

# Unscented Importance Sampling for Parameter Calibration of Carbon Sequestration Systems

---

Mirhamed Sarkarfarshi\*, Robert Gracie\*\*+

\* PhD Candidate, Department of Civil and Environmental Engineering, University of Waterloo  
200 University Avenue West, Waterloo, Ontario, Canada N2L 3G1

\*\* Assistant Professor, Department of Civil and Environmental Engineering, University of Waterloo  
200 University Avenue West, Waterloo, Ontario, Canada N2L 3G1

+ Corresponding author: [rgracie@uwaterloo.ca](mailto:rgracie@uwaterloo.ca) 519-888-4567 x 38827

## Abstract

A Bayesian importance sampling method is developed to efficiently and accurately calibrate the parameters of non-linear and non-Gaussian system models. The UIS consists of two stages. The first stage uses the latest monitoring data to generate a Gaussian approximation of the true posterior distribution of the uncertain parameters and utilizes the measurement update stage of the Unscented Kalman Filter (UKF) to approximate the posterior. The second stage of UIS uses a mixture of approximate posterior computed in the first stage and a heavy tailed distribution as the proposal distribution for Bayesian importance sampling. UIS is repeated whenever new monitoring data becomes available. Two case studies were developed to study the UIS method and to compare it UKF and Importance Sampling (IS) methods: a non-linear analytical system model and synthesized CO<sub>2</sub> injection model using a numerical multi-phase flow simulator. In analytical case study, it is shown that UIS is more accurate than both UKF and traditional IS with static proposal and the relative accuracy of the UIS over traditional IS increases with dimensionality of the parameter space. The higher accuracy of UIS compared to UKF and traditional IS with static proposal is also shown in the CO<sub>2</sub> injection case study. It is also shown that increasing number of samples and a defensive mixture distribution with a mixture ratio between 0.1 and 0.25 enhances the performance of UIS.

Keywords: *CO<sub>2</sub> sequestration, Uncertainty Mitigation, Bayesian Update, Importance Sampling*

# 1. Introduction

Risk assessment and management are inevitable components of any CO<sub>2</sub> Sequestration (CS) project (Kopp, et al., 2010)(Walton, et al., 2004). One challenge often faced when performing risk assessment for CS and other engineered systems is that system models tend to be extremely computationally expensive and involve many uncertain parameters. The behavior of a CS system will be uncertain at the start of injection and so CS sites will be continuously monitored (Lawton, et al., 2010)(Lawton, 2010). Methods for using monitoring data to reduce system model uncertainty over the lifetime of a project are therefore desirable. The primary contribution of this article is development of a new Bayesian calibration method, Unscented Importance Sampling (UIS), which continuously, accurately, and efficiently calibrates the uncertain and non-Gaussian parameter distributions of complex and non-linear system models using continuously streaming monitoring data. Accuracy, computational efficiency and compatibility with non-linear system models and non-Gaussian probability distributions are the key features of the UIS. The UIS is described in terms of a general system model, and is applied to a non-linear analytical system model and a multiphase flow system model of CO<sub>2</sub> injection into a heterogeneous aquifer.

There are various sources of uncertainty in CS systems - e.g., the uncertain geological properties of the injection site (Alshuhail, et al., 2009)(Lavoie & Keith, 2010)(Hongjun, et al., 2010), modelling strategies (Nordbotten, et al., 2012), uncertain monitoring data (Sato, 2011) and uncertain consequences of unwanted incidents (Li & Fall, 2013) - negatively impact the quality of risk assessment (Sarkarfarshi, et al., 2014). Consequently, the effectiveness of risk management and decision making is impaired. In CS, parameter uncertainty is believed to be one of the most significant sources of uncertainty among the other sources. In CS projects, data from core samples is limited and sparse, because drilling wells for obtaining core samples penetrate the caprock of the target formation and creates potential leakage pathways for supercritical CO<sub>2</sub>. Remote measurements - e.g. by seismic surveys - of physical parameters of the subsurface are subject to significant uncertainty. Consequently, knowledge of system parameters is likely to be sparse and uncertain at the beginning of a project and expert opinion is likely to play a significant role in characterizing initial parameter distributions.

Risk and uncertainty assessment in a CS context is considered in several works; however, few have had a quantitative basis. Notable quantitative risk and uncertainty analysis studies include risk assessment methods for site selection (Kopp, et al., 2010), significance of consequences of failure (Price & Oldenburg, 2009) and stochastic plume migration and leakage forecast (Walton, et al.,

2004) (Kopp, et al., 2010)(Oladyshkin, et al., 2011). The emphasis for these earlier assessments was pre-injection and they may be suboptimal for risk assessment and management over the lifetime of a CS project since monitoring data is not incorporated directly to calibrate parameter distributions and to make model forecasts.

Data acquired from site monitoring provides valuable information about the current behaviour of a CS system. Several existing monitoring techniques are applicable to CS such as 3D and 4D seismic (Chadwick, et al., 2004), gravity (Nooner, et al., 2007), electromagnetic (Hoversten, et al., 2004), time lapse satellite imaging (INSAR) (Raikes, et al., 2008) and fluid sampling (Freifeld, et al., 2005). To date there has been a limited number of works focusing on how such monitoring data can be incorporated into the calibration of CS models. (Johnson & White, 2012) used Markov Chain Monte Carlo (MCMC) for inverting seismic signals to refine the permeability fields; (Bhowmik, et al., 2011) stochastically calibrated the permeability field in a synthesized CS system using pressure data from injection and monitoring wells; Finally, (Espinet & Shoemaker, 2013) and (Tavakoli, et al., 2013) compared several deterministic and stochastic calibration algorithms, respectively, using synthesized case studies in CS systems. Due to limitations of the abovementioned works, e.g. being either deterministic, computationally demanding or using assumptions such as linearity of the models or Gaussian structure of the probability distributions, more research is required to study, adopt and developed efficient methods for calibration of CS systems which are accurate, computationally efficient and able to calibrate.

For purpose of parameter uncertainty mitigation of CS models, UIS is developed and presented in this manuscript. The UIS methodology presented in this manuscript contributes to computationally efficient and stochastic calibration of non-linear models such as those used in CS systems without assuming linear models and Gaussian probability distributions. UIS is a stochastic calibration algorithm, which incorporates both expert opinion and noisy time-lapse monitoring data. In UIS, the proposal distribution of Bayesian importance sampling method (Marshall, 1956) is continuously updated utilizing the latest monitoring data to improve sample quality and mitigate the demand for larger sample size. The proposal density is constructed by mixing a computationally efficient version of Kalman Filters and a heavy-tailed *defensive distribution* (Hesterberg, 1995). The defensive distribution has two roles. The first role is to ensure the robustness of the importance sampling step, and the second role is to allow the algorithm seek parameters that are not assigned a high probability initially but might gain large likelihood and improve the sample quality. Mixture ratio between Kalman filter output and the defensive distribution is a tuning parameter of UIS. While a version of the Kalman Filter is used for constructing the proposal density, actual probability

distributions of the parameters are not required to be Gaussian. As a result, UIS is more versatile for parameter estimation than most Kalman Filter-based methods.

In Section 2, Bayesian parameter calibration and UIS are described mathematically. In Section 3, application of UIS is presented in a simple analytical case study. In Section 4, UIS is applied on a more sophisticated CS case study, the sensitivity of UIS to tuning parameters is investigated and results are discussed. Conclusions are presented in Section 5.

## 2. Bayesian parameter estimation

Bayesian frameworks allow updating prior parameter data and expert opinion using site monitoring data. In this section, several solutions to Bayes formula - e.g., the measurement update stage of Kalman Filter (KF), Extended Kalman Filter (EKF), Unscented Kalman Filter (UKF) and Importance Sampling (IS) – are presented to lay a theoretical foundation for the discussion of the UIS. Next, the importance of the proposal distribution in IS is illustrated, and finally, the UIS algorithm is presented, including underlying formulation for recursively obtaining better proposal distributions for IS incorporating all available monitoring data.

### 2.1 Bayes Rule

Assume that  $\theta$  is the model parameter vector. For simplicity, assume that the parameter space,  $\Omega_\theta$ , is continuous<sup>1</sup>. Let  $\pi(\theta|D)$  denote the probability distribution of  $\theta$  conditioned by observations  $D$ . According to the Bayes rule,  $\pi(\theta|D)$  is proportional to the product of the unconditioned probability distribution of  $\theta$ ,  $\pi(\theta)$ , and the likelihood of  $D$  given  $\theta$ ,  $L(D|\theta)$ , i.e.

$$\pi(\theta|D) = c^{-1} \cdot L(D|\theta)\pi(\theta) \quad (1)$$

where  $c$  is a normalizing constant and is given by

$$c = \int_{\Omega_\theta} L(D|\theta)\pi(\theta)d\theta \quad (2)$$

and where  $\pi(\theta)$  is the *prior* distribution, which represents the prior knowledge of  $\theta$  while the observations  $D$  are not accounted for (Rougier, 2008). The prior is a subjective judgement of reality and can be assigned based on initial measurements, data from similar sites, expert opinion,

---

<sup>1</sup> For a discrete parameter space, all integrals will be replaced by summation.

parameter bounds, etc (Liu, et al., 2010). The likelihood, unlike probability which predicts occurrence of an unknown event based on known parameters, allows us to predict unknown parameters using known observations (Hall, et al., 2012). The *posterior* distribution,  $\pi(\theta|D)$ , represents the probability distribution of  $\theta$  accounting for the observations and can be interpreted as the calibrated version of the prior given  $D$ .

Let  $Y^t = [Y(x_i, T_t)|i = 1 \dots m_t]$  denote the true behaviour of the CS system (e.g. the CO<sub>2</sub> saturation distribution in a saline aquifer) at time  $T_t$  and at locations  $x_1$  to  $x_{m_t}$  where  $m_t$  is number of measurements at time  $T_t$ . The  $t$  index in  $m_t$  indicates that the number of measurements can be different at each time step (or calibration cycle). Let  $Z^t$  denote measurements of  $Y^t$  at time  $T_t$  and at locations  $x_1$  to  $x_{m_t}$  with a random measurement noise,  $e$ , as shown in equation (3),

$$Z^t = Y^t + e \quad e \sim g_e(e) \quad (3)$$

where  $g_e(e)$  denotes the distribution of the measurement error. It is fairly standard and realistic to assume  $e$  is not correlated with  $Y^t$  nor time (Rougier, 2008).

Let  $f^t(\theta) = [f(x_i, T_t, \omega, \theta)|i = 1 \dots m_t]$  denote a deterministic system model of the system behaviour at time  $T_t$  and at locations  $x_1$  to  $x_{m_t}$ . The system model is parameterized by a set of deterministic parameters  $\omega$  and a set of uncertain parameters  $\theta = [\theta_i|i = 1 \dots n]$ , where  $\theta_i$  denotes uncertain model parameter  $i$ . Let  $\pi(\theta)$  denote an  $n$ -dimensional probability distribution that reflects the uncertainty in  $\theta$ .

The difference between the model and true system behaviour,  $\epsilon = Y^t - f^t$ , is called the *model discrepancy* (Rougier, 2008). It is assumed that the model discrepancy is also a random variable with distribution,  $g_\epsilon(\epsilon)$ , and that it is independent from  $\theta$  and time (Kaipio & Somersalo, 2007). Thus, we can re-write equation (3) as follows:

$$Z^t = f^t(\theta) + \epsilon + e \quad e \sim g_e(e) \quad \epsilon \sim g_\epsilon(\epsilon) \quad (4)$$

Let  $\varepsilon = \epsilon + e$  and let  $g_\varepsilon(\varepsilon)$  denote the distribution of  $\varepsilon$ . In fact,  $\varepsilon$  reflects the uncertainty of model and measurement when the ideal  $\theta$  is used (Li, et al., 2004). The likelihood of the measurement  $Z^t$  given  $\theta$  can be then calculated as

$$L(Z^t|\theta) = g_\varepsilon(\varepsilon) = g_\varepsilon(Z^t - f^t(\theta)) \quad (5)$$

Unlike  $g_e(e)$ ,  $g_\epsilon(\epsilon)$  is less likely to be known *a priori*. The modeling discrepancy,  $\epsilon$ , can be ignored only if it is negligible compared to measurement noise; otherwise, it should be prescribed. When the modeling discrepancy is not negligible,  $g_\epsilon$  can be estimated by a least squares method (Stedinger, et al., 2008) or alternatively, the hyper-parameters of  $g_\epsilon$  can be included in the uncertain parameter vector  $\theta$ .

