NPL
National Physical Laboratory

NPL REPORT MS 3

Sensitivity analysis, optimisation, and sampling methods applied to continuous models: three test problems

Trevor Esward, Clare Matthews, Louise Wright, Xin-She Yang

November 2010

National
Measurement
System

NPL Report MS 3

Sensitivity analysis, optimisation, and sampling methods applied to continuous models: three test problems

Trevor Esward, Clare Matthews, Louise Wright, Xin-She Yang
Mathematics and Scientific Computing Group

November 2010

ABSTRACT

This report covers the initial stages of work in the SSfM project Tools for Continuous Modelling and Simulation. The aims of the project are to improve the understanding and use of continuous modelling packages and to improve the confidence of model users in the quality of model results. The initial stages have investigated the application of sensitivity analysis, optimisation, and sampling methods to finite element models.

The work has applied a variety of sensitivity analysis, optimisation and sampling methods to three small finite element models that cover a range of physical problems, complexities and interactions between model parameters. The aim of the work was to identify methods that repeatably give accurate results with minimal computational effort, and to provide advice on algorithm selection.

Due to the varying nature of the test problems, this stage of the work produced both model specific and more general conclusions that will be applied to two larger case studies in the next stage of the project.

NPL Report MS 3

ISSN 1754–2960

National Physical Laboratory,
Hampton Road, Teddington, Middlesex, United Kingdom TW11 0LW

We gratefully acknowledge the financial support of the UK Department for
Business, Innovation and Skills (National Measurement Office)

Approved on behalf of NPLML by Jonathan Williams,
Knowledge Leader for the Optical Technology and Scientific Computing team

Contents

1 Introduction

1.1 Scope

This report introduces key aspects of the Software Support for Metrology (SSfM) project *Tools for Continuous Modelling and Simulation*. The Software Support for Metrology programme is funded by the National Measurement Office (part of the Department for Business, Innovation and Skills).

The general aim of government-supported National Measurement System (NMS) programmes is to provide and develop, at the national level, an infrastructure that ensures measurement in the UK is valid, fit for purpose, consistent and internationally recognised. SSfM is a programme that underpins the NMS, focusing on the use of mathematics and computing in metrology. Its purpose is to achieve a balance between research and development, whilst also extending the range of techniques and applications available to meet the continually changing needs of metrology.

The results of the programme are promulgated through reports, published papers, best and good practice guides (available from the NPL website), newsletters, workshops, presentations and other awareness activities, and through the support that the SSfM team gives to other parts of the NMS. Solutions to problems specific to one field of measurement are generalised to provide techniques that are more widely applicable.

The purpose of the *Tools for Continuous Modelling and Simulation* project is to support the understanding and use of continuous modelling packages and to improve the confidence of model users in the quality of model results. We regard continuous modelling as encompassing modelling methods based on the solution of differential and partial differential equations, using such techniques as finite element, finite difference, boundary element and computational fluid dynamics methods. Continuous modelling is widely applied in many branches of science and engineering and the increased usability of software, especially arising from the development of pre- and post-processing tools linked to powerful modern computers, means that methods are readily accessible to non-expert modellers.

There are two common uses of continuous modelling within NMS programmes. The first is to design new measurement systems or experiments, and the second is to estimate unknown properties of existing devices or to assist in interpreting the results of experiments. In applications of the first type, the purpose of the model is to generate a design for a new product or piece of equipment with a particular performance. This application is particularly common in industry where continuous modelling, and particularly finite element (FE) modelling, is widely

used for product design. In applications of the second type, models of equipment, objects, or experiments require specification of model parameters, such as material or dimensional properties, that are not accurately known, but experimental data collected from the system being modelled may be available. For both these application types optimisation methods can be used to estimate values of model parameters to match model results to measured values or to produce a desired performance.

An understanding of the sensitivity of the model results to the various model parameters enables robust designs to be identified and allows the metrologist to identify and control the critical parameters without wasting time and effort controlling the unimportant ones.

1.2 Project outputs

There are four main outputs from the *Tools for Continuous Modelling and Simulation* project.

- The first output is a companion document to this report, *Sensitivity analysis, optimisation, and sampling methods applied to continuous models* [6], which sets out a general methodology for sensitivity analysis and optimisation applied to continuous models.
- The second output is this report, which describes how the general methodology was applied to three specific finite element problems. It demonstrates the methodology in action and gives much more detail than is included in the first report. Whilst this report can be understood without reference to the other outputs, the methodology used to generate the results is only given in summary form and the companion report provides a more detailed explanation that may help the reader's understanding.
- Extensive additional numerical results from the three finite element problems analysed in the second report have been made available on the SSfM web site. This information is provided for readers who wish to understand in more detail the analysis of the three test problems reported here.
- Two case studies have been carried out in which optimisation has been applied to two challenging metrology problems that are currently the subject of research at NPL. The first is a transient heat flow problem in which the thermal properties of solid materials are measured using a laser flash technique. The second is the measurement of charge mobility in organic light emitting diodes. Both problems are modelled using finite difference techniques and experimental results are in both cases used in the objective

function in an optimisation process. The case studies are described in the two reports [7, 8].

1.3 The audience for this report

The audience for this report are as users of continuous modelling, including those who employ models as tools for design, who want to determine unknown properties by matching model results to measurement data, or who wish to associate uncertainties with their results (uncertainties due to errors in model parameters, rather than discretisation errors or other errors that may be associated with the choice of solution method). It is expected that the audience has read the companion report [6] and wishes to know how the conclusions outlined in that report were reached. The methods we discuss in this report are applicable to all forms of continuous modelling, not simply to finite element models. As explained above, we illustrate the use of the methods in both finite element and finite difference models.

1.4 Content and structure of report

This report details the results of applying the methodology outlines in the companion report [6] to a set of three small finite element problems. The work reported here compares a range of different sensitivity measures, different optimisation algorithms, and different sampling methods on grounds of computational efficiency, robustness, and repeatability.

The structure of this report is as follows. The next section summarises the methodology and terminology described in depth in the companion report [6]. Section 2.1 defines the terminology used throughout both reports, and section 2.2 summarises the methodology used for generating the results reported here. Section 2.3 outlines some of the decisions made when choosing and setting up the finite element models and explains the reasons for the choices that have been made. Section 2.4 lists the sensitivity, optimisation, and sampling methods used and provides references for a deeper explanation of each method.

The remainder of the report focusses on the three finite element problems. Section 3 describes a static linear elasticity problem, section 4 a natural frequency problem, and section 5 a transient thermal problem.

Each of these sections has the same structure. First the problem is defined, including a complete definition of the finite element model and an objective function for the optimisation. Next the results of the sensitivity analysis are

presented, including a comparison of the different measures available and the effects of sample size on the measures. The next subsection describes the results of the optimisation, including a comparison of different methods and identification of any unusual features of the problem. The sampling methods are compared in the next subsection, inclusing tests of repeatability and an examination of the effects of sample size. The final subsection presents the conclusions to be drawn from the results for the problem.

The final section 6 draws together the conclusions from the individual test problems.

1.5 What this document does not do

All finite element models described in this work have been implemented within SOFEA [28], a finite element toolbox written in Matlab [26]. One of the reasons for this choice was the ease with which this finite element code can be interfaced with the wide range of optimisation algorithms available within Matlab.

This document does not provide advice on wrapping algorithms around specific continuous modelling software packages, nor does it explain how to automate repetitive tasks in specific packages. The majority of popular continuous modelling packages now offer automated running by means of a scripting language, but the languages and approaches vary too much between software packages for generic advice to be useful.

Similarly, we do not give advice about individual implementations of optimisation algorithms. Two different implementations of one algorithm have been tested during the work, and the variation in results that was observed illustrates that implementation can make a difference to performance. Since there are a wide range of packages available and our experience of these packages applied to finite element models is limited, no package-specific advice is given here.

For further assistance on such topics, consult either the appropriate software manual or the software suppliers. The authors of this report are happy to investigate specific examples but cannot offer general advice beyond what is given in the deliverables listed in section 1.2.

2 Methodology and terminology

2.1 Terminology

Consider a mathematical model of the form

$$\mathbf{y} = \mathbf{g}(\mathbf{x}; \alpha), \tag{1}$$

where $\mathbf{y} = (y_k; k = 1, 2, \ldots, K)^{\mathrm{T}}$, $\mathbf{x} = (x_n; n = 1, 2, \ldots, N)^{\mathrm{T}}$ and $\alpha = (\alpha_j; j = 1, 2, \ldots, J)^{\mathrm{T}}$. The values of the α_j are known and fixed. The values of the x_n can vary from one model evaluation to another. The x_n are called **model parameters**. The y_k, which may be a subset of a larger set of results of the model, are called the **results of interest**.

The purpose of a sensitivity analysis is to determine how much a variation in a model parameter affects a result of interest. Locally this is an estimate of the value of $\partial y_k / \partial x_n$, the **sensitivity coefficient**, but in this work the sensitivity is usually estimated globally using **sensitivity indices**.

The aim of the optimisation process is to minimise the difference between the results of interest y_k and a set of **target results** $Y_k, k = 1, 2, \ldots, K$, by altering the model parameters. The approach is to minimise an **objective function**, typically of one of the following two forms (expressed in terms of absolute and relative differences, respectively):

$$\sum_{k=1}^{K} (Y_k - g_k(\mathbf{x}; \alpha))^2, \tag{2}$$

$$\sum_{k=1}^{K} \left(1 - \frac{g_k(\mathbf{x}; \alpha)}{Y_k}\right)^2. \tag{3}$$

Other forms of objective function can be used, such as weighted sums of squares or the maximum difference between the results of interest and the target results, but only the forms above have been used in the work in this project.

The values of the model parameters that minimise the objective function are called the **optimal solution**. The solution that an algorithm identifies as the optimal solution is called the **converged solution**. The problems reported here have been designed so that the minimum value of the objective function is zero, and the values of the model parameters that produce this minimum are known. This set of model parameters is called the **reference solution**.

The sampling methods used for uncertainty analysis are rated on the accuracy of their estimates of the mean, standard deviation, and distribution function

associated with the results of interest. Assessment of this accuracy requires a **reference distribution**. These distributions have been assembled by carrying out a large-scale calculation using random sampling to generate sets of model parameters.

2.2 Summary of the methodology

The methodology consists of seven main steps:

- Define the problem in general terms, specifying the model parameters, the results of interest, and the reference solution.
- Define a finite element model of the problem, including specification of a mesh, a set of target results, and an objective function.
- Evaluate the model repeatedly to generate a reference distribution and reference values of the sensitivity indices.
- Carry out repeated evaluations of the sensitivity indices using small sample sizes.
- Carry out repeated optimisation runs using different initial guesses of the parameter values.
- Carry out repeated tests of the sampling methods using small sample sizes.
- Compare different sensitivity indices, different optimisation methods, and different sampling methods. Evaluate each based on repeatability, computational efficiency, and quality of results (the closer to the reference solution/distribution the better).

The methodology is described in more detail in section 2 of the companion report [6]. The aim of the methodology is to provide a fair framework for evaluating and comparing different methods.

2.3 Choice and definition of test problems

The test problems reported here have been chosen with the following requirements in mind:

- The problems should be drawn from different areas of physics and should be based on different differential equations.

- The model parameters should include geometric quantities, material properties, and boundary condition data.
- The results of interest for at least one problem should include different types of quantity (e.g. stresses and displacements in the same objective function).
- The problems should be sufficiently small and simple that the corresponding finite element (FE) model runs in less than a minute.
- If possible, the problems should have analytical solutions.

The first three requirements ensure that many of the features of real optimisation problems occur in the test problems. The fourth requirement means that the reference distributions can be generated in a reasonable amount of time, and that the optimisation algorithms can be run repeatedly to test their dependence on the initial estimates of the model parameters. The final requirement makes identification of an appropriate finite element mesh simple. It is easier to judge when the finite element mesh is producing sufficiently accurate results if the true solution to the original problem is known.

All the finite element models have been created using SOFEA [28], a set of Matlab-based finite element routines. The SOFEA package was chosen because the use of Matlab means that pre-processing of the models and post-processing of the results can be automated easily. Matlab includes a range of optimisation algorithm implementations, and the NAG Matlab toolbox expands the range of available algorithms further. The use of a Matlab-based finite element model makes creating the interface between the FE model and the optimisation algorithm straightforward. All the SOFEA models have been validated against solutions to the same problems using other packages (typically Abaqus or Comsol).

All the SOFEA models have been created to have a mesh with a fixed distribution of elements, and a set element geometry. Constraints have been placed on the geometric parameters to ensure that this approach does not lead to excessive element distortion. This approach was used to eliminate any uncertainties associated with the variation in mesh, although the use of a mesh that produces converged results should minimise such problems. The approach is easy to apply for simple geometries, but in real problems it is likely that automatic meshing will be used, which is less easy to control.

2.4 Methods and algorithms used in this work

A more detailed description of the various algorithms and methods used in the work reported here can be found in the companion report [6]. This section of the

report aims to give a brief explanation of the methods and summarise their key features, but the reader is advised to consult the references and the companion report for a more detailed explanation.

Three sensitivity indices were used: Morris one-at-a-time (OAT) indices (see chapter 4 of [21]), Sobol indices, and Jansen indices (chapter 8 of [21]). All three are global sensitivity indices, so the model parameters are varied over the full space of feasible values, and are not an approximation of local behaviour.

The Morris indices give a measure of the impact of changing the value of one factor at a time. The values of the Morris indices depend on the values of the results of interest, making intercomparison tricky, and so the indices have been rescaled to lie between 0 and 1. The same scaling factors have been used throughout for a given output quantity, so the indices for the same result of interest from different sample sizes will be directly comparable.

The Sobol and Jansen indices are variance-based methods. The fullest form of the technique generates a series of indices. Each index approximates the contribution of a given model parameter (first-order index) or a given interaction between multiple model parameters (higher-order index) to the overall variance of the result of interest. The indices used in this report are total sensitivity indices, which are the sum of the first and higher-order indices over all interactions in which a given model parameter participates. Theoretically, Sobol and Jansen indices take values between 0 and 1 and so no scaling is needed before comparison.

The optimisation algorithms can be divided into two main classes. Four of the algorithms are deterministic and use objective function values to determine the next point to be tested. Such algorithms are generally better suited to problems where the objective function is a smooth continuous function with a single unique minimum. Two of the algorithms are probabilistic and use objective function values to identify good solutions and randomness to maintain the diversity of the search.

The deterministic algorithms used are two implementations of sequential quadratic programming (SQP), the Levenburg-Marquardt algorithm (LM) and the Nelder-Mead algorithm (NM).

The Nelder-Mead algorithm [16] is a simplex search algorithm that uses function values alone to create a search, rather than requiring or approximating function derivatives. The implementation used in the work reported here is Matlab's `fminsearch`. This implementation does not handle constraints, and so where necessary the constraints were implemented within the objective function as heavy penalties.

The Levenburg-Marquardt [4, 5] algorithm was used in a form designed to minimise objective functions that have a sum of squares form. The algorithm uses finite differences to approximate local derivatives and uses a damping parameter to produce a search direction that is a blend of the steepest descent and Gauss-Newton directions. The implementation used was the Matlab Optimization Toolbox function `lsqnonlin`.

Sequential quadratic programming methods [10, 11, 12, 19, 20] solve nonlinear constrained optimisation problems. The methods use a set of conditions that are necessary and sufficient for optimality to generate a sequence of quadratic programming subproblems that can be solved more easily. The two implementations used are Matlab's `fmincon` and the NAG Matlab toolbox's `e04us`, which is designed to solve sum of squares problems.

The two probabilistic algorithms were both implemented within Matlab by project team members. The simulated annealing (SA) algorithm [15] is inspired by the annealing process of metals and its progress depends on a cooling schedule. Throughout the work reported here, the geometric cooling schedule $T_n = \alpha^n T_0$ has been used. Two sets of values have been used for α: slow cooling had $\alpha = 0.95 \sim 0.975$ and fast cooling had $\alpha = 0.8 \sim 0.9$.

The particle swarm optimisation (PSO) algorithm [2, 13, 14] is based on "swarm intelligence" seen in nature in large groups of animals. In the implementation used here, the progress of the algorithm is controlled by two learning parameters, set throughout at $\alpha = \beta = 2$, and a parameter γ that controls the randomness such that all random numbers in the nth iteration lie in the interval $[0, \gamma^n]$. Two values have been used for γ: all runs labelled PSO1 in tables have $\gamma = 1$, and all runs labelled PSO2 have $\gamma = 0.9$.

The three sampling methods used were Monte Carlo sampling (MCS, random sampling), Latin Hypercube sampling (LHS), and importance sampling IS). It has been assumed that all model parameters are independent. For a sample size M, MCS draws M random samples from the probability distribution function (pdf) of the model parameters. LHS divides each parameter's distribution into r regions of equal probability, samples M/r times within each region, and combines the samples of each parameter into M sets. IS divides the joint pdf of the parameters (i.e. the product of their individual pdfs) into r regions and samples M/r times within each region. Hence for IS and LHS it is necessary to state r as well as M to define the approach fully. See chapter 6 of [21] for more details of these methods and their assumptions and biases.

3 Problem 1: Elliptical membrane

The first small test problem used for investigations was based on a NAFEMS linear elasticity benchmark (LE1 in [9]). Since this test was the first one undertaken, the results of the test and conclusions drawn from the results were used to refine the investigation procedure for the two subsequent tests. The refinements will be described in the conclusions to this section.

3.1 Problem definition

The FE model simulates an elliptical membrane with a confocal elliptical hole under uniform outward pressure on its outer edge. The geometry and loading of the test problem are as shown in figure 1.

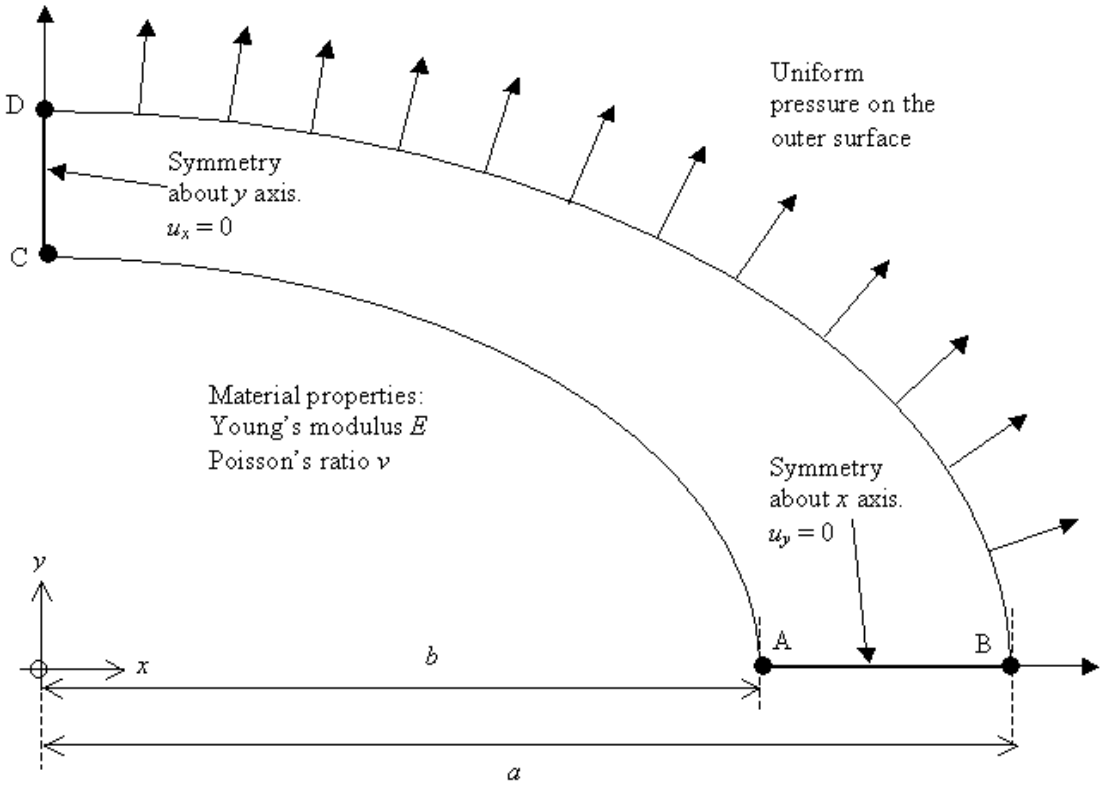

Figure 1: Geometry and loading of the first small test problem. Labelled points are points at which results of interest are calculated.

Four model parameters for the FE model were considered: the Young's modulus E and Poisson's ratio ν of the material, and the horizontal outer and inner semi-axes (a and b respectively, as shown in figure 1) of the membrane. The vertical semi-axes of the ellipses were fixed such that the elliptical edges of the membrane were confocal with foci at $(\pm\sqrt{3}, 0)$ m for all values of a and b. The sensitivity and optimisation tasks were carried out simultaneously (and so the sensitivity of the objective function to the model parameters was not considered or used as a screening method for the parameters) and all four model parameters were varied during the optimisations.

The optimisation test problem was defined to include a constraint since many common design problems include constraints. During the optimisation runs, the model parameters were constrained such that their values lay between the limits given in the fourth column of table 1, with the horizontal inner and outer semi-axes additionally constrained such that $a - b > 0.25$ m. To investigate the use of sampling methods for uncertainty and sensitivity calculations, probability distributions were assigned to each of the four parameters. These distributions are listed in table 1. The distributions for a and b were chosen so that the constraint $a - b > 0.25$ m was always satisfied. Note that the units of each quantity are only given because the original NAFEMS problem specified units: the problem, as with the majority of FE models, has the same solution as long as a consistent set of units is chosen.

Fixed numbers of elements were used in the locally radial and angular directions of the ellipse, independent of the geometric parameters. The mesh was chosen to balance the conflicting needs of achieving convergence of the FE model solution and doing so in a reasonable time. The chosen mesh had 30 linear quadrilateral plane stress elements in the radial direction and 60 in the angular direction, with the result that the stress at the point (2, 0) had converged to within 5 % of its target value as given in the LE1 benchmark, and the average model run time on a desktop PC (dual 1.86 GHz processors, 1 GB RAM) was approximately 20 seconds.

Parameter	Reference solution value	Distribution for sensitivity analysis	Bounds for optimisation
E	210 GPa	Normal, μ = 210 GPa, s = 10.7 GPa	[150, 300] GPa
ν	0.3	Normal, μ = 0.3, s = 0.0765	[0.15, 0.45]
a	3.25 m	Uniform on [2.5, 4] m	[1.75, 6] m
b	2 m	Uniform on [1.75, 2.25] m	[1.75, 6] m

Table 1: Reference solution values, distributions, and bounds for each of the model parameters for the first test problem. Here, μ is the expectation and s is the standard deviation of a normal distribution.

The objective function S was defined as

$$S = \sqrt{\sum_{i=1}^{N} \left(1 - \frac{y_{calc,i}}{y_{targ,i}}\right)^2} \tag{4}$$

where $y_{targ,i}$ is the ith target result and $y_{calc,i}$ is the corresponding value calculated by the model. Comparing this function with the form given in (3), note that $N = K$, $y_{calc,i} = g_k(\mathbf{x}; \alpha)$, $\mathbf{x} = (E, \nu, a, b)^{\mathrm{T}}$, and $y_{targ,i} = Y_k$.

The FE model result quantities used were the stresses and displacements listed in table 2, and the target values for the quantities are also given in the table, where u denotes a displacement, σ denotes a stress, and x and y are the co-ordinate directions. The target values were calculated using the fixed mesh SOFEA model rather than using the solution given in the original definition of LE1 since the full analytical solution to LE1 has not been identified. The optimisation problem was to find the values of the four model parameters that minimise the value of S.

Corner	Quantity	Value	Corner	Quantity	Value
A	u_x	−0.101 626 2 mm	A	σ_{yy}	88.946 207 MPa
B	u_x	−0.073 455 77 mm	B	σ_{yy}	−0.956 161 MPa
C	u_y	0.549 013 64 mm	C	σ_{xx}	−6.548 688 MPa
D	u_y	0.545 749 9 mm	D	σ_{xx}	31.706 638 MPa

Table 2: Target values for the first test problem.

