Quantifying Prediction Fidelity in Ocean CirculationModels

Mohamed Iskandarani
Rosenstiel School of Marine and Atmoshperic Science
University of Miami
4600 Rickenbacker Causeway, Miami, FL 33149
phone: (305) 421-4045 fax: (305) 421-4696 email:miskandarani@rsmas.miami.edu

Omar Knio
Department of Mechanical Engineering and Materials Science
Duke University
144 Hudson Hall, Box 90300, Durham, NC 27708
phone: (919) 660 5344 fax: (919) 660-8963 email:omar.knio@duke.edu

Ashwanth Srinivasan
Rosenstiel School of Marine and Atmoshperic Science
University of Miami
4600 Rickenbacker Causeway, Miami, FL 33149
phone: (305) 421-4763 fax: (305) 421-4696 email:asrinivasan@rsmas.miami.edu

William C. Thacker
Cooperative Institute for Marine and Atmospheric Studies University of Miami
4600 Rickenbacker Causeway, Miami, FL 33149
phone: (305) 361-4323 fax: (305) 421-4696 email:carlisle.thacker@noaa.gov

Award Number: N00014-101-0498
http://www.onr.navy.mil/sci_tech/32/reports/annual/

LONG-TERM GOALS

The development, validation and application of robust uncertainty quantification methods to ocean modeling, forecasting, and parameter estimation.

OBJECTIVES

This project explores the use of Polynomial Chaos (PC) expansions for improving our understanding of uncertainties in Ocean General Circulation Models (OGCM). Given adequate initial and boundary conditions, most OGCMs can be used to forecast the evolution of the oceanic state consistent with known physical laws. Reliable ocean forecasts, however, require an objective, practical and accurate methodology to assess the inherent uncertainties associated with the model *and* data used to produce these forecasts. OGCMs uncertainties stem from several sources that include: physical approximation of the equations of oceanic motion; discretization and modeling errors; an incomplete set of sparse (and often noisy) observations to constrain the initial and boundary conditions of the model; and uncertainties in surface momentum and buoyancy fluxes.

Our objective is the development of an uncertainty quantification methodology that is efficient in representing the solution's dependence on the stochastic data, that is robust even when the solution depends discontinuously on the stochastic inputs, that can handle non-linear processes, that propagates the full probability density functions without apriori assumption of Gaussianity, and that can be applied adaptively to probe regions of steep variations and/or bifurcation in a high-dimensional parametric space. In addition we are interested in developing utilities for decision support analysis; specifically, we plan to demonstrate how PC representations can be used effectively to determine the non-linear sensitivity of the solution to particular components of the random data, identify dominant contributors to solution uncertainty, as well as guide and prioritize the gathering of additional data through experiments or field observations.

APPROACH

Our approach to uncertainty quantification (UQ) relies on PC expansions (Le Maître and Knio (2010)) to investigate uncertainties in simulating the oceanic circulation. We have opted to use the HYbrid Coordinate Ocean Model (HYCOM) as our simulation engine because it has been developed as the next generation model for the US Navy and has been adopted by NOAAs National Center for Environmental Prediction. HYCOM is equipped with a suite of sequential assimilation schemes that will be used to investigate how UQ may be beneficial to data assimilation. Details about the model, its validation, and sample applications can be found in Bleck (2002); Halliwell (2004); Chassignet et al. (2003, 2006) as well as by visiting http://www.hycom.org. Below we present the main ideas of the PC expansion before we summarize our efforts since the last annual report of Sep 2011.

PC expansions express the dependency of the solution on the uncertain parameter as a series of the form: $u(\boldsymbol{x},t,\boldsymbol{\xi}) = \sum_{k=0}^{P} \hat{u}_k(\boldsymbol{x},t)\Psi_k(\boldsymbol{\xi})$, where $u(\boldsymbol{x},t,\boldsymbol{\xi})$ is a model solution that depends on space $\boldsymbol{x}$, time t and the uncertain parameters $\boldsymbol{\xi}$; $\Psi_k(\boldsymbol{\xi})$ is a suitably chosen orthogonal basis; and $\hat{u}_k(\boldsymbol{x},t)$ are the expansion coefficients. Here, u can represent a variable expressed directly in the model such as sea surface temperature or velocity at a specified point, or a derived quantity such as the mean surface cooling under a hurricane track. In the jargon of UQ u is referred to as a Quantity of Interest or an observable.

