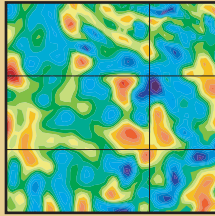# Quantifying
Uncertainties
in Ocean
Predictions

BY PIERRE F.J. LERMUSIAUX,

CHING-SANG CHIU,

GLEN G. GAWARKIEWICZ,

PHIL ABBOT,

ALLAN R. ROBINSON,

ROBERT N. MILLER,

PATRICK J. HALEY,

WAYNE G. LESLIE,

SHARANYA J. MAJUMDAR,

ALEX PANG, AND

FRANCOIS LEKIEN

> *From observations and conservation laws,*
> *ocean scientists formulate models that aim*
> *to explain and predict dynamics of the sea.*

A multitude of physical and biological processes occur in the ocean over a wide range of temporal and spatial scales. Many of these processes are nonlinear and highly variable, and involve interactions across several scales and oceanic disciplines. For example, sound propagation is influenced by physical and biological properties of the water column and by the seabed. From observations and conservation laws, ocean scientists formulate models that aim to explain and predict dynamics of the sea. This formulation is intricate because it is challenging to observe the ocean on a sustained basis and to transform basic laws into generic but usable models. There are imperfections in both data and model estimates. It is important to quantify such uncertainties to understand limitations and identify the research needed to increase accuracies, which will lead to fundamental progress.

There are several sources of uncertainties in ocean modeling. First, to simplify models (thereby reducing computational expenses), explicit calculations are only performed on a restricted range of spatial and temporal scales (referred to as the "scale window")

(Nihoul and Djenidi, 1998). Influences of scales outside this window are neglected, parameterized, or provided at boundaries. Such simplifications and scale reductions are a source of error. Second, uncertainties also arise from the limited knowledge of processes within the scale window, which leads to approximate representations or parameterizations. Third, ocean data are required for model initialization and parameter values; however, raw measurements are limited in coverage and accuracy, and they are often processed with the aim of extracting information within a predetermined scale window. Initial conditions and model parameters are thus inexact. Fourth, models of interactions between the ocean and Earth system are approximate and ocean boundary conditions are inexact. For example, effects of uncertain atmospheric fluxes can dominate oceanic uncertainty. Fifth, miscalculations occur due to numerical implementations. All of the above leads to differences between the actual values (unknown) and the measured or modeled values of physical, biological, and geo-acoustical fields and properties.

To reduce uncertainties, the sources of information (the various data and dynamical models) are combined by data assimilation (DA) (Robinson et al., 1998; Robinson and Lermusiaux, 2002). Data assimilation is challenging and expensive to carry out, but optimal in the sense that each type of information is weighted in accord with its uncertainty. Of course, should optimal estimates fail to be accurate, a priori assumptions about uncertainties are revised, and models and data sets improved.

Any comprehensive ocean prediction (e.g., Mooers, 1999; Pinardi and Woods, 2002) should include uncertainty estimates. Predicted uncertainties consist of the integration in time of initial errors and of errors introduced during model integration. Uncertainty is defined in terms of the probability density function (PDF) of the error in the estimate. Error refers to the difference between the truth and the estimate. Uncertainties are often represented by low-order characteristics of the error PDF (e.g., the moments or confidence intervals). Because ocean fields are four-dimensional, straightforward uncertainty representations are also fields, with structures in time and space. Variability and uncertainty are related but different (e.g., Lermusiaux,

2002). For any estimate, the portion of variability that contains errors contributes to uncertainty. The variability that is unresolved is purely uncertainty. For example, the historical temperature variability maps shown on Figure 1 are moments of a variability PDF. The standard deviations (Figure 1b) are uncertainty amplitudes for the mean (Figure 1a) if the historical data are the sole information used to estimate this mean.

Although uncertainties have been at the heart of ocean investigations for a long time, realistic uncertainty predictions are recent. Early attempts in the context of DA are described in Malanotte-Rizzoli (1996). The first real-time uncertainty predictions using an advanced DA scheme in a full-featured nonlinear model were carried out for

the Strait of Sicily in 1996 (Lermusiaux, 1999). The scheme utilized was Error Subspace Statistical Estimation (ESSE, Lermusiaux et al., 2002). Related Monte-Carlo ensemble schemes (Evensen, 1994; Miller et al., 1999) are now being used in other regions. Generalized inverse schemes can account for all uncertainty sources (an excellent example is Egbert et al., 1994), but avoid computing uncertainty fields to gain computational speed. In atmospheric studies, ensemble forecasting has been utilized for uncertainty predictions for some time (e.g., Toth and Kalnay, 1993; Molteni et al., 1996; Ehrendorfer, 1997) and realistic ensemble DA has been carried out recently (Whitaker et al., 2004; Houtekamer et al., 2005; Szunyogh et al., 2005). Climate uncertainty forecasting has been initiated, often based on simple perturbations of selected parameters and initial conditions (Murphy et al., 2004; Stainforth et al., 2005).

The present study describes and illustrates the mechanics and computations involved in modeling and predicting uncertainties for ocean science and its modern applications. It is an outgrowth of the U.S. Office of Naval Research's (ONR) Capturing Uncertainty in the Tactical Environment Initiative (ONR, 2001),

The evolution of error covariances depends on four factors: (1) the initial error condition, (2) the deterministic dynamics that increase or reduce errors..., (3) the stochastic forcings that model errors in the deterministic model..., and (4) the impact of data that reduces variance.

Pierre F.J. Lermusiaux (*pierrel@pacific.deas.harvard.edu) is Research Associate, Harvard University, Cambridge, MA, USA*. **Ching-Sang Chiu** *is Professor, Naval Postgraduate School, Monterey, CA, USA*. **Glen G. Gawarkiewicz** *is Associate Scientist, Woods Hole Oceanographic Institution, Woods Hole, MA, USA*. **Phil Abbot** *is President, Ocean Acoustical Services and Instrumentation Systems, Inc., Lexington, MA, USA*. **Allan R. Robinson** *is Professor, Harvard University, Cambridge, MA, USA*. **Robert N. Miller** *is Professor, Oregon State University, Corvallis, OR, USA*. **Patrick J. Haley** *is Project Scientist, Harvard University, Cambridge, MA, USA*. **Wayne G. Leslie** *is Senior Project Scientist, Harvard University, Cambridge, MA, USA*. **Sharanya J. Majumdar** *is Research Assistant Professor, University of Miami, Miami, FL, USA*. **Alex Pang** *is Professor, University of California, Santa Cruz, CA, USA*. **Francois Lekien** *is Research Associate, Princeton University, Princeton, NJ, USA*.
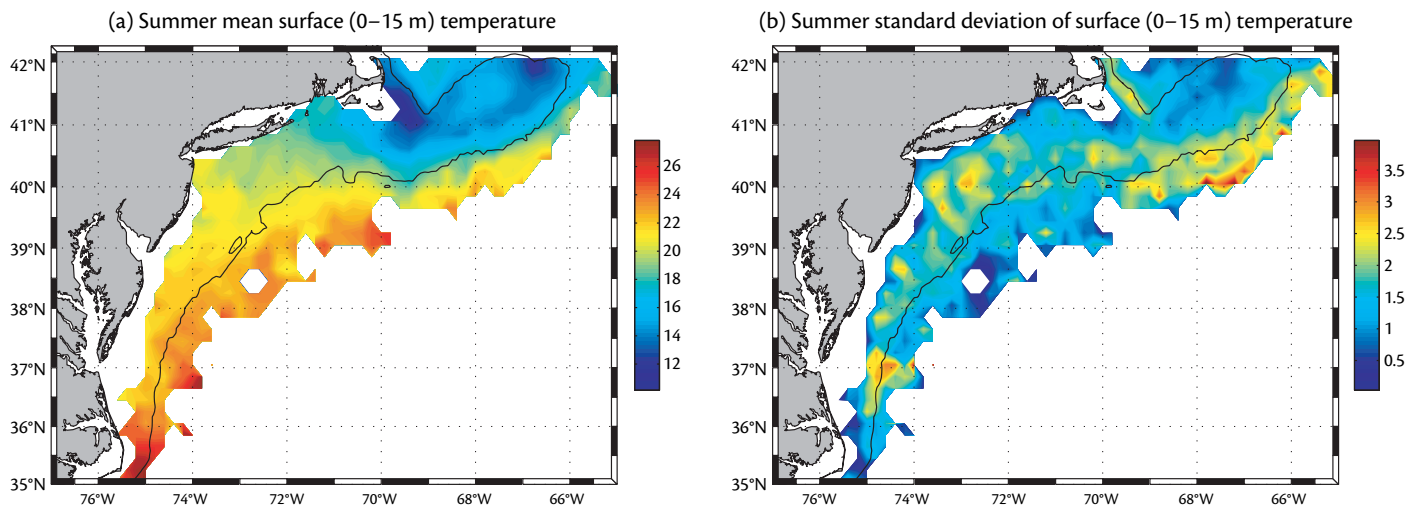
Figure 1. Measurement-based variability estimates for the summer season in the Mid-Atlantic Bight (MAB) continental shelf and slope region, within 0-m to 15-m depths. (a) Mean of the temperature data, in °C. (b) Standard deviation of the temperature data, in °C. The data used to compute these maps are historical raw temperature profiles from a variety of data sources (Linder and Gawarkiewicz, in press). The maps are representations of the variability—the mean and standard deviation of a variability PDF.

