Lawrence Berkeley National Laboratory

Recent Work

Title

A metric for evaluating conformance robustness during geologic CO2 sequestration operations

Permalink

<https://escholarship.org/uc/item/7373h2dj>

Authors

Harp, DR
Oldenburg, CM
Pawar, R

Publication Date

2019-06-01

DOI

10.1016/j.ijggc.2019.03.023

Peer reviewed

A Metric for Evaluating Conformance Robustness During Geologic CO₂ Sequestration Operations

Dylan R. Harp^{a,*}, Curtis M. Oldenburg^b, Rajesh Pawar^a

^a*Earth and Environmental Sciences Division, Los Alamos National Laboratory, Los Alamos, NM, 87544*

^b*Lawrence Berkeley National Laboratory, Berkeley, CA, 94720*

Abstract

A metric for quantifying the robustness of a designation of conformance of a geologic CO₂ sequestration (GCS) project during its operational phase is developed and demonstrated. Conformance in this context is a measure of the degree to which the sequestration system is understood and can be accurately modeled along with the degree to which the storage system is performing as designed. The robustness of conformance quantifies the degree to which parameter values can deviate from their current nominal estimates and still produce model forecasts that meet the performance criteria for the GCS operation. We develop and demonstrate the approach on a simplified scenario to illustrate the concept using a single uncertain parameter (homogeneous reservoir permeability) and a single performance criterion (critical pressure at a monitoring well in the reservoir; i.e., one that may displace brine from the reservoir to an overlying drinking water aquifer for example). Increased confidence in conformance assessment as more monitoring data are obtained is incorporated through the standard error of the coefficient (reservoir permeability in the case presented here), which we designate as the concordance metric. As more monitoring data become available during the course of the GCS operation, the standard error of the coefficient decreases (in general), thereby leading to increased conformance robustness as a larger deviation from nominal is required to fail to meet performance criteria. Increasing conformance robustness over time builds confidence that a GCS project will continue to meet performance criteria during the life-span of the project, thereby supporting designations of conformance. A lack of conformance robustness provides a critical warning that the performance criteria of the GCS operation are not robust against probabilistic and non-probabilistic uncertainty in model conceptualization and/or model parameters.

Keywords: geologic carbon sequestration, CO₂ storage, conformance, concordance, performance

*Corresponding author: Dylan R. Harp (dharp@lanl.gov)

1. Introduction

A successful geologic CO₂ sequestration (GCS) operation will result in long-term storage of CO₂ with acceptably low or no measurable impacts to health, safety, the environment, or existing resources. Beyond the impacts of failed storage, factors that can lead to an unsuccessful GCS operation include lower (or possibly higher) than anticipated reservoir permeability and/or porosity or unanticipated reservoir compartmentalization leading to an inability to inject a desired quantity of CO₂, and excessive pressure build-up leading to unwanted fracturing or induced seismicity (Sminchak et al., 2001; Lucier et al., 2006; Cappa and Rutqvist, 2011; Dempsey et al., 2014). These GCS operation failures can result in releasing CO₂ back to the atmosphere (negating the purpose of GCS) (Lewicki et al., 2007), contaminating resources such as overlying hydrocarbon reserves and shallow aquifers (Keating et al., 2010; Little and Jackson, 2010; Trautz et al., 2012; Navarre-Sitchler et al., 2013; Bacon et al., 2016; Keating et al., 2016), and damage to public perception of GCS (Palmgren et al., 2004; Curry et al., 2005; Miller et al., 2007; Wilson et al., 2008; Elliot et al., 2012).

One of the keys to ensuring a successful GCS operation (or the mitigation of a failed GCS operation) is monitoring. Refer to Harbert et al. (2016) for an extensive review of GCS monitoring strategies and approaches. The value of monitoring data in assessing the performance of GCS operations cannot be overstated. However, direct measurements of pressures, CO₂ saturations, pH, etc. are necessarily sparse due to the large areas of concern associated with GCS operations and the expense associated with developing monitoring wells. Although indirect measurements, such as seismic monitoring data and other geophysical survey data, can provide more continuous spatial coverage, they are highly uncertain and affected by many factors requiring expert interpretation. As a result, the data in and of themselves are not sufficient to establish confidence in GCS operational performance. Models are used to fill the gaps in understanding of system performance that would otherwise be based on observational data alone.

Fully utilizing the information produced with a monitoring program requires a formal approach to integrate monitoring data into models in a manner that provides indications of the risks associated with the GCS operation. While there are many existing GCS related publications describing and demonstrating alternative monitoring techniques (Cheng et al., 2010; Dafflon et al., 2012; Doetsch et al., 2013; Lin and Huang, 2015), uncertainty quantification approaches (Sun et al., 2013; Trainor-Guitton et al., 2013; Harp et al., 2017), and data assimilation approaches (Kumar, 2010; Chen et al., 2018), there are few examples describing how to incorporate monitoring data as it becomes available during a GCS operation to support designations of conformance. Bielicki et al. (2016) present an approach to quantify the monetized leakage risk of GCS operations, including the impact of remediating leaks, but the approach is intended to evaluate

potential sites, not assimilate monitoring data during a GCS operation. One of the few existing examples is Chadwick and Noy (2015) who demonstrate that as more seismic data is collected and used to refine models, conformance can be demonstrated as the models are able to predict new data more accurately. In this paper, we discuss how data and modeling can be coupled to support robust designations of conformance in GCS operations. The approach uses a similar decision analysis as in Harp et al. (2017), which proposed an approach to select a GCS pressure management strategy, but in this case applies and extends the decision analysis to designations of conformance over the course of the GCS operation.

2. Background and Definitions

The term *conformance* has many different meanings in different contexts (Oldenburg, 2018). One definition pertains to a regulatory designation indicating that a GCS operation is performing and will continue to perform within acceptable levels of risk and within the bounds of its permit and related legal requirements. As such, a GCS operation can only be deemed to be in conformance by a regulator. And yet we have observed that the term conformance is used more broadly in the GCS community as a metric of understanding of the system as demonstrated by agreement between models and monitoring data, along with performance of the system (Chadwick and Noy, 2015; Oldenburg, 2018). The first component of conformance is quantified by the similar term *concordance* which is a measure of the coherence of modeling results with observed data (e.g., agreement between modeled and measured pressures). The second part of conformance is quantified by *performance*, which indicates whether or not the GCS operation is meeting design expectations (i.e., meeting performance criteria). Performance criteria can include ensuring that a desired quantity of CO₂ is injected, plume stability or migration is as expected (established through measurements of pressure, CO₂ saturation or geophysics), levels of induced seismicity related to CO₂ injection are acceptable, or that CO₂ or brine leakage are not occurring (Harp et al., 2017).

