Probabilistic Assessment of Above Zone Pressure Predictions at a Geologic Carbon Storage Site

Argha Namhata[1, 2, *], Sergey Oladyshkin[3], Robert M. Dilmore[2], Liwei Zhang[2], David V. Nakles[1]

1 Department of Civil & Environmental Engineering, Carnegie Mellon University, Pittsburgh, PA 15217, USA

2 U.S. Department of Energy, National Energy Technology Laboratory, 626 Cochrans Mill Road, Pittsburgh, Pennsylvania 15236, USA

3 Department of Stochastic Simulation and Safety Research for Hydrosystems (IWS/SRC SimTech), University of Stuttgart, Germany

** Corresponding Author: Email: anamhata@andrew.cmu.edu* (A. Namhata), *Tel: +1 (412) 370-4671.*

ABSTRACT

Carbon dioxide (CO_2) storage into geological formations is regarded as an important mitigation strategy for anthropogenic CO_2 emissions to the atmosphere. This study first simulates the leakage of CO_2 and brine from a storage reservoir through the caprock. Then, we estimate the resulting pressure changes at the zone overlying the caprock also known as Above Zone Monitoring Interval (AZMI). A data-driven approach of arbitrary Polynomial Chaos (aPC) Expansion is then used to quantify the uncertainty in the above zone pressure prediction based on the uncertainties in different geologic parameters. Finally, a global sensitivity analysis is performed with Sobol indices based on the aPC technique to determine the relative importance of different parameters on pressure prediction. The results indicate that there can be uncertainty in pressure prediction locally around the leakage zones. The degree of such uncertainty in prediction depends on the quality of site specific information available for analysis. The scientific results from this study provide substantial insight that there is a need for site-specific data for efficient predictions of risks associated with storage activities. The presented approach can provide a basis of optimized pressure based monitoring network design at carbon storage sites.

INTRODUCTION

Capture and geologic storage of carbon dioxide (CO_2) is considered as one of a portfolio of solutions for the reduction of anthropogenic greenhouse gas emissions. The increasing emphasis on the commercialization and implementation of CO_2 capture and storage (CCS) has led to the development of system-wide mathematical models for the quantitative assessment of system performance and the risk associated with it. A major technical and regulatory concern that has gained attention of the research community is the unanticipated leakage of CO_2 and brine from deep storage reservoirs to overlying geologic formations such as the Above Zone Monitoring Interval (AZMI) and groundwater aquifers[1,2,3], through preferential migration pathways such as, wellbores, faults, fractures and presence of high-permeability zones in the caprocks. While the leaked CO_2 and brine may possess threat to the environmental receptors, it is also possible that it would attenuate pressure and CO_2 saturation. The knowledge of changes in pressure helps for adequate management of the reservoir. Hence, it becomes important to monitor the above zone pressure, since it may provide potentially useful source of information about

seal performance and subsurface pressure response to CO_2 and brine leakage from the storage reservoir.

The United States Department of Energy (DOE) through its National Risk Assessment Partnership (NRAP)[4,5], is conducting research to develop and demonstrate science-based methodologies to quantify the environmental risks associated with long-term geologic storage of CO_2[6]. Central to this research is the development of Integrated Assessment Models, or IAMs to describe site-scale performance of geologic storage systems. These IAMs are system-based models that simulate and couple the primary sub-system components of the storage system, i.e., storage reservoir, migration pathways (i.e., seals, wellbores, faults and fractures), groundwater, and atmosphere, with the goal of predicting potential leakage performance/storage security through the period of active CO_2 injection, and post-injection site care. Since the integration of fully characterized numerical models of individual sub-system into an IAM is both challenging and computationally expensive, the NRAP approach calls for modeling the sub-system components using simplified reduced order characterizations, or reduced order models (ROMs), that are much more computationally efficient[3,5,7-13].

One such ROM development effort has focused on characterizing the CO_2 and brine leakage through the primary sealing layer to the interval directly overlying that seal (the AZMI). A ROM developed and previously reported by Namhata and coworkers[14] predicts spatially-varying changes in pressure through time in response to that fluid leakage. The AZMI ROM is believed to provide a useful approximation of real-world response, but also includes a number of conceptual and quantitative uncertainties. Insufficient or lack of information related to geological properties represents one important source of parameter uncertainty that may lead to significant uncertainties in model predictions, with potential to mask the influence of secondary physical processes[15]. Because full-physics numerical simulation models are computationally expensive, it may require several hours to days to complete a single, deterministic realization; as such, exploring uncertainty/variability in system performance using such models and brute-force Monte Carlo simulation is generally considered intractable[16-20]. This makes it favorable to use advanced stochastic tools to model uncertainties of complicated processes involved in the geologic storage of carbon modeling. It also holds true for a coupled ROM approach. Application of advanced stochastic tools to predict uncertainties in coupled ROM systems like the reservoir-

caprock-AZMI coupled model used in this study will be computationally efficient over a complex Monte Carlo like analysis.

In this study, we use a recently developed data-driven uncertainty quantification approach, called the arbitrary polynomial chaos (aPC) expansion that provides a massive stochastic model reduction[15,21] to analyze the uncertainties in predictive ability of the AZMI ROM. aPC has certain advantages over more conventional polynomial chaos methods. This approach provides a more robust convergence[21] in comparison to the classical methods (e.g., Wiener, 1938; Ghanem and Spanos, 1993; Le Maître and Knio, 2010)[22-24] once underlying distributions of uncertain parameters dictated by real-world data; it also allows for use of arbitrary probability distributions of uncertain parameters[21]. The more complex the system is, the greater will be the associated uncertainty of the system models. Uncertainty of any parameter in the modeling procedure propagates through the model to impact the model predictions. Hence, it is important to rank the influence of the model input parameters on the output space. This aids in better understanding the system behavior, adding value to the task of analyzing model uncertainties and sensitivities.

Sensitivity analysis is widely used to identify the contribution of uncertainty sources within the modeling process[21] and that in turn helps in improving the understanding of model behavior[25]. We performed global sensitivity analysis (GSA) using variance-based Sobol sensitivity index parameterization[25]. The motivation to GSA over a local sensitivity analysis approach is that local analysis is unable to cover the non-linear variation of model responses over the entire range of probability distributions of the input parameters[26]. The aim of GSA is to quantify the relative importance of each individual input parameter on model output prediction, and rank those parameters by importance. The aPC-based response surface used in the uncertainty quantification is based on orthonormal polynomials whose properties are well exploited[27]. The goal of this study is to probabilistically assess the role of various geologic parameters in AZMI pressure predictions.

RESULTS AND DISCUSSIONS

Above Zone Model Setup

In this modeling effort, we aim to study the migration of subsurface fluids (here, CO_2 and brine) to the AZMI and the resulting changes in pressure. The modeled system comprises three components: reservoir, caprock and the AZMI. The calculations in reservoir and caprock are necessary to model the pressure changes in the AZMI. This study demonstrates the application of above zone pressure modeling using the AZMI ROM by using the Kimberlina CO_2 storage site (California, USA) as an example [28-31]. The reservoir-scale CO_2 migration model developed by Wainwright et al., 2013[30] is based on a geological study in the Southern San Joaquin Basin, California. The model uses geologic and hydrogeologic data obtained from many oil fields in that region. The model domain extends 71.3 km in the eastern direction and 91.6 km in the northern direction as shown in Figure 1. The simulation assumes that CO_2 injection is conducted in the center of the domain into the 400-m thick and at about 2750 m deep Vedder formation (depth is at top). The Vedder formation is quite permeable which should allow large industrial scale fluid injectivity. The injection well location is also marked in the Figure 1. Since we intend to use the reservoir simulator results in seal ROM, NSealR[7], the reservoir area is converted into 100 by 100 grid block system, for consistency. The location of the conceptual injection well is at coordinate (34, 46) in that reduced-resolution spatial domain. The overlying Temblor–Freeman shale with a nominal thickness of 200 m is considered a suitable caprock for stratigraphic containment of the supercritical CO_2 injected into the underlying Vedder formation. This storage formation site model is used because there are considerable data available. In this paper, we use the modified Kimberlina model of Wainwright et al., 2013[30] as used by Pawar and co-authors[4] to simulate the reservoir pressure and saturations. A hypothetical scenario is assumed where 5 million tons of CO_2 is injected per year for a period of 50 years. There are several faults known to be present in the reservoir. Fault zone properties are quite uncertain; however, there are qualitative observations that most fault zones are less conductive than the adjacent sandstone formations[17]. In the reservoir simulations, the potential for leakage of CO_2 and/or brine through permeable faults and/ or fractures has been ignored, i.e., the seal is characterized with realizations of uniform, but varying permeability. Also, the potential for fault reactivation in response to fluid injection is not addressed. Since there is no information about the fault zones, we just assumed that there are no faults transcending into the caprock and/or above layers. Supplementary Figs. S1 and S2 show the pressure and saturation at the reservoir-seal interface, respectively. It can be seen from Fig. S1 that the pressure increase spreads from the injection zone to the boundaries of the domain during

