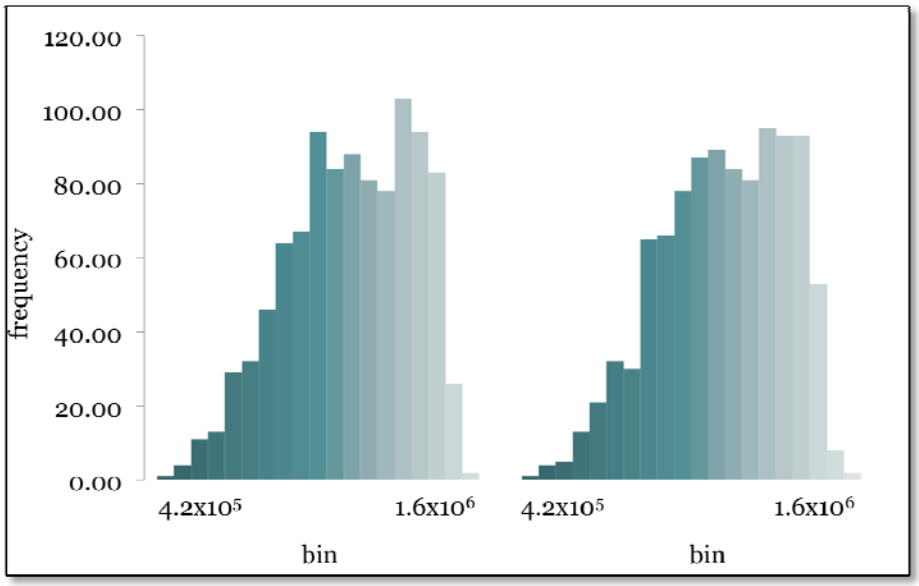
nrap

National Risk
Assessment Partnership



Reduced-Order Model Development for CO_2 Storage in Brine Reservoirs

17 October 2012

U.S. DEPARTMENT OF ENERGY

NETL

Office of Fossil Energy

NRAP-TRS-III-005-2012

Disclaimer

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference therein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed therein do not necessarily state or reflect those of the United States Government or any agency thereof.

Cover Illustration: Histogram of the model output from the Monte Carlo simulation with 1000 samples, using (a) HFM and (b) ROM.

Suggested Citation: Zhang, Y.; Pau, G. *Reduced-Order Model Development for CO_2 Storage in Brine Reservoirs*; NRAP-TRS-III-005-2012; NRAP Technical Report Series; U.S. Department of Energy, National Energy Technology Laboratory: Morgantown, WV, 2012; p 20.

An electronic version of this report can be found at: www.netl.doe.gov/nrap.

Reduced-Order Model Development for CO_2 Storage in Brine Reservoirs

Yingqi Zhang[1] and George Pau[1]

[1]Earth Sciences Division, Lawrence Berkeley National Lab, 1 Cyclotron Road, Berkeley, CA 94720

NRAP-TRS-III-005-2012

Level III Technical Report Series

17 October 2012

This page intentionally left blank

Table of Contents

1. EXECUTIVE SUMMARY 1

2. INTRODUCTION 2

3. METHODOLOGY 3
3.1 GAUSSIAN PROCESS REGRESSION 3
3.2 RADIAL BASIS FUNCTION 4

4. TEST PROBLEM 5
4.1 PROBLEM DESCRIPTION 5
4.2 RESULTS AND DISCUSSION 7
4.3 APPLICATION TO UNCERTAINTY QUANTIFICATION 8

5. APPLICATION TO KIMBERLINA SITE 10

6. DISCUSSION AND FUTURE WORK 12
6.1 IMPROVEMENT OF GP MODEL PERFORMANCE 12
6.2 LIMITATION OF CURRENT APPROACH 12
6.3 OTHER ROM APPROACHES 13
6.4 NEXT STEP 13

7. REFERENCES 14

List of Figures

Figure 1: Flow diagram for the adaptive sampling procedure. Sn is the sample set evaluated by the high-fidelity model to build the ROM, SS is a set of potential sampling points in which to search for the parameter point with the largest estimated approximation error. 4

Figure 2: Liquid saturation at (a) 10 days; and (b) 40 days for log(k)=-15, n=2 and log(1/α)=6. X is the radial distance to the tunnel. The first data point (red dot) shows the liquid saturation of the tunnel element. Interested model output is at X=1.77 m, 2cm away from the tunnel.....6

Figure 3: Liquid saturation at (a) 10 days; and (b) 40 days for log(k)=-14.75, n=2 and log(1/α)=6. X is the radial distance to the tunnel. The first data point (red dot) shows the liquid saturation of the tunnel element. Interested model output is at X=1.77 m, 2cm away from the tunnel.....6

Figure 4: Capillary pressure at 2 cm from the tunnel as a function of the three parameters. 6

Figure 5: Samples used to build a GP-based ROM, selected by the adaptive sampling scheme based on predictive variance. 7

Figure 6: Histogram of the model output from the Monte Carlo simulation with 1000 samples, using (a) HFM and (b) ROM. 9

Figure 7: The relative error from the validation data set using both lookup table (blue line) and the GP partially adaptive model (red line). 11

List of Tables

Table 1: The approximation error for log(k) in the range [-19, -16].....7

Table 2: The approximation error for log(k) in the range [-19, -14].....7

Table 3: Comparison of UQ results between a HFM and its corresponding ROM.....8

Table 4: Comparison of the relative error e_{rel}(p) between the ROMs/LUT and the HFM for different sample sizes 11

Acronyms and Abbreviations

Term	Description
GP	Gaussian process
GPML	Gaussian process regression and classification
HFM	High-fidelity model
isoSE	Isometric squared exponential
LBNL	Lawrence Berkeley National Laboratory
LOOV	Leave one out validation
LUT	Lookup table
PEST	Model-independent parameter estimation (uncertainty analysis software)
RBF	Radial basis function
ROM	Reduced-order model
SA	Sensitivity analysis
TF	Temblor-Freeman
TOUGH2	Transport of unsaturated groundwater and heat (LBNL reservoir simulator)
TOUGH-MP	Massively parallel version of TOUGH2
TP	Thermocouple psychrometers
UQ	Uncertainty quantification
VS	Vedder sand

Acknowledgments

This work was completed as part of National Risk Assessment Partnership (NRAP) project. Support for this project came from the DOE Office of Fossil Energy's Crosscutting Research program. The authors wish to acknowledge Robert Romanosky (NETL Strategic Center for Coal) and Regis Conrad (DOE Office of Fossil Energy) for programmatic guidance, direction, and support.