The choice of basis function is dictated primarily by the probability density function of the uncertain input data, $p(\boldsymbol{\xi})$, which enters all aspects of the UQ computations. These can be done much more efficiently if the basis vectors are orthonormal with respect to $p(\boldsymbol{\xi})$. Hence the basis functions are Legendre, Hermite, or Laguerre polynomials when the input uncertainty is described by uniform, Gaussian, or Gamma distributions, respectively.

The computation of the stochastic modes is best achieved by the so-called Non-Intrusive Spectral Projection (NISP) method since we would like to avoid modifying the original OGCM code. Taking advantage of the orthonormality of the basis, NISP works by projecting the solution u on the basis function Ψ_k via inner products, and by replacing the integrals with quadrature formula. The coefficients can then be computed simply by running the model at specified values of the uncertain parameters, storing the desired observable, and post-processing via a simple matrix-vector multiplication. No modification to the OGCM need be performed.

The investigative team at Duke University consisted of Dr. Omar M. Knio, and his post-doctoral associates, Drs. Alen Alexanderian and Ihab Sraj, and graduate student, Justin Winokur; they have concentrated on advancing the technical and theoretical aspects of the Uncertainty Quantification efforts. Drs. Mohamed Iskandarani (lead PI), Ashwanth Srinivasan and William C. Thacker (University of Miami) have focused on formulating the oceanographic uncertainty problems, the modification to the HYCOM code and its actual execution, and the preparation of the necessary data to carry out the research agenda. Dr. Matthieu Le Henaff assisted us with the Gulf of Mexico configuration, and we have held discussions with Dr. François Counillon as to the applicability of PCs in Ensemble Kalman Filters based data assimilation. In FY12, we collaborated with Dr. Shuyi Chen on the inverse modelling problem discussed below; her research group produced the high-resolution space time atmospheric fields needed to force HYCOM during typhoon Fanapi. In FY12, we also collaborated with Dr. Youssef Marzouk's team at MIT on the development and implementation of sparse, adaptive, pseudo-spectral quadratures.

WORK COMPLETED

The main thrust of our work consisted of two primary lines of investigations: a science application focused on identifying key uncertain drag parameters using ITOP observations, and a major technical improvement to Polynomial Chaos methods revolving around adaptive sampling.

The science application consisted of an inverse modeling problem that capitalized on the observational data obtained during typhoon Fanapi, and which coincided with the extensive field campaign "Impact of Typhoon on Ocean over the Pacific" (ITOP), to constrain key unknown wind drag parameters. It is well known that the wind drag coefficient does not keep increasing monotonically with wind speed, and that it saturates (figure 1); the saturation value $C_D^{\max}$, however, and the wind speed at which it occurs, $V_{\max}$, are not well-known and are difficult to measure in the field. Moreover, some observations suggest that, for wind speeds greater than $V_{\max}$, the drag coefficient decreases linearly with slope m. Our goal was to use the ITOP data in order to constrain the aforementioned unknown values. The inverse problem was cast in a Bayesian Inference framework, so that the outcomes of the analysis are not only optimal values for the unknown parameters, but posterior probability density functions that measure our confidence in the inferred values. The Bayesian Inference requires the use of Monte Carlo sampling methods and these are prohibitively expensive since they require tens of thousands of HYCOM runs. Our solution was to build a PC-based surrogate (Marzouk et al. (2007)), using an ensemble with 67 members only, for the HYCOM temperature and use it in lieu of the model; the surrogate is the key step that makes these computations practical and feasible. A number of error metrics were monitored, figure 2, in order to ensure that the surrogate is indeed a faithful representation of HYCOM. The atmospheric forcing fields were obtained from a hindcast, triply-nested WRF simulation to replicate the atmospheric conditions during Fanapi at high space-time resolution; partnering with Dr. Shuyi Chen's group was instrumental for this phase (and which turned out to be more time-consumming than anticipated). The HYCOM realizations were initialized from a data-assimilated 1/12° global HYCOM simulation in order to position oceanic features and thermal fronts correctly. The Bayesian Inference analysis revealed a $C_D^{\max}$ of about 2.3×10^{-3}, that occurs upward of $V_{\max} = 34$ m/s, and that the ITOP data was inconclusive with regard to the slope m beyond saturation (figures 3 and 4). This work has been summarized in a manuscript that is currently under review (Sraj et al. (2012)).

