

**NPL REPORT
DEM-ES-017**

**Deriving uncertainties when
using simulation in
metrology**

T J Esward and L Wright

NOT RESTRICTED

March 2007



The DTI drives our ambition of 'prosperity for all' by working to create the best environment for business success in the UK. We help people and companies become more productive by promoting enterprise, innovation and creativity.

We champion UK business at home and abroad. We invest heavily in world-class science and technology. We protect the rights of working people and consumers. And we stand up for fair and open markets in the UK, Europe and the world.

The National Physical Laboratory is operated on behalf of the DTI by NPL Management Limited, a wholly owned subsidiary of Serco Group plc

Deriving uncertainties when using simulation in metrology

T J Esward and L Wright
Mathematics and Scientific Computing Group

March 2007

ABSTRACT

This report for the Software Support for Metrology programme considers the use of simulations in metrology. It presents three case studies (simulation of piston-cylinder pressure balances using finite element methods, Monte Carlo methods for evaluating the uncertainties associated with the measurements of the absolute thermodynamic temperatures of blackbodies, and molecular dynamics methods) that address the use of simulations to study or derive uncertainties in metrology applications, and uncertainties in simulations themselves. Each case study provides practical advice that can help ensure effective use of simulation techniques in metrology.

Note that most figures in this document are in colour and may best be viewed in the electronic version of this report.

© Crown copyright 2007.
Reproduced with the permission of the Controller of HMSO
and Queen's Printer for Scotland.

ISSN 1744-0475

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

Extracts from this report may be reproduced provided the source is
acknowledged and the extract is not taken out of context.

We gratefully acknowledge the financial support of the UK Department of
Trade and Industry (National Measurement System Directorate).

Approved on behalf of the Managing Director, NPL,
by Jonathan Williams, Knowledge Leader for the Electrical and Software team

Contents

1	Introduction	1
1.1	Motivation for the work	1
1.2	Simulation in metrology	2
1.3	Virtual testing	3
1.4	Molecular modelling	4
1.5	Existing SS/M resources for simulation, modelling and visualisation	5
1.6	Structure of this report	5
2	Current perspectives on simulation and its uses	7
2.1	Introduction	7
2.2	Defining simulation: some observations from the scientific literature on simulation	7
2.3	Simulation-Based Engineering Science: a study by the U.S.A.'s National Science Foundation	9
2.3.1	Definition of simulation	10
2.3.2	Benefits of simulation	10
2.3.3	Challenges	11
2.3.4	Verification and validation	11
2.3.5	Importance of visualisation in simulation	12
2.4	Karniadakis, Uncertainty quantification in simulation science . .	13
2.5	Co-ordinate measurement metrology	14
2.5.1	Baldwin et al: Application of simulation software to co-ordinate measurement uncertainty evaluation	14
2.6	Comparing simulation results with experiment	16
2.6.1	Measures of agreement between computation and experiment: validation metrics	16
3	Simulating pressure balances using ANSYS finite element software	18
3.1	Introduction	18
3.2	How a piston-cylinder pressure balance works	18
3.3	The effective area of a pressure balance	19
3.4	NPL's application of finite element modelling to pressure balances	20
3.4.1	The purpose of finite element simulations in pressure balance metrology	21
3.5	Analytical methods for understanding pressure balances	22
3.6	Why numerical methods are preferred to analytical methods . . .	23
3.7	Key issues for pressure balance simulations	24
3.7.1	Material properties	25
3.7.2	Geometry	26
3.7.3	Boundary conditions	27
3.7.4	Modelling simplifications	27
3.8	Simulating pressure balances using ANSYS finite element software	28
3.8.1	The ANSYS Parametric Design Language	28
3.8.2	Pre-processing	30
3.8.3	The solution and post-processing	30
3.8.4	Analysing the results	32
3.8.5	Modelling the fluid	32

3.9	Typical outputs from the finite element simulations	33
3.9.1	Comparison between finite element results and simple analytical models	34
3.10	Controlled clearance balances	35
3.10.1	Evaluating the uncertainty associated with an estimate of the model output	39
3.11	Reducing uncertainties with a new controlled-clearance design . .	43
3.12	Some limitations of finite element methods	46
3.12.1	Finite element simulations for pressure balance metrology: the pros and cons summarised	47
3.13	Molecular modelling and pressure balances	48
4	Monte Carlo method of investigating experimental uncertainties in a parallel computing environment	49
4.1	Introduction	49
4.2	The physical problem	49
4.2.1	Determining thermodynamic temperature	49
4.3	Experimental design	50
4.4	Simulation using Monte Carlo methods	51
4.4.1	Monte Carlo methods in uncertainty evaluation	51
4.5	Distributed computing and Monte Carlo methods	53
4.5.1	Distributed computing and random number generation . .	55
4.6	Implementing the parallel version of the model on the NPL Grid	56
4.6.1	Implementation in Matlab: the issues	56
4.6.2	Investigation of an alternative seeding strategy	57
4.6.3	Determining the number of work units and the work unit size	57
4.6.4	Grid statistics	57
4.7	Initial results	60
4.7.1	Investigating the effect of the input uncertainties	60
4.7.2	Comparing the two seeding strategies	61
4.8	Conclusions	61
5	Uncertainties in molecular dynamics simulations	64
5.1	Molecular modelling and metrology	64
5.2	Introduction to molecular dynamics simulations	65
5.3	Model building	65
5.3.1	The example problem	65
5.3.2	The inter-atomic potential	66
5.3.3	Initial conditions	66
5.3.4	Non-dimensionalisation	68
5.4	Model solving	69
5.4.1	Numerical solution method	69
5.4.2	Implementation of the potential function	71
5.5	Post processing	72
5.6	Sources of uncertainty	74
5.7	Results	75
5.7.1	Effects of varying time step	75
5.7.2	Effects of varying cut-off radius and skin thickness	77
5.7.3	Effects of varying the initial conditions	81

5.7.4	Handling correlation between states	84
5.8	Conclusions	86
6	Conclusion	92
6.1	Simulating pressure balances	92
6.2	Tackling a large Monte Carlo problem in temperature metrology	93
6.3	Uncertainties in molecular dynamics simulations	93
6.4	A caveat	94
	References	95

1 Introduction

This report has been produced for the National Measurement System Directorate as part of the 2004-2007 Software Support for Metrology (SS/M) programme. It considers *simulations* and their application to problems in metrology and aims to address by means of case study examples two related topics:

- The use of simulations to study or derive uncertainties in metrology applications, that is, to identify and quantify items that should appear in the uncertainty budgets of metrology experiments;
- Uncertainties in simulations themselves: how do we associate an uncertainty with the outputs of simulations or how reliable are the outputs of simulations?

1.1 Motivation for the work

The report considers the use of simulations in metrology and includes guidance on the evaluation of uncertainty contributions associated with simulations and the use of simulation software packages. The need for the work was identified during a study of new developments in mathematics and computing that was carried out during the previous SS/M programme [15]. This study reviewed a report produced by the Smith Institute for Industrial Mathematics and System Engineering [26] that considered developments in industrial mathematics that might be required over the next ten years. The report identified a strong requirement for more support for industry in the use of simulation to enable many design options to be studied without incurring the high costs of prototype development. This message found strong resonance amongst the metrology community during the consultation phase of formulation of the 2004-2007 SS/M programme. Discussions at the appraisal stage of formulation of the programme decided that the highest priority was to focus on the related uncertainty issues. This would then lay the foundations for longer-term work on simulation within SS/M to meet key future industrial requirements, as identified by the Smith Institute study, to the extent that they can be expected to arise in metrology. It was also considered important to make available to the potential user base for simulation within metrology, information that will enable users to recognise and tackle questions related to uncertainties and simulations.

The Smith Institute roadmap identified simulation as an essential component of modern science and engineering, particularly given the development of cheaper and more powerful computers, so that simulation is used not only to support but often to replace experimentation. The traditional manner in which scientists and engineers attempt to validate their theories and the results of model predictions is the comparison of theoretical predictions with experimental results. When no experiment exists or when the modelling and simulation is done to replace experimental work, this practical approach to

validation is no longer available. However, there is a continuing need for greater physical realism in simulations, the Institute argued, but many simulations lack sufficient realism because good models do not exist, or because model data are uncertain or incomplete. However, as the report's authors point out, improving the realism of a simulation usually leads to a longer time being needed to carry it out. Expensive models on fine scales often need to be replaced with effective coarse-scale descriptions when computation on the fine scale is not feasible. Thus, hierarchical models may be of use, so that a small, fine-scale simulation may be used to verify a much larger coarse-scale model. This topic will be investigated in more detail in the 2007-2010 SS/M programme.

In industry traditional approaches to new product development have involved the use of the “make and test” philosophy, which necessitates expensive and time-consuming activities of prototype manufacture and testing, followed by a number of design iterations. It has long been recognised by industry that competitiveness could be substantially improved if it were possible to develop computer-based methods of modelling, simulating and optimising the manufacture and performance of products before prototypes are manufactured.

1.2 Simulation in metrology

Simulation of experiments is also becoming more common within the metrology community. The main reasons for this development are listed below.

- Simulations can be used to assist in experimental design.
- They can provide information about behaviour inside an object or structure that is not accessible experimentally.
- Simulations can assist in understanding the physical processes that take place in the course of an experiment.
- Theoretical assumptions and simplifications can be tested and their consequences evaluated. This can assist in identifying and quantifying terms in experimental uncertainty budgets.
- A common use of simulations is to investigate whether useful predictions can be made from measurement data.
- Simulations also allow one to study how sensitive the output of a model is to changes in input data.

The key mathematical issues that arise from a consideration of the use of simulation in metrology were identified in [15] as:

- How does one validate simulations of complex experimental systems, especially when those systems may be unique, as may be the case in some state-of-the-art metrology applications?

- Minimising sources of uncertainty when realising primary measurement standards can be an extremely challenging task. If simulations are to be used in helping to evaluate uncertainties, what contribution to uncertainty budgets do simulations themselves make?
- Simulations frequently rely on the output of complex commercial software packages. What contribution to the “simulation uncertainty” arises from the use of these software packages?

1.3 Virtual testing

A specific set of activities that form part of the discipline of simulation are those that are associated with virtual testing. Such methods have proved to be particularly popular among materials scientists and engineers who want to avoid the need to carry out extensive physical tests on large numbers of samples. Processes can be evaluated and optimised entirely by exercising computer models. Substantial savings in time, in the purchase and disposal of materials, in human resources and money can be achieved by reducing the number of physical tests performed by private industry. A successful example of such an approach within a national measurement institution is the Virtual Cement and Concrete Testing Laboratory of the United States’ National Institute of Standards and Technology (NIST). The virtual laboratory is based on a computer model for the hydration and microstructure development of cement-based systems developed over many years of research at NIST. More information can be found on the NIST web site at <http://www.bfrl.nist.gov/861/vcctl/>.

NPL’s material scientists have developed models to allow virtual testing of the stress-strain behaviour of composite materials. An important aspect of the models is the prediction of the onset of damage and the evolution of material properties as damage increases. These models allow the engineer to assess the suitability of the material for a particular component, like a wing member, and to fine tune values of material parameters.

One key concern when using virtual testing methodologies is the ease in which the effects of changes in material parameters can be examined. If the mathematical model is of the finite element (FE) type then there will be an inevitable wait while the solution is calculated and there will be some difficulty in changing material parameters, owing to the complexity of finite element software. There are now alternative mathematical models that will predict laminate behaviour quickly enough to be truly interactive and with the addition of a graphical interface, user friendly interaction with the model input parameters is possible. More information about NPL’s work here can be found at <http://materials.npl.co.uk/netshare/guest/virtualtesting/KNvirtualtesting.html>.

1.4 Molecular modelling

Much simulation in engineering disciplines and in measurement science has been based on continuous modelling methods, i.e., differential equations, partial differential equations, and integral equations, and solution techniques such as finite element, finite difference and boundary element methods. However, driven by the need to simulate at the nanoscale and to be able to use better and more precisely defined materials properties as inputs to simulations, metrologists are turning increasingly to molecular modelling to assist them in their work. More details about the role of mathematics in nanoscale metrology at NPL can be found in [21]. This report therefore includes a case study on the use of molecular dynamics methods. The case study is intended to be an illustration of potential sources of uncertainty for those who are new to molecular modelling, rather than an examination of issues that concern experienced modellers.

Such models typically calculate the individual trajectories (positions and velocities) of a large number of atoms or molecules, and then use functions of the kinetic energy and potential energy of the entire system to calculate bulk properties such as temperature and pressure. The trajectories of the atoms are obtained by calculating the force acting on each atom, applying Newton's second law of motion to obtain the acceleration of the atom, and using finite difference techniques over a small time step to calculate the new position and velocity of the atom. The forces on the atoms are caused by interaction with other atoms, described by a potential energy of interaction, and possibly by external fields. The simplest models have no external fields and only consider the interactions between pairs of atoms. Molecular dynamics simulations require initial conditions because they are transient models. The initial positions and velocities of the atoms within the system must be specified because the simulation uses a second-order time derivative (i.e., acceleration). The initial positions and velocities have an associated potential energy and kinetic energy, defining the initial total energy of the system. In equilibrium simulations, it is assumed that the total energy remains constant, and the system evolves until it eventually reaches an equilibrium state where temperature and pressure are approximately constant. This temperature may be a free parameter fixed by the energy of the initial conditions, or it may be a fixed parameter controlled by rescaling the atomic velocities using a multiplicative factor. In both cases, the equilibrium temperature (and hence pressure) value will not be achieved precisely: instead, the temperature will fluctuate around its equilibrium value. From the descriptions above, it is clear that molecular models involve a number of approximations and assumptions that will lead to uncertainties being associated with the results. The contributions to the uncertainty include:

- uncertainty arising from the application of finite difference methods to a continuum problem,
- uncertainty arising from the simplifying assumptions and numerical techniques used during model development (e.g. cut-off radius, rescaling of the velocities for fixed temperature), and

- uncertainty associated with bulk properties (e.g. temperature, pressure) owing to their oscillation around the equilibrium value, and the effects of the initial conditions on this value.

Each of these effects has been examined in more detail in the simulation case study that is described in chapter 5 of this report.

1.5 Existing SSfM resources for simulation, modelling and visualisation

The Software Support for Metrology programme has produced a wide range of reports and guides on simulation, modelling, uncertainties and visualisation, all of which are relevant to the subject of this report. They contain detailed advice and suggestions to ensure good practice and readers should consult them if they require more information than can be provided here. There are two good practice guides, one on continuous modelling [23] and another on data visualisation [25], and several reports on a range of topics including uncertainty evaluation in continuous modelling, testing and validating continuous modelling software and visualisation and uncertainty [16, 17, 18, 22, 36, 59]. The guides and reports can be downloaded from the Software Support for Metrology web site at <http://www.npl.co.uk/ssfm/download/>. In addition, we offer a training course on the simulation of experiments that provides practical advice on developing, solving and validating models, based on the contents of these guides and reports.

1.6 Structure of this report

Chapter 2 reviews some important recent papers in simulation science that help to define what simulation is and to identify how one can test and validate simulation results, so that users can have confidence in the outputs of simulations.

Three case studies are then described that exemplify aspects of simulation and uncertainty in metrology. The first case study is presented in chapter 3. This sets out an account of the use of finite element modelling in the simulation of piston-cylinder pressure balances. An important parameter that describes the way that pressure balances behave under load is the distortion coefficient. Finite element simulations are currently the preferred method of evaluating uncertainties of distortion coefficients and this chapter describes in detail the advantages and disadvantages of the simulation techniques that are employed in pressure balance metrology. Chapter 4 describes the application of Monte Carlo methods to the study of an experimental design problem in metrology. The purpose of the Monte Carlo analysis is to quantify the sensitivity of the output of a simulation to changes in inputs, as a mean of identifying the most productive areas in which to concentrate experimental design work to minimise measurement uncertainties in the experiment itself. The final case study in

chapter 5 introduces the use of atomic scale simulations in metrology, and demonstrates that even apparently simple simulations can present substantial challenges to the modeller. Conclusions are then summarised in chapter 6.

2 Current perspectives on simulation and its uses

2.1 Introduction

This chapter reviews a range of publications on simulation. The aim here is to introduce the reader to some of the debates on simulation and the validation of simulations. These provide the basis for understanding the work reported in the case study chapters 3, 4 and 5 of this report.

2.2 Defining simulation: some observations from the scientific literature on simulation

It is necessary first to attempt to define simulation, or at least to distinguish it from mathematical modelling in general. In a text that is now more than 40 years' old Chorafas [7] makes the following points:

Simulation is essentially a working analogy. . . . Simulation involves the construction of a working mathematical or physical model presenting similarity of properties or relationships with the natural or technological system under study. In this manner we can pre-operate a system without actually having a physical device to work with, and we can pre-decide on the optimization of its characteristics.

He then summarises the uses to which simulation has been put:

- For the purposes of experimentation or evaluation, that is, trying to predict the consequences of changes in policy, conditions, or methods without having to spend money or take the risk of actually making the change in real life;
- As a means of learning about new systems in order to redesign or refine them;
- As a tool in familiarising personnel with a system or a situation which may, as yet, not exist in real life;
- For the verification or demonstration of a new idea, system or approach;
- As a means of projecting into the future and thus providing quantitative bases for planning and forecasting.

These bear some similarity to the list set out in section 1.2 in the brief discussion of uses of simulation in metrology.

Chorofas then goes on to argue that the problem of mathematically describing a practical situation in precise and meaningful terms is often referred to as “model-making”. In his view a mathematical model or simulation should be:

- simple enough for manipulation and understanding by those who would use it,
- representative enough, in the total range of the implications it may have, and
- complex enough to accurately represent the system under study.

Naylor et al [42] in a textbook on computer simulation techniques make a series of useful observations about simulation:

- Simulation makes it possible to study and experiment with the complex internal interactions of a given system whether it be a firm, an industry, an economy or some subsystem of these.
- Through simulation one can study the effects of certain informational, organisational, and environmental changes on the operation of a system by making alterations in the model of the system and observing the effects of these alterations on the model output.
- Detailed observation of the system being simulated may lead to a better understanding of the system and to suggestions for improving it, which otherwise would not be obtainable.
- Simulation can be used as a device for teaching.
- The experience of designing a computer simulation model may be more valuable than the actual simulation itself. The knowledge obtained in designing a simulation study frequently suggests changes in the system being simulated.
- Simulation of complex systems can yield valuable insight into which variables are more important than others in the system and how these variables interact.
- Simulation can often be used to experiment with new situations about which we have little or no information so as to prepare for what may happen.
- Simulations can serve as a “pre-service test” to try out new policies and decision rules for operating a system, before running the risk of experimenting on the real system.
- Simulations provide a convenient way of breaking down a complicated system into sub-systems each of which may then be modelled by an analyst or team which is expert in that area.
- Monte Carlo simulations can be performed to verify analytic solutions.

- Simulation enables one to study dynamic systems in either real time, compressed time, or expanded time.
- When new components are introduced into a system, simulation can be used to help foresee bottlenecks and other problems that may arise in the operation of the system.
- Simulation makes generalists out of specialists. Analysts are forced into an appreciation and understanding of all facets of the system, with the result that conclusions are less apt to be biased by particular inclinations and less apt to be unworkable within the system framework.

These authors then state that,

No substantial part of the universe is so simple that it can be grasped and controlled without abstraction. Abstraction consists in replacing the part of the universe under consideration by a model of a similar but simpler structure. Models ... are thus a central necessity of scientific procedure.

Baird [2] in his thesis on simulation of co-ordinate measurement machines summarises the key aspects of the modelling and simulation process. The process is essentially one of abstraction, the elimination of irrelevant detail based on information gathered. Three aspects of abstraction are identified:

- Generalisation: requires a model to be simplified to represent a more general class of systems.
- Approximation: applied for cases in which a simpler description is used in place of a more detailed one.
- Perspectivisation: specific aspects only are considered.

Vvedensky [56], in a review of multiscale modelling of nanostructures, points out that experimental methods of materials design and synthesis are costly and time-consuming, leading to long development times. Modelling and simulation can be used to optimise processing parameters and to interpret results, and that this is especially important where phenomena are too fast or too complex to image. In his own field of research he argues that simulations based on large-scale quantum mechanical calculations and molecular dynamics can be stopped and examined at any time.

2.3 Simulation-Based Engineering Science: a study by the U.S.A.'s National Science Foundation

In May 2006 a report entitled "Simulation-based Engineering Science: revolutionizing engineering science through simulation" was published by the U.S.A.'s National Science Foundation Blue Ribbon Panel on Simulation-based

Engineering Science [44]. The complete report can be accessed on-line at http://www.nsf.gov/pubs/reports/sbes_final_report.pdf¹. The report reviewed the potential impact on science and engineering of advances in simulation technology and identified challenges and barriers to further advances. It also made specific recommendations to the United States' government on future funding for research in simulation.

Set out below are extensive extracts from the report as these help provide the context both for discussions later in this chapter and for the case studies in the remainder of this report.

2.3.1 Definition of simulation

To supplement the general definitions of simulation that were included in section 2.2, set out below is the definition employed by the authors of the National Science Foundation report

“Simulation refers to the application of computational models to the study and prediction of physical events or the behaviour of engineered systems. The development of computer simulation has drawn from a deep pool of scientific, mathematical, computational, and engineering knowledge and methodologies. With the depth of its intellectual development and its wide range of applications, computer simulation has emerged as a powerful tool, one that promises to revolutionize the way engineering and science are conducted in the twenty-first century.

Computer simulation represents an extension of theoretical science in that it is based on mathematical models. Such models attempt to characterize the physical predictions or consequences of scientific theories. Simulation can be much more, however. For example, it can be used to explore new theories and to design new experiments to test these theories. Simulation also provides a powerful alternative to the techniques of experimental science and observation when phenomena are not observable or when measurements are impractical or too expensive.

Simulation-Based Engineering Science (SBES) is defined as the discipline that provides the scientific and mathematical basis for the simulation of engineered systems.”

2.3.2 Benefits of simulation

The report emphasises that, unlike most theory, which relies on restricted idealised systems, simulation deals with real systems and therefore provides unprecedented access to real-world conditions. There are economic benefits to simulation, too, as it suffers none of the limitations of iterative experimental design and test processes, which may be hampered by cost constraints,

¹Accessed 3rd October 2006.

unrealistic parameter ranges, and restrictions imposed by treaties or health and environmental concerns.

The authors state that:

“For these reasons, computer simulation is credited with numerous triumphs in the twentieth century. It has become indispensable, for example, in assessments of vehicle crashworthiness. It is fundamental to the generation of predictive models of the weather, climate change, and the behaviour of the atmosphere. Its importance in broad areas of engineering analysis and design are well known. It has become essential to product manufacturing. Its achievements in biomedical applications are widely discussed. Systems design in defence, communication, and transportation also rely on computer simulation.”

2.3.3 Challenges

The report recognises that multiscale and multiphysics modelling will be important for the future development of simulation in science and engineering.

“Formidable challenges stand in the way of progress in SBES research. These challenges involve resolving open problems associated with multiscale and multi-physics modelling, real-time integration of simulation methods with measurement systems, model validation and verification, handling large data, and visualization. Significantly, one of those challenges is education of the next generation of engineers and scientists in the theory and practices of SBES.”

The consequences for metrology of the new developments in multiscale and multiphysics modelling are also reviewed in a recent SS/M report by Esward and Sokhan [21].

2.3.4 Verification and validation

The National Science Foundation report then discusses at some length the need for verification and validation of simulation software, recognising that if industry is to replace testing with simulation, then simulation tools must undergo robust verification and validation procedures for effectiveness.

In relation to benchmark tests for simulation software the authors comment:

“Although some efforts have been made at providing validation benchmark problems for linear analysis, non-linear simulation software has not been subjected to extensive validation procedures. In fact, we find considerable controversy as to what appropriate

validation procedures are, how broadly they apply, and whether they are even feasible. Clearly, a basic understanding of verification and validation procedures is urgently needed. After all, to be useful, the simulation tools used by industry and defence agencies must provide reliable results. Furthermore, since many real-world phenomena are not deterministic, statistical methods that can quantify uncertainty will be needed.”

The difficulties inherent in attempting to validate simulation software are summarised in the report.

“Beginning with the conceptual understanding of certain physical events of interest and with scientific theories that explain them (the target physical phenomena or engineering system identified for study), the analyst (the modeller, scientist or engineer) constructs a mathematical model of the event. The mathematical model is a collection of mathematical constructions, equations, inequalities, constraints, etc., that represent abstractions of the reality, and are dictated by the theory or theories characterizing the events. The analyst then develops a computational model of the event. The computational model is a discretized approximation of the mathematical model, and its purpose is to implement the analysis on a computer. Validation is the subjective process that determines the accuracy with which the mathematical model depicts the actual physical event. Verification is the process that determines the accuracy with which the computational model represents the mathematical model. In simple terms, validation asks, “Are the right equations solved?” while verification asks, “Are the equations solved correctly?”. . . . Ultimately, the most confounding aspect of V&V has to do with uncertainty in the data characterizing mathematical models of nature. In some cases, parameters defining models are determined through laboratory tests, field measurements, or observations, but the measured values of those parameters always vary from one sample to another or from one observation to the next. Moreover, the experimental devices used to obtain the data can introduce their own errors because of uncontrollable factors, so-called noise, or errors in calibration. For some phenomena, little quantitative information is known, or our knowledge of the governing physical processes is incomplete or inaccurate. In those cases; we simply do not have the necessary data needed to complete the definition of the model.”

2.3.5 Importance of visualisation in simulation

The report’s authors argue that the ability to visualise simulation results is fundamental to our understanding of simulation outputs because it reduces and refines data streams rapidly, and allows large volumes of data to be

understood rapidly because human beings can “comprehend visual representations of data much more rapidly than they can digest the raw numbers or text.”

However, the report goes on to recognise that scientists need to associate uncertainty statements with both measurements and simulation results and that methods of representing uncertainty in visualisations are not well-developed. The authors argue:

“Visualization research must continually respond to and address the needs of the scientific community. For example, the ability to visualize measures of error and uncertainty will be fundamental to a better understanding of three-dimensional simulation data. This understanding will allow the validation of new theoretical models, improve the interpretation of data, and facilitate decision-making. With few exceptions, however, visualization research has ignored the need for visual representation of errors and uncertainty for three-dimensional visualizations . . . We need to create an SBES visualization framework for uncertainty and to investigate new visual representations for characterizing error and uncertainty.”

Advice from the Software Support for Metrology programme on the uses of visualisation in metrology can be found in [25, 59].

2.4 Karniadakis, Uncertainty quantification in simulation science

Karniadakis and Glimm in a recent paper in the Journal of Computational Physics [28] make some useful observations about uncertainty quantification in simulation. They argue that two key aspects of uncertainty are numerical uncertainty and physical uncertainty. The former includes spatial and temporal discretisation errors, errors in solvers, geometric discretisation and artificial boundary conditions. Physical uncertainty includes errors due to unknown or imprecise material properties, boundary and initial conditions, random geometric roughness, equations of state and constitutive laws. Uncertainty can also be characterised as epistemic (arising from insufficient data) or irreducible. Their examples are:

- Epistemic uncertainty: given current rapid advances in quantitative imaging technologies, for example, the rock permeability of a reservoir could be measured much more accurately in the future - this is an example of epistemic uncertainty.
- Irreducible uncertainty: in many cases uncertainty is irreducible beyond some scale or level, for example, background turbulence - there are no absolutely quiet wind tunnels and the atmosphere and the ocean are inherently noisy environments.

2.5 Co-ordinate measurement metrology

Co-ordinate measurement metrology is an important activity for many national measurement institutions such as the UK's National Physical Laboratory (NPL) [24] and Germany's Physikalisch-Technische Bundesanstalt (PTB). Co-ordinate measuring machines (CMMs) are valuable and versatile metrology tools which are very popular throughout industry. They can be based on contact or non-contact measuring systems and can be used to measure a wide variety of features and dimensions on components, assemblies and parts. More information about NPL's co-ordinate measurement metrology work can be found at

<http://www.npl.co.uk/length/dmet/science/cmm-technology.html>. To understand how uncertainties can be evaluated in co-ordinate measuring machines, the concept of the Virtual CMM has been developed by some laboratories and software providers. A virtual CMM is a simulation of a CMM's operation, often accompanied by visualisation, that allows the whole of a measurement process to be modelled, from the choice of probe and position of probing points to the uncertainties to be expected from a particular measurement on a particular component. Information about the PTB's virtual CMM can be found at <http://euromet.ptb.de/en/publikationen/news/html/news032/artikel/03206.htm>

2.5.1 Baldwin et al: Application of simulation software to co-ordinate measurement uncertainty evaluation

Baldwin et al [3] consider task specific uncertainties in co-ordinate measurement metrology and how these can be embodied in virtual CMM software. They point out that an attractive feature of CMMs is their extreme versatility but this leads to difficulty in stating measurement uncertainty so that task-specific uncertainty evaluation is necessary but difficult.

The authors begin with a general review of methods of uncertainty evaluation, pointing out that there are five main ways in which uncertainties can be assessed.

- Sensitivity analysis: list each uncertainty source, its magnitude and effect on results and its correlation with other uncertainty sources, then combine them in a manner that accounts for the sensitivity factor of each source - this is the standard approach of the Guide to the Expression of Uncertainty in Measurement [5] and is useful if one can obtain a model of the process.
- Expert judgement: this may be the only available method if a mathematical model or measurement data are not available.
- Substitution: repeated measurement of a calibrated master yields a range of errors and therefore uncertainty.
- Computer simulation: this is where a virtual model of a CMM is created including all measurement effects and is employed, as explained above,

to examine task-specific results to determine bias and measurement variability.

