Mathematische Methoden der Unsicherheitsquantifizierung

Oliver Ernst

Professur Numerische Mathematik

Sommersemester 2016

Mathematik!
TU Chemnitz

Contents

Contents

Contents

What is Uncertainty Quantification? (UQ)

- What is uncertainty quantification (UQ) about?

- What is uncertainty?
- How can uncertainty be described?
- How can the effects of uncertainty be treated and quantified?
- Why isn't this a class on statistics and probability theory?
- Methods for solving the resulting mathematical problems.

uncertain: *Not able to be relied on; not known or definite.*

Oxford Collegiate Dictionary

uncertain: *not exactly known or decided; not definite or fixed*

Merriam Webster Online Dictionary

What is Uncertainty Quantification? (UQ)
Auf Deutsch?

unsicher: *gefahrvoll, gefährlich, keine Sicherheit bietend*
gefährdet, bedroht
das Risiko eines Misserfolges in sich bergend, keine [ausreichenden] Garantien bietend; nicht verlässlich; zweifelhaft unzuverlässig
einer bestimmten Situation nicht gewachsen, eine bestimmte Fähigkeit nicht vollkommen, nicht souverän beherrschend nicht selbstsicher
(etwas Bestimmtes) nicht genau wissend
nicht feststehend; ungewiss

Duden Online

What is Uncertainty Quantification? (UQ)

Auf Deutsch?

unsicher: *gefahrvoll, gefährlich, keine Sicherheit bietend*
gefährdet, bedroht
das Risiko eines Misserfolges in sich bergend, keine [ausreichenden] Garantien bietend; nicht verlässlich; zweifelhaft unzuverlässig
einer bestimmten Situation nicht gewachsen, eine bestimmte Fähigkeit nicht vollkommen, nicht souverän beherrschend nicht selbstsicher
(etwas Bestimmtes) nicht genau wissend
nicht feststehend; ungewiss

Duden Online

ungewiss: *fraglich, nicht feststehend; offen*
unentschieden, noch keine Klarheit gewonnen habend
(gehoben) so [beschaffen], dass nichts Deutliches zu erkennen, wahrzunehmen ist; unbestimmbar

Duden Online

What is Uncertainty Quantification? (UQ)

A poetic description

There are known knowns;
there are things we know we know.

We also know there are known unknowns;
that is to say, we know there are some things we do not know.

But there are also unknown unknowns – the ones we don't know we don't know.

Donald Rumsfeld, U.S. Secretary of Defense
DoD News Briefing; Feb. 12, 2002

What is Uncertainty Quantification? (UQ)

Historical orientation

- Life is full of uncertainty.
  Pre-modern coping mechanisms: religion, divinatory practices, fatalism.

What is Uncertainty Quantification? (UQ)

Historical orientation

- Life is full of uncertainty.
  Pre-modern coping mechanisms: religion, divinatory practices, fatalism.
- Late 17th, early 18th century: rise of insurance business (Lloyd's Coffee House, London) shipping, fire, life, etc.

What is Uncertainty Quantification? (UQ)

Historical orientation

- Life is full of uncertainty.
  Pre-modern coping mechanisms: religion, divinatory practices, fatalism.
- Late 17th, early 18th century: rise of insurance business (Lloyd's Coffee House, London) shipping, fire, life, etc.
- Simultaneously, development of probability theory, Bernoullis, De Moivre, Laplace, ...

What is Uncertainty Quantification? (UQ)

Historical orientation

- Life is full of uncertainty.
  Pre-modern coping mechanisms: religion, divinatory practices, fatalism.
- Late 17th, early 18th century: rise of insurance business (Lloyd's Coffee House, London) shipping, fire, life, etc.
- Simultaneously, development of probability theory, Bernoullis, De Moivre, Laplace, ...
- Predictive mathematical modelling, a success story since Newton and Galileo. Quantities of interest as functions of known quantities, solutions of algebraic, differential integral equations.

What is Uncertainty Quantification? (UQ)

Historical orientation

- Life is full of uncertainty.
  Pre-modern coping mechanisms: religion, divinatory practices, fatalism.
- Late 17th, early 18th century: rise of insurance business (Lloyd's Coffee House, London) shipping, fire, life, etc.
- Simultaneously, development of probability theory, Bernoullis, De Moivre, Laplace, . . .
- Predictive mathematical modelling, a success story since Newton and Galileo. Quantities of interest as functions of known quantities, solutions of algebraic, differential integral equations.
- Laplace: nature is deterministic, probability a description of our ignorance. Quantum theory (1930s) showed nature exhibits intrinsic randomness.

What is Uncertainty Quantification? (UQ)

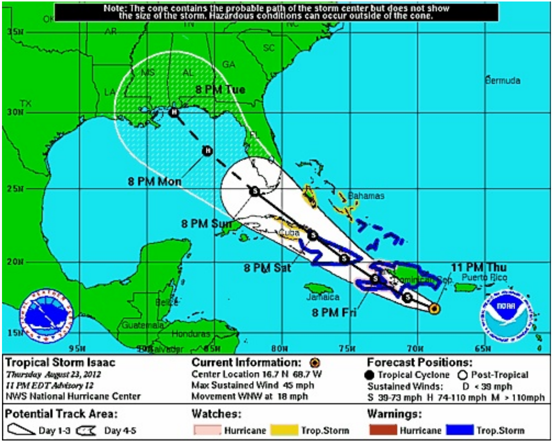
Historical orientation

- Life is full of uncertainty.
  Pre-modern coping mechanisms: religion, divinatory practices, fatalism.
- Late 17th, early 18th century: rise of insurance business (Lloyd's Coffee House, London) shipping, fire, life, etc.
- Simultaneously, development of probability theory, Bernoullis, De Moivre, Laplace, ...
- Predictive mathematical modelling, a success story since Newton and Galileo. Quantities of interest as functions of known quantities, solutions of algebraic, differential integral equations.
- Laplace: nature is deterministic, probability a description of our ignorance. Quantum theory (1930s) showed nature exhibits intrinsic randomness.
- Today, computational science can simulate/predict many phenomena, often main limitation is now data uncertainty, rather than rounding error, discretization error, simulation cost.

- Life is full of uncertainty.
  Pre-modern coping mechanisms: religion, divinatory practices, fatalism.
- Late 17th, early 18th century: rise of insurance business (Lloyd's Coffee House, London) shipping, fire, life, etc.
- Simultaneously, development of probability theory, Bernoullis, De Moivre, Laplace, ...
- Predictive mathematical modelling, a success story since Newton and Galileo. Quantities of interest as functions of known quantities, solutions of algebraic, differential integral equations.
- Laplace: nature is deterministic, probability a description of our ignorance. Quantum theory (1930s) showed nature exhibits intrinsic randomness.
- Today, computational science can simulate/predict many phenomena, often main limitation is now data uncertainty, rather than rounding error, discretization error, simulation cost.
- We need mathematical, numerical and algorithmic techniques for quantifying uncertainty in scientific and engineering computations.

(Increasingly?) many aspects of modern life involve uncertainty.

- **Social systems:** military, finance, insurance industry, elections
- **Environmental systems:** weather, climate, seismics, subsurface geophysics
- **Engineering systems:** automobiles, aircraft, bridges, structures
- **Biological systems:** health and medicine, pharmaceuticals, gene expression, cancer research
- **Physical systems:** quantum physics, radioactive decay

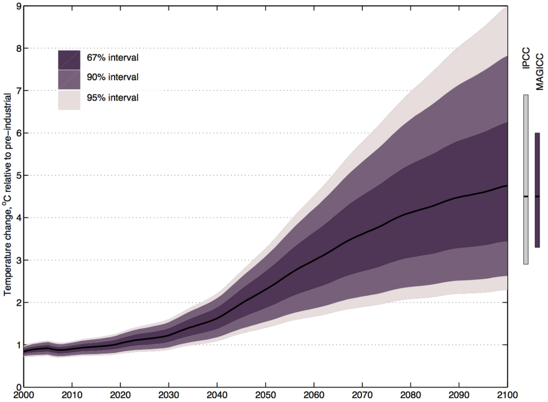


Source: National Hurricane Center, USA

Predicted storm path with uncertainty cones.

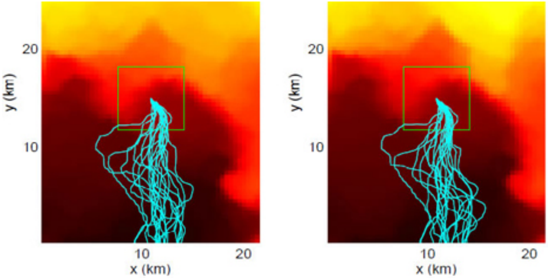
What is Uncertainty Quantification? (UQ)

Uncertainty in Modern Life



Source: Brodman & Karoly, 2013

Global-mean temperature change for a business-as-usual emission scenario, relative to pre-industrial. Black line: median, shaded regions 67% (dark), 90% (medium) and 95% (light) confidence intervals.



Source: K. A. Cliffe, 2012

Sample paths of groundwater-borne contaminant particles emanating from an underground radioactive waste disposal site.

Radioactive decay

- Radium-226: half-life of 1602 years
- Decays into Radon gas (Radon-222) by emitting alpha particles.
- Over a period of 1602 years, half the radium atoms in a given sample will decay.
- But we cannot say which half!

This kind of uncertainty seems to be 'built in' to the physical world.

What is Uncertainty Quantification? (UQ)

Examples

Rolling a die (or several dice)

- Cube, 6 faces, numbered 1–6.
- One or more thrown onto a table.
- For “fair dice”, expect to see the numbers 1–6 appear equally often, provided the dice are thrown sufficiently many times.

How does this differ from radioactive decay?

Is this uncertainty also ’built-in’ to the physical world, or is it just that we don’t know how to calculate what will happen when the dice are thrown?

Screening/testing for disease

- Incidence of disease among general population: 0.01 %
- Test has true positive rate (sensitivity) of 99.9 %.
- Same test has true negative rate (specificity) of 99.99 %.
- What is the chance that someone who tests positive actually has the disease?

Screening/testing for disease

- Incidence of disease among general population: 0.01 %
- Test has true positive rate (sensitivity) of 99.9 %.
- Same test has true negative rate (specificity) of 99.99 %.
- What is the chance that someone who tests positive actually has the disease?

Answer (relative probabilities, conditional probabilities, Bayes' formula)

$$\begin{aligned}\mathbf{P}(\text{desease}|\text{pos}) &= \frac{\mathbf{P}(\text{pos}|\text{disease}) \cdot \mathbf{P}(\text{disease})}{\mathbf{P}(\text{pos}|\text{disease}) \cdot \mathbf{P}(\text{disease}) + \mathbf{P}(\text{pos}|\text{no disease}) \cdot \mathbf{P}(\text{no disease})} \\ &= \frac{0.999 \cdot 0.0001}{0.999 \cdot 0.0001 + (1 - 0.9999) \cdot (1 - 0.0001)} \\ &\approx 0.4998\end{aligned}$$

Alternative phrasing (of same answer using natural frequencies)

- Think of random sample 10,000 people.
- Of these, on average 1 will have the disease, 9,999 will not.
- The person who has the disease will almost certainly test positive.
- of the 9,999 healthy people, on average one will test (falsely) positive.
- Thus, roughly one out of every two positive patients actually has the disease.

