Multi-objective tuning of GENIE Earth system models

A. R. Price[1], I. I. Voutchkov[1], N. R. Edwards[2], J. K. Hughes[3], D. J. Lunt[3], T. M. Lenton[4], P. J. Valdes[3], S. J. Cox[1] and the GENIE team[5]

(1) School of Engineering Sciences, University of Southampton, Southampton SO17 1BJ, UK
(2) Earth Sciences, The Open University, Milton Keynes, UK
(3) School of Geographical Sciences, University of Bristol, Bristol BS8 1SS, UK
(4) School of Environmental Sciences, University of East Anglia, Norwich NR4 7TJ, UK
(5) http://www.genie.ac.uk/people

GENIE
Grid ENabled Integrated Earth system model
www.genie.ac.uk

NATURAL ENVIRONMENT RESEARCH COUNCIL

GENIE*fy*: Grid ENabled Integrated Earth system model *for* the Community

This article can be found at www.e-science.soton.ac.uk/posters

Abstract

We present the results of applying a multi-objective Non-dominated Sorting Genetic Algorithm (NSGA-II) to tune two models from the GENIE framework. The genie-eb-go-gs (3D frictional geostrophic ocean model, 2D energy moisture balance model and 2D sea-ice) and genie-ig-fi-fi-ml (3D atmosphere, 2D fixed ocean and sea-ice and land surface) models are tuned to appropriate target data sets by minimising multiple measures of model-data mismatch across different physical fields. Grid computing is exploited to perform the large number of concurrent simulations that comprise the generations of the algorithm. Recent advances in the method use Response Surface Modelling (RSM) to provide surrogate models of the underlying objective functions. These RSMs can be searched much more cheaply and extensively to provide considerable performance improvements in the optimisation.

Introduction

GENIE is a Grid-enabled modelling framework that allows the flexible coupling of constituent models of varying resolution, dimensionality and comprehensiveness to form new integrated ESMs. Grid technology is a key enabler for the flexible coupling of constituent models, subsequent execution of the resulting ESMs and the management of the data that they generate.

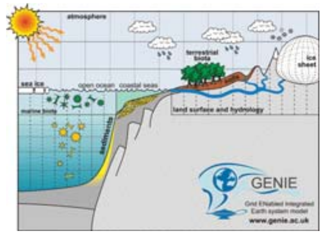

Figure 1: Schematic of the components modelled in GENIE.

In order to simulate at the multi-decadal time scale and beyond, climate models rely heavily on parameterisations of physical processes that occur on comparatively small time and spatial scales. A key concern in climate modelling is therefore to find appropriate values for these parameters so that a reasonable climatology is simulated. This is of particularly importance within the GENIE modelling framework where component codes, that are often developed independently, are coupled together to form new Earth system models. In order to produce stable and sensible model output it is almost always necessary to re-tune the parameters of the coupled system. However, as with many design problems, the nonlinear response of a model to its parameters and the often conflicting tuning objectives make this a difficult problem to solve.

The general problem of optimising a set of model parameters in order to improve a number of possibly conflicting design objectives is typically approached in one of two ways. One can create a single objective measure of design quality by computing a weighted sum of the individual objectives and seek to find the set of variables that minimise or maximise this measure. Many sophisticated algorithms can be applied to a single objective problem but the weighting factors can be critical in the performance of the optimisation. Alternatively, multi-objective methods can be employed to seek a Pareto set of non-dominated solutions; designs that are superior when all objective measures are considered but that may be inferior when a subset of those objectives are considered. Such a solution set can inform the user of competition in the design goals and allows domain expertise to be applied to select the most appropriate parameter sets for further study.

Multi-objective Optimisation

Tuning studies of GENIEfy models have typically involved minimising a single objective function composed of a weighted sum of the RMS errors between model fields s_i and equivalent observational data S_i across physical fields (i) in the model. The weights w_i are often arbitrarily chosen and usually reflect the number of grid cells contributing to the calculation for each model field.

$$f(x) = \sqrt{\frac{1}{N}\sum_{i=1}^{N} w_i \frac{(s_i(x) - S_i)^2}{\hat{\sigma}_i^2}}$$

These weightings necessarily lead to some bias in the objective function. The multi-objective approach is not impacted by a choice of weighting factors because Pareto optimal solutions are sought. The single objective function can be split into its constituent components.

$$f_i(x) = \sqrt{\frac{(s_i(x) - S_i)^2}{\hat{\sigma}_i^2}}, \qquad i = 1, \ldots, N$$

