Comput Geosci (2013) 17:671–687
DOI 10.1007/s10596-013-9350-6

ORIGINAL PAPER

Bayesian updating via bootstrap filtering combined with data-driven polynomial chaos expansions: methodology and application to history matching for carbon dioxide storage in geological formations

Sergey Oladyshkin · Holger Class · Wolfgang Nowak

Received: 25 October 2012 / Accepted: 8 April 2013 / Published online: 27 April 2013

Abstract Model calibration and history matching are important techniques to adapt simulation tools to real-world systems. When prediction uncertainty needs to be quantified, one has to use the respective statistical counterparts, e.g., Bayesian updating of model parameters and data assimilation. For complex and large-scale systems, however, even single forward deterministic simulations may require parallel high-performance computing. This often makes accurate brute-force and nonlinear statistical approaches infeasible. We propose an advanced framework for parameter inference or history matching based on the arbitrary polynomial chaos expansion (aPC) and strict Bayesian principles. Our framework consists of two main steps. In step 1, the original model is projected onto a mathematically optimal response surface via the aPC technique. The resulting response surface can be viewed as a reduced (surrogate) model. It captures the model's dependence on all parameters relevant for history matching at high-order accuracy. Step 2 consists of matching the reduced model from step 1 to observation data via bootstrap filtering. Bootstrap filtering is a fully nonlinear and Bayesian statistical approach to the inverse problem in history matching. It allows to quantify post-calibration parameter and prediction uncertainty and is more accurate than ensemble Kalman filtering or linearized methods. Through this combination, we obtain a statistical method for history matching that is accurate, yet has a computational speed that is more than sufficient to be developed towards real-time application. We motivate and demonstrate our method on the problem of CO_2 storage in geological formations, using a low-parametric homogeneous 3D benchmark problem. In a synthetic case study, we update the parameters of a CO_2/brine multiphase model on monitored pressure data during CO_2 injection.

Keywords History matching · Arbitrary polynomial chaos · Bayesian updating · Bootstrap filter · CO_2 storage · Uncertainty quantification

S. Oladyshkin (✉) · H. Class · W. Nowak
SRC Simulation Technology, Institute for Modelling Hydraulic and Environmental Systems (LH2), University of Stuttgart, Pfaffenwaldring 61, 70569 Stuttgart, Germany
e-mail: Sergey.Oladyshkin@iws.uni-stuttgart.de

1 Introduction

1.1 Modeling carbon dioxide storage

We would like to motivate our work on the example of modeling CO_2 storage in geological formations. CO_2 storage is currently being discussed intensively as an interim technology with a high potential for mitigating CO_2 emissions (e.g., [32]). In recent years, great research efforts have been directed towards understanding the processes in CO_2 storage. The multiphase flow and transport processes involved are strongly nonlinear. They include phase changes in the region of the critical point and effects such as gravity-induced fingering and convective mixing as well as geochemical and geomechanical processes, etc.

In order to describe the space–time evolution of injected CO_2 plumes, to analyze the influence of potentially leaky abandoned wells, and to investigate possible geomechanical failure of reservoirs, (semi-) analytical solutions have been derived by Nordbotten et al. and others [53, 54]. A comparison study of various simplifying semi-analytical models with complex numerical simulation tools was performed by Ebigbo et al. [12]. The analysis in [4] focused on the effects of large-scale CO_2 leakage through low-permeability layers. Various optimization strategies for monitoring surface

leakage using near-surface measurement approaches were developed in [9]. These studies are cited here merely to provide a few examples. More detailed reviews are provided in, e.g., [32], [7], and [12].

Modeling underground CO_2 storage involves many conceptual and quantitative uncertainties [29]. Class et al. [7] published a benchmark study that compares a number of mathematical and numerical models of varying complexity. However, the lack of information on subsurface properties (porosity, permeability, etc.) may lead, depending on the specific question at hand, to parameter uncertainties up to a level where parameter uncertainties dominate or even override the influence of secondary physical processes. In [57, 60], the authors of the current study pointed out that the efforts invested in improved physical conceptualization, numerical codes, and stochastic modeling can easily be overwhelmed by error through human subjectivity in data interpretation at a very early stage of modeling, and they proposed a purely data-driven approach to overcome this problem.

In the development of CO_2 injection as a large-scale interim solution, our ability to quantify its uncertainties and risks will play a key role. Sensitivity analyses (e.g., [4], [35]) for CCS have been applied up to the present. Fault-tree analyses have been used to identify risks through different factors [89], but have not yielded quantitative risk information. A significant part of the applied and scientific community still refrains from explicitly considering uncertainty in modeling, although the corresponding arguments are discussed and rejected one by one in [68]. A key hindrance to quantitative risk assessment is that current numerical simulation models are often inadequate for stochastic simulation techniques based on brute-force Monte Carlo simulation and related approaches (e.g., [47], [69]), because even single deterministic simulations may require parallel high-performance computing.

This triggered an urgent need to develop reasonably fast stochastic approaches for probabilistic risk assessment of CO_2 sequestration. In [58], the authors of the current study have pioneered the application of massive stochastic model reductions based on the polynomial chaos expansion (PCE) to CO_2 storage and developed a novel approach to join robust design ideas with the PCE technique. Statistical uncertainty has been combined with scenario uncertainty in the recent work [84] to estimate the risk related to the brine migration resulting from CO_2 injection into saline aquifers. In the current paper, we suggest a history matching (or model calibration, parameter inference) framework based on Bayesian updating, which is both accurate and sufficiently fast to be applied to complex large-scale models. Putting history matching into the Bayesian updating context allows to combine available data with prior expert judgment and to quantify parametric uncertainty after calibration. In this context, our approach will help to override the subjectivity of prior assumptions on parameter distribution by incorporating observed data.

1.2 Inverse modeling and history matching

Inverse modeling and history matching to past production data are extremely important issues in order to improve the quality of prediction. When cast into a stochastic/statistical framework, the results also provide information on the remaining parameter and prediction uncertainty (e.g., [43]). Also, the statistical framework allows to assess the worth of monitoring data and then to optimize future monitoring strategies (e.g., [38]). This would help to raise studies on CO_2 plume monitoring (e.g., [9]) to the level of formal optimization.

The accuracy of inversion or history matching depends on the quality of the established physical model (including, e.g., seismic, geological, and hydrodynamic characteristics; fluid properties; etc.) and on the accuracy of the involved parameter calibration, stochastic inversion, or data assimilation techniques. The quality will also depend on the computational efficiency of all involved methods [63], because too large computational time would have to be mitigated by compromises in numerical or inversion accuracy.

The principal challenge of history matching is that the full reservoir behavior has to be reproduced from local measurements [22]. The reliability of history matching is increasing with the number of available observations. Moreover, besides the proper definition of the involved geological and physical models, the list of uncertain parameters has to be defined adequately for further investigation [36]. Compromises in this task are indispensable. On the one hand, a too short list of parameters leads to a very parsimonious model that allows robust parameter inference, but simply cannot accurately reproduce the observation data. On the other hand, a too long list of parameters leads to a very complex task which cannot be afforded, computationally, and one can run into under-determined or otherwise ill-posed problems [15]. In this stage, an expert opinion on relevant parameters that are useful for calibration, probably combined with sensitivity analyses (see, e.g., [8, 59, 72]) or regularization choices, is very important.

History matching is very well known in the field of reservoir engineering [46]. Traditionally, an iterative manual process of trial and error (see, e.g., [90]) is applied to adjust the reservoir geological model in order to reproduce past observations of oil or gas production. Such a manual approach is very popular among experts in reservoir engineering and demands a very strong understanding of geology and processes. However, the nontrivial and nonlinear interaction of the matched parameters can complicate the history matching procedure a lot [64]. Instead of the

manual technique, formal optimization methods such as gradient search or the adjoint method (see [20, 42, 70]) can be applied. Unfortunately, the mentioned optimization approaches often lead to high computational costs and cannot be easily applied for complex real-world tasks. The state of the art and its recent progress for history matching are presented in detail in the review paper [63].

Another important point about history matching techniques is that they can produce nonunique solutions [62], which means that several virtual models and parameter sets can match the observation data equally well. In fact, this problem is common to most inverse problems [80, 81]. Stochastic approaches can handle such type of uncertainty occurring during the matching procedure without the need to introduce regularization or to artificially restrict the parameter space. Their result is a probability distribution of possible parameter sets instead of a single best estimation. As an improvement of classical optimization, several successful stochastic approximation methods have been adapted for the history matching problem [3, 21, 40].

However, stochastic approaches are more expensive than classical optimization-based (deterministic) calibration techniques, because they need to explore the full range of possible model outcomes with many model runs. In particular, this requires to draw samples from the conditional distribution of the parameters as equally likely calibrated parameter sets. This can be done, e.g., via Markov chain Monte Carlo [27], bootstrap filtering [77], GLUE [38], or rejection sampling [77]. A particular reason for the high computational costs is that, when exploring the full range of possible model outcomes, many model runs are rejected upon comparison with the data. Therefore, the high computational costs of forward modeling get multiplied by large factors in stochastic calibration methods.

The overall efficiency challenge of history matching or inverse modeling can be subdivided into to two principal parts. The first part of the challenge consists in the acceleration of forward modeling itself, using reduced forward models (e.g., response surfaces, surrogate models, low-parametric representation, etc.). The second part of the challenge consists in developing efficient searching algorithms or integration rules for inverse modeling.

To accelerate forward modeling, fast streamline forward simulations have been used in [11]. An optimal representation of uncertain heterogeneous parameter fields in terms of a Karhunen–Loève expansion is presented in [74], leading to a reduced number of parameters to be treated. Later, a kernel-based principal component analysis was presented in [73] as an alternative to the standard Karhunen–Loève expansion. That method is capable to parametrize large-scale non-Gaussian, nonstationary random fields and to reproduce complex geological structures. In low-parametric inverse problems, surrogate models can be constructed with response surface techniques. For example, experimental design techniques were applied in the paper [94] to build an accurate proxy (surrogate) model using polynomials of second order. So-called proxy functions based on polynomials combined with multidimensional kriging have been used to approximate the output of a flow simulator in [65]. Parametric response surfaces have been used in [17] to approximate the reservoir production forecasts and obtain their probabilistic distribution by propagating the remaining posterior uncertainty of input parameters. Recently, a distance-based stochastic technique has been presented in [75].

The ensemble Kalman filter (EnKF) method is one of the simplest yet most successful ways to transfer Bayesian theory (see, e.g., [77]) to practice for model updating and forecasting. The EnKF [14] is derived from a first-order second-moment approximation of error propagation for Bayesian updating. The special issue in *Computational Geosciences* [51] was fully devoted to (ensemble) Kalman filtering for model updating. As a practically successful attempt to accelerate inverse modeling, the EnKF recently received a lot of attention for history matching (e.g., [1, 52]). It is a comparatively cheap method that can generate reasonable history-matched models for real fields. Due to its foundation on first-order second-moment analysis, it is optimal only if all involved model parameters, model states, and data follow a joint multi-Gaussian distribution and if the used ensemble is sufficiently large for accurate covariance estimation. For example, the paper [45] addresses the problem of matching production data by the EnKF for updating facies delineations in reservoir models. The recent article [67] offers an alternative way of computing a linear Bayesian estimator, which allows updating of non-Gaussian quantities. It has been shown that, although no formal linearization is involved, the EnKF contains an implicit form of linearization [55]. Thus, the EnKF has a theoretical limitation which does not allow it to deal with strongly nonlinear problems. For example, the papers [87] and [95] pointed out that the EnKF suffers from nonlinearity and inconsistency in matching. The univariate deviation from Gaussianity can be removed by empirical data transformation (e.g., [31, 76]), but the multivariate structure still poses a problem.