the 50 year CO_2 injection period, and the pressure increase gradually decreases once the injection stops. Fig. S2 shows that when compared to pressure, the CO_2 plume is more localized, and the observable increase in CO_2 saturation can only be seen right above the injection zone. After the reservoir simulation, we use the NRAP Seal Barrier ROM, NSealR[7] to compute the migration of CO_2 and brine through the seal to overlying AZMI formation through intrinsic permeability and/or the presence of natural/induced fractures in the seal. NSealR uses a two-phase, relative permeability approach with Darcy's law for one-dimension (1-D) flow computations of CO_2 through the horizon in the vertical direction. The reservoir pressure and saturation generated using the Kimberlina model[30] is used as an input to NSealR to produce CO_2 and brine flux from top of the seal in a 100 by 100 uniform grid format. The CO_2 flux through the 200-m thick Temblor–Freeman shale calculated using NSealR is shown in Supplementary Fig. S3. Although the case considered assumes a single thickness seal, NSealR allows for spatially varying thickness and effective permeability.

The AZMI ROM used in this study to predict above zone pressure changes due to leakage through the primary seal has been previously described by Namhata et al., 2016[14]. A hypothetical AZMI system for Kimberlina is defined for the model analysis. This conceptual base case system consists of a 10-m thick AZMI layer overlying the aforementioned Temblor–Freeman shale of thickness 200 m. The AZMI formation features have been derived from the existing Olcese sandstone which overlies the Temblor-Freeman shale. It is assumed that the AZMI is initially fully saturated with brine. The reference parameters used for the Kimberlina site in this model are taken from Wainwright et al., 2013[30] and are shown in Table 1. Relative permeability has important implications for fluid flow in subsurface geological systems. The Brooks-Corey[32] model has been used to define relative permeability in the AZMI ROM. The relative permeability curve used for this study has been shown in the Supplementary Fig. S4.

Figure 2 presents the changes in pressure response in the AZMI over time generated using flux from the seal for the simulation periods previously discussed. The highest increase in pressure in the AZMI is observed above the injection point at the end of injection (i.e., 50 years), with the maximum predicted increase in pressure of 0.185 MPa. The change in pressure gradually decreases away from the injection point location. After CO_2 injection stops, the rate of increase in pressure abruptly stops with the termination of

injection, followed by an initial fast rate of pressure decrease that gradually slows and approaches a stable, but positive net pressure change by the end of the simulation (i.e., following a post-injection period of 150 years). Supplementary Fig. S5 shows the time evolution of pressure change for the base case.

Uncertainty Quantification of Above Zone Pressure

Many geologic parameters are influential in predicting the CO_2 and brine flow dynamics in the AZMI ROM. Due to lack of information stemming from a limited ability to make direct measurements, parameters such as porosity and permeability, to name two, often remain uncertain[33,34]. These uncertainties can have a substantial effect on the output of the ROM. Thus, a quantitative analysis of the impact of these uncertainties on the predictive capabilities of the model was performed and is presented in this section.

A model-based uncertainty analysis, though efficient, requires statistical data for all of the model input parameters, which increases the demand on data availability or results in highly subjective assumptions to deal with missing data[15]. Uncertainties in complex systems can be efficiently and accurately quantified using stochastic models based on an approach using data-driven polynomial chaos expansion (PCE) methods[35-38]. Thus, the uncertainty quantification of the AZMI ROM was performed using the arbitrary Polynomial Chaos (aPC) approach[15,21]. In aPC, the statistical moments are the only source of information required to define the stochastic parameters. Hence, accurate descriptions of the probability density functions (PDF) of the uncertain parameters are not required to perform the analysis.

Statistical Distribution of Input Parameters

The data-driven aPC method, as described in the Methods section, only requires information on finite number of moments, and does not explicitly require the shapes of probability density functions. The arbitrary distributions can be either discrete, continuous, or discretized continuous and can be specified either through a few statistical moments, analytically as PDF/CDF, numerically as a histogram, or theoretically through the even more general format of a probability measure[15].

In this study, the uncertainty analysis was performed for five input parameters: AZMI permeability (k_{AZMI}), AZMI porosity (Φ_{AZMI}), thickness of AZMI (H_{AZMI}), caprock permeability ($k_{caprock}$) and caprock thickness ($H_{caprock}$). Though the study is site specific, caprock and AZMI thickness are considered to be a part of the list of uncertain parameters. The purpose of using these parameters is to have an idea of how the uncertainty will be in a generalized setup in case there is not enough information about the formation thickness. Also, the thickness of formation usually doesn't tend to be uniform throughout storage sites and hence this assessment will be helpful in finding out the effect of thickness on pressure buildup. Figure 3 demonstrates the stochastically generated distributions of the parameters that have been used in the analysis. The data distribution pattern is generated based on data available from the US National Petroleum Council Public Database[38,39].

AZMI Output Statistics

We analyze these statistical moments for the AZMI ROM for the simulation period of 200 years. Mean and standard deviation of changes in pressure response above the AZMI over time is shown in Figures 4 (a) and 5 respectively, based on the uncertainty of the five input parameters. A total of $56 = (5+3)!/(5! \times 3!)$ detailed simulations (see equation (3)) is carried out to generate the uncertainties in model outputs based on aPC framework. It can be seen from Figure 4 (a) that the mean of pressure buildup above the AZMI from the aPC simulations is approximately 0.50 MPa higher than that of the base case scenario in Figure 2. Since the highest-pressure buildup above the AZMI occurs right above the injection well location, we checked the variation in pressure change output from the entire set of simulations to that of the calculated mean. The analysis is shown in Figure 4 (b) by plotting the range of predictions from the 56 simulations. We also estimate the probability of detecting a pressure build-up above the injection well using cumulative distribution plot. Figure 6 shows the probability distribution of pressure build-up above the injection point. This result can be used to predict the risk associated with CO_2 leakage at the AZMI. If the system is required to be assessed based on a threshold AZMI pressure, the probability of failure can be calculated based on such results. It can be seen from Figure 5 that the areal extent of standard deviation increases as more CO_2 is injected into the system. Then the deviation starts to decrease considerably over time. The maximum deviation in pressure change is also above the injection well location. It is consistent with the fact that the greater

the amount of leakage, the greater will be the pressure buildup. Thus it becomes important to gather high-accuracy data of the geologic properties. Uncertainties of input parameters can lead to significant deviations in model outputs. Hence, the need for site-specific data is an essential requirement for efficient model predictions which validates what geoscientists know. Larger variation in input parameters of a model will lead to large deviations in outputs, which will lead to failure in understanding of the storage system and hence predict the containment risk properly. This work shows a parameter based uncertainty. Model based uncertainty analysis (e.g., Goodman et al., 2013[40]) can also be carried out by comparing predictions for similar simulation settings using other coupled reservoir-caprock-AZMI models to account for the uncertainty in model representativeness. The large uncertainties in the AZMI ROM prediction makes it important to analyze the role of each individual parameter on the output space. Therefore, we do a sensitivity analysis of the AZMI ROM, shown in the following section.