The ultimate aim is to produce a set of solutions which, for at least one criterion, are better than the rest. Such a set is called a Pareto front and is known to be the set of non-dominated solutions in the objective space. Scientists can then discuss each of these solutions, compare between requirements and select the best trade-off. This allows several goals to be considered simultaneously and evenly distributed, aiming to obtain as many solutions as possible, evenly distributed and widely spread around the objective space. Multi-objective algorithms can be classified according to these three criteria, which define the quality of the Pareto front. We exploit the NSGA-II algorithm, a well known algorithm based on a GA routine, with further enhanced GA mutation and crossover techniques to refine the algorithm towards optimal performance. Genetic Algorithms require a population of data points to be evaluated over multiple generations to reach an optimal solution. The expense of these evaluations sets practical limits on the scale of study that can be achieved. Recent developments have therefore sought to reduce the number of calls to the expensive function evaluations by introducing surrogate modelling techniques that are popular in single objective optimisation.

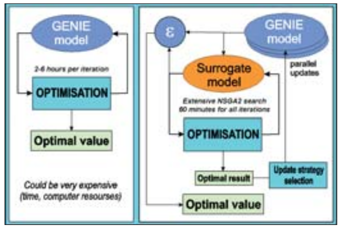

Figure 2: Surrogate models for optimization.

In the single objective world, approaches using surrogate models are fairly well established and have proven to successfully mitigate the problem of computational expense (Figure 2). The key idea is to build a computationally cheaper model; one that can be extensively searched instead of the expensive one. When an optimal solution is found, it needs to be verified on the expensive model and then added to the data pool on which the cheaper model is being trained. The procedure iterates until a good agreement between the two models is achieved or an acceptable solution is found. More and more companies have adopted surrogate assisted optimization techniques and some are making steps to incorporate this approach in their design cycle as a standard. This makes optimization not only useful, but usable and affordable.

Surrogate Modelling

The optimisation is started by performing an initial space filling sampling of the objective function. The Kriging hyper-parameters are tuned to provide the best fit of the surrogate models to the data. A full NSGA-II search is then performed over the Response Surface Models selecting update points to promote diversity in the population using the following strategies

- Select random points to aid escape from local minima
- Take evenly spaced points from the RSM Pareto front
- Perform a small secondary NSGA-II to promote diversity
- Maximum mean squared error on the RSM Pareto front
- Maximum expected improvement on the RSM Pareto front

Once all update points have been selected for evaluation the optimiser submits the compute jobs to the specified compute resource. The results are subsequently added to the existing data pool of direct function evaluations. The best points from the data set in terms of closeness to the last Pareto front and separation in objective space are chosen. This ensures that the cost of building the RSMs is kept constant and provides the points from the global data pool likely to give the best representation of the underlying functions. The data pool of real function evaluations is ranked and the Pareto front extracted. New Krig models are tuned and the whole process is repeated until convergence criteria are satisfied.

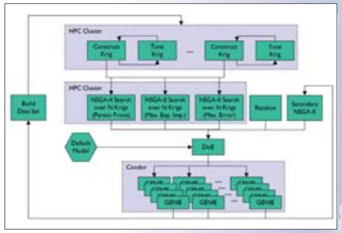

Figure 3: Multi-objective optimisation process.

It should be noted, that the evaluation of the large number of update points is made practical by the utilization of Grid computing. In particular, for studies of GENIEfy models, we exploit the Condor workload management system for high throughput computing. A particular strength of Condor is its ability to pool idle workstations within an institution and harness what would otherwise be wasted CPU cycles. This is an ideal resource since we typically need to perform a large number (50-400) of relatively short simulations concurrently. We target a large Condor pool comprising in excess of 1,100 nodes available at the University of Southampton.

Grid Software

A principle aim of the GENIEfy project is to exploit Grid computing technology to ease the construction, execution and management of Earth System Models. The software deployed to meet these needs has been built upon output from the GEODISE project (http://www.geodise.org). Functions are supplied in the Matlab problem solving environment that can be used to manage ESM simulations on Compute and Data grids. The OPTIONS Design Search and Optimisation package provides the GA algorithm. To configure the study a user simply writes a submission and retrieval function within the Matlab environment. The submission script accepts as input an array of parameter values which it uses to configure and instantiate a model run on a specified Grid resource. The function returns a data structure describing the job which the optimiser can use to monitor the progress of the simulation. The retrieval function processes the job upon completion and returns the multiple objective function evaluations for that point.

Figure 4: User-supplied job submission script.