Alternative phrasing (of same answer using natural frequencies)

- Think of random sample 10,000 people.
- Of these, on average 1 will have the disease, 9,999 will not.
- The person who has the disease will almost certainly test positive.
- of the 9,999 healthy people, on average one will test (falsely) positive.
- Thus, roughly one out of every two positive patients actually has the disease.

In [Gigerenzer, 1996]: Medical practitioners were given the following information regarding mammography screenings for breast cancer:

incidence: 1 %; sensitivity: 80 %; specificity: 90 %.

When asked to quantify the probability of the patient actually having breast cancer given a positive screening result (7.5%), 95 out of 100 physicians estimated this probability to lie above 75%.
See also [Gigerenzer et al., 1998] for similar observations in AIDS counseling.

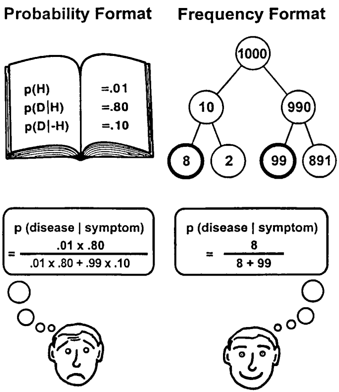


FIGURE 1. Bayesian computations are simpler when information is represented in a frequency format (right) than when it is represented in a probability format (left). p(H) = prior probability of hypothesis. H (breast cancer), p(D|H) = probability of data D (positive test) given H, and p(D|−H) = probability of D given −H (no breast cancer).

Sometimes the description of uncertainty is crucial for its transparent communication.

Source: Gigerenzer, 1996

Modeling biological systems

- From one view, biology is just very complicated physics and chemistry.
- But even the simplest biological systems are far too complicated to be understood from basic principles at the moment.
- Models are constructed that attempt to capture the essential features of what is happening, but often there are competing models and they may all fail in some way or other to predict the observed phenomena.
- In short, we don't really know what the model is!

How does this situation differ from the previous two?

Climate change

> *The weight of evidence makes it clear that climate change is a real and present danger. The Exeter conference was told that whatever policies are adopted from this point on, the Earth's temperature will rise by 0.6°F within the next 30 years. Yet those who think climate change just means Indian summers in Manchester should be told that the chances of the Gulf stream - the Atlantic thermohaline circulation that keeps Britain warm - shutting down are now thought to be greater than 50%*
>
> *The Guardian, 2005*

> *Most of the observed increase in globally-averaged temperatures since the mid-20th century is very likely due to the observed increase in anthropogenic GHG concentrations. It is likely there has been significant anthropogenic warming over the past 50 years averaged over each continent (except Antarctica).*
>
> *IPCC Fourth Assessment*
> *Summary for Policymakers.*

What do these statements mean?

Unknown unknowns

- Obviously can't give a current example.
- A good example ist the state of Physics at the end of the 19th century.

  There is nothing new to be discovered in physics now. All that remains is more and more precise measurement.

  Lord Kelvin, 1900

- Quantum mechanics and relativity theory were unknown unknowns.

It is easy to underestimate uncertainty.

What is Uncertainty Quantification? (UQ)

Political Implications

Questions:[1]

1. How do we account for all the uncertainties in the complex models and analyses that inform decision makers?
2. How can those uncertainties be communicated simply but quantitatively to decision makers?
3. How should decision makers use those uncertainties when combining scientific evidence with more socio-economic considerations?
4. How can decisions be communicated so that the proper acknowledgment of uncertainty is transpartent?

[1] posed on entry at the 2006 EPSRC Ideas Factory on the topic *Scientific Uncertainty and Decision Making for Regulatory and Risk Assessment Purposes*.

Contents

> *The principle of indifference asserts that if there is no known reason for predicating of our subject one rather than another of several alternatives, then relatively to such knowledge the assertions of each of these alternatives have an equal probability.*
>
> *J. M. Keynes, 1921*

- Rule for assigning epistemic probabilities: in the absence of further information, use the uniform distribution.
- Taken as intuitively obvious by [J. Bernoulli, Ars Conjectandi, 1713].
- Later used by [Laplace, 1774] to define classical probability.
- Originally known as *"Principle of insufficient reason"*, (cf. Leibniz' "principle of sufficient reason"[2]), current term due to [Keynes, 1921].
- Leads to contradictions (numerous paradoxes in literature).

[2] "For every fact F, there must be an explanation why F is the case."

Given probability space $(\Omega, \mathfrak{A}, \mathbf{P})$, $A, B \in \mathfrak{A}$, $\mathbf{P}(B) > 0$, then the conditional probability of A given B is defined by

$$\mathbf{P}(A|B) := \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)}.$$

Given probability space $(\Omega, \mathfrak{A}, \mathbf{P})$, $A, B \in \mathfrak{A}$, $\mathbf{P}(B) > 0$, then the conditional probability of A given B is defined by

$$\mathbf{P}(A|B) := \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)}.$$

Solving for $\mathbf{P}(A \cap B)$, exchanging roles of A and B, assuming $\mathbf{P}(A) > 0$, gives

$$\mathbf{P}(A|B) = \frac{\mathbf{P}(B|A)\,\mathbf{P}(A)}{\mathbf{P}(B)}$$ Bayes' rule [Bayes, 1763]

Given probability space $(\Omega, \mathfrak{A}, \mathbf{P})$, $A, B \in \mathfrak{A}$, $\mathbf{P}(B) > 0$, then the conditional probability of A given B is defined by

$$\mathbf{P}(A|B) := \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)}.$$

Solving for $\mathbf{P}(A \cap B)$, exchanging roles of A and B, assuming $\mathbf{P}(A) > 0$, gives

$$\mathbf{P}(A|B) = \frac{\mathbf{P}(B|A)\,\mathbf{P}(A)}{\mathbf{P}(B)}$$

Bayes' rule [Bayes, 1763]

- A: unobservable state of nature, with prior probability $\mathbf{P}(A)$ of occurring;
- B: observable event, probability $\mathbf{P}(B)$ known as evidence;
- $\mathbf{P}(B|A)$: probability that A causes B to occur (likelihood);
- $\mathbf{P}(A|B)$: posterior probability of A knowing that B has occurred.
- Terms: inverse probability, Bayesian inference.

Given partition $\{A_j\}_{j\in\mathbb{N}}$ of Ω into exhaustive and exclusive disjoint events, de Morgan's rule and countable additivity give, assuming all $\mathbf{P}(A_j) > 0$,

$$\mathbf{P}(B) = \sum_{j\in\mathbb{N}} \mathbf{P}(B|A_j)\,\mathbf{P}(A_j) \qquad \text{(law of total probability)},$$

leading to another variant of Bayes' rule:

$$\mathbf{P}(A_k|B) = \frac{\mathbf{P}(B|A_k)\,\mathbf{P}(A_k)}{\sum_{j\in\mathbb{N}} \mathbf{P}(B|A_j)\,\mathbf{P}(A_j)},$$

giving posterior probability of each A_k after observing B.

Given real-valued random variables X, Y with probability density functions (pdfs)

- $f_X(x), f_Y(y)$: density of X, Y at value x, y,
- $f_{X|Y}(x|y)$: density of $(X|Y)$ at x having observed $Y = y$,
- $f_{Y|X}(y|x)$: analogously.

Then Bayes' theorem states that

$$f_{X|Y}(x|y) = \frac{f_{Y|X}(y|x)\, f_X(x)}{f_Y(y)} = \frac{f_{Y|X}(y|x)\, f_X(x)}{\int f_{Y|X}(y|x)\, f_X(x)\, \mathrm{d}x}.$$

- $f_{Y|X}(y|x)$ is now called the likelihood function.
- $\int f_{Y|X}(y|x)\, f_X(x)\, \mathrm{d}x$ is calles the normalizing factor or marginal.
- Short form:

$$f_{X|Y} \propto f_{Y|X}\, f_X.$$

Problem:

- Given $A \in \mathfrak{A}$, suppose $p := \mathbf{P}(A) \in [0,1]$ is unknown.
- Assume A has occurred in k out of n independent and identical trials.
- For $0 \leqslant p_1 < p_2 \leqslant 1$, what is the probability that $p \in (p_1, p_2)$?

Problem:

- Given $A \in \mathfrak{A}$, suppose $p := \mathbf{P}(A) \in [0,1]$ is unknown.
- Assume A has occurred in k out of n independent and identical trials.
- For $0 \leqslant p_1 < p_2 \leqslant 1$, what is the probability that $p \in (p_1, p_2)$?

Solution:

$$\mathbf{P}(p_1 < p < p_2) = \frac{(n+1)!}{k!(n-k)!} \int_{p_1}^{p_2} p^k (1-p)^{n-k} \, \mathrm{d}p.$$

- Classical probability (Bernoulli, Laplace): given probability $p = \mathbf{P}(A)$, how many independent trials n are necessary to be "morally certain" that A occurs $k = pn$ times?
- Bayes: Given occurrence rates and notion of prior probability for A, what is $\mathbf{P}(A)$?

Problem: A box contains a large number N of black and white balls. We draw n balls with replacement, of which k turn out to be black, $n-k$ white. What is the conditional probability that the next draw will yield a black ball?

Problem: A box contains a large number N of black and white balls. We draw n balls with replacement, of which k turn out to be black, $n-k$ white. What is the conditional probability that the next draw will yield a black ball?

Solution:

$$\mathbf{P}(\text{next draw black}|k \text{ of } n \text{ previous draws black}) = \frac{k+1}{n+2}.$$

Problem: A box contains a large number N of black and white balls. We draw n balls with replacement, of which k turn out to be black, $n-k$ white. What is the conditional probability that the next draw will yield a black ball?

Solution:

$$\mathbf{P}(\text{next draw black} | k \text{ of } n \text{ previous draws black}) = \frac{k+1}{n+2}.$$

Problem: [Laplace, 1814] What is the probability that the sun will rise tomorrow, given that it has risen on each day of the past 5000 years?

Problem: A box contains a large number N of black and white balls. We draw n balls with replacement, of which k turn out to be black, $n-k$ white. What is the conditional probability that the next draw will yield a black ball?

Solution:

$$\mathbf{P}(\text{next draw black}|k \text{ of } n \text{ previous draws black}) = \frac{k+1}{n+2}.$$

Problem: [Laplace, 1814] What is the probability that the sun will rise tomorrow, given that it has risen on each day of the past 5000 years?