Sensitivity Analysis of Modeling Parameters

Assessment of the relative importance of the input parameters on the AZMI ROM output is required to understand the degree of their individual impact on the model predictions. This assessment is performed using a global sensitivity analysis with Sobol indices that are based on the aPC technique as described in Ashraf et al., 2013[26] and Oladyshkin et al., 2012[27], as described in the Methods section. As discussed in these previous works, the global aPC-based sensitivity analysis obtains global sensitivity information at low computational costs.

Quantitative sensitivity information for the AZMI ROM is extracted from the polynomial response surface. The Sobol indices (equation (10)) and the total Sobol indices (equation (11)) calculations are being done for the AZMI modeling scenario. The results are based on the 3rd order aPC expansion being assessed by fifty-six detailed simulations carried out for the uncertainty analysis. Model sensitivity analysis is performed for the five previously described input parameters (i.e., AZMI permeability, AZMI porosity, AZMI thickness, caprock permeability and caprock thickness) that have been used to quantify the model uncertainty. The test evaluated the impact of these parameters on the model output--pressure buildup response. The total Sobol sensitivities of input parameters on the ROM output are summarized in Figure 7. The figure presents the sensitivity results above the

injection point and a point approximately 4 km down south from the injection point at the end of injection period (= 50 years) and at the end of simulation (= 200 years). Table 2 represents the ranked 2nd order Sobol indices for the five uncertain parameters at the end of injection (= 50 years). The Sobol sensitivity calculation is shown for the 2nd order expansion and not the 3rd. The results are shown just to represent how parameter-parameter interaction can play a role in sensitivity calculation.

Figure 8 shows the time profile of the total Sobol indices (S_{Ti}) measured using equation (11), quantifying the contribution of a modeling parameter on the uncertainty of the pressure: (a) above the injection well, i.e. coordinate (34,46) and, (b) 4 km away from injection well southwards, i.e., coordinate (30,44). The sensitivity is normalized by variance at each time step. It should be noted that the sum of S_{Ti} of each parameter need not be equal to one, suggesting the presence of parameter-parameter interaction effects[41]. The AZMI permeability (k_{AZMI}) is clearly the most influential parameter, with higher total Sobol index corresponding to higher pressure buildup. If the permeability of the AZMI is high, pressure should easily dissipate resulting in lower pressure buildup in the AZMI. If the formation has a higher porosity, it means it can store more CO_2 per unit volume of the porous medium. This allows the incoming CO_2 to later accumulate the porous space and causing pressure change in the area. When the flow physics changes from injection to a gravity-dominated system, we observe a distinct change in the sensitivity patterns. During the injection period, H_{AZMI}, $k_{caprock}$ and $H_{caprock}$ are more dominant than k_{AZMI} and Φ_{AZMI}. The reason is that the incoming CO_2 takes time to mobilize and accumulate in the AZMI. Initially the model is largely dominated by the incoming flux through the seal which is dependent on $k_{caprock}$ and $H_{caprock}$ and the pressure buildup is also positively affected by the thickness of the AZMI, H_{AZMI}. Higher permeability of caprock leads to higher CO_2 and brine mobility, which leads to higher pressure buildup in the AZMI from incoming CO_2 and brine. The sensitivity of the pressure output with respect to higher AZMI permeability jumps up, right after stopping the injection.

METHODS

Arbitrary Polynomial Chaos Expansion

Assuming a physical model,

$$\Omega = f(\omega) \tag{1}$$

where,

$\omega = \{\omega_1, \dots \dots, \omega_M\} \in \mathbb{R}^M, M \geq 1$ is a vector of uncertain parameters (model inputs), and,

$\Omega = \{\Omega_1, \dots \dots, \Omega_N\}^T \in \mathbb{R}^N, N \geq 1$ is a vector of model outputs of interest.

The model output is a random variable if the parameter vector ω is uncertain. In our study the model output is a function of saturation and pressure. Polynomial chaos theory has a long history and according to Wiener, 1938[22], Ω can be expressed in the following form:

$$\Omega(\omega) = \sum_{i=1}^{M} c_i \psi_i(\omega) \tag{2}$$

where, c_i's are coefficients quantifying the dependence of model output on its input and $\psi_i(\omega)$ are orthogonal polynomial forming basis in the input probability space.

Since the AZMI ROM is space-time dependent, the model output is written as $\Omega\,(X, \omega)$ where the vector $X = \{x, y, t\}$ consists of two space coordinates and time. Hence, coefficients, c_i is determined for each point in space and time, i.e., $c_i(X)$.

In practice, this PCE is truncated at a finite number of basis functions, ψ_i. The number of the terms *M* in equation (2) depends on the total number of input parameters *N* and the order *d* of the expansion, i.e., the highest degree of polynomial basis functions, according to the following:

$$M = \frac{(N + d)!}{(N!\, d!)} \tag{3}$$

In the current study, we choose 3rd order aPC expansion. We use a 3rd order of expansion since it has a freedom to describe non-monotonic behaviors in comparison to the 2nd order. Also, the choice of 3rd problem is supported by the work of Oladyshkin and co-workers[42] where the authors have shown convergence analysis of aPC-based Sobol analysis concluding that all expansion beyond 2nd can capture non-linearity of a model. The detailed model description is provided in the Supplementary section.

Uncertainty Quantification

Uncertainty analysis using PCE can be typically characterized using two methods: intrusive and non-intrusive. In the present context, non-intrusive probabilistic collocation method (PCM)[36,43] is used, since it evaluates the coefficients in model expansion using a small number of model simulations and requires no manipulation using partial differential equation[7]. The method requires computing model Ω with *M* different sets of parameters ω that are called collocation points. In the current study, we use the recent version of PCM as described in Oladyshkin et al., 2011a[17] to compute the collocation points.

The model outputs $\Omega(\omega)$ are directly based on the model and the specified distribution of input parameters. The mean value (μ) and standard deviation (σ) of $\Omega(\omega)$ are given by the following analytical relations:

$$\mu(\Omega) = c_1 \,, \qquad \sigma(\Omega) = \sqrt{\sum_{j=2}^{N} c_j^2} \tag{4}$$

Likewise, all other moments of Ω can be obtained analytically, based on expansion coefficients and the moments of input parameters. The uncertainty outputs are space-time dependent, hence they are written as $\mu(x, y, t)$ and $\sigma(x, y, t)$.

Sobol Sensitivity Indices

A variance-based sensitivity analysis approach by calculation of Sobol Sensitivity Indices[25] was used. Studies on the combination of PCE techniques with Sobol indices have been performed in several previous studies[27,44,45]. The basic idea behind this approach is to replace the analyzed system with an approximating function that permits the calculation of numerical and mathematical benefits of a sensitivity analysis[26]. Since the calculation of output variances from statistics of input variables of polynomials is relatively fast, polynomials are used for the approximation. For the AZMI modeling scenario, the solution is approximated by orthogonal polynomials with ascending polynomial degree.
Let us assume we break the system output into components as follows:

$$\Omega = \Omega_0 + \sum_i \Omega_i + \sum_i \sum_{j>i} \Omega_{ij} + \cdots \tag{5}$$

where, indices i and j show dependency on two or more variables. If we consider the input vector ω to have n component ω_i for $i=1, \ldots\ldots, n$, then, $\Omega_i = f_i\,(\omega_i)$ and $\Omega_{ij} = f_{ij}\,(\omega_i, \omega_j)$. Saltelli et al., 2008[41] defined the higher order sensitivity index, or Sobol index[46], representing the significance of variation in output generated from the joint uncertainty in several input variables, i.e., from the interaction of uncertain parameters, as:

$$S_{ij} = \frac{\mathbb{V}\left[E\left(\Omega \middle| \omega_i, \omega_j\right)\right] - \mathbb{V}[E(\Omega|\omega_i)] - \mathbb{V}\left[E\left(\Omega \middle| \omega_j\right)\right]}{\mathbb{V}[\Omega]} \tag{6}$$

where, $\mathbb{V}\left[E\left(\Omega \middle| \omega_i, \omega_j\right)\right]$ is the variance of output expectations for a gi ven value of inputs ω_i and ω_j. If all the indices containing a given variable ω_i are added, we get total Sobol index[15]:

$$S_{Ti} = S_i + \sum_{j \neq i} S_{ij} + \sum_{j \neq i} \sum_{k \neq i} S_{ijk} + \cdots \tag{7}$$

The total Sobol index is a sensitivity measure to rank parameters according to their influence on model output. The higher the index, the greater is the effect of the corresponding input parameter on the model output. Sobol indices are calculated analytically[27] from the expansion coefficients of the aPC, shown in equation (2).

