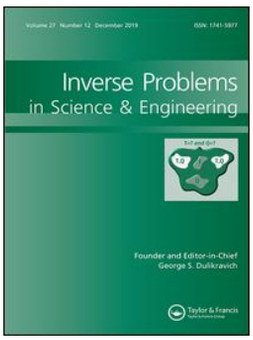

Inverse Problems in Science and Engineering

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/gipe20

Accelerated Bayesian inference-based history matching of petroleum reservoirs using polynomial chaos expansions

Sufia Khatoon, Jyoti Phirani & Supreet Singh Bahga

To cite this article: Sufia Khatoon, Jyoti Phirani & Supreet Singh Bahga (2021): Accelerated Bayesian inference-based history matching of petroleum reservoirs using polynomial chaos expansions, Inverse Problems in Science and Engineering, DOI: 10.1080/17415977.2021.1973455

To link to this article: https://doi.org/10.1080/17415977.2021.1973455

Published online: 07 Sep 2021.

INVERSE PROBLEMS IN SCIENCE AND ENGINEERING
https://doi.org/10.1080/17415977.2021.1973455

Accelerated Bayesian inference-based history matching of petroleum reservoirs using polynomial chaos expansions

Sufia Khatoon[a], Jyoti Phirani[b] and Supreet Singh Bahga[a]

[a]Department of Mechanical Engineering, Indian Institute of Technology, Delhi, India; [b]Department of Chemical Engineering, Indian Institute of Technology, Delhi, India

ABSTRACT

The forecast for oil production from an oil reservoir is made with the aid of reservoir simulations. The model parameters in reservoir simulations are uncertain whose values are estimated by matching the simulation predictions with production history. Bayesian inference (BI) provides a convenient way of estimating parameters of a mathematical model, starting from a probable distribution of parameter values and knowing the production history. BI techniques for history matching require Markov chain Monte Carlo (MCMC) sampling methods, which involve large number of reservoir simulations. This limits the application of BI for history matching in petroleum reservoir engineering, where each reservoir simulation can be computationally expensive. To overcome this limitation, we use polynomial chaos expansions (PCEs), which represent the uncertainty in production forecasts due to the uncertainty in model parameters, to construct proxy models for model predictions. As an application of the method, we present history matching in simulations based on the black-oil model to estimate model parameters such as porosity, permeability, and exponents of the relative permeability curves. Solutions to these history matching problems show that the PCE-based method enables accurate estimation of model parameters with two orders of magnitude less number of reservoir simulations compared with MCMC method.

ARTICLE HISTORY
Received 24 March 2020
Accepted 19 August 2021

KEYWORDS
Inverse problems; history matching; Bayesian inference; polynomial chaos

AMS SUBJECT CLASSIFCATIONS
62F; 65C; 60J

1. Introduction

The ever-increasing demand for petroleum necessitates efficient management of existing oil and gas fields and development of new fields. The forecast of oil production from an oil reservoir is made with the aid of reservoir simulations, which predict the flow of oil, gas, and water through porous media [1]. Accurate production forecasts using reservoir simulations require precise knowledge of the properties of the reservoirs [2]. The subsurface of a petroleum reservoir is invariably complex and heterogeneous [3]. Because measurements are limited to discrete locations, geological properties and initial conditions are usually determined by solving inverse problems using the historical production data.

CONTACT Supreet Singh Bahga bahga@mech.iitd.ac.in

Solving inverse problems to estimate reservoir properties by matching the predictions of reservoir models to the production data is called history matching.

The traditional method employed in industry to perform history matching is based on manual selection of model parameters to match production history [4]. The manual approach requires experience and often becomes unfeasible for complex models that involve a significant number of uncertain model parameters. Nowadays, the process of history matching has been automated [5], wherein the inverse problems are solved to estimate the model parameters by using the production data. Various methods to solve the inverse problems can be broadly classified into (i) deterministic methods and (ii) stochastic methods. A particular method can be classified as deterministic or stochastic depending upon the type of objective function of the inverse problem. The deterministic methods predict single values of model parameters such that the mismatch between predicted and observed data is minimized [6]. On the other hand, stochastic methods model the uncertainties in model parameters with appropriate probability distributions and predict their estimated values with associated uncertainty. The deterministic methods such as the gradient descent method require the computation of derivatives, which involves an enormous programming effort. Though single values of model parameters, as estimated by deterministic methods may be useful, they are not sufficient to manage a reservoir as they do not provide an estimation of risks. Also for many parameters and non-linear complex models, gradient methods tend to obtain a local minimum instead of finding a global minimum.

Various algorithms have been developed and successfully implemented for history matching in a petroleum reservoir. Genetic algorithms [7] and Evolutionary strategies [8] are some state-of-the-art techniques that have been reported in the literature to perform history matching. While the Genetic algorithms honour the complex geological constraints of the reservoir model, they are slow in convergence to obtain a global minimum and require a large number of deterministic simulations for evaluation of the forward model. To avoid the slow rate of convergence of the Genetic algorithms and the Evolutionary strategies, Bayesian inference (BI) based approaches have been developed in petroleum reservoir history matching. In BI, one starts with a prior distribution of model parameters, which is updated using the Bayes' rule, taking into account the production data, to get a posterior distribution of the model parameters [9]. The most general method to implement BI is through the MCMC method, which is applicable to Gaussian and non-Gaussian prior distributions. However, due to the prohibitive computational cost of the MCMC method, computationally efficient approaches of performing Bayesian inference, such as the Ensemble Kalman Filter (ENKF) [4,10], have been developed and such methods are based on various simplifying assumptions. For example, the ENKF method implements BI by assuming Gaussian priors [11]. In certain cases, however, ENKF leads to an overestimation of uncertainty in the production forecasts [12]. Also, few other modified form of ENKF have been used to perform history matching, these techniques are derived from Bayesian inference and are formulated based on the linear or Gaussian approximations. A detailed review of the various methods used in literature to perform history matching has been done by Oliver et al. [8] with the limitations of each method. Bayesian inference gives a quantitative assessment of uncertainty by providing a posterior probability distribution of model parameters instead of providing a single value of model parameters. The posterior distribution can be used to estimate mean, standard deviation, and other higher-order moments of the model parameters [13].

The existing Bayesian inference methods to solve inverse problems have a wide range of applications such as in heat transfer [14], climate modelling [15], and geophysics [16]. The most commonly used methods to sample from the posterior distribution are the sequential Monte Carlo approach [17] and the Markov Chain Monte Carlo (MCMC) method [18]. The Bayesian inference along with MCMC is considered to be the standard method for history matching in reservoir modelling [12,19] but it cannot be used directly for heavy reservoir simulation models due to the requirement of a large number of simulations which will be computationally expensive. In the context of reservoir simulation, various studies have been done to show the use of the MCMC method to generate realizations of the permeability fields [18]. However, MCMC methods involve a large number of simulation runs to sample the posterior distribution. In practical circumstances, where every single assessment of a reservoir model is computationally expensive, this limits the applicability of Bayesian inference to obtain a global convergence for history matching. While faster methods such as the Randomized maximum likelihood (RML), can also be used to sample from the posterior distribution, it is an approximation to MCMC. RML is based on the assumption that the data is linearly related to the model. Moreover, given the fact that evaluations with PCE are extremely fast, the use of MCMC along with PCE does not significantly increase the computational overhead.

The drawback of high computational cost in Bayesian inference using MCMC can be overcome by using proxy models to accelerate model predictions. One such method to generate a proxy model to solve inverse problems is the experimental design (ED) method [20]. In the ED method, a response surface is fit to the model predictions from a smaller set of deterministic simulations, and this response surface acts as a proxy for the reservoir model. The drawback of the ED method is that it does not take into consideration the probability distribution of model parameters but gives equal weight to all deterministic simulations, making it less reliable to act as a proxy model [20]. Moreover, as shown by Lawal [21], the use of ED methods for proxy modelling in the petroleum industry can violate fundamental physics and do not preserve the correlation between otherwise related variables.

The limitations mentioned above of the ED method have been overcome using the polynomial chaos expansion (PCE) method, in which a stochastic process is expressed by a series of orthogonal polynomials of random variables [22]. In the PCE method, the model parameters are considered to be uncertain random variables. New stochastic dimensions are used for each of the uncertain parameters to define their probability density functions. PCEs are used to characterize the dependence of model predictions on uncertain model parameters through these stochastic dimensions. The coefficients of PCEs can be obtained by projecting multiple solutions of the deterministic simulations onto the polynomial chaos basis [20,23]. The PCE method to form a proxy model guarantees convergence as the degree of the polynomials in the PCE proxy is increased [7]. While Bayesian Linear Regression can also be used, the coefficients of PCE easily provide uncertainty in model parameters and help us to identify individual and coupled contributions of the uncertainty in model parameters to the total uncertainty in model predictions [24]. While Gaussian processes [25] proxy models are also widely used for the development of surrogate models, they are based on the Gaussian hypothesis and generally do not utilize prior distribution information while PCE-based proxy models are based on orthonormal polynomials which are chosen based on prior distribution and are more computationally efficient [26]. For

an accurate sampling of the posterior, we have used MCMC which is the most general method for sampling and does not require approximations for distributions. Hence, it can be used as a proxy model in petroleum engineering. While PCE method has been used as a proxy model in various field such as structural mechanics [27], fluid flow [28], reacting flows [24,29] and petroleum engineering [30,31] for uncertainty quantification, their potential for accelerating Bayesian inference in petroleum engineering has not been fully leveraged.

Marzouk et al. [9] have assessed the utility of the PCE method to solve the inverse transient diffusion problem to estimate source location using Bayesian inference. In petroleum engineering, Bazargan et al. [32] have shown the feasibility of PCE-based proxy model for estimating heterogeneous reservoir permeability using Bayesian inference. So far, Bayesian inference has been used to predict various reservoir parameters such as porosity and permeability [4,33]. In the current work, we show that PCE-based proxy model can be used to accelerate Bayesian inference to estimate a variety of model parameters such as geological properties like porosity and permeability and the flow properties like the exponents of the relative permeability curves of oil and water. We also show the utility of the PCE method to determine sensitivities of model predictions to model parameters, which helps us in choosing the relevant production data to predict the model parameters accurately. The sensitivity analysis of model predictions to the model parameters also helps us in deciding the parameters that need to be measured with more accuracy. Hence, it is important to note that besides accelerating the Bayesian inference method, the PCE proxy also helps in finding the sensitive parameters to perform history matching. Moreover, the inverse problems that we consider in the current work are new and these can be used as benchmark problem for Bayesian inference based history matching.

The novelty of our work is to accelerate forward computation to perform history matching in petroleum reservoirs rather than solving the full deterministic models. To achieve the same, we have used Bayesian inference with MCMC along with PCE. We have used a proxy model using PCE to directly obtain the forward model instead of running the deterministic simulations. Our approach is different from other methods in which we can draw millions of samples from the posterior distribution and obtain the forward model efficiently to calculate the likelihood using the PCE proxy. We have also shown the convergence of the moments in a model problem obtained from the PCE proxy with the moments obtained from the standard Monte Carlo method. We have proposed a general framework that can be used for any reservoir model irrespective of the probability distribution of the model parameters. We have shown the applicability of our framework for various reservoir models, including a reservoir scale field-based model PUNQ-S3.

In the present work, we show the application of Bayesian inference to predict model parameters using three different reservoir models. In the first problem, we estimate the value of permeability and porosity in the standard SPE1CASE2 problem of SPE's Comparative Solution Projects [34]. Using the PCE proxy, which we validate with MC simulations, we perform sensitivity analysis to show the effect of variability in porosity and permeability on the model predictions. We choose the most sensitive model parameters based on the sensitivity analysis and predict a better posterior estimate of porosity and permeability by solving an inverse problem using historical production data. For the second case, we choose a water flooding problem with a high degree of heterogeneity to show the application of Bayesian inference to predict the exponents of relative permeability curves of oil and

water. Finally, for the third case, we choose a large dimensional real field-based reservoir model PUNQ-S3 with 26 uncertain model parameters.

We start by introducing the governing equations for the reservoir simulation models in Section 2. Next, we describe the history matching methodology using Bayesian inference followed by the stochastic methodology to form a PCE proxy. In Section 3, we then explain the deterministic simulations of the reservoir models. Next, we generate the PCE-based proxy and validate the proxy with Monte Carlo simulations. Using the PCE proxy, we perform a sensitivity analysis to estimate the most sensitive model parameters. Finally, we present history matching to predict those uncertain model parameters on which the model predictions are most sensitive. At last, we conclude our work in Section 4.

