Advances in Water Resources 56 (2013) 49–60

Contents lists available at SciVerse ScienceDirect

Advances in Water Resources

journal homepage: www.elsevier.com/locate/advwatres

Assessing leakage detectability at geologic CO_2 sequestration sites using the probabilistic collocation method

Alexander Y. Sun [a,*], Mehdi Zeidouni [a], Jean-Philippe Nicot [a], Zhiming Lu [b], Dongxiao Zhang [c]

[a] Bureau of Economic Geology, Jackson School of Geosciences, The University of Texas at Austin, TX, USA
[b] EES-16, Los Alamos National Laboratory, Los Alamos, NM, USA
[c] College of Engineering, Peking University, Beijing, China

ARTICLE INFO

Article history:
Received 5 June 2012
Received in revised form 20 November 2012
Accepted 25 November 2012
Available online 5 December 2012

Keywords:
Carbon sequestration and storage
Leakage detection
Probabilistic collocation method
Detectability
Signal-to-noise ratio
Uncertainty quantification

ABSTRACT

We present an efficient methodology for assessing leakage detectability at geologic carbon sequestration sites under parameter uncertainty. Uncertainty quantification (UQ) and risk assessment are integral and, in many countries, mandatory components of geologic carbon sequestration projects. A primary goal of risk assessment is to evaluate leakage potential from anthropogenic and natural features, which constitute one of the greatest threats to the integrity of carbon sequestration repositories. The backbone of our detectability assessment framework is the probability collocation method (PCM), an efficient, nonintrusive, uncertainty-quantification technique that can enable large-scale stochastic simulations that are based on results from only a small number of forward-model runs. The metric for detectability is expressed through an extended signal-to-noise ratio (SNR), which incorporates epistemic uncertainty associated with both reservoir and aquifer parameters. The spatially heterogeneous aquifer hydraulic conductivity is parameterized using Karhunen–Loève (KL) expansion. Our methodology is demonstrated numerically for generating probability maps of pressure anomalies and for calculating SNRs. Results indicate that the likelihood of detecting anomalies depends on the level of uncertainty and location of monitoring wells. A monitoring well located close to leaky locations may not always yield the strongest signal of leakage when the level of uncertainty is high. Therefore, our results highlight the need for closed-loop site characterization, monitoring network design, and leakage source detection.

1. Introduction

Carbon capture and storage is an active R&D area being studied across the world for reducing greenhouse gas emission. A primary goal of geologic carbon sequestration is to ensure that injected CO_2 can be safely contained in host formations for extensive performance periods. Unintended CO_2 migration from storage formations can occur as a result of natural and anthropogenic impacts, such as overpressure in cap rocks and leakage through geologic faults and abandoned wells. Therefore, risk analysis and management play critical roles in all stages of CO_2 sequestration projects for mitigation of potential health and environmental risks. Many countries have or are in the process of mandating risk management for site license applications. For example, the European Union explicitly requires that a proposed CO_2 sequestration site shall "show no significant risk of leakage and no significant environmental or health risks under the proposed conditions of use" [1]. Under its Underground Injection Control program, the US Environmental Protection Agency (EPA) has proposed specific rules on CO_2 injection wells to protect underground sources of drinking water (USDW). In particular, the proposed EPA rules have extensive requirements to ensure that wells used for geologic sequestration are appropriately sited, constructed, tested, monitored, funded, and closed [2]. Recently, the US Department of Energy releases best practice manuals on risk analysis and management activities related to CO_2 storage projects [3,4].

Another important and closely related goal of risk management is to optimize monitoring networks for timely detection of CO_2 leakage under both model and data uncertainty. The first and foremost concern of regulators is how efficiently a proposed monitoring network can detect leakage signals when they first appear. *Detectability* refers to the capability of an observer or a piece of equipment to differentiate between *noise* and *signal plus noise* during an arbitrary observation interval or sampling event [5]. Assessment of detection probability invariably requires quantitative information on signal-to-noise ratios (SNRs). In a series of papers [6–8], Taguchi used SNR as a metric for optimal product design, which he referred to as "the operation of choosing settings for the design parameters of a product or manufacturing process to reduce sensitivity to noise". A now widely used Taguchi SNR for product design is the logarithm of the ratio between mean and

* Corresponding author. Tel.: +1 512 475 6190; fax: +1 512 471 0140.
E-mail address: alex.sun@beg.utexas.edu (A.Y. Sun).

0309-1708/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.advwatres.2012.11.017

standard deviation of observed signals; alternatively, the ratio can be used directly without taking logarithm. The optimization principles behind Taguchi's design are to (1) minimize product sensitivity to variations transmitted from components and (2) minimize product sensitivity to environmental fluctuations. We adopt the same principles for the design of CO_2 monitoring networks under parametric uncertainty.

In the context of the current work, noise is broadly interpreted as uncertainty caused by (1) natural variability in a system (aleatory uncertainty) and/or (2) lack of complete knowledge about system characteristics (epistemic uncertainty). Common examples of the latter are (1) hydrogeologic (or reservoir) properties of monitoring intervals, including both the injection zone itself and aquifers above the injection zone (i.e., the above-zone monitoring interval (AZMI)), and (2) the type of leakage source and its properties [9]. An inverse relationship generally exists between cost and sensitivity of leakage-detection techniques, which inherently depend on four macroscales pertaining to (1) the region needs to be monitored, (2) the region affected by leakage flux, (3) the main leakage zone (e.g., scale of sources), (4) the footprint of the monitoring equipment used [10,11], as well as on pore- and other microscales. Physical and chemical processes associated with these scales can all potentially affect SNR. Macroscales 1 and 4 are usually either known or can be practically considered deterministic. Macroscales 2 and 3, however, are rather dynamic and uncertain. The lack of information on system properties and length scales identified here, especially during the planning stage of many CO_2 sequestration projects, often leads to a situation in which uncertainty quantification (UQ) becomes the dominant question, overriding the influence of secondary physical processes [12].

