

**Reliability and reproducibility in computational science**

Implementing verification, validation and uncertainty quantification in silico

**24 January 2020**

**Abstracts**

**Reproducibility, computability and the scientific method**

Professor Peter V. Coveney, UCL, UvA

We discuss the central importance of reproducibility in the pursuit of scientific knowledge, understanding, and prediction. Unfortunately, at this time it remains true that much computational science is not reproducible, for a number of reasons. Our current work seeks to redress some of these shortcomings through the development of a software toolkit that can provide support for verification, validation and uncertainty quantification, which is applicable and now being used in multiple different domains. In many application domains that one encounters in high performance computing contexts, one must confront both systematic and random errors; in the latter case, one commonly has to deal with chaotic dynamical systems. The question of how to sample these reliably and effectively is integral to the credibility of reported results. We have recently shown that such dynamical systems, when represented on digital computers, manifest a new pathology of the IEEE floating point numbers, which cannot be removed regardless of the level of precision used. In the case of a very simple but representative dynamical system, the generalised Bernoulli map, for which exact results are known, we find that the errors produced by floating point arithmetic can sometimes be catastrophic while, in more generic cases, the errors are of order unity. This indicates that many computational studies of chaotic systems, such as arise in turbulence and molecular dynamics, are likely to contain significant errors which hitherto are unknown to the community of practitioners. I will conclude with some suggestions for how to address this pathology.

**Towards Validated Multiscale Simulations for Fusion**

Dr Onnie Luk, Institute for Plasma Physics, Max-Planck-Gesellschaft

Harnessing energy produced by the thermonuclear fusion reaction has the potential to provide a clean and inexpensive source of energy to Earth. However, throughout the past seven decades, physicists learned that creating our very own fusion energy source is very difficult to achieve. This is mainly due to the fact that the physics problem spans into disparate temporal and spatial scales, and it cannot be fully understood with single scale models. We constructed a component-based, multiscale fusion workflow to model fusion plasma inside the core of a tokamak. To ensure the simulation results agree with experimental values, the model needs to undergo the process of verification, validation, and uncertainty quantification (VVUQ). This talk will go over the VVUQ work carried out in the multiscale fusion workflow, with the help of the EasyVVUQ software library developed by the VECMA project.

**Semi-intrusive Uncertainty Quantification for Reliable Simulations**

Dr Anna Nikishova, University of Amsterdam

There is growing alarm about the reproducibility crisis in Computational Science: lack of possibility to repeat the experiment with the same method, code and input data and obtain the same output of the simulation. Projects like VECMA have in their goal to establish standardised techniques of Validation, Verification and Uncertainty Quantification (UQ) that would ensure that a computational model produces reliable and reproducible results with estimates of the accuracy and precision, therefore, ensure that the response of the simulations can be applied in the decision making process. In this talk, I will make a survey on reliability and reproducibility in computational science. Then, I will focus on UQ and will discuss how to perform efficiently UQ for coupled multiscale models. I will give an overview of the semi-intrusive UQ methods, which are based on metamodeling and sensitivity analysis and will demonstrate their efficiency on a biomedical application.

**Uncertainty and unresolved processes**

Professor Daan Crommelin, CWI Amsterdam

In multiscale dynamical systems such as atmosphere and oceans, physical processes over a wide range of space and time scales play a role. From planetary-scale circulation patterns to small-scale turbulence and cloud formation, there are many processes active, interacting with each other in a nonlinear way. In numerical simulations of such systems, not all of these processes can be explicitly resolved, as resolving all scales is computationally not feasible. As a result, processes at small scales remain unresolved, and their feedback onto the larger scales is represented in a simplified manner. This creates a source of uncertainty in simulations. In my talk I will discuss this type of uncertainty and how it can be represented by means of stochastic modelling.

**Reducing Uncertainty in Agent-Based Simulations through Real-Time Data Assimilation**

Professor Nick Malleson, University of Leeds

Agent-based modelling has been shown to be a valuable method for modelling systems whose behaviour is driven by the interactions of distinct entities. However, the methodology faces a fundamental difficulty: there are no established mechanisms for dynamically incorporating real-time data into models. This work begins to address this gap by demonstrating how data assimilation techniques can be used to allow data to stream into running models to reduce the uncertainty in their estimates of the current state of the target system. This allows for better predictions of the present and potentially more accurate short-term forecasts. The methods are illustrated using a case study of pedestrian movements. By laying the groundwork for the real-time simulation of crowd movements, this work has implications for the management of complex environments such as transportation hubs, hospitals, shopping centres, etc.

**Importance of GP emulation and history matching**

Dr Victoria Volodina, University of Exeter and The Alan Turing Institute

Computer simulators (models) are widely used to study of interest across a wide range of fields. Due to complexity of these models it is commonly time-consuming and expensive to produce multiple runs of a simulator. As a solution Gaussian process (GP) emulators, a type of surrogate, are used to produce high-fidelity prediction together with uncertainty for computer model output. There is an interest to solve an inverse problem, which relates simulator to the real data and deduces the parameter values needed to run the simulator. One of the approaches to inverse problem is history matching. History matching combined with GP emulator is performed by ruling out regions of input parameter space that are considered unlikely to give model output consistent with real world data. In this talk we demonstrate how these two important concepts in Uncertainty Quantification (UQ) are implemented in mogp\_emulator software library.