Towards Bayesian Ocean Physical-Biogeochemical-Acidification Prediction and Learning Systems for Massachusetts Bay

P. J. Haley, Jr.[a], A. Gupta[a], C. Mirabito[a], P. F. J. Lermusiaux[a,†]
[a] *Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA*
[†]Corresponding Author: pierrel@mit.edu

***Abstract*—Better quantitative understanding and accurate data-assimilative predictions of the three-dimensional and time-dependent ocean acidification (OA) processes in coastal regions is urgently needed for the protection and sustainable utilization of ocean resources. In this paper, we extend and showcase the use of our MIT-MSEAS systems for high-resolution coupled physical-biogeochemical-acidification simulations and Bayesian learning of OA models in Massachusetts Bay, starting with simple empirical and equilibrium OA models. Simulations are shown to have reasonable skill when compared to available in situ and remote data. The impacts of wind forcing, internal tides, and solitary waves on water transports and mixing, and OA fields, are explored. Strong wind events are shown to reset circulations and the OA state in the Bay. Internal tides increase vertical mixing of waters in the shallow regions. Solitary waves propagating off Stellwagen Bank coupled with lateral turbulent mixing provide a pathway for exchange of surface and deep waters. Both of these effects are shown to impact biological activity and OA. A mechanism for the creation of multiple subsurface chlorophyll maxima is presented, involving wind-induced upwelling, internal tides, and advection of near surface fields. Due to the measurement sparsity and limited understanding of complex OA processes, the state variables and parameterizations of OA models are very uncertain. We thus present a proof-of-concept study to simultaneously learn and estimate the OA state variables and model parameterizations from sparse observations using our novel dynamics-based Bayesian learning framework for high-dimensional and multi-disciplinary estimation. Results are found to be encouraging for more realistic OA model learning.**

***Index Terms*—coupled physical-biogeochemical ocean modeling, ocean acidification, Massachusetts Bay, skill evaluation, uncertainty quantification, Bayesian learning, Gaussian mixture models**

I. Introduction

Monitoring, quantifying, and predicting the three-dimensional and time-dependent ocean acidification processes, from the atmospheric exchanges and river discharges to the ocean interior, and over days to decades, remains a fascinating observational, theoretical, and modeling challenge. This challenge is the long-term driver of our "Bayesian Intelligent Ocean Modeling and Acidification Prediction Systems" (BIOMAPS) research. Ocean acidification (OA), or the progressive decrease in pH of seawater, is caused primarily by excess atmospheric CO_2 and is linked to climate change [1]–[3]. Its chemical perturbations are expected to be larger in coastal regions than on global average [4], [5]. In the Gulf of Maine and Massachusetts Bay regions, the shellfish growth and reproduction are affected by coastal acidification, with negative impacts on crustaceans (lobsters, crabs) and both wild and farmed mollusks (scallops, oysters, clams, mussels), hence also on major industries and employment sources [6]. Improving the monitoring, modeling, and forecasting of regional OA is urgent.

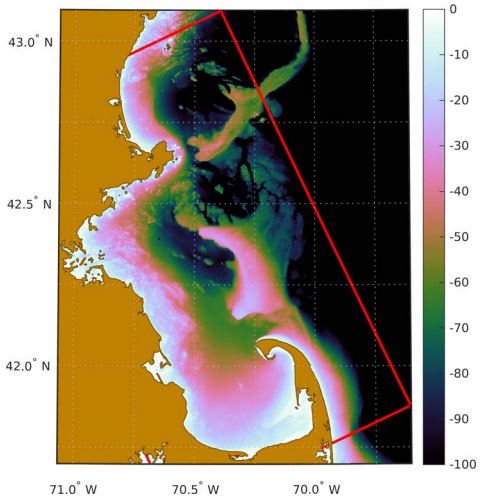


Fig. 1. MSEAS-PE 333 m resolution modeling domain (boundary shown in red) over bathymetry (m).

The overarching goal of our research is to develop and demonstrate principled Bayesian intelligent ocean modeling and acidification prediction systems that discriminate among and infer new OA models, rigorously learning from data-model misfits and accounting for uncertainties, so as to better monitor, predict, and characterize OA over time scales of days to months in the Massachusetts Bay and Stellwagen Bank regions.

This paper is organized as follows. In Section II, we set up our modeling system for coupled physical-biogeochemical simulations for Massachusetts Bay. We examine the skill in our simulations by comparison to available in situ and remote data. We then examine the impact of atmospheric and tidal forcings and interactions with topography on the physical and biogeochemical fields. In Section III, we lay out our overall strategy for coupling to OA models. We also examine the first steps using coupled empirical and equilibrium OA models with our coupled physical-biogeochemical simulations and briefly explore the impacts of forcings and interactions on the OA fields. In Section IV, we use use a PDE-based Bayesian learning framework to showcase a learning experiment that simultaneously infers augmented state variables and parameters of empirical relationships between an OA field and the

Global Oceans 2020: Singapore - U.S. Gulf Coast | 978-1-7281-5446-6/20/$31.00 ©2020 IEEE | DOI: 10.1109/IEEECONF38699.2020.9389210

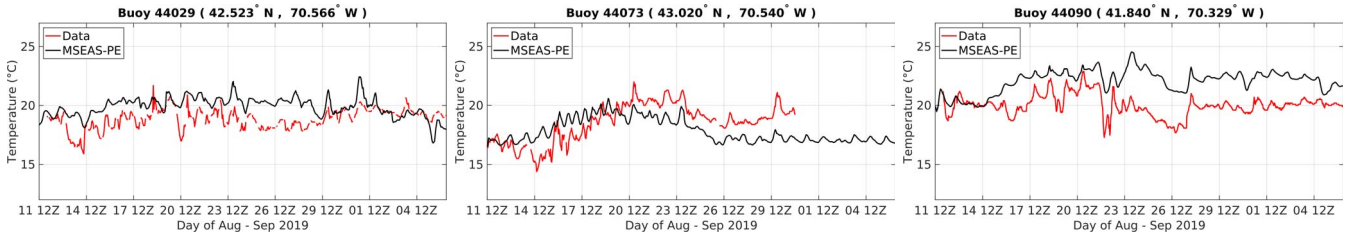


Fig. 2. Comparison of MSEAS-PE simulated hindcast temperature (°C; black lines) and independent NDBC buoy temperatures (red lines) between August 11 12Z and September 7 0Z, 2019, at three buoys: 44029 (off Gloucester, MA; left), 44073 (near Isles of Shoals; center; stopped recording on August 31), and 44090 (Cape Cod Bay; right).

non-hydrostatic physics of flows past a seamount idealizing Stellwagen Bank. The principled Bayesian learning [7]–[10] that we employed combines the Dynamically Orthogonal (DO) methodology [11], [12] for reduced-dimension stochastic evolution with a Gaussian mixture model (GMM)-DO filtering scheme [13], [14]. Finally, we present our conclusions in Section V.

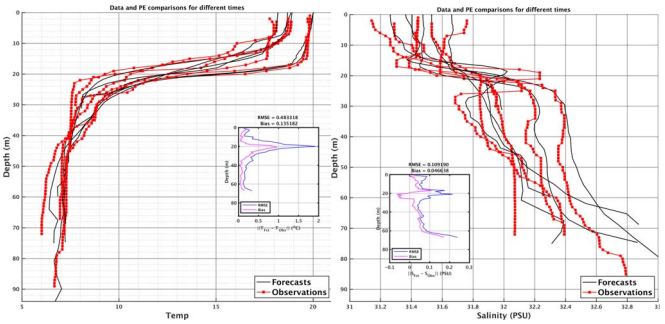
II. Ocean Physical-Biogeochemical Modeling

To simulate the Mass. Bay ocean dynamics, we employed the non-linear free surface primitive equation model of our Multidisciplinary Simulation, Estimation, and Assimilation System (MSEAS) [15], [16], using tidal and atmospheric forcing. This modeling system has been used around the world's oceans [16]–[26]. Applications include monitoring [27]; real-time acoustic predictions and DA [28]–[31]; environmental predictions and management [32]–[34]; relocatable rapid response [35], [36]; path planning for autonomous vehicles [37]–[40]; and, adaptive sampling [41]–[43]. MSEAS has been tested and validated in many real-time forecasting exercises [16], [17], [21], [22], [25], [26], [30], [44]–[47] including Mass. Bay [32], [48], [49]. Using ensemble methods [41], we issued large-ensemble forecasts at high-resolution for 3D underwater-GPS exercises [50]. MSEAS also includes finite-element codes for non-hydrostatic dynamics [51], [52] and a stochastic modeling framework [53], [54].

