Request for Computing Resources in Support of NSF Grant 0851065
Title A Feasibility Study for Understanding Climate Uncertainty with an Ocean Focus
PI: Robin Tokmakian, Dept. of Oceanography, Naval Postgraduate School, Monterey, CA 93950[redacted]
NSF Program Officer: [redacted]

1. Overview

Uncertainty in climate simulations arises from various aspects of the end-to-end process of modeling the Earth's climate. Figure 1 illustrates the parts of the climate model system that contribute to the total uncertainty. First, there is uncertainty from the structure of the climate model components (e.g. ocean/ice/atmosphere). Even the most complex models are deficient, not only in the complexity of the processes they represent, but in which processes are included in a particular model. These structural uncertainties occur in how a climate model is designed, for example, which processes are included. These structural and numerical (e.g. resolution) uncertainties are the important reasons for differences between models, for example in runs for the Intergovernmental Panel on Climate Change (IPCC) assessments. Next, additional uncertainties arise from the inherent error in the initial and boundary conditions of a simulation. Initial conditions are the state of the weather or climate at the beginning of the simulation, and typically come from observations. Boundary conditions are conditions that are not computed by the model. Finally, there is the uncertainty associated with the values of parameters in the model. These parameters may represent physical constants or effects, such as ocean mixing, or non-physical aspects of modeling and computation. The uncertainty in these input parameters propagates through the non-linear model to give uncertainty in the outputs. This last contribution to the total uncertainty is the focus of this research.

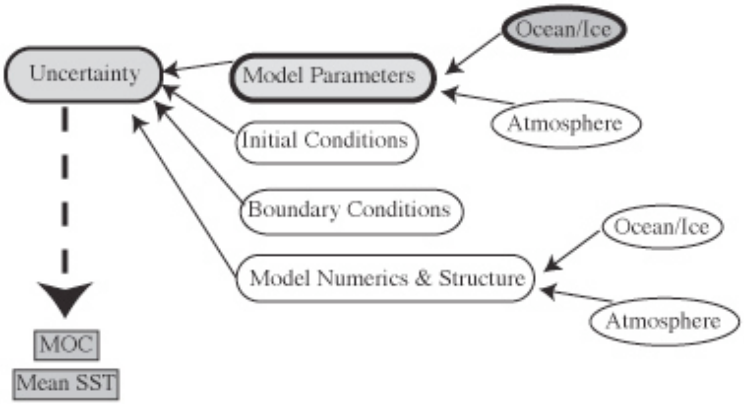

Figure 1. The various contributions to the total uncertainty in estimates of, for example, Meridional Overturning Circulation (MOC) or global mean temperature (mean SST). The shaded grey regions are the areas of interest to this effort.

The models in 2020 will no doubt be better than today's models, but they will still be imperfect, and development of uncertainty analysis technology is a critical aspect of understanding model realism and prediction capability. *Smith* [2002] and *Cox and Stephenson* [2007] discuss the need for methods to quantify the uncertainties within complicated systems so that limitations or weaknesses of the climate model can be understood. In making climate predictions, we need to have available both the most reliable model or simulation and a methods to quantify the reliability of a simulation.

If quantitative uncertainty questions are to be answered with complex simulations such as AOGCMs, then the only known path forward is based on model ensembles that characterize behavior with alternative parameter settings [*Rougier*, 2007; *O'Hagan* 2006; *Christie et al*. 2005]. Such analyses are being applied to a number of 'simulations of phenomena' problems, for example in cosmology [*Heitmann et al*. 2006]. The need for uncertainty analysis is multiplied as models become more complex to incorporate additional phenomenology, for example, in ocean biogeochemistry processes of IPCC class of climate models expected in the 5th Assessment Report. The development of these complex models into climate models partly involves how to use and understand a large set of parameters that represent physical systems in a model that cannot capture all the temporal and spatial scales observed in nature. This work represents the first steps in what will be an increasingly complex, and yet crucial, undertaking. The relevance and feasibility of using "Statistical Analysis of Computer Code Output" (SACCO) methods for examining uncertainty in ocean circulation due to parameter specification will be demonstrated.

To address the feasibility of using such methods to examine uncertainty in AOGCMs, we are proposing an initial study using just the ocean/ice component models of an AOGCM to begin to quantify the uncertainty and to understand how to apply these methods to complex geophysical problems. This will lead to future projects using a fully coupled AOGCM to understand the importance of the oceans to the Earth's climate. A dataset will be generated that can be used also by external researchers in developing methods for understanding large complex model uncertainty. This will be one of the first times this state of the art statistical methodology will be applied to a component of a climate AOGCM.

Science Objectives

This project's objective is to evaluate the feasibility of using advanced statistical methods to understand uncertainty in ocean metrics in the ocean component of an AOGCM (e.g. the POP/CICE components [*Smith and Gent*, 2002] of the Community Climate System Model, CCSM3.0 or later version) arising from our ignorance of the input parameters. This increase in knowledge about parameters can lead to suggestions for improvements in the model and in data collection.

