Uncertainty Quantification and Ocean Model Analysis

James R. Gattiker,
Sham Bhat, CCS-6;
Matthew W. Hecht, CCS-2

Uncertainty quantification (UQ) methods are a tool for analysis of the performance of simulation models. This research summary describes how the application of UQ tools to the POP2 ocean model can be used to quantify model improvements, to understand where next-generation models can be improved, and to determine the settings of parameters and uncertainty in model applications.

Uncertainty quanitification (UQ) is concerned with learning about model performance from the comparison between model output and accepted performance targets. This research uses LANL's Bayesian analysis tool GPM/SA (Gaussian Process Models for Simulation Analysis) [1] to study the Community Earth System Model's ocean component, the LANL-developed POP2 [2]. Two products of the analysis are discussed: (1) a method for quantitatively assessing model performance for qualification of model improvements, and (2) an analysis of distributions on the parameters of key model settings in the face of conflicting measures of model performance.

Complex physical models often use parameterizations to overcome limitations. The Gent-McWilliams isopycnal transport and mixing scheme (GM) [3] corrects for the transport by ocean eddies that cannot be resolved in the current generation of climate simulations, run at grid resolutions comparable to our 0.8-degree grid simulations (Fig. 1b). The GM parameterization is arguably the most significant parameterization in ocean models used for long climate simulations, greatly improving the representation of ocean heat transports [4]. The current GM formulation relies on two uncertain parameter settings, an overall magnitude and a stratification-keyed weakening coefficient called tapering. The specific application goal in this project is to apply UQ analysis to the uncertain parameters of GM, determining its quality and best settings.

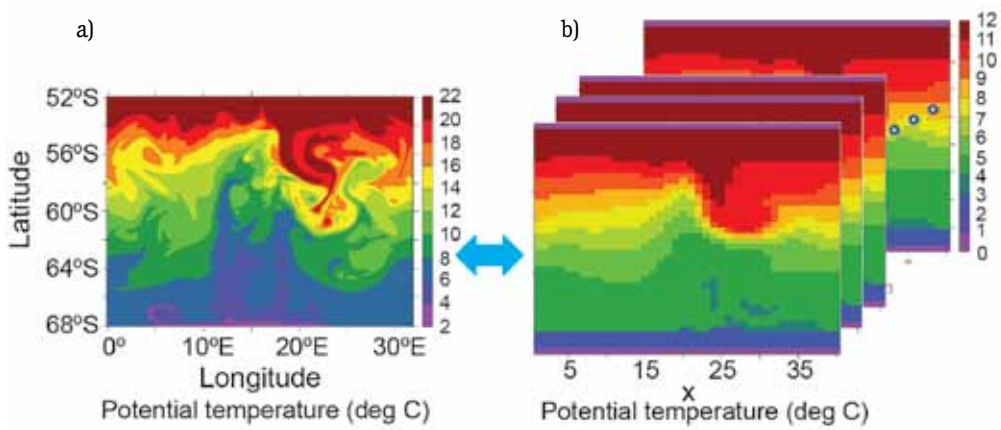

Fig. 1. Uncertainty quantification compares a) a reference high-resolution solution to b) many examples of low-resolution simulations at various settings of uncertain parameters.

Ocean model analysis challenges include: (1) sparse Earth system observations, especially for the deep ocean, making direct model-observation comparison problematic; and (2) the ocean state contains long time scales, so the amount of computation required to reach a statistical steady state is daunting. To address these challenges, a test-bed ocean configuration is adopted whose scope and computation are considerably reduced from the full Earth system. The "channel model" is designed to study eddy-induced transport, mimicking the Antarctic Circumpolar Current of the Southern Ocean. The reduced scope makes the computation of high-resolution simulations explicitly resolving eddies and their mixing effects feasible. A high-resolution simulation can then be taken as "truth" for the purposes of studying the GM parameters. Figure 1 shows the channel model's sea surface temperature in the high-resolution result, as it is to be compared to an ensemble of low-resolution model runs. The ensemble is designed to span the plausible ranges of the GM parameters. This ensemble allows the construction of a fast statistical emulator of the model's response that can be used to support inferences in the comparison, including parameter distributions and measures for how well the model fits the dataset.

A key element in comparing the simulations to their targets is the specific set of metrics to be used for the comparison. The metrics define the question we are asking of the models and the analysis. Figure 2 shows four metrics used for analysis, from a larger set of measures suggested by domain experts. Each member of this set is an extracted property versus depth, averaged horizontally over the simulation domain, and over a time window of five years of simulation, after equilibration. Temperature and salinity profiles show direct impacts of transport and mixing, the density in turn is dependent on these, and the vertical heat transport summarizes the movement of heat. Because we are

For more information contact James R. Gattiker at gatt@lanl.gov.

