Ocean Modelling 135 (2019) 40–55

Contents lists available at ScienceDirect

Ocean Modelling

journal homepage: www.elsevier.com/locate/ocemod

Quantifying wavelengths constrained by simulated SWOT observations in a submesoscale resolving ocean analysis/forecasting system

Joseph M. D'Addezio[a,*], Scott Smith[b], Gregg A. Jacobs[b], Robert W. Helber[b], Clark Rowley[b], Innocent Souopgui[c], Matthew J. Carrier[b]

[a] University of Southern Mississippi, School of Ocean Sciences and Technology, Division of Marine Science, MS, USA
[b] Naval Research Laboratory, Ocean Dynamics and Prediction, MS, USA
[c] University of New Orleans, Department of Physics, LA, USA

ARTICLE INFO

Keywords:
OSSE
SWOT
Altimeter
3DVAR
Mesoscale
Submesoscale

ABSTRACT

Using a suite of Observing System Simulation Experiments (OSSEs), the utility of simulated Surface Water Ocean Topography (SWOT) observations is estimated in a high-resolution (1 km) ocean analysis/forecasting system. Sampling a Nature Run provides observations for the OSSEs and the realism of the Nature Run is established by comparison to climatological data and an independent ocean analysis/forecast system. Each OSSE experiment assimilated different sets of simulated observations including traditional nadir altimeters, satellite sea surface temperature (SST), in situ profile data, and SWOT. OSSE evaluation metrics include area-averaged errors and wavenumber spectra with the latter providing much finer differentiation between experiments. 100 m temperature, sea surface height (SSH), and mixed layer depth (MLD) errors across the observed wavenumber spectra were reduced by up to 20% for OSSEs assimilating the simulated SWOT observations. The minimum constrained wavelength was found to be 130 km when both nadir altimetry and SWOT observations were used. The experiment using only nadir altimetry produced a value of 161 km. This 31 km gain in skill of predictable scales suggests that ocean forecasts can expect substantial gains in capability when utilizing the forthcoming SWOT data. Experimentation with the analysis decorrelation length scale suggests that emerging multi-scale assimilation methodologies will provide additional advancements in predictive skill.

1. Introduction

Numerous studies have documented the utility of altimetric observations of the ocean surface in generating mesoscale predictive skill (e.g. Smedstad et al., 2003; Ananda et al., 2006; Jacobs et al., 2014a; Le Traon et al., 2015). However, the current suite of nadir looking altimeters produces an ocean surface representation that is limited by spatial and temporal sampling. Since Jason-2, at least two altimeter sensors have been in orbit at any one time with their merged observations producing two dimensional grids of sea surface height (SSH) that can resolve wavelengths of approximately 150 km and greater (Ducet et al., 2000; Fu and Ubelmann, 2014). These estimates are significantly coarser than the features present ocean prediction systems can resolve. This observational deficit may soon be alleviated with the launch of the Surface Water Ocean Topography (SWOT) satellite mission in 2021 (Fu and Ubelmann, 2014). The sensor is expected to collect along- and across-track SSH observations at very high resolutions (~1 km), allowing oceanographers to study previously under observed phenomena on a global scale. This observational capability poses important questions to the ocean prediction community: what increases in forecast skill can be expected by assimilating these data into a high-resolution ocean analysis/forecasting system? Will the SWOT observations allow the assimilation system to constrain spatial scales below that which is currently achievable using a constellation of nadir altimeters? The objective of this study is to determine how SWOT observations can change ocean predictability when compared to results derived using only currently operating nadir altimetry.