which involved scientists from physical oceanography, ocean modeling, marine geosciences, ocean acoustics, signal processing, and sonar engineering. Detailed mathematical and computational aspects are given in the references. ESSE is used to exemplify interdisciplinary data-assimilative uncertainty estimation and prediction, focusing on regional applications.

MODELING AND PREDICTING UNCERTAINTIES IN THE OCEAN

Uncertainty estimation begins with the identification of significant variability that is not represented. This is the input or prior information. Uncertainty predictions can then be obtained from evolution equations (Jazwinski, 1970) for the error PDFs of the model state and parameters. When observations are made, these PDFs are combined with the new data and their PDFs. However, it is impractical to solve such PDF equations for discrete ocean-model variables because of the large number $O(10^5-10^7)$ of grid points. Ocean uncertainty estimation

has thus focused on: (1) the conditional mean, which is the minimum error variance estimate, and (2) error variances and covariances, which are simple but essential components of the error statistics. (The variance is the square of the standard deviation, which is a measure of the average deviation from the mean. Error covariances measure the extent to which errors in two variables vary together.)

The evolution of error covariances depends on four factors: (1) the initial error condition, (2) the deterministic dynamics that increase or reduce errors by internal advection, diffusion, or reaction, and by external forcing, (3) the stochastic forcings that model errors in the deterministic model and increase error variance, and (4) the impact of data that reduces variance. Each of these factors is normally important. Care is thus required when approximate equations are used to evolve error covariances. For example, a passive tracer equation would only capture part of factor 2.

Ocean uncertainty forecasts can be

used to qualify the prediction, assimilate data, or estimate predictability limits. Today, most uncertainty forecasting schemes are based on ensemble Monte-Carlo approaches and reductions of the high-dimensional error space to a low-dimensional subspace that contains the essential uncertainty. The schemes first aim to account for the largest uncertainties in each source of information: dynamical model, measurement model/data, initial and boundary conditions, and parameters. With these uncertainty inputs, they then predict the largest uncertainties (the error subspace) of the dynamical state and reduce them by DA. Mathematically, it is the DA criterion that sets the choice of the subspace. The suboptimal truncation of errors in the full space is then optimal. For a minimum error variance, the subspace is defined by dominant modes of the error covariance matrix. Computational components involved in such modeling and prediction of uncertainties are outlined next and illustrated with ESSE.

Deterministic Models and Their Approximations

Most physical ocean models are derived from the classic Navier-Stokes equations for fluid dynamics in a rotating frame of reference. These equations are deterministic: they always give the same output for a given input. Practical assumptions are used to limit the range of modeled scales. A common reduction, the Primitive-Equations (PE) model (Pedlosky, 1987), is used here within the Harvard Ocean Prediction System (HOPS, 2004). Acoustic models are also derived from Navier-Stokes and are usually based on a wave equation for the sound pressure (Kuperman, 2004). Efficient acoustic model approximations include modal decompositions (Chiu et al., 1996) and linearizations. Even though much progress has been made in marine ecosystem modeling (e.g., Hofmann and Friedrichs, 2002), deterministic biological equations as fundamental as Navier-Stokes are not yet available. For lower trophic levels, most models are based on advection-reaction-diffusion equations. They differ in their structure, the number of state variables employed, and the parameterizations used. Details on the models used in this manuscript are in Lermusiaux et al. (2002) and Lermusiaux and Chiu (2002).

Deterministic physical, biological, or acoustical models commonly compute future conditions based on given initial conditions. They also play an important role in the prediction of uncertainties. They allow explaining the deterministic evolution of the initial errors. However, approximations to fundamental equations lead to errors in these models, which need to be taken into account.

Stochastic Forcing and Models of Uncertainties in Deterministic Dynamical Models

To represent the dominant components of processes neglected or not well represented in deterministic models, stochastic error models are starting to be used. For example, Figure 2 illustrates statistical effects of sub-mesoscale processes not resolved in a mesoscale-resolution PE model. The model errors are modeled using unbiased random noise with an exponential decorrelation in time. In space, the amplitude of the noise varies only in the vertical direction and has about a two-grid point correlation in every spatial direction. For each prognostic equation, the noise variance at a given depth is set to be a small fraction (25 percent) of the amplitude of the terms involved in the dominant dynamical balance at that depth.

What is modeled with a deterministic or stochastic equation depends on knowledge and on the scale window of interest. Generally, processes that are well known in this window are modeled deterministically. All other processes, inside or outside of the scale window, should be modeled with a stochastic component. These stochastic components can be either additive (added as a new term to the deterministic model) or multiplicative (e.g., inside an original term of the model). Additive forcing, uncorrelated with the deterministic variables, is useful in ocean models, but it should be auto-correlated in time and space because the statistics of many natural processes can be approximated this way (Gardiner, 1983; Lermusiaux et al., 2002).

Boundary Condition Uncertainties

Open boundary conditions in regional modeling are a large source of uncertainties, in part because their estimation is not always well posed (Bennett, 1992, and references therein). Moreover, exchanges between the ocean and atmosphere are often computed based on atmospheric forcing fluxes obtained from an independent atmospheric model. Inaccuracies also arise in surface and coastal boundary conditions such as parameterizations of boundary layers, fluxes exchanged at coastlines, or river discharge inputs. Most boundary condition uncertainties are modeled with simple stochastic forcing and can be underestimated to limit numerical instabilities. For example, ESSE currently uses white noise models or time-correlated noise models at boundaries. At the ocean surface, more advanced atmospheric flux uncertainty models are definitely needed, for example, to account for flow-dependent uncertainties.