We will:

1. Run a designed experiment, using the POP/CICE parameter space ("perturbed physics", *Stainforth et al.* 2005), of 100 years. We estimate that we will need an ensemble of approximately 100 members (and less if possible) at a resolution (~3°) defined by available computer time to create a representative set of climate simulations. The parameters to be examined will include those related to three physical processes: the diffusivity of tracers along isopycnal surfaces, how wind mixes with depth, and the vertical diffusivity of tracers (parameterizations incorporated in the K-profile vertical mixing [*Large et al,* 1994] and Gent-McWilliams isopycnal mixing schemes [*Gent and McWilliams*, 1990]). This task is the focus of the request for NCAR computing resources.
2. Develop a statistical analysis of uncertainty based on emulator techniques to model the simulation response and model inadequacy to explore uncertainties and sensitivities in the ocean component over the defined parameter space using the climate simulations' output and the statistical model.
3. Explore the model uncertainties of ocean/ice simulations (using the same computer code) at a higher resolution using inferences from the lower resolution model

and using similar emulation techniques to task 2 that allow for the estimation of the discrepancy in model response between models of different complexity.

4. Technical Aspects

4.1. Tasks necessary to advance

The work being proposed has two goals. The first is to perform an uncertainty/ sensitivity analysis on a relatively low-resolution current generation ocean/ice model consistent with components of an AOGCM. This will demonstrate the feasibility of the methodology for use with a fully coupled AOGCM. The experiment will involve running an ensemble of approximately 100 model simulations in a designed experiment. These ensembles can then be used to test a wide set of questions relating to the model's response and would be available to the rest of the scientific community to address other questions. From this ensemble, a statistical emulator is created (Task 2) enabling the further effort in the analysis of a list of ocean metrics (Task 1). Included in this task is the provision of a set of tools that can be easily used by the community for related uncertainty efforts. The second goal is to gain understanding in how to relate and explore uncertainties in a complex system of the same family but at higher resolution (Task 3). This will involve running smaller ensembles at higher resolution and propagating information about uncertainty up the chain of models.

The successful application of the statistical methods in this project will have a substantial impact on the requirements for future computational resources, demonstrating that an emphasis on computing parametric ensembles is necessary.

4.1.1 Task 1: Run a designed experiment, using the POP/CICE parameter space, of 100 years for a full ensemble of 100 at a resolution 3°.

Traditional use of complex simulation studies, after some level of validation, is to run a single simulation at the highest possible capability. Although this is intuitively compelling, it does not give insight into the range of model results. Running an ensemble of alternatives does require the reduction of simulation fidelity from the best available. Not to diminish the role of examining the best available model runs for purposes of research and as a path to evaluating model improvements, but for the purpose of uncertainty analysis in system behavior and in answering policy questions based on these models, there is no alternative [*Smith* 2002]. We are proposing to use the ocean and ice components of the CCSM3.0 or later version [*Collins et al.* 2005] because of our familiarity with the ocean component; it has been developed as a community model; and it is a current state-of-the-art model.

Ocean model specification: The ocean component of the model is based on the early efforts of *Bryan and Cox* [1967] and *Bryan* [1969a, b]. CCSM3 ocean component model uses the Parallel Ocean Model v. 1.4.3, a z-level model [*Smith and Gent*, 2002]. The prognostic variables are velocity, potential temperature, salinity, sea surface height, ideal age (time since ocean parcel was at the surface), and any number of specific tracers included for a unique simulation. The resolution of the ocean component is either about 3° or 1° degree and somewhat smaller at the equator. The variability that occurs on scales smaller than the models' resolution (eddy scales and smaller) is parameterized using the *Gent and McWilliams* [1990] scheme. The dynamics of the mixed layer for the model use the K-profile parameterization (KPP) [*Large et al,* 1994]. *Ice model specification:* The dynamics of the ice uses complex elastic-viscous-plastic (EVP) rheology of *Hunke and Dukowicz* [1997]. The EVP method explicitly solves for the ice stress tensor. The thermodynamics portion use variations of the *Bitz and Lipscomb* [1999] thermodynamics [*Briegleb et al.* 2002]. This code accounts for more of the physical processes within the ice, including the melting of internal brine regions and conserves energy. *Atmosphere specification:* We will force the ocean/ice system with products from derived from the ECMWF ERA40 re-analysis product or a similar product. An annual climatology will be created to drive to ocean/ice system in a consistent manner. Without being coupled to an atmosphere model, this feasibility study will not include aspects of feedbacks between the ocean and atmosphere. However, understanding the uncertainty in just the ocean and how it takes up heat is an important first step. *Initialization:* The model runs would be initialized from a previous run, spun-up simulation. We anticipate some model drift initially, but not as great as from starting from climatology (only T/S initialized). Test runs will be done first to examine model drift using a few different parameter values prior to starting the production of the simulations and flux corrections may be necessary to be incorporated to prevent excessive drift.