Therefore, we believe that further advancements can be achieved by using more accurate nonlinear approaches to Bayesian updating, when combined with sufficiently accurate (and more expansive) model reduction techniques. In fact, techniques based on response surfaces or other surrogate models have lately been combined with more accurate versions of Bayesian updating. In [33] and [48], the polynomial chaos expansion has been applied to accelerate Bayesian updating via Markov chain Monte Carlo (MCMC)

methods. Even the EnKF has been combined with polynomial chaos expansion for accelerated and accurate computation of the required covariance matrices (see [71], [30], and also related work [67]).

To summarize, we would like to point out that the methodology for history matching has progressed from manual procedures to automatic optimization approaches. During the last years, we see a transition from classical optimization to statistical inference based on Bayesian principles for uncertainty quantification. Also, we can observe that, due to the high computational demands of reservoir models especially in the stochastic framework, the use of surrogate models in history matching received a quickly increasing attention in the very recent years.

The most promising novel approach seems to be the combination of PCE-based response surfaces with EnKF filtering or with more accurate but more expensive implementations of Bayesian updating, such as MCMC. However, more accurate Bayesian updating approaches incorporate higher order stochastic information on the input parameters (e.g., in the form of high-order statistical moments) [26, 37, 88], which goes beyond the scope of the Gaussian assumption. This leaves the optimality range of the EnKF. Also, the Gaussian assumption in stochastic inversion typically goes along with (possibly implicit) linearization of model dependencies.

Our key motivation is that more accurate Bayesian updating also requires to work with high-order approximations of the involved models than linear ones. Using a more accurate updating rule for higher moments would be inadequate when the involved surrogate model is too inaccurate (e.g., by linearizing strongly nonlinear model dependencies). As a consequence, model reduction techniques also need to retain the nonlinearity of models at sufficiently high orders.

1.3 Approach and contributions

The goal of this work is to further advance statistical (Bayesian) model calibration, here working on the example of history matching. We wish to find an even better compromise between computational efficiency and accuracy of the inversion. The resulting framework will take into consideration the nonlinearity of the forward model and of inversion and will provide a cheap but highly accurate tool for reducing prediction uncertainty. We will also aim at a consistent use and processing of high-order statistical moments for the considered uncertain model parameters. Our approach is to combine fully accurate Bayesian updating via bootstrap filtering (or, alternatively, via MCMC) with a model reduction based on the arbitrary polynomial chaos expansion (aPC) technique [57, 60, 61] (see details in Section 2).

The aPC is a so-called data-driven generalization of the PCE technique (see details in Section 2.1). It provides the necessary flexibility to handle arbitrary distributions of model parameters, including high-order statistical moments. Due to this flexibility, it can also handle the conditional parameter distributions that occur after Bayesian updating (see Section 2.2), if one would desire to renew the PCE-based model reduction after Bayesian updating. A second advantage of the aPC approach is that it provides improved convergence with respect to increasing order of expansion [60] of the surrogate to the original model in comparison to more classical PCE techniques, when applied to input distributions that fall outside the range of classical PCE. Thus, the aPC allows efficient allocation of computational resources for constructing of the surrogate model.

Bootstrap filtering (BF) is the most direct yet simple numerical implementation of Bayes' theorem, based on brute-force Monte-Carlo. It approximates the conditional probability density function (PDF) by a sufficiently large ensemble of realizations, and it is exact at the limit of infinite ensemble size. It can account for arbitrary nonlinear model equations and for arbitrary distribution shapes in comparison to (ensemble) Kalman filters. Hence, BF is a perfect match for combination with the aPC technique. This combination is superior to the previous combinations of the classical PCE with (ensemle) Kalman filtering, because the latter two components are both optimal only for (multi-) Gaussian distributions.

Technically, the conditional ensemble is obtained by simple rejection sampling [77] applied to an ensemble of parameter vectors drawn from the prior PDF, i.e., to the parameter distributions assumed before incorporating the calibration or history matching data. However, an accurate representation of conditional statistics for model parameters and responses demands a sufficiently large size of the simulated conditional sample. This can be a problem if the used data set is large and accurate, because the acceptance probability in the rejection sampling approaches to zero. For that reason, BFs can be feasible for very complex models, only when extremely fast evaluation techniques for the response surface are available. For the polynomial response surface resulting from the aPC, this is fulfilled in cases where small acceptance probability still poses a problem; BF can simply be exchanged for more efficient MCMC methods, e.g., [83].

Overall we expect that (1) thanks to the computational efficiency caused by the improved convergence of the aPC and (2) due to the accuracy of bootstrap filtering, Bayesian updating for history matching or more general data assimilation can be developed further towards real time even for complex or large-scale simulation models. For example, we demonstrate in Section 4 that, for our computational example, it takes only about 20 s of CPU time to perform the actual Bayesian updating step for history matching.

In Section 4, we demonstrate a straightforward implementation of aPC-based BF on the benchmark problem of

CO_2 leakage taken form [7] and described in Section 3. We apply our method to match the model repose to synthetic pressure data monitored during CO_2 injection. However, the surrogate model may be very inaccurate and may lead to wrong results, when the prior information is strongly offset against reality. This is caused by the fundamental property of all PCE techniques that the error of approximation is lowest where the (prior) probability density is high, i.e., large errors may occur in low-probability regions. For that case, we introduce an advanced iterative approach for aPC-based BF in Section 5. It allows to perform Bayesian updating even in the case where the prior assumptions on model parameters are far from reality, such that the response surface has to be reiterated in order to be accurate in the relevant regions of high posterior probability. The iteration can be seen as an extension of statistical inversion based on successive linearization to successive high-order expansion. Statistical inversion via successive linearization can be found, e.g., in extended Kalman filtering, in iterative ensemble Kalman filters, or in the quasi-linear geostatistical approach [14, 28, 34, 55].

2 Bootstrap filtering on the polynomial chaos expansion

As outlined in Section 1.3, we propose an advanced stochastic inversion framework for history matching based on a recent generalization of the PCE and bootstrap filtering. Hence, our framework consists of two main steps: a massive but high-order accurate stochastic model reduction via the aPC (Section 2.1) and accurate numerical Bayesian updating of the reduced model via BF (Section 2.2).

2.1 Response surface based on the arbitrary polynomial chaos expansion

PCE techniques can be seen as an efficient and mathematically optimal approach to construct the response surface of a model with uncertain parameters. A response surface can be constructed in different ways. The most straightforward way is to construct it directly on a dense Cartesian grid of input parameters together with an adequate interpolation rule. Due to the curse of dimension [2] for tensor grids, this approach has an extremely high computational effort if more than a single parameter is of interest. Alternatively, conceptually straightforward numerical Monte Carlo (MC) simulation techniques can be used to screen the parameter space. Their convergence is independent of the number of uncertain parameters. However, they are also computationally demanding since the statistical accuracy of their predictions increases only with the square root of the number of realizations used.

The PCE introduced by Wiener [88], generally, can be viewed as an efficient approximation to full-blown stochastic modeling (e.g., exhaustive MC). The basic idea is to approximate the response surface of a model with the help of an orthonormal polynomial basis in the parameter space [26, 37]. In simple words, the dependence of model output on all relevant input parameters is approximated by projection onto a high-dimensional polynomial. This projection can be interpreted as an advanced approach to statistical regression. Alternatively, PCE can be interpreted as a smart and polynomial-based interpolation and extrapolation rule of model output between and beyond different parameter sets.

The key attractive features of all PCE techniques are the high-order approximation of the model [16, 19, 25, 26, 96] combined with its computational speed when compared to alternatives such as MC [58]. This advantage holds mostly for low-parametric problems. The exact break-even point against Monte Carlo, however, depends on model nonlinearity, on the dimension of the parameter space, and on the chaos expansion order, so it varies from problem to problem.

Formally, let $\boldsymbol{\omega} = \{\omega_1, ..., \omega_N\}$ represent the vector of N input parameters for some model $\Omega = f(\boldsymbol{\omega})$. The model $\Omega(\omega)$ may be an explicit or implicit expression (e.g., a partial or ordinary differential equation or a coupled system). We wish to investigate the influence of all parameters $\boldsymbol{\omega}$ on the model output Ω. That the model output may be time and space dependent, $\Omega = f(\omega, t, \mathbf{x})$, where $\mathbf{x} = (x_1, x_2, x_3)$. According to polynomial chaos theory [88], the model output Ω can be approximated by polynomials $\Psi_j(\boldsymbol{\omega})$:

$$\Omega(\mathbf{x}, t; \boldsymbol{\omega}) \approx \sum_{j=0}^{M} c_j(\mathbf{x}, t)\Psi_j(\boldsymbol{\omega}) = \widetilde{\Omega}(\mathbf{x}, t; \boldsymbol{\omega}). \tag{1}$$

The number M of polynomials Ψ_j and corresponding coefficients c_j depends on the total number of analyzed input parameters (N) and on the order d of the polynomial representation as discussed below. The coefficients $c_j(\mathbf{x}, t)$ in Eq. 1 quantify the dependence of the model output Ω on the input parameters $\boldsymbol{\omega}$ for each desired point in space $\mathbf{x}$ and time t, resulting in a surrogate model $\widetilde{\Omega}$.

Assuming that the input parameters within $\boldsymbol{\omega}$ are statistically independent, the multidimensional basis Ψ_k can be constructed as a simple product of the corresponding univariate polynomials $\P_j$ (e.g., [26]):

$$\Psi_k(\boldsymbol{\omega}) = \prod_{j=1}^{N} P_j^{(\alpha_j^k)}(\omega_j), \quad \sum_{j=1}^{N} \alpha_j^k \leq M, \quad k = 1 \ldots N. \tag{2}$$

Here, α_j^k is a multivariate index that contains the combinatoric information on how to enumerate all possible products of individual univariate basis functions. In other words, the index α can be seen as $M \times N$ matrix, which contains

the corresponding degree (e.g., 0, 1, 2, etc.) for parameter number j in expansion term k. The set of polynomials $\{P_j^{(0)}, \ldots, P_j^{(d)}\}$ forms an an orthogonal basis of degree d in the space of parameter ω_j. For example, the polynomials $P_j^{(k)}(\omega_j)$ are Hermite polynomials, if the parameters ω_j are Gaussian-distributed (e.g., [88]). Generally, the polynomial $P_j^{(k)}(\omega_j)$ of degree k in an individual parameter ω_j can be written as a simple linear combination of the different powers i of ω_j:

$$P_j^{(k)}(\omega_j) = \sum_{i=0}^{k} p_{i,j}^{(k)} \omega_j^i, \quad k = 0 \ldots d, \quad j = 0 \ldots N, \quad (3)$$

where $p_{i,j}^{(k)}$ is the coefficient for the power $i = 0 \ldots k$ within the polynomial $P_j^{(k)}(\omega_j)$.