To illustrate these terms, in Figure 1 we present a graphic modified slightly from Oldenburg (2018) showing how observational data and history matching are used to quantify concordance, and how concordance plus performance leads to conformance as we use the term here. As shown, observations and history matching are inherently based on the past, while forecasts can be made using models including forecasts of performance. As shown, concordance plus a forecast of continued performance combine to make a forecast of conformance.

While concordance metrics are important model diagnostic tools, they can only indirectly support designations of conformance by building confidence that we understand the past behavior of the system. On the other hand, performance metrics are crucial in directly supporting determinations of conformance. While

Figure 1: Graphical illustration of key terms and their meanings relative to conformance as used in this paper, along with temporal context.

past performance is important, projections (i.e., forecasts) of future performance are critical for building confidence in GCS operations and thereby allowing a forecast of conformance (Oldenburg, 2018). However, uncertainty in projections of future performance must be considered in a conformance analysis. In this paper, we demonstrate how goodness-of-fit metrics (e.g., the standard error of the coefficient) derived from a concordance evaluation can be used to quantify the uncertainty in future projections of performance. We present a conformance robustness metric that couples concordance and future performance as illustrated in Figure 1. The conformance robustness quantifies the ability to be incorrect in our current understanding of the system and still meet performance criteria in the future. In other words, we know that our model will be incorrect or deficient to some degree. As a result, we want to know how incorrect can our model be (how far can our nominal estimates of the system deviate from the actual system) and the GCS operations can still be considered to be performing as good (or better) than designed. Nominal estimates of the system (properties, boundary or initial conditions, etc.) can be based on expert opinion, deduced from measurements, or a combination of the two. They are the current best guess and the values that would be used to make a decision today given the currently available data.

Our conformance robustness analysis approach utilizes concepts from information gap decision theory (IGDT) (Ben-Haim, 2006). IGDT provides a formal approach to quantify the robustness of alternative decisions, where robustness captures the ability for nominal estimates to deviate from actual values and yet still meet performance criteria. IGDT captures non-probabilistic uncertainty due to gaps in information and/or understanding through set-based uncertainty models in which the level that our model can be incorrect (e.g., poor conceptual model or inaccurate parameter estimates) has no limit. Our conformance robustness metric incorporates probabilistic uncertainty into the IGDT analysis in the form of the standard error of the coefficient, thereby allowing the analysis to consider both probabilistic and non-probabilistic uncertainty. In other words, uncertainties beyond probabilistic uncertainty (i.e., surprises, unanticipated events, incorrect conceptual model, etc.) can be considered in the conformance robustness metric. We feel that accounting for occurrences beyond probabilistic uncertainty when evaluating conformance is important because failures will likely be caused by unanticipated incidents, not by well-characterized and understood uncertainties. This is not to imply that our conformance robustness metric will predict unforeseen or unanticipated events, but that the metric will consider incidents beyond well-characterized probabilistic uncertainty captured by the standard error of the coefficient.

The need to consider inaccurate conceptual models and events beyond anticipation is motivated by considering two CO₂ sequestration projects, Snøhvit and In Salah. At Snøhvit upon initial injection, there

was higher than anticipated pressurization of the target injection interval (Grude et al., 2014). At In Salah following several years of injection, there was faster-than-expected CO₂ flow from an injector to a well with a damaged wellhead (Oldenburg et al., 2011; White et al., 2014). Both of these cases are examples where the conceptual models upon which the injection design was based were flawed, leading to poor performance and an unanticipated incident during CO₂ injection. In the case of Snøhvit, a backup plan to inject into an alternative interval was implemented which remedied the situation.

At Snøhvit, a conformance robustness analysis would have been able to consider the implications on performance criteria of lower-than-anticipated injectivity into the Tubåen formation and motivate more data collection to reduce uncertainty in injectivity and potentially avoid injection into the Tubåen at this particular location. At In Salah, a conformance robustness analysis would have been able to consider the effects of higher than expected permeability between the injection well and the damaged well. If such a scenario had been anticipated, perhaps the damaged well would have been used as an official observation well to monitor for the possibility of the high-permeability pathway, which could have possibly avoided the problem of the wellhead hardware being unlawfully removed and thereby damaging the wellhead.

The conformance robustness analysis framework presented here allows such inaccurate conceptual models and unanticipated events to be considered. When a formal consideration of the possibility of unanticipated events to occur and an evaluation of how robust a design is to accommodate their occurrence while still allowing the system to perform acceptably, it is expected that decisions about operations can be made to adjust to the new information. In many cases, the need for decision support extends far beyond optimizing operations to ensuring that operations continue when unanticipated events occur.

Implied in a designation of conformance is the condition that the details of the GCS system are *understood* (e.g., pressures in the reservoir, extent of CO₂ plume, etc.). While it is not possible to quantitatively establish that future predictions of the system will be accurate, concordance, which is inherently based on observations and therefore based on hindsight, builds confidence in our understanding of the past behavior of the system. The degree to which we understand the system is idealized here in our simple example case study as the current standard error of the coefficient (reservoir permeability parameter estimate in this case), but in practice would likely be more complicated involving conceptual model variations, etc., as opposed to simple parameter updating.

While how well the past behavior of the system is understood is a primary consideration in conformance, it is not sufficient for establishing conformance robustness. Establishing robustness requires that we consider the implications that our understanding of the system is flawed, incorrect, or in some way deficient. Of course,

uncertainty can work both to the detriment and to the advantage of GCS system performance. Consider the case where our understanding of the system is flawed in a way that leads to meeting performance criteria at levels greater than expected (more CO₂ can be injected, pressures are less than anticipated, etc.). In this case, our understanding of the system is flawed, but the GCS operation is still performing well and meets criteria for a forecast of conformance. However, the pernicious side of uncertainty (e.g., unexpected induced seismicity occurs, the CO₂ plume encounters an unknown legacy well and contaminates a shallow aquifer, etc.) leads towards failing to meet GCS operational performance criteria. The conformance robustness metric described below focuses on the pernicious side of uncertainty which goes beyond our confidence in how well we understand the past behavior of the system. In other words, the conformance robustness analysis is performed on top of the conformance analysis, and can be considered a type of sensitivity analysis of the conformance.