- Measurement history: this can be useful if large number of measurements over time are available.

They comment on sensitivity analysis, as follows:

“Sensitivity analysis is rated questionable as regards tractability and comprehensiveness, due to the need for explicit information on the standard deviation and sensitivity factor for every uncertainty source and on the correlation between every pair of uncertainty sources. In some cases sensitivity coefficient calculation is impossible, since the measuring process cannot always be analytically described. We rate it as poor as regards cost due to the labour-intensive nature of the process and poor from a versatility perspective since much of the analysis is unique for each individual application.”

However, they ignore the fact that the attempt to understand specific effects is what drives new advances in metrology, especially for Type B sources of uncertainty.

The authors then provide a sample list of CMM measurement influence variables. This shows the the range of information and detail that is required for effective simulation of a complex engineering device such as a CMM.

- CMM geometry: rigid body errors, quasi-static errors, dynamic errors, scale resolution.
- Sensor system: probe type, stylus configuration, calibration strategy, stylus bending, approach velocity, probe repeatability, lobing, index head repeatability, scanning force and speed.
- Environment: thermal effects, external vibration, humidity, atmospheric pressure, power and other utility variations, lighting, ventilation system.
- Workpiece factors: systematic form error, distortion by fixturing.
- Sampling strategy: number and location of sampling points.
- Data analysis: fitting algorithm choice.

They also emphasise the economic aspect of uncertainty analysis: lower uncertainties may imply higher costs, and larger uncertainties may be achievable at lower cost.

2.6 Comparing simulation results with experiment

A recent paper by Oberkampf and Barone [43] discusses the difficulty of comparing simulation results with experiment, especially where experimental results are used as part of a model validation process. We include it here for the benefit of readers who may wish to develop their own methodologies and metrics for comparing experimental results with simulation outputs.

2.6.1 Measures of agreement between computation and experiment: validation metrics

Oberkampf and Barone consider methods by which models and simulations may be validated. They argue that the logical first step of code verification and numerical error estimation should be completed before model validation activities are conducted or at least before actual comparisons of computational results against experimental results are carried out. It is necessary to have convincing evidence that the computational results obtained from the code reflect the physics assumed in the models implemented in the code and that the results are not distorted or polluted by coding errors or large numerical solution errors. Thus all validation activities should include evidence of code verification and numerical error estimation. To achieve the most value from the validation “experiment” as they call it, there should be in-depth, forthright, and frequent communication between the modellers and the experimentalists during the planning and design of the experiment. They state that validation from a practical or engineering perspective is *not* a philosophical statement of truth. They distinguish between verification and validation in the following way: verification deals with mathematics, validation deals with physics.

Validation should be based on agreed metrics that allow quantitative statements to be made about the model and model results being tested. They propose the following general approach.

- A metric should either explicitly state an estimate of the numerical error in the SRQ (system response quantity) of interest resulting from the computational simulation or exclude the numerical error in the SRQ of interest only if the numerical error was previously estimated, by some reasonable means, to be small.
- A metric should be a quantitative evaluation of predictive accuracy of the SRQ of interest, including all of the combined modelling assumptions, physics approximations, and previously obtained physical parameters embodied in the computational model.
- A metric should include either explicitly or implicitly an estimate of the error resulting from the post-processing of the experimental data to obtain the same SRQ that results from the computational model.
- A metric should incorporate or include in some explicit way an estimate of the measurement errors in the experimental data for the SRQ that are the basis of the comparison with the computational model.

- A metric should depend on the number of experimental measurements that are made of given SRQ of interest.
- A metric should exclude any indications either explicit or implicit of the level of adequacy of agreement between computational and experimental results. This lies outside the requirements of a validation metric and may involve value judgements.

3 Simulating pressure balances using ANSYS finite element software

3.1 Introduction

This chapter of the report presents a detailed case study of the application of simulation in pressure balance metrology. It provides an example of how simulation can provide information about the behaviour of pressure balances that can be incorporated directly into measurement uncertainty budgets. The case study also shows how the simplifications built into the simulation methods (both analytical and finite element) traditionally employed in this area of metrology were found to be inadequate for certain types of pressure balance behaviour, which led to a recognition of the need to employ atomistic simulation methods for the first time in pressure balance metrology. A further outcome from the simulation process was the discovery of a pressure balance geometry that had the potential to reduce substantially uncertainties associated with pressure balance measurements.

3.2 How a piston-cylinder pressure balance works

In mechanical metrology piston-cylinder pressure balances are used to generate accurately-known static pressures in the pressure range from 0.1 MPa to 1 GPa. Pressure balances consist of a piston located within a closely fitting cylinder. The radius of the piston is typically a few millimetres and the external radius of the cylinder may be several times that of the piston. The clearance between the piston and cylinder is of the order of one micrometre or less. The gap between the piston and cylinder is filled with fluid, which flows along the gap under the influence of a pressure gradient. This pressure gradient arises because the lower end of the piston and cylinder assembly is connected to a pressurised reservoir of fluid (in practice this may be another pressure balance) and the top of the assembly is exposed either to atmospheric pressure or to a vacuum. The force on the piston generated by the application of the pressurised fluid is balanced by weights applied to the top of the piston.

Pressure balances are used to generate reference pressures from the fundamental relationship

$$p = \frac{Mg}{A_{\text{eff}}}, \quad (1)$$

where M is the total mass of piston plus all weights placed on it, g is the local acceleration of free fall, and A_{eff} is the *effective area* of the piston-cylinder unit. (See figure 1 for a photograph of a pressure balance manufactured by Ruska and figure 2 for a schematic diagram of a piston-cylinder system.)

The accurate determination of the effective area is a difficult problem since it depends on many parameters including surface properties of the piston and cylinder material, temperature and pressure itself, and is the primary aim of the current research. The measuring element in the balance is a cylindrical

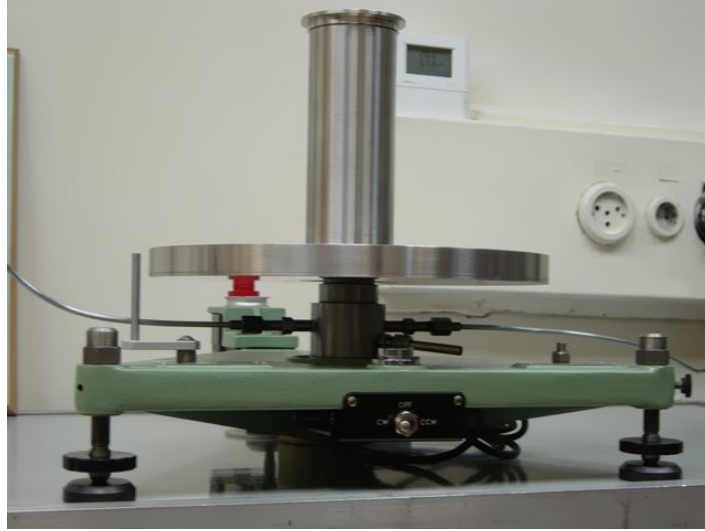


Figure 1: Photograph of a piston-cylinder pressure balance manufactured by Ruska. Source <http://www.inm.ro/en/?page=labs&lab=lab02>.

piston rotating freely in a closely matching cylinder, lubricated by the fluid that escapes from the space below the piston through the small clearance between piston and cylinder bore. The gravitational force exerted by the piston mass and a concentric stack of annular weights acts on the piston in the direction of its axis. This force is balanced by the force exerted by the pressure acting on the “effective” bottom surface of the piston, that is, on the effective area of the piston-cylinder unit. To enable the measurement of absolute pressures, the piston-cylinder unit and the weights are mounted inside a bell jar that can be evacuated. Allowance must be made for the residual pressure p_{res} under the bell jar. If $p_{\text{res}} = p_{\text{ambient}}$, the pressure balance measures gauge pressures, and buoyancy corrections must be applied to the floating parts.

3.3 The effective area of a pressure balance

The canonical work on pressure balance metrology is *The Pressure Balance: theory and practice* [9]. The authors show that for simple piston-cylinder geometries, the application of elastic distortion theory leads to the following equation for the effective area of a pressure balance:

$$A_{\text{eff}} = A_0(1 + \lambda P), \quad (2)$$

where A_{eff} is the effective area of the piston-cylinder combination, A_0 is the effective area at atmospheric pressure, P is the applied pressure and λ is the distortion coefficient. Where the radii of the piston and cylinder vary as a function of position along the engagement length of the cylinder, but circular symmetry about a common vertical axis is retained, Dadson et al [9] derive equations that allow the distortion coefficient to be calculated by means of

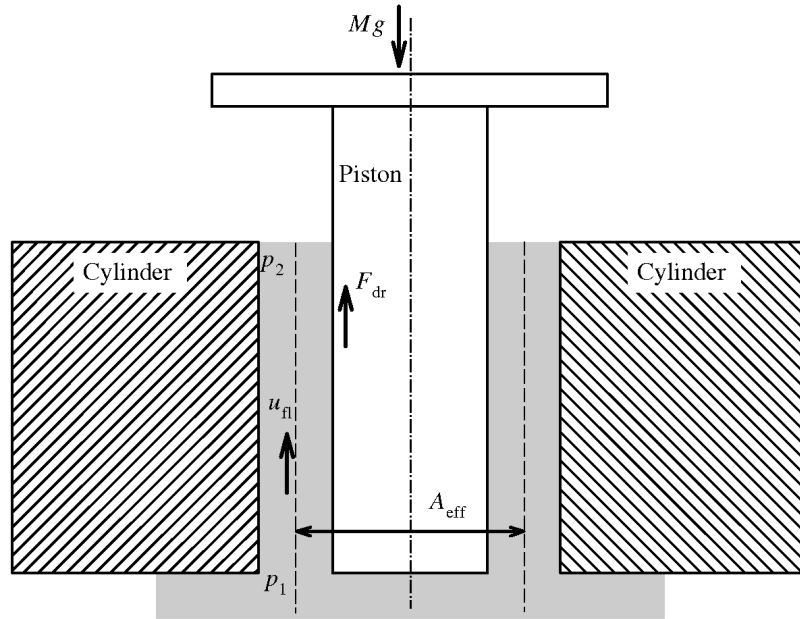


Figure 2: Schematic of a simple piston-cylinder assembly of the pressure balance. The fluid flow u_{fl} in the clearance between piston and cylinder induces a drag force F_{dr} on the vertical surface of the piston.

integration along the engagement length (the active part of the balance where the piston engages with the cylinder). NPL has for many years used finite element methods to predict the pressure distribution along the engagement length of the piston and cylinder, and consequently the local distortion of the piston and cylinder components. The distortion coefficient is obtained by using the outputs of the modelling as inputs to the relevant equations set out in [9].

The history of NPL's use of finite element methods to simulate pressure balances goes back to the end of the 1980s [13, 40, 47, 49, 50, 51, 52, 53] and readers are referred to these publications for more detail about our work over the past 15 to 20 years.

3.4 NPL's application of finite element modelling to pressure balances

NPL's preferred approach to the finite element modelling of pressure balances is to employ ANSYS finite element software to calculate both the pressure distribution in the fluid and the distortion of the individual pressure balance components. A suite of macros in the ANSYS Parametric Design Language has been developed to handle all pre-processing, problem solution, and post-processing aspects of the finite element analysis. The method has been standardised so that it is only necessary to produce a single, new

pre-processing macro to model a new pressure balance or to investigate a new pressure balance geometry. All other aspects of the modelling method remain the same. We have validated our methods by comparison with the predictions of other laboratories in the course of EUROMET projects and by comparisons with models of the same balance produced using alternative finite element methods. For simple piston-cylinder structures we also compare our distortion coefficient predictions with the results of analytical models.

3.4.1 The purpose of finite element simulations in pressure balance metrology

The main purpose of simulation is to investigate the sensitivity of the output of finite element models of pressure balances to changes in input parameters. Distortion coefficients are predicted for a reference model using finite element methods. Input parameters are then varied and new predictions of the distortion coefficient are obtained. Systematic investigation of the input parameter space leads to the evaluation of the sensitivity of the output to variations in specific input parameters. In this way an “uncertainty budget” can be developed for the model which identifies the range of predicted values which might be expected when input parameters cannot be determined exactly or are themselves expressed as a value with an associated uncertainty.

The inputs which are of most interest include the properties of the materials which form the piston and cylinder. The materials are typically tungsten carbide with added cobalt, or steel. In some cases a piston and cylinder may be manufactured from tungsten carbide, but the cylinder may be encased in a steel sleeve so that both materials may be present in the same structure. For the purposes of finite element modelling, the Young’s modulus and Poisson’s ratio of the material must be known. Other important inputs are geometrical (the radii of the piston and cylinder and the size of the gap between them), boundary conditions (where the structure is free to move and where it is constrained), and loading (where and with what accuracy are forces applied to the structure). For oil-operated balances (which is the case for all balances discussed in this report) the viscosity and density of the fluid as a function of pressure must also be known. Finally, the mathematical modelling techniques themselves contribute uncertainties to the output predictions. Examples of sources of “mathematical modelling uncertainty” include variations in finite element model mesh densities, the choice of elements to represent a particular structure, and the choice of equations to represent the pressure-dependent behaviour of the fluid. By combining the results for each input parameter which is varied, it is possible to evaluate both an estimate of the distortion coefficient and the uncertainty associated with the estimate for any particular pressure balance structure which is modelled.

3.5 Analytical methods for understanding pressure balances

Although it was published nearly 25 years ago, the most comprehensive account of the theory and practice of pressure balances is the text by Dadson, Lewis and Peggs [9]. In addition, an earlier paper by Dadson, Greig and Horner [8] provides a useful summary of the basic principles underlying pressure balance calibration. The principles and experimental techniques described are still relevant today.

Much of the early research on the theory of piston-cylinder pressure balances was concerned with the derivation of simple equations to allow the effective area of a balance to be calculated from knowledge of the materials from which the balance was manufactured and the radii of its components. In the case of what are known as balances of the “simple” type (that is, the applied pressure acts on the lower cross-section of the inner component or piston, which rotates relative to the outer component or cylinder), Dadson, Lewis and Peggs (p 100 [9]) conclude, after having made simplifying assumptions concerning axial symmetry and the pressure-dependent behaviour of the distorted piston and cylinder components, that the effective area, A_P , of a simple piston-cylinder balance in free-deformation mode (that is, there is no external constraint on the cylinder arising from pressures applied to it to limit its movement) is given by:

$$A_{\text{eff}} = A_0 \left\{ 1 + \frac{2}{2E}(3\sigma + 1) + \frac{P}{2E'} \left\{ \frac{(1 + \sigma')R'^2 + (1 - \sigma')R^2}{R'^2 R^2} \right\} \right\}, \quad (3)$$

where E , E' and σ , σ' are the respective values of Young’s modulus and Poisson’s ratio for the piston and cylinder, R is the internal radius of the cylinder and R' is the external radius of the cylinder. A_0 is the undistorted area.

In the case of controlled clearance balances (these have an additional external pressure applied to the outside surface of the cylinder to counteract the distortion caused by the pressure gradient along the engagement length between piston and cylinder), Dadson et al.(p 110 [9]) provide a simplified effective area equation:

$$A_{\text{eff}} = A_0 \left\{ 1 + \frac{P}{E}(3\sigma - 1) \right\}. \quad (4)$$

It is of interest to note that expression 4 predicts that the effective area is independent of material properties in the case where the Poisson’s ratio of the material is 0.33, and is then determined solely by the geometry of the assembly.

The analytical expressions for the effective area of a pressure balance that have been presented above apply only for pressure balances that do not have complicated geometries and in cases in which the distortion caused by the pressure gradient along the engagement length is small. Note that they rely only on knowledge of the properties of the materials from which the balance

has been manufactured and limited geometrical information about the undistorted balance. Nevertheless, they often provide good agreement with the predictions of finite element simulations, as will be shown later in this case study, and can act as a good first check on the likely effective area of a pressure balance under load.

3.6 Why numerical methods are preferred to analytical methods

To derive the effective area of a particular piston-cylinder combination in a form that can be exploited by numerical methods, one needs to be able to calculate the forces acting on the piston due to the fluid pressure. These are (p 108 [9]):

1. The force due to the applied line pressure acting on the cross-sectional area of the lower end of the piston, allowing for the change of cross-section at that end due to the combined effects of axial and radial compression;
2. The vertical component of the force due to the pressure (which normally decreases progressively from the applied pressure, P , to zero or to atmospheric pressure) acting on the sides of the piston;
3. The upward frictional forces due to the movement of the fluid between the piston and cylinder surfaces.

Assuming axial symmetry and laminar flow of a viscous pressure transmitting fluid, the above three forces lead to an expression for the effective area of a piston-cylinder pressure balance in the following form [9] (p 97):

$$A_{\text{eff}} = \pi r_0^2 \left\{ 1 + \frac{h_0}{r_0} + \frac{u_0 + U_0}{r_0} + \frac{1}{r_0 P} \int_0^l (p - p_2) \frac{d(u + U)}{dx} dx \right\}, \quad (5)$$

where r_0 is the radius of the undistorted piston, h_0 is the undistorted gap width, u and U are the differences between the distorted and undistorted radii of the radii piston and cylinder respectively along the engagement length, u_0 and U_0 are the values these differences take at the applied pressure end of the engagement length, $0 \leq x \leq l$ is distance along the engagement length, P is the line pressure, p is the pressure at each x value, and p_2 is the pressure at the top of the engagement length (atmosphere or vacuum, for example).

To evaluate expression 5 it is necessary to determine p , u and U along the engagement length of the piston and cylinder. However, the pressure distribution along this gap, and therefore also u and U , are dependent on the distortions of the piston and cylinder assembly and on the pressure dependence of fluid viscosity. A closed-form analytical solution of equation 5 is not possible unless one makes some simplifications, such as assuming that the x -dependent part of the radial distortions of the structure at each axial

position is proportional to the actual applied pressure at the same position. This approach can only be applied in limited circumstances [9] (pp 99-100). A general solution, which takes into account the actual structure of the piston and cylinder assembly and the manner in which it is loaded or restrained can only be achieved by numerical methods, and it is here that numerical simulations, in our case finite element methods, can make a major contribution. Note, however, that the finite element method is used simply to evaluate the integral term and the values of u_0 and U_0 in equation 5. All the modelling assumptions, including axial symmetry and laminar flow of a viscous fluid, which went into the derivation of this equation, are still present in the finite element modelling results. Once the effective area A_{eff} has been evaluated using expression 5, the distortion coefficient λ is evaluated using expression 2.

NPL uses the finite element method to predict the pressure distribution along the engagement length of the piston and cylinder, and consequently the local distortion of the piston and cylinder components, and then derives the distortion coefficient by using the outputs of the modelling as inputs to equation 5 and other related equations. Currently, we employ an uncoupled solution technique, so that the fluid-flow finite element problem is solved to derive a pressure distribution, which is then used as the stress input to the finite element calculation of the piston and cylinder strain. The distortion results are utilised in the next iteration of the fluid-flow calculation and this process is repeated until convergence is achieved.

It is worth pointing out that all the numerical and analytical approaches to pressure balance distortion simulation that have been reported in the literature to date have been based on the assumption of axial symmetry. No attempt has been made to develop a three-dimensional solution to the distortion problem. However, the limitations of the current methodology and the feasibility of moving to a full three-dimensional solution are reviewed later in this case study, especially in relation to incorporating measurements of piston-cylinder roundness and straightness into finite element modelling.

3.7 Key issues for pressure balance simulations

There are four important limitations to effective simulations of pressure balances. These are:

- Material properties;
- Geometry;
- Boundary conditions;
- Modelling simplifications.

Each of these is considered separately.

3.7.1 Material properties

As can be understood from expressions 3 and 4 above, even the simplest analytical equations for predicting distortion coefficients require knowledge of the Young's modulus and Poisson's ratio of the piston and cylinder materials. This information is naturally also necessary for finite element models. In addition, to model the fluid and derive the pressure distribution along the engagement length, the density and viscosity of the pressure transmitting fluid must be known as functions of pressure.

The materials employed for the piston-cylinder structures which are analysed in this report are tungsten carbide with added cobalt, and steel. In the case of controlled-clearance balances to be operated at relatively high pressures, a tungsten carbide cylinder is often encased in a steel sleeve. For operation at high pressures (100s of megapascals) tungsten carbide is preferred as its Young's modulus is three times that of steel, making it less susceptible to large distortions at high pressure. However, the addition of cobalt to the tungsten carbide contributes an extra source of uncertainty to the determination of the appropriate Young's modulus value to use for mathematical modelling purposes. Doi et al [11] have published data on the variation of Young's modulus of tungsten carbide cobalt with the volume fraction of cobalt. Figure 3 presents their results.

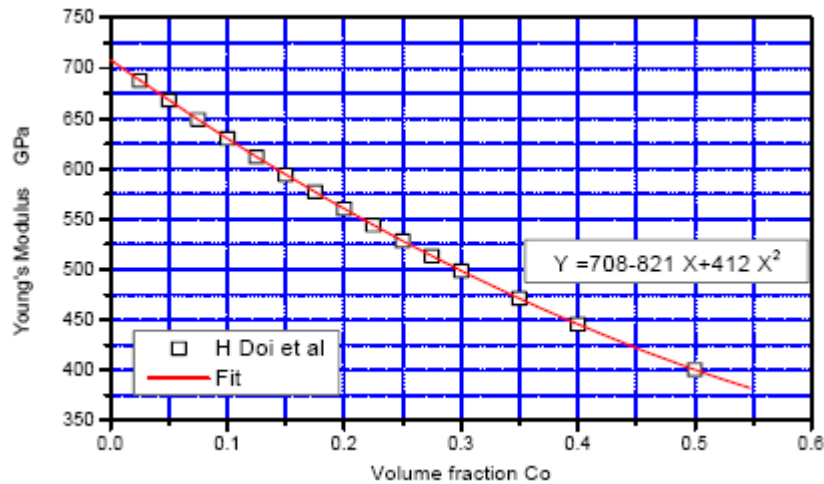


Figure 3: Young's modulus of tungsten carbide cobalt as a function of volume fraction of cobalt.

It is important to understand that data points in this figure are averaged results and that for a given volume fraction of cobalt, materials from different manufacturers may display variations in the value of Young's modulus. For the most reliable modelling results, and especially when comparing model predictions with experiment, it is essential that material properties are known as accurately as possible. It would be preferable if properties could be determined at the time the balance is manufactured. This requirement may, however, pose some difficulties as it is the practice of manufacturers to prepare

a cylinder and then to find the most suitable piston to fit the cylinder in question from a batch of pistons, which may not have identical properties. Measurement of all components' properties and accurate record keeping are required to ensure traceable data for each piston and cylinder pair.

Measurement of the properties of a completed assembly carries the risk that the components may be damaged during the measurement process. However, Sabuga [48] has shown that it is possible to make measurements of the Young's modulus and Poisson's ratio of pistons to an accuracy of 1% using strain gauges and without damage to the piston undergoing the test.

Data on the viscosity and density of pressure-transmitting fluids are often difficult to obtain over wide pressure ranges, as one is entirely dependent on what empirical data have been published in the literature. For all the analyses described in this case study, the pressure transmitting fluid has been assumed to be diethyl-hexyl-sebacate (DHS) where the dynamic viscosity, η , was taken to be:

$$\eta = 0.021554 \left\{ 1 + 1.90036 \times 10^{-3} p_1 \right\}^{8.81} \text{ Pa s}, \quad (6)$$

where p_1 is the pressure in MPa [40]. This formula appears to cover the pressure range up to approximately 600 MPa. Above that value, data are less readily available, although NPL has occasionally attempted to model balances up to 1 GPa with this equation.

Molinar et al [40] also quoted an equation for the pressure-dependent density of DHS. However, the data for this equation appear to have been obtained from compressibility measurements. It is not clear whether such data are appropriate for the case when the fluid is free to flow along the gap under the influence of the pressure gradient. In such a case it may be preferable to treat the fluid as incompressible. In the analyses described in this case study the density of DHS was held constant at 912.67 kg m^{-3} , although a brief investigation of the effect of Molinar et al's equation was carried out. The effect on distortion coefficient predictions of employing a pressure-dependent density term was shown to be negligible for the balances simulated.

3.7.2 Geometry

The simplified expressions 3 and 4 require one to know the radii of balance components. Equation 5 requires that the undistorted piston radius and gap width be known. In all cases these are single-valued terms taken to be characteristic of the component in question. Although NPL commonly has its dimensional metrologists determine in detail the roundness and straightness of pressure balance components, their results are reduced to a single value with an associated uncertainty for each of the piston and cylinder radii. In most finite element modelling of pressure balances it is the case that the undistorted piston and cylinder are taken to be ideally round and straight. It should be noted that equation 5 implicitly assumes an initially ideally round and straight undistorted piston and cylinder, or at least that single values can be stated for

the undistorted piston radius and gap width, r_0 and h_0 . However, the integral term in equation 5 does not itself rely directly on this assumption, but requires only that the changes from the undistorted values be a known function of position along the engagement length and that the structure be regarded as axisymmetric.

The shape of the piston and cylinder in the region of the engagement length - any steps, relieved sections, counterbores, tapers and so on - affect the manner in which the assembly distorts along its engagement length and also the values to be used in equation 5 for u_0 and U_0 , the distortions at the applied pressure end of the piston-cylinder gap. Only finite element modelling can take this structural detail into account.

3.7.3 Boundary conditions

In any real device it will not be possible to deliver the applied line pressure directly to the bottom of the piston-cylinder gap and to nowhere else in the pressure-balance structure. Inevitably there will be surfaces which experience loading from the applied pressure, including those portions of the piston and cylinder which extend below the engagement length. The effects of these loads on the manner in which the balance distorts must also be understood. In the case of controlled-clearance designs, the exact definition of the external region over which the jacket pressure is applied becomes important, as does the effect of oil seals in this part of the structure.

The second aspect of boundary condition determination concerns the definition of surfaces to which movement restraints are to be applied. In axisymmetric structures such restraints limit either axial or radial movement or both axial and radial movement at particular points of the pressure-balance structure or over particular surfaces within the structure. The identification of such surfaces often requires both the study of the balance's engineering drawings and also an understanding of the manner in which the balance is used in practice.

Once again, the finite element technique provides an effective way of including pressure loading and restraint boundary conditions in the mathematical model and of investigating the sensitivity of the predictions of the model to changes in these boundary conditions.

3.7.4 Modelling simplifications

Whether one uses the simplified equations 3 and 4 or employs numerical methods, including finite element analysis, to solve the integral term in equation 5, there are a number of features of real piston-cylinder balances which are ignored. It has already been pointed out that, in all the cases with which we are concerned, axial symmetry of the piston-cylinder structure is assumed. In addition, for oil-operated balances, all fluid flow is assumed to be the laminar flow of a viscous fluid in response to a one-dimensional pressure

gradient.

A further simplification relates to the movement of the balance. The models assume that the piston and cylinder remain stationary, both in terms of their positions or heights relative to each other (that is, the effect of the piston falling or the cylinder rising is ignored), and that the cylinder and piston do not rotate. In reality, either the piston or cylinder is rotated with respect to the other component. The motivation for this form of operation is to overcome friction arising from direct contact between the floating and fixed elements of a pressure balance structure. Typically, the floating element, including the attached load, is rotated steadily by some form of mechanical drive. However, for the highest precision work free rotation is preferred to avoid the introduction of extraneous forces into the system, especially at the lower end of the pressure range. A full discussion of rotation effects can be found in Dadson et al [9] (pp 34-45).

Modelling the effects of rotation is possible, and is a task commonly undertaken by manufacturers and designers of journal bearings, which also consist of a piston with a lubricated cylinder. However, it has to the authors' knowledge never been applied to the modelling of pressure balances. In any case, the motivation for journal bearing designers is different, in that they are concerned chiefly with lubrication questions, whereas the users of piston-cylinder pressure balances are interested in the details of structural distortion.

Ignoring rotation may also have consequences for a full understanding of the behaviour of the pressure-transmitting fluid. Diethyl-hexyl-sebacate is a long-chain molecule and the flow of such molecules will depend on their orientation. It has been suggested that for a large molecule to flow under a pressure gradient, it must first rearrange its segments into an easy flow configuration and then move in segments of a fixed size [30]. Note also that the direction of easy flow for a long-chain molecule must be different for the rotating and non-rotating case, as the preferred easy-flow direction must be in the direction of the governing flow regime. It may be that different equations for pressure-dependent viscosity are required for the rotating and non-rotating cases.

3.8 Simulating pressure balances using ANSYS finite element software

The pressure balance simulations reported in this case study were carried out using ANSYS finite element software. In this section of the report we summarise how NPL approaches this task.

3.8.1 The ANSYS Parametric Design Language

Relatively straightforward finite element analyses, such as the solution of a linear stress problem, can often be carried out using the graphical user

interface routines of finite element software. These routines typically allow the user to “draw” the various engineering components on the VDU screen in a manner analogous to computer-aided design software, to assign loads and restraints graphically, to select the solution methods required from drop-down menus, and to display results graphically. For standard types of analysis, or when many iterations of the model are not required, this is often the most efficient means of carrying out a finite element analysis.