Up to the present, all implementations of PCE require the random variables in which one expands to be statistically independent. However, PCE can also be applied to correlated variables, if correlation can be removed (or minimized) by adequate linear or nonlinear transformation. Only in a few specific cases (including linear correlation), adequate linear (or nonlinear) transformations allow expanding statistical by dependent variables in independent random variables. The advantages of using an expansion basis that is orthonormal directly in the physical random variables without further need for transformation are manifold. They include direct accessibility of the moments of Ω from the expansion coefficients c_j in Eq. 1, better performance of integration rules, and the fact that additional transformation would include new nonlinear terms in the expansion [60]. In particular, the technique presented can be extended to many classes of heterogeneous systems, where spatially correlated heterogeneous parameter fields can be decomposed into their uncorrelated principal components using a truncated Karhunen–Loève (KL) expansion (e.g., [41]), if heterogeneity does not span over too many scales.

In the current paper, we will apply a most recent generalization of the PCE technique known as the aPC (see [57, 60, 61]). Highly similar ideas can be found in [23, 78, 91, 92]. Compared to earlier PCE techniques, the aPC adapts to arbitrary probability distribution shapes of input parameters and, in addition, can even work with unknown distribution shapes when only a few statistical moments can be inferred from limited data or from expert elicitation. The arbitrary distributions for the framework can be either discrete, continuous, or discretized continuous. They can be specified either analytically (as probability density/cumulative distribution functions), numerically as histogram or as raw data sets. Thus, exact probability density functions do not have to be known and do not even have to exist. The available information can directly and most purely be used in stochastic analysis, when using our data-driven formulation of aPC.

The aPC approach provides improved convergence with respect to increasing order of expansion (e.g., [60]) in comparison to classical PCE techniques, when applied to input distributions that fall outside the range of classical [88] or generalized version [93] of the PCE. aPC deals with a highly parsimonic and yet purely data-driven description of uncertainty. The new freedom opens the path to accessing with the chaos expansion technique even those applications where data samples of limited size merely allow the inference of a few moments, and one would not be able to construct a probability density function without introducing subjective assumptions and hence dangerous sources of bias. The necessity to adapt to arbitrary distributions in practical tasks is discussed in more detail in [57]. A future incentive to work with the aPC is that, during sequential Bayesian updating for nonlinear problems, parameter distributions generally change their shapes from updating step to updating step. Thus, they will almost surely fall out of the optimality range of many previous PCE techniques.

According to [60], an orthogonal polynomial basis up to order d can be constructively defined for any arbitrary probability measure, given that ω_j has finite statistical moments (e.g., mean, variance, skewness, etc) up to order $2d-1$. The coefficients $p_{i,j}^{(k)}$ within the basis polynomial in Eq. 3 can be defined from the following matrix equation:

$$\begin{bmatrix} \mu_{0,j} & \mu_{1,j} & \ldots & \mu_{k,j} \\ \mu_{1,j} & \mu_{2,j} & \ldots & \mu_{k+1,j} \\ \ldots & \ldots & \ldots & \ldots \\ \mu_{k-1,j} & \mu_{k,j} & \ldots & \mu_{2k-1,j} \\ 0 & 0 & \ldots & 1 \end{bmatrix} \begin{bmatrix} p_{0,j}^{(k)} \\ p_{1,j}^{(k)} \\ \ldots \\ p_{k-1,j}^{(k)} \\ p_{k,j}^{(k)} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \ldots \\ 0 \\ 1 \end{bmatrix}. \quad (4)$$

Here, $\mu_{i,j}$ are the i^{th} noncentral (raw) statistical moments of the random variable ω_j. It becomes evident from Eq. 4 that statistical moments are the only form of information required on the input distributions. Other possible contribution methods are discussed in [23, 78, 91, 92]. If the matrix in Eq. 4 has a poor condition for high-order expansions, adequate problem scaling will be helpful.

The above orthogonal polynomial basis can be used directly for analysis. However, a normalized basis has further useful properties and can be obtained as

$$\widehat{P}_j^{(k)} = \frac{P_j^{(k)}}{\left\| P_j^{(k)} \right\|}, \quad \left\| P_j^{(k)} \right\|^2 = \int_{\omega_j \in \Lambda} \left[P_j^{(k)}(\omega) \right]^2 d\Gamma(\omega_j), \quad (5)$$

where $\| P_j^k \|$ is the normalizing constant of the polynomial P_j^k for space of events Λ (where $\omega_j \in \Lambda$) with probability measure Γ. In practical applications of the aPC, the integration in Eq. 5 is performed analytically based in the moments of ω up to order $2d$, which avoids (like Eq. 4) the necessity to know exactly the probability measure [60].

To determine the unknown coefficients $c_j(\mathbf{x}, t)$ of the expansion in Eq. 1, many different methods exist. We choose to apply the nonintrusive probabilistic collocation method (PCM for short, see, e.g., [41, 58]). Non-intrusive methods do not require modifications in the system of governing equations. Thus, they require no corresponding changes in simulation codes, which is very valuable for modeling complex flow or transport processes with commercially available software. The PCM is based on a minimal and optimally chosen set of model evaluations, each with a defined set of model parameters (called collocation points) that is related to the roots of the polynomial basis via the optimal integration theory [82].

From the practical point of view, the computational costs of our framework are dominated by the model calls required in construction of the surrogate model, i.e., by aPC combined with PCM. In the PCM technique, the number M of model evaluations is equal to the number M of coefficients, which can be obtained from the following equation:

$$M = (N + d)!/(N!d!) - 1. \tag{6}$$

The order d of expansion is typically found as a compromise between the accuracy required by the application, the number of parameters N included in the analysis, the computational costs of individual model evaluations and the available computer power.

The challenge here is to find a compromise between computational effort and a reasonable approximation of the physical processes. Sometimes, a much smaller number M can be found by a sparse selection [5] of the relevant terms in Eq. 1. Also, an initial sensitivity analysis can be applied to keep the number of analyzed parameters and their respective degree of expansion at a low and efficient level (see, e.g., [59, 72]). Because the number M quickly increases with the degree d of expansion when the number of parameters N is large, the framework is most efficient for low-parametric systems. In order to translate these advantages to spatially heterogeneous systems, a combination with the KL expansion (e.g., [41]) to truncate the representation of spatial heterogeneity can be very useful.

2.2 Bootstrap filter

The aPC method in Section 2.1 reduces the original simulation model $\Omega(\omega)$ to a response surface (surrogate model) $\widetilde{\Omega}(\omega)$ that approximates the original model. Because the reduced model is merely a polynomial and has exploitable properties due to the orthogonal basis, it is vastly faster than the original one and offers a large playground for stochastic analysis [24, 57, 85], risk assessment [58, 84], and global sensitivity analysis [10, 57, 79]. In this Section, we will apply Bayesian updating (see, e.g., [77]) to match the surrogate model $\widetilde{\Omega}(\omega)$ to available measurements in a data vector $\mathbf{y}$ of state variables or to other past or real-time observations of system behavior. In this context, Bayes' theorem is

$$f(\omega|\mathbf{y}) = \frac{f(\mathbf{y}|\omega)\, f(\omega)}{f(\mathbf{y})}, \tag{7}$$

where $f(\omega)$ is the joint prior PDF for the vector of model parameters ω, $f(\mathbf{y})$ is the prior probability of $\mathbf{y}$ used as normalization constant, $f(\mathbf{y}|\omega)$ is the conditional PDF of $\mathbf{y}$ for given ω, i.e., the likelihood of the parameters, and $f(\omega|\mathbf{y})$ is the conditional PDF of ω for given $\mathbf{y}$, which we seek to approximate swiftly and accurately.

Under the common assumption that the measurement errors are independent and Gaussian, i.e., $\mathbf{y} = \Omega(\omega) + \epsilon$, $\epsilon \sim N(0, \mathbf{R})$, the likelihood function becomes

$$f(\mathbf{y}|\omega) \propto \exp\left[-0.5\,(\mathbf{y} - \Omega(\omega))^T\, \mathbf{R}^{-1}\, (\mathbf{y} - \Omega(\omega))\right], \tag{8}$$

where $\mathbf{R}$ is the diagonal (co)variance matrix of measurement errors. We will adapt this assumption for the remainder of this study, although other arbitrary error models could be considered as well.

The most direct numerical implementation of Bayes' theorem (Eq. 7) is known as bootstrap filtering (e.g., [77]). Formally, we draw a number N_p of realizations of parameter vectors ω_i from the prior PDF $f(\omega)$:

$$\omega \sim f(\omega), \quad i = \overline{1, N_\mathrm{p}} \tag{9}$$

where N_p is a sufficiently large number. The correction form $f(\omega)$ to $f(\omega|\mathbf{y})$ in Eq. 7 is represented by assigning importance weights w_i to each realization ω_i:

$$w_i = \frac{f(\mathbf{y}|\omega_i))}{\max(f(\mathbf{y}|\omega_i))}. \tag{10}$$

Then, realization ω_i is accepted as a legitimate ensemble member of the posterior distribution, if $w_i \geq u_i$ is fulfilled for a random number u_i drawn from the uniform distribution $u(0, 1)$. $\max(f(\mathbf{y}|\omega_i))$ is the largest individual values $f(\mathbf{y}|\omega_i)$ appearing across all realizations $i = \overline{1, N_\mathrm{p}}$. The convergence for a given value N_p can be assessed, e.g, via weighted Jackknife of bootstrapping techniques [13, 39].

3 Scenario definition for history matching: problem of CO_2 leakage

We will consider a relatively simple benchmark model to illustrate and thoroughly test the concept presented in the current paper. We use the benchmark leakage problem of injected CO_2 into overlying formations through a leaky well defined by Class et al. [7] as described in Section 3.1. Within our approach, the physical complexity of the problem can be increased to an arbitrary extent because of the nonintrusive black-box conception of the proposed framework. How the approach scales with larger numbers of uncertain parameters has already been discussed in Section 2.1. Statistical

assumptions, generation of the synthetic data sets, and the test cases are defined in Section 3.2

3.1 Physical and numerical setup of the CO_2 leakage problem

Figure 1 illustrates a 2D section of the 3D domain under consideration. The top of the figure is not the ground surface, but the transition to yet another aquitard. CO_2 is injected into a deep aquifer, spreads within the aquifer, and, upon reaching a leaky abandoned well, rises to a shallower aquifer. CO_2 is injected at a constant rate of 8.87 kg/s, which corresponds to 1,600 m^3/d at reservoir conditions. The leaky well is at the center of the domain, and the injection well is 100 m away. Both aquifers are 30 m thick and the separating aquitard has a thickness of 100 m. The initial conditions in the fully saturated domain include a hydrostatic pressure distribution which is dependent on the brine density. The aquifers are initially filled with brine. The initial pressure at the bottom of the domain (at a 3,000-m depth) is 3.086×10^7 Pa.

The benchmark problem assumes that fluid properties such as density and viscosity are constant; all processes are isothermal; CO_2 and brine are two separate and immiscible phases; mutual dissolution is neglected; the pressure conditions at the lateral boundaries are constant over time; the formation is isotropic, rigid, and chemically inert; and capillary pressure is negligible. More details and modeling parameters can be found in [7].

