

Possible structure for the meat of the report:

Stage of design process

- Description
- Example of uncertainties
- Uncertainty Approaches
 - Bayesian strategy?
 - Frequentist strategy?
 - IP strategy?

OR

Stage of design process

- Description
- Example of uncertainties

Analysis framework

- Flowchart showing the inter-relation between the different approaches

Introduction

The engineering design process is, whether explicitly acknowledged or not, about managing uncertainty. Any system we design will be required to operate in a range of environmental conditions, or there may be some variability induced in the systems performance as a result of imperfections in the manufacturing process. This is the first type of uncertainty engineers encounter in the design process. Sometimes called aleatory uncertainty. It's often possible to characterise this uncertainty with probability theory, and in most cases this type of uncertainty is considered irreducible. Not irreducible in the sense an engineer cannot modify the design or manufacture process to reduce variability, but irreducible in the sense that any more information we gain about this uncertainty does not reduce it.

The second type of uncertainty engineers have to deal with in the design process is what they do not know. Sometimes called epistemic uncertainty. Epistemic uncertainty is encountered in many places within the design process. This might be as a result of some boundary conditions of the model being imprecisely known, or some material property that is not known. Or this type of uncertainty may be introduced because of some limitation in the precision of measuring devices used to gather data. Epistemic uncertainty, conversely, is often considered reducible through the investment of empirical effort. For example, we may not precisely know some system boundary condition, but we could conduct some experiments to try and improve our knowledge about it. Or, we may conduct further tests on materials used in our construction to better characterise their properties. Or, we may use better measuring devices to improve the precision and accuracy of data we extract.

These two types of uncertainty are well defined, and whilst in practice they may be difficult to separate, much of their calculus and their effects are well understood. The third type of uncertainty in design is about the decisions of the engineer. This is sometimes called volitional uncertainty. This topic is less well explored within the scientific literature, and in

some ways it's more difficult to pin down a definition, or often to recognise its effect within practical applications. When engineers design complex solutions and systems they often have a number of competing objectives. It's most often the case that one design solution will not be optimal against all objectives. This is an obvious case where the engineer needs to add some subjective weight to particular objectives in order to make a decision. In some cases it may be fine to not explicitly account of the effects of this kind of decision, for example the paint colour of a product. The paint colour, in most cases, will have no impact at all on any other decisions, in this way it is isolated from others and accounting for engineers volition here is useless. In more complex scenarios however, the outcome of one decision may have a significant impact on other subsequent decisions. In these scenarios these decisions are not isolated from one another, so tracking the dependence and interaction between decisions may be helpful within the design process.

This report will explore the engineering design process with a particular focus on the nature and implications of the three types of uncertainty previously described. These uncertainties are often difficult to characterise in isolation -- it's hard to separate them -- and their combined implications require subtle analyses to quantify in complex situations. We will touch on some methodologies developed within the uncertainty quantification community, and discuss strategies to combine them in more complex solution sequences to help answer the very difficult questions engineers typically ask in the engineering design process.

Interpretation

- Narratives about possibilities
- Decision and visualisation
- Visualisation
 - Risk analytic visualisations
 - Farmer curves
 - Tangled vs fans
 - Higher dimensions
-

The general design problem

Most engineering design challenges share some commonality. We almost always begin by identifying the task we need a system, or a solution to complete. This typically involves the specification of a specific mission, or set of objectives. Once these objectives have been defined an engineer might typically look at other products that do a similar task, this might be a previous generation of the same system, or the product of a competitor.

- Mission
- Review other systems
- Computational Model
- Prototype and experiment
- Calibrate model
- Refine the design / enhance robustness / reliability
- Manufacture / Implement
- Validation

<<Maybe we can introduce some mathematical notation here for the general design problem. Where we might have some design variables d , sources of aleatory uncertainty a , and epistemic uncertainties e .