The role of UQ in risk analysis and management is thus twofold. First, it helps in identifying the dominant system and environmental variables that contribute to system response variability, an analysis that is also important for subsequent activities such as data collection and monitoring network design. Second, UQ yields bounding scenarios for system outputs and, therefore, provides direct inputs to risk-informed performance assessment. UQ can become more powerful when coupled with data assimilation techniques and applied to site management adaptively. However, to be suitable for data fusion or real-time decision support, a UQ technique must be highly efficient. The purpose of this work is to investigate the use of the probabilistic collocation method (PCM), an efficient stochastic-response surface method for assessing detectability under parameter uncertainty.

Models have been used extensively to simulate the migration of injected CO_2 through various leakage pathways [13–17]. However, a number of challenges potentially exist when these models are applied to UQ using conventional techniques. For example, perturbation methods and stochastic-moment approaches require deriving and solving a set of coupled stochastic partial differential equations (PDEs) corresponding to various uncertain input variables, which is a nontrivial task for complex, nonlinear processes. In addition, the assumptions underlying these approaches largely restrict their applicability to small parameter variability [18]. Whereas Monte Carlo methods do not impose strong assumptions on the variability of uncertain variables and do not require modifications of existing codes (i.e., nonintrusive), they are computationally demanding and become intractable for large-scale problems without access to parallel or distributed computing facilities.

In recent years, a new breed of UQ techniques—the polynomial chaos expansion (PCE) method and stochastic collocation (SC) method—have received broad attention in engineering-reliability analyses [19–24]. Both UQ techniques belong to the so-called stochastic-response surface methods, and both represent parametric uncertainties as an expansion of orthogonal polynomials of independent random variables and propagate them to quantify model-output uncertainty. Exponential convergence rates can be achieved by both methods for a wide range of probabilistic analysis problems [21].

The classical PCE method, pioneered by Ghanem and Spanos [24], is based on the homogeneous chaos theory of Wiener [25]. The method starts with a spectral expansion of input uncertain variables, through which the variables are projected onto a stochastic space spanned by a set of complete orthogonal polynomials. A main consequence of this spectral expansion is that the uncertain variables are represented as a deterministic part (i.e., coefficients of expansion) and a stochastic part (i.e., polynomial chaos basis). Galerkin projection is then applied to each polynomial chaos basis, thereby replacing stochastic PDEs with a coupled deterministic system of equations, from which the coefficients of expansion can be solved for. Ghanem and Spanos [24] worked with Hermite polynomial chaos, which is optimal for Gaussian random variables. Xiu and Karniadakis [21] later introduced generalized polynomial chaos expansion using the Wiener–Askey scheme, such that a number of commonly used continuous and discontinuous probability distribution functions (PDFs) could be accommodated. The classical PCE method requires developing and solving a coupled system of deterministic, ordinary differential equations, a procedure that can be cumbersome and nontrivial when the problem at hand is complex and nonlinear.

The SC method, like its deterministic counterpart in the finite element method, seeks to construct a response surface using a prescribed set of collocation points, but in stochastic space rather than physical space. In the SC method, the expansion coefficients are nothing but model outputs corresponding to each of the collocation points. The quantities to be solved for are expansion polynomials, which themselves are based on Lagrange interpolation polynomials. The SC method is closely related to PCE because collocation points that offer high accuracy are also zeros of orthogonal polynomial bases used in PCE [20]. By design, the SC method is nonintrusive and leads naturally to uncoupled deterministic systems, as opposed to coupled system of equations resulting from Galerkin projection. The efficiency and accuracy of SC depend largely on the number and location of collocation points. Several methods exist for generating collocation points, such as the tensor product grid and Smolyak sparse grid. The tensor product method is suitable only for low-dimensional systems because of the large number of collocation points it generates. The sparse grid approach uses a subset of the points generated by the tensor product method and therefore can be significantly more efficient [19,26,27].

A variant of the SC method is the probabilistic collocation method (PCM), which combines features of both the classical PCE and SC methods. Like PCE, it prescribes a set of orthogonal polynomial bases. The expansion coefficients, however, are obtained by solving a linear system of equations, in which the right-hand-side data vector (or matrix for transient problem) consists of model responses at collocation points. The PCM, originally introduced by Tatang et al. [28], is attractive for engineering-reliability analyses because it is nonintrusive and its implementation is relatively straightforward. The PCM has been used to solve both single- and multiphase flow and mass transport problems in porous media [29–33]. Recently, Oladyshkin et al. [12] used PCM in risk assessment of a hypothetical CO_2 sequestration site, in which the authors considered the impact of uncertain parameters (reservoir porosity and permeability and leaky-well permeability) and design parameters (injection rate and size of screening interval) on predicted injection-zone variables (i.e., cap-rock pressure and CO_2 leakage rate). Oladyshkin et al. [12] assumed that all uncertain variables are independent and uniform in physical space; the uncertain variables were represented using Hermite polynomial chaos, which is not optimal for non-Gaussian random variables, as we have pointed out before. Walter et al. [34] later applied a similar approach to study