The technical development concerned experimentation with various adaptive sparse quadrature methods to improve the efficiency and accuracy of calculating the polynomial chaos coefficients. This calculation is the most CPU-intensive portion of the forward UQ problem, and directly impacts the accuracy of the resulting PC expansions. In our previous work (Alexanderian et al. (2012)) an ensemble of 385 members were required to explore a four-dimensional parameter space where each variable was expanded in 5th-order PC expansion. Subsequent analysis, however, revealed that the quantity of interest (surface temperature) was insensitive to 3 out of the 4 parameters, that the response for 2 of the four parameters is almost linear, and that 7-th order polynomials would be preferable in one direction. This was a clear indication that some form of adaptation is required. The database of 385 members was hence expanded to 513 members to provide a reference solution against which adaptive schemes can be tested. Furthermore, a database of 256 simulations were also computed based on Latin Hypercube Sampling to provide a separate and random set of sampling points that is distinct from the training points (the 513 member set). The database consisted then of 769 independent HYCOM realizations. Two adaptive methods were tested: one simply truncates the PC representation anisotropically in each parameter direction (Gerstner and Griebel (2003)); the other relies on a pseudo-spectral construction that accomodates arbitrary admissible sparse quadrature grids (Constantine et al. (2012); Conrad and Marzouk (2012)). This second approach proved to be superior to the first one and is the one that was adopted for our inverse modeling work. Figure 5 compares the error metrics of the unadapted, the dimensional truncation, and the pseudo-spectral adaptations: the latter achieves a lower error levels (and maintains it in time) using a smaller number of realizations than the other approaches. The description and analysis of the adaptive algorithms are currently being written for publication. We should also note that the HYCOM database is currently being used as a reference solution by numerous groups to test new adaptive quadratures.

In addition to the aforementioned work we have been applying PC expansions to study the impact of uncertain initial conditions on an ocean forecast in the Gulf of Mexico. Modes of variability were identified from a long-running HYCOM simulation, multiplied by stochastic amplitudes and then added as perturbations to the initial state of the ocean forecast. We have performed an ensemble run of these perturbations and we are in the process of analyzing the results. We have also initiated a discussion with our colleagues on the application of PC methods to Lagrangian drifter studies, particularly those associated with the Deep Water Horizon oil spill of 2010. These discussions are still in the early stages.

The work associated with this project has been publicized at several conferences, workshops, seminars and invited talks, including:

- P. Conrad, J. Winokur, I. Sraj, A. Alexanderian, M. Iskandarani, A. Srinivasan, Y. Marzouk, O. Knio (2012) Sparse Adaptive Polynomial Chaos Representations for Ocean General Circulation Models, presented at 9th AIMS Conference, Orlando, FL, July 1-5, 2012. (invited)
- J. Winokur, P. Conrad, I. Sraj, A. Alexanderian, M. Iskandarani, A. Srinivasan, Y. Marzouk, O.M. Knio (2012) A Priori Testing of Adaptive Sampling and Sparse PC Representations for Ocean General Circulation Models, presented at 2012 SIAM International Conference on Data Mining, Anaheim, CA, April 26-28, 2012. (invited)
- O.M. Knio (2012) Polynomial Chaos Approaches to Multiscale and Data Intensive Computations, presented at SIAM Conference on Uncertainty Quantification, Raleigh, NC, April 2-5, 2012. (plenary)
- J. Winokur, A. Alexanderian, I. Sraj, M. Iskandarani, A. Srinivasan, C. Thacker, O. Knio (2011) Quantifying Parametric Uncertainty in Ocean General Circulation Models: A Sparse Quadrature Approach, presented at the DFD11 Meeting of the American Physical Society.
- Polynomial Chaos Approaches to Multiscale and Data Intensive Computations, UNC, March 23,

2012.

- Uncertainty Quantification Challenges in Modeling Complex Systems, Winter Enrichment Program, KAUST, January 15, 2012.
- Organized a minisymposium "Sensitivity Analysis, Data Assimilation and Uncertainty Quantification in Ocean Modeling" at the Ocean Science 2012 meeting in collaboration with Ibrahim Hoteit, and Bruce Cornuelle.
- "Bayesian Inference of Drag Coefficient Parameters using AXBT data from Typhoon Fanapi", University Miami, MPO seminar, Sep 5 2012.
- "Uncertainty Analysis and Quantification of the HYCOM SST Response to Hurricane Ivan Using Polynomial Chaos Expansions" Ocean Sciences Meeting 2012, Utah Feb 24 2012.
- "Propagating Oceanographic Uncertainties Using the Method of Polynomial Chaos Expansion" Ocean Sciences Meeting 2012, Utah Feb 22 2012.
- "Application of Polynomial Chaos Expansions for Uncertainty Quantification in Oceanic Simulations", University Miami, MPO seminar, Sep 7 2011.