We have the mission :

$G(d|a,e) > T$

.. I think it might help when we come to discuss the methods to plug things together.>>

Data driven engineering

Increasingly, engineers rely on data to support their design process. Data driven engineering is

Data however takes many forms, and can be a radically different qualities, it's not always the perfect data, extracted with a perfectly representative sample, in a similar situation to that for which the system will be required to operate, or indeed with accurate sensor or measurement devices.

- Data
 - Good Data = satisfies all our assumptions
 - Bad Data = interferes with some or all of our assumptions
 - No Data = How do we make things up, but be conservative still.

Data structure

Numbers have different representations. And these can come from the uncertainty classes.

- Whittling
- Data fitting
- Performance based characterisations
 - Worst-case
 - C-boxes
 - k-s

Model uncertainty

Uncertainty about the model of uncertainty (distribution shape)

Developing Uncertainty Language

Data structures

Class **Data**

- Data storage
 - Table or matrix
- Units
 - String
- Ensemble
- Uncertainty about each entry
- Nature of uncertainty/justification/pedegree/source
 - > Visualisations

Class **Dependence**

- Inheritance
- Independence
- Matrix of dependence
- Correlation
- Copula
- Empirical copula
 - > Visualisations

Auto correlations

Lists

Series

Fields

Design decisions

Trees

Parametric

- Lists
- Ranges
- Nesting

Constructors

- Defaults
- Data centric approach
 - Bayesian strategies
 - Bayesian approach <Inference>
 - ABC
 - Robust Bayesian approach
 - Distribution free
 - Bootstrapping
 - ECDF
 - Maximum-Likelihood
 - With constraint
 - Without constraint
 - Method-of-Moments
 - Basic Idea (motivation)
 - Implementation
 - Caveats

- References

- Fiducial

- Confidence boxes
- Dependence
- Whittling

- Joint data
- Maximum-Entropy
- Expert elicitation

	Measurements	Point-values	Intervals
Do you have measurements of the quantity? (3.5)	Entire population	empirical distribution	
Of the entire population? (3.5.2, 3.5.3)	Abundant random sample data		
Of a sample from a population? (3.5.4)	Random sample data		
Can you limit the range? (3.5.6.1)	Opportunity samples		
Can you limit the possible range of the quantity? (3.4.1)	Anecdotes		
Do you know (bounds on) the mean? (3.4.1.2, 3.4.1.5)			
Do you know (bounds on) the mode? (3.4.1.4)	Constraint information		
Do you know (bounds on) the median or other quantiles? (3.4.1.3)	Quantiles		
Do you know (an upper bound on) the variance? (3.4.4)	Mean, standard deviation & range		
Do you know (bounds on) the mean? (3.4.4.1)	Mean & standard deviation		
Do you know either the minimum or maximum? (3.4.4.2)	Median & range		
Do you know both the minimum and maximum? (3.4.4.3)	Mode & range		
Can you put upper limits on the probability density? (3.4.8)	Range only		
Can you put lower limits on the probability density? (3.4.8)	Distribution family		
Is the quantity necessarily positive? (3.4.11.1)	Positive		
Is the distribution symmetric? (3.4.11.3)	Symmetric		
Is the distribution unimodal?	Unimodal		
Do you know (bounds on) the mode? (3.4.11.2)	Convex		
Do you know the mean and at least one endpoint of the range? (3.4.11.2)	Concave		

Do you know the mean and variance? (3.4.11.2)	Hazard function sign
Is the distribution function convex? (3.4.11.4)	
Is the distribution function concave? (3.4.11.5)	Expert opinions
Is the hazard function increasing or decreasing? (3.4.11.6)	
Can you express the quantity in a model in terms of other better known quantities? (3.2)	No information
Do you specify a set of prior expectations for the distribution of the quantity?	
Can you construct a set of likelihood functions? (3.3)	
Can you mechanistically justify a particular distribution shape?	
Can you bound the parameters? (3.1)	
Do you have multiple estimates for the quantity? (4)	
Are all the estimates reliable? (4.3)	

Calculus of uncertainty

Unary operators

Binary operations (min,max,power)

Transformations

Logical operators

Back calculations

Control structures (if, while, for)

Simulation of uncertainty

Simulation based and other non-intrusive methods.