NRAP funding was provided to Lawrence Berkeley National Laboratory (LBNL) under U.S. Department of Energy Contract No. DE-AC02-05CH11231. The authors wish to thank Haruko Wainwright for providing forward model realizations at the Kimberlina site. Technical review by Stefan Finsterle (LBNL) is greatly appreciated.

1. EXECUTIVE SUMMARY

The objective of this study was to develop a reduced-order model (ROM) method that can be used in the risk assessment of geological carbon sequestration. In this context, the developed ROM does not have to be a simplification or reduction of a high-fidelity forward model. Instead, a response surface for the entire parameter space can be built based on a limited number of high-fidelity forward simulations for selected parameter values. This response surface can then be used to approximate the model output for other parameter values. The focus of the ROM development of this work was within such a framework. The approaches considered included Gaussian process (GP) regression and radial basis functions (RBFs). An adaptive sampling scheme was included in the GP model to improve the accuracy of the ROM with fewer high-fidelity evaluations. GP also has the advantage to provide an estimation uncertainty. Both algorithms were tested for example problems against a lookup table combined with linear interpolation. GP ROM performed better in most of the cases examined, but the accuracy of the model could be further improved. Both approaches (GP and RBF) were implemented into the inverse modeling framework iTOUGH2 (Finsterle, 2010) and could be used for sensitivity analysis and uncertainty quantification.

The implemented GP was applied to approximate the pressure output at certain locations in a hypothetical CO_2 storage project at the Kimberlina site (Zhou et al., 2011). High-fidelity forward simulations were conducted for a 50-year CO_2 injection scenario, using a model initially developed in Birkholzer et al. (2010). Again, GP gave a better approximation than the lookup table with linear interpolation, especially when the number of high-fidelity simulations used to build ROMs or lookup tables was limited. At the current stage, the search of the samples for building the ROM in the adaptive sampling scheme is limited to the available forward simulations. As such, the full potential of an adaptive sampling scheme is not realized. As future work, the adaptive sampling scheme will be re-applied within the iTOUGH2 framework, and expect an even more accurate GP ROM.

2. INTRODUCTION

The objective of this study is to develop a reduced order model (ROM) method based on a model input-output relationship. The developed ROM will be used in risk assessment of geological carbon sequestration where uncertainty quantification is needed.

Subsurface model simulations are becoming more and more time-consuming due to increased complexity in the processes being considered and the fact that both small-scale and large-scale effects need to be accounted for in the same model. Moreover, sensitivity analysis, parameter estimation by inverse modeling, and uncertainty quantification require many such high-fidelity simulations, making the computational demands even more challenging if not prohibitive. One possibility is to use parallel computing. The other alternative is to approximate the high-fidelity model with a computationally much more efficient surrogate model, which is referred to as a reduced-order model (ROM). These ROMs can either be a simplification of the forward model or an approximation of the response surface where a simple input-output relationship is established (Razavi et al., 2012). Even when ROMs are used, parallel computing can still be useful in the high-fidelity simulations to construct the ROM when each high-fidelity simulation is very expensive.

In this study, ROMs were developed based on the response surface method. Particularly, a number of forward simulations were run for selected parameter values, and building a relationship between input parameters and the output of interest, using two methodologies: Gaussian process (GP) regression (Rasmussen and Williams, 2006) and radial basis function (RBF) (Buhmann, 2003). This relationship was then used to approximate the model output for other parameter values and to perform the corresponding analysis, e.g., uncertainty quantification.

3. METHODOLOGY

Two methods were implemented into iTOUGH2 (Finsterle, 2010) to approximate the response surface of a subsurface model: GP regression (Rasmussen and Williams, 2006) and RBF (Buhmann, 2003). These methods are described in the following sections.

3.1 GAUSSIAN PROCESS REGRESSION

The problem can be stated as follows: given a scalar function $f(\mathbf{p})$, where $\mathbf{p}=\{p_1, \ldots, p_n\}$ is a parameter vector of length n, $f(\mathbf{p})$ by $g(\mathbf{p})$ can be approximated using only known solutions of $f(\mathbf{p})$ for $\mathbf{p}$ in a sample set $\boldsymbol{S_N}=\{q_1, \ldots, q_N\}$ of size N. A Gaussian process regression first assumes the relation between $\mathbf{p}$ and $f(\mathbf{p})$ can be described by a Gaussian process characterized by its mean function, $m(\mathbf{p})$, and covariance function, $k(\mathbf{p}, \mathbf{p}')$ (Rasmussen and William, 2006):

$$m(\mathbf{p}) = E[f(\mathbf{p})] \tag{1}$$

$$k(\mathbf{p},\mathbf{p}') = E[(f(\mathbf{p}) - m(\mathbf{p}))(f(\mathbf{p}') - m(\mathbf{p}'))] \tag{2}$$