Combining equations (5) and (1), the posterior distribution of  $\pi(\theta)$  can be obtained as

$$\pi(\theta|Z^t) = \frac{g_\epsilon(Z^t - f^t(\theta))\pi(\theta)}{\int_{\Omega_\theta} g_\epsilon(Z^t - f^t(\theta))\pi(\theta)d\theta} \quad (6)$$

The posterior distribution in equation (6) at time  $T_t$  becomes the prior distribution for time  $T_{t+1}$ . For simplicity of notation, we denote the initial prior distribution by  $\pi^0(\theta)$ , the posterior distribution at time  $T_1$  by  $\pi^1(\theta) = \pi(\theta|Z^1)$ , the posterior distribution at time  $T_2$  by  $\pi^2(\theta)$ , so on and so forth.

## 2.2 Kalman Filter

Analytical solution to (6) is available only if  $\pi(\theta)$  and  $g_\epsilon(Z^t - f^t(\theta))$  follow specific pairs of distribution known as *conjugate families*, which are reported widely in statistics literature. Due to the nonlinearity of  $f^t(\theta)$  in the CS models it is unlikely that  $\pi(\theta)$  and  $g_\epsilon(Z^t - f^t(\theta))$  belong to a conjugate family even if  $\pi(\theta)$  and  $g_\epsilon(\theta)$  belong to a conjugate family.

When both  $\pi(\theta)$  and  $g_\epsilon(\epsilon)$  are Gaussian and  $f$  is linear, i.e.  $f(\theta) = H\theta$  and  $g_\epsilon = N(0, \Sigma_\epsilon)$ , the posterior distribution in (6) is also Gaussian and the mean ( $\mu^a$ ) and the covariance ( $\Sigma^a$ ) of the posterior distribution are

$$\begin{aligned} \mu^a &= \mu^b + K(Z^t - f^t(\mu^b)) \\ \Sigma^a &= \Sigma^b - K.(\Sigma(f^t) + \Sigma_\epsilon).K^T \end{aligned} \quad (7)$$

where  $\mu^b$  and  $\Sigma^b$  are the prior mean and covariance,  $\Sigma(\theta, f^t)$  is the prior covariance between  $\theta$  and  $f^t(\theta)$  (calculated as  $\Sigma^b H^T$ ),  $\Sigma(f^t)$  is the prior covariance of  $f^t$  (calculated as  $H\Sigma^b H^T$ ), and  $K$  is the optimal Kalman gain calculated as

$$K = \Sigma(\theta, f^t).(\Sigma(f^t) + \Sigma_\epsilon)^{-1} \quad (8)$$

Equation (7) and (8) are the measurement update stages of a Kalman filter (Kalman, 1960). While KFs are usually used for estimation of state of dynamic systems, measurement update stage of KF can also be used for parameter estimation, as described in (Wan & Merwe, 2001).

If  $f$  is non-linear, it can be linearized about  $\mu^b$  using a first order Taylor series expansion, as done in Extended Kalman Filtering (EKF) formulation. An alternative solution is linearizing  $f(\theta)$  in a process so called an Unscented Transformation (UT)(Terejanu, n.d.), by linear regression between system model outputs, given  $2n + 1$  deterministically chosen parameter sets. UT is basically a method for obtaining the statistics of a non-linear function (system model) of a random variable using the statistics of a finite number of points (Wan & Merwe, 2001) and it is the basis of the Unscented Kalman Filter (UKF) (Julier & Uhlmann, 1996). Unlike the EKF which is first order accurate, UKF is at least second order accurate for non-Gaussian probability densities and third order accurate for Gaussian probability densities (Chen, 2003). Another clear advantage of the UKF over the EKF is that it can be used with any non-linear function (system model) and does not require the calculation of the Jacobian or Hessian of the function (Chen, 2003), which is non-trivial when the system model is a numerical.

Now we describe a simple version of UT for obtaining statistics (mean, covariance and cross covariance) of  $f(\theta)$  when  $\theta \sim \pi(\theta)$  and the mean and covariance of  $\pi(\theta)$  are  $\mu(\theta)$  and  $\Sigma(\theta)$ , respectively. For a parameter space of size  $n$ ,  $2n + 1$  weighted points are deterministically chosen so that their mean is equal to  $\mu(\theta)$  and their covariance is equal to  $\Sigma(\theta)$ . These points are known as *sigma points* and are calculated as

$$\begin{aligned}
 \chi_0 &= \mu(\theta) \\
 -1 &< \omega_0 < 1 \\
 \chi_i &= \mu(\theta) + \left( \sqrt{\frac{n}{1-\omega_0} \Sigma(\theta)} \right)_i \quad i = 1, 2, \dots, n \\
 \chi_i &= \mu(\theta) - \left( \sqrt{\frac{n}{1-\omega_0} \Sigma(\theta)} \right)_i \quad i = n + 1, n + 2, \dots, 2n \\
 \omega_i &= \frac{1-\omega_0}{2n} \quad i = 1, 2, \dots, 2n
 \end{aligned} \tag{9}$$

where  $\left(\sqrt{\frac{n}{1-\omega_0}}\Sigma(\theta)\right)_i$  is the  $i^{\text{th}}$  row of the square root of  $\frac{n}{1-\omega_0}\Sigma(\theta)$ , obtained by Cholesky decomposition and  $\omega_i$  are the weights.  $\omega_0$  is the weight of the first Sigma point and determines how scattered the sigma points are about  $\chi_0$ . The sigma points are then propagated through the non-linear function  $f$ ,

$$Y_i = f(\chi_i) \quad i = 0, 1, \dots, 2n \quad (10)$$

The mean, covariance and cross covariance of the sigma points,  $\chi_i$ , and the propagated sigma points,  $Y_i$ , are calculated as

$$\begin{aligned} \mu(\chi) &= \sum_{i=0}^{2n} \omega_i \chi_i \\ \mu(Y) &= \sum_{i=0}^{2n} \omega_i Y_i \\ \Sigma(Y) &= \sum_{i=0}^{2n} \omega_i [Y_i - \mu(Y)][Y_i - \mu(Y)]^T \\ \Sigma(\chi, Y) &= \sum_{i=0}^{2n} \omega_i [\chi_i - \mu(\chi)][Y_i - \mu(Y)]^T \end{aligned} \quad (11)$$

Substituting (11) into (7), the posterior mean and covariance of  $\theta$  is approximated as

$$\begin{aligned} \mu^a &= \mu(\theta) + K \left( Z^t - f^t(\mu^b) \right) \\ \Sigma^a &= \Sigma(\theta) - K \cdot \left( \sum_{i=0}^{2n} \omega_i [Y_i - \mu(Y)][Y_i - \mu(Y)]^T + \Sigma_\varepsilon \right) \cdot K^T \end{aligned} \quad (12)$$

where the Kalman gain is obtained by

$$K = \left( \sum_{i=0}^{2n} \omega_i [\chi_i - \mu(\chi)][Y_i - \mu(Y)]^T \right) \cdot \left( \sum_{i=0}^{2n} \omega_i [Y_i - \mu(Y)][Y_i - \mu(Y)]^T + \Sigma_\varepsilon \right)^{-1} \quad (13)$$



Equations (12) and (13) are the measurement update stage of an UKF. For simplicity of notation, we denote the posterior Gaussian distribution  $N(\mu^a, \Sigma^a)$  obtained from (12) and (13) by  $UT(\pi(\theta), Z^t)$ .

An Alternative KF for non-linear systems is the Monte-Carlo approach to KF, known as Ensemble Kalman Filter (EnKF) (Evensen, 1994). EnKF has found application when number of variables are very large, e.g. in weather forecast (Houtekamer, et al., 2005), reservoir engineering (Aanonsen & Reynolds, 2009)(Naevdal, et al., 2005) and hydrology (Reichle, et al., 2002)(Shu, et al., 2005). In a recent study, (Tavakoli, et al., 2013) demonstrated how EnKF and other ensemble-based algorithms such as Ensemble Smoothers (ES) (van Leeuwen & Evensen, 1996) can be applied on CO<sub>2</sub> sequestration. Similar to UKF, EnKF does not require calculating the Jacobian of the models. For very high dimensional state vectors, EnKF is computationally more efficient than UKF (Kim, 2011). However, EnKF could be less accurate than UKF (Ambadan & Tang, 2009)(Mesbah, et al., 2011).

### 2.3 Importance Sampling

The KF-based solutions are not exact and can be inadequate if  $f$  is highly non-linear or the probability distributions are non-Gaussian (Leisenring & Moradkhani, 2011). A common solution to this issue is drawing random samples from the posterior distribution and describing its properties (e.g. moments and quintiles) with the properties of the samples (Rougier, 2008). Since the posterior distribution in (6) is usually unavailable, indirect sampling approaches such as IS or Markov Chain Monte Carlo (MCMC) can be used. In IS, we sample from a proposal distribution which is close to the posterior and is easy to sample from and weight the samples accordingly (Marshall, 1956). The proposal distribution encourages samples from areas of more importance, and the weighting ensures that the importance sampling estimator is unbiased. Equation (14) demonstrates the general formulation for IS,

$$p(\theta) \approx \sum_{i=1}^N W_i \delta(\theta - \bar{\theta}_i) \quad \bar{\theta}_i \sim q(\theta) \quad (14)$$

where  $p(\theta)$  is the probability distribution we want to sample from,  $q(\theta)$  is the proposal distribution,  $\bar{\theta}_i$  are the samples.  $W_i$  are normalized *importance weights*, evaluated as

$$W_i = \frac{w_i}{\sum_{i=1}^N w_i} \quad i = 1 \dots N \quad (15)$$

where  $w_i = p(\bar{\theta}_i)/q(\bar{\theta}_i)$ . The statistics of  $p(\theta)$  can then be described using the statistics of the samples  $\{\bar{\theta}_{i=1..N}\}$  and the importance weights  $\{W_{i=1..N}\}$ . For instance, the mean and covariance of  $p(\theta)$  can be approximated by

$$\mu(p(\theta)) \approx \sum_{i=1}^N W_i \bar{\theta}_i \quad (16)$$

$$\Sigma(p(\theta)) \approx \sum_{i=1}^N W_i (\bar{\theta}_i - \mu(p(\theta))) (\bar{\theta}_i - \mu(p(\theta)))^T$$

respectively. Next, we want to sample from the posterior distribution in (6). After sampling  $\{\bar{\theta}_i\}$  from  $q(\theta)$ , the un-normalized weights are calculated as

$$w_i = c^{-1} \frac{g_\varepsilon(Z^t - f^t(\bar{\theta}_i)) \pi^0(\bar{\theta}_i)}{q(\bar{\theta}_i)} \quad i = 1 \dots N \quad (17)$$

and the normalized weights are calculated as

$$W_i = \frac{c^{-1} \frac{g_\varepsilon(Z^t - f^t(\bar{\theta}_i)) \pi^0(\bar{\theta}_i)}{q(\bar{\theta}_i)}}{c^{-1} \sum_{j=1}^N \frac{g_\varepsilon(Z^t - f^t(\bar{\theta}_j)) \pi^0(\bar{\theta}_j)}{q(\bar{\theta}_j)}} \quad i = 1 \dots N \quad (18)$$

It can be seen that  $c^{-1}$  will be cancelled from numerator and denominator in (18). Thus, the normalizing constant does not to be computed to approximate the posterior distribution using IS.