CONCLUSIONS

This work presents the application of reduced order models (ROMs) to predict the pressure response in the Above Zone Monitoring Interval (AZMI) and flux response above the caprock using the hypothetical Kimberlina CO_2 storage site (California, USA) as a base case example problem. We presented a data-driven arbitrary polynomial chaos expansion (aPC) method for uncertainty and sensitivity analysis of above zone pressure predictions. The data-driven approach provides a response surface based on a global orthonormal polynomial basis for arbitrary distributions. The method does not require extensive statistical knowledge for the data analysis. Thus, the aPC approach provides ability to model complex systems with unknown probability distribution functions, when only data sets of limited size or prior knowledge is available. The primary goal was to demonstrate the application and feasibility of aPC-based methods in the context of realistic CO_2 injection scenarios. We implemented this method with the base case Kimberlina storage scenario. Five uncertain parameters with assumed uncertainty distributions are used to compute the mean of above zone pressure buildup and the associate deviations in prediction related to the model uncertainties. The results show large uncertainties in the above zone pressure prediction, making it important to analyze the role of each individual parameter on the output space. Also, it emphasizes the need for site-specific data for efficient model predictions. The above zone pressure sensitivity to different geological parameters is then evaluated and quantified using Sobol indices. The results have shown that the most influential parameter for the pressure buildup responses is the permeability of the AZMI layer. The other parameters have almost equal influence on the predictions with different trends over time. Since, in general, the involved input uncertainties are hypothetical and just for demonstration purposes, the implications of this study are limited to the probabilistic assumptions made in this study and might not truly represent the actual CO_2 storage system.

REFERENCES

1. Bacon, D.H.; Dai, Z.; Zheng, L. Geochemical impacts of carbon dioxide, brine, trace metal and organic leakage into an unconfined, oxidizing limestone aquifer, *Energy Procedia*, **63**, 4684-4707, (2014).
2. Carroll, S. et al. Key Factors for Determining Groundwater Impacts Due to Leakage from Geologic Carbon Sequestration Reservoirs, *International Journal of Greenhouse Gas Control*, 29, 153–168, (2014).
3. Xiao, T.; McPherson, B.; Pan, F.; Esser, R.; Jia, W. Potential Chemical Impacts of CO_2 Leakage on Underground Source of Drinking Water (USDWs) Assessed by Quantitative Risk Analysis. *International Journal of Greenhouse Gas Control*, **50**, 305-316, (2016).
4. Pawar, R. et al. The National Risk Assessment Partnership's integrated assessment model for carbon storage: A tool to support decision making amidst uncertainty, *International Journal of Greenhouse Gas Control,* **52**, 175-189, (2016).
5. NETL (National Energy Technology Laboratory) National Risk Assessment Partnership (NRAP) Project Summary, (2011). http://www.netl.doe.gov/publications/factsheets/rd/R%26D179.pdf, 15/01/2014
6. NETL, "National Risk Assessment Partnership (NRAP) - FY14 Project Management Plan", Field Work Proposal Number: NRAP_FY14, National Energy Technology Laboratory, Office of Research and Development, March 31 (2014).
7. Zhang, Y.; Sahinidis, N.V. Uncertainty Quantification in CO_2 Sequestration Using Surrogate Models from Polynomial Chaos Expansion, *Industrial & Engineering Chemistry Research*, **52**, 9, (2013).
8. Dai, Z. et al. Probabilistic evaluation of shallow groundwater resources at a hypothetical carbon sequestration site, *Scientific Reports*, **4**: 4006, (2014).
9. Shahkarami, A.; Mohaghegh, S.D.; Gholami, V.; Haghighat, A.; Moreno, D. Modeling pressure and saturation distribution in a CO_2 storage project using a Surrogate Reservoir Model (SRM). *Greenhouse Gas Science and Technology*, **4**:1–27, (2014).
10. Huerta, N. J.; Vasylkivska, V. S. Well Leakage Analysis Tool (WLAT) User's Manual; NRAP-TRS-III-XXX-2015; NRAP Technical Report Series; U.S. Department of Energy, National Energy Technology Laboratory: Morgantown, WV, (2015).

11. Lindner, E. NSealR—A Brief User's Guide for Gen3; NRAP-TRS-III-0XX-2015; NRAP Technical Report Series; *U.S. Department of Energy*, National Energy Technology Laboratory: Morgantown, WV; p 118, (2015).
12. Bacon, D., N. Qafoku, Z. Dai, E. Keating, C. Brown, Modeling the Impact of Carbon Dioxide Leakage into an Unconfined, Oxidizing Carbonate Aquifer, *International Journal of Greenhouse Gas Control*, **44**, 290-299, (2016).
13. Keating, E. H., Harp, D. H., Dai, Z., Pawar R. J., Reduced Order Models for Assessing CO_2 Impacts in Shallow Unconfined Aquifers, *International Journal of Greenhouse Gas Control*, **46**, 187-196, (2016).
14. Namhata, A.; Zhang, L.; Dilmore, R.M.; Oladyshkin, S.; Nakles, D.V. Modeling pressure changes due to migration of fluids into the Above Zone Monitoring Interval of a Geologic Carbon Storage Site, *International Journal of Greenhouse Gas Control*, **In Press**, (2016), doi: 10.1016/j.ijggc.2016.11.012
15. Oladyshkin S., Class H., Helmig R. and Nowak W. A concept for data-driven uncertainty quantification and its application to carbon dioxide storage in geological formations. *Advances in Water Resources*, **34**, 1508–1518, (2011a).
16. Maltz, F. H.; Hitzl, D. L. Variance reduction in Monte Carlo computations using multi-dimensional hermite polynomial. *Journal of Computational Physics*; **2**, 345–76, (1979).
17. Robert, C. P.; Casella, G. Monte Carlo methods. New York: *Springer*, (2004).
18. Oladyshkin, S.; Class, H.; Nowak, W. 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. *Comput. Geosci.* **17**, 671–687, (2013).
19. Jia, W.; McPherson, B.; Pan, F.; Xiao, T.; Bromhal, G. Probabilistic Analysis of CO_2 Storage Mechanisms in a CO_2-EOR Field Using Polynomial Chaos Expansion. *International Journal of Greenhouse Gas Control*, **51**, 218-229, (2016).
20. Oladyshkin, S., Schröder, P., Class, H., Nowak, W. Chaos Expansion based Bootstrap Filter to Calibrate CO_2 Injection Models. *Energy Procedia*, **40**, 398-407, (2013).
21. Oladyshkin, S., Nowak, W. Date-driven uncertainty quantification using the arbitrary polynomial chaos expansion. *Reliability Engineering & System Safety*, **106**, 179–190, (2012).
22. Wiener N. The homogeneous chaos. *American Journal of Mathematics*, 60: 897–936, (1938).