Mass balances of the two phases and the multiphase version of Darcy's law give the following system of differential equations:

$$\begin{aligned} -\phi \frac{\partial S_\mathrm{w}}{\partial t} - \nabla \cdot \left\{ \frac{k_\mathrm{rw}}{\mu_\mathrm{w}} \mathbf{K} \cdot (\nabla p - \rho_\mathrm{w} \mathbf{g}) \right\} - q_\mathrm{w} = 0, \\ \phi \frac{\partial S_\mathrm{g}}{\partial t} - \nabla \cdot \left\{ \frac{k_\mathrm{rg}}{\mu_\mathrm{g}} \mathbf{K} \cdot (\nabla p - \rho_\mathrm{g} \mathbf{g}) \right\} - q_\mathrm{g} = 0, \end{aligned} \tag{11}$$

which is constrained by the equation

$$S_\mathrm{w} + S_\mathrm{g} = 1. \tag{12}$$

The subscripts w and g stand for the brine (water) phase and the CO_2-rich (gas) phase, respectively. The primary variables in Eq. 11 are the gas-phase saturation S_g and pressure p. S_w is the brine-phase saturation. The relative permeabilities k_rw and k_rg are secondary variables and linear functions of S_w and S_g ($k_\mathrm{rw} = S_\mathrm{w} = 1 - S_\mathrm{g}$, $k_\mathrm{rg} = S_\mathrm{g}$), $\mathbf{g}$ is the gravity vector, $\mathbf{K}$ is the absolute permeability tensor, ϕ is porosity, and q_w and q_g are sources/sinks.

The CO_2 leakage rate is defined in the benchmark study as the total CO_2 mass flux integrated over a horizontal control plane midway between the top and bottom aquifer, divided by the injection rate, in percent. In the current study, the benchmark problem is simulated using DuMuX, a multiscale, multiphysics toolbox for the simulation of flow and transport processes in porous media [18]. The goal of the simulation is to quantify the leakage rate which depends on the pressure buildup in the aquifer due to injection, on the proprieties of formation and the leakage well, and on the plume evolution.

3.2 Setup for the history matching task

In our study, we will consider three uncertain parameters: reservoir absolute permeability, reservoir porosity, and permeability of the well. For simplicity, we feature here only a scalar permeability, homogeneous within the entire reservoir. The prior assumptions on the parameters (see section below) will be updated using our framework, such that model predictions match a synthetic time series of pressure monitored daily at the top of the well (Fig. 1) during the 100 days of injection.

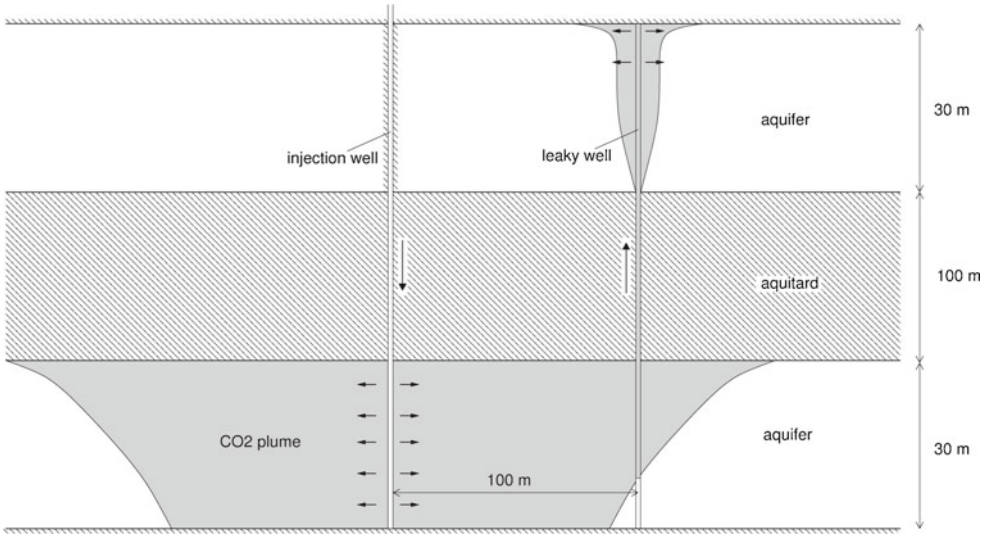

Fig. 1 Benchmark problem for CO_2 leakage through an abandoned well, taken from [7]

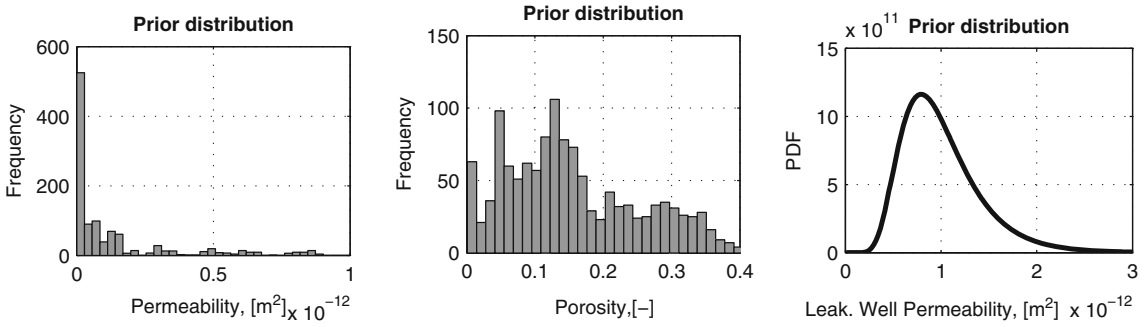

Fig. 2 Prior distribution of model parameters: absolute permeability, porosity, and leakage well permeability

We assume prior distributions of the uncertain model parameters as illustrated in Fig. 2. For reservoir permeability and porosity, these distributions are represented by raw data sets taken from the US National Petroleum Council public database (which includes 1,270 reservoirs, see also [35]). For the well, we assume as expert knowledge a log-normal PDF with the location parameter $\mu = -27.7$ and the scale parameter $\sigma = 0.4$. The presented aPC-based framework does not have any restrictions on the shapes of prior distributions; thus, it can handle the statistical information provided here without further modification.

As a specific example, the positivity of permeability and the (0,1)-boundedness of porosity are directly enforced through their assigned probability distributions, without any additional necessary action. In our test case, we use a third-order aPC expansion, because it demonstrates the advantages of nonlinear information and expert opinion processing over linear methods (e.g., the entire family of Kalman filters). Typically, a third-order expansion has the freedom to handle nontrivial and possibly nonmonotonic behaviors of model output. However, a second-order analysis has been shown to be a reasonable approximation for aPC applied to this benchmark problem in [57]. In general, the degree of expansion can be chosen according to the specific needs and available computer resources in any specific application, and it can also be increased adaptively during the matching procedure (see details in Section 5). We prefer here a third-order expansion, because most parameter-state relations in the featured model are monotonic, and second-order polynomials do not offer the possibility to approximate monotonic behavior over large parameter ranges.

To illustrate the proposed method, we will generate two reference simulations (case 1 and case 2), which provide synthetic reference time series of pressure data at the monitoring well. In case 1, the parameter values (Table 1) used in the reference simulation are reasonably probable within the assumed prior distributions (see Fig. 2). In contrast to case 1, case 2 will use parameter reference values (see porosity in Table 1) in a pressure response which is extremely improbable according to the prior. The resulting reference pressure time series and how they compare to the prior distribution of pressure can be seen in Fig. 3 (case 1) and Fig. 9 (case 2). Test case 1 is treated in Section 4, and test case 2 is treated in Section 5.

We will consider a random synthetic measurement error (i.e., standard deviation) of 1 bar. Thus, in our case studies, the original model can never reproduce exactly the data, even when the "true" parameter set used in data generation is applied, which is typical in real-world applications. Please note that, due to the overdetermined and stochastic problem character, the individual data points will not be matched precisely, but the solution will be a best compromise between prior information and a best fit of the physical model to the cloud of data points.

4 Straightforward aPC-based bootstrap filtering

We will now apply the straightforward aPC-based BF described in Section 2 to case 1 of the benchmark problem of CO_2 leakage described in Section 3. That means, we will perform history matching to the synthetic pressure data marked as squares in the right plot of Fig. 3. First, we construct a surrogate model by projecting the original model (Eq. 11) onto a polynomial response surface via the aPC at third order (see details in Section 2.1). Then, we apply the BF as described in Section 2.2. Figure 3 illustrates the unmatched pressure histogram after 100 days of

Table 1 Reference values: case 1 and case 2

	Reservoir absolute permeability	Reservoir porosity	Leaky well permeability
Case 1	$1.43 \cdot 10^{-13}$ m^2	0.023 [–]	$1.09 \cdot 10^{-12}$ m^2
Case 2	$1.43 \cdot 10^{-13}$ m^2	0.259 [–]	$1.09 \cdot 10^{-12}$ m^2

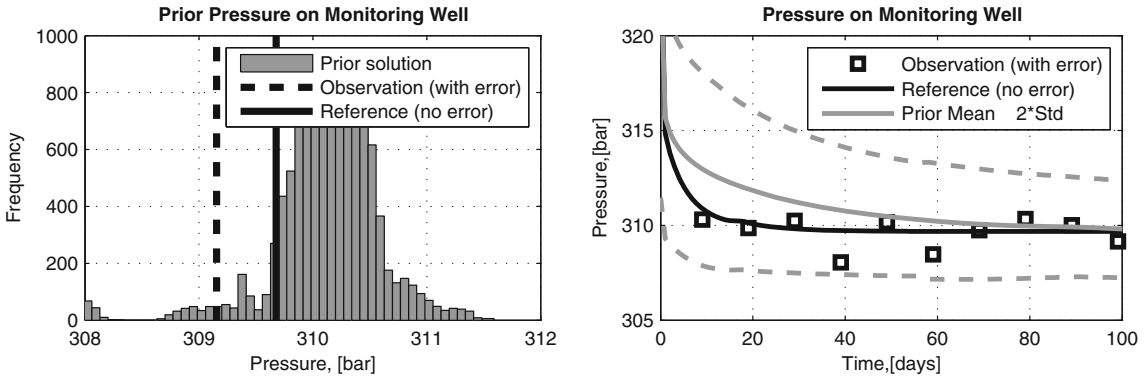

Fig. 3 Prior pressure at monitoring well: pressure histogram after 100 days of injection (*left plot*) and correspondence to observation values during 100 days (*right plot*), for case 1

injection (left plot) and the correspondence of predicted values (mean ± 2 standard deviations) to the synthetic observed values during the first 100 days (right plot).

We draw a number of $N_p = 100{,}000$ particles from the prior distributions. This number was assured to be sufficiently large to yield stable posterior statistics upon filtering. Because the data base we use for defining the prior distributions in not as large as 100,000 data sets, we use a PDF estimation technique based on Kernel smoothing (see, e.g., [86]) with minimal Kernel width to obtain $N_p = 100{,}000$ particles. Then, the obtained surrogate model is evaluated for each particle, and each particle is reweighed according to its correspondence to observation values (see 7 and 8).