1) Physical Ocean Dynamics: We set up our MSEAS ocean modeling system for Massachusetts Bay and completed analyses and forecasts. The present modeling domain (Fig. 1) off the northeast US coast has a 333 m horizontal resolution and 100 vertical levels with optimized level depths (e.g., in deeper water, higher resolution near the surface or large vertical derivatives, while at coasts, evenly spaced to minimize vertical numerical CFL-restrictions). The bathymetry was obtained from the 3 arc second USGS Gulf of Maine digital elevation model [55]. The sub-tidal initial and boundary conditions were downscaled from 1/12-degree Hybrid Coordinate Ocean Model (HYCOM) analyses [56], using our optimization for our higher resolution coastlines and bathymetry [57]. Importantly, local corrections were made using feature models and synoptic CTDs of opportunity. Tidal forcing was computed from the high resolution TPXO8-Atlas from OSU [58], [59], by reprocessing for our higher resolution bathymetry/coastline and quadratic bottom drag (a nonlinear extension of [60]). The atmospheric forcing consisted of hourly analyses/forecasts of wind stresses, net heat flux, and surface freshwater flux from the 3 km North American Mesoscale Forecast System (NAM) [61].

We validated the MSEAS ocean simulation against independent data, including NOAA NDBC buoy data [62]. In Fig. 2, we show the comparison of near-surface temperature data from the buoys (red curves) to the MSEAS-PE simulated temperature interpolated to the buoy positions/depths (black curves). Given the uncertainties in the 3 km atmospheric forcing and those arising from the unresolved processes in the downscaled IC/BCs, we do not expect tight matching in these point comparisons. Nonetheless, we find that the MSEAS-PE produces similar daily cycle excursions (both in amplitude and frequency). We also see that the general trends and events do align well and that the mismatches (typically between 0 and 2 °C) are what we expect for the given uncertainties. We also note that by September, for the buoy in central Mass. Bay, the accuracy becomes very good.

Having employed the CTD profiles in the downscaled IC/BCs, they are no longer fully independent validation data. However, we can still employ them in limited skill comparisons. In Fig. 3, we compare the CTD profiles to the simulated temperature and salinity fields at the time the data were observed. We see that overall the corrections to the ICs are maintained by the simulation, although there is some drift appearing in the deep salinity. Comparisons with SST and HF-radar data (not shown) also indicate acceptable simulations.



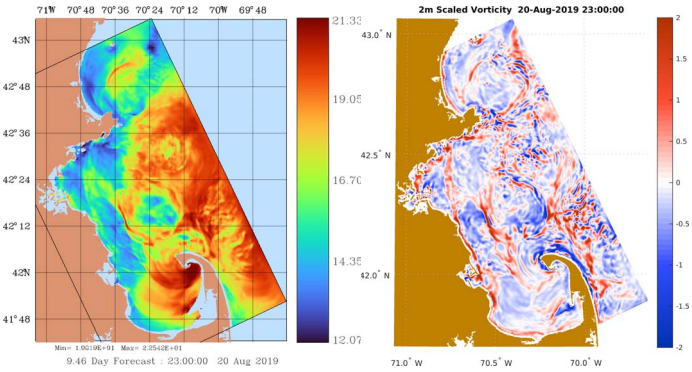
(a) Temperature (b) Salinity

Fig. 3. MSEAS forecast skill against NMFS CTD data.

The circulation in Mass. Bay [48] is commonly from north

to south and remotely driven from the Gulf of Maine coastal current and mean wind stress. However, it varies seasonally and in response to wind events.

During Aug.–Sep. 2019, several wind events modified the coastal circulation. Initially, the flow in the thermocline entered Mass. Bay from the north by Cape Ann and proceeded southward to the west of Stellwagen Bank then around Cape Cod Bay before it exited by Race Point. Following the wind event of August 24–26, the 30 m southward flow was displaced east of Stellwagen Bank (bypassing Mass. Bay). During Aug. 26–31, the 30 m flow reestablished itself in Mass. Bay. These overall conditions are modified by tides and resulting internal tides and solitary waves generated by the bathymetry, especially from Stellwagen Bank. The surface expressions of these waves can be clearly seen in the 10 m temperature (Fig. 4a) and 2 m vorticity (Fig. 4b) fields. These internal tides and solitary waves bring surface waters down in depth where turbulent mixing at depth can transport waters and material out of the wave. They also enhance the overall vertical turbulent mixing of material in the shallow regions [63]. The impacts of these transports on the biology will be examined in §II-2.



(a) 10 m Temperature (b) 2 m Vorticity

Fig. 4. MSEAS coupled physical-biogeochemical simulated fields in Massachusetts Bay show strong internal tides and topographic effects.

2) Biogeochemical Ocean Dynamics: Our biological model is a generalized advective-diffusive-reaction model for the lower trophic levels, described in [64]. For our Mass. Bay simulations we configured this generalized model in a 7 component system: nitrate (NO_3), ammonium (NH_4), chlorophyll (Chl), zooplankton (Z), detritus (D) and two phytoplankton (P) components tracking NO_3 and NH_4 uptake. This configuration and the parameters governing the biological reactions follow those in [32] which also simulated Mass. Bay during late summer.

The biological fields and parameters were initialized using in situ and historical data, and tuned to the August conditions. Historical a data for NO_3 and Chl were obtained from the World Ocean DataBase [65] and objectively analyzed to produce 3D fields for August and September (Fig. 5). The fields that were not observed or only partially observed (NH_4, Z, D, P) were dynamically computed using the reaction terms of our biogeochemical equations in a weak constraint form, with the parameters from [32]. This ensures bioegeochemical initial conditions that are dynamically-balanced.

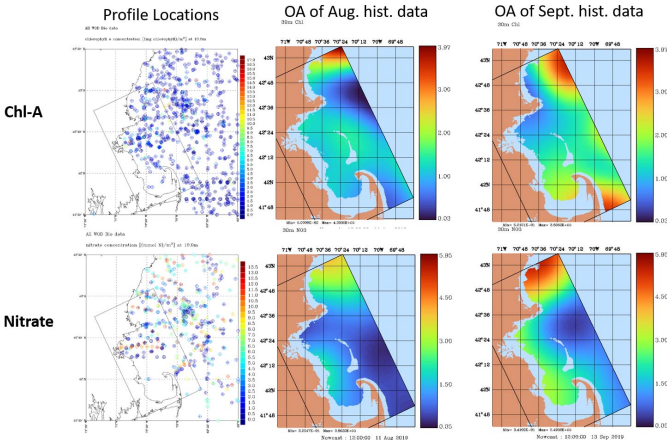
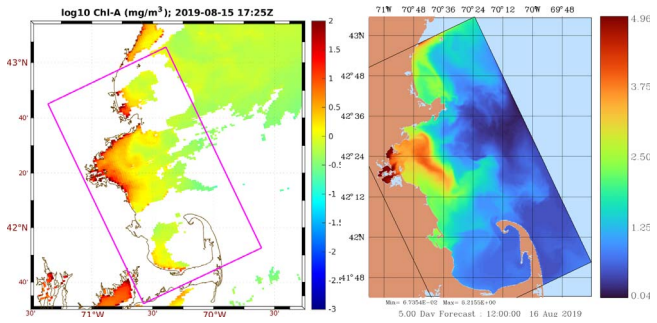


Fig. 5. Initial conditions for chlorophyll and nitrate are obtained from objective analyses of historical data (World Ocean DataBase).