Funding Acknowledgments
DOE, Office of Science, BER, Climate Science for a Sustainable Energy Future Project

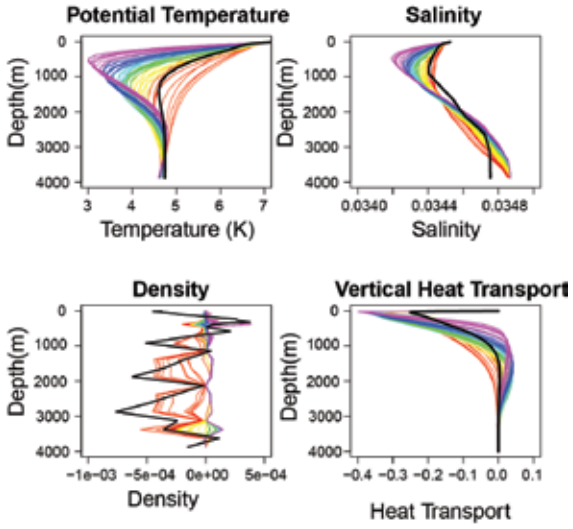

Fig. 2. Four metrics for comparison: Potential Temperature, Salinity, Density (residual from simulations' mean), and Vertical Heat Transport. Black is high-resolution, color is a composite related to the sum of parameters for presentation.

attempting to approximate with the GM scheme the effects of computationally complex fluid dynamics, it is not expected that we can meet the resolved simulation results in every detail. Qualitatively, these metrics appear in some cases to be very informative, and in other cases show difficulty.

UQ analysis accounts for uncertainty in the emulated model response, uncertainty in the parameters controlling the match between the low-resolution model and its target, uncertainty in the general quality of fit between these models, and uncertainty regarding irreducible structural differences. These aspects cannot be considered in isolation, they are all part of the same problem of comparing simulations to a reference target, in this case the low-resolution parameterized model to the high-resolution target. The various interdependent aspects of uncertainty must be considered even when the focus is on one particular aspect. The UQ approach uses Gaussian process emulator models, an additive Gaussian process for structural discrepancy, and parameters for overall model fit expressed as a target precision. The result of analysis is a joint probability distribution of parameters, both those in the statistical model and those controlling the behavior of GM.

In our first goal to demonstrate the capability to qualify model improvements, we are interested in the summary measure of the ability of the model with GM to fit the target. We compare this measure for an older one-parameter GM implementation, the two parameters of the current GM, and the value of considering a third parameter related to the shape of the GM tapering. The improved model (one parameter to two parameters) is clearly quantifiable, while exposing the addition of a third parameter does not represent an advance in model fit. In this case the additional parameter isn't independently contributing to quality fit with respect to the metrics.

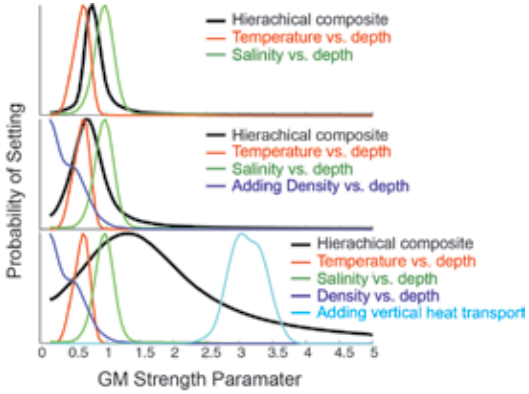

Fig. 3. Summary of parameter distributions with respect to different metric sets.

The complete target of a UQ analysis is to understand what parameter settings are implied by the data, or perform "inverse" analysis–again, taking into account the various sources of uncertainty. Each metric results in a distribution on parameters, shown in the colored curves in the axes of Fig. 3. A novel aspect of this analysis is the methodology used to combine these various results. The composite distribution should attempt to unify the results where they are feasibly similar, but when there is unresolvable conflict should admit appropriate uncertainty. An appropriate model for this is a Bayesian hierarchical model across relevant metrics with a prior preferring consistency. Figure 3 shows how unification proceeds while adding indicators. The top axis shows that the implications of the temperature and salinity metrics are consistent. The middle axis shows that although the results of density are not independently identical, that information can be unified with the previous result. However, vertical heat transport represents information in conflict, hence its inclusion results in a large increase in uncertainty. In the problem context, we know that GM cannot improve all possible measures. The result of the UQ analysis communicates to the domain expert where these various metrics can be improved by GM consistently and where there is substantial discrepancy, providing valuable input to the next cycle of model improvements. Today, this information assists in understanding how to set the values of the current generation of applied models to achieve the best performance.

[1] Higdon, D. et al., *J Am Stat Assoc* **103**(482), 570-583 (2008).

[2] Smith, R. et al., "The Parallel Ocean Program (POP) Reference Manual," LANL Technical Report LA-UR-10-01853 (2010).

[3] Gent, P.R. and J.C. McWilliams, *J Phys Oceanogr* **20**, 150 (1990).

[4] Gent, P.R. et al., *J Clim* **11**, 1287 (1998).