled output based on the forecast step (Equation 2.12) and the prior parameter-output matrix. The Kalman gain is defined as,

$$\mathbf{K} = \mathbf{C}_z\mathbf{H}^T(\mathbf{H}\,\mathbf{C}_z\,\mathbf{H}^T + \mathbf{C}_d)^{-1}, \tag{2.30}$$

where $\mathbf{C}_d$ is the covariance of the data, $\mathbf{C}_z$ is the covariance matrix of the system and $\mathbf{H}$ is the measurement operator which used to obtain the prior modelled output from the parameter-output vector $\mathbf{z}$,

$$\mathbf{Hz} = \mathbf{y} \tag{2.31}$$

It can be shown that Equation 2.29 represents a weighted update of the prior based on the data mismatch ($\mathbf{D} - \mathbf{Y}$) where the weight depends on the relative errors in the prior and the data as indicated by the definition of $\mathbf{K}$ in Equation 2.30: for a prior with small uncertainty (i.e. small values of the elements of matrix $\mathbf{C}_z$) and data with a large uncertainty (i.e. large values of the elements of $\mathbf{C}_d$) the updates will be small. For the for a prior with large uncertainty and data with a small uncertainty the reverse is true. In EnKF, the covariance matrix of the system is calculated with,

$$\mathbf{C}_z = \frac{\Delta\mathbf{Z}_{prior}\,\Delta\mathbf{Z}_{prior}{}^{\boldsymbol{T}}}{N_r - 1}, \tag{2.32}$$

where N_r is the number of ensemble members and the $\Delta\mathbf{Z}_{prior}$ is the perturbation matrix calculated with,

$$\Delta\mathbf{Z}_{prior} = \mathbf{Z}_{prior} - \overline{\mathbf{Z}_{prior}}, \tag{2.33}$$

with,

$$\overline{\mathbf{Z}_{prior}} = [\bar{\mathbf{z}}^1\ \bar{\mathbf{z}}^2 \cdots \bar{\mathbf{z}}^i \cdots \bar{\mathbf{z}}^{N_P}], \tag{2.34}$$

where each element in $\overline{\mathbf{Z}_{prior}}$ contains the mean calculated based on the sampled distribution $\mathbf{Z}_{prior}$.

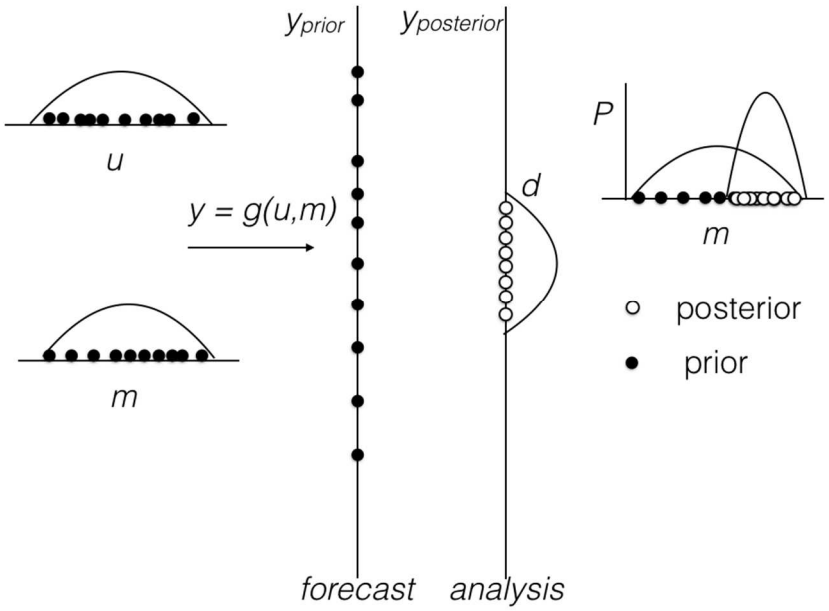


Figure 2.6: Schematic illustration of EnKF applied for parameter estimation. First step is to make ensemble members of $\mathbf{m}$ and $\mathbf{u}$, in this example there is one uncertain model parameter and one measured input u. In the forecast step predicted measurements of the ensemble members are calculated with the forward model, see Equation 11. In the Analysis step the ensemble members are updated based on the Kalman gain, see Equation 29 and 30. The posterior

distribution of m could be described as Gaussian distribution, and should show a smaller spread (when parameter m is correlated with y).

In Figure 2.6 an example of the EnKF is displayed for model parameter estimation. The difference between the ensemble-based EnKF and other Kalman filtering techniques such as KF and EKF is the calculation of the covariance matrix of the system (Equation 2.32). The Kalman filter (KF) is a sequential data assimilation method that works for linear models and assumes a Gaussian probability distribution of the prior [22]. EKF is applied for nonlinear systems, where the linear update step (Equation 2.29) is based on linearization of the nonlinear equations, which is computationally more demanding and more challenging to implement than EnKF [13]. EnKF is an ensemble based method which takes samples of the prior distributions for an approximation of $\mathbf{C}_z$ based on the prior distribution $\mathbf{Z}_{prior}$. EnKF makes use of the same analysis step as the original Kalman filter (Equation 2.29), however with an approximated $\mathbf{C}_z$ [21].

The matrix $\mathbf{P_z}$ gives correlation coefficients between uncertain model parameters $\mathbf{m}$ and modelled output $\mathbf{y}$ approximated in $\mathbf{C}_z$. The correlation between $\mathbf{m}$ and $\mathbf{y}$ is calculated with,

$$\mathbf{P}_z(z_i, z_j) = \frac{\mathbf{C_Z}(z_i,z_j)}{\sqrt{\mathbf{C_z}(z_i,z_i)\mathbf{C}_z(z_j,z_j)}}, \tag{2.35}$$

where $\mathbf{C}_z$ is the covariance matrix of the system, z_i and z_i are parameters of the parameter-output vector. The resulting matrix $\mathbf{P}_z$, contains correlation coefficient between -1 and 1 indicating the correlations. These correlations are used in the analysis step to update the sampled uncertain model parameters, see Figure 2.6.

A limitation of the EnKF is the requirement of a Gaussian distribution of the initial model parameters, which makes EnKF unsuitable for updating non-Gaussian distributions such as multi-model distributions [3]. The EnKF does not perform well for non-Gaussian distribution because the ensemble members are updated based on the mean and variance of the model in the analysis step [21]. When the model is highly non-linear, applying EnKF could lead to unsatisfactory results due to application of linear update on a nonlinear model. This could be solved with multiple data assimilations updates within a data assimilation step, see e.g. [23].

The main difference between the particle filter and the ensemble Kalman filter is the approach of approximation of the posterior distribution. Both data assimilation work with a forward step and an analysis step. In the analysis step of the PF, the mismatch between modelled output $\mathbf{y}$ and measurements $\mathbf{d}$ is used to assign weights to each particle. Whereas, EnKF updates the prior distribution based on statistical information of the system, i.e. ensemble members are updated based on the mismatch of $\mathbf{y}$ and $\mathbf{d}$ and the correlation of $\mathbf{y}$ with the uncertain model parameters $\mathbf{m}$. The practical difference between EnKF and PF is that the PF can cope with non-linear system and non-Gaussian distributions.

2.3 EnKF and PF applied to a production system

In this thesis, the application of the EnKF and PF methods is tested with twin experiments. A twin experiment is a means to test the performance of a data assimilation method in the absence of real data. Synthetic data generation gives control over the available

measurements. In this thesis, two production phases are introduced which differ in the available measurements and goal of applying data assimilation. The two production phases are defined as:

- 'Test separator phase' – in this production phase pressure and flow rate measurements are available. Flow rates of oil, gas and water are measured with a test separator. The pressures are measured at the tubing head and the bottom hole, see Figure 2.7. At start of production the reservoir pressure and downstream pressure of the choke are considered known. The primary goal of the test separator phase is to calibrate uncertain model parameters.
- 'Flow rate estimation phase' – in this production phase only pressure measurements are available. The pressure is measured at the tubing head and the bottom hole. The pressure at the downstream part of the choke is considered known and assumed to be below the critical pressure, i.e. the downstream pressure of the choke has no influence on the upstream pressure. The reservoir pressure is considered as an uncertain parameter. The choke diameter could change due to erosion, and should be considered as uncertain. The primary goal of this production phase is 'soft sensing' of the flow rates.

Furthermore, the following assumptions are made throughout this report:

- Single-phase oil reservoir, reservoir pressure well above bubble point.
- Multi-phase (oil and gas) flow in the well and choke.
- Semi-steady state decline of the pressure (constant pressure decline of the reservoir pressure).
- Fluid properties, such as density and viscosity, at standard conditions are fixed parameters.
- The downstream pressure of the choke has no influence on the upstream pressure of the choke, i.e. the ratio between the downstream and upstream pressure is lower than the critical pressure ratio.

The production system is described untill the downstream part of the choke. Due to the assumption of critical flow, the pressure at the downstream part of the choke has no influence on the pressures in the production system.

2.3.1 Production phase 1 – test separator phase

The production system of Figure 2.1 is displayed in Figure 2.7 regarding data assimilation terminology. During the 'test separator phase' the flow rates of oil, gas and water are measured with a test separator in addition to the pressure measurements. The measurements of the oil rate and gas rate will be used as uncertain input parameters. Samples of the inputs are made based on Equation 2.6. The three outputs in the configuration are,

$$\mathbf{y} = [y_1\ y_2\ y_3]^T = \left[p_{tf,choke}\ p_{tf,tubing}\ p_{wf}\right]^T, \tag{2.36}$$

with the tubing head and bottom hole pressures predicted based on uncertain model parameters $\mathbf{m}$ and uncertain inputs $\mathbf{u}$ and a forward model $\mathbf{g}$. The predicted pressure at the tubing head could be calculated by modelling the pressure drop over the choke with,

$$y_1 = g_{choke}(\mathbf{u}_{choke}, \mathbf{m}_{choke}), \tag{2.37}$$

where y_1 is the modelled output at the tubing head $p_{tf,choke}$. Note that the flow line pressure p_{fl} is not considered as input, in the critical flow regime, p_{fl} does not have an influence on

the modelled output pressure y_1. A different way to predict the tubing head pressure is by estimating the pressure drop over the tubing with,

$$y_2 = g_{tubing}(\mathbf{u}_{tubing}, \mathbf{m}_{tubing}), \tag{2.38}$$

where y_2 is the predicted output of the tubing head pressure $p_{tf,tubing}$, with measured oil rate, gas rate and bottom hole pressure d_2 as inputs $\mathbf{u}_{tubing}$ and uncertain model parameters $\mathbf{m}_{tubing}$. The pressure drop over the near-well bore element could be described with,

$$y_3 = g_{res}(\mathbf{u}_{res}, \mathbf{m}_{res}), \tag{2.39}$$

where y_3 is the predicted bottom hole pressure p_{wf}, with the measured oil rate and gas rate as uncertain inputs $\mathbf{u}_{res}$, the reservoir pressure as fixed input and $\mathbf{m}_{res}$ as uncertain model parameters.

In EnKF and PF, realizations of the predicted output $\mathbf{y}$ are based on samples of uncertain model parameter $\mathbf{m}$ and $\mathbf{u}$. In the 'test separator phase' the extended parameter-output vector for data assimilation is described by,

$$\mathbf{z} = [\mathbf{m}^T \ \mathbf{y}^T]^T = [\mathbf{m}^T \ p_{tf,choke} \ p_{tf,tubing} \ p_{wf}]^{\mathrm{T}}, \tag{2.40}$$

with predicted measurements $\mathbf{y}$ (Equation 2.34) and uncertain model parameters,

$$\mathbf{m} = \left[\mathbf{m}_{choke}^{\boldsymbol{T}} \ \mathbf{m}_{tubing}^{\boldsymbol{T}} \ \mathbf{m}_{res}^{\boldsymbol{T}}\right]^{\boldsymbol{T}}. \tag{2.41}$$

Samples of $\mathbf{z}$ are updated in some averaged sense by EnKF and PF by minimizing the mismatch between predict pressures $\mathbf{y}$ and actual measurements of the pressures,

$$\mathbf{d} = [d_1 \ d_1 \ d_2]^T = \left[p_{tf,obs} \ p_{tf,obs} \ p_{wf,obs}\right]^T, \tag{2.42}$$

with measured pressures at the tubing head $p_{tf,obs}$ and bottom hole $p_{wf,obs,}$ see Figure 2.7. The uncertain model parameters of Equation 2.40 are yet to be defined for the specific elements. Selection of the model parameters depends on the uncertainty of the parameters and the effect on the predicted output. The prior distribution and selection of uncertain model parameters in the choke model $\mathbf{m}_{choke}$ is described in Chapter 3. The selection of the uncertain model parameters in multi-phase flow model g_{tubing} is described in Chapter 4. The modelling of uncertain model parameter $\mathbf{m}_{res}$ for a single-phase oil reservoir is described in Chapter 5. Data assimilation is applied for the 'test separator phase' in Chapter 6, after definition of the uncertain model parameters and their initial distribution in the previous chapters.

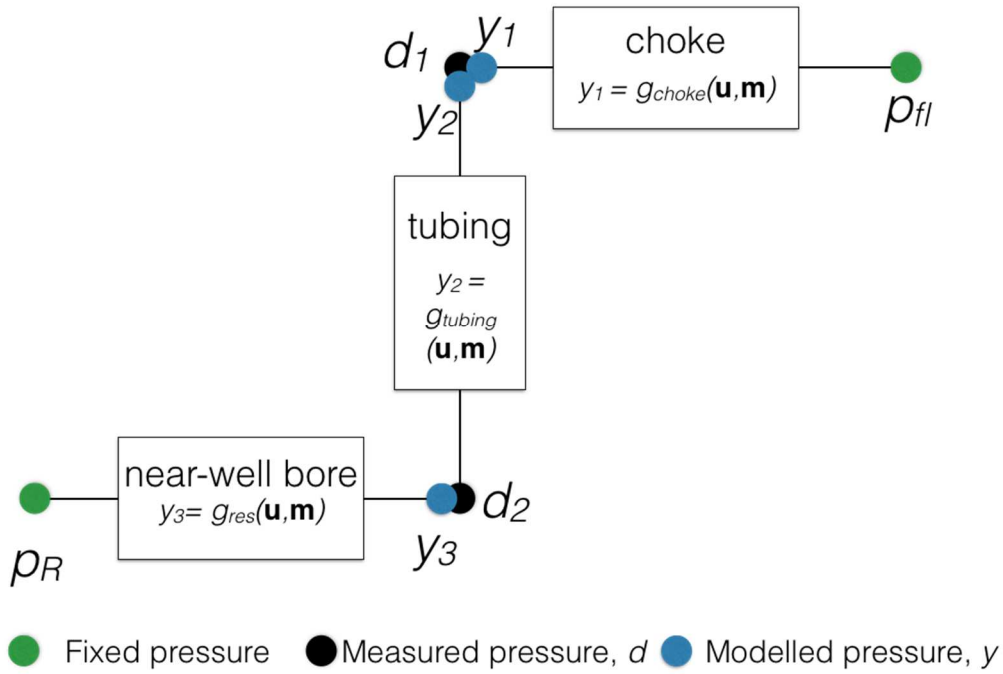


Figure 2.7: Production system of Figure 1 for data assimilation purposes for production phase 1 'test seperator phase'. The reservoir pressure p_R and the flow line pressure p_{fl} are fixed. The tubing head pressure p_{tf} and bottom hole pressure p_{wf} are measured, and are represented as measurments d_1 and d_2 respectively. Based on uncertain model parameters and input parameters different realizations of $\mathbf{y}$ are calculated based on the different elements, see Equations 2.37, 2.38 and 2.39.

2.3.2 Production phase 2 – flow rate estimation

During (semi-steady state) production, the flow rates of oil and gas, the reservoir pressure and the choke diameter are subject to change and could be considered as uncertain model parameters. A forward model is required to predict the model parameters in the next time step. The calibrated pdfs of the model parameters from the test separator phase are used as inputs in $\mathbf{u}$,

$$\mathbf{u} = \left[A\ B\ C\ c_{grav}\ c_{fric}\ c_J\ R_{sb}\ V_o\ \right]^T. \tag{2.43}$$

Possible uncertain model parameters which should be estimated are,

$$\mathbf{m} = \left[q_{o,sc}\ p_{R,av}\ d_{ch}\ \right]^T, \tag{2.44}$$

3 TUNING OF MODEL PARAMETERS OF THE CHOKE

The choke is a designed flow restriction, where a pressure drop occurs. In a petroleum system, a wellhead choke connects the flow line to the tubing head of a well. It has two main functions, (1) create a pressure drop by dissipating large amounts of energy over a small distance and (2) regulate the flow rate to produce in a controlled manner [1]. Controlling the fluid rate prevents excessive water and gas production, sand production and erosion due to high fluid velocities [24]. The pressure drop in the choke for multi-phase flow could be modelled in several ways, e.g. with empirical models and theoretical models. Empirical models are based on relationships of curve-fitted data with different parameters. Theoretical models are based on the underlying physics and thermodynamics [1]. In this chapter, the data assimilation methods ensemble Kalman filter (EnKF) and particle filter (PF) are used to update model parameters of the extended Gilbert model for multi-phase flow in chokes.

For multi-phase flow and single-phase gas flow the choke performance can be divided into two flow regimes, critical flow and sub-critical flow. During critical flow, a pressure disruption occurs between the upstream and downstream pressure, i.e. the downstream pressure does not influence the upstream pressure [25]. Tuning of the model parameters is only applied assuming critical flow, i.e. when the pressure ratio of the downstream and upstream pressure is below the critical ratio.

3.1 Choke models

There are different multi-phase models available for critical flow and sub-critical flow to model the pressure drop in the choke. The Gilbert correlation is an empirical model based on experimental data to describe the pressure drop for critical flow and has three input parameters: the choke diameter d_{ch}, the producing gas-liquid ratio R_{gl} and the liquid production rate $q_{l,sc}$ [11]. After the introduction of the Gilbert correlation, several attempts have been made to estimate coefficients A, B and C of the Gilbert correlation based on experimental data [26-32]. A general form of these correlations can be formulated as the Gilbert equation,

$$p_{tf} = -Aq_{l,sc}\frac{(ER_{gl})^B}{(Fd_{ch})^C} + D, \tag{3.1}$$

where,

p_{tf} is the tubing head pressure at the inlet of the choke [Pa],
d_{ch} is the choke diameter [m],
$q_{l,sc}$ is the liquid flow rate at standard conditions [m^3/s],
R_{gl} is the producing gas-liquid ratio [m^3/m^3],
A, B and C are experimentally-determined coefficients for the different empirical models,
D is the atmospheric pressure 1.01×10^5 [Pa],
E and F are constants.

In Equation 3.1, the tubing head pressure p_{tf} does not depend on the downstream pressure and p_{tf} has a linear relationship with the production rate $q_{l,sc}$. The pressure drop in the choke is in reality more difficult to predict with a model, because p_{tf} depends on more factors than specified in Equation 3.1. In some cases Equation 3.1 is modified to have a better fit with field data, with parameters to account for free water, sediment and emulsion (BS & W) [33, 34] and the temperature at the tubing head [29, 34, 35]. The Gilbert equation is only valid for the critical flow regime; for sub-critical flow the Gilbert equation can be extended heuristically with a third order polynomial function to estimate p_{tf} at sub-critical conditions [1]. There are also more theoretical models available describe the pressure drop for multi-phase flow in chokes. The theoretical model derived by [1] which is based on [27, 36-38], leads to a near-linear relationship of the tubing head pressure and the flow rate in the critical regime, see Figure 3.1.

Possible reasons of a mismatch between the extended Gilbert model and observed data are: (1) incorrect model parameters A, B and C, (2) larger choke diameter due to choke erosion, (3) incorrect critical pressure ratio used, (4) change in parameter which is not defined in the extended Gilbert model, such as the temperature, and (5) measurement errors. See Figure 3.1, for different choke performance curves based on different correlations but calculated for the same choke diameter d_{ch} and producing gas-liquid ratio R_{gl}.

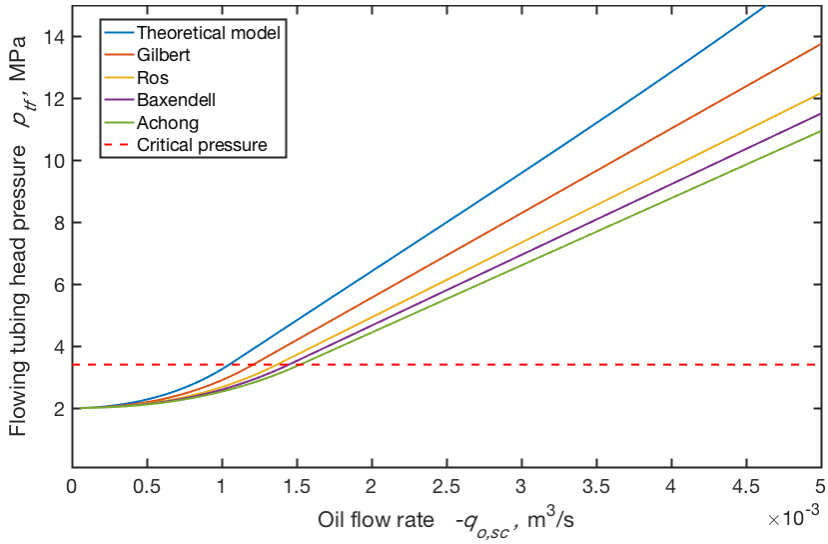


Figure 3.1: Choke performance curves based on extended Gilbert equation with different correlations, and based on the theoretical model described in [1].

3.2 Data assimilation for the choke model

Data assimilation requires three elements: measurements, a model and a data assimilation method. In production phase 1 – 'test separator phase', the distribution of uncertain model parameters A, B and C of the Gilbert model (Equation 1) are updated with the stochastic data assimilation methods ensemble Kalman filter (EnKF) and particle filter (PF). An initial prior distribution of the model parameters is required to initialize the data assimilation process.

3.2.1 Prior distribution of model parameters

In past researches the coefficients A, B and C have been estimated based on different experimental data [26-32], which illustrates that these model parameters are not

applicable to all chokes and fluid properties and thus can be considered uncertain. Based on the various available correlations a mean and variance of the model parameters can be determined, see Appendix A.1. Correlations used to obtain the prior distribution of the model parameters are from Gilbert [11], Ros [27], Baxendell [26] and Achong [28]. In Appendix A.2 a sensitivity study is conducted to determine the correlation between the model parameters and the tubing head pressure. The sensitivity analysis is based on a base case which uses the mean of the model parameters described in Appendix A.1.