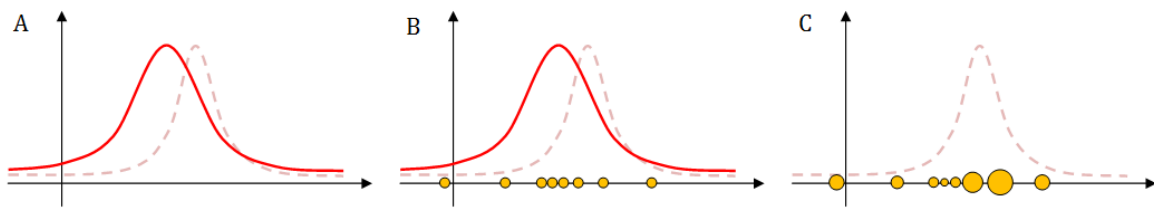
If monitoring data from times  $T_1$  to  $T_t$  are available, equation (18) is modified to

$$W_i = \frac{\frac{\pi^0(\bar{\theta}_i)}{q(\bar{\theta}_i)} \prod_{l=1}^t g_\varepsilon(Z^l - f^l(\bar{\theta}_i))}{\sum_{i=1}^N \frac{\pi^0(\bar{\theta}_i)}{q(\bar{\theta}_i)} \prod_{l=1}^t g_\varepsilon(Z^l - f^l(\bar{\theta}_i))} \quad i = 1 \dots N \quad (19)$$

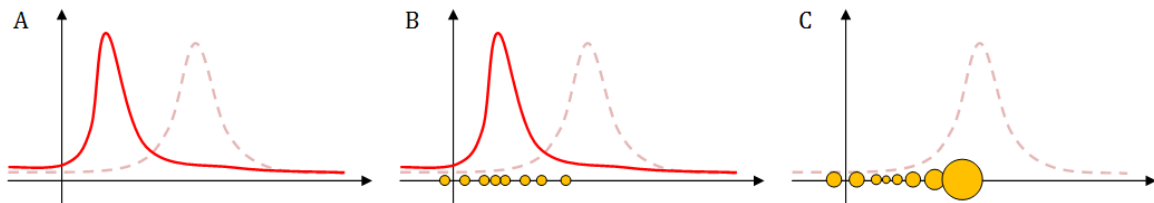
## 2.4 Unscented Importance Sampling

Proper choice of the proposal distribution is the key to an efficient importance sampler. If the proposal distribution is not close to the true posterior, most samples will be drawn from unimportant areas which results in decreased accuracy and computational efficiency. The

significance of a proper choice of the proposal distribution is illustrated in Fig. 1 and Fig. 2. Fig. 1 illustrates sampling from a proposal distribution which is close to the actual symmetric sampling distribution and Fig. 2 shows a poor choice for a proposal distribution. It can be seen that samples in Fig. 1 are better distributed and the weights are closer in magnitude compared to those illustrated in Fig. 2. In Fig. 2C however, all samples are located on the left side of mean of the actual distribution, making the sample biased. Consequently, more samples are required in case illustrated in Fig. 2 compared to that shown Fig. 1 in order to ensure that samples are well distributed throughout the important areas of the actual distribution. The optimal choice of the proposal distribution for sampling from  $\pi^t(\theta)$  is  $q(\theta) \propto \pi^t(\theta)$  (Smith, et al., 1997) which is not practical because if  $\pi^t$  is computable then there is no need for importance sampling (Owen & Zhou, 1998)(Merwe, et al., 2000). Our effort is therefore focused on finding a proposal distribution close to  $\pi^t(\theta)$ , to enhance the accuracy and computational efficiency of the sampler.



**Fig. 1** Importance sampling using a proposal distribution (solid line) close to the actual sampling distribution (dashed line). Samples in (B) are drawn from the proposal distribution and are weighted accordingly in (C). Size of the samples indicates sample weights.



**Fig. 2** Importance sampling using a poor proposal distribution (solid line) which is not close to the actual sampling distribution (dashed line). Samples in (B) are drawn from the proposal distribution and are weighted accordingly in (C). Size of the samples indicates sample weights. Since the sample is clearly biased, more samples are required which negatively affects the computational efficiency of the sampler.

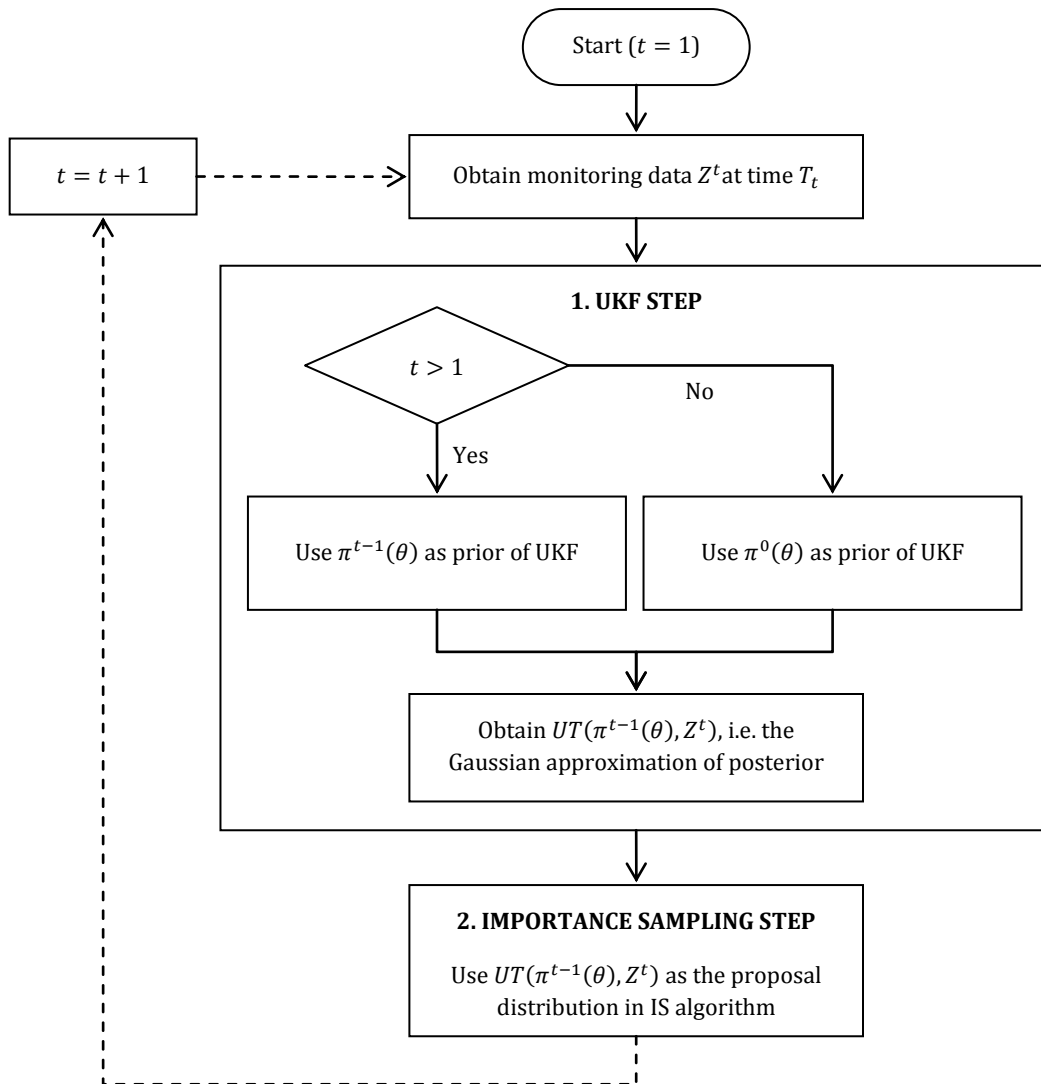
An intuitive and common choice of the proposal distribution is the prior distribution itself (Merwe, et al., 2000), which reduces (18) to the likelihood function and (19) to product of the likelihoods. This “sample from the prior, weight by the likelihood” (Rouger, 2008) approach, however, does not contain the latest information from measurements.

Revising the proposal distribution periodically using the information from measurements can improve the quality of the samples. For instance  $\pi^{t-1}(\theta)$  can be used as the proposal distribution at time  $T_t$ , since it is likely to be closer to  $\pi^t(\theta)$  than the initial proposal. However,  $\pi^{t-1}(\theta)$  does not contain the latest monitoring information from  $Z^t$ . Another solution is iteratively improving the proposal distribution as in Population Monte Carlo (PMC) methods (Cappé, et al., 2004). In the PMC, each iteration involves repeating the importance sampling with a proposal distribution generated by a weighted sum of proposals of the previous iterations, while the proposal weights are adaptively adjusted. PMC and methods developed later on the same basis such as D-kernel PMC (Douc, et al., 2007) and M-PMC (Cappé, et al., 2008) showed to be promising in enhancing the performance of IS. However, the iterative nature of these algorithms requires running the system model several times more compared to the original IS which makes them computationally much more demanding than the original IS and consequently, less favourable for CS applications.

Alternatively, an approximation of  $\pi^t$  which is obtained from a computationally faster solution can be used as the proposal distribution. We propose using the measurement update stage of the UKF for this purpose and call this approach *Unscented Importance Sampling (UIS)*. A flowchart of UIS is illustrated in Fig. 3. As shown in the flowchart, each update cycle in UIS includes two major steps:

1. UKF step: the posterior distribution of the IS step from the previous cycle,  $\pi^{t-1}(\theta)$ , is utilized as the prior of the UKF measurement update at time  $T_t$ . For the initial cycle, the initial prior,  $\pi^0(\theta)$ , is used instead. Next, the Gaussian posterior,  $UT(\pi^{t-1}(\theta), Z^t)$ , is obtained using the measurement update stage of the UKF, as described by (12) and (13). Thus, the UKF step uses  $Z^t$  while previous monitoring data are accounted for in the prior.
2. IS step:  $UT(\pi^{t-1}(\theta), Z^t)$  from UKF step is used as the proposal distribution for the Bayesian importance sampling as described by (19). In this step, the initial prior is used as the prior in Bayes theorem and all monitoring data ( $Z^1$  to  $Z^t$ ) are accounted for in the weights of the samples.

When new monitoring data becomes available, the algorithm will be repeated. The stopping criteria are flexible and application-specific. For instance, criteria such as reaching a certain number of calibration cycles or reaching certain parameter variance can be used.



**Fig. 3** Flowchart of Unscented Importance Sampling (UIS)

By substituting proper equations into each block of Fig. 3, the complete algorithm of UIS is obtained and is summarized in Algorithm 1.

Algorithm 1: Algorithm of unscented importance sampling

1. Set time  $t = t + 1$  where  $t$  is monitoring time step (start with  $t = 0$ )
2. Obtain monitoring data  $Z^t$ .
- 3. UKF step**
  - 3.1. Set the prior of UKF step  $\pi_b^t(\theta) = N(\mu_b^t, \Sigma_b^t)$  where  $\mu_b^t = \sum_{i=1}^N W_i^{t-1} \bar{\theta}_i^{t-1}$  and  $\Sigma_b^t = \sum_{i=1}^N W_i^{t-1} (\bar{\theta}_i^{t-1} - \mu_b^t) (\bar{\theta}_i^{t-1} - \mu_b^t)^T$  when  $t > 1$  and  $\mu_b^t = \mu(\pi^0(\theta))$  and  $\Sigma_b^t = \Sigma(\pi^0(\theta))$  when  $t = 1$ .
  - 3.2. Select Sigma points and their weights as follows.
$$\chi_0 = \mu_b^t \quad -1 < \omega_0 < 1 \quad \omega_i = \frac{1 - \omega}{2n} \quad i = 1, 2, \dots, 2n$$

$$\chi_i = \mu_b^t \pm \left( \sqrt{\frac{n}{1 - \omega_0} \Sigma_b^t} \right)_i \quad + (i = 1, 2, \dots, n) \quad - (i = n + 1, n + 2, \dots, 2n)$$
  - 3.3. Propagate sigma points through non-linear system model and obtain  $Y_i = f(\chi_i)$  where  $i = 0, 1, \dots, 2n$
  - 3.4. Calculate statistics of sigma points and propagated sigma points as follows.
$$\mu(\chi) = \sum_{i=0}^{2n} \omega_i \chi_i \quad \mu(Y) = \sum_{i=0}^{2n} \omega_i Y_i$$

$$\Sigma(Y) = \sum_{i=0}^{2n} \omega_i [Y_i - \mu(Y)][Y_i - \mu(Y)]^T \quad \Sigma(\chi, Y) = \sum_{i=0}^{2n} \omega_i [\chi_i - \mu(\chi)][Y_i - \mu(Y)]^T$$
  - 3.5. Calculate mean and covariance of the Gaussian posterior of UKS step as  $\mu^a = \mu_b^t + K(Z^t - f^t(\mu_b^t))$  and  $\Sigma^a = \Sigma_b^t - K \cdot \Sigma(\Sigma(Y) + \Sigma(g_\varepsilon)) \cdot K^T$  where  $K = \Sigma(\chi, Y) \cdot (\Sigma(Y) + \Sigma(g_\varepsilon))^{-1}$
  - 3.6. Set  $UT(\pi_b^t(\theta), Z^t) = N(\mu^a, \Sigma^a)$
- 4. IS step**
  - 4.1. Set the proposal distribution  $q^t(\theta) = UT(\pi_b^t(\theta), Z^t)$
  - 4.2. Sample  $\{\bar{\theta}^t\}$  from  $q^t(\theta)$
  - 4.3. Simulate the model for each sample from time zero to  $T_t$ .
  - 4.4. Calculate un-normalized importance weights by  $w_i^t = \pi^0(\bar{\theta}_i^t) / q^t(\bar{\theta}_i^t) \prod_{i=1}^t g_\varepsilon(Z^i - f^i(\bar{\theta}_i^t))$  and normalize them to obtain  $\{W^t\}$
  - 4.5. Approximate  $\pi^t(\theta)$  with  $\sum_{i=1}^N W_i^t \delta(\theta - \bar{\theta}_i^t)$

```
Repeat steps 1 to 4 for each calibration cycle until stopping criteria
are met.
```

This algorithm is inspired from the Unscented Particle Filtering (UPF) (Julier & Uhlmann, 1996) and ensures that the proposal distribution improves with time and contains information from all the monitoring data up to  $T^t$ . The final posterior distribution in the UIS is generated in the importance sampling step. Thus, unlike the KF family, no assumption of the model structure or distribution types is necessary. Moreover, the sampling based approach in both steps (UKF and IS) lets various model types (e.g., different numerical models) be incorporated in a straightforward way. The abovementioned characteristics make the UIS flexible for parameter calibration in various applications.