Parametric Uncertainties

To motivate the need for representing parametric uncertainties, consider the fit of the mixing-layer depth in a parameterization of the transfer of wind stress to the ocean's surface boundary layer (Lermusiaux, 2001). Figure 3 illustrates such a fit of the mixing-layer depth factor to Seasoar data collected during the shelfbreak PRIMER experiment and to atmospheric fluxes obtained from adjusted model fields (Baumgartner and Anderson, 1999). The fitted factor varies in time (solid curve on top of Figure 3). Its uncertainty is represented by the histogram around the mean fit (Figure 3, bottom). As the top panel shows, in the
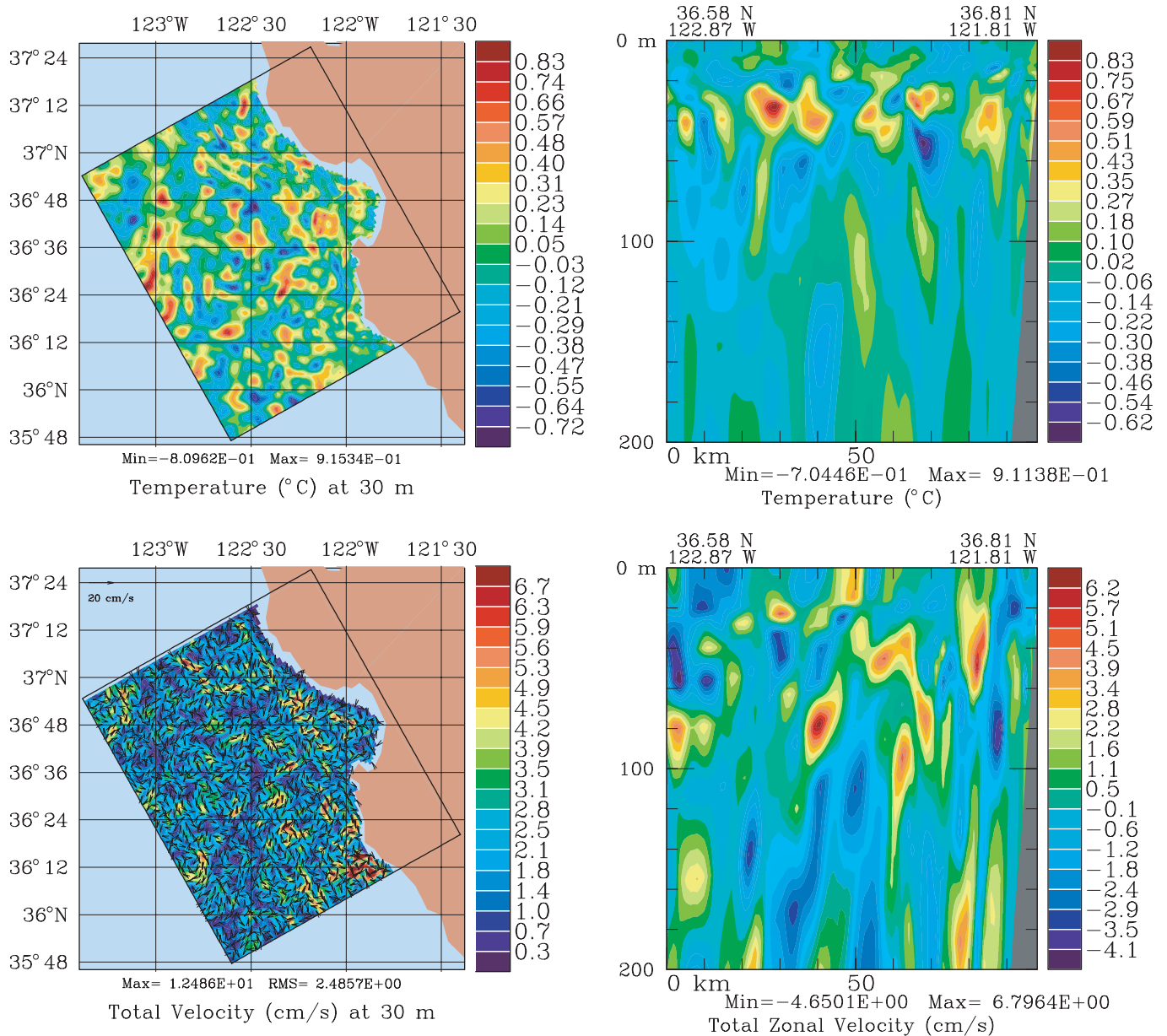
Figure 2. Illustration of the effects of adding random forcings, correlated in time and space, to a PE model. Shown are differences between a deterministic and stochastic PE model simulation over the Monterey Bay and California Current System region, after one day of integration. Differences in horizontal maps of T and $\|\boldsymbol{u}_n\|$ are shown at 30-m depth (top and bottom left) and in cross sections (from offshore to the coast in Monterey Bay) of T and u, from 0–200 m depth. The amplitudes of the random forcings were set to a fraction (25 percent) of the average geostrophic balance at each depth ($\|\text{geostrophy}(z)\|$), with a half-day decorrelation in time and one-to-two grid point correlation in space. Geostrophy is not always the sole component of the dominant PE balance in the region. Future random forcing models will for example include impacts of atmospheric forcing in the balance.

ideal case, the value of the Ekman factor should be adapted in real time to the wind and mixed-layer depth data.

Most ocean model parameters are kept constant in time and space. Ide-ally, uncertainties of influential parameters should be modeled in a prediction or error budget. Parameter values can be estimated directly by DA. Priors are then assigned for each parameter and posteriors are the result of the DA. In biological estimation, such quantitative parameter estimation can be necessary to achieve meaningful results (e.g., Spitz et al., 2001).
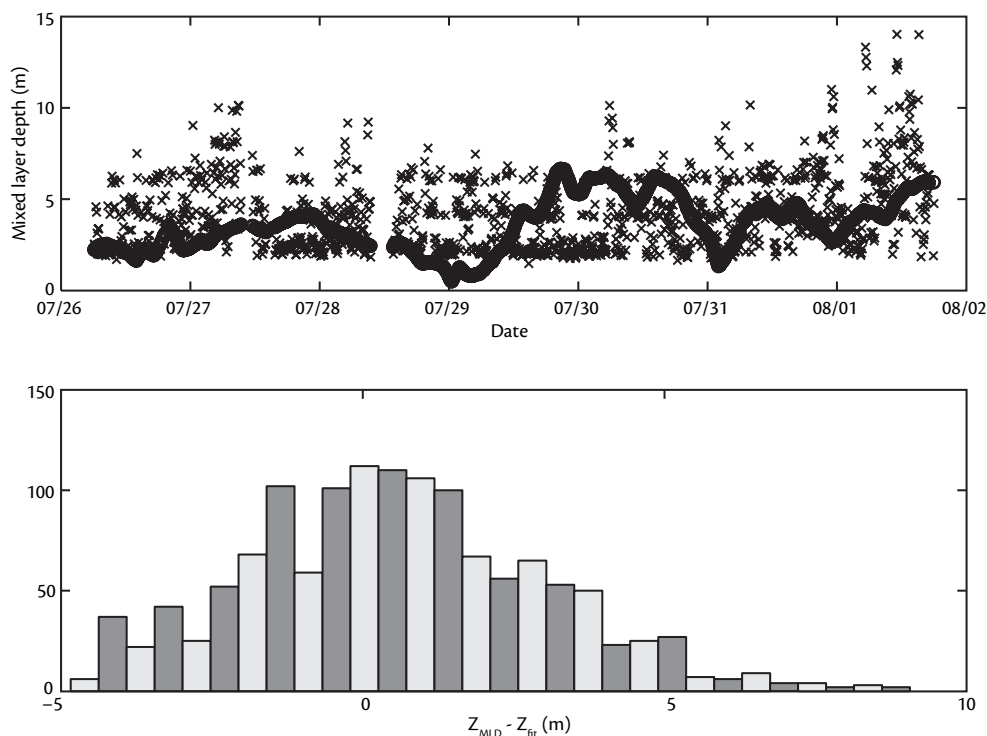
Fitted Ekman Factor: 0.0586

Figure 3. Uncertainties in multiple model parameters. Example of the mixing-layer depth factor over the MAB shelfbreak front region. This factor is the proportionality "Ekman factor (E_k)" linking the Ekman depth to the turbulent friction velocity and the Coriolis frequency. Top: Fitted mixed-layer depth (dark continuous curve) and estimated mixed-layer depths (crosses) computed for every Seasor ocean data profile and the corresponding local atmospheric model of wind-stress (Baugmarter and Anderson, 1996) at the profile location and time. Bottom: Distribution of the misfit between the: (1) mixed-layer depths estimated from the wind-stress and the ocean data; and (2), the fitted mixed-layer depth (dark curve on top). The fitted Ekman factor is 0.0586.