The objective function (4) differs from those that would typically be used in optimisation for design problems, since it has a known unique global minimum of $S = 0$ and a corresponding unique set of "correct solution values" (given in table 1), and the target values in table 2 have been given to a high precision. In a typical design problem the required performance may not be achievable (i.e. the minimum of S may not be zero) and the target values will be given to a lower precision. The function used here, and in all the other small case studies, has been deliberately defined with a known minimum so that optimisation algorithms can be assessed by calculating the percentage of optimisation runs that produce converged values of the model parameters that are within 0.05 % of the correct solution values. The values in table 2 (and in the equivalent tables for the other case studies) are given to a high precision to promote the convergence to the correct solution values for model parameters that only have a small effect on the value of S.

3.2 Sensitivity results

An initial qualitative sensitivity analysis was performed to identify the most important model parameters and to highlight possible interactions between parameters for both the objective function S and the individual results of interest listed in table 2. A quadratic regression analysis identified the geometric model parameters as the most important, while the Young's modulus and Poisson's ratio have relatively little effect on any of the model results. As an example, consider figure 2. The model results for the stress at mesh point A, from a large-scale (i.e. many samples) Monte Carlo simulation, shows a smooth variation with the lengths of the inner and outer semi-axes (figure 2a). In this plot the remaining two model parameters take values drawn randomly from their distributions at each point shown. The quadratic regression model based on data from the large-scale Monte Carlo simulation captures the main features of the data (figure 2b), particularly the lack of sensitivity of the results to the remaining two model parameters. Within this plot the remaining model parameters are fixed at their mean values. The similarities between the two plots illustrate that the effects of the material property parameters is so small that the results obtained by fixing the material properties to their mean values are very similar to those obtained by allowing their values to vary.

Another illustration of the lack of sensitivity to material property parameters is shown in figure 3. This figure shows the results of the quadratic regression model for the stress at mesh point A as a scatter plot. The lack of colour variation along the vertical axis shows that the Poisson's ratio has much less effect on the stress value than the lengths of the semi-axes.

Results for all eight results of interest show a similar level of agreement between a quadratic regression model and data, a smooth response surface, and also a similar dependence on the lengths of the inner and outer semi-axes. For large values of the outer semi-axis (typically greater than the mean value), the value of the inner semi-axis has little effect on the model results. When the outer semi-axis is smaller, the inner length becomes more important. The most extreme result values are always observed under a relatively small outer- and large inner semi-axis, creating a thin ellipse. The stress at mesh point A exceeds 600 MPa for this case of a thin ellipse.

Values of the objective function, calculated over all points of the Monte Carlo simulation, demonstrate the lack of sensitivity of the objective function to the parameters describing material properties (figure 4). The objective function appears to be smooth over the sample space, including within the region of the target value (figure 4b). At the plotted scale, the function appears not have local minima, which will affect the performance of the various optimisation algorithms.

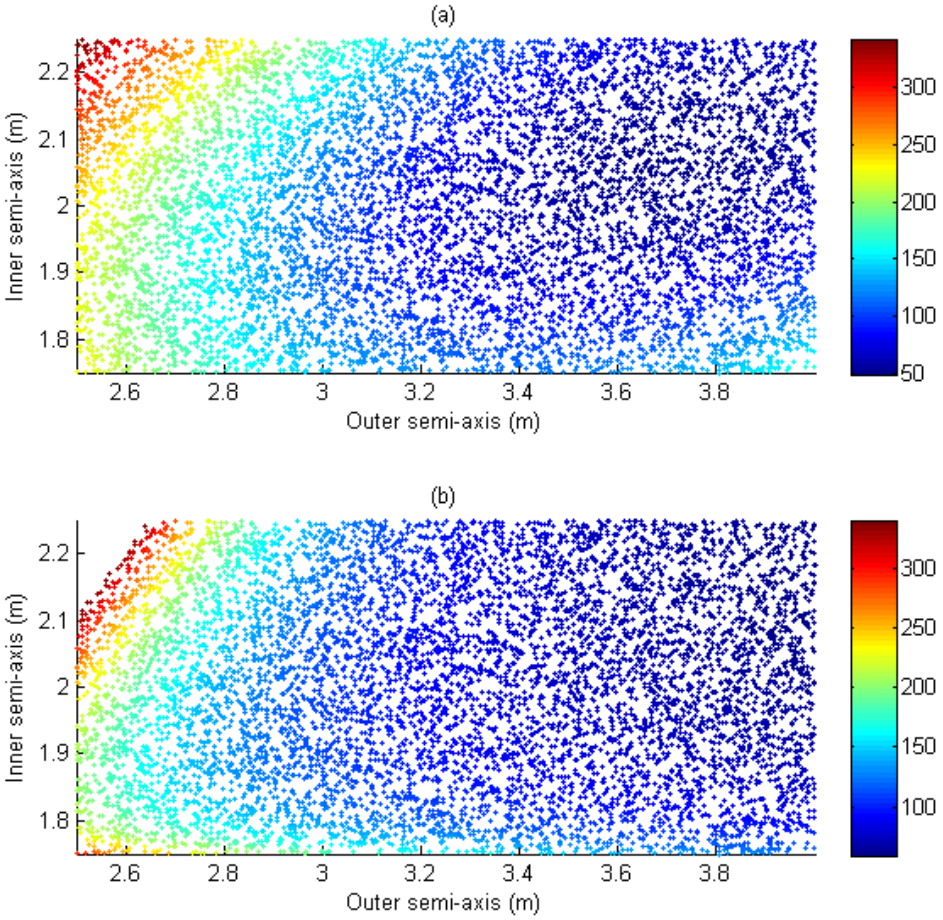

Figure 2: Monte Carlo output (a) and quadratic regression model (b) for stress at mesh point A (MPa). The extreme values of stress (> 600 MPa) that occur for a thin ellipse have been omitted to allow an easier comparison of the two plots.

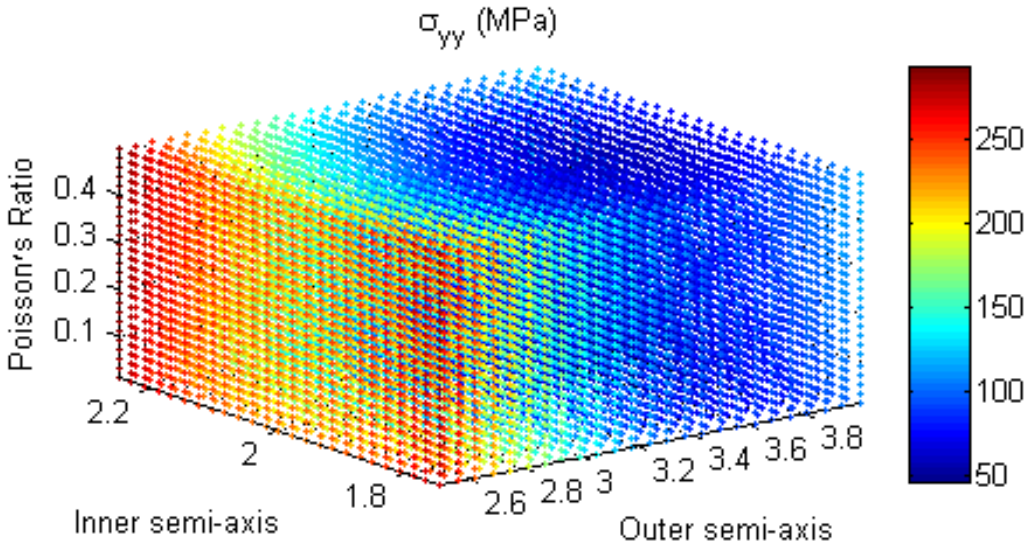

Figure 3: Output of the quadratic regression model illustrating the lack of variation of the stress at point A with Poisson's ratio. Other stresses and displacements show a similar lack of dependency on the material property parameters.

A large-scale test (i.e. many samples) for each of the three sensitivity methods was run to compare the methods and to provide a reference set of sensitivity indices. These tests were based on a sample size of around 10^4. Smaller-scale tests based on fewer samples were run multiple times to compare the quality of results against the reference indices and also to assess the variability of the indices generated from an identical sample size. The results of these tests indicate how many runs are required to identify the most important model parameters.

3.2.1 Sobol sensitivity measure

The sensitivity indices calculated from the large-scale Sobol test were consistent with qualitative results from the Monte Carlo simulation (figure 5). The values of all indices related to the Young's modulus and the Poisson's ratio were less than 10^{-5}. These values show that variation of these two model parameters has very little effect on the model results. The sensitivity indices for the inner and outer semi-axes imply the greater dependence of results is on the outer semi-axis (> 0.8 for all results), and that this dependence is particularly high for stress at the inner mesh points A and C (0.93 and 0.92 respectively). The indices for the inner semi-axis lie in the interval 0.2 - 0.6, and are greater for displacement than for

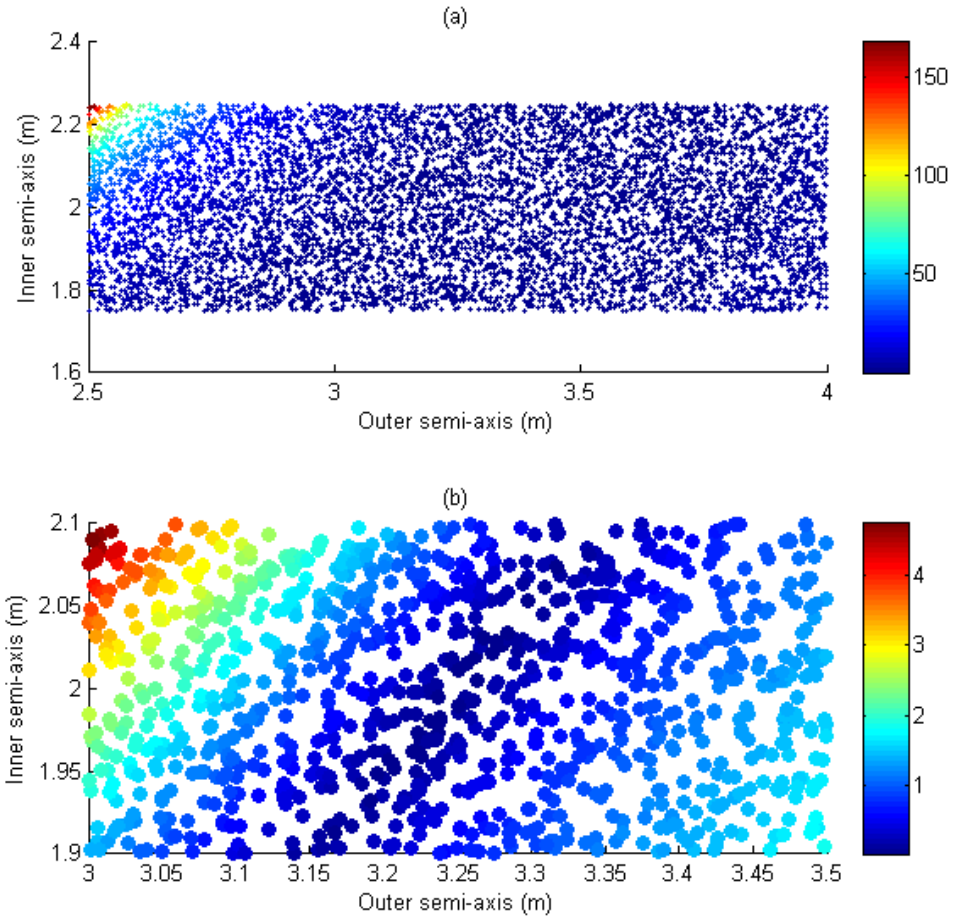

Figure 4: Objective function calculated from results of the Monte Carlo simulation, including (a) all data points and (b) those produced from inner and outer semi-axes within 0.1 m and 0.25 m of their mean target values, respectively.

stress. Although the indices should strictly lie between 0 and 1, for insufficiently large sample size or a parameter with very little influence, the method can give a negative value. This behaviour is observed in the Sobol test on the elliptical membrane problem due to the negligible influence of Young's modulus on the results, particularly displacement (figure 5). Figure 5 shows the results of the large-scale test for the stress and displacement at point A. The results at points B, C and D showed similar dependency on the model parameters.

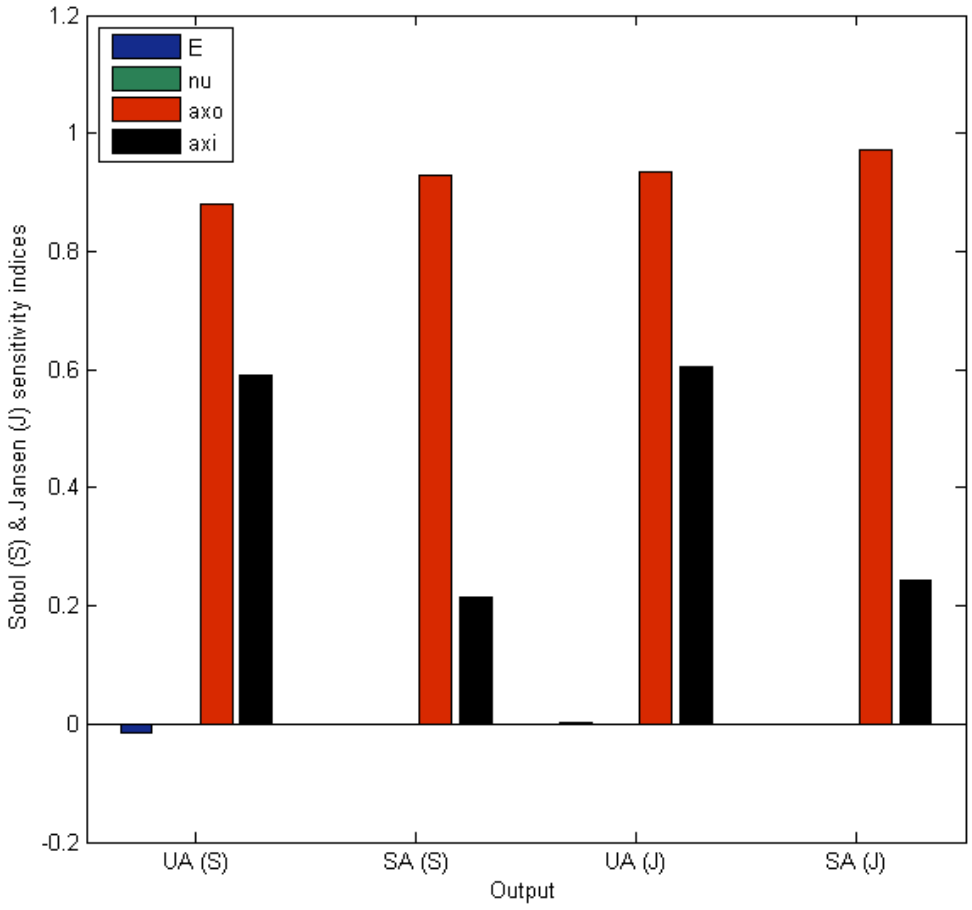

Figure 5: Sobol (S) and Jansen (J) sensitivity indices for each of the four model parameters of the elliptical membrane problem. Model results shown are the displacement at mesh point A (UA) and stress as mesh point A (SA). Stresses and displacements at other points showed similar behaviour.

3.2.2 Jansen sensitivity measure

The Jansen indices from a large-scale test are again consistent with the previous sensitivity analysis (figure 5). Results obtained at point A are shown, and those at the other points showed similar dependencies. The results are consistently greater than the Sobol indices. However, the indices vary in proportion for the two methods. For example, similarly to the Sobol method, the Jansen index for the outer semi-axis is the largest over all results, with a maximum value for the stress at mesh point A. The inner semi-axis has larger index values for displacement than for stress, with a minimum for the stress at mesh point A.

The total sensitivity index is calculated as a ratio of the mean of a squared difference to a variance and therefore, unlike the Sobol measure, cannot take negative values.

3.2.3 Morris OAT measure

Calculations of the mean absolute value and the standard deviation for elementary effects for the displacement and stress results are consistent with the Monte Carlo sensitivity analysis and the previous sensitivity measures (figure 6). The outer semi-axis is the most important parameter for stress and displacement at all four mesh points, with the two material property parameters consistently the least important. Similar values for the standard deviations for each result's response to the inner and outer semi-axes (figure 6) suggests some degree of interaction and co-dependence on these two model parameters.

3.2.4 Sample size analysis

The three sensitivity measures were calculated for sample sizes of 10, 25, 100 and 250. A mean and standard deviation of the sensitivity of each model result to the model parameters for each of the sensitivity measures was calculated over ten runs with each sample size. These means and standard deviations are plotted against sample size in figure 7.

As described in Section 3.2.1, the Sobol sensitivity measure can take negative values. The Sobol and Jansen measures are based on mean values for the product and difference respectively of model results for a change to a single model parameter value. With a small sample size, a single value can have a large effect on these mean values and consequently result in values of the sensitivity measures that lie outside the interval [0,1]. With the smallest sample size, 10 % of the Sobol indices and 30 % of the Jansen indices took values greater than one, with values on

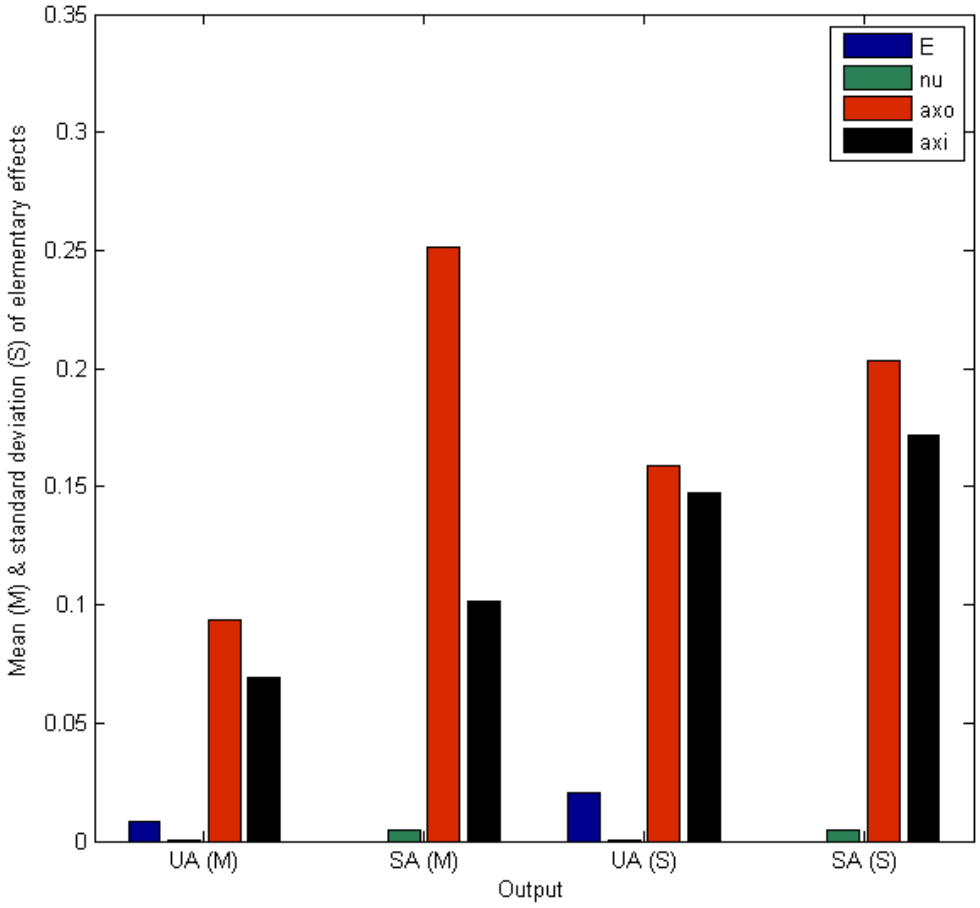

Figure 6: Mean (M) and standard deviation (S) of elementary effects calculated for the Morris OAT sensitivity measure, for each of the four model parameters of the elliptical membrane problem. Model results shown are the displacement at mesh point A (UA) and the stress at mesh point A (SA). Stresses and displacements at other points showed similar behaviour.

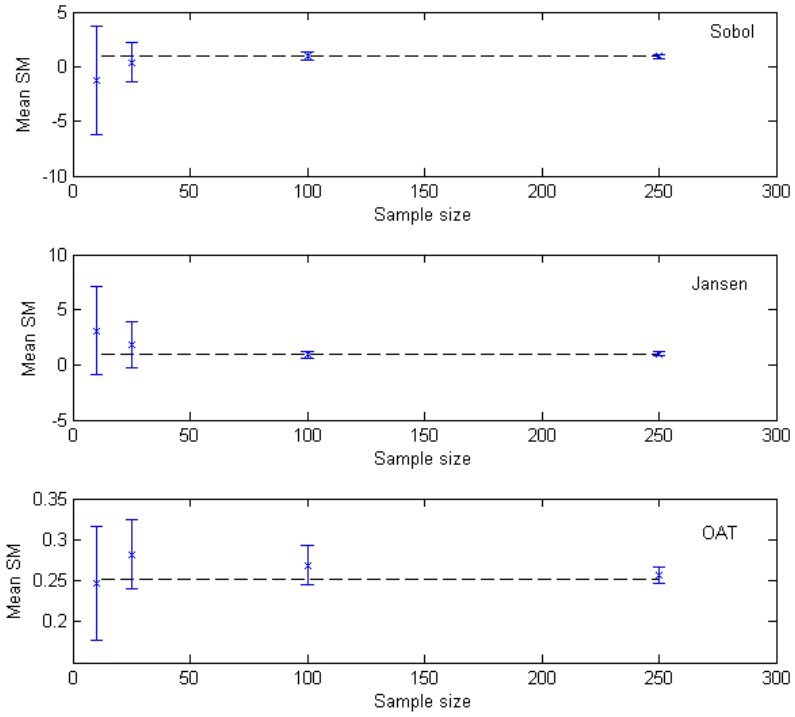

Figure 7: Mean Sobol, Jansen and Morris OAT sensitivity indices for the sensitivity of the stress at mesh point A to the outer semi-axis. The means are calculated over 10 runs with sample sizes of 10, 25, 100 and 250. Error bars indicate one standard deviation.

the order of 100 occuring in some cases. A further 40 % of the Sobol indices took negative values. With the largest sample size, 5 % and 10 % of the Sobol and Jansen indices respectively still took values greater than one. However, these were restricted to the cases shown to be most sensitive by the large-scale runs, such as the sensitivity of stress to the the outer semi-axis, and were generally very close to one ($1 < \mathrm{TS} < 2$, where TS is the total sensitivity index). Less than 1 % of the Sobol indices that did not take a negative or zero value in the large-scale run produced a negative value when a sample size of 250 was used.

The occurence of negative indices and values greater than one for the Sobol and Jansen measures lead to a significantly larger standard deviation of indices across the 10 runs than that for the Morris OAT indices. The standard deviation for the Morris OAT measure remains smallest of the three tests for the larger sample sizes. The Morris OAT indices generally provide the same qualitative information as the reference indices even with the smallest sample size (figure 8).

A very small number of model runs is sufficient to identify the most important model parameters and the relative effect of these parameters on the model results. The Morris OAT indices give the correct ordering of parameters for a sample size of 10. While the Sobol and Jansen measures also often put the parameters in the correct order of importance for small sample sizes, the occurence of values outside the interval [0,1] make interpretation of these indices more difficult.

3.3 Optimisation results

Initially all four model parameters were used in the optimisation because the optimisation and sensitivity analyses were carried out simultaneously, and so no information from the sensitivity analysis could be used to guide the optimisation work.

All algorithms were run 50 times starting from randomly-generated initial estimates of teh model parameters. The SA algorithm used two separate values of α as stated in section 2.4.

It should be noted that a key parameter affecting the success rate is the comparative magnitudes of the model parameters. Some of the algorithms that were investigated do not consider the magnitude of a quantity when calculating what the next estimate of its value should be. In particular, some algorithms that attempt to approximate gradients using finite differences do not seem to take the magnitude of the parameter into account when determining the step size, so a parameter taking a value 10^6 may only be perturbed by a value of 10^{-6}, which is very unlikely to affect the function value.