Knowing $f(\mathbf{q})$, where $\mathbf{q} \in \boldsymbol{S_N}$, the joint distribution of the $f(\mathbf{q})$ and $g(\mathbf{p})$ based on the above prior is then:

$$\begin{bmatrix} f \\ g \end{bmatrix} \sim N\left(m(\mathbf{q}), \begin{bmatrix} K(\mathbf{q},\mathbf{q}) & K(\mathbf{q},\mathbf{p}) \\ K(\mathbf{p},\mathbf{p}) & K(\mathbf{p},\mathbf{p}) \end{bmatrix} \right) \tag{3}$$

For any given $\mathbf{p}$, the GP regression procedure gives the expected value and variance of the approximating function $g(\mathbf{p})$:

$$E[g(p)] = K(p,q)K(q,q)^{-1} f(q) + m(q) \tag{4}$$

$$\sigma^2[g(p)] = K(p,p) - K(p,q)K(q,q)^{-1}K(p,q) \tag{5}$$

The isometric squared exponential (isoSE) function is the covariance function examined in this study:

$$K(\mathbf{p},\mathbf{p}') = \sigma_f \exp(-\frac{1}{2}(\frac{|\mathbf{p}-\mathbf{p}'|}{l})^2) + \sigma_n \delta_{p,p'} \tag{6}$$

In the above definitions, σ_f and σ_n are known as the hyperparameters. These are determined by maximizing the marginal Gaussian likelihood function, which is equivalent to minimizing the following negative log marginal likelihood (Rasmussen and William, 2006).

$$-\log(P(f|\mathbf{p})) = (1/2)(f^T K f + \log(|K|) + n\log(2\pi)) \tag{7}$$

This ROM was built based in part on the Gaussian Process Regression and Classification (GPML) Toolbox version 3.1 (http://gaussianprocess.org/gpml/code). However, the adaptive sampling scheme was developed to minimize the number of forward high-fidelity simulations needed to obtain a ROM with desired accuracy. The adaptive sampling scheme works as follows:

1. An initial ROM $g_1(\mathbf{p})$ is built using simulation at one selected parameter point or a set of parameter points (S_n) obtained through statistical methods
2. An estimate e_n of the approximation error, $g(\mathbf{p})$ -$f(\mathbf{p})$, is determined and evaluated over a large sample set (S_S) in the parameter space. This estimate should not require high fidelity simulations for this sample set
3. The maximum error and the parameter point where this maximum error occurs are determined
4. If the maximum error is less than the error tolerance specified, a ROM is obtained with the desired accuracy; otherwise a high-fidelity simulation is run at this parameter point, it is added to the sample set Sn and the above procedure is repeated until the approximation error is acceptable or number of sample points reaches the maximum allowable number.

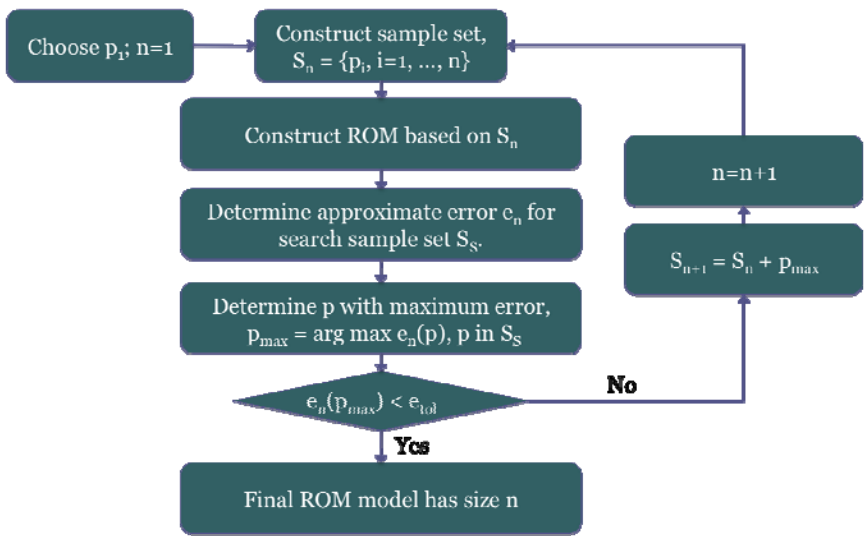


Figure 1: Flow diagram for the adaptive sampling procedure. Sn is the sample set evaluated by the high-fidelity model to build the ROM, SS is a set of potential sampling points in which to search for the parameter point with the largest estimated approximation error.

3.2 RADIAL BASIS FUNCTION

An RBF is a real-valued function whose value depends only on the distance from a point p' (Equation 8). In the context of a ROM, a radial basis function is used to interpolate among points where high-fidelity simulations are performed to approximate the response surface (Equation 9). RBF can have many different functional forms. The Gaussian function (Equation 10) was used to build a ROM for the Kimberlina site. More details can be found in Forrester and Keane (2009).

$$\phi(p, p') = \phi(\|p - p'\|) \tag{8}$$

$$g(p) = \sum_{i=1}^{n} \omega_i \phi(\|p - p^i\|) \tag{9}$$

$$\varphi(r) = \exp(-\frac{r^2}{2\sigma^2}) \tag{10}$$

4. TEST PROBLEM

4.1 PROBLEM DESCRIPTION

The proposed GP approach was tested for a two-phase flow model developed by Finsterle and Pruess (1995). In order to determine the macro-permeability of crystalline rocks, starting on November 26, 1991, a series of ventilation tests was conducted at the Grimsel Rock Laboratory, Switzerland. The experimental site was located in mildly deformed granodiorite that was considered homogeneous on the scale of interest. A section of a tunnel of radius 1.75 m was sealed off and ventilated, causing the formation near the tunnel wall to partially dry out despite the fact that the tunnel is located far below the water table, where considerable hydrostatic pressures prevail. This dry-out occurs because the evaporation rate and vapor transport is higher than the flow rate with which liquid water converges towards the tunnel in this formation of very low permeability. Thermocouple psychrometer (TP) sensors were installed at six different depths to measure negative water potentials in the partially saturated region as a function of time. This analysis focuses on the water potential at 2 cm from the tunnel wall. The total moisture inflow to the tunnel was obtained from measurements of the moisture extracted from the circulated air in a cooling trap. These tests were interpreted using a two-phase, radial flow model implemented using TOUGH2 (Pruess et al., 1999); details of the model are described in Finsterle and Pruess (1995). In the model, the relative permeability function and capillary pressure function of van Genuchten were revised (Finsterle, 2007) for preventing capillary pressure from decreasing towards negative infinity as the effective saturation approaches zero.