2. Methodology

In this section, we discuss the Bayesian inference approach using the PCE proxy for parameter estimation in the reservoir simulation model. We show the application of the method to three different three-dimensional reservoir models, which we discuss in detail in Section 3. First we review the governing equations used in reservoir simulations. Thereafter, we discuss the application of Bayesian inference method for history matching, followed by a discussion on the use of PCE proxy to accelerate Bayesian inference method.

2.1. Governing equations for fluid flow

The governing equations for fluid flow in reservoir simulations are given by the black-oil model. This model considers three phases, oil, water, and gas. The oil and gas are two hydrocarbon components, and the third component being water. The model assumes that no mass transfer takes place between the water phase and the other two phases. Also, the oil component is considered to have low volatility; hence, it remains in the oil phase. However, the gas is assumed to be miscible in oil and immiscible in water. Gas miscibility in oil depends on pressure according to a specified phase behaviour. The governing equations of black-oil model consist of the following mass conservation equations for gas, oil, and water components [35],

$$\frac{\partial(\phi\rho_g S_g + \phi\rho_{go} S_o)}{\partial t} + \nabla \cdot (\rho_g \boldsymbol{v_g} + \rho_{go}\boldsymbol{v_o}) = q_g, \tag{1}$$

$$\frac{\partial(\phi\rho_o S_o)}{\partial t} + \nabla \cdot (\rho_o \boldsymbol{v_o}) = q_o, \tag{2}$$

$$\frac{\partial(\phi\rho_w S_w)}{\partial t} + \nabla \cdot (\rho_w \boldsymbol{v_w}) = q_w. \tag{3}$$

Here ρ denotes the density, $\boldsymbol{v}$ the velocity, q is the source or sink term which correspond to injection and production wells. ϕ is the porosity of the reservoir and S denotes the saturation of various phases. In Equations (1)–(3), $\nabla\cdot$ is the divergence operator in three spatial dimensions (x, y, z) and t refers to time. In these equations, the subscripts g, o, and w refer to the gas, oil, and water phases respectively and go refers to the gas dissolved in oil phase. The saturations of gas, oil and water phases are given by,

$$S_g + S_o + S_w = 1. \tag{4}$$

The flow velocities of various phases in Equations (1)–(3) are given by the Darcy's law,

$$\boldsymbol{v_i} = -\frac{Kk_{ri}}{\mu_i}(\nabla p_i - \rho_g \boldsymbol{g} \nabla z), \quad i = g, o, w. \tag{5}$$

In Equation (5), K is the absolute permeability, k_{ri} is the relative permeability of the i-th phase, p is the pressure, $\boldsymbol{g}$ is the acceleration due to gravity term, z is the depth of the reservoir, and μ is the viscosity. The individual phase pressures p are related to two independent capillary pressures inside a reservoir which are defined as,

$$p_{c_{ow}} = p_o - p_w, \tag{6}$$

$$p_{c_{go}} = p_g - p_o. \tag{7}$$

The third capillary pressure is a combination of the other two. That is,

$$p_{c_{gw}} = p_g - p_w = p_{c_{go}} + p_{c_{ow}}. \tag{8}$$

The capillary pressures depend on saturation and are determined experimentally. In our simulations, the capillary pressure is neglected and hence, $p_o = p_w$ and $p_g = p_o$. The relative permeability k_{ri} also depends on saturation and is given by various models or calculated empirically. One such model is the Brooks-Corey [36] model. Using the Brooks–Corey model, the relative permeabilities of water (k_{rw}) and oil (k_{row}) are given by

$$k_{rw}(S_w) = k^o_{rw}(S_{wn})^{N_w}, \tag{9}$$

$$k_{row}(S_w) = (1 - S_{wn})^{N_{ow}}, \tag{10}$$

respectively, where

$$S_{wn} = S_{wn}(S_w) = \frac{(S_w - S_{wir})}{(1 - S_{wir} - S_{orw})}. \tag{11}$$

Here, S_w is the water saturation, S_{wir} is the irreducible water saturation, S_{orw} is the irreducible oil saturation, k^o_{rw} is the endpoint of water relative permeability curve, and, N_w and N_{ow} are empirical constants.

Equations (1)–(7) are a set of nine equations using which nine variables S_g, S_o, S_w, v_g, v_o, v_w, p_g, p_o and p_w can be solved along with the prescribed initial and boundary conditions. For all our simulations, we use 'Flow' simulator of Open Porous Media (OPM) [1] to solve the deterministic reservoir simulation equations. 'Flow ' is an open-source, black-oil reservoir simulator, which is a fully implicit and capable of running industry-standard reservoir simulation models. We have compared the simulation output of 'Flow' simulator with commercially available 'Eclipse' software, and the results match well as shown in Figure A2 in Appendix 2.

2.2. History matching using Bayesian inference

The parameters of the black-oil model are often unknown and are estimated by solving inverse problem knowing the production history. In the present work, we use Bayesian

inference based history matching to estimate the model parameters. To illustrate the use of Bayesian inference for history matching, we here consider a forward problem,

$$\boldsymbol{d} = G(\boldsymbol{Z}). \tag{12}$$

Here, the function G represents the reservoir model, $\boldsymbol{d}$ denotes the predicted value of observable data such as flow rates and well-pressures obtained from the reservoir model corresponding to model parameters $\boldsymbol{Z}$ such as porosity, permeability, etc. In the Bayesian framework, the unknown model parameters $\boldsymbol{Z}$ are considered to be random variables, and probability is used to describe the parameter's true values. The unknown model parameters are assumed to have a prior probability density of $\pi(\boldsymbol{Z})$ before observing the data $\boldsymbol{d}$. The a priori information of unknown model parameters is based on available geological or experimental data, physical limits of the model parameters. For example, a prior distribution of porosity can be assumed knowing that porosity lie between 0 and 1 and taking into account well logs or reservoir cores. Bayesian inference entails prediction of posterior probability density of model parameters $\pi(\boldsymbol{Z}|\boldsymbol{d}_{obs})$ taking into account the observed data $\boldsymbol{d}_{obs}$. The observed data $\boldsymbol{d}_{obs}$ differs from the predicted data $\boldsymbol{d}$ due to measurement uncertainties. The measurement uncertainties are assumed to be additive,

$$\boldsymbol{d_{obs}} = \boldsymbol{d} + \boldsymbol{e} = G(\boldsymbol{Z}) + \boldsymbol{e}, \tag{13}$$

where $\boldsymbol{e}$ denotes measurement errors that are independent and identically distributed. Therefore, the probability density of observed data given the model parameters is $\pi(\boldsymbol{d}_{obs}|\boldsymbol{Z}) = \pi_e(\boldsymbol{d}_{obs} - G(\boldsymbol{Z}))$, where π_e is the probability density of measurement error; $\pi(\boldsymbol{d}_{obs}|\boldsymbol{Z})$ is called the likelihood function. The measurement errors for different observations can be assumed to be uncorrelated and hence the likelihood function can be taken as

$$\pi(\boldsymbol{d}_{obs}|\boldsymbol{Z}) = \prod_i \varphi_i(d_{obs,i} - d_i). \tag{14}$$

Here, φ_i denotes the probability density of i-th component of $\boldsymbol{e}$, and d_i and $d_{obs,i}$ respectively the i-th component of predicted and observed data. Usually, the measurement errors are Gaussian and independently distributed, therefore likelihood function can be simplified to

$$\pi(\boldsymbol{d}_{obs}|\boldsymbol{Z}) \propto \prod_i \exp\left(-\frac{1}{2}\sum \frac{(d_{obs,i} - d_i)^2}{\sigma^2}\right). \tag{15}$$

Here σ is the standard deviation of the measurement noise.

The inverse problem of history matching deals with the calculation of the probability density of model parameters given observed data $\pi(\boldsymbol{Z}|\boldsymbol{d}_{obs})$. To this end, we employ the Bayes' rule

$$\pi(\boldsymbol{Z}|\boldsymbol{d}_{obs}) = \frac{\pi(\boldsymbol{d}_{obs}|\boldsymbol{Z})\pi(\boldsymbol{Z})}{\int \pi(\boldsymbol{Z}|\boldsymbol{d}_{obs})\pi(\boldsymbol{Z})d\boldsymbol{Z}}. \tag{16}$$

We use the Metropolis-Hastings algorithm [37] to sample the posterior density $\pi(\boldsymbol{Z}|\boldsymbol{d}_{obs})$. The Metropolis-Hastings algorithm is a Markov chain Monte Carlo (MCMC) method based on the principle that the sampling is done from a Markov chain in such a way that the sampled distribution converges to the target distribution at equilibrium state. In the

MCMC method, the sampling is initiated with an initial value of $\boldsymbol{Z^o}$. Then, the next value $\boldsymbol{Z^*}$ is sampled using a proper proposal distribution, which is a conditional distribution on the current state. Thereafter, the ratio of the probability of the new proposed sample to the probability of the current state is calculated as

$$\alpha = \frac{\pi(\boldsymbol{Z^*})}{\pi(\boldsymbol{Z^{t-1}})}. \tag{17}$$

If the new sample proposed is more likely than the existing sample ($\alpha > 1$), the sample is accepted, else the sample is accepted with probability α. This procedure is repeated until a stationary state is achieved, and the Markov Chain approaches the target distribution after a sufficient burn-in period.

2.3. The PCE-based proxy for acceleration of Bayesian inference

The MCMC method discussed in the previous section involves a large number of computations of the forward model (Equation (12)), which makes this method computationally expensive. In the current work, we use PCE-based proxy model to accelerate the MCMC method. We first obtain the PCE-based proxy model for the exact forward problem, for a prior distribution of model parameters $\boldsymbol{Z}$. Thereafter, the PCE proxy is used in the MCMC method, instead of solving the exact forward problem, while sampling from the posterior distribution. The use of the PCE proxy enables sampling from the posterior distribution in significantly less time.

In the current work, the model parameters such as porosity, permeability, and the exponents of the relative permeability curves are considered to be uncertain random variables. To create a proxy model for the model predictions in terms of the uncertain model parameters, we use a PCE-based non-intrusive stochastic method [29]. First, we express the random variables in terms of additional stochastic dimensions. Next, deterministic simulations based on the black-oil model are performed for different samples of the random model parameters. The results of these deterministic simulations are used to construct the PCE proxy of model predictions such as oil production rate, gas production rate, and water production rate. These PCE proxies are used to accelerate the inverse problem by calculating the model predictions using its proxy, instead of solving the forward problem. These PCE proxies can also be used for finding sensitivity of model predictions to variability in model parameters. We now explain the mathematical formulation of the PCE-based proxy.

2.3.1. Mathematical formulation of the PCE proxy

We use a non-intrusive approach to form a PCE proxy for our model predictions. For a detailed discussion of the method, one can refer to Reagan et al. [29], which we briefly summarize here. In the PCE method, we propagate the variability in model parameters to model predictions by introducing additional stochastic dimensions $\boldsymbol{\xi} = [\xi_1, \xi_2, \xi_3, \ldots \xi_d]$, corresponding to d uncertain model parameters. $\xi \sim N(0, 1)$ is selected as a random variable with standard normal probability density function for normally distributed model parameters. Therefore, a normally distributed model parameter can be written in terms of the random variable ξ as

$$w = \mu_w + \sigma_w \xi, \tag{18}$$

and, a log-normally distributed model parameter w can be expressed in terms of ξ as

$$w = \exp(\mu_w + \sigma_w \xi) \tag{19}$$

where μ_w is the mean value of parameter w and σ_w is the standard deviation. The model predictions such as oil production rate, gas production rate and water production rate, denoted by f_i are expressed using PCEs [23] as

$$f_i(x, y, z, t, \boldsymbol{\xi}) = \sum_{k=0}^{\infty} f_{i,k}(x, y, z, t)\Psi_k(\boldsymbol{\xi}), \tag{20}$$

where, $f_{i,k}$ represents the coefficients of the i-th model predictions at that location and time and Ψ_k are the basis functions of the PCE. The choice of the basis functions depends on the nature of the distribution of random variables. The basis functions are orthogonal to each other with the probability density as the weighing function. In the present case, the random variables are assumed to follow normal and log-normal distribution. Therefore, the polynomials chosen are Hermite polynomials [29].