In the case of pressure-balance modelling, the finite element analysis simply provides the raw data needed for the solution of equation 5, and for the work reported here, the solution method is an iterative one, as the fluid flow problem and the structural strain problem are solved in an uncoupled manner. The method requires one to make an initial guess at the pressure distribution along the gap (the simplest approach being to assume an initial linear pressure distribution) and to use this pressure distribution as input to a solution of the structural deformation problem. The structural deformation solution is then used to make a new prediction of the pressure distribution in the fluid, which is again used as input to the structural problem, which is then re-solved. This process continues until convergence is reached. It is also the case that knowledge of the structural deformation is simply a means towards the end of calculating the distortion coefficient and the piston fall-rate (the piston moves downward as fluid leaks out at the top of the engagement length), and such calculations are naturally not part of standard finite element software routines.

Furthermore, we need not only to solve an uncoupled problem iteratively but also need to be able to make changes to model input parameters and geometries without having to re-work the whole model from the beginning. One of the disadvantages of finite element methods is that they only solve a particular problem and do not provide solutions to a set or class of problems. To make small structural, materials property, or boundary condition changes to a finite element model, such as are required when performing sensitivity analyses, can be time-consuming if it is always necessary to re-work a model completely using the graphical user interface. The re-working of a model can be avoided if non-graphical methods of data generation and analysis are employed. ANSYS provides the facility to define and carry out analyses using the ANSYS Parametric Design Language (APDL). This high-level language has a FORTRAN-like structure and allows the user to define and re-define parameters, to generate macros, initiate solution routines, and write macros to analyse results, in the same manner as one writes a typical piece of line-by-line computer code. In our case we employ APDL to control all three aspects of the finite element analysis: pre-processing of the initial definition of the problem, the solution of both the structural and fluid flow problem, and the post-processing stages of evaluating convergence criteria, calculating the predicted distortion coefficient and displaying pressure profiles and gap profiles graphically.

The use of APDL allows us to prepare a series of macros to carry out each stage of the analysis. The solution and post-processing stages are common to all the finite element analyses we carry out. To prepare a new model it is only necessary to generate new pre-processing macros. In practice, the

pre-processing macros are all very similar and the differences between balances relate only to geometry, materials properties and boundary conditions. The structure of the macros themselves remains the same from one analysis to another.

The main macros needed for each stages of the analysis process are described below.

3.8.2 Pre-processing

The pre-processing stages are typically separated into two macros. The first sets up the global variables for the analysis in question - materials properties, co-ordinates of key points, and line pressure value, for example. The second macro turns the above data into a meshed model with all the elements and boundary conditions defined.

Once these two macros have been run then one may begin the solution of the finite element problem. At this stage ANSYS itself will have carried out a number of important checks to ensure that key features of the modelling are valid. It performs geometry validation, including checking that areas and volumes are closed, that the element types are consistent with the required analysis, and warnings are produced if aspect ratios of elements are unreasonable.

For any new model these two macros are the most crucial, as they are the ones which need the most work and which define the key parameters of the model. They will be unique to each model, whereas many of the macros which are used during the solution and post-processing stages may be common to all or to most finite element analyses. In some cases, for example, when analysing a piston-cylinder whose geometry bears no relation to devices which have been modelled previously, it may be preferable to combine these two macros together into one file, rather than follow the two macro system.

3.8.3 The solution and post-processing

In the case of the ANSYS models of piston-cylinder pressure balances, it is difficult to separate the solution and post-processing stages as the finite element problem is solved many times iteratively. At each cycle of the iteration some mathematical analysis is carried out to define the input for the next cycle of the iteration. This process continues until the model being analysed converges to a solution. The mathematical analysis routines, which are written by NPL and not part of the ANSYS software package, are technically post-processing routines in finite element terms.

Separate macros carry out the following tasks:

1. Apply the pressure loads to the relevant boundary;
2. Start the ANSYS solver and test results for convergence after each loop;

3. Set up the necessary post-processing mathematics, including calculation of distortion coefficient and pressure loads for the next iteration;
4. Write the pressures and displacements at each node along the engagement length to a text file after each iteration;
5. Write the results of the calculations of effective area, distortion coefficient, and fall rate to a file after each iteration;
6. Graphical display of results, including charts of the profile of the distorted gap, the pressure along the engagement length, the distorted piston profile, the distorted cylinder profile, viscosity of fluid along the gap, effective area as a function of iteration number, distortion coefficient as a function of iteration number, the fall rate as a function of iteration number, and the value of the convergence test parameter. In addition it is possible produce two-dimensional colour plots of the distorted structure, x -direction displacements of all components in the distorted structure, and stress and strain distributions in the distorted structure.

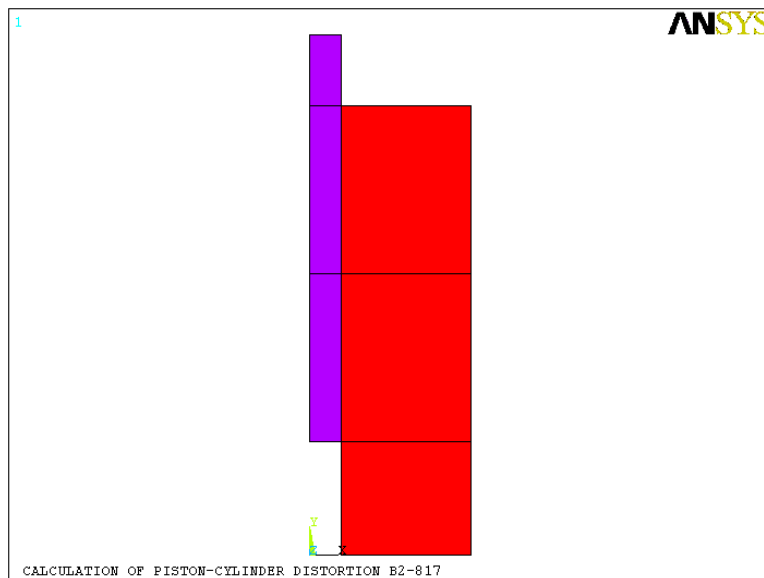


Figure 4: Main blocks of axisymmetric finite element representation of piston and cylinder of simple balance (NPL balance ref. no B817).

Figure 4 shows how the structure of a balance can be reduced to the main features of the piston and cylinder when balances are of simple construction, and where gaps are of the order of 1 micrometre and applied pressures are of the order of 200 MPa or less. On the scale of this image it is not possible to see the gap between the piston and the cylinder.

3.8.4 Analysing the results

Although the purpose of the modelling is to produce a prediction of the distortion coefficient for the pressure balance being modelled, it is essential to be able to follow the history of a particular solution and also to examine detailed results of predicted pressure distributions and structural distortions along the engagement length. The purpose of writing the pressure and displacement data to text files at stages four and five is to allow such examinations to be performed at any time after an analysis has been completed, without the need to repeat the analysis itself. As it may take 100 iterations or more and several hours for an analysis to converge to a solution, it is important to retain as many results of the analysis as possible to avoid the need to repeat calculations.

The use of text files also allows data to be imported into other software packages such as Excel or Matlab, thus facilitating alternative methods of data analysis, including the production of graphical output. Another benefit is that it renders the user independent of ANSYS's own graphics capabilities. All the figures in this case study that show pressure distributions or gap profiles have been produced using Matlab or Excel.

3.8.5 Modelling the fluid

To model the fluid flow and therefore the pressure distribution along the gap between the piston and cylinder we employ an analogy between fluid flow and thermal flow, which allows us to use thermal finite elements within our ANSYS model. The model of fluid flow we employ in the gap is steady-state viscous laminar flow of a fluid moving laterally between two plates whose length and width are much greater than the distance separating them. This can also be applied to flow in the gap between a piston and cylinder provided the gap is small. As Drysdale et al [12] (p 116) point out, in such a case the total fluid flow across a unit width perpendicular to the flow direction is equal to the rate of decrease of pressure in that direction multiplied by the constant $h^3/12$, where h is the gap width and μ is the coefficient of viscosity. In the case of one-dimensional heat flow, the heat flux is equal to the rate of decrease of temperature in the direction of interest multiplied by KA , where K is the thermal conductivity and A is the cross-sectional area. Thus, if $h^3/12$ can be set equal to KA in value, the expressions that govern fluid flow and heat flow are identical in form.

The LINK32 ANSYS finite element is a uniaxial conducting bar element with a single degree of freedom, temperature, at each node point and can be used for plane or axisymmetric analyses. The element is defined by two nodes, a cross-sectional area and the material properties. For steady-state solutions, density and specific heat of the bar material are ignored, so that only the thermal conductivity and the cross-sectional area of the element are required. Thus this element can be used to calculate pressure distributions in the case under consideration provided the thermal conductivity value is replaced by the appropriate value for the fluid flow problem. The "thermal" finite element

problem is solved independently of the structural finite element problem and its results are used as the input to the next iteration of the structural problem. The solution of the structural finite element analysis then provides the information on the gap width, and therefore area, needed for the next iteration of the “thermal” analysis.

The solution and post-processing macros also calculate the rate at which the piston falls as a result of fluid flow. For this purpose it is simply necessary to determine the volume flow rate of the fluid through the annular gap using equation 7 given by Molinar et al [40], in which the piston is considered to fall at the rate at which the volume of fluid it displaces equals the volume of fluid flowing through the gap:

$$v_F = \frac{Q_m}{\pi r^2 \rho(p_1)}, \quad (7)$$

where v_F is the piston fall rate and Q_m is the mass flow in the gap, given by equation 8:

$$Q_m = \frac{\pi r h^3(x) \rho(p)}{6\eta(p)} \frac{dp}{dx}. \quad (8)$$

Note that although the piston fall rate is calculated, the position of the piston relative to the cylinder is not adjusted, but remains the same from one iteration to the next. This is another example of the simplifications that are employed in finite element models of pressure balances.

From what has been described above, it is clear that our current modelling approach can only be regarded as strictly valid in the case where the fluid flow regime and the pressure distribution in the fluid can be described using the simplifications built into the model. These are of course those assumptions which are also adopted by Dadson et al [9]. In cases where gaps become very small indeed so that molecular effects become important, or it is necessary to model rotation or three-dimensional flow, alternative methods of fluid modelling would have to be adopted for the most accurate results. There may be limited advantage for computational purposes only in describing fluid properties by an “effective viscosity” coefficient in some cases, but such an approach would necessarily obscure the underlying physics of the problem. In fact, it will be seen that the current finite element models are relatively insensitive to changes in fluid viscosity and this seems to be a natural consequence of the assumptions built into the modelling method itself, which is also true of the simplified equations 3 and 4 quoted earlier in this case study.

3.9 Typical outputs from the finite element simulations

Some typical outputs from the modelling process indicate both the advantages and limitation of finite element modelling of pressure balances. We present below results of modelling two balances, identified by NPL’s metrologists by the reference numbers B816 and B817, which have very similar geometry but are made of different materials. Figure 5 shows a finite element prediction of the pressure profile along the engagement length of a simple piston and

cylinder made of tungsten carbide and cobalt (B816). This material is preferred for pressure balances as it is much stiffer than steel, which is an alternative pressure balance material. Ten profiles are shown for applied line pressures (that is, the pressures that are applied at the bottom of the piston) over the range from 14 MPa to 140 MPa. The engagement length (the distance over which the piston engages the cylinder) is 25 mm. For this pressure balance, the cylinder extends approximately 8.5 mm below the end of the piston and the base of the cylinder is used as the reference point for measurements of distance along the engagement length. The gap between the piston and cylinder when the balance is unpressurised is 1 μm . Note the approximately linear pressure profiles. Figure 6 shows the manner in which the gap between the piston and the cylinder varies as a function of applied line pressure over the 14 MPa to 140 MPa range. The figure shows that at the lowest line pressure the gap hardly changes, remaining close to 1 μm throughout the engagement length. Note, too, that the gap profiles are approximately linear, as were the pressure profiles.

NPL possesses a pressure balance that is identical in basic geometry to the balance that has just been described but is made of steel rather than tungsten carbide and has a smaller gap (B817). Young's modulus for steel can be one third of the Young's modulus for the tungsten carbide and cobalt combination. Figure 7 and 8 show the pressure and gap profiles respectively over the same pressure range as figures 5 and 6. In the case of the steel balance there are clear departures from linearity in both the pressure and gap profiles at all applied pressures, including the lowest pressures.

3.9.1 Comparison between finite element results and simple analytical models

The results of the finite element analysis are used to calculate the distortion coefficient for the balance using expressions 2 and 5. These can be compared with the predictions of the analytical equation 3. This was done for the tungsten carbide-cobalt (B816) and steel (B817) models presented above for a range of Young's modulus values for the materials in question. The comparisons between the finite element and analytical results for the two balances (and therefore the two materials) are given in tables 1 and 2, together with the ratios of the finite element results to the results of the analytical model calculations.

From table 1 we see that for all values of Young's modulus that were studied for the tungsten carbide case the difference between the finite element predictions of the distortion coefficient and the analytical model predictions are 0.3%, with the finite element results consistently 0.3% higher than the results from the analytical model. Note that as was observed in figures 5 and 6, finite element modelling of this balance shows an almost linear relationship between pressure and position along the engagement length and between gap width and position.

For balance B817, which is of identical geometry to balance B816 but is made

of steel, a less stiff material than tungsten carbide, the results in table 2 show that the difference between the finite element results and the analytical results are close to 6% for the lowest values of Young's modulus and are greater than 4% for the highest values, with the finite element results being consistently higher than the analytical modelling results. In this case figures 7 and 8 showed substantial departures from linear behaviour both for the pressure and gap width as a function of position along the engagement length. Similar observations apply to a comparison of the two modelling approaches when Poisson's ratio is varied, and interested readers are referred to [13] for further information.

It appears from the results presented here that the results obtainable from the analytical model are adequate for balances of simple geometry where the material is relatively stiff and the pressure and gap width profiles are linear. The further the balance departs from these conditions, the greater the discrepancy between the finite element results and the analytical results.

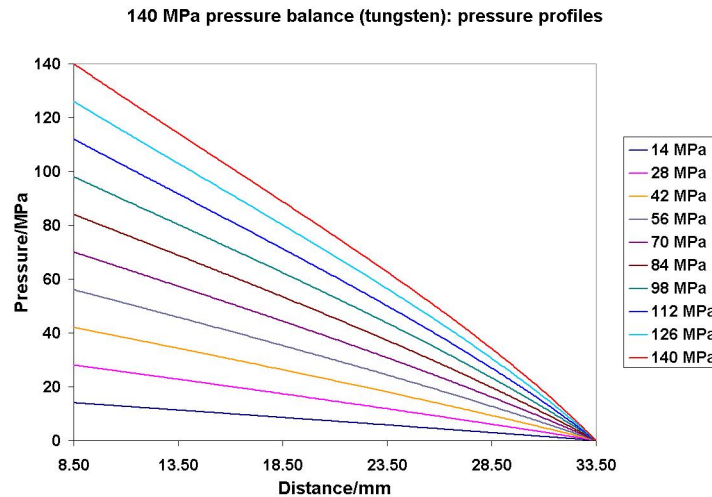


Figure 5: Pressure as a function of position along engagement length for ten input line pressures: simple tungsten carbide piston-cylinder design.

3.10 Controlled clearance balances

One way in which the designers of pressure balances aim to control the manner in which a balance distorts when it is pressurised is to apply an external pressure (commonly referred to as the jacket pressure) to the outside of the cylinder, which acts to compress the balance and to limit the amount of distortion. The aim is to control the maximum size of the gap between the piston and the cylinder and for this reason such balances as referred to as “controlled clearance” balances. Typically an external pressure of between 10 % and 25 % of the line pressure applied to the balance is used as pressure applied to the cylinder's outer face. The consequences of this can be seen in

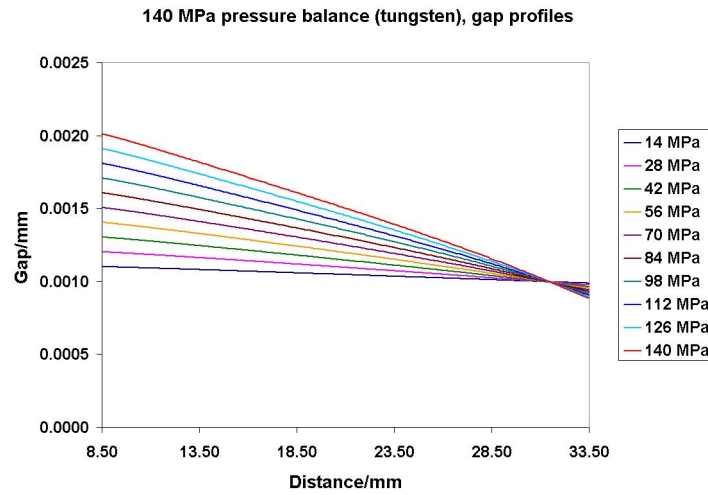


Figure 6: Gap width as a function of position along engagement length for ten input line pressures: simple tungsten carbide piston-cylinder design.

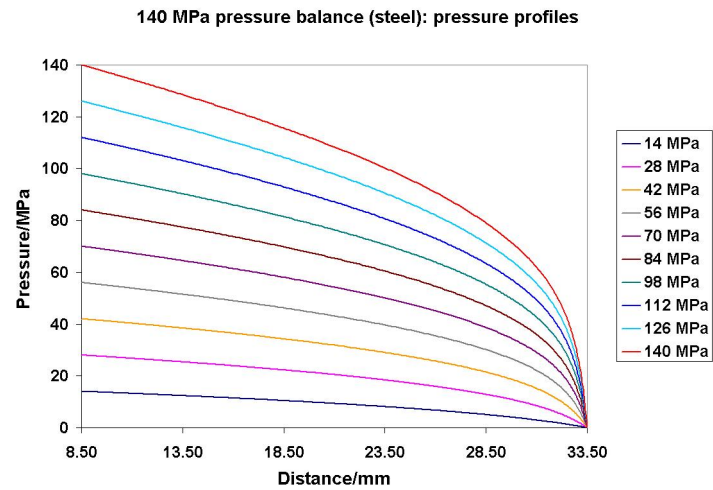


Figure 7: Pressure as a function of position along engagement length for ten input line pressures: simple steel piston-cylinder design.

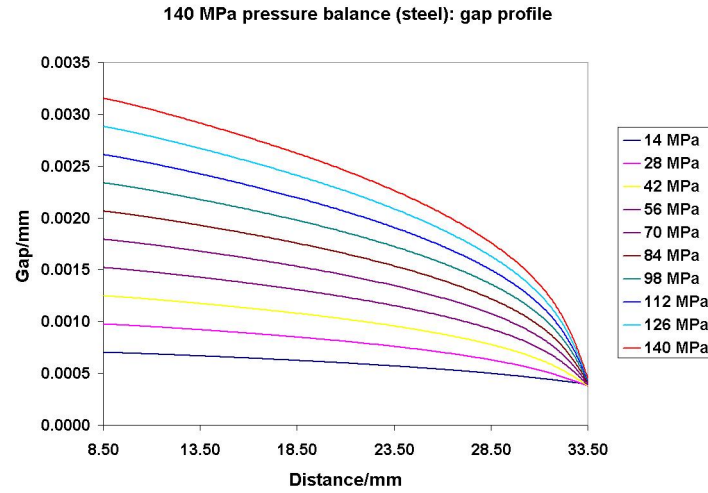


Figure 8: Gap width as a function of position along engagement length for ten input line pressures: simple steel piston-cylinder design.

<i>Young's modulus/ GPa</i>	<i>Distortion coefficient /ppm/MPa</i>		<i>Ratio FEA/Simple</i>
	<i>FEA</i>	<i>Simple</i>	
493	1.0144	1.0115	1.003
508	0.9844	0.9816	1.003
523	0.9561	0.9534	1.003
538	0.9293	0.9269	1.003
553	0.9041	0.9017	1.003
568	0.8801	0.8779	1.003
583	0.8575	0.8553	1.003
598	0.8359	0.8339	1.002
613	0.8154	0.8135	1.002
628	0.7959	0.7940	1.002
643	0.7773	0.7755	1.002
658	0.7596	0.7578	1.002
673	0.7427	0.7409	1.002

Table 1: B816 tungsten carbide balance: compare finite element and simple analytical model as Young's modulus is varied at 140 GPa line pressure.

<i>Young's modulus/ GPa</i>	<i>Distortion coefficient /ppm/MPa</i>		<i>Ratio FEA/Simple</i>
	<i>FEA</i>	<i>Simple</i>	
170	3.9728	3.7568	1.057
180	3.7453	3.5481	1.056
190	3.5417	3.3614	1.054
200	3.3587	3.1933	1.052
210	3.1932	3.0412	1.050
220	3.0430	2.9030	1.048
230	2.9059	2.7768	1.046
240	2.7804	2.6611	1.045
250	2.6648	2.5546	1.043

Table 2: B817 steel balance: compare finite element and simple analytical model as Young's modulus is varied at 140 GPa line pressure.

figures 9 and 10. These show results for a balance that was modelled as part of a EUROMET intercomparison of finite element modelling methods. More information about this project can be found in [13, 49, 50]. Controlled clearance balances can also be operated without an externally-applied pressure. This is known as free deformation mode. The figures show a comparison between the controlled clearance mode and the free deformation mode at a line pressure of 400 MPa for a balance that was modelled as part of the EUROMET 463 project. Note how the controlled clearance mode of operation reduces the gap width, as shown in figure 10. In section 3.10.1 below, we discuss the evaluation of the uncertainty associated with an estimate of the distortion coefficient of this balance.

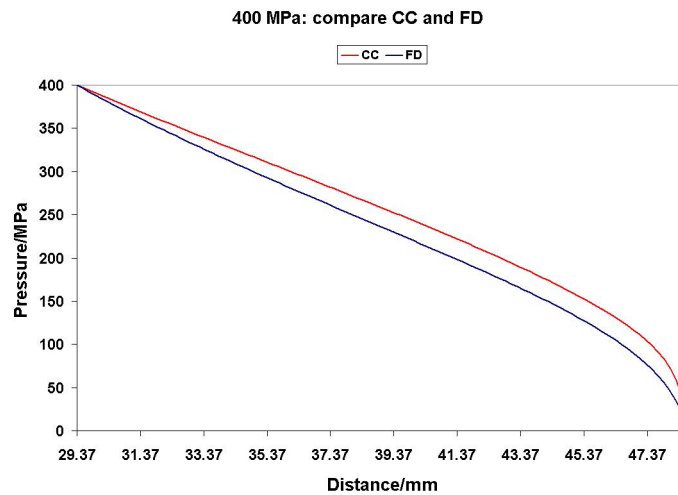


Figure 9: EUROMET 463 project: pressure as a function of position along engagement length: free deformation (FD) and controlled clearance (CCC) operation compared.

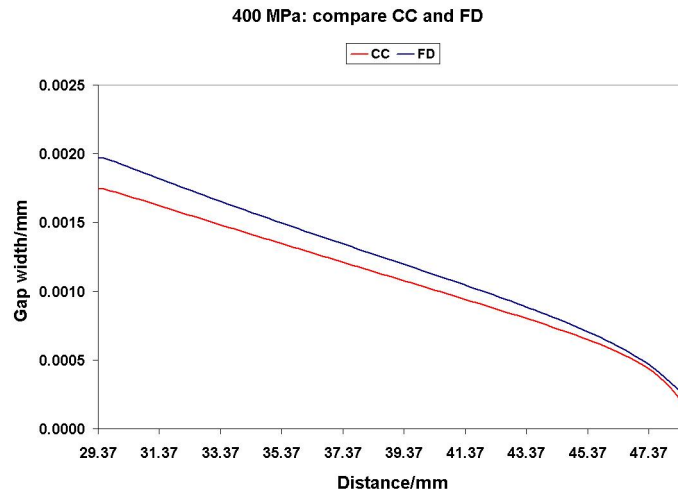


Figure 10: EUROMET 463 project: gap width as a function of position along engagement length: free deformation (FD) and controlled clearance (CC) operation compared.

3.10.1 Evaluating the uncertainty associated with an estimate of the model output

In the course of the work for the EUROMET project no 463 on comparisons of finite element models of pressure balances, as reported in [13, 49, 50], project participants were required to evaluate uncertainties in their model outputs, although the project protocol did not set out any guidance as to how this might be done. The main result from the models that were being compared was an estimate of the piston-cylinder distortion coefficient and its associated uncertainty for a defined input line pressure.

Figure 11 shows a schematic diagram of the structure of the cylinder for the EUROMET 463 modelling intercomparison and figure 12 shows how the main blocks of the model appear in an axisymmetric ANSYS representation of the piston and cylinder. The mesh that was used to solve the model is shown in figure 13. For this balance, which is operated by the Physikalisch-Technische Bundesanstalt (PTB) and used in controlled clearance mode with pressures up to 1 GPa, and whose undistorted gap width is only 0.32 micrometres, it was not possible to obtain reliable results with a simplified model of the type shown in figure 4, and it was necessary to work with a more detailed representation of the balance's geometry.

As a result of some preliminary comparisons of finite element modelling results with experimental determinations of the pressure balance's distortion coefficient and fall rate, it appeared that the piston and cylinder of the balance had been manufactured from tungsten carbide with different properties. The PTB were able to measure the Young's modulus and Poisson's ratio of the balance piston and had determined these to be 543 ± 7 GPa and $0.238 \pm$

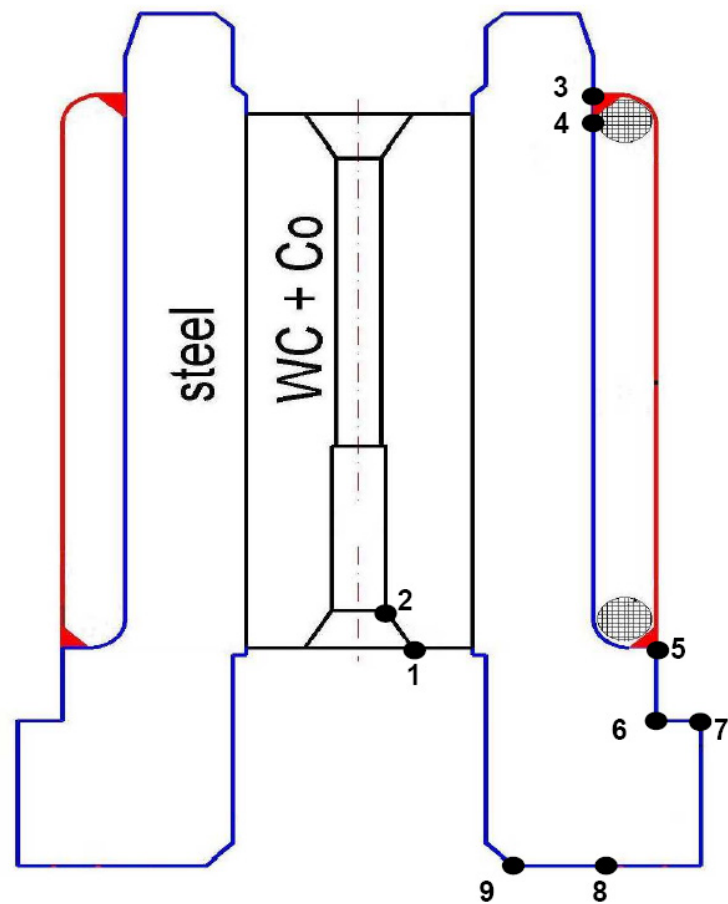


Figure 11: Schematic of structure of cylinder for EUROMET 463 modelling intercomparison. Numbered positions indicate surfaces to which specific boundary conditions are to be applied - see text for details.

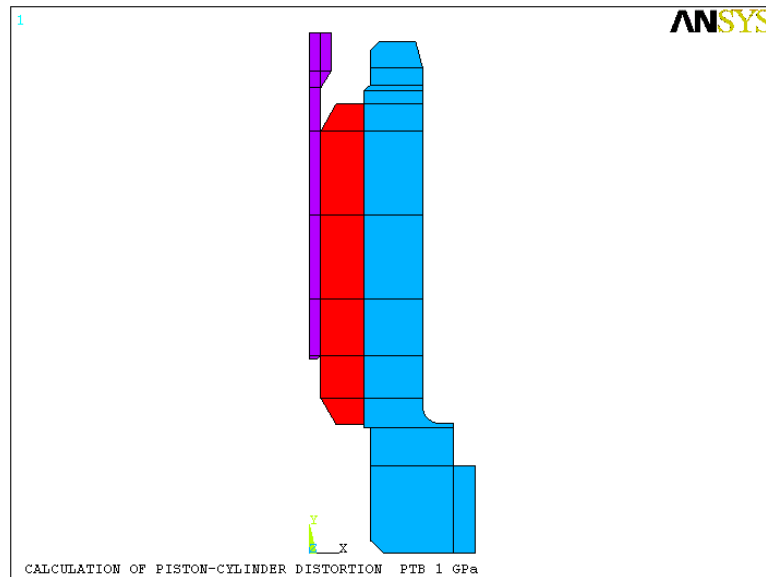


Figure 12: Main blocks of axisymmetric finite element representation of piston and cylinder for EUROMET 463 modelling intercomparison.