At first, we are interested in the statistical distribution of the uncertain parameters and of the pressure at the monitoring well before and after matching, in order to assess the accuracy of the inversion. The resulting posterior distributions of model parameters are presented in Figs. 4 and 5 against the reference values. The quality of matching the pressure data is shown as PDF in Fig. 6. Figure 6 also shows how the quality of matching can be improved when increasing the number of available observation values. Overall, we observe very satisfactory and consistent results: increasing the degree of expansion helps to better catch nonlinearity of the original model and shows convergence of distribution peaks and mean values to reference values in Fig. 4, as much as allowed by the data considered for history matching. The peaks of the posterior parameter PDFs in Fig. 6 (for third order) and for the posterior pressure values in Fig. 5 (also third order) are close to the reference values, with improved accuracy for an increasing number of considered data values. The posterior pressure PDF at the monitoring well converges to the noise-free reference value. In Fig. 6, the time-dependent posterior mean of pressure at the monitoring well converges, already with just few data values, to a best-fit curve through the noisy data, almost coinciding completely with the nose-free reference curve. Finally, the noisy data points fall into the 95 % confidence interval of the posterior pressure distribution (see Fig. 6), indicating reasonable error bounds.

In addition, the multivariate posterior probability density function can be constructed from the posterior ensemble for the model parameters (see the left plot in Fig. 7). Here, we used the multivariate kernel density estimation based on

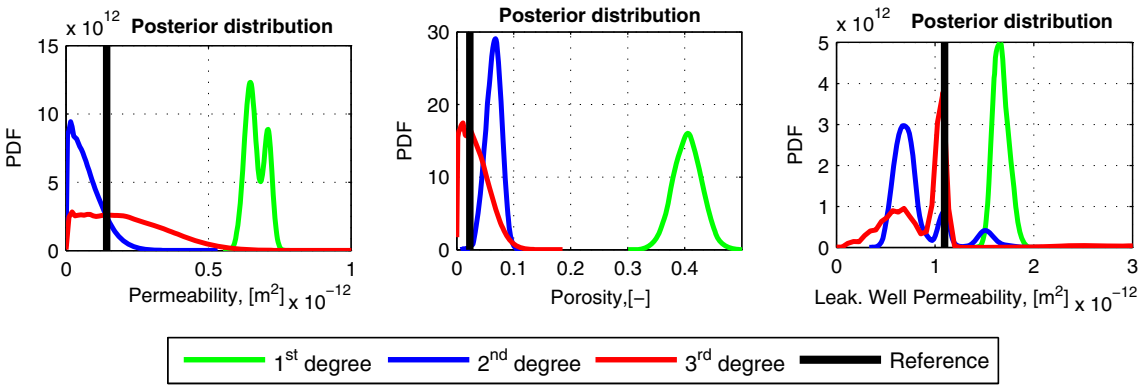

Fig. 4 Posterior distribution of model parameters obtained by first, second, and third degree of expansion using 100 measurement values: absolute permeability (*left*), porosity (*center*), and leakage well permeability (*right*) in case 1

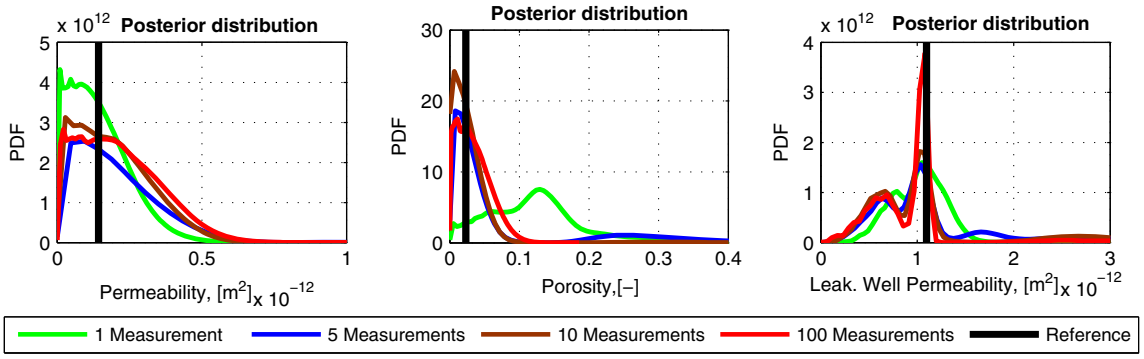

Fig. 5 Posterior distribution of modeling data using 1, 5, 10, and 100 measurements at third order of expansion: absolute permeability, porosity, and leakage well permeability in case 1

FIGTree algorithm [49]. The black point in the left plot of Fig. 7 indicates the peak of the posterior multivariate PDF. As a final convergence check, the original model has been simulated within the corresponding parameter values. The root square error of the obtained pressure from the reference is presented in the right plot of Fig. 7. It shows an acceptable approximation error that is small compared to the 95 % confidence interval of the data noise (±2 bars for 95 % confidence interval). It can be improved even further by adjustment of the response surface in the domain of high posterior probability density, as it will be discussed in the next section.

Once the virtual model is matched to observation values, the forecast of storage behaviors can be investigated. Diverse physical quantities, such as saturation, pressure distribution within the overall domain and the CO_2 leakage rate can be predicted. We will illustrate the reservoir behavior after 1,000 days by the cumulative probability function (CDF) of the CO_2 leakage rate, which is an integrative and very important characteristic of the overall benchmark problem for probabilistic risk assessment. The CDF represents the probability that the CO_2 leakage after 1,000 days is less than or equal to a particular value. Figure 8 shows the corresponding prior and posterior CDFs. It also illustrates how the matching with more and more data improves the prediction. This illustrates how the proposed framework could be used within worth-of-data and information analysis (e.g., [38, 56]).

In order to illustrate the efficiency of the presented history matching framework, we will provide some details regarding the involved CPU time. Here, we considered a simple benchmark model which demands approximately 10 min of CPU time without parallelization on a contemporary PC. To obtain the response surface ($N = 3$ and $d = 3$), we performed 20 runs of the original benchmark model, resulting in an overall simulation time of 200 min. One evaluation of the obtained response surface for history matching demands approximately 0.2 ms. In this example, we used $N_p = 100{,}000$ particles for bootstrap filtering. Hence, for the overall procedure of history matching, we

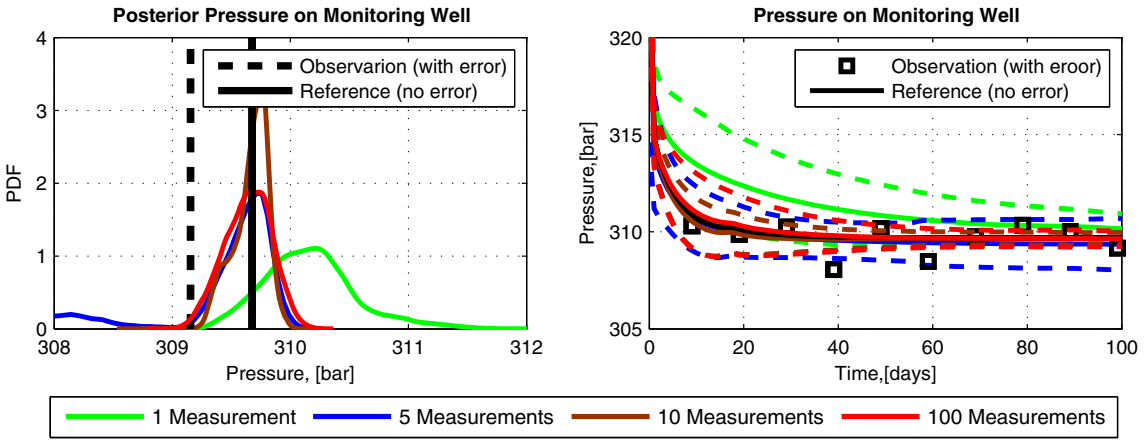

Fig. 6 Posterior pressure at monitoring well: pressure distribution after 100 days of injection using 1, 5, 10, and 100 measurements (*left plot*) and matching of pressure (mean ± 2 standard deviations) to all observation values during 100 days (*right plot*) for case 1. All results obtained at third order

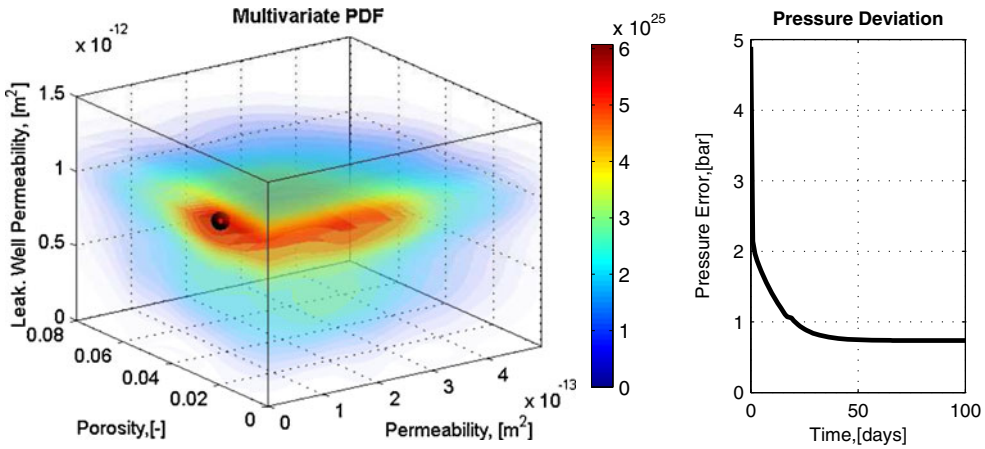

Fig. 7 Multivariate posterior probability density function based on 100 measurements (*left plot*) and deviation of posterior peak PDF pressure from the reference (*right plot*) for case 1. The large values for probability density on the color scale in the left plot are caused by the small magnitudes of permeability

used about 3.34 h (20×600 s $+100{,}000 \times 0.0002$ s, i.e., it takes only about 20 s only to perform the actual Bayesian updating step) of CPU time. Alternatively, direct application of bootstrap filtering to the original benchmark model would demand approximately 695 days ($100{,}000 \times 600$ s) of CPU time. It is evident that such direct application of bootstrap filtering (or similar methods) seems to be not feasible even when involving parallelization, especially for realistic and complex models. Thus, the presented framework provides an extremely efficient way for handling the task of history matching, with very accurate statistical methods. Without the PCE-based response surface, nobody would ever even consider applying an approach like bootstrap filtering on this type of model.

5 Iterative aPC-based bootstrap filtering

In this section, we will reconsider the problem of history matching for the case where the true properties of the system are very far from the prior (case 2 defined in Section 3.2, see Fig. 9). This is a very challenging problem for updating, because the assumed prior distribution does not adequately cover the domain of interest in hindsight. Thus, the aPC-based response surface used in BF is fitted to a distant and poorly chosen region within the parameter space. That means, it represents an expansion around a point or distribution far away from the region of interest and hence cannot be expected to be accurate. Theoretically, the approximation could be improved significantly using a higher degree of expansion, leading to a very high number of model evaluations. Unfortunately, this would ask for a very high computational power, as can be seen from Eq. 6. As an alternative, we will reiterate the response surface, keeping it at third order.