For one uncertain normally distributed parameter, the basis functions $\Psi_k = \psi_k$ are one-dimensional (1-D) Hermite polynomials.

$$\psi_0 = 1, \quad \psi_1 = \xi, \quad \psi_2 = \xi^2 - 1, \quad \psi_3 = \xi^3 - 3\xi. \tag{21}$$

For practical purpose, the infinite series in Equation (20) is truncated to a fixed order P which is the order of the highest polynomial required to construct the PCE. Therefore, for the case of one uncertain parameter, the total number of terms in PCE is $P + 1$. In general, for d dimensions and highest order of polynomial as p, the total number of terms in PCE in Equation (20) is $P + 1$ given as

$$P + 1 = \frac{(p + d)!}{p!d!}. \tag{22}$$

The PCE basis functions $\Psi_k(\boldsymbol{\xi})$ are d-dimensional Hermite polynomials which are the product of one dimensional Hermite polynomials. The Hermite polynomials involve multi-index m_k^i.

$$\Psi_k(\boldsymbol{\xi}) = \prod_{i=1}^{n} \psi_{m_k^i}(\xi_i). \tag{23}$$

The first few 2-D Hermite polynomials for two uncertain model parameters ($d = 2$), are given by

$$\begin{aligned}
\Psi_o(\boldsymbol{\xi}) &= \psi_o(\xi_1)\psi_o(\xi_2) = 1, \\
\Psi_1(\boldsymbol{\xi}) &= \psi_1(\xi_1)\psi_o(\xi_2) = \xi_1, \\
\Psi_2(\boldsymbol{\xi}) &= \psi_o(\xi_1)\psi_1(\xi_2) = \xi_2, \\
\Psi_3(\boldsymbol{\xi}) &= \psi_2(\xi_1)\psi_o(\xi_2) = \xi_1^2 - 1, \\
\Psi_4(\boldsymbol{\xi}) &= \psi_1(\xi_1)\psi_1(\xi_2) = \xi_1\xi_2, \\
\Psi_5(\boldsymbol{\xi}) &= \psi_o(\xi_1)\psi_2(\xi_2) = \xi_2^2 - 1.
\end{aligned} \tag{24}$$

We compute the coefficients $f_{i,k}$ of the corresponding PCE in Equation (20) to describe how the model predictions are dependent on the uncertain model parameters. To this end, we truncate Equation (20) to P terms and estimate the coefficients $f_{i,k}$ using the least squares regression method [23]. In this method, $\boldsymbol{\xi}^{(j)}, j = 1, \ldots, N$ realizations are chosen using Smolyak quadrature [38] on which deterministic calculations are performed. The PCE coefficients can be obtained with reasonable accuracy using fewer quadrature points on Smolyak grid even for high dimensional problems. For example, if the number of uncertain parameters is $d = 6$, the number of deterministic models required is 85, for $d = 7$, the number of deterministic models required is 113, for $d = 8$, the number of deterministic models required is 145, and for $d = 9$, 181 deterministic simulations are required to obtain the PCE coefficients with comparable accuracy. Thereafter the coefficients $f_{i,k}$ are determined from

$$f_i(x, y, z, t, \boldsymbol{\xi}^{(j)}) = \sum_{k=0}^{P} f_{i,k}(x, y, z, t)\Psi_k(\boldsymbol{\xi}^{(j)}), \tag{25}$$

where $j = 1, \ldots, N$, to minimize the sum of residuals. Denoting the terms in Equation (25) as,

$$\boldsymbol{f} = \begin{bmatrix} f_i(x, y, z, t, \boldsymbol{\xi}^{(1)}) \\ \vdots \\ f_i(x, y, z, t, \boldsymbol{\xi}^{(N)}) \end{bmatrix}, \hat{\boldsymbol{f}} = \begin{bmatrix} f_{i,o}(x, y, z, t) \\ \vdots \\ f_{i,P}(x, y, z, t) \end{bmatrix}$$
$$\text{and} \quad \boldsymbol{Z} = \begin{bmatrix} \psi_o(\boldsymbol{\xi}^{(1)}) & \ldots & \psi_P(\boldsymbol{\xi}^{(1)}) \\ \vdots & \ddots & \vdots \\ \psi_o(\boldsymbol{\xi}^{(N)}) & \ldots & \psi_P(\boldsymbol{\xi}^{(N)}) \end{bmatrix}, \tag{26}$$

Equation (25) can be written in a compact matrix notation as

$$\boldsymbol{f} = \boldsymbol{Z}\hat{\boldsymbol{f}}. \tag{27}$$

The coefficients are computed using the least squares approach by

$$\hat{\boldsymbol{f}} = (\boldsymbol{Z}^{\boldsymbol{T}}\boldsymbol{Z})^{-\mathbf{1}}\boldsymbol{Z}^{\boldsymbol{T}}\boldsymbol{f}. \tag{28}$$

Solving for the coefficients of PCE yields the proxy model for the actual reservoir model. Thereafter, we use the proxy model in Equation (15) to predict the production variable d_i, which accelerates the solution of inverse problem. In addition, knowing the PCEs we can also compute the various statistical moments such as mean and standard deviation in the model predictions. The mean is given by the zeroth-order coefficient $\hat{f}_{i,0}$ because the expectation $\langle \Psi_k \rangle = 0$ for $k > 0$ and the standard deviation is calculated using the higher-order coefficients. The standard deviation of i-th model prediction is calculated by,

$$\sigma_i^2 = \langle (\hat{f}_i - \langle \hat{f}_i \rangle)^2 \rangle = \sum_{k=1}^{P} \hat{f}_{i,k}^2 \langle \Psi_k^2 \rangle. \tag{29}$$

On grouping the terms in Equation (29) corresponding to the same stochastic dimension, the contribution of each uncertain model parameter to the overall variability in model predictions can be acquired.

3. Results and discussion

In this section, we present the application of the accelerated Bayesian inference method using the PCE proxy on three model problems. First, we demonstrate the application of the method for history matching on a three-dimensional, three layered black-oil reservoir model to predict the porosity and absolute permeability. Thereafter, we present history matching of a three-dimensional, fifteen layered water flooding black-oil model to predict the exponents of relative permeability curves of oil and water. Finally, we demonstrate the applicability of our framework on a large-dimensional real field-based PUNQ-S3 reservoir model to perform history matching and predict 26 model parameters.

3.1. Case 1. SPE1CASE2 reservoir model

First, we discuss the reservoir model on which we demonstrate the application of PCE-based Bayesian inference to predict the values of porosity and permeability. As a model problem, we consider the standard SPE1CASE2 model of SPE's Comparative Solution Projects [34], which is a three-dimensional black-oil reservoir simulation model. The geometry of the reservoir, the stratification, and reservoir properties are illustrated in Figure 1. The reservoir is discretized into 10×10 grid blocks in x and y directions. All the grid blocks are of 1000×1000 feet in x and y directions. The reservoir is 100 feet deep in the z-direction. There are three layers in the z-direction and each layer represents a grid block. The top, middle and bottom layers are 20, 30 and 50 feet deep respectively. All the layers have the same porosity but distinct permeability. There is one injection well located at grid block $(1, 1, 1)$ and one production well at $(10, 10, 3)$ as shown in Figure 1.

Initially, the reservoir is under-saturated with gas. The initial oil and water saturations are 0.88 and 0.12, respectively. Gas is injected at the injection well at a constant injection rate of 100 mmscf/day. Oil is produced at the production well at a constraint of maximum production rate of 20,000 stb/day, and the minimum flowing bottom hole pressure for the production well is 1000 psia. All reservoir boundaries, except the wells, are impermeable.

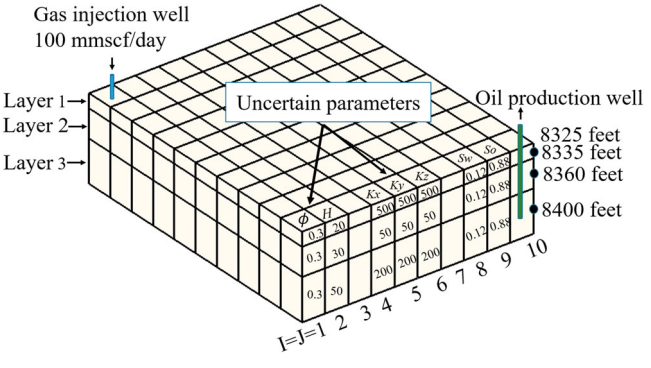

Figure 1. Reservoir geometry showing the stratification and reservoir properties for the SPE1CASE2 problem. The values of porosity and permeability at each grid block for the standard SPE1CASE2 problem are shown in the figure. In the current work, we consider the values of porosity of the reservoir and permeability of the top layer as uncertain and estimate them using history matching.

The initial reservoir pressure is 4800 psia at a depth of 8400 feet. The simulations are run for a period of 10 years, and data are saved at an interval of one month.

3.2. Case 1. SPE1CASE2 problem description

To perform history matching, it is important to understand the flow physics of the problem and the parameters on which the simulation predictions are more sensitive. Therefore, we first run a deterministic simulation on the reservoir model to present a brief discussion of the physics of the problem. For this deterministic simulation, we took the permeabilities of top, middle, and bottom layers as 500 mD, 50 mD, and 200 mD, respectively. The porosity of each layer is 0.3.

The phase behaviour properties for the simulation are taken from the standard problem [34]. The gas is injected at a constant rate of 100 mmscf/day in the injection well, and oil is produced at a maximum oil production rate (OPR) of 20,000 stb/day. Figure 2(a) shows the deterministic OPR ($\sigma = 0$) for a time period of 10 years. As the gas is injected at a constant rate of 100 mmscf/day into the injection well, the oil is displaced horizontally towards the production well. When the oil reaches the production well, the production of oil starts at the maximum rate of 20,000 stb/day for a time period of 52 months. As the oil production takes place from the reservoir, the pressure inside the reservoir continuously falls and reaches the minimum flowing bottom hole pressure of 1000 psia. Due to the fall in pressure inside the reservoir, the OPR starts decreasing from its maximum rate. Meanwhile, the injected gas also reaches the other end of the top layer, and due to continuous injection, it moves towards the bottom layer and reaches the production well at about $t = 43$ months. Figure 2(b) shows the gas production rate (GPR) for the deterministic simulation. Along with the production of oil, concurrently, gas is also produced from the production well. Initially, the gas that is dissolved in the reservoir oil is produced, but once the injected gas reaches the production well, the GPR increases suddenly at $t = 43$ months, as shown in Figure 2(b).

3.3. Validation of PCE method with Monte Carlo simulations

After getting an insight into the physics of the fluid flow in the reservoir, we first validate our PCE methodology by comparing it with Monte Carlo (MC) simulations. For that, we solved the forward problem with a prior distribution of porosity and layer permeabilities. For our calculations, we assumed that the porosity of the whole reservoir and permeability of each layer are independent, log-normally distributed random variables. The mean value of the porosity of each layer was assumed to be 0.3, and the permeability of the top, middle, and bottom layers were taken as 500 mD, 50 mD, and 200 mD respectively as labelled in Figure 1. We assumed a 20% standard deviation around the mean values of porosity and permeability of each layer of the reservoir. First, we performed MC simulations and calculated the mean and standard deviation of model predictions. To obtain convergence of the MC method, we required 10,000 deterministic simulation runs. Next, we used the PCE method to find the mean and standard deviation of model predictions and compared the two. We represented the model predictions using fourth-order PCEs which were sufficiently accurate in describing the effect of variability of the model parameters on the model predictions. The fourth-order PCE for the four random variables consisted of 70

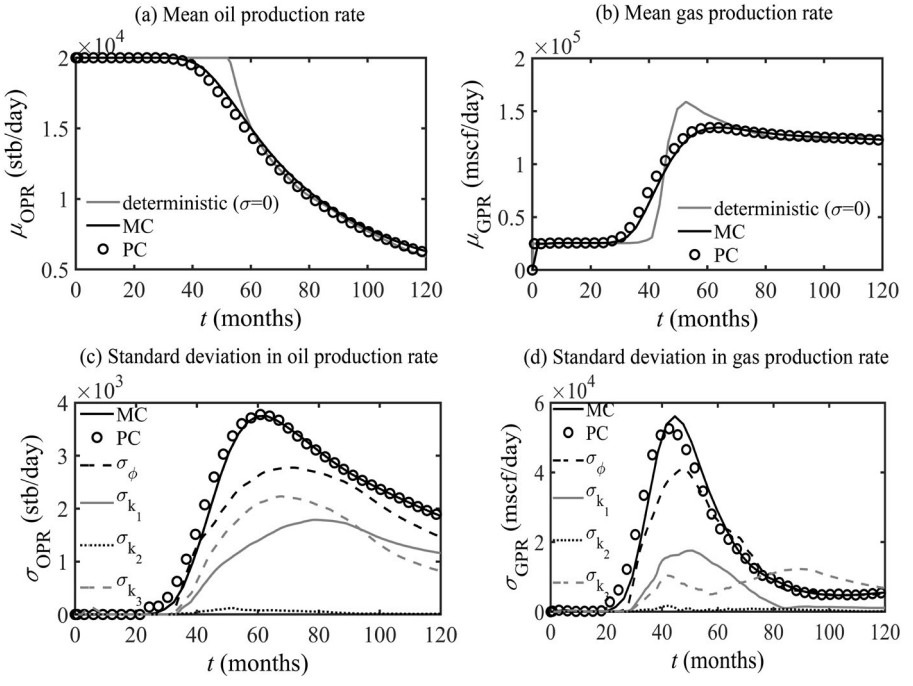

Figure 2. Comparison of mean and standard deviation in production variables predicted by 10,000 MC simulations and 385 PCE simulations for SPE1CASE2 problem. (a) and (b) show the mean values of OPR and GPR versus time. (c) and (d) show the standard deviation in OPR and GPR versus time. The mean and standard deviation converge well for fourth-order PCE simulations with 385 collocation points for four uncertain parameters with 10,000 MC simulations. The contribution of individual uncertainties in the model parameters are also shown in (c) and (d). The effect of porosity and permeability is shown, where ϕ is the porosity, k_1, k_2, and k_3 are the permeability of the top, middle and bottom layers respectively. OPR is most sensitive to porosity followed by the permeability of bottom-most layer, topmost layer and least sensitive to the middle layer. GPR is most sensitive to porosity followed by permeability of topmost layer, bottom-most layer and least sensitive to the middle layer.

coefficients. The computation of these 70 coefficients required 385 simulations using the stochastic collocation method based on Smolyak quadrature [38].