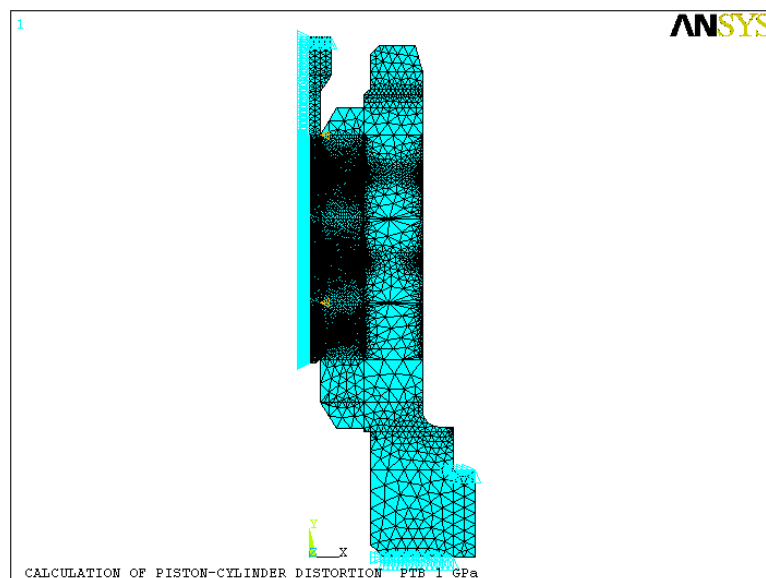


Figure 13: Mesh used for the EUROMET 463 balance.

0.002 respectively [48] (where the quoted uncertainties are stated at the 95 % confidence level). The properties of the cylinder were taken to be those which had originally been provided by the manufacturer, Désgranges and Huot, namely a Young's modulus of 630 GPa and a Poisson's ratio of 0.22. The uncertainties associated with the cylinder properties were taken to be ± 10 GPa for Young's modulus and ± 0.002 for Poisson's ratio. The cylinder sleeve is made from steel and its properties were considered to be 200 ± 7 GPa for Young's modulus and 0.29 ± 0.02 for Poisson's ratio. The gap width in the ideally round and straight case was taken to be $0.32 \pm 0.03 \mu\text{m}$. In controlled clearance mode, calculations of distortion coefficient and piston fall rate were to be carried out at applied pressures of 100, 250, 400, 600, 800 and 1000 MPa, with the jacket pressures exactly one tenth of these values. In free deformation mode the applied pressures to be investigated were 100, 250, 500 and 600 MPa. In addition to the calculations required by the PTB, the effect of varying boundary conditions, of using a pressure-dependent density (equation 2 from [41]) and of coarsening the finite element mesh were also investigated.

Figure 11 identifies the location of lines upon which specific boundary conditions were to be applied. Lines 6 to 7 and 8 to 9 were to be constrained from movement in both the axial and radial directions. The applied pressure should be applied on line 1 to 2, as well as to the piston and to higher points on the cylinder in the gap region. In the controlled clearance case, the jacket pressure was to be applied along the line from point 4 to 5. In addition, the effect of extending the region of jacket pressure application to point 3 was also investigated, as was moving point 4 lower to coincide with the location of the top of the piston-cylinder engagement length.

Following the analysis of the case in which the piston and cylinder are taken as having different properties, the calculations were repeated on the assumption that the cylinder had the same properties as the piston, that is, a Young's modulus of 543 ± 7 GPa and a Poisson's ratio of 0.238 ± 0.002 . The inputs to the base line models were then varied in accordance with the requirements of the PTB project co-ordinator. In the case of the materials properties for the piston, cylinder and cylinder sleeve, additional analyses were carried out for two values of each property, representing the extremes of the uncertainty range which had been quoted. Thus, in the case of the piston, the base line Young's modulus value was 543 GPa, and the two additional cases which were calculated were 536 GPa and 550 GPa. A similar approach was adopted for the other materials properties and components. In the case of the gap width, which was taken to be $0.32 \mu\text{m}$ for the base line model, additional gap widths of $0.29 \mu\text{m}$ and $0.35 \mu\text{m}$ were also analysed. In each case, two approaches to defining the revised gap were investigated. The first was to apply the $0.03 \mu\text{m}$ adjustment needed wholly to the piston radius and the second was to apply it wholly to the cylinder internal radius.

Other variations to the 400 MPa base line model results were to vary the applied pressure from 100 MPa to 1000 MPa in the controlled clearance mode, and from 100 MPa to 600 MPa in the free deformation mode, to investigate changes in boundary conditions, to test the effect of employing a pressure-dependent density term and to vary the finite element mesh density.

The boundary condition variations which were implemented were to omit the application of the line pressure on the line between points 1 and 2 in both the controlled clearance and free deformation cases, and to vary the range over which the jacket pressure was applied in the controlled clearance case by applying it additionally between lines 3 and 4, by removing it below point 4 to a location level with the top of the engagement length, and by removing it in the region of the lower oil seal close to point 5 in figure 11. To test pressure-dependent density effects, the density equation from [40] was implemented. Finally, the effect of coarser and denser finite element meshes was investigated. Although one would ideally like to use as fine a mesh as possible, such an approach is not feasible when run times are long and many repetitions of the analysis are required. Thus the mesh chosen for the base line model is dense enough to require run times of several hours, but not dense enough to produce intractable run times for repeat analyses.

The standard uncertainties from each source of modelling uncertainty, determined from repeat runs of the model with the inputs varied over the ranges set out above, were summed in quadrature and the expanded uncertainty obtained by taking a k value of 2 (corresponding to a 95 % coverage probability under a Gaussian assumption). Separate uncertainty calculations were carried out for the controlled clearance and free deformation modes and the results are set out in tables 3 and 4. It is important to note that the standard uncertainties predicted for the controlled clearance cases are substantially larger than those for the free deformation case. This is because variations in the boundary conditions related to the application of the jacket pressure in the controlled clearance mode are the major contributor to the uncertainty budget. Further analysis of these results can be found in [13].

Apart from the uncertainty contribution from the mesh density term, every other contribution to the uncertainty budget arises from a limitation on the information which is known about the pressure balance. If material properties or boundary conditions could be defined more precisely, then the evaluated uncertainties would be reduced. Conversely, the less that is known about the true properties of a balance, the larger the uncertainties associated with the predicted distortion coefficient and fall rate become. To gain the most benefit from finite element modelling, especially when its results are to be used in calibration certificates and in calculations of calibration uncertainties, substantial experimental work is needed to obtain the most precise quantification possible of materials properties and boundary conditions. Note, also, that even if inputs to the model can be precisely defined, the evaluated uncertainties for controlled clearance mode will always exceed those of the free deformation mode of operation.

3.11 Reducing uncertainties with a new controlled-clearance design

Simulations allow one to explore possibilities that it would be difficult to achieve experimentally. NPL has used finite element methods to study the design of controlled clearance pressure balances and as a result of a systematic

<i>EUROMET 463 project D&H 7594 balance in CC mode</i> <i>Differing cylinder and piston properties</i>		
<i>Source of uncertainty</i>	<i>Range of uncertainty of quantity</i>	<i>Std uncertainty $\pm$ppm/MPa</i>
Piston: E	536 GPa to 550 GPa	0.0021
Piston: σ	0.236 to 0.240	0.0031
Cylinder: E	620 GPa to 640 GPa	0.0062
Cylinder: σ	0.218 to 0.222	0.0008
Sleeve: E	193 GPa to 207 GPa	0.0007
Sleeve: σ	0.27 to 0.31	0.0027
Gap width	0.29 μ m to 0.35 μ m	0.0007
Pressure variation	100 MPa to 1000 MPa	0.0044
Boundary conditions	-	0.0242
Oil: density as $f(p)$	-	0.0001
Oil: viscosity	expt varies from 8.80 to 8.82	0.0000
Mesh density	61 to 141 nodes along gap	0.0019
<i>Combined standard uncertainty</i>		0.0259
<i>Expanded uncertainty</i>		0.0517
<i>Distortion coefficient = 0.3666 ± 0.0517 ppm/MPa</i>		

Table 3: EUROMET 463 balance, controlled clearance mode, differing cylinder and piston properties, distortion coefficient uncertainty, showing uncertainties from individual components and the combined uncertainty.

<i>EUROMET 463 project D&H 7594 balance in FD mode</i> <i>Differing cylinder and piston properties</i>		
<i>Source of uncertainty</i>	<i>Range of uncertainty of quantity</i>	<i>Std uncertainty $\pm$ppm/MPa</i>
Piston: E	536 GPa to 550 GPa	0.0020
Piston: σ	0.236 to 0.240	0.0032
Cylinder: E	620 GPa to 640 GPa	0.0094
Cylinder: σ	0.218 to 0.222	0.0010
Sleeve: E	193 GPa to 207 GPa	0.0004
Sleeve: σ	0.27 to 0.31	0.0001
Gap width	0.29 μ m to 0.35 μ m	0.0001
Pressure variation	100 MPa to 1000 MPa	0.0005
Boundary conditions	-	0.0044
Oil: density as $f(p)$	-	0.0001
Oil: viscosity	expt varies from 8.80 to 8.82	0.00003
Mesh density	61 to 141 nodes along gap	0.0010
<i>Combined standard uncertainty</i>		0.0111
<i>Expanded uncertainty</i>		0.0223
<i>Distortion coefficient = 0.7651 ± 0.0223 ppm/MPa</i>		

Table 4: EUROMET 463 balance, free deformation mode, differing cylinder and piston properties, distortion coefficient uncertainty, showing uncertainties from individual components and the combined uncertainty.

investigation of the design space (variation in materials properties, piston-cylinder geometry, applied external pressure and boundary conditions) was able to design a pressure balance of a particular geometry that could potentially achieve a negligibly small distortion coefficient in certain operating conditions. This work was reported in Measurement Science and Technology [20] and readers are referred to that paper for a full description of the balance. We have since been co-operating with a manufacturer of balances with the aim of producing a working version of the proposed balance.

When pressure balances are operated at high pressures, the elastic distortion of the piston and cylinder produces effective areas for pressure balances that differ substantially from the area of the undistorted balance. As pressure balance metrology relies on the accurate determination of the effective area of balances, it would be of benefit if piston-cylinder combinations could meet the following three requirements:

1. The effective area should be predictable;
2. The design should be insensitive to variations in the geometrical and material properties of the balance;
3. The effective area should be as close as possible in value to the area of the undistorted balance, that is, the pressure distortion coefficient should be negligible.

In general the results of the simulations carried out for this design process showed the benefits of employing the controlled-clearance mode of operation for piston-cylinder balances in the pressure balance geometry which was modelled. Piston fall rates for the controlled-clearance case were half that obtained in free deformation mode. A linear relationship between the distortion and the ratio of jacket pressure to applied pressure was demonstrated and once this relationship is known it is straightforward to determine by linear interpolation the ratio of jacket pressure to applied pressure which produces a distortion coefficient of zero. Once this optimal operating point has been established, the controlled-clearance distortion coefficient was shown to be ten times less sensitive to variations in Young's modulus than the free deformation distortion coefficient. Small variations in the geometry of the cylinder in the region of the engagement length were also shown to have a minimal effect on distortion coefficients and piston fall rates.

There was a much greater sensitivity to variations in Poisson's ratio but the authors concluded that this was likely not to be problematic, simply because it is a ratio. Once the optimum jacket pressure for a particular balance material and geometry had been established, the benefits of operation at this point would be retained. Thus, the most important question for the manufacturer and user of such a balance would be the realization of a ratio of jacket pressure to applied pressure which would give a zero distortion coefficient. Material properties considerations and small geometrical variations then become secondary.

3.12 Some limitations of finite element methods

Although finite element simulations have been shown to be useful for understanding the manner in which pressure balances distort, especially at high pressures, they have a number of limitations:

1. To date they have only been used to model balances axisymmetrically. No researchers have reported attempts to model balances in three dimensions without the assumption of axial symmetry. Piston-cylinder pressure balances are very similar to journal bearings in their construction and fully three dimensional simulation has been applied successfully to journal bearings - see, for example, <http://www.vibroacoustics.co.uk/>.² However, the pressure balance community has not adopted such methods for pressure balance simulation.
2. Pressure balance modelling has traditionally begun by considering that the undistorted piston and cylinder are perfectly round and perfectly straight. Although attempts have been made to take into account the real shapes of pistons and cylinders these have always been based on deriving some averaged profile for the piston and cylinder that is then used as the basis for an axisymmetric model. Accounts of such approaches can be found in [49, 50].
3. Current simulation techniques also ignore the relative movement of the cylinder and piston, as they assume that both components remain stationary

NPL's experience of trying to base modelling of pressure balances on measured piston and cylinder profiles is described in [13]. The key conclusions from this work were:

Current approaches to the finite element modelling of pressure balances all rely on the assumption of axial symmetry and all ignore the relative movement of the piston and cylinder, and therefore, of the pressure-transmitting fluid. In addition, the equations which are used to derive the distortion coefficient also assume axial symmetry and indeed ideal roundness and straightness, as single values are used for the piston radius and the gap width. Measured three-dimensional profile data have to be forced to fit into these prior assumptions. The most straightforward way of doing this is to reduce measured profile data obtained from dimensional measurements of the undistorted piston and cylinder to a series of axisymmetric slices, usually by averaging measured profiles to produce an averaged shape for the piston and the cylinder. The slices are stacked along the same perpendicular axis to form the piston and cylinder shapes. However, as has been shown in the case of the EUROMET 463 project, the result of this

²Accessed on-line 4th July 2006.

process is a piston and cylinder configuration which is not representative of the real pressure balance.

The process of building the axisymmetric slices into a finite element modelling is time consuming, as one cannot take full advantage of automatic mesh generation facilities to mesh large areas of the structure rapidly, and it is necessary to derive equations, such as splines or polynomials, from curve fitting to the measured profiles so that radius values can be derived for the piston and cylinder at the nodal positions required by the finite element model.

It also appears that distortion coefficient predictions obtained from measured profiles may not differ substantially from those obtained for the ideally round and straight case. In cases in which the measured profiles do not depart greatly from the ideal shape, this is to be expected. The methods adopted by Dadson, Lewis and Peggs [9], in which ideal geometry is assumed, have been applied to a wide range of real devices over several decades, and for practical purposes have been shown to be adequate for almost all types of oil-based balance. If small departures from roundness and straightness led to large changes in distortion coefficient, this would have tended to invalidate their approach, which has gained general acceptance in the metrology community. In addition, if the use of measured piston and cylinder profiles as inputs to FEA models is to be adopted by NMIs, then it is necessary to reach agreement on how this is to be done in the context of axisymmetric models and their limitations. . . .

3.12.1 Finite element simulations for pressure balance metrology: the pros and cons summarised

The NPL report cited earlier [13] concludes by summarising the NPL's experience of simulating pressure balances by finite element methods. It is worth quoting these in full as they are relevant to a wide range of simulations of engineering components and equipment.

- For oil-operated tungsten carbide balances of simple structure in free deformation mode, there is little advantage in using finite element modelling over more simple analytical equations for predicting distortion coefficients.
- For balances made from less stiff materials, such as steel, finite element methods are to be preferred, as the simple analytical equations based on elastic theory are less reliable in this case.
- For balances with complex geometries, and which are operated at high pressures and with small gaps in controlled clearance mode, finite element analysis can provide good predictions of distortion coefficients and fall rates.

- For tungsten carbide balances in particular, variations in gap width and in line pressure have little effect on the predicted distortion coefficients. This is not so for the stainless steel balance which was modelled, which showed a dependence of the distortion coefficient on both these parameters.
- If the outputs of finite element models are to be relied upon for deriving distortion coefficient uncertainties for calibration certificates, it is essential that accurate and precise measurement data exist for Young's modulus and Poisson's ratio of all component materials. Uncertainties in knowledge of material properties are the dominant factor in the distortion coefficient uncertainties derived from mathematical modelling.
- To model controlled clearance balances accurately it is necessary to be able to define the boundary conditions in a manner which represents how the jacket pressure is actually applied. Exact representation of the pressure loading in the region of oil-seals and at the top of the engagement length is essential for reliable results.
- Attempts to represent the real three-dimensional shapes of pistons and cylinders in axisymmetric models are problematic. They cannot reflect the true shape of the piston and cylinder and ignore the effect of relative motion of the piston and cylinder, and therefore of the fluid. For cases in which the measured profiles do not show large departures from the ideally round and straight case there may be little benefit to be gained from persisting with axisymmetric models, as in any case, distortion coefficients are calculated using theory which assumes ideal roundness and straightness of the initial undistorted piston and cylinder.
- Variations in fluid properties, i.e. viscosity as a function of pressure, and density, have very little effect on distortion coefficient predictions. Fall rate predictions are more sensitive to these parameters.

3.13 Molecular modelling and pressure balances

The simulation work presented so far in this case study has concentrated on the modelling of piston-cylinders operated at relatively high pressures (at least tens of megapascals) and with the use of oils as the pressure balance fluid. Recently, NPL has pioneered the use of molecular modelling methods for gas-operated pressure balances. The flow of gases along narrow channels cannot easily be represented by the idealised flow models that have been presented earlier in this case study. As a result, distortion coefficient predictions for gas-operated balances at pressures up to 20 MPa have not been reliable. In collaboration with Imperial College London NPL embarked on a detailed study of flows in gas-operated balances using molecular dynamics methods. This work has been described at great length in another report produced as part of the current Software Support for Metrology programme and will not be reproduced here [21]. Readers who wish to know more about the topic are referred to that report.

4 Monte Carlo method of investigating experimental uncertainties in a parallel computing environment

4.1 Introduction

This chapter of the report summarises a case study on the use of Monte Carlo methods to understand the effects of key parameters on the design of an experiment in thermal metrology. NPL experimentalists are developing an alternative to traditional methodologies for measuring the temperature of high-temperature blackbodies and wish to concentrate their design work on those aspects of the apparatus and experimental technique that will lead to clear improvements in the achievable measurement uncertainties. The complexity of the governing equations and the time that it takes to solve a single instance of the simulation in question were such that the only way to carry out sufficient Monte Carlo trials was to recast the problem in parallel rather than serial form and to use the NPL distributed computing system. A description of the distributed computing system and its use at NPL can be found in [19].

4.2 The physical problem

The problem is one that has arisen in NPL's optical detectors group which is part of the optical radiation team and is described briefly below.

4.2.1 Determining thermodynamic temperature

Using the technique of filter radiometry the thermodynamic temperature of a blackbody source is determined routinely with high precision. The technique is based on the measurement of the radiance of a blackbody source within a narrow wavelength range. From this measurement, and Planck's law, the temperature of the blackbody source is determined directly. Different National Metrology Institutes (NMIs) have different designs of primary filter radiometer and characterise these instruments in different ways. This is highly valuable in improving the understanding of systematic effects. The double wavelength technique is an alternative approach that provides the absolute thermodynamic temperatures of two blackbodies at different temperatures by measuring the ratio of their spectral radiances using a measurement system comprising optics, a monochromator and a photodiode detector. If the technique can be realised experimentally with uncertainties comparable to those associated with filter radiometry, it would provide an independent validation of filter radiometry measurements. NPL is currently undertaking work to build and test this new method [58].

The spectral radiance at wavelength λ of a blackbody and temperature T is

described by Planck's function,

$$L(\lambda, T) = \frac{c_{1L}/n^2}{\lambda^5 (\exp(c_2/(n\lambda T)) - 1)},$$

where c_{1L} and c_2 are known physical constants, and n is the refractive index of air. The measurement of radiance using a monochromator does not provide an estimate of spectral radiance corresponding to a single wavelength λ , but instead has the affect of integrating spectral radiance over a (usually, small) interval of wavelengths containing λ . Thus, the measured value provided by a monochromator system is an estimate of the quantity

$$L_m(\lambda, T) = R(\lambda) \int \frac{c_{1L}/n^2}{\ell^5 (\exp(c_2/(n\ell T)) - 1)} s(\lambda, \ell) d\ell,$$

where $s(\lambda, \ell)$ describes the bandpass function of the monochromator, and $R(\lambda)$ is a calibration factor that is approximately constant with temperature. The bandpass function is regarded as zero outside an interval $[\lambda - \delta\lambda, \lambda + \delta\lambda]$ centred on λ with half-width $\delta\lambda$, in which case the integration is taken over this interval.

Consider two blackbodies at temperatures T_1 and T_2 respectively. In the double wavelength technique, the ratio of the radiances of the two blackbodies is measured at two wavelengths λ_1 and λ_2 . Thus, measured values x_i , $i = 1, 2$, are obtained as estimates of the quantities

$$X_i = \frac{L_m(\lambda_i, T_1)}{L_m(\lambda_i, T_2)} = \frac{\int \frac{c_{1L}/n^2}{\ell^5 (\exp(c_2/(n\ell T_1)) - 1)} s_i(\lambda_i, \ell) d\ell}{\int \frac{c_{1L}/n^2}{\ell^5 (\exp(c_2/(n\ell T_2)) - 1)} s_i(\lambda_i, \ell) d\ell}, \quad i = 1, 2. \quad (9)$$

In terms of the measured values x_i , $i = 1, 2$, the equations (9) define two equations that may be solved to obtain estimates $\hat{T}_i$, $i = 1, 2$, of the temperatures of the blackbodies. The equations are non-linear and consequently appropriate numerical techniques are used to solve them.

4.3 Experimental design

To design experimental apparatus that can realise the proposed new method, it is necessary at the design stage to be able to predict the likely effect on measurement uncertainty of key aspects of the experimental apparatus. If the sensitivity of the measurement to variations in the properties of the apparatus and experimental set-up can be determined, one can concentrate during the design and manufacturing process on those components that are likely to have most effect in reducing the combined measurement uncertainty.

To allow a simulation of the effects of uncertainties equation 9 was re-cast into the following form:

$$X_1 = e_1 \left(\frac{\int_{\lambda_1 - \delta\lambda}^{\lambda_1 + \delta\lambda} (L(\lambda_1 + e_5, T_1 + e_{7,x_1}) + (s_1(\lambda_1 + e_5, \ell + e_{11}(\ell)) \times e_9(\ell)) e_{3,top} d\ell)}{\int_{\lambda_1 - \delta\lambda}^{\lambda_1 + \delta\lambda} (L(\lambda_1 + e_5, T_2 + e_{8,x_1}) + (s_1(\lambda_1 + e_5, \ell + e_{11}(\ell)) \times e_9(\ell)) e_{3,bottom} d\ell)} \right)$$

$$X_2 = e_2 \left(\frac{\int_{\lambda_1 - \delta\lambda}^{\lambda_1 + \delta\lambda} (L(\lambda_2 + e_6, T_1 + e_{7,x_2}) + (s_1(\lambda_1 + e_6, \ell + e_{12}(\ell)) \times e_{10}(\ell)) e_{4,top} dl}{\int_{\lambda_1 - \delta\lambda}^{\lambda_1 + \delta\lambda} (L(\lambda_2 + e_6, T_2 + e_{8,x_2}) + (s_1(\lambda_1 + e_6, \ell + e_{12}(\ell)) \times e_{10}(\ell)) e_{4,bottom} dl} \right)$$

where the e_i values correspond to the different sources of uncertainty given in table 5.

One sees from table 5 that the experimentalist has identified a dozen sources of uncertainty in the measurement process and has been able to make estimates of their approximate size. However, given the complicated nature of the equations that describe the measurement process it is not immediately obvious which uncertainty components are likely to have the largest effect on the measurement results and thus it is not possible to identify easily those aspects of the experimental design on which to concentrate so as to minimise uncertainties associated with the results. Note also that the values in table 5 are likely to be revised as NPL undertakes further work and gains a better understanding of the various effects.

4.4 Simulation using Monte Carlo methods

Because of the complexity of the mathematical model it was decided to use Monte Carlo methods to evaluate the uncertainties associated with the estimated temperatures. This approach accounts fully for the non-linearity of the model and nature of the probability distributions associated with each source of uncertainty (in the first instance, these were assumed to be Gaussian).

The model was implemented using the Matlab programming language. However, a single evaluation of the model requires the solution of two non-linear equations with each equation being defined by two definite integrations. On typical personal computers it takes between 30 and 90 seconds to carry out a single evaluation of the model. Even these relatively short run times for a single trial render intractable serial methods of applying Monte Carlo methods to uncertainty evaluation.

4.4.1 Monte Carlo methods in uncertainty evaluation

The Joint Committee for Guides in Metrology has a working group (WG1) “Expression of uncertainty in measurement” with the task of promoting the use of the Guide to the Expression of Uncertainty in Measurement (commonly known amongst metrologists as the GUM) and preparing supporting documents for its broad application [4]. The first document to be produced by the working group, GUM Supplement 1 [5, 6], is concerned with the propagation of probability distributions as the basis of the evaluation of measurement uncertainty. It describes a Monte Carlo method as a general numerical implementation of the propagation of probability distributions that

Index	Source of uncertainty	Approximate size
1	Multiplicative effect in the X_1 -ratio, from e.g. non-linearity	0.01 %
2	Multiplicative effect in the X_2 -ratio	0.005 %
3	Additive effect in the X_1 -ratio (top and bottom, uncorrelated), from e.g. measurement noise	1 pA cf signal of 85 mA for Re-C and 0.4 nA for Cu
4	Additive effect in the X_2 -ratio (top and bottom, uncorrelated)	1 pA cf signal of 1.2 mA for Re-C and 0.3 mA for Cu
5	Systematic wavelength effect in the X_1 -ratio (wavelength offset common to both measurements)	0.01 nm
6	Systematic wavelength effect in the X_2 -ratio	0.05 nm
7	Re-C fixed point temperature variation from the X_1 -ratio measurement to the X_2 -ratio	20 mK
8	Cu fixed point temperature variation from the X_1 -ratio measurement to the X_2 -ratio measurement	5 mK
9	Noise in the determination of the band-pass function for the X_1 -ratio	0.5 %
10	Noise in the determination of the band-pass function for the X_2 -ratio	1 %
11	Uncorrelated wavelength offsets in individual measurements for the bandpass function determination, X_1 -ratio	0.01 nm
12	Uncorrelated wavelength offsets in individual measurements for the bandpass function determination, X_2 -ratio	0.02 nm

Table 5: Sources of uncertainty for double wavelength technique and estimates of their approximate size, where the values stated are standard uncertainties.

can be expected to provide results that are free of the approximations involved in applying the GUM uncertainty framework and, therefore, an uncertainty evaluation that is reliable for a wide range of measurement problems. However, unlike the GUM uncertainty framework, the method is computationally intensive and requires implementation on a computer.

There are two operations that are central to an implementation of a Monte Carlo method. The first is to make random draws from the probability distributions assigned to the input quantities in the measurement model to obtain values for these quantities. The second is to evaluate the measurement model for these values of the input quantities to obtain a value of the output quantity. These operations, that together constitute a single Monte Carlo trial, are repeated a large number of times. GUM Supplement 1 describes an ‘adaptive’ Monte Carlo procedure that undertakes a number of Monte Carlo trials to ensure that the results obtained from the procedure have stabilised in a statistical sense. Often, of the order of 1,000 trials will be adequate to obtain a reliable estimate of the output quantity and the associated standard uncertainty as the expectation and standard deviation of the set of values of the output quantity. However, depending on the nature of the probability distribution for the output quantity, of the order of 10^6 trials may be required to obtain a coverage interval for the output quantity corresponding to a coverage probability of 95 %. For higher coverage probabilities, many more trials may be needed.

For many measurement problems, undertaking 10^5 or 10^6 trials will be feasible, with results provided within a few seconds or perhaps minutes. However, for the measurement problem described, a single model evaluation may take 90 s. It follows that 1,000 Monte Carlo trials would take 90,000 s (or 25 hours) and, consequently, even for this small number of trials the calculation cannot be regarded as feasible.

4.5 Distributed computing and Monte Carlo methods

A Monte Carlo calculation is an example of a data parallel application in which the same computational task (that of model evaluation) is to be applied to different data (values of the input quantities). In consequence, it is a computational task that is suitable for a distributed computing system.

Distributed computing in its most general form is the co-ordination of a (usually large) number of computers that are remote from each other to access and process information [29]. An example is where financial or accounting information contained in databases is stored and processed at one central location, but the information is accessed and modified by users working remotely. A recent application of distributed computing is in collaborative initiatives in which the spare processing capacity of a large number of desktop computers is used to tackle computationally intensive problems, with the collaboration often being coordinated across the Internet or an Intranet. A well-known example of such an initiative is *SETI@home* in which desktop computers are used to analyse radio telescope data in the SETI project

(Search for Extraterrestrial Intelligence).

Distributed computing is an example of parallel processing [39] but for which, and unlike general parallel processing architectures, there is no communication between, or sharing of memory by the individual processors. In a distributed computing system, the computational task is divided into a number of ‘work units’ that are allocated to processors (computers) in the system by a central system server. There is no guarantee that a given computer will be available to complete a work unit by a particular time and, consequently, the completion of the overall computational task must not depend on communication between the work units or sharing of information. Indeed, the process by which work units are allocated to the different computers can be quite sophisticated, with sometimes the same work unit allocated to different computers and the result for that work unit taken as that provided by the computer that completes the work unit first. For this reason, distributed computing systems are most effective when applied to data parallel applications, in which the computations undertaken by the different work units are the same but are applied to different data. The *SETI@home* initiative is an example of a data parallel application.

Within metrology, the use of comprehensive and realistic models of measurement that require extensive computing resources for their solution is increasing. In several areas of metrology the computational requirements of such models are so demanding that there is a strong requirement for distributed computing across a network of desktop computers. For this reason, the UK’s National Physical Laboratory (NPL) has invested in a distributed computing system [19] comprising several hundred of the desktop computers within the organisation used by scientists and administrators or employed to control experimental apparatus. The system has been used to solve problems involving electron transport in quantum dots, adiabatic energy transfer in caesium atoms, Monte Carlo calculations for ionising radiation applications, and free-field calibrations of underwater electroacoustic transducers [60].