3. Approach

To describe our approach, we start with a description of the synthetic GCS scenario that we use to demonstrate the approach. We then define the components of the conformance robustness analysis, including the (1) system model, (2) uncertainty model, (3) concordance metric, and (4) performance criterion. Finally, we present our conformance robustness metric derived from these components.

3.1. Synthetic GCS scenario

In order to focus attention on the details of the conformance robustness metric, we use a simplified CO₂ sequestration scenario sketched in Figure 2. The scenario involves injecting CO₂ into a homogeneous reservoir at 950 m depth overlain by an impervious caprock layer. The geothermal gradient is assumed to be 0.03°C/m, a typical value for North America, and the average ground surface temperature is assumed to be 20°C, resulting in 48.5°C at the top of the reservoir. The average ground surface pressure is assumed to be 0.1 MPa and the hydraulic gradient is assumed to be a typical value of approximately 9.78 kPa/m, resulting in a pressure at the top of the reservoir of 9.4 MPa. Based on these conditions, the density of the supercritical CO₂ in the reservoir would be approximately 328 kg/m³ (Linstrom, 2003). Scenarios where CO₂ is injected at a constant rate of 0.75, 1.0, and 1.25 Mt/y are considered. A well is monitoring pressure in the radially symmetric storage reservoir at approximately 141 m from the injection well. While at any time there is a best-fit approximation of the homogeneous reservoir permeability, its value is uncertain and the primary (only) source of uncertainty in our scenario. While the model is extremely simple in this demonstration so

Figure 2: Schematic of synthetic CO2 sequestration scenario (not to scale).

that we can focus on the details of the conformance robustness approach, the approach can easily be applied to more complex and realistic models.

3.2. System model

We simulate pressures at the monitoring well using a semi-analytical solution of reservoir pressures (Celia et al., 2011) as implemented in NRAP-Open-IAM (King et al., 2018), a framework being developed within the US Department of Energy to evaluate risks associated with GCS. However, the use of this model is not a requirement of the approach, and any other model of reservoir pressures with permeability as an input parameter could be used. The system model in our analysis is defined as the current maximum predicted pressure at the monitoring well defined as

$$P_{F,max}(k) = \max_{i \in M-N} P_{N+i}(k), \quad (1)$$

where k is the uncertain reservoir permeability, N is the current number of collected measurements, and M is the total number of measurements that will ultimately be collected during the GCS operation.

3.3. Concordance metric

As stated above, concordance metrics quantify the level of coherence between measured and modeled quantities, such as pressures, temperatures, resistivities, concentrations, saturations, etc. Many forms of concordance metric exist, including root-mean-squared error (RMSE), sum-of-squared error (SSE), and coefficient of determination (R^2). As each new measurement is obtained, the concordance metric can be up-

dated. Specifically, in our simplified demonstration here, as each new measurement is obtained, the reservoir permeability is re-estimated using all the currently available measurements. We use a Levenberg-Marquardt optimization scheme (Marquardt, 1963), a least-squares gradient-based inverse approach, to re-estimate the reservoir permeability. We use the standard error of the coefficient (reservoir permeability parameter estimate) as our concordance metric defined as

$$\sigma_c = \sqrt{\sigma_e^2 (\mathbf{J}'\mathbf{J})^{-1}} = \sqrt{\frac{\sum_{i=1}^N (P_i - \hat{P}_i)^2}{N-1}} (\mathbf{J}'\mathbf{J})^{-1}, \quad (2)$$

where σ_e is the standard error of the estimate (standard error of the simulated past pressures to the pressure measurements), N is the number of measurements, P_i is the i th measured pressure at the monitoring well, $\hat{P}_i$ is the i th modeled pressure at the monitoring well, and $\mathbf{J}$ is the Jacobian matrix (matrix of derivatives of pressure observations with respect to parameters), where the $(\mathbf{J}'\mathbf{J})^{-1}$ term in Equation 2 is a linear approximation of the covariance matrix. The standard error of the coefficient will decrease over time as more pressure measurements become available, indicating increased confidence in the inference of the reservoir permeability, and hence, increased confidence in the predictions of the system model. Therefore, increased concordance is quantified by decreasing values of the standard error of the coefficient, referred to as the *concordance metric* here.

While similarities can be drawn between the parameter and standard error of the coefficient updating described above and Bayesian inference, it is important to note that the approach above is not truly Bayesian. The approach instead utilizes non-linear regression (Levenberg-Marquardt) to update the permeability coefficient and a linear approximation of the covariance matrix to estimate the standard error of the coefficient. There is no implementation of Bayes' Law in the formulation. However, a Bayesian inference scheme could be utilized here to derive a concordance metric with little or no change to the rest of the conformance robustness analysis approach.

3.4. Performance criterion

We assume that the performance criterion of interest in our synthetic scenario is that pressures not exceed a critical value at the monitoring well. At any time during the injection, while we will know past pressure measurements (and if the GCS operation has been meeting performance criteria in the past), we will not know whether or not future pressure measurements will exceed the critical value. Therefore, our performance criterion is defined as

$$P_{F,max}(k) < P_c, \quad (3)$$

where $P_{F,max}$ is the system model defined in Equation (1), and P_c is the critical pressure. The critical pressure threshold can be considered to be the pressure above which expert opinion based on site-specific information has deemed that the risk of induced seismicity and/or abandoned wellbore leakage are unacceptable for the GCS operation. In other cases, the critical pressure may be defined by regulation, for example Class-VI Area of Review criteria in the United States that sets the area in which leakage of brine or CO₂ may endanger an underground source of drinking water. In the analyses presented here, we choose representative values for P_c to demonstrate the conformance robustness analysis without considering or contriving the details that would be required to select this value in practice. In fact, multiple performance criteria are likely to be required to ensure satisfactory performance of a GCS operation (Chadwick and Noy, 2015), and multiple criteria can be accommodated within the IGDT framework (Harp et al., 2017).

3.5. Uncertainty model

IGDT utilizes set-based uncertainty models, as opposed to distribution-based uncertainty models used in probabilistic analyses. In our case, the uncertainty model is defined as

$$U(h) = \left\{ k : 0 \leq \frac{\log_{10}(\tilde{k}) - \log_{10}(k)}{\sigma_c} \leq h \right\}, h \geq 0, \quad (4)$$

where $\tilde{k}$ is the nominal (least-squares fit) reservoir permeability, σ_c is the standard error of the coefficient defined as the concordance metric above (Equation (2)), and h is referred to as the *horizon of uncertainty* and defines the threshold for nested sets of increasing deviation of reservoir permeability from nominal (normalized by the current standard error of the coefficient). As h increases, larger deviations from nominal are included in the set (in units of standard errors of the coefficient), and thereby more uncertainty is considered. The value of h is not selected, but can instead be varied to investigate the repercussions of different levels of deviation from nominal. Utilizing Equation 4 within IGDT involves ensemble evaluations of the system model for different levels of h to identify the maximum and/or minimum quantity of interest (pressure at the monitoring well in this case).

