Risk Assessment for CO$_2$ Sequestration - Uncertainty Quantification based on Surrogate Models of Detailed Simulations

Yan Zhang
Advisor: Nick Sahinidis
Department of Chemical Engineering
Carnegie Mellon University
3/11/2012
Modeling CO$_2$ Underground Plume

For $\alpha = \text{CO}_2$, brine

\[ \frac{\partial \phi}{\partial t} \rho_\alpha S_\alpha + \nabla \cdot (\rho_\alpha \mathbf{v}_\alpha) = V_\alpha \rho_\alpha; \]

Mass balance

\[ \mathbf{v}_\alpha = -\frac{K k_\alpha}{\mu_\alpha} (\nabla p_\alpha + \rho_\alpha g \nabla z) \]

Darcy’s law
Uncertainty Analysis Options

- Numerical model + Monte Carlo simulation
- Stochastic response surface methods
  - **Surrogate model**
    - Approximate the output of the original model by polynomial functions of the uncertain parameters
  - **Monte Carlo simulation with surrogate model**

The primary goal of this work is to provide an assessment of surrogate-based uncertainty analysis for CO$_2$ injection into a saline aquifer
Polynomial Chaos Expansion (PCE)

* A physical model

\[ Y = M(x) \]

\[ x = \{x_1, \ldots, x_M\}^T \in \mathbb{R}^M, M \geq 1 \quad \text{porosity, permeability, etc} \]

\[ Y = \{y_1, \ldots, y_N\}^T \in \mathbb{R}^N, N \geq 1 \quad \text{CO}_2 \text{ mass, CO}_2 \text{ plume dynamic, etc} \]

* Random \( x \) \rightarrow random model response \( y \)

* Assuming finite variance, chaos representation

\[ y = M(x) = \alpha_0 + \alpha_1 B_1(x) + \alpha_2 B_2(x) + \alpha_3 B_3(x) + \ldots \]

\( \{B_1, B_2, B_3, \ldots\} \): orthogonal polynomial basis functions for \( x \)
Orthogonal Polynomial Basis

- **Orthogonal polynomials for a single random** \( x \)
  \[
  \varphi_d(x) = x^d + \text{lower-degree terms}, \quad d = 0, 1, 2, \ldots
  \]
  \[
  f(x): \text{probability density function} \quad \int_{\Omega} \varphi_d(x) \varphi_c(x) f(x) dx = 0, \quad d \neq c
  \]

- **Multivariate polynomial basis for** \( \mathbf{x} = \{x_1, \ldots, x_M\}^T \in \mathbb{R}^M, \: M \geq 2 \)
  - **Independent**
    \[
    B_d(x_1, \ldots, x_M) = \prod_{i=1}^{M} \varphi_{d_i}(x_i), \quad \sum_{i=1}^{M} d_i = d
    \]
  - **Dependent: Nataf Transformation**
    \[
    \Phi(z_i) = F(x_i) \quad \text{correlated variables} \rightarrow \text{correlated standard normals}
    \]
    \[
    z = L\xi \quad \text{correlated normals} \rightarrow \text{uncorrelated standard normals}
    \]
    \[
    y = M(\xi)
    \]
Coefficient Estimation

\[ y = M_d(x) = \alpha_0 + B_1 \alpha_1 + \ldots + B_d \alpha_d \]

**Regression**

\[ X = \{x^1, \ldots, x^{N_p}\}: \text{a set of } N_p \text{ realizations of random input vectors} \]

\[ y = \{y^1, \ldots, y^{N_p}\}^T = \{y^i = M_d(x^i), i = 1, \ldots, N_p\}^T \]

\[ B \alpha = y \]

\[ \alpha = (B^T B)^{-1} B^T y \]

**Number of points to choose:**

\[ N_t = \frac{(M+d)!}{M!d!} \quad d: \text{degree of expansion, } M: \text{inputs} \]