3.2.2 Stochastic data assimilation application on the Gilbert choke model for the test separator phase

In the 'test separator phase' measurements are available of the gas-liquid ratio $R_{gl,obs}$, oil flow rate $q_{l,sc,obs}$ and the tubing head pressure $p_{tf,obs}$. The predicted measurement of the tubing head pressure $p_{tf,choke}$ is calculated with,

$$y = g_{choke}(\mathbf{u}_{choke}, \mathbf{m}_{choke}), \tag{3.2}$$

where y is $p_{tf,choke}$, g_{choke} is the Gilbert correlation (Equation 3.1), $\mathbf{u}$ are inputs,

$$\mathbf{u}_{choke} = [u_1\ u_2]^{\boldsymbol{T}} = \left[q_{l,sc}\ \ R_{gl}\right]^T, \tag{3.3}$$

and $\mathbf{m}$ are model parameters,

$$\mathbf{m}_{choke} = [m_1\ m_2\ m_3]^{\boldsymbol{T}} = [A\ \ B\ \ C]^T, \tag{3.4}$$

with A, B and C uncertain model parameters of the Gilbert correlation (Equation 3.1). The pdfs of A, B and C are specified in Appendix B.1. The extended parameter-output vector is defined as,

$$\mathbf{z} = [\mathbf{m}^T\ y^T]^{\boldsymbol{T}} = \left[A\ B\ C\ p_{tf,choke}\right]^T. \tag{3.5}$$

The goal of applying data assimilation in the test separator phase is to update the pdfs of the three model parameters of the choke model based on pressure and flow rate measurements. The updated posterior distributions should give better predictions, i.e. with lower variances, compared to the prior distributions of choke performance curves.

3.3 Twin experiment

A twin experiment is performed to test the performance of EnKF and PF during the 'test separator phase' to match the choke performance curve based on measurements. Synthetic data is created with a theoretical model described in [1]. The main assumptions of the twin experiments are: (1) two-phase flow in the choke with fixed gas-oil ratio and zero water cut, (2) the downstream flow line pressure is fixed and known, i.e. p_{fl} is not a stochastic variable, (3) the choke is operating under critical flow conditions, (4) the choke diameter is fixed and known, (5) measurements are available of oil rate $q_{o,sc,}$ gas rate $q_{g,sc}$ and tubing head pressure p_{tf} with a known error distribution.

The input parameters to create synthetic data are given in Appendix C.1. The theoretical based model is used to create synthetic data; this means there are no 'truth' values for the model parameters. Validation of the updated model parameters is done by calculating realizations of the output p_{tf} based on the posterior distribution of model parameters and compare the corresponding output with the 'true' output for the same flow rates and choke diameter.

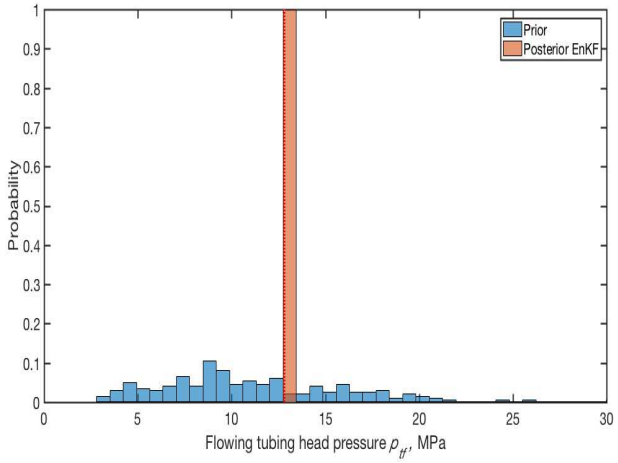
3.4 Results

Data assimilation is applied to update the model parameters defined in $\mathbf{z}$. Probability distribution functions are displayed as normalized histograms based on samples of $\mathbf{Z}_{prior}$ and $\mathbf{Z}_{posterior}$. In this part the prior and posterior pdfs of the output and model parameters

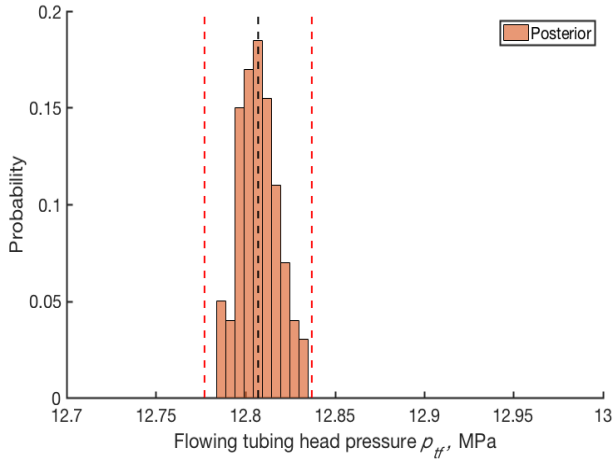
are compared. In the next section the posterior pdfs of model parameters will be validated in two ways: (1) with the prediction of p_{tf} for the same inputs and (2) by comparing choke performance curves based on *A, B* and C from the prior pdfs and posterior pdfs, and based on the 'truth model'.

3.4.1 Posterior pdfs of outputs

The posterior pdfs of the outputs based on the EnKF with 200 ensemble member show that the variance is within the measurement variance, see Figure 3.2. The posterior pdf of the output of the PF is obtained with residual resampling, see Figure 3.3. Filter degeneracy can be observed from Figure 3.3b, there are not enough particles with weights to represent the Gaussian distribution which was expected (see Equation 2.25). More information about resampling with the PF for this problem is discussed in Appendix A.5.2. Note that the measurement variance is increased compared to the EnKF, due the limited non-zero particles obtained otherwise. In Appendix A.5.1, the minimum number of particles is discussed by analysing the number of non-zero particles for different

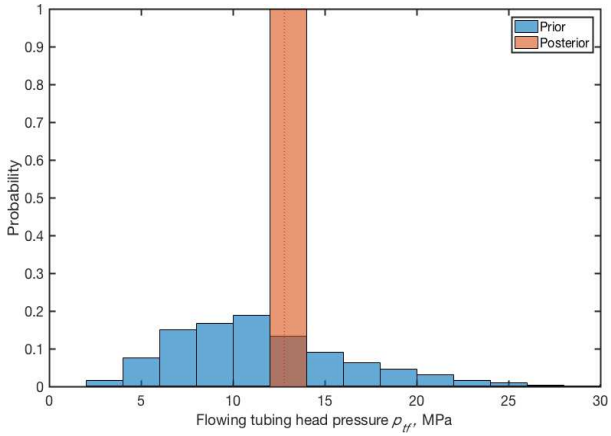


(a)Prior and posterior distributions of the output.

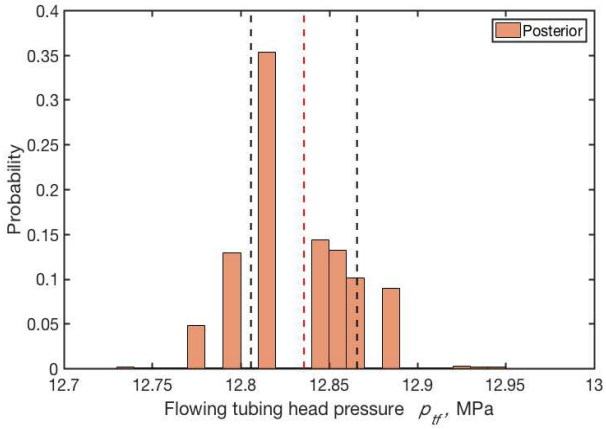


(b) Close-up of posterior distribution of the output, dotted line indicates measurement and variance of measurements.

Figure 3.2: Prior and posterior distribution of tubing head pressure, based on EnKF with 200 ensemble members.



(a) Prior and posterior distribution of the output.



(b) Close-up of posterior distribution of (a), dotted line indicates measurement and variance of measurements.

Figure 3.3: Prior and posterior distribution of tubing head pressure, based on PF with residual resampling with 1000 particles.

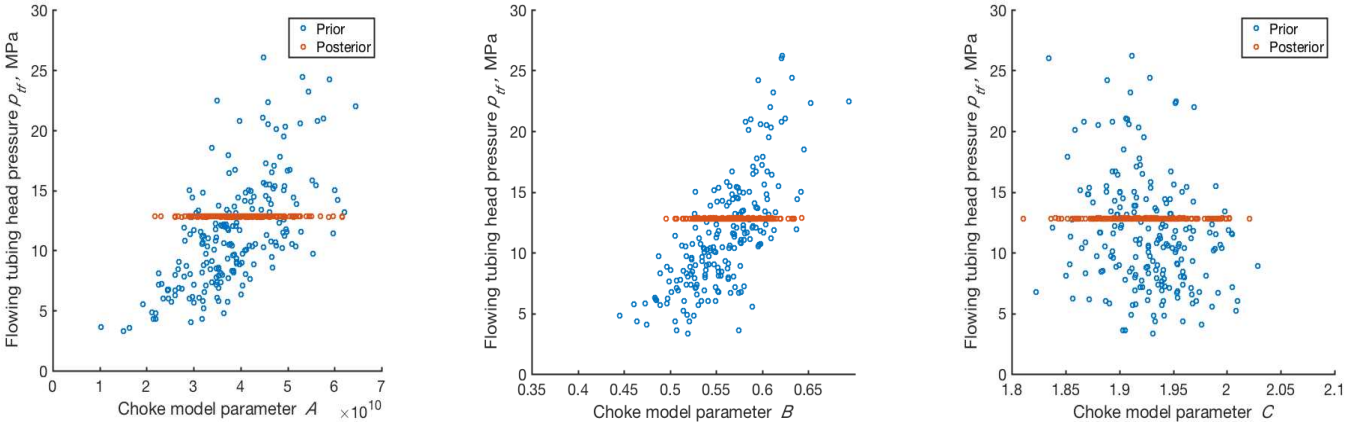
measurement variances and number of particles. Based on the results of Appendix A.5.1, a sample size of 1000 particles is determined.

3.4.2 Posterior pdfs of model parameters

The estimated correlation matrix based on the prior ensemble, see Equation 2.34, indicates the correlation between elements in parameter output **z,** i.e. correlation parameter-parameter and parameter-output. The correlation for 200 ensemble members is shown in Table 3.1. Because of independent sampling of model parameters, the correlation between model parameters should be approaching zero. The correlation matrix for different ensemble sizes appears in Appendix A.4. Table 3.1 indicates a positive correlation of A and B with y and shows a negative correlation of C and y.

Table 3.1: Correlation matrix based on N_p = 200.

	$m_1 = A$	$m_2 = B$	$m_3 = C$	$y = p_{tf,choke}$
$m_1 = A$	1.00	0.03	0.00	0.58
$m_2 = B$	0.03	1.00	0.03	0.73
$m_3 = C$	0.00	0.03	1.00	-0.25
$y = p_{tf,choke}$	0.58	0.73	-0.25	1.00



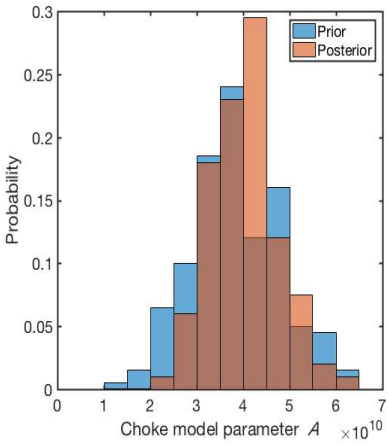
(a) Prior and posterior A vs p_{tf}. (b) Prior and posterior B vs p_{tf}. (c) Prior and posterior C vs p_{tf}.

Figure 3.4: Prior and posterior cross-plots of model parameters **m** and output p_{tf}, based on EnKF with 200 ensemble members

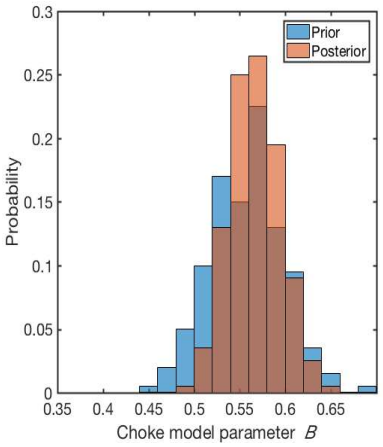
In Figure 3.4, the correlation between the uncertain model parameters and the predicted output is displayed. The model parameters and output are not strongly correlated, see Figure 3.4. Hence the variance of the model parameters is not reduced significantly, see Figure 3.5. However, this does not imply that the variance of the predicted output will also not be reduced significantly, because combinations of model parameters are updated which jointly may produce posterior simulated output close to the "truth". The posterior model parameters are validated in Section 3.5.

The posterior pdfs of the model parameters are also derived with three implementations of the particle filter, (1) SIR with residual resampling, (2) SIR with regularization and (3) auxiliary particle filter (ASIR) with regularization. The pdfs in Figure 3.6 are based on SIR with residual resampling implementation of the PF with 1000 particles. Multiple combinations of parameters have high probabilities based on the PF; the posterior

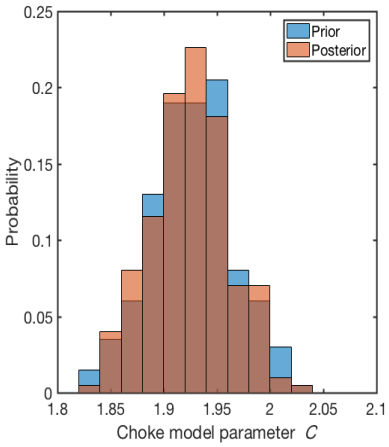
distribution does not indicate the prior Gaussian distribution which was used to make the initial samples of model parameters.



(a) Prior and posterior histogram *A.*

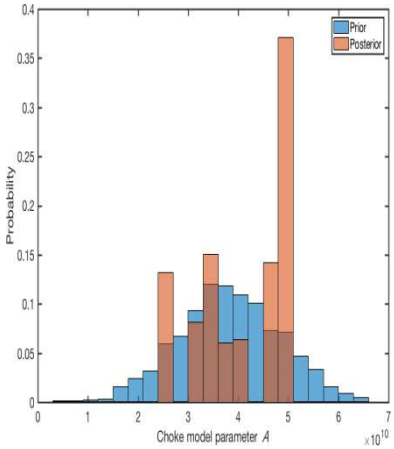


(b) Prior and posterior histogram *B.*

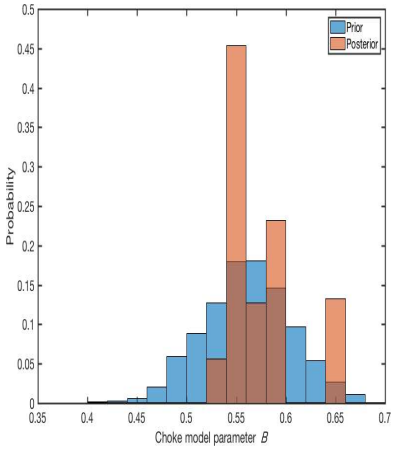


(c) Prior and posterior histogram *C.*

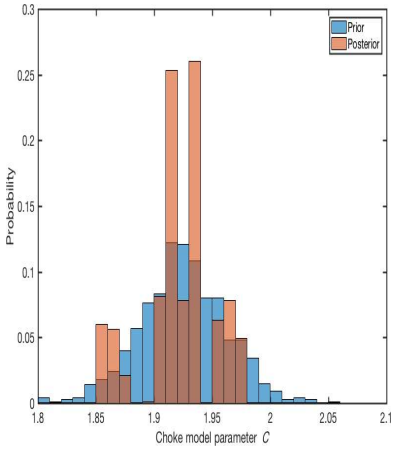
Figure 3.5: Prior and posterior distribution with EnKF and 200 ensemble members



(a) Prior and posterior histogram *A.*

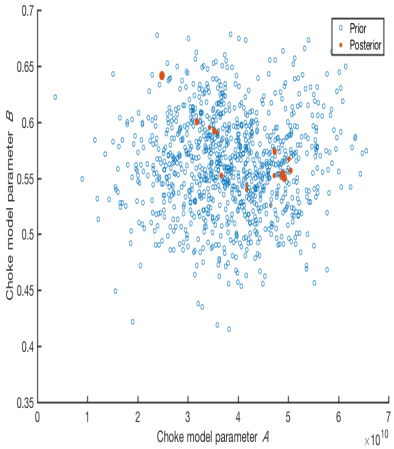


(b) Prior and posterior histogram *B.*

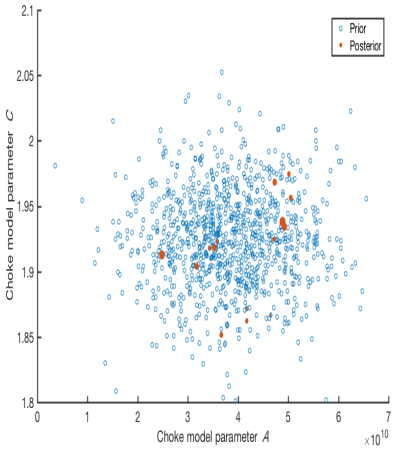


(c) Prior and posterior histogram *C.*

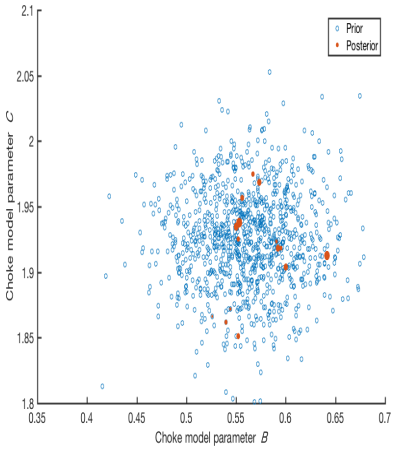
Figure 3.6: Prior and posterior pdfs of model parameters based on PF with 1000 particles with residual resampling.



(a) Prior and posterior correlation between *A* and *B.*



(b) Prior and posterior correlation between *A* and *C.*



(c) Prior and posterior correlation between *B* and *C.*

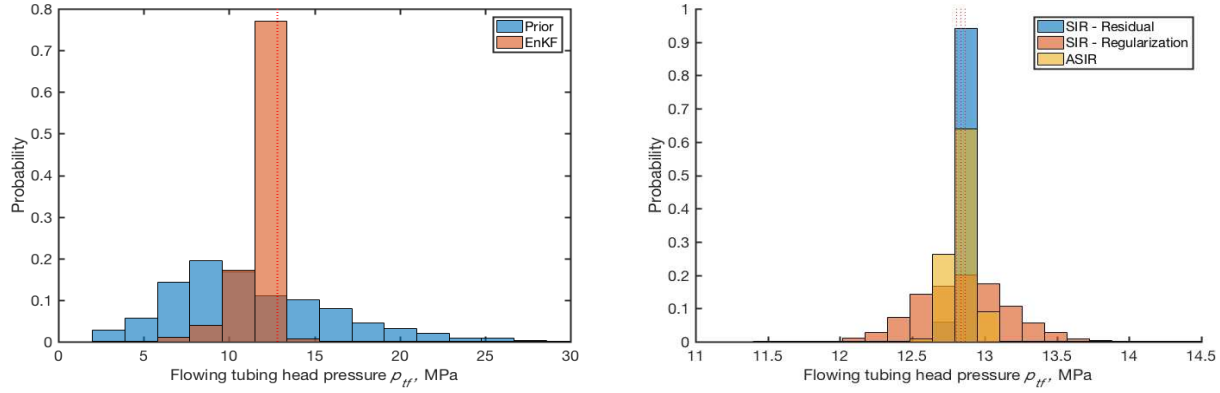
Figure 3.7: Prior and posterior cross plots of model parameters based on PF with 1000 particles (SIS).

Cross plots of the model parameters show that the particles with high weights are observed all over in the model parameters space, Figure 3.7. The results of the different implementation of the particle filter are discussed in Appendix A.5.2.

3.5 Validation

The updated distributions of model parameters are validated in two ways. Predicting pressure measurements and validate the obtained distribution of the output by observing the variance of the predictions and the mismatch from the 'true' tubing head pressure. The second method of validating the posterior pdfs of model parameters is to calculate realizations of choke performance curves and compare these with the choke performance curve based on the 'true' model.

In Figure 3.8, the tubing head pressure p_{tf} is calculated based on the model parameters of the prior pdfs and posterior pdfs based on EnKF (Figure 3.8a) and three implementations of the PF (Figure 3.8b); note that they are not on the same scale. The pdfs of model parameters updated with the particle filter give better predictions of the tubing head pressure. The variance of the predicted output of the particle filter depends on the resampling method. The largest variance reduction is observed with residual resampling because no sampling noise is added. With regularization as resampling technique, the variance in the output is still reduced significantly compared to the predicted measurement based on prior model parameters and EnKF posterior model parameters. The ASIR gives a better approximation compared to SIR regularization at the price of having to run the samples twice in the model, see Appendix A.5.2 for detailed explanation.