Figure 2(a,b) shows the mean values of OPR and GPR with time. Figure 2(c,d) shows the variation of the standard deviation of OPR and GPR with time. These figures show that the results obtained from fourth-order PCE for 385 deterministic simulations match well with 10,000 MC simulations. Therefore, PCE can be used as a proxy model in reservoir simulations. We also observe in Figure 2(a ,b) that the mean values of production show more dispersion in contrast to the sudden change in rates in the deterministic simulation. Due to the effect of variability, the discontinuities disappear in the mean production profiles. This is in agreement with the simulation results of Zhang and Tchelepi [39], who showed similar dispersion in mean production profiles for the Buckley–Leverett problem.

3.4. Sensitivity analysis of simulation predictions to model parameters

Having validated the PCE method with MC simulations, we present a sensitivity analysis of model predictions to model parameters. In Figure 2(c), we show the standard deviation

in the OPR due to uncertainty in porosity and layer permeabilities. In the present problem, the OPR is most sensitive to the porosity of the reservoir. This is expected because the production potential directly depends on the total hydrocarbon content in the reservoir which depends on the porosity. After porosity, the OPR is sensitive to permeability of the bottom-most layer, followed by the permeability of the topmost layer with an almost negligible contribution of permeability of the middle layer.

Figure 2(d) shows the sensitivity of the GPR to the porosity and permeability of each layer. Similar to the OPR, the GPR is also most sensitive to the porosity of the reservoir. In contrast to the OPR, after porosity, the GPR is more sensitive to the permeability of the topmost layer as compared to the bottom-most layer. The observations from sensitivity analysis suggest that, to predict the permeability of the bottom-most layer we need to perform history matching on the OPR, whereas to predict permeability of the topmost layer, we need to perform history matching on GPR. Also, the standard deviation of GPR is higher than OPR, which tells that GPR is more sensitive to model parameters than OPR.

3.5. History matching

We now present Bayesian inference along with PCE proxy for history matching. As shown by the sensitivity analysis, the GPR is more sensitive to the model parameters than the OPR. Hence, we took the GPR as the measured data to predict the posterior values of permeability and porosity. Also, the sensitivity analysis shows that the GPR is more sensitive to the porosity of the reservoir and permeability of the topmost layer. Therefore, we limit our inverse problem to predicting porosity and permeability of topmost layer. To get the synthetic measured data to perform history matching, we performed a deterministic simulation on a fixed value of porosity and permeability which we call as the true value and added noise to it. The simulation period to obtain the history data was for 10 years and simulation for another 10 years was performed to evaluate the prediction performance of the model. The model parameter values used to obtain the synthetic reference case of GPR are listed in Table 1. The GPR obtained using the true values of model parameters are shown in Figure 3. We added Gaussian noise with $\sigma = 5000$ mscf/day to the true values of GPR, as shown in Figure 3, and considered the resulting data as the production history. In the petroleum industry, measured data is not available at all times. Therefore to limit the data, we took the measured data at time intervals of five months. To get the proxy model of GPR for solving the inverse problem, we considered an uncorrelated prior probability distributions of porosity and permeability which were log-normally distributed with mean values of 0.3 and 500 mD, respectively, with a standard deviation of 20% around the mean values as shown in Figure 3. Based on this prior probability distribution, we first obtained the PCE proxy for the production forecast, i.e. GPR. We represented the model predictions using fourth-order PCEs which consisted of fifteen coefficients for the two random variables. The convergence of the PCE coefficients is presented in Figure A1 in Appendix A. We chose fifteen coefficients, which were sufficient to obtain the PCE proxy. The computation of these fifteen coefficients required 89 simulations using the stochastic collocation method based on Smolyak quadrature [38]. From Figure 3, we observe that there is a mismatch between the GPR obtained on the prior mean values on which the PCE proxy is developed and the measured data.

Table 1. The parameters used for history matching in the SPE1CASE2 model are shown. The measured data was obtained by adding noise to the true values. The mean and standard deviation of the prior and posterior distribution is shown here. Prior values were used to get the proxy model. The proxy model was used to calculate the posterior mean by solving the inverse problem using Bayesian inference.

			Prior			Posterior estimate	
Parameters	Layer	True value	mean	standard deviation	range	mean	standard deviation
Porosity	1 – 3	0.27	0.3	$\pm$ 0.06	0.12 – 0.66	0.26	$\pm$ 0.01
Permeability (mD)	1	450	500	$\pm$100	215 – 1114	436.28	$\pm$ 21.08

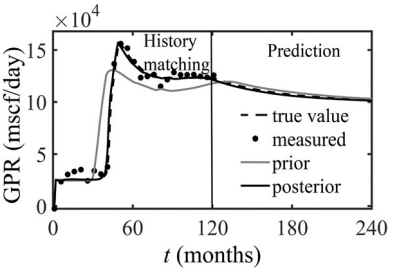

Figure 3. Comparison of predicted values of GPR versus time with the measured data and predictions based on the true values for SPE1CASE2 problem. The measured data was obtained by adding noise on GPR obtained using the true values for the first 10 years. The GPR predicted with prior mean values of model parameters does not match with the measured data. However, the temporal evaluation of GPR predicted by the posterior mean values shows good agreement with the model predictions based on measured data.

To obtain the posterior estimate of model parameters, we used the Metropolis-Hastings algorithm to get 50,000 samples from the posterior distribution of porosity and permeability. We used the PCE proxy for calculation of GPR at the sampled values of model parameters, which accelerated the sampling from the posterior distribution. Once we obtained the posterior distribution, we calculated the posterior mean by taking average of all the samples and then calculated the standard deviation of the model parameters. Table 1 shows the posterior mean estimate and the standard deviation of the model parameters. We get a closer estimate of the true values in posterior distribution, even though our prior estimates were not close to the true values. Figure 3 shows that the GPR obtained using the posterior mean values in the proxy model is very close to the measured data as compared to the GPR obtained on the prior values. Figure 3 also shows that the prediction performance for the next 10 years matches well with the true GPR.

Figure 4 shows the prior and posterior probability density of porosity and permeability obtained using the Bayesian inference method. Starting with a broader prior distribution of the model parameters, we obtained a narrower posterior distribution. The posterior values lie closer to the true values as compared to the prior values, as shown in Table 1. After history matching, the model parameter values converge to the true values and the predicted variances also decrease after history matching as shown in Figure 4 and its inset. The shape of the posterior distribution also gives information on how the porosity and permeability values must be tweaked to match model predictions with production data. From Figure 4, we observe that both porosity and permeability must increase or decrease simultaneously

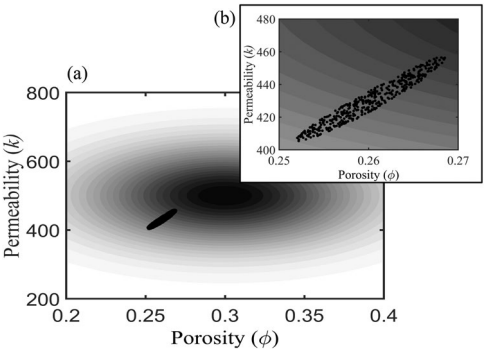

Figure 4. Comparison of prior and posterior probability density of porosity of reservoir and permeability of the topmost layer for SPE1CASE2 problem. (a) shows the prior probability of porosity of the reservoir and permeability of the topmost layer. The black colour corresponds to high values and the white colour corresponds to low values. The inset (b) shows the zoomed-in view of the samples obtained from the posterior distribution. Starting from a broad prior distribution, Bayesian inference results in a narrow distribution of permeability of the topmost layer and porosity of the reservoir.

to match the history data. This can be interpreted as follows. For a fixed value of porosity, when the permeability is increased the injected gas can flow easily through the reservoir, hence, the breakthrough of gas occurs earlier and more gas is produced. However, for a fixed permeability, if porosity is increased, it means that more oil is initially present in the reservoir. Hence, maximum oil production takes place from the reservoir for a longer time and the breakthrough of the gas is delayed. Increasing porosity, therefore leads to reduction in GPR. Therefore, to produce the same amount of gas, an earlier breakthrough must take place for which the permeability of the layer should be higher.

3.6. Case 2. SPE9 reservoir model and problem description

After successfully performing Bayesian inference to predict the porosity and permeability of a reservoir model, we now consider a history matching problem of estimating other parameters such as exponents of relative permeability curves of oil and water. To this end, we consider SPE9 [40] reservoir model, which is a model of moderate size (9000 grid blocks) and with a high degree of heterogeneity provided by a geostatically based permeability field. The reservoir geometry is shown in Figure 5. The reservoir model is based on a dipping reservoir with 25 randomly placed production wells and one water injection well placed at grid points $I = 24$, $J = 25$, and K value for the connecting grid blocks varies from 11 to 15. The cells dip in the x direction at an angle of 10°. The production wells are present in layers 2, 3 and 4 only. There are fifteen layers in the reservoir with constant porosity values in each layer, and a geostatically generated permeability field on a cell by cell basis. The simulation time to perform history matching is 900 days and the performance of the history matched model is evaluated for the next 900 days. The water injector is set to a maximum rate of 5000 stb/day with a maximum bottom hole pressure of 4000 psia. The maximum oil production rate for all production wells is set at 1500 stb/day, and the minimum flowing bottomhole pressure for all production wells is 1000 psia. At 300 days, the maximum oil production rate is lowered to 100 stb/day for all wells. At 360 days, the

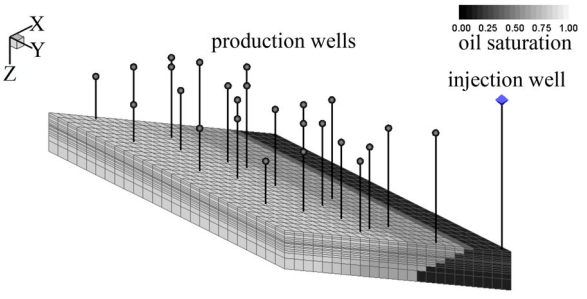

Figure 5. Geometry of water flooding SPE9 reservoir model along with the initial oil saturation. This problem involves a dipping reservoir with 25 randomly placed producing wells and one water injection well. Here we perform history matching on this reservoir model to estimate the values of exponents of the Corey model for the relative permeability curves of oil (N_{ow}) and water (N_w) by solving an inverse problem using history matching.