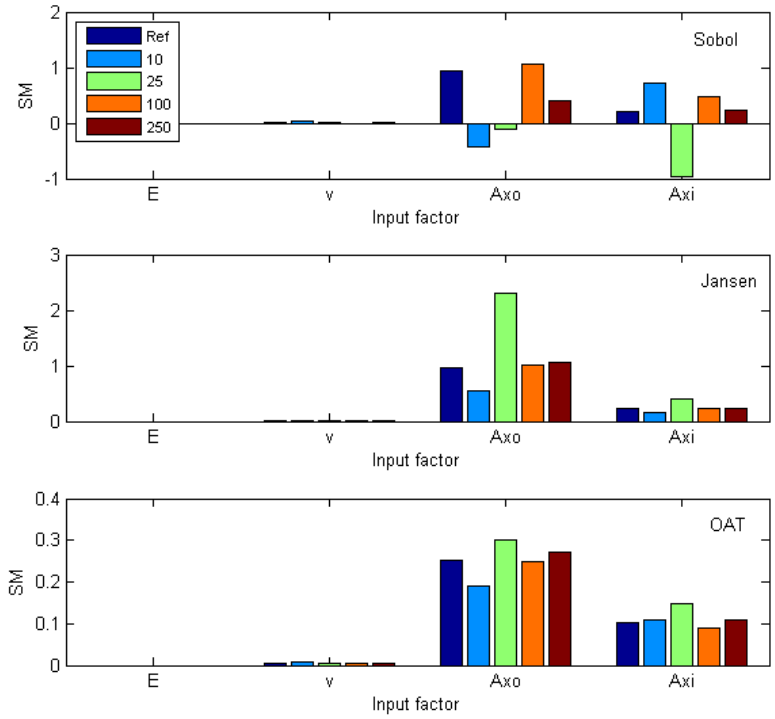

Figure 8: Sobol, Jansen and Morris OAT sensitivity measures for the sensitivity of stress at mesh point A to Young's Modulus (model parameter 1), Poisson's ratio (2), and the outer (3) and inner (4) semi-axes. The sensitivity is calculated from a single run with each sample size 10, 25, 100 and 250. Dashed lines indicate the value of the appropriate index obtained from the large-scale test.

This lack of normalisation within the algorithms led to a redefinition of the problems such that all parameters were scaled to have values $O(1)$, so for instance E was treated by the optimisation algorithms as though its reference value was 2.1 and was scaled so that the FE programme received the correct value of 210 GPa. This approach improved the success rate of several of the algorithms.

Table 3 compares the results of the optimisations using the different methods. The percentage success is the percentage of converged solutions equal to the reference solution. The mean and standard deviation of the number of objective function evaluations are taken across all runs including those that did not converge to the correct solution (note that often convergence to an incorrect solution occurred more quickly than convergence to the reference solution).

Optimisation method	Mean of the number of function evaluations	St dev of the number of function evaluations	Percentage success
SA (slow)	3375	1237	100
SA (fast)	869	406	100
PSO (1)	2140	821	100
PSO (2)	1973	819	100
NM	410	115	60
SQP1	233	80	100
SQP2	214	40	100
LM	59	14	100

Table 3: Results of the optimisation runs using all four model parameters.

It is clear from table 3 that the metaheuristic algorithms were significantly more computationally expensive than the traditional optimisatin methods. It is likely that the unimodality of the problem gives the traditional optimisation algorithms an advantage over the metaheuristic algorithms. Figure 4 showed the smoothness of the objective function and the uniqueness of its minimum. Metaheuristic algorithms tend to outperform traditional algorithms on more challenging problems such as multimodal problems or those featuring discontinuities in the objective function. It is also interesting to note that the standard PSO performs better than standard SA, while SA with fast cooling rate outperforms the improved PSO. These results may also be affected by the unimodal nature of the objective function.

The smoothness and unimodality of the objective function has led to the success of the Levenberg-Marquardt algorithm. This algorithm is designed for solving minimisation problems that feature sums of squares, such as the form given in (4), and to some extent the method relies on the existence of a continuous gradient, i.e. a smooth function. The method is described in full in the references [4, 5] and in the companion to this report [6] but it is similar to the steepest descent method.

The steepest descent method uses a finite difference approximation to the gradients if they are not available analytically, and evaluates the next estimate along the line of steepest descent. This method works best for smooth functions with a unique minimum.

The performance of the sequential quadratic programming (SQP) algorithm is surprising since it is also designed for "sum of squares" problems. A detailed examination of the values of the model parameters at which the function is evaluated shows a surprisingly large amount of repetition for SQP1 in particular. An example taken from the first run using this method is shown in figure 9. This figure shows the amount by which each model parameter was perturbed from its initial value for the first 97 function evaluations on a logarithmic scale (note that any value plotted as 10^{-6} is actually zero). The figure shows that one parameter is varied at a time, and that the magnitude of variation of each parameter is the same. This leads to unnecessary repetition. Of the 97 runs shown here, only 17 distinct sets of values of the model parameters are used. The initial stages of other runs looked similar. In addition, the SQP1 algorithm did not always appear to start its calculations from the requested starting estimate. Sometimes model parameters would be started from the upper or lower bounds as given in table 1 rather than the randomly-chosen initial estimate. There was no clear pattern to the occurrences of this event, and no obvious cause.

It is not clear why the success rate of the Nelder-Mead algorithm is so low. There is no clear link between the initial estimate and the success or failure of the run, although many of the final wrong solutions lie along a straight line in the (a, b) plane. A detailed examination of these solutions may lead to more information: it may be that they are local minima at a level not identified in the previous plots of the objective function surface, but this has not been checked. It is not clear why the other algorithms do not exhibit simlar behaviour.

It was showed in section 3.2 that the geometric parameters have much more of an effect on the objective function value than the material properties. This result means that it is important to determine the geometric parameters accurately, and that the material properties can be neglected. Hence the optimisation runs were repeated using only the geometric parameters a and b to see how focussing on the most important parameters affected the number of function evaluations and the success rate. The results of the optimisation runs with two parameters are shown in table 4.

The numbers of function evaluations required for convergence are significantly smaller for the runs only using a and b. There are two reasons for this reduction. One is that a search space of two parameters is smaller than one of four parameters, and in particular any algorithm approximating derivatives numerically will have to make half as many derivative calculations as as result. The other

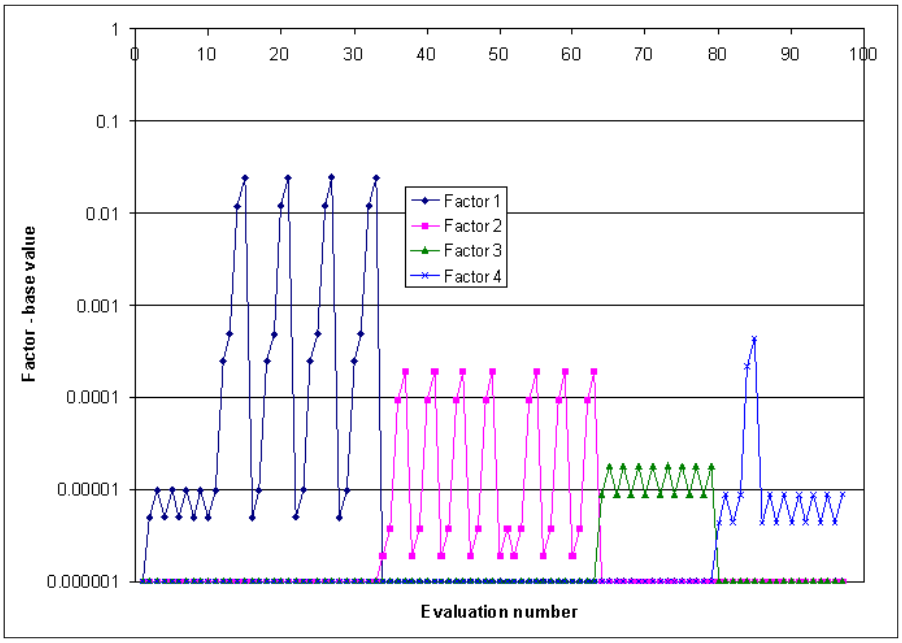

Figure 9: Absolute value of the deviation of the model parameters of the first 97 function evaluations in a typical optimisation run using the SQP algorithm from their baseline values. Values are plotted on a logarithmic scale and values plotted as 10^{-6} are actually zero.

Optimisation method	Mean of the number of function evaluations	St dev of the number of function evaluations	Percentage success
SA (slow)	1729	618	100
SA (fast)	492	205	100
PSO (1)	1619	725	100
PSO (2)	1054	712	100
NM	94	29	98
SQP1	120	40	100
SQP2	112	15	100
LM	30	8	100

Table 4: Results of the optimisation runs using the two geometric parameters.

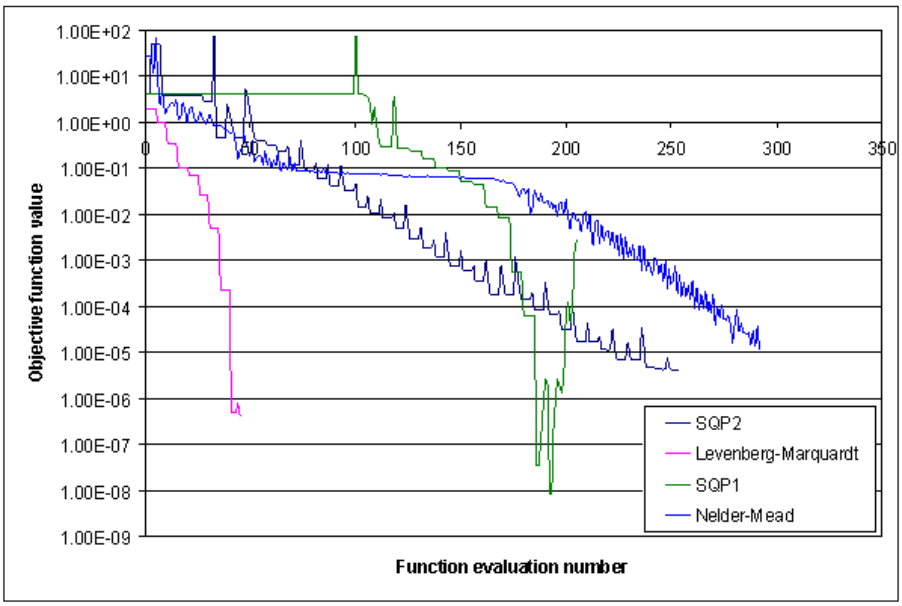

Figure 10: Objective function value versus function evaluation number for the traditional optimisation algorithms

reason is that the elimination of the less important parameters means that the algorithms are less likely to get stuck in regions where a parameter having little effect on the objective function makes the minimum difficult to identify.

Figure 10 illustrates the convergence behaviour of the traditional methods. The figure plots objective function value against function evaluation number. The figure shows the value of every function evaluation, so not all values are improvements on previous values. The values were generated during an optimisation varying all four model parameters.

There are a number of points to note from this figure. The Levenberg-Marquardt algorithm converges quickly, and its function evaluations occur in groups whose values are almost equal (indistinguishable at the scale of the graph). This effect occurs because the method evaluates a point, perturbs slightly to estimate derivatives in the region of the point, then repeats this procedure at a new point. The SQP1 algorithm shows no visible difference over its first 100 evaluations due to the repetition effect mentioned above, but following the repetitive section it converges more quickly than the other SQP implementation (SQP2), which progresses in steps similar to those of the Levenburg-Marquardt algorithm.

Figure 11 shows the objective function value against function evaluation number for the meta-heuristic algorithms. Other studies in the literature [17] have shown

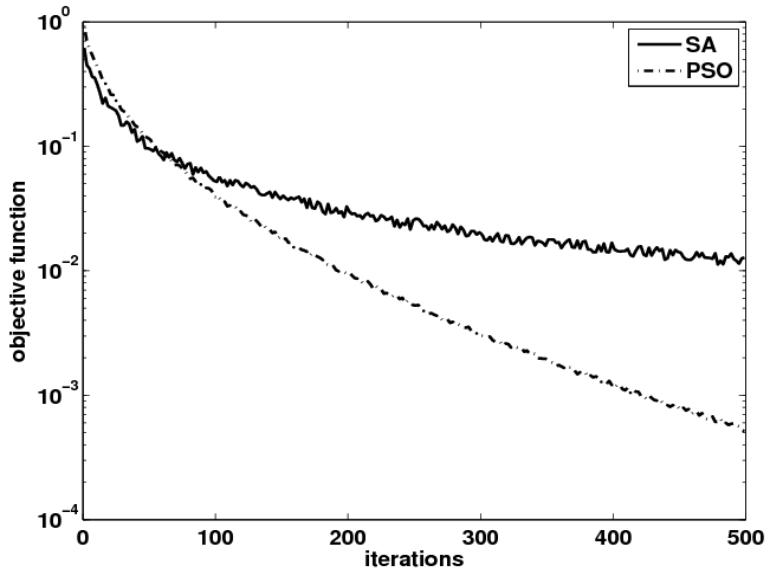

Figure 11: As fifure 10, but for the metaheuristic algorithms SA and PSO.

that SA is generally a quick starter but converges slowly, while PSO is a slow starter and converges more quickly when approaching the optimum. This pattern is reflected by the results shown in figure 11.

3.4 Performance of sampling methods

The large-scale Monte Carlo run of the elliptical membrane test problem can be used to assess the quality of results produced from simulations on much smaller sample sizes, and for different sampling methods.

For the elliptical membrane problem sample sizes of 16, 81, 256, 625 and 1296 were used. These values are all fourth powers, meaning that exactly one sample could be taken within each subdivision for the Latin Hypercube and importance sampling methods. These values enabled a simple, objective method for dividing the sample space equally for importance sampling because they are all fourth powers and so the cumulative density function (cdf) of each model parameter could have its axis divided into regions of equal probability and the sampling regions could be defined as hypercubes formed by the product of one subdivision from each axis. For each of the methods and sample sizes, the mean and standard deviation of the values of each quantity of interest are compared with the reference results obtained from the large-scale Monte Carlo simulation. To assess repeatability, ten runs were carried out for each method and sample size.

3.4.1 Accuracy

The Latin hypercube method generally produces the best approximation to the reference solution, in terms of mean and standard deviation, at the smallest sample size. For example, for a typical run with the smallest sample size of 16, the Latin hypercube method produces a discrepancy of 1.6 % from the reference value for the mean stress at mesh point A (figure 12). Random and importance samples of the same size produce mean values within 6.5 % and 8.1 % of the reference value respectively. The Latin hypercube and importance sampling methods provide a mean value within 1 % of the reference value for sample sizes of 256 and greater. The Latin hypercube method also provides the best approximation to the reference standard deviation at the smallest sample size, giving a value within 4.5 % of the reference compared to 49.3 % and 76.0 % for the Monte Carlo and importance methods respectively (figure 13).

For small sample sizes, Latin hypercube and importance sampling provide a more representative cover of the sampling space (and are therefore more likely to produce outlying values from long-tailed distributions than random sampling). As a result, distributions of values produced by Latin hypercube and importance samples tend to give a better approximation to the respective reference distribution at small sample sizes (e.g. figure 14).

3.4.2 Repeatability

For small sample sizes (sample size = 16, 81), the variance of the mean values produced for each result, over ten runs, is consistently smallest for the Latin Hypercube method. The standard deviation of the mean stress at each of the mesh points is almost 50 % less for the Latin Hypercube runs than the Monte Carlo simulations with a sample size of 16. The Latin Hypercube sampling also gives the smallest variance of standard deviation of the stress at each mesh point for a sample size of 16. The variance of standard deviation of displacements is similar over all three sampling methods at the minimum sample size.

The variance of the mean and standard deviation of the model results decreases very quickly with an increase in sample size (figures 15 and 16). For example, the standard deviation of the mean stress at mesh point A decreases by 62 %, 75 %, and 72 % with an increase in sample size from 16 to 81 for random, Latin Hypercube and importance samples respectively.

Plotting the cumulative probability for a result provides an additional comparison of the three different sampling methods. For example, for a sample size of 81, the Latin Hypercube runs show a tighter fit to the reference probability distribution

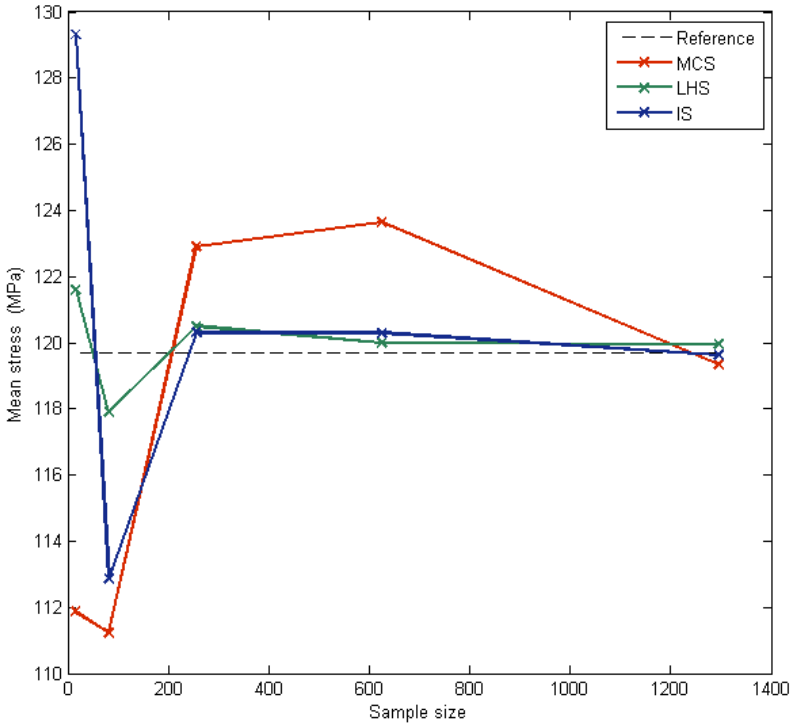

Figure 12: Mean value of stress at mesh point A for single runs with sample sizes 16, 81, 256, 625 and 1296 using Monte Carlo (MCS), Latin Hypercube (LHS) and importance (IS) sampling. The dashed line shows the reference value calculated from the large-scale Monte Carlo simulation. The solid lines are added to provide clearer presentation.

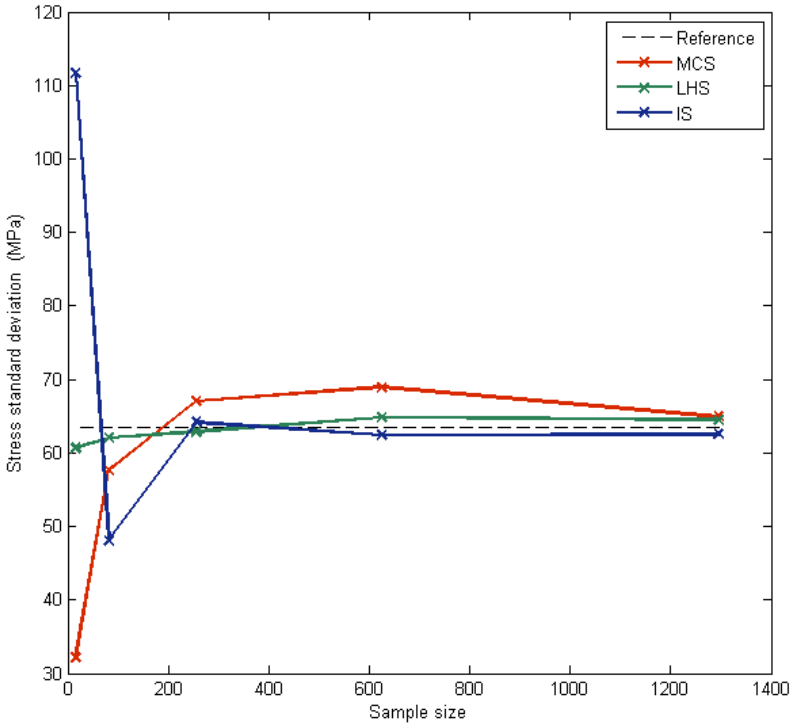

Figure 13: As figure 12, but for the standard deviation of stress at mesh point A.

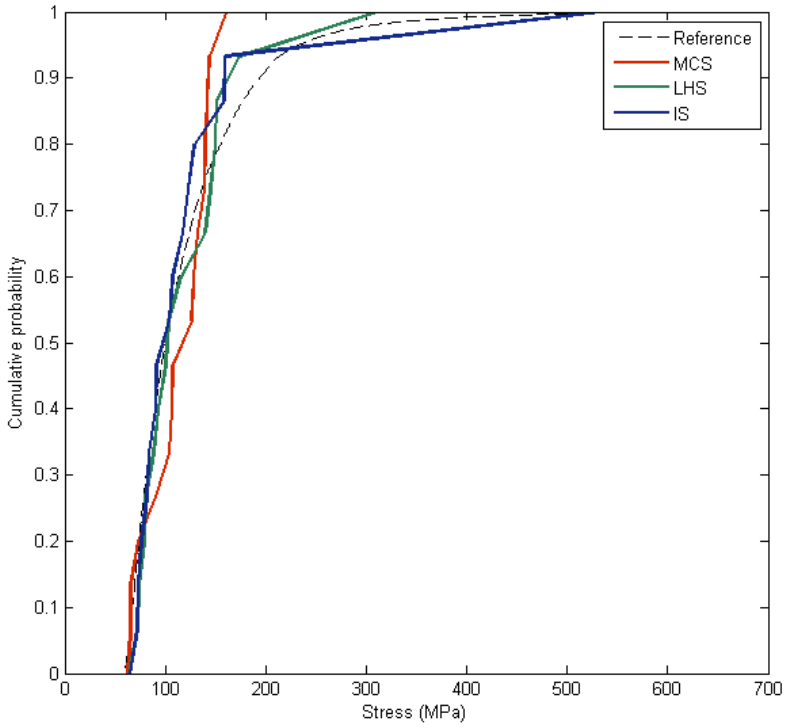

Figure 14: Cumulative probability of the stress at mesh point A, provided by a sample size of 16 for the Monte Carlo, Latin Hypercube and importance sampling methods. The dashed curve shows the reference distribution from the large scale Monte Carlo simulation.

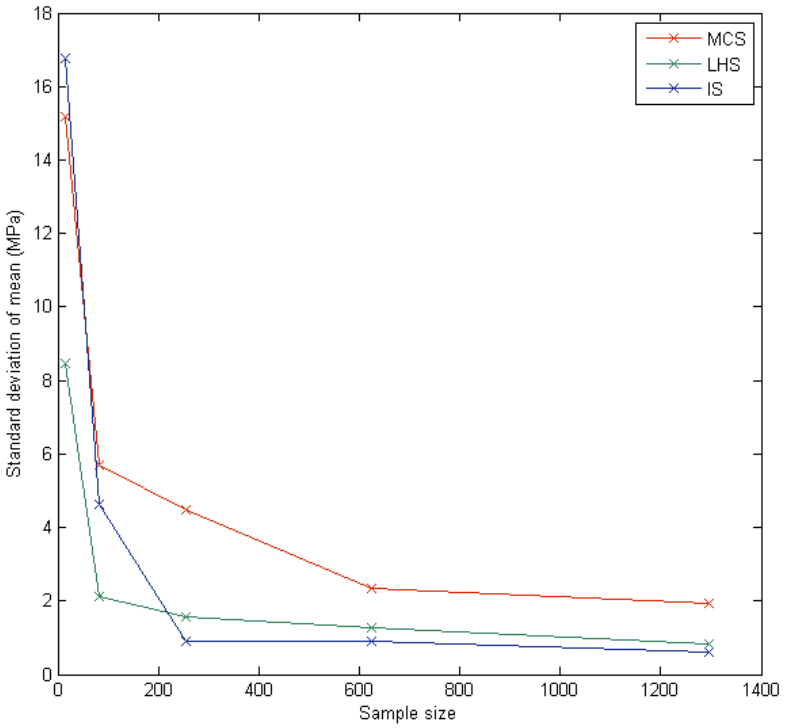

Figure 15: Standard deviation, over 10 runs, of the mean values of stress at mesh point A, for increasing sample size of random (MCS), Latin Hypercube (LHS) and importance (IS) sampling methods.