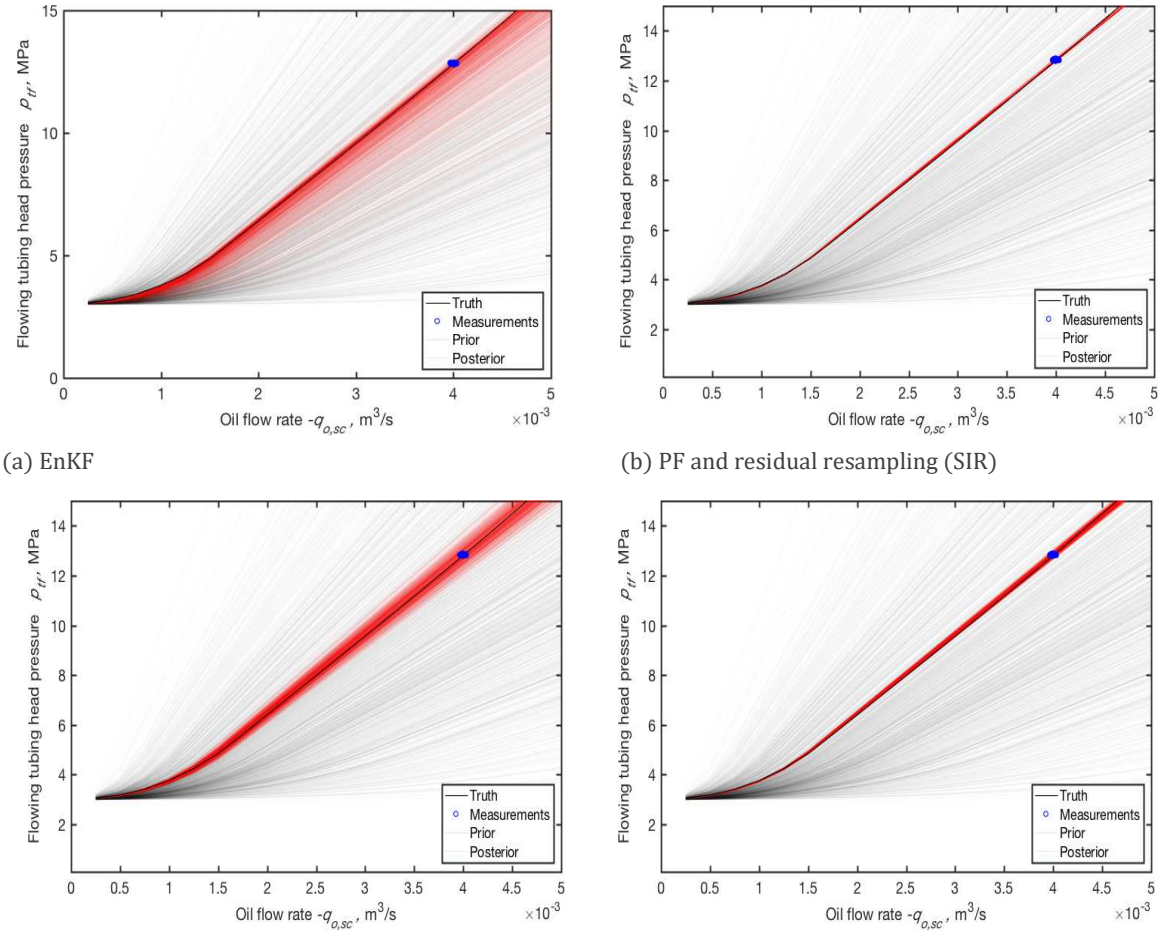
(a) EnKF (b) PF

Figure 3.8: Predictions based on model parameters pdfs of (a) prior and EnKF and (b) PF with SIR and ASIR.

Choke performance curves based on the prior and posterior model parameters are plotted in Figure 3.9. As observed in Figure 3.8, the EnKF has the highest variance in the resulting choke performance curve. The choke performance curves based on the posterior model parameters obtained with the particle filter are displaying lower variance compared to EnKF. The effect of the resampling method in the particle filter is like Figure 3.8 also observable in Figure 3.9.

The results show that the model parameters can be updated to give better predictions of the tubing head pressure, however, tuning three model parameters based on one output proved to be difficult due the large sample space of model parameters which resulted in an ill-posed problem. EnKF was not able to update the model parameter in a sufficient way (Figure 3.8a), the prior pdfs is based on independent samples from a

Gaussian distribution, such that the correlation between the model parameters is not considered in the sampling process. The PF required more samples to tune the model parameters, which is due the variance in the predictions based on the prior model parameters and the comparably low measurement variance leading to filter degeneracy. This is partially solved with implementation the auxiliary particle filter, see Appendix A.5.2.



(a) EnKF

(b) PF and residual resampling (SIR)

(c) PF with regularisation (SIR)

(d) Auxiliary PF with regularisation (ASIR)

Figure 3.9: Prior and posterior pdfs of model parameters with EnKF and PF, based on 1000 realizations.

4 TUNING OF MULTI-PHASE FLOW MODEL PARAMETERS

Multiple phases, such as oil, gas, water and/or solids are produced in oil wells [39]. Modelling of the fluid flow in pipes in production systems involves conservation of mass, momentum and energy [24]. There are many multi-phase correlations available for flow in pipes, which can be divided into two categories: (1) homogenous models and (2) mechanistic models [40]. In mechanistic models the multiple phases are treated as separate fluids [14]; they are also referred as semi-empirical models [1]. Homogenous models treat the multiple phases as a mixture with certain properties, such as hold-up, mixture velocity and represent a form of empirical models.

The multiple phases present in wells make modelling of the pressure drop complicated [24]. One of the complicating factors is the different flow regimes or flow patterns which are a typical feature in multi-phase flow in wells. Another complicating factor is the description of the difference in velocity between the phases, also known as slip, which causes liquid hold-up [1]. Most multi-phase models make use of empirical expressions based on experiments to describe the liquid hold-up based on fluid properties, flow rates, pipe diameter and inclination [1]. The available models lead to different predictions of pressure and flow rates. With data assimilation a better match between model (prediction) and reality (measurement) can be obtained.

Various applications of data assimilation methods in multi-phase well flow model are already available. The ensemble Kalman filter (EnKF) and extended Kalman filter (EKF) have been used for several applications in multi-phase flow modelling, e.g. to update model parameters of the drift-flux model [5, 6] and for flow rate estimation [8]. Other applications involving drift-flux models and pressure and/or temperature measurements are found for under-balanced drilling [41, 42], estimation of zonal production rates [8, 9, 42], and multi-phase flow estimation in horizontal wells [14, 43]. The most recent developments are for transient well flow models which are able to track sudden changes in flow rates [4, 7, 44, 45].

The focus in this chapter to tune two model parameters of a homogenous steady-state multi-phase well flow model with the data assimilation methods EnKF and PF. The model used to describe the pressure drop in the well is the Mukherjee and Brill correlation [12] extended with a gravity correction factor c_{grav} and a friction correction factor c_{fric}. The correction factors are modelled as uncertain model parameters. Twin experiments are performed with the synthetic data based on the drift-flux model [1, 46].

4.1 Multi-phase flow models in wells

The element equation for multi-phase well flow can be written as,

$$p_{out} = g_{tubing}(p_{in}, q_{g,sc}, q_{o,sc}, q_{w,sc}), \tag{4.1}$$

where,

g_{tubing} is the multi-phase flow model in wells, p_{in} is the pressure at the inlet which in this case is the tubing head pressure (FTHP) or the bottom hole pressure (FBHP), p_{output} is the pressure at the outlet, either FBHP or FTHP, and $q_{g,sc}$, $q_{o,sc}$ and $q_{w,sc}$ are the flow rates of oil, gas and water at standard conditions. There are different multi-phase flow models

available varying in complexity; for a detailed overview of available multi-phase well flow models, see [24], and for a basic explanation of the concepts, see [1]. For steady-state flow the description of the total pressure gradient in the well,

$$\left(\frac{dp}{ds}\right)_{tot} = \left(\frac{dp}{ds}\right)_{grav} + \left(\frac{dp}{ds}\right)_{fric} + \left(\frac{dp}{ds}\right)_{acc}, \tag{4.2}$$

consist of three components: head, friction and acceleration losses [1, 24]. The predominant factor for pressure loss in a production well is the elevation/head loss component followed by the friction component with usually a negligible acceleration term. For empirical/homogeneous models the pressure drop is described in terms of mixture density and velocity,

$$\left(\frac{dp}{ds}\right)_{tot} = -\rho_m g \sin\theta - \frac{\rho_m}{2d} f\, v_m\, |v_m| - \rho_m v_m \frac{dv_m}{ds}, \tag{4.3}$$

where,
ρ_m is the mixture density [kg m^{-3}], which depends on empirical hold-up correlations,
g is the gravitational constant 9.81 [m s^{-2}],
θ is the pipe inclination, where for vertical wells $\sin\theta = 1$,
d is the diameter of the well [m],
f is the friction factor,
v_m is the mixture velocity, which depends on slip and hold-up correlations. The pressure loss due the elevation change depends on the mixture density. The pressure loss due to friction depends on a two-phase friction factor, pipe diameter, mixture velocity and mixture density. More complicated (semi-empirical) models are available. However, the description of these model is outside the scope of this research.

The pressure drop measured in the field and the modelled pressure drop based on a multi-phase model do not necessarily match. Calculation of the pressure drop with empirical multi-phase models requires several assumption for the friction factor, mixture density and mixture velocity. This empirical nature of the equations leads to different intake curves with the same inputs for well and fluid properties. A typical feature of multi-phase flows in wells is that different flow patterns occur which influence the pressure drop. In this research, the pressure drop is calculated with the Mukherjee and Brill correlation, which takes different flow patterns and slip into consideration. For a detailed description of the correlation see [12].

4.2 Data assimilation for multi-phase well flow models

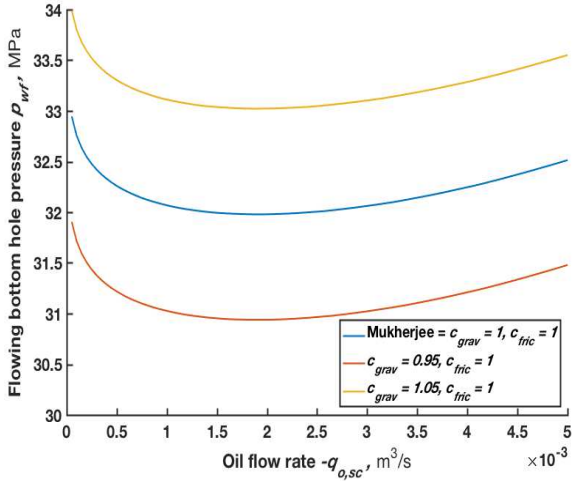
As mentioned above, there are many ways to compute the gravity and friction component, which explains the mismatch between observations and predictions. The mismatch between predicted output and measured output depends on the model error and errors in input parameters. The predicted output based on the Mukherjee & Brill correlation is corrected by applying correction factors on indiviual components of Equation 4.2,

$$\left(\frac{dp}{ds}\right)_{tot} = c_{grav}\left(\frac{dp}{ds}\right)_{grav} + c_{fric}\left(\frac{dp}{ds}\right)_{fric} + \left(\frac{dp}{ds}\right)_{acc}, \tag{4.4}$$

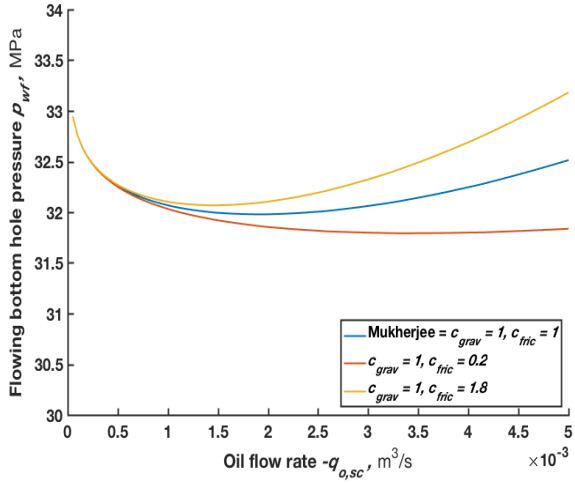
with, c_{grav} the gravity correction factor and c_{fric} the friction correction factor. Note that multiple combinations of the correction factors can lead to the same output.

4.2.1 Defining prior distributions of gravity and friction correction factors

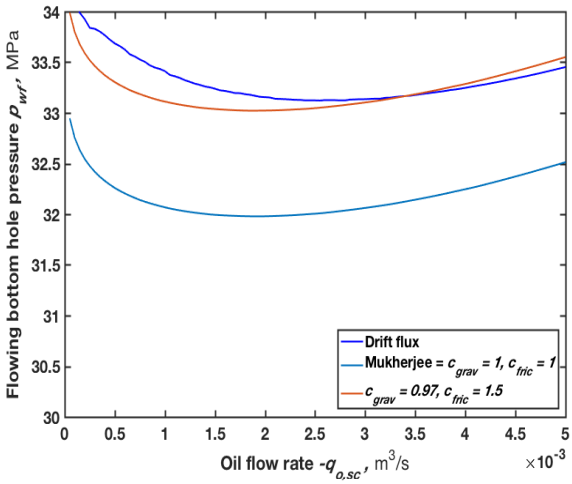
The initial distributions of the gravity correction factor c_{grav} and the friction correction factor c_{fric} need to be defined. The correction factors in this report are applied to the Mukherjee and Brill correlation [12]. However, the approach of applying the correction factors may also be used with other homogenous multi-phase pipe flow models. The mean of c_{grav} and c_{fric} is set at 1, which corresponds to the Mukherjee and Brill correlation. The correction factor of the gravity component causes a translation of the tubing intake curve, i.e. the tubing intake curve moves up or down based on the value of c_{grav}, see Figure 4.1a. The friction component has more effect on the part where friction forces dominate, i.e. at higher flow rates, see Figure 4.1b. With the friction correction factor and gravity correction factor, predicted tubing intake curves and the 'truth' intake curve can be matched, see Figure 4.1c, where the tubing intake curve based on the drift flux model is approximated with the tubing intake curve calculated with Equation 4.4 and Mukherjee and Brill correlation.



(a) Effect of the gravity correction factor on the tubing intake curve



(b) Effect of friction correction factor on the tubing intake curve



(c) Manual matching of Mukherjee and Brill correlation with Drift flux model with correction factors

Figure 4.1: Effect of the gravity and friction correction factors on tubing intake curve

4.2.2 Stochastic data assimilation application applied to multi-phase well models for test separator phase

In the 'test separator phase' measurements are available of the gas-liquid ratio $R_{\text{gl,obs}}$, oil flow rate $q_{\text{l,sc,obs}}$, the tubing head pressure $p_{tf,obs}$ and the bottom hole pressure $p_{\text{wf,obs}}$. The predicted measurement of the tubing head pressure $p_{tf,tubing}$ is calculated with,

$$y = g_{tubing}\left(\mathbf{u}_{tubing}, \mathbf{m}_{tubing}\right), \tag{4.5}$$

where y is $p_{tf,tubing}$, g_{tubing} is the Mukherjee and Brill correlation extended with correction factors (Equation 4.4), $\mathbf{u}$ are inputs,

$$\mathbf{u} = [u_1\ u_2\ u_3]^{\boldsymbol{T}} = \left[q_{o,sc}\ q_{g,sc}\ p_{wf,obs}\right]^T, \tag{4.6}$$

and $\mathbf{m}$ are model parameters,

$$\mathbf{m} = [m_1\ m_2]^{\boldsymbol{T}} = \left[c_{grav}\ c_{fric}\right]^T. \tag{4.7}$$

The extended parameter-output vector is defined as,

$$\mathbf{z} = [\mathbf{m}^T\ y^T]^{\boldsymbol{T}} = \left[c_{grav}\ c_{fric}\ \ p_{tf,tubing}\right]^T. \tag{4.8}$$

The pressure drop in this element can be calculated in two ways: with the bottom hole pressure as input or with the tubing head pressure as input, see Equation 4.1.

4.3 Twin experiment

Synthetic data is created based on the drift flux model, see [46]. The Mukherjee & Brill correlation gives a different tubing intake curve given the same input values of fluid properties, well properties and flow rates, see Figure 4.2c.

The twin experiments are conducted with tubing head pressure p_{tf} fixed, and predict the well bore pressure p_{wf}, i.e. tubing intake curves are plotted.

'True' data is created based on the black oil Standing correlations [47] and drift flux model [46]. With the Glaso black oil correlations [48] as input and the Mukherjee and Brill correlation modified with correction factors, the bottom hole pressure is predicted. Appendix B.3 gives a detailed overview of the inputs of the models. See Figure 4.3 for the difference between the 'true' tubing intake curve and the 'base' intake curve which is based on the mean of the prior distributions, i.e. Mukherjee and Brill correlation.

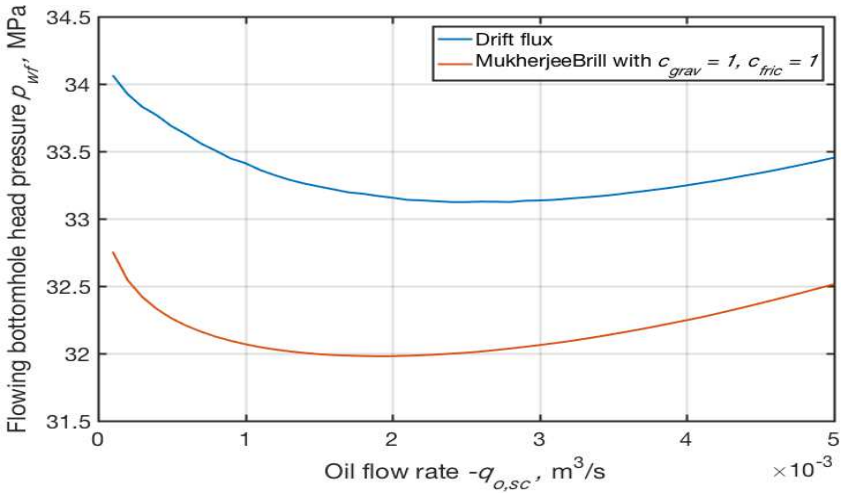


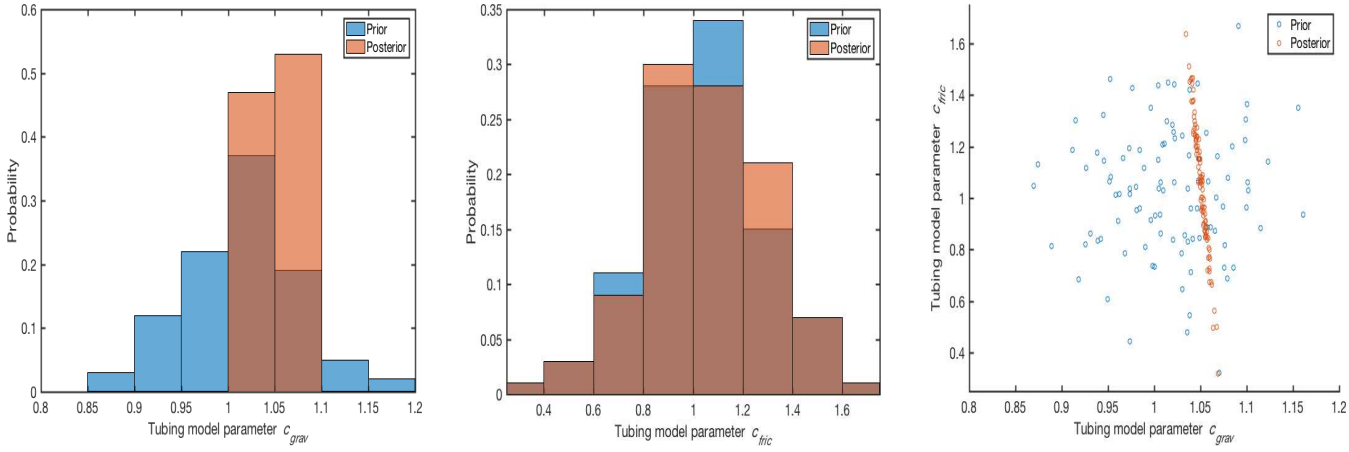
Figure 4.2: Tubing intake curve calculated based on 'truth model' with drift flux model and intake curve based on Mukherjee and Brill correlation.

4.4 Results

Data assimilation is applied to update the gravity correction factor and the friction factor. The posterior pdfs of the model parameters in this section are based on EnKF and the auxiliary particle filter (ASIR). Results are obtained with 100 and 500 realizations. In the next section the posterior pdfs of model parameters will be validated in two ways: (1) with the prediction of p_{wf} for the same inputs and (2) comparing tubing intake curves based on model parameters from the prior pdfs and posterior pdfs, and based on the 'truth model'.

4.4.1 Results - 100 realizations

The posterior distribution of the gravity correction factor shows a reduction in variance, see Figure 4.3a. The variance of the friction correction factor is not reduced, see Figure 4.3b. However, Figure 4.3c, shows as a clear correlation between c_{grav} and c_{fric}.



(a) Prior and posterior histogram c_{grav}.

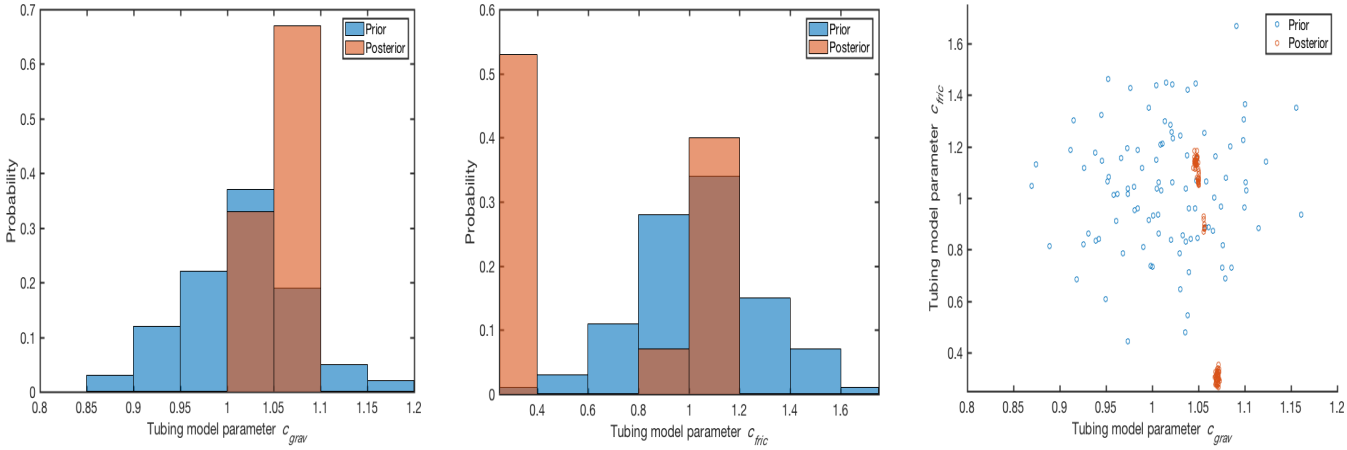
(b) Prior and posterior histogram c_{fric}.

(c) Cross-plot between c_{grav} and c_{fric}.

Figure 4.3: Prior and posterior pdfs of model parameters based on EnKF with 100 ensemble members.

The posterior distribution obtained with ASIR and 100 particles displays a different result compared to the EnKF, see Figure 4.3 and 4.4. However, Figure 4.4c shows that the cross

plot of the posterior model parameters (Figure 4.4c) is similar to Figure 4.3c, only with less model parameter space covered.



(a) Prior and posterior histogram c_{grav}.