The Monte Carlo calculation described here for the evaluation of measurement uncertainty is an example of a data parallel application that is suitable for implementation on a distributed computing system. Let n work units be defined, each providing M/n values of the output quantity Y of the measurement model corresponding to M/n Monte Carlo trials. The results from the individual work units are aggregated to give M values of Y from which is obtained an approximation to the probability distribution for Y as if the M values had been provided by a single Monte Carlo calculation comprising M trials. In an ideal case, n would be chosen to be the number of computers available in the distributed computing network. Then, if the work units were run simultaneously on identical machines, the calculation of the M values of Y would be faster by a factor of (about) n compared to the calculation on a single machine. If the principal cost of the calculation is in obtaining the M values, rather than the post-processing of these values, the savings can be significant.

4.5.1 Distributed computing and random number generation

A number of algorithms for generating pseudo-random numbers are available that provide sequences of numbers that pass comprehensive tests of ‘randomness’. However, in the context of a distributed computing system, there is the further requirement that the sequences generated on different processors (computers) in the system do not overlap. With an existing long period generator, even with randomly chosen seeds, there is a small risk that two sequences will overlap, and the results provided by the Monte Carlo calculation will be compromised.

The ability to generate pseudo-random numbers from a rectangular distribution is fundamental in its own right, and also as the basis for generating numbers from any probability distribution. In the latter regard, the quality of the numbers generated from a non-rectangular distribution depends on that of the numbers generated from the rectangular distribution and on the properties of the algorithm employed. Test suites for random number generators are publicly available, including the DIEHARD [37] and TestU01 [35] test suites. The TestU01 package is comprehensive with many individual tests but also three batteries of tests, of increasing ‘difficulty’, referred to as ‘Small Crush’, ‘Crush’, and ‘Big Crush’. Big Crush uses 2^{38} random numbers and can take over 24 hours to execute, depending upon the speed of the generator being tested. Several random number generators that pass the Big Crush battery of statistical tests are available [57]. These include Mersenne twister [38], MRG32k3a [33], TAOCP [31], CLCG4 [34], Wichmann–Hill [57], and MRG63k3a [33].

The Wichmann–Hill generator used in this work is constructed as a combination of four multiplicative congruent generators. The four constituent generators are defined by primes p_i and multipliers m_i , $i = 1, \dots, 4$, and take the form

$$J_{k+1}^i = m_i \times J_k^i \bmod p_i, \quad i = 1, \dots, 4, \quad (10)$$

where the integer J_{k+1}^i is the state of the i th generator after k applications of the generator starting from J_1^i . The four constituent generators are combined by taking each state J_k^i as the fraction of its prime p_i , summing them and taking the fractional part of the result. The use of 64-bit integer arithmetic to undertake the calculations can be avoided, and a variant of the generator suitable for 32-bit integer arithmetic is available [57]. The generator has a period of (about) 2.65×10^{36} , which is adequate for the majority of applications, and passes both the DIEHARD and TESTU01 tests. It is particularly simple to code (compared with the other generators listed above) with a state that is easy to handle (comprising only the four integers J_k^i , $i = 1, \dots, 4$). A disadvantage of the generator is that it is not as fast to execute as some of the others listed.

The Wichmann–Hill generator is one of the few generators for which consideration has been given to its use within a (highly parallel) distributed computing system [57].

4.6 Implementing the parallel version of the model on the NPL Grid

The detailed implementation of the distributed version of the model is described in [10, 14]. However, some key aspects of the process are highlighted below.

4.6.1 Implementation in Matlab: the issues

A serial version of the code to solve the model using Monte Carlo methods had already been developed to allow some initial investigations of the experimental design problem. This had been written in Matlab and used Matlab's own built-in random number generation routines. It is possible to generate executable files from Matlab that are suitable for use in a distributed computing system, and which does not require a Matlab licence to be installed on individual machines in the grid. This is achieved by the use of the Matlab compiler to produce an executable file from a Matlab m-file and the installation of the Matlab Component Runtime (MCR) on each machine on the grid. The MCR is a stand-alone set of shared libraries that enables the execution of encrypted m-files created using the Matlab compiler. Thanks to MCR, a computer on the grid is able to run a Matlab executable without opening a Matlab window or calling Matlab itself. Readers who require more detailed information about the process are referred to [19].

The ability to use Matlab as the basis of the parallel version of the software meant that the main software engineering task necessary to complete the production of the parallel code was the replacement of Matlab's random number generator with the parallel version of the Wichmann-Hill procedure. Validated versions of the Wichmann-Hill algorithm already existed in C and FORTRAN and these could be used to help validate the Matlab version. A key issue here was the declaration of the variables to be used in the random number generation process. There is no type specification in Matlab; however it was necessary to ensure that long integers (32 bits) were used. Hence, several versions of the code were made using:

- no type declaration; the numbers were doubles by default;
- *int32* initialisation for the variables;
- *int32* initialisation for all numbers, variables and parameters.

According to [57], obtaining the two first random numbers correctly with the initial seed [1111] up to the 15th digit is sufficient to conclude that the constants have been correctly set and that the random number generator can be used. The test methodology to compare the three Matlab versions with the C version was as follows:

- Compare the two first random numbers obtained with an initial seed [1111].

- Identify the number of elements which were less than 0.5 over the 50 following random numbers.
- Record the CPU time required to compute 1,000 and 10,000 random numbers.

The tests revealed that the three Matlab versions give the same results as the C version to a high level of accuracy (15 or 16 decimal places). In addition, the Matlab version using no type declaration was slightly faster than the other two versions and was used in the parallel Monte Carlo analysis.

4.6.2 Investigation of an alternative seeding strategy

Another issue of interest is the question of the extent to which it is possible to identify the performance benefits that arise from the use of a random number generator that has been specially developed for parallel implementations, as opposed to the use of generators that do not take into account the fact that it is necessary to be confident that the separate sequences of numbers generated on each machine on the grid can be regarded as independent. Therefore we performed some analyses using the the clock of each computer to generate the seed for the random number generator, rather than obeying the Wichmann-Hill rules. The initial results of this study are described briefly in section 4.7.2 of this report.

4.6.3 Determining the number of work units and the work unit size

To ensure confidence in the quality of the results generated by the Monte Carlo analysis it was decided that the minimum number of iterations required should be 10,000. This number of trials would require approximately 900,000 seconds (250 hours) on a typical machine on the NPL grid. Tests of run times for individual work units, combined with limitations arising from the fact that not all machines that were available for grid calculations had the MCR installed and that machines varied in their speed, led to the conclusion that 150 work units, each performing 100 Monte Carlo trials, would allow the demonstration of significant time savings from the use of distributed computing. This is confirmed by the timing data set out in table 14, in which the actual run time and the equivalent CPU time are given for two jobs with 150 work units of 100 trials. Both jobs were completed by the grid in less than four hours but comprised a total equivalent CPU time of approximately 7.5 days, a performance gain of a factor of 45 from the use of the NPL grid. As more machines with the MCR installed become available this performance will improve.

4.6.4 Grid statistics

It is possible to learn more about the manner in which the grid carried out the calculations for the two jobs analysed in table 14 by studying the time lines for

Job Number	Workunits	Grid time required	Equivalent CPU time
1540	150	$\approx$ 3h 54m	7 days 12h 14m 50s
1541	150	$\approx$ 3h 49m	7 days 11h 9m 18s

Figure 14: Statistics for job numbers 1540 and 1541. Each job had 150 work units undertaking 100 trials each.

each work unit. These are shown in figure 15.

This figure shows the duration of each work unit for the jobs numbered 1540 and 1541, with job 1540 presented in the left hand column of the figure. The results have been sorted by time of completion with the work units that were completed earliest shown at the bottom of each column of results. The x -axis of the graphs show the elapsed time in seconds since the job was created on the distributed computing systems. The results lead to a number of interesting observations about the performance of the system.

- For job 1540 the fastest work unit (372499) was computed in approximately 2,500 seconds (41 minutes) and the longest (372504) in approximately 7,700 seconds (2hours 8 minutes). In theory, each work unit requires the same CPU time. The difference comes from the varying speeds of the PCs on the grid and also from the amount of CPU time made available by the agent PC to the computation. This depends on how much the agent PC is being used for computation while the background distributed task is running. Local requirements take precedence over calculations for distributed work units. Figure 16, which is an enlargement of part of figure 15 shows only a small section of the work units for job 1540, and makes clear the substantial variation in work unit processing times that occurs in the grid system.
- Several stages of the computation exist. For job 1540, 76 PCs on the grid accept a work unit soon after the job is made available to the grid. Once the fastest machines complete their first work unit, they accept a second unit. 15 machines are fast enough to be able to complete three work units.
- It is likely that at the time job 1540 was initiated only 76 of the available machines on the grid had the Matlab Component Runtime software installed.
- Qualitatively, job 1541 is similar to job 1540. However, by the time the second job began, another 20 machines had the Matlab Component Runtime software installed. Thus a larger proportion of the 150 available work units were allocated soon after the job was created on the grid system.
- Given the larger take-up of work units at the start, fewer of the fast machines got the chance to accept a third job. Nevertheless, the time to completion for both jobs was very similar. Job 1541 took five minutes

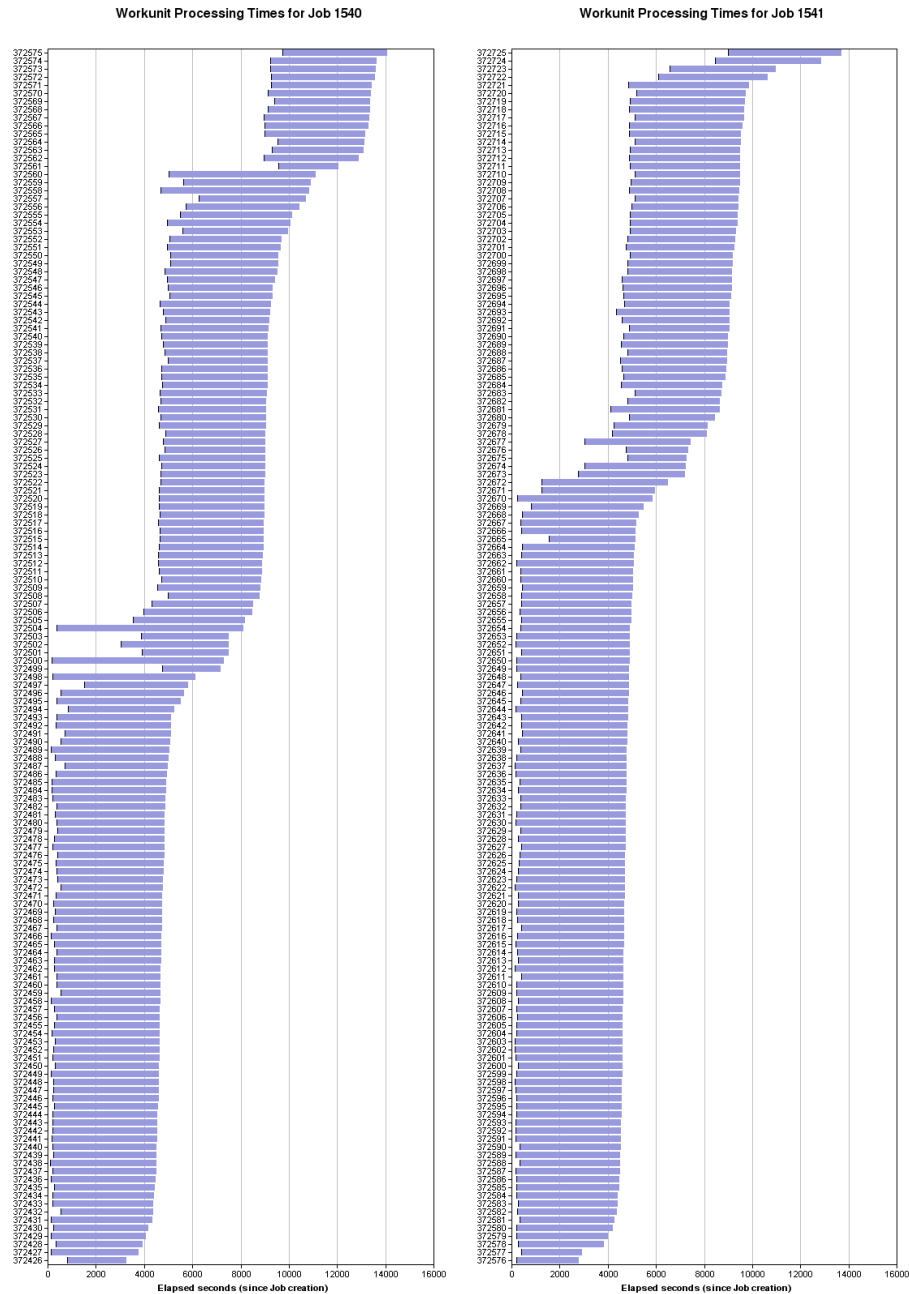


Figure 15: Workunit timeline graph of job 1540 and 1541, provided by the grid software. This figure is best viewed in enlarged format in the electronic version if the reader wishes to see specific work unit numbers.

less than job 1540 to complete, which represents 2% of the run time needed for job 1540.

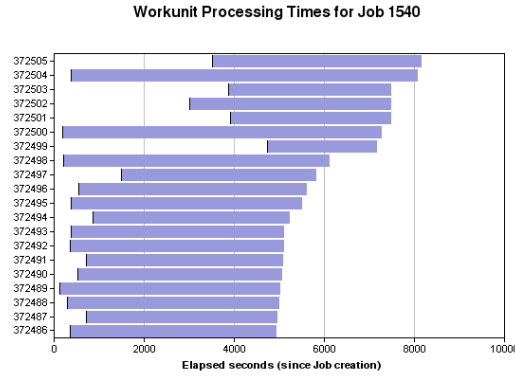


Figure 16: Enlarged section of workunit time line graph of job 1540 showing the differences of time of computation for various workunits. Unit 372499 is seen to be completed much more quickly than unit 372504.

The experience outlined above underlines the importance of good experimental design for parallel Monte Carlo calculations. If the aim is to maximise the gain in run time over the serial process then careful attention must be paid to the balance of work units and number of iterations per work unit. Suppose that only 75 work units had been generated, each with 200 iterations, then the time to complete each unit would have been doubled, but even the fastest machines would have only been able to complete one work unit as all the available work units would already have been allocated to other machines, some of which would have been much slower. The time to complete all the work units would then be governed by the slowest machines on the grid.

4.7 Initial results

4.7.1 Investigating the effect of the input uncertainties

Initial tests of the simulation outputs were made to help understand how best to present the results to the metrologists. The estimated temperatures of the two blackbodies were calculated and T_2 was plotted against T_1 . Figure 17 shows what was observed when all effects in table 5 except one were held constant, that is, one term was varied in the Monte Carlo analysis and the other 11 effects were kept at fixed values. This analysis was performed first for item 1 in the table (the multiplicative effect in the X_1 ratio) and then for item 8 (the copper fixed point temperature variation). Note, first, that in both cases and indeed in all the cases we have examined so far, the temperatures T_1 and T_2 are highly correlated. Secondly, it is clear that varying term 8 in the uncertainty budget produces a much greater range of variation in the temperature outputs than is observed for term 1 in the budget. This suggests

that the metrologist may wish to concentrate efforts in experimental design in reducing uncertainties arising from term 8 rather than term 1. It is observations such as this which demonstrate the benefit of this approach to experimental design, as the outputs give clear indications of which sources of uncertainty may dominate. These cannot be calculated straightforwardly from the governing equations, given their non-linear nature.

4.7.2 Comparing the two seeding strategies

As was mentioned in section 4.6.2 we have begun an investigation of possible differences between model outputs when different seeding strategies are chosen. The use of the Wichmann-Hill parallel algorithm ensures that one can be confident that the sequences of random numbers generated by each agent PC on the grid are independent. With other popular seeding strategies one cannot be sure that this is the case. In cases where one uses a PC's own clock as a source of seeds, it is possible on a distributed system that a number of jobs start at approximately the same time, leading to starting seeds that may be similar or perhaps identical. Relying on the in-built random number generators in software may also be problematic, unless one is sure of the quality of that generator and also considers whether the software's own seeding strategy is adequate. We are currently investigating ways of comparing seeding strategies and the results will be the subject of future SS/M publications. An example of outputs from two different seeding strategies are in fact jobs 1540 and 1541 that were discussed in section 4.6.4. These jobs both solved the same Monte Carlo problem (that is, the input uncertainties that were to be varied and their distributions were identical) but job 1540 employed the Wichmann-Hill parallel processing strategy and job 1541 relied on the clock of the agent PC for its seed. The results of the temperature predictions for these two jobs are shown in figure 18. Visual inspection of the two plots is insufficient for identifying differences between the two seeding strategies. Methods of identifying differences are currently the subject of research.

4.8 Conclusions

This case study has demonstrated the benefits of simulations using Monte Carlo methods for a specific experimental design problem that is concerned with minimising uncertainties in a metrology experiment. The problem could not have been tackled using conventional GUM approaches owing to the non-linearity of the governing equations. In addition, the solution time for the mathematical model of the experiment is so long that Monte Carlo studies would be uneconomic on a single PC running a serial Monte Carlo analysis. Parallel Monte Carlo methods on a distributed computing system are shown to be effective in providing solutions to the Monte Carlo problem over a relatively short time scale, that allows metrologists to use simulation outputs to guide their experimental design work. They are likely to become a vital tool in metrology simulations that employ the Monte Carlo method.

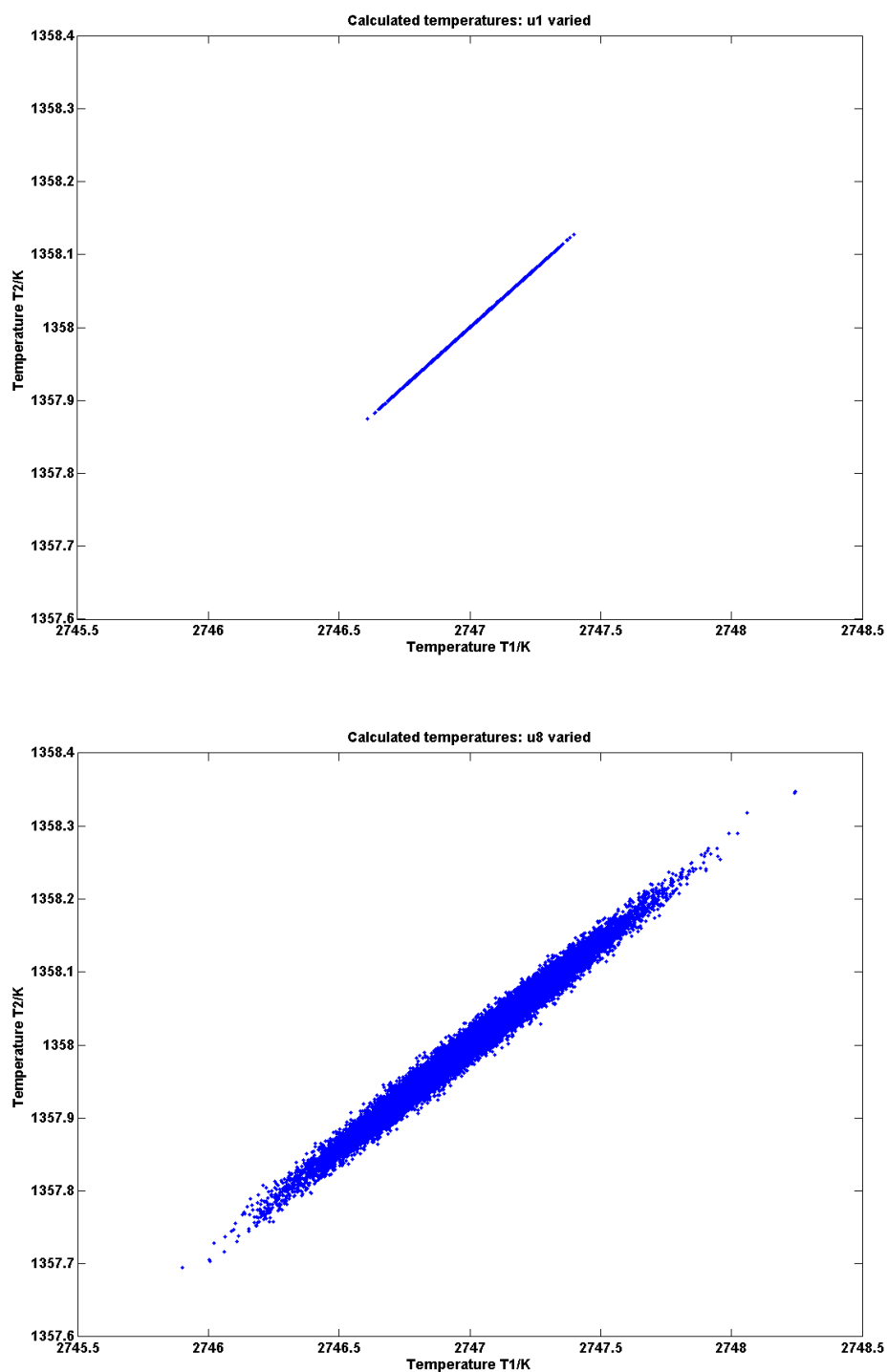


Figure 17: Calculated temperatures, for two separate uncertainty contributions.

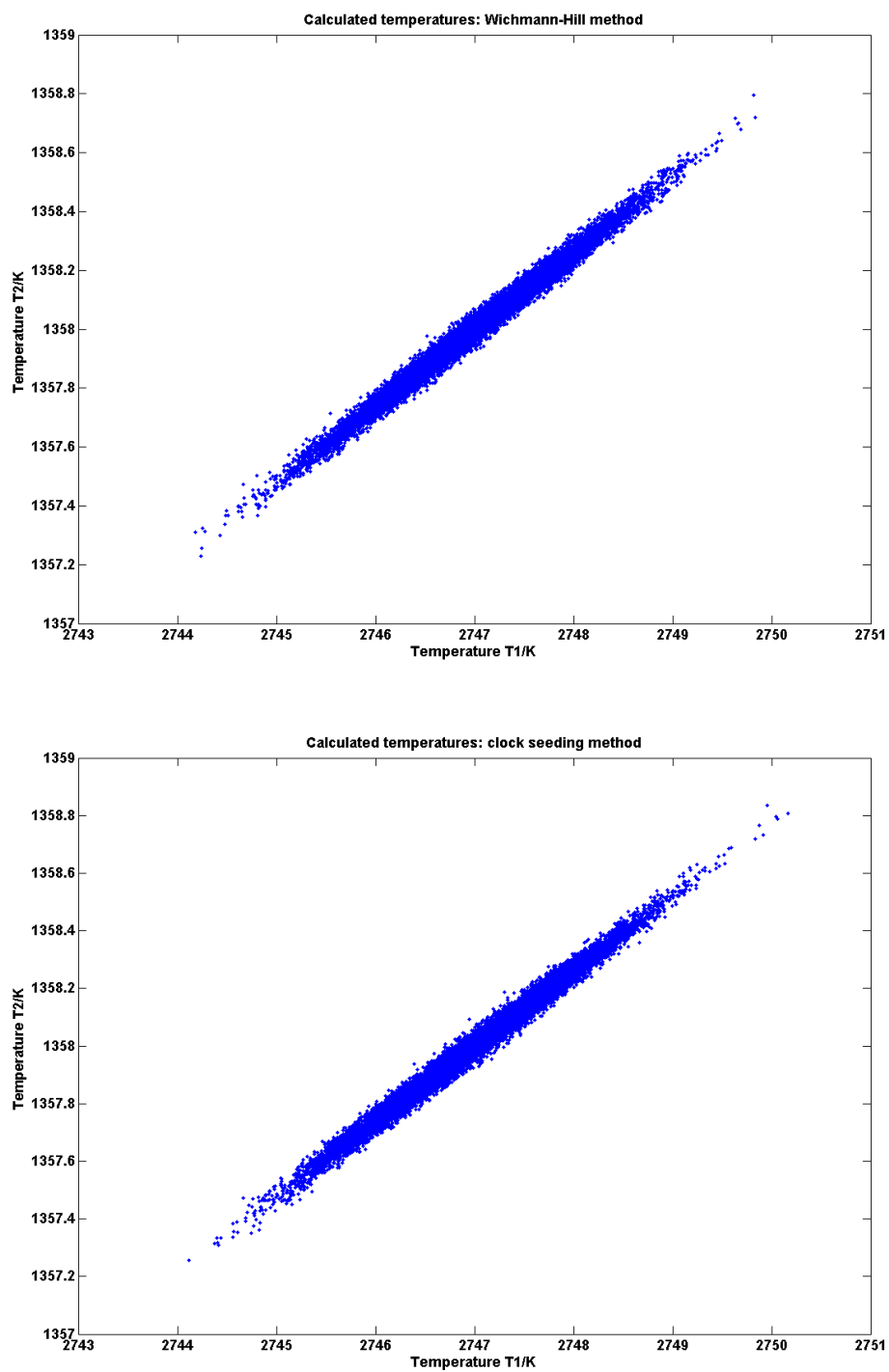


Figure 18: Calculated temperatures, with the two seeding strategies.

5 Uncertainties in molecular dynamics simulations

5.1 Molecular modelling and metrology

Modelling at the nanoscale offers the opportunity to use calculation as a means to determine and characterise parameters which are impossible, unsafe or uneconomic to measure. This makes molecular modelling a vital tool for metrology as it adapts to the challenges of new technologies such as biotechnology and nanotechnology, and seeks to extend the performance at the limits of the continuum approach.

To take an example from the apparently mature area of pressure metrology (which was the subject of the case study in section 3 of this report), in the case of gas flows in pressure balance standards, there is currently no method of evaluating the uncertainties resulting from transition flow (the regime that links hydrodynamic and kinetic descriptions of the flows and surface texture effects), so simple continuum models and perfect surfaces are universally assumed. This lack of scientific understanding is a fundamental limitation to the metrology, making the continually challenging industrial demands on the accuracy of standards increasingly difficult, if not impossible, to achieve.

Similar challenges exist in the newer areas of metrology. For instance, the need to know and understand the orientation of biological molecules on surfaces is crucial to a range of techniques from atomic force microscopy to methods of monitoring biological molecules through macroscopic measurement. Further, in the field of microfluidics, conventional techniques such as finite element analysis and computational fluid dynamics are unable to model the fluid interactions with complex materials required, for example, to design and calculate the behaviour of biosensors and other novel sensor technologies.

Whilst molecular modelling has been developed in universities and by specific NMS programmes for a number of years, the issue of confidence in the modelling results remains of key importance for the adoption and usefulness of the techniques in a metrology context. Confidence in modelling at the nanoscale requires the reconciliation of modelling results across the range of length (and time) scales involved.

As part of our research into nanoscale modelling methods, we carried out a case study to investigate some sources of uncertainty in molecular dynamics models. The case study is described below. As was pointed out in the introduction, the case study is intended to be an illustration of potential sources of uncertainty for those who are new to molecular modelling, rather than an examination of issues that concern experienced modellers.

5.2 Introduction to molecular dynamics simulations

Molecular dynamics simulations typically calculate the individual trajectories (positions and velocities) of a large number of atoms or molecules (the system), and then use the time evolution of the kinetic energy and potential energy functions of the entire system to calculate bulk properties such as temperature and pressure. For simplicity, throughout the following, “atoms” will be used to mean “atoms or molecules” when referring to the particles that make up a model.

The trajectories of the atoms are obtained by calculating the force acting on each atom, applying Newton’s second law to obtain the acceleration of the atom, and using finite difference techniques over a small time step to calculate the new position and velocity of the atom (trajectories cannot be calculated analytically in the general case for more than two atoms). The forces on the atoms are caused by interaction with other atoms, described by a potential energy of interaction, and possibly by external fields. The simplest models have no external fields and only consider the interactions between pairs of atoms. In reality, there may be three- and four-way interactions that affect the potential energy, but these are often neglected or are included in the two-way interactions in some average way giving “effective” two-way forces.

This chapter looks at the uncertainties associated with the results of a very simple molecular model. The aim of the work is not to model a particular physical situation: instead the aim is to look at some of the uncertainties associated with the results of molecular dynamics models by using a straightforward example. The points considered in this chapter will also be of relevance to more general cases.

5.3 Model building

This section discusses some of the aspects of defining the model. Since a very simple example is being used, some common sources of uncertainty such as boundary conditions are not discussed here. Instead, the aspects that have been investigated using simulation runs are presented. The section also introduces the non-dimensionalisation that is used to make the model as general as possible. The non-dimensionalisation also reduces the likelihood of numerical errors by avoiding the need to manipulate very small or very large numbers.

5.3.1 The example problem

The example problem used in this work simulates the motion of 1372 atoms in a cube of side length 10.85 (in non-dimensionalised units). Born-von Karman periodic boundary conditions are applied in each direction, so that the cube is part of an array that extends infinitely in each direction. Whilst this is physically unreasonable, it is a good approximation to the behaviour of atoms in the middle of a large domain. The atoms are subject to interatomic forces,

described by a potential function outlined in section 5.3.2, but there are no external forces being applied and there are no explicit interfaces in the system. The results of interest are the variation over time of the thermodynamic temperature and pressure. In addition, one specific atom (the first in the computer-generated list) has had its trajectory studied in detail. The calculation of thermodynamic properties from atomic trajectories is described in section 5.3.3.

