Geoscientific Model Development
An Interactive Open Access Journal of the European Geosciences Union

Geosci. Model Dev., 6, 1715-1728, 2013
www.geosci-model-dev.net/6/1715/2013/
doi:10.5194/gmd-6-1715-2013

The potential of an observational data set for calibration of a computationally expensive computer model

D. J. McNeall[1], P. G. Challenor[2], J. R. Gattiker[3], and E. J. Stone[4]
[1]Met Office Hadley Centre, Exeter, UK
[2]College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter, UK
[3]Los Alamos National Laboratory, Los Alamos, New Mexico, NM 87545, USA
[4]School of Geographical Sciences, University of Bristol, Bristol, UK

Abstract. We measure the potential of an observational data set to constrain a set of inputs to a complex and computationally expensive computer model. We use each member in turn of an ensemble of output from a computationally expensive model, corresponding to an observable part of a modelled system, as a proxy for an observational data set. We argue that, given some assumptions, our ability to constrain uncertain parameter inputs to a model using its own output as data, provides a maximum bound for our ability to constrain the model inputs using observations of the real system.

The ensemble provides a set of known parameter input and model output pairs, which we use to build a computationally efficient statistical proxy for the full computer model, termed an emulator. We use the emulator to find and rule out "implausible" values for the inputs of held-out ensemble members, given the computer model output. As we know the true values of the inputs for the ensemble, we can compare our constraint of the model inputs with the true value of the input for any ensemble member. Measures of the quality of constraint have the potential to inform strategy for data collection campaigns, before any real-world data is collected, as well as acting as an effective sensitivity analysis.

We use an ensemble of the ice sheet model Glimmer to demonstrate our measures of quality of constraint. The ensemble has 250 model runs with 5 uncertain input parameters, and an output variable representing the pattern of the thickness of ice over Greenland. We have an observation of historical ice sheet thickness that directly matches the output variable, and offers an opportunity to constrain the model. We show that different ways of summarising our output variable (ice volume, ice surface area and maximum ice thickness) offer different potential constraints on individual input parameters. We show that combining the observational data gives increased power to constrain the model. We investigate the impact of uncertainty in observations or in model biases on our measures, showing that even a modest uncertainty can seriously degrade the potential of the observational data to constrain the model.

Citation: McNeall, D. J., Challenor, P. G., Gattiker, J. R., and Stone, E. J.: The potential of an observational data set for calibration of a computationally expensive computer model, Geosci. Model Dev., 6, 1715-1728, doi:10.5194/gmd-6-1715-2013, 2013.