In the case that the true posterior is multimodal, the presented version of UIS is still applicable. However, the performance of the algorithm diminishes because the Gaussian proposal cannot accurately represent multi-modal distributions. For such cases, the Gaussian approximation of the posterior in the UKF step can be replaced by a Gaussian mixture model.

#### 2.4.1 Defensive Importance Sampling in UIS

In IS, the proposal distribution needs to have a heavier tail than the actual sampling distribution for the solution to be robust (Owen & Zhou, 1998). To ensure the tails of the proposal distribution  $q(\theta)$  are heavier than the sampling distribution, we can mix  $q(\theta)$  with a heavy-tailed distribution  $\hat{q}(\theta)$  such as a uniform distribution in a process called “defensive mixture sampling” (Hesterberg, 1995) or “defensive importance sampling” (Owen & Zhou, 1998),

$$q^*(\theta) = (1 - \eta)q(\theta) + \eta\hat{q}(\theta) \quad (20)$$

where  $\eta$  is the defensive mixture ratio and  $0 < \eta < 1$ . Moreover, the defensive mixture distribution searches the areas of the parameter space with very small probability density for potentially better parameter sets. A defensive mixture distribution can be added to UIS by modifying the sub-step 4.1 of Algorithm 1 as

$$q^*(\theta) = (1 - \eta)UT(\pi_b^t(\theta), Z^t) + \eta\hat{q}(\theta) \quad (21)$$

where  $\eta$  is the mixture ratio and is one of the tuning parameters. A very small  $\eta$  does not affect the sample population and a very large  $\eta$  is likely to waste numerous model runs by weighting samples

with negligible posterior weight. In the literature,  $0.1 < \eta < 0.5$  is recommended (Hesterberg, 1995)(Owen & Zhou, 1998).

### 3. Analytical Case Study

Before the CS case study, a simplified case study is presented to demonstrate the application of UIS and compare it with two other algorithms, the measurement update stage of UKF and ordinary IS. This case study uses a significantly nonlinear analytical function as the system model and enables us to investigate the performance of UIS for various numbers of uncertain parameters with a reasonable computational effort.

#### 3.1 Case Description

Assume a deterministic system described by

$$Y^t(x) = \sum_{i=1}^{M-1} a_i^2 \cos(a_{i+1})^2 \sin(x) + a_{i+1}^2 \cos(a_i)^2 \sin(t) \quad (22)$$

where  $M$  is the total number of  $a_i$ ,  $x$  is dimensionless location varying between 1 and 20,  $t$  is dimensionless time and  $a_i = 8$  for odd  $i$  and  $a_i = 10$  for even  $i$ .  $Y^t(x)$  is modeled by a scalar function denoted by  $f^t(\theta_1, \theta_2, x)$  with a similar structure,

$$f^t(\theta, x) = \sum_{i=1}^{M-1} \theta_i^2 \cos(\theta_{i+1})^2 \sin(x) + \theta_{i+1}^2 \cos(\theta_i)^2 \sin(t) \quad (23)$$

where  $\theta_i$  are the uncertain model parameters and  $\theta = [\theta_i]_{i=1\dots M}$  is the uncertain parameters vector. In four time steps,  $t = 1, 2, 3, 4$ ,  $Y^t(x)$  is measured at two locations,  $x = 7$  and  $x = 13$ . These measurements are denoted by  $Z^t(7)$  and  $Z^t(13)$  for time step  $t$ . Measurements are subject to a Gaussian noise denoted by  $e$ , i.e.,  $Z^t(x) = Y^t(x) + e$  where  $e \sim g_e(e)$  and  $g_e(e)$  is Gaussian with a mean of zero and a standard deviation of 100. Since the model structure matches the true system structure, model discrepancy is equal to zero. Thus, we can write  $Z^t(x) = f^t(\theta_1, \theta_2, x) + e$ .

The prior distribution of  $\theta$  is assumed to be Gaussian denoted by  $\pi^0(\theta)$  with diagonal elements of 12 and a covariance matrix with diagonal elements of 1, and off-diagonal elements of zero.



At each of the four time steps, new monitoring data is obtained. Our objective is to find the posterior distribution of  $\theta$  at each time step, denoted by  $\pi^t(\theta)$ , where  $t$  is 1 to 4, by using Algorithm 1. We call this process a *calibration cycle* for each of the time steps.

### 3.2 Calibration Scenarios and Comparison Metrics

We solved this problem using the measurement update stage of UKF, ordinary IS with 200 samples, and UIS with 200 samples and no defensive mixture distribution. Three sizes of the uncertain parameter vector are considered  $M = 10$ ,  $M = 20$  and  $M = 50$ . Table 1 summarizes all calibration scenarios considered in this case study.

Table 1: Calibration scenarios in the simplified case study

Scenario Name	Method	Number of parameters ( $M$ )	Sample count ( $N$ )	Simulations per cycle
M10-UKF	UKF	10	-	21
M10-IS	IS	10	200	200
M10-UIS	UIS	10	200	221
M20-UKF	UKF	20	-	41
M20-IS	IS	20	200	200
M20-UIS	UIS	20	200	241
M50-UKF	UKF	50	-	101
M50-IS	IS	50	200	200
M50-UIS	UIS	50	200	301

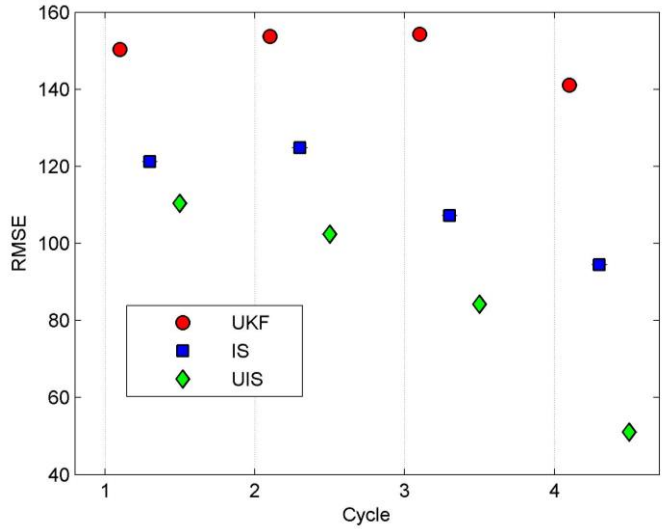
To compare the accuracy of the methods, the Root Mean Square Error (RMSE) of the model output at time  $t = 4$  is calculated at the end of each calibration cycle, using the expected value (mean) of the posterior distribution of  $\theta$ ,

$$RMSE^t = \sqrt{\frac{\sum_{x=1}^{20} [Y^{t=4}(x) - \mu^t(f^{t=4}(x, \theta))]^2}{20}} \quad (24)$$

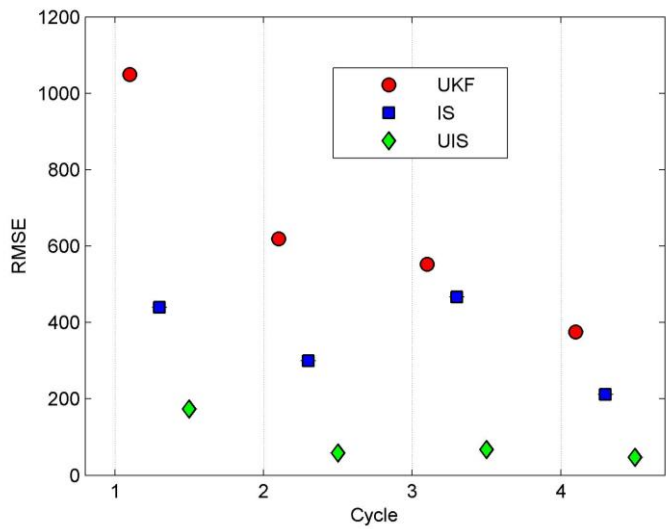
where  $\mu^t(f^{t=4}(x, \theta))$  is the mean of the posterior distribution of  $f^{t=4}(x, \theta)$  obtained at calibration cycle  $t$ .

### 3.3 Results and Discussion

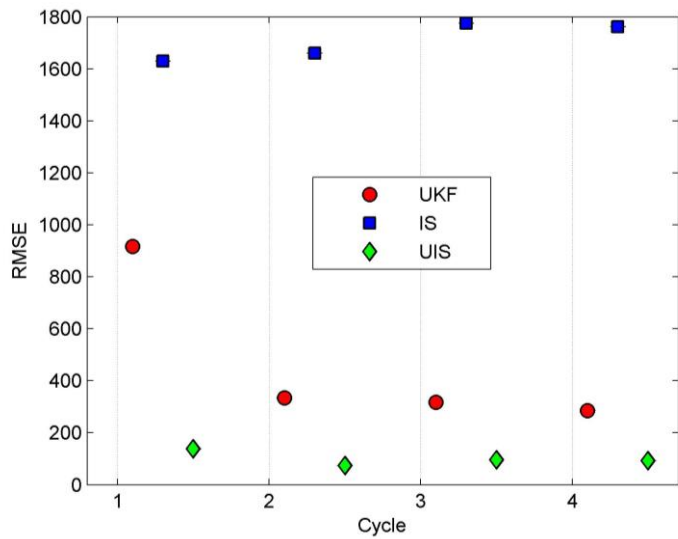
Fig. 4 plots the RMSE of all scenarios listed in Table 1. It can be seen that the RMSE of UIS is considerably smaller than the RMSE of the UKF measurement stage and IS in all cases. In the M10 and M20 cases, IS was more accurate than UKF but worse than UIS in terms of accuracy. In the M50 case, IS performed worse than both UIS and UKF.



(a) M10



(b) M20



(c) M50

**Fig. 4:** The RMSE of model output at the end of each calibration cycle for the (a) M10, (b) M20 and (c) M50 cases. The RMSE of UIS was well below RMSE of both UKF measurement stage and IS in all cycles of all scenarios.

From Table 1, it can be seen that the measurement update stage of UKF demanded the smallest number of model simulation. However, the UIS was always more accurate than the UKF. For instance in the case of 20 uncertain variables, it can be seen from Fig. 4(b) that the error measured by the RMSE of M20-UIS was more than 10 times smaller than the error obtained in M20-UKF. The reason why UIS is more accurate than either IS or UKF is that UIS samples from a proposal distribution that is closer to the posterior, without assuming linearity of the model and Gaussian distributions. Thus, UIS is able to capture the true posterior more accurately than measurement update stage of UKF and IS without a dramatic increase in number of simulations.

It can also be seen from Table 1 that UIS used more simulations (10.5%, 20.5% and 50.5% in M10, M20 and M50 cases, respectively) than ordinary IS but obtained notably more accuracy in all scenarios. It can also be seen that relative accuracy of UIS over the ordinary IS method increase with number of parameters. For instance, the RMSE of M10-UIS was less than two times smaller than the RMSE of M10-IS, while the RMSE of M50-UIS was more than 30 times smaller than the RMSE of M50-IS. The reason for this increasing relative accuracy is that UIS adaptively revised the proposal distribution while ordinary IS sampled from the initial prior in all cycles. Thus, when the number of uncertain parameters increase, IS was unable to estimate the larger-dimensional posterior distributions with 200 samples while UIS samples from important areas of posterior distributions and yields a much smaller RMSE with the same number of samples.

For more in-depth comparison of UIS with UKF and IS, including an investigation of effect of tuning parameters on the performance of UIS, a more sophisticated CS case study is presented in the next section.

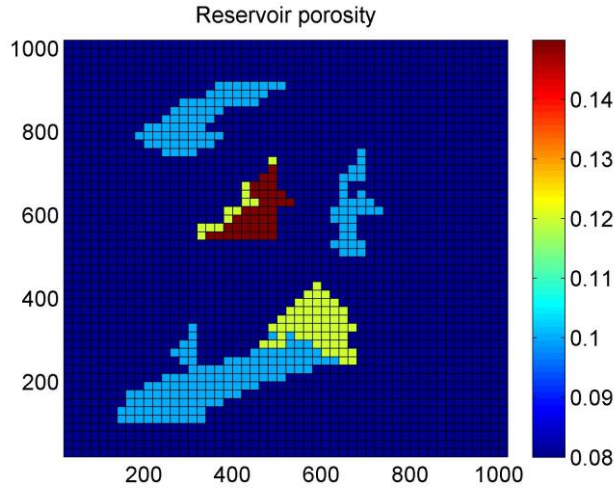
## 4. CS Case Study

A synthesized case study of CO<sub>2</sub> injection in a deep saline aquifer is generated and used for two purposes; first, comparing the effectiveness of UIS with measurement update of UKF and IS with a static proposal; and second, studying the effect of the number of samples ( $N$ ) and defensive mixture ratio ( $\eta$ ) on the performance of UIS. Nine calibration scenarios are designed for this task.