Class simulation

- Input Inheritance
- Model wrapper
- Propagation method

Uncertainty and design <Examples>

Uncertainty Characterisation <Where does uncertainty come from in design?>

Here we identify different sources of uncertainty.

Design conception (design/volitional uncertainty) (vagueness)

- Mission design / Motivation
- Mixing and matching of constrained option design spaces
 - Standardisation
 - Evidence about performance (Extrapolation/ Experimental set ups/), prior information

Review

Literature/ Copying / second generation design (Inferential)

- Characterising uncertainty when we have historical experience

Experiments (Inferential)

- Measurement uncertainty
- Statistical sampling
- Sensor calibration (example)
- Experimental design (Vol / Power Analysis /)

Operationally latent variable (No Data)

- Inference
- Expert elicitation
- Reparameterisation
- Additional modelling
 - Fitting
 - Expert elicitation
 - Vagueness? Is it useful in this case
 - Is it aleatory - can it be vague about whether or not it is vague

Model construction

- Probabilistic numerics
- Stochastic perturbation
- Interval
- Bounding studies
- Multi-fidelity models
 - computational uncertainty (mesh size, rep size, roundoff errors, etc.)
 - Numerical error
 - Model error (Vagueness ?)
 - Predictive capability

Model calibration

Deterministic -> optimisation

Stochastic -> updating
Stochastic+imprecise -> hierachicle updating
Fuzzy ->

Risk- Reliability-based design

Environmental sourced of uncertainty

Manufacturer or assembly tolerances

- Confidence Level (vagueness)
- Analyses involving uncertainty
All of these tools need to appreciate uncertainty. Here we are describing how the uncertainty described in the characterisation impact in the case of employing each of these analyses.
 - BASIC IDEA
 - 1 EXAMPLE
 - Tee-UP other document
- Calibration
- Sensitivity
- Validation
- Propagation (Refer to framework 1)
- Reliability analysis
- Optimisation
- Iterative design strategies
 - Decision theory
 - Set based design
- Certification / Safety
 - Event tree/ Fault tree

The first step is the hardest: specifying input distributions

23 July 2014, Liverpool, United Kingdom
Institute for Risk and Uncertainty, University of Liverpool
10 am, Mason Bibby Common Room, Harrison Hughes Building

A fundamental task in probabilistic risk analysis is selecting an appropriate distribution or other characterization with which to model each input variable within the risk calculation. Currently, many different and often incompatible approaches for selecting input distributions are commonly used, including the method of matching moments and similar distribution fitting strategies, maximum likelihood estimation, Bayesian methods, maximum entropy criterion, among others. We compare and contrast six traditional methods and six recently proposed methods for their usefulness in risk analysis in specifying the marginal inputs to be used in probabilistic assessments. We apply each method to a series of challenge problems involving synthetic data,

taking care to compare only analogous outputs from each method. We contrast the use of constraint analysis and conditionalization as alternative techniques to account for relevant information, and we compare criteria based on either optimization or performance to interpret empirical evidence in selecting input distributions. Despite the wide variety of available approaches for addressing this problem, none of the methods seems to suffice to handle all four kinds of uncertainty that risk analysts must routinely face: sampling uncertainty arising because the entire relevant population cannot be measured, mensurational uncertainty arising from the inability to measure quantities with infinite precision, demographic uncertainty arising when continuous parameters must be estimated from discrete data, and model structure uncertainty arising from doubt about the prior or the underlying data-generating process.

Woodpile -

Second framework document ideas

Good Data

Bad Data

Little Data

No Data

Expert elicitation

Uncertainty Characterisation

Measurement?

- Poorly done by most literature
- Review of standards
- Type A & B

Data sets - Data Fusion

- What if data comes from
- Data fusion
 - Frechet
 - Build a model

Literature

- Chronic under reporting of uncertainty

No Data

Computational Methods

Decision Theory

Probabilistic

Bayesian

IP

Info-gap