RESULTS

The inverse modeling problem provided a very good and practical example of how PC methods can be used effectively in estimating hard to measure parameters. The results showed that the wind drag coefficient saturates at around 2.3×10^{-3} when the wind speed is about 34 m/s, and that for higher wind speeds the data is incloncusive whether the drag coefficient decreases or remains constant. The wind speed estimate is biased towards the end of the interval explored, and suggests that the maximum wind speed is at least the inferred value. One limitation, unfortunately, is that the ITOP data did not include AXBT drops at higher wind speeds.

The inverse modeling success illustrate the great potential the present methodology may hold for other pressing UQ and parameter estimation problems of interest, e.g. the impact of air-sea interaction uncertainty on hurricane intensification. The present pilot study focused on a single component of the coupled ocean-atmosphere system for simplicity. Other air-sea interaction uncertain parameters, such as the heat exchange coefficient, are also of interest, but their influence is likely to be more dramatic on the atmosphere than on the ocean and thus necessitates the use of an atmospheric model. We certainly hope to explore the impact of these uncertainties on atmospheric simulations in the future.

Polynomial Chaos methods can be categorized as ensemble methods, and the size of the ensemble that one can afford is the main hurdle to their application to large and complex models. The adaptive strategy has shown that our exploration of a four-dimensional space with equal weighing of all parameters was wasteful of CPU cycles and storage, and that an ensemble of 70 realizations would have been enough. The adaptation also provides a continuous monitoring of error metrics to ensure that the resulting PC representation is accurate. This adaptive strategy will significantly enhance our ability to increase the size of the problem pursued while maximizing the efficiency of our computations.