Three parameters are considered uncertain in the analysis: the logarithm of the absolute permeability, log(k), and the van Genuchten parameters n and log(1/α). The range of each parameter is [-19,-14], [2,3] and [5,6], respectively. Initially, the tunnel is dry, i.e., 100% gas saturation. The rest of the model domain is fully water saturated. As time goes on, liquid close to the tunnel (left side of the domain) evaporates into the tunnel due to reduced relative humidity maintained by ventilation; a drying front develops, propagating radially out from the tunnel wall. At the same time, the hydrostatic pressure in the far field (right side of the model domain) drives water towards the tunnel, which is fixed at atmospheric pressure. Results of these two competing processes are demonstrated in Figures 2 and 3, which show the liquid saturation of the model domain at 10 and 40 days, with n=2, log(1/α)=6, and log(k)=-15 (Figure 2) and log(k)=-14.75 (Figure 3). For the case with lower permeability, i.e., log(k)=-15, the evaporation front is able to propagate into the formation against the prevailing pressure gradient in the liquid phase. However, an increase in log(k) to -14.75 enables higher liquid flow towards and into the tunnel, which eventually leads to a resaturation of the formation, as shown in Figure 3. (Note that filling the tunnel with liquid water does not correspond to the actual test conditions, which included moisture removal from the sealed-off tunnel section).

As a result of these two competing processes, the capillary pressure at 2 cm appears to be smooth for log(k) between [-19, -16], but shows a sudden change for log(k) between [-16, -14] (Figure 4). This non-linear behavior makes it interesting and challenging for a ROM approximation. To investigate how ROM performance is affected by this sudden change in capillary pressure output, two cases were considered: one in which log(k) may vary between [-19, -16]; the other one in which log(k) may vary between [-19, -14].

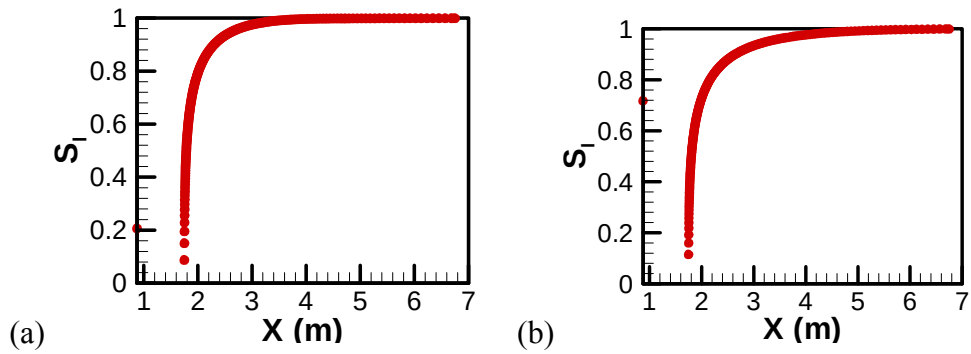


Figure 2: Liquid saturation at (a) 10 days; and (b) 40 days for log(k)=-15, n=2 and log(1/α)=6. X is the radial distance to the tunnel. The first data point (red dot) shows the liquid saturation of the tunnel element. Interested model output is at X=1.77 m, 2cm away from the tunnel.

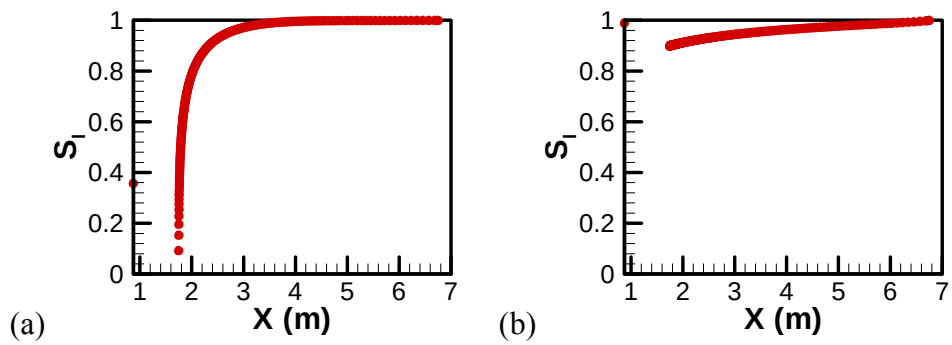


Figure 3: Liquid saturation at (a) 10 days; and (b) 40 days for log(k)=-14.75, n=2 and log(1/α)=6. X is the radial distance to the tunnel. The first data point (red dot) shows the liquid saturation of the tunnel element. Interested model output is at X=1.77 m, 2cm away from the tunnel.