As discussed above, in general, as more measurements are obtained, the value of standard error of the coefficient will decrease. Dividing the deviation from nominal log permeability (numerator in the center of the inequality in Equation (4)) by the standard error of the coefficient captures the increased confidence in the nominal (least-squares estimated) permeability as more measurements are obtained. In other words, in this uncertainty model, the range of permeability included in one unit of uncertainty (h) decreases (in general) as the estimate of the permeability improves.

Figure 3: Schematic of set-based uncertainty model expressing uncertainty around a nominal estimate of reservoir permeability using nested sets defined by the horizon of uncertainty h .

In Figure 3, we provide a schematic representation of the uncertainty model (Equation (4)). Unlike probabilistic models of uncertainty, the frequencies/probabilities of the occurrence of all events are not defined and there is no upper limit on the horizon of uncertainty (h is unbounded). Set-based uncertainty models, such as Equation (4), describe cases where we have a good idea of a nominal value (i.e., the value that we would use if we had to make a decision today); however, given a lack of information, deviations from nominal are expected. The lack of information also precludes the ability to assign probabilities of occurrence of all potential events (i.e., reservoir permeabilities in our example) or provide definite limits for the potential magnitude of deviations from nominal.

Bayesian approaches utilize expert knowledge to define prior probabilistic distributions in order to fill the gap left in the available data. While this approach is widely used (Varis and Kuikka, 1997; Johnson et al., 2010; Kuhnert et al., 2010), confidence in these analyses can be undermined by the heuristics and biases known to exist during expert elicitation of the probability of events (Tversky and Kahneman, 1974; Kynn, 2008; Morgan, 2014). Due to the sensitivity of Bayesian analyses to the choice of prior distributions (Berger, 1990; Lambert et al., 2005; Lopes and Tobias, 2011; Müller, 2012), the lack of well-informed prior distributions can limit the confidence that can be placed on the results of such analyses. Robust Bayesian approaches (Berger, 1990; Insua and Ruggeri, 2012) have been developed to deal with this, but they require a sensitivity analysis of alternative prior distributions entailing multiple Bayesian analyses which are often computationally infeasible for analyses of complex systems, such as GCS operations. Because failures of GCS operations will likely be due to events beyond anticipation (e.g., pressure or CO_2 plume encountering an unknown abandoned wellbore or fault), set-based uncertainty models attempt to capture the potential for events beyond anticipation (i.e., the horizon of uncertainty has no upper bound). This is not to say that

unanticipated events are explicitly considered, but that the horizon of uncertainty has not limit, and can consider deviations from nominal to totally unanticipated. IGDT is an important alternative approach for GCS due to the often-limited data availability at GCS sites and the expense associated with additional data collection.

A unique feature of our non-probabilistic uncertainty model is that it incorporates a probabilistic statistic, the standard error of the coefficient, as a normalization factor. This allows the uncertainty model to capture the increased probabilistic confidence in the parameter estimate while the uncertainty model can still consider non-probabilistic uncertainty beyond the envelope of probabilistic uncertainty.

3.6. Conformance robustness metric

The conformance robustness metric requires four components: (1) the system model (Equation (1)), (2) the concordance metric (Equation (2)), (3) the performance criterion (Equation (3)), and (4) the uncertainty model (Equation (4)), all defined in the subsections above. Based on these components, the conformance robustness is derived as a robustness function from IGDT as

$$\hat{h}(P_c, \theta) = \max \left\{ h : \left(\max_{k \in U(h)} P_{F,max}(k) \right) \leq P_c \right\}, \quad (5)$$

where θ is a vector of GCS operational decisions (e.g., CO₂ injection rate, brine extraction rate) and all the other variables have been defined above. Evaluation of the term within the round brackets ($\max_{k \in U(h)} P_{F,max}(k)$) requires ensemble evaluations of the system model ($P_{F,max}(k)$, Equation 1) sampling over the uncertainty in the reservoir permeability according to the uncertainty model for different values of h (i.e., $k \in U(h)$). Based on these ensemble evaluations, values of the conformance robustness metric are determined by identifying the maximum value of h associated with an ensemble where none of the predicted pressures exceed the critical pressure P_c , as required by Equation 3. In plain language, the conformance robustness (Equation (5)) quantifies *how wrong can the nominal estimate of reservoir permeability be and still not have future measurements exceed the critical pressure at the monitoring well*. This is accomplished through the maximization of h as defined by Equation 5.

The uncertainty model (Equation (4)) quantifies deviation from nominal reservoir permeability in units of the current standard error of the coefficient. Of course, this does not account for the noise in the measurements, or assumes that the noise is simply Gaussian measurement error and that the least-square predictions are accurate representations of future reservoir pressures. Lower, non-zero values of conformance robustness are still potentially risky as they are still within the range of the probabilistic uncertainty of the reservoir

Figure 4: Subfigure (a) is a plot of decreasing least-square standard errors of the coefficient (i.e., the concordance metric) as more pressure measurements are obtained over time. Subfigure (b) is a plot of the least-squares permeability estimate (circles) and least-squares standard errors of the coefficient (vertical bars) as more measurements are obtained over time. In subfigure (b), the dashed horizontal line is the synthetic true reservoir permeability.

permeability estimate. However, our analysis is not limited to only consider probabilistic uncertainty, and we are therefore able to quantify the conformance robustness for events well beyond what would be included in probabilistic uncertainty alone (i.e., events beyond anticipation or conceptual model inadequacies). In our results, we indicate the portions of the conformance robustness curves associated with 1, 2, and 3 standard errors of the coefficient (log permeability parameter estimate) from nominal, and propose that confidence in conformance is achieved at higher values of robustness beyond the envelope of probabilistic uncertainty (considered as three standard errors of the coefficient here (99.7 % of the probabilistic uncertainty)). The larger the allowable deviations from nominal beyond probabilistic uncertainty, the more robust is the conformance.