Solution: $n = 5000 \cdot 365.2426 = 1,826,213$, $k = n$,

$$\mathbf{P}(\text{sunrise tomorrow}) = \frac{1,826,214}{1,826,215} \approx 0.9999995.$$

> *But this number is incomparably greater for him who, recognizing in the totality of phenomena the principal regulator of days and seasons, sees that nothing at the present moment can arrest the course of it.*
>
> *P.-S. Laplace, Essai Philosophique sur les Probabilités,1814*

[Taleb & Blythe, 2011], [B. Russell, 1912], [Gigerenzer, 2014]

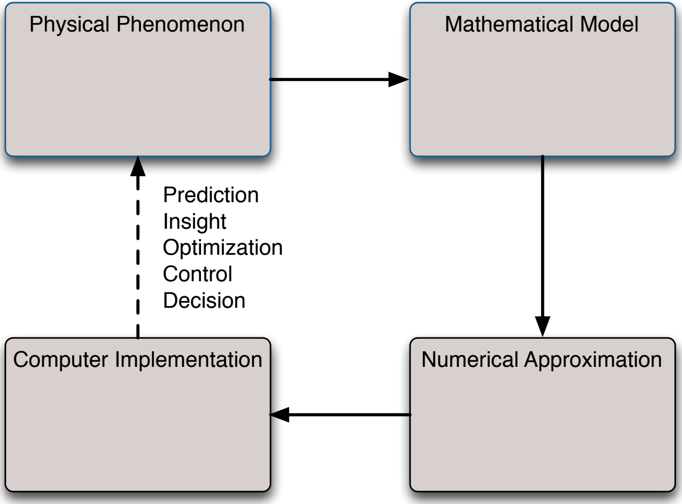
- A turkey is fed by the farmer every day for many months.
- The turkey applies the rule of succession and feels more confident with every passing day.
- ... until Thanksgiving.
- The turkey had too much confidence in his model of uncertainty; he was missing important information (unknown unknowns).
- Fundamental question in epistemology (the theory of knowledge), known as the Problem of Induction [D. Hume, 1748]
- [K. Popper, 1959] postulated that induction is not possible, that scientific theories can only be falsified.
- The turkey illusion is the belief that a risk can be calculated when it cannot.
- [F. Knight, 1921]: distinction between known risk ("risk") and unknown risk ("uncertainty"). Uncertainty in this sense requires more tools than probability.

- Prakash Gorroochurn: Classic Problems of Probability. Wiley 2012 (Chapter 14)
- Gerd Gigerenzer: Risk Savvy – How to Make Good Decisions. Penguin 2014.
- Nassim Taleb: The Black Swan – The Impact of the Highly Improbable. Penguin 2007.
- Claudia Klüppelberg et al. (eds.): Risk – A Multidisciplinary Introduction. Springer 2014.
- Joseph Y. Halpern: Reasoning About Uncertainty. MIT Press 2003.

Contents

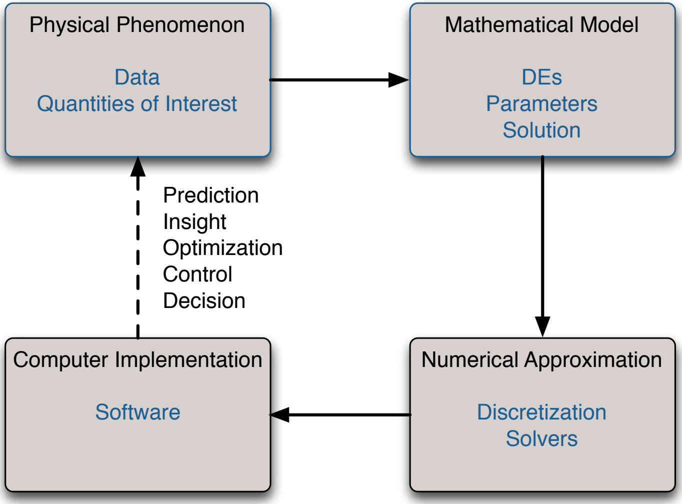
UQ and Scientific Computing

UQ and the scientific computing paradigm



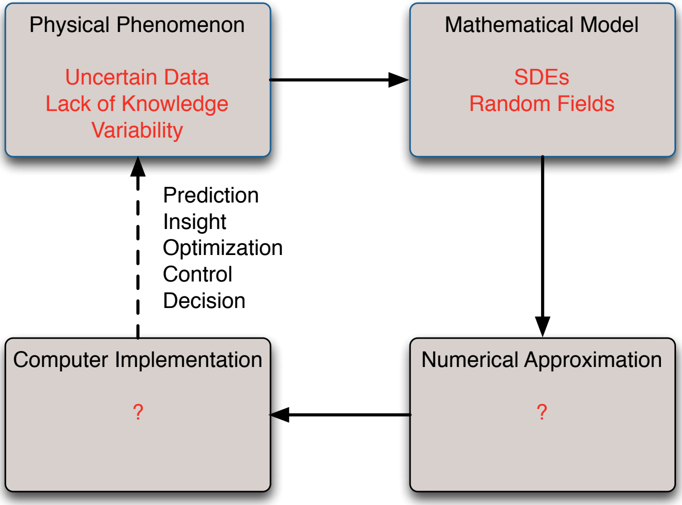
UQ and Scientific Computing

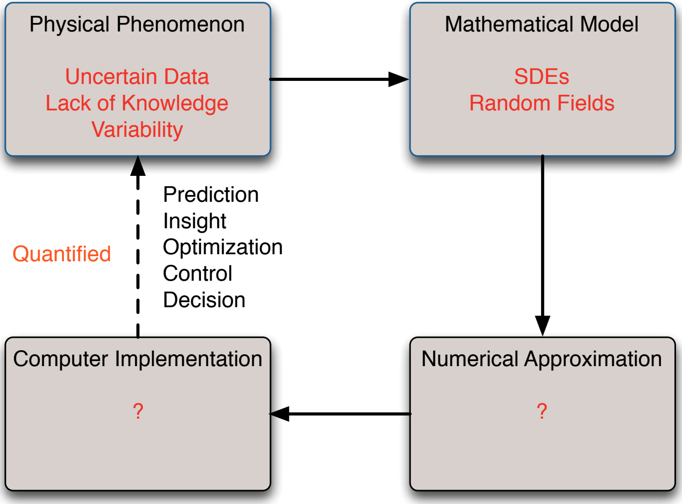
UQ and the scientific computing paradigm



UQ and Scientific Computing

UQ and the scientific computing paradigm




Physical Phenomenon
Uncertain Data
Lack of Knowledge
Variability
Mathematical Model
SDEs
Random Fields
Prediction
Insight
Optimization
Control
Decision
Quantified
Computer Implementation
?
Numerical Approximation
?

What confidence can be assigned to a computer prediction of complex phenomena?

Validation: The determination of whether a mathematical model adequately represents the pysical or engineering phenomenon under study.
"Are we solving the right problem?"

Is this even possible? (cf. Carl Popper)

Verification: The determination of whether an algorithm and/or computer code correctly implements a given mathematical model.
"Are we solving the problem correctly?"

- code verification (software engineering)
- solution verification (a posteriori error estimation)

Aleatoric Uncertainty: (variability) Uncertainty due to true intrinsic variability; cannot be reduced by additional experimentation, improvement of measuring devices etc.

Examples:

- rolling a die
- wind stress on a structure
- production variations

Epistemic Uncertainty: Uncertainty due to lack of knowledge/incomplete information.
Examples:

- turbulence modeling assumptions
- surrogate chemical kinetics
- the probability distribution a random quantity follows

Note: This distinction is not always meaningful or possible.

Contents

The most popular model problem in the UQ community has become the second-order elliptic PDE with an uncertain coefficient function:

$$-\nabla\cdot(a\nabla u) = f \quad + \text{ domain } D \subset \mathbb{R}^d \quad + \text{ BC.}$$

Rather than the solution u (whatever that may be), typical UQ problems center on some functional Q of the solution, e.g. its value at a point in the computational domain, average over a subdomain, flux across a boundary etc. Such a functional is known as a quantity of interest (QoI).

Examples:

$$Q(u) = u(\boldsymbol{x}_0), \qquad Q(u) = \frac{1}{|D_0|}\int_{D_0} u(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$$

Introduce input set V containing all possible inputs a and associated output set $W = \{Q(u(a)) : a \in V\}$ as well as mapping $G : V \to W$ of inputs to outputs.

In what way might uncertainty in the coefficient a be addressed?

Introduce an ϵ-ball around a given function a_0 (in a suitable norm).

Examples:

$$V_\infty := \{a \in L^\infty(D) : \|a - a_0\|_{L^\infty(D)} \leqslant \epsilon\},$$
$$V_1 := \{a \in W^{1,\infty}(D) : \|a - a_0\|_{W^{1,\infty}(D)} \leqslant \epsilon\},$$
$$V_{\text{const}} := \{a : a \text{ is constant in } D, |a - a_0| \leqslant \epsilon\}.$$

Worst case analysis: determine uncertainty interval

$$I = \left[\inf_{a\in V} Q(u(a)), \sup_{a\in V} Q(u(a))\right].$$

The uncertainty range of Q is then the length of I.

This is a generalization of interval analysis.

Worst-case analysis is a common approach in robust optimization.

Idea: Some values (functions) $a \in V$ are more likely than others.

Purely probabilistic approach:

- Introduce probability measure on V.
- (Measurable) mapping $G : V \to W$ induces probability measure on W. ("uncertainty propagation").
- Big issue: choice of distribution, too much subjective information?
- Ststistical tradition of 'eliciting probability distributions from expert opinion'.
- Some classical guidelines: Laplace's principle of insufficient reason, maximum entropy (information theory), etc.
- Choosing distribution based on data is point of departure for Bayesian inverse problem.

Generalizes probabilistic model (also called Dempster-Shafer theory)

- Finite or countable family $\mathfrak{F}$ of events.
- Set function $m : \mathfrak{F} \to [0, 1]$ (belief function, support function) giving likelihood information for each event, satisfies

$$\sum_{A \in \mathfrak{F}} m(A) = 1, \qquad m(\varnothing) = 0,$$

  but, unlike probability measures, need not satisfy $A \subset B \Rightarrow m(A) \leqslant m(B)$.
- Belief and plausability functions for admissible events C

$$\operatorname{bel}(C) = \sum_{A \in \mathfrak{F}, A \subset C} m(A), \qquad \operatorname{pl}(C) = \sum_{A \in \mathfrak{F}, A \cap C \neq \varnothing} m(A).$$

  provide lower and upper bounds, respectively, on likelihood of event C.
- Likelihood function dependent on expert opinion.

Deterministic approach introduced by [Zadeh, 1965].

- Generalizes "$\in$" relation of classical set theory: for $C \subset V$, in place of exhaustive alternatives $x \in C$ and $x \notin C$, introduces membership function

$$\mu_C : V \to [0,1]$$

expressing truth degree of statement $a \in C$.