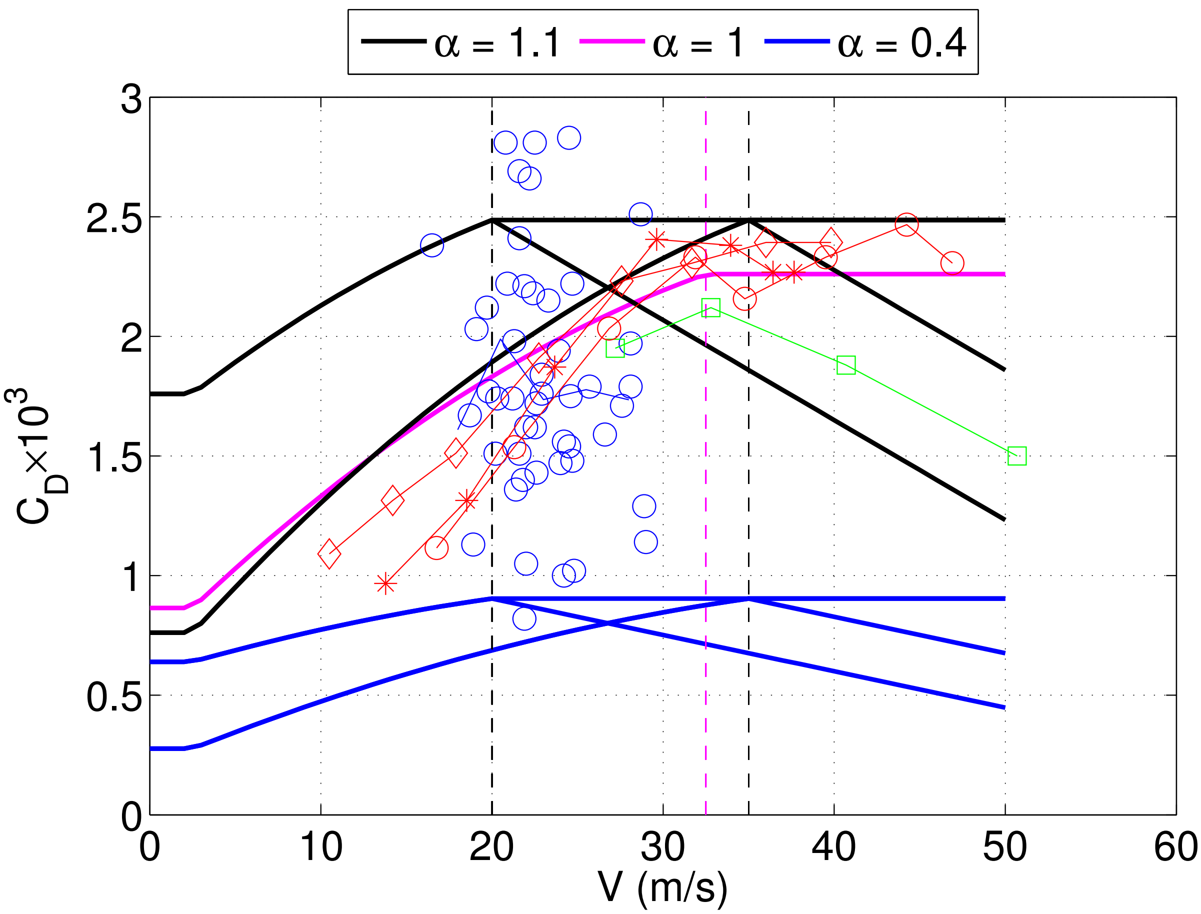

Figure 1: Observation-based estimates of drag coefficient C_D compared with our wind-speed-dependent representation. Thin blue (French et al. (2007)), red (Donelan et al. (2004)), and green (Powell et al. (2003)) lines through corresponding colored points indicate the observation-based estimates. Our representation is a modification of that of Kara et al. (2002): the parameter $V_{\max}$ is used to adjust the wind speed at which the drag coefficient saturates; for speeds less than $V_{\max}$ the parameter α is used to adjust the size of the drag coefficient while preserving the shape of the wind-speed dependence; and for speeds greater than $V_{\max}$ the parameter m (slope of drag coefficient after saturation) allows for the possibility of decreasing drag with increasing wind. The blue and black curves illustrate the eight cases determined by putting the parameters at the extremes of their allowed ranges are shown in blue ($\alpha = 0.4$) and black ($\alpha = 1.1$) with $V_{\max}$ either 20 or 35 m/s and m either -3.8$\times 10^{-5}$ or 0. The unperturbed HYCOM parameterization of C_D (Kara et al., 2002) is shown in magenta ($\alpha = 1$, $V_{\max} = 32.5\ m/s$ and $m = 0$).

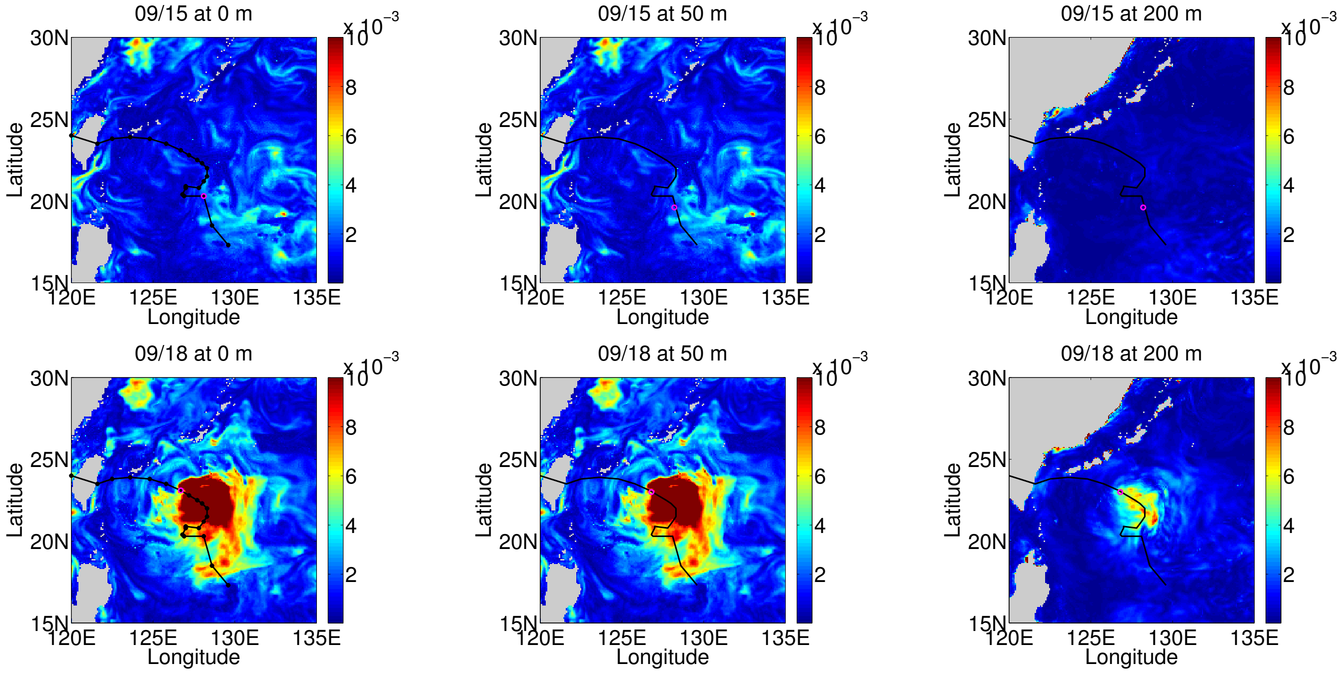

Figure 2: Relative normalized error between realizations and the corresponding PC surrogates at different depths: surface (left); 50 m (center); and 200 m (right). Top row: 00:00 UTC Sep 15; bottom row: 00:00 UTC Sep 18.