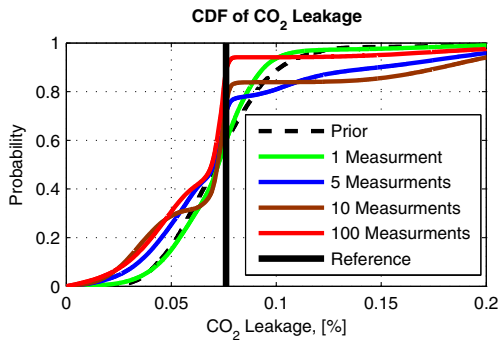

Fig. 8 Cumulative density function of CO_2 leakage: prior and posterior prediction after 1,000 days of injection using 1, 5, 10, and 100 measurements for case 1

We will initially perform the same steps as in Section 4, i.e., we will project the original model (Eq. 11) onto a response surface via aPC and then apply BF on the obtained surrogate model. Let us denote this initial step as zeroth iteration. The corresponding results are presented by green lines in Fig. 10 for the posterior distribution of model parameters and for the updated pressure at the monitoring well in Fig. 11.

We will improve this initial zeroth prediction by applying an iterative approach. The idea we will pursue here is to improve the response surface in the parameter domain where the respective previous step indicated a high posterior probability density, because this is the alleged (best current guess for the) parameter region of interest. For that, in each iteration, we will include new integration points for the projection onto the orthonormal basis. These additional integration points for projection are located in the space of uncertain parameters, where the current iteration of the posterior density is largest. For all new integration points, we will run the original model (Eq. 11) to obtain the corresponding model outputs Ω. Then, we will perform a new projection of the model onto the orthonormal basis (see Section 2.1) using all cumulatively available collocation

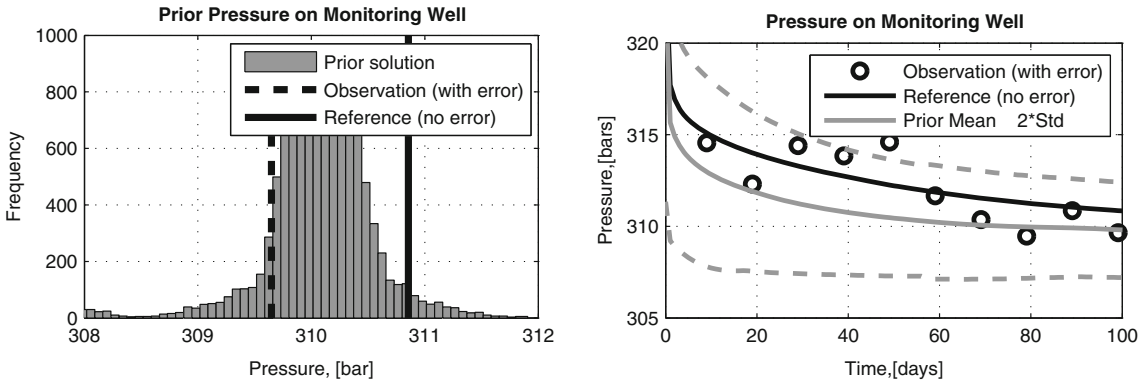

Fig. 9 Prior pressure at monitoring well: pressure histogram after 100 days of injection (*left plot*) and correspondence to observation values during 100 days (*right plot*) for case 2

points within the least-squares collocation method (e.g., [6, 50]) instead of simple collocation. The extension of the original probabilistic collocation method by least-squares collocation during the iterations is very useful for complex applied tasks, because it keeps the black-box property. Also, the initial PCM in the zeroth iteration can be understood as the special case of least squares at the limit of a determined problem.

The residual of the least-squares collocation method in each integration point is not equal to zero in comparison to the classical probabilistic collocation technique, where the residuals are zero at all collocation points. However, it provides an integration rule which minimizes the overall residuals [44] in all collocation points. In that way, we can cumulatively incorporate all available (i.e., original and additional) integration points. Thus, the updated response surface contains more accurate information about the system behavior in all alleged regions of interest. In practice, it is improved in the region where the posterior distributions during iteration assign a strong weighting. This bears some conceptual similarity to approaches like the shifted Hermite chaos (e.g., [66]) for evaluation of failure probabilities. The proposed iteration scheme can be very useful for history matching because it can help to avoid the dangerous consequences of specifying a misleading prior distribution. Our method does not change the meaning or importance of such prior information in Bayesian updating, but merely protects the response surface from being expanded in an inappropriate alleged region of interest.

In each iteration step, the number of new integration points can vary from 1 up to any desired number of such points. In our concept, the modeler can specify the order of local refinement (zeroth, first, second, etc.). According to that order, a new polynomial basis can be constructed using the posterior distribution obtained from the previous iteration. Then, a new full tensor set of optimal integration points can be computed from the respective roots of the new polynomial basis, again based on the optimal integration theory [82].

Obviously, using the full-tensor grid would be very costly, especially when the number of parameters and the order of local refinement are high. To reduce the number of model evaluations, we suggest again using the principles behind the PCM method. This means to add just a limited number of points from the suggested tensor grid, selected to cover the high probability region of the current posterior. Thus, the modeler has some freedom to choose the number of evaluations of the original model according to the

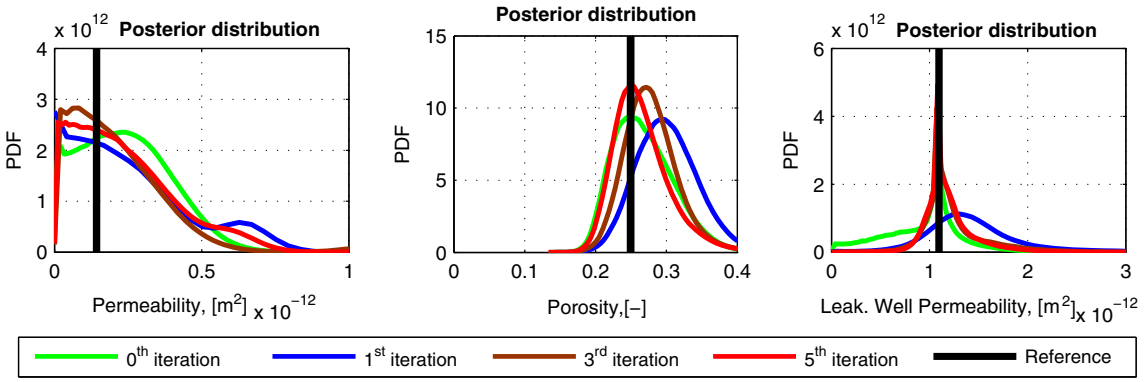

Fig. 10 Posterior distribution of modeling data: absolute permeability, porosity, and leakage well permeability

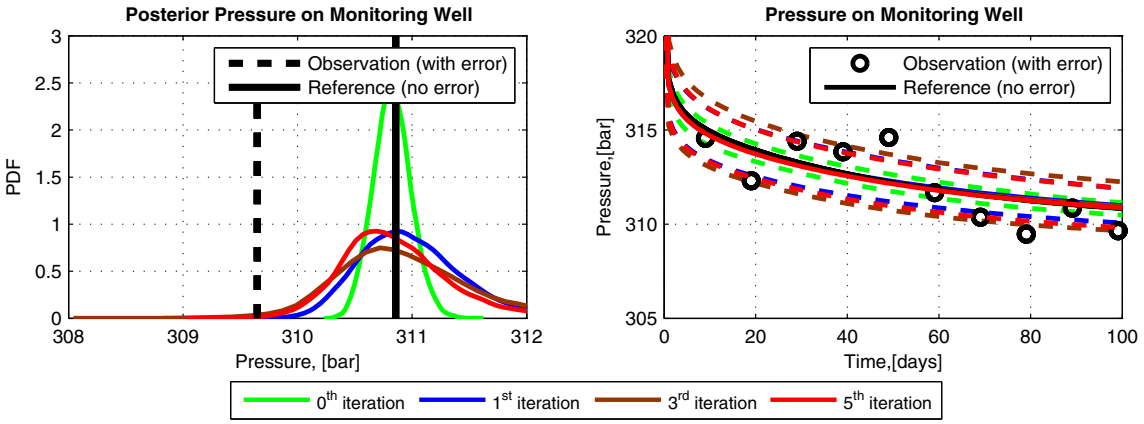

Fig. 11 Posterior pressure at monitoring well: pressure distribution after 100 days of injection (*left plot*) and corresponding matching of pressure (mean $\pm$ 2 standard deviations) to observation values during 100 days (*right plot*)

available computational resources. Evidently, a large number of additional integration points will lead to a more robust projection in the region of the possible posterior. However, it is not guaranteed that a large effort for local refinement will immediately lead to final results, because the refinement can be placed poorly in the parameter space at early iteration steps. Thus, we suggest to use a moderate number of new integration points in each iteration step.

Once the response surface is improved by adding new collocation points, the history matching using bootstrap filtering is applied again. This procedure is repeated until the matched model matches the observation values with a desired precision or until no significant changes can be observed between the iterations anymore. Overall, the iterative accumulation of collocation points leads to an efficient estimation of the posterior distribution. This refinement idea could also be extended to different methods which use response surfaces for inverse modeling (e.g., EnKF or MCMC combined with PCE, etc.).

In case study 2, we choose a first-order refinement for three parameters and add four highly probable integration points out of the eight available ones in each iteration step. The results of our iterations are presented in Figs. 10 and 11 by green, blue, brown, and red lines for the zeroth, first, third, and fifth iterations, respectively. According to our observations, the posterior improves quite fast (with only a few iterations). In our case study, we stopped at the fifth iteration, because all posterior distributions were stabilized, i.e., the posterior distributions in the ith iteration did not differ significantly from the posteriors in the $(i-1)$th iteration.

After matching the model via the iterative version of aPC-based bootstrap filtering, we perform a forecast. As in Section 4, we predict the CDF of the CO_2 leakage rate after 1,000 days. The posterior CDF (after the fifth iteration) is significantly different from the prior CDF (see Fig. 12) and also differs strongly from the results after the zeroth and first iterations.

The question on how to optimize the overall projection in the iteration approach and, in more general, under the changing probability measure from prior to posterior is a more general and wider challenge than what can be addressed within the scope of the study. Such a procedure should find a balance between the number of iterations and the number of new integration points per iteration. If the prior assumption is a strong offset against the posterior, then the computational effort for a very accurate expansion according to the prior distribution would be spent almost in vain. To overcome this drawback, the iterative framework presented here could first perform a cheap expansion using a low order. Then, using the flexibility of aPC, one could construct expansions of higher orders during the iterations, as the knowledge on the posterior improves and the increasing set of collocation points becomes more trustworthy to provide an accurate high-order response surface in the current alleged posterior region. Altogether, this poses significant challenges for future research.