- Important tool: α-cut of set C defined by

$$C^\alpha := \{a \in V : \mu_C(a) \geqslant \alpha\}$$

giving set characterization of uncertainty.

- Mapping G then again propagates fuzziness of input set V to output set W.

Contents

A Case Study: Radioactive Waste Disposal

- An area where UQ has played a central role in the past 25 years is the assessment of strategies and sites for the long-term storage of radioactive waste.
- Uncertainties arise from technological complexity as well as the long time scales to be considered.
- Many leading industrial countries (USA, UK, Germany) have scrapped previous plans for national long-term disposal sites and are re-evalutating their strategies.
- We consider a basic UQ problem which occurs in site assessment studies.

- Radioactive waste is produced in large part by power plants, in which the heat from controlled nuclear fission is used to produce electric power. (Other sources: medical, weapon production, non-nuclear industries)
- Exposure to high radiation levels seriously harmful to humans and animals; long-term exposure to low-level radiation can cause cancer and other long-term health problems.
- Classification
  - high-level waste (HLW): highly radioactive, produces heat, small quantities.
  - intermediate-level waste (ILW): still very radioactive, does not produce heat.
  - low-level waste (LLW): low radiactivity; packaging material, protective clothing, soil, concrete etc. which has been exposed to radioactivity.
- Quantities in storage (source: IAEA database, http://newmdb.iaea.org)
  - Germany: 120,000 m^3 (2007)
  - France: 90,000 m^3 (2007)
  - UK: 350,000 m^3 (2007)
  - USA: 540,000 m^3 (2008)

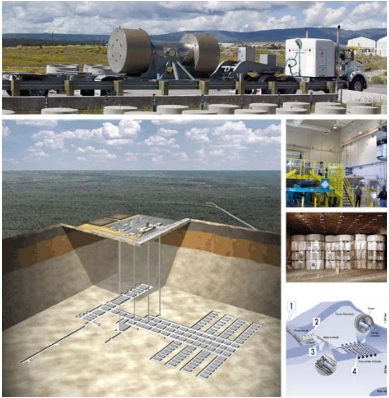
Since this problem has received serious consideration ($\approx$ 1970s), several options have been discussed

- Surface storage: current universal solution, not long-term, risky.
- Disposal at sea: banned by international treaty (London Convention)
- Disposal in space: too dangerous, prohibitive cost (but permanent solution).
- Transmutation: not yet proven technology, would mitigate but not solve the problem.
- Deep geological disposal
  - Favored by nearly all countries with a radioactive waste disposal program.
  - Storage in containers in tunnels, several hundred meters deep, in stable geological formations.
  - Issue: retrievable or not?
  - No human intervention required after final closure of repository.
  - Several barriers: chemical, physical, geological.
  - Substantial engineering challenge (containment must be assured for at least 10,000 years).
  - Main escape route for radionuclides: groundwater pathway.

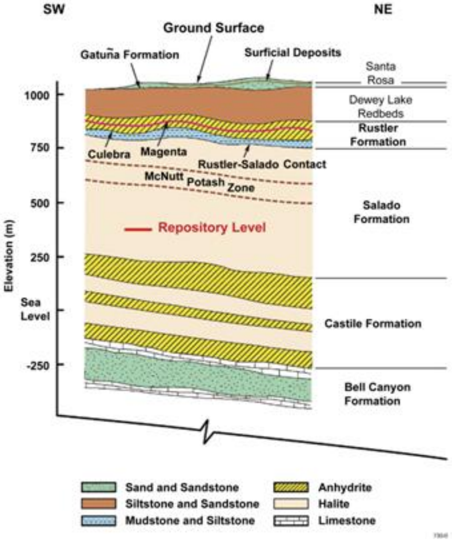
A Case Study: Radioactive Waste Disposal

WIPP

- US DOE repository for radioactive waste situated near Carlsbad, NM.
- Fully operational since 1999.
- Extensive site characterization and performance assessment since 1976, also in course of compliance certification and recertification by US EPA (every 5 years).
- Large amount of publicly available data.
- http://www.wipp.energy.gov



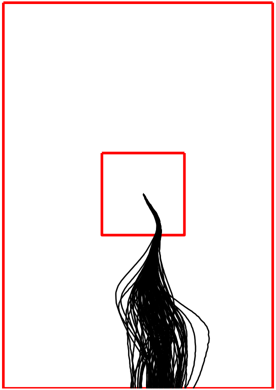
- Repository located at depth of 655 m within bedded evaporites, primarily halite (salt).
- The most transmissive rock in the region is the Culebra Dolomite.
- In the event of an accidental breach, Culebra would be the principal pathway for transport of radionuclides away from the repository.



A Case Study: Radioactive Waste Disposal

WIPP UQ scenario

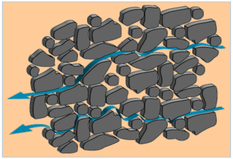
- One scenario at WIPP is a release of radionuclides by means of a borehole drilled into the repository.
- Radionuclides are released into the Culebra Dolomite and then transported by groundwater.
- Travel time from release point in the repository to the boundary of the region is an important quantity.
- Flow is two-dimensional to a good approximation.



- The simplest mathematical model for flow through a porous medium (as in groundwater through an aquifer) is given by Darcy's law

$$\boldsymbol{q} = \frac{-\boldsymbol{k}}{\mu}\nabla p,$$

in which $\boldsymbol{q}$ denotes the volumetric flux or Darcy velocity (discharge per unit area in [m/s]), $\boldsymbol{k}$ the permeability tensor,a material parameter describing how easily water flows through the given medium, μ the dynamic viscosity of the fluid and p the hydraulic head of the fluid.



- The hydraulic conductivity tensor is defined as $K := \boldsymbol{k}\rho g/\mu$, where g is the acceleration due to gravity and ρ the fluid density.
- The actual pore velocity with which the fluid particles move through the pores is obtained as $\boldsymbol{u} = \boldsymbol{q}/\phi$, where $\phi \in [0, 1]$ denotes the porosity of the medium.

Stationary Darcy flow	$\boldsymbol{q} = -K\nabla p$	$\boldsymbol{q}$: Darcy flux
		K : hydraulic conductivity
		p : hydraulic head
mass conservation	$\nabla\cdot \boldsymbol{u} = 0$	$\boldsymbol{u}$: pore velocity
	$\boldsymbol{q} = \phi\boldsymbol{u}$	ϕ : porosity
transmissivity	$T = Kb$	b : aquifer thickness
particle transport	$\dot{\boldsymbol{x}}(t) = -\dfrac{T(\boldsymbol{x})}{b\phi}\nabla p(\boldsymbol{x})$	$\boldsymbol{x}$: particle position
	$\boldsymbol{x}(0) = \boldsymbol{x}_0$	$\boldsymbol{x}_0$: release location

Quantity of interest: particle travel time to reach WIPP boundary (actually, its $\log_{10}$).

Primal form of Darcy equations:

$$\nabla\cdot[T(\boldsymbol{x})\nabla p(\boldsymbol{x})] = 0, \quad \boldsymbol{x} \in D, \qquad p = p_0 \text{ along } \partial D.$$

Model T as a random field (RF) $T = T(\boldsymbol{x}, \omega)$, $\omega \in \Omega$, with respect to underlying probability space $(\Omega, \mathfrak{A}, \mathbf{P})$.

Modeling Assumptions: (standard in hydrogeology)

- T has finite mean and covariance

$$\begin{aligned} \overline{T}(\boldsymbol{x}) &= \mathbf{E}\left[T(\boldsymbol{x}, \cdot)\right], & \boldsymbol{x} \in D, \\ \mathbf{Cov}_T(\boldsymbol{x}, \boldsymbol{y}) &= \mathbf{E}\left[\left(T(\boldsymbol{x}, \cdot) - \overline{T}(\boldsymbol{x})\right)\left(T(\boldsymbol{y}, \cdot) - \overline{T}(\boldsymbol{y})\right)\right], & \boldsymbol{x}, \boldsymbol{y} \in D. \end{aligned}$$

- T is lognormal, i.e., $Z(\boldsymbol{x}, \omega) := \log T(\boldsymbol{x}, \omega)$ is a Gaussian RF.
- $\mathbf{Cov}_Z$ is stationary and isotropic, i.e., $\mathbf{Cov}_Z(\boldsymbol{x}, \boldsymbol{y}) = c(\|\boldsymbol{x} - \boldsymbol{y}\|_2)$, and of Matérn type.

$$c(\boldsymbol{x}, \boldsymbol{y}) = c_{\boldsymbol{\theta}}(r) = \frac{\sigma^2}{2^{\nu-1}\,\Gamma(\nu)} \left(\frac{2\sqrt{\nu}\,r}{\rho}\right)^{\nu} K_\nu\left(\frac{2\sqrt{\nu}\,r}{\rho}\right), \quad r = \|\boldsymbol{x} - \boldsymbol{y}\|_2$$

K_ν : modified Bessel function of order ν

Parameters $\boldsymbol{\theta} = (\sigma^2, \rho, \nu)$

σ^2 : variance

ρ : correlation length

ν : smoothness parameter

$$c(\boldsymbol{x}, \boldsymbol{y}) = c_{\boldsymbol{\theta}}(r) = \frac{\sigma^2}{2^{\nu-1}\,\Gamma(\nu)} \left(\frac{2\sqrt{\nu}\,r}{\rho}\right)^{\nu} K_\nu\left(\frac{2\sqrt{\nu}\,r}{\rho}\right), \quad r = \|\boldsymbol{x} - \boldsymbol{y}\|_2$$

K_ν : modified Bessel function of order ν

Parameters $\boldsymbol{\theta} = (\sigma^2, \rho, \nu)$

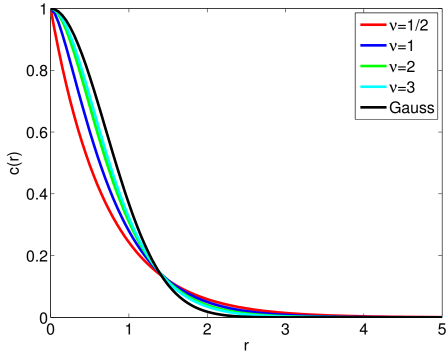
σ^2 : variance

ρ : correlation length

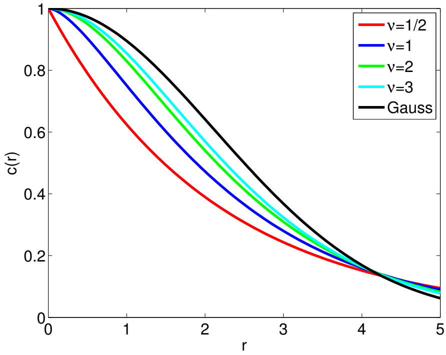
ν : smoothness parameter

Special cases:

$\nu = \frac{1}{2}$:	$c(r) = \sigma^2 \exp(-\sqrt{2}r/\rho)$	exponential covariance
$\nu = 1$:	$c(r) = \sigma^2 \left(\frac{2r}{\rho}\right) K_1\left(\frac{2r}{\rho}\right)$	Bessel covariance
$\nu \to \infty$:	$c(r) = \sigma^2 \exp(-r^2/\rho^2)$	Gaussian covariance



$\rho = 1$



$\rho = 3$

Smoothness: Realizations $Z(\cdot, \omega)$ are k times differentiable $\Leftrightarrow \nu > k$.