Measurement Models

The platforms and sensors used today in the ocean provide a wide range of observations on physical, biological, acoustical, and geological features, from microstructures to climate (Dickey, 2003; Griffiths et al., 2002). Melding these data with dynamical models requires measurement models that link the dynamical model variables and parameters to the observations. These models include uncertainties because sensors, data processing, and relationships among data and dynamical variables are approximate. Even though uncertainties in measurement models can be complex, simple data noise models are often used. For example, the ESSE system uses measurement errors that are correlated only in the vertical and that have amplitudes as a function of depth only.

Uncertainty Initialization

Dynamics and historical data are used for uncertainty initialization. In ESSE, the dominant error modes are initialized in two steps: what is observed first, then what is not observed. The "observed portions" are those that can be estimated from differences between a background state and historical data. Synoptic data can also be used to specify uncertainties corresponding to unresolved scales in the background. The "non-observed" portions are then computed by dynamical inference. First, the observed portions are used to perturb the background. An ensemble of model integrations is then carried out to adjust the non-observed portions to the observed ones. The result is an ensemble of complete uncertainty samples from which one can estimate dominant modes of

the initial error covariance. This approach can be generalized to multiple scales (Lermusiaux, 2002).

To obtain an ensemble of states, the background state is perturbed by a combination of the initial error modes, with dynamical constraints. A white noise model is also used to model uncertainties truncated by projection in the subspace. A detailed exposition of other ensemble generation methods can be found in Miller and Ehret (2002).

Data Assimilation and Uncertainty Reduction

In DA, the data, measurement models, and dynamical models are combined in accord with their prior uncertainty estimates using a criterion that determines the weight of each source of information. DA can provide melded estimates of the

state, parameters, and even of the model of the ocean. Various methods for DA in meteorology and oceanography can be found in Bennett (1992), Wunsch (1996), Robinson et al. (1998), and Kalnay (2003). Schemes are derived from estimation theory, control theory (variational approaches), or optimization theory. Estimation theory schemes solve a forward/filtering problem or a smoothing/inverse problem. Control theory schemes solve a smoothing problem. Almost all schemes are linked to a minimization of an error norm, the DA criterion. Optimization theory schemes directly minimize such a criterion or cost function. For real DA, most methods are based on least-squares norms and focus on the conditional mean and error covariance matrix.

Posterior uncertainties and data-model misfits can be used for two essential DA feedbacks. First, adaptive sampling estimates the types and locations of the observations that are most needed (Bishop et al., 2001; Lermusiaux, 1999). Second, adaptive modeling identifies the model properties that need the most improvements (Lermusiaux et al., 2004). These feedbacks lead to improved understanding and can be most powerful when uncertainty estimates are available.

METHODOLOGIES AND INTERDISCIPLINARY APPLICATIONS

Error Subspace Statistical Estimation

There are seven components in ESSE: error subspace initialization, state and uncertainty prediction, minimum error variance, adaptive error correction, smoothing, and adaptive sampling. The

ESSE uncertainty initialization is based on data, model, and multi-scale decompositions. The prediction is obtained from a central forecast and an ensemble of nonlinear model integrations that include random forcing. Currently, vertically correlated noise models represent uncertainties in measurements and white noise models in boundary conditions and parameters. Computations are distributed on a set of computers. Quantitative criteria control the ensemble size. When they are satisfied, data and models are combined by minimum error variance in the error subspace. Data residuals are then used for adaptive error corrections. To update past estimates, smoothing via ESSE is run backward in time. Adaptive sampling plans are predicted (Lermusiaux, 1999, 2001) using uncertainty estimates.

ESSE has been developed for, and applied to, fundamental research and real-time operations. The ESSE filtering/smoothing schemes permit physical, bio-logical, and acoustical DA with four-dimensional interdisciplinary covariances. Physical data then influence the biology and acoustics, and vice versa. Computational complexities in ESSE arise from the diversity of ocean geometries, data properties, deterministic and stochastic

parameters, and distributed workflows. However, many computations involve linear algebra, which allows the use of efficient community packages.

Estimation of Uncertainties in Secondary Variables: Coherent Structures

Nowcasts and forecasts are commonly used to infer secondary or diagnostic quantities, such as energy, vorticity, bioluminescence, acoustic travel time, drifter paths, or other Lagrangian indicators. For example, engineers and scientists interested in the trajectories of particles or vehicles in the ocean compute Direct Lyapunov Exponent (DLE) and Lagrangian Coherent Structures (LCSs) fields from velocity fields (Lekien et al., 2005, and references therein). The LCSs are here defined as ridges in the DLE field. They represent mobile separatrices, which divide the flow into regions of distinct motions and can indicate non-obvious boundaries in complex,

time-varying flows. Figure 4a shows the DLE field and their ridges during an upwelling event in Monterey Bay. The LCSs' ridges clearly separate regions of different properties. For example, consider the strong LCS between the cyclonic circulation in the Monterey Bay and the Cali-

> The representation, attribution and propagation of four-dimensional oceanic uncertainties presents many interesting challenges and requires increased theoretical and applied research efforts.

fornia Current System, or the circular LCS enclosing fluid trapped in an eddy. The utility of such estimates depends on their robustness to uncertainties. Figure 4b shows uncertainties transferred from the ocean state to the DLE state. Major LCSs are, in this case, regions of small relative uncertainties, making them robust descriptors of the flow.

Uncertainty Predictions for Acoustical and Physical Fields in a Shelfbreak Front

The main hydrographic feature near the Mid-Atlantic Bight shelfbreak is a mesoscale front of temperature, salinity, and hence sound-speed, separating the shelf and slope water masses (Figure 5a). The frontal system is variable on multiple temporal and spatial scales. Atmospheric forcing, Gulf Stream rings, river inflows, and buoyancy-driven flows, as well as tides and internal waves, affect its dynamics. The main *in situ* data were collected during July 26–August 4, 1996, mostly over an intensive acoustic domain (Figure 5b), as part of the ONR Shelf-break PRIMER Experiment (Lynch et al., 1997). The ocean model was substantially tuned to achieve useful physical-acoustical simulations (Robinson and Lermusiaux, 2004). ESSE is started on July 8 from a National Marine Fisheries Service survey. The central forecast on July 24 is illustrated on Figures 5a and 5b by horizontal maps of temperature (T) at 10 m (note the large meanders). Fields at 10-m depth are strongly influenced by both atmospheric forcing and internal ocean dynamics (Figure 5a), with the former imprinting its larger scales on the latter. The zoom around the acoustic region (Figure 5b), overlaid with horizontal current vectors, illustrates that larger-scale ocean context (Figure 5a) is necessary to understand the regional acoustic context. The corresponding error standard deviation maps (Figures 5c–d) show that at 10-m depth, the largest uncertainties in a 16-day prediction without DA are dominant around the surfacing location of the
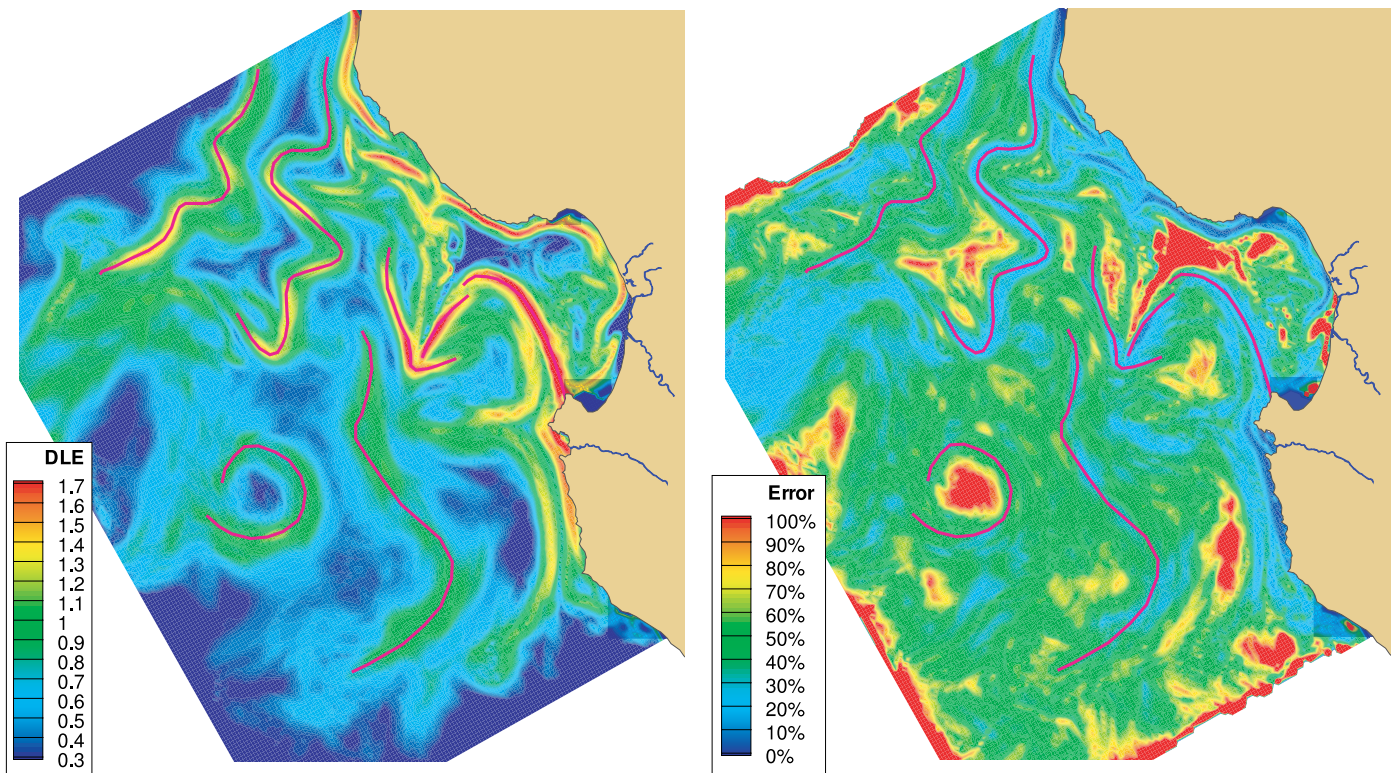