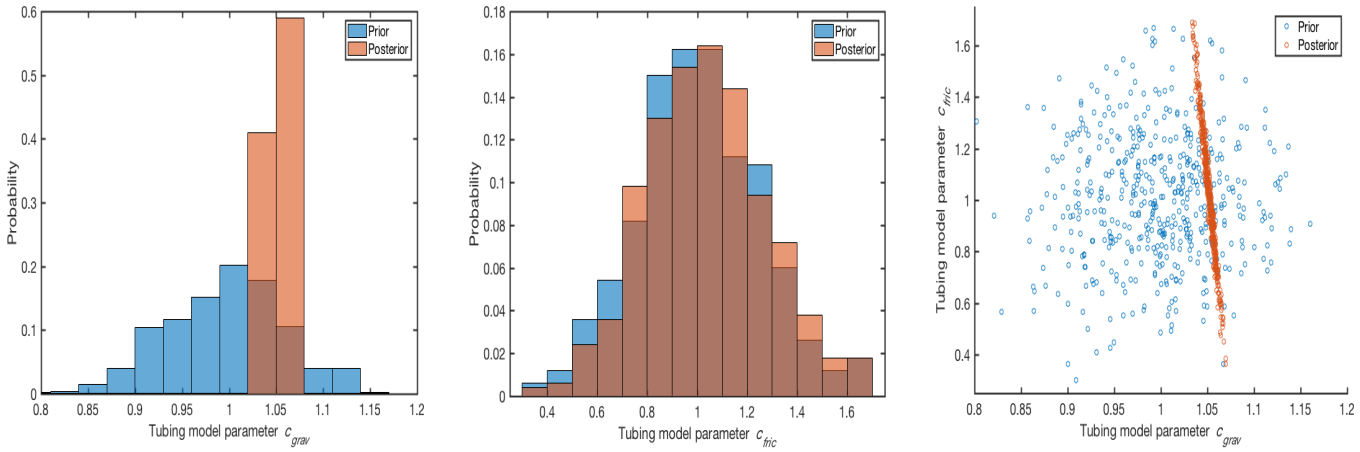
(b) Prior and posterior histogram c_{fric}.

(c) Cross-plot between c_{grav} and c_{fric}.

Figure 4.4: Prior and posterior pdfs of model parameters based on the auxiliary PF with 100 particles.

4.4.2 Results – 500 realizations

The resulting cross plot of model parameters, in Figure 4.5c, show similar trends compared to 100 realizations with EnKF, see Figure 4.3c. The posterior pdfs obtained with ASIR show a distributions, resembling the correlation between model parameters much better compared to the 100 realizations case, see Figure 4.6c.

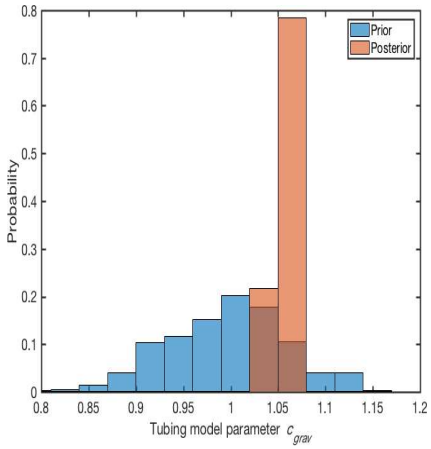


(a) Prior and posterior histogram c_{grav}.

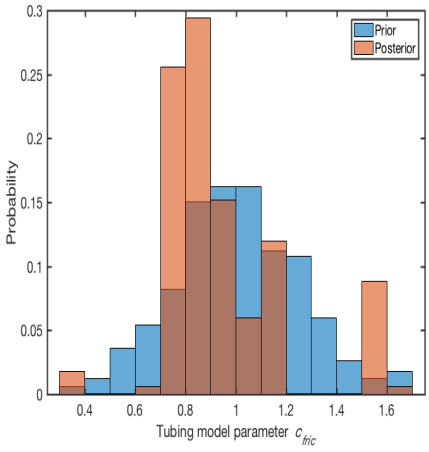
(b) Prior and posterior histogram c_{fric}.

(c) Cross-plot between c_{grav} and c_{fric}.

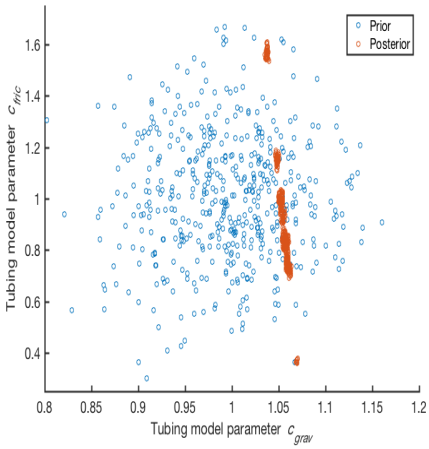
Figure 4.5: Prior and posterior pdfs of model parameters based on EnKF with 500 ensemble members.



(a) Prior and posterior histogram c_{grav}.



(b) Prior and posterior histogram c_{fric}.



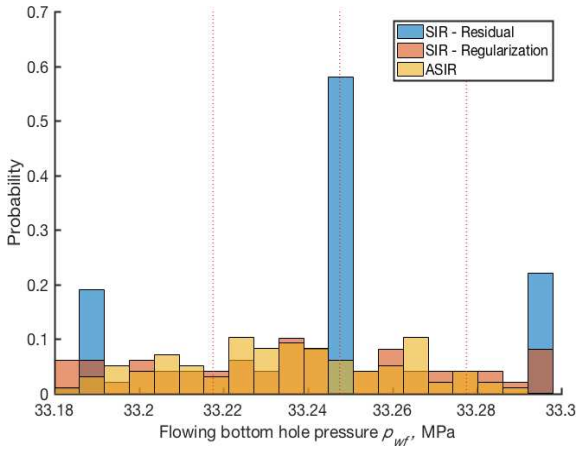
(c) Cross-plot between c_{grav} and c_{fric}.

Figure 4.6: Prior and posterior pdfs of model parameters based on the auxiliary PF with 500 particles.

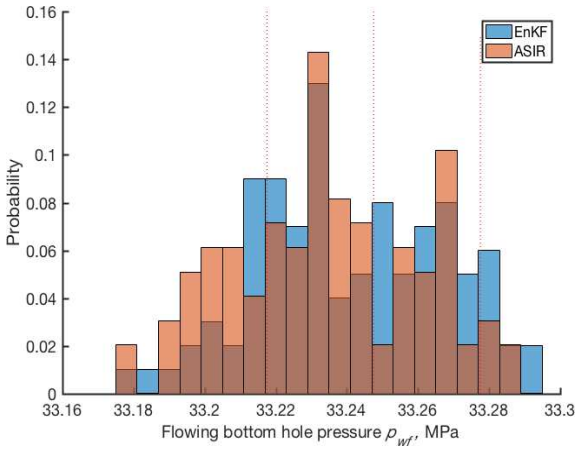
4.5 Validation

The posterior pdfs of model parameters are validated in the two ways described in Section 3.5. The predicted bottom hole pressure is calculated with the posterior model parameters. The pdfs in Figure 4.7 are obtained with 100 samples, and show three peaks. These peaks indicate filter degeneracy, and because residual resampling will generate a distribution close the SIS distribution, these peaks are picked up with residual resampling. In this example, the use of regularization as resampling technique becomes clear, by adding noise to the resampled particles a more accurate distribution is obtained, see Figure 4.7a, the extreme probabilities are not visible anymore. The implemented ASIR produces similar results EnKF, however the cross-plots in Figure 4.6c show that gaps are present in the model parameter space.

Tubing intake curves based on prior and posterior model parameters in Figure 4.9 show that the ensemble variance near the measurement decreases. An important note is that ensemble collapse can happen with the current prior distribution, see Appendix B.2.

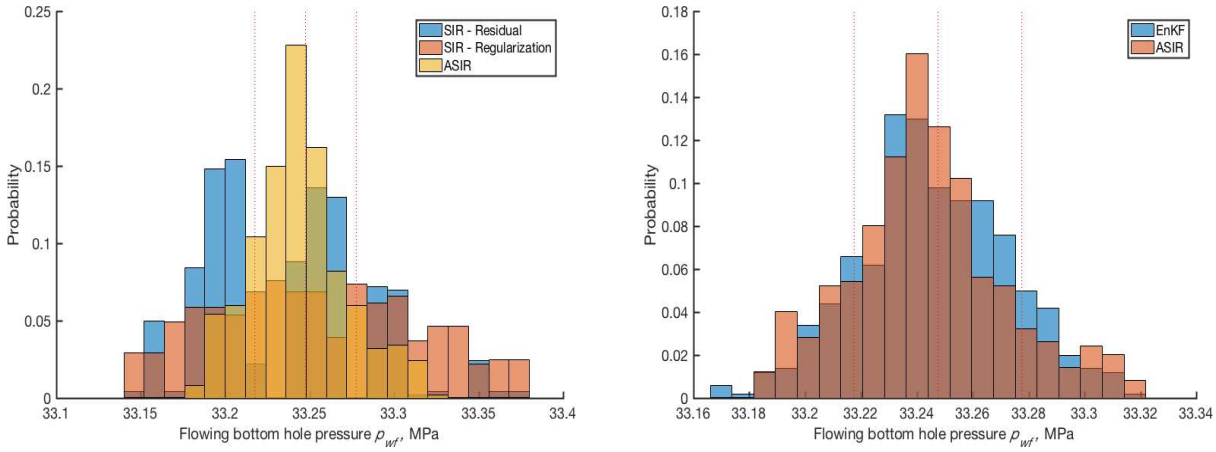


(a) PF



(b) EnKF

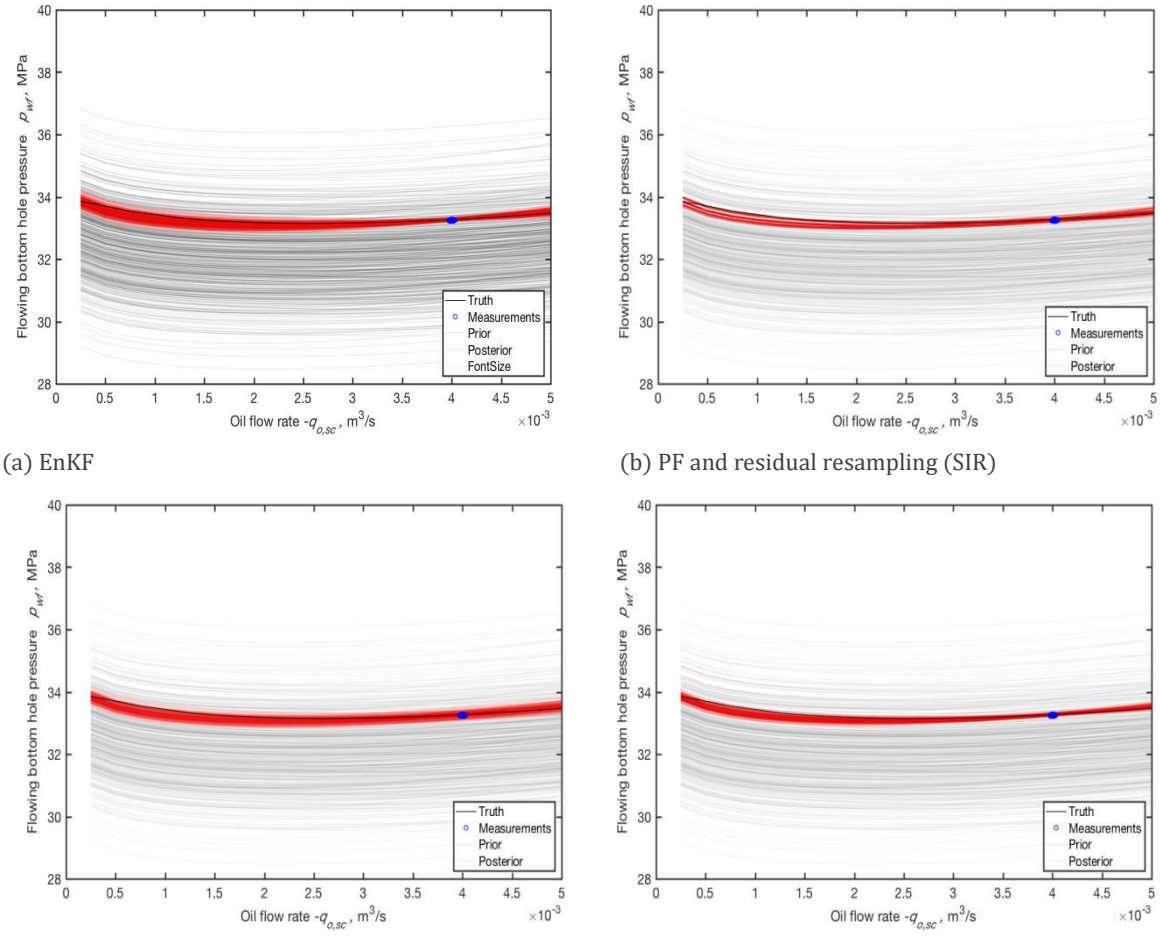
Figure 4.7: Predictions based on 100 samples of model parameters based on (a) PF with SIR with residual, regularization and ASIR (b) EnKF and ASIR.



(a) PF (b) EnKF

Figure 4.8: Predictions based on 500 samples of model parameters based on (a) PF with SIR with residual, regularization and ASIR (b) EnKF and ASIR.

With multiple measurements, at different operating conditions, better estimation of the combination of the correction factors can probably be estimated.



(a) EnKF (b) PF and residual resampling (SIR)

(c) PF with regularisation (SIR) (d) Auxiliary PF with regularisation (ASIR)

Figure 4.9: Tubing intake curves based on samples from prior and posterior based on 500 realizations compared to the 'truth' tubing intake curve.

5 TUNING OF NEAR-WELL BORE AREA MODEL PARAMETERS

The driving force for the inflow in the well bore is the difference between the reservoir pressure and the bottom hole pressure. The inflow in the well could be restricted by several factors, like reservoir properties, fluid properties and completion of the well [1]. All these factors determine the inflow performance relationship (IPR), which can be described as the relationship between the flow rate and the pressure in the well bore and reservoir.

For single phase oil reservoirs, the IPR is linear. The slope of the IPR depends on reservoir properties such as permeability, skin and the reservoir radius and fluid properties such as density and viscosity. The IPR for an oil reservoir above bubble point is also referred as the productivity index (PI). In this chapter, the predicted PI is corrected with a PI correction factor c_J.

5.1 The inflow performance relationship

The IPR for single phase oil reservoirs with the reservoir pressure above the bubble point pressure can be expressed as the productivity index (PI),

$$J = \frac{-q_{o,sc}}{p_{R,ref} - p_{wf}}, \tag{5.1}$$

where,

$p_{R,ref}$ is the reference reservoir pressure [Pa], which refers to the average reservoir pressure $p_{R,av}$ or the pressure at the drainage boundary of the well p_R,

p_{wf} is the bottom hole pressure[Pa],

$q_{o,sc}$ is the oil flow rate at standard conditions [$m^3\ s^{-1}$],

J is the productivity index [$m^3\ s^{-1}\ Pa^{-1}$].

For single phase radial steady-state oil flow near a well, the bottom hole pressure can be approximated by an analytical expression [49],

$$p_{wf} = p_{R,ref} + \frac{\mu_o B_o q_{o,sc}}{2\pi k h}\left[\ln\left(\frac{r_e}{r_w}\right) - f_R + S\right], \tag{5.2}$$

where,

μ_o is the viscosity of the oil [Pa s],

B_o is the formation volume factor of the oil [$m^3\ m^{-3}$],

k is the permeability of the reservoir rock [m^2],

h is the reservoir height [m],

r_e is the drainage radius of the reservoir [m],

r_w is well radius [m],

f_R is a correction factor which depends on the reference reservoir pressure $p_{R,ref}$ and the flow regime [-],

S is skin factor [-].

The factor f_R depends on the flow regime, i.e. steady state flow or semi-steady state flow and the reference reservoir pressure, i.e. $p_{R,ref} = p_{R,av}$ or $p_{R,ref} = p_R$, resulting in a constant value for f_R,

$$f_R = \begin{cases} 0 \text{ for steady-state flow with } p_{R,ref} = p_R, \\ \frac{1}{2} \text{ for steady-state flow with } p_{R,ref} = p_{R,av}, \\ \frac{1}{2} \text{ for semi-steady-state flow with } p_{R,ref} = p_R. \\ \frac{3}{4} \text{ for semi-steady-state flow with } p_{R,ref} = p_{R,av}. \end{cases} \quad (5.3)$$

Based on Equation 5.1 and 5.2 an analytical "inflow performance relationship" can be derived as,

$$J = \frac{-q_{o,sc}}{p_{R,av} - p_{wf}} = \frac{2\pi kh}{\mu B_o \left[\ln\left(\frac{r_e}{r_w}\right) - f_r + S\right]}, \quad (5.4)$$

for single phase oil flow. In addition to the analytical model, a numerical model is also available for a homogeneous oil reservoir, which takes the pressure dependency of the viscosity in account [1]. In Figure 5.1, an IPR is plotted with a linear relationship of oil rate and pressure above the bubble point pressure p_b, and a non-linear one below p_b.

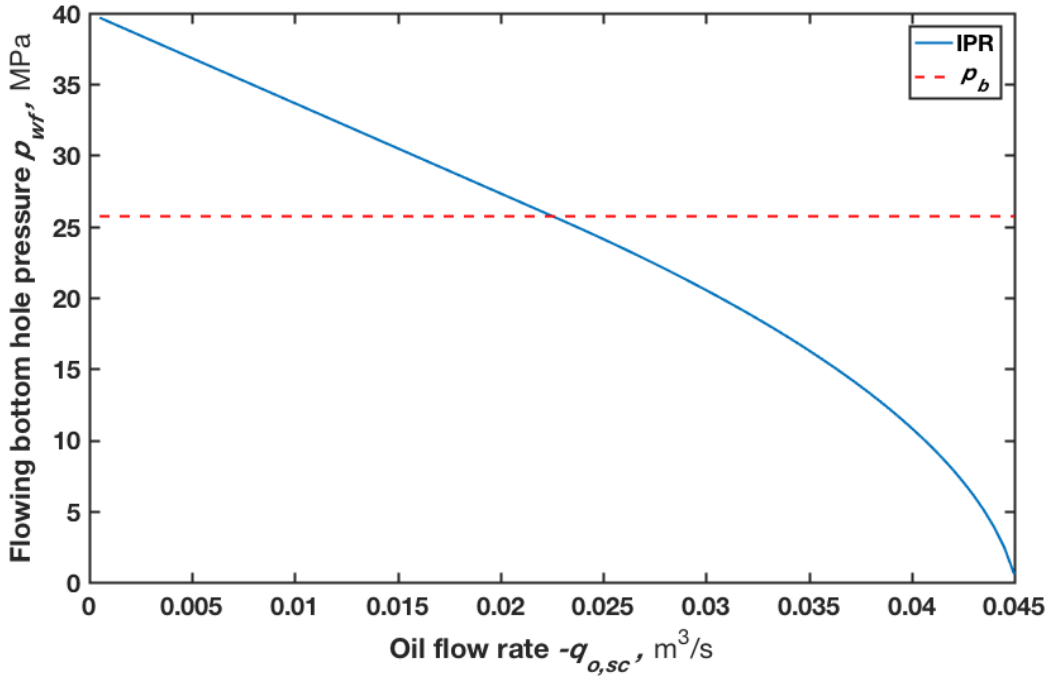


Figure 5.1 Inflow performance relationship (IPR) with bubble point pressure p_b, illustrating single phase flow when bottom hole pressure above p_{wf} and multi-phase below p_b.

The bubble point pressure p_b, depends on the type of oil, and the temperature, and can be calculated with correlations which are valid for certain ranges. Thus p_b could be considered as an uncertain model parameter. Equation 5.2 is not applicable for multi-phase flow, i.e. when the pressure in the reservoir is lower than the bubble point pressure p_b, in which case the Vogel equation can be used [50].

5.2 Data assimilation

There are many uncertain parameters during production such as permeability, skin and drainage area. However, the effect of all these parameters is to change the PI, and hence only a single PI correction factor is considered as a stochastic parameter. Equation 5.4 is extended with,

$$c_J\, J = \frac{-q_{o,sc}}{p_{R,av}-p_{wf}} = c_J \left(\frac{2\pi kh}{\mu B_o \left[\ln\left(\frac{r_e}{r_w}\right) - \frac{3}{4} + S \right]} \right), \tag{5.5}$$

where c_J is the PI correction factor. Note that semi-steady-state conditions and an average reservoir pressure are considered in Equation 5.5. In the 'test separator phase' measurements are available of the gas-liquid ratio $R_{gl,obs}$, oil flow rate $q_{o,sc,obs}$ and the bottom hole pressure $p_{wf,obs}$. The reservoir pressure is assumed to be a known parameter in the test separator phase. The predicted measurement of the bottom hole pressure p_{wf} is calculated with,

$$y = g_{res}(\mathbf{u}_{res}, \mathbf{m}_{res}), \tag{5.6}$$

where y is the bottom hole pressure p_{wf}, g_{res} is the analytical model describing the pressure in the near-well bore area (Equation 5.1), and $\mathbf{u}$ are inputs,

$$\mathbf{u} = [u_1\ u_2]^{\boldsymbol{T}} = \left[q_{\mathrm{o,sc}}\ \ q_{g,sc}\right]^T, \tag{5.7}$$

while in this case m_{res} only consist of one parameter

$$m_{res} = c_J. \tag{5.8}$$

The extended parameter-output vector is defined as,

$$\mathbf{z} = [m_{res}\ y]^{\boldsymbol{T}} = \left[c_J\ \ p_{wf}\right]^T. \tag{5.9}$$

5.3 Twin experiment

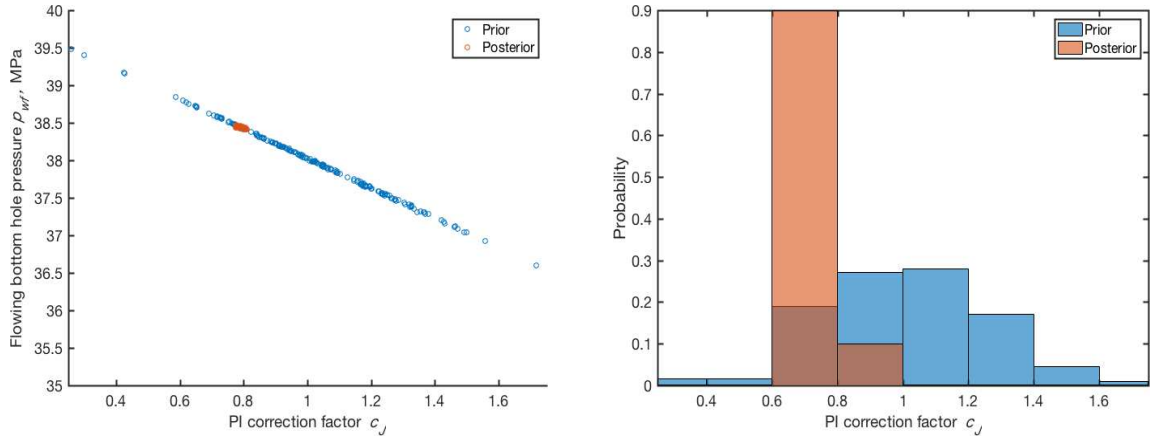
This twin experiment is conducted with different values for r_e, k and S in the 'true' model and the model used for data assimilation, see Table 5.1. The PI correction factor is estimated with EnKF and PF.