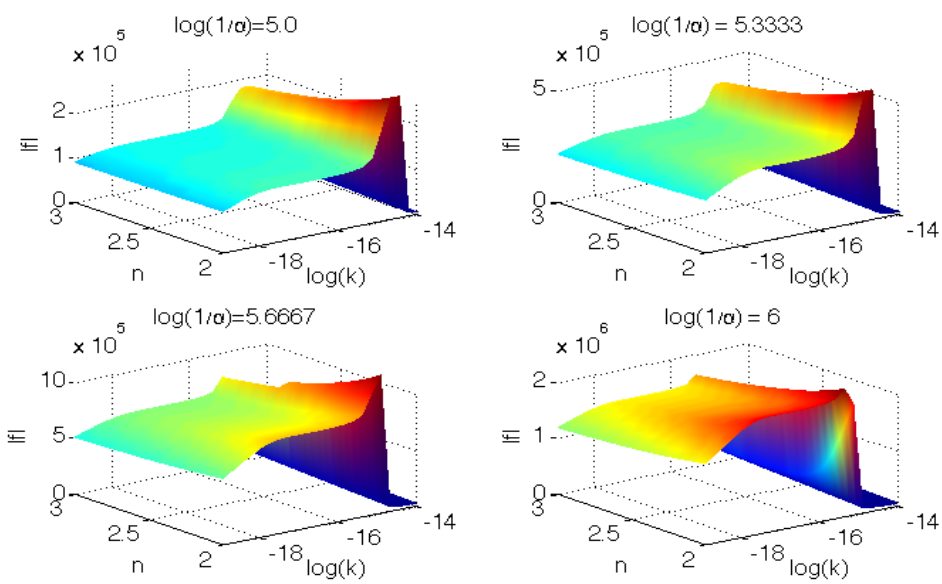


Figure 4: Capillary pressure at 2 cm from the tunnel as a function of the three parameters.

4.2 RESULTS AND DISCUSSION

The relative error (Equation 11) is used to quantify the accuracy of the approximation:

$$e_{\text{rel}}(p) = \frac{|f(p) - g(p)|}{|f(p)|} \tag{11}$$

for **p** within a test sample set (the search sample set, S_S, was used), and to determine maximum, mean and standard deviation of $e_{\text{rel}}(\mathbf{p})$. To evaluate $e_{\text{rel}}(\mathbf{p})$, $f(\mathbf{p})$ for all **p** in S_S needs to be evaluated

For this test problem, the search sample set S_S is 21 × 21 × 21 (i.e., 21 uniformly distributed samples for each parameter). The estimate e_n of the approximation error used in the adaptive sampling scheme is the predictive variance, given by Equation 5. The actual number of samples used to build the ROM is 27. The locations of these samples are shown in Figure 5. This number was chosen for the convenience of comparing results to a linear interpolation result using a tensor grid of 3 × 3 × 3. The comparisons of approximation errors for the two methods are shown in Tables 1 and 2.

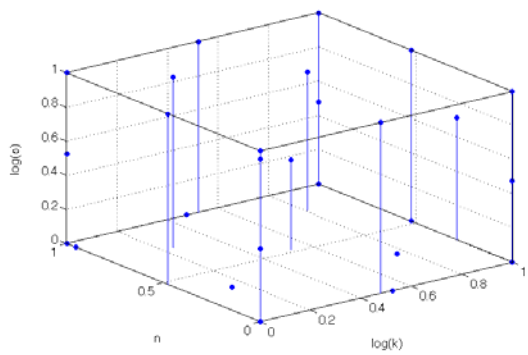


Figure 5: Samples used to build a GP-based ROM, selected by the adaptive sampling scheme based on predictive variance.

Table 1: The approximation error for log(*k*) in the range [-19, -16]

Approach	$e_{\text{rel}}(p)$		
	Maximum	Mean	Standard Deviation
GP model	0.034	0.007	0.007
Linear interpolation	0.091	0.016	0.017

Table 2: The approximation error for log(*k*) in the range [-19, -14]

Approach	$e_{\text{rel}}(p)$		
	Maximum	Mean	Standard Deviation
GP model	0.95	0.32	0.25
Linear interpolation	1.12	0.27	0.30

For this particular test problem, if the model output was smooth over the entire parameter space, the GP model showed a significantly better performance than linear interpolation; if the model output had a sudden change, the performance of the two approaches was approximately the same. Although the mean of the error was 0.27 using a linear interpolation, a little less than a GP model, the standard deviation was higher, which means a higher uncertainty.

The deterioration of GP in the non-smooth case was partially due to the covariance function used (isoSE), which is not suitable for non-smooth functions. There are several potential solutions to improve the GP model performance. A different covariance function can be used, for example one where the hyperparameter is dependent on p (Plagemann, 2008). The accuracy can also be improved by determining the hyperparameters – using a different global optimization algorithm (currently, multistart steepest descent algorithm is used). A different error measure can also lead to a different performance. For example, instead of just using the variance as the error measure for selecting subsequent samples, an "approximate actual error", e.g., obtained by comparing a ROM to a simplified model (e.g., a model with coarser grid), can be combined with the variance as the error measure. Finally, different ROM models can be used in different regions of the parameter space; partition of the parameter space can be guided by the physics of the modeled process. These approaches are currently under investigation.

4.3 APPLICATION TO UNCERTAINTY QUANTIFICATION

The purpose of developing such a ROM was to substitute a time-consuming high-fidelity model by a computationally much more efficient surrogate model in an inverse analysis or sampling-based uncertainty quantification (UQ) analysis, where many forward model evaluations were needed. A UQ based was implemented on ROM into iTOUGH2 and UQ performed for the same sample problem (log(k) in [-19, -16] was considered). Thirty samples were used to build this ROM.