4. Results

In order to demonstrate the conformance robustness analysis, we generated synthetic pressure measurements using the same analytical solution used for the system model given a *true* reservoir permeability of $1 \times 10^{-11} \text{ m}^2$ and added Gaussian noise with 0.5 MPa standard deviation to the pressures. The conformance robustness analysis is initiated after two pressure measurements have been collected, where pressure measurements are collected every 15 days. In Figure 4a, standard errors of the coefficient (the concordance metric defined above in Equation (2)) are plotted over time as more and more pressure measurements are collected. Figure 4b contains the same information presented as error bars around the permeability estimate over time. In both plots, the standard error of the coefficient decreases, in general, over time as more measurements are

Figure 5: Pressure predictions (blue lines) and pressure band associated with one standard error of the coefficient (blue shaded regions) when 2, 10, and 18 pressure measurements are available. The best-fit simulated pressures (gray lines) and associated one standard error of the coefficient (gray shaded regions) are shown for reference. The currently available pressure measurements are shown as circles in each plot.

obtained and confidence in the permeability estimate improves. A perfect match is not obtained over time because noise has been added to the synthetic pressure measurements as described above.

In Figure 5, we present the effect of the uncertainty of the reservoir permeability estimate on pressure predictions. The three plots in Figure 5 present uncertainty bands for one standard error of the coefficient when 2, 10, and 18 measurements are available at 15, 135, and 255 days after injection starts, respectively. The pressure uncertainty bands are generated by running the model with plus and minus one standard error of the coefficient, which, from a purely statistical (probabilistic) point of view, would be associated with 68.2 % of the measurements. As more measurements become available, the width associated with plus/minus one standard error of the coefficient decreases. The result on the conformance robustness is that our nominal estimate of the reservoir permeability can be wrong by a greater number of standard errors of the coefficient (h in Equation 4 and 5) and the GCS project is still forecasted to meet the performance criterion. Or, in other words, greater numbers of deviations from nominal permeability estimates are possible before the critical pressure is exceeded, resulting in increased conformance robustness. In short, our nominal understanding can be more incorrect, and the project is still forecasted to be successful. This indicates the increase in confidence in the GCS project conformance over time as more data become available.

There is little change in the best-fit reservoir permeability (Figure 4b) and resulting pressure predictions

Figure 6: Least-squares predicted pressures (solid black lines) estimated over the course of the operation as more and more measurements are collected. The start of each line indicates that its predictions were made after the previous measurement was collected. Circles are the measured pressures and colored dotted lines indicate alternative critical pressures used in the analysis here.

(Figure 5) due to the fact that Gaussian noise is added to the synthetically generated pressure measurements using the same model used in the system model in the example. In practice, it is likely (almost guaranteed) that system predictions will evolve more drastically as the model update process may involve conceptual model updates as more information is obtained, as opposed to the simple parameter updating used here for illustrative purposes. As mentioned previously, we use a simpler CO₂ injection scenario than would likely be adequate in practice so that the conformance robustness concepts can be presented as clearly as possible.

In Figure 6, we present pressure predictions produced as new measurements are obtained as solid black lines. Each line starts at its first predicted pressure, where the preceding measured pressures (circles) in each case were used to estimate the permeability for the predictions. The maximum pressure prediction for each line represents the system model defined above in Equation (1) evaluated with the current least-squares (nominal) permeability estimate ($\tilde{k}$). The pressure predictions change over time as the permeability estimate is updated using newly obtained pressure measurements, as illustrated in Figure 5. The critical pressure thresholds evaluated below (10.0, 10.1, and 10.2 MPa) are indicated in Figure 6 for reference. The critical pressure thresholds define the performance criterion defined above (P_c in Equation (3)). While in practice, injection times would be longer, a shorter injection time is sufficient with the simplified GCS scenario used here to illustrate the robustness conformance concept.

In Figure 7, we plot the conformance robustness for the three critical pressures and an injection rate of

Figure 7: Evaluation of the robustness against future projections exceeding a critical pressure at the monitoring well (P_c) evaluated over time as additional pressure measurements are collected. Three critical pressures are evaluated (different colored lines). Conformance robustness is defined as the potential to deviate from nominal reservoir permeability in units of the standard error of the coefficient and not exceed the critical pressure at the monitoring well. The gray horizontal bars indicate standard errors of the coefficient from the nominal (best-fit) permeability estimate. The probability of a future measurement in each region is indicated.

1 Mt/y. The regions of the conformance robustness between 0 and 1, 1 and 2, 2 and 3, and greater than 3 standard errors of the coefficient are indicated with gray shading in the figure. As indicated in the plot, these regions are associated with reservoir permeability probabilities of 34.1, 13.6, 2.1, and 0.1%, respectively. However, the actual range in permeability deviations associated with these bands changes over time as the standard error of the coefficient is updated as more measurements are obtained. The values of conformance robustness indicate the number of (current) standard errors of the coefficient that the nominal estimate of the permeability can deviate from the true value and the GCS operation will not exceed the critical pressure at the monitoring well (refer to Equation 5 for details). In other words, larger values indicate that our nominal estimate of the reservoir permeability can be more incorrect (in units of the current probabilistic uncertainty, i.e., standard errors of the coefficient)) and the GCS operation is still forecasted to meet performance criteria.

Note that the conformance robustness will only be revealed over time as more measurements are obtained and the reservoir permeability estimate and standard error of the coefficient are updated. In Figure 7, a critical pressure of 10.2 MPa achieves conformance robustness beyond three standard errors of the coefficient at around 180 days (i.e., the green line exceeds a conformance robustness of 3 at 180 days), while 10.1 MPa (orange line) requires around 450 days to do the same, and 10.0 MPa (blue line) is never able to achieve

significant conformance robustness throughout the GCS operation and remains within one standard error of the coefficient or is zero.

A conformance robustness of zero indicates that the maximum pressure prediction exceeds the critical pressure. While the robustness for all three critical pressures starts below 1, the trend for 10.1 (orange line) and 10.2 (green line) MPa as more measurements are obtained indicate that confidence in conformance is increasing, while for 10.0 MPa (blue line), it is quickly apparent that the GCS operations is not likely to meet the performance criterion throughout the operation assuming current operating conditions. This example illustrates that while it may be difficult to demonstrate a high degree of confidence in GCS conformance early in the project, over time as more measurements become available, it will become clear whether you are trending towards conformance or not.

Reviewing Figure 6, the lack of robustness when P_c is 10.0 MPa is apparent as many pressure measurements (and one forecast) exceed this critical value. In Figure 7, when P_c is 10.1 MPa, the GCS operation is more robust, but remains within the range of probabilistic uncertainty (3 standard errors of the coefficient) for most of the GCS operation. It is only when P_c is 10.2 MPa that significant conformance robustness is achieved during the GCS operation.