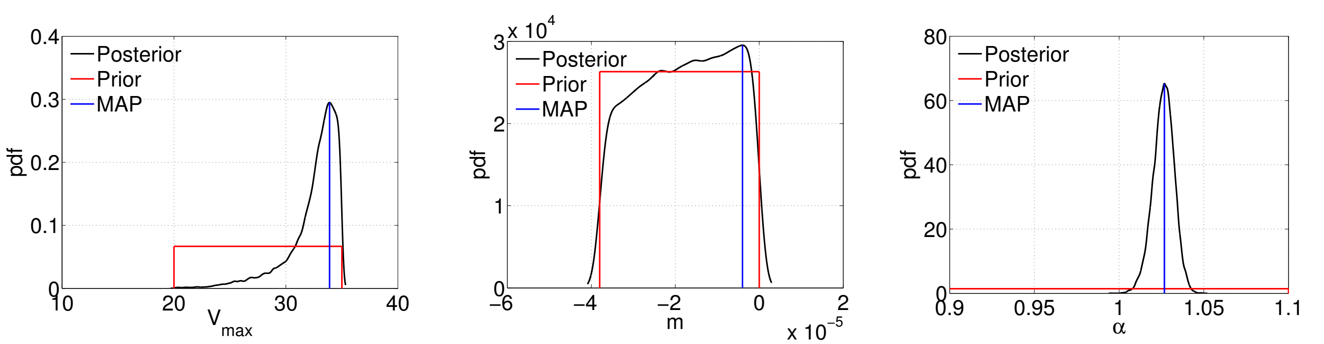

Figure 3: Posterior distributions for the drag parameters (top) and the variance between simulations and observations (bottom). The posterior pdf of V_{max} ***(left), exhibits a well-defined peak at around 34 m/s, the posterior of*** m ***is similar to the prior and indicate no information gain from the data (upper center), and finally the multiplicative factor shows a sharp peak at*** 1.03***.***

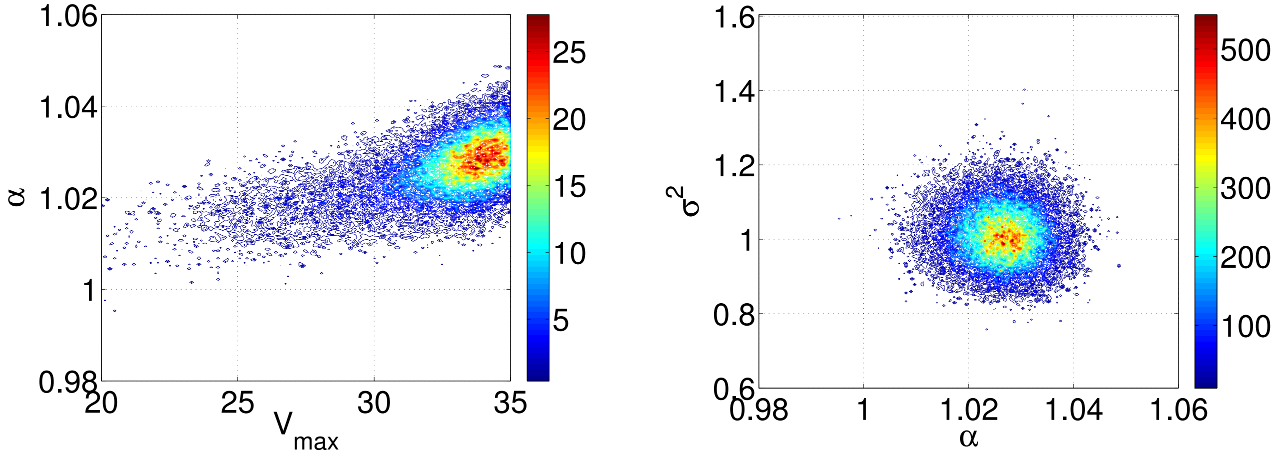

Figure 4: Left: joint posterior distribution of α (left) and V_{max}; right: joint posterior of α and σ^2, generated for Sep 17-Sep 18.