Figure 4. (Left) Direct Lyapunov Exponent (DLE) field during an upwelling event in Monterey Bay (August 26–29, 2003) derived from a velocity forecast. Ridges in the DLE field are highlighted by purple lines and represent Lagrangian Coherent Structures (LCSs), which divide the flow between regions of qualitatively different dynamics. The DLE and LCSs were computed using MANGEN (http://www.mangen.info), HOPS, and ESSE. (Right) Relative error in the DLE field computed from the Error Subspace Statistical Estimation (ESSE) velocity ensemble. In relative terms, the LCSs repel the uncertainties inherited from the ocean state.

front and its meanders, over 15–20 km on each side of the mean front. This distance is about twice the expected internal Rossby radius of deformation of the front. Importantly, at depths closer to the core of the tilted front (30–40 m), uncertainties are larger and more uniform in the horizontal, reflecting the frontal tilt and more turbulent nature of the internal dynamics of the front.

After DA on July 24, ESSE ocean physics uncertainties were transferred on July 26 to acoustical uncertainties across the shelfbreak along the main acoustic vertical section (its position is on Figure 5).
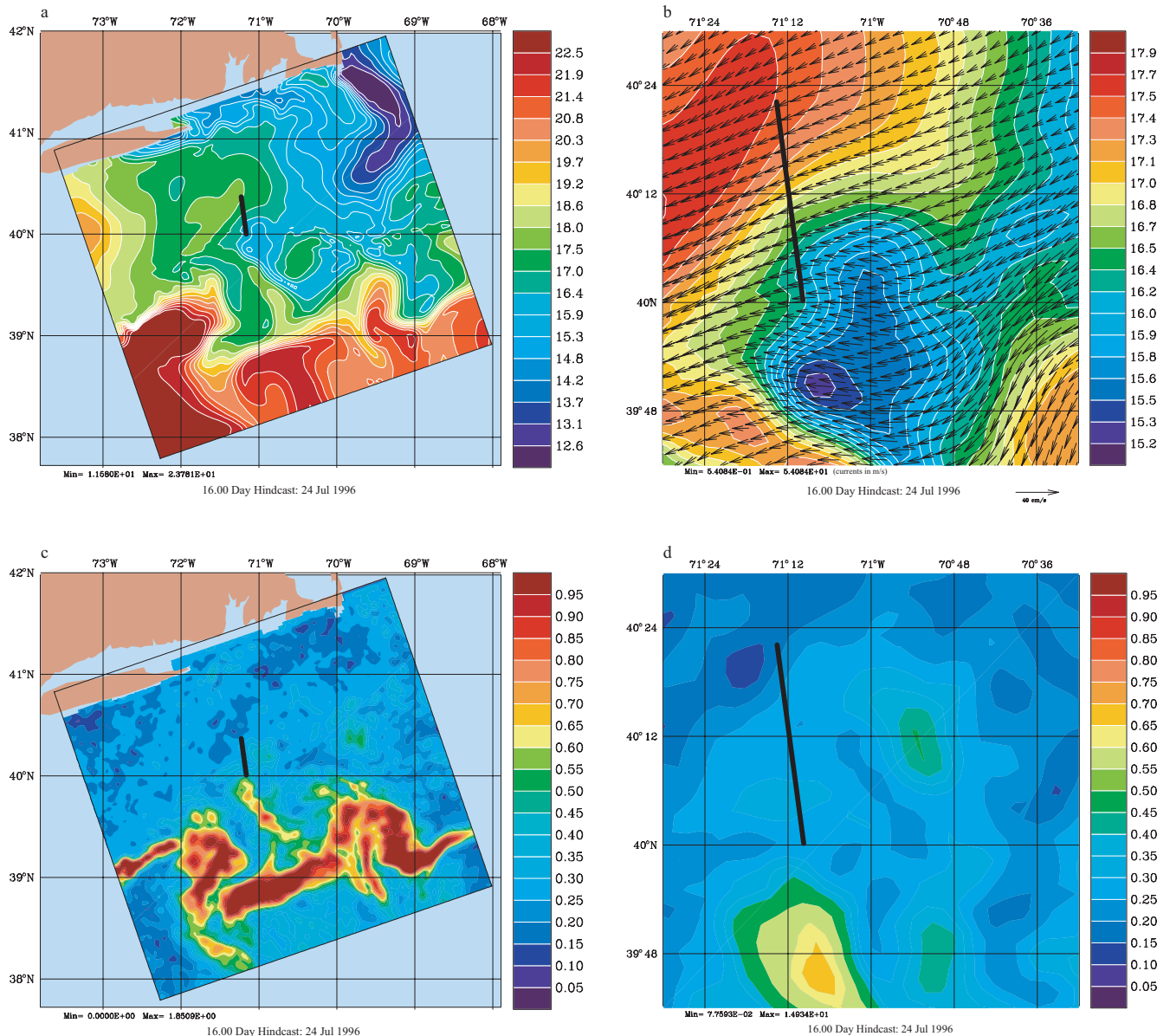

Figure 5. Simulated temperature field at 10-m depth (5a–b) over the MAB shelfbreak front region and its error standard deviation estimate (5c–d), as computed by Harvard Ocean Prediction System (HOPS) and Error Subspace Statistical Estimation (ESSE). Panels (5a, 5c) cover the whole 387 km by 360 km numerical ocean domain, at 3-km resolution in the horizontal. Panels (5b, 5d) are zooms over the PRIMER acoustic domain (89 km by 85 km), overlaid with horizontal velocities u_h on Panel 5b. The ensemble ESSE simulation starts on July 8, 1996, from historical and feature model data and an error covariance estimate. The plotted fields are 16-day hindcasts for July 24. No *in situ* data were available in the acoustic region during these first 16 days. The position of the main acoustic section, from (40.002°N, 71.163°W) to (40.368°N, 71.226°W) is also shown, on each panel.