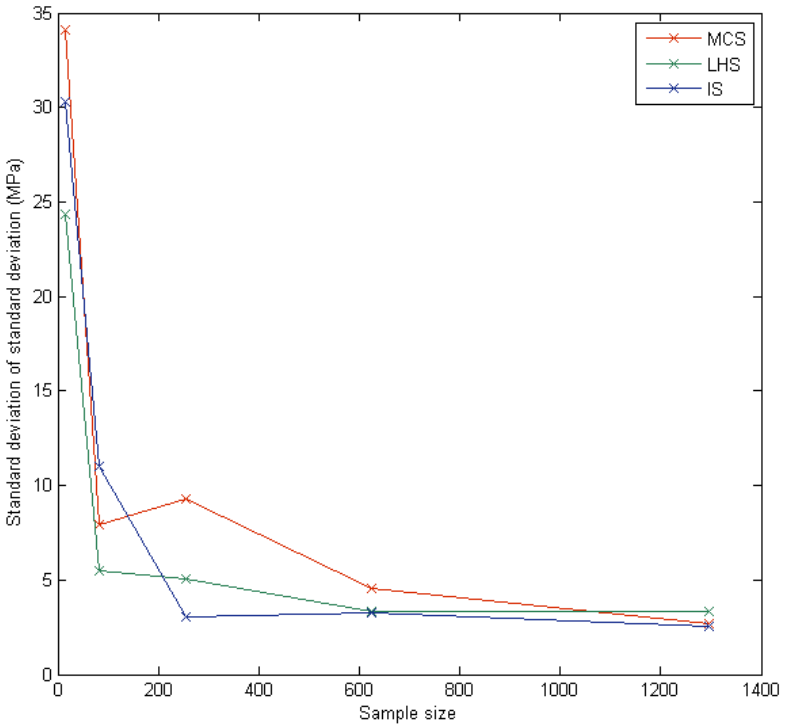

Figure 16: Standard deviation, over 10 runs, of the standard deviations of stress at mesh point A, for increasing sample size of random (MCS), Latin Hypercube (LHS) and importance (IS) sampling methods.

(black line) of stress at mesh point A than the Monte Carlo runs and the importance sampling runs (figure 17).

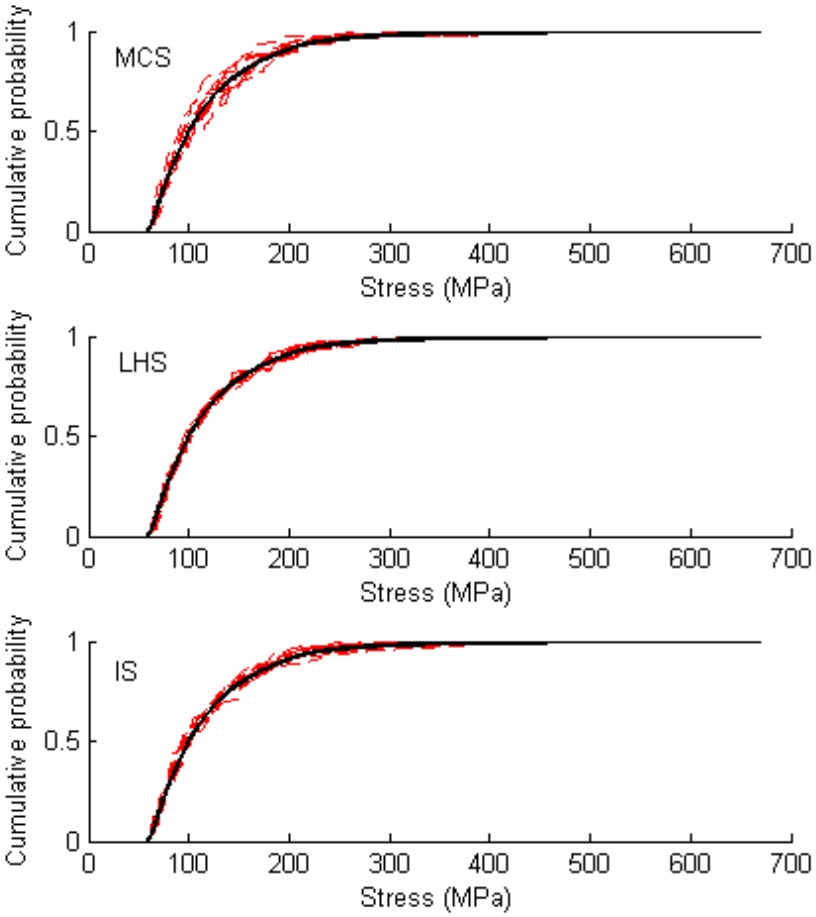

Figure 17: Cumulative probability of stress at mesh point A for each of the ten runs, with sample size 81, for random (MCS), Latin Hypercube (LHS) and importance (IS) sampling.

3.5 Conclusions

- Carrying out a sensitivity analysis provides valuable information that can be used in subsequent optimisation analyses. The results allow the user to choose a subset of model parameters, which decreases the run time of the optimisation. The results also provide information about the nature of the

objective function, which can help the user to choose the most suitable optimisation algorithm for the problem.

- Sensitivity indices can be used to identify the most important parameters affecting the value of a quantity. The indices examined here gave consistent results and, in particular, the Morris OAT indices gave valid results for quite small sample sizes.
- Normalisation of model parameters so that all parameters are approximately $O(1)$ ensures that all parameters are treated the same way even when the optimisation algorithm has been implemented poorly. This approach can lead to particular improvements if the optimisation algorithm uses finite differences to approximate gradients.
- For the objective function used in this test problem, the traditional optimisation algorithms have shown better convergence behaviour than the metaheuristic optimisation algorithms. This difference is likely to be due to the smooth unimodal nature of the objective functions.
- The Latin Hypercube and importance sampling methods appear to give better accuracy and better repeatability than Monte Carlo sampling for small sample sizes when determining mean values and standard deviations of the distribution functions associated with model results.

4 Problem 2: Vibration of a thin disk

4.1 Problem definition

The second test problem simulates the natural vibrational modes of an elastic disk fully fixed at its circumference. Initially, five model parameters were considered for the model: a, the radius of the disk, t, the thickness of the disk, E, the Young's modulus of the elastic material, ν, the Poisson's ratio of the elastic material, and ρ, the density of the elastic material. These quantities were chosen because the analytical solution to the problem involves all five parameters.

In the derivation of the analytical solution, it is assumed that the stresses σ_{zz}, σ_{rz} and $\sigma_{\theta z}$ are all zero, where z is the through-thickness direction with $z = 0$ being the base of the disk, r is a raidal coordinate with $r = 0$ lying at the centre of the disk, and θ being an angular coordinate. Note that since the disk is uniform it is impossible to fix $\theta = 0$, and this has been considered during the subsequent work by not considering results that are dependent on the angular coordinate. The assumptions about the stresses are valid provided that t/a is small. The analytical expression for the fundamental frequencies of the disk is

$$f_{mn} = \frac{\lambda_{m,n}^2 t}{2\pi a^2}\left(\frac{E}{12\rho(1-\nu^2)}\right)^{1/2}, \quad m, n, = 0, 1, 2, \ldots \tag{5}$$

where $\lambda_{m,n}$ is the mth solution to

$$J_n(\lambda_{m,n})I_{n+1}(\lambda_{m,n}) + J_{n+1}(\lambda_{m,n})I_n(\lambda_{m,n}) = 0, \tag{6}$$

where J_n is the Bessel function of the first kind of order n, and I_n is the modified Bessel function of the first kind of order n. The stresses and displacements at each frequency can also be calculated analytically (up to a constant of proportionality). It should be noted that f_{mn} is proportional to the dimensionless quantity $\sqrt{Et^2/(a^4\rho(1-\nu^2))}$, and so any optimisation based on using the frequencies as target results will determine a set of parameter values that have the correct value for this quantity, but will not necessarily be able to identify the reference solution values for the individual parameters. Hence extra information must be included in the target results to achieve a unique reference solution.

Two stresses and a displacement were chosen as the required extra information to enable an optimisation to determine unique values for the model parameters. The stresses were σ_{xx} at $r = 0$ for the first frequency and σ_{xy} at $r = 0$ for the fourth frequency, and the displacement was $\sqrt{u_x^2 + u_y^2}$ for the second frequency. The analytical expressions for these quantities are:

$$\sigma_{xx} = B_1\frac{Et(1+C_{1,0})\lambda_{1,0}^2}{4a^2(1+\nu)(1-2\nu)} \tag{7}$$

$$\sqrt{u_x^2 + u_y^2} = B_2 \frac{t(1 - C_{1,1}\lambda_{1,1})}{2a} \quad (8)$$

$$\sigma_{xy} = B_3 \frac{Et\lambda_{1,2}^2(1 - C_{2,1})}{8a^2(1 + \nu)} \cos 2\theta \sin 2\theta \quad (9)$$

where $B_k, k = 1, 2, 3$, is a scaling constant and $C_{m,n} = J_n(\lambda_{m,n})/I_n(\lambda_{m,n})$. It should be noted that the expression for σ_{xy} is ill-defined because it has a dependence on angle, and angle is ill-defined at the origin. The value was used in the optimisation because the FE model gave a well-defined, repeatable result that, when included in the objective function, enabled optimisation algorithms to find the reference solution.

The analytical solution was used to calculate sensitivity coefficients that could be compared to the sensitivity measures calculated numerically. Since the Morris OAT sensitivity measure is based on normalised model parameters, the derivatives of the results with respect to a normalised version of each model parameter was calculated (denoted by a hat over the relevant parameter). The normalisation constant in each case (denoted by a bar over the relevant parameter) was taken to be the mean value of each of the distributions given in table 7. The formulae for the sensitivity coefficients are given in table 5.

Result of interest	Sensitivity to normalised parameter				
	$\hat{E}$	$\hat{\nu}$	$\hat{\rho}$	$\hat{a}$	$\hat{t}$
f_i	$f_i/(2\hat{E})$	$f_i \frac{\hat{\nu}\bar{\nu}^2}{(1-\bar{\nu}^2\hat{\nu}^2)}$	$-f_i/(2\hat{\rho})$	$-2f_i/\hat{a}$	$f_i/\hat{t}$
$\sqrt{u_x^2 + u_y^2}$	0	0	0	$-\sqrt{u_x^2 + u_y^2}/\hat{a}$	$\sqrt{u_x^2 + u_y^2}/\hat{t}$
σ_{xx}	$\sigma_{xx}/\hat{E}$	$\sigma_{xx} \frac{(\bar{\nu}+4\hat{\nu}\bar{\nu}^2)}{((1+\bar{\nu}\hat{\nu})(1-2\bar{\nu}\hat{\nu}))}$	0	$-2\sigma_{xx}/\hat{a}$	$\sigma_{xx}/\hat{t}$
σ_{xy}	$\sigma_{xy}/\hat{E}$	$-\sigma_{xy} \frac{\bar{\nu}}{(1+\bar{\nu}\hat{\nu})}$	0	$-2\sigma_{xy}/\hat{a}$	$\sigma_{xy}/\hat{t}$

Table 5: Analytical sensitivity coefficients with respect to normalised model parameters.

The formulae in table 5 are evaluated in full for the mean values of the model parameters in table 6. The values suggest that for all the result quantities, the radius is the most important model parameter (the negative sign indicates that a larger value of a leads to a lower value of the result quantity).

The mesh for the disk consisted of quadrilateral tetrahedra with 11 elements in the radial direction and one element through the thickness direction. The analytical expressions for the stresses and the displacements are either independent of z or are linearly dependent on z, and so a single quadratic element should be sufficient to capture through-thickness variations correctly. The meshing routine was altered so that all disks had the same mesh, resized according to the values of the

Result of interest	Sensitivity to normalised parameter				
	$\hat{E}$	$\hat{\nu}$	$\hat{\rho}$	$\hat{a}$	$\hat{t}$
f_i	$f_i/2$	$0.1f_i$	$-f_i/2$	$-2f_i$	f_i
$\sqrt{u_x^2+u_y^2}$	0	0	0	$-\sqrt{u_x^2+u_y^2}$	$\sqrt{u_x^2+u_y^2}$
σ_{xx}	σ_{xx}	$1.27\sigma_{xx}$	0	$-2\sigma_{xx}$	σ_{xx}
σ_{xy}	σ_{xy}	$-0.23\sigma_{xy}$	0	$-2\sigma_{xy}$	σ_{xy}

Table 6: Formulae given in table 5 evaluated at the mean values of the normalised model parameters.

parameters a and t. This approach was taken in order to ensure that the mesh was always well-defined and that there was always a node at the centre.

The reference solution values and associated distributions used for the initial sensitivity analysis are listed in table 7. Bounds for use in optimisation runs are given for all parameters, although not all are necessarily used in optimisation runs. The effects of constraining the parameters were examined by carrying out one set of runs with no bounds on the feasible parameter values ans a second set with bounds as listed in table 7.

Parameter	Reference solution value	Distribution for sensitivity	Bounds for optimisation
E (MPa)	10^5	Normal, $\mu = 10^5$, s = 5100	[0.99, 1.1]$\times 10^5$
ν	0.3	Normal, $\mu = 0.3$, s = 0.0765	[0.15, 0.45]
ρ (kg m^{-3})	10^3	$\mu = 10^3$, s = 51	[.99, 1.1]$\times 10^3$
a (m)	0.025	Uniform on [0.01, 0.04]	[0.01, 0.04]
t (m)	0.001	Uniform on [0.0005, 0.004]	[0.0005, 0.004]

Table 7: Reference solution values, distributions, and bounds for each of the parameter values for the second test problem.

The objective function S was defined as

$$S = \sqrt{\sum_{i=1}^{N} \left(1 - \frac{y_{calc,i}}{y_{targ,i}}\right)^2} \tag{10}$$

where $y_{targ,i}$ is the ith target result and $y_{calc,i}$ is the corresponding value calculated by the model. The target results used were the first 10 frequencies (only six of which are distinct), listed in table 8, two stresses, and a displacement. The analytical values of the frequencies are also listed in table 8, and the difference between the two values, relative to the analytical value, is less than 0.6 % throughout.

Frequency number	Analytical value	FE model target result value
1	7.872311894	7.922888797
2	16.38325456	16.46519639
3	16.38325456	16.47109072
4	26.87623053	26.99632673
5	26.87623053	27.00586146
6	30.64763204	30.78959675
7	39.32372637	39.39581573
8	39.32372637	39.54032405
9	46.87455189	47.00991741
10	46.87455189	47.03192986

Table 8: Fundamental frequencies of the disk calculated analytically and using the finite element model. Both sets of calculations used the reference solution values given in table 7 for the model parameters.

The values used for the other target results were $\sigma_{xx} = 8992887.1$, $\sqrt{u_x^2 + u_y^2} = 4.1057928721$, and $\sigma_{xy} = 10742103$. As with the ellipse test problem, the values are given to a large number of figures to so that the reference solution will be the set of values that minimise the function (truncation of the values might make a different set of parameter values optimal). The use of normalisation coefficients in the analytical expressions for these quantities as given above makes determination of their values analytically difficult to attempt and so no values have been determined. Initial investigations suggested that these parameters should be sufficient to enable optimisation algorithms to converge to the reference solution.

4.2 Sensitivity results

Qualitative analysis of the large-scale Monte Carlo results showed little sensitivity of the target frequencies, stresses and displacement to the material property parameters Young's modulus, Poisson's ratio and density. The more dominant model parameters are the geometrical parameters describing the disk radius and thickness. Scatter plots highlight the dependence of result values on the disk geometry and show a smooth relationship in the region of the reference solution for the objective function (figure 18) as well as the target frequencies, stresses and displacement (figure 19). The shear stress at the origin for the fourth frequency shows some dependence on the material properties, illustrated by the presence of isolated points and the difference in colour between adjacent points in some areas of figure 20(b).

Extreme values of the results occur for a disk radius close to the lower limit of 0.01 m (figure 19(a)).

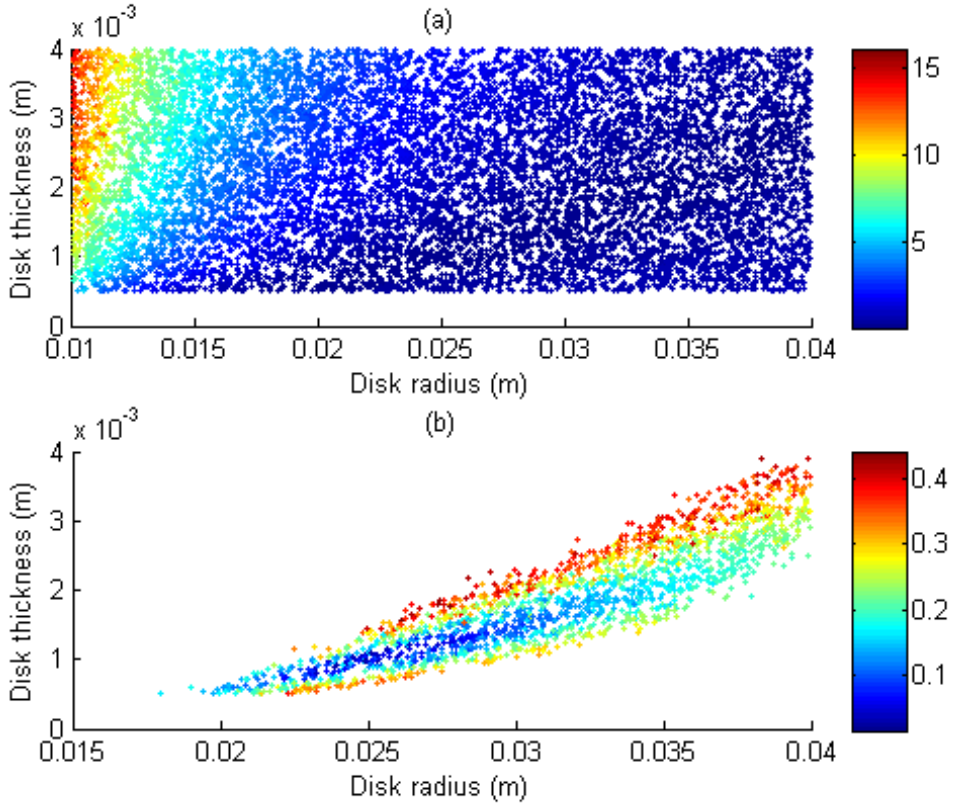

Figure 18: Large-scale Monte Carlo results for variation of the objective function with disk radius and thickness, showing (a) all data points and (b) data points for which all results lie within 50% of their target value. Material properties vary randomly with each point across their respective distributions.

4.2.1 Sobol sensitivity measure

The large-scale Sobol sensitivity test requires $10^4 \times 6$ runs of the vibration model to produce a sensitivity index for each of the five model parameters. The resulting set of indices are consistent with the analysis of the Monte Carlo simulation, and are listed in table 9. The three model parameters describing material properties, Young's modulus, Poisson's ratio and density, all have a maximum value of their sensitivity for the shear stress for the fourth frequency, σ_{xy}. However, the index for this stress is only 0.07 for sensitivity to Poisson's ratio, and less than 0.005 for the

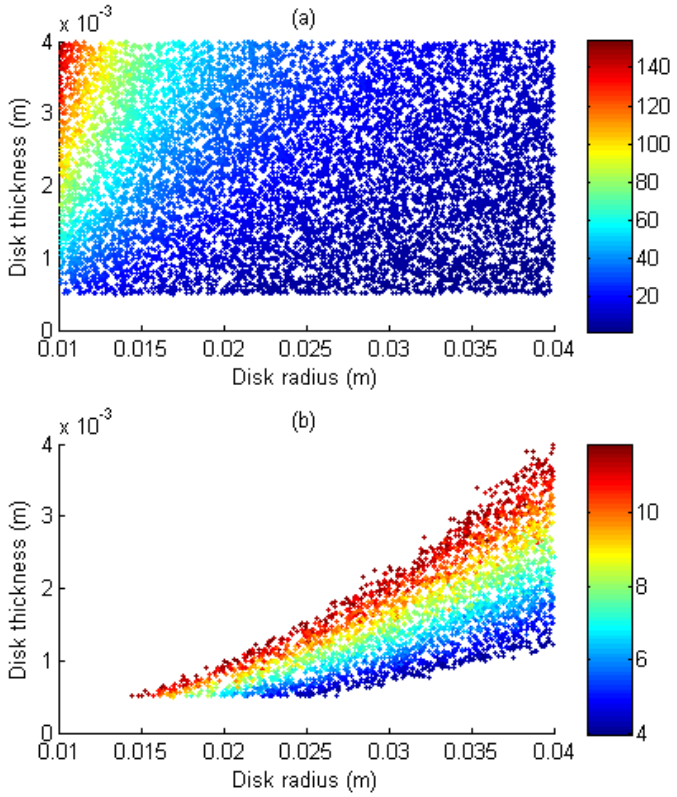

Figure 19: Large-scale Monte Carlo results for variation of first fundamental frequency with disk radius and thickness, showing (a) all data points and (b) results within 50 % of their target value. Material properties vary randomly with each point across their respective distributions.

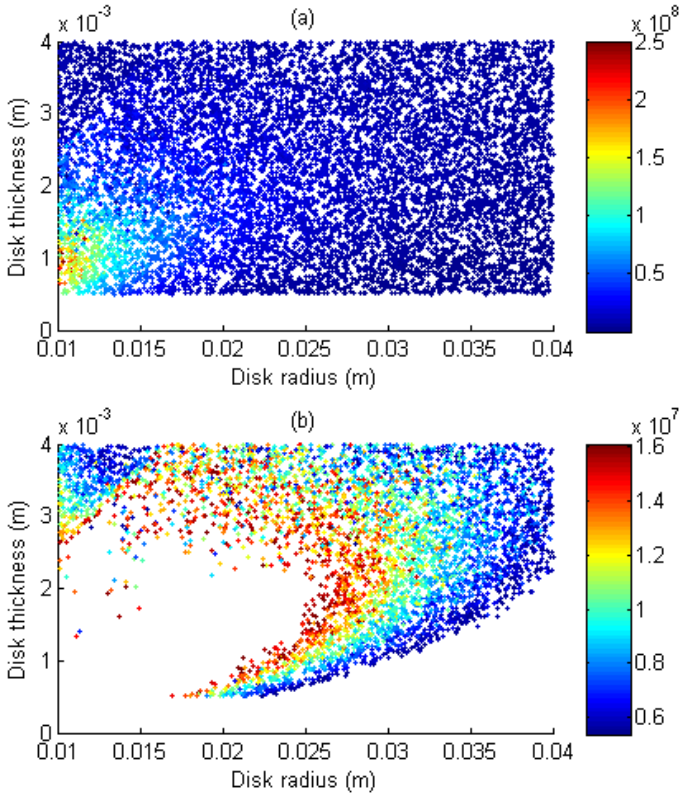

Figure 20: Large-scale Monte Carlo results for variation of the shear stress at the origin for the fourth frequency, with disk radius and thickness, showing (a) all data points and (b) results within 50 % of their target value. Material properties vary randomly with each point across their respective distributions.

remaining two material properties. Many indices give a negative result for Young's modulus, suggesting very little sensitivity.

	F1	F2	F4	F6	F7	F9	σ_{xx}	σ_{xy}	U
E	-0.002	-0.002	-0.001	-0.001	-0.001	-0.001	0.001	0.003	0
ν	0	0	0	0	0	0.002	0.010	0.072	0.001
ρ	0.003	0.003	0.003	0.003	0.004	0.004	0.002	0.005	0.005
a	0.806	0.807	0.807	0.808	0.820	0.827	0.959	0.893	0.963
t	0.302	0.283	0.262	0.255	0.227	0.213	0.102	0.513	0.063

Table 9: Sobol indices for sensitivity of six distinct frequencies (F1, F2,...,F6), normal and shear stresses (σ_{xx}, σ_{xy}) and displacement (U= $\sqrt{u_x^2 + u_y^2}$ evaluated at $r = 0$) results from the vibration model to model parameters Young's modulus (E), Poisson's ratio (ν), density (ρ), disk radius (a) and disk thickness (t). .

For the two parameters describing disk geometry, Sobol indices suggest that all results have the greatest sensitivity to disk radius. Again, the result with the greatest sensitivity to disk thickness is the shear stress (0.51). The displacement at the origin for the second frequency has the greatest sensitivity, across all results, to disk radius (0.96). The sensitivity indices for this displacement suggest that the result depends almost entirely on the radius of the disk. Plotting the displacement against radius only (figure 21), with all other model parameters varying with each point, implies a greater dependence on the radius than a similar plot for the shear stress, which the Sobol test suggests is also sensitive to disk thickness.

4.2.2 Jansen sensitivity measure

The results from the large-scale Jansen test provide the same overall picture as the Sobol test (table 10). In particular, (i) there is a lack of sensitivity of all results to the material properties, (ii) disk radius is the most important model parameter, with indices suggesting almost all sensitivity of the displacement is due to radius, (iii) disk thickness is of greater importance than the material properties, but has less influence on all results than the radius of the disk, with shear stress being the result most sensitive to thickness.