### 4.1 Case Description

A horizontal and rigid saline aquifer is assumed, initially saturated with brine. The aquifer is 20 m thick and its base is located at depth of 2 km. The aquifer domain is assumed to be 1020 m  $\times$  1020 m. Four homogeneous facies are assumed within the aquifer, distributed as shown in Fig. 5. The

porosities of the facies are assumed to be  $\phi_1 = 0.18$ ,  $\phi_2 = 0.13$ ,  $\phi_3 = 0.15$  and  $\phi_4 = 0.21$ , respectively. For all four facies, the porosity-permeability correlation is assumed to be  $\log_{10}(k_i) = 7 + 7 \log_{10}(\phi_i)$  where  $k_i$  is the intrinsic permeability of facies  $i$  in millidarcies and  $\phi_i$  is the porosity of facies  $i$  as a bulk volume fraction (SPE International, n.d.).

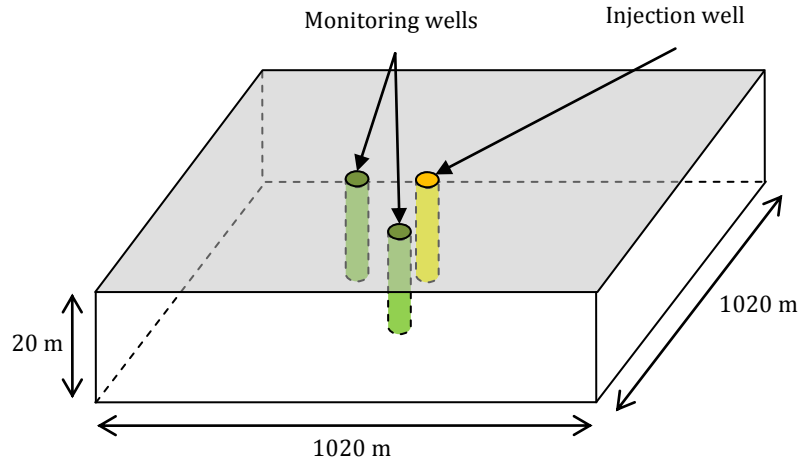


**Fig. 5** True porosity distribution in the reservoir

Initial salinity of the brine is assumed to be 50,000 ppm. The  $\text{CO}_2$  injection rate ( $Q$ ) is assumed to be 0.035 Mt/yr and the injection is assumed to take place continuously for 5 years through a vertical well located at the center of the domain and penetrating the whole thickness of the aquifer.

All parameters are assumed to be known except the porosity and the permeability of the four facies. The porosities are considered to be the uncertain parameters and the permeabilities are obtained using the porosity-permeability correlation given above. The prior distribution of the uncertain parameter vector  $([\phi_1, \phi_2, \phi_3, \phi_4]^T)$  is assumed to be a multivariate Gaussian with mean of  $[0.15, 0.15, 0.15, 0.15]^T$ , diagonal covariance matrix elements of  $0.02^2$  and off-diagonal covariance matrix elements of zero.

Pressure and saturation at the end of each year is measured at the injection well and two other monitoring wells, located 110 m away from the injection well in X and Y directions, respectively. Fig. 6 illustrates a schematic of the aquifer dimensions, the injection wells and the monitoring wells. The saturation measurement error is assumed Gaussian with a mean of zero and a standard deviation of 0.02. The pressure measurement error is also assumed Gaussian with a mean of zero and a standard deviation of  $5 \times 10^5$  Pa.



**Fig. 6** Schematic of the aquifer dimensions, injection well and the monitoring wells.

## 4.2 System Model

The system model is governed by the mass conservation equations for the multiphase and multicomponent flow in a porous medium of porosity  $\phi$ , which are given by

$$\frac{\partial(\phi \sum_i X_\beta^i \rho_i S_i)}{\partial t} + \sum_i \nabla \cdot (X_\beta^i \rho_i q_i) = Q_i \quad (25)$$

where  $X_\beta^i$  is the mass fraction of component  $\beta$  in phase  $i$ ,  $\rho_i$  is the density of phase  $i$ ,  $S_i$  is the saturation of phase  $i$  and  $Q_i$  is the source/sink volumetric flux of phase  $i$ . The volume flux of phase  $i$ ,  $q_i$ , is given by Darcy's law,

$$q_i = \frac{k k_{ri}}{\mu_i} \nabla (P_i - \rho_i g) \quad (26)$$

where  $i$  is the phase index,  $k$  is the intrinsic permeability,  $g$  is the gravity and  $k_{ri}$  is the relative permeability of phase  $i$ .  $P_i$  is the pressure of phase  $i$  and is obtained by

$$P_i = P_{ref} - P_{ci} \quad (27)$$

where  $P_{ref}$  is the reference pressure and  $P_{ci}$  is the capillary pressure. Relative permeability and capillary pressure of each phase are assumed functions of saturation.

Schlumberger’s ECLIPSE reservoir simulator is used for solving above flow equations to obtain pressure and the saturation of each phase (CO<sub>2</sub> or water) and the molar fraction of each component (salt). GASWAT keyword in ECLIPSE allows simulating multiphase equilibrium of gas and aqueous phases using a modified Peng-Robinson equation of state (Soreide & Whitson, 1992)(Schlumberger, n.d.). No flow boundary conditions are specified on top and bottom of the reservoir, open boundary conditions are specified on reservoir boundaries and an adaptive implicit solution procedure is used (Tavakoli, et al., 2013). A more detailed description of the flow equations and the solution options in ECLIPSE is beyond the scope of this article and can be found in the ECLIPSE technical manual (Schlumberger, n.d.).

### 4.3 Calibration Scenarios and Comparison Metrics

Nine calibration scenarios are described in Table 2. All scenarios use UIS as calibration method, except scenario #1 which uses the measurement update step of UKF and scenario #2 which uses IS with prior distribution as the proposal. The 5<sup>th</sup> scenario is chosen as the base case for UIS where 200 samples are taken and defensive mixture distribution is not used. Scenarios 3, 4 and 6 are similar to the base case, except for number of samples. Scenarios 7, 8 and 9 have different defensive mixture ratios compared to the base case, with a uniform mixture distribution bounded between 0.01 and 0.5 for the porosities of all facies. For all scenarios,  $\omega_0$  in the UT stage is assumed to be 0.

Table 2: Calibration scenarios in CS case study

<b>Scenario # and name</b>	<b>Method</b>	<b>Sample count (N)</b>	<b>Defensive mixture ratio (<math>\eta</math>)</b>	<b>Simulations per cycle</b>
1: UKF	UKF	-	-	9
2: IS	IS	200	0	200
3: N50	UIS	50	0	59
4: N100	UIS	100	0	109
5: BASE	UIS	200	0	209
6: N500	UIS	500	0	509
7: $\eta$ 0.1	UIS	200	0.1	209
8: $\eta$ 0.25	UIS	200	0.25	209
9: $\eta$ 0.5	UIS	200	0.5	209

Scenarios above are constructed so that three set of comparisons can be conducted:

1. Comparing measurement stage of UKF, IS (with the prior as proposal) and UIS (scenarios 1,2 and 3)
2. Comparing the effect of number of samples in UIS (scenarios 3, 4, 5 and 6)
3. Comparing the effect of defensive mixture ratio (scenarios 5, 7 and 8)

Three metrics are chosen for these comparisons:

- A. The first metric is the normalized Root Mean Squared Error (RMSE) of saturation and pressure fields during the first 5 years of injection, evaluated after each calibration cycle. RMSE of each field is first obtained using the expectation of the corresponding error field, and is normalized by the variance of monitoring error corresponding to that field. The sum of the normalized RMSE of both fields is then used as the comparison metric.
- B. The second metric is the ratio of effective samples to the total samples and is denoted by  $R_{eff}$  for scenarios 2 to 10, where effective samples size is defines as

$$N_{eff} = \frac{1}{\sum_{i=1}^N W_i^2} \quad (28)$$

and

$$R_{eff} = \frac{N_{eff}}{N} \quad (1)$$

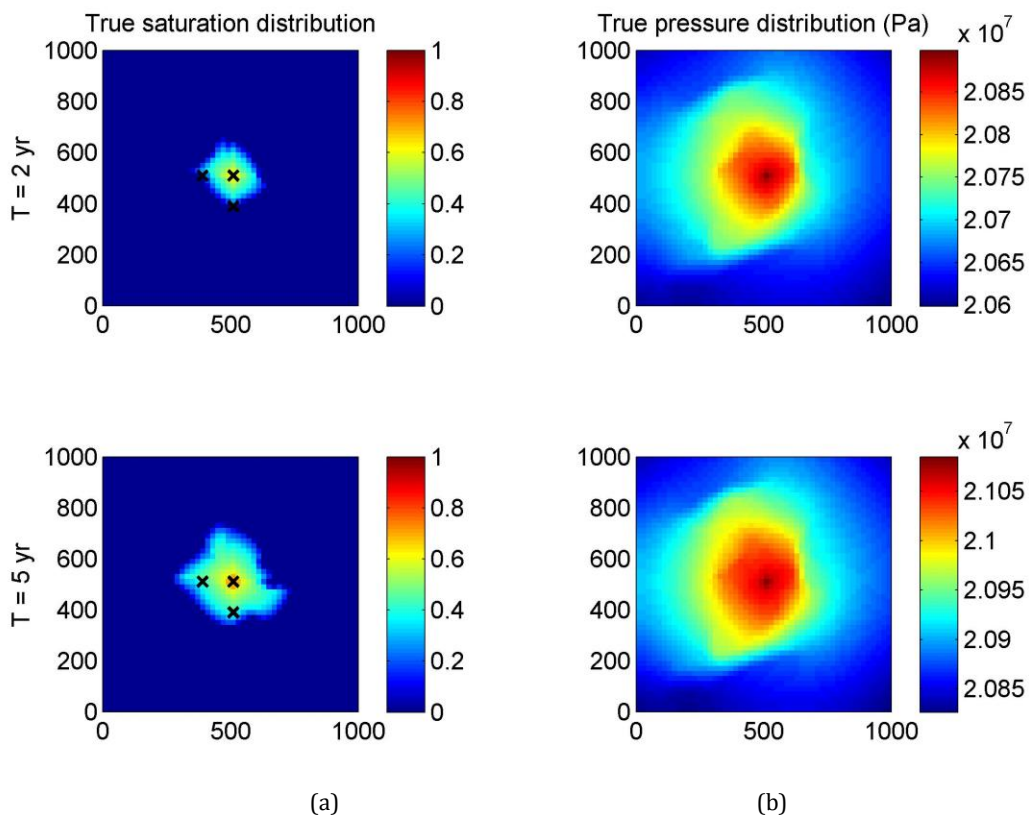
$N_{eff}$  is a measure for sample quality of the particle filters (Arulampalam, et al., 2002). When all importance weights are equal,  $N_{eff}$  is equal to  $N$  meaning that the samples are drawn from the true posterior. When most samples are drawn from non-important areas of the true posterior, most  $W_i$  become very close to zero while a few  $W_i$  carry most of the weight. In such cases,  $N_{eff}$  becomes closer to one indicating that just a few samples are representing the whole distribution and sample quality is poor. In the extreme case that one sample carries all the weight (i.e. one  $W_i$  equals to one and all other  $W_i$  equal zero),  $N_{eff}$  becomes equal to one. This situation is known as *filter degeneracy* of the particle filters (Arulampalam, et al., 2002) and should be avoided. The adaptive nature of UIS is expected to increase  $N_{eff}$  since the proposal distribution is adaptively adjusted to resemble to true posterior approximately.



- C. The third metric is the expectation of the absolute error and standard deviation of the error for both the saturation and pressure fields. These fields are plotted in Section 4.5 and discussed qualitatively.

## 4.4 Results

The synthetic “True” system behaviour generated using true system characteristics and is perturbed by the monitoring noise as described in Section 4.1. Fig. 7 is a snapshot of the true saturation and pressure distribution of CO<sub>2</sub> within the reservoir after 2 and 5 years of injection. Black “X”s in the saturation plots indicate monitoring locations. The CO<sub>2</sub> plume reached the monitoring wells by the second year of injection. Thus, the first calibration cycle is mostly based on the pressure measurements because the pressure plume spreads considerably faster than the saturation plume.



**Fig. 7** Snapshots of true (a) effective saturation and (b) pressure distribution (Pa) of the reservoir after 2 and 5 years of injection. Black “X”s on saturation plots indicates monitoring locations.