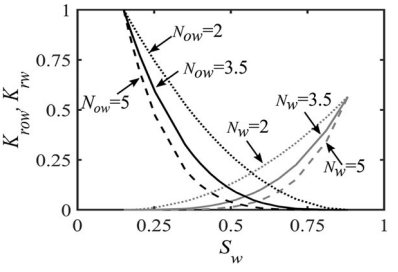

Figure 6. The dependence of relative permeability curves on the values of exponents of the Corey model given by Equations (9)–(11) for SPE9 problem. The decreasing curves correspond to the relative permeability of oil (K_{row}) for varying exponents N_{ow}. For a fixed value of water saturation (S_w), increasing the value of N_{ow} reduces the relative permeability of oil. The other set of curves correspond to the relative permeability of water (K_{rw}). Similar to oil, increasing the value of N_w reduces the relative permeability of water. For these curves, we used the values of connate water saturation to be 0.18 and residual oil saturation to be 0.88.

maximum oil production rate is again raised to 1500 stb/day for all wells until the end of simulation. The phase behaviour properties of the oil, water, and gas are taken from Killough et al. [40]. The relative permeability curves are given by Brooks-Corey correlation given by Equations (9) to (11).

In this inverse problem, we predict the exponents (N_{ow} and N_w) of Brooks–Corey correlations of the relative permeability curves of oil and water. Figure 6 shows the effect of varying the exponents of relative permeability curves of oil and water given by the Brooks-Corey correlations. For higher values of N_w in Equation (9), the relative permeability curve of water shifts to the right of the curve, as shown in Figure 6. Therefore, for the same saturation of water, the relative permeability of water decreases. This leads to a decrease in the flow rate of water, hence a decrease in water production rate. Likewise, an increase in the exponent N_{ow} of Equation (10), causes the relative permeability curve of oil to shift to the left of the curve, as shown in Figure 6. Hence, for the same saturation of water, the relative permeability of oil decreases with an increase in N_{ow}, which slows down the flow of oil and decreases the FOPR.

Table 2. The parameters used for history matching in the SPE9 model. The measured data was obtained by adding noise to the simulations performed using the true values. Prior values were used to get the proxy model. The mean and standard deviation of the prior values are shown here. The proxy model was used to calculate the posterior values by solving the inverse problem using Bayesian inference.

		Prior			Posterior estimate	
Parameters	True value	mean	standard deviation	range	mean	standard deviation
N_w	1.5	3.5	± 1.5	0.9 – 6	1.63	± 0.12
N_{ow}	1.5	3.5	± 1.5	0.9 – 6	1.56	± 0.29

3.7. Sensitivity of simulation predictions to model parameters

Before performing Bayesian inference, we look at the sensitivity of the model predictions to the model parameters. For our calculations, we considered both N_w and N_{ow} as uncorrelated normally distributed random variables with mean values of 3.5 and standard deviation of 1.5, respectively. For all our computations, we represented the model predictions using fourth-order PCEs. The fourth-order PCE consisted of fifteen coefficients for two random variables. The computation of these fifteen coefficients required 25 simulations using the stochastic collocation method based on Smolyak quadrature [38]. We obtained the mean and standard deviation for field oil production rate (FOPR) and field water production rate (FWPR).

Figure 7(a,b) shows the mean FOPR and mean FWPR, respectively. The mean profiles are obtained as per the control set on the production well as discussed in Section 3.6. Figure 7(c,d) shows the standard deviation in FOPR and FWPR, respectively. Similar to the first problem, the standard deviation in model predictions is highest at the time when the production rate is suddenly reduced at 300 days, as shown in Figure 7(c,d). From Figure 7(c), we can also observe that FOPR is more sensitive to N_{ow} whereas FWPR is more sensitive to N_w. This is because a change in exponent alters the relative permeability curve of that particular fluid. Therefore, to predict the value of the two constants, if we perform history matching on the FOPR, we expect to get a better estimate of N_{ow}, whereas if we perform history matching on the FWPR, we get a better prediction of N_w. Therefore, to predict accurate values for both the parameters, we take the measured data of both FOPR and FWPR for history matching.

3.8. History matching

As suggested by the sensitivity analysis, we chose production data of FOPR and FWPR for history matching. We chose a true value of N_w and N_{ow} as shown in Table 2. The FOPR and FWPR for true values of N_w and N_{ow} is shown in Figure 8(a,b) respectively. To get the measured data, we added Gaussian random noise with $\sigma = 1000$ stb/day for FOPR and $\sigma = 200$ stb/day for FWPR. The measured values of FOPR and FWPR are shown in Figure 8(a,b) respectively. We took the prior mean and standard deviation of N_w and N_{ow} as shown in Table 2. The FOPR and FWPR for prior values of N_w and N_{ow} is shown in Figure 8(a,b) respectively. We observe that the prior values of FOPR and FWPR do not match with the measured data. Based on the prior distribution of N_w and N_{ow}, we obtained

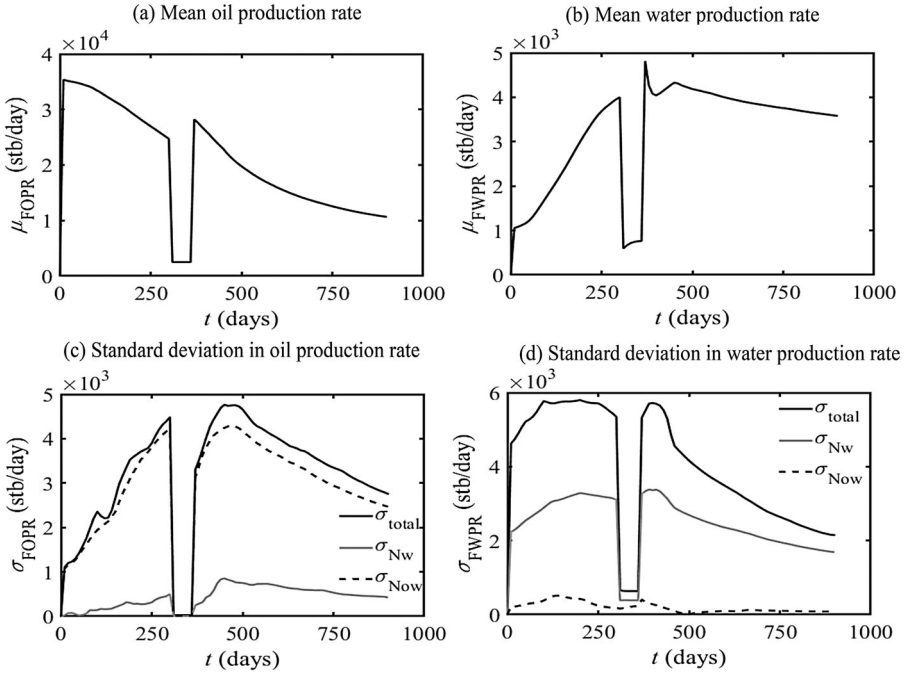

Figure 7. Sensitivity of FOPR and FWPR on N_w and N_{ow}. (a) and (b) show the mean values of FOPR and FWPR versus time. The oil and water production rate is obtained as per the boundary conditions set on the production wells. (c) shows the standard deviation in FOPR versus time and sensitivity of FOPR to N_w and N_{ow}. N_{ow} is the dominant factor to the overall variation in FOPR at all times. (d) shows standard deviation in FWPR versus time and sensitivity of FWPR to N_w and N_{ow}. N_w is the dominant factor to the overall variation in FWPR at all times.

the PCE proxy for the forward model as discussed in Section 2.2. Thereafter, we obtained 100,000 samples of the posterior distribution using MCMC, and predicted the exponents of the relative permeability curves. Data presented in Table 2 shows that the predicted posterior mean values of N_w and N_{ow} are very close to the true value. This shows the efficacy of the method in predicting various uncertain parameters. Figure 8 shows the posterior FOPR and FWPR obtained after solving the inverse problem for the history matching period and the prediction period. It shows that the FOPR and FWPR obtained on the posterior model parameters are very close to the measured history data and the predicted data matches well with the true data.

3.9. Case 3. PUNQ-S3 reservoir model and problem description

After successfully performing history matching on simplistic reservoir models, we now consider a real field-based reservoir model PUNQ-S3, operated by Elf Exploration Production [4]. We chose the PUNQ-S3 test case because the performance of various history matching techniques have been reported in literature for this model. This model contains $19 \times 28 \times 5$ grid blocks, out of which 1761 blocks are active. The reservoir field is bounded to the east and south by a fault and links to a strong aquifer in the north and west. The field requires no injection well because of the strong pressure support from the aquifer and six

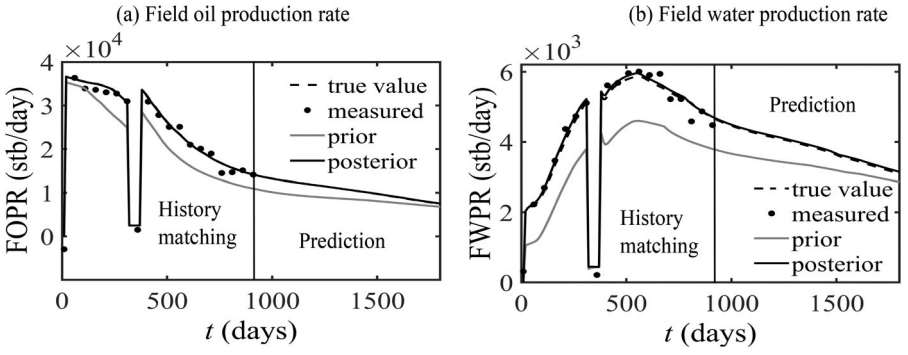

Figure 8. Comparison of predicted values of FOPR and FWPR versus time using prior mean and posterior mean values of model parameters with the measured data for SPE9 problem. The measured data were obtained by adding noise on the model predictions obtained using the true values. (a) shows that the FOPR predicted with prior mean values of model parameters does not match with the measured data. However, after history matching, the FOPR predicted by the posterior mean values shows good agreement with the model predictions based on the measured data. Also, the prediction data for the next 10 years match well with the true data. Similarly, (b) shows that the FWPR predicted with prior mean values of model parameters does not match with the measured data. However, after history matching, the FWPR predicted by the posterior mean values shows good agreement with the model predictions based on the measured data.

producer wells are present. The producer wells are present around the gas-oil contact in the centre of the reservoir as shown in Figure 9. In this test case, the porosity and permeability fields are taken from the original reservoir models. However, the relative permeability and capillary pressure curves are taken from a modified model of the PUNQ-S3 model presented by Zhang et al.[41]. In the modified model, the power-law model is used to represent the relative permeability and capillary pressure curves. In this model problem, the production schedule of the wells has been taken from the original model which involves a simulation period of 16.5 years. The production history of the first 8 years is used to predict the model parameters and predict the production for the next 8.5 years. The first 8 years include the first year of well testing, followed by 3 years of shut-in period and 4 years of the production period. The oil production rate of 150 m^3/day is fixed during the production period and a two-week shut-in period is applied each year to collect the shut-in pressure data.

The two-phase oil-water relative permeability and capillary pressure curves in the modified models are represented as

$$k_{rw}(S_w) = a_w \left(\frac{S_w - S_{wi}}{1 - S_{wi} - S_{orw}} \right)^{n_w}, \tag{30}$$

$$k_{row}(S_w) = a_{ow} \left(\frac{1 - S_w - S_{orw}}{1 - S_{wi} - S_{orw}} \right)^{n_{ow}}, \tag{31}$$

and

$$P_{cow}(S_w) = P^*_{cow} \left(\frac{1 - S_w - S_{orw}}{1 - S_{wi} - S_{orw}} \right)^{n_{cow}} + P_{ceow}. \tag{32}$$

Here, S_{wi} is the irreducible water saturation, S_{orw} is the irreducible oil saturation, a_w and a_{ow} are the endpoints of water relative permeability curve and oil relative permeability curve respectively. The exponents n_w and n_{ow} are empirical constants that denote the curvature of the relative permeability curve of water and oil respectively. In Equation (32), P_{ceow} is the entry capillary pressure of the oil-water system, n_{cow} is the exponent that determines the curvature and $P^*_{cow} + P_{ceow}$ is the capillary pressure at the irreducible water saturation.

Likewise, two-phase gas-oil relative permeability and capillary pressure curves can be expressed as

$$k_{rg}(S_g) = a_g \left(\frac{S_g - S_{gc}}{1 - S_{org} - S_{wi} - S_{gc}} \right)^{n_{rg}}, \tag{33}$$

$$k_{rog}(S_g) = a_{og} \left(\frac{1 - S_{org} - S_{wi} - S_g}{1 - S_{org} - S_{wi} - S_{gc}} \right)^{n_{rog}}, \tag{34}$$

and

$$P_{cog}(S_g) = P^*_{cog} \left(\frac{S_g - S_{gc}}{1 - S_{org} - S_{wi} - S_{gc}} \right)^{n_{cog}} + P_{ceog}. \tag{35}$$

Here, S_{gc} is the critical gas saturation, S_{org} is the residual oil saturation in the oil–gas system, a_g and a_{og} are the endpoint of the gas relative permeability and oil relative permeability

Table 3. The parameters used for history matching in the PUNQ-S3 model are shown. The measured data was obtained by adding noise to the true values. The mean and standard deviation of the prior and posterior distribution is shown here. Prior values were used to get the proxy model. The proxy model was used to calculate the posterior estimates by solving the inverse problem using Bayesian inference.