4.2.3 Morris OAT measure

Mean absolute values for the scaled elementary effects calculated for the OAT measure again show little sensitivity to the material properties and highlight disk radius as the most important parameter (table 11).

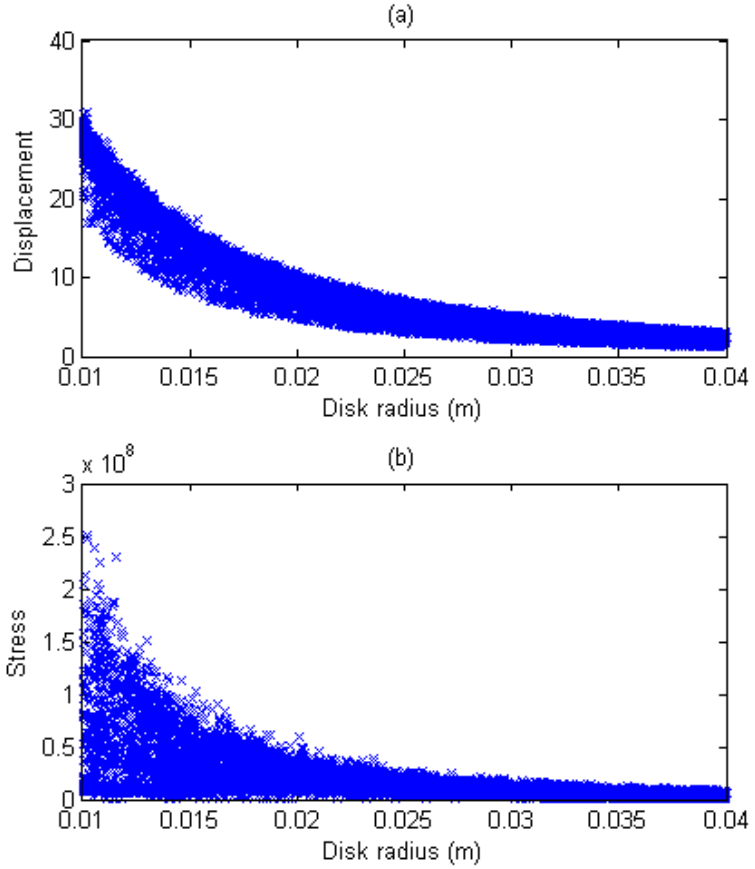

Figure 21: Variation of (a) in-plane displacement and (b) shear stress with disk radius in the Monte Carlo simulation for the vibration model. The remaining four model parameters vary across their distributions with each point.

	F1	F2	F4	F6	F7	F9	σ_{xx}	σ_{xy}	U
E	0.001	0.002	0.002	0.002	0.002	0.002	0.004	0.004	0
ν	0.002	0.001	0.001	0.001	0.001	0.001	0.016	0.069	0.001
ρ	0.002	0.002	0.002	0.002	0.002	0.002	0.001	0.001	0.002
a	0.841	0.839	0.832	0.833	0.833	0.843	0.967	0.863	0.968
t	0.294	0.273	0.248	0.24	0.21	0.196	0.091	0.517	0.053

Table 10: As table 9, but showing values of the Jansen indices.

	F1	F2	F4	F6	F7	F9	σ_{xx}	σ_{xy}	U
E	0.026	0.030	0.038	0.037	0.042	0.044	0.034	0.022	0.001
ν	0.025	0.024	0.028	0.026	0.028	0.029	0.064	0.078	0.017
ρ	0.026	0.03	0.039	0.037	0.043	0.044	0.017	0.012	0.042
a	0.377	0.423	0.52	0.493	0.523	0.528	0.346	0.178	0.582
t	0.231	0.250	0.292	0.27	0.258	0.254	0.097	0.131	0.132

Table 11: As table 9, but showing mean of the absolute values of the Morris OAT indices.

	F1	F2	F4	F6	F7	F9	σ_{xx}	σ_{xy}	U
E	0.026	0.030	0.038	0.036	0.042	0.043	0.033	0.022	0.000
ν	0.025	0.024	0.028	0.025	0.027	0.018	0.063	-0.072	-0.014
ρ	-0.026	-0.030	-0.038	-0.037	-0.042	-0.043	-0.017	-0.011	-0.042
a	-0.377	-0.422	-0.520	-0.493	-0.523	-0.527	-0.346	-0.173	-0.582
t	0.231	0.250	0.292	0.270	0.258	0.254	0.097	-0.094	0.130

Table 12: As table 11, but showing mean values of the Morris OAT indices.

Finding the mean of the absolute value of the scaled elementary effects makes comparison between model parameters easier, but also assumes that it is the magnitude of effect, rather than the sign, that is important. Calculating the mean value indicates the signed effect of an model parameter on result values, and differences between the mean absolute value and the mean signed value can provide further information about the nature of sensitivity to a particular parameter. For example, the mean value for the elementary effects associated with disk radius are negative for all results, implying that results decrease as the radius increases. In the case of disk thickness, positive mean values are recorded for all results except the shear stress. These properties can be seen in scatter plots of the data (figures 19 and 20).

For the elementary effects associated with disk radius, and those describing sensitivity of the frequencies, displacement and normal stress to disk thickness, there is very little difference (generally of the order 0.01 % - 1 %) between the magnitudes of the mean (shown in table 12) and mean absolute value (shown in table 11) of the ratios. This similarity implies a uniformity in the effect of an increase (or decrease) in an model parameter on the sign of change (i.e. positive or negative) of a result (figure 19). The magnitude of the mean value of elementary effects describing the sensitivity of the shear stress to disk thickness is 39 % smaller than the mean absolute value. A significantly smaller value for the mean is produced when an increase in the model parameter can lead to both a positive or negative change in the result value, depending on the local behaviour of the

function at a specific point in the model space (figure 20). The standard deviation is also greater for the sensitivity of the shear stress to disk thickness than for the remaining results. The combination of all of these statistics provides a deeper understanding of the behaviour and sensitivity of the objective function.

It is interesting to compare the values given in table 11 to the values in table 6. Whilst the two tables seem to give the same order for the importance of the parameters (a and then the rest), they differ in the extent to which t is more important than the other parameters. Additionally the analytical values for the importance of E, ρ and ν seem larger than the calculated values. There are several possible explanations for this disparity. The first is that the calculated and analytical values are not comparable. The analytical formulae are expressions of local sensitivity, whereas the Morris OAT method calculates global sensitivity over the entire space of model parameters. It may be possible to average or integrate the local sensitivity to obtain a global sensitivity but this has not been done. Another possible reason is that the normalisation coefficients used in the calculation of the displacements and stresses are, within the FE model, dependent on the material properties (though this would not affect the sensitivity of the frequencies). A third possibility is that the FE model contains some artefact that the analytical model does not, although every effort has been made to ensure that the two models are equivalent.

4.2.4 Sample size analysis

The series of tests with smaller sample sizes produce similar findings to those described in Section 3.2.4 for the elliptical membrane problem. For the smallest sample sizes the Sobol and Jansen indices often take values greater than one or negative values for the Sobol tests. The magnitudes of these outlying indices are generally much smaller than for the elliptical membrane problem, where very large values were caused by large differences between model results for a wide and thin ellipse. These negative and large index values result in a larger standard deviation of indices across the 10 runs for the Sobol and Jansen methods than the Morris OAT method. All three measures provide accurate indices relative to their respective reference values for the largest sample size of 250 (figure 22).

The smallest sample sizes can provide an accurate qualitative impression of the sensitivities of the model results to the model parameters. For example, mean indices over the 10 runs for each sensitivity measure show a much smaller sensitivity of all results to Young's Modulus than to the drum radius (figures 23 and 24). There is good qualitative agreement between the large scale and small scale runs. For a given model parameter, the most sensitive result according to the reference run will have the highest index in the small-scale run, and for a given

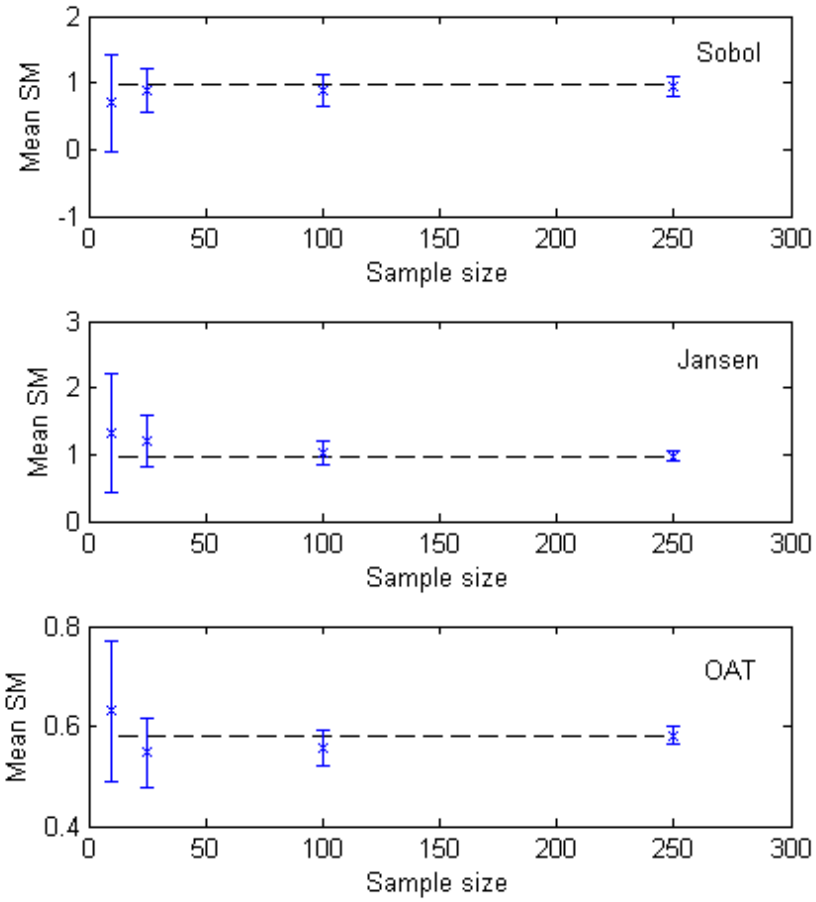

Figure 22: Mean Sobol, Jansen and Morris OAT sensitivity indices for the sensitivity of the in-plane displacement to the disk radius. The means are calculated over 10 runs with sample sizes of 10, 25, 100 and 250. Error bars indicate one standard deviation. Dashed lines represent the mean values found by the large-scale runs.

result the model parameter to which it is most sensitive will be same in the large-scale and small-scale runs.

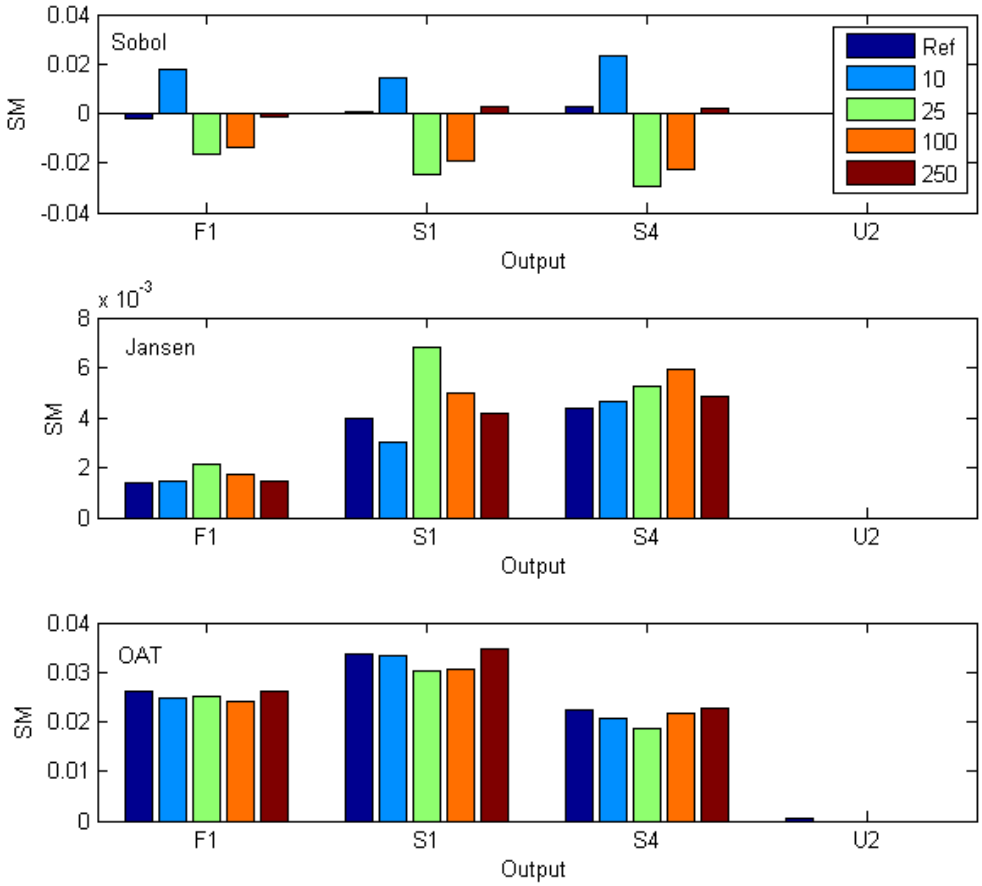

Figure 23: Mean Sobol, Jansen and Morris OAT sensitivity measures for the sensitivity of the first frequency (F1), stress for F1 (S1), shear stress for F4 (S4) and in-plane displacement for F2 (U2), to Young's Modulus. The mean sensitivity is calculated over 10 runs with each sample size 10, 25, 100 and 250.

4.3 Optimisation results

The optimisation problem has five model parameters, and the objective function involves more terms than that of the problem described in section 3. This increase in complexity is likely to make the optimisation problem more challenging and

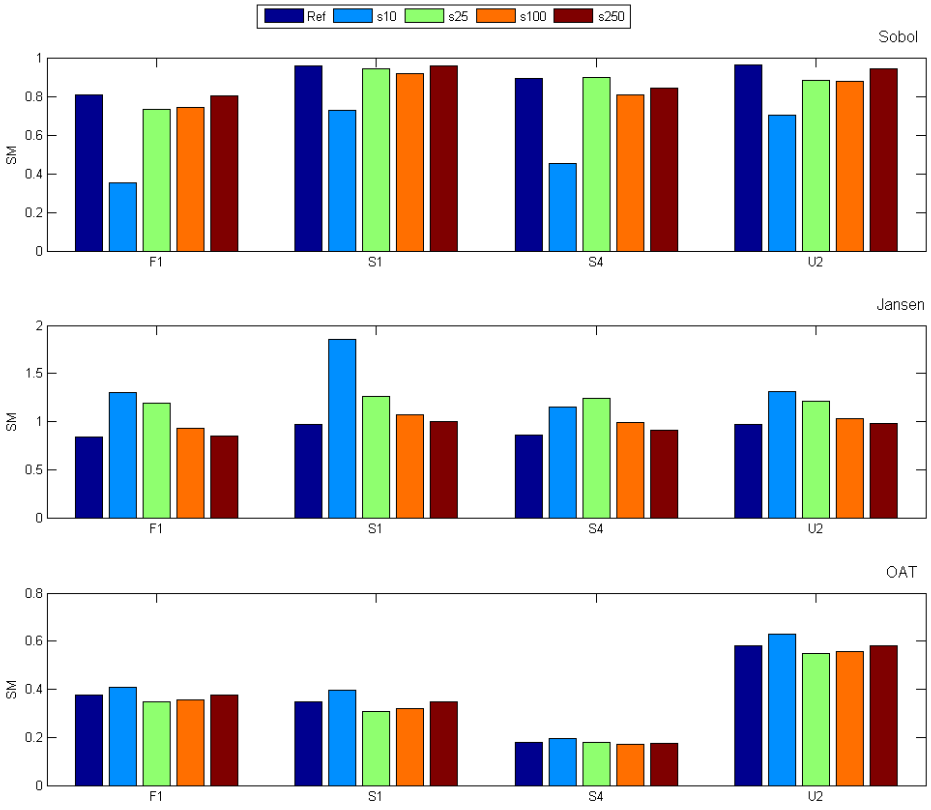

Figure 24: As figure 23 but for sensitivity to disk radius.

hence more computationally expensive to solve, although the problem is still small compared to the majority of optimisation problems. A set of initial runs had been carried out with a degree of success using the Nelder-Mead algorithm and all five model parameters, but after the results of the sensitivity analysis became available, the optimisation work focussed on determining values of the most important parameters: a, t and ν, with initial runs focussing on a and t.

An initial set of runs was carried out allowing a and t to vary without limiting the values they could take. The FE model had been designed such that negative values of radius and thickness would not lead to an error, but would create a mirror image of the expected geometry. This mirror image would have the same fundamental frequencies and the same stresses and displacements at the origin, and so $a = -0.025, t = -0.001$ would be accepted as the correct solution. Not all the algorithms were used at this stage: the poor performance of the SQP2 method led to the SQP1 algorithm not being tested on this example (though it was applied to subsequent examples using this model).

Following these runs, two further sets of runs were carried out: one varying a and t whilst imposing limits on the values, and one varying a, t and ν within their bounds. In both cases the limits are those given in table 7. The results of all three sets of optimisations are listed in table 13.

It is clear from this table that, as before, using fewer model parameters leads to fewer function evaluations before convergence. The use of limiting values has improved the convergence rate for the meta-heuristic algorithms, and has improved the success rate for algorithms that were initially performing poorly. The reduction in the number of evaluations with the reduction from three model parameters to two is less than might be expected for the meta-heuristic algorithms. This feature indicates that the algorithms scale well and may be more efficient for problems with a larger number of model parameters than is illustrated here.

Despite the increase in the number of results of interest used in the objective function compared to the elliptical membrane problem, the sensitivity analysis has shown the objective function to be smooth and unimodal, which explains the comparative performances of the traditional algorithms against the meta-heuristic algorithms. Table 13 shows that standard PSO is better than standard SA, but SA with a fast cooling schedule has virtually the same rate of convergence as the improved PSO.

Comparing tables 4 and 13 shows that many of the methods converge more quickly for the disk problem than the elliptical membrane problem, with the PSO algorithm showing particular improvement. This is unexpected since it is thought that the disk problem is more computationally challenging. The improvement may be due to a better understanding of the fine-tuning of the PSO parameters, or it

may be that the correlation between some of the results used in the objective function makes convergence faster. The fine-tuning requires minimum effort once an understanding of the effects of the parameters has been gained.

Optimisation method	Number of parameters	Limits? (Y/N)	Mean of the number of function evaluations	St dev of the number of function evaluations	Percentage success
SA (slow)	2	N	1988	772	100
SA (slow)	2	Y	1455	623	100
SA (slow)	3	Y	2593	790	100
SA (fast)	2	N	615	249	100
SA (fast)	2	Y	517	240	100
SA (fast)	3	Y	689	256	100
PSO1	2	N	781	430	100
PSO1	2	Y	710	392	100
PSO1	3	Y	1293	523	98
PSO2	2	N	624	347	98
PSO2	2	Y	568	314	100
PSO2	3	Y	809	377	96
NM	2	N	68	10	94
NM	2	Y	73	15	100
NM	3	Y	157	42	96
SQP1	2	N	-	-	-
SQP1	2	Y	266	57	54
SQP1	3	Y	339	104	32
SQP2	2	N	49	55	36
SQP2	2	Y	85	12	100
SQP2	3	Y	173	24	98
LM	2	N	21	4	98
LM	2	Y	22	3	100
LM	3	Y	37	38	90

Table 13: Optimisation results for the disk problem.

It is clear from table 13 that some of the success rates are very low. In particular, the SQP2 algorithm applied to the two parameter problem without limits is very poor (36 % success rate) and the SQP1 algorithm on the two and three parameter problem with parameter limits is only slightly better (54 % success rate). The low success rate of these cases has also affected the repeatability of the SQP2 method very poorly: the standard deviation of the number of function evaluations has become greater than the average number of function evaluations. Large standrad deviations have occurred because the runs that find a wrong solution typically stop

after around 10 iterations, whereas the successful runs take about 100 iterations. Conversely, the high standard deviation of the LM method with three parameters was caused by the solutions that found the wrong answer taking a long time to converge.

The cases with low success rates were investigated further by looking at whether the starting position affected the performance of the algorithm. Figure 25 plots the 50 initial estimates used for the SQP2 runs. The runs that converged to the reference solution are plotted in pink, and those that did not are plotted in blue. The reference solution is plotted as a large black point. The figure shows two clear groups of points, suggesting that the initial position has a strong effect on the success or failure of the run. A similar figure was obtained for the SQP1 runs when limits were applied to a and t, with successful and unsuccessful runs in roughly comparable positions, although there was some overlap between the two sets of points for that case. It is likely that the problems with the SQP2 algorithm were caused by the lack of limiting values since the algorithm is designed for constrained problems (see [11] for more details), and the final solutions had very large values of a and t. From examination of figure 18, it seems that the successful and unsuccessful points lie on different sides of the valley shown in figure 18(b). It is not clear why this would lead the algorithm to converege to such extreme values since the local gradient behaviour appears smooth.

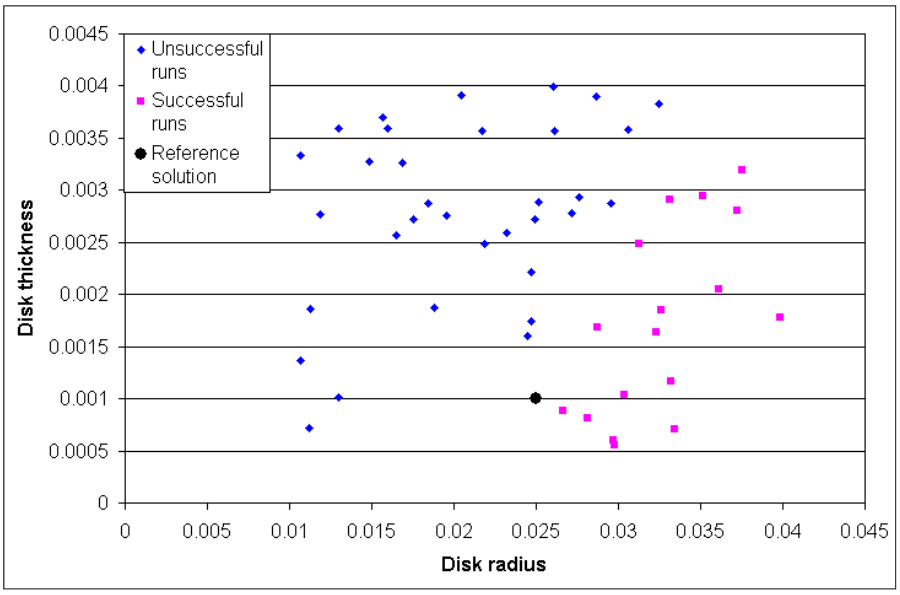

Figure 25: Positions of successful and unsuccessful runs of the SQP2 algorithm.

4.4 Performance of sampling methods

For the vibration problem sample sizes of 32, 64, 243, 486 and 1024 were used. These values are multiples of fifth powers and therefore enabled a simple method for dividing the sample space equally for importance sampling. The sample sizes 32, 243 and 1024 all used one sample within each sub-division, and the sample sizes 64 and 486 used two samples per subdivision.

4.4.1 Accuracy

With the smallest sample size of 32, the Latin Hypercube and importance sampling methods consistently outperform the Monte Carlo random sampling in terms of accuracy of sample mean compared to the reference mean values (e.g., figure 26). For sample sizes 486 and 1024, accuracy to the reference mean is similar for all three sampling methods.

Based on the average difference of the sample means from the reference values, a smaller sample size is needed for accuracy to within 5 % of the reference value for Latin Hypercube sampling than for Monte Carlo or importance sampling (table 14).

	F1	σ_{xx}	σ_{xy}	U
MCS	243	486	486	243
LHS	32	64	243	32
IS	243	468	243	243

Table 14: Minimum sample size, for each sampling method, producing an average (percentage) difference between the sample and reference mean values of less than 5 %. Mean values are compared for the first frequency (F1), two stresses (σ_{xx} and σ_{xy}) and the displacement (U).