Table 5.1: Parameters for twin experiment in near-well bore element

Variable	*Symbol*	*Unit*	*'Truth' value*	*'Model' value*
Reservoir data:				
Reservoir height	h_R	m	30	25
Permeability	k	m^2	0.9×10^{-13}	1.1×10^{-13}
Skin	S	–	0	3
Reservoir radius	r_e	m	400	800
Fluid properties:				
Black oil correlation			Standing[47]	Glaso[48]
Oil density	$\rho_{o,sc}$	kg m^{-3}	890	880

5.4 Results

All methods give similar results when the number of realizations is sufficient, which is much lower compared to the realizations used in Chapter 4 and 5. In Figure 5.2 the update of the PI correction factor is displayed.



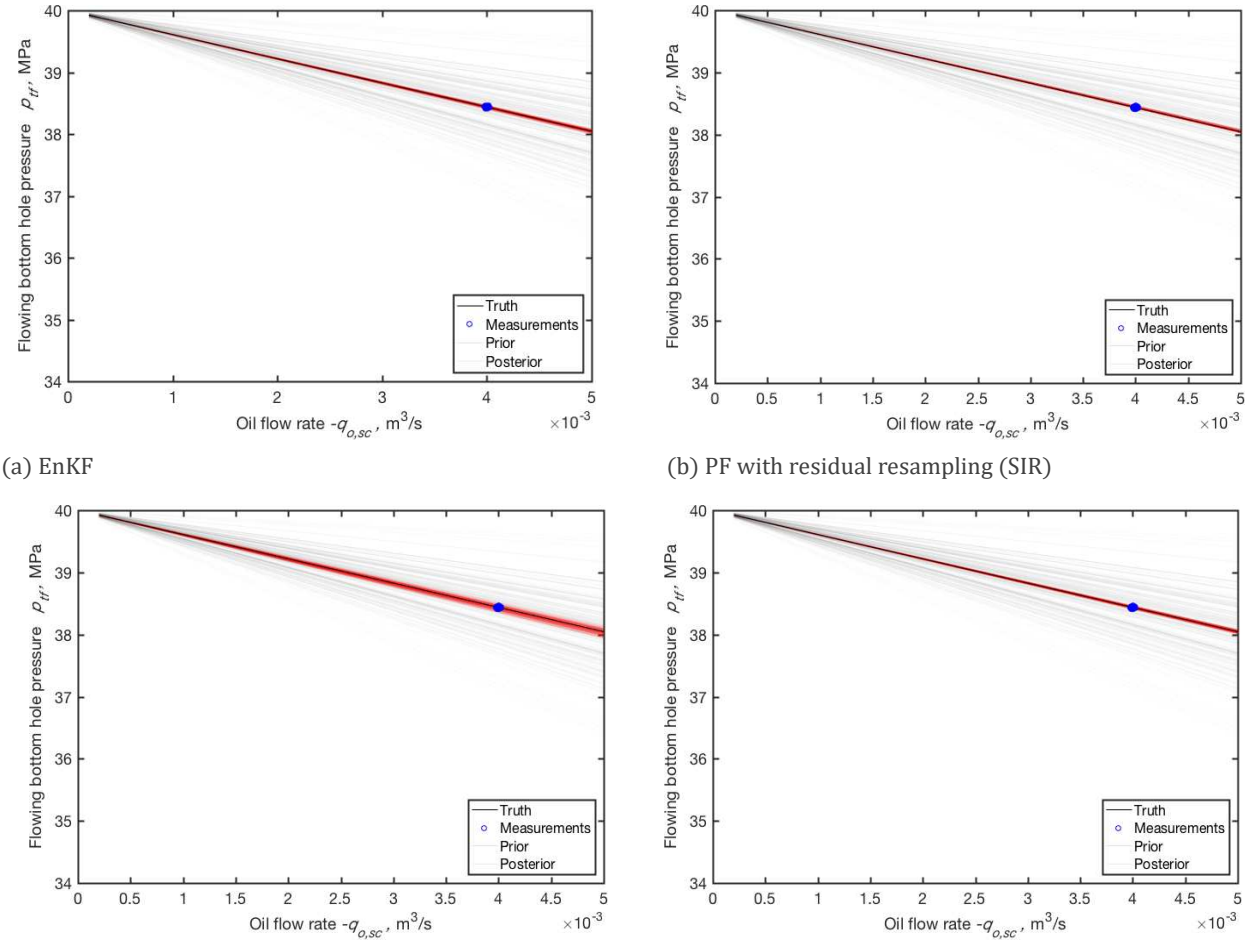
(a) Cross plots of model parameter and output

(b) Probability distribution of PI correction factor.

Figure 5.2: PI correction update with EnKF and 100 realizations.

5.5 Validation

The IPR based on the posterior pdfs shows that the PI can be accurately estimated; see Figure 5.3. With the available measurements of flow rate, well bore pressure and assumed fixed reservoir pressure the PI can be calculated with Equation 5.1. The PI correction factor gives an idea of the mismatch between the analytical expression and the measured PI.



(a) EnKF

(b) PF with residual resampling (SIR)

(c) PF with regularisation (SIR)

(d) ASIR

Figure 5.3: Realizations of IPR based on different data assimilation methods.

6 MODEL CALIBRATION DURING TEST SEPARATOR PHASE

Stochastic model parameters of the production system are updated during the test separator phase, see Section 2.3.1. The initial distributions and the choice of the uncertain model parameters are discussed in the Chapter 3, 4 and 5 for the choke, tubing and near-well bore element respectively.

6.1 Data assimilation on a production system

The stochastic data assimilation methods ensemble Kalman filter (EnKF) and particle filter (PF) are used to update probability density functions (pdfs) of model parameters in the test separator phase. The uncertain model parameters in the parameter-output vector

$$\mathbf{z} = [\mathbf{m}^{\mathrm{T}}\ \mathbf{y}^{\mathrm{T}}]^{\mathrm{T}}, \tag{6.1}$$

can now be specified with the chosen parameters described in Chapters 3, 4 and 5 as,

$$\mathbf{m} = \left[\mathbf{m}_{choke}^{\mathrm{T}}\ \mathbf{m}_{tubing}^{\mathrm{T}}\ \mathbf{m}_{res}^{\mathrm{T}}\right]^{\mathrm{T}} = \left[A\ B\ C\ c_{grav}\ c_{fric}\ c_J\right]^{\mathrm{T}}, \tag{6.2}$$

resulting in an extended parameter-output vector,

$$\mathbf{z} = \left[A\ B\ C\ c_{grav}\ c_{fric}\ c_J\ p_{tf,choke}\ p_{tf,tubing}\ p_{wf}\right]^{\mathrm{T}}, \tag{6.3}$$

which contains model parameters and predicted outputs for all the elements in one extended parameter-output vector.

In the previous chapters, only one predicted output was used in the data assimilation procedure. In the application to a production system three predicted measurements are used, i.e. N_y is equal to three. Data assimilation with the ensemble Kalman filter is applied with 200 realizations, motivated by computation time, and results from previous chapter.

Assigning weights to a particle $\mathbf{z}$ proved to be difficult with the particle filter. As illustration, with measurements errors increased 50 times and 15000 realizations, only one particle was assigned a weight, i.e. severe filter degeneracy occurred. The current production system and uncertain model parameters permit a different approach of solving this data assimilation problem with the particle filter (PF) in the test separator phase. Instead of applying data assimilation simultaneously to all elements of a production system,

$$\mathbf{z} = [\mathbf{m}^{\mathbf{T}}\ \mathbf{y}^{\mathbf{T}}] = \left[\mathbf{m}_{choke}\ \mathbf{m}_{tubing}\ \mathbf{m}_{res}\ y_1\ y_2\ y_3\right]^{\mathrm{T}}, \tag{6.4}$$

data assimilation is applied to discrete elements of the production system:
choke,

$$\mathbf{z}_{choke} = \left[\mathbf{m}_{choke}^{\mathbf{T}}\ y_1\right]^{\mathbf{T}} = \left[A\ B\ C\ p_{tf,choke}\right]^{\mathbf{T}}, \tag{6.5}$$

tubing,

$$\mathbf{z}_{tubing} = \left[\mathbf{m}_{tubing}^{\mathbf{T}}\ \ y_2\right]^{\mathrm{T}} = \left[c_{grav}\ c_{fric}\ p_{tf,tubing}\right]^{\mathrm{T}}, \tag{6.6}$$

and reservoir,

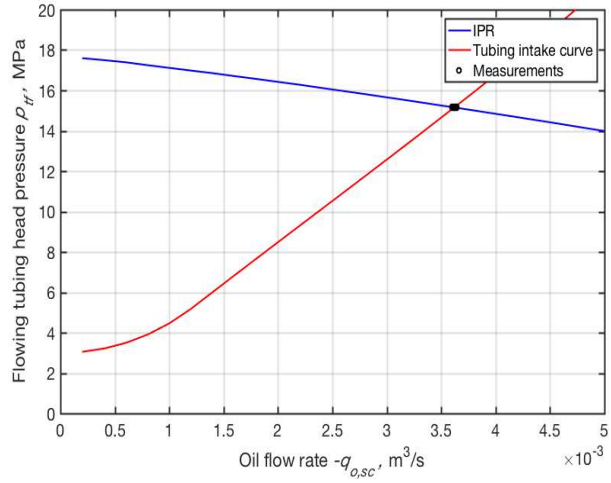
$$\mathbf{z}_{res} = [\mathbf{m}_{res}^{\mathbf{T}}\ y_3] = [c_J\ p_{wf}]^{\mathrm{T}}. \tag{6.7}$$

This approach could be justified, because the uncertain model parameters of each element are independent, i.e. not having any correlation with each other. However, in this approach weights are assigned to each element separately, i.e. applying Equation 2.24 three times (note, the number of realizations is not affected in this approach and the number of calculations of the modelled outputs remains the same).

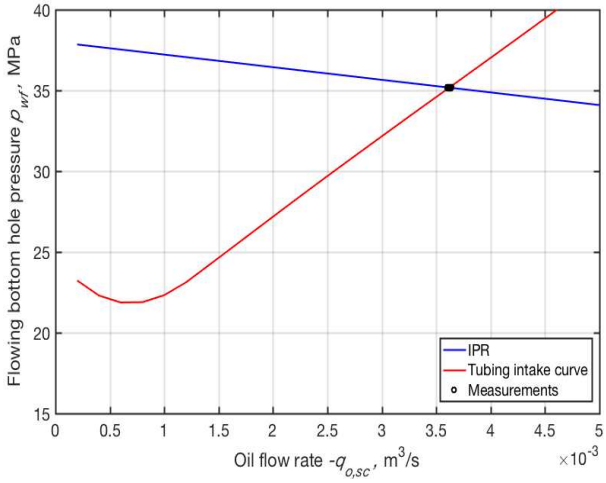
6.2 Twin experiment

In a twin experiment, synthetic data is created based on a truth model. With the pdfs of model parameters and chosen models, prior realizations of performance curves are made.

6.2.1 Truth model



(a) Combined plot of 'true' choke performance and tubing performance curves and synthetic measurements of p_{tf}



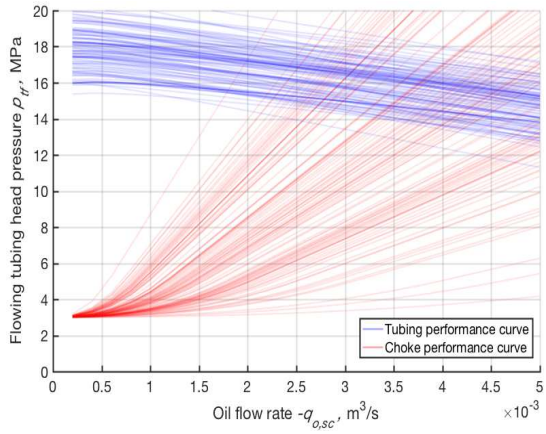
(b) Combined plot of 'true' inflow performance and tubing intake curves and synthetic measurements of p_{wf}

Figure 6.1: 'True' performance curves at tubing head and bottom hole node.

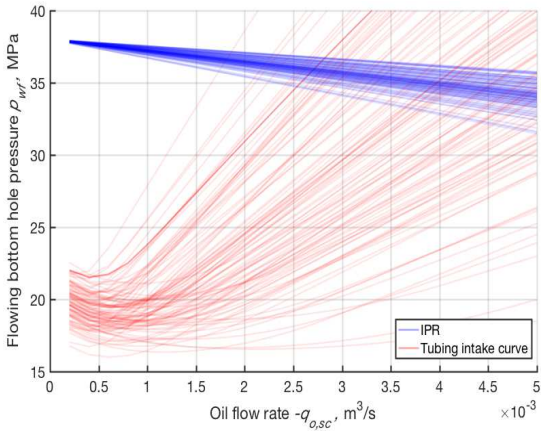
A 'true' model is used to create synthetic data. Synthetic data for the 'test separator phase' is created based on deterministic inputs and models for the choke, tubing and near-well bore element to obtain the performance curves. With the nodal analysis procedure described in Chapter 2, the operating flow rate and pressures are obtained which will be used as synthetic measurements, see Figure 6.1. See Appendix D for inputs of the model.

6.2.2 Prior ensemble

Realizations of 'performance curves' can be plotted at the tubing head and bottom hole nodes of the production system, see Figure 6.2. The prior distributions of the model parameters are given in Appendix D.



(a) Combined plot of realizations of choke performance and tubing performance curves at tubing head

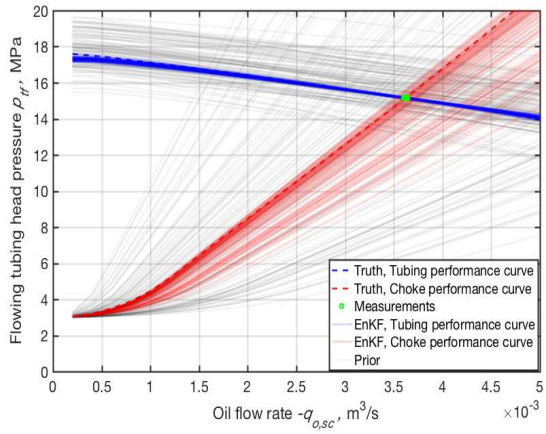


(b) Combined plot of realizations of inflow performance relationship and tubing intake curves at bottom hole node

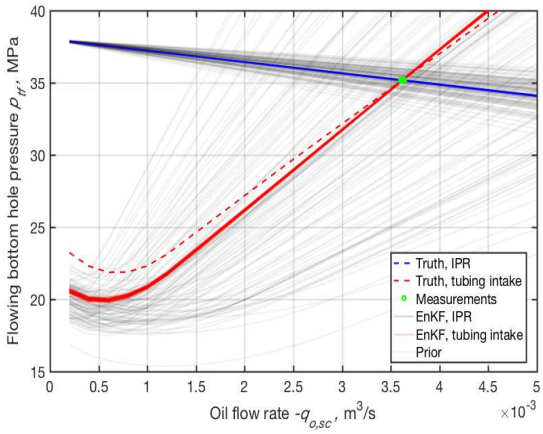
Figure 6.2: Ensemble of performance curves based on distributions of model parameters of choke, tubing and near-well bore element.

6.3 Results

Realizations of the performance curves can be calculated based on the posterior model parameters. The posterior pdf of model parameters is obtained with ASIR and EnKF. EnKF results are obtained with 200 realizations, the ASIR results are obtained with 1000 realizations of model parameters. More realizations in PF were needed to avoid filter degeneracy. Realizations of the performance curves calculated with the EnKF based calibrated model parameters are displayed in Figure 6.3. The choke performance curves show a large variance compared to the measurement variance, like the results observed in Chapter 3. Realizations of the tubing performance curves based on EnKF indicate that the applied method is successful in minimizing the variance in the predicted output by tuning model parameters. However, the tubing intake curves in Figure 6.3b show that there is a mismatch between predicted and the 'true' tubing intake curve. Not containing the 'true' solution in the posterior ensemble, could be the result of a wrong 'initial' prior distribution or too much reduction of the variance in uncertain model parameters. Both reasons seem to apply in this case: as observed in Figure 6.3b, the prior distribution of the tubing intake curve does not contain the 'true' intake curve in the ensemble at low flow rates, i.e. there is not enough variance in the gravity correction factor. Furthermore, for high flow rates, the ensemble of tubing intake curves show that the 'true' intake curve is not in the ensemble. The estimated PI correction factor gives a good approximation of the PI, see Figure 6.3b.



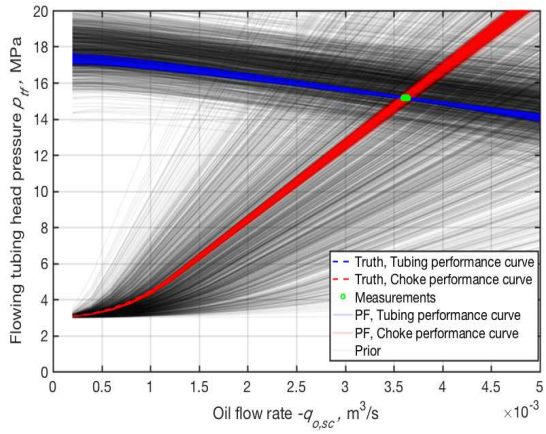
(a) Combined plot of realizations of choke performance and tubing performance curves at tubing head



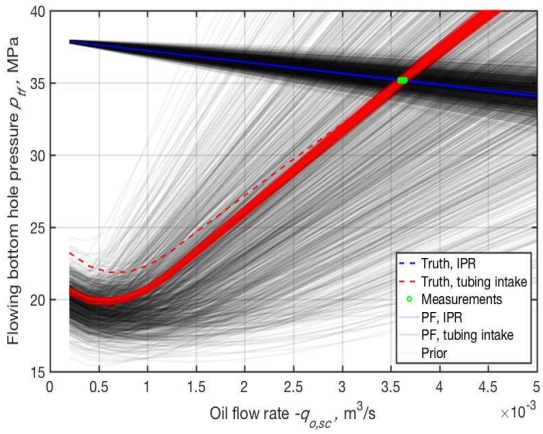
(b) Combined plot of realizations of inflow performance relationship and tubing intake curves at bottom hole node

Figure 6.3: Ensemble of performance curves based on prior and updated distributions with EnKF of model parameters of choke, tubing and near-well bore element.

The auxiliary particle filter (ASIR) is able to update the model parameters to obtain an ensemble of the choke performance curve with lower variance compared to the EnKF, see Figure 6.4a. The tubing intake curves display better results, however at lower flow rates the estimated tubing intake curves deviated from the 'truth'. The ASIR implementation in the current approach needs 20 times more model calculations compared to the EnKF.



(a) Combined plot of realizations of choke performance and tubing performance curves at tubing head



(b) Combined plot of realizations of inflow performance relationship and tubing intake curves at bottom hole node

Figure 6.4: Ensemble of performance curves based on prior and updated distributions with PF of model parameters of choke, tubing and near-well bore element.

The goal of the 'test separator phase' is to tune the uncertain model parameters. The uncertain model parameters can subsequently be used as stochastic input in the 'flow rate estimation phase'. However, we do not address this subsequent step in this thesis.

7 Conclusion

The focus of this thesis is to investigate the application of the data assimilation methods ensemble Kalman filter (EnKF) and particle filter (PF) to a production system with twin experiments. The goal for applying data assimilation for the steady-state models of the production system in this report is tuning of uncertain model parameters during the test separator phase. The results show that the posterior distributions of the model parameters lead to ensembles of performance curves with lower variances compared to the performance curves based on the prior distribution. EnKF and PF can be applied to estimate uncertain model parameters in a petroleum production system, however further research is required to reach the stage where the calibrated model parameters will be used as stochastic inputs in the flow rate estimation phase.

7.1 Discussion

The biggest challenge is selection of model parameters and the prior distribution based on these model parameters. This discussion is split into two main categories: (1) individual elements and (2) general remarks.

In each element the number of stochastic model parameters is different, three in the choke, two in the tubing and one in the near-well bore. A higher number of uncertain parameters leads to an increase of the parameter-output dimension, leading to increasing complexity of estimation of distributions of model parameters for higher order problems.

Parameters of the original proposed Gilbert correlations are adjusted by many researches, see Chapter 3, which motivated the choice of the parameters A, B and C as uncertain parameters. The current approach is to take independent samples of prior distribution of the model parameters, not taking into account the relationship of the parameters. The results of the results with particle show that combinations of the model parameters will have the same output, i.e. this is an ill-posed problem. In hindsight, this is not surprising because only a single steady-state measurement is used to estimate three parameters. Better results can therefore be expected by either reducing the number of parameters to one, or increasing the number of measurements (separator tests at different flow rates) to at least three.

The posterior distribution of the two correction factors in the well flow model is also difficult to estimate, i.e., this problem is also an ill-posed problem. The case study in Appendix B.2 show that the prior distribution is important to avoid ensemble collapse, i.e. an ensemble of tubing intake curves without the truth curve. Also here, better results can be expected by either reducing the number of parameters to one, or increasing the number of separator tests at different flow rates to at least two.

The inflow performance relationship for single-phase oil flow is described with a linear relationship between FBHP and oil rate. With the proposed correction factor, an initial estimate of the PI is obtained, with the reservoir pressure considered as a deterministic parameter in the test separator phase. Unfortunately, individual production parameters such as the skin factor, permeability and drainage radius cannot be corrected/estimated. Doing so would require a pressure transient test.

The EnKF needs a lower number of realizations compared to the particle filter. Applying data assimilation to the parameter-output vector for a production system consisting of all six model parameters and three outputs was computationally not feasible with the

particle filter, leading to splitting of the parameter-output vector for the different elements. The EnKF was able to update the parameter-output with 100 realizations of model parameters, whereas the PF required 1000 realizations (with ASIR twice as much model calculations, i.e. 20 times more compared to the EnKF).