23. Ghanem, R., Spanos, P. A stochastic galerkin expansion for nonlinear random vibration analysis. *Probabilistic Engineering Mechanics*; **8**, 255–64, (1993).
24. Le Maître, O., Knio, O. Spectral Methods for Uncertainty Quantification: With Applications to Computational Fluid Dynamics. *Springer*, New York, (2010).
25. Sobol, I. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates, *Mathematics and Computers in Simulation*, **55**, 271–280, (2001).
26. Ashraf, M.; Oladyshkin, S.; Nowak, W. Geological storage of CO_2: Global sensitivity analysis and risk assessment using arbitrary polynomial chaos expansion, *International Journal of Greenhouse Gas Control*, **19**, 704-719, (2013).
27. Oladyshkin S.; de Barros, F. P. J.; Nowak W. Global sensitivity analysis: a flexible and efficient framework with an example from stochastic hydrogeology. *Advances in Water Resources*, **37**, 10-22, (2012).
28. Birkholzer, J.T.; Zhou, Q.; Cortis, A.; Finsterle, S. A sensitivity study on regional pressure buildup from large-scale CO_2 storage projects. *Energy Procedia*, **4**, 4371–4378, (2011).
29. Zhou, Q., Birkholzer, J.T., Wagoner, J.L., Modeling the potential impact of geologic carbon sequestration in the southern San Joaquin basin, California. *The Ninth Annual Carbon Capture & Sequestration*, Pittsburgh, PA, (2011).
30. Wainwright, H.M., Finsterle, S., Zhou, Q., Birkholzer, J.T., Modeling the performance of large-scale CO_2 storage systems: a comparison of different sensitivity analysis methods. *International Journal of Greenhouse Gas Control*, **17**, 189–205, (2013).
31. Zhang, Y.; Liu, Y.; Pau, G.; Oladyshkin, S.; Finsterle, S. Evaluation of multiple reduced-order models to enhance confidence in global sensitivity analysis. *International Journal of Greenhouse Gas Control,* **49**, 217-226, (2016).
32. Brooks, R. H.; Corey, A. T. Properties of Porous Media Affecting Fluid Flow. Journal of the Irrigation and Drainage Division, *Proceedings of the American Society of Civil Engineers*, **92** (IR 2), 61–88, (1966).
33. Dai, Z. et al. CO_2 Accounting and Risk Analysis for CO_2 Sequestration at Enhanced Oil Recovery Sites. *Environmental Science & Technology*, **50**, 7546-7554, (2016).
34. Pan, F. et al. Uncertainty Analysis of Carbon Sequestration in an Active CO_2-EOR Field. *International Journal of Greenhouse Gas Control*, **51**, 18-28, (2016).
35. Eldred, E. S.; Burkardt, J. Comparison of non-intrusive polynomial chaos and stochastic collocation methods for uncertainty quantification. *47th AIAA Aerospace*

Sciences Meeting including the New Horizons Forum and Aerospace Exposition, Orlando, FL, January 5−8, AIAA paper 2009-0976, (2009).

36. Li, H.; Zhang, D. Probabilistic collocation method for flow in porous media: Comparisons with other stochastic methods. *Water Resources Research*, **43**, 44−56, (2007).
37. Xiu, D.; Karniadakis, G. E., Modeling uncertainty in flow simulations via generalized polynomial chaos, *Journal of Computational Physics*,**187**, 137−167, (2003).
38. Kopp, A., Class, H., Helmig, R. Investigations on CO_2 storage capacity in saline aquifers. Part 1: dimensional analysis of flow processes and reservoir characteristics. *International Journal of Greenhouse Gas Control*; **3**, 263–76, (2009).
39. NPC (National Petroleum Council). U.S. National Petroleum Council Public Database. http://www.netl.doe.gov, (1984).
40. Goodman, A. et al. Comparison of methods for geologic storage of carbon dioxide in saline formations. *International Journal of Greenhouse Gas Control,* 18, 329–342, (2013).
41. Saltelli, A.; Ratto, M.; Andres, T. Global Sensitivity Analysis: The Primer, *John Wiley & Sons*, (2008).
42. Oladyshkin, S.; Class, H.; Helmig, R.; Nowak, W. An integrative approach to robust design and probabilistic risk assessment for storage in geological formations. *Computational Geosciences*. **15**, 565−577, (2011b).
43. Oladyshkin S, Class H, Helmig R, Nowak W. Highly efficient tool for probabilistic risk assessment of ccs joint with injection design, Computational Methods in Water Resources (CMWR), *XVIII International Conference on Water Resources*, Barcelona, Spain, (2010).
44. Crestaux, T.; Le Maitre, O.; Martinez, J.-M. Polynomial chaos expansion for sensitivity analysis, *Reliability Engineering and System Safety*, **94(7)**, 1161–1172, (2009).
45. Sudret, B. Global sensitivity analysis using polynomial chaos expansions, *Reliability Engineering and System Safety*, **93(7),** 964–979, (2008).
46. Sobol, I.M. On sensitivity estimation for nonlinear mathematical models, Matem. Mod. **2** (1), 112–118, (1990).

ACKNOWLEDGMENTS

This work was completed as part of the National Risk Assessment Partnership (NRAP) project. Support for this project came from the U.S. Department of Energy's (DOE) Office of Fossil Energy's CCS and Crosscutting Research Programs, by the Department of Civil and Environmental Engineering and the Bertucci fellowship program at Carnegie Mellon University, and by training fellowship through the Oak Ridge Institute for Science & Education (ORISE). The authors would like to thank Grant Bromhal, Traci Rodosta, Robert Romanosky, M. Kylee Rice, and Steven Seachman of NETL and Mark Ackwiecz and Regis Conrad of U.S. DOE, Fossil Energy for their technical direction and Programmatic guidance; Mitchell Small and Athanasios Karamalidis of Carnegie Mellon University; Seth King and Ernest Lindner of AECOM, Inc. at NETL, and Ya-Mei Yang of NETL/ORISE at NETL for their valuable technical comments. The authors would also like to thank Dr. Haruko Wainwright, Dr. Erika Gasperikova and Dr. Tom Daley of the Lawrence Berkeley National Laboratory for providing valuable technical direction.

AUTHOR CONTRIBUTIONS STATEMENT

A.N. did the modeling setup, carried out the model simulations, performed data analysis and wrote the draft manuscript. S.O. and L.Z. helped in the modeling formulation, discussed the results, reviewed and revised the manuscripts. R.M.D. and D.V.N. helped in conceptualizing the model, discussed the results, and revised the manuscript.

ADDITIONAL INFORMATION

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Table 1. Reference parameters used for this study: horizontal permeability (k_h), anisotropy ratio (k_v/ k_h), porosity (Φ), pore compressibility (β_p), van Genuchten parameters (α, m), Brooks-Corey parameter (γ), bubbling pressure (P_b), residual brine saturation (S_{rb}) and residual CO_2 saturation (S_{rc}).

Parameter	Reservoir	Caprock	AZMI
k_h(mD)	depth dependent*	0.002	0.1
k_v/ k_h	0.2	0.5	0.5
Φ	depth dependent*	0.338	0.32
β_p (10^{-10} Pa^{-1})	4.9	14.5	14.5
α (10^{-5} Pa^{-1})	13	0.42	-
m	0.457	0.457	-
γ	-	-	2
P_b	-	0.01	0.02
S_{rb}	0.30	0.45	0.35
S_{rc}	0.25	0.40	0.30

**depth dependent values are taken from Wainwright et al. (2013)[30]*

Table 2: Second-order Sobol indices for five parameters: [1] AZMI permeability (k_{AZMI}), [2] AZMI porosity (Φ_{AZMI}), [3] thickness of AZMI (H_{AZMI}), [4] caprock permeability ($k_{caprock}$) and [5] caprock thickness ($H_{caprock}$) at coordinates above the injection well, i.e. coordinate (34,46) and, 4 km away from injection well southwards, i.e., coordinate (30,44) at the end of injection (= 50 years).