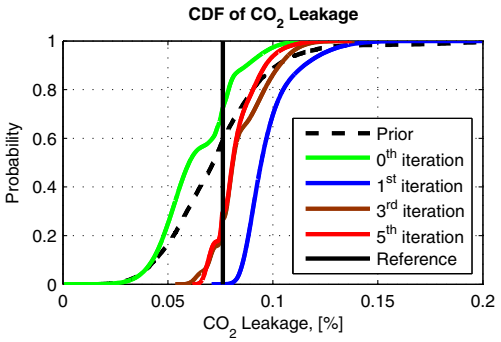

Fig. 12 Cumulative density function of CO_2 leakage: prior and posterior prediction during iterations after 1,000 days

6 Summary and conclusions

The present paper deals with history matching of mathematical models to observation values. We propose an advanced framework for history matching based on response surface attained via the PCE and strict Bayesian principles. The reduced model represented by the response surface captures the model's dependence on its three physically most relevant parameters. For constructing the response surface, we used the aPC, a recent generalization of PCE theory, because it has very good convergence properties and tends to represent efficiently the underlying model according to the available statistical information. The aPC allows accommodating for a wide range of prior distributions of model parameters. The applied technique for projection onto the polynomial basis is totally nonintrusive, i.e., it is black-box compatible with arbitrary commercial simulation codes. Then, we perform Bayesian updating via bootstrap filtering in order to match the obtained reduced polynomial model to past or real-time observations of system behavior. Basically, with bootstrap filtering, we follow a direct implementation of Bayes' theorem.

The combination of high-order expansion and bootstrap filtering accounts for the nonlinearity of both the forward model and of the inversion. It takes into consideration higher order statistical moments in comparison to (ensemble) Kalman filters. Hence, our method is more accurate than linearized inversion rules or related second-order moment approaches based on implicit lineralizations or on multi-Gaussian assumptions. The usually high computational costs of accurate filtering become very feasible in our suggested method by combining it with a response surface framework. Thanks to the computational efficiency of the aPC, Bayesian updating for history matching becomes an interactive task and could even be applied to real-time problems with complex large-scale, real-world models in future works. The key contribution of our approach is that the response surface can be prepared in expansive offline computation and can serve for Bayesian updating to new incoming data within seconds.

The efficiency and power of Bayesian updating strongly depend on the accuracy of prior information. In our aPC-based methodology, the model parameter distributions can be determined from arbitrary available information (modeler's experience, expert opinion, general prior information, or field data) and reflect the uncertainty or expected range of variation of input parameters. The polynomial basis of the aPC is able to adapt to arbitrary shapes of these parameter distributions. The presented methodology approximates the original model best where the prior probability density of the parameters is highest. In the case where the prior guess is highly inaccurate and strongly offsets against the posterior, we suggest to use an iterative procedure, which helps to overcome this drawback even with small costs. We propose to account for the posterior on each iteration step and increase the precision of expansion around the current iteration of the posterior distribution.

A direct and straightforward application example of the proposed methodology is illustrated using a CO_2 benchmark problem. In this example, we found highly satisfying accuracy and computational efficiency. However, our methodology is not restricted to this example, as both the polynomial chaos expansion and bootstrap filtering require no specific properties for the forward model or the inversion task. The only restriction is that the forward model has to be approximated well by the aPC.

Acknowledgments The authors would like to thank the German Research Foundation (DFG) for its financial support of the project within the Cluster of Excellence in Simulation Technology (EXC 310/1) at the University of Stuttgart. The authors would like to express their thanks to Andeas Geiges at the Department of Hydromechanics and Modeling of Hydrosystems, University of Stuttgart, for his help with the multivariate kernel density estimation.

References

1. Aanonsen, S.I., Naevdal, G., Reynolds, A.C., Valls, B.: The ensemble Kalman filter in reservoir engineering—a review. SPE J. **14**(3), 393–412 (2009)
2. Babuska, I., Nobile, F., Tempone, R.: A stochastic collocation method for elliptic partial differential equations with random input data. SIAM J. Numer. Anal. **45**(3), 1005–1034 (2007)
3. Bangerth, W., Klie, H., Wheeler, M., Stoffa, P., Sen, M.: On optimization algorithms for the reservoir oil well placement problem. Comput. Geosci. **10**(3), 303–319 (2006)
4. Birkholzer, J.T., Zhou, Q., Tsang, Ch.-F.: Large-scale impact of CO_2 storage in deep saline aquifers: a sensitivity study on pressure response in stratified systems. Int. J. of Greenh. Gas Control. **3**, 181–194 (2009)
5. Blatman, G., Sudret, B.: Efficient computation of global sensitivity indices using sparse polynomial chaos expansions. Reliab. Eng. Syst. Saf. **95**, 1216–1229 (2010)
6. Chen, J.-S., Wang, L., Hu, H.-Y., Chi, S.-W.: Subdomain radial basis collocation method for heterogeneous media. Int. J. Numer. Meth. Engng. **80**, 163–190 (2009)
7. Class, H., Ebigbo, A., Helmig, R., Dahle, H., Nordbotten, J.N., Celia, M.A., Audigane, P., Darcis, M., Ennis-King, J., Fan, Y., Flemisch, B., Gasda, S., Jin, M., Krug, S., Labregere, D., Naderi, A., Pawar, R.J., Sbai, A., Sunil, G.T., Trenty, L., Wei, L.: A benchmark study on problems related to CO_2 storage in geologic formations. Comput. Geosci. **13**, 451–467 (2009)
8. Cominelli, A., Ferdinandi, F., de Montleau, P.C., Rossi, R.: Using gradients to refine parameterization in field-case history-matching projects. SPE Reserv. Evalu. Eng. **10**(3), 233–240 (2007)
9. Cortis, A., Oldenburg, C., Benson, S.M.: The role of optimality in characterizing CO_2 seepage from geologic carbon sequestration sites. Int. J. of Greenh. Gas Control. **2**, 640–652 (2008)
10. Crestaux, T., Le Maitre, O., Martinez, J.-M.: Polynomial chaos expansion for sensitivity analysis. Reliab. Eng. Syst. Saf. **94**(7), 1161–1172 (2009)

11. Daoud, A.M.I.: Automatic history matching in Bayesian framework for field-scale applications. Dissertation. Texas A&M University (2004)
12. Ebigbo, A., Class, H., Helmig, R.: CO_2 leakage through an abandoned well: problem-oriented benchmarks. Comput. Geosci. **11**(2), 103–115 (2007)
13. Efron, B., Tibshirani, R.J.: An Introduction to the Bootstrap (Monographs on Statistics & Applied Probability). Chapman & Hall, London (2010)
14. Evensen, G.: Data Assimilation: the Ensemble Kalman Filter, Berlin (2006)
15. Ewing, R.E., Pilant, M.S., Wade, G.J., A.T. Watson: Estimating parameters in scientific computation. IEEE Comput. Sci. Eng. **1**(3), 19–31 (1994)
16. Fajraoui, N., Ramasomanana, F., Younes, A., Mara, T.A., Ackerer, P., Guadagnini, A.: Use of global sensitivity analysis and polynomial chaos expansion for interpretation of non-reactive transport experiments in laboratory-scale porous media. Water Resour. Res. (2011). doi:10.1029/2010WR009639
17. Feraille, M., Marrel, A.: Prediction under uncertainty on a mature field. Oil and Gas Science and Technology-Revue dIFP Energies nouvelles **67**(2), 193–206 (2012)
18. Flemisch, B., Fritz, J., Helmig, R., Niessner, J., Wohlmuth, B. In: Ibrahimbegovic, A., Dias, F. (eds.): ECCO3MAS Thematic Conference on Multi-scale Computational Methods for Solids and Fluids, Cachan, France, 28–30 November 2007
19. Foo, J., Karniadakis, E.G.: Multi-element probabilistic collocation method in high dimensions. J. Comput. Phys. **229**(5), 1536–1557 (2010)
20. Gao, G., Reynolds, A.C.: An improved implementation of the LBFGS algorithm for automatic history matching. SPE J. **11**(1), 5–17 (2006)
21. Gao, G., Reynolds, A.C.: A stochastic optimization algorithm for automatic history matching. SPE J. **12**(2), 196–208 (2007)
22. Gavalas, G.R., Shah, P.C., Seinfeld, J.H.: Reservoir history matching by Bayesian estimation. SPE J. **16**(6), 337–350 (1976)
23. Ghanem, R., Doostan, A.: On the construction and analysis of stochastic models: characterization and propagation of the errors associated with limited data. J. Comput. Phys. **217**, 63–81 (2006)
24. Ghanem, R., Spanos, P.: A stochastic Galerkin expansion for nonlinear random vibration analysis. Probab. Eng. Mec. **8**, 255–264 (1993)
25. Ghanem, R., Spanos, P.D.: Polynomial chaos in stochastic finite elements. J. Appl. Mech. **57**, 197–202 (1990)
26. Ghanem, R.G., Spanos, P.D.: Stochastic finite elements: a spectral approach. Springer (1991)
27. Gilks, W.R., Richardson, S., Spiegelhalter, D.J.: Markov chain Monte Carlo in practice. Chapman & Hall, London (1996)
28. Gu, Y., Oliver, D.S.: An iterative ensemble Kalman filter for multiphase fluid flow data assimilation. SPE J **12**(4), 438–446 (2007)
29. Hansson, A., Bryngelsson, M.: Expert opinions on carbon dioxide capture and storage: a framing of uncertainties and possibilities. Energy Policy **37**, 2273–2282 (2009)
30. He, J., Sarma, P., Durlofsky, L.J., Chen, W.: Use of reduced-order models for improved data assimilation within an EnKF context. Presented in Reservoir Simulation Symposium, The Woodlands, Texas, USA, SPE 141967 (2011)
31. Hendricks Franssen, H.-J., Kinzelbach, W.: Ensemble Kalman filtering versus sequential self-calibration for inverse modelling of dynamic groundwater flow systems. J. Hydrol. **365**(3–4), 261–274 (2009)
32. IPCC: Special Report on Carbon Dioxide Capture and Storage. Technical Report, Intergovernmental Panel on Climate Change (IPCC), Prepared by Working Group III. Cambridge University Press, Cambridge (2005)
33. Jin, B.: Fast Bayesian approach for parameter estimation. Int. J. Numer. Meth. Engng. **76**, 230–252 (2008)
34. Kitanidis, P.K.: Quasi-linear geostatistical theory for inversing. Water Resour. Res. **31**(10), 2411–2419 (1995)
35. Kopp, A., Class, H., Helmig, H.: Investigations on CO_2 storage capacity in saline aquifers—part 1: dimensional analysis of flow processes and reservoir characteristics. Int. J. Greenh. Gas Control. **3**, 263–276 (2009)
36. Helmig, C., Seinfeld, J.H.: Identification of parameters in distributed parameter-systems by regularization. SIAM J. Control Optim. **23**(2), 217–241 (1985)
37. Le Maitre, O., Knio, O.: Spectral Methods for Uncertainty Quantification: with Applications to Computational Fluid Dynamics, New York (2010)
38. Leube, P., Geiges, A., Nowak, W.: Bayesian assessment of the expected data impact on prediction confidence in optimal sampling design. Water Resour. Res. **48**, W02501 (2012)
39. Leube, P.C., Nowak, W., Schneider, G.: Temporal moments revisited: why there is no better way for physically-based model reduction in time. Water Resour. Res. (2012). doi:10.1029/2012WR011973
40. Li, H., Sarma, P., Zhang, D.: A comparative study of the probabilistic collocation and experimental design methods for petroleum reservoir uncertainty quantification. SPE J. **16**, 429–439 (2011). SPE–140738–PA–P
41. Li, H., Zhang, D.: Probabilistic collocation method for flow in porous media: comparisons with other stochastic methods. Water Resour. Res. **43**, 44–48 (2007)
42. Li, R., Reynolds, A.C., Oliver, D.S.: Sensitivity coefficients for three-phase flow history matching. J. Can. Pet. Technol. **42**(4), 70–77 (2003)
43. Lia, O., Omre, H., Tjelmeland, H., Holden, L., Egeland, T.: Uncertainties in reservoir production forecasts. AAPG Bull. **81**(5), 775–802 (1997)
44. Liou, Ch.-L., Lin, Ch.-H.: Applications of the methods of weighted residuals in system science. Int. J. Syst. Sci. **22**(9), 1509–1525 (1991)
45. Liu, N., Oliver, D.S.: Ensemble Kalman filter for automatic history matching of geologic facies. J. Pet. Sci. Eng. **47**, 147–161 (2005)
46. Makhlouf, E.M., Chen, M.L., Wasserman, W.H., Seinfeld, J.H.: A general history matching algorithm for three-phase, three-dimensional petroleum reservoirs. SPE Adv. Technol. **1**(2), 83–91 (1993)
47. Maltz, F.H., Hitzl, D.L.: Variance reduction in Monte Carlo computations using multi-dimensional Hermite polynomials. Journal Comput. Phys. **2**, 345–376 (1979)
48. Marzouk, Y., Najm, H., Rahn, L.: Stochastic spectral methods for efficient Bayesian solution of inverse problems. J. Comput. Phys. **224**(2), 560–586 (2007)
49. Morariu, V.I., Srinivasan, B.V., Raykar, V.C., Duraiswami, R., Davis, L.S.: Automatic online tuning for fast gaussian summation. Adv. Neural. Inform. Process. Syst. **21**, 1113–1120 (2009)
50. Moritz, H.: Least-squares collocation. Rev. Geophys. Space Phys. **16**(3), 421–430 (1978)
51. Naevdal, G., Hanea, R.G., Oliver, D.S., Valles, B.: Ensemble Kalman filter for model updating—a special issue. Comput. Geosci. **15**(2), 223–224 (2011)
52. Naevdal, G., Johnsen, L.M., Aanonsen, S.I., Vefring, E.H.: Reservoir monitoring and continuous model updating using ensemble Kalman filter. SPE J. **10**(1), 66–74 (2005)
53. Nordbotten, J., Celia, M., Bachu, M.: Injection and storage of CO_2 in deep saline aquifers: analytical solution for CO_2 plume evolution during injection. Transp. Porous Media **58**(3), 339–360 (2005)
54. Nordbotten, J.M., Kavetski, D., Celia, M.A., Bachu, S.: A semi-analytical model estimating leakage associated with CO_2 storage