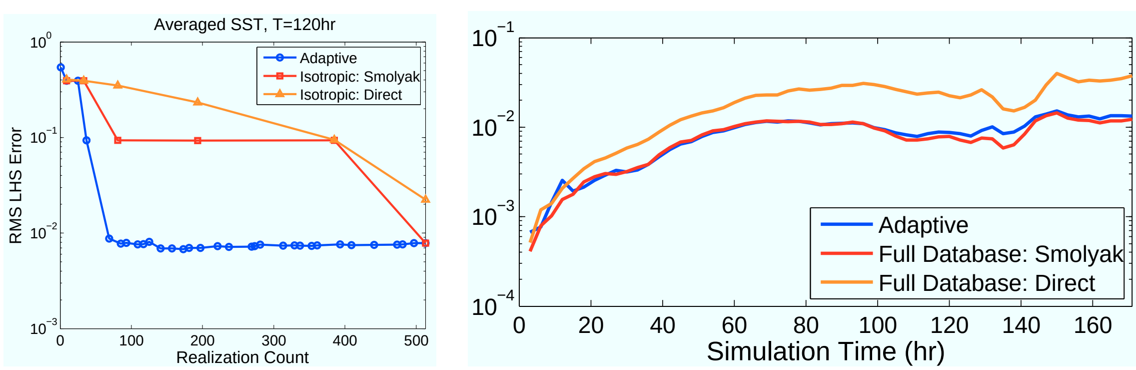

Figure 5: Left: LHS error of adaptive and isotropic truncations (with Smolyak pseudo-spectral and direct projections) for area-averaged SST at t=120hr; the dramatic reduction in the adaptive case occurs early (iteration 69) because the parameter needing the most attention has been identified. Right: Time evolution of the LHS error for the different quadrature and truncation schemes (direct projection, pseudo-spectral and adaptive).The adaptive truncation is based on iteration 5 at T=60hr and uses 69 realizations with 59 polynomials. The full 513 database curves have 402 polynomials for the Smolyak pseudo-spectral construction and 168 polynomials for the direct projection.

IMPACT/APPLICATIONS

The present project presents an approach to characterize the entire response surface of an ocean model to uncertainties in its input data. This has implications for the fields of parameter estimation, and data assimilation, particularly for ensemble Kalman filter based approaches. The methodology developed here will be of use either for the efficient update of the covariance matrices and/or quantifying the errors incurred by small size ensembles. We are currently exploring these ideas.

TRANSITIONS

RELATED PROJECTS

Dr. Ashwanth Srinivasan was partially supported by an NSF-RAPID grant (NSF OCE-1048697) for his work on the oil-fate model.

PUBLICATIONS

A. Alexanderian, O. Le Maître, H. Najm, M. Iskandarani, and O. Knio. Multiscale stochastic preconditioners in non-intrusive spectral projection. *Journal of Scientific Computing*, 50(2):306–340, 2011. ISSN 0885-7474. doi: 10.1007/s10915-011-9486-2. URL http://dx.doi.org/10.1007/s10915-011-9486-2.

A. Alexanderian, J. Winokur, I. Sraj, A. Srinivasan, M. Iskandarani, W. C. Thacker, and O. M. Knio. Global sensitivity analysis in an ocean general circulation model: a sparse spectral projection approach. *Computational Geosciences*, 16(3):757–778, 2012. doi: 10.1007/s10596-012-9286-2. URL http://dx.doi.org/10.1007/s10596-012-9286-2.

I. Sraj, M. Iskandarani, A. Srinivasan, W. C. Thacker, J. Winokur, A. Alexanderian, C.-Y. Lee, S. S. Chen, and O. M. K. (2012). Bayesian inference of dependence of drag coefficient on wind speed using axbt data from typhoon fanapi. *Monthly Weather Review*, 2012. submitted.

A. Srinivasan, J. Helgers, C. B. Paris, M. LeHenaff, H. Kang, V. Kourafalou, M. Iskandarani, W. C. Thacker, J. P. Zysman, N. F. Tsinoremas, and O. M. Knio. Many task computing for modeling the fate of oil discharged from the deep water horizon well blowout. In *Many-Task Computing on Grids and Supercomputers (MTAGS), 2010 IEEE Workshop on*, pages 1–7, November, 2010. IEEE. doi: 10.1109/MTAGS.2010.5699424. URL http://dx.doi.org/10.1109/MTAGS.2010.5699424.