<table>
<thead>
<tr>
<th></th>
<th>[d = 1]</th>
<th>[d = 2]</th>
<th>[\ldots]</th>
<th>[d = 6]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>3</td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>21</td>
<td></td>
<td>462</td>
</tr>
<tr>
<td>10</td>
<td>11</td>
<td>66</td>
<td></td>
<td>8008</td>
</tr>
</tbody>
</table>
Stepwise PCE Approximation

**Forward selection:** for all degree-\(d\) basis functions, add one term to \(A\) and retain the term if \(Q^2\) increases

**Backward elimination:** for all lower degree (\(< d\)) basis functions, remove one term from \(A\) if \(Q^2\) does not decrease much

S. Weisberg, Applied Linear Regression, 3rd edition, 2005
Assessment of PCE approximation

- **Empirical error**

\[
R^2 = 1 - \frac{1}{N_p} \sum_{i=1}^{N_p} \frac{(y^i - M_d(x^i))^2}{\sigma^2(y)}
\]

\(R^2 \rightarrow 1\) as \(N_t\) increases \(\Rightarrow\) overfitting

- **Leave-one-out cross validation**

\(N_p\) points

Training set

Test set

\(X\)

\((N_p-1)\) points

\(X \setminus \{x^i\}\)

Metamodel

\(M_{d, X \setminus i}\)

1 point

\(x^i\)

Predicted residual

\(\Delta^i = y^i - M_{d, X \setminus i}(x^i)\)
Assessment of PCE approximation

- **Empirical error**

  \[
  R^2 = 1 - \frac{1}{N_p} \sum_{i=1}^{N_p} (y^i - M_d(x^i))^2 \quad \sigma^2(y)
  \]

  $R^2 \to 1$ as $N_t$ increases $\Rightarrow$ overfitting

- **Leave-one-out cross validation**

  \[
  Q^2 = 1 - \frac{1}{N_p} \sum_{i=1}^{N_p} (\Delta^i)^2 \quad \sigma^2(y)
  \]

  \[
  \Delta^i = y^i - M_d(X \setminus \{x^i\})(x^i)
  \]
Case Study

- CO₂ injection into a deep saline aquifer
- Simulated using TOUGH2 developed by LBNL since 1980

\[ \text{output} = M \text{ (porosity, permeability)} \]

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pressure at well</td>
<td>110 bar</td>
</tr>
<tr>
<td>Temperature</td>
<td>37 °C</td>
</tr>
<tr>
<td>Salinity</td>
<td>3.2 wt.-% NaCl</td>
</tr>
<tr>
<td>Injection rate</td>
<td>0.1585 kg/s</td>
</tr>
</tbody>
</table>

Output:
- Gas saturation (space, time)
- Pressure (space, time)
- CO₂ mass distribution

ECO2N manual, 2005
Specify Input Random Variables

- National Petroleum Council Public Database
- Non-standard marginal distributions

- Strong correlation $\rho = 0.8$
- Transform to normal variables
Derive Orthogonal Polynomial Basis

- **Hermite polynomials for a standard normal random variable**
  \[
  \begin{align*}
  \varphi_0(\xi) &= 1 \\
  \varphi_1(\xi) &= \xi \\
  \varphi_2(\xi) &= \xi^2 - 1 \\
  \varphi_3(\xi) &= \xi^3 - 3\xi \\
  \vdots
  \end{align*}
  \]
  \[
  \begin{align*}
  \varphi_{d+1} &= \xi\varphi_d - d\varphi_{d-1}
  \end{align*}
  \]