Given the lack of synoptic biological data for our initialization, we do not expect our simulated biological fields to exactly track the 2019 conditions. However we can do a qualitative assessment for our skill. In Fig. 6, we use MODIS Aqua satellite Chl from Aug 15 to compare to our surface Chl field. We observe qualitative agreement with enhanced Chl near Boston Harbor and along the northern coast of Mass. Bay. We also observe an intrusion of elevated Chl descending southwards from northern edge of Mass. Bay into the interior of Mass. Bay. As expected, the exact details do not match but the qualitative agreement suggests that the biological model is behaving properly. Additional tuning of model parameters and initial conditions would be needed for improved agreement.



(a) MODIS Aqua Chl (b) MIT-MSEAS 0m Chl

Fig. 6. Evaluation of biological component of MSEAS coupled physical-biogeochemical simulated fields in Massachusetts Bay. We observe qualitative agreement between (a) MODIS Aqua satellite chlorophyll on Aug 15 and (b) surface chlorophyll from MIT-MSEAS simulation. Both show an intrusion of chlorophyll from the northern edge of Massachusetts Bay.

The high-resolution simulations clearly indicate significant biogeochemical responses to internal tides, wind forcing, and local circulations. To showcase these responses, maps of the main biological fields (Chl, Z, NO_3, NH_4 and D) at 10 m are presented in Fig. 7. Wind-induced upwelling has led to

Chl growth along the coast in the northern half of Mass. Bay and along the coast to the north and subsequent transport of elevated Chl into the middle of Mass. Bay. In response, we also see growth in Z, NH_4 and D concentrations. The effects of internal tides are especially visible in the ripples in the Chl, Z, and NH_4 fields in the northern half of Mass. Bay.

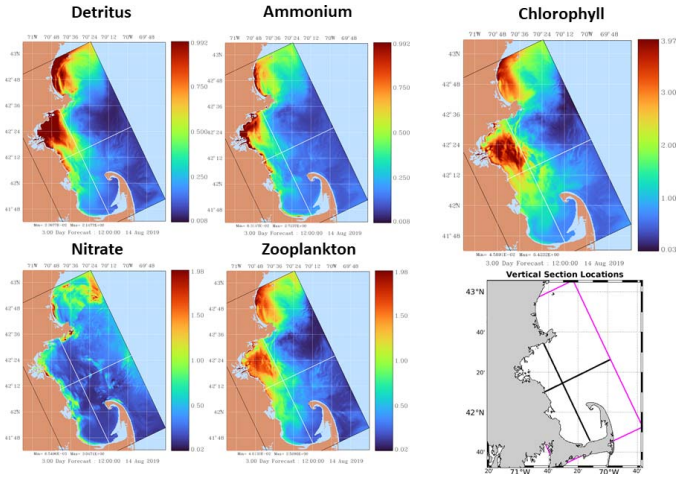


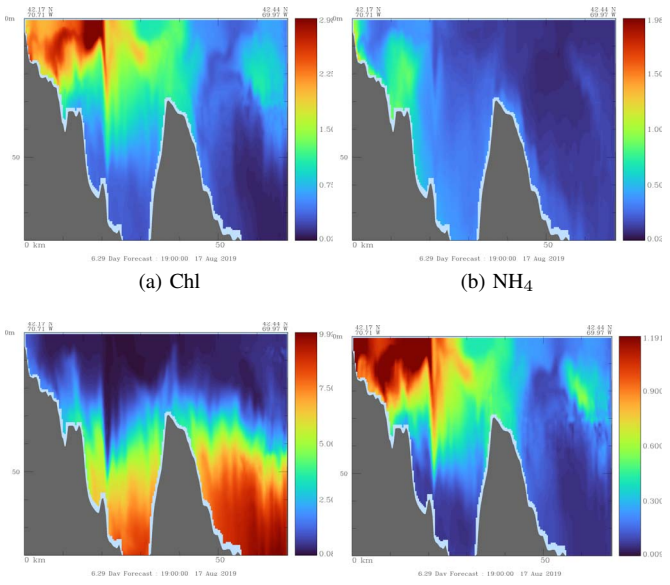
Fig. 7. MSEAS coupled physical-biogeochemical simulated fields in Massachusetts Bay. Significant biogeochemical responses to internal tides, wind forcing, and local circulations are visible in the elevated biological activity near the coasts and in the ripples in Chl, Z, and NH_4 in the northern half of Mass. Bay.

In Fig. 8, we look at an east-west cross-section across Mass. Bay, through the middle of Stellwagen Bank (location given in lower right of Fig. 7). Internal tides and solitary waves are generated along the western side of Stellwagen Bank and propagate westwards towards the coast. The effect of one such solitary wave is clearly visible in the Chl (Fig. 8a) and Z (Fig. 8d) fields. The solitary wave briefly brings the surface fields down to depth where turbulent mixing can bring them into contact with nutrients. On the longer term, the internal tides enhance the vertical mixing of deep nutrients upward, which can be best seen in the NH_4 field (Fig. 8b) above the 30 m ridge on the western side of the section.

Fig. 9 presents a north-south cross-section through the middle of Mass. Bay (location given in lower right of Fig. 7). Initially, winds to the northeast induce upwelling along the northern coast of Mass. Bay, feeding Chl, Z growth. Later, internal tides bring Chl into contact with deeper water, causing the Chl subsurface maximum to deepen. Z grows at this deeper Chl maximum and releases NH_4 at this depth, helping to sustain the deeper subsurface maximum. Subsequently in the near surface, water with enhanced Chl, Z, NH_4 is advected above the deep subsurface maximum leading to double maxima in the water column. In the middle of the section, the effects of the strong internal tides/solitary waves are again visible in Chl and Z.

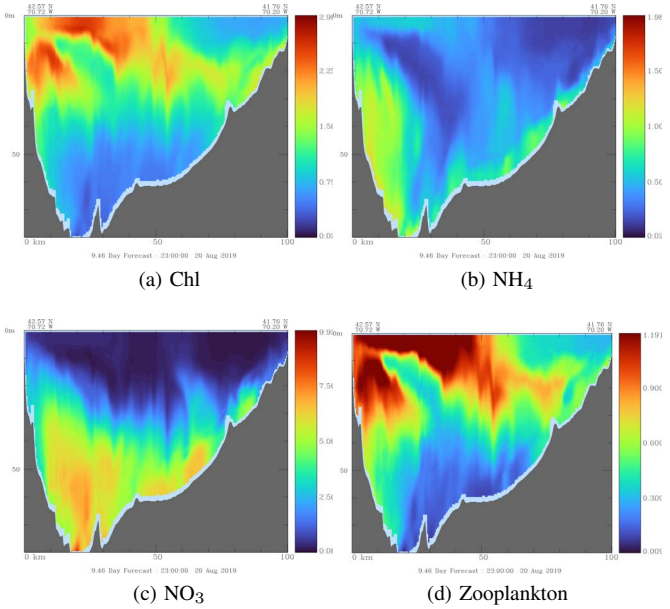
III. Ocean Acidification Modeling

Our long-term goal is to utilize and augment published regressive and partial-differential-equation models to generate