Covariance function of RF $Z \in L^2_{\mathbf{P}}(\Omega; L^\infty(D))$

$$c(\boldsymbol{x}, \boldsymbol{y}) = \mathbf{Cov}_Z(\boldsymbol{x}, \boldsymbol{y}) := \mathbf{E}\left[\Big(Z(\boldsymbol{x}, \cdot) - \overline{Z}(\boldsymbol{x})\Big)\Big(Z(\boldsymbol{y}, \cdot) - \overline{Z}(\boldsymbol{y})\Big)\right], \ \boldsymbol{x}, \boldsymbol{y} \in D,$$

is symmetric in $\boldsymbol{x}, \boldsymbol{y}$, positive semidefinite, and continuous on $D \times D$ if continuous along 'diagonal' $\{(\boldsymbol{x}, \boldsymbol{x}) : \boldsymbol{x} \in D\}$.

The covariance operator

$$C = C_Z : L^2(D) \to L^2(D), \qquad (Cu)(\boldsymbol{x}) = \int_D u(\boldsymbol{y}) c(\boldsymbol{x}, \boldsymbol{y}) \, \mathrm{d}\boldsymbol{y}$$

is therefore selfadjoint, compact, nonnegative. Its eigenvalues $\{\lambda_m\}_{m \in \mathbb{N}}$ form a nonincreasing sequence accumulating at most at 0.

Denoting eigenfunctions by $\{Z_m\}_{m\in\mathbb{N}}$ there exists sequence of RV

$$\{\xi_m\}_{m\in\mathbb{N}} \subset L^2_{\mathbf{P}}(\Omega), \quad \mathbf{E}[\xi_m] = 0, \quad \mathbf{E}[\xi_k \xi_m] = \delta_{k,m},$$

such that the expansion

$$Z(x,\omega) = \overline{Z}(\boldsymbol{x}) + \sum_{m=1}^{\infty} \sqrt{\lambda_m}\, Z_m(\boldsymbol{x})\, \xi_m(\omega)$$

converges in $L^2_{\mathbf{P}}(\Omega; L^\infty(D))$.

[Karhunen, 1947], [Loève, 1948]

For normalized eigenfunctions $Z_m(\boldsymbol{x})$,

$$\mathbf{Var}_Z(\boldsymbol{x}) := c(\boldsymbol{x}, \boldsymbol{x}) = \sum_{m=1}^{\infty} \lambda_m Z_m(\boldsymbol{x})^2,$$

Total variance:

$$\int_D \mathbf{Var}_Z(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} = \sum_{m=1}^{\infty} \lambda_m \underbrace{(Z_m, Z_m)_D}_{=1} = \text{trace } C.$$

For constant variance (e.g., stationary RF),

$$\mathbf{Var}_Z \equiv \sigma^2 > 0, \qquad \sum_m \lambda_m = |D|\, \sigma^2.$$

Interpretation: M first covariance eigenmodes form best rank-M approximation to C in sense of retaining maximal amount of variance.

Truncate KL expansion after M leading terms:

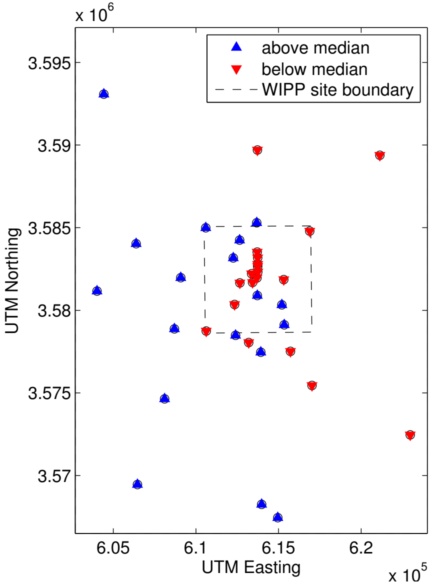
$$Z^{(M)}(\boldsymbol{x},\omega) = \overline{Z}(x) + \sum_{m=1}^{M} \sqrt{\lambda_m}\, Z_m(\boldsymbol{x})\, \xi_m(\omega).$$

Truncation error

$$\mathbf{E}\left[\|Z - Z^{(M)}\|^2_{L^2(D)}\right] = \sum_{m=M+1}^{\infty} \lambda_m.$$

Choose M to retain sufficient fraction $\delta \in (0,1)$ of total variance, i.e.,

$$\frac{\mathbf{E}\left[\|Z - Z^{(M)}\|^2_{L^2(D)}\right]}{\mathbf{E}\left[\|Z\|^2_{L^2(D)}\right]} = \frac{\sum_{m=M+1}^{\infty} \lambda_m}{\sum_{m=1}^{\infty} \lambda_m} = 1 - \frac{\sum_{m=1}^{M} \lambda_m}{|D|\sigma^2} < \delta.$$



- transmissivity measurements at 38 test wells
- head measurements, used to obtain boundary data via statistical interpolation (kriging)
- constant layer thickness of $b = 8\text{m}$
- constant porosity of $\phi = 0.16$
- SANDIA Nat. Labs reports
  [Caufman et al., 1990]
  [La Venue et al., 1990]

Merge transmissivity data with statistical model:

(1) Point estimates of parameters σ, ρ and ν via restricted maximum likelihood estimation (REML).

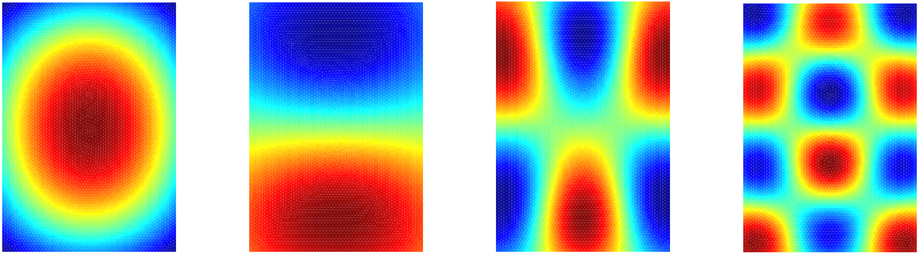
(2) Condition resulting covariance structure of $\log T$ on transmissivity measurements. (Low-rank modification of covariance operator.)

(3) Approximate $\log T$ by truncated Karhunen-Loève expansion.

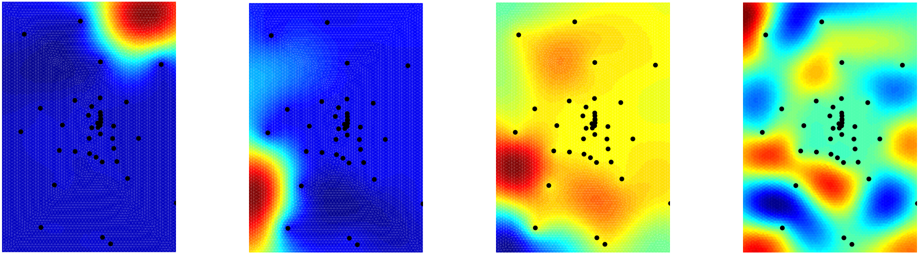
A Case Study: Radioactive Waste Disposal

WIPP KL modes conditioned on 38 transmissivity observations

unconditioned, $m = 1, 2, 9, 16$



conditioned, $m = 1, 2, 9, 16$



A Case Study: Radioactive Waste Disposal

Deterministic parametric representation

- Parametrize input RF by vector of independent Gaussian RV $\{\xi_m\}_{m=1}^M =: \boldsymbol{\xi}$.
- If ξ_m has density ρ_m and image $\Gamma_m := \xi_m(\Omega)$, then (Doob-Dynkin lemma)

$$L^2_{\mathbf{P}}(\Omega) \simeq L^2_\rho(\Gamma), \qquad \text{where} \qquad \Gamma := \times_{m=1}^{\infty} \Gamma_m, \;\; \rho = \prod_m \rho_m.$$

- Replace $Z(\boldsymbol{x},\omega)$, $p(\boldsymbol{x},\omega)$ ... with $Z(\boldsymbol{x},\boldsymbol{\xi})$, $p(\boldsymbol{x},\boldsymbol{\xi})$.

BVP becomes purely deterministic with (possibly) high-dimensional parameter space:

$$\begin{aligned} \nabla\cdot[T(\boldsymbol{x},\boldsymbol{\xi})\nabla p(\boldsymbol{x},\boldsymbol{\xi})] &= 0, && \boldsymbol{x}\in D, \quad \mathbf{P}\text{-a.s.}, \\ p(\boldsymbol{x},\boldsymbol{\xi}) &= p_0(\boldsymbol{x}), && \boldsymbol{x}\in \partial D, \quad \mathbf{P}\text{-a.s.}, \end{aligned}$$

where

$$\log T(\boldsymbol{x},\boldsymbol{\xi}) = \overline{Z}(\boldsymbol{x}) + \sum_{m=1}^{M} \sqrt{\lambda_m}\, Z_m(\boldsymbol{x})\, \xi_m.$$

- Generate sufficiently large ensemble of log-travel times $s(\boldsymbol{\xi}) = \log_{10} t(\boldsymbol{\xi})$
- Compute empirical CDF to quantify uncertainty in travel time.

Three sampling methods:

(1) Monte Carlo (MC) sampling of RV $\boldsymbol{\xi} \to s(\boldsymbol{\xi})$.
(N_{MC} solutions of PDE)

(2) Stochastic collocation (SC) → RF representation of velocity field $\boldsymbol{u}_{N_{SC}}(\boldsymbol{x}, \boldsymbol{\xi})$, use this to sample $s(\boldsymbol{\xi})$.
(N_{SC} solutions of PDE)

(3) Gaussian process emulator: N_{DP} MC samples of $s(\boldsymbol{\xi})$ used to calibrate surrogate of mapping $\boldsymbol{\xi} \to s(\boldsymbol{\xi})$, use this surrogate to sample $s(\boldsymbol{\xi})$.
(N_{DP} solutions of PDE)

- Draw independent random samples $\{\boldsymbol{\xi}_j\}_{j=1}^{N_{MC}}$ of $\boldsymbol{\xi}$.
- Solve determinisic PDE for each conductivity $\exp(Z_M(\boldsymbol{x}, \boldsymbol{\xi}_j))$.
- Solve ODE for each flow field $\boldsymbol{u}(\boldsymbol{x}, \boldsymbol{\xi}_j)$ and compute $s(\boldsymbol{\xi}_j)$.