Sobol index	Value at (34,46)	Rank at (34,46)	Value at (30,44)	Rank at (30,44)
S_1	0.215	4	0.285	2
S_2	0.171	5	0.467	1
S_3	0.397	1	0.165	3
S_4	0.268	3	0.081	6
S_5	0.303	2	0.073	7
S_{1-2}	0.042	8	0.102	4
S_{1-3}	0.055	7	0.037	8
S_{1-4}	0.003	10	0.003	11
S_{1-5}	0.001	14	0.002	13
S_{2-3}	0.017	9	0.085	5
S_{2-4}	0.001	13	0.004	10
S_{2-5}	0.001	15	0.003	12
S_{3-4}	0.002	11	0.001	14
S_{3-5}	0.001	12	0.001	15
S_{4-5}	0.073	6	0.006	9

FIGURES

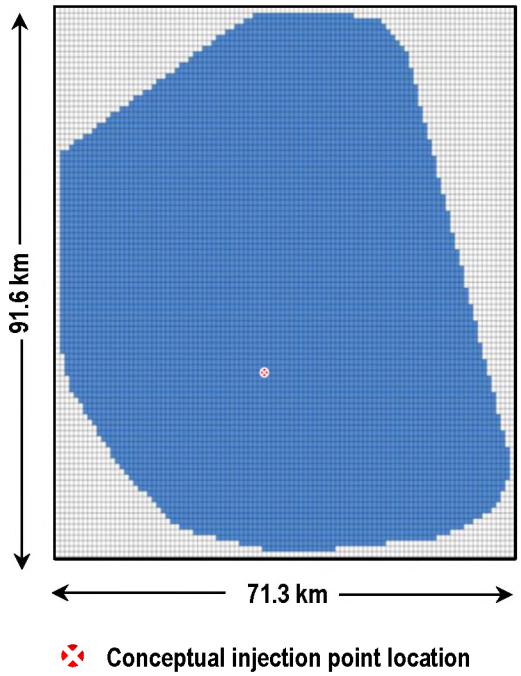

Figure 1: Plan view of the model domain (in blue) with numerical grid. The red point is the location of the conceptual injection well (coordinate (34,46)).

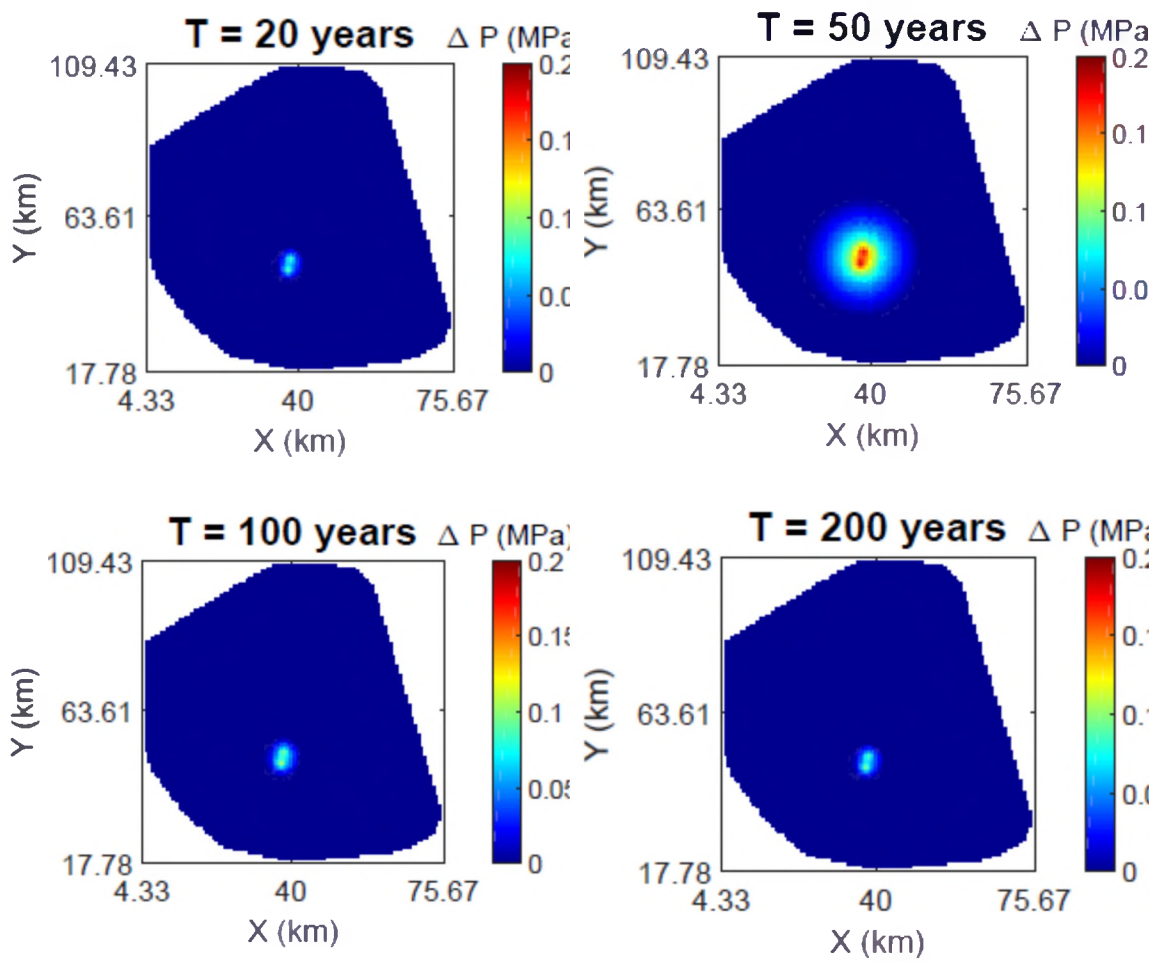

Figure 2: Changes in pressure response (in MPa) at the top of the AZMI at 20, 50, 100 and 200 years after the start of CO_2 injection (base case results).