5.3.2 The inter-atomic potential

The example problem uses a Lennard-Jones potential to describe the potential energy of the interaction between two uncharged atoms. This potential, of the form

$$v_{LJ}(r) = 4\varepsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right], \quad (11)$$

where r is the distance between the two atoms and ε and σ are material constants, is commonly used in computer simulations to describe van der Waal's interactions. The force on each atom is the gradient of the potential with respect to the atom's position. The form of the potential's dependence on r reflects the key features of inter-atomic forces: repulsion at short separations, attraction at larger separations, and the force acts along a line linking the two interacting atoms. Typical units for the parameters are nm for σ and K for ε/k_B where k_B is Boltzmann's constant, $k_B = 1.380 \times 10^{-23}$ J/K. The Lennard-Jones potential is particularly good for simulation of noble gases: values of $\sigma \sim 0.34$ nm and $\varepsilon/k_B \sim 120$ K lead to good agreement between model predictions and the experimental properties of liquid argon. The implementation used in this work has chosen non-dimensional scales such that ε and σ are used as units of energy and length respectively. The non-dimensionalisation is explained in more detail in section 5.3.4.

5.3.3 Initial conditions

As with any initial value problem governed by a set of differential equations, molecular dynamics simulations require initial conditions. The initial positions and velocities of the atoms within the system must be specified because the model uses a second-order time derivative (i.e., acceleration).

The initial positions and velocities have an associated potential energy $\mathcal{V}$ and kinetic energy $\mathcal{K}$, defining the initial total energy of the system $\mathcal{E}$. For a three-dimensional system consisting of N atoms numbered 1 to N , the instantaneous temperature $\mathcal{T}$ and pressure $\mathcal{P}$ at any time are given by

$$\mathcal{T} = \frac{2\mathcal{K}}{3Nk_B} = \frac{1}{3Nk_B} \sum_{i=1}^N m_i \dot{\mathbf{x}}_i \cdot \dot{\mathbf{x}}_i, \quad (12)$$

$$\mathcal{P} = \rho k_B \mathcal{T} - \frac{1}{3V_0} \sum_{i=1}^N \mathbf{x}_i \cdot \left(\sum_{j=i+1}^N \nabla_{\mathbf{x}_j} (v_{LJ}\{|\mathbf{x}_i - \mathbf{x}_j|\}) \right), \quad (13)$$

where the atomic positions are $\mathbf{x}_i$, the atomic velocities are $\dot{\mathbf{x}}_i$, V_0 is the volume, ρ is the number density (i.e., number of atoms per unit volume, N/V_0) and m_i is the mass of atom i . In the work described here, only one type of atom is used and so $m_i = m, i = 1, 2, \dots, N$. The averages of $\mathcal{T}$ and $\mathcal{P}$ over time asymptotically approach the thermodynamic temperature and pressure.

Two types of simulation have been run during this work: fixed-energy (free-temperature) and fixed-temperature. In the fixed energy simulations, the total energy remains constant and the system evolves until it eventually reaches an equilibrium state where the instantaneous temperature and pressure fluctuate around their thermodynamic values. In the fixed-temperature simulations, all the velocities are rescaled using a multiplicative factor so that the expression for $\mathcal{T}$ above gives the user's required temperature, i.e., if $\bar{T}$ is the required thermodynamic temperature and $T(t_n)$ is the temperature calculated at time step n , then all the velocities are rescaled by a factor

$$\chi = \sqrt{(\bar{T}/T(t_n))}. \quad (14)$$

In the fixed-energy case, the temperature and pressure will fluctuate around the equilibrium values, but the total energy remains fixed. In the fixed temperature case, the temperature is constant and the total energy and the pressure fluctuate about their average values. The fluctuations can be initially large as the system settles to equilibrium, and when the system approaches equilibrium they become stationary. According to Onsager's regression hypothesis [45, 46] these fluctuations can be used to estimate transport properties of the substance but this has not been attempted here. Any trend in the "fixed" values other than small fluctuations is indicative of a programming error or an instability caused by an unsuitable choice for one of the parameters (often an unsuitable choice of time step).

For the work presented here, two types of initial conditions have been used. One set of conditions has assumed that the atoms are initially at rest. The initial positions are generated by distributing the atoms in a face-centred cubic arrangement (i.e., at the corners and face-centres of an array of cubes) and perturbing each of the co-ordinates by an amount $2d(p - 0.5)$ where d is a small user-specified perturbation distance and p is a uniformly distributed random number between 0 and 1. Since the atoms are at rest, the initial energy of the system (and hence the equilibrium temperature and pressure) is defined by the potential energy, which is dependent on d . An example of the initial locations of the atoms with $d = 0.1$ is shown as a three-dimensional plot in figure 19.

The second type of initial conditions was used only for a set of fixed temperature runs. The initial positions and velocities were the final positions and velocities generated at the end of a run using the same input parameters at a different (lower, fixed) temperature. These runs were carried out to investigate the effects of the initial conditions on the fluctuations. It is expected that the initial large fluctuations of the energy and pressure will die away more quickly for runs where the initial temperature and the fixed temperature are close together than for the runs starting from the perturbed lattice.

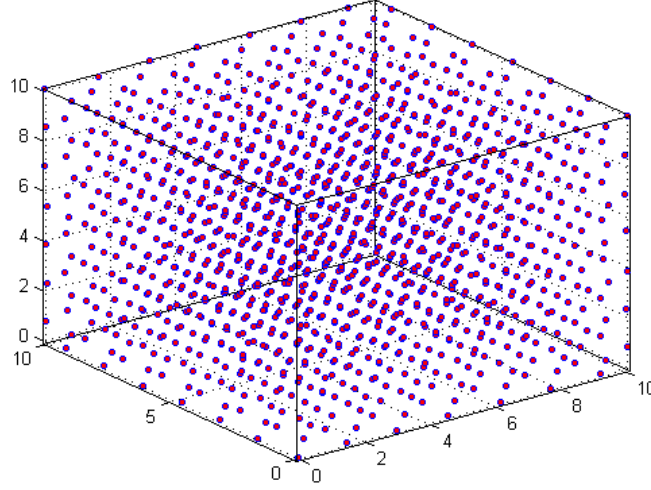


Figure 19: A three-dimensional plot of the initial atomic locations, created by perturbing a face-centred lattice independently in each of the three co-ordinate directions.

5.3.4 Non-dimensionalisation

The implementation used for the example in this report uses dimensionless units. These units are common for several reasons, the main reason being that the calculated results can then apply to any substance that can be described by a Lennard-Jones potential. Another reason is that the speed of calculation can be increased slightly if the non-dimensionalisation is chosen to eliminate some parameters from the model. The non-dimensionalisation chosen in this work was:

$$\hat{\mathbf{x}}_i = \mathbf{x}_i / \sigma, \quad (15)$$

$$\hat{\rho} = \rho \sigma^3, \quad (16)$$

$$\hat{T} = k_B T / \varepsilon, \quad (17)$$

$$\hat{E} = E / \varepsilon, \quad (18)$$

$$\hat{P} = P \sigma^3 / \varepsilon, \quad (19)$$

$$\hat{t} = t [\varepsilon / m \sigma^2]^{1/2}, \quad (20)$$

where $\hat{y}$ is the non-dimensionalised version of the quantity y . The volume, velocities, and accelerations are also non-dimensionalised by dividing by the appropriate combination of time and length scales.

Some researchers prefer to avoid non-dimensionalisation, seeing the choices listed above as an arbitrary set of units and no improvement on SI units. They have been used in the work described here in order to avoid linking the work to a particular physical system, and because in some cases non-dimensionalisation could provide benefits in terms of numerical accuracy.

These scalings should be borne in mind when considering the results: the figures as presented are not in K, nm, and J, because to present the numbers as dimensional qualities would tie the results to a specific set of atoms, and the work presented here is a general investigation better suited to a more generic approach. The “hats” in equations (15) to (20) have been omitted from all subsequent equations for clarity.

5.4 Model solving

The model has been solved using a simple freeware molecular dynamics package called md3. The source code is available from <http://www.ud.infn.it/~ercolessi/md/f90/>, as is source code for a program that generates the crystal lattice for the initial conditions. This section describes some of the assumptions and approximations made within the code, focusing on the assumptions that the model runs have investigated.

5.4.1 Numerical solution method

The problem to be solved can be written

$$m_i \ddot{\mathbf{x}}_i(t) = \mathbf{F}_i(t, \mathbf{x}_j(t), j = 1, 2, \dots, N), \quad i = 1, 2, \dots, N, \quad (21)$$

where $\mathbf{F}_i(t, \mathbf{x}_j(t), j = 1, 2, \dots, N)$ is the force acting on atom i due to all the other atoms and any external conditions. As stated above, this problem cannot be solved analytically for more than two atoms, and so numerical techniques must be used to calculate an approximate solution. The problem is a large system of coupled second-order non-linear ordinary differential equations. There is a very large range of techniques available to solve coupled second-order ordinary differential equations, ranging from simple one-step explicit techniques to complicated methods designed for stiff systems [32]. The methods calculate the positions at a set of discrete times $0 = t_0 < t_1 < \dots < t_M$. Often these times are equally spaced so that $t_m - t_{m-1} = \Delta t, m = 1, 2, \dots, M$, where Δt is called the time step. A uniform time step has been used throughout the work presented here.

Writing the numerical approximation to $\mathbf{x}_i(t)$ as $\mathbf{r}_i(t_m, m = 1, 2, \dots, M)$, the numerical approximations at a given time typically result in a system of the form

$$\sum_{j=1}^{3N} A_{ij} r_j = F_i(r_j, j = 1, 2, \dots, 3N), \quad i = 1, 2, \dots, 3N, \quad (22)$$

where the scalars $r_j, j = 1, 2, \dots, 3N$, are the components of the vectors $\mathbf{r}_i, i = 1, 2, \dots, N$ and the scalars F_i are the corresponding components of the force.

Several factors affect the choice of method used for this particular problem. The first is that there are generally a large number of atoms in models of this type, and so the associated matrix $A_{ij}, i, j = 1, 2, \dots, 3N$ is also large and so

some solution methods can become computationally expensive. The second is that the dependence of the force $\mathbf{F}_i(t, \mathbf{x}_j, j = 1, 2, \dots, N)$ on the atomic positions $\mathbf{x}_j, j = 1, 2, \dots, N$, is generally non-linear, which rules out any implicit methods because the inversion of non-linear formulae is often difficult and computationally expensive, particularly when each force depends on several atomic pair interactions. Finally, the velocities are required at each of the times t_m so that the kinetic energy $\mathcal{K}$ and hence the instantaneous temperature $\mathcal{T}$ can be calculated. This requirement makes many members of the leapfrog class of methods, a set of methods that calculate positions and velocities at different times, inefficient.

The method used in the work presented here is one of the most commonly used methods for molecular modelling problems. It is straightforward to program and requires neither inversion nor iteration, and it gives a good approximation to the solution for reasonably large time steps. The method is called the Verlet algorithm [54, 55] and uses the following steps:

$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \Delta t \mathbf{v}_i(t) + \frac{1}{2} \Delta t^2 \mathbf{a}_i(t), \quad i = 1, 2, \dots, N, \quad (23)$$

$$\mathbf{a}_i(t + \Delta t) = \frac{1}{m_i} \mathbf{F}_i(t + \Delta t, \mathbf{r}_j(t + \Delta t), j = 1, 2, \dots, N), \quad i = 1, 2, \dots, N, \quad (24)$$

$$\mathbf{v}_i(t + \Delta t) = \mathbf{v}_i(t) + \frac{1}{2} \Delta t (\mathbf{a}_i(t + \Delta t) + \mathbf{a}_i(t)), \quad i = 1, 2, \dots, N, \quad (25)$$

where $\mathbf{v}_i$ is the numerical approximation to $\dot{\mathbf{x}}_i$ and $\mathbf{a}_i$ is the numerical approximation to $\ddot{\mathbf{x}}_i$. The atomic accelerations, velocities and positions at time t are used to calculate the positions at time $t + \Delta t$. The new positions are then used to calculate new accelerations, and the accelerations at times t and $t + \Delta t$ are used to calculate the velocities. Hence if the initial positions, velocities, and accelerations are known, the trajectories of the atoms can be calculated for subsequent times.

The particle positions and velocities can be expanded as Taylor series in time in order to estimate the truncation errors caused by using this algorithm. Substituting Taylor expansions into equation (25), it can be shown that

$$\dot{\mathbf{x}}_i(t + \Delta t) = \dot{\mathbf{x}}_i(t) + \frac{\Delta t}{2} (\ddot{\mathbf{x}}_i(t + \Delta t) + \ddot{\mathbf{x}}_i(t)) - \frac{\Delta t^3}{6} \frac{d^4 \mathbf{x}_i}{dt^4} + O(\Delta t^4), \quad i = 1, 2, \dots, N, \quad (26)$$

so the truncation error is proportional to Δt^3 . Similarly, using equation (23) gives

$$\mathbf{x}_i(t + \Delta t) = \mathbf{x}_i(t) + \Delta t \dot{\mathbf{x}}_i(t) + \frac{1}{2} \Delta t^2 \ddot{\mathbf{x}}_i(t) + \frac{\Delta t^3}{6} \frac{d^3 \mathbf{x}_i}{dt^3} + O(\Delta t^4), \quad i = 1, 2, \dots, N, \quad (27)$$

which also has a truncation error proportional to Δt^3 .

It is expected that these truncation errors in the expressions for the atomic positions and velocities will lead to fluctuations in the quantities held constant in the runs (i.e., the energy for the fixed energy runs and the thermodynamic temperature for the fixed temperature runs). In particular, it is expected that the fluctuations (i.e., the standard deviation of the instantaneous values) of the “fixed” quantity are proportional to Δt^2 .

5.4.2 Implementation of the potential function

The Lennard-Jones potential is used in molecular simulations because it is computationally considerably less expensive than many more accurate potentials. This ease of computation makes the potential good for getting approximate results and for testing code prior to implementing a more complicated potential. The potential can be truncated at a certain distance, the cut-off radius r_c , beyond which the potential is zero so that

$$v_{LJc}(r) = \begin{cases} 4\epsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^6 \right] - 4\epsilon \left[\left(\frac{\sigma}{r_c}\right)^{12} - \left(\frac{\sigma}{r_c}\right)^6 \right], & r < r_c, \\ 0, & r \geq r_c. \end{cases} \quad (28)$$

The potential discontinuity, introduced by the truncation, is removed by subtracting a constant quantity for all values of r , making the function continuous at $r = r_c$. The cut-off radius affects the modelled forces because the forces are taken to be zero at larger separations than r_c (meaning that the forces are discontinuous at $r = r_c$), but the potential shift itself does not affect the forces (and hence the equations of motion), because the forces are calculated from the gradient of the potential, which is unaffected by the shift.

The main benefit of this simplification is that interactions between pairs of atoms only need to be considered if they are separated by a distance less than r_c , so the calculations can be carried out more quickly since only a limited number of interacting pairs needs to be considered. A list of the pairs of atoms that are separated by less than the cut-off radius is created within the program and updated periodically to ensure that all required interactions are considered. A further reduction in calculation time can be achieved by using tabulated values for the dependence of the potential on r , and using interpolation techniques to calculate the interactions at intermediate values. This approximation makes little difference to the calculation time when using the Lennard-Jones potential, but it is a generally applicable technique and can significantly decrease calculation time for potentials that use more complicated functions of r .

The main disadvantage with the cut-off radius formulation is that its first derivative is discontinuous at $r = r_c$, which could lead to numerical instabilities in the force calculations, but the effects of the discontinuity should be minimised if the cut-off radius is sufficiently large. This drawback is particularly problematic for interactions between particles that have an orientation as well as a position, the most common example being dipole-dipole interactions. As soon as particles are separated by a distance greater than their cut-off radius, their orientations cease to be correlated which can cause problems for subsequent interactions.

The calculation of the potential (equation 28) requires all atoms separated by a distance less than r_c to interact. It would be inefficient to calculate all interatomic separations at every time step, since it is unlikely that atoms separated by a large distance at the end of one time step will be close enough to interact at the end of the next time step. Instead the software code maintains a list of which atoms are “close” to one another, and updates the

list at intervals throughout the computation. “Close” in this context means that the interatomic separation is less than a distance $r_c + r_s$ where r_s is called the “skin”. All atoms that are close to one another are used in the computation of the forces. The displacement of each atom since the last update of the list is calculated at every time step, and when the sum of the two largest of these displacements is greater than r_s , the list is updated again.

Variation of the skin should not affect the numerical results, since atoms between the cut-off radius and the skin are not included in the calculation, but it may affect run-time. There are two factors connected with the skin that will affect the run time: the first is the length of the close list, and the second is the frequency of updating of that list. The larger the skin, the longer the list will be since more atoms will be separated by a distance less than $r_c + r_s$. The longer the list is, the longer computation of the forces will take, since more pairs of atoms will be used in the computation.

Since the list is updated whenever the sum of two displacements exceeds r_s , a smaller skin will lead to more frequent updating. It is expected that the increase in frequency of updating of the list will have a greater effect on the run time than the increase in the length of the force computations, because updating the list involves calculating the separation of every pair of atoms, whereas the increased list length will increase the time spent on each force calculation by a small amount. The optimum value for the skin thickness is usually found empirically by analysing a series of short runs.

5.5 Post processing

Once a molecular dynamics run is completed, a number of useful calculations can be carried out during post-processing. In addition to the calculation of thermophysical properties, the results can be used to generate information to improve the computational efficiency of subsequent usage of the models and the processing of their results. In particular, analysis of the results can give information about correlation and the reduction of data storage requirements.

In general, the Verlet algorithm uses a large number of small time steps to calculate the atomic trajectories. It is impractical to output the position of every atom within the system at every time step for further analysis, since this would lead to huge amounts of data to be stored and subsequently post-processed. Since a simple time-stepping algorithm is used to calculate the atomic positions, each atomic position is dependent on the position and velocity at the end of the previous time step, and so the positions in consecutive time steps are very strongly correlated. In order to reduce the storage requirements, it would be useful to output only those results that are sufficiently separated in time that they could be regarded as uncorrelated. Strictly, all the results at different times will be correlated, but since the evolution of the atomic positions is caused by interactions of a large number of atoms, it is reasonable to expect the values of the thermodynamic variables of the system (temperature, pressure) at two times separated by a sufficiently large interval to be uncorrelated.

This expectation leads to the need to define what a sufficiently large time interval is. This problem can be approached by defining the normalised auto-correlation function of a quantity $\mathcal{A}$, with mean $\bar{\mathcal{A}}$, by writing

$$c_{\mathcal{A}\mathcal{A}}(m) = \frac{1}{M-m} \sum_{k=1}^{M-m} (\mathcal{A}(t_k) - \bar{\mathcal{A}})(\mathcal{A}(t_{k+m}) - \bar{\mathcal{A}}). \quad (29)$$

This function treats the values of $\mathcal{A}$ at times separated by an interval $t_{k+m} - t_k$ as two separate variables and calculates their correlation. It is expected that the correlation will decrease as m increases. However, as m approaches M , the number of terms in the sum (29) will decrease, and so the quality of the correlation estimate will decline. According to [1], the correlation time can be calculated from the integral of the auto-correlation function. Since the correlation function can take negative values, the root-mean-square integral is used, so that if t_c is the correlation time and the maximum value of m to be used is m_0 , applying the trapezium rule gives

$$t_c \approx \left\{ \sum_{m=0}^{m_0-1} \frac{\Delta t}{2} [(c_{\mathcal{A}\mathcal{A}}(m))^2 + (c_{\mathcal{A}\mathcal{A}}(m+1))^2] \right\}^{1/2}. \quad (30)$$

The block averaging technique [1] aims to evaluate the uncertainty associated with the calculated mean values of the thermodynamic quantities. The following argument applies to temperature and pressure values, but only mentions temperature for brevity.

It is assumed that since the instantaneous temperatures are calculated by summing contributions from a large number of atoms that are interacting in a complicated way, the temperature values can be regarded as samples from a set of (possibly correlated) identical distributions with mean μ and variance σ^2 . Then if the temperature values were independent, by the central limit theorem, the average (i. e., thermodynamic) temperature would be normally distributed with mean μ and variance σ^2/M where M is the number of steps in the calculation. The standard deviation $\sigma/\sqrt{M}$ is an estimate of the uncertainty associated with the calculated mean value.

Since the temperature values are clearly not independent, it is suggested [1] that the variance of a quantity $\mathcal{A}$ can be given by

$$\sigma^2(\bar{\mathcal{A}}) = 2m_{\mathcal{A}}\sigma^2/M, \quad (31)$$

where $m_{\mathcal{A}}$ is the correlation length (a time step number, to be determined as explained below) associated with the quantity $\mathcal{A}$, and σ^2 is the variance of the temperature calculated over the full run of M steps. This formula is reached [27] by considering the case where the values used to calculate the mean consist of a set of blocks of $2m_{\mathcal{A}}$ identical values, and adjusting the formula for the case where there is no correlation in a suitable manner.

In order to determine $m_{\mathcal{A}}$, the M time steps are split into blocks of length L , where $M = L \times B$, and the block averages of the thermodynamic quantities

are calculated. For instance, for a quantity $A(t)$,

$$\bar{\mathcal{A}}_b(L) = \frac{1}{L} \sum_{k=1}^L A(t_{bL+k}), \quad b = 0, 1, \dots, B-1, \quad (32)$$

gives a set of block averages. The block averages are then used to calculate a variance

$$\sigma^2(\bar{\mathcal{A}}_b) = \frac{1}{B} \sum_{b=0}^{B-1} (\bar{\mathcal{A}}_b(L) - \bar{\mathcal{A}})^2. \quad (33)$$

The argument used in deriving equation (31) could also be applied to the block averages, so that

$$\sigma^2(\bar{\mathcal{A}}_b) = 2m_{\mathcal{A}}\sigma^2(\mathcal{A})/L, \quad (34)$$

and so a plot of $\sigma^2(\bar{\mathcal{A}}_b)/\sigma_{\mathcal{A}}^2$ against $1/L$ should be a straight line with gradient $2m_{\mathcal{A}}$. As with the correlation calculations, $\sigma^2(\bar{\mathcal{A}}_b)/\sigma_{\mathcal{A}}^2$ will diverge from this behaviour for large values of L because the number of terms in the sum (33) will be small. Additionally, for small values of L , the block averages will be correlated and so expression (31) will not be a good estimate. The analysis is also affected by the finite precision and limited accuracy of the positions and velocities calculated by the model, which lead to a finite value of σ^2 .

5.6 Sources of uncertainty

From the descriptions above, it is clear that molecular models involve a number of approximations and assumptions that will lead to uncertainties being associated with the results. The contributions to the uncertainty include:

- uncertainty arising from the application of finite difference methods to a continuum problem,
- uncertainty arising from the simplifying assumptions and numerical techniques used during model development (e.g., truncation of the potential),
- uncertainty associated with bulk properties (e.g., thermodynamic temperature and pressure) due to their oscillation around the equilibrium value, and
- uncertainty caused by the use of approximations during the post-processing of results.

Each of these effects has been examined in more detail in the simulation runs that follow.

The question of whether or not the Lennard-Jones potential is a good description of the true inter-atomic potential has been ignored here because this work is not intended to be a model of a particular system: rather it is a

general exploration of the uncertainties that affect molecular dynamics models. As molecular dynamics techniques become more widely used, the number of potential energy models of inter-atomic interactions is increasing. There is much work to be done in identifying the best potential for each problem (in fact, it is likely that the main problem when choosing a potential is a lack of measurement data or other knowledge to support the choice).

5.7 Results

The results presented in this section examine the variation in the calculated results produced and the computational effort required when one input quantity is varied and the others are held fixed. The quantities that have been varied are:

- the time step Δt ,
- the cut-off radius r_c ,
- the skin thickness r_s ,
- the initial perturbation d , and
- the initial temperature (for fixed temperature runs only).

In addition to these values, a single run was carried out that used the analytic expression for the truncated Lennard-Jones potential and its derivative rather than tabulated values. The default values of the quantities listed above were $\Delta t = 1 \times 10^{-3}$, $r_c = 2.5$, $r_s = 0.2$, $d = 0.05$, and all runs were carried out to a total time of $t_M = 5$ unless stated otherwise. The variations of each input quantity will be described in the relevant subsection.

Note that unless stated otherwise, all plots of energy are averaged total energy where the average has been taken across all atoms, so total energy is 1372 times the displayed value.

5.7.1 Effects of varying time step

The variation of Δt shows the effect of using a finite difference method to solve the equations (21). The time step values used in the runs presented here are $\Delta t = 1 \times 10^{-4}$, 2×10^{-4} , 5×10^{-4} , 8×10^{-4} , 1×10^{-3} , 2×10^{-3} , 5×10^{-3} , 8×10^{-3} , 1×10^{-2} , 2×10^{-2} .

According to Allen and Tildesley's book [1] a rule of thumb for the choice of a stable time step is that Δt should be at least an order of magnitude smaller than the Einstein period $t_E = 2\pi/\omega_E$, where the Einstein frequency ω_E is given by

$$\omega_E^2 = \frac{\sum_{i=1}^N \mathbf{F}_i \cdot \mathbf{F}_i / (m_i^2 \sum_{i=1}^N \mathbf{v}_i \cdot \mathbf{v}_i)}{1} = \frac{1}{N} \sum_{i=1}^N \nabla_{\mathbf{r}_i}^2 \left[\sum_{i \neq j} v_{LJc}(\mathbf{r}_i, \mathbf{r}_j) \right], \quad (35)$$

and the second expression can be used for particles at rest. The Einstein period has been calculated at the final time t_M using the results of the run with $\Delta t = 1 \times 10^{-4}$, and the calculated value was 0.34. Hence it would be expected that all the time steps listed above would be sufficiently small to be stable, and examination of the energy results has shown this to be the case (as stated in section 5.3.3, an unstable time step would be expected to lead to a drift in the total energy).

The effects of the varying time step on the trajectory of the first atom are shown in figure 20. The figure plots the projection of the atomic trajectory onto the x - y plane for four different time steps. The starting position for the four lines is the same, because the initial conditions are identical. The trajectories for $\Delta t = 1 \times 10^{-4}$ and $\Delta t = 1 \times 10^{-3}$ are indistinguishable, but there is some separation between the final positions for these time steps and for $\Delta t = 1 \times 10^{-2}$ and $\Delta t = 5 \times 10^{-3}$. As a further insight into how the time step affects the trajectories, the distances between the position of the first atom calculated using $\Delta t = 1 \times 10^{-4}$ and the position of the first atom using other time steps was calculated at 100 time points during the run. These distances are plotted against time in figure 21. If the distances are plotted on a logarithmic scale, the lines representing different time steps are approximately equally spaced. The average of these distances is plotted against time step in figure 22, and the results can be fitted well by a straight line.

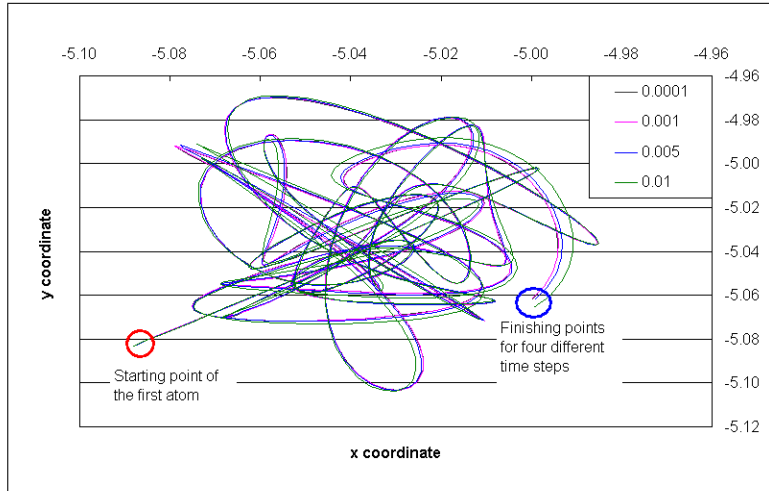


Figure 20: Plot of the projection onto the x - y plane of the trajectory of the first atom calculated using four different time steps. The start and end points are circled.

In general, the temperature and pressure values were not strongly affected by the choice of time step after the first few steps. The average relative difference between the run with $\Delta t = 1 \times 10^{-4}$ and each of the other runs is shown in table 6. The values in this table were reached by comparing the solutions at 125 equally-spaced times throughout the run and calculating the difference relative to the value obtained using $\Delta t = 1 \times 10^{-4}$. For all runs, the relative error decreased in size as the simulation time increased.

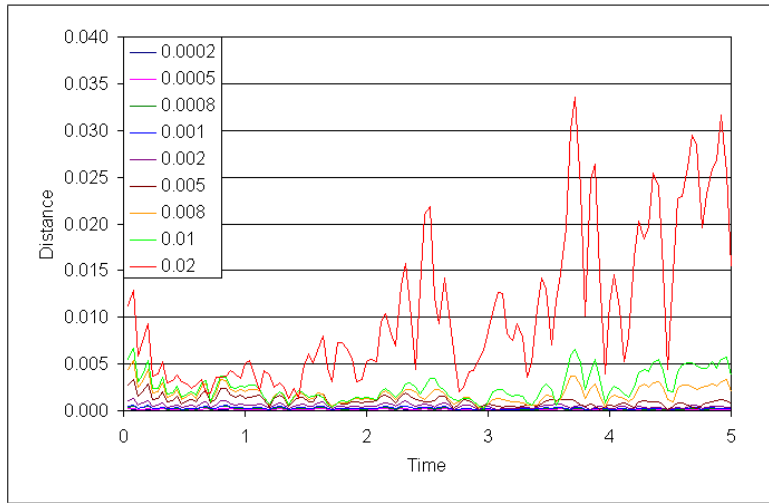


Figure 21: Plot of the distance separating the trajectory calculated using a time step of 1×10^{-4} and the trajectories using other time steps.