			Prior			Posterior estimate	
Parameters	Layer	True value	mean	standard deviation	range	mean	standard deviation
a_w	–	0.72	0.67	±0.067	0.56–0.77	0.73	±0.03
n_w	–	4.32	4	±0.4	2.8–3.8	4.5	±0.22
a_{ow}	–	0.93	0.93	±0.086	0.73–0.98	0.82	± 0.06
n_{ow}	–	4.32	4	±0.4	3.3–4.5	4.3	± 0.18
n_{cow}	–	4.32	4	±0.4	3.4–4.6	4.2	±0.03
a_g	–	0.32	0.3	±0.03	0.25-0.34	0.47	± 0.05
n_{rg}	–	4.32	4	±0.4	3.4–4.6	4.1	± 0.25
a_{og}	–	2.16	2	±0.2	1.7–2.3	1.8	±0.34
n_{rog}	–	4.32	4	±0.4	2.7–3.7	4.3	±0.05
n_{cog}	–	4.32	3	±0.4	2.5–3.4	4.5	±0.23
P_{ceog}	–	6.48	6	±0.6	5.1–6.9	6.2	±0.43
Porosity	1	1	0.8	±0.3	0.5–1.1	1.01	±0.01
Porosity	2	1	0.8	±0.3	0.5–1.1	0.94	±0.04
Porosity	3	1	0.8	±0.3	0.5–1.1	0.95	±0.07
Porosity	4	1	0.8	±0.3	0.5–1.1	0.95	±0.07
Porosity	5	1	0.8	±0.3	0.5–1.1	0.97	±0.02
x-Permeability	1	1	0.8	±0.3	0.5–1.1	0.97	±0.09
x-Permeability	2	1	0.8	±0.3	0.5–1.1	0.98	± 0.08
x-Permeability	3	1	0.8	±0.3	0.5–1.1	0.98	± 0.08
x-Permeability	4	1	0.8	±0.3	0.5–1.1	0.98	± 0.02
x-Permeability	5	1	0.8	±0.3	0.5–1.1	0.98	± 0.03
z-Permeability	1	1	0.8	±0.3	0.5–1.1	0.98	± 0.03
z-Permeability	2	1	0.8	±0.3	0.5–1.1	0.98	± 0.03
z-Permeability	3	1	0.8	±0.3	0.5–1.1	0.98	± 0.04
z-Permeability	4	1	0.8	±0.3	0.5–1.1	0.98	± 0.05
z-Permeability	5	1	0.8	±0.3	0.5–1.1	0.99	± 0.05

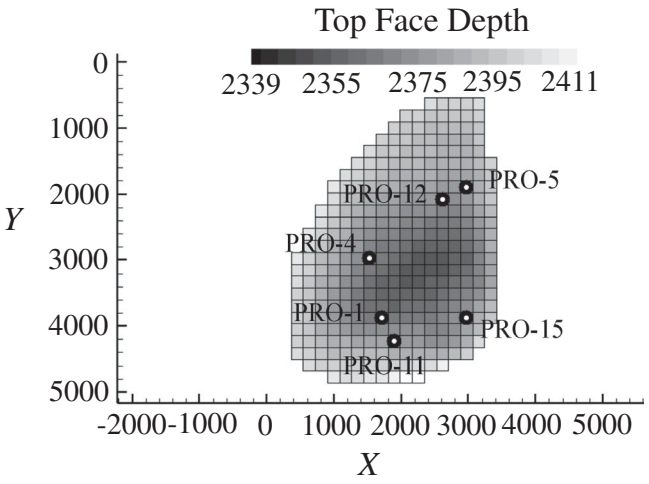

Figure 9. Geometry of the PUNQ-S3 reservoir model showing the well locations and the top structure depth (in metres).

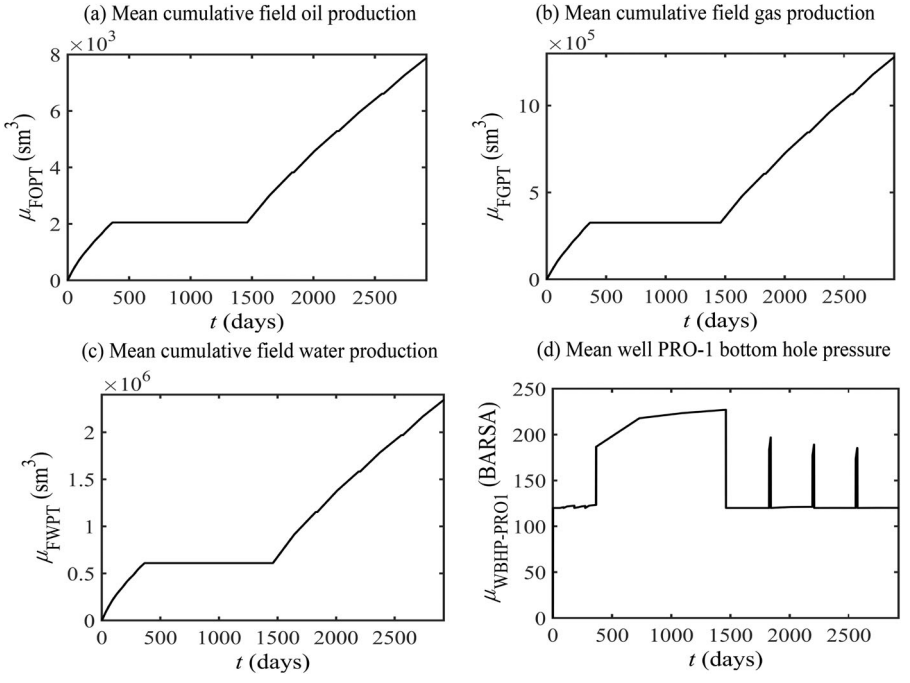

Figure 10. Mean values of the production parameters (a), (b) and (c) show the mean values of cumulative FOPT, FGPT and FWPT versus time for 8 years of production period. (d) show the mean bottom hole pressure of well PRO-1 with time during the 8 years of production period.

curve for the oil–gas system. The exponents n_{rg}, n_{rog} and n_{cog} determine the curvature of the curves. $P^*_{cog} + P_{ceog}$ is the capillary pressure at maximum gas saturation for the oil–gas system.

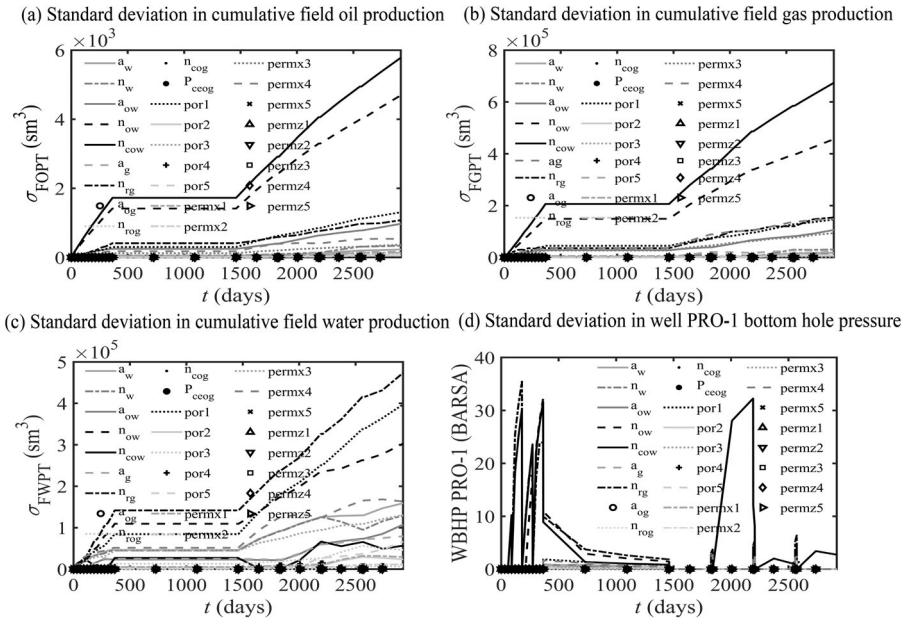

Figure 11. Sensitivity of FOPT, FGPT, FWPR and WBHP-PRO1 on the uncertain model parameters. (a), (b) and (c) show the standard deviation in FOPT, FGPT and FWPT versus time and their sensitivity to the model parameters. n_{cow} is the dominant factor to the overall variation in FOPT and FGPT at all times whereas FWPT is most sensitive to n_{rg}. (d) shows standard deviation in WBHP of well PRO-1 versus time and sensitivity of WBHP to the model parameters.

To perform history matching and predict model parameters in the PUNQ-S3 model we chose 26 uncertain model parameters. The first set of uncertain parameters are the constants in the relative permeability and capillary pressure curves of the oil-water and gas–oil system given in Equations (30)–(35). These constants are a_w, n_w, a_{ow}, n_{ow}, n_{cow}, a_g, n_{rg}, a_{og}, n_{rog}, n_{cog}, P_{ceog}. For the next set of model parameters, we used the porosity and permeability field of the original reservoir model for all 5 layers. To assume uncertainty in the porosity and permeability fields, we assumed porosity and permeability multipliers in x and z directions to be uncertain. So, we have fifteen uncertain parameters corresponding to the five porosity multipliers in the five layers, five permeability multipliers in the x direction and five permeability multipliers in the z direction for the five layers.

3.10. Sensitivity analysis of simulation predictions to model parameters

Before solving the inverse problem, we analyse the sensitivity of model predictions to the model parameters. The mean values and standard deviation of the normally distributed uncertain parameters are shown in Table 3. In this case, we represented the model parameters using first-order PCEs. The first-order PCE consisted of 27 coefficients for 26 random variables. The computation of these 27 coefficients required 53 simulations using the stochastic collocation method based on Smolyak quadrature. We obtained the mean and standard deviation for the cumulative field oil production (FOPT), cumulative field gas production (FGPT), cumulative field water production (FWPT)) and the bottom hole pressure of well PRO-1 (WBHP PRO-1).

Figure 10(a–d) shows the mean FOPT, FGPT, FWPT and mean BHP of well PRO-1, respectively. The mean profiles are obtained as per the control set on the production wells as discussed in Section 3.9. The simulation starts with 1 year of well testing period followed by 3 years of shut-in period where the cumulative production of oil, gas and water remains constant and then followed by 4 years of the production period. The bottom hole pressure is set at a minimum constraint of 12.0 MPa. Figure 11(a–d) shows the sensitivity of FOPT, FGPT, FWPT and WBHP PRO-1 to the different uncertain model parameters. The main parameters contributing to the uncertainty in the model predictions are n_{cow}, n_{ow}, the porosity of the top layer, n_{rg}, and permeability of the fourth layer. The porosity of the top layer contributes more to the total uncertainty as the original reservoir model has the highest porosity in the top layer. The permeability of the fourth layer has the maximum contribution as all the wells are present in this layer in the reservoir model. Therefore, to predict the production from the reservoir, the most sensitive model parameters need

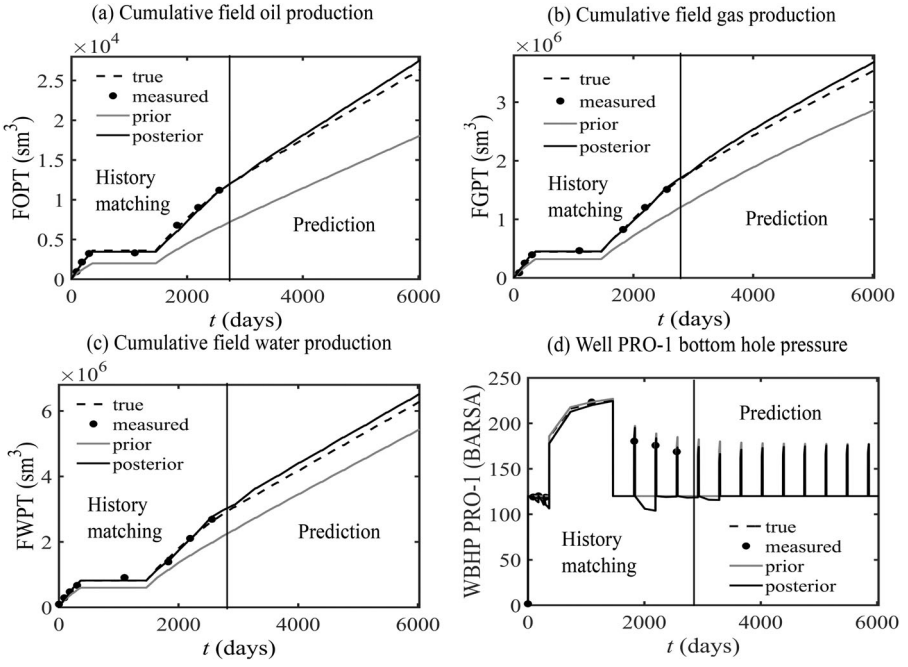

Figure 12. Comparison of predicted values of FOPT, FGPT, FWPT and WBHP-PRO1 versus time using prior mean and posterior mean values of model parameters with the measured data for PUNQ-S3 problem. The measured data were obtained by adding noise on the model predictions obtained using the true values for the first 8 years of production. (a) shows that the FOPT predicted with prior mean values of model parameters does not match with the measured data. However, after history matching, the FOPT predicted by the posterior mean values shows good agreement with the model predictions based on the measured data. Also the predicted FOPT matches well with the true value of FOPT on which noise was added to obtain the history data. Similarly, (b) shows that the FGPT predicted with prior mean values of model parameters does not match with the measured data. However, after history matching, the FGPT predicted by the posterior mean values shows good agreement with the model predictions based on the measured data. (c) and (d) show that the predicted FWPT and WBHP-PRO1 do not match with their prior values but match well with the measured data and the posterior matches well with the true data during prediction period.