(a) Chl (b) NH_4 (c) NO_3 (d) Zooplankton

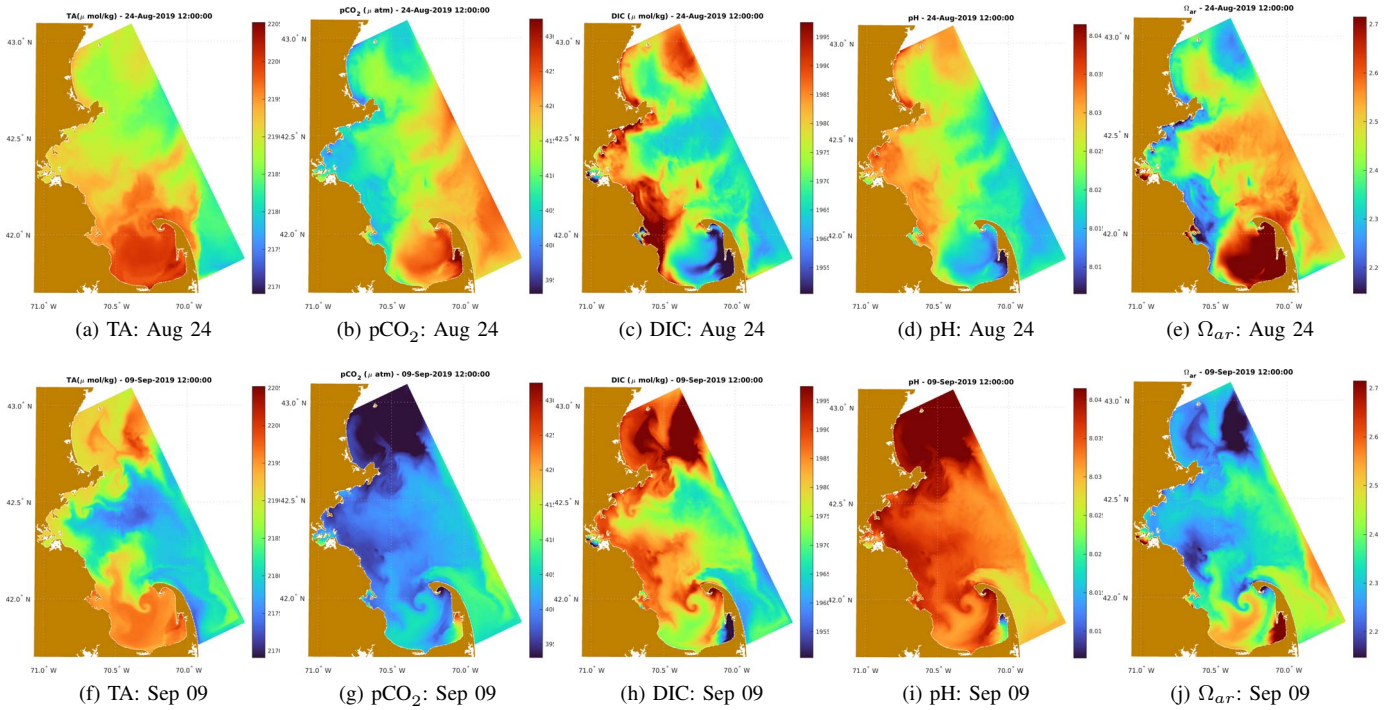
Fig. 8. MSEAS simulated impacts of internal tides on biology. Strong internal tides and solitary waves are generated off Stellwagen Bank.



(a) Chl (b) NH_4 (c) NO_3 (d) Zooplankton

Fig. 9. MSEAS simulated winds and tides off Cape Ann (northern end of section) combine to create deep subsurface maxima. Shallower subsurface maxima subsequently advected over deeper maxima.

initial conditions and forecasts for the state variables of the OA and CO_2 system in the Mass. Bay region. The OA carbonate regressive and differential models would then be coupled with our physical-biogeochemical models. OA models could be configured with different biological parameterizations and regressive models. Learning would then be needed to infer which carbonate variables are best modeled as state variables and which are best modeled via regressive or equilibrium models. Learning would also be used to estimate which



(a) TA: Aug 24 (b) pCO_2: Aug 24 (c) DIC: Aug 24 (d) pH: Aug 24 (e) Ω_{ar}: Aug 24

(f) TA: Sep 09 (g) pCO_2: Sep 09 (h) DIC: Sep 09 (i) pH: Sep 09 (j) Ω_{ar}: Sep 09

Fig. 10. Ocean acidification (OA) fields diagnosed from our coupled physical-biogeochemical MSEAS simulations and empirical and equilibrium OA models. (a)-(e): OA fields during a typical ocean state (Aug. 24). (f)-(j): OA fields one day after the passage of the remnants of Hurricane Dorian (Sep. 9).

reaction functions and parameters best fit the region.

We start by using a combination of empirical and equilibrium models to diagnose the surface ocean acidification fields from the MSEAS simulated physical and biogeochemical fields. Using empirical formulae for the Gulf of Maine and the North American east coast [66]–[68], we estimate the (a) surface total alkalinity (TA) from the surface salinity and (b) surface partial pressure of carbon dioxide (pCO_2) from the surface temperature, salinity and chlorophyll. With these fields, we then use the equilibrium model CO2SYS [69] to estimate the surface pH, dissolved inorganic carbon (DIC), and saturation state of argonite (Ω_{ar}).

We briefly examine the diagnosed OA fields in Fig. 10. The fields in a typical ocean state (Aug. 24) are presented in Fig. 10a–10e. During this period, there is some westward transport of water from the Gulf of Maine into northern Mass. Bay which seems to be reflected in DIC (Fig. 10c) and Ω_{ar} (Fig. 10e). We also note a general inverse relation between pH and Ω_{ar} (as expected). The effects of internal tides are reasonable visible in pCO_2 (Fig. 10b), DIC (Fig. 10c) and Ω_{ar} (Fig. 10e) near the coast around 42.25 N. One day after the remnants of Hurricane Dorian pass through (Sep. 9), the OA state is changed (Fig. 10f–10j). Presumably in response to the deep wind mixing, we simulate an increase in pH (Fig. 10i) and a degradation of Ω_{ar} (Fig. 10j). Throughout these events the argonite remains in a supersaturated state ($\Omega_{ar} > 1$). In addition to resetting the coastal circulations [48], our results thus show that strong wind events also reset the OA state in the Bay.

IV. Data-driven modeling and learning of OA models

The OA models have many sources of uncertainties [70], [71]. Due to the semi-empirical methodology of developing OA models, there is uncertainty associated with the parameters, functional forms, and the level of complexity of OA models. A set of parameter values may only be valid in certain ocean regions, or may be subject to seasonal variability. Observations are integral to the formation of these models, but are in general used for data fitting in order to find appropriate parameter values or functional forms of these models in offline mode, or for interpolation/extrapolation of data using the models.

Hence, in this work, we showcase a simulated learning experiment which employs a novel PDE-based Bayesian learning framework that simultaneously infers the augmented state variables and the unknown parameters. This method could further be extended to learn unknown functional forms as shown in [9]. The experimental setup consists of a 2-dimensional domain (Fig. 11) with a seamount representing an idealized sill or strait, inspired by the Stellwagen Bank. For the ocean physical fields, we numerically integrate the stochastic nonhydrostatic Navier-Stokes equations with salinity (S) as a passive tracer, and consider uncertainty in the initial conditions (more details can be found in [9]). For the OA fields, the total alkalinity (TA) is assumed to be linearly dependent on the salinity, given by $TA(\boldsymbol{x}, t; \omega) = TA_o + b(\omega)S(\boldsymbol{x}, t; \omega)$, where this system belongs to a domain $\boldsymbol{x} \in \mathcal{D}$ at time t, and ω is a realization index belonging to a measurable space Ω. We

consider the slope of the linear relationship ($b(\omega)$) to be an uncertain parameter to be inferred from the data.

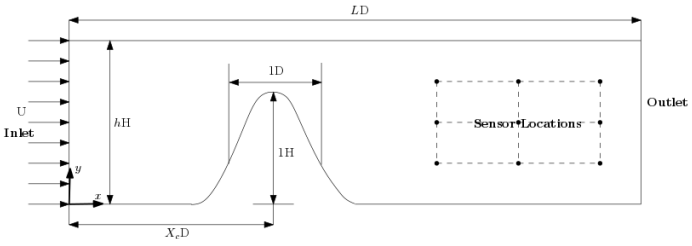


Fig. 11. Two-dimensional spatial domain of the nonhydrostatic flow past a seamount. All lengths are given in terms of length scales (D & H) of the seamount, described by height $H \exp^{(x-X_c)^2}$. Observation locations are marked by dots, downstream of the seamount. Adapted from [9].