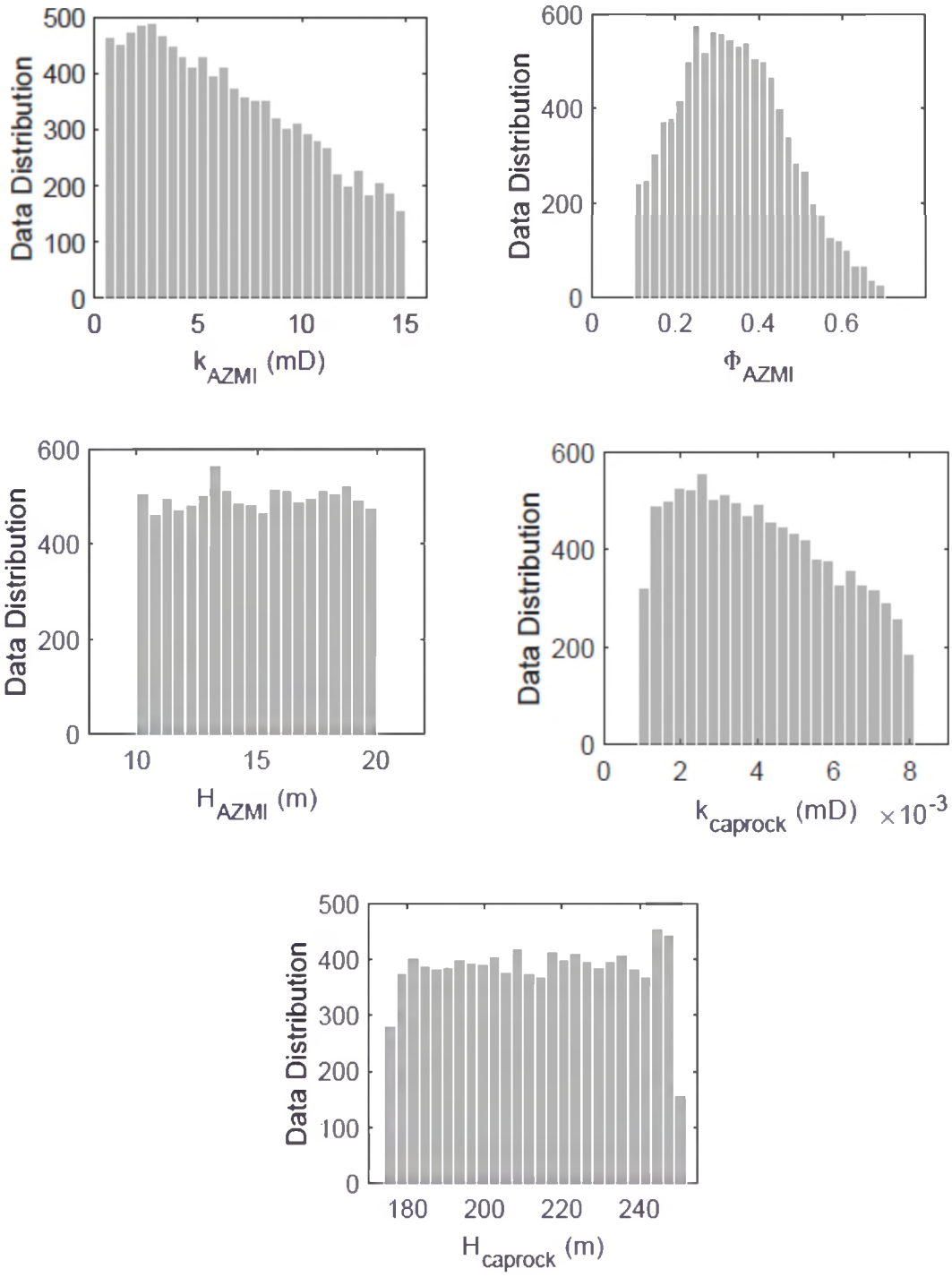

Figure 3: Distribution of AZMI permeability (k_{AZMI}), AZMI porosity (Φ_{AZMI}), thickness of AZMI (H_{AZMI}), caprock permeability ($k_{caprock}$) and caprock thickness ($H_{caprock}$) for aPC uncertainty analysis.

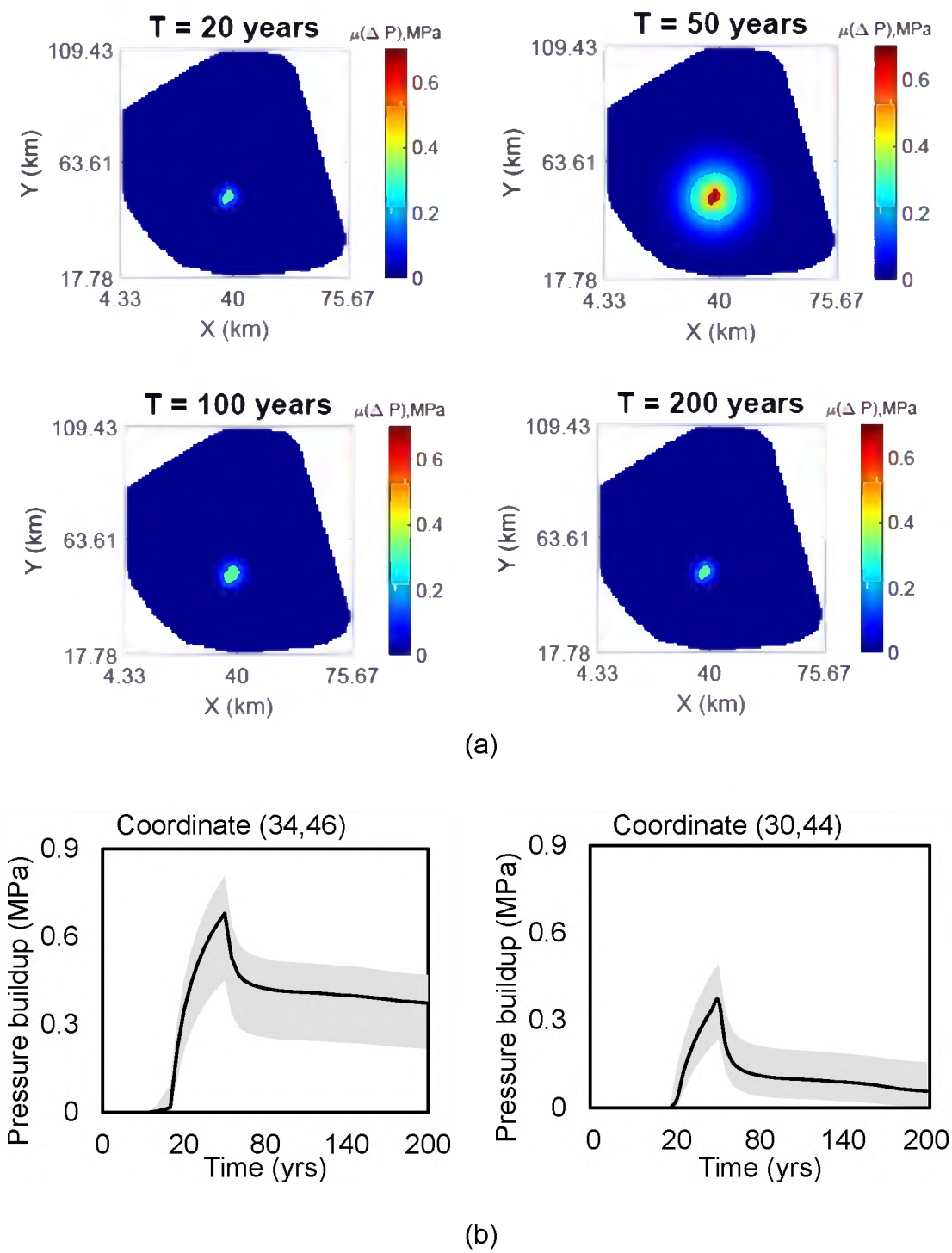

Figure 4: Plots showing (a) mean change in pressure response (in MPa) at the top of AZMI over time based on aPC analysis, (b) mean pressure change (black line) and range of pressure change from 56 simulations (grey shaded region) above injection point, i.e., coordinate (34,46) and 4 km away from injection well southwards, i.e., coordinate (30,44) over time.

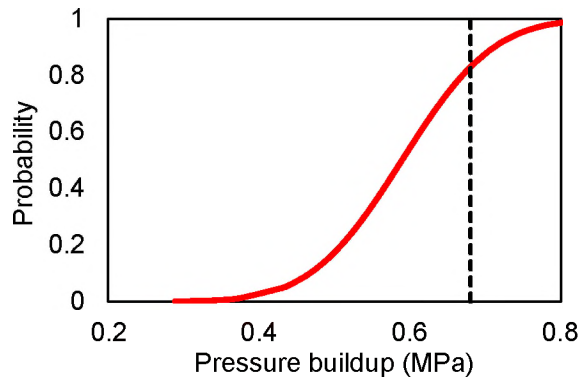

Figure 5: Probability of detection of pressure build-up (shown in red line) at the top of the AZMI above injection point, i.e., coordinate (34,46) at the end of injection. The black dotted line shows the mean of pressure build-up based on aPC analysis.