When different results of interest are analysed, no single sampling method performs consistently better or worse than others tested in terms of accuracy compared to the reference standard deviation (e.g. figure 27).

4.4.2 Repeatability

Across all sample sizes, Latin Hypercube sampling produces the smallest standard deviation, over 10 runs, of each of the mean frequencies (e.g. figure 28). For the remaining three results, Latin Hypercube sampling again produces the smallest

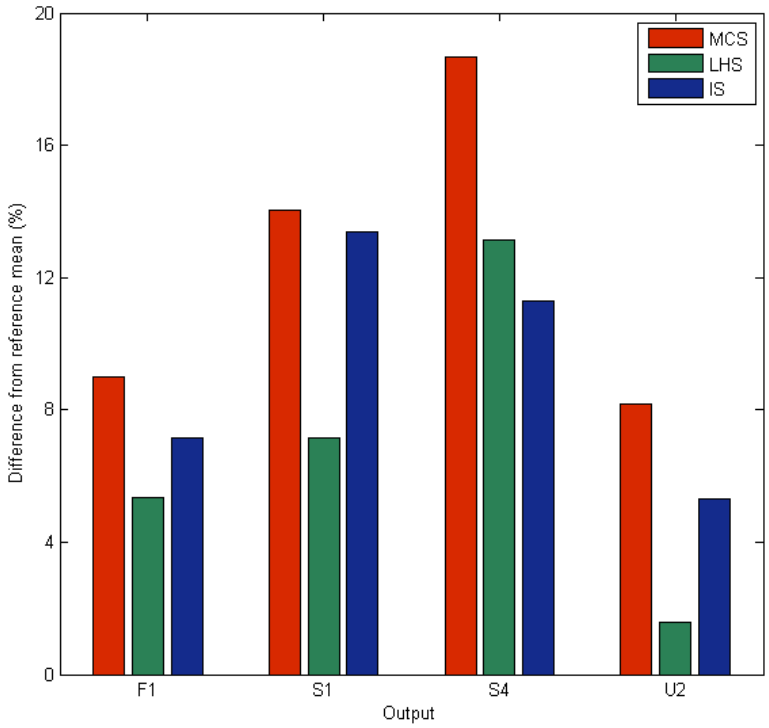

Figure 26: Mean percentage difference between mean first frequency (F1), stress (S1), shear stress (S4) and displacement (U2) and respective reference mean values, for sample size 32 and Monte Carlo (MCS), Latin Hypercube (LHS) and importance (IS) sampling methods.

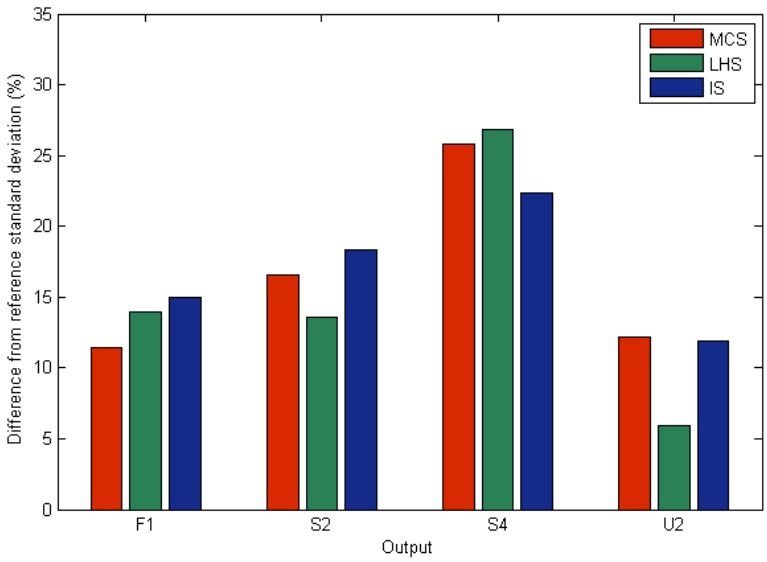

Figure 27: Mean percentage difference between standard deviation of first frequency (F1), stress (S1), shear stress (S4) and displacement (U2) and respective reference standard deviation values, for sample size 32 and Monte Carlo (MCS), Latin Hypercube (LHS) and importance (IS) sampling methods.

standard deviation at the smaller sample sizes. For sample sizes of 243 and greater, all three methods perform similarly for estimating the shear stress and displacement. Latin Hypercube sampling remains the most consistent for estimates of the stress for the first frequency.

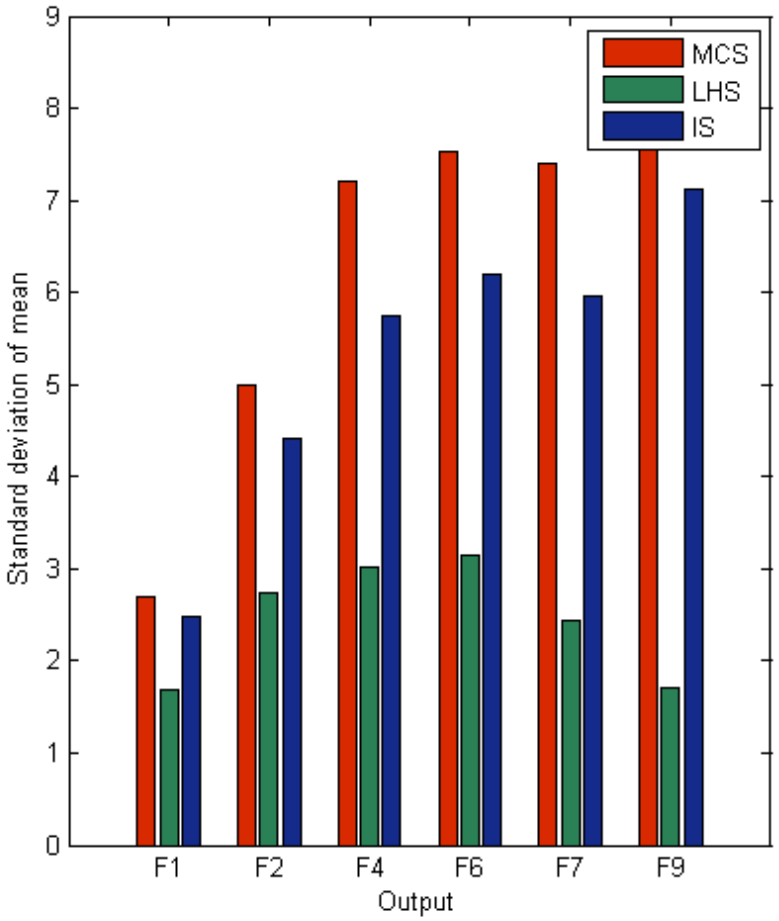

Figure 28: Standard deviation over 10 runs of the mean value for distinct frequencies F1, F2, F4 F6, F7, F9, for sample size 32 and Monte Carlo (MCS), Latin Hypercube (LHS) and importance (IS) sampling methods.

For all three sampling methods, the improvement in repeatability of mean values with increasing sample size is greatest while the sample size is small (e.g. figure 29). For example, from the Latin Hypercube tests the standard deviation of the mean shear stress decreases by 52 % with an increase in sample size from 32 to 64. The subsequent decrease when sample size is increased to 243 is 65 %. This

equates to a 1.6 % decrease in standard deviation per unit increase in sample size from 32 to 64, and a 0.4 % decrease per unit increase in sample size from 64 to 243. Monte Carlo and importance sampling show similar improvements with increasing sample size.

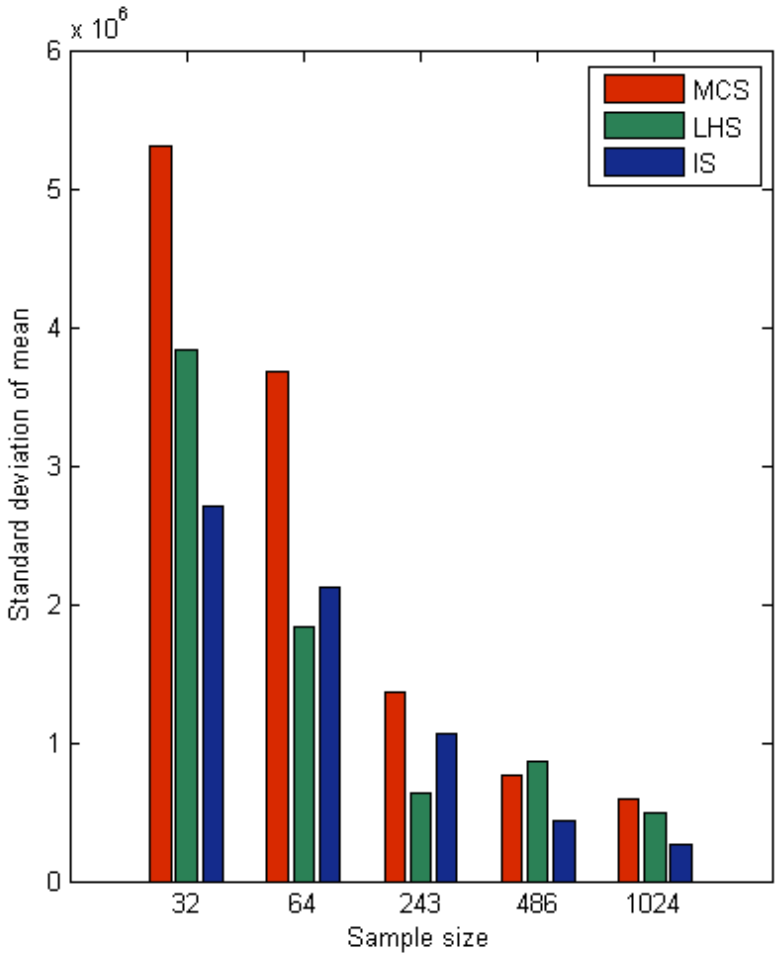

Figure 29: Standard deviation over 10 runs of the mean shear stress for sample sizes 32, 64, 243, 486 and 1024, with Monte Carlo (MCS), Latin Hypercube (LHS) and importance (IS) sampling methods.

At the smallest sample size, all three sampling methods produce similar standard deviations, over ten runs, of the standard deviation of each model result. For larger sample sizes, the Latin Hypercube tests generally produce the lowest standard deviation across the ten runs.

Observations made from cumulative probability plots for each of the ten runs, with a given sample size and sampling method, are consistent with those outlined above. Observations are that Latin Hypercube tests produce the smallest spread in results; variance of results decreases with increasing sample size for all three sampling methods and the improvement with sample size increase diminishes with each increase (e.g. figure 30).

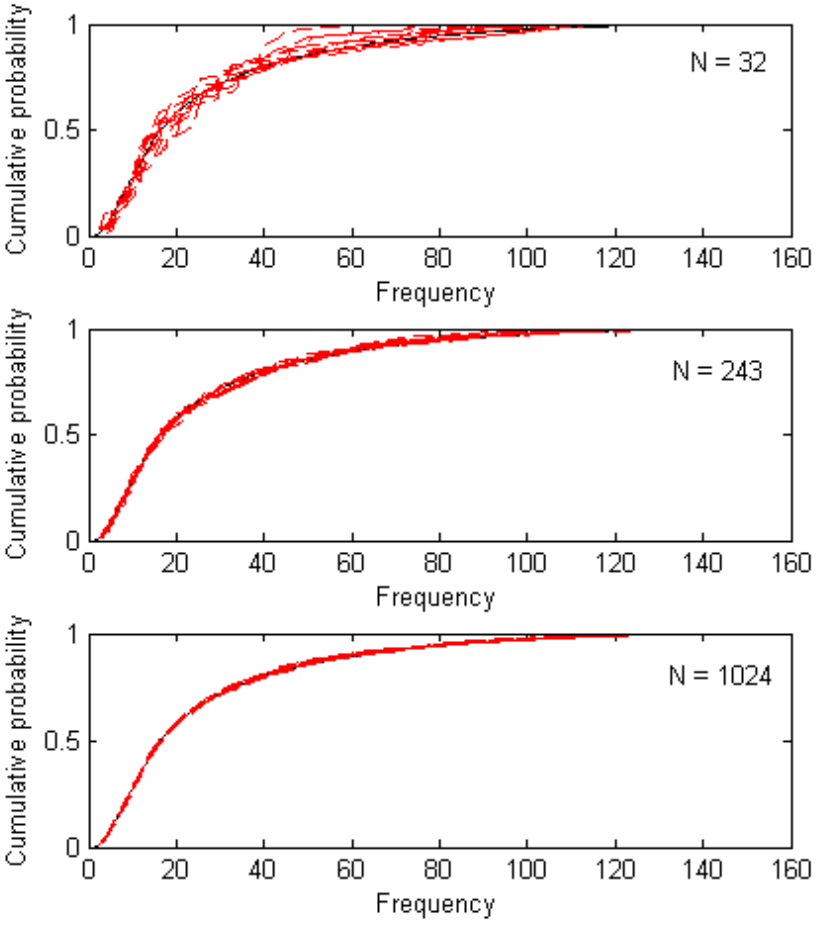

Figure 30: Cumulative probability of first frequency for each of ten runs, with sample size 32, 243 and 1024, for Latin Hypercube sampling.

4.5 Conclusions

- Using a combination of the various statistics derived from the Morris OAT method it is possible to build a more informed picture of the model sensitivities than by just examining the mean values alone.
- Whilst sensitivity indices as used here give a good picture of global sensitivity, they do not always capture local behavior.
- Placing limiting values on model parameters during an optimisation can improve both convergence rate and success rate.
- The choice of starting values for the model parameters can make a large difference to the success of an optimisation run. Knowledge of the response surface may suggest that some regions of the feasible space are more suitable than others. If possible, repeated runs should be carried out starting from different estimates.
- For small sample sizes, the LHS method outperforms the other sampling methods in terms of accuracy and repeatability.

5 Problem 3: Thermal problem

5.1 Problem definition

The final small test problem is a transient thermal problem based on NAFEMS benchmark T3 [9]. The FE model simulates a rod with a fixed temperature boundary condition at one end and a time-dependent boundary condition at the other end, as shown in figure 31.

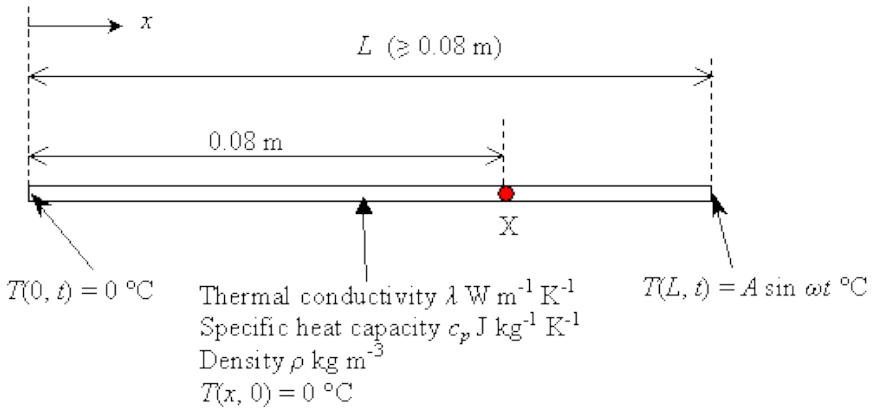

Figure 31: Sketch of the transient thermal problem.

The model parameters considered were the length of the rod, L, the amplitude A and frequency ω associated with the transient boundary condition, and the thermal conductivity of the rod λ. The density was fixed at $\rho = 7\,200\,\mathrm{kg\,m^{-3}}$ and the specific heat capacity at $c_p = 440.5\,\mathrm{J\,kg^{-1}\,K^{-1}}$. The reference solution for these quantities and the associated distributions are given in table 15. Optimisations were run with and without bounds to compare success rates and convergence times.

The mesh was designed such that the element size was 1/1500 between $x = 0$ and $x = 0.08$, and that the space between $x = 0.08$ and $x = L$ was filled with elements of approximately this size. This approach was taken in order to guarantee the presence of a node at $x = 0.08$ and because a mesh of this type gave converged results for various choices of model parameters. A trapezoidal method was used for time integration due to its combination of efficiency and stability. A time step of 0.5 s was used throughout. An analytical solution is available for this problem, and the FE model results were compared to this solution when the mesh was being developed. The analytical solution is

$$T(x,t) = AB(x)\sin(\omega t + \phi(x)) + \frac{4\pi A k^2}{L^2}\sum_{n=1}^{\infty}\frac{n(-1)^{n+1}}{\alpha_n^4 + 4k^4}\sin(\alpha_n x)e^{-\kappa\alpha_n^2 t} \tag{11}$$

where

$$\kappa = \frac{\lambda}{\rho c_p}, \tag{12}$$

$$\alpha_n = \frac{n\pi}{L}, \tag{13}$$

$$k = \left(\frac{\omega}{2\kappa}\right)^{1/2}, \tag{14}$$

$$B = \left|\frac{\sinh kx(1+i)}{\sinh kL(1+i)}\right|, \tag{15}$$

$$\phi = \arg\left(\frac{\sinh kx(1+i)}{\sinh kL(1+i)}\right). \tag{16}$$

These last two quantities can be more usefully written as

$$B = \left(\frac{\cosh 2kx - \cos 2kx}{\cosh 2kL - \cos 2kL}\right)^{1/2}, \tag{17}$$

$$\phi = \tan^{-1}\left(\frac{\sin k(x-L)\sinh k(x+L) - \sinh k(x-L)\sin k(x+L)}{\cos k(x-L)\cosh k(x+L) - \cosh k(x-L)\cos k(x+L)}\right). \tag{18}$$

Parameter	Reference solution value	Distribution for sensitivity analysis	Bounds for optimisation
L/m	0.1	Uniformly distributed on [0.08, 0.12]	[0.08, 0.12]
A/K	100	Normal, $\mu = 100$, $s = 31$	[90, 110]
ω/s^{-1}	$\pi/40$	Normal, $\mu = \pi/40$, $s = 0.011$	[0.05, 0.1]
λ/W m^{-1} K^{-1}	35	Normal, $\mu = 35$, $s = 5.82$	[25, 45]

Table 15: Reference solution values, distributions, and bounds for each of the model parameters for the third test problem.

The results of interest are the temperature at the point marked X in figure 31 at various times. For the optimisation, the objective function S was defined as

$$S = \sqrt{\sum_{i=1}^{N}\left(1 - \frac{y_{calc,i}}{y_{targ,i}}\right)^2} \tag{19}$$

where $y_{targ,i}$ is the ith target result and $y_{calc,i}$ is the corresponding value calculated by the model. The $y_{targ,i}$ were the temperature at the point marked X every two seconds between 2 s and 32 s inclusive, so $N = 16$. The sensitivity analysis looked

at the value of S and at the values of the temperature at X every 8 seconds between 8 and 32, denoted throughout as T_8, T_{16}, T_{24},and T_{32}. The target results are given in table 16, along with the analytical solution at the same times for comparison.

Time /s	Analytical value /°C	FE model value /°C	Time /s	Analytical value /°C	FE model value /°C
2	0.005936448	0.008874314	18	18.36171042	18.51476996
4	0.239489939	0.246447744	20	21.97072289	22.13983203
6	1.114014718	1.132172212	22	25.41786898	25.59866734
8	2.742780671	2.780124462	24	28.58173531	28.76938332
10	5.055283876	5.116217321	26	31.35279847	31.54247705
12	7.921236493	8.007333326	28	33.63597623	33.8225918
14	11.19359414	11.30426187	30	35.35157965	35.53010655
16	14.72250589	14.85590565	32	36.43688026	36.60248732

Table 16: Analytical and numerical results for the temperature at point X for various times. The numerical results are used as target results in the objective function.

It is clear from the formulae in (11) to (18) that analytical expressions for the local sensitivity coefficients are extremely complicated, and so they have not been determined. Instead, the partial derivatives of each of the temperatures listed in table 16 in the region of the reference solution with respect to normalised versions of the model parameters have been determined numerically using finite difference approximations applied to the analytical solution. The normalised versions of the parameters are denoted as $\hat{L}$, $\hat{A}$, $\hat{\omega}$, and $\hat{\lambda}$. Normalised versions have been used for ease of comparison with the sensitivity indices (which are also calculated with normalised parameters).The results are presented in table 17 as factors multiplying the temperature result at each time, since this makes it clear how the influence of each parameter on the temperature varies as time increases.

	Sensitivity to			
Result	$\hat{\lambda}$	$\hat{L}$	$\hat{A}$	$\hat{\omega}$
T_8	1.9T8	−19T8	1.0T8	0.97T8
T_{16}	1.2T16	−12T16	1.0T16	0.80T16
T_{24}	0.81T24	−8.1T24	1.0T24	0.44T24
T_{32}	0.56T32	−5.5T32	1.0T32	−0.22T32

Table 17: Numerically determined sensitivity coefficients with respect to normalised parameters $\hat{L}$, $\hat{A}$, $\hat{\omega}$, and $\hat{\lambda}$ in the region of the reference solution. The reference solution is $\hat{L} = \hat{A} = \hat{\omega} = \hat{\lambda} = 1$.

Table 17 shows that

- the solution is linearly dependent on A at all times (as is obvious from the analytical solution),
- the rod length L is the most important parameter but it becomes less important as time passes, and
- the thermal conductivity λ and the frequency ω become less important as time passes.

5.2 Sensitivity results

Data from the large-scale Monte Carlo simulation for the thermal problem were analysed using simple scatter plots, to assess the relative importance of the four model parameters on the temperature results. The model results show the greatest sensitivity to the length of the rod L and the amplitude for the boundary condition A. Of the two parameters, rod length appears to be more important, particularly for lengths close to 0.08 m, as is shown in figure 32(a) by the steep gradients in the region close to $L = 0.08$ m. However, comparing scatter plots showing the temperature after 8 s and 32 s as a function of length and amplitude suggests that the amplitude becomes more important with time as it affects the result for increasingly longer rod lengths (figures 33 and 34). This dependency is also indicated by the difference in colour of adjacent points in figure 32(b), particularly close to the target solution ($L = 0.1$ m, $A = 100$°C).

The scatter plot showing the temperature after 32 s, against L and A, suggests that there is also some sensitivity to the remaining two model parameters, especially when the rod length is relatively small (figure 34). Values for the thermal conductivity λ and boundary condition frequency ω vary across each point in the scatter plot. Further analysis suggests that ω is the more important of the two parameters, particularly for large values of ω and a relatively short rod length. For example there a two data points representing a much cooler temperature than the surrounding points in the top left corner of the upper plot in figure 34, and both have values for ω greater than two standard deviations above its mean value.

5.2.1 Sobol sensitivity measure

The large-scale Sobol sensitivity indices are consistent with the qualitative findings from the Monte Carlo simulation, described above, and are listed in table 18. The indices show rod length as the most important model parameter, followed by amplitude A. The temperature becomes more sensitive to the amplitude and less sensitive to rod length with time. Again, as suggested by the Monte Carlo analysis, the frequency ω is generally more important than the thermal

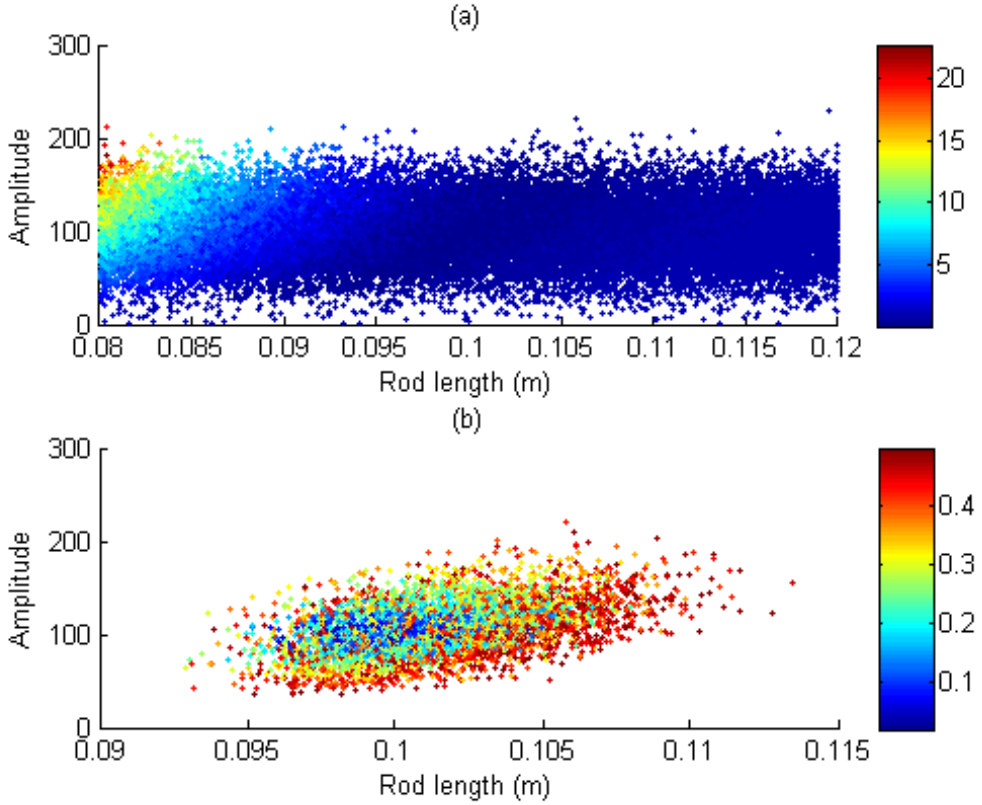

Figure 32: Large-scale Monte Carlo results for variation of the objective function with rod length and boundary condition amplitude, showing (a) all data points and (b) data points for which $S < 0.5$.