The parameter values associated with this experiment are provided in Table I. The dynamical ocean physics and uncertain OA models are implemented in a finite volume framework [53], using the Dynamically Orthogonal (DO) methodology [11], [12], [72] for efficient reduced-dimension uncertainty propagation. To perform inference of the augmented state variables and parameters, we make use of a filtering algorithm developed by combining the DO method with a Gaussian mixture model (GMM) based Kalman filter [13], [14], [73], [74].

TABLE I
VALUES OF THE PARAMETERS USED IN THE COUPLED NONHYDROSTATIC PHYSICAL-BIOLOGICAL-FISH MODEL. FOR THE NON-DIMENSIONALIZATION, THE SCALINGS USED ARE: $H = 50\ m$, $D = 1\ km$, AND TIME-SCALE OF $12.5\ d$.

Parameters	**Values**
Domain	
Horiz. length scale, D (km)	1
Vert. length scale, H (m)	50
Domain length, L (non-dim.)	20
Domain height, h (non-dim.)	2
Seamount center, X_c (non-dim.)	7.5
Model	
Inlet velocity, U (cm/s)	1
Eddy viscosity, ν_E (m^2/s)	0.01
Diffusion constants in horiz. and vert., $\mathcal{K}_x$ & $\mathcal{K}_z$ $(m^2/s$; same for all tracers, except fish density)	0.01 & 0.001
Reference Salinity, S_o (PSU)	35
TA relationship intercept, TA_o $(\mu mol/kg)$	190
TA relationship slope, b $(\mu mol/kg/PSU)$	unif{50, 75}

In this dynamics-based Bayesian learning framework, we perform simultaneous estimations of the unknown state variables (velocities u & v, S, and TA), and the slope parameter (b) using only very sparse observations. We employ so-called "identical twin experiment" [75] in which observations made from a simulated truth are generated using a deterministic run with a particular set of parameter values which lie within the uncertain realization space. The data are sparse in both space and time, with the observations only available for the TA field at nine locations every two non-dimensional times, starting at time $T = 3$ and ending at $T = 9$.

Figure 12 shows the prior of the system at $T = 3$, that is, just before the first set of observations are available. There are many differences between the mean and true fields of all the state variables. The blue dotted line in the probability plots of the slope parameter marks the true value. The prior probability of this parameter is considered to be uniform within a certain range. A vortex also starts to develop in the wake of the seamount indicative of the relative complexity of our non-hydrostatic dynamical setup.

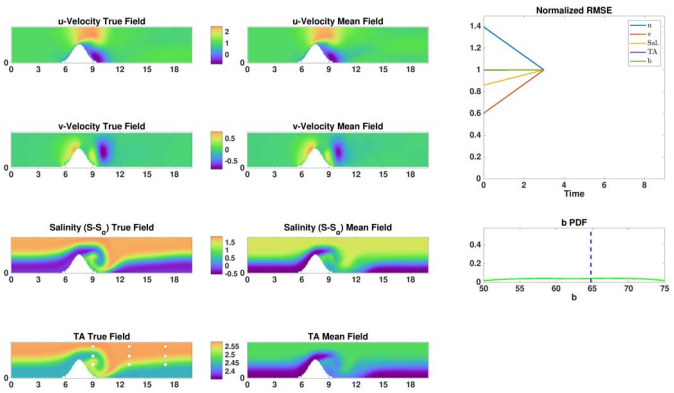


Fig. 12. Prior state of the simulated stochastic non-hydrostatic ocean physics and OA dynamics used in the learning experiment, at $T = 3$ (i.e. just before the 1st observational episode). The first two columns consist of the true (left) and mean (right) field of the state variables. In the third column, the first plot shows the variation of RMSE with time for various stochastic state variables and the parameter. The remaining plot contains the probability distribution of the uncertain OA slope parameter (green curve). The white circles on the total alkalinity (TA) true field marks the nine observation locations. TA is normalized with $1000 \mu mol/kg$.

In Figure 13, we provide the posterior of the system after dynamics-based Bayesian learning from four observational episodes, i.e. at $T = 9$. By observing the TA field, we are not only able to correct that field itself along with the corresponding slope parameter, but also the velocity and salinity fields. We use the variation of Root Mean Square Error (RMSE) over time to judge performance. RMSE is the L_2 distance between the mean of the random variables in the stochastic run and the simulated truth. The RMSE value for each of the variables at every time is normalized by the corresponding RMSE value just before the first assimilation step. Hence, our findings are corroborated by the decrease in RMSE for the parameters and state variables and by the narrowing of the posterior distribution of the OA slope parameter around the true value.

V. CONCLUSION

In this paper, we extended and showcased our MSEAS systems for high resolution coupled physical-biogeochemical-acidification simulations and Bayesian learning of OA models in Massachusetts Bay, starting with simple empirical and equilibrium OA models. These simulations were shown to have reasonable skill when compared to available in situ and remote data. The impacts of wind forcing, internal tides, and solitary waves on water transports and mixing, and OA fields, were explored. Strong wind events were confirmed to reset

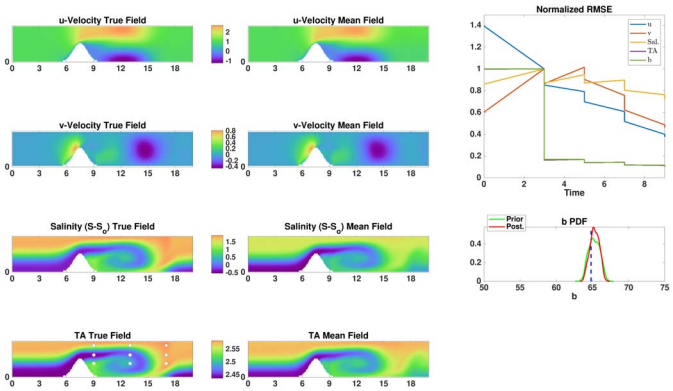


Fig. 13. As Fig. 12, but for the posterior state of the simulated non-hydrostatic ocean physics and OA dynamics used in the learning experiment, at $T = 9$ (i.e. just after the 4^{th} observational episode).

circulations and to also alter the OA state in the Bay. Internal tides were shown to increase vertical mixing of waters and material in the shallow regions. Large solitary waves propagating off Stellwagen Bank coupled with lateral turbulent mixing provided a pathway for exchange of surface and deep waters and material. Both of these effects impacted biological activity and OA fields. A mechanism for the creation of multiple subsurface maxima was presented, involving wind-induced upwelling, internal tides and advection of near surface fields. An outline was presented for coupling with ocean acidification models of varying complexity. Results from initial coupling of empirical and equilibrium OA models with our physical-biogeochemical modeling system were presented. We finally presented a proof-of-concept study to simultaneously learn and estimate the OA state variables and model parameterizations from sparse observations using our novel dynamics-based Bayesian learning framework for high-dimensional and multi-disciplinary estimation. Specifically, we numerically integrated the stochastic nonhydrostatic Navier-Stokes equations with salinity (S) as a passive tracer and total alkalinity assumed to be linearly dependent on the salinity, with uncertainty in the initial conditions and in the OA slope parameter. Using the GMM-DO Bayesian estimation, we learn the uncertain slope parameter along with all the state variables even though the observed data was sparse. This framework could be extended to discover new OA models or discriminate among the existing models.

Acknowledgments

We thank all members of the MSEAS group, past and present. We are grateful to Sea Grant and NOAA for support under the grant NA18OAR4170105 (BIOMAPS) to MIT. We thank the HYCOM team for their ocean fields, NMFS (Tamara Holzwarth-Davis and Paula Fratantoni) for their survey CTD data, NCEP (Matthew Pyle, Eric Rogers, Geoff DiMego, and Arun Chawla) for their help and support for atmospheric forcing forecasts, NOAA NDBC for supplying buoy data, JHU APL for processed SST images, and CORDC for HF Radar data.