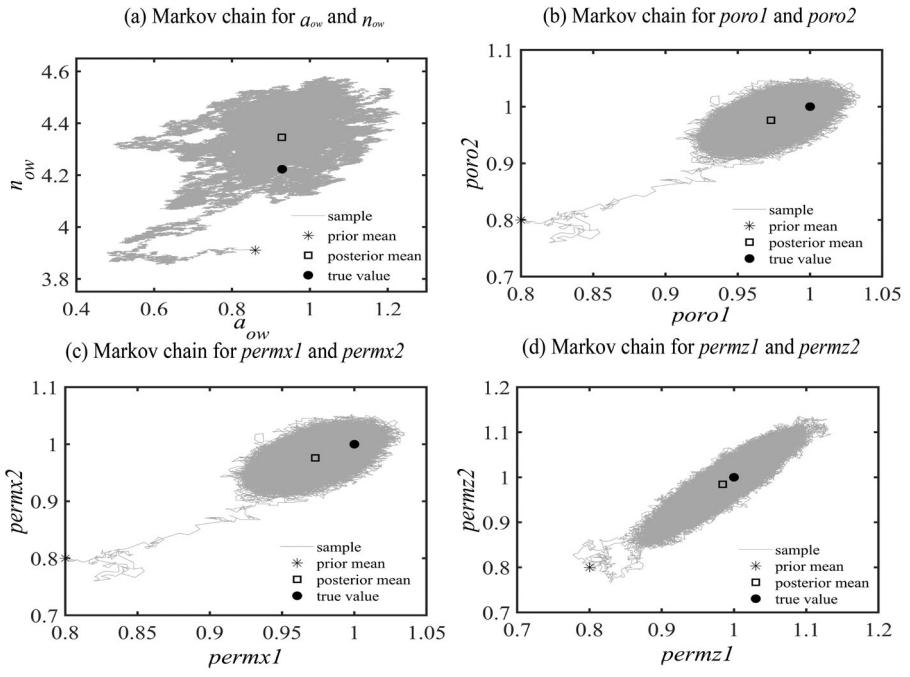

Figure 13. Convergence of 1 million MCMC samples for the PUNQ-S3 problem. (a) Markov chain convergence for a_{ow} and n_{ow}. (b) Markov chain convergence for porosity multiplier of first (*poro*1) and second (*poro*2) layer. (c) Markov chain convergence for *x* permeability multiplier of first (*permx*1) and second (*permx*2) layer. (d) Markov chain convergence for *z* permeability multiplier of first (*permz*1) and second (*permz*2) layer.

to be estimated accurately. To estimate multiple model parameters, we need multiple history data. Therefore, we have used the production data of oil, gas, water and, bottom-hole pressure for all the wells to predict the model parameters.

3.11. History matching

As suggested by the sensitivity analysis, we chose production data of FOPT, FGPT, FWPT, and bottom-hole pressure for all the wells to perform history matching. First, we chose a true value of all the model parameters as shown in Table 3. To generate the data against which we perform history matching, we performed deterministic simulations using the model parameters shown in Table 3, which we consider as the true values of these parameters. The FOPT, FGPT, FWPT and WBHP PRO-1 for true values of model parameters are shown in Figure 12(a–d) respectively. The production data was generated by adding Gaussian random noise to the production data from the deterministic simulation. The standard deviation for the noise is $\sigma = 1\%$ of the maximum value of the corresponding production data. The measured values of FOPT, FGPT, FWPT, and WBHP PRO-1 are shown in Figure 12(a–d) respectively. We took the prior mean and standard deviation of the 26 uncertain model parameters as shown in Table 3. The FOPT, FGPT, FWPT, and WBHP PRO-1 for prior values of the model parameters are shown in Figure 12(a–d) respectively.

We observe that the production data simulated using the prior values of model parameters does not match the true production data. Based on the prior distribution of the model parameters, we obtained the PCE proxy for the forward model as discussed in Section 2.2. Thereafter, we obtained one million samples of the posterior distribution using MCMC which were enough to provide samples from the converged posterior and predicted the model parameters. The convergence of the Markov chain for few model parameters are shown in Figure 13(a–d). The data presented in Table 3 show that the predicted posterior mean values of the model parameters are much closer to the true values than the prior values. Moreover, the uncertainty range associated with the predicted model parameters is much lower than the prior model parameters. This shows the efficacy of the method in predicting multiple uncertain parameters. Figure 12(a–d) shows the posterior FOPT, FGPT, FWPT and WBHP PRO-1 obtained after solving the inverse problem. It shows that the FOPT, FGPT, FWPT, and WBHP PRO-1 obtained on the posterior model parameters are very close to the measured history data and are accurate in predicting the production data.

4. Conclusion

In this work, we have described the use of PCEs to accelerate Bayesian inference for history matching in petroleum reservoir simulations. We have shown that using PCE of model predictions for uncertain model parameters, as a proxy for actual reservoir simulation significantly accelerates the solution of inverse problem using Bayesian inference. As an application of this method, we considered three reservoir model problems based on the black-oil model. In the first problem, we demonstrated PCE-based approach to estimate porosity and permeability in SPE1CASE2 problem. In the second case, we considered a water flooding problem and showed the application of the method to estimate the exponents of the relative permeability curves of oil and water. For the third case, we chose a real field-based reservoir model to predict large number of model parameters. Our history matching results show that PCEs not only act as excellent proxy models for the full reservoir models but also enable the proper selection of most sensitive model parameters that must be tuned to match model predictions with production history.

The use of PCE-based proxy model to solve the inverse problem for history matching replaces the full reservoir simulation at every step of MCMC sampling with an analytical expression. The computational cost for evaluating the PCE-based proxy model is negligible compared with the actual reservoir simulation. Moreover, construction of PCE for model predictions requires two orders of magnitude less number of reservoir simulations compared with that required by MCMC sampling. Consequently, the use of PCEs for Bayesian inference leads to significant reduction in computational time for solving the inverse problem of history matching. Although, we have shown the applicability of the method for history matching in black-oil gas injection and water flooding problems, this approach can also be applied for various other problems in history matching, such as surfactant polymer flooding [7], history matching of fractured reservoirs [42], and CO_2 enhanced oil recovery [43]. In addition to the estimation of spatially constant model parameters, as shown in this work, this approach can also be used for spatially varying model parameters such as porosity and permeability.

The proposed approach has been successfully applied to three reservoir models. Parameter prediction of 26 uncertain model parameters has been made on a real field-based reservoir model PUNQ-S3. These examples illustrate the applicability of our approach for global history matching of petroleum reservoirs. However, a limitation of the present method is that its accuracy and computational time are dependent on the dimension of the model and the sensitivities of model predictions to each uncertain model parameter. For high dimensional problems, where hundreds of or thousands of local grid properties are to be adjusted to fine-tune the reservoir model, a large number of simulations are required to form an accurate PCE proxy. Moreover, the accuracy of estimation of less sensitive parameters would be low compared with more sensitive parameters.

While we have demonstrated the current work on black-oil models, the framework is general for all models, and many different model parameters can be predicted. If the proposed methodology is applied to a compositional model, the coefficients in the equation of state can be the uncertain parameters. In the polymer flooding problem, the uncertain parameters which can be estimated are the viscosity and the shear rate behaviour of the polymer solution and the polymer properties such as permeability reduction [7]. For fractured reservoirs, fracture-matrix transfer function parameters and initial saturation distribution are some of the uncertain parameters which can be estimated.

Note

1. The Open Porous Media project, http://www.opm-project.org

Acknowledgments

Sufia Khatoon wishes to thank Oil and Natural Gas Corporation (ONGC) India for financial assistance under the Pan IIT-ONGC Collaborative Research Program and the Industrial Research and Development (IRD) Unit of Indian Institute of Technology, Delhi for providing the financial support.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by IIT Delhi [IRD].

References

[1] Fanchi JR. Principles of applied reservoir simulation. 3rd ed. Amsterdam, The Netherlands: Gulf Professional Publ.; 2005.
[2] Zhang D, Li L, Tchelepi H. Stochastic formulation for uncertainty analysis of two-phase flow in heterogeneous reservoirs. SPE J. 2000;5:60–70.
[3] Haldorsen H, Brand P, Macdonald C. Review of the stochastic nature of reservoirs. Mathematics in oil production. Oxford: Claredon; 1987. p. 109–210.
[4] Gu Y, Oliver DS. History matching of the punq-s3 reservoir model using the ensemble kalman filter. SPE J. 2005;10:217–224.
[5] Gu Y, Oliver DS. The ensemble kalman filter for continuous updating of reservoir simulation models. J Energy Resour Technol. 2006;128:79–87.
[6] Kaleta MP, Hanea RG, Heemink AW, et al. Model-reduced gradient-based history matching. Comput Geosci. 2011;15:135–153.

[7] Naik P, Aramideh S, Ardekani AM. History matching of surfactant-polymer flooding using polynomial chaos expansion. J Petroleum Sci Eng. 2019;173:1438–1452.
[8] Oliver DS, Chen Y. Recent progress on reservoir history matching: a review. Comput Geosci. 2011;15:185–221.
[9] Marzouk YM, Najm HN, Rahn LA. Stochastic spectral methods for efficient bayesian solution of inverse problems. J Comput Phys. 2007;224:560–586.
[10] Zhang Y, Oliver DS. History matching using the ensemble kalman filter with multiscale parameterization: A field case study. SPE J. 2011;16:307–317.
[11] Bi H, Ma J, Wang F. An improved particle filter algorithm based on ensemble kalman filter and markov chain monte carlo method. IEEE J Sel Top Appl Earth Obs Remote Sens. 2014;8:447–459.
[12] Emerick AA, Reynolds AC. Combining the ensemble Kalman filter with Markov chain Monte Carlo for improved history matching and uncertainty characterization. SPE Reservoir Simulation Symposium; 2011.
[13] Elsheikh AH, Jackson MD, Laforce TC. Bayesian reservoir history matching considering model and parameter uncertainties. Math Geosci. 2012;44:515–543.
[14] Wang J, Zabaras N. A bayesian inference approach to the inverse heat conduction problem. Int J Heat Mass Transf. 2004;47:3927–3941.
[15] Jackson C, Sen MK, Stoffa PL. An efficient stochastic bayesian approach to optimal parameter and uncertainty estimation for climate model predictions. J Clim. 2004;17:2828–2841.
[16] Malinverno A. Parsimonious bayesian markov chain monte carlo inversion in a nonlinear geophysical problem. Geophys J Int. 2002;151:675–688.
[17] Doucet A, De Freitas N, Gordon N. An introduction to sequential monte carlo methods. Sequential Monte Carlo methods in practice. Springer; 2001, p. 3–14.
[18] Oliver DS, Cunha LB, Reynolds AC. Markov chain monte carlo methods for conditioning a permeability field to pressure data. Math Geol. 1997;29:61–91.
[19] Iglesias MA, Law KJ, Stuart AM. Evaluation of gaussian approximations for data assimilation in reservoir models. Comput Geosci. 2013;17:851–885.
[20] Heng L, Sarma P, Zhang D. A comparative study of the probabilistic-collocation and experimental-design methods for petroleum-reservoir uncertainty quantification. SPE J. 2011;16:429–439.
[21] Lawal KA. modelling subsurface uncertainties with experimental design: some arguments of non-conformists. Nigeria Annual International Conference and Exhibition; 2009.
[22] Wiener N. The homogeneous chaos. Amer J Math. 1938;60:897–936.
[23] Le Maître O, Knio OM. Spectral methods for uncertainty quantification: with applications to computational fluid dynamics. Berlin: Springer Science & Business Media; 2010.
[24] Jha AK, Bahga SS. Uncertainty quantification in modeling of microfluidic t-sensor based diffusion immunoassay. Biomicrofluidics. 2016;10:014105.
[25] Rana S, Ertekin T, King GR. An efficient assisted history matching and uncertainty quantification workflow using gaussian processes proxy models and variogram based sensitivity analysis: Gp-vars. Comput Geosci. 2018;114:73–83.
[26] Yan L, Duan X, Liu B, et al. Gaussian processes and polynomial chaos expansion for regression problem: linkage via the rkhs and comparison via the kl divergence. Entropy. 2018;20:191.
[27] Ghanem RG, Spanos PD. Stochastic finite element method: response statistics. Stochastic finite elements: a spectral approach. Springe; 1991, p. 101–119.
[28] Najm HN. Uncertainty quantification and polynomial chaos techniques in computational fluid dynamics. Annu Rev Fluid Mech. 2009;41:35–52.
[29] Reagana MT, Najm HN, Ghanem RG, et al. Uncertainty quantification in reacting-flow simulations through non-intrusive spectral projection. Combust Flame. 2003;132:545–555.
[30] Khatoon S, Phirani J, Bahga SS. polynomial chaos based solution to inverse problems in petroleum reservoir engineering. Fluids Engineering Division Summer Meeting; Vol. 59087, 2019, p. V005T05A076.