The analysis in Figure 7 illustrates the trajectories of alternative performance criteria. However, in practice, pressure performance criteria would most likely be determined by site-specific information and expert opinion, and an evaluation of alternative critical pressures would not be warranted. However, for GCS operational performance criteria, such as injecting a desired quantity of CO₂, the robustness of achieving alternative injection quantities may be evaluated.

Along these lines, Figure 8 evaluates the conformance robustness of alternative injection rates for the same problem using $P_c = 10.2$ MPa. The analysis considers 0.75, 1.0, and 1.25 Mt/y injection rates. As expected, conformance robustness increases with decreasing injection rate as smaller injection rates result in lower pressure in the reservoir. Once again, as in Figure 7, conformance robustness is low at early times; but as more monitoring data are obtained, the trend in conformance robustness quickly becomes clear, where for 0.75 and 1.0 Mt/y, it is apparent that conformance robustness is increasing, while for 1.25 Mt/y, it becomes quickly apparent that it is not increasing quickly and that the risk of failing to meet the performance criterion (not exceeding a critical pressure while meeting an injection goal) during the GCS operation is significant as the trend will likely remain within the probabilistic envelope of robustness.

Figure 8: Evaluation of the robustness against future projections exceeding a critical pressure at the monitoring well (P_c) evaluated over time as additional pressure measurements are collected (i.e., conformance robustness). Three injection rates are evaluated (different colored lines) given $P_c = 10.2$ MPa. Conformance robustness is defined as the potential to deviate from nominal reservoir permeability in units of the standard error of the coefficient and not exceed the critical pressure at the monitoring well. The gray horizontal bars indicate standard errors of the coefficient from the nominal (best-fit) permeability estimate. The probability of a future measurement in each region is indicated.

5. Discussion

While significant efforts have resulted in strategies to optimize aspects of CO₂ sequestration operations (Birkholzer et al., 2012) and the development of novel injection strategies (Buscheck et al., 2016a,b), less attention has been applied to ensuring that GCS operations are robust in achieving their objectives. As the GCS community strives to achieve Gigatonne-scale operations, it is imperative to ensure that we have formal approaches to make confident (robust) assessments and/or confirmations of conformance. Field-scale GCS operations such as Snøhvit and In Salah have demonstrated that we need to be prepared to evaluate the potential for unanticipated events and inadequacies in our conceptual models. In this paper, we have proposed and demonstrated a formal approach to evaluate the robustness of an assessment of conformance at a GCS operation.

To apply a conformance robustness analysis in practice making use of a complex numerical model of the GCS operation, multiple uncertainties, multiple performance criteria and longer time scales, the same steps as outlined in the Section 2 would be followed for each performance criterion (Harp et al., 2017). To incorporate multiple uncertainties, the uncertainty model would be multi-dimensional where the sets would include the normalized deviations from nominal for each uncertainty. Examples of such IGDT uncertainty models can

be found in Ben-Haim (2006). The results of these more complex analyses may reveal non-obvious trade-offs and nuances in conformance robustness (Harp et al., 2017). The overall conformance robustness for multiple performance criteria is the minimum of the individual performance criteria conformance robustnesses.

Examples of how to develop robustness functions in a more complicated GCS scenario can be found in Harp et al. (2017). It is important to note that there will be some key differences in the robustness functions in Harp et al. (2017), where the robustness functions are developed to support decisions regarding pressure management strategies, compared to the robustness function presented here, where the robustness function is designed to support designations of conformance. Specifically, the robustness function here includes the standard error of the coefficient so that increased confidence in the model as more monitoring data become available is captured, while this is not included or necessary in Harp et al. (2017).

An example of a more complex and practical performance criteria would be that the CO₂ plume stabilizes within a specified time frame. Within the current framework, this would be incorporated using (1) a *system model* of plume migration, (2) a *concordance metric* defining the discrepancy between the simulated and observed plume location (where the plume location may be observed through monitoring wells or geophysics), (3) *performance criteria* defining the critical time frame within which the plume migration should have stopped, or some other measure of plume stability is met based on pressure, CO₂ saturation, geophysical, or other measurements, and (4) *uncertainty model* that captures all relevant uncertainties that may effect the plume stability, such as permeabilities, porosities, strata dip angle, heterogeneities, etc. Once these four components have been defined, a *conformance robustness metric* can be derived in a similar fashion as presented here. This is not to trivialize the process of developing these components, which will require expertise and skill, but to demonstrate that even in more complex scenarios, the framework presented here provides a formal approach to evaluate conformance robustness.

The evaluation of the conformance robustness (Equation (5)) requires ensemble evaluations of system models involving numerical simulations such as would likely be used for a GCS operation. Robustness functions can be analytically derived in some cases when system models are represented using analytical solutions (Ben-Haim, 2006), however, in practice analytical solutions may not be sufficient to capture important details of a GCS system. The ensemble evaluations are required in order to quantify the deviation from nominal at which the performance criteria are still met. In practice this can be challenging for computationally intensive models, but is required in order to quantify robustness. However, the ensemble analysis will not require probabilistic convergence, and in general, will require fewer simulations than required for Bayesian analyses (thousands to millions of simulations to reach probabilistic convergence) and far fewer than would

be required for a Robust Bayesian analysis (i.e., multiple Bayesian analyses).

6. Summary

In this paper, we developed and described a metric to quantify the robustness of conformance. In summary, the conformance robustness:

1. quantifies the ability for the best-fit model to be incorrect and still meet performance criteria (the definition of conformance robustness as defined here),
2. accounts for increased confidence in conformance over time as monitoring data are obtained by incorporating the standard error of the coefficient (defined as the concordance metric here) in the uncertainty model,
3. incorporates probabilistic uncertainty by normalizing the non-probabilistic uncertainty model by the standard error of the coefficient (concordance metric),
4. considers non-probabilistic uncertainty using a set-based uncertainty model that allows deviations from nominal beyond the envelope of probabilistic uncertainty to be considered,
5. can consider multiple sources of uncertainty and performance criteria (not shown here).

Acknowledgments

This work was supported by the US Department of Energy through the Los Alamos National Laboratory. Los Alamos National Laboratory is operated by Triad National Security, LLC, for the National Nuclear Security Administration of U.S. Department of Energy (Contract No. 89233218CNA000001). 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 Coal Research program. The authors wish to acknowledge Traci Rodosta and Mary K. Underwood (NETL Strategic Center for Coal) and Mark Ackiewicz and Darin Damiani (DOE Office of Fossil Energy) for programmatic guidance, direction, and support. We also thank Abdullah Cihan and Robert M. Dilmore for constructive reviews of this manuscript.