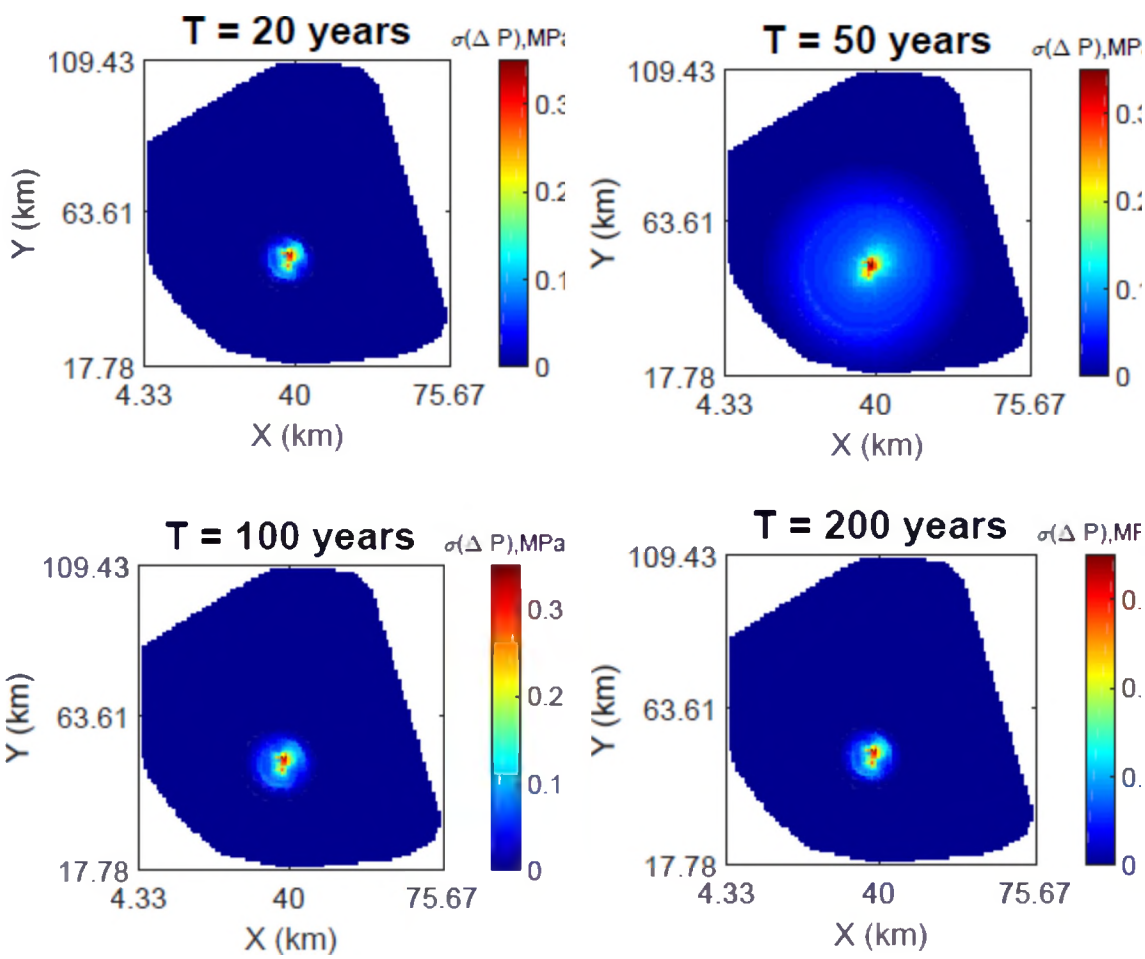

Figure 6: Estimation of standard deviation of the change in pressure response (in MPa) prediction by AZMI ROM at the top of AZMI.

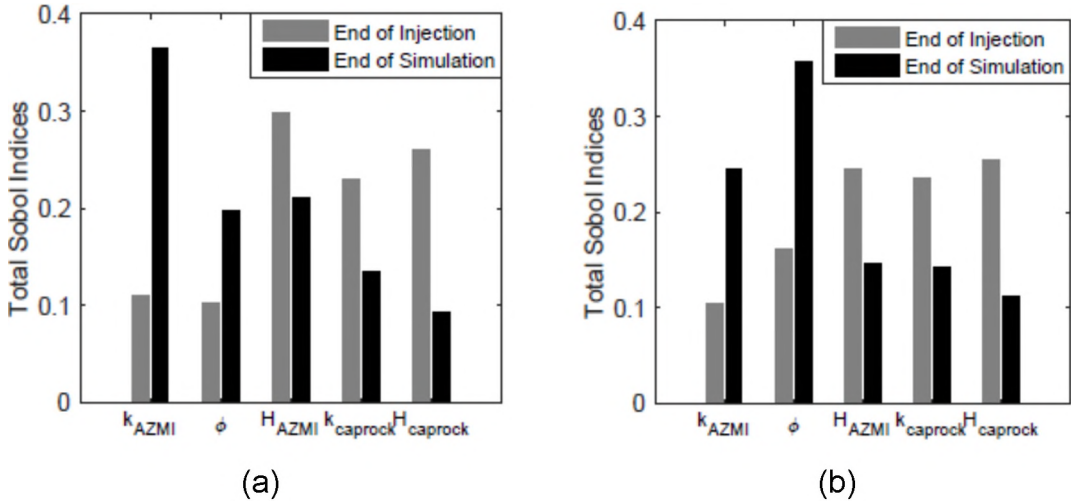

Figure 7: Sensitivity of AZMI ROM output for changes in pressure with respect to the uncertain parameters: AZMI permeability (k_{AZMI}), AZMI porosity (Φ_{AZMI}), thickness of AZMI (H_{AZMI}), caprock permeability ($k_{caprock}$) and caprock thickness ($H_{caprock}$) at: (a) above the injection well, i.e. coordinate (34,46) and, (b) 4 km away from injection well southwards, i.e., coordinate (30,44).

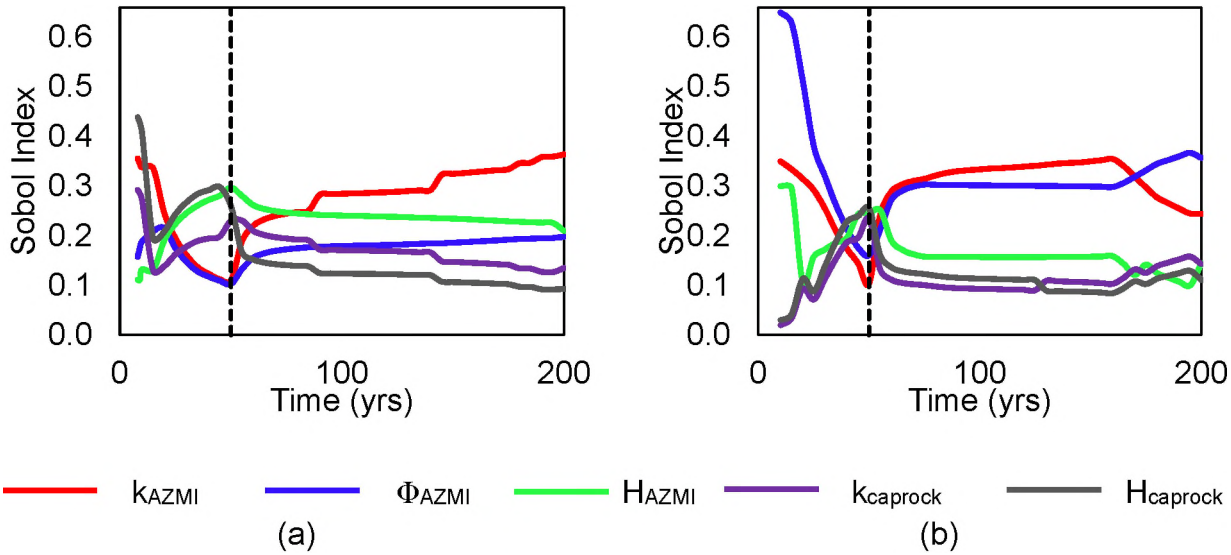

Figure 8: Sobol sensitivity results for AZMI ROM outputs over time at (a) above the injection well, i.e. coordinate (34,46) and, (b) 4 km away from injection well southwards, i.e., coordinate (30,44).