Time is fixed and an acoustic broadband Transmission Loss (TL) field is computed for each ocean realization of the ESSE ensemble, using the coupled normal-mode model of Chiu et al. (1996). The 450-Hz sound source is at 300-m depth, near the deepest point on the slope. The mean and standard deviation of the coupled physical-acoustical fields along the section are shown in Figure 6. The mean sound speed field is characteristic of the shelfbreak front in summer (tilted front and surface thermocline). The TL field shows the sound attenuation in the surface mixed layer over the shelf and the funneling of sound in the sub-surface duct (colder shelf waters). The largest error standard deviations in the sound speed on July 26 are in the core of the front (30–40-m depth and range of 2–7 km), along the tilted front, and in the surface thermocline on the shelf. The largest error standard deviations in TL are close to the source, near the foot of the front and on the shelf. At the receiver vertical line array near 41-km range, they vary from 2 to 3 dB.
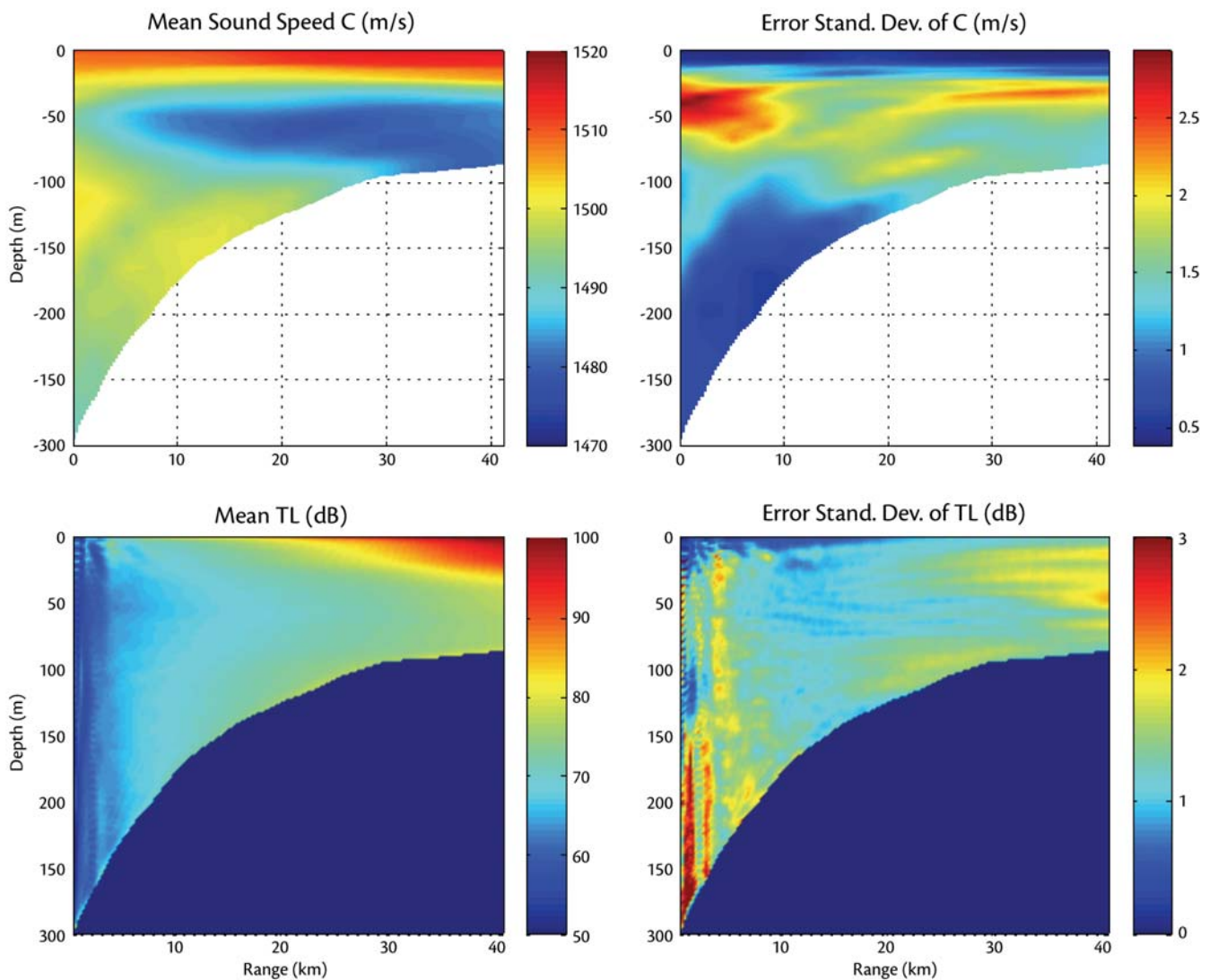

Figure 6. Cross-sections on July 26, 1996 in the mean and error standard deviation of the sound-speed (top) and broadband TL (bottom), as estimated by ESSE along the main acoustic section (western section of the PRIMER experiment). Note that if we also model uncertainties in the bottom attenuation coefficient (not shown), we find that the mean TL remains similar, but the TL error standard deviation increases substantially in the surface mixed layer above the shelf. This is a result of the uncertain attenuation at each acoustic bottom bounce and shows the importance of seabed uncertainties in acoustic predictions.

Physical-Biogeochemical Uncertainties in Massachusetts Bay

Uncertainty predictions are carried out for Massachusetts Bay's ecosystem, using HOPS, ESSE, and physical and biological data collected during summer 1998 (Beskitepe et al., 2003). The physics hindcast is initialized for August 20, 1998. Biogeochemical fields are also initialized for that period. The initial physical-biogeochemical covariance is estimated in two steps. Vertical EOFs of profiles of temperature (T), salinity (S), chlorophyll a (Chl), nitrate (NO_3), and ammonium (NH_4) are first multiplied with dominant eigenmodes of horizontal correlation functions, to lead to a three-dimensional eigen-decomposition of the T, S, Chl, NO_3, and NH_4 covariance matrix. This "observed" decomposition is then used to perturb the initial state and estimate the "non-observed" uncertainty by dynamical model integration. ESSE is then started from this dynamically adjusted error subspace on August 25. A hindcast of 600 perturbed runs, forced with physical stochastic noise, is then carried out for September 2, 1998.

The hindcast is illustrated on Figure 7 by uncertainties of Chl at 20-m depth (around the sub-surface Chl maxima). Shown are the mean Chl at 20 m, its error standard deviation estimate, and eight Chl histograms (PDF estimates) at various locations. Such PDF estimates fully characterize uncertainty. Mean amplitudes (top right) are largest along the coastline in Cape Cod Bay, in response to wind-driven upwelling, and south of Stellwagen Bank (marker 5) that is an accumulation region where whales are often found in late summer. Uncertainties (top right) are largest in the center and

mouth of the Bay and near recent coastal upwellings. Near Stellwagen Bank, maximum uncertainties are more at the edges than at the peaks of the Chl maxima.

This location is due to uncertainties in the burgeoning fall blooms and advective features (stronger currents are also along these edges). Looking at the PDF estimates, PDFs 1, 5, 6, 7, and 8 are steeper than a Gaussian of identical standard deviation, while PDFs 3 and 4 are closer to a Gaussian. PDF 2 is a bit flatter than a Gaussian because it combines two PDF peaks from nearby locations (south: lower Chl in a Gulf of Maine inflow; north: higher Chl in the eddy field of the coastal current). PDFs 5 and 7 are skewed towards lower Chl values because they are near the low Chl jet exiting Massachusetts Bay from the center of Cape Cod Bay. PDF 8, east of Cape Cod, is skewed towards positive values because it is near the high Chl content of the Gulf of Maine coastal current, flowing in and out of Massachusetts Bay.