Calibration is conducted for all scenarios listed in Table 2. After each calibration cycle for each scenario, the model is stochastically simulated (i.e. Monte Carlo simulation) up to year 5 using the weighted samples from the posterior distribution of that cycle. The expectation and standard deviation of the error between true and forecasted saturation and pressure fields is then generated for each scenario. In all scenarios, the saturation error expectation and the saturation uncertainty (standard deviation) was more significant near the plume boundaries. The pressure error expectation and uncertainty were more complex and did not follow such a standard pattern.

The RMSE of all scenarios and  $R_{eff}$  for scenarios 2 to 9 were calculated after each cycle and are listed in Table 3. This table will be used as the reference for further comparisons in Section 4.5. The overall trend of the table suggests that UIS performed better than both UKF and IS, increasing number of samples increases the performance of UIS, and a moderate defensive mixture ratio increases the performance of UIS at the cost of a slight decrease in  $R_{eff}$ .

Table 3: RMSE and ratio of effective samples for all scenarios.

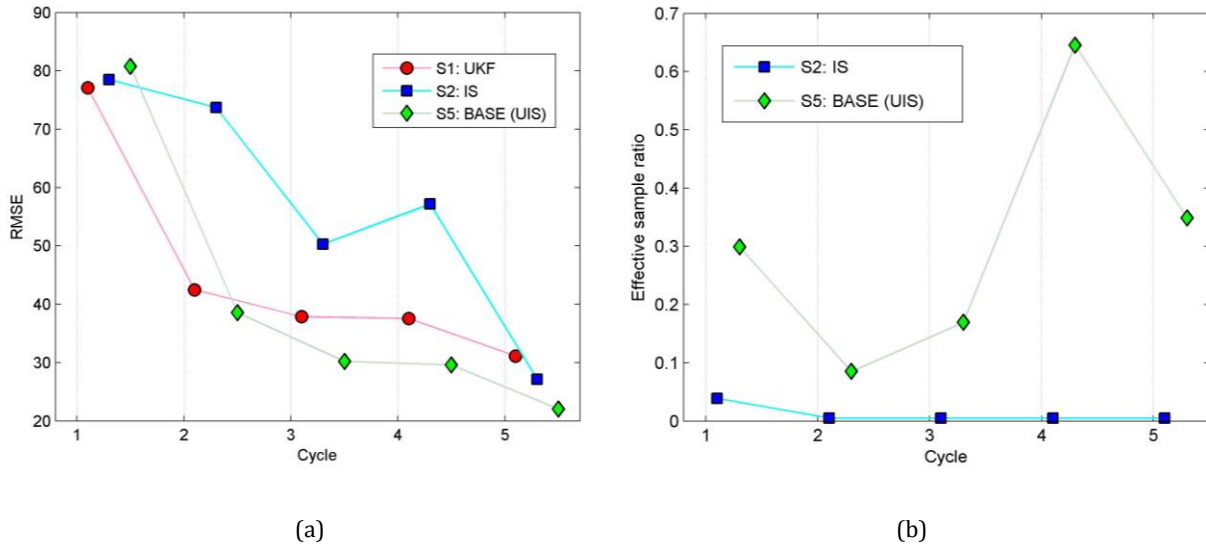
Scenario	Cycle 1		Cycle 2		Cycle 3		Cycle 4		Cycle 5	
	RMSE	$R_{eff}$	RMSE	$R_{eff}$	RMSE	$R_{eff}$	RMSE	$R_{eff}$	RMSE	$R_{eff}$
1: UKF	77.12	-	42.49	-	37.86	-	37.57	-	31.12	-
2: IS	78.48	0.04	73.68	0.005	50.30	0.005	57.18	0.005	27.14	0.005
3: N50	77.74	0.03	65.27	0.09	46.68	0.13	40.84	0.042	32.81	0.05
4: N100	76.59	0.50	37.16	0.25	35.11	0.25	28.44	0.13	23.94	0.18
5: BASE	80.72	0.30	38.55	0.09	30.21	0.17	29.60	0.65	22.05	0.34
6: N500	75.42	0.26	34.91	0.15	20.64	0.34	28.45	0.79	21.88	0.44
7: $\eta$ 0.1	83.61	0.19	33.89	0.17	20.88	0.50	20.00	0.57	23.71	0.27
8: $\eta$ 0.25	75.50	0.28	22.83	0.061	31.42	0.16	29.01	0.46	21.17	0.29
9: $\eta$ 0.5	78.05	0.21	39.62	0.13	29.74	0.13	28.05	0.16	22.04	0.15

## 4.5 Discussion

Now we compare the calibration scenarios to assess the performance of UIS against UKF measurement update and IS, and also investigate the effect of the tuning parameters (number of samples and defensive mixture ratio) on the performance of UIS. We also discuss the computational cost of these algorithms.

### 4.5.1 Comparing Calibration Algorithms

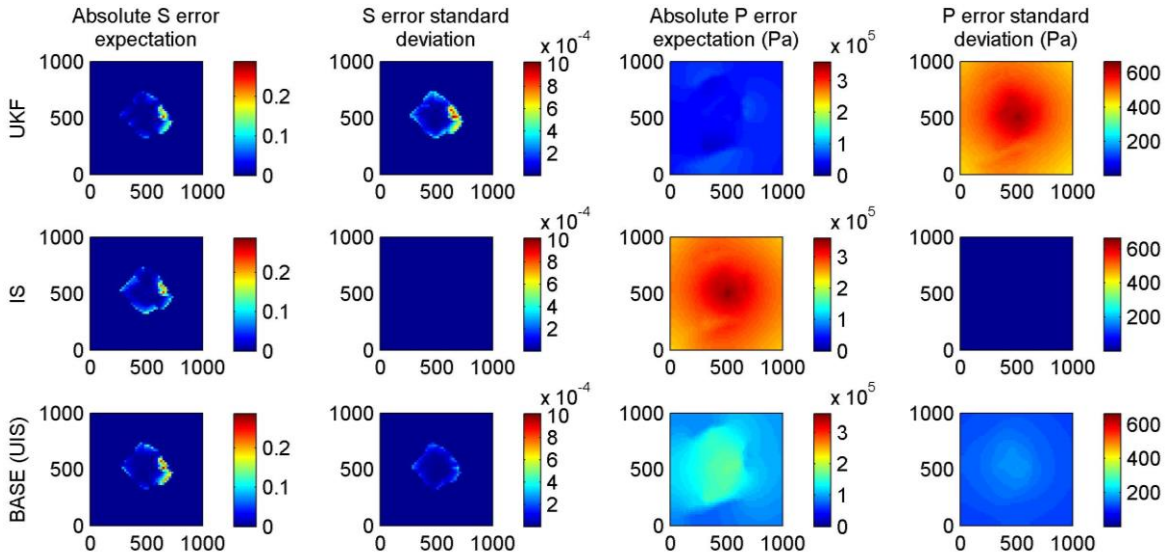
The RMSE for UKF, IS and the base case of UIS (scenarios 1, 2 and 5) and the effective ratio of the samples for IS and the base case of UIS are extracted from Table 3 and plotted in Fig. 8. It can be seen that UIS had a smaller RMSE compared to both UKF and IS. With the exception of the last year, UKF update stage performed better than IS. This is because the static proposal distribution in IS was considerably wider compared to the real posterior. Thus, most samples were drawn from non-important areas of the posterior probability density, resulting in a large ratio of samples having weights of zero or close to zero. Thus, the posterior distribution was represented by very few samples. This shortcoming is apparent in Table 3 and Fig. 8b. It can be seen that after the second year, number of effective samples for IS is very close to one. The number of samples should be increased to increase  $N_{eff}$  for this case, which will also increase the computational burden. UIS on the other hand used the same number of samples but adaptively adjusted the proposal with the Gaussian approximation of the real posterior. This approach resulted in significantly larger  $R_{eff}$  compared to IS, which is more desirable.



**Fig. 8** (a) RMSE and (b)  $R_{eff}$  for scenarios 1 (UKF), 2 (IS with the prior used as proposal) and 5 (Base case for UIS). Data point connections and the small offset between data points of the same cycle are for clarity. UIS demonstrate smaller RMSE compared to UKF and IS, and has significantly larger ratio of effective samples compared to IS.

Fig. 9 shows the absolute error expectation and error standard deviation of the saturation and pressure fields at the end of year five and after five calibration cycles for UKF, IS and UIS. In the IS scenario, the uncertainty in both the saturation and the pressure fields were almost zero and the error expectation for both fields were noticeably higher than UKF and UIS, because of having one

sample carrying most of the importance weight. The uncertainty in both fields were higher in UKF than in UIS. It is worth noting that while sampling-based solutions tend to underestimate the uncertainty, the UKF results do not necessarily represent the true uncertainty either because of the model linearization and the Gaussian assumptions in UKF method.

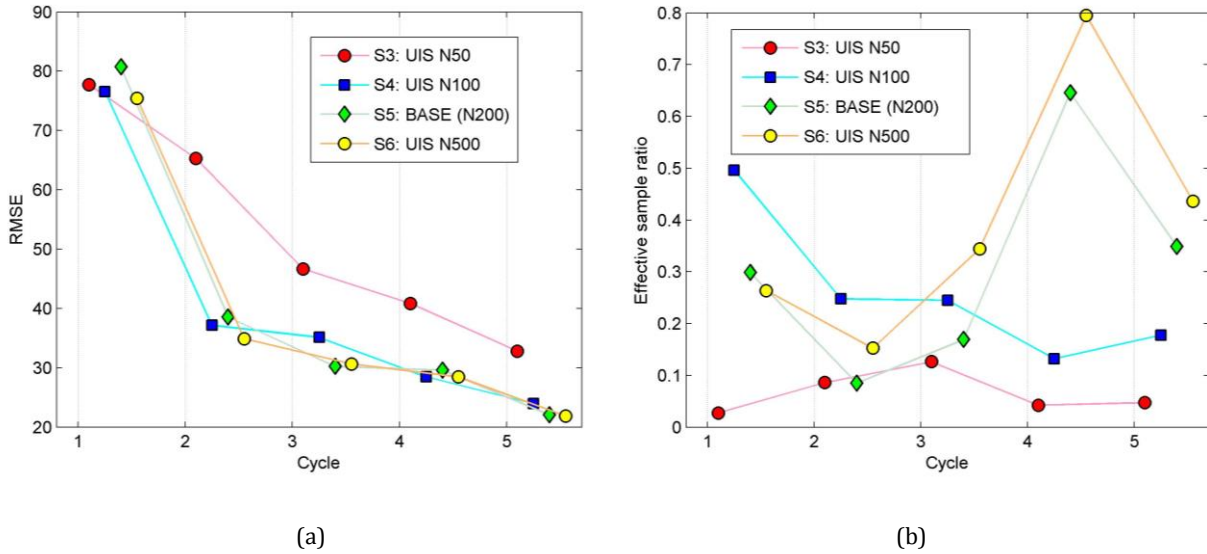


**Fig. 9** Absolute error expectation and standard deviation of saturation (S) and pressure (P) fields at the end of year 5 and after 5 calibration cycles for UKF, IS and BASE (UIS) scenarios.

The accuracy of UKF is expected to deteriorate more compared to that of UIS when the distributions (prior and the likelihood) become non-Gaussian. The performance of UIS in case of non-Gaussian densities will not be affected as much as UKF, since no assumptions is made on the type of distributions and UKF stage of UIS generates the proposal not the posterior.