- Draw independent random samples $\{\boldsymbol{\xi}_j\}_{j=1}^{N_{MC}}$ of $\boldsymbol{\xi}$.
- Solve determinisic PDE for each conductivity $\exp(Z_M(\boldsymbol{x}, \boldsymbol{\xi}_j))$.
- Solve ODE for each flow field $\boldsymbol{u}(\boldsymbol{x}, \boldsymbol{\xi}_j)$ and compute $s(\boldsymbol{\xi}_j)$.

How many samples do we need for a desired sampling error of

$$\mathbf{P}\left(\sup_{\boldsymbol{x}\in\mathbb{R}} |\hat{F}_{N_{MC}}(\boldsymbol{x}) - F(\boldsymbol{x})| \leqslant 0.01 \right) \geqslant 0.95 \; ?$$

Here F denotes the true CDF of s and $F_{N_{MC}}$ the empirical CDF obtained by N_{MC} samples. By Donsker's theorem we have

$$\sqrt{N_{MC}} \sup_{\boldsymbol{x}\in\mathbb{R}} |\hat{F}_{N_{MC}}(\boldsymbol{x}) - F(\boldsymbol{x})| \xrightarrow[N_{MC}\to\infty]{d} \sup_{\boldsymbol{x}\in[0,1]} |B(\boldsymbol{x})|,$$

where B is a standard Brownian Bridge on $[0, 1]$.

- Draw independent random samples $\{\boldsymbol{\xi}_j\}_{j=1}^{N_{MC}}$ of $\boldsymbol{\xi}$.
- Solve determinisic PDE for each conductivity $\exp(Z_M(\boldsymbol{x}, \boldsymbol{\xi}_j))$.
- Solve ODE for each flow field $\boldsymbol{u}(\boldsymbol{x}, \boldsymbol{\xi}_j)$ and compute $s(\boldsymbol{\xi}_j)$.

How many samples do we need for a desired sampling error of

$$\mathbf{P}\left(\sup_{\boldsymbol{x}\in\mathbb{R}} |\hat{F}_{N_{MC}}(\boldsymbol{x}) - F(\boldsymbol{x})| \leqslant 0.01\right) \geqslant 0.95 \text{ ?}$$

Here F denotes the true CDF of s and $F_{N_{MC}}$ the empirical CDF obtained by N_{MC} samples. By Donsker's theorem we have

$$\sqrt{N_{MC}} \sup_{\boldsymbol{x}\in\mathbb{R}} |\hat{F}_{N_{MC}}(\boldsymbol{x}) - F(\boldsymbol{x})| \xrightarrow[N_{MC}\to\infty]{d} \sup_{\boldsymbol{x}\in[0,1]} |B(\boldsymbol{x})|,$$

where B is a standard Brownian Bridge on $[0, 1]$.

This yields $N_{MC} \approx 20,000$. Can we do better than solving 20k PDEs?

Evaluate $v : \boldsymbol{\Gamma} \to V$ at collocation points $\boldsymbol{\Xi} := \{\boldsymbol{\xi}_j\}_{j=1}^{N_{SC}} \subset \boldsymbol{\Gamma}$, approximate $v_w \approx v$ in N_{SC}-dim. function space $\mathscr{V}_{\boldsymbol{\xi}}(\boldsymbol{\Gamma}; V)$.

Here: **Smolyak sparse tensor collocation**

$$v_w = \sum_{|\boldsymbol{i}| \leqslant w} \left[\bigotimes_{m=1}^{M} \Delta_{i_m}^{(m)} \right] v,$$

where $\Delta_0^{(m)} = 0$, $\Delta_k^{(m)} = I_k^{(m)} - I_{k-1}^{(m)}$ for $k \in \mathbb{N}$ and

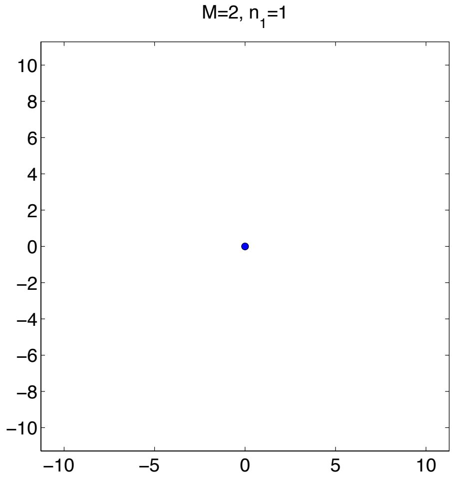
$$\left(I_k^{(m)} f \right)(\xi) := \sum_{\xi_j \in \Xi_k^{(m)}} f(\xi_j)\, \ell_j(\xi), \qquad \text{for } f : \Gamma_m \to V,$$

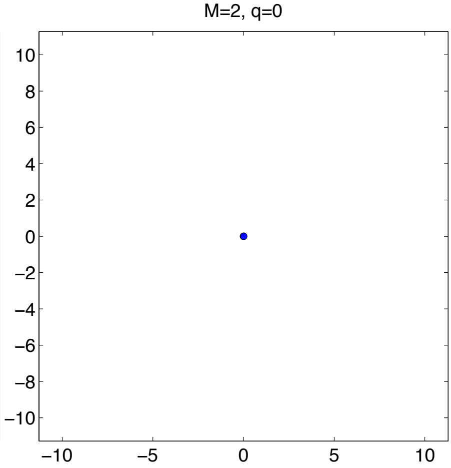
ℓ_j Lagrange polynomials associated with (1D) nodal sets $\Xi_k^{(m)} \subset \Gamma_m$.

Here: $\Xi_k^{(m)}$ are the $(2^{(k-1)} + 1)$th **Gauss-Hermite nodes**, $\Xi_1^{(m)} = \{0\}$.

A Case Study: Radioactive Waste Disposal

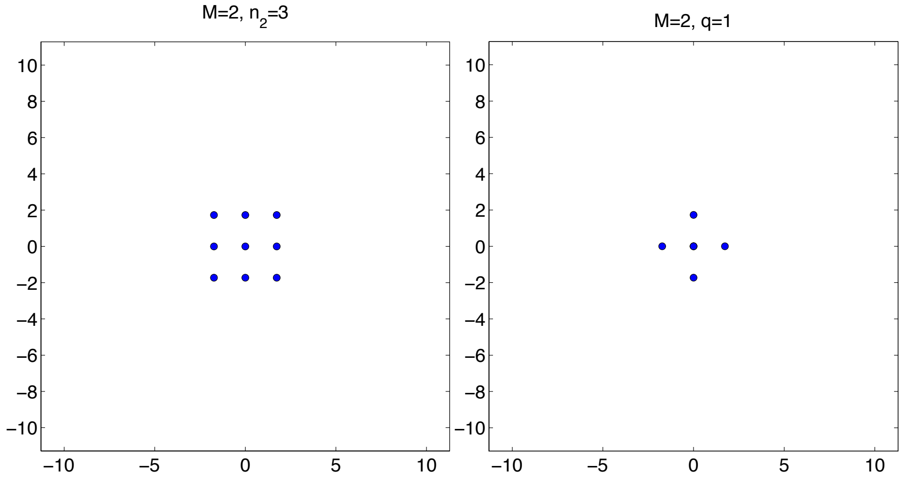
Smolyak sparse grid based on Gauss-Hermite nodes





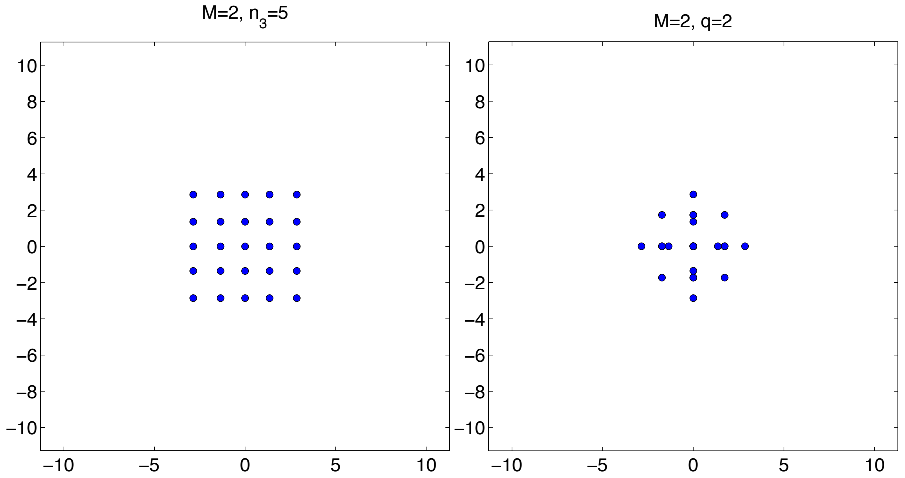
A Case Study: Radioactive Waste Disposal

Smolyak sparse grid based on Gauss-Hermite nodes



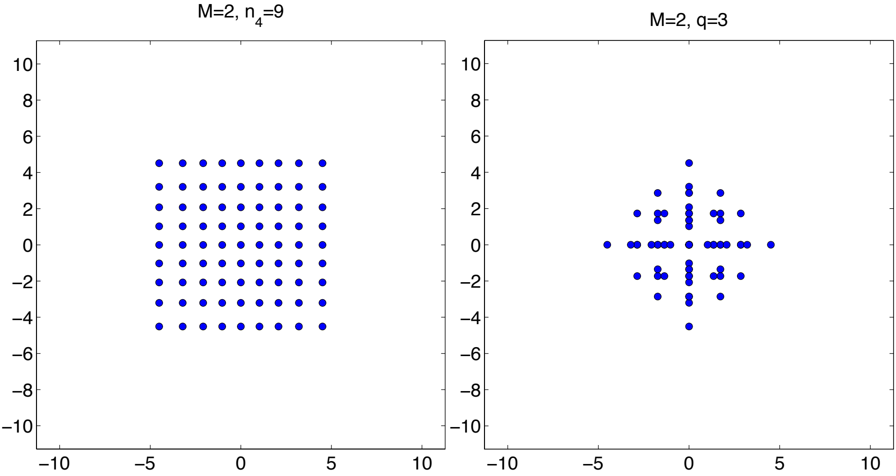
A Case Study: Radioactive Waste Disposal