[31] Li H, Zhang D. Efficient and accurate quantification of uncertainty for multiphase flow with the probabilistic collocation method. SPE J. 2009;14:665–679.
[32] Bazargan H, Christie M, Elsheikh AH, et al. Surrogate accelerated sampling of reservoir models with complex structures using sparse polynomial chaos expansion. Adv Water Resour. 2015;86:385–399.
[33] Gao G, Zafari M, Reynolds AC. Quantifying uncertainty for the PUNQ-S3 problem in a Bayesian setting with RML and EnKF. SPE reservoir simulation symposium; Vol. 11, 2006, p. 506–515.
[34] Odeh AS. Comparison of solutions to a three-dimensional black-oil reservoir simulation problem. J Petroleum Technol. 1981;33:13–25.
[35] Peaceman DW. Fundamentals of numerical reservoir simulation. Vol. 6, Amsterdam: Elsevier; 2000.
[36] Corey AT. The interrelation between gas and oil relative permeabilities. Producers Monthly. 1954;19:38–41.
[37] Chib S, Greenberg E. Understanding the metropolis-hastings algorithm. Am Stat. 1995;49:327–335.
[38] Smolyak SA. Quadrature and interpolation formulas for tensor products of certain classes of functions. Doklady Akademii Nauk. 1963;148:1042–1045.
[39] Zhang D, Tchelepi H. Stochastic analysis of immiscible two-phase flow in heterogeneous media. SPE J. 1999;4:380–388.
[40] Killough J. Ninth SPE Comparative solution project: A re-examination of black-oil simulation. paper SPE 29110, proceedings of the 13th SPE Symposium on Reservoir Simulation; San Antonio, TX: February 12–15, 1995.
[41] Zhang Y, Fan Z, Yang D, et al. Simultaneous estimation of relative permeability and capillary pressure for punq-s3 model with a damped iterative-ensemble-kalman-filter technique. SPE J. 2017;22:971–984.
[42] Cui H, Kelkar MG. Automatic history matching of naturally fractured reservoirs and a case Study. SPE Western Regional Meeting; Irvine, CA.: March 30–April 1, 2005.
[43] Alfi M, Hosseini SA. Integration of reservoir simulation, history matching, and 4D seismic for CO2-EOR and storage at Cranfield. Fuel. 2016;175:116–128.

Appendix 1. Convergence of PCE coefficients

We evaluated the convergence of PCE coefficients for each case study to choose the order of the polynomial. We chose the minimum order PCE which was sufficient for the respective problem. We have added the figures below showing the convergence of PCE coefficients for all the test cases. Figure A1(a) shows that fifteen PCE coefficients were sufficient for GPR for the SPEICASE2 problem. Figure A1(b,c) shows that fifteen coefficients were sufficient to obtain the convergence of FOPR and FWPR in SPE9 problem. Figure A1(d–f) shows that twenty-six coefficients were sufficient to obtain the convergence of FOPT, FGPT and FWPT in the PUNQ-S3 problem.

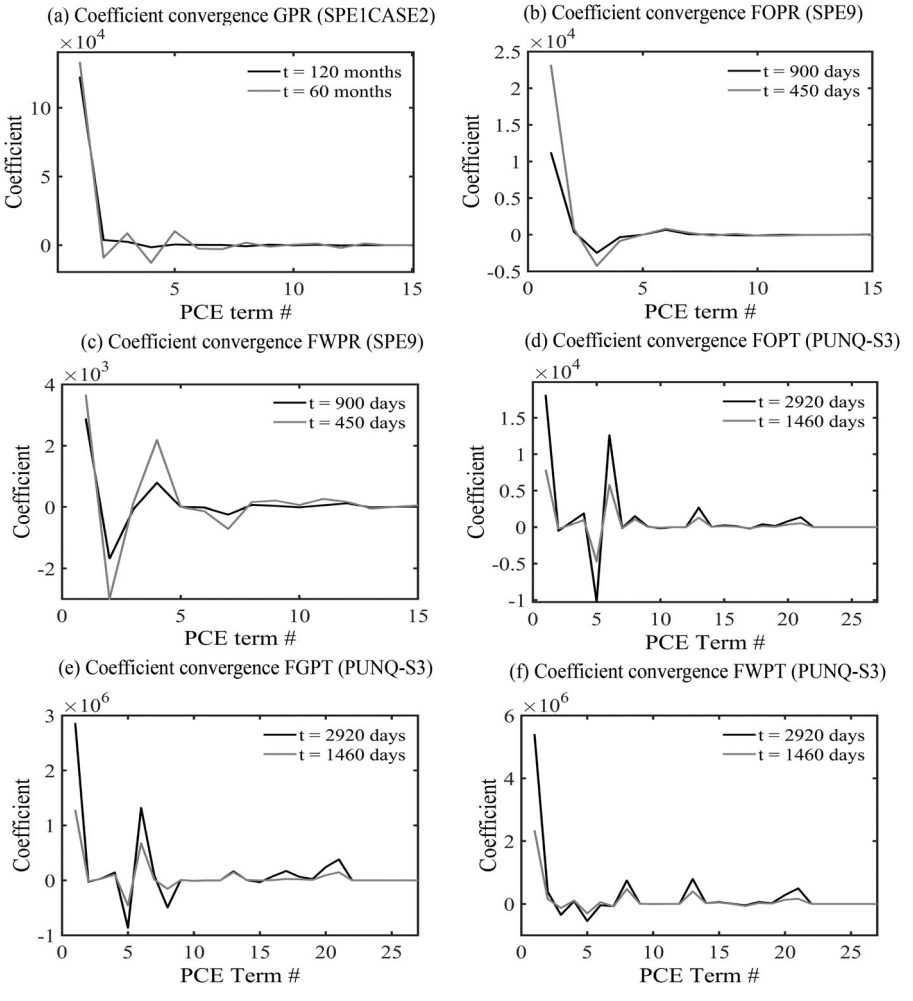

Figure A1. Convergence of PCE coefficients for each test case at two different time instants during the simulation. (a) and (b) show the convergence of PCE coefficients for GPR and FOPR for the SPE1CASE2 problem. (c) shows the convergence of PCE coefficients for FWPR for the SPE9 problem. (d), (e) and (f) show the convergence of PCE coefficients for FOPT, FGPT, and FWPT for the PUNQ-S3 problem.

Appendix 2. Validation of 'Flow' simulator with 'Eclipse'

Here, we present the validation of the 'Flow' simulator with commercially available 'Eclipse' software. We have run the simulations in both the software and the results match well which are presented below. Figure A2(a,b) shows the FOPR and FGPR simulation output using the 'Flow' simulator and the 'Eclipse' software for SPE1CASE2 model problem. Figure A2(c,d) shows the FOPR and FWPR simulation output using 'Flow' and 'Eclipse' for SPE9 model problem.

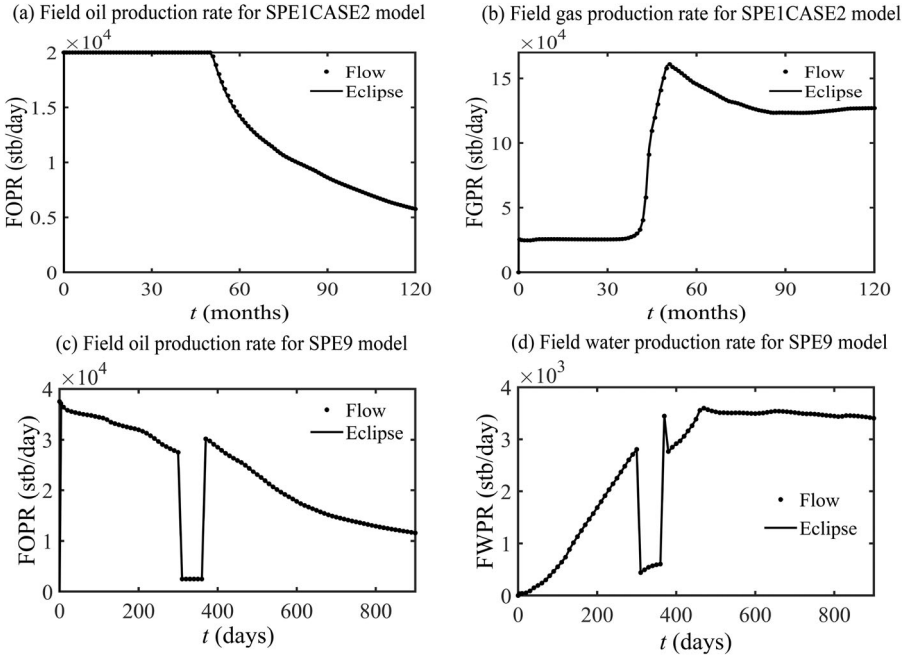

Figure A2. Comparison of the simulation outputs from the 'Flow' simulator and the 'Eclipse' software. The results match well for the two reservoir models. (a) and (b) show the FOPR and FGPR simulation output using the two simulators for the SPE1CASE2 model. (c) and (d) show the FOPR and FWPR simulation output using the two simulators for the SPE9 model.

Appendix 3. Grid convergence

We have performed grid convergence tests on the model problem of SPE1CASE2 for two grid sizes of 1000×1000 and 3000×3000. Figure A3(a,b) shows the simulation output of FOPR and FGPR for the two grid sizes.

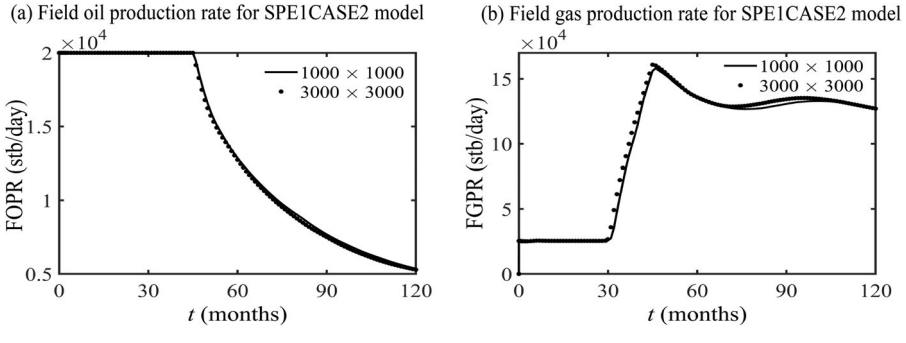

Figure A3. Grid convergence for SPE1CASE2 model for 1000×1000 and 3000×3000 grid size. (a) shows the grid convergence for FOPR and (b) shows the grid convergence for FGPR.