ESPRESSO: Earth Science Problems for the Evaluation of Strategies, Solvers and Optimizers

Andrew Valentine¹, Jiawen He², Juerg Hauser³, and Malcolm Sambridge²

¹Department of Earth Sciences, Durham University, Durham, UK (andrew.valentine@durham.ac.uk)
²Research School of Earth Sciences, The Australian National University, Canberra, Australia
³Commonwealth Scientific and Industrial Research Organisation, Canberra, Australia

Many Earth systems cannot be observed directly, or in isolation. Instead, we must infer their properties and characteristics from their signature in one or more datasets, using a variety of techniques (including those based on optimization, statistical methods, or machine learning). Development of these techniques is an area of focus for many geoscience researchers, and methodological advances can be instrumental in enhancing our understanding of the Earth.

In our experience, progress is substantially hindered by the absence of infrastructure facilitating communication between sub-disciplines. Researchers tend to focus on one area of the earth sciences — such as seismology, hydrology or oceanography — with only slow percolation of ideas and innovations from one area to another. Indeed, silos often exist even within these subfields. Testing new ideas on new problems is challenging as it requires the acquisition of domain knowledge, an often difficult and time-consuming endeavour with uncertain returns. Key questions that arise include: What is a relevant field data set, and how has it been processed? Which simulation package is most appropriate to predict the data? What would a 'good' model look like and what should it be able to resolve? What is the current best practice?

To address this, we introduce the ESPRESSO project — a collection of Earth Science Problems for the Evaluation of Strategies, Solvers and Optimisers. It aims to provide access to a suite of ‘test problems’, spanning a wide range of inference and inversion scenarios. Each test problem defines appropriate dataset(s) and simulation routines, accessible within a standardised Python interface. This will allow researchers to rapidly test new techniques across a spectrum of problems, share domain-specific inference problems and ultimately identify areas where there may be potential for fruitful collaboration and development. ESPRESSO is envisaged as an open, community-sourced project, and we invite contributions from across the geosciences.