7.2 Recommendations

- Use multiple flow rates and pressure measurements in the 'test separator phase, or a smaller number of parameters, to obtain better estimates of the parameters in the choke and multi-phase well flow models.
- Make a predictive model for pressure decline and prediction of flow rates. With EnKF and PF, the calibrated model parameters as inputs and the predicted output of the pressure at the tubing head and bottom hole the flow rate and average reservoir pressure can be estimated. Eventually parameters for near well-bore damage and choke erosion could be estimated.
- Perform twin experiments with different measurement variances and available data. What will the effect be on the distributions of model parameters if the only measurement is the tubing head pressure?
- Make use of the temperature dependency for choke flow, see [34], and for multi-phase flow in pipes, for flow rate estimation; see e.g. [8].
- Use mechanistic multi-phase well flow models as 'true' models in twin experiments. Also use the output of commercial software such as PIPESIM as 'truth'.
- Test the stochastic nodal analysis concept with actual production data.

A DATA ASSIMILATION ON THE EXTENDED GILBERT EQUATION

A.1 Prior distribution

Table A.1: Prior Gaussian distribution of uncertain model parameters A, B and C based on correlations from Gilbert [11], Ros [27], Baxendell [26] and Achong [28].

Model parameter	*A*	*B*	*C*
Correlation:			
Gilbert	3.75 x 10^{10}	0.546	1.89
Ros	6.52 x 10^{10}	0.500	2.00
Baxendell	3.58 x 10^{10}	0.546	1.93
Achong	1.43 x 10^{10}	0.650	1.88
Statistical information:			
Mean, μ	3.82 x 10^{10}	0.561	1.93
Standard deviation, σ_{calc}	2.09 x 10^{10}	0.064	0.055
Standard deviation used, σ	0.5 x 2.09 x 10^{10}	0.7 x 0.064	0.7 x 0.0.055

If the standard deviation is calculated based on the correlations, negative values for A could be sampled from the distribution. The standard deviation from Table A.1 is lower than the standard deviation based on the four correlations.

For model parameter A, two Gaussian distributions are displayed in Figure 1, the first distribution based on $\mu = \mu_A$ and $\sigma = \sigma_A$ and the second distribution based on on $\mu = \mu_A$ and $\sigma = 0.5 \times \sigma_A$. The first normal distribution has negative values for A which is not possible (leads to positive flow rates). This is the reason of taking $0.5 \times \sigma_A$ as the standard deviation for the prior distribution of parameter A.

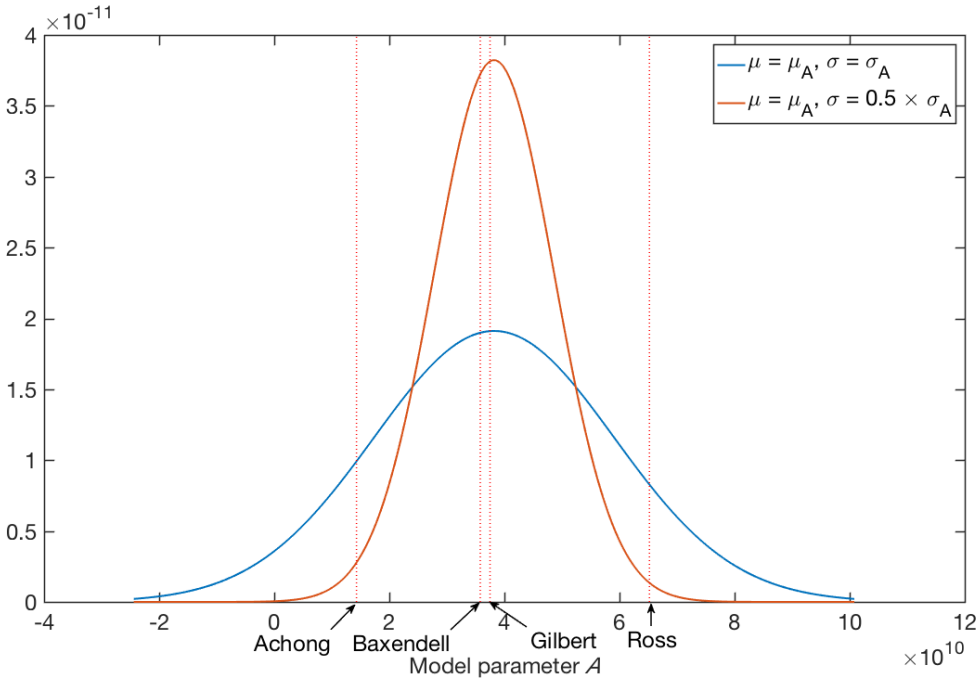


Figure A.1: probability density function (pdf) of prior distribution of model parameter A. The prior distribution is a Gaussian distribution based on the mean and standard deviation calculated based on the Achong, Baxendell, Gilbert and Ros correlation. Based on $N(\mu_{A,}\sigma_A)$ negative values could occur for realizations A. The distribution based on $N(\mu_{A,}0.5\ \sigma_A)$ is used as prior distribution to sample parameter A in stochastic data assimilation.

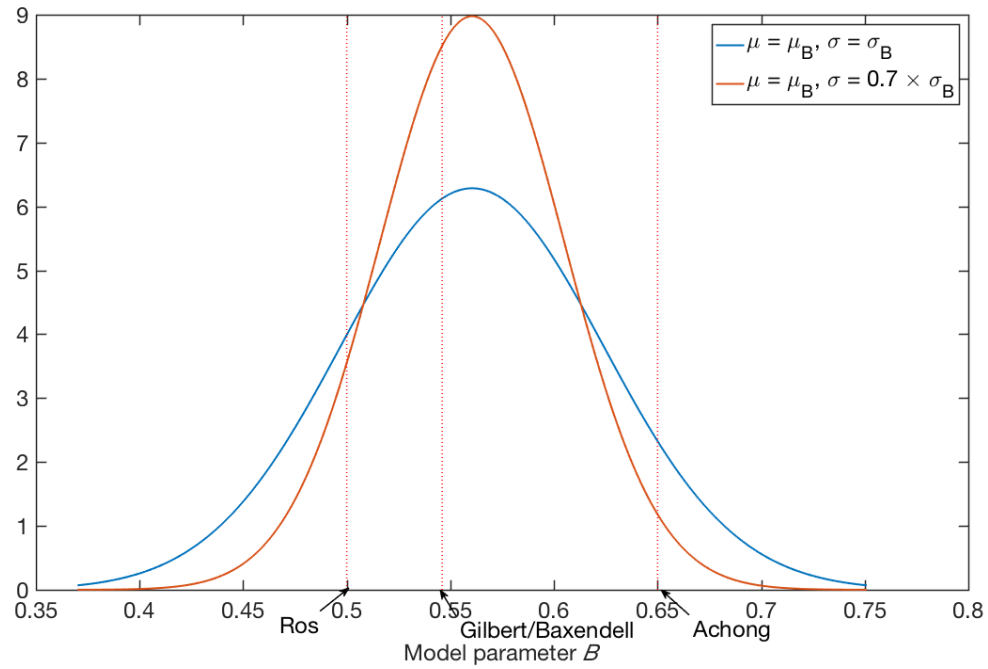


Figure A.2: probability density function (pdf) of prior distribution of model parameter B. The prior distribution is a Gaussian distribution based on the mean and standard deviation calculated based on the Achong, Baxendell, Gilbert and Ros correlation. Based on $N(\mu_{B,}\sigma_B)$ the spread in the prior is too large and does not take Ros and Achong values as boundary values. The distribution based on $N(\mu_{B,}0.7\ \sigma_A)$ is used as prior distribution to sample parameter B in stochastic data assimilation.

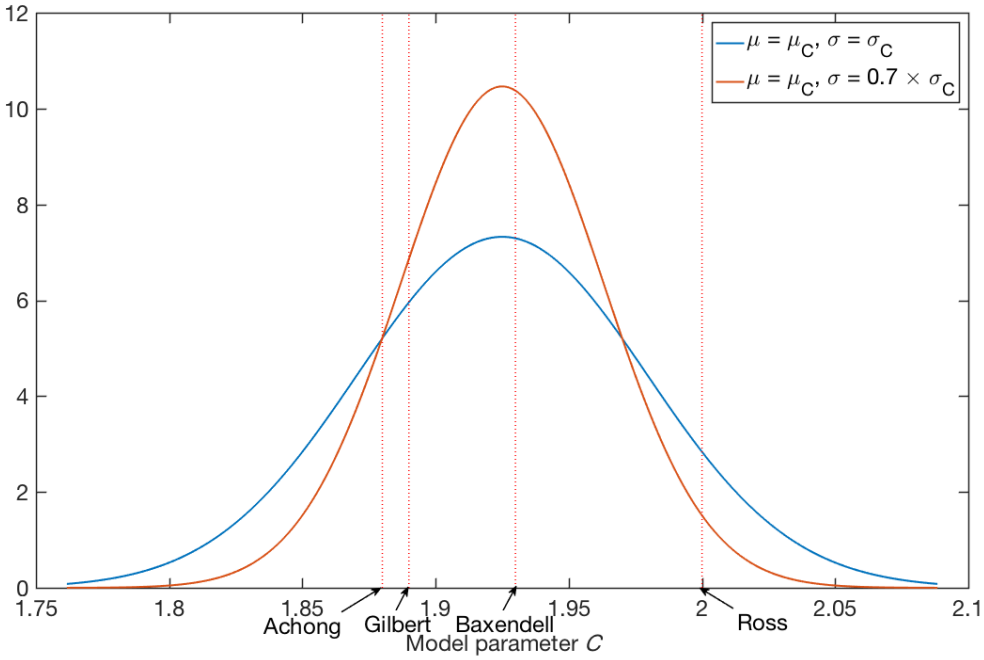


Figure A.3 probability density function (pdf) of prior distribution of model parameter C. The prior distribution is a Gaussian distribution based on the mean and standard deviation calculated based on the Achong, Baxendell, Gilbert and Ros correlation. Based on $N(\mu_{C,}\sigma_C)$ the spread is too high. The distribution based on $N(\mu_{C,}0.7\ \sigma_C)$ is used as prior distribution to sample parameter C in stochastic data assimilation

A.2 Sensitivity analysis

Based on the prior distributions the sensitivity of the model can be analysed by inspecting the effect of changing a parameter to the tubing head pressure p_{tf}. The sensitivity analysis is performed by first defining a base case with fixed inputs d_{ch}, p_{fl}, $q_{l,sc}$ and R_{gl} and model parameters A, B and C. The model parameters are fixed with the mean values which can be found in Table A.1. With the base case parameters, the base case output of p_{tf} is calculated with the extended Gilbert equation. In the sensitivity analysis the correlation of the model parameters with the output is obtained. The results of the sensitivity analysis are displayed in Figure A.4.

Table A.2: Base case parameters for sensitivity analysis of the extended Gilbert equation by changing model parameters A, B, and C with the standard deviation defined in Appendix A.1

Fixed variables	*Symbol*	*Unit*	*Base*	*Lower bound*	*Upper bound*
Choke diameter	d_{ch}	1/64th in	28	-	-
Downstream pressure of the choke	p_{fl}	MPa	2	-	-
Liquid production rate	$q_{l,sc}$	m^3/s	-0.004	-	-
Producing gas-liquid ratio	R_{gl}	m^3/ m^3	150	-	-
Water cut	$f_{w,sc}$	-	0	-	-
Model parameters					
Proportionality constant	A	-	μ_A	$\mu_A - 3\,\sigma_A$	$\mu_A + 3\,\sigma_A$
Gas-liquid ratio exponent	B	-	μ_B	$\mu_B - 3\,\sigma_B$	$\mu_B + 3\,\sigma_B$
Choke size exponent	C	-	μ_C	$\mu_C - 3\,\sigma_C$	$\mu_C + 3\,\sigma_C$

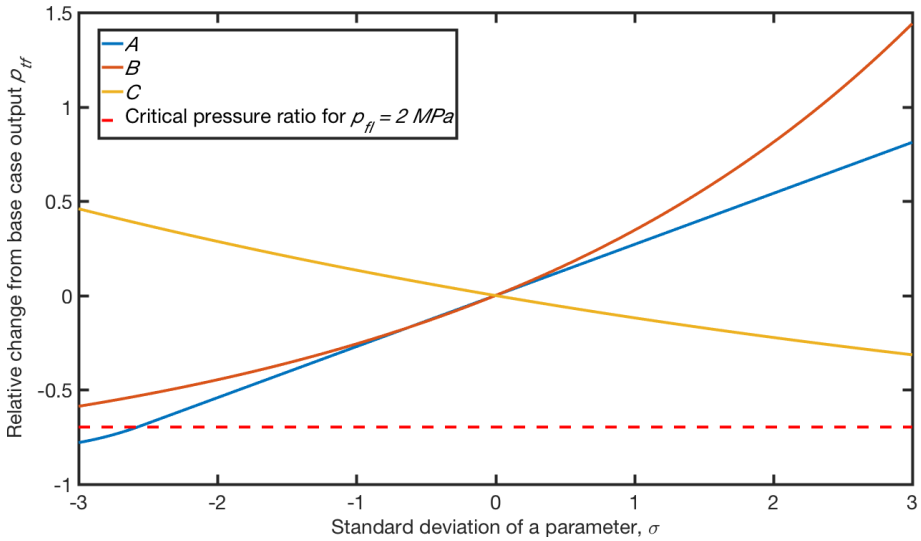


Figure A.4: sensitivity analysis of model parameters A, B and C of the extended Gilbert equation. Base case values in Table 2.

A.3 Input tables

A.3.1 Truth model

Table A.3: 'truth' data for twin experiment based on theoretical choke model.

Variable	*Symbol*	*Unit*	*Value*
	Production data:		
Production phase			test separator phase
Water cut	$f_{w,sc}$	–	0
Oil rate	$q_{o,sc}$	m^3/s	- 0.004
Producing liquid gas-oil ratio	R_{gl}	m^3/ m^3	150
Flowline pressure	p_{fl}	Pa	2 MPa
	Choke data:		
Choke size	d_{ch}	$1/64^{th}$ in	28
Specific heat capacity of liquid at	c_l	J/(kg K)	0.80
Specific heat capacity of gas at constant	c_p	J/(kg K)	0.25
Specific heat capacity of gas at constant	c_v	J/(kg K)	0.20
Choke discharge coefficient	C_d	-	0.7
	Temperature data:		
Tubing head temperature	T_{tf}	°C	80
	Fluid properties:		
Oil model			Glaso correlation [48]
Gas density at standard conditions	$\rho_{g,sc}$	kg/m^3	0.9
Oil density at standard conditions	$\rho_{o,sc}$	kg/m^3	890
Water density at standard conditions	$\rho_{w,sc}$	kg/m^3	1000
	Output data:		
Tubing head pressure	p_{tf}	MPa	12.8

A.3.2 Input parameters for data assimilation

Table A.4: parameters for twin experiment 1, choke.

Variable	*Symbol*	*Unit*	*Value*
Stochastic data assimilation data:			
Number of realizations	N_r		1000
Number of measured inputs	N_u		2
Number of uncertain parameters	N_m		3
Number of measured/model outputs	N_y		1
Length of extended state vector	N_z		4
Fixed production data:			
Production phase			test separator phase
Water cut	$f_{w,sc}$	–	0
Flowline pressure	p_{fl}	Pa	2 MPa
Measured production data (inputs):			
Measured oil rate	$q_{o,sc}$	m^3/s	- 0.004
Measurement error oil rate	$\epsilon_{q_{o,sc}}$	m^3/s	$N(0, 1 \times 10^{-5})$
Measured gas rate	$q_{g,sc}$	m^3/s	- 0.060
Measurement error oil rate	$\epsilon_{q_{g,sc}}$	m^3/s	$N(0, 1 \times 10^{-3})$
Choke data:			
Choke size	d_{ch}	$1/64^{th}$ in	28
Uncertain choke parameters:			
Mean of A	μ_A	-	3.82×10^{10}
Standard deviation of A	σ_A	-	1.04×10^{10}
Mean of A	μ_B	-	0.561
Standard deviation of A	σ_B	-	0.044
Mean of A	μ_C	-	1.93
Standard deviation of A	σ_C	-	0.038
Proportionality constant	A	-	$N(\mu_A, \sigma_A)$
Gas-oil ratio exponent	B	-	$N(\mu_B, \sigma_B)$
Choke size exponent	C	-	$N(\mu_C, \sigma_C)$
Measured output:			
Tubing head pressure	p_{tf}	MPa	12.8
Measurement error of tubing head	$\epsilon_{p_{tf}}$	MPa	$N(0, 0.1)$

A.4 Correlation matrix based on prior ensemble

The effect of ensemble size on the correlation matrix of the Gilbert choke model is stated down below, with higher realizations lower model parameter – model parameter correlations are observerd.

Table D.1: Correlation matrix based on N_r = 50.

	$m_1 = A$	$m_2 = B$	$m_3 = C$	$y = p_{tf,choke}$
$m_1 = A$	1.00	-0.08	-0.06	0.51
$m_2 = B$	-0.08	1.00	-0.12	0.74
$m_3 = C$	-0.06	-0.12	1.00	-0.39
$y = p_{tf,choke}$	0.51	0.74	-0.39	1.00

Table D.2: Correlation matrix based on N_r = 100.

	$m_1 = A$	$m_2 = B$	$m_3 = C$	$y = p_{tf,choke}$
$m_1 = A$	1.00	-0.04	0.15	0.48
$m_2 = B$	-0.04	1.00	-0.03	0.77
$m_3 = C$	0.15	-0.03	1.00	-0.25
$y = p_{tf,choke}$	0.48	0.77	-0.25	1.00

Table D.3: Correlation matrix based on N_r = 200.

	$m_1 = A$	$m_2 = B$	$m_3 = C$	$y = p_{tf,choke}$
$m_1 = A$	1.00	0.03	0.00	0.58
$m_2 = B$	0.03	1.00	0.03	0.73
$m_3 = C$	0.00	0.03	1.00	-0.25
$y = p_{tf,choke}$	0.58	0.73	-0.25	1.00

Table D.4: Correlation matrix based on N_r = 1000.

	$m_1 = A$	$m_2 = B$	$m_3 = C$	$y = p_{tf,choke}$
$m_1 = A$	1.00	0.01	-0.02	0.61
$m_2 = B$	0.01	1.00	-0.04	0.69
$m_3 = C$	-0.02	-0.04	1.00	-0.32
$y = p_{tf,choke}$	0.61	0.69	-0.32	1.00

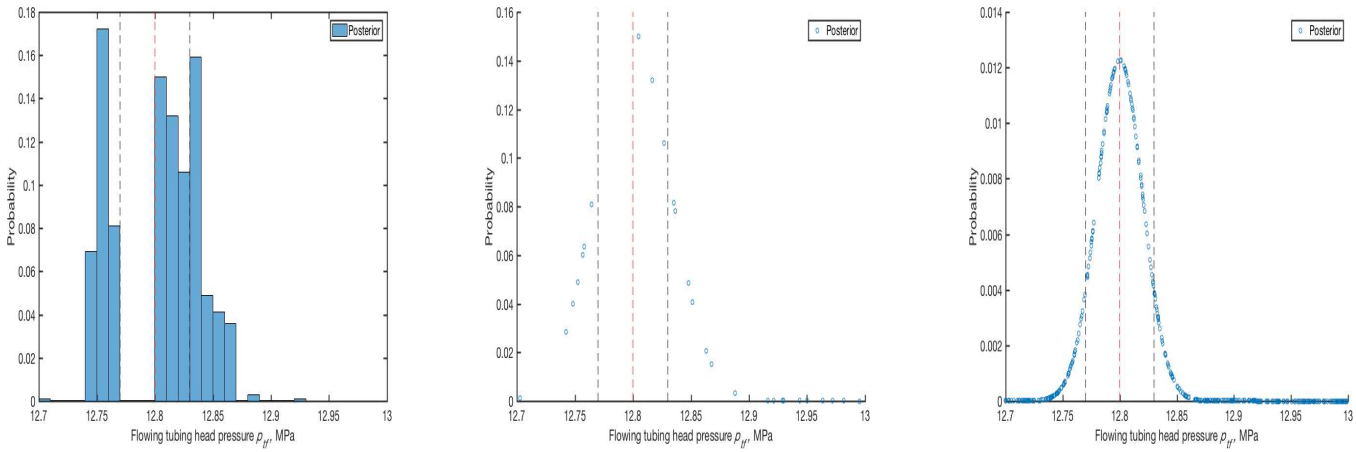
Table D.5: Correlation matrix based on N_r = 5000.

	$m_1 = A$	$m_2 = B$	$m_3 = C$	$y = p_{tf,choke}$
$m_1 = A$	1.00	0.00	0.03	0.62
$m_2 = B$	0.00	1.00	-0.03	0.69
$m_3 = C$	0.03	-0.03	1.00	-0.28
$y = p_{tf,choke}$	0.62	0.69	-0.28	1.00

A.5 Particle filter

A.5.1 Number of realizations

In Table A.2, the number of non-zero weighted particles for different number particles is displayed. The results of Figure A.5b are obtained with a 10x bigger measurement error than defined for EnKF, this due the limited non-zero particles obtained otherwise. This can be quantified by calculating the effective numbers of particles detailed explanation. In Figure A.5b the posterior distribution with SIS is displayed. Based on the posterior distribution generated with regularization and another model run for the same data assimilation step weights are obtained, shown in Figure A.5c (which are the weights for ASIR). Filter degeneracy compared to Figure A.5b is reduced. A better approximation of the Gaussian distribution is observed, because the variance in the predicted measurements is smaller. This motivated the choice to reduce the measurement variance compared to the first data assimilation step form 10x to 4x used in the ASIR weighting step (Equation 2.24)



(a) Based on residual resampling

(b) SIS

(c) SIS based on model parameters obtained with SIR

Figure A.5: Effect of resampling to obtain more effective particles

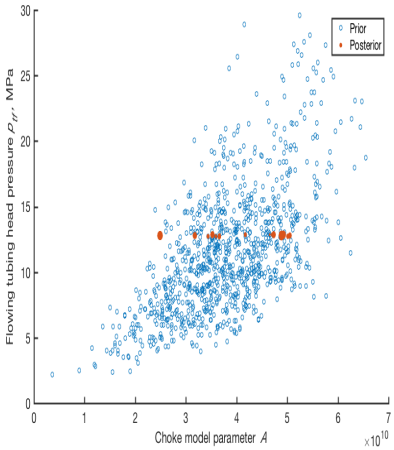
Table A.2: Number of particles and effect of measurement error on non-zero particles.