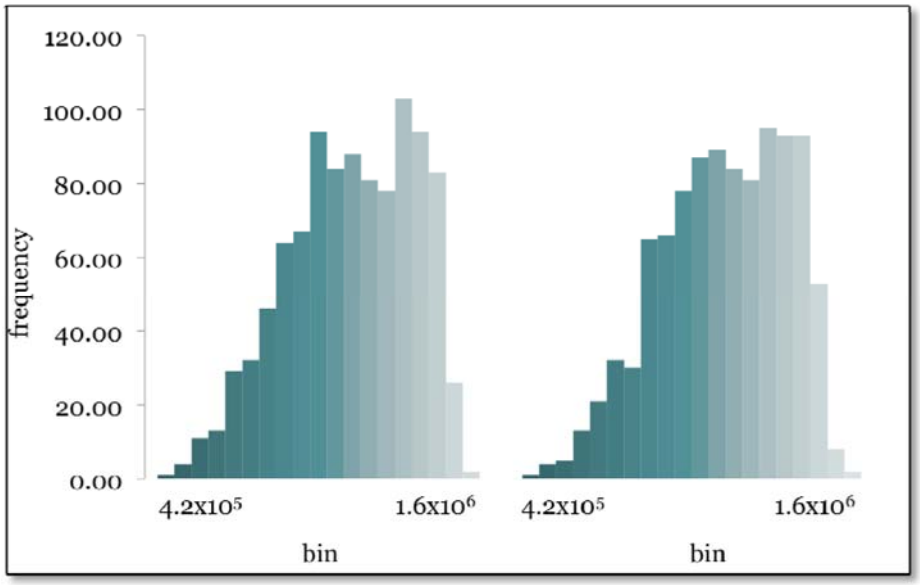
For comparison purposes, a UQ was also performed using the high-fidelity model (HFM). The Monte Carlo simulation was performed with a sampling size of 100 and 1000 for both the ROM and HFM. The mean and variance of the model output from each Monte Carlo simulation are listed in Table 3.

Table 3: Comparison of UQ results between a HFM and its corresponding ROM

	HFM-1000	ROM-1000	HFM-100	ROM-100
Mean	1.11e6	1.15e6	1.14e6	1.17e6
Standard deviation	2.4e5	2.4e5	2.3e5	2.3e5

The ROM seems to be able to re-produce the standard deviation of the UQ analysis. The error of the mean estimation using the ROM was about 3%. For this particular example, 100 samples seem to be sufficient for uncertainty quantification.

The histograms using 1000 samples are plotted in Figures 6 (a) and (b). The ROM appears to re-produce the histogram of the model output relatively well.



(a) (b)

Figure 6: Histogram of the model output from the Monte Carlo simulation with 1000 samples, using (a) HFM and (b) ROM.

5. APPLICATION TO KIMBERLINA SITE

The developed ROM was applied to a hypothetical storage project at the Kimberlina site in the Southern San Joaquin Basin in California, where CO_2 was assumed to be injected and stored. The ultimate goal was to be able (1) to predict the pressure and CO_2 plume evolution at locations of interest (or anywhere in the field that may have leakage), and (2) to perform a prediction uncertainty analysis. A description of the forward model can be found in Birkholzer et al. (2010). The forward simulation was performed using the parallel version of TOUGH2: TOUGH-MP. A sensitivity analysis has been performed for nine input parameters (permeability, porosity and compressibility of three formations) and it was determined that three of them – permeability and porosity of the Vedder Sand (VS-k and VS-ϕ, respectively) and permeability of the caprock (TF-k) – were the most sensitive parameters (Wainwright et al., 2012). Within the same work, an UQ was performed with 245 Monte Carlo realizations of these three parameters. This study used the existing available simulations as the search sample set Ss. The model output of interest was the pressure at a location 1.8 km updip from the injection location at 50 years.

Both the GP and RBF method were applied and compared both to results obtained by linearly interpolation from a lookup table (which can be considered a ROM in itself). Three fixed sample sets S_n on a tensor grid were used to build lookup tables: 8 (2 x 2 x 2), 27 (3 x 3 x 3), 54 (6 x 3 x 3) for VS-k x VS-ϕ x TF-k. The same samples were used to build both GP-and RBF-based ROMs. In addition, a GP model with the adaptive sampling procedure was built. However, since the search set was limited to the 245 available realizations, which were obtained through high-fidelity TOUGH2-MP simulations, the potential of the adaptive sampling could not be fully realized; this compromise was referred to as "GP partially adaptive".

The sample set with 54 samples was given by Wainwright and Birkholzer (2012). The 27 sample set is a subset of those 54, in which VS-k took the first, fourth, and sixth of the samples in the 54 sample set. The 8 sample set, a subset of the 27 sample set, takes the two bounding values of each parameter. The validation data used to calculate the error is the rest of the samples after exclusion of the samples used to build the ROM, i.e., if 27 samples are used to build the ROM or lookup table, the validation data set contains 245-27=218 samples.

To compare these four methods (lookup table, GP with the same samples as lookup table, RBF with the same samples as the lookup table, and GP partially adaptive), the max., mean, and standard deviation of the relative error were calculated, which are listed in Table 4. The following observations were made:

- When the sample size is small, GP gives a smaller mean and maximum approximation error, as well as a smaller standard deviation of this error, which implies a smaller uncertainty. Clearly, it performs better than RBF and lookup table. GP with partial adaptive sampling scheme is better than GP on a fixed grid. An even better performance could be achieved if the full adaptive scheme is applied. This assertion can be validated using iTOUGH2-PEST (Finsterle and Zhang, 2011) that has the adaptive sampling procedure implemented and has the capability to call TOUGH-MP (Zhang et al., 2008) through the PEST protocol (Doughty, 2007).
- When the sample size is large, the high density of interpolation points improves the performance of the lookup table. However, the performance of the GP model with the partial adaptive scheme deteriorates. Due to a lack of variation in the response surface, GP does not

require 54 samples. Instead, it leads to a covariance function that is poorly conditioned, leading to larger numerical error in the approximation. This further illustrates the advantage of using an adaptive sampling procedure that avoids unnecessary forward simulations, and that can intrinsically handle insensitive parameters.

- RBF performs a little better than the lookup table (LUT) when the sample size is small, but not as good as the GP model in general.