References

[1] J. C. Orr, V. J. Fabry, O. Aumont, L. Bopp, S. C. Doney, R. A. Feely, A. Gnanadesikan, N. Gruber, A. Ishida, F. Joos, R. M. Key, K. Lindsay, E. Maier-Reimer, R. Matear, P. Monfray, A. Mouchet, R. G. Najjar, G.-K. Plattner, K. B. Rodgers, C. L. Sabine, J. L. Sarmiento, R. Schlitzer, R. D. Slater, I. J. Totterdell, M.-F. Weirig, Y. Yamanaka, and A. Yool, "Anthropogenic ocean acidification over the twenty-first century and its impact on calcifying organisms," *Nature*, vol. 437, no. 7059, pp. 681–686, 2005.

[2] S. C. Doney, W. M. Balch, V. J. Fabry, and R. A. Feely, "Ocean acidification: a critical emerging problem for the ocean sciences," *Oceanography*, vol. 22, no. 4, pp. 16–25, 2009.

[3] J. T. Mathis, S. R. Cooley, K. K. Yates, and P. Williamson, "Introduction to this special issue on ocean acidification: The pathway from science to policy," *Oceanography*, vol. 28, no. 2, pp. 10–15, 2015.

[4] R. A. Feely, C. L. Sabine, J. M. Hernandez-Ayon, D. Ianson, and B. Hales, "Evidence for upwelling of corrosive" acidified" water onto the continental shelf," *science*, vol. 320, no. 5882, pp. 1490–1492, 2008.

[5] D. K. Gledhill, M. M. White, J. Salisbury, H. Thomas, I. Mlsna, M. Liebman, B. Mook, J. Grear, A. C. Candelmo, R. C. Chambers *et al.*, "Ocean and coastal acidification off new england and nova scotia," *Oceanography*, vol. 28, no. 2, pp. 182–197, 2015.

[6] S. C. Talmage and C. J. Gobler, "Effects of past, present, and future ocean carbon dioxide concentrations on the growth and survival of larval shellfish," *Proceedings of the National Academy of Sciences*, vol. 107, no. 40, pp. 17 246–17 251, 2010.

[7] P. G. Y. Lu, "Bayesian inference of stochastic dynamical models," Master's thesis, Massachusetts Institute of Technology, Department of Mechanical Engineering, Cambridge, Massachusetts, February 2013.

[8] P. G. Y. Lu and P. F. J. Lermusiaux, "Pde-based bayesian inference of high-dimensional dynamical models," Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA, MSEAS Report 19, 2014.

[9] A. Gupta, P. J. Haley, D. N. Subramani, and P. F. J. Lermusiaux, "Fish modeling and Bayesian learning for the Lakshadweep Islands," in *OCEANS 2019 MTS/IEEE SEATTLE*. Seattle: IEEE, Oct. 2019, pp. 1–10.

[10] A. Gupta, "Bayesian inference of obstacle systems and coupled biogeochemical-physical models," Master's thesis, Indian Institute of Technology Kanpur, Kanpur, India, 2016.

[11] T. P. Sapsis and P. F. J. Lermusiaux, "Dynamically orthogonal field equations for continuous stochastic dynamical systems," *Physica D: Nonlinear Phenomena*, vol. 238, no. 23–24, pp. 2347–2360, Dec. 2009.

[12] F. Feppon and P. F. J. Lermusiaux, "Dynamically orthogonal numerical schemes for efficient stochastic advection and Lagrangian transport," *SIAM Review*, vol. 60, no. 3, pp. 595–625, 2018.

[13] T. Sondergaard and P. F. J. Lermusiaux, "Data assimilation with Gaussian Mixture Models using the Dynamically Orthogonal field equations. Part I: Theory and scheme." *Monthly Weather Review*, vol. 141, no. 6, pp. 1737–1760, 2013.

[14] ——, "Data assimilation with Gaussian Mixture Models using the Dynamically Orthogonal field equations. Part II: Applications." *Monthly Weather Review*, vol. 141, no. 6, pp. 1761–1785, 2013.

[15] P. J. Haley, Jr. and P. F. J. Lermusiaux, "Multiscale two-way embedding schemes for free-surface primitive equations in the "Multidisciplinary Simulation, Estimation and Assimilation System"," *Ocean Dynamics*, vol. 60, no. 6, pp. 1497–1537, Dec. 2010.

[16] P. F. J. Lermusiaux, P. J. Haley, W. G. Leslie, A. Agarwal, O. Logutov, and L. J. Burton, "Multiscale physical and biological dynamics in the Philippine Archipelago: Predictions and processes," *Oceanography*, vol. 24, no. 1, pp. 70–89, 2011, Special Issue on the Philippine Straits Dynamics Experiment.

[17] P. F. J. Lermusiaux, C.-S. Chiu, G. G. Gawarkiewicz, P. Abbot, A. R. Robinson, R. N. Miller, P. J. Haley, Jr, W. G. Leslie, S. J. Majumdar, A. Pang, and F. Lekien, "Quantifying uncertainties in ocean predictions," *Oceanography*, vol. 19, no. 1, pp. 92–105, 2006.

[18] W. G. Leslie, A. R. Robinson, P. J. Haley, Jr, O. Logutov, P. A. Moreno, P. F. J. Lermusiaux, and E. Coelho, "Verification and training of real-time forecasting of multi-scale ocean dynamics for maritime rapid

environmental assessment," *Journal of Marine Systems*, vol. 69, no. 1, pp. 3–16, 2008.

[19] R. Onken, A. R. Robinson, P. F. J. Lermusiaux, P. J. Haley, and L. A. Anderson, "Data-driven simulations of synoptic circulation and transports in the Tunisia-Sardinia-Sicily region," *Journal of Geophysical Research: Oceans*, vol. 108, no. C9, 2003.

[20] P. J. Haley, Jr., P. F. J. Lermusiaux, A. R. Robinson, W. G. Leslie, O. Logoutov, G. Cossarini, X. S. Liang, P. Moreno, S. R. Ramp, J. D. Doyle, J. Bellingham, F. Chavez, and S. Johnston, "Forecasting and reanalysis in the Monterey Bay/California Current region for the Autonomous Ocean Sampling Network-II experiment," *Deep Sea Research Part II: Topical Studies in Oceanography*, vol. 56, no. 3–5, pp. 127–148, Feb. 2009.

[21] A. Gangopadhyay, P. F. Lermusiaux, L. Rosenfeld, A. R. Robinson, L. Calado, H. S. Kim, W. G. Leslie, and P. J. Haley, Jr., "The California Current system: A multiscale overview and the development of a feature-oriented regional modeling system (FORMS)," *Dynamics of Atmospheres and Oceans*, vol. 52, no. 1–2, pp. 131–169, Sep. 2011, Special issue of Dynamics of Atmospheres and Oceans in honor of Prof. A. R. Robinson.

[22] S. R. Ramp, P. F. J. Lermusiaux, I. Shulman, Y. Chao, R. E. Wolf, and F. L. Bahr, "Oceanographic and atmospheric conditions on the continental shelf north of the Monterey Bay during August 2006," *Dynamics of Atmospheres and Oceans*, vol. 52, no. 1–2, pp. 192–223, Sep. 2011, Special issue of Dynamics of Atmospheres and Oceans in honor of Prof. A. R. Robinson.

[23] M. E. G. D. Colin, T. F. Duda, L. A. te Raa, T. van Zon, P. J. Haley, Jr., P. F. J. Lermusiaux, W. G. Leslie, C. Mirabito, F. P. A. Lam, A. E. Newhall, Y.-T. Lin, and J. F. Lynch, "Time-evolving acoustic propagation modeling in a complex ocean environment," in *OCEANS - Bergen, 2013 MTS/IEEE*, 2013, pp. 1–9.