Bacon, D. H., Qafoku, N. P., Dai, Z., Keating, E. H., Brown, C. F., 2016. Modeling the impact of carbon dioxide leakage into an unconfined, oxidizing carbonate aquifer. *International Journal of Greenhouse Gas Control* 44, 290–299.

Ben-Haim, Y., 2006. *Info-gap decision theory: decisions under severe uncertainty*. Elsevier.

421 Berger, J. O., 1990. Robust Bayesian analysis: sensitivity to the prior. *Journal of statistical planning and*
422 *inference* 25 (3), 303–328.

423 Bielicki, J. M., Pollak, M. F., Deng, H., Wilson, E. J., Fitts, J. P., Peters, C. A., 2016. The leakage risk
424 monetization model for geologic co₂ storage. *Environmental science & technology* 50 (10), 4923–4931.

425 Birkholzer, J. T., Cihan, A., Zhou, Q., 2012. Impact-driven pressure management via targeted brine ex-
426 tractionconceptual studies of CO₂ storage in saline formations. *International Journal of Greenhouse Gas*
427 *Control* 7, 168–180.

428 Buscheck, T. A., Bielicki, J. M., White, J. A., Sun, Y., Hao, Y., Bourcier, W. L., Carroll, S. A., Aines, R. D.,
429 2016a. Pre-injection brine production in CO₂ storage reservoirs: An approach to augment the development,
430 operation, and performance of CCS while generating water. *International Journal of Greenhouse Gas*
431 *Control* 54, 499–512.

432 Buscheck, T. A., White, J. A., Carroll, S. A., Bielicki, J. M., Aines, R. D., 2016b. Managing geologic co 2
433 storage with pre-injection brine production: a strategy evaluated with a model of co 2 injection at snøhvit.
434 *Energy & Environmental Science* 9 (4), 1504–1512.

435 Cappa, F., Rutqvist, J., 2011. Impact of CO₂ geological sequestration on the nucleation of earthquakes.
436 *Geophysical Research Letters* 38 (17).

437 Celia, M. A., Nordbotten, J. M., Court, B., Dobossy, M., Bachu, S., 2011. Field-scale application of a
438 semi-analytical model for estimation of CO₂ and brine leakage along old wells. *International Journal of*
439 *Greenhouse Gas Control* 5 (2), 257–269.

440 Chadwick, R. A., Noy, D. J., 2015. Underground CO₂ storage: demonstrating regulatory conformance by
441 convergence of history-matched modeled and observed CO₂ plume behavior using sleipner time-lapse
442 seismics. *Greenhouse Gases: Science and Technology* 5 (3), 305–322.

443 Chen, B., Harp, D. R., Lin, Y., Keating, E. H., Pawar, R. J., 2018. Geologic co 2 sequestration monitoring
444 design: A machine learning and uncertainty quantification based approach. *Applied energy* 225, 332–345.

445 Cheng, A., Huang, L., Rutledge, J., 2010. Time-lapse vsp data processing for monitoring co 2 injection. *The*
446 *Leading Edge* 29 (2), 196–199.

447 Curry, T. E., Reiner, D. M., Ansolabehere, S., Herzog, H., 2005. How aware is the public of carbon capture
448 and storage? MIT.

- Dafflon, B., Wu, Y., Hubbard, S. S., Birkholzer, J. T., Daley, T. M., Pugh, J. D., Peterson, J. E., Trautz, R. C., 2012. Monitoring co₂ intrusion and associated geochemical transformations in a shallow groundwater system using complex electrical methods. *Environmental science & technology* 47 (1), 314–321.
- Dempsey, D., Kelkar, S., Pawar, R., Keating, E., Coblenz, D., 2014. Modeling caprock bending stresses and their potential for induced seismicity during CO₂ injection. *International Journal of Greenhouse Gas Control* 22, 223–236.
- Doetsch, J., Kowalsky, M. B., Doughty, C., Finsterle, S., Ajo-Franklin, J. B., Carrigan, C. R., Yang, X., Hovorka, S. D., Daley, T. M., 2013. Constraining co₂ simulations by coupled modeling and inversion of electrical resistance and gas composition data. *International Journal of Greenhouse Gas Control* 18, 510–522.
- Elliot, T. R., Dammel, J., Buscheck, T. A., Rohmer, J., Celia, M. A., et al., 2012. Promising synergies to address water, sequestration, legal, and public acceptance issues associated with large-scale implementation of CO₂ sequestration. *Mitigation and Adaptation Strategies for Global Change* 17 (6), 569–599.
- Grude, S., Landrø, M., Dvorkin, J., 2014. Pressure effects caused by CO₂ injection in the Tubåen Fm., the snøhvit field. *International Journal of Greenhouse Gas Control* 27, 178–187.
- Harbert, W., Daley, T. M., Bromhal, G., Sullivan, C., Huang, L., 2016. Progress in monitoring strategies for risk reduction in geologic co₂ storage. *International Journal of Greenhouse Gas Control* 51, 260–275.
- Harp, D. R., Stauffer, P. H., OMalley, D., Jiao, Z., Egenolf, E. P., Miller, T. A., Martinez, D., Hunter, K. A., Middleton, R. S., Bielicki, J. M., et al., 2017. Development of robust pressure management strategies for geologic CO₂ sequestration. *International Journal of Greenhouse Gas Control* 64, 43–59.
- Insua, D. R., Ruggeri, F., 2012. *Robust Bayesian Analysis*. Vol. 152. Springer Science & Business Media.
- Johnson, S. R., Tomlinson, G. A., Hawker, G. A., Granton, J. T., Feldman, B. M., 2010. Methods to elicit beliefs for Bayesian priors: a systematic review. *Journal of Clinical Epidemiology* 63 (4), 355–369.
- Keating, E. H., Fessenden, J., Kanjorski, N., Koning, D. J., Pawar, R., 2010. The impact of CO₂ on shallow groundwater chemistry: observations at a natural analog site and implications for carbon sequestration. *Environmental Earth Sciences* 60 (3), 521–536.
- Keating, E. H., Harp, D. H., Dai, Z., Pawar, R. J., 2016. Reduced order models for assessing CO₂ impacts in shallow unconfined aquifers. *International Journal of Greenhouse Gas Control* 46, 187–196.