Δt	Mean difference	Δt	Mean difference	Δt	Mean difference
2×10^{-4}	2.0×10^{-4}	5×10^{-4}	7.9×10^{-4}	8×10^{-4}	1.4×10^{-3}
1×10^{-3}	1.8×10^{-3}	2×10^{-3}	3.7×10^{-3}	5×10^{-3}	9.4×10^{-3}
8×10^{-3}	1.5×10^{-2}	1×10^{-2}	1.9×10^{-2}	2×10^{-2}	4.3×10^{-2}

Table 6: Average relative difference between temperature values calculated using a time step of 1×10^{-4} and various other time steps.

As was stated in section 5.4.1, the energy fluctuations are expected to be proportional to Δt^2 . The standard deviation of the calculated energy value has been calculated, and is plotted against time step in figure 23 on log-log axes. The deviations from the mean value for the runs with the smallest time steps were too small to produce a good estimate of the fluctuation at the output accuracy used, so not all values have been plotted. The straight line fit shown has a gradient of 1.95, indicating that the fluctuations are approximately proportional to Δt^2 , as expected.

5.7.2 Effects of varying cut-off radius and skin thickness

The cut-off radius is expected to affect both the numerical results and the run time of the model. The larger the cut-off radius, the better the approximation to the full Lennard-Jones potential, but more pairs of atoms will interact which will increase the run time. As was mentioned in section 5.4.2, the skin thickness is not expected to affect the results, but it is expected to affect the run time. Six cut-off radii have been used: 1.1, 1.8, 2.5, 3.2, 3.9, and 4.6. The use of periodic boundary conditions and 1372 atoms meant that the cut-off radius could be 4.65 at most before the sphere of influence of each atom began to self-intersect. The skin thickness was varied in steps of 0.05 between 0.05 and 0.5.

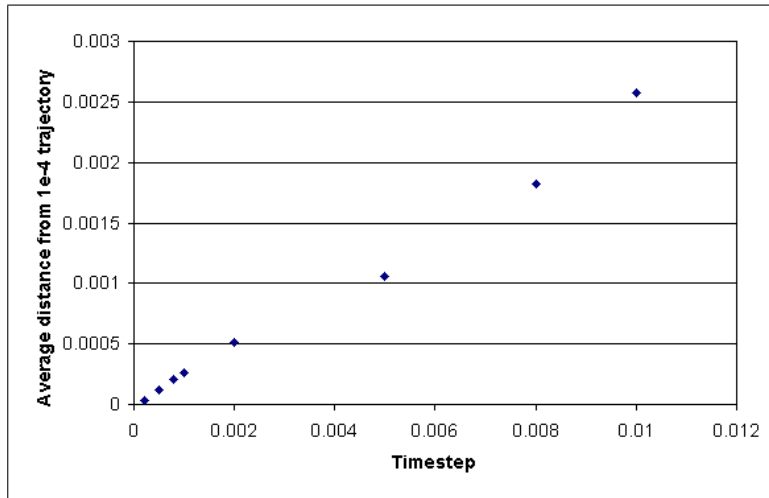


Figure 22: Average separation of each trajectory from that calculated using a time step of 1×10^{-4} plotted against time step.

Figure 24 plots the run time in seconds (calculated from the computer clock: all model runs were carried out with the minimum possible number of other applications running) against the skin thickness. There is no clear trend to the results as the skin thickness increases. It was argued in section 5.4.2 that two processes affect the run time when the skin thickness is increased: the close list will be updated less frequently, but it will be longer and thus will take more time to update. If the atoms do not move around very much, for instance if velocities are low, it is likely that there will be little change to the run time if the skin thickness is altered. This may be the explanation for why there is no clear trend in the results shown in figure 24.

In order to test this idea, another set of runs was carried out with a larger value of d , the perturbation of the lattice conditions. If the initial lattice is perturbed by a larger amount, then the initial potential energy will be increased and hence at subsequent times the total energy and hence the kinetic energy and the average velocity will be increased. Therefore it would be expected that the use of a varying skin thickness would have more of an effect on the run times of a model that starts from a lattice that has been perturbed to a greater degree.

Figure 25 shows the run time versus skin thickness for the larger value of d . The plot shows more of a general trend: as the skin thickness increases, the run time decreases. This was expected because it is likely that frequent updates and computation with a short list will be more computationally expensive than infrequent updates and computation with a long list.

Figure 26 shows the variation in run time when the cut-off radius is changed. There is a clear power law relationship between cut-off radius and run time, and further analysis shows that the exponent is approximately 2.6. This value is approximately equal to the exponent linking a radius to the number of

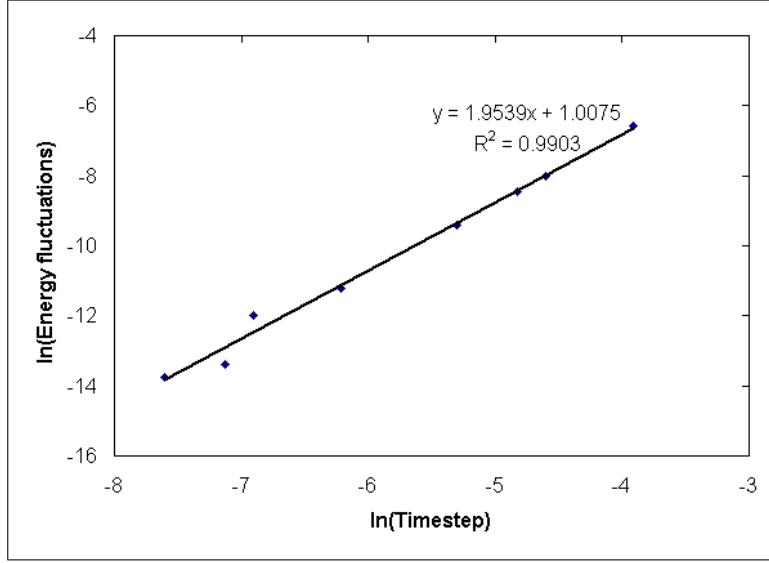


Figure 23: Energy fluctuation plotted against time step on log-log axes. The straight line fit shown indicates that the fluctuations are proportional to the square of the time step.

lattice atoms contained within a sphere of that radius (approximately $N \sim R^3$).

Figure 27 shows the evolution over time of the x co-ordinate of the position of the first atom for different cut-off radii. Initially all of the curves agree, but as time progresses the curves begin to diverge. After 5000 time steps, the curves with $r_c = 1.1$ and $r_c = 1.8$ are quite different from the other radii.

Figures 28 and 29 show the instantaneous temperature and pressure values respectively for the runs. The temperature runs for the different cut-off radii are broadly similar: all runs are oscillating about roughly equal mean values, and the values are virtually indistinguishable for $r_c \geq 2.5$. In contrast, each pressure curve in figure 29 oscillates about a different mean value. This observation means that the pressure results are more sensitive to the cut-off radius than are the temperature results. Figure 30 shows the thermodynamic temperature and pressure plotted against the cut-off radius. The values converge as the radius increases.

The slow convergence of the thermodynamic pressure can be explained by looking at the effects of changing the radius on the calculated accelerations. Suppose that if the cut-off radius is r_1 then the acceleration of the i^{th} molecule at the n^{th} time step is $\mathbf{a}_i^n$, and that if the cut-off radius is $r_2 \neq r_1$, the acceleration becomes $\mathbf{a}_i^n + \Delta\mathbf{a}_i^n$. Then it can be shown that the corresponding new values of $\mathbf{x}_i$ and $\mathbf{v}_i$ are

$$\mathbf{v}_i^n + \Delta t \left(\frac{1}{2} \Delta\mathbf{a}_i^n + \sum_{k=1}^{n-1} \Delta\mathbf{a}_i^k \right), \quad (36)$$

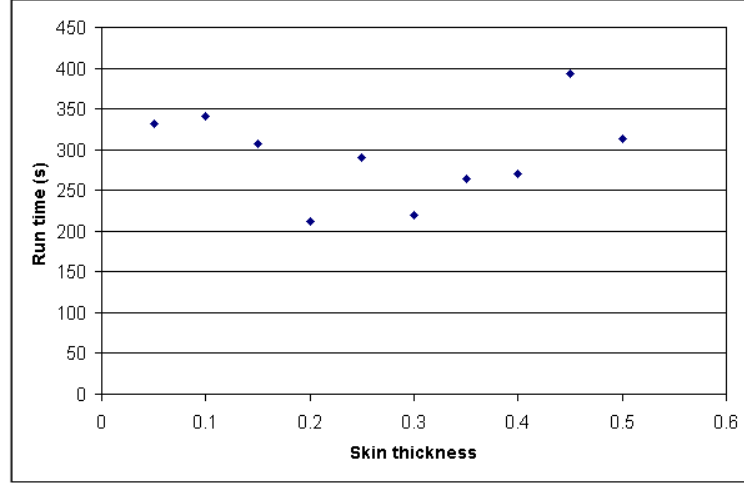


Figure 24: Plot of run time in seconds against skin thickness for $d = 0.05$.

$$\mathbf{x}_i^n + \Delta t^2 \left(\sum_{m=1}^n \sum_{k=1}^m \Delta \mathbf{a}_i^k \right). \quad (37)$$

Hence the change in the velocity is proportional to the time step, and the change in the position is proportional to the square of the time step. The non-dimensionalised versions of the expressions for the instantaneous temperature and pressure can be written

$$\mathcal{T} = \frac{1}{3N} \sum_{i=1}^N \mathbf{v}_i \cdot \mathbf{v}_i, \quad (38)$$

$$\mathcal{P} = \frac{1}{3V_0} \left(\sum_{i=1}^N \mathbf{v}_i \cdot \mathbf{v}_i - \sum_{i=1}^N \mathbf{x}_i \cdot \mathbf{a}_i \right), \quad (39)$$

and so if the cut-off radius is altered, the values of these expressions change to

$$\mathcal{T} = \frac{1}{3N} \sum_{i=1}^N (\mathbf{v}_i + O(\Delta t)) \cdot (\mathbf{v}_i + O(\Delta t)), \quad (40)$$

$$\mathcal{P} = \frac{1}{3V_0} \left(\sum_{i=1}^N (\mathbf{v}_i + O(\Delta t)) \cdot (\mathbf{v}_i + O(\Delta t)) - \sum_{i=1}^N (\mathbf{x}_i + O(\Delta t^2)) \cdot (\mathbf{a}_i + \Delta \mathbf{a}_i) \right) \quad (41)$$

The $\Delta \mathbf{a}_i$ term in the expression for $\mathcal{P}$ will dominate the $O(\Delta t)$ and higher-order terms for a small enough time step. Hence the convergence of the thermodynamic pressure with respect to a changing cut-off radius will be slower than that of the thermodynamic temperature. As r_c increases, $|\Delta \mathbf{a}|$ will tend to zero because the effect of each additional atomic interaction will be smaller, and so the sum (41) will converge.

Allen and Tildesley [1] provide a method for correcting the total energy and the thermodynamic pressure to take long-range interactions neglected due to

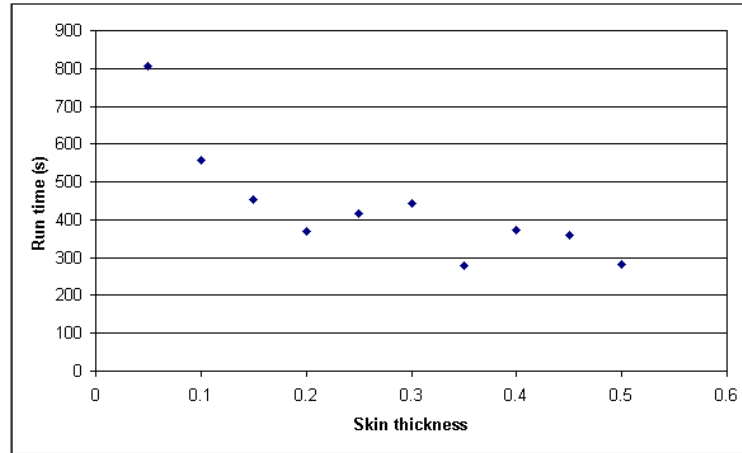


Figure 25: Plot of run time in seconds against skin thickness for $d = 0.15$.

the cut-off radius into account. The method assumes that all separations between two atoms that are greater than the cut-off radius are equally likely, and calculates the contribution to the energy and pressure from the neglected interactions. In reality, some atomic separations are more likely than others, although the variation is slight for large separations. The reference supplies expressions for these corrections for the Lennard-Jones potential, and the corrections have been applied to the results calculated here. Figures 31 and 32 show the corrected and uncorrected values for energy and pressure respectively. The corrected values show improved convergence behaviour, but for small cut-off radii it is clear that the uniformity assumption is insufficient to approximate the true behaviour.

The results of the extra run carried out with full calculation of the potential rather than approximation via interpolation tables were very similar to the results obtained using interpolation. The energy and temperature values differed by at most 0.004%, and the trajectories of the first atom showed similar levels of agreement. The run time was 250 seconds, whereas two other runs carried out with the same parameters and interpolation had run times of 250 and 210 seconds, so the run time for full calculation of the potential lies within the scatter of the other results. It is likely that the use of interpolation would have more effect on the run time if a different (more complicated) potential were being used. Interpolation is unlikely in general to affect the results, provided a suitable interpolation grid is used for the steepness of the potential.

5.7.3 Effects of varying the initial conditions

The perturbed lattice conditions are parameterised by the lattice perturbation d . This initial perturbation defines the total energy, which should be approximately constant throughout the simulation run. The initial perturbation has been varied between 0.025 and 0.25 in steps of 0.025.

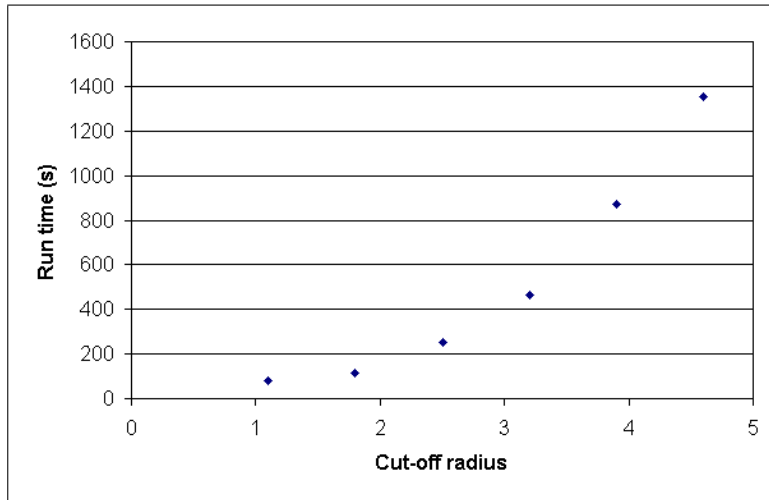


Figure 26: Plot of run time against cut-off radius.

Additionally, a set of runs has been carried out at fixed temperatures of 0.2, 0.4, 0.6, 0.8, and 1.0. For each fixed temperature, one run has been carried out starting from a perturbed lattice with all the atoms at rest, and runs have been carried out starting from the final results of every lower fixed temperature. For example, four runs have been carried out at a temperature of 0.8: one run starting from a lattice, one run each starting from the results of runs at 0.2, 0.4, and 0.6.

Figure 33 shows the variation in energy over time for three different values of d (values of d less than 0.2 showed the same behaviour as $d = 0.2$). The energy values have been normalised by the mean energy for each value of d so that all graphs would fit onto a single set of axes. It is clear that two of the runs are not showing the expected constant energy behaviour: these runs had the largest two values of d . It is likely that these problems are caused by the use of an unstable time step. Section 5.7.1 gave a rule of thumb for determining the stable time step: the time step should be at least an order of magnitude smaller than $2\pi/\omega_E$ where ω_E was defined in section 5.7.1. This quantity has been calculated for all values of d from the initial positions, and the highest two values of d give $2\pi/\omega_E = 0.06$ and $2\pi/\omega_E = 0.03$. According to the rule of thumb, these values should not cause problems for a time step of 1×10^{-3} .

Rerunning the jobs with a time step of 1×10^{-4} has removed the energy drift problem, suggesting that the rule of thumb quoted in section 5.7.1 is not a sufficiently tight upper bound on the time step in this case. For a system in equilibrium, the best way of determining whether a time step is stable is to run the model with a small number of time steps (e.g. 50 to 100 steps) and check for any drift in the energy value. The remainder of the results quoted in this section have been calculated using the appropriate stable time step.

In general, increasing the perturbation increases the total energy of the system, and hence its thermodynamic temperature and pressure. Figure 34 plots the

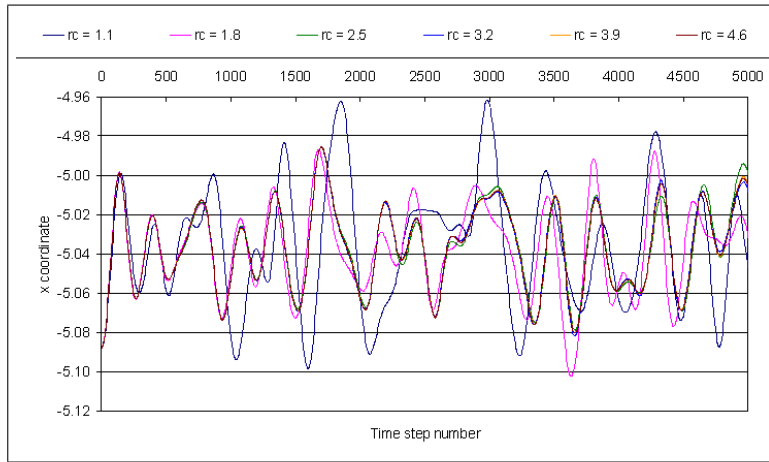


Figure 27: Plot of the evolution of the x coordinate of the first atom over time for various cut-off radii.

logarithms of the mean temperature and pressure calculated for each value of d . Both plots are almost linear, suggesting that the dependence of temperature and pressure on the initial perturbation may be exponential (particularly for larger perturbations). The total energy shows a similar dependence on d .

The fixed temperature runs have instantaneous energy and pressure values that oscillate about a mean value. Figure 35 shows the calculated instantaneous pressure values for a fixed temperature of 1.0. The five curves represent the five different initial conditions, with $T_0 = 0$ on the graph legend indicating that the run started from a perturbed lattice. The figure shows that the initial conditions only affect the first 500 or so steps: after these steps the size of the oscillations about the mean seems to be independent of the initial state. The size of the initial oscillation is dependent on the temperature difference between the fixed temperature of the run and the initial temperature: the run starting from a perturbed lattice shows a larger oscillation than the runs started from higher temperatures, and the cooler initial temperatures give rise to larger oscillations than the hotter ones.

Table 7 reinforces this impression. The data in the table shows the fixed temperature value, the initial temperature (again, 0 is used to mean that the run started from a perturbed lattice), the mean pressure, and two calculated standard deviations. The full standard deviation σ_{full} is the standard deviation of all 5000 instantaneous pressures, and the partial standard deviation σ_{part} is the steps numbered 501 to 5,000. It is clear that for each fixed temperature, the partial standard deviations vary less with the initial conditions than the full standard deviations. The full standard deviations also depend in the expected way upon the initial temperature: temperatures that are close to the fixed temperature give lower standard deviations than those that are further away.

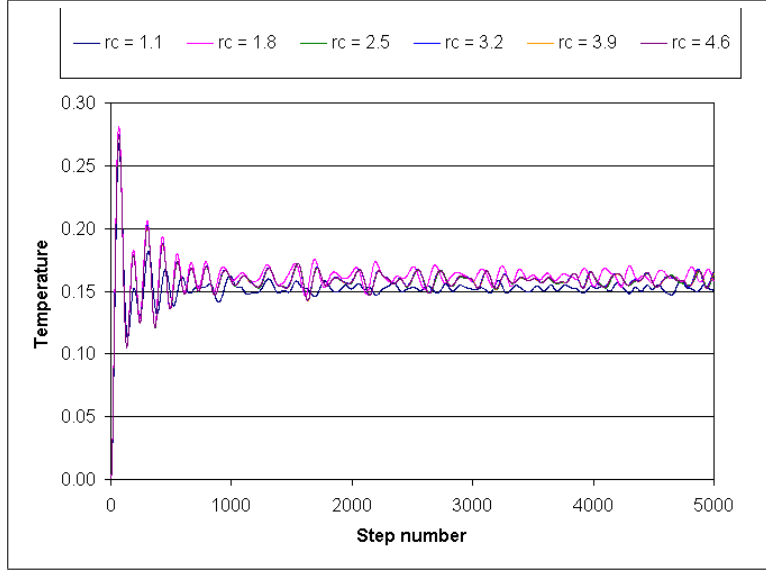


Figure 28: Plot of the evolution of the instantaneous temperature over time for various cut-off radii.

5.7.4 Handling correlation between states

This section examines the post-processing algorithms described in section 5.5. In particular, it looks how the auto-correlation function (29) and the block average standard deviation (33) vary with the number of terms used in the summations in the two equations, and whether the dependence on block length is as predicted.

In order to look at the behaviour of these quantities, an extra simulation was carried out. The time step was set to $\Delta t = 1 \times 10^{-4}$, and the total simulation time was set to 51.051, a total of $M = 510510$ steps. This run time was chosen because $510510 = 2 \times 3 \times 5 \times 7 \times 11 \times 13 \times 17$, and this choice of prime factors leads to a large number of possible block sizes (128 possible sizes, including 1 and 510510).

Normalised auto-correlation functions were calculated for the temperature and pressure for $m = 0, 50, 100, \dots, 255250 (\approx M/2)$. Ideally, the auto-correlation function would decay to zero as m increased. However, for reasons explained in section 5.5, this does not happen in practice which is why only half of the available values of m were used. For large values of m , there are too few terms used in the calculation to produce a good estimate. Figure 36 shows the normalised auto-correlation functions for the temperature and pressure. The two functions are virtually identical, and show the expected behaviour. The auto-correlation function starts at 1.0 and fluctuates with decreasing amplitude about 0. It is not clear why the long-term behaviour seems to consist of a high-frequency oscillation with amplitude determined by a much lower-frequency wave. This behaviour makes it difficult to obtain a correlation time from integration because each burst of oscillations increases the sum in

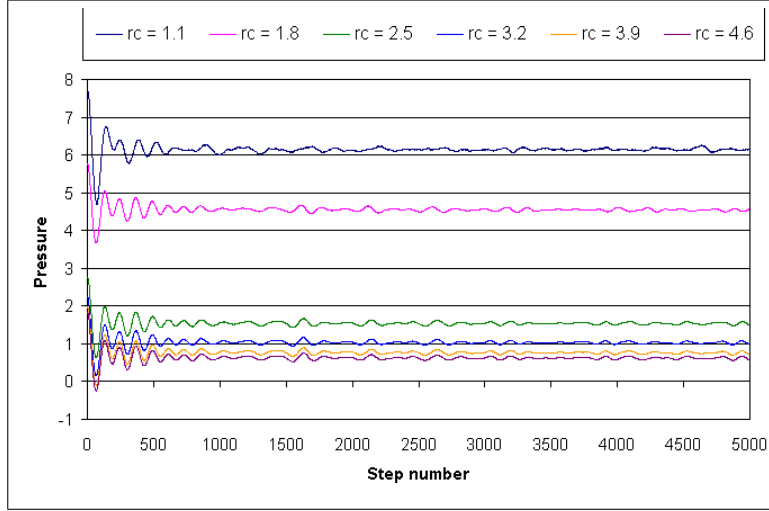


Figure 29: Plot of the evolution of the instantaneous pressure over time for various cut-off radii.

expression 30. The value obtained using the points available was about 0.23 for temperature and for pressure, which seems likely to be an underestimate.

Figure 37 shows the auto-correlation function of the total displacement of the first atom from its starting position. It is clear that this function does not decay to zero as quickly as the autocorrelation functions of the temperature and pressure, and that the long-term oscillatory behaviour has a greater amplitude. This effect may be because the temperature and pressure are calculated by averaging over many atoms, and the averaging process may reduce the correlation time.

Figure 38 shows $\sigma^2(\bar{\mathcal{A}}_b)/\sigma_{\mathcal{A}}^2$ plotted against $1/L$ for the temperature results. Note that $1/L$ is plotted to a logarithmic scale. This plot illustrates the behaviour for large and small blocks well: large blocks have a small variance and small blocks have a variance close to the full sample variance. Pressure results exhibit similar behaviour. It is clear that these results could not be fitted with a straight line as required by equation (34), the results from the large and small blocks being ignored. Additionally, it has not been possible to fit a straight line to any logically-chosen subset of the results and so no value of $m_{\mathcal{T}}$ has been obtained.

An alternative approach outlined in the references [1] is to plot

$$s = \frac{L\sigma^2(\bar{\mathcal{A}}_b)}{\sigma_{\mathcal{A}}^2} \quad (42)$$

and look for a limit as $L \rightarrow \infty$. Figure 39 shows s plotted against L . The plot has a clear peak, but does not seem to have a limit as L increases. It is not clear what is causing this behaviour. It may be that the number of molecules is too small to create the required independence of the temperature values, or it may be that a mistake has been made in the implementation, but the block

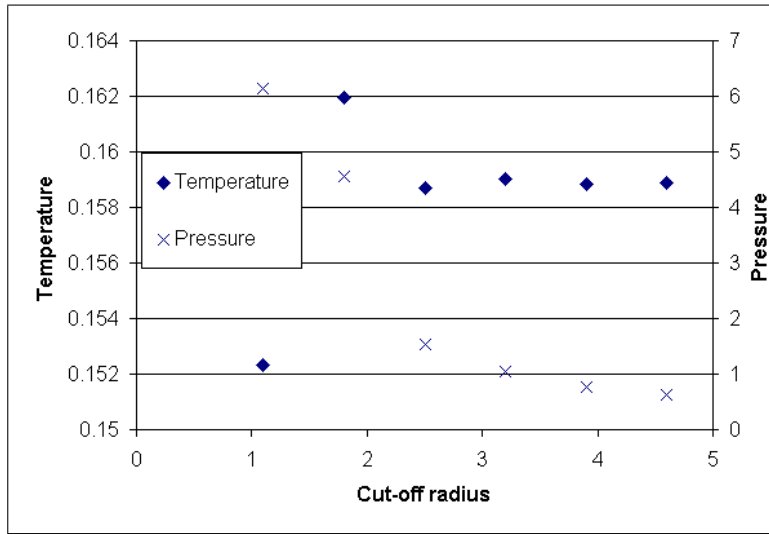


Figure 30: Plot of the thermodynamic temperature and pressure against cut-off radius.

averaging technique does not seem to behave as the references have described.

5.8 Conclusions

This chapter of the report has examined some of the uncertainties associated with molecular dynamics simulations. In particular, it has looked at the effects of the use of a finite difference method, the effects of the various assumptions used to simplify the interatomic interactions, the effects of varying the initial conditions, and how post-processing the data can help to estimate uncertainties and correlations associated with the results.

The choice of a finite difference method for time integration means that a stable time step has to be used. Whilst there are rules of thumb that indicate orders of magnitude for suitable time steps, the most reliable method is to run the model with a small number of time steps and to check that the results are behaving as expected, i.e., the quantities that are supposed to be constant are constant. The truncation errors caused by the finite step size did not have a significant effect on the results over the range of stable steps that was used.

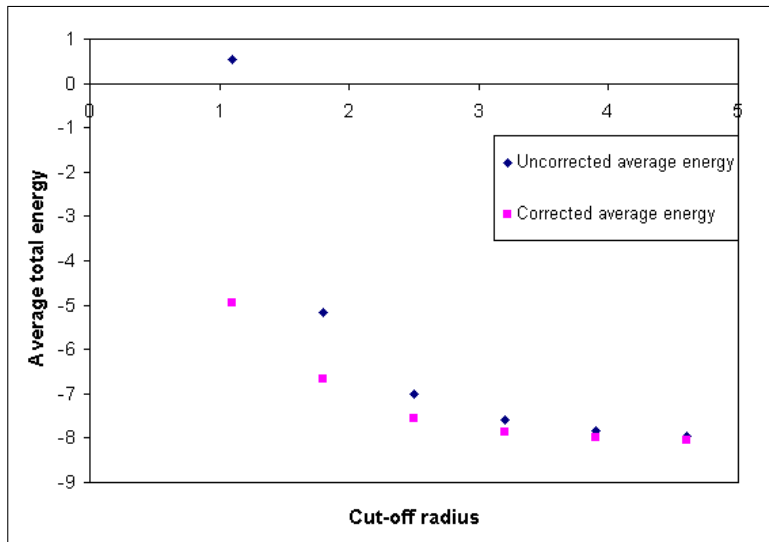


Figure 31: Corrected and uncorrected average total energy.