	$N_p = 100$	$N_p = 500$	$N_p = 1000$	$N_p = 5000$
$1\ \epsilon_{p_{\mathrm{tf}}}$	1	1	6	15
$5\ \epsilon_{p_{\mathrm{tf}}}$	1	3	11	60
$10\ \epsilon_{p_{\mathrm{tf}}}$	1	7	18	74

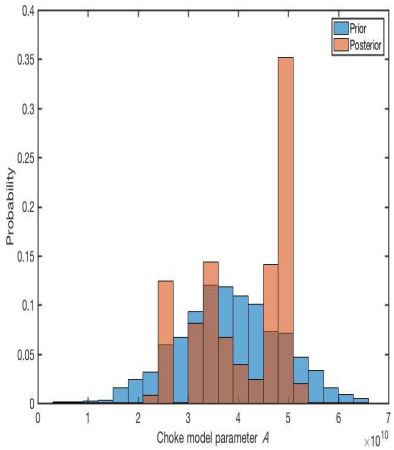
A.5.2 Resampling applied on the Gilbert equation

Available resampling methods for the SIR implementation of PF are residual resampling and regularization. The auxiliary particle filter is applied with regularization. Based on the weights of the particle filter (Figure A.5a) and regularization the posterior distribution of model parameters is calculated (Figure A.6b). Based on the updated posterior model parameters (Figure A.6b) the output is calculated for the same data assimilation step, hence another run of the model for each particle is required. Weights are again calculated (Figure A.6c), due the smaller variance in the predicted output, more particles with weights will be observed, (see Figure A.5c). By resampling again, the posterior distribution of model parameters is observed. Note the difference between Figures A.6a and A.6c, the variance in predicted measurement is clearly reduced.

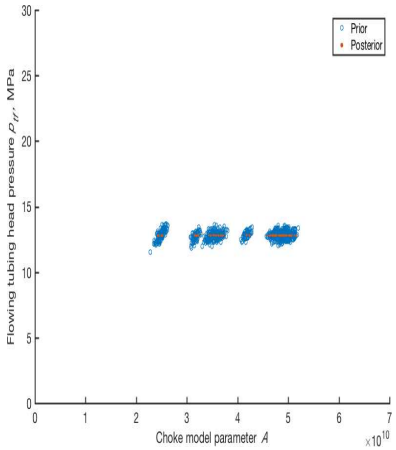
The difference implementation of the particle filter lead to different distributions of the model parameters, see Figure A.7, A.8 and A9. The update is also plotted with for cross plot of model parameters, see Figure A.10, A.11, and A.12.



(a) Assign weights based on predicted measurement and actual measurements (SIS)

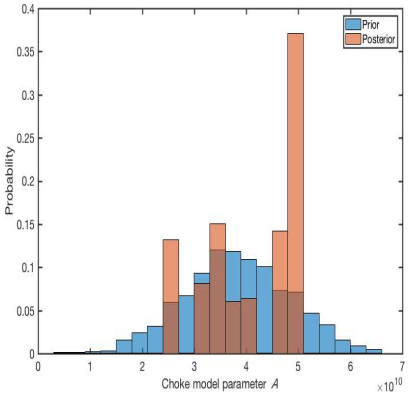


(b) Resample with Gaussian noise, (SIR with regularization)

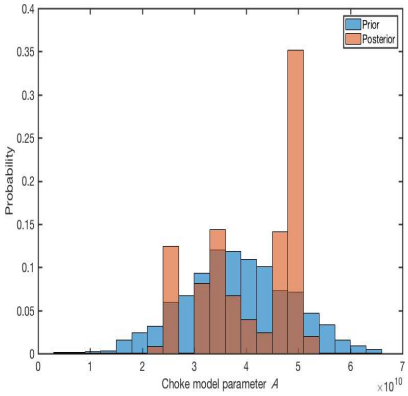


(c) Cross plot A and y, for the ASIR

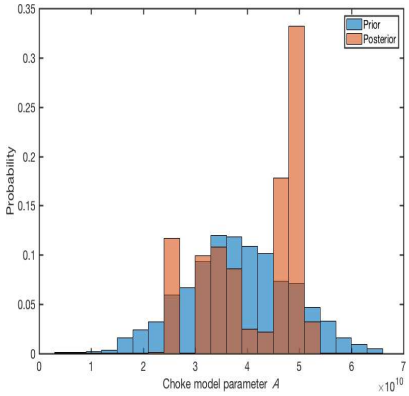
Figure A.6: The auxiliary particle filter (ASIR) workflow, till (b), same as SIR. The model parameters obtained with SIR are used to predict the output and redo steps (a) and (b). See (c) for predicted output based on SIR.



(a) SIR - residual

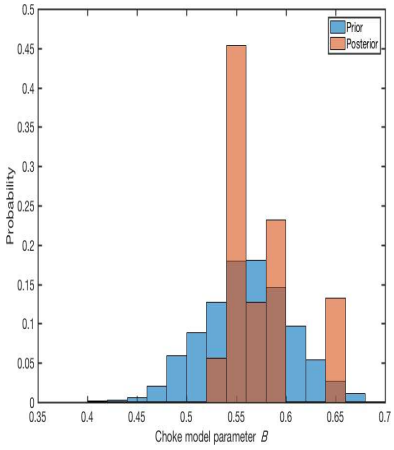


(b) SIR - regularization

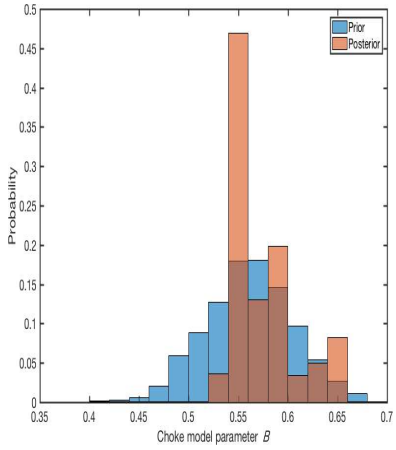


(c) ASIR

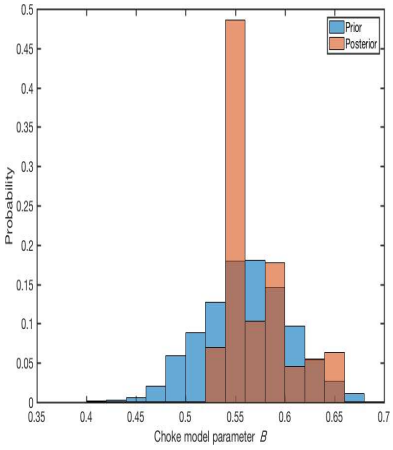
Figure A.7: Prior and posterior distribution based on PF of model parameter A.



(a) SIR - residual

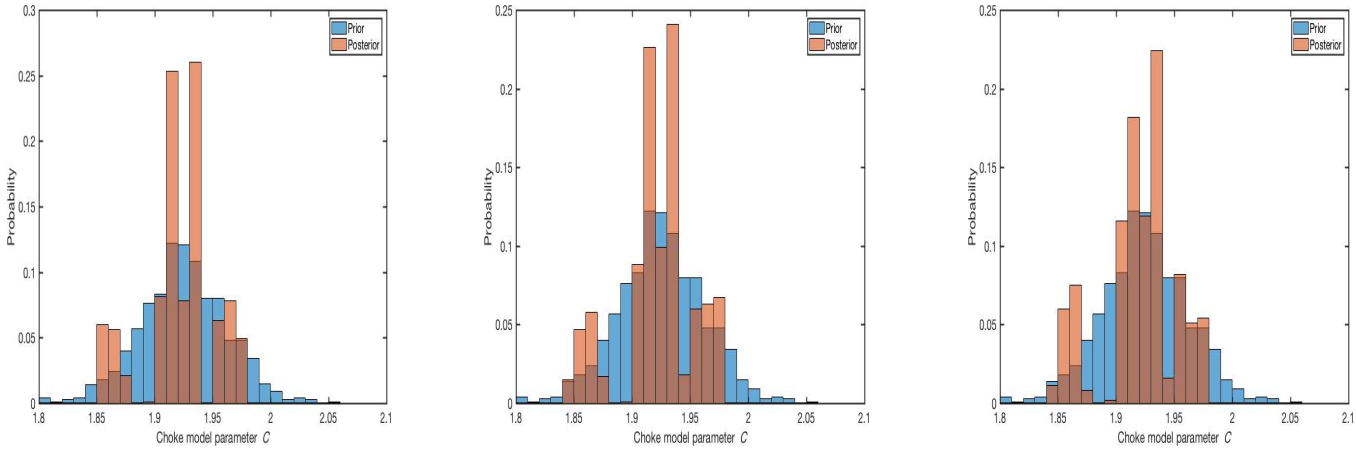


(b) SIR - regularization



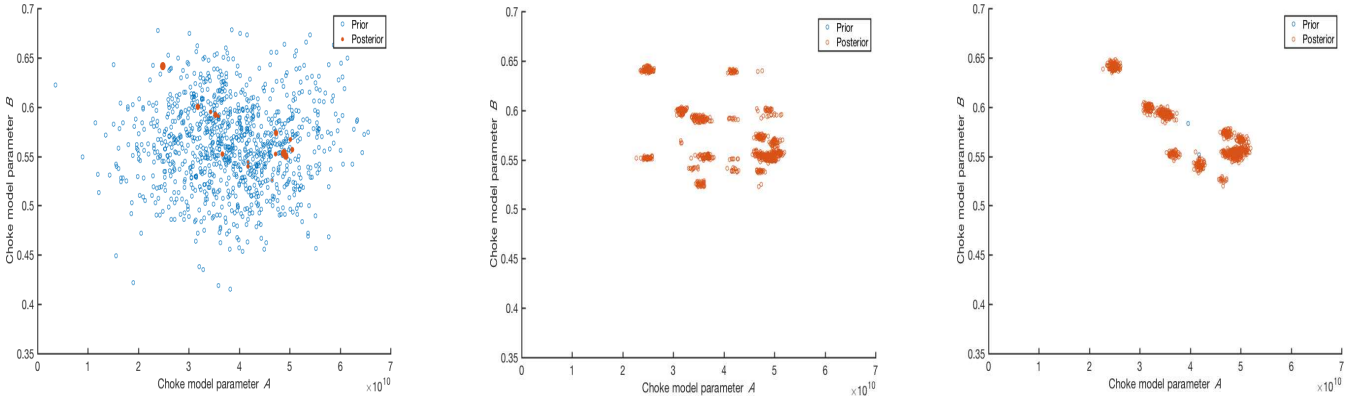
(c) ASIR

Figure A.8: Prior and posterior distribution of model parameter B based on PF.



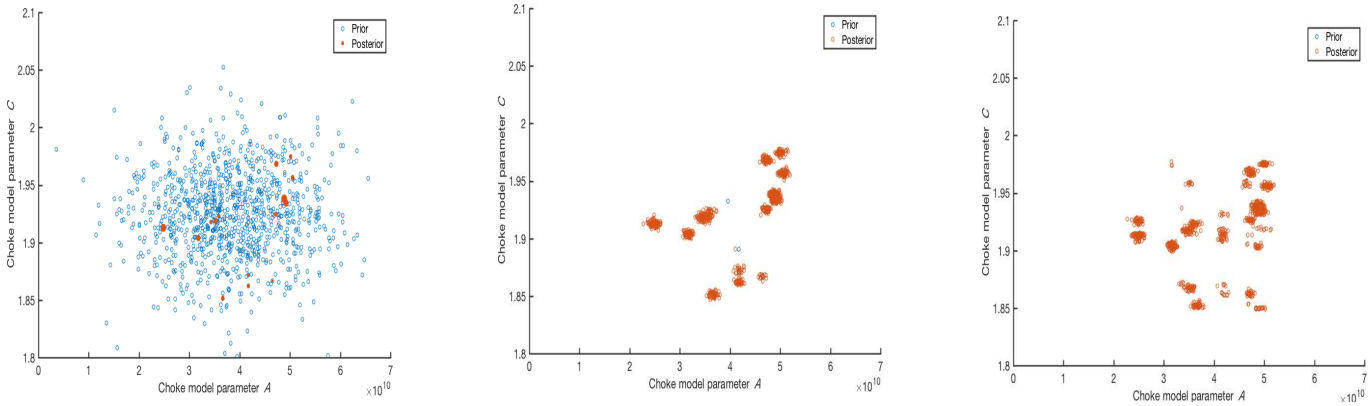
(a) SIR - residual (b) SIR - regularization (c) ASIR

Figure A.9: Prior and posterior distribution of model parameter C based on PF.



(a) cross plot A and B (b) Resampling with regularization (C) ASIR

Figure A.10: Resampling with ASIR illustrated with correlation of A and B, (a) weighting particles, (b) based on weights, posterior distribution is obtained with regularization of particles and (c) the ASIR step, see Figure A.6.



(a) cross plot A and C (b) Resampling with regularization (C) ASIR

Figure A.11: Resampling with ASIR illustrated with correlation of A and C, (a) weighting particles, (b) based on weights, posterior distribution is obtained with regularization of particles and (c) the ASIR step.

A.5.3. Pseudocode Resampling methods

Pseudo code SIR - Residual

$\mathbf{m}_{prior}$ = prior vector of model parameter
$\mathbf{w}$ = weights of particles
N_r = number of particles
$\mathbf{n}$ = integer part of $N_r\mathbf{w}$

$\mathbf{ind}$ = find($w > 1/N_r$) % Find all particles with integers
$\mathbf{n}$ = floor($\mathbf{w}\, N_r$) % Get integer part of particles of index
$\mathbf{n}_{cum}$ = cumsum($\mathbf{n}(\mathbf{ind})$) % Cumulative sum vector of $\mathbf{n}$ based on indices, to get cumulative values of how much particle to copy

% Duplicate particles
$\mathbf{m}_{posterior}(1{:}\mathbf{n}_{cum}(1)) = \mathbf{m}_{prior}(\mathbf{ind}(1))$ % Copy particles
for i = 2:length($\mathbf{n}_{cum}$)
    $\mathbf{m}_{posterior}(\mathbf{n}_{cum}(i\text{-}1)+1{:}\mathbf{n}_{cum}(i)) = \mathbf{m}_{prior}(\mathbf{ind}(i))$ % Copy particles
end

% Resample remaining particles with systemic resampling
$n_{residual} = N_r - \text{sum}(\mathbf{n})$ % Remaining particles
$P_{residual} = N_r\mathbf{w} - \mathbf{n}$ % Update probabilities, remove already copied particles
$\mathbf{w}_{residual} = \frac{P_{residual}}{sum(P_{residual})}$ % Normalize weights
$\mathbf{m}_{posterior}(\text{sum}(\mathbf{n})+1{:}N_r) = \text{uniform\_resample}(\mathbf{m}_{prior}, \mathbf{w}_{residual}, n_{residual})$

Pseudo code SIR – Regularization
% same till duplicate particle
σ= standard deviation of the model parameter, needs to be specified
$\mathbf{m}_{posterior}(1{:}\mathbf{n}(1)) = \mathbf{m}_{prior}(\text{index}(1)) + \sigma\ \text{randn}(1{:}\mathbf{n}(1))$
for i = 2:length($\mathbf{n}$)
    $\mathbf{m}_{posterior}(\mathbf{n}_{cum}(i\text{-}1)+1{:}\mathbf{n}_{cum}(i)) = \mathbf{m}_{prior}(\text{index}(i)) + \sigma\ \text{randn}(\mathbf{n}_{cum}(i\text{-}1)+1{:}\mathbf{n}_{cum}(i))$
end
$\mathbf{m}_{posterior}(\text{sum}(\mathbf{n})+1{:}N_r) = \text{uniform\_resample}(\mathbf{m}_{prior}, \mathbf{w}_{residual}, n_{residual})$

B MULTI-PHASE FLOW MODELLING THROUGH WELLS

B.1 Sensitivity analysis

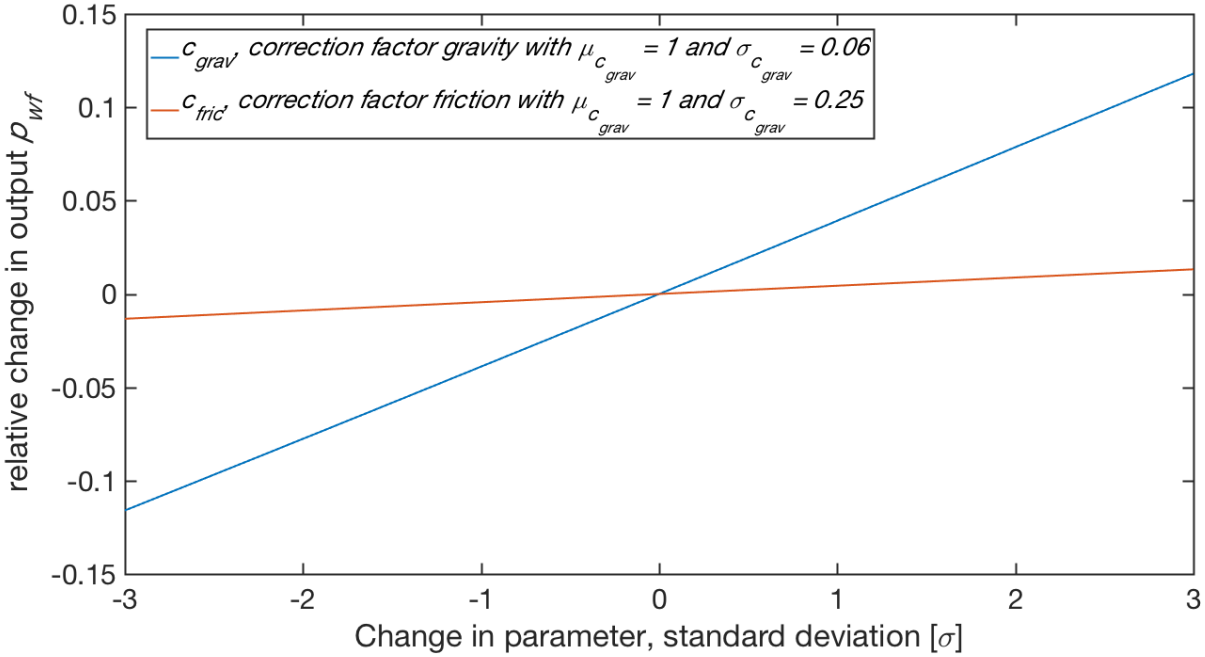


Figure B.1: sensitivity analysis of model parameters A, B and C of the extended Gilbert equation. Base case values in Table 2.

B.2 Case study – effect of prior distribution of model parameters on posterior pdf

In this appendix, the effect of the prior distribution of the uncertain parameters is discussed with a case study. Method used is the EnKF with 200 ensemble members.

See Appendix B.3 for a detailed overview of the inputs of the models. See Figure 4.3 for difference between 'truth' tubing intake curve and 'base' intake curve which will be as mean in the modelling of priors.

Table B.1: parameters for twin experiment in multi-phase well flow models.

Variable	*Symbol*	*Unit*	*'Truth' value*	*'Base' value*
Well data:				
Pipe roughness	e_w	m	30×10^{-6}	30×10^{-5}
Multi-phase flow model			Drift flux	Mukherjee&Brill
Fluid properties:				
Black oil correlation			Standing[47]	Glaso[48]

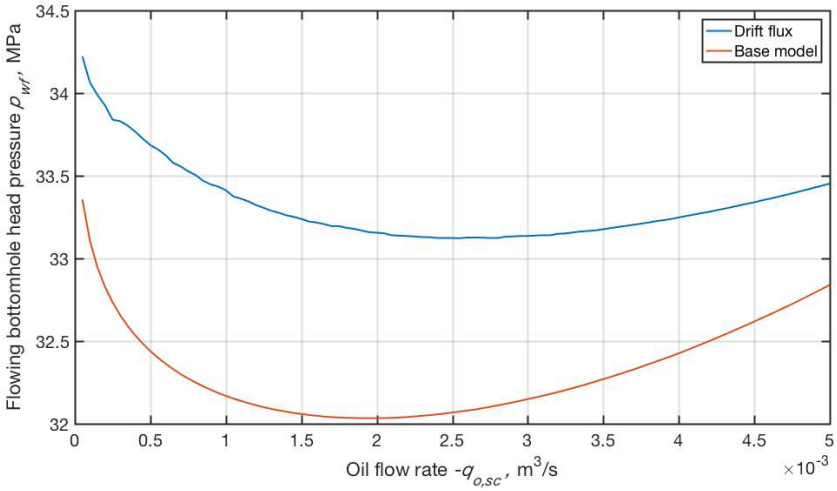


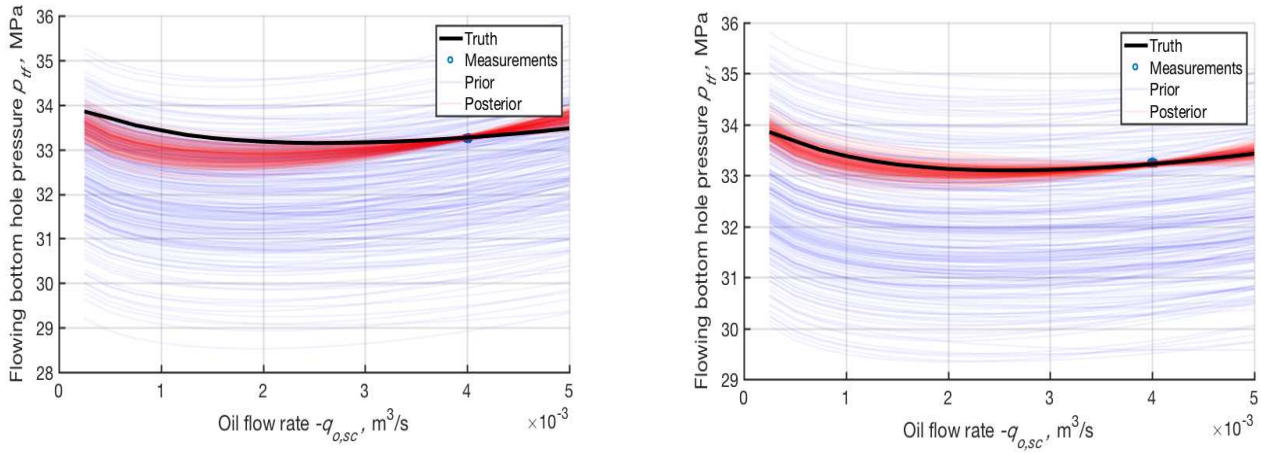
Figure B.2: Tubing intake curves calculated based on Table 4.1 and inputs values from Appendix C.

This twin experiment is conducted twice, with two different prior distributions of the model parameters. See Table B.2 for details about the prior distributions.