Smolyak sparse grid based on Gauss-Hermite nodes



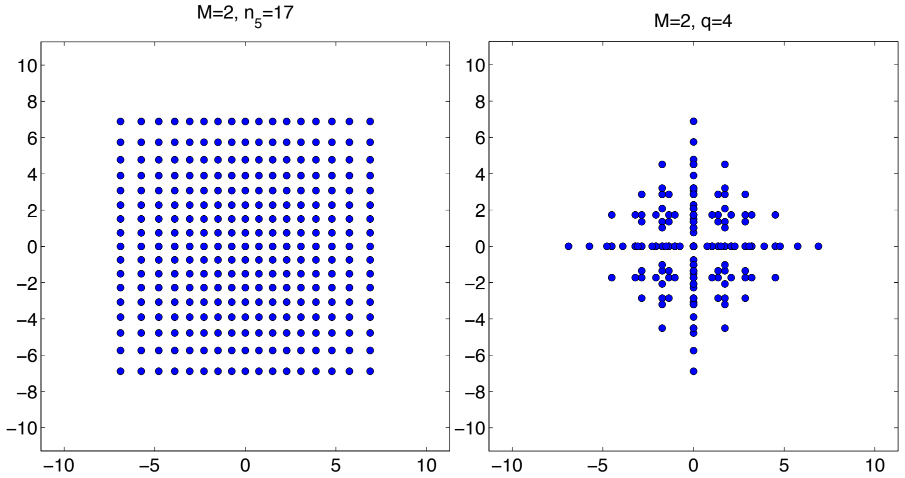
A Case Study: Radioactive Waste Disposal

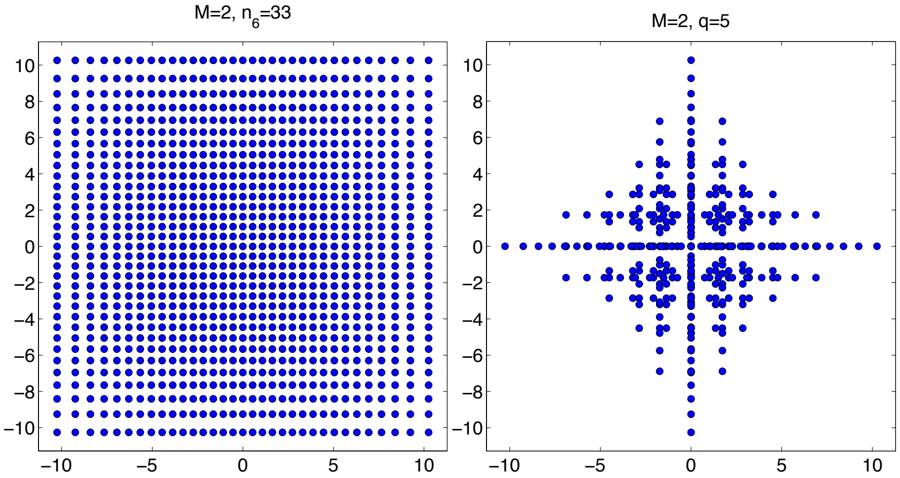
Smolyak sparse grid based on Gauss-Hermite nodes



A Case Study: Radioactive Waste Disposal

Smolyak sparse grid based on Gauss-Hermite nodes





M=2, n6=33
M=2, q=5
10
8
6
4
2
0
−2
−4
−6
−8
−10
−10
−5
0
5
10

An emulator is a statistical approximation to the output of a computer code

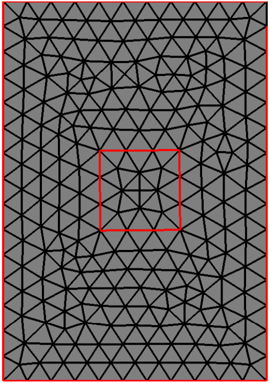
$$\boldsymbol{y} = f(\boldsymbol{x}).$$

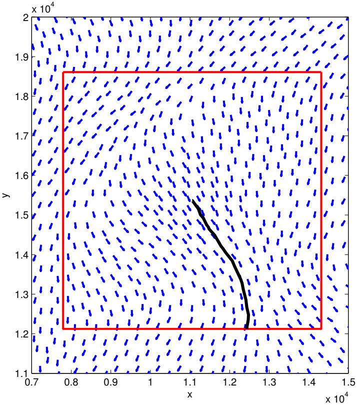
Basic idea:

(1) Represent the code, $f(\cdot)$ as a Gaussian stochastic process.
(2) Run model for sample of design inputs $\boldsymbol{x}$ and observe outputs $\boldsymbol{y}$.
(3) Condition GP on observed outputs $\boldsymbol{y}$.
(4) Emulator provides a distribution function for the output of the computer code.
(5) Use emulator as a surrogate for computer model when performing MC analysis.

[Kennedy & O'Hagan, 2001], [Stone, 2011]

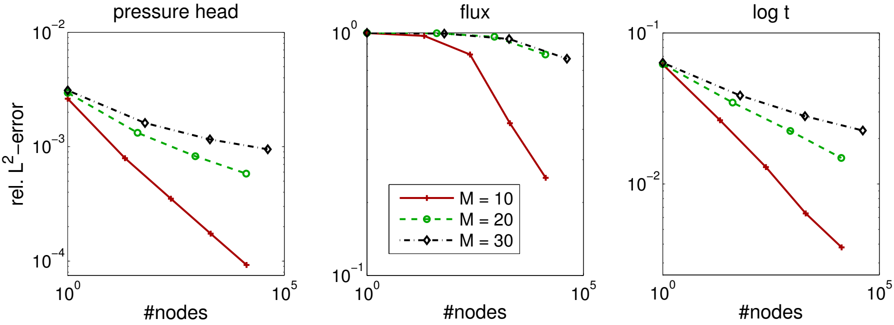
- Mixed FE discretization:
  lowest order RT elements for $\boldsymbol{u}$, pcw. constants for p.
- Fixed mesh, 29 208 triangles, (73 234 DOF)
- Flow divergence-free $\Rightarrow$ discrete fluxes pcw. constant, (makes particle trajectory calculation trivial).





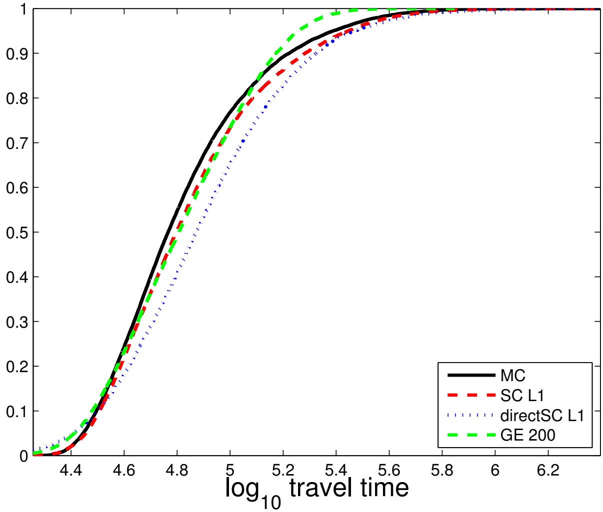
Error w.r.t. MC reference calculation with $N_{MC} = 20,000$.

Errors measured in $L^2_\rho(\mathbf{\Gamma}; L^2(D))$, $L^2_\rho(\mathbf{\Gamma}; H(\mathrm{div}, D))$ resp. $L^2_\rho(\mathbf{\Gamma}; \mathbb{R})$ norms.



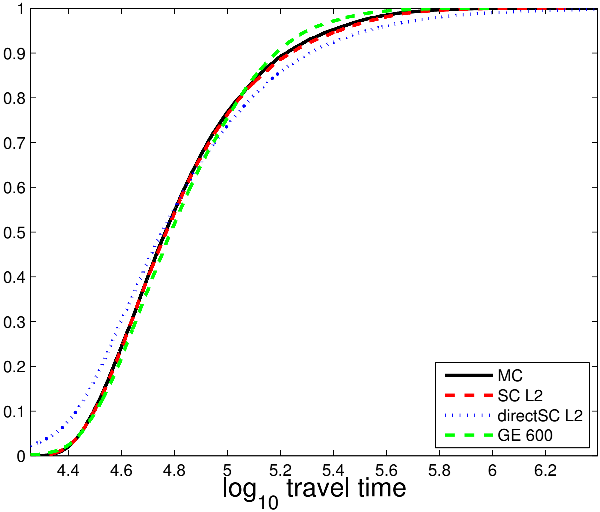
A Case Study: Radioactive Waste Disposal

Travel Time CDFs for $M = 20$ KL modes



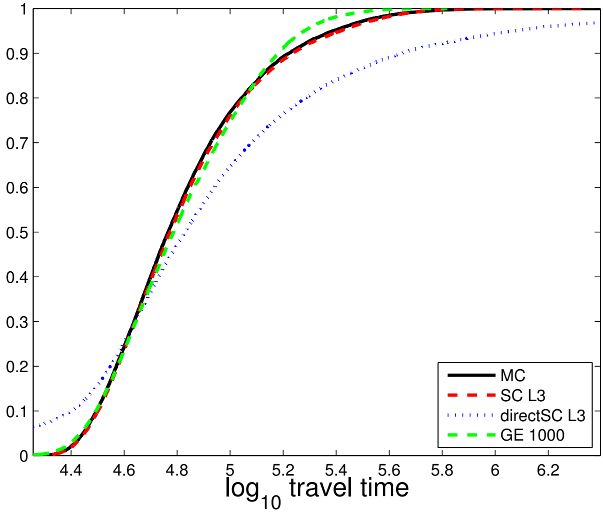
A Case Study: Radioactive Waste Disposal

Travel Time CDFs for $M = 20$ KL modes



A Case Study: Radioactive Waste Disposal

Travel Time CDFs for $M = 20$ KL modes

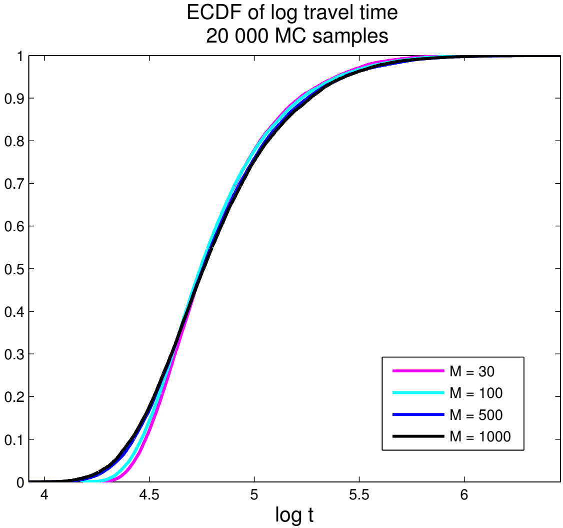


Statistical test to determine whether the random evaluations of the surrogates were drawn from the same distribution as pure MC.

Significance level: $\alpha = 0.05$. KL length: $M = 20$.

Surrogate	$N_{surrogate}$	KS-test (1K)	KS-test (20K)
SC	41	✗	✗
	881	✓	✓
	13201	✓	✓
GPE	200	✗	✗
	400	✓	✗
	600	✓	✗
	1000	✓	✗

Basically the same results for $M = 10$ and $M = 30$.