[24] S. M. Kelly and P. F. J. Lermusiaux, "Internal-tide interactions with Gulf Stream and Middle Atlantic Bight shelfbreak front," *Journal of Geophysical Research: Oceans*, vol. 121, pp. 6271–6294, 2016.

[25] P. F. J. Lermusiaux, P. J. Haley, Jr., S. Jana, A. Gupta, C. S. Kulkarni, C. Mirabito, W. H. Ali, D. N. Subramani, A. Dutt, J. Lin, A. Shcherbina, C. Lee, and A. Gangopadhyay, "Optimal planning and sampling predictions for autonomous and Lagrangian platforms and sensors in the northern Arabian Sea," *Oceanography*, vol. 30, no. 2, pp. 172–185, Jun. 2017, special issue on Autonomous and Lagrangian Platforms and Sensors (ALPS).

[26] P. F. J. Lermusiaux, D. N. Subramani, J. Lin, C. S. Kulkarni, A. Gupta, A. Dutt, T. Lolla, P. J. Haley, Jr., W. H. Ali, C. Mirabito, and S. Jana, "A future for intelligent autonomous ocean observing systems," *Journal of Marine Research*, vol. 75, no. 6, pp. 765–813, Nov. 2017, the Sea. Volume 17, The Science of Ocean Prediction, Part 2.

[27] P. F. J. Lermusiaux, P. J. Haley, Jr, and N. K. Yilmaz, "Environmental prediction, path planning and adaptive sampling: sensing and modeling for efficient ocean monitoring, management and pollution control," *Sea Technology*, vol. 48, no. 9, pp. 35–38, 2007.

[28] J. Xu, P. F. J. Lermusiaux, P. J. Haley Jr., W. G. Leslie, and O. G. Logutov, "Spatial and Temporal Variations in Acoustic propagation during the PLUSNet-07 Exercise in Dabob Bay," in *Proceedings of Meetings on Acoustics (POMA)*, vol. 4. Acoustical Society of America 155th Meeting, 2008, p. 11.

[29] F.-P. A. Lam, P. J. Haley, Jr., J. Janmaat, P. F. J. Lermusiaux, W. G. Leslie, M. W. Schouten, L. A. te Raa, and M. Rixen, "At-sea real-time coupled four-dimensional oceanographic and acoustic forecasts during Battlespace Preparation 2007," *Journal of Marine Systems*, vol. 78, no. Supplement, pp. S306–S320, Nov. 2009.

[30] P. F. J. Lermusiaux, J. Xu, C.-F. Chen, S. Jan, L. Chiu, and Y.-J. Yang, "Coupled ocean–acoustic prediction of transmission loss in a continental shelfbreak region: Predictive skill, uncertainty quantification, and dynamical sensitivities," *IEEE Journal of Oceanic Engineering*, vol. 35, no. 4, pp. 895–916, Oct. 2010.

[31] T. F. Duda, Y.-T. Lin, W. Zhang, B. D. Cornuelle, and P. F. J. Lermusiaux, "Computational studies of three-dimensional ocean sound fields in areas of complex seafloor topography and active ocean dynamics," in *Proceedings of the 10th International Conference on Theoretical and Computational Acoustics*, Taipei, Taiwan, 2011.

[32] Ş. T. Beşiktepe, P. F. J. Lermusiaux, and A. R. Robinson, "Coupled physical and biogeochemical data-driven simulations of Massachusetts Bay in late summer: Real-time and post-cruise data assimilation," *Journal of Marine Systems*, vol. 40–41, pp. 171–212, 2003.

[33] G. Cossarini, P. F. J. Lermusiaux, and C. Solidoro, "Lagoon of Venice ecosystem: Seasonal dynamics and environmental guidance with uncertainty analyses and error subspace data assimilation," *Journal of Geophysical Research: Oceans*, vol. 114, no. C6, Jun. 2009.

[34] J. Coulin, P. J. Haley, Jr., S. Jana, C. S. Kulkarni, P. F. J. Lermusiaux, and T. Peacock, "Environmental ocean and plume modeling for deep sea mining in the Bismarck Sea," in *Oceans 2017 - Anchorage*, Anchorage, AK, Sep. 2017.

[35] M. Rixen, P. F. J. Lermusiaux, and J. Osler, "Quantifying, predicting, and exploiting uncertainties in marine environments," *Ocean Dynamics*, vol. 62, no. 3, pp. 495–499, 2012.

[36] M. De Dominicis, S. Falchetti, F. Trotta, N. Pinardi, L. Giacomelli, E. Napolitano, L. Fazioli, R. Sorgente, P. J. Haley, Jr., P. F. J. Lermusiaux, F. Martins, and M. Cocco, "A relocatable ocean model in support of environmental emergencies," *Ocean Dynamics*, vol. 64, no. 5, pp. 667–688, 2014.

[37] O. Schofield, S. Glenn, J. Orcutt, M. Arrott, M. Meisinger, A. Gangopadhyay, W. Brown, R. Signell, M. Moline, Y. Chao, S. Chien, D. Thompson, A. Balasuriya, P. F. J. Lermusiaux, and M. Oliver, "Automated sensor networks to advance ocean science," *Eos Trans. AGU*, vol. 91, no. 39, pp. 345–346, Sep. 2010.

[38] T. Lolla, P. F. J. Lermusiaux, M. P. Ueckermann, and P. J. Haley, Jr., "Time-optimal path planning in dynamic flows using level set equations: Theory and schemes," *Ocean Dynamics*, vol. 64, no. 10, pp. 1373–1397, 2014.

[39] T. Lolla, P. J. Haley, Jr., and P. F. J. Lermusiaux, "Time-optimal path planning in dynamic flows using level set equations: Realistic applications," *Ocean Dynamics*, vol. 64, no. 10, pp. 1399–1417, 2014.

[40] P. F. J. Lermusiaux, T. Lolla, P. J. Haley, Jr., K. Yigit, M. P. Ueckermann, T. Sondergaard, and W. G. Leslie, "Science of autonomy: Time-optimal path planning and adaptive sampling for swarms of ocean vehicles," in *Springer Handbook of Ocean Engineering: Autonomous Ocean Vehicles, Subsystems and Control*, T. Curtin, Ed. Springer, 2016, ch. 21, pp. 481–498.

[41] P. F. J. Lermusiaux, "Adaptive modeling, adaptive data assimilation and adaptive sampling," *Physica D: Nonlinear Phenomena*, vol. 230, no. 1, pp. 172–196, 2007.

[42] K. D. Heaney, G. Gawarkiewicz, T. F. Duda, and P. F. J. Lermusiaux, "Nonlinear optimization of autonomous undersea vehicle sampling strategies for oceanographic data-assimilation," *Journal of Field Robotics*, vol. 24, no. 6, pp. 437–448, 2007.

[43] K. D. Heaney, P. F. J. Lermusiaux, T. F. Duda, and P. J. Haley, Jr., "Validation of genetic algorithm based optimal sampling for ocean data assimilation," *Ocean Dynamics*, vol. 66, pp. 1209–1229, 2016.

[44] G. Gawarkiewicz, S. Jan, P. F. J. Lermusiaux, J. L. McClean, L. Centurioni, K. Taylor, B. Cornuelle, T. F. Duda, J. Wang, Y. J. Yang, T. Sanford, R.-C. Lien, C. Lee, M.-A. Lee, W. Leslie, P. J. Haley, Jr., P. P. Niiler, G. Gopalakrishnan, P. Velez-Belchi, D.-K. Lee, and Y. Y. Kim, "Circulation and intrusions northeast of Taiwan: Chasing and predicting uncertainty in the cold dome." *Oceanography*, vol. 24, no. 4, pp. 110–121, 2011.

[45] P. F. J. Lermusiaux, P. J. Haley, Jr., G. G. Gawarkiewicz, and S. Jan, "Evaluation of multiscale ocean probabilistic forecasts: Quantifying, predicting and exploiting uncertainty," *Ocean Dynamics*, 2020, to be submitted.