Ocean Parameter	**Nominal**	**Min**	**Max**	**Additional References ***
K-profile background vert. diffusivity cm^2/s	0.524	0.05	5.0	*e.g. Li et al.,*[2001],Table D2, *Brierley* [2007].
K-profile background vert. diffusivity variation cm^2/s	0.3	0.	0.5	
K-profile depth: background vertical diffusivity: cm	1.0e5	0.5e5	2.5e5	
G-M diffusion coef. bolus cm^2/s	0.8e7	0.6e7	1e7	*Danabasoglu and Marshall* [2007]; *Collins et al* [2007[]
G-M max slope bolus	0.3	0.01	0.3	*Danabasoglu and Marshall* [2007

Table 1: Initial list of parameters and ranges (*Nominal values: *CCSM3/POP User's Guide* 2004

Experimental design of the ensemble: Careful consideration of the design of the parameter space needs to occur prior to the creation of the ensemble members. *Collins et al.* [2007], in a perturbed physics experiment in which they varied one parameter at a time, looked at the changes in parameter values for three physical processes: diffusivity of tracers along isopycnal surfaces, how wind mixes with depth, and the vertical diffusivity of tracers. We will look at the same processes in the POP/CICE system that incorporates different mixing schemes structure (K-profile vertical mixing and Gent-McWilliams isopycnal mixing scheme) including more parameters to vary. The parameters used will include those listed in Table 1 that show the ranges of possible values. The work of *Collins et al.* [2007] had only a small number of ensemble members and no runs in which non-linear interactions were examined by co-varying parameters. We will model the non-linear interactions between these parameters and will include other parameters as described (and evaluated by experts as to their importance) in *Brierley* [2007]. Table 1 lists the most important of these parameters in terms of influence on the processes of interest.

Collaborators: [redacted] [redacted] [redacted] [redacted] [redacted] [redacted] [redacted] [redacted]
[redacted]

[redacted]

Computer resources:

Resolution	AOGCM Model years/day= Y	Number of ensemble members times Number of model years = N	No. of Processors	Processor-hrs/Myr	Total GAUs
~3° 40 levels	Y=133 Myrs/day (T62_gx3v5 coupled ocean/ice only)	100 ensemble members*100 yrs = 10,000	22 [1 node or 32] (8 pop2 4 cice 8 cpl)	1.97 for a 5 day test ~0.18pes/yr*1.4*32 = 8.06 GAUs/Myr	**81,000** (from a 10yr test run: 1.8hrs/10yr*32 PE*1.4systemfact= 80.6 GAUs for 10yrs)
~1° 40 levels	Y=27.23 Myrs/day (T62_gx1v3 coupled ocean/ice only)	10 ensemble members*100 yrs = 1,000	116 [4 nodes or 128] 96 pop2 16 cice 4 cpl	46.08 for a 5 day test 0.219hr/yr*128PE*1.4 = 29.4 GAUs/Myr	**29,400** (from a 3 month test run: 0.219hr/yr*128 PE*1.4 = 29.4 GAUs/Myr)

Table 2: NCAR-CCSM3.1 - bluefire –test and performance runs using CCSM3.1

1. The computer resources required to run 100 simulations limits us to the use of the 3° resolution configuration initially. Table 2 lists the timings of the model as run with a test code setup on bluefire using a CCSM3.1 coupled ocean/cice simulation. The code is a primitive equation code and has been extensively tested by the Community Climate System Model Group. The g3v5 code has 100x116 and 40 levels (464,000 grid points). The g1v3 code is 320x384 and 40 levels (4,915,200 grid points). The total number of GAUs requested is **81,000 + 30,000 = 111,000 GAUs**. It is anticipated that the first priority will be the 3° degree simulations and at minimum this will require 81,000 GAUs initially. The amount of computer time necessary to complete the research will be re-examined periodically and will be minimized when possible. Most likely, the PI will use CCSM4 for the ensemble runs when it is available for release (2009).

2. A full Monte Carlo ensemble set would require 1000 or so members. The use of an emulator reduces the number to around 100. If, in the process of the experiment, fewer members are required, the number will be reduced. The number is required by the nature of the uncertainty problem. High performance computing resources at NPS may also be available for a small number of the ensemble members.

3. There are no special requirements

4. Storage Space for project : g3x5: 3Gbyte/Myr; 300Gbytes/100Myr *100 ensembles = **30TB;** for the g1x3: 5Gbyte/Myr; 500Gbytes/100Myr* 10 ensembles = **10TB.** Monthly fields will be stored of consisting of 10 prognostic variables for the ice and ocean fields. Some limited diagnostics will also be stored.