CONCLUSIONS

The computational aspects of data-driven modeling and prediction of uncertainties were outlined and exemplified by

regional interdisciplinary applications. The representation, attribution and propagation of four-dimensional oceanic uncertainties presents many interest-

These (research) efforts include ocean observation campaigns dedicated to uncertainty modeling, interdisciplinary data assimilation, ocean stochastic modeling, new computational methods, adaptive modeling, and adaptive sampling research.

ing challenges and requires increased theoretical and applied research efforts. These efforts include ocean observation campaigns dedicated to uncertainty modeling, interdisciplinary data assimilation, ocean stochastic modeling, new computational methods, adaptive modeling, and adaptive sampling research.
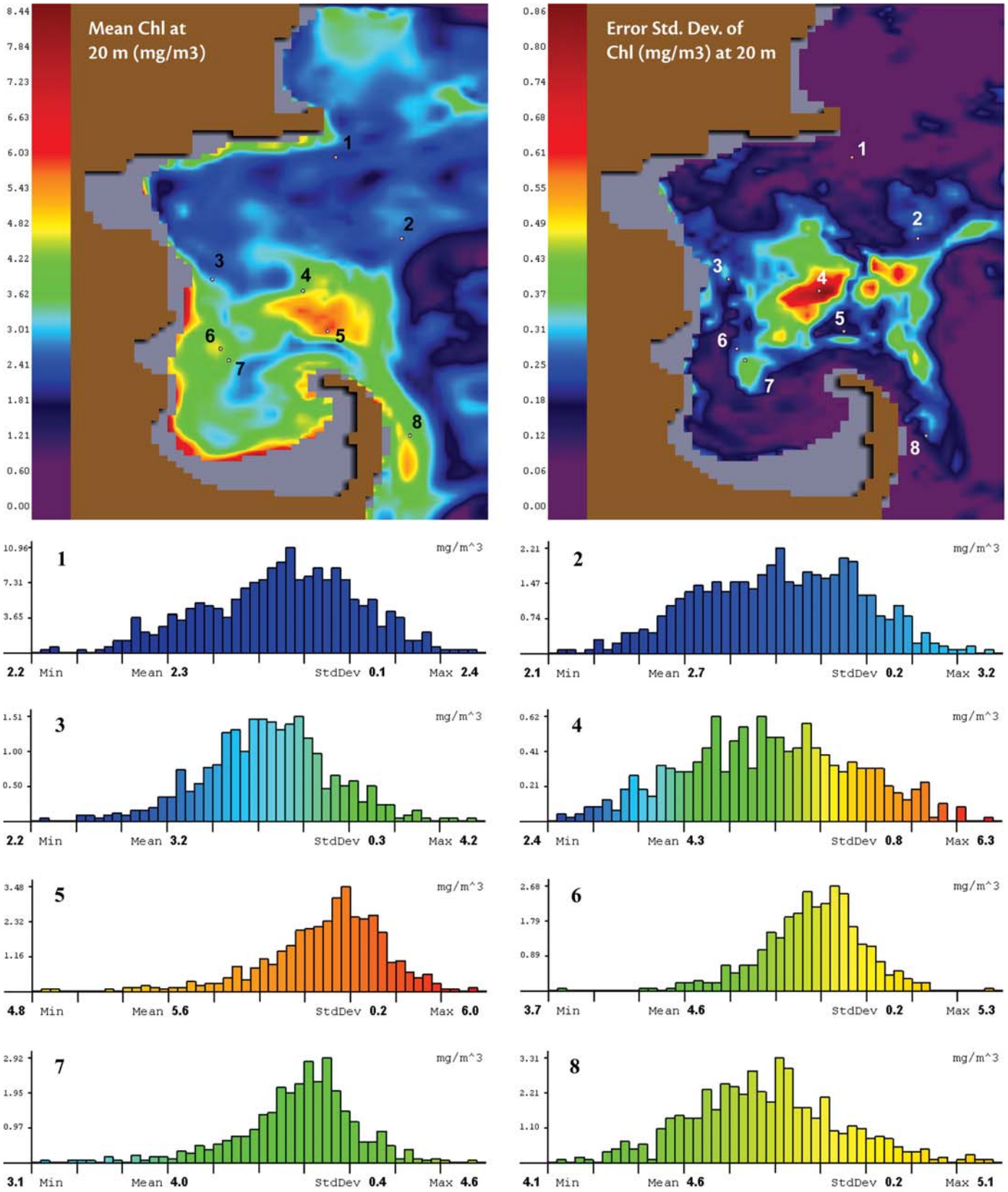
Figure 7. Chlorophyll *a* (Chl) mean and uncertainties at 20-m depth in the Mass Bay region on September 2, 1998, as hindcast by 600 Error Subspace Statistical Estimation (ESSE) ensemble members. ESSE was initialized on August 25, 1998. (Top left/right) Mean/Error Standard Deviation of Chl. (Bottom) Eight PDF estimates (normalized histograms, numbered 1 to 8) corresponding to the eight marked locations on the horizontal maps. Bars on the histograms are colored according to the center Chl value. The minimum, mean, standard deviation, and maximum values are given on each histogram (illustration by R.G. Hero, University of California, Santa Cruz).

REFERENCES

Baumgartner, M.F., and S.P. Anderson. 1999. Evaluation of regional numerical weather prediction model surface fields over the Middle Atlantic Bight. *Journal of Geophysical Research* 104(18):141–158.

Bennett, A.F. 1992. *Inverse Methods in Physical Oceanography*. Series by Cambridge Monographs on Mechanics and Applied Mathematics. Cambridge University Press, United Kingdom.

Besiktepe, S.T., P.F.J. Lermusiaux, and A.R. Robinson. 2003. Coupled physical and biogeochemical data-driven simulations of Massachusetts Bay in late summer: Real-time and post-cruise data assimilation. Special issue on "The use of DA in coupled hydrodynamic, ecological and bio-geo-chemical models of the oceans," M. Gregoire, P. Brasseur, and P.F.J. Lermusiaux, eds. *Journal of Marine Systems* 40:171–212.

Bishop, C.H., B.J. Etherton, and S.J. Majumdar. 2001. Adaptive sampling with the Ensemble Transform Kalman Filter: Part I, Theoretical Aspects. *Monthly Weather Review* 129:420–436.

Chiu, C.-S., J.H. Miller, and J.F. Lynch. 1996. Forward coupled-mode propagation modeling for coastal acoustic tomography. *Journal of the Acoustical Society of America* 99(2):793–80.

Dickey, T. 2003. Emerging ocean observations for interdisciplinary data assimilation systems. *Journal of Marine Systems* 40–41:5–48

Egbert, G.D., A. Bennett, and M.G.G. Foreman. 1994. *TOPEX/POSEIDON* tides estimated using a global inverse model. *Journal of Geophysical Research* 99(C12):24,821–24,852.

Ehrendorfer, M. 1997. Predicting the uncertainty of numerical weather forecasts: A review. *Meteorologische Zeitschrift* 6(4):147–183.

Evensen, G. 1994. Inverse methods and data assimilation in nonlinear ocean models. *Physica D* 77:108–129.

Gardiner, C.W. 1983. *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*. 2nd edition. Springer Verlag, New York, NY, 442 pp.

Griffiths, G., S. Felding, and H.S.J. Roe. 2002. Biological-physical-acoustical interactions. In: *The Sea: Biological-Physical Interactions in the Sea*, Volume 12, A.P. Robinson, J.J. McCarthy, and B.J. Rothschild, eds. John Wiley and Sons, New York, NY.

HOPS (Harvard Ocean Prediction System.) *Harvard Ocean Prediction System*. [Online] Available at: http://oceans.deas.harvard.edu/HOPS/HOPS.html [accessed 02-19-06].

Hofmann, E.E., and M.A.M. Friedrichs. 2002. Predictive modeling for marine ecosystems. In: *The Sea: Biological-Physical Interactions in the Sea*, Volume 12, A.R. Robinson, J.J. McCarthy, and B.J. Rothschild, eds. John Wiley and Sons, New York, NY, 537–565.