Table 4: Comparison of the relative error $e_{rel}(p)$ between the ROMs/LUT and the HFM for different sample sizes

Number of Samples	ROM Type	Max	Mean	Std Dev
8	Lookup table	0.36	0.16	0.08
	GP fixed grid	0.29	0.12	0.07
	GP partially adaptive	0.24	0.08	0.06
	RBF fixed grid	0.38	0.13	0.08
27	Lookup table	0.18	0.05	0.05
	GP fixed grid	0.13	0.03	0.03
	GP partially adaptive	0.11	0.04	0.03
	RBF fixed grid	0.19	0.04	0.04
54	Lookup table	0.12	0.03	0.03
	GP fixed grid	0.13	0.03	0.02
	GP partially adaptive	0.17	0.03	0.03
	RBF fixed grid	0.23	0.09	0.05

A ROM was built using 22 samples, where the corresponding lookup table was not from a tensor grid but based on a tetrahedral grid (unstructured grid) with triangulation (linear interpolation of the vertices of the triangulation) and the 22 samples were determined through the adaptive sampling procedure. It takes about 4 hours to run one forward simulation for the Kimberlina model with 64 K grid blocks to 100 years on a machine with 12 cores. For a ROM evaluation it takes seconds on a PC. For a large-scale simulation like Kimberlina, the computational saving is tremendous for a UQ or sensitivity analysis (SA). Figure 7 shows the relative error using both lookup table (blue line) and the GP partially adaptive model (red line). Clearly, the GP model performs much better for a non-tensor grid.

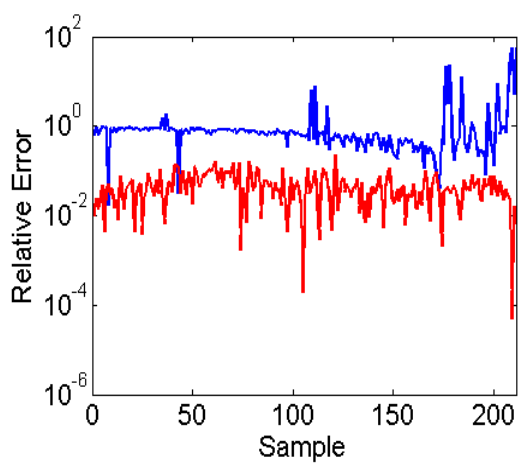


Figure 7: The relative error from the validation data set using both lookup table (blue line) and the GP partially adaptive model (red line).

6. DISCUSSION AND FUTURE WORK

6.1 IMPROVEMENT OF GP MODEL PERFORMANCE

As discussed earlier, there are a few potential ways to improve the performance of a ROM based on GP regression:

- Implement different covariance functions. There are many different covariance function formulations in the literature that model different forms of response surfaces. These options will be implemented and their performance determined. In addition, an automated selection scheme will allow different covariance functions to be examined during the construction of the ROM. Performance of a covariance function can be determined through the leave one out validation (LOOV) procedure, thus avoiding the need for a large validation sample set.
- A different type of global optimization algorithm (such as evolutionary algorithms or simulated annealing) may help to find hyperparameters that give a more accurate ROM.
- A different estimate of the approximation error can be used for the adaptive sampling scheme. An example is that of a physics-based ROM, e.g., a coarser model can be used as another approximation of the model response. The difference between the two ROMs can be combined with the variance (current estimate of the approximation error) to select subsequent samples for updating the GP ROM. Alternatives for the physics-based ROM will be explored.
- The parameter space can be subdivided into smaller domains, each with its own ROM. If within each domain the model output is smooth, the sample size for each ROM could be very small. As a result, this may have the potential of improving the efficiency of each ROM by reducing the number of samples needed for each ROM. To realize the full potential of this approach, the parameter space should not be partitioned randomly or even uniformly. Rather, it should be guided by the modeling process.

6.2 LIMITATION OF CURRENT APPROACH

In the current implementation of the ROMs, the approximation is done for one model output of interest, e.g., in the Kimberlina case, it is pressure at one point in space and time. However, if the location of interest is uncertain or if there are multiple points of interest (e.g., a potential leakage location, which could be anywhere in the model domain), an approximate solution would be needed for multiple points, or even for the entire model domain at multiple temporal points. For example, a total number of N pressure output in space and time, a total number of N ROMs will need to be built. If it takes Mi (i = 1, N) HFM simulations to build the ith ROM with a pre-specified accuracy, the maximum total number of HFM simulations is $\sum_i^N M_i$. This can be a large computational effort and as such this is a limitation of the current approach. However, the actual number of necessary simulations may be much less since some of the HFM simulations may be used in multiple ROM constructions. It may even be possible to have one limited set of samples, and to construct the multiple ROMs using this same set of samples. An approach that reconstructs the field solution by evaluating on M ROMs, where $M << N$, is currently being investigating. This investigation includes coupling the ROMs based on response surface approach with function approximation methods such as empirical interpolation method and

proper orthogonal decomposition, and consideration of various spatial and temporal locations as additional parameters for GP approximation.

6.3 OTHER ROM APPROACHES

Fuzzy logic may be considered to approximate the model response. The basic idea for such a method is that, based on limited simulations, model responses can be divided into regions so within each region the model output changes smoothly; a set of rules can be generated for each region and used as an approximation.

6.4 NEXT STEP

Our next efforts will focus on limitations of the current approach. More specifically, for the Kimberlina site, (1) a full adaptive sampling scheme will be applied to improve GP model performance; and (2) an approach will be developed that requires minimum effort to approximate the entire field solution.

7. REFERENCES

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*, **2010**.

Buhmann, M.D. *Radial Basis Functions: Theory and Implementations*; Cambridge University Press, 2003.

Doughty C. Modeling geologic storage of carbon dioxide: comparison of non-hysteretic and hysteretic characteristic curves. *Energy Conversion and Management* **2007,** *48*(6), 1768-1781.

Finsterle, S. *iTOUGH2 User's Guide;* Report LBNL-40040; Lawrence Berkeley National Laboratory: Berkeley, CA, 2010.