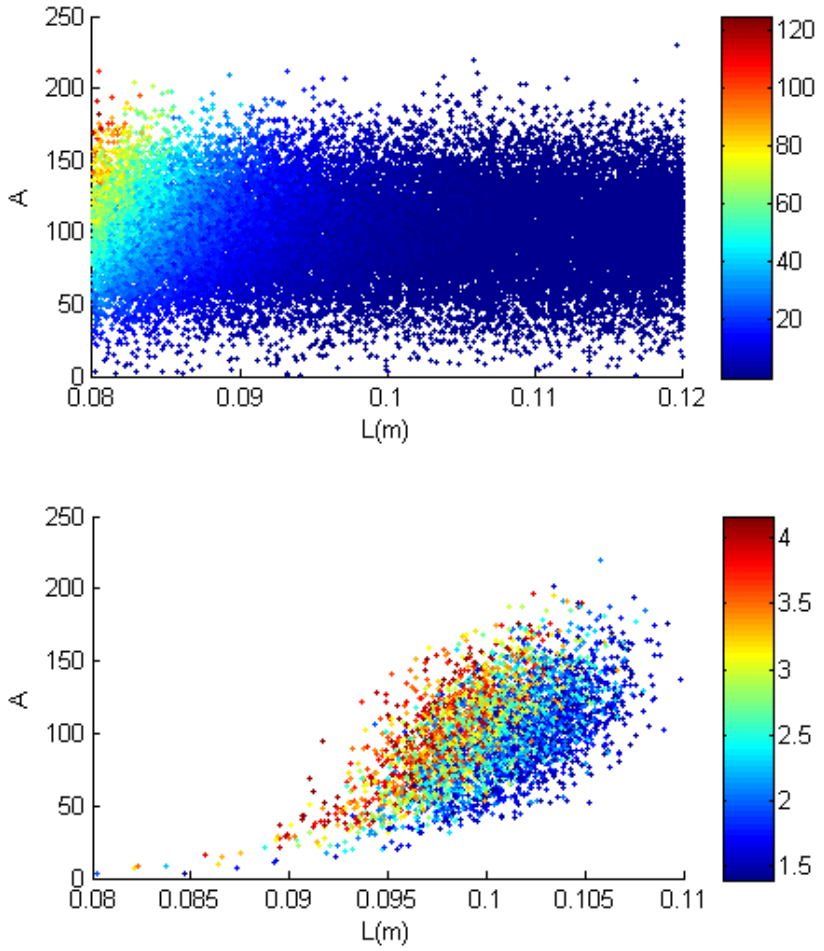

Figure 33: Temperature after 8 s at a point 0.08 m along a rod of length L, heated by a sinusoidal transient boundary condition with amplitude A. The frequency for the boundary condition and the thermal conductivity vary with each data point. The lower figure shows data points that lie within 50 % of target value of 2.8 °C.

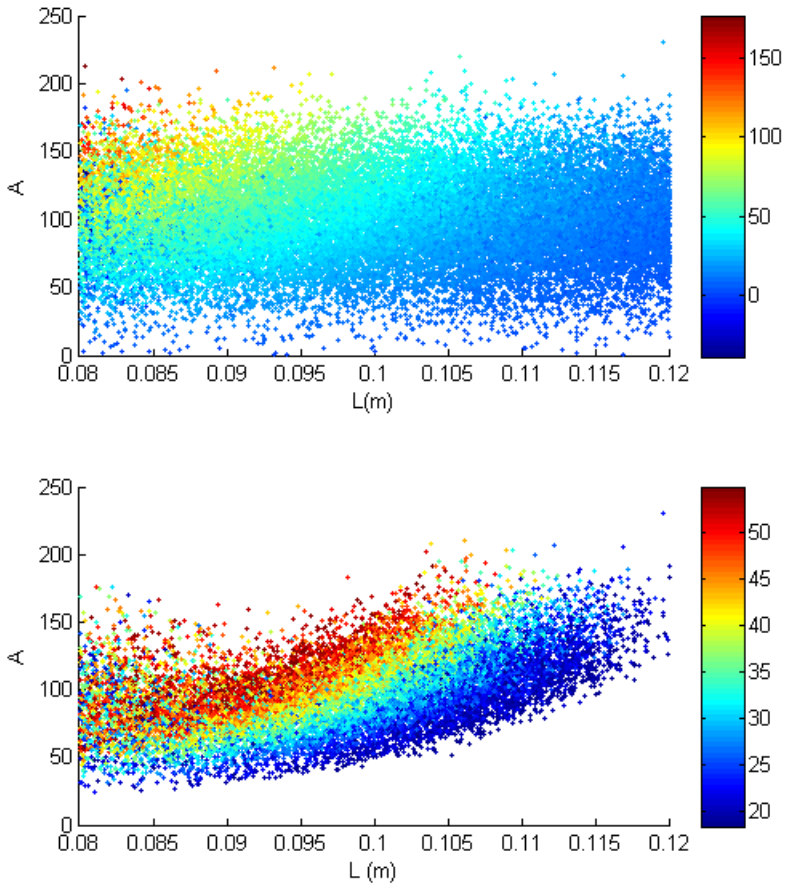

Figure 34: As figure 33 except after 32 s and for a target value of 36.6 °C

conductivity. However, the temperature becomes slightly more sensitive to the thermal conductivity with time.

	T_8	T_{16}	T_{24}	T_{32}	S
λ	0.0060	0.0081	0.011	0.015	0.0058
L	0.967	0.93	0.88	0.74	1.00
ω	0.020	0.0099	0.010	0.17	0.021
A	0.12	0.16	0.21	0.28	0.11

Table 18: Sobol indices for sensitivity of rod temperature after 8, 16, 24 and 32 s, and the objective function S (19) to parameters thermal conductivity (λ), rod length (L), and boundary condition frequency (ω) and amplitude (A).

5.2.2 Jansen sensitivity measure

The results from the large-scale Jansen test are again consistent with the Monte Carlo analysis and give similar values to the Sobol indices (table 19).

	T_8	T_{16}	T_{24}	T_{32}	S
λ	0.0056	0.0075	0.010	0.015	0.0045
L	0.95	0.92	0.86	0.72	0.98
ω	0.022	0.012	0.0088	0.16	0.023
A	0.12	0.16	0.21	0.28	0.10

Table 19: As table 18, but for Jansen indices.

5.2.3 Morris OAT measure

The Morris OAT test again identifies rod length L and amplitude A as the two most important model parameters. Conversely to the Sobol and Jansen tests, the sensitivity to the rod length increases with time, along with that for the amplitude. The Morris OAT test suggests that temperature at the final time point has become slightly more sensitive to amplitude than to rod length. Data points for temperature close to the target value after 32 s (figure 34) suggest a greater sensitivity to the amplitude than that implied by the Sobol and Jansen indices. An additional difference between the Sobol and Jansen indices, and the Morris OAT results is that the latter shows ω rather than λ to be the least important parameter. Qualitative analysis of the Monte Carlo runs, as well as the Sobol and Jansen tests, suggests that the opposite may be true.

	T_8	T_{16}	T_{24}	T_{32}	S
λ	0.050	0.075	0.098	0.12	0.024
L	0.37	0.47	0.53	0.53	0.19
ω	0.0070	0.0072	0.0052	0.024	0.0038
A	0.19	0.28	0.40	0.56	0.092

Table 20: As table 18, but for mean absolute elementary effects from the Morris OAT analysis.

Comparing these results to the local sensitivity calculations given in table 17 is of interest. Whilst the rod length L is found to be the most important parameter in both sets of results, it appears that the frequency ω is more important to local sensitivity in the region of the target solution than it is to global sensitivity.

5.2.4 Sample size analysis

The series of smaller scale tests for the thermal problem provides information on the three different sensitivity tests consistent with that obtained for the previous two problems.

The Morris OAT sensitivity measures tend to have a smaller variation over multiple runs with small sample sizes, and all three tests produce accurate sensitivity values, relative to their respective reference values, for the largest sample size of 250 (figures 35 and 36). A typical run of the Morris OAT and Jansen tests produces index values in the correct relative order, even for the smallest sample size of 10 (figure 37). The possibility of the occurence of negative values when the Sobol test is run with a small sample size means that this is not always the case for Sobol indices.

5.3 Optimisation Results

The sensitivity analysis suggested that the parameters L and A were most important, and that the parameters λ and ω had less effect on the value of the objective function. Accordingly, three sets of optimisation runs were carried out: one set with all four model parameters varying (which was not completed as will be explained below), one varying λ, A and ω (only using traditional algorithms) and one allowing L and A to vary but keeping λ and ω fixed. The results are shown in table 21.

The traditional algorithms struggled to complete the runs varying all four model

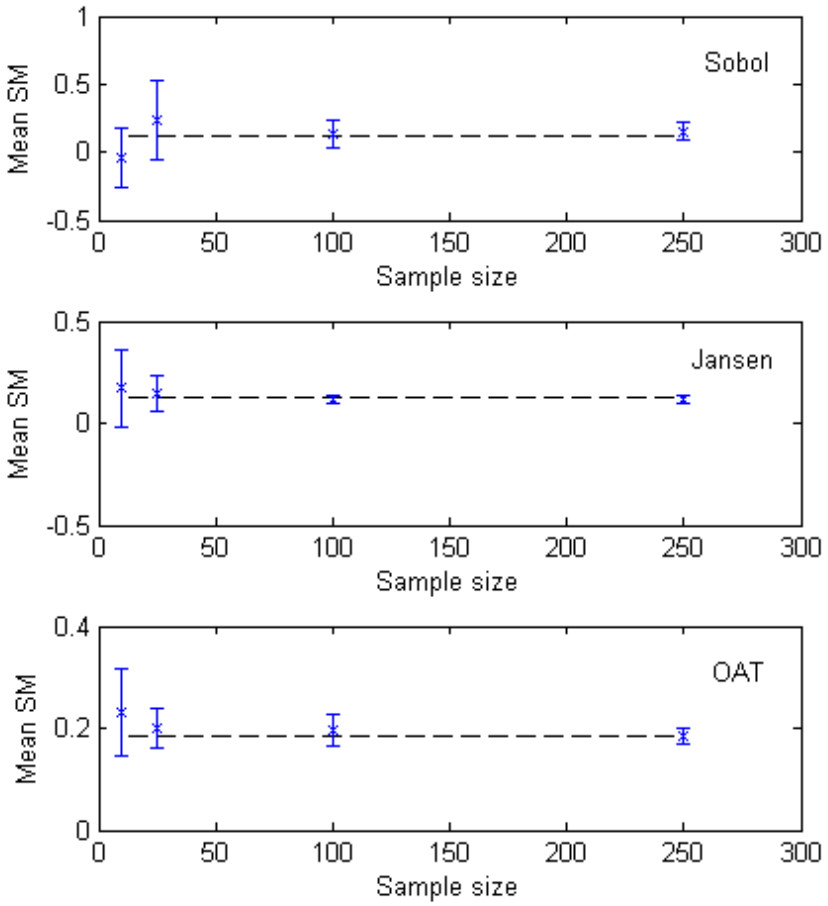

Figure 35: Mean Sobol, Jansen and Morris OAT sensitivity indices for the sensitivity of the temperature after 8 s to the boundary condition amplitude. The means are calculated over 10 runs with sample sizes of 10, 25, 100 and 250. Error bars indicate one standard deviation.dashed lines indicate value calculated from large-scale tests

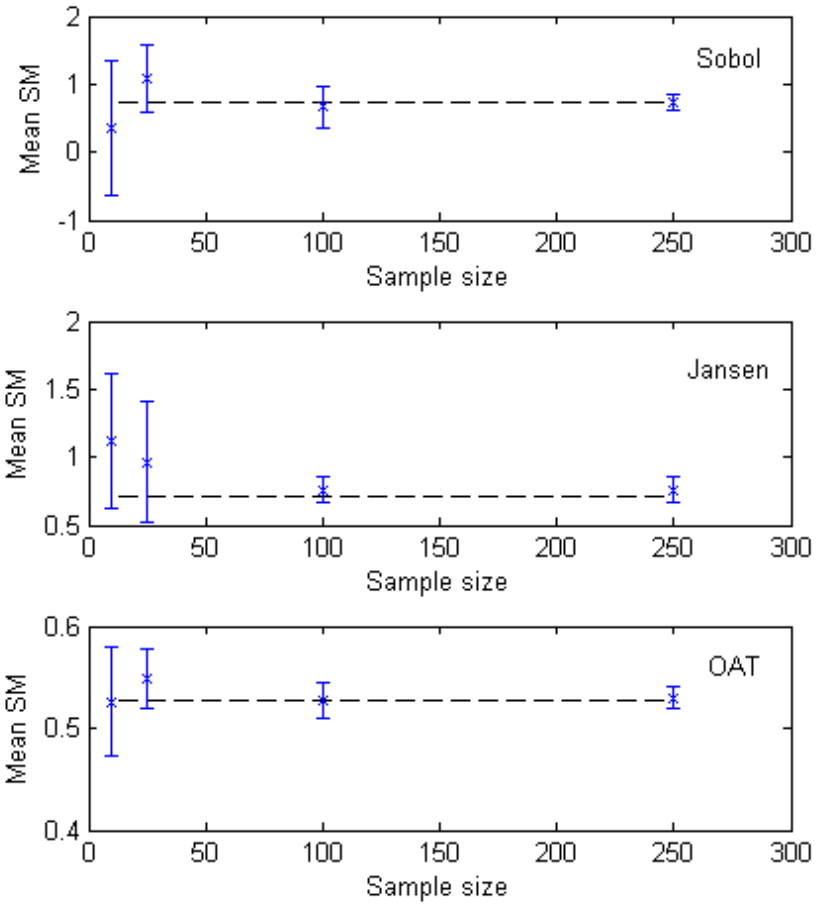

Figure 36: As for figure 35, but for the sensitivity of the temperature after 32 s to the rod length.

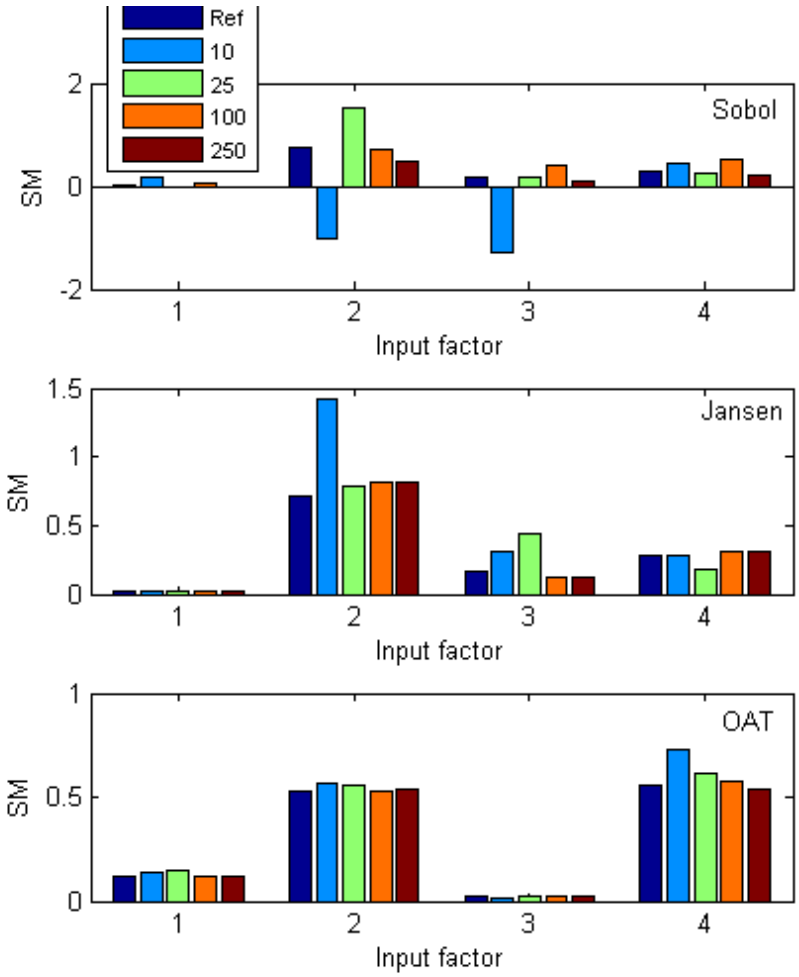

Figure 37: Sobol, Jansen and Morris OAT sensitivity indices for the sensitivity of the temperature after 32 s to the model parameters: (1) thermal conductivity, (2) rod length, and boundary condition (3) frequency and (4) amplitude. Indices are calculated over a single run with sample sizes of 10, 25, 100 and 250.

parameters. Of the two algorithms that were tried (NM and LM), the former had a zero percent success rate and the latter had a 22 % success rate. After these algorithms had performed so poorly it was decided that applying the other traditional algorithms to the problem would be a waste of effort. Examining the failed runs showed that in general the algorithms had determined ω, L, and A accurately but could not converge to the correct value of λ in the majority of cases. The meta-heuristic algorithms did not share this problem: all of the sets of runs using SA and PSO had a success rate of at least 90 %.

A second set of runs were carried out using only the traditional algorithms to investigate whether λ always caused problems. All four traditional algorithms were applied to a problem that allowed λ, A and ω to vary. These runs were considerably more successful: the lowest success rate was 84 % for the SQP2 runs. Finally all algorithms were applied to a problem where A and L were allowed to vary.

Optimisation method	Number of parameters	Mean of the number of function evaluations	St dev of number of function evaluations	Percentage success
SA (slow)	2	4705	923	98
SA (slow)	4	7057	1154	100
SA (fast)	2	1268	707	100
SA (fast)	4	1920	897	90
PSO1	2	2769	542	100
PSO1	4	4073	812	100
PSO2	2	1457	490	100
PSO2	4	2476	735	100
NM	2	99	9	100
NM	3	150	17	100
NM	4	292	136	0
SQP1	2	203	11	84
SQP1	3	229	11	100
SQP2	2	105	35	86
SQP2	3	155	42	84
LM	2	40	20	98
LM	3	20	2	100
LM	4	232	106	22

Table 21: Optimisation results for the thermal problem.

The objective function and constraints of the thermal problem have the same basic form as those of the other two problems. The problem involves four model parameters and the objective function is based on evaluating the same quantity at 16 different times. On the surface this problem seems as though it should not be

significantly more difficult to solve than the other small problems. The failure rate of the traditional algorithms shows that this is not the case. The use of a black-box function has made the shape of the response surface difficult to predict, even though the objective function is a sum of squares.

As with the previous problems, the reduction in the number of model parameters generally leads to a reduction in the number of function evaluations required for convergence and an increase in success rates. An interesting exception to this is the behaviour of the Levenburg-Marquardt algorithm for two and three model parameters. The runs with two parameters require on average twice as many function evaluations as those with three parameters, and the standard deviation is ten times as high. This behaviour is not seen for any other algorithm and it is not immediately explicable. The success rate for the algorithm is good in both cases and so convergence to a wrong solution is not the cause of the problem.

The stochastic algorithms are still significantly less efficient than the traditional algorithms for equivalent problems, suggesting that the response surface is comparatively smooth and unimodal. A detailed examination shows that there is some non-smoothness close to the optimum, as shown in figure 38. The local behaviour shows a sharp dependence on L but very little dependence on A. The traditional gradient-based methods are likely to struggle in this region.

A detailed examination of the runs of the Levenburg-Marquardt method shows that the algorithm gets close to the correct value of L quite quickly but takes longer to converge to the correct value for L, so the seemingly contradictory behaviour is explained by the local shape of the response surface. This is also reflected in the sensitivity coefficient values given in table 17. The values for λ, A and ω are broadly similar for all times, suggesting that the surface gradients will be approximately equal around the correct solution, whereas the values for L are between 5 and 50 times as large as those for A, suggesting the kind of sharp valley shown in figure 38.

Matlab's implementation of the Nelder-Mead algorithm [16] includes several parameters that control when the algorithm stops, including a tolerance on the step size of the model parameters called TolX, and a tolerance on the maximum change in the function value called TolFun (see the Matlab help files or the companion report [6] for more details). In order to investigate the effects of this tolerance on the convergence, two sets of 50 optimisation runs allowing A and L to vary were carried out: one set of runs fixed TolX to be 10^{-2} and the other fixed it to be 10^{-4}. Both sets of runs left TolFun set to its default value for the Nelder-Mead algorithm, 10^{-4}. Each set of runs started from the same 50 initial positions to allow direct comparison between runs. The main points of interest were how the value of TolX affected the accuracy of the converged solutions relative to the reference solution, and how the value of TolX affected the number of

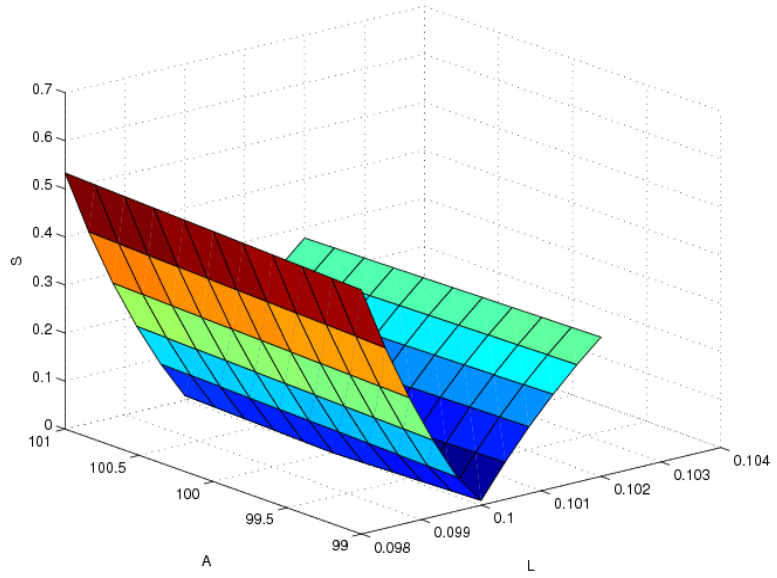

Figure 38: Non-smoothness of the response surface near the optimum.

function evaluations taken before the algorithm stopped.

In all cases the runs with TolX= 10^{-2} stopped before the runs with TolX= 10^{-4}, suggesting that TolX controlled the stopping (rather than TolFun). All results obtained with TolX= 10^{-2} were correct to four significant figures (maximum relative error of 8×10^{-5}) and all those from TolX= 10^{-4} were correct to six significant figures (maximum relative error of 8×10^{-7}). Comparing pairs of runs that started from the same point, the smaller value of TolX led to an average increase in the number of function evaluations of 29 (the average number of function evaluations for the two values of TolX being 68 for 10^{-2} and 97 for 10^{-4}).

5.4 Performance of sampling methods

For the thermal problem, the three different sampling methods were tested using sample sizes of 16, 81, 256, 625 and 1296. These are all fourth powers, meaning that exactly one sample could be taken within each subdivision for the Latin Hypercube and importance sampling methods. Ten runs were carried out for each method and sample size.