Houtekamer, P.L., H. L. Mitchell, G. Pellerin, M. Buehner, M. Charron, L. Spacek, and B. Hansen. 2005. Atmospheric data assimilation with an ensemble Kalman filter: Results with real observations. *Monthly Weather Review* 133:604–620.

Jazwinski, A.H. 1970. *Stochastic Processes and Filtering Theory*. Academic Press, San Diego, CA.

Kalnay, E. 2003. *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, MA, 341 pp.

Kuperman, W.A. 2004. *Underwater Acoustics.* In: *Encyclopedia of Physical Science and Technology*.. Elsevier Science Ltd., City, State, 317–338.

Lekien, F., C. Coulliette, A.J. Mariano, E.H. Ryan, L.K. Shay, G. Haller, and J.E. Marsden. 2005. Pollution release tied to invariant manifolds: A case study for the coast of Florida. *Physica D* 210(1–2):1–20.

Lermusiaux, P.F.J. 1999. Estimation and study of mesoscale variability in the Strait of Sicily. *Dynamics of Atmospheres and Oceans* 29:255–303.

Lermusiaux, P.F.J. 2001. Evolving the subspace of the three-dimensional multiscale ocean variability: Massachusetts Bay. *Journal of Marine Systems* 29(1-4; special issue):385–422.

Lermusiaux, P.F.J., and A.R. Robinson. 2001. Features of dominant mesoscale variability, circulation patterns and dynamics in the Strait of Sicily. *Deep Sea Research* 48(9):1,953–1,997.

Lermusiaux, P.F.J. 2002. On the mapping of multivariate geophysical fields: Sensitivity to size, scales and dynamics. *Journal of Atmospheric and Oceanic Technology* 19:1,602–1,637.

Lermusiaux, P.F.J., and C.-S. Chiu. 2002. Four-dimensional data assimilation for coupled physical-acoustical fields. In: *Acoustic Variability*, SACLANTCEN Conference Proceedings, N.G. Pace and F.B. Jensen, eds. Kluwer Academic Press, New York, NY, 417–424.

Lermusiaux, P.F.J., A.R. Robinson, P.J. Haley, and W.G. Leslie. 2002. Advanced interdisciplinary data assimilation: Filtering and smoothing via Error Subspace Statistical Estimation. In: *The OCEANS 2002 MTS/IEEE*. Holland Publications, Escondido, CA, 795–802.

Lermusiaux, P.F.J, C. Evangelinos, R. Tian, P.J. Haley, J.J. McCarthy, N.M. Patrikalakis, A.R. Robinson, and H. Schmidt. 2004. Adaptive coupled physical and biogeochemical ocean predictions: A Conceptual Basis. In: *Computational Science—ICCS 2004*, F. Darema, ed. Lecture Notes in Computer Science 3038. Springer Verlag, New York, NY, 685–692.

Linder, C.A., and G.G. Gawarkiewicz. In press. Climatological estimation of environmental uncertainty over the Middle Atlantic Bight Shelf and Slope. *IEEE Journal of Oceanic Engineering*.

Lynch, J.F., G.C. Gawarkiewicz, C.-S. Chiu, R. Pickart, J.H. Miller, K.B. Smith, A.R. Robinson, K.H. Brink, R. Beardsley, B. Sperry, and G. Potty. 1997. 'Shelfbreak PRIMER—An integrated acoustic and oceanographic field study in the Middle Atlantic Bight. In: Proceedings of the International Conference on Shallow Water Acoustics, April, 21–25, Beijing, China, 205–212.

Lynch, J., A. Newhall, B. Sperry, G. Gawarkiewicz, P. Tyack, and C.-S. Chiu. 2001. Spatial and temporal variations in acoustic propagation characteristics at the New England shelfbreak front. *IEEE Journal of Oceanic Engineering* 28:129–150.

Malanotte-Rizzoli, P., ed. 1996. *Modern Approaches to Data Assimilation in Ocean Modeling*. Elsevier, Amsterdam, The Netherlands.

Miller, R.N., and L.L. Ehret. 2002. Ensemble generation for models of multimodal systems. *Monthly Weather Review* 130(9):2,313–2,333.

Miller, R.N., E.F. Carter, and S.L. Blue. 1999. Data assimilation into nonlinear stochastic models. *Tellus* 51A:167–194.

Molteni, F., R. Buizza, T.N. Palmer, and T. Petroliagis. 1996. The new ECMWF ensemble prediction system: methodology and validation. *Quarterly Journal of the Royal Meteorological Society* 122:73–119.

Mooers, C.N.K., ed. 1999. *Coastal Ocean Prediction*. AGU Coastal and Estuarine Studies Series. American Geophysical Union, Washington, DC, 523 pp.

Murphy, J.M., D.M. Sexton, D.N. Barnett, G.S. Jones, M.J. Webb, M. Collins, and D.A. Stainforth. 2004. Quantification of modeling uncertainties in a large ensemble of climate change simulations. *Nature* 430: 768–772.

Nihoul, J.C.J., and S. Djenidi. 1998. Coupled physical, chemical and biological models. In: *The Sea: Processes and Methods*, Volume 10, K.H. Brink, and A.R. Robinson, eds. John Wiley and Sons, New York, NY.

ONR (Office of Naval Research). 2001. Uncertainty DRI. [Online] Available at: http://www.onr.navy.mil/sci_tech/ocean/321_sensing/cuwg/proceedings.asp [accessed 02-19-06].

Pedlosky, J. 1987. *Geophysical Fluid Dynamics*. Second edition, Springer Verlag, New York, NY, 710 pp.

Pinardi, N., and J. Woods, eds. 2002. *Ocean Forecasting: Conceptual Basis and Applications*. Springer Verlag, New York, NY, 472 pp.

Robinson, A.R., P.F.J. Lermusiaux, and N.Q. Sloan. 1998. Data assimilation. *In The Sea: Processes and Methods*, Volume 10, K.H. Brink, and A.R. Robinson, eds. John Wiley and Sons, New York, NY, 541–594.

Robinson, A.R., and P.F.J. Lermusiaux. 2002. Data assimilation for modeling and predicting coupled physical-biological interactions in the sea. In: *The Sea: Biological-Physical Interactions in the Sea*, Volume 12, A.P. Robinson, J.J. McCarthy, and B.J. Rothschild, eds. John Wiley and Sons, New York, NY, 475–536.

Robinson, A.R., and P.F.J. Lermusiaux. 2004. Prediction systems with data assimilation for coupled ocean science and ocean acoustics. In: *Proceedings of the Sixth International Conference on Theoretical and Computational Acoustics*, A. Tolstoy, et al., eds. World Scientific Publishing Company Inc., Hackensack, NJ, 325–342.

Spitz, Y.H., J.R. Moisan, and M.R. Abbott. 2001. Configuring an ecosystem model using data from the Bermuda-Atlantic Time Series (BATS). *Deep-Sea Research II* 48:1,733–1,768.

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, and M.R. Allen. 2005. Uncertainty in predictions of the climate response to rising levels of greenhouse gases. *Nature* 433:403–406.

Szunyogh, I., E.J. Kostelich., G. Gyarmati, D.J. Patil, B.R. Hunt, E. Kalnay, E. Ott, and J.A. Yorke. 2005. Assessing a local ensemble Kalman filter: Perfect model experiments with the National Centers for Environmental Prediction global model. *Tellus* 57A:528–545.

Toth, Z., and E. Kalnay. 1993. Ensemble forecasting at NMC: The generation of perturbations. *Bulletin of the American Meteorological Society* 74:2,317–2,330.

Whitaker, J. S., G.P. Compo, X. Wei, and T.M. Hamill. 2004. Reanalysis without radiosondes using ensemble data assimilation. *Monthly Weather Review* 132:1,190–1,200.

Wunsch, C. 1996. *The Ocean Circulation Inverse Problem*. Cambridge University Press, MA.