#### 4.5.2 Effect of Number of Samples on UIS

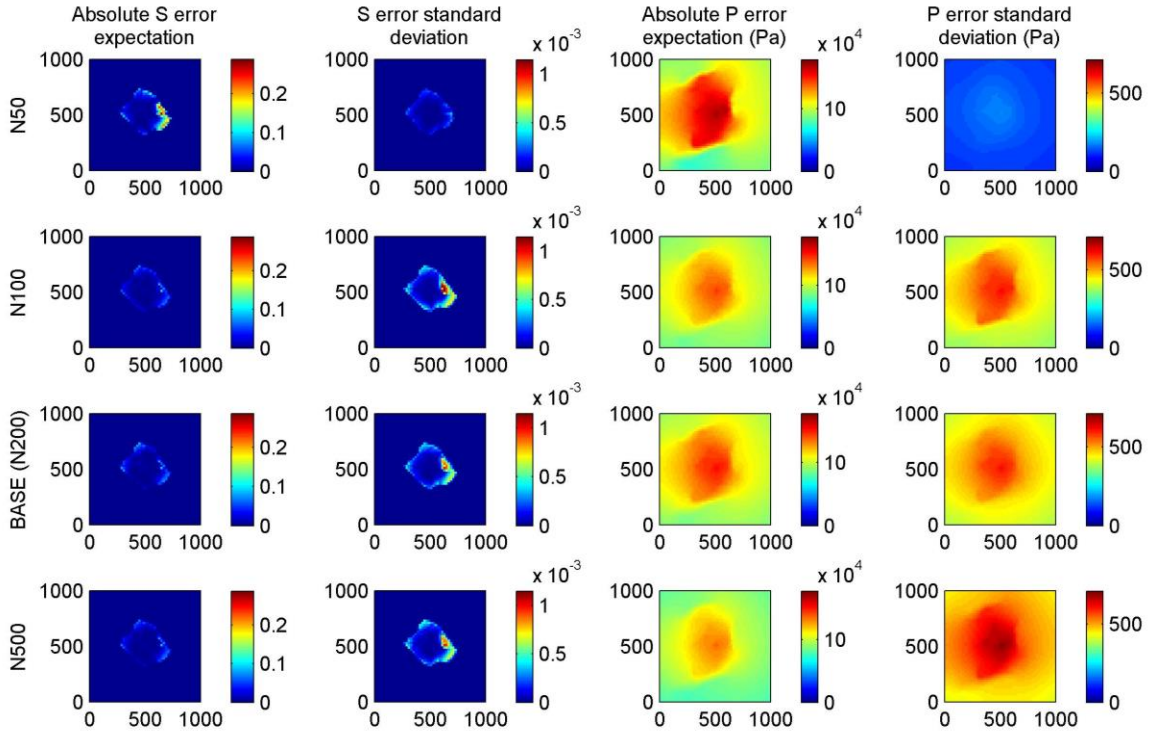
Fig. 10 illustrates the RMSE and  $R_{eff}$  for UIS with 50, 100, 200 and 500 samples (scenarios 3, 4, 5 and 6), extracted from Table 3. The general trend of the RMSE in Fig. 10a shows that increasing the number of samples increases the accuracy of UIS, which was not un-expected in a Monte-Carlo based method. However, the difference between scenarios deteriorates when the number of samples increases, e.g., the difference between RMSE of N50 and N100 scenarios is more significant compared to the difference between RMSE of N100, N200 and N500 scenarios. Taking into account both accuracy and computational efficiency, N100 scenario maintained the best balance among rest of the scenarios.



**Fig. 10** (a) RMSE and (b)  $R_{eff}$  for UIS with 50, 100, 200 and 500 samples. Data point connections and the small offset between data points of the same cycle are for clarity. Increasing number of samples decreased RMSE in general. The difference between the case with 200 samples and 500 samples was not significant. Ratio of effective samples increased with number of samples in general.

According to Fig. 10b, the effective sample ratio increases with number of samples. This is because the weights are more evenly distributed among the samples when more samples are drawn from the posterior distribution. As a result, the variance of the weights decreases and  $R_{eff}$  increases with  $N$ .

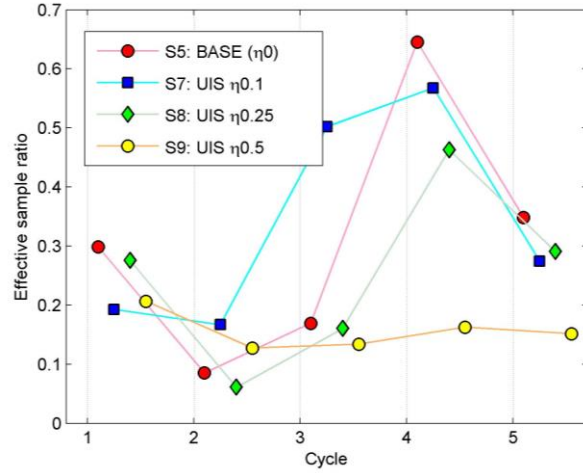
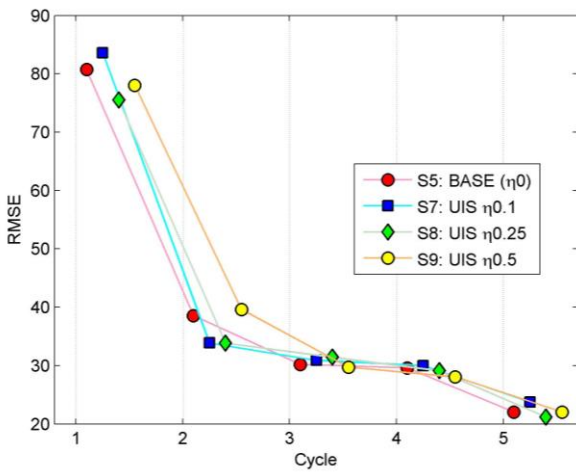
Fig. 11 shows the absolute error expectation and the error standard deviation of the saturation and pressure fields at the end of year five and after five calibration cycles for UIS with 50 to 500 samples. The uncertainty of both the pressure and the saturation fields are noticeably underestimated and the error expectations are noticeably higher in the N50 scenario compared to other scenarios, due to a very small sample size. Scenarios with more samples are expected to more accurately represent the true uncertainty in the pressure and saturation fields. Moreover, it can be seen that the error expectation is generally decreasing in both fields with increasing the sample size.



**Fig. 11** Absolute error expectation and standard deviation of saturation (S) and pressure (P) fields at the end of year 5 and after 5 calibration cycles for UIS scenarios with 50, 100, 200 and 500 samples.

### 4.5.3 Effect of Defensive Mixture Ratio in UIS

The effect of defensive mixture ratios of 0, 0.1, 0.25 and 0.5 (scenarios 5, 7, 8 and 9) on the RMSE and  $R_{eff}$  of UIS are plotted in Fig. 11, extracted from Table 3. It can be seen that the  $\eta_{0.1}$  and  $\eta_{0.25}$  scenarios resulted in smaller RMSE compared to the two other scenarios.  $R_{eff}$  for  $\eta_{0.1}$  and  $\eta_{0.25}$  scenarios was slightly smaller than the base case; however, increasing  $\eta$  to 0.5 decreased  $R_{eff}$  significantly. Increase in the defensive mixture ratio is generally expected to increase the robustness of UIS at the price of a slight decrease in  $R_{eff}$ , provided the proposal is close the real posterior. When this is not the case, e.g. when the system model is poor or the initial prior deviates from the posterior significantly, the defensive distribution helps search a wider range of the parameter space and decrease the chance of degeneracy of the UIS.

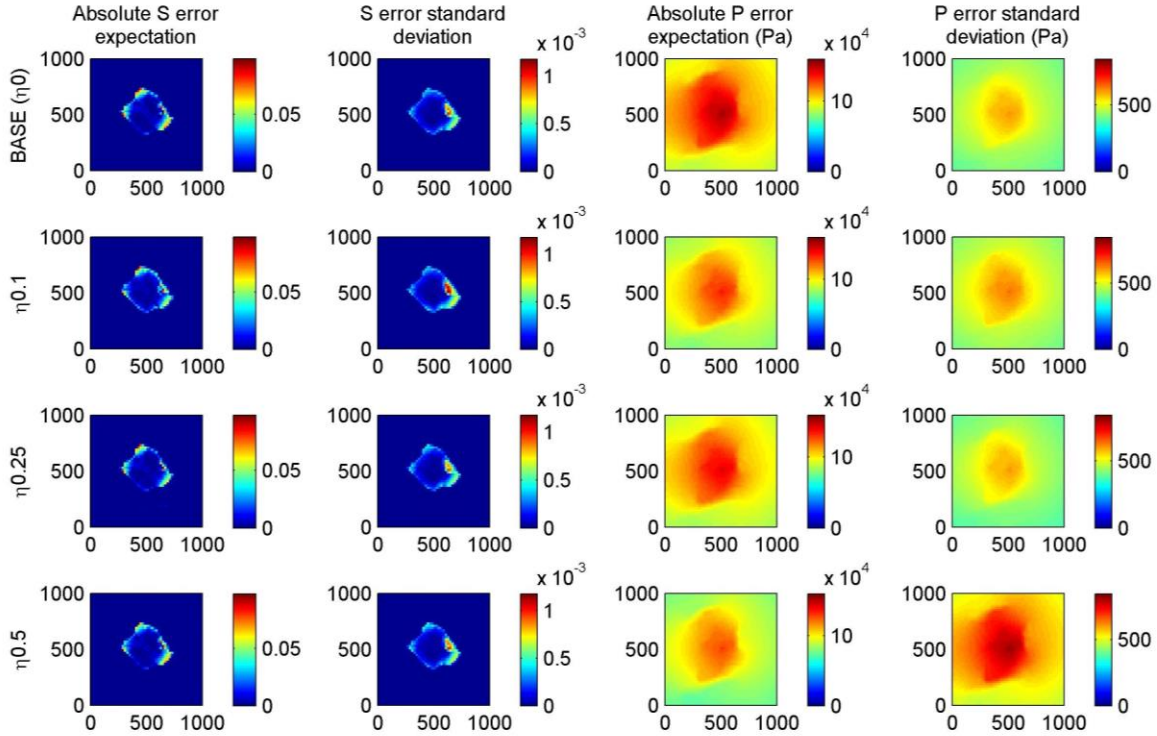


(a)

(b)

**Fig. 12** (a) RMSE and (b)  $R_{eff}$  for UIS with defensive mixture ratio of 0, 0.1, 0.25 and 0.50. Data point connections and the small offset between data points of the same cycle are for clarity. Increasing defensive ratio up to 0.25 slightly decreased the ratio of the effective samples, but the decline for  $\eta = 0.5$  is more significant. The RMSE for  $\eta = 0.25$  presented the best overall RMSE.

Fig. 13 shows the absolute error expectation and the error standard deviation of the saturation and the pressure fields at the end of year five and after five calibration cycles for UIS with 0 to 0.5 defensive mixture ratios. The error expectations for both the saturation and the pressure fields decreased with increasing defensive mixture ratio in general.



**Fig. 13** Absolute error expectation and standard deviation of saturation (S) and pressure (P) fields at the end of year 5 and after 5 calibration cycles for UIS scenarios with defensive mixture ratio of zero, 0.1, 0.25 and 0.5. All defensive mixture densities were uniform bounded between 0.01 and 0.5.

According to current results, we recommend using a defensive mixture ratio between  $\eta = 0.1$  and  $\eta = 0.25$  with a uniform defensive distribution. While larger defensive mixture ratios might result in slightly smaller errors in some cases (e.g. when the prior is very poorly chosen), they tend to decrease  $R_{eff}$  noticeably.

#### 4.5.4 Computational Cost

A factor that should be considered when comparing the algorithm accuracies is the computational cost. The most computationally costly stage of all the algorithms discussed above was model simulation. The difference between the computational cost of optimal size UIS (e.g. N100) was almost half of IS (109 model simulation per cycle versus 200 model simulation per cycle) while the differences in accuracies were significant. The computational cost of UKF and optimal size UIS (N100), however, was significant (109 model simulations per cycle versus 9 model simulations per cycle). UKF and in general, KF-based algorithms are computationally more efficient than sampling-based algorithms such as IS and UIS. The reason for this computational efficiency is the linearity and Gaussian assumptions in the KF-based algorithms which makes them computationally less costly. Sampling-based algorithms on the other hand require more samples in order to capture the



true posterior distribution more accurately without these assumptions. Thus, the difference between the computational costs of N100 UIS and UKF was not un-expected. The accuracy of UIS however was notably better than UKF, even though the prior distribution and likelihood were Gaussian in this case study.

## 5. Conclusion

The Bayesian-based Unscented Importance Sampling (UIS) method for mitigating parameter uncertainty is presented and applied to geological CO<sub>2</sub> sequestration. The methodology accounts for expert opinion via the prior distribution at the beginning of the calibration and periodically revises the joint probability distribution of uncertain parameters utilizing noisy monitoring data.

UIS is applied to a synthesized CO<sub>2</sub> injection case study in order to be benchmarked against Importance Sampling with a static proposal and the measurement update stage of an Unscented Kalman Filter. The case study is also used to investigate the impact of the number of samples and the defensive mixture ratio on the performance of UIS. Nine scenarios are designed for these comparisons. In the case study presented, UIS outperforms IS and UKF. It is demonstrated that increasing the number of samples enhance UIS performance. Using a moderate defensive mixture ratio increased the performance slightly with the price of slight decrease in efficient number of the samples. It is suggested that defensive mixture densities become more significant when the prior distribution is chosen more poorly. In the case study, UIS performed best when a defensive mixture ratio between  $\eta = 0.1$  and  $\eta = 0.25$  was used, coupled with a uniform defensive mixture distribution.

The benefits of UIS can be summarized as:

- Stochastic parameter calibration: UIS utilizes monitoring data to periodically revise the parameter probability distributions and mitigate the parameter uncertainty. Moreover, revised parameter distributions can demonstrate which areas of a geosequestration site are poorly understood and require heavier monitoring. Thus, it can also contribute to an adaptive monitoring program.
- Model flexibility: no assumptions are made on the system model. Thus, UIS can be used with linear/nonlinear/analytical/semi-analytical/numerical system models.