5. Optimization of code. The CCSM group has optimized the code.

6. Code Performance: As given by Table 2, the performance is at 90-93% of the requested resources (timing tests provided by CCSM group used to determine processor allocation for optimum performance).

References

Bitz, C.M. and Lipscomb, 1999, W. H., An Energy-Conserving Thermodynamic Model of Sea Ice *J.of Geophys. Re.,*104, 15,669-16,677.

Brierley, C., 2007, Ocean model uncertainty and time-dependent climate projections, PhD thesis, University of Reading.

Briegleb, B.P., E.C. Hunke, C.M. Bitz, W.H. Lipscomb, and J.L. Schramm, 2002, *Description of the Community Climate System Model Version 2 Sea Ice Model*, 60 pp, (available at http://www.ccsm.ucar.edu/models/ccsm2.0/csim).

Bryan, K. and Cox, M. D., 1967, A numerical investigation of the oceanic general circulation, Tellus, XIX, 54–80.

Bryan, K., 1969a, A numerical method for the study of the circulation of the world ocean, J*. of Computational Physics*, 4, 347–376.

Bryan, K., 1969b, Climate and the ocean circulation III: The ocean model, *Monthly Weather Review*, 97, 806–824.

Christie, M.A., J. Glimm, J. W. Grove, D. M. Higdon, D. H. Sharp, and M. M. Wood-Schultz, 2005, Error Analysis and Simulations of Complex Phenomena, *Los Alamos Science* 29, 6 http://www.lanl.gov/orgs/t/t13/research/lascience/ds05.pdf.

Collins WD, CM Bitz, ML Blackmon, GB Bonan, CS Bretherton, JA Carton, P Chang, SC Doney, JJ Hack, TB Henderson, JT Kiehl, WG Large, DS McKenna, BD Santer, RD Smith, 2005, The Community Climate System Model: CCSM3. *J Climate*, 18.

Collins, M., C. M Brierley, M. MacVean, B.B.B. Booth, and G. R. Harris, 2007, The sensitivity of the rate of transient climate change to ocean physics perturbations, *J. of Climate,* 20, 2315–2320, DOI: 10.1175/JCLI4116.1.

Cox, P. and D. Stephenson, Climate Change: A Changing Climate for Prediction, 2007, *Science* 317 (5835), 207, DOI: 10.1126/science.1145956.

Danabasoglu and Marshall. "Effects of Vertical Variations of Thickness Diffusivity in an Ocean General Circulation Model, 2007, *Ocean Modelling*, 18, 122-141.

Gent, P. and J.C. McWilliams, 1990, Isopycnal mixing in ocean circulation models, *J. Phys. Oceanogr.*, 20, 150-155.

Heitmann, K., D. Higdon, C. Nakhleh and S. Habib, 2006, Cosmic Calibration, *The Astrophysical Journal*, 646, L1-L4.

Hunke, E. C., and J. K. Dukowicz, 1997, An elastic-viscous-plastic model for sea ice dynamics, *J. of Phys. Oceanography*, 27, 1849-1867.

Large, W., J.C. McWilliams, and S.C. Doney, 1994, Oceanic vertical mixing: A review and a model with a nonlocal boundary mixing parameterization, *Rev. Geophys.,* 32, 363- 403.

Li, X., Y. Chao, J. McWilliams, and L. Fu, A comparison of two vertical-mixing schemes in a pacific ocean general circulation model, 2001, Journal of Climate, 14, 1377-1398.

O'Hagan, A., 2006, Bayesian analysis of computer code outputs: a tutorial. *Reliability Engineering and System Safety,* **91**, 1290–1300.

Rougier, J. C., 2007: Probabilistic Inference for Future Climate Using an Ensemble of Climate Model Evaluations, *Climatic Change*, 81, 247-264.

Smith L., 2002, What might we learn from climate forecasts? *Proc. Nat'l Academy of Sciences*, Vol. 99, suppl. 1, 2487-2492 doi:10.1073/pnas.012580599.

Smith, R. D. and P. R. Gent, 2002, Reference manual for the Parallel Ocean Program (POP), ocean component of the Community Climate System Model (CCSM2.0 and 3.0). Technical Report LA-UR-02-2484, Los Alamos National Laboratory, available online at http://www.ccsm.ucar.edu/models/ccsm3.0/pop.

Stainforth, D. A., T. Aina, C. Christensen, M. Collins, N. Faull, D. J. Frame, J. A. Kettleborough, S. Knight, A. Martin, J. M. Murphy, C. Piani, D. Sexton, L. A. Smith, R. A. Spicer, A. J. Thorpe, M. R. Allen, 2005, Uncertainty in predictions of the climate response to rising levels of greenhouse gases, *Nature*, 443, 403-407.