Within the submission function it is simple to change the Grid resource on which the simulations will be performed by modifying the resource data structure. The user provides details of the local runtime environment and then scripts the assignment of the tuneable parameters to the appropriate fields of the configuration data structure. A single function call *gc_jobsubmit* is invoked to submit the compute job to the specified resource. The optimiser monitors the returned job handle and invokes the post-processing script upon completion which returns the multiple objective measures of fitness to data.

genie-eb-go-gs (C-GOLDSTEIN)

The genie-eb-go-gs model (3D frictional geostrophic ocean model, 2D energy moisture balance model and 2D sea-ice) has twelve principal transport parameters whose values are poorly constrained but which strongly influence the model climatology. Four physical data sets consisting of 3D ocean temperature (OCNTEMP) and salinity fields (OCNSAL) and 2D atmospheric surface air temperature (SURFTEMP) and humidity (SURFHUM) fields are employed as targets for the tuning. These data sets are an average over approximately five decades (spanning the period 1948-2002) and obtained from Levitus and the National Centers for Environmental Prediction (NCEP). Each simulation is run for 2000 model years to ensure equilibrium before the objective functions are evaluated. These calculations typically require about 40 minutes of CPU time on an average desktop.

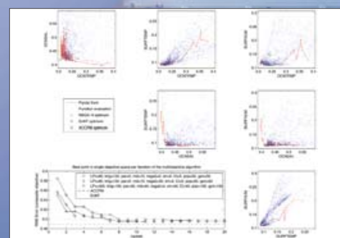

Figure 5: Function evaluations projected onto 2D objective space for each pair of objectives. Bottom left: Progress of the NSGA-II algorithm.

Comparative studies have been performed with and without the use of surrogate modelling. High quality solutions on the Pareto Front are obtained when using response surface modelling and the number of model years simulated can be reduced by an order of magnitude (Table 1). The multi-objective method is competitive with the EnKF and Proximal-ACCPM algorithms which have been applied to the equivalent single objective optimisation problem (Figure 5)

	N_{POP}	N_{GEN}	LP_T	Updates	Model Years	Total Model Years
NSGA-II + RSM	50	50	60	30 x 20	2000	1,320,000
NSGA-II + RSM	50	50	60	60 x 20	2000	2,520,000
NSGA-II + RSM	100	150	120	200 x 15	2000	6,400,000
NSGA-II	50	100	-	-	2000	10,000,000

Table 1: Comparison of years simulated with different methods.

genie-ig-fi-fi-ml (IGCM)

The genie-ig-fi-fi-ml (3D atmosphere, 2D fixed ocean and sea-ice and land surface) has forty-one parameters that are ill-constrained by observations or theory. Two objective functions are defined to improve surface energy fluxes and surface hydrology in the model. Each simulation is performed over 11 model years, taking an average of the last 10 years of output, with each evaluation requiring ~2 hours of CPU time on an average desktop. The target data is taken from output of the HadCM3 model to build traceability with the 3D IGCM atmosphere.

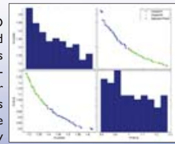

Figure 6: Pareto front from the IGCM study.

There is clear competition between the energy fluxes and the hydrological properties of the model. Improvements in one objective can only be achieved at the expense of the other objective. Applying an a posteriori weighting we select a solution for further analysis (Figure 7).

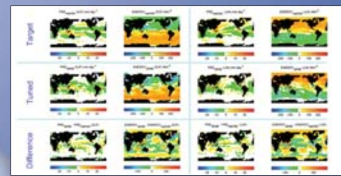

Figure 7: Tuning results for 3D atmosphere component.

Conclusions

Multi-objective optimisation methods avoid the need for formulating a single composite objective function when studying an optimisation problem with competitive design goals. We have applied the multi-objective NSGA-II method using surrogate models to two models from the GENIE Earth system modelling framework. Twelve parameters of the GENIE-1 model (C-GOLDSTEIN) and forty-one parameters of the IGCM atmosphere model have been tuned to improve their fit to observational climate data. Grid computing has been exploited to perform concurrent runs of the underlying function to provide updates to the data set from which the meta-models are generated. The optimisation tools are integrated within a familiar working environment and tuning exercises for GENIE models are quickly and easily configured.

Acknowledgements

The GENIE and GENIEfy projects are funded by the Natural Environment Research Council. Development of the OptionsNSGA2 software has been sponsored by Rolls Royce under the VIVACE project.

References

Price, A. R., Voutchkov, I. I., Pound, G. E., Edwards, N. R., Lenton, T. M., Cox, S. J. and the GENIE team, Multiobjective tuning of Grid-enabled Earth System Models using a Non-dominated Sorting Genetic Algorithm (NSGA-II). Proceedings of the 2nd International Conference on eScience and Grid Computing 2006.

School of Engineering Sciences, University of Southampton, Southampton SO17 1BJ
Tel: +44 (0)2380 598362; Fax: +44 (0)2380 594813; Email: andrew@soton.ac.uk www.e-science.soton.ac.uk