Table B.2: Prior distribution for twin experiment in multi-phase well flow models.

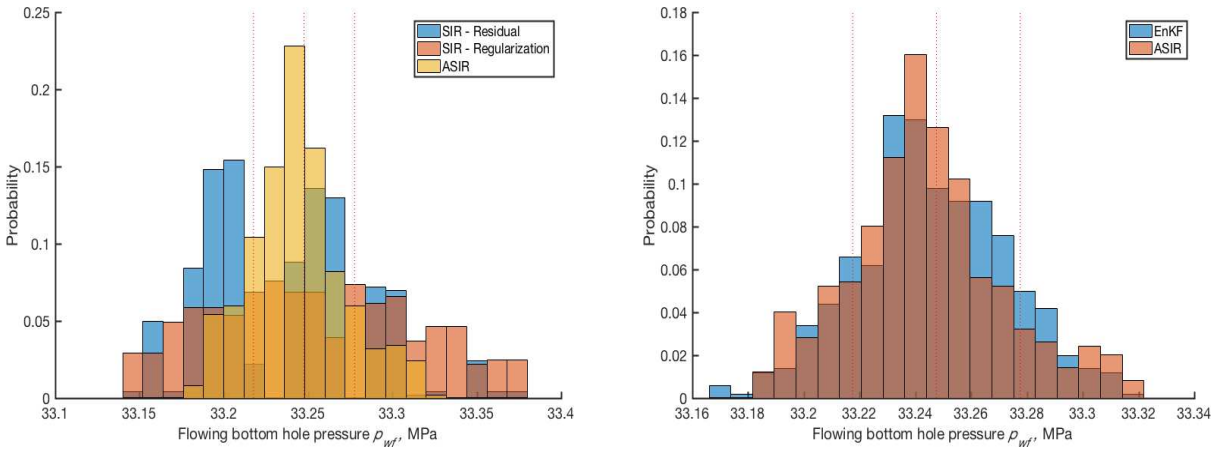
Variable	*Symbol*	*Case 1*	*Case 2*
Correction factors:			
Gravity correction factor	c_{geav}	*N(1, 0.05)*	*N(1, 0.05)*
Friction correction factor	c_{fric}	*N(1, 0.3)*	*N(0.6,0.2)*

With this case study with a different pipe roughness between 'truth' and 'base' model, see Table B.1, the friction term will be overestimated. By manipulating the prior distribution, the posterior tubing intake curve will be more in the middle of the ensemble, see Figure B.3a and Figure B.3b. This example shows that there are chances for ensemble collapse, i.e., the model parameters to predict the 'truth' curve will be not in the ensemble after calibration.



(a) Prior case 1 (b) Prior case 2

Figure B.3: Tubing intake curves based on prior and posterior distribution Table B.1 and B.2.



(a) PF (b) EnKF

Figure B.4: Predicted measurements based on posterior distribution.

B.3 Input tables

Table B.3: Data for 'truth'.			
Variable	*Symbol*	*Unit*	*Value*
Production data:			
Water cut	$f_{w,sc}$	–	0
Producing GOR	R_{go}	m^3/m^3	150
Tubing head pressure	p_{tf}	Pa	12 MPa
Oil rate at standard conditions	$q_{o,sc}$	m^3/s	-0.004
Fluid data:			
Oil model	Black oil with Standing correlation [47]		
Solution GOR at bubble point pr.	R_{sb}	m^3/m^3	150
Gas density at standard conditions	$\rho_{g,sc}$	kg/m^3	0.90
Oil density at standard conditions	$\rho_{o,sc}$	kg/m^3	860
Well data:			
AHD	L_{tot}	m	3000
Wellbore inclination	α	deg	0
Tubing inside diameter	d_t	m	0.07600
Tubing roughness	e_t	m	30×10^{-6}
Multi-phase flow correlation	Drift flux model		
Flow regime	Semi-steady-state		
Output:			
Bottom hole pressure	p_{wf}	MPa	33.25

Table B.4: Data for 'model'.			
Variable	*Symbol*	*Unit*	*Value*
Production data:			
Water cut	$f_{w,sc}$	–	0
Tubing head pressure	p_{tf}	MPa	12.8
Measurement error of tubing head	$\epsilon_{p_{tf}}$	MPa	$N(0, 0.1)$
Measured production data (inputs):			
Measured oil rate	$q_{o,sc}$	m^3/s	- 0.004
Measurement error oil rate	$\epsilon_{q_{o,sc}}$	m^3/s	$N(0, 1 \times 10^{-5})$
Measured gas rate	$q_{g,sc}$	m^3/s	- 0.060
Measurement error oil rate	$\epsilon_{q_{g,sc}}$	m^3/s	$N(0, 1 \times 10^{-3})$
Fluid data:			
Oil model	Black oil with Standing correlation		
Solution GOR at bubble point pr.	R_{sb}	m^3/m^3	150
Gas density at standard conditions	$\rho_{g,sc}$	kg/m^3	0.90
Oil density at standard conditions	$\rho_{o,sc}$	kg/m^3	890
Stochastic parameters data:			
Gravity correction factor	c_{grav}		$N(1, 0.06)$
Friction correction factor	c_{fric}		$N(1, 0.25)$
Well data:			
AHD	L_{tot}	m	3000
Wellbore inclination	α	deg	0
Tubing inside diameter	d_t	m	0.07600
Tubing roughness	e_t	m	30×10^{-6}
Output:			
Bottom hole pressure, FBHP	p_{wf}	MPa	33.2
Measurement error FBHP	$\epsilon_{p_{wf}}$	MPa	$N(0, 0.01)$

C INPUT DATA NEAR-WELL BORE

Table C.1: Data for 'truth'.			
Variable	*Symbol*	*Unit*	*Value*
Production data:			
Water cut	$f_{w,sc}$	–	0
Producing GOR	R_{go}	m^3/m^3	150
Reservoir pressure	$p_{R,av}$	Pa	40 MPa
Oil rate at standard conditions	$q_{o,sc}$	m^3/s	-0.004
Fluid data:			
Oil model		Black oil with Standing	
Solution GOR at bubble point pr.	R_{sb}	m^3/m^3	150
Gas density at standard conditions	$\rho_{g,sc}$	kg/m^3	0.90
Oil density at standard conditions	$\rho_{o,sc}$	kg/m^3	890
Reservoir data:			
Reservoir height	h_R	m	10
Permeability	k	m^2	1×10^{-13}
Forchheimer coefficient	β	1/m	0
Skin	S	–	0
Reservoir drainage radius	r_e	m	400
Well bore radius	r_w	m	0.1778
Flow regime		Semi-steady-state	
Output:			
Bottom hole pressure, FBHP	p_{wf}	MPa	38.12

Table C.2: Data for 'model'.

Variable	*Symbol*	*Unit*	*Value*
Production data:			
Water cut	$f_{w,sc}$	–	0
Reservoir pressure	$p_{R,av}$	Pa	40 MPa
Measured oil rate	$q_{o,sc}$	m^3/s	- 0.004
Measurement error oil rate	$\epsilon_{q_{o,sc}}$	m^3/s	$N(0, 1 \times 10^{-5})$
Measured gas rate	$q_{g,sc}$	m^3/s	- 0.060
Measurement error oil rate	$\epsilon_{q_{g,sc}}$	m^3/s	$N(0, 1 \times 10^{-3})$
Fluid data:			
Oil model	Black oil with Standing correlation [47]		
Solution GOR at bubble point pr.	R_{sb}	m^3/m^3	150
Gas density at standard conditions	$\rho_{g,sc}$	kg/m^3	0.90
Oil density at standard conditions	$\rho_{o,sc}$	kg/m^3	890
Reservoir data:			
Reservoir height	h_R	m	10
Permeability	k	m^2	1×10^{-13}
Forchheimer coefficient	β	1/m	0
Skin	S	–	0
Reservoir drainage radius	r_e	m	400
Well bore radius	r_w	m	0.1778
Flow regime	Semi-steady-state		
Productivity index correction factor	c_J		$N(1, 0.25)$
Output:			
Bottom hole pressure, FBHP	p_{wf}	MPa	38.12
Measurement error FBHP	$\epsilon_{p_{wf}}$	MPa	$N(0, 0.01)$

D PRODUCTION SYSTEMS

Table D.1: Data for 'truth'.

Variable	*Symbol*	*Unit*	*Value*
Production data:			
Water cut	$f_{w,sc}$	–	0
Producing GOR	R_{go}	m^3/m^3	150
Reservoir pressure	$p_{R,av}$	Pa	40 MPa
Oil rate at standard conditions	$q_{o,sc}$	m^3/s	-0.004
Fluid data:			
Oil model	Black oil with Standing correlation [47]		
Solution GOR at bubble point pr.	R_{sb}	m^3/m^3	150
Gas density at standard conditions	$\rho_{g,sc}$	kg/m^3	0.90
Oil density at standard conditions	$\rho_{o,sc}$	kg/m^3	860
Choke data:			
Choke size	d_{ch}	$1/64^{th}$ in	28
Choke model	Theoretical model, see A.3.		
See Table A.3 for other inputs			
Well data:			
AHD	L_{tot}	m	3000
Wellbore inclination	α	deg	0
Tubing inside diameter	d_t	m	0.07600
Tubing roughness	e_t	m	30×10^{-6}
Multi-phase flow correlation	Drift flux model		
Reservoir data:			
Reservoir height	h_R	m	10
Permeability	k	m^2	1×10^{-13}
Skin	S	–	0
Reservoir drainage radius	r_e	m	400
Well bore radius	r_w	m	0.1778
Flow regime	Semi-steady-state		
Output:			
Tubing head pressure	p_{tf}	MPa	15.15
Bottom hole pressure	p_{wf}	MPa	35.17

Table D.2: parameters for twin experiment, Chapter 6.

Variable	*Symbol*	*Unit*	*Value*
Stochastic data assimilation data:			
Number of realizations	N_r		200 EnKF, 1000 PF
Number of measured inputs	N_u		2
Number of uncertain parameters	N_m		6
Number of measured/model outputs	N_y		3
Length of extended state vector	N_z		9
Fixed production data:			
Production phase		test separator phase	
Water cut	$f_{w,sc}$	–	0
Reservoir pressure	$p_{R,av}$	Pa	40 MPa
Flowline pressure	p_{fl}	Pa	2 MPa
Measured production data (inputs):			
Measured oil rate	$q_{o,sc}$	m^3/s	- 0.0036
Measurement error oil rate	$\epsilon_{q_{o,sc}}$	m^3/s	$N(0, 1 \times 10^{-5})$
Measured gas rate	$q_{g,sc}$	m^3/s	- 0.5429
Measurement error oil rate	$\epsilon_{q_{g,sc}}$	m^3/s	$N(0, 1 \times 10^{-3})$
Fluid data:			
Oil model	Black oil with Glaso correlation [47]		
Solution GOR at bubble point pr.	R_{sb}	m^3/m^3	150
Gas density at standard conditions	$\rho_{g,sc}$	kg/m^3	0.90
Oil density at standard conditions	$\rho_{o,sc}$	kg/m^3	860
Choke data:			
Choke size	d_{ch}	$1/64^{th}$ in	28
Well data:			
AHD	L_{tot}	m	3000
Wellbore inclination	α	deg	0
Tubing inside diameter	d_t	m	0.07600
Tubing roughness	e_t	m	30×10^{-6}
Multi-phase flow correlation	Mukherjee and Brill (1985a)		
Reservoir data:			
Reservoir height	h_R	m	15
Permeability	k	m^2	1×10^{-13}
Reservoir drainage radius	r_e	m	800
Well bore radius	r_w	m	0.1778
Uncertain parameters:			

Mean of A	μ_A	-	3.82×10^{10}
Standard deviation of A	σ_A	-	1.04×10^{10}
Mean of B	μ_B	-	0.561
Standard deviation of B	σ_B	-	0.044
Mean of C	μ_C	-	1.93
Standard deviation of C	σ_C	-	0.038
Proportionality constant	A	-	$N(\mu_A, \sigma_A)$
Gas-oil ratio exponent	B	-	$N(\mu_B, \sigma_B)$
Choke size exponent	C	-	$N(\mu_C, \sigma_C)$
Gravity correction factor	c_{grav}		$N(1, 0.06)$
Friction correction factor	c_{fric}		$N(1, 0.25)$
Productivity index correction factor	c_J		$N(1, 0.25)$
Measured output:			
Tubing head pressure	p_{tf}	MPa	15.15
Bottom hole pressure	p_{wf}	MPa	35.17
Measurement error tubing head pressure	$\epsilon_{p_{tf}}$	MPa	$N(0, 0.01)$
Measurement error bottom hole pressure	$\epsilon_{p_{wf}}$	MPa	$N(0, 0.01)$

BIBLIOGRAPHY

[1] J. D. Jansen, *Nodal analysis of oil and gas wells – system modeling and numerical implementation*. SPE, Richardson. In production., 2017.

[2] M. Ghil and P. Malanotte-Rizzoli, "Data assimilation in meteorology and oceanography," *Advances in geophysics,* vol. 33, pp. 141-266, 1991.

[3] D. S. Oliver and Y. Chen, "Recent progress on reservoir history matching: a review," *Computers & Geosciences,* vol. 15, no. 1, pp. 185-221, 2011.

[4] R. J. Lorentzen, A. S. Stordal, and N. Hewitt, "An auxiliary adaptive Gaussian mixture filter applied to flowrate allocation using real data from a multiphase producer," *Computers & Geosciences,* vol. 102, pp. 34-44, 2017.

[5] R. J. Lorentzen, G. Nævdal, and A. C. Lage, "Tuning of parameters in a two-phase flow model using an ensemble Kalman filter," *International Journal of Multiphase Flow,* vol. 29, no. 8, pp. 1283-1309, 2003.

[6] H. Bloemen, S. Belfroid, W. Sturm, and F. Verhelst, "Soft sensing for gas-lift wells," in *SPE Annual Technical Conference and Exhibition*, 2004: Society of Petroleum Engineers.

[7] R. J. Lorentzen, A. S. Stordal, X. Luo, and G. Naevdal, "Estimation of production rates by use of transient well-flow modeling and the auxiliary particle filter: Full-scale Applications," *SPE Production & Operations,* vol. 31, no. 02, pp. 163-175, 2016.

[8] K. M. Muradov and D. R. Davies, "Zonal Rate Allocation in Intelligent Wells," 2009/1/1/, 2009.

[9] R. Malakooti, K. Muradov, D. Davies, and A. Kuznetsov, "Flow control optimisation to maximise the accuracy of multi-phase flow rate allocation," in *SPE Bergen One Day Seminar*, 2015: Society of Petroleum Engineers.

[10] M. Leskens and S. Belfroid, "Real-Time Estimation of Well Inflow Parameters Using Only Pressure and Temperature Measurements," in *26th Benelux Meeting on Systems and Control*, 2007, p. 30.

[11] W. Gilbert, "Flowing and gas-lift well performance," in *Drilling and production practice*, 1954: American Petroleum Institute.

[12] H. Mukherjee and J. Brill, "Pressure drop correlations for inclined two-phase flow," *Journal of energy resources technology,* vol. 107, no. 4, pp. 549-554, 1985.

[13] G. Evensen, "Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics," *Journal of Geophysical Research: Oceans,* http://doi.org/10.1029/94JC00572 vol. 99, no. C5, pp. 10143-10162, 1994.

[14] A. N. Gryzlov, "Model-based estimation of multi-phase flows in horizontal wells," TU Delft, Delft University of Technology, 2011.

[15] A. Doucet, N. De Freitas, and N. Gordon, "Sequential monte carlo methods in practice," ed: Springer-Verlag, 2001.

[16] P. J. Van Leeuwen, "Particle filtering in geophysical systems," *Monthly Weather Review,* vol. 137, no. 12, pp. 4089-4114, 2009.

[17] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Transactions on signal processing,* vol. 50, no. 2, pp. 174-188, 2002.

[18] J. S. Liu and R. Chen, "Sequential Monte Carlo methods for dynamic systems," *Journal of the American statistical association,* vol. 93, no. 443, pp. 1032-1044, 1998.
[19] M. K. Pitt and N. Shephard, "Filtering via simulation: Auxiliary particle filters," *Journal of the American statistical association,* vol. 94, no. 446, pp. 590-599, 1999.
[20] G. Evensen, "Data assimilation: the ensemble Kalman filter," ed: Springer Science & Business Media, 2009.
[21] S. I. Aanonsen, G. Nævdal, D. S. Oliver, A. C. Reynolds, B. Vallès, and others, "The ensemble Kalman filter in reservoir engineering--a review," *Spe Journal,* vol. 14, no. 03, pp. 393-412, 2009.
[22] R. Faragher, "Understanding the basis of the kalman filter via a simple and intuitive derivation [lecture notes]," *IEEE Signal processing magazine,* vol. 29, no. 5, pp. 128-132, 2012.
[23] A. A. Emerick and A. C. Reynolds, "History matching time-lapse seismic data using the ensemble Kalman filter with multiple data assimilations," *Computational Geosciences,* vol. 16, no. 3, pp. 639-659, 2012.
[24] J. P. Brill and H. Mukherjee, *Multiphase flow in wells*. Henry L. Doherty Memorial Fund of AIME, Society of Petroleum Engineers Richardson, TX, 1999.
[25] B. Guo, X. Liu, and X. Tan, "Chapter 5 - Choke Performance," in *Petroleum Production Engineering (Second Edition)*Boston: Gulf Professional Publishing, 2017, pp. 111-128.
[26] P. Baxendell, "Bean performance-lake wells," *Shell Internal Rep,* 1957.
[27] N. C. Ros, "An analysis of critical simultaneous gas/liquid flow through a restriction and its application to flowmetering," *Applied Scientific Research,* vol. 9, no. 1, pp. 374-388, 1960.
[28] I. Achong, "Revised bean performance formula for Lake Maracaibo wells," *Internal Company Report,* 1961.
[29] M. Safar Beiranvand and M. Babaei Khorzoughi, "Introducing a new correlation for multiphase flow through surface chokes with newly incorporated parameters," *SPE Production & Operations,* vol. 27, no. 04, pp. 422-428, 2012.
[30] M. K. Alrumah and M. S. Bizanti, "New Multiphase Choke Correlations for Kuwait Fields," 2007/1/1/, 2007.
[31] A. Elgibaly and I. Nashawi, "New correlations for critical and subcritical two-phase flow through wellhead chokes," *Journal of Canadian Petroleum Technology,* vol. 37, no. 06, 1998.
[32] D. J. Sadiq, "Predication of Oil Flow Rate through Choke at Critical Flow for Iraqi Oil Wells," *Journal of Petroleum Research and Studies,* vol. 212, no. 6, pp. 53-79, 2012.
[33] M. Safar Beiranvand, P. Mohammadmoradi, B. Aminshahidy, B. Fazelabdolabadi, and S. Aghahoseini, "New multiphase choke correlations for a high flow rate Iranian oil field," *Mechanical Sciences,* vol. 3, no. 1, pp. 43-47, 2012.
[34] M. B. Khorzoughi, M. S. Beiranvand, and M. R. Rasaei, "Investigation of a new multiphase flow choke correlation by linear and non-linear optimization methods and Monte Carlo sampling," *Journal of Petroleum Exploration and Production Technology,* vol. 3, no. 4, pp. 279-285, 2013.
[35] H. R. Nasriani, M. MoradiDowlatAbad, and A. Kalantariasl, "A New Correlation for Prediction of Critical Two-phase Flow through Wellhead Chokes," in *78th EAGE Conference and Exhibition 2016*, 2016.
[36] T. K. Perkins, "Critical and subcritical flow of multiphase mixtures through chokes," *SPE Drilling & Completion,* vol. 8, no. 04, pp. 271-276, 1993.

[37] R. Sachdeva, Z. Schmidt, J. Brill, and R. Blais, "Two-phase flow through chokes," in *SPE Annual Technical Conference and Exhibition*, 1986: Society of Petroleum Engineers.
[38] E. M. Alsafran and M. G. Kelkar, "Predictions of two-phase critical-flow boundary and mass-flow rate across chokes," *SPE Production & Operations,* vol. 24, no. 02, pp. 249-256, 2009.
[39] B. Guo, X. Liu, and X. Tan, "Chapter 4 - Wellbore Flow Performance," in *Petroleum Production Engineering (Second Edition)*Boston: Gulf Professional Publishing, 2017, pp. 83-109.
[40] M. J. Carrizales, J. Jaramillo, and D. Fuentes, "Prediction of multiphase flow in pipelines: literature review," *Ingeniería y Ciencia,* vol. 11, no. 22, pp. 213-233, 2015.
[41] A. Nikoofard, U. J. F. Aarsnes, T. A. Johansen, and G.-O. Kaasa, "State and Parameter Estimation of a Drift-Flux Model for Underbalanced Drilling Operations," *IEEE Transactions on Control Systems Technology,* 2017.
[42] R. J. Lorentzen, K. K. Fjelde, J. Frøyen, A. C. V. M. Lage, G. Nævdal, and E. H. Vefring, "Underbalanced and Low-head Drilling Operations: Real Time Interpretation of Measured Data and Operational Support," 2001/1/1/, 2001.
[43] A. Gryzlov, W. Schiferli, and R. Mudde, "Soft-sensors: Model-based estimation of inflow in horizontal wells using the extended Kalman filter," *Flow Measurement and Instrumentation,* vol. 34, pp. 91-104, 2013.
[44] X. Luo, R. J. Lorentzen, A. S. Stordal, and G. Nævdal, "Toward an enhanced Bayesian estimation framework for multiphase flow soft-sensing," *Inverse Problems,* vol. 30, no. 11, p. 114012, 2014.
[45] R. J. Lorentzen, A. Stordal, G. Nævdal, H. A. Karlsen, and H. J. Skaug, "Estimation of Production Rates With Transient Well-Flow Modeling and the Auxiliary Particle Filter," 2014.
[46] H. Shi *et al.*, "Drift-flux modeling of two-phase flow in wellbores," *Spe Journal,* vol. 10, no. 01, pp. 24-33, 2005.
[47] M. B. Standing, "A Pressure-Volume-Temperature Correlation For Mixtures Of California Oils And Gases," 1947.
[48] O. Glaso, "Generalized Pressure-Volume-Temperature Correlations," 1980.
[49] L. Dake, "Fundamentals of Reservoir Engineering, No. 8. Amsterdam: Developments in Petroleum Science," ed: Elsevier Science BV, 1978.
[50] A. M. Ansari, N. D. Sylvester, O. Shoham, and J. P. Brill, "A Comprehensive Mechanistic Model for Upward Two-Phase Flow in Wellbores," 1990.