- King, S., Vasylykivska, V., Bacon, D., Harp, D., Chen, B., Mansoor, K., Onishi, T., Y, Y., Y, Z., E, K., 2018. NRAP-Open-IAM (version alpha-0.5) [computer software] <http://edx.netl.doe.gov/nrap>.
- Kuhnert, P. M., Martin, T. G., Griffiths, S. P., 2010. A guide to eliciting and using expert knowledge in Bayesian ecological models. *Ecology letters* 13 (7), 900–914.
- Kumar, A., 2010. Time-lapse seismic modeling and production data assimilation for enhanced oil recovery and co2 sequestration. Ph.D. thesis, Texas A & M University.
- Kynn, M., 2008. The heuristics and biases bias in expert elicitation. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 171 (1), 239–264.
- Lambert, P. C., Sutton, A. J., Burton, P. R., Abrams, K. R., Jones, D. R., 2005. How vague is vague? a simulation study of the impact of the use of vague prior distributions in mcmc using winbugs. *Statistics in medicine* 24 (15), 2401–2428.
- Lewicki, J., Oldenburg, C., Dobeck, L., Spangler, L., 2007. Surface CO₂ leakage during two shallow subsurface CO₂ releases. *Geophysical Research Letters* 34 (24).
- Lin, Y., Huang, L., 2015. Quantifying subsurface geophysical properties changes using double-difference seismic-waveform inversion with a modified total-variation regularization scheme. *Geophysical Supplements to the Monthly Notices of the Royal Astronomical Society* 203 (3), 2125–2149.
- Linstrom, P., 2003. NIST Standard Reference Database Number 69. NIST Chemistry WebBook(retrieved August 27, 2018).
- Little, M. G., Jackson, R. B., 2010. Potential impacts of leakage from deep CO₂ geosequestration on overlying freshwater aquifers. *Environmental science & technology* 44 (23), 9225–9232.
- Lopes, H. F., Tobias, J. L., 2011. Confronting prior convictions: On issues of prior sensitivity and likelihood robustness in Bayesian analysis. *Annu. Rev. Econ.* 3 (1), 107–131.
- Lucier, A., Zoback, M., Gupta, N., Ramakrishnan, T., 2006. Geomechanical aspects of CO₂ sequestration in a deep saline reservoir in the Ohio River Valley region. *Environmental Geosciences* 13 (2), 85–103.
- Marquardt, D. W., 1963. An algorithm for least-squares estimation of nonlinear parameters. *Journal of the society for Industrial and Applied Mathematics* 11 (2), 431–441.

503 Miller, E., Bell, L. M., Buys, L., 2007. Public understanding of carbon sequestration in Australia: Socio-
504 demographic predictors of knowledge, engagement and trust. *International Journal of Emerging Technolo-*
505 *gies and Society* 5 (1), 15–33.

506 Morgan, M. G., 2014. Use (and abuse) of expert elicitation in support of decision making for public policy.
507 *Proceedings of the National Academy of Sciences*, 201319946.

508 Müller, U. K., 2012. Measuring prior sensitivity and prior informativeness in large Bayesian models. *Journal*
509 *of Monetary Economics* 59 (6), 581–597.

510 Navarre-Sitchler, A. K., Maxwell, R. M., Siirila, E. R., Hammond, G. E., Lichtner, P. C., 2013. Elucidating
511 geochemical response of shallow heterogeneous aquifers to CO₂ leakage using high-performance computing:
512 Implications for monitoring of CO₂ sequestration. *Advances in Water Resources* 53, 45–55.

513 Oldenburg, C. M., 2018. Are we all in concordance with the meaning of the word conformance, and is our
514 definition in conformity with standard definitions? *Greenhouse Gases: Science and Technology* 8 (2),
515 210–214.

516 Oldenburg, C. M., Jordan, P. D., Nicot, J.-P., Mazzoldi, A., Gupta, A. K., Bryant, S. L., 2011. Leakage
517 risk assessment of the In Salah CO₂ storage project: Applying the certification framework in a dynamic
518 context. *Energy Procedia* 4, 4154–4161.

519 Palmgren, C. R., Morgan, M. G., Bruine de Bruin, W., Keith, D. W., 2004. Initial public perceptions of
520 deep geological and oceanic disposal of carbon dioxide. *Environmental Science & Technology* 38 (24),
521 6441–6450.

522 Sminchak, J., Gupta, N., Byrer, C., Bergman, P., 2001. Issues related to seismic activity induced by the
523 injection of CO₂ in deep saline aquifers. Tech. rep., National Energy Technology Laboratory, Pittsburgh,
524 PA, and Morgantown, WV (United States).

525 Sun, A. Y., Zeidouni, M., Nicot, J.-P., Lu, Z., Zhang, D., 2013. Assessing leakage detectability at geologic
526 co₂ sequestration sites using the probabilistic collocation method. *Advances in water resources* 56, 49–60.

527 Trainor-Guitton, W. J., Ramirez, A., Yang, X., Mansoor, K., Sun, Y., Carroll, S., 2013. Value of information
528 methodology for assessing the ability of electrical resistivity to detect co₂/brine leakage into a shallow
529 aquifer. *International Journal of Greenhouse Gas Control* 18, 101–113.

530 Trautz, R. C., Pugh, J. D., Varadharajan, C., Zheng, L., Bianchi, M., Nico, P. S., Spycher, N. F., Newell,
531 D. L., Esposito, R. A., Wu, Y., et al., 2012. Effect of dissolved CO₂ on a shallow groundwater system: A
532 controlled release field experiment. *Environmental science & technology* 47 (1), 298–305.

533 Tversky, A., Kahneman, D., 1974. Judgment under uncertainty: Heuristics and biases. *Science* 185 (4157),
534 1124–1131.

535 Varis, O., Kuikka, S., 1997. BeNe-EIA: A Bayesian approach to expert judgment elicitation with case studies
536 on climate change impacts on surface waters. *Climatic Change* 37 (3), 539–563.

537 White, J. A., Chiaramonte, L., Ezzedine, S., Foxall, W., Hao, Y., Ramirez, A., McNab, W., 2014. Geome-
538 chanical behavior of the reservoir and caprock system at the In Salah co₂ storage project. *Proceedings of*
539 *the National Academy of Sciences*, 201316465.

540 Wilson, E. J., Morgan, M. G., Apt, J., Bonner, M., Bunting, C., Gode, J., Haszeldine, R. S., Jaeger, C. C.,
541 Keith, D. W., McCoy, S. T., et al., 2008. Regulating the geological sequestration of CO₂. *Environmental*
542 *Science & Technology* 42 (8), 2718–2722.