Polynomial Chaos for the Geosciences
Diane Donovan, Steve Tyson, Brodie Lawson, Marvin Tas, Bevan Thompson

Abstract
Uncertainty in input parameters is a serious challenge to traditional simulation-based methods for model prediction in geoscience and engineering. For example given limited subsurface information, significant additional computational costs are incurred when applying Monte Carlo methods to predict subsurface flows in aquifers and reservoirs and to study the effect of upsampling of parameters.

This project seeks to apply Polynomial Chaos Expansion techniques to quantify uncertainty in model inputs, providing significantly faster model predictions with reduced computational costs.

Some Questions
- Can Polynomial Chaos Expansions (PCE) take flow rates or pressure drop data and determine porosity/permeability?
- Can PCE translate cumulative probability or confidence intervals for model inputs to model outputs?
- How will this translate to better workflows for multi-phase subsurface flows?
- How will this translate into enhanced predictions for reservoir capacity and recovery?

A Problem and Method
A simple model to describe contaminant concentration, C, in groundwater given advection and dispersion, at any time t, and point x in space.

\[ \frac{\partial C}{\partial t} + u \frac{\partial C}{\partial x} = D \frac{\partial^2 C}{\partial x^2} \]

where D is the dispersion coefficient, and u is the advective velocity. The boundary and initial conditions are

\[ C(x, 0) = C_0 \quad \text{and} \quad \frac{\partial C}{\partial t} \bigg|_{t=0} = 0 \]

Solving for C(x,t) can be done analytically under idealized conditions. However, for more complex cases, numerical methods are required.

PCE captures an expansion of the solution of the PDE, which is a higher-dimensional function, in terms of orthogonal polynomials.

\[ C(x, t) \approx \sum_{i=1}^{n} c_i \phi_i(x) \phi_i(t) \]

where \( c_i \) are coefficients that can be calculated, and \( \phi_i(x) \) are the basis polynomials.

A Comparison of Methods for Uncertainty Propagation
For example, the dependence of the penetration distance of contaminant C on the hydraulic conductivity K and organic carbon partition coefficient Koc can be characterized by a parameter map (on the right) built from an ‘off the shelf solver’. The map can be used to identify parameter sets that minimize the predicted value of desirable (undesirable) quantities.

Monte Carlo
- Requires a very fine grid to capture all the variability
- Relatively simple to implement

PCE
- Can be used to approximate the response surface
- Faster than Monte Carlo

Intrusive PCE
- Can be used to approximate the response surface
- Faster than Monte Carlo

Non-intrusive PCE
- Can be used to approximate the response surface
- Faster than Monte Carlo

Application – Parameters from Data
As new field data becomes available models need to be recalibrated. Here the problem is quantifying the error between model predictions and new data. In this case, the parameter space of uncertain inputs must still be mapped, but the focus is on using PCE to predict parameter values that minimise error.

Accurate estimates of these physical properties are very important for correctly predicting total site output.

Application – CDF Generation
The distributions of the uncertain input parameters can be used to generate statistical information about the solution. For example, it might be desirable to calculate confidence intervals for contaminant spread – specifically how much contaminant will leach past a certain point. Since the PCE is built from the distributions of the uncertain input variables, the CDF’s of model outputs are easily generated.

The Monte Carlo CDF cannot be seen because the higher order PCE predictions match almost exactly. Generation of the highest order PCE required only 17.6 seconds, as opposed to 388.5 seconds for Monte Carlo.