W. C. Thacker, A. Srinivasan, M. Iskandarani, O. M. Knio, and M. L. Henaff. Propagating boundary uncertainties using polynomial expansions. *Ocean Modelling*, 43–44:52–63, 2012. doi: 10.1016/j.ocemod.2011.11.011. URL http://dx.doi.org/10.1016/j.ocemod.2011.11.011.

REFERENCES

R. Bleck. An oceanic general circulation model framed in hybrid isopycnic cartesian coordinates. *Ocean Modelling*, 4(2):55–88, 2002.

E. P. Chassignet, L. T. Smith, G. R. Halliwell, and R. Bleck. North atlantic simulations with the hybrid coordinate ocean model (hycom): Impact of the vertical coordinate choice, reference pressure, and thermobaricity. *Journal of Physical Oceanography*, 33(12):2504–2526, DEC 2003. ISSN 0022-3670.

E. P. Chassignet, H. E. Hurlburt, O. M. Smedstad, G. R. Halliwell, P. J. Hogan, A. J. Wallcraft, and R. Bleck. Ocean prediction with the hybrid coordinate ocean model (hycom). In E. Chassignet and J. Verron, editors, *Ocean Weather Forecasting: An Integrated View of Oceanography*, pages 413–426, PO BOX 17, 3300 AA DORDRECHT, NETHERLANDS, 2006. SPRINGER. ISBN 1-4020-3981-6. International Summer School of Oceanography, Agelonde, FRANCE, SEP 20-OCT 01, 2004.

P. R. Conrad and Y. M. Marzouk. Adaptive smolyak pseudospectral expansions. In *Proceedings of SIAM conference on Uncertainty Quantification*, pages 74–74, Raleigh NC, April 2012. SIAM.

P. G. Constantine, M. S. Eldred, and E. Phipps. Sparse pseudospectral approximation method. *Computer Methods in Applied Mechanics and Engineering*, 28(24):11751–11762, 2012. doi: 10.1016/j.cma.2012.03.019.

M. A. Donelan, B. K. Haus, N. Reul, W. J. Plant, M. Stiassnie, H. C. Graber, O. B. Brown, and E. S. Saltzman. On the limiting aerodynamic roughness of the ocean in very strong winds. *Geophysical Research Letters*, 31(L18306):1–5, September 2004. doi: doi:10.1029/2004GL019460. URL http://dx.doi.org/10.1029/2004GL019460.

J. R. French, W. M. Drennan, J. A. Zhang, and P. G. Black. Turbulent fluxes in the hurricane boundary layer. part i: Momentum flux. *Journal of the Atmospheric Sciences*, 64:1089–1102, 2007. doi: doi:10.1175/JAS3887.1.

T. Gerstner and M. Griebel. Dimension-adaptive tensor-product quadrature. *Computing*, 71(1):65–87, Sep 2003. doi: 10.1007/s00607-003-0015-5.

G. R. Halliwell. Evaluation of vertical coordinate and vertical mixing algorithms in the hybrid-coordinate ocean model (hycom). *Ocean Modelling*, 7(3-4):285–322, 2004. ISSN 1463-5003. doi: 10.1016/j.ocemod.2003.10.002. URL http://dx.doi.org/10.1016/j.ocemod.2003.10.002.

A. B. Kara, P. A. Rochford, and H. E. Hurlburt. Air-sea flux estimates and the 1997-1998 enso event. *Boundary-Layer Meteorology*, 103:439–458, 2002. ISSN 0006-8314. doi: 10.1023/A:1014945408605. URL http://dx.doi.org/10.1023/A:1014945408605.

O. P. Le Maître and O. M. Knio. *Spectral Methods for Uncertainty Quantification*. Springer-Verlag, 2010.

Y. M. Marzouk, H. N. Najm, and L. A. Rahn. Stochastic spectral methods for efficient bayesian solution of inverse problems. *Journal of Computational Physics*, 224(2):560–586, JUN 10 2007. ISSN 0021-9991. doi: {10.1016/j.jcp.2006.10.010}.

M. D. Powell, P. J. Vickery, and T. A. Reinhold. Reduced drag coefficient for high wind speeds in tropical cyclones. *Nature*, 422:279–283, March 2003. doi: doi:10.1038/nature01481. Wind measurement and wind stress calculation in (observed) hurricane conditions.

HONORS/AWARDS/PRIZES

- 13 invited lectures, of which 2 were plenary.
- O. Knio was named Distinguished Professor, July 1, 2012.

- O. Knio was elected member of the Editorial Board of SIAM/ASA Journal on Uncertainty Quantification.