- **Basis for the two-dimensional independent input vector \((\xi_1, \xi_2)\)**
  \[
  \begin{align*}
  B_0(\xi) &= \varphi_0(\xi_1)\varphi_0(\xi_2) = 1 \\
  B_1(\xi) &= \varphi_1(\xi_1)\varphi_0(\xi_2) = \xi_1 \\
  B_1(\xi) &= \varphi_0(\xi_1)\varphi_1(\xi_2) = \xi_2 \\
  B_2(\xi) &= \varphi_2(\xi_1)\varphi_0(\xi_2) = \xi_1^2 - 1 \\
  B_2(\xi) &= \varphi_1(\xi_1)\varphi_1(\xi_2) = \xi_1\xi_2 \\
  B_2(\xi) &= \varphi_0(\xi_1)\varphi_2(\xi_2) = \xi_2^2 - 1 \\
  \vdots
  \end{align*}
  \]
Select an experimental design $\xi$ (Latin Hypercube Sampling)
Build PCE Model

- **Select an experimental design** $\xi$ (Latin Hypercube Sampling)

- **Reverse of Nataf transformation**
  \[
  z = L\xi \\
  x_i = F^{-1}(\Phi(z_i))
  \]

- **Model output evaluations**
- **Stepwise regression for coefficients and degree**
PCE Model Example 1

100 simulations
$Q^2=0.98$
2nd order expansion

Mass of $CO_2$ in aquifer after 30 days of injection (kg)
PCE Model Example 2

100 simulations
$Q^2 = 0.98$
$4^{th}$ order expansion

Mass of $\text{CO}_2$ in caprock after 30 days of injection (kg)
Use PCEs for Uncertainty Analysis

**Input:**
Randomly sampled from probability distribution function of independent parameters

**Monte Carlo simulation**

**Output:**
Statistical analysis

<table>
<thead>
<tr>
<th>Sample #</th>
<th>$x_1$</th>
<th>$x_2$</th>
<th>...</th>
<th>$x_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Permeability</td>
<td>Porosity</td>
<td>...</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Outputs evaluated using PCEs</th>
</tr>
</thead>
<tbody>
<tr>
<td>$y_1$</td>
</tr>
<tr>
<td>Mass</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
Correlated Sampling Using Copulas

![Graph showing correlation between permeability and porosity](image)

1. 1000 Correlated Samples
2. NPC database

- Permeability (m²)
- Porosity

Log-log scale for better visualization of the relationship between permeability and porosity.
MC Simulation Result 1

The graph shows the empirical cumulative distribution function (CDF) of the fraction of total injected CO$_2$ in an aquifer after 30 days of injection. Two models are compared:

- **TOUGH2 model** represented by a solid blue line.
- **2nd order surrogate model** represented by a dashed red line.

The x-axis represents the fraction of total injected CO$_2$ ranging from 0.92 to 1, while the y-axis represents the empirical CDF ranging from 0 to 1.
MC Simulation Result 2

Empirical CDF

Fraction of total injected CO$_2$ in caprock after 30 days of injection

- TOUGH2 model
- 4th order surrogate model
Pressure Map with TOUGH2
Pressure Map with PCE Model

Mean value of pressure contour map obtained with PCE models (Pa)

Caprock
Maximum Caprock Pressure

Probability of overpressure = 0.01
Gas Saturation Map with TOUGH2
Gas Saturation Map with PCE Model

Mean value of gas saturation contour map obtained with PCE models
Optimal Injection under Uncertainty

Optimal injection rates under various realizations of uncertain parameters

Maximize \( \mathbb{E}_{\xi_1, \xi_2 \in \Omega} \left( \sum_{i=1}^{986} S_{g_i}(z, \xi_1, \xi_2) \right) \)
subject to \( p(z, \xi_1, \xi_2) \leq p_{\text{limit}} \)

Nonconvex stochastic NLP solved with BARON using scenario-based approach
Conclusions

- Polynomial chaos expansion model was iteratively built with stepwise regression for correlated uncertain parameters
- Results obtained with PCE-based surrogate model match well the results obtained with TOUGH2
- Developed a nonconvex stochastic NLP model for finding optimal injection rate under model uncertainty
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 herein 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 herein do not necessarily state or reflect those of the United States Government or any agency thereof.