Bayesian Emulation of Complex Computer Models with Structured Partial Discontinuities

J. Owen
Durham University

Supervisor(s): Dr. Ian Vernon (Durham University) and Prof. Michael Goldstein (Durham University)

Ph.D. expected duration: Oct. 2017 - Sep. 2021

Adress: Durham University, Mathematical Sciences & Computer Science Building, Upper Mountjoy Campus, Stockton Road, Durham University, DH1 3LE

Email: jonathan.owen@durham.ac.uk

Abstract:

Complex mathematical computer models are used across many scientific disciplines and industry to improve the understanding of the behaviour of physical systems and provide decision support. A major limitation is their long evaluation times, thus performing a full uncertainty quantification that frequently requires a large number of evaluations becomes prohibitively computationally expensive. Instead, computer models are often substituted by emulators which provide a statistical approximation, yielding predictions for as yet unevaluated parameter settings along with a statement of the corresponding uncertainty. Crucially emulators are extremely fast to evaluate and thus facilitate the large number of model evaluations required in full uncertainty quantification calculations.

Emulators are typically Bayesian constructs which benefit from underlying judgements regarding the form of the computer model output as a function of the inputs. It is very common to judge the function to be continuous, as well as smooth or potentially possessing several derivatives. Issues arise for Gaussian process based emulators of the described form when the function has various types of structured discontinuity. The problem we address is the emulation of functions that possess a finite set of discontinuities of possibly complex form, where the endpoint locations of each discontinuity may lie within the input space; the discontinuities do not necessarily bisect the input space. Figure 1 illustrates a simple 2-dimensional example of such a problem. Standard Gaussian Process emulation methodology struggles to handle such discontinuities, specifically due to endpoints being within the input space, since it is hard to define valid covariance functions that allow for such behaviour. Below we discuss a motivating example from the oil industry (also shown in fig. 2) where these discontinuities are substantial and hence must be directly addressed.

We propose embedding the computer model input space on a hypersurface within a higher dimensional space. The choice of this embedding hypersurface is relatively flexible, but is torn along the locations of the existing structured discontinuities. In the higher dimensional space there are no discontinuities; these are entirely handled by the torn embedding hypersurface, hence a Gaussian process emulator is constructed over this higher dimensional space with free choice from the typical large class of emulator forms. Obtaining an emulator over the original input space which respects the discontinuities is achieved via a projection. However, a naïve implementation results in a distorted and warped induced covariance function and hence emulator over the original input space. This is a consequence of the non-linear, warped and potentially stretched embedding transformation. We present a linearised correction by deriving a non-stationary covariance function on the higher dimensional input space that accounts for the effect of the non-linear embedding. This requirement is relatively weak, ensuring the user still has the full flexibility to choose the lower-dimensional induced covariance function.

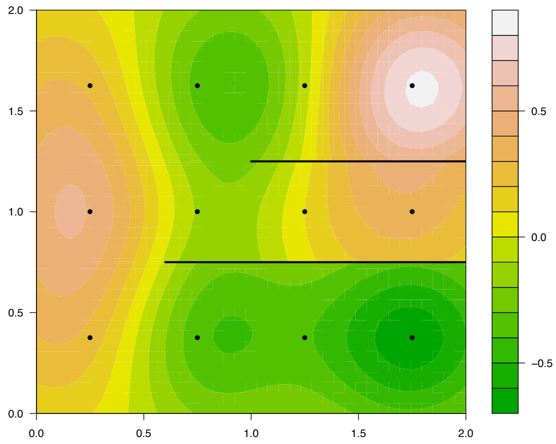

Figure 1: Toy example of the emulator adjusted expectation for a function with a 2-dimensional input space. The black lines depict the discontinuities and the black points the model run locations.

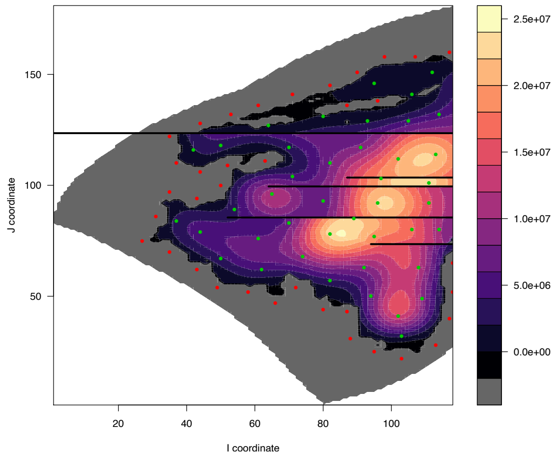

Figure 2: TNO OLYMPUS well placement NPV emulator adjusted expectation over the OLYMPUS reservoir map. Green points denote the reservoir model run locations and the black lines the geological fault discontinuities.

The described methodology is applied to a toy model with 2-dimensional input space and two partial discontinuities of different lengths, both possessing one endpoint on the boundary whilst the other is within the input space. Figure 1 shows the emulator adjusted expectation surface and demonstrates that it respects the discontinuities (black lines) whilst remaining continuous everywhere else, as would any individual realisation from this emulator.

This research is motivated by a problem from the oil industry which is seen in the TNO OLYMPUS Well Placement Optimisation Challenge. The aim is to devise a drilling strategy including the location, type and order of wells for a new oil reservoir in order to maximise the Net Present Value (NPV – the total discounted difference between the revenue and expenditure) over the field lifetime. There exists numerous geological faults which induce discontinuities in the NPV surface as a function of well location. Application of the described methodology to the first producer well placement is illustrated in fig. 2, which shows the NPV emulator adjusted expectation surface possessing discontinuities across the black lines representing the geological faults. The green points are the reservoir model run locations.

The proposed emulation methodology is able to handle general forms of structured partial discontinuity whilst still maintaining a flexible choice of emulator form. This methodology has been successfully applied to the TNO OLYMPUS Challenge and forms part of an ongoing well placement project. It is also fully generalisable to other scientific and industrial applications possessing such structured discontinuities.

Short biography – Jonathan Owen is a PhD candidate at Durham University researching in Bayesian uncertainty analysis and decision support for complex models of physical systems with application to oil field development optimisation. The work is funded by an iCASE studentship funded by EPSRC and Emerson, and supported by the Smith Institute.