Empirical CDF of log t based on KL approximations of different length

Recall parametrized Darcy flow model (primal form):

$$\nabla\cdot\big(T(\boldsymbol{x},\boldsymbol{\xi})\nabla p(\boldsymbol{x},\boldsymbol{\xi})\big) = 0, \quad \boldsymbol{x}\in D, \boldsymbol{\xi}\in\Gamma, \qquad p = p_0 \text{ along } \partial D.$$

- Observations $\{p(\boldsymbol{x}_k\}_{k=1}^K$ of PDE solution at locations $\boldsymbol{x}_k \in D$. Denote by observation operator $\boldsymbol{Q} : H^1(D) \to \mathbb{R}^K$.
- Forward map (parameter-to-observable map):

$$\boldsymbol{\xi} \xmapsto{\text{parametrization}} T(\cdot,\boldsymbol{\xi}) \xmapsto{\text{PDE}} p(\cdot,\boldsymbol{\xi}) \xmapsto{\text{observation}} \boldsymbol{Q}(p(\cdot,\boldsymbol{\xi})) =: \boldsymbol{y}.$$

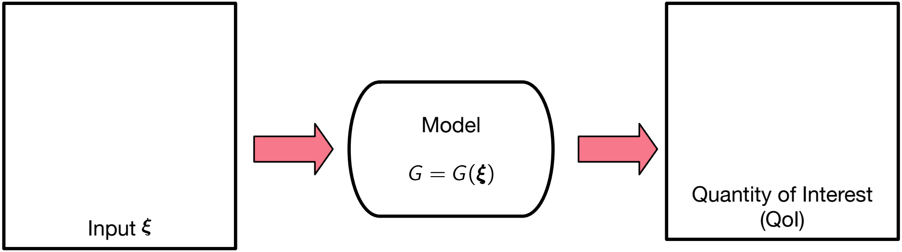
  Summarize as parameter identification problem: given data $\boldsymbol{y}$, determine parameter $\boldsymbol{\xi}$ such that

$$\boldsymbol{y} = G(\boldsymbol{\xi}).$$

- **Note:** depending on parametrization, $\boldsymbol{\xi}$ may be finite or infinite-dimensional.

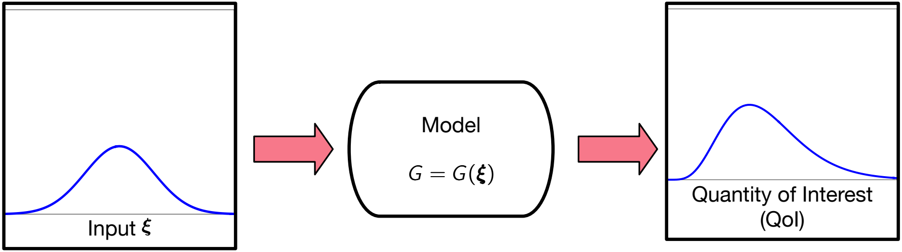
A Case Study: Radioactive Waste Disposal

Uncertainty propagation vs. Bayesian inversion



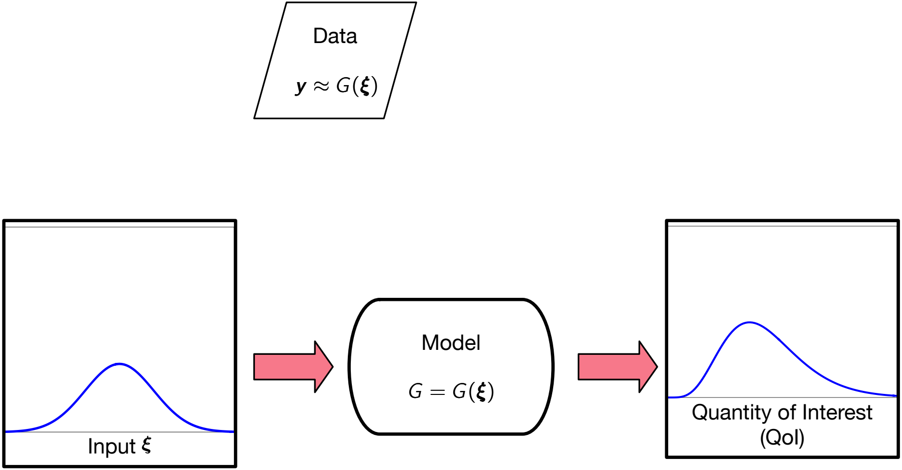
A Case Study: Radioactive Waste Disposal

Uncertainty propagation vs. Bayesian inversion



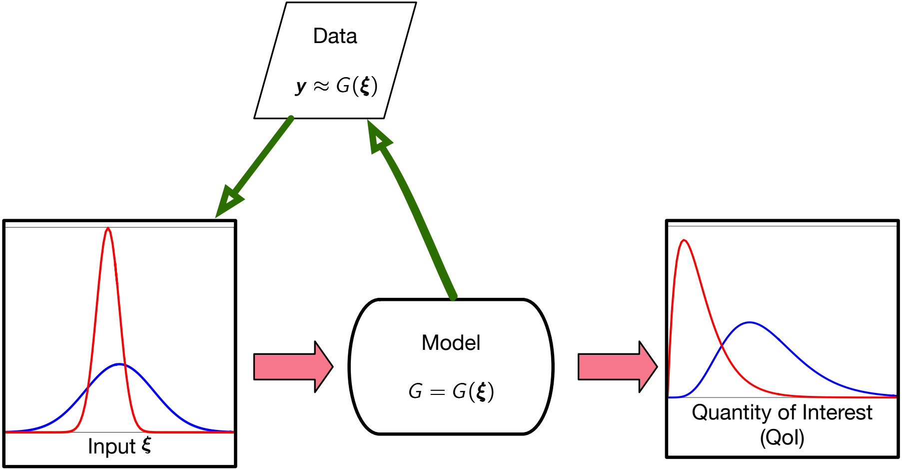
A Case Study: Radioactive Waste Disposal

Uncertainty propagation vs. Bayesian inversion



A Case Study: Radioactive Waste Disposal

Uncertainty propagation vs. Bayesian inversion



Given forward map $G : \mathscr{H} \to \mathbb{R}^k$, $\mathscr{H}$ a separable Hilbert space, and noisy data

$$\boldsymbol{y} = G(\boldsymbol{\xi}) + \boldsymbol{\varepsilon}, \qquad \boldsymbol{\varepsilon} \in \mathbb{R}^K,$$

we wish to infer $\boldsymbol{\xi} \in \mathscr{H}$.

Bayesian approach:

- $\boldsymbol{\varepsilon}$ is a realization of random noise
- $\boldsymbol{\xi}$ is a realization of a prior probability measure μ_0 on $\mathscr{H}$

Given forward map $G : \mathscr{H} \to \mathbb{R}^k$, $\mathscr{H}$ a separable Hilbert space, and noisy data

$$\boldsymbol{y} = G(\boldsymbol{\xi}) + \boldsymbol{\varepsilon}, \qquad \boldsymbol{\varepsilon} \in \mathbb{R}^K,$$

we wish to infer $\boldsymbol{\xi} \in \mathscr{H}$.

Bayesian approach:

- $\boldsymbol{\varepsilon}$ is a realization of random noise
- $\boldsymbol{\xi}$ is a realization of a prior probability measure μ_0 on $\mathscr{H}$
- "Inversion": Condition $\boldsymbol{\xi} \sim \mu_0$ on the event that

$$G(\boldsymbol{\xi}) + \boldsymbol{\varepsilon} = \boldsymbol{y},$$

obtain conditional probability measure $\mu^{\boldsymbol{y}}$ on $\mathscr{H}$

$$\mu^{\boldsymbol{y}}(A) := \mathbf{P}(\boldsymbol{\xi} \in A \,|\, G(\boldsymbol{\xi}) + \boldsymbol{\varepsilon} = \boldsymbol{y}), \qquad A \in \mathfrak{B}(\mathscr{H}).$$

- Random noise is multivariate Gaussian: $\varepsilon \sim N(0, \Sigma)$
- Prior measure is Gaussian measure on $\mathscr{H}$: $\mu_0 = N(0, C_0)$
- Forward map $G : \mathscr{H} \to \mathbb{R}^k$ is continuous and $\forall \alpha > 0 \; \exists K_\alpha < \infty$:

$$|G(\boldsymbol{\xi})| \leqslant K_\alpha \exp(\alpha \|\boldsymbol{\xi}\|^2_{\mathscr{H}}).$$

- $\boldsymbol{\xi} \sim \mu_0$ and $\varepsilon \sim N(0, \Sigma)$ are independent

- Random noise is multivariate Gaussian: $\varepsilon \sim N(0, \Sigma)$
- Prior measure is Gaussian measure on $\mathscr{H}$: $\mu_0 = N(0, C_0)$
- Forward map $G : \mathscr{H} \to \mathbb{R}^k$ is continuous and $\forall \alpha > 0 \; \exists K_\alpha < \infty$:

$$|G(\boldsymbol{\xi})| \leqslant K_\alpha \exp(\alpha \|\boldsymbol{\xi}\|^2_{\mathscr{H}}).$$

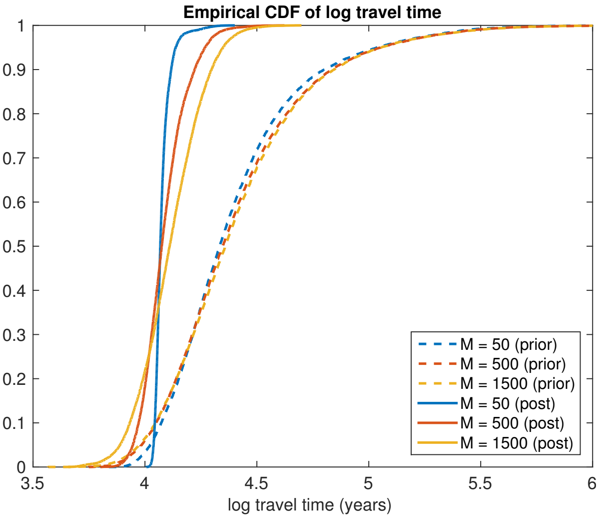
- $\boldsymbol{\xi} \sim \mu_0$ and $\varepsilon \sim N(0, \Sigma)$ are independent

Then the conditional probability measure $\mu^{\boldsymbol{y}}$ is given by Bayes' rule:

Theorem 1.1 (Bayes' rule [Stuart (2010)],[Dashti & Stuart (2016)])

The posterior measure $\mu^{\boldsymbol{y}}$ is given by

$$\mu^{\boldsymbol{y}}(\mathrm{d}\boldsymbol{\xi}) \propto \exp(-\Phi(\boldsymbol{\xi}; \boldsymbol{y}))\, \mu_0(\mathrm{d}\boldsymbol{\xi}), \qquad \Phi(\boldsymbol{\xi}; \boldsymbol{y}) = \frac{1}{2}|\boldsymbol{y} - G(\boldsymbol{\xi})|^2_{\Sigma^{-1}}.$$



Empirical CDF of log travel time
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
3.5
4
4.5
5
5.5
6
log travel time (years)
M = 50 (prior)
M = 500 (prior)
M = 1500 (prior)
M = 50 (post)
M = 500 (post)
M = 1500 (post)