- Distribution flexibility: no assumptions are made on the type of probability distributions (prior/posterior/likelihood). Any type of continuous/discrete/formal/empirical probability distribution can be used as the prior or likelihood.
- Accuracy and computational efficiency: It is shown that UIS can be more accurate than importance sampling with a static proposal distribution. It is also shown that UIS is more accurate than the UKF measurement stage, even with a Gaussian prior and likelihood. As a result, to obtain certain level of accuracy UIS is likely to require a considerably smaller number of samples compared to IS with a static proposal.

Current work is subject to limitations that can be addressed in further researches, namely:

- Current formulation does not account for systematic model error, e.g., error autocorrelation. While this is a fairly usual approach in Bayesian parameter calibration, it is possible to revisit the formulation in future works to include error autocorrelation.
- As described in section 2.4, current algorithm could underperform for multi-modal posterior distributions. In those cases, a bank of UKFs can replace the single UKF for generating the proposal distribution. However, the added computational demand for such cases should be also taken into account.
- When a very large dimensional parameter space is used, the computational demand of UKF becomes unfavourable. In those cases, Ensemble Kalman Filter (EnKF) can replace the UKF phase of the UIS method since EnKF does not require the sample size to increase linearly with the dimensionality of the parameter space.

## Acknowledgement

The authors acknowledge funding from the Natural Science and Engineering Research Council of Canada (NSERC) through Carbon Management Canada (CMC) and from NSERC's Discovery Grant program. Lastly, we would like to thank our colleagues Drs. James Craig and Maurice Dusseault for their moral support and academic insights.

## Bibliography

- Aanonsen, S. I. & Reynolds, A. C., 2009. The Ensemble Kalman Filter in Reservoir Engineering--a Review. *SPE Journal*, 14(3), pp. 393 - 412.
- Alshuhail, A., Lawton, D. & Isaac, H., 2009. *Seismic Characterizations of the Nisku Formation*, Calgary: University of Calgary.
- Ambadan, J. T. & Tang, Y., 2009. Sigma-Point Kalman Filter Data Assimilation Methods for Strongly Nonlinear Systems. *Journal of the Atmospheric Sciences*, Volume 66, p. 261–285.
- Arulampalam, M. S., Maskell, S., Gordon, N. & Clapp, T., 2002. A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking. *IEEE TRANSACTIONS ON SIGNAL PROCESSING*, 50(2), pp. 174-188.
- Bhowmik, S., Mantilla, C. A. & Srinivasan, S., 2011. Tracking CO<sub>2</sub> plume migration during geologic sequestration using a probabilistic history matching approach. *Stochastic Environmental Research and Risk Assessment*, Volume 25, p. 1085–1090.
- Cappé, O., Douc, R., Guillin, A. & Marin, J. M. R. C. P., 2008. Adaptive Importance Sampling in General Mixture Classes. *Statistics and Computing*, 18(4), pp. 447-459.
- Cappé, O., Guillin, A., Marin, J. M. & Robert, C. P., 2004. Population Monte Carlo. *Journal of Computational and Graphical Statistics*, 13(4), pp. 907-929.
- Chadwick, R. et al., 2004. 4D seismic imaging of an injected CO<sub>2</sub> plume at the Sleipner Field, Central North Sea. In: Davies R. (ed.) 3D Seismic Technology: Application to the Exploration of Sedimentary Basins. *Geological Society*, pp. 311-320.
- Chen, Z., 2003. Bayesian Filtering: From Kalman Filters to Particle Filters, and Beyond. *Statistics*, pp. 1-69.
- Douc, R., Guillin, A., Marin, J. -M. & Robert, C. P., 2007. Minimum variance importance sampling via Population Monte Carlo. *ESAIM: Probability and Statistics*, Volume 11, pp. 427-447.
- Espinet, A. J. & Shoemaker, C. A., 2013. Comparison of Optimization Algorithms for Parameter Estimation of Multiphase Flow Models with Application to Geological Carbon Sequestration. *Advances in Water Resources*, Volume 54, p. 133–148.

- Evensen, G., 1994. Sequential data assimilation with nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research*, 99(C5), p. 143–162.
- Freifeld, B. et al., 2005. The U-Tube: A novel system for acquiring borehole fluid samples from a deep geologic CO<sub>2</sub> sequestration experiment. *Journal of Geophysical Research*, Volume 110.
- Hall, B., Statisticat & LLC, 2012. *Bayesian Inference*, s.l.: s.n.
- Hesterberg, T., 1995. *Weighted average importance sampling and defensive mixture distributions*, Stanford: Stanford University.
- Hongjun, Z., Xinwei, L., Yanfang, C. & Xiaoliang, Z., 2010. Sensitivity analysis of CO<sub>2</sub> sequestration in saline aquifers. *Petroleum Science*, Volume 7, pp. 372-378.
- Houtekamer, P. L. et al., 2005. Atmospheric Data Assimilation with an Ensemble Kalman Filter: Results with Real Observations. *Monthly Weather Review*, 133(3), pp. 604-620.
- Hoversten, G. et al., 2004. Cross-well electromagnetic seismic imaging: An examination of coincident surveys at a steam flood project. *Geophysics*, Volume 69, p. 406–414.
- Johnson, J. & White, D., 2012. History Matching and Performance Validation. In: *Best Practices for Validating CO<sub>2</sub> Geological Storage: Observations and Guidance from the IEAGHG Weyburn-Midale CO<sub>2</sub> Monitoring and Storage Project*. 254: Geoscience Publishing, p. 215.
- Julier, S. & Uhlmann, J. K., 1996. A General Method for Approximating Nonlinear Transformations of Probability Distributions.
- Kaipio, J. P. & Somersalo, E., 2007. Statistical inverse problems: Discretization, model. *Journal of Computational and Applied Mathematics*, 198(2), p. 493–504.
- Kalman, R., 1960. A new approach to linear filtering and prediction problems. Volume 82, pp. 35-45.
- Kim, I. S., 2011. *Large Scale Data Assimilation with Application to the Ionosphere-thermosphere*. s.l.:Proquest, Umi Dissertation Publishing.
- Kopp, A. et al., 2010. A contribution to risk analysis for leakage through abandoned wells in geological CO<sub>2</sub> storage. *Advances in Water Resources*, Volume 33, pp. 867-879.

- Lavoie, R. & Keith, D., 2010. *Executive Summary of Wabamun Area CO2 Sequestration Project (WASP)*, Calgary: University of Calgary.
- Lawton, D., 2010. Carbon capture and storage: opportunities and challenges for geophysics. *CSEG Recorder*, 35(6), pp. 7-10.
- Lawton, D. et al., 2010. *Recommendations for Injection and Storage Monitoring, WABAMUN AREA CO2 SEQUESTRATION PROJECT (WASP)*, Calgary: s.n.
- Leisenring, M. & Moradkhani, H., 2011. Snow water equivalent prediction using Bayesian data assimilation methods. *Stochastic Environmental Research and Risk Assessment*, 25(2), pp. 253-270.
- Liu, X., Cardiff, M. A. & Kitanidis, P. K., 2010. Parameter estimation in nonlinear environmental problems. *Stochastic Environmental Research and Risk Assessment*, 24(7), pp. 1003-1022.
- Li, X., Cardiff, M. A. & Kitanidis, P. K., 2004. Parameter estimation in nonlinear environmental problems. *Stochastic Environmental Research and Risk Assessment*, 24(7), pp. 1003-1022.
- Li, Z. & Fall, M., 2013. *A modeling tool for assessment of potential groundwater contamination in response to CO2 leakage from geological disposal of CO2*. s.l., s.n.
- Marshall, A., 1956. *The use of multi-stage sampling schemes in Monte Carlo computations*. New York, Wiley, p. 123-140.
- Merwe, R. v. d., Doucet, A., Freitas, N. d. & Wan, E., 2000. *The Unscented Particle Filter*, s.l.: s.n.
- Mesbah, A., Huesmana, A., Kramerb, H. & Van den Hofa, P., 2011. A comparison of nonlinear observers for output feedback model-based control of seeded batch crystallization processes. *Journal of Process Control*, 21(4), p. 652-666.
- Naevdal, G., Aanonsen, S. I., Vefring & EH, 2005. Reservoir Monitoring and Continuous Model Updating Using Ensemble Kalman Filter. *SPE Journal*, 10(1), pp. 66 - 74.
- Nooner, S. et al., 2007. Constraints on the in situ density of CO2 within the Utsira Formation from time-lapse seafloor gravity measurements. *International Journal of Greenhouse Gas Control*, Volume 1, p. 198-214.
- Nordbotten, J. et al., 2012. Uncertainties in practical simulation of CO2 storage. *International Journal of Greenhouse Gas Control*, Volume 9, pp. 234-242.

- Oladyshkin, S., Class, H., Helmig, R. & Nowak, W., 2011. An integrative approach to robust design and probabilistic risk assessment for CO<sub>2</sub> storage in geological formations. *Computational Geoscience*, Volume 15, pp. 565-577.
- Owen, A. & Zhou, Y., 1998. *Safe and effective importance sampling*, s.l.: Stanford University.
- Price, P. & Oldenburg, C., 2009. The consequences of failure should be considered in siting geological carbon sequestration projects. *International Journal of Greenhouse Gas Control*, Volume 3, pp. 658-663.
- Raikes, S. et al., 2008. *Integration of 3D Seismic With Satellite Imagery At In Salah CO<sub>2</sub> Sequestration Project, Algeria*. Las Vegas, Nevada, s.n.
- Reichle, R. H., McLaughlin, D. B. & Entekhabi, D., 2002. Hydrologic Data Assimilation with the Ensemble Kalman Filter. *Monthly Weather Review*, 130(1), pp. 103-116.
- Rougier, J., 2008. Formal Bayes Methods for Model Calibration with Uncertainty. In: *Applied Uncertainty Analysis for Flood Risk Management*. s.l.:Imperial College Press.
- Sarkarfarshi, M., Malekzadeh, F. A., Gracie, R. & Dusseault, M. B., 2014. Parametric sensitivity analysis for CO<sub>2</sub> geosequestration. *International Journal of Greenhouse Gas Control*, Volume 23, pp. 61-71.
- Sato, K., 2011. Value of information analysis for adequate monitoring of carbon dioxide storage in geological reservoirs under uncertainty. *International Journal of Greenhouse Gas Control*, Volume 5, pp. 1294-1302.
- Schlumberger, n.d. *ECLIPSE Technical Description Version 2013.2*. s.l.:s.n.
- Shu, Q., Kemblowski, M. W. & McKee, M., 2005. An application of Ensemble Kalman Filter in integral-balance subsurface modeling. *Stochastic Environmental Research and Risk Assessment*, Volume 19, p. 361-374.
- Smith, P. J., Shafi, M. & Gao, H., 1997. Quick Simulation: A Review of Importance Sampling Techniques in Communications Systems. *IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS*, 15(4), pp. 597-613.
- Soreide, I. & Whitson, C., 1992. Peng Robinson predictions for hydrocarbons, CO<sub>2</sub>, N<sub>2</sub> and H<sub>2</sub>S with pure water and NaCl brine. *Fluid Phase Equilibria*, Volume 77, pp. 217-240.

SPE International, n.d. *Permeability determination*. [Online]  
Available at: [http://petrowiki.org/Permeability\\_determination#Determining\\_permeability](http://petrowiki.org/Permeability_determination#Determining_permeability)

Stedinger, J., Vogel, R., Lee, S. & Batchelder, R., 2008. Appraisal of the generalized likelihood uncertainty estimation (GLUE) method. *Water Resources Research*, Volume 44, p. W00B06.

Tavakoli, R. et al., 2013. Comparison of ensemble filtering algorithms and null-space Monte Carlo for parameter estimation and uncertainty quantification using CO<sub>2</sub> sequestration data. *WATER RESOURCES RESEARCH*, Volume 49, pp. 1-20.

Terejanu, G. A., n.d. *Unscented Kalman Filter Tutorial*. Buffalo(NY): University at Buffalo.

van Leeuwen, P. J. & Evensen, G., 1996. Data assimilation and inverse methods in terms of a probabilistic formulation. *Monthly Weather Review*, Volume 124, p. 2898–2913.

Walton, F. B., Tait, J. C., LeNeveu, D. & Sheppard, M. I., 2004. *Geological storage of CO<sub>2</sub>: A statistical approach to assessing performance and risk*. Vancouver, Canada, s.n.

Wan, E. & Merwe, R. v. d., 2001. *The Unscented Kalman Filter*. s.l.:Wiley Publishing.