in large-scale multi-layered geological systems with multiple leaky wells. Environ. Sci. Technol. **43**(3), 743–749 (2009)
55. Nowak, W.: Best unbiased ensemble linearization and the quasi-linear Kalman ensemble generator. Water Resour. Res. **45**, W04431 (2009)
56. Nowak, W., Rubin, Y., de Barros, F.P.J.: A hypothesis-driven approach to optimal site investigation. Water Resour. Res. (2012). doi:10.1029/2011WR011016
57. Oladyshkin, S., Class, H., Helmig, R., Nowak, W.: A concept for data-driven uncertainty quantification and its application to carbon dioxide storage in geological formations. Adv. Water Resour. **34**, 1508–1518 (2011)
58. Oladyshkin, S., Oladyshkin, H., Helmig, R., Nowak, W.: An integrative approach to robust design and probabilistic risk assessment for CO_2 storage in geological formations. Comput. Geosci. **15**(3), 565–577 (2011)
59. Oladyshkin, S., de Barros, F.P.J., Nowak, W.: Global sensitivity analysis: a flexible and efficient framework with an example from stochastic hydrogeology. Adv. Water Resour. **37**, 10–22 (2011)
60. Oladyshkin, S., Nowak, W.: Data-driven uncertainty quantification using the arbitrary polynomial chaos expansion. Reliab. Eng. Syst. Saf. **106**, 179–190 (2012)
61. Oladyshkin, S., Nowak, W.: Polynomial Response Surfaces for Probabilistic Risk Assessment and Risk Control via Robust Design. Novel Approaches and Their Applications in Risk. InTech, Manhattan (2012)
62. Oliver, D.S., Reynolds, A.C., Liu, N.: Inverse theory for petroleum reservoir characterization and history matching. Cambridge (2008)
63. Oliver, D.S., Chen, Y.: Recent progress on reservoir history matching: a review. Comput. Geosci. **15**, 185–221 (2011)
64. Oliver, D.S., Reynolds, A.C., Bi, Z., Abacioglu, Y.: Integration of production data into reservoir models. Pet. Geosci. **7**, 65–73 (2001)
65. Slotte, P.A., Smrgrav, E.: Response surface methodology approach for history matching and uncertainty assessment of reservoir simulation models. Europec/EAGE Conference and Exhibition, SPE 113390, Rome, Italy, 9–12 June 2008
66. Wever, U., Paffrath, M.: Adapted polynomial chaos expansion for failure detection. J. Comput. Phys. **226**(1), 263–281 (2007)
67. Pajonk, O., Rosic, B.V., Litvinenko, A., Matthies, H.G.: A deterministic filter for non-Gaussian Bayesian estimation applications to dynamical system estimation with noisy measurements. Physica. D. **241**, 775–788 (2012)
68. Pappenberger, F., Beven, K.J.: Ignorance is bliss: or seven reasons not to use uncertainty analysis. Water Resour. Res. **42**(5), 1–8 (2006)
69. Robert, C.P., Casella, G.: Monte Carlo methods. Springer, New York (2004)
70. Rodrigues, J.R.P.: Calculating derivatives for automatic history matching. Comput. Geosci. **10**, 119–136 (2006)
71. Saad, G., Ghanem, R.: Characterization of reservoir simulation models using a polynomial chaos-based ensemble Kalman filter. Water Resour. Res. **45**, W04417 (2009)
72. Saltelli, A., Ratto, M., Andres, T.: Global Sensitivity Analysis: the Primer. Wiley, New York (2008)
73. Sarma, P., Durlofsky, L.J., Aziz, K.: Kernel principal component analysis for efficient, differentiable parameterization of multipoint geostatistics. Math. Geosci. **40**(1), 3–32 (2008)
74. Sarma, P., Durlofsky, L.J., Aziz, K., Chen, W.H.: Efficient real-time reservoir management using adjoint-based optimal control and model updating. Comput. Geosci. **10**(1), 3–36 (2006)
75. Scheidt, C., Caers, J., Chen, Y., Durlofsky, L.J.: A multi-resolution workflow to generate high-resolution models constrained to dynamic data. Comput. Geosci. **15**(3), 545–563 (2011)
76. Schoeniger, A., Nowak, W., Hendricks Franssen, H.-J.: Parameter estimation by ensemble Kalman filters with transformed data: Approach and application to hydraulic tomography. Water Resour. Res. **45**, W04431 (2012)
77. Smith, A.F.M., Gefland, A.E.: Bayesian statistics without tears: a sampling-resampling perspective. Am. Stat. **46**(2), 84–88 (1992)
78. Soize, C., Ghanem, R.: Physical systems with random uncertainties: chaos representations with arbitrary probability measure. SIAM J. Sci. Comput **26**(2), 395–410 (2004)
79. Sudret, B.: Global sensitivity analysis using polynomial chaos expansions. Reliab. Eng. Syst. Saf. **93**(7), 964–979 (2008)
80. Sun, N.Z.: Inverse Problems in Groundwater Modeling (Theory and Applications of Transport in Porous Media), New York (1999)
81. Tarantola, A.: Inverse Problem Theory and Methods for Model Parameter Estimation. SIAM, Philadelphia (2005)
82. Villadsen, J., Michelsen, M.L.: Solution of Differential Equation Models by Polynomial Approximation. Prentice-Hall, Upper Saddle River (1978)
83. Vrugt, A.J., ter Braak, C.J.F., Diks, C.G.H., Robinson, B.A., Hyman, J.H., Higdon, D.: Accelerating Markov chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling. Int. J. Nonlinear Sci. Numer. Simul. **10**(3), 271–288 (2009)
84. Walter, L., Binning, P., Oladyshkin, S., Flemisch, B., Class, H.: Int. J. Greenh. Gas. Control. **9**(495–506) (2012)
85. Wan, X., Karniadakis, E.G.: Multi-element generalized polynomial chaos for arbitrary probability measures. SIAM J. Sci. Comput. **28**(3), 901–928 (2006)
86. Wand, M.P., Jones, M.C.: Kernel smoothing. Monographs on Statistics and Applied Probability 60. Chapman & Hall, Boca Raton (1995)
87. Wang, Y., Li, G., Reynolds, A.C.: Estimation of depths of fluid contacts by history matching using iterative ensemble Kalman smoothers. SPE J. **15**(2), 509–529 (2010)
88. Wiener, N.: The homogeneous chaos. Am. J. Math **60**, 897–936 (1938)
89. Wildenborg, A.F.B., Leijnse, A.L., Kreft, E., Nepveu, M.N., Obdam, A.N.M., Orlic, B., Wipfler, E.L., van Kesteren, W., van der Grift, B., Gaus, I., Czernichowski-Lauriol, I., Torfs, P., Wojcik, R.: Risk assessment methodology for CO_2 storage: the scenario approach. In: Thomas, D.C., Benson, S.M. (eds.) Carbon Dioxide Capture for Storage in Deep Geologic Formations. Elsevier, London (2005)
90. Williams, M.A., Keating, J.F., Barghouty, M.F.: The stratigraphic method: a structured approach to history-matching complex simulation models. SPE Reserv. Evalu. Eng. **1**(2), 169–176 (1998)
91. Witteveen, J.A.S., Bijl, H.: Modeling arbitrary uncertainties using Gram-Schmidt polynomial chaos. 44th AIAA Aerospace Sciences Meeting and Exhibit, Reno, Nevada, AIAA–2006–896 (2006)
92. Witteveen, J.A.S., Sarkar, S., Bijl, H.: Modeling physical uncertainties in dynamic stall induced fluidstructure interaction of turbine blades using arbitrary polynomial chaos. Comput. Struct. **85**, 866–878 (2007)
93. Xiu, D., Karniadakis, E.G.: Modeling uncertainty in flow simulations via generalized polynomial chaos. J. Comput. Phys. **187**, 137–167 (2003)
94. Zabalza-Mezghani, I., Manceau, E., Feraille, M., Jourdan, A.: Uncertainty management: from geological scenarios to production scheme optimization. J. Pet. Sci. Eng. **44**, 11–25 (2004)
95. Zafari, M., Reynolds, A.C.: Assessing the uncertainty in reservoir description and performance predictions with the ensemble Kalman filter. SPE J. **12**(3), 382–391 (2007)
96. Zhang, D., Lu, Z.: An efficient, high-order perturbation approach for flow in random media via Karhunen-Loève and polynomial expansions. J. Comput. Phys. **194**, 773–794 (2004)