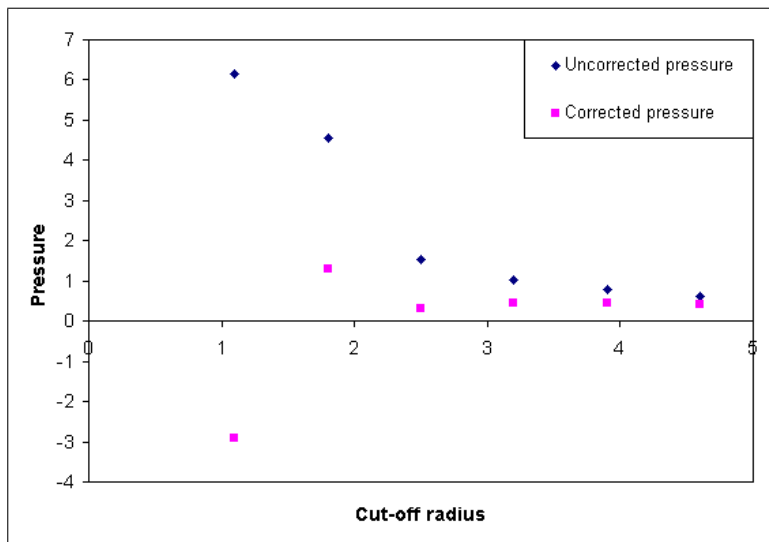


Figure 32: Corrected and uncorrected thermodynamic pressure.

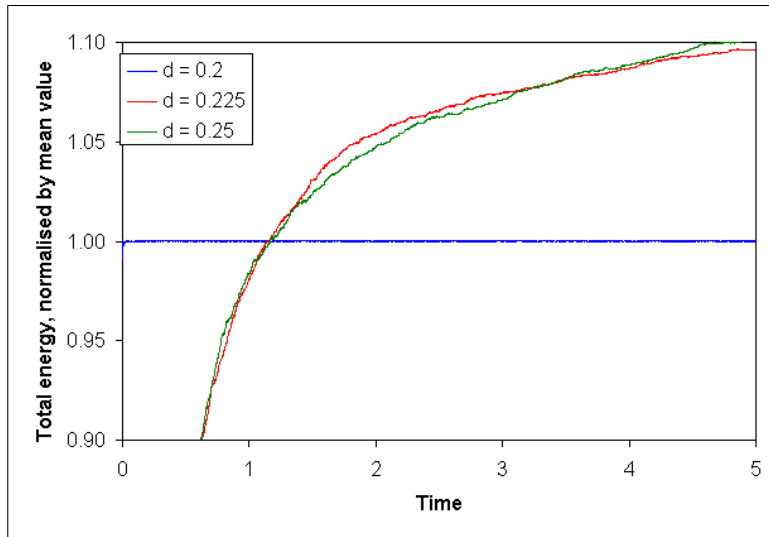


Figure 33: Plot of the evolution of the total energy over time for various values of d .

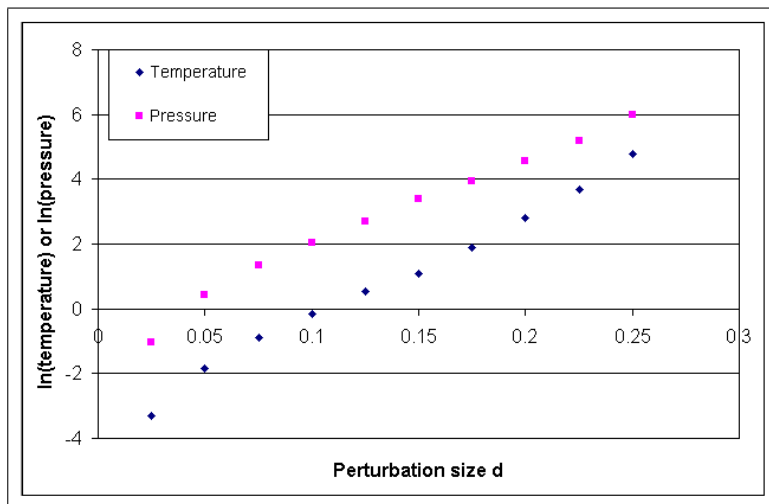


Figure 34: Plot of the logarithms of the thermodynamic temperature and pressure for various values of d .

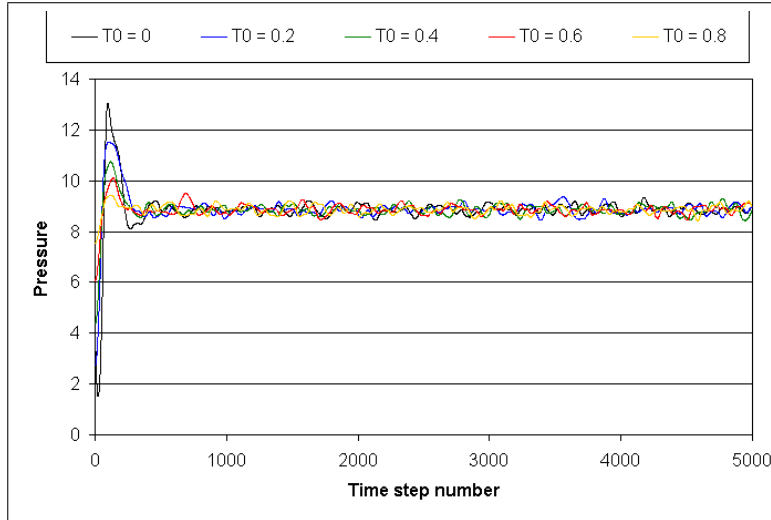


Figure 35: Instantaneous pressure values for a fixed temperature of 1.0. Different curves have been created by starting from different initial conditions.

Fixed temperature	Initial temperature	Mean pressure	σ_{full}	σ_{part}
0.4	0	3.76	0.366	0.079
0.4	0.2	3.77	0.185	0.069
0.6	0	5.51	0.542	0.102
0.6	0.2	5.51	0.341	0.096
0.6	0.4	5.51	0.175	0.109
0.8	0	7.20	0.700	0.140
0.8	0.2	7.21	0.491	0.143
0.8	0.4	7.20	0.303	0.129
0.8	0.6	7.21	0.198	0.162
1.0	0	8.86	0.839	0.155
1.0	0.2	8.86	0.626	0.168
1.0	0.4	8.86	0.435	0.168
1.0	0.6	8.87	0.294	0.167
1.0	0.8	8.86	0.191	0.157

Table 7: Calculated mean pressure, full standard deviation, and partial standard deviation for the fixed-temperature runs. Partial standard deviations were calculated by ignoring the first 500 time steps.

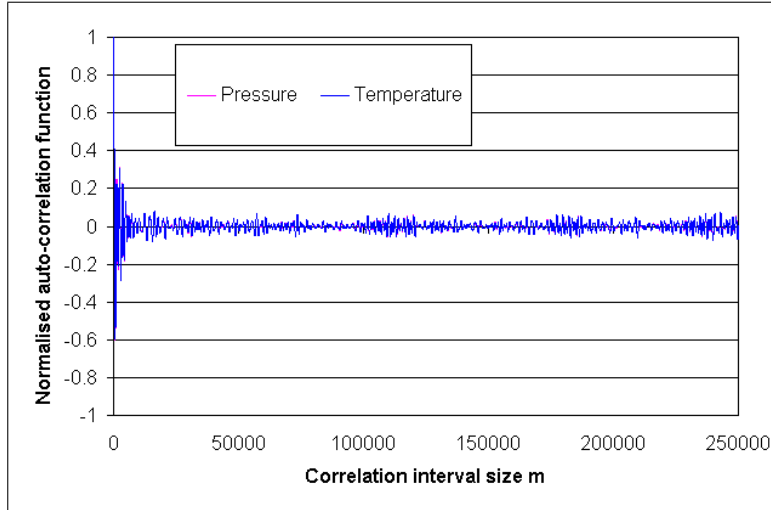


Figure 36: Normalised auto-correlation functions for temperature and pressure plotted against correlation interval m .

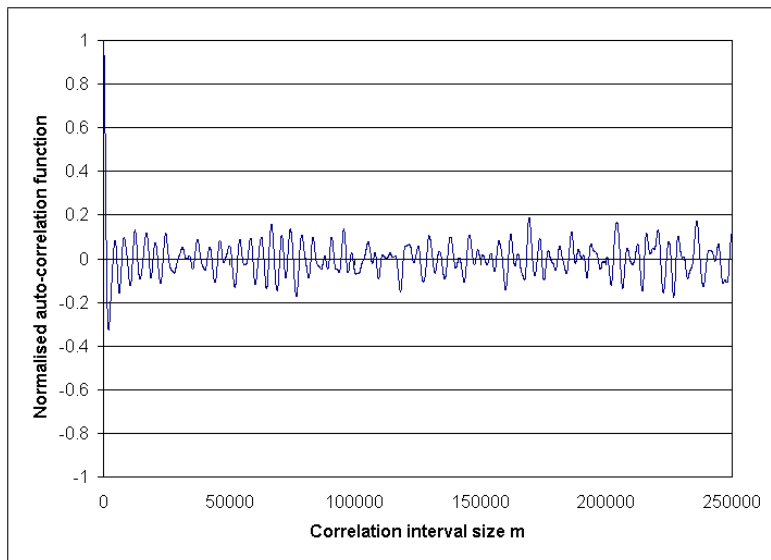


Figure 37: Normalised auto-correlation function for displacement of the first atom plotted against correlation interval m .

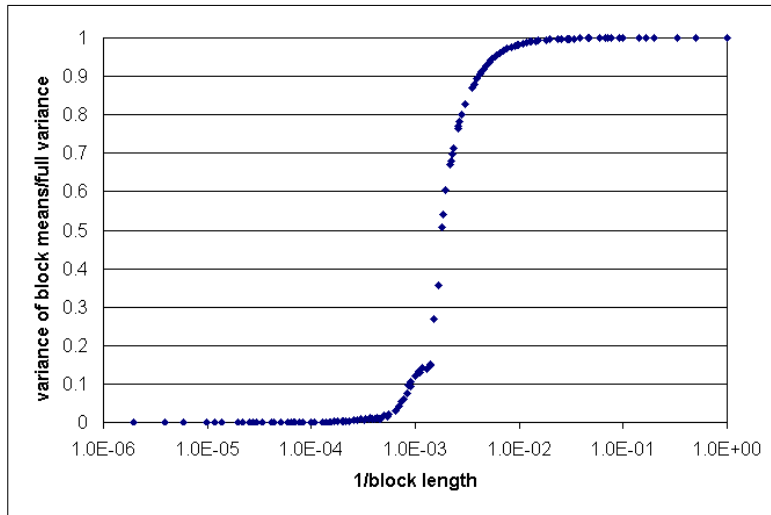


Figure 38: Plot of $\sigma^2(\bar{A}_b)/\sigma_A^2$ against $1/L$ for the temperature results.

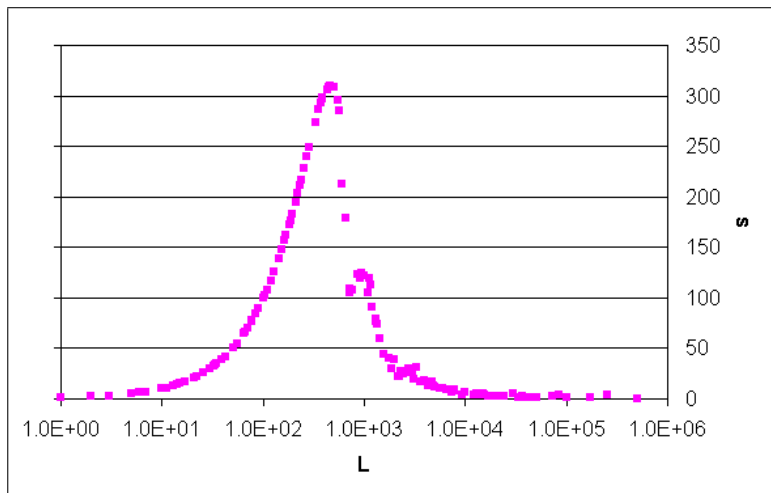


Figure 39: Plot of s against L for the temperature results.

6 Conclusion

The aim of this report is to study simulations and their application to problems in metrology, in particular:

- The use of simulations to study or derive uncertainties in metrology applications, that is, to identify and quantify items that should appear in the uncertainty budgets of metrology experiments;
- Uncertainties in simulations themselves: how do we associate an uncertainty with the outputs of simulations or how reliable are the outputs of simulations?

This has been carried out by means of three case studies from three areas of metrology, which illustrate different aspects of the two main topics identified above.

6.1 Simulating pressure balances

The pressure balance case study showed how finite element methods have been applied to piston-cylinder balances to derive distortion coefficients and their associated uncertainties. In addition, it considered how to quantify uncertainties that arise from the simulation process itself.

Traditionally, the distortion coefficient and its associated uncertainty were determined experimentally or by using simple analytical and semi-analytical models of pressure balance behaviour. However, in recent years finite element methods have become the preferred means of estimating the uncertainty in the distortion coefficient. Uncertainty budgets now typically incorporate a term for the uncertainty associated with the distortion coefficient that is based on finite element simulations of the behaviour of a balance under load.

Developing the model is a challenging process for the following reasons:

- The model must combine two physical processes - the manner in which the pressure balance's structure distorts under stress, and how the pressure balance's pressure-transmitting fluid flows along the small gap between the piston and cylinder as a result of the pressure gradient in the balance.
- One component (either the piston or the cylinder) rotates during the operation of the balance.
- Some balances have complicated geometries, and in addition, definition of boundary conditions can be difficult, especially when external pressures are applied to the cylinder in an attempt to control the manner in which it distorts.

- Material properties are often not known or difficult to obtain. This is often the case for balances made from tungsten carbide and cobalt alloys. Young's modulus is highly dependent on the cobalt concentration in the tungsten carbide matrix, which must be known exactly for accurate modelling.

As a result, simulations have tended to rely on simplifications that seek to capture the main features of balance behaviour without making models intractable for practical use. To date all researchers have modelled balances axisymmetrically. No researchers have reported attempts to model balances in three dimensions without the assumption of axial symmetry. Pressure balance modelling has traditionally begun by considering that the undistorted piston and cylinder are perfectly round and perfectly straight. Although attempts have been made to take into account real shapes of pistons and cylinders these have always been based on deriving some averaged profile for the piston and cylinder that is then used as the basis for an axisymmetric model. Current simulation techniques also ignore the relative movement of the cylinder and piston, as they assume that both components remain stationary.

6.2 Tackling a large Monte Carlo problem in temperature metrology

The second case study introduced simulations using the Monte Carlo method for uncertainty evaluation in a parallel processing environment. The motivation for the work was a study of a technique to estimate the absolute thermodynamic temperatures of two blackbodies by measuring the ratios of their radiances. The uncertainties associated with the estimates of the temperatures were evaluated, and the outputs of the simulations were used to inform the experimental realisation of the technique. The difficulty in computing model sensitivity coefficients means that the use of the GUM uncertainty framework as an approach to the evaluation of uncertainty is not feasible, and a Monte Carlo method was used for this purpose. The computational effort required to evaluate the measurement model is such that without parallel or distributed computing methods the evaluation of uncertainty using a Monte Carlo method would not be practical. To ensure that the results provided by a Monte Carlo method implemented on a distributed computing system are reliable, consideration was given to the approach to generating pseudorandom numbers, which constitutes a key component of the Monte Carlo procedure. The case study provided advice on the choice of random number generator for parallel Monte Carlo calculations and on the practicalities of implementing a distributed version of the simulation software.

6.3 Uncertainties in molecular dynamics simulations

The third case study examined uncertainties associated with molecular dynamics simulations. It looked at the effects of the use of a finite difference

method, the effects of the various assumptions used to simplify the inter-atomic interactions, the effects of varying the initial conditions, and how post-processing the data can help to evaluate uncertainties and correlations associated with the results.

The use of finite difference methods for time integration requires the use of a stable time step. There are rules of thumb that indicate orders of magnitude for suitable time steps, but we have found that the most reliable method is to run the model with a small number of time steps and to check that the results are behaving as expected, i.e., the quantities that are supposed to be constant are constant. The truncation errors caused by the finite step size did not have a significant effect on the results over the range of stable steps that was employed in the case study.

6.4 A caveat

The case studies presented have shown that simulation can be a powerful tool for metrologists to assist in understanding sources of uncertainty and for experimental design. However, one must also be aware of the limitations of simulation. To end this report we quote some words of J.D. Bjorken from a talk given at the 75th anniversary celebration of the Max-Planck Institute of Physics, Munich, Germany, in December, 1992, and quoted in: *Beam Line*, Winter 1992, Vol. 22, No. 4. They refer specifically to Monte Carlo simulations but are of course applicable more widely.

... the Monte Carlo simulation has become the major means of visualization of not only detector performance but also of physics phenomena. So far so good. But it often happens that the physics simulations provided by the Monte Carlo generators carry the authority of data itself. They look like data and feel like data, and if one is not careful they are accepted as if they were data. ... I am prepared to believe that the computer-literate generation (of which I am a little too old to be a member) is in principle no less competent and in fact benefits relative to us in the older generation by having these marvellous tools. They do allow one to look at, indeed visualize, the problems in new ways. But I also fear a kind of "terminal illness", perhaps traceable to the influence of television at an early age. There the way one learns is simply to passively stare into a screen and wait for the truth to be delivered. A number of physicists nowadays seem to do just this.

Acknowledgements

The work described here was supported by the National Measurement System Directorate of the UK Department of Trade and Industry as part of its NMS Software Support for Metrology programme.

The authors would also like to acknowledge vital contributions, particularly to the Monte Carlo case study, from Peter Harris and Ian Smith of NPL's Mathematics and Scientific Computing Group, from Alix de Ginestous, who tackled this topic for her M Sc dissertation at Heriot-Watt University, and from Emma Woolliams and her colleagues Rainer Winkler and Saber Salim from NPL's Optical Detectors Group, whose thermal metrology problem was simulated for this case study.

References

- [1] M.P. Allen and D.J. Tildesley. *Computer Simulation of Liquids*. Oxford University Press, 1993.
- [2] P.J.S. Baird. *Mathematical modelling of the parameters and errors of a contact probe system and its application to the computer simulation of coordinate measuring machines*. PhD thesis, Dept of Manufacturing and Systems Engineering, Brunel University, UK, 1996.
- [3] J.M. Baldwin, K.D. Summerhays, D.A. Campbell, and R.P. Henke. Application of simulation software to coordinate measurement uncertainty evaluation. In *NCSL International Workshop and Symposium*, pages 4E-1, 2004.
- [4] W. Bich, M.G. Cox, and P.M. Harris. Evolution of the 'Guide to the expression of uncertainty in measurement'. *Metrologia*, 43:S161-S166, 2006.
- [5] BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, and OIML. Guide to the Expression of Uncertainty in Measurement, 1995. ISBN 92-67-10188-9, Second Edition.
- [6] BIPM, IEC, IFCC, 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, 2006. Draft (Paris, France: Joint Committee for Guides in Metrology).
- [7] D.N. Chorafas. *Systems and Simulation*. Academic Press, New York, USA, 1965.
- [8] R.S. Dadson, R.G.P. Greig, and A. Horner. Developments in the accurate measurement of high pressures. *Metrologia*, 1:55-67, 1965.
- [9] R.S. Dadson, S.L. Lewis, and G.N. Peggs. *The Pressure Balance: theory and practice*. HMSO, London, UK, 1982.

- [10] A. de Ginestous. Monte Carlo methods for uncertainty analysis in metrology and their application in a parallel processing environment. Master's thesis, Heriot Watt University, Edinburgh, UK, 2006.
- [11] H. Doi, Y. Fujiwara, K. Miyake, and Y. Oosawa. Systematic investigation of elastic moduli of WC-Co alloys. *Met. Trans. A*, pages 1417–1425, 1970.
- [12] C.V. Drysdale, A.H. Ferguson, A.E.M. Geddes, A.H. Gibson, F.R.W. Hunt, H. Lamb, A.G.M. Mitchell, and G.I. Taylor. *The mechanical properties of fluids*. Blackie and Son, U.K., 1925.
- [13] T.J. Esward. Finite element modelling of oil-operated piston cylinder pressure balances: determining uncertainties in distortion coefficients using the finite element method. Technical Report CMAM 85, National Physical Laboratory, Teddington, UK, 2002.
- [14] T.J. Esward, A. de Ginestous, P.M. Harris, I.M. Smith, M.J. Stevens, B.A. Wichmann, and E.R. Woolliams. A Monte Carlo method for uncertainty evaluation implemented on a distributed computing system, 2007. In preparation: for submission to Metrologia.
- [15] T.J. Esward, A.B. Forbes, and A.J. Mansfield. New Directions - future directions in mathematics and scientific computing: recommendations for the software support for metrology programme 2004-2007. Technical Report CMSC 16/03, National Physical Laboratory, Teddington, UK, 2003.
- [16] T.J. Esward, E.G. Johnson, and L. Wright. Testing continuous modelling software. Technical Report CMSC 41/04, National Physical Laboratory, Teddington, UK, 2004.
- [17] T.J. Esward, K. Lees, D. Sayers, and L. Wright. Testing continuous modelling software: three case studies. Technical Report CMSC 42/04, National Physical Laboratory, Teddington, UK, 2004.
- [18] T.J. Esward, G. Lord, and L. Wright. Model validation in continuous modelling. Technical Report CMSC 29/03, National Physical Laboratory, Teddington, UK, 2003.
- [19] T.J. Esward, N.J. McCormick, K.M. Lawrence, and M.J. Stevens. Software Support for Metrology Good Practice Guide no. 17: distributed computing for metrology applications. Technical Report DEM-ES GPG17, National Physical Laboratory, Teddington, UK, 2007.
- [20] T.J. Esward, R.C. Preston, and P.N. Gélat. Piston-cylinder pressure balance design with a negligible distortion coefficient. *Measurement Science and Technology*, 14(6):796–806, 2003.
- [21] T.J. Esward and V.P. Sokhan. Mathematical Modelling for Metrology at the Nanoscale. Technical Report DEM-ES 005, National Physical Laboratory, Teddington, UK, 2006.
- [22] T.J. Esward and L. Wright. Guide to the use of finite element and finite difference software. Technical Report CMSC 30/03, National Physical Laboratory, Teddington, UK, 2003.

- [23] T.J. Esward and L. Wright. Software Support for Metrology Good Practice Guide Continuous Modelling. Technical report, National Physical Laboratory, Teddington, UK, 2005.
- [24] A.B. Forbes and P.M. Harris. Simulated Instruments and Uncertainty Estimation. Technical Report CMSC 01/00, National Physical Laboratory, Teddington, UK, 2000.
- [25] J. Gilby and J. Walton. Software Support for Metrology Good Practice Guide no. 13 Data Visualisation. Technical Report DEM-ES 009, National Physical Laboratory, Teddington, UK, 2006.
- [26] Smith Institute. Mathematics: giving industry the edge. Technical report, Smith Institute for Industrial Mathematics and System Engineering, Guildford, UK, 2002.
- [27] G. Jacucci and A. Rahman. Comparing the efficiency of Metropolis Monte Carlo and molecular dynamics methods for configuration space sampling. *Nuovo Cimento D4*, pages 341–356, 1984.
- [28] G.E. Karnadiakis and J. Glimm. Uncertainty quantification in simulation science. *Journal of Computational Physics*, 217(1):1–4, 2006.
- [29] R. Khanna, editor. *Integrating Personal Computers in a Distributed Client-Server Environment*. Prentice Hall, New Jersey, USA, 2005.
- [30] Min Gon Kim and J. Lamb. The effects of pressure and temperature on the viscosities of liquids composed of long-chain molecules. *Journal of the Korean Physical Society*, 18 (2):109–122, 1985.
- [31] D.E. Knuth. *The Art of Computer Programming, vol. 2, Seminumerical Algorithms*. Addison Wesley, 2002.
- [32] J.D. Lambert. *Numerical Methods for Ordinary Differential Systems*. John Wiley, 1991.
- [33] P. L’Ecuyer. Good parameters and implementations for combined multiple recursive random number generators. *Operations Research*, 47(1):159–164, 1999.
- [34] P. L’Ecuyer and T.H. Andres. A random number generator based on the combination of four LCGs. *Mathematics and Computers in Simulation*, 44:99–107, 1997.
- [35] P. L’Ecuyer and R. Simard. TestU01: A software library in ANSI C for empirical testing of random number generators, 2005. Montréal, Canada: Université de Montréal IRO Available at <http://www.iro.umontreal.ca/~simardr/>.
- [36] G.J. Lord and L. Wright. Uncertainty evaluation in continuous modelling. Technical Report CMSC 31/03, National Physical Laboratory, Teddington, UK, 2003.

- [37] G. Marsaglia. The Marsaglia random number CDROM, with the Diehard battery of tests of randomness, 1985. Florida, US: Florida State University) Available at <http://www.cs.hku.hk/~diehard>.
- [38] M. Matsumoto and T. Nishimura. Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Trans. Model. Comput. Simul.*, 8(1):3–30, 1998.
- [39] J.J. Modi. *Parallel Algorithms and Matrix Computation*. Oxford University Press, Oxford, UK, 1990.
- [40] G. Molinar, W. Sabuga, G. Robinson, and J.C. Legras. Comparison of methods for calculating distortion in pressure balances up to 400 MPa - EUROMET project 256. *Metrologia*, 35 (5):739–759, 1998.
- [41] G.F. Molinar, R. Maghenzani, M. Urbanski, R. Wisniewski, G. Buonanno, and M. Dell’Isola. Elastic distortions of a mixed (simple and re-entrant) piston-cylinder unit for pressure measurements up to 2.6 GPa. *Metrologia*, 36 (6):585–590, 1999.
- [42] T.H. Naylor, J.L. Balintfy, D.S. Burdick, and Kong Chu. *Computer Simulation Techniques*. John Wiley, New York, USA, 1966.
- [43] W.L. Oberkampf and M.F. Barone. Measures of agreement between computation and experiment: validation metrics. *Journal of Computational Physics*, 217:5–36, 2006.
- [44] National Science Foundation Blue Ribbon Panel on Simulation-Based Engineering Science. Simulation-Based Engineering Science. Technical report, National Science Foundation, U.S.A, 2006. On-line http://www.ices.utexas.edu/events/SBES_Final-Report.pdf accessed 22nd September 2006.
- [45] L. Onsager. Reciprocal relations in irreversible processes I. *Physical Review*, 37:405–426, 1931.
- [46] L. Onsager. Reciprocal relations in irreversible processes II. *Physical Review*, 38:2265–2279, 1931.
- [47] G. Robinson. Second NPL analysis of the distortion of pressure balances as part of EUROMET project 256. Technical Report MOM-120, National Physical Laboratory, Teddington, UK, 1996.
- [48] W. Sabuga. Determination of the elastic properties of piston gauges using the strain gauges method. In *Proceedings of the 17th International Conference on Mass, Torque and Pressure Measurements, September 2001*, pages 380–385, Istanbul, 2001.
- [49] W. Sabuga, G. Molinar, G. Buonanno, T.J. Esward, J.C. Legras, and L. Yagmur. Finite element method used for calculation of the distortion coefficient and associated uncertainty of a PTB 1 GPa pressure balance - EUROMET project 463. *Metrologia*, 43:311–325, 2006.

- [50] W. Sabuga, G. Molinar, G. Buonanno, T.J. Esward, T. Rabault, and L. Yagmur. Calculation of the distortion coefficient and associated uncertainty of PTB and LNE 1 GPa pressure balances using finite element analysis-EUROMET project 463. In *Proceedings of the 4th CCM International Conference on pressure metrology from ultra-high vacuum to very high pressures*, volume 42 (6), pages S202–S206, Berlin, 2005.
- [51] N.D. Samaan. *Mathematical modelling of instruments for pressure metrology*. PhD thesis, City University, London, UK, 1990.
- [52] N.D. Samaan. Computer-aided modelling of pressure balances. *Metrologia*, 30 (6):641–644, 1994.
- [53] N.D. Samaan and F. Abdullah. Characterisation of a 1 GPa piston-cylinder unit. *Measurement*, 23 (3):179–183, 1998.
- [54] L. Verlet. Computer “experiments” on classical fluids I: thermodynamical properties of Lennard-Jones molecules. *Physical Review*, 159 (1):98–103, 1967.
- [55] L. Verlet. Computer “experiments” on classical fluids ii: equilibrium correlation functions. *Physical Review*, 165 (1):201–214, 1968.
- [56] D.D. Vvedensky. Multiscale modelling of nanostructures. *J. Phys.: Condensed Matter*, 16:R1537–R1576, 2004.
- [57] B.A. Wichmann and I.D. Hill. Generating good pseudo-random numbers. *Computational Statistics and Data Analysis*, 51(3):1614–1622, 2006.
- [58] E.R. Woolliams, R. Winkler, S.G.R. Salim, and P.M. Harris. The double wavelength technique—an alternative technique to determine thermodynamic temperature, 2007. Submitted to *it International Journal of Thermophysics*.
- [59] L. Wright. Visualisation, uncertainties and continuous modelling. Technical Report CMSC 39/04, National Physical Laboratory, Teddington, UK, 2004.
- [60] L. Wright. Simulation of underwater acoustic radiation from axisymmetric bodies with arbitrary boundary conditions on a distributed computing system, 2007. Submitted to *J. Acoust. Soc. Am.*