Finsterle, S. *iTOUGH2 Sample Problems;* Report LBNL-40042 Rev.; Lawrence Berkeley National Laboratory: Berkeley, CA, 2007.

Finsterle, S.; Pruess, K. Solving the estimation-identification problem in two-phase flow modeling, *Water Resour. Res.* **1995,** *31* (4), 913–924.

Finsterle, S.; Zhang, Y. Solving iTOUGH2 simulation and optimization problems using the PEST protocol. *Environmental Modeling and Software* **2011,** *26*, 959–968.

Forrester, A.I.J.; Keane, A.J. Recent advances in surrogate-based optimization. *Progress in Aerospace Sciences* **2009,** *45* (1-3), 50-79.

Plagemann C.; Kersting, K.; Burgard, W. Nonstationary Gaussian Process Regression Using Point Estimates of Local Smoothness. *Lecture Notes in Computer Science*; *In Machine Learning and Knowledge Discovery in Database;* 2008; *5212*, 204-219.

Pruess, K.; Oldenburg, C.; Moridis, G. *TOUGH2 User's Guide, Version 2.0;* Report LBNL-43134; Lawrence Berkeley National Laboratory: Berkeley, CA, 1999.

Rasmussen, C.E.; Williams, C.K.I. *Gaussian Processes for Machine Learning*; MIT Press, 2006.

Razavi, S.; Tolson, B. A.; Burn, D. H. Review of surrogate modeling in water resources. *Water Resour. Res.* **2012,** *48*, W07401.

Wainwright, H.; Finsterle, S.; Zhou, Q.; Birkholzer, J. *Modeling the Performance of Large-Scale CO_2 Storage Systems: A Comparison of Different Sensitivity Analysis Methods;* NRAP-TRS-III-002-2012; NRAP Technical Report Series*;* U.S. Department Energy, National Energy Technology Laboratory: Morgantown, WV*,* 2012.

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, May 2011.

Zhang, K.; Wu, Y.S.; Pruess, K. *User's Guide for TOUGH2-MP – A Massively Parallel Version of the TOUGH2 Code*; Report LBNL-315E; Lawrence Berkeley National Laboratory: Berkeley, CA, 2008.

nrap
National Risk
Assessment Partnership

NRAP is an initiative within DOE's Office of Fossil Energy and is led by the National Energy Technology Laboratory (NETL). It is a multi-national-lab effort that leverages broad technical capabilities across the DOE complex to develop an integrated science base that can be applied to risk assessment for long-term storage of carbon dioxide (CO_2). NRAP involves five DOE national laboratories: NETL-RUA, Lawrence Berkeley National Laboratory (LBNL), Lawrence Livermore National Laboratory (LLNL), Los Alamos National Laboratory (LANL), and Pacific Northwest National Laboratory (PNNL). The NETL-RUA is an applied research collaboration that combines NETL's energy research expertise in the Office of Research and Development (ORD) with the broad capabilities of five nationally recognized, regional universities—Carnegie Mellon University (CMU), The Pennsylvania State University (PSU), the University of Pittsburgh (Pitt), Virginia Tech (VT), and West Virginia University (WVU)—and the engineering and construction expertise of an industry partner (URS Corporation).

NRAP Technical Leadership Team

Jens Birkholzer
LBNL Technical Coordinator
Lawrence Berkeley National Laboratory
Berkeley, CA

Grant Bromhal
NETL Technical Coordinator
Lead, Reservoir Working Group
Office of Research and Development
National Energy Technology Laboratory
Morgantown, WV

Chris Brown
PNNL Technical Coordinator
Lead, Groundwater Working Group
Pacific Northwest National Laboratory
Richmond, WA

Susan Carroll
LLNL Technical Coordinator
Lawrence Livermore National Laboratory
Livermore, CA

Laura Chiaramonte
Lead, Natural Seals Working Group
Lawrence Livermore National Laboratory
Livermore, CA

Tom Daley
Lead, Monitoring Working Group
Lawrence Berkeley National Laboratory
Berkeley, CA

George Guthrie
Technical Director, NRAP
Office of Research and Development
National Energy Technology Laboratory
Pittsburgh, PA

Rajesh Pawar
LANL Technical Coordinator
Lead, Systems Modeling Working Group
Los Alamos National Laboratory
Los Alamos, NM

Tom Richard
Deputy Technical Director, NRAP
The Pennsylvania State University
NETL-Regional University Alliance
State College, PA

Brian Strazisar
Lead, Wellbore Integrity Working Group
Office of Research and Development
National Energy Technology Laboratory
Pittsburgh, PA

Sean Plasynski
Deputy Director
Strategic Center for Coal
National Energy Technology Laboratory
U.S. Department of Energy

Jared Ciferno
Director
Office of Coal and Power R&D
National Energy Technology Laboratory
U.S. Department of Energy

Robert Romanosky
Technology Manager
Office of Coal and Power R&D
National Energy Technology Laboratory
U.S. Department of Energy

Regis Conrad
Director
Division of Cross-cutting Research
Office of Fossil Energy
U.S. Department of Energy

nrap
National Risk
Assessment Partnership

NRAP Executive Committee

Cynthia Powell
Director
Office of Research and Development
National Energy Technology Laboratory
U.S. Department of Energy

Alain Bonneville
Laboratory Fellow
Pacific Northwest National Laboratory

Donald DePaolo
Associate Laboratory Director
Energy and Environmental Sciences
Lawrence Berkeley National Laboratory

Melissa Fox
Chair, NRAP Executive Committee
Program Manager
Applied Energy Programs
Los Alamos National Laboratory

Julio Friedmann
Chief Energy Technologist
Lawrence Livermore National
Laboratory

George Guthrie
Technical Director, NRAP
Office of Research and Development
National Energy Technology Laboratory