[46] A. Agarwal and P. F. J. Lermusiaux, "Statistical field estimation for complex coastal regions and archipelagos," *Ocean Modelling*, vol. 40, no. 2, pp. 164–189, 2011.

[47] Y. Pan, P. J. Haley, Jr., and P. F. J. Lermusiaux, "Interactions of internal tides with a heterogeneous and rotational ocean," *Journal of Fluid Mechanics*, 2020, sub-judice.

[48] P. F. J. Lermusiaux, "Evolving the subspace of the three-dimensional multiscale ocean variability: Massachusetts Bay," *Journal of Marine Systems*, vol. 29, no. 1, pp. 385–422, 2001.

[49] P. F. J. Lermusiaux, A. R. Robinson, P. J. Haley, and W. G. Leslie, "Advanced interdisciplinary data assimilation: Filtering and smoothing via error subspace statistical estimation," in *Proceedings of The OCEANS 2002 MTS/IEEE conference*. Holland Publications, 2002, pp. 795–802.

[50] P. F. J. Lermusiaux, C. Mirabito, P. J. Haley, Jr., W. H. Ali, A. Gupta, S. Jana, E. Dorfman, A. Laferriere, A. Kofford, G. Shepard, M. Goldsmith, K. Heaney, E. Coelho, J. Boyle, J. Murray, L. Freitag, and A. Morozov, "Real-time probabilistic coupled ocean physics-acoustics forecasting and data assimilation for underwater GPS," in *OCEANS 2020 IEEE/MTS*. IEEE, Oct. 2020, in press.

[51] M. P. Ueckermann and P. F. J. Lermusiaux, "Hybridizable discontinuous Galerkin projection methods for Navier–Stokes and Boussinesq equations," *Journal of Computational Physics*, vol. 306, pp. 390–421, 2016.

[52] M. P. Ueckermann, C. Mirabito, P. J. Haley, Jr., and P. F. J. Lermusiaux, "High order hybridizable discontinuous Galerkin projection schemes for non-hydrostatic physical-biogeochemical ocean modeling," *Ocean Dynamics*, 2020, to be submitted.

[53] M. P. Ueckermann and P. F. J. Lermusiaux, "2.29 Finite Volume MATLAB Framework Documentation," Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA, MSEAS Report 14, 2012. [Online]. Available: http://mseas.mit.edu/?p=2567

[54] M. P. Ueckermann, P. F. J. Lermusiaux, and T. P. Sapsis, "Numerical schemes for dynamically orthogonal equations of stochastic fluid and ocean flows," *Journal of Computational Physics*, vol. 233, pp. 272–294, Jan. 2013.

[55] E. R. Twomey and R. P. Signell, "Construction of a 3-arcsecond digital elevation model for the Gulf of Maine," U.S. Geological Survey, Open-File Report 2011–1127, 2013. [Online]. Available: https://pubs.usgs.gov/of/2011/1127/

[56] J. A. Cummings and O. M. Smedstad, *Variational Data Assimilation for the Global Ocean*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 303–343.

[57] P. J. Haley, Jr., A. Agarwal, and P. F. J. Lermusiaux, "Optimizing velocities and transports for complex coastal regions and archipelagos," *Ocean Modeling*, vol. 89, pp. 1–28, 2015.

[58] G. D. Egbert and S. Y. Erofeeva, "Efficient inverse modeling of barotropic ocean tides," *Journal of Atmospheric and Oceanic Technology*, vol. 19, no. 2, pp. 183–204, 2002.

[59] ——, "OSU tidal inversion," http://volkov.oce.orst.edu/tides/tpxo8_atlas.html, 2013.

[60] O. G. Logutov and P. F. J. Lermusiaux, "Inverse barotropic tidal estimation for regional ocean applications," *Ocean Modelling*, vol. 25, no. 1–2, pp. 17–34, 2008. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1463500308000851

[61] National Centers for Environmental Prediction (NCEP), "North American Mesoscale Forecast System (NAM)," https://www.emc.ncep.noaa.gov/index.php?branch=NAM, Sep. 2019.

[62] National Data Buoy Center (NDBC), https://www.ndbc.noaa.gov/, Sep. 2019.

[63] P. F. J. Lermusiaux, M. Doshi, C. S. Kulkarni, A. Gupta, P. J. Haley, Jr., C. Mirabito, F. Trotta, S. J. Levang, G. R. Flierl, J. Marshall, T. Peacock, and C. Noble, "Plastic pollution in the coastal oceans: Characterization and modeling," in *OCEANS 2019 MTS/IEEE SEATTLE*. Seattle: IEEE, Oct. 2019, pp. 1–10.

[64] P. F. J. Lermusiaux, C. Evangelinos, R. Tian, P. J. Haley, Jr, J. J. McCarthy, N. M. Patrikalakis, A. R. Robinson, and H. Schmidt, "Adaptive coupled physical and biogeochemical ocean predictions: A conceptual basis," in *Computational Science - ICCS 2004*, ser. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2004, vol. 3038, pp. 685–692.

[65] National Centers for Environmental Information (NCEI), "World ocean database 2018," Jun. 2020. [Online]. Available: https://www.nodc.noaa.gov/OC5/WOD/pr_wod.html

[66] W.-J. Cai, X. Hu, W.-J. Huang, L.-Q. Jiang, Y. Wang, T.-H. Peng, and X. Zhang, "Alkalinity distribution in the western north atlantic ocean margins," *Journal of Geophysical Research: Oceans*, vol. 115, no. C8, 2010. [Online]. Available: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2009JC005482

[67] S. R. Signorini, A. Mannino, R. G. Najjar Jr., M. A. M. Friedrichs, W.-J. Cai, J. Salisbury, Z. A. Wang, H. Thomas, and E. Shadwick, "Surface ocean pco2 seasonality and sea-air co2 flux estimates for the north american east coast," *Journal of Geophysical Research: Oceans*, vol. 118, no. 10, pp. 5439–5460, 2013. [Online]. Available: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/jgrc.20369

[68] R. van Hooidonk, "East coast ocean acidification product suite," 2020. [Online]. Available: https://www.coral.noaa.gov/accrete/east-coast-oaps.html

[69] E. R. Lewis and D. W. R. Wallace, "Program Developed for CO_2 System Calculations," Oak Ridge National Laboratory, Oak Ridge, Tennessee, Tech. Rep. CDIAC-105, 1998. [Online]. Available: https://www.nodc.noaa.gov/ocads/oceans/CO2SYS/co2rprt.html

[70] D. Cressey, "Seawater studies come up short," *Nature*, vol. 524, no. 7563, pp. 18–19, 2015.

[71] T. L. Frölicher, K. B. Rodgers, C. A. Stock, and W. W. Cheung, "Sources of uncertainties in 21st century projections of potential ocean ecosystem stressors," *Global Biogeochemical Cycles*, vol. 30, no. 8, pp. 1224–1243, 2016.

[72] T. P. Sapsis and P. F. J. Lermusiaux, "Dynamical criteria for the evolution of the stochastic dimensionality in flows with uncertainty," *Physica D: Nonlinear Phenomena*, vol. 241, no. 1, pp. 60–76, 2012.

[73] T. Lolla and P. F. J. Lermusiaux, "A Gaussian mixture model smoother for continuous nonlinear stochastic dynamical systems: Theory and scheme," *Monthly Weather Review*, vol. 145, pp. 2743–2761, Jul. 2017.

[74] ——, "A Gaussian mixture model smoother for continuous nonlinear stochastic dynamical systems: Applications," *Monthly Weather Review*, vol. 145, pp. 2763–2790, Jul. 2017.

[75] P. F. J. Lermusiaux, "Data assimilation via Error Subspace Statistical Estimation, part II: Mid-Atlantic Bight shelfbreak front simulations, and ESSE validation," *Monthly Weather Review*, vol. 127, no. 7, pp. 1408–1432, Jul. 1999.