5.4.1 Accuracy

Over multiple trials, the Latin Hypercube and importance sampling methods produce more accurate results with the smallest sample size in terms of the reference mean value. For example, over ten runs the average difference between the small-sample (sample size 16) and reference mean for the temperature after 32 s is 7 % and 8 % for the Latin Hypercube and importance sampling methods respectively (e.g. figure 39). The average difference for the Monte Carlo sampling is 14 %. In relation to the reference standard deviation for the same model result, all three sampling methods perform similarly, with average differences of 18 %, 16 % and 13 % for the Monte Carlo, Latin Hypercube and importance sampling respectively (e.g. figure 40).

For the largest sample size of 1296, all three methods perform similarly, with differences from the reference mean and standard deviation typically 0.5 % - 2 %.

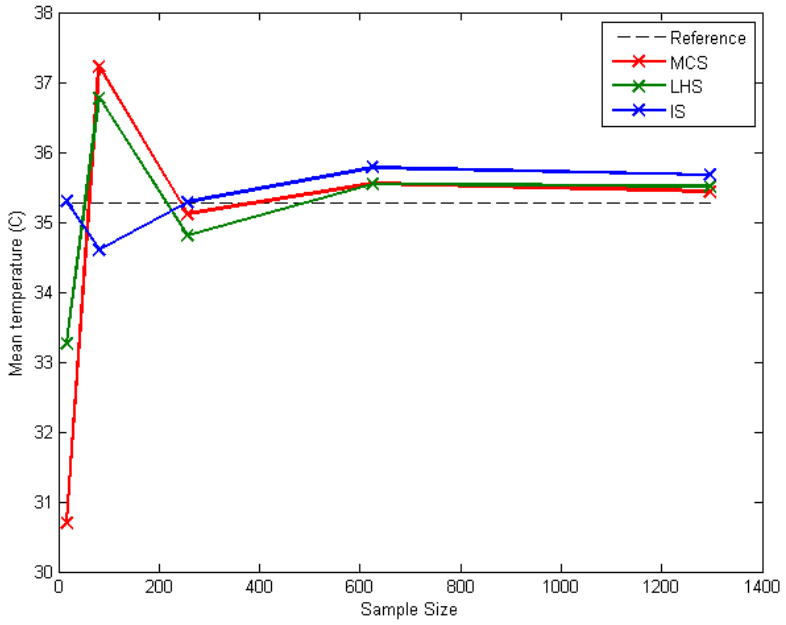

Figure 39: Mean value of the temperature after 32 s for single runs over sample sizes 16, 81, 256, 625 and 1296 using Monte Carlo (MCS), Latin Hypercube (LHS) and importance (IS) sampling. The dashed line shows the reference value calculated from the large scale Monte Carlo simulation. The solid lines are added to provide clearer presentation.

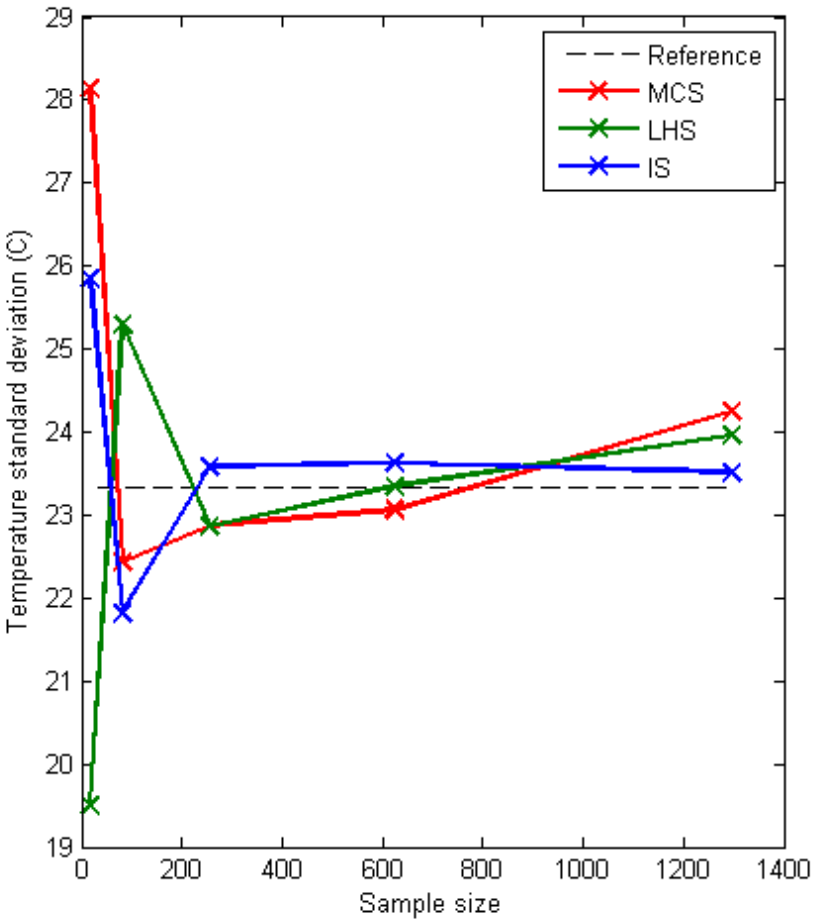

Figure 40: As figure 39, but for the standard deviation of the temperature after 32 s.

5.4.2 Repeatability

For the smallest sample size of 16, the standard deviation over 10 runs of the mean temperatures produced for each time point is consistently smallest for the Latin Hypercube method (e.g. figure 41). For this sample size, the standard deviation of the mean result values using Monte Carlo sampling is more than double that from the Latin Hypercube trials. For the remaining, larger sample sizes, importance sampling performs similarly to Latin Hypercube sampling, both consistently outperforming Monte Carlo sampling. Importance sampling shows a significant improvement in the standard deviation of mean temperature when the sample size is increased from 16 to 81, for all model results. The standard deviations decrease by 84 %, 85 %, 83 % and 72 % with the increase in sample size, for the mean temperature after 8, 16, 24 and 32 s respectively.

While Latin Hypercube sampling also gives the smallest variance, over 10 runs, of the standard deviations of temperature after 8 s, the results over all results are less consistent than those for the variance of the means. For example, at the smallest sample size, importance sampling goes from having the largest variance for temperature after 8 s, to having the smallest variance for the temperature after 32 s (figures 42 and 43).

Plots of the cumulative distribution function for the model results again show that the importance and Latin Hypercube sampling provide a closer fit to the reference distribution for all sample sizes (e.g. figure 44).

5.5 Conclusions

- Metaheuristic algorithms can find solutions where other algorithms fail to converge to the correct solution.
- Convergence criteria can strongly affect the performance of an algorithm but it can be difficult to establish the details of how the criteria affect when the algorithm stops.

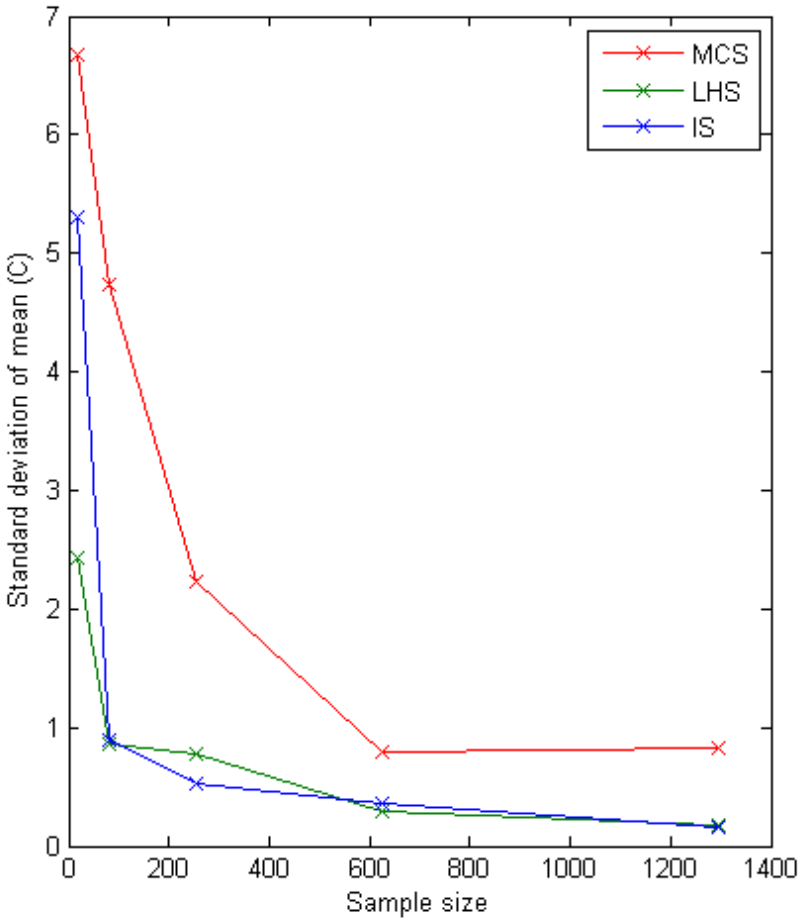

Figure 41: Standard deviation, over 10 runs, of the mean temperature after 16 s for sample sizes 16, 81, 256, 625 and 1296 using Monte Carlo (MCS), Latin Hypercube (LHS) and importance (IS) sampling. The solid lines are added to provide clearer presentation.

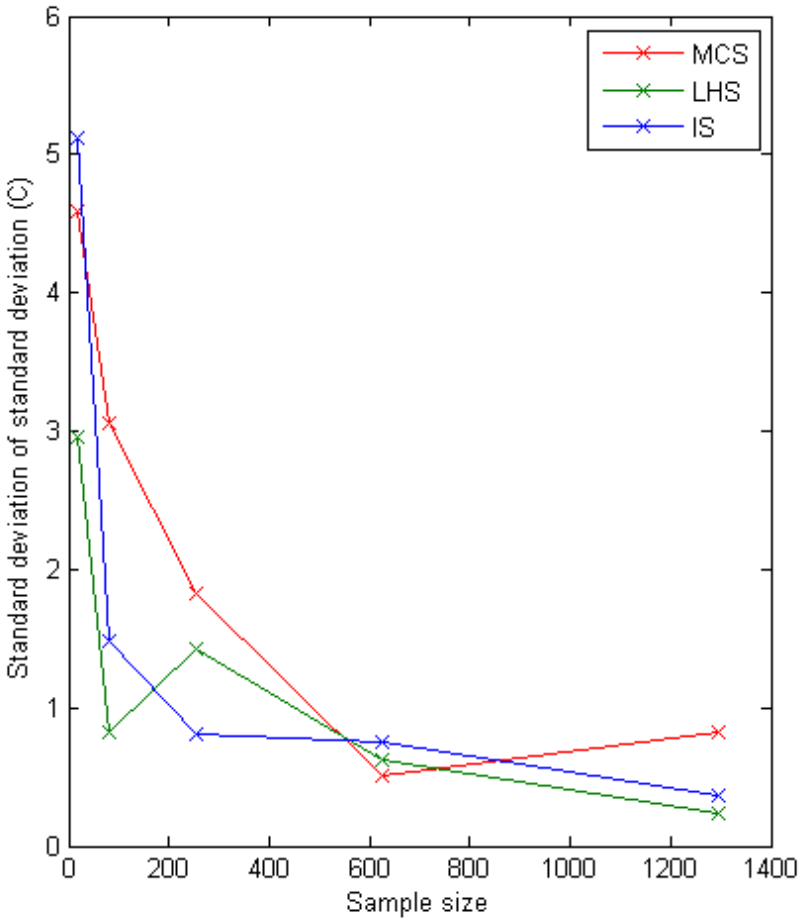

Figure 42: Standard deviation, over 10 runs, of the standard deviation of temperature after 8 s for sample sizes 16, 81, 256, 625 and 1296 using Monte Carlo (MCS), Latin Hypercube (LHS) and importance (IS) sampling. The solid lines are added to provide clearer presentation.

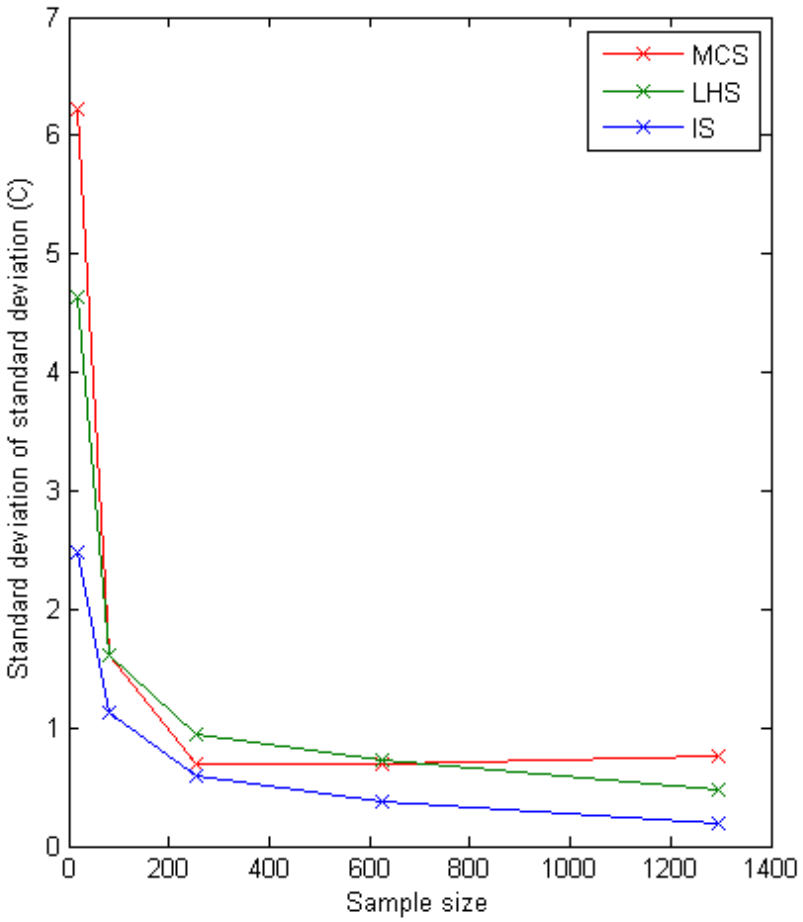

Figure 43: As for figure 42 but for the temperature after 32 s.

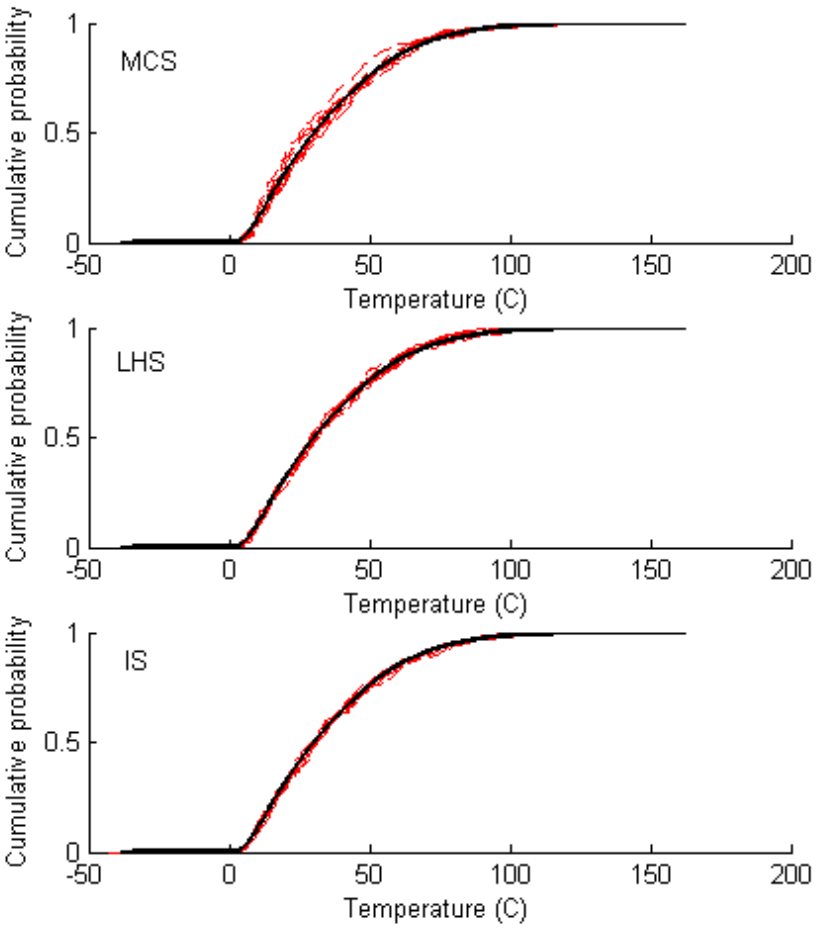

Figure 44: Cumulative probability of temperature after 32 s for each of ten runs with sample size 256, for Monte Carlo (MCS), Latin Hypercube (LHS) and importance (IS) sampling.

6 Conclusions

- Carrying out a sensitivity analysis of a model brings multiple benefits. It allows the key model parameters to be identified in advance of optimisation work, which reduces computational effort and often increases success rates. It also provides information about the smoothness and continuity of the objective function, which enables the user to choose a suitable optimisation algorithm.

- Basing a regression model on the data from a sensitivity analysis can give enough qualitative information to identify which parameters can be neglected. Coloured scatter plots have proved a valuable tool for visualising the results of a full sensitivity analysis and associated regression models.

- Sensitivity indices provide a good measure of global sensitivity, but it should be noted that this is not the same as local sensitivity and so may not be an accurate measure of local behaviour and dependencies. The choice of global or local measures needs to be made based on what information is to be obtained: choosing key parameters may require a global approach where an uncertainty analysis will require local information.

- The three types of index examined in this work (Sobol total sensitivity indices, Jansen total sensitivity indices and Morris one-at-a-time [OAT] indices) gave consistent orderings of importance of the model parameters for all problems studied.

- Of the three sensitivity indices, Morris OAT indices offer more information about the nature of the dependence of a model result on the model parameters than the others, since the mean value of the indices provides an indication of the degree of linear dependence on the parameters and their standard deviation is related to the amount of nonlinear dependence (including interaction with other model parameters).

- Sobol indices can sometimes be uninformative as they can take negative values if the sample size is small or if the model parameter is insignificant. Jansen indices cannot take negative values as they are computed from a sum of squares. Both indices can take values larger than one for small sample sizes.

- For the problems studied here, a comparatively small number of model runs was required to compute indices that rank the model parameters correctly in order of importance. In general, the Morris OAT indices were better than the other indices for small sample size, and had better repeatability than the other indices over repeated runs.

- The performance of optimisation algorithms is very strongly affected by the nature of the objective function. The problems investigated here have all had objective functions based on a sum of squares, and have been smooth and unimodal. These features have meant that algorithms designed for smooth unimodal functions, such as the Levenburg-Marquardt algorithm, have performed well and those designed for more challenging problems, such as metaheuristic algorithms, have not been so efficient. The choice of algorithm should be guided by as much knowledge as possible of the objective function's characteristics. The case studies associated with this work [7, 8] include functions that are neither smooth nor unimodal and may provide a better illustration of the strengths of metaheuristic algorithms.

- As would be expected, problems with a small number of parameters converge more quickly than those with a large number of parameters and are often easier to solve (i.e. algorithms have better success rates on smaller problems). Similarly, constraining parameters, even if it is just by imposing simple limits on their values, reduces computational expense and increases success rates.

- If using a black-box optimisation algorithm within a software package, it is worth simplifying the problem as much as possible by normalising all parameters to be $O(1)$ as far as is possible. Black-box algorithms can exhibit unexpected behaviour (e.g. unnecessary repetition of evaluations).

- The success rate of traditional algorithms can be strongly affected by the choice of the initial estimate. Metaheuristic algorithms are less prone to this problem because one of their key features is that they attempt to maintain a diverse search space throughout.

- The convergence rate of metaheuristic algorithms can be improved by adjusting their internal parameters, but there are no set rules for choosing the best parameter values, and the choice of values is dependent on the nature of the objective function. Values that are good for a smooth unimodal function may restrict the search of a more complicated surface too rapidly.

- For a small sample size, Latin Hypercube sampling and importance sampling give better estimates of the mean, standard deviation, and distribution function of a model result than random sampling. The former methods also have better repeatability over multiple trials.

References

[1] C. Blum and A. Roli. Metaheuristics in combinatorial optimization: Overview and conceptural comparison. *ACM Comput. Surv.*, 35:268–308, 2003.

[2] E. Bonabeau, M. Dorigo, and G. Theraulaz. *Swarm Intelligence: From Natural to Artificial Systems.* Oxford University Press, 1999.

[3] A. Chatterjee and P. Siarry. Nonlinear inertia variation for dynamic adapation in particle swarm optimization. *Comp. Oper. Research*, 33:859–871, 2006.

[4] T. F. Coleman and Y. Li. An interior trust region approach for nonlinear minimization subject to bounds. *SIAM Journal on Optimization*, 6:418–485, 1996.

[5] T. F. Coleman and Y. Li. On the convergence of reflective newton methods for large-scale nonlinear minimization subject to bounds. *Mathematical Programming*, 67:189–224, 1994.

[6] T. J. Esward, C. E. Matthews, L. Wright, and X.-S. Yang. Sensitivity analysis, optimisation, and sampling methods applied to continuous models. NPL Report MS 2, National Physical Laboratory, 2010.

[7] L. A. Chapman, C. E. Matthews, S. J. Roberts, L. Wright, and X.-S. Yang. Tools for continuous modelling case study 1: the laser flash experiment. NPL Report MS 4, National Physical Laboratory, 2010.

[8] T. J. Esward, S. Knox, and L. Wright. Tools for continuous modelling case study 2: organic light-emitting diodes. NPL Report MS 5, National Physical Laboratory, 2010.

[9] G. A. O. Davies, R. T. Fenner, and R. W. Lewis. *NAFEMS Background to Benchmarks.* NAFEMS, Glasgow, 1993.

[10] P.E. Gill, W. Murray, and M. H. Wright. *Practical optimization.* Academic Press Inc, 1981.

[11] P. E. Gill, W. Murray, and M. A. Saunders. An SQP algorithm for large-scale constrained optimization. *SIAM Journal on Optimization*, 12:979–1006, 2002.

[12] S. P. Han. A globally convergent method for nonlinear programming. *Journal of Optimization Theory and Applications*, 22:297–309, 1977.

[13] J. Kennedy and R. C. Eberhart. Particle swarm optimization. In *Proc. of IEEE International Conference on Neural Networks*, pages 1942–1948. IEEE Press, 1995.

[14] J. Kennedy, R. Eberhart, and Y. Shi. *Swarm intelligence*. Academic Press, 2001.

[15] S. Kirkpatrick, C. D. Gellat, and M. P. Vecchi. Optimisation by simulated annealing. *Science*, 220:671–680, 1983.

[16] J. C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright. Convergence properties of the Nelder-Mead simplex method in low dimensions. *SIAM Journal of Optimization*, 9:112–147, 1998.

[17] J. W. Mann and G. D. Smith. A comparison of heuristics for telecommunications traffic routing. In *Modern Heuristic Search Methods*. John Wiley & Sons, 1996.

[18] M. Mitchell. *An introduction to genetic algorithms*. MIT Press, Cambridge, Mass., USA, 1996.

[19] M. J. D. Powell. *A fast algorithm for nonlinearly constrained optimization calculations*. Springer Verlag, 1978.

[20] M. J. D. Powell. The convergence of variable metric methods for nonlinearly constrained optimization calculations. In *Nonlinear Programming 3*. Academic Press, 1978.

[21] A. Saltelli, K. Chan, and E. M. Scott. *Sensitivity Analysis*. Wiley, 2001.

[22] D. H. Wolpert and W. G. Macready. No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation*, 1:67–82, 1997.

[23] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, AND OIML *Evaluation of measurement data — Supplement 1 to the "Guide to the expression of uncertainty in measurement" — Propagation of distributions using a Monte Carlo method.* Joint Committee for Guides in Metrology, JCGM 101: 2008.

[24] M. G. Cox and B. R. L. Siebert. The use of a Monte Carlo method for evaluating uncertainty and expanded uncertainty. *Metrologia* , 43:S178–88, 2006 .

[25] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, AND OIML *Guide to the expression of uncertainty in measurement* Joint Committee for Guides in Metrology JCGM 100: 2008.

[26] MATLAB www.mathworks.co.uk/products/matlab/

[27] NAG http://www.nag.co.uk/

[28] SOFEA http://hogwarts.ucsd.edu/ pkrysl/sofea/publish.html