A number of studies have tested the usefulness of simulated SWOT observations for more narrowly focused hydrological and hydrographic modeling (Andreadis et al., 2007; Durand et al., 2008; Biancamaria et al., 2010; Yoon et al., 2012; Pedinotti et al., 2014; Munier et al., 2015; Oubanas et al., 2018a, 2018b). Carrier et al. (2016) showed that within a regional ocean 4DVAR data assimilation framework, simulated SWOT observations significantly improved analysis/forecast skill with respect to both the placement of mesoscale features and surface velocities. The work presented here extends the resolution of the model

* Corresponding author at: The University of Southern Mississippi, Division of Marine Science, 1020 Balch Blvd., Stennis Space Center, MS 39529, USA.
E-mail address: joseph.daddezio@usm.edu (J.M. D'Addezio).

https://doi.org/10.1016/j.ocemod.2019.02.001
Received 14 June 2018; Received in revised form 15 November 2018; Accepted 4 February 2019
Available online 08 February 2019
1463-5003/

system from 6 km to 1 km and quantifies the bulk error magnitudes and resolved spatial scales. This is a critical expansion of experimentation as a horizontal model resolution of 1 km allows the simulation to explicitly resolve submesoscale processes with wavelengths O (10 km) (Capet et al., 2008). These small-scale dynamics are important to accurately represent and predict in numerical models because they have strong impact on near surface dynamics (McWilliams, 2016), net ocean-atmosphere heat fluxes (Su et al., 2018), and biogeochemical processes (Levy et al., 2012). The current set of satellite altimeters cannot observe submesoscale features, and therefore regional models that explicitly resolve submesoscales are unable to constrain the placement and evolution of the phenomenon. SWOT will provide an important path forward towards bridging this gap between model resolution and observational capabilities. This requires testing to quantify the range of wavelengths the SWOT observations will allow the assimilation system to constrain. Ocean forecasting systems have been designed for mesoscale prediction (Cummings et al., 2009) and it is unclear if mesoscale assumptions translate to submesoscale predictive skill even when utilizing high-resolution data streams that explicitly resolve the phenomenon.

The material is organized by first describing the construction of an Observing System Simulation Experiment (OSSE) (e.g. Halliwell et al., 2014) in Section 2. A simulated truth, hereafter called NATURE, was generated using a submesoscale resolving, non-assimilative model run. NATURE was sampled at real observation locations and times for remotely sensed sea surface temperature (SST), in situ profiles, and satellite altimetry. SWOT sampling was derived using the SWOT simulator. OSSEs deviate from NATURE through initial condition perturbation. Different sets of simulated observations were assimilated into each OSSE experiment. The errors in the OSSEs are evaluated in Section 3 by comparison of area-averaged statistics and by wavenumber spectra of errors. Section 4 considers the potential usefulness of emerging data assimilation techniques based on results derived by adjusting the analysis decorrelation length scale in an attempt to further utilize the small-scale variability observed by SWOT in Section 3.

2. Methods

Ascertaining the utility of a remote sensing platform that is not currently producing data requires the construction of a controlled laboratory environment in which simulated data, consistent with expected sampling, can be tested. This section documents the methodology of the experiments: the numerical ocean model configuration used by NATURE and each of the OSSE experiments (Section 2.1), NATURE spin up and validation (Section 2.2), setup of the OSSEs (Section 2.3), and the data assimilation methodology (Section 2.4).

2.1. Model system setup

The Navy Coastal Ocean Model (NCOM) (Barron et al., 2006) provided simulated three dimensional ocean fields, and we applied the system within the western Pacific Ocean (Fig. 1a); a region that contains a strong western boundary current, extensive continental shelves, deep ocean basins, strong seasonal surface mixed layer variability, strong mesoscale features in the deep ocean, submesoscale processes, and interactions between all of these. NCOM integrates the primitive equations forward in time using the hydrostatic and Boussinesq approximations. It features a hybrid σ/z vertical grid allowing for higher vertical resolution in the near surface environment as well as terrain following capabilities in the coastal ocean. A standard z vertical grid is used at deeper layers and the transition from z to upper layer sigma coordinates is a user input. For experiments here, 50 vertical layers were used to a maximum depth of 4000 m with the σ/z transition occurring at layer 25, corresponding to a depth of 120 m. At rest, the surface sigma layer at its thickest is about 1 m, and layer thicknesses progressively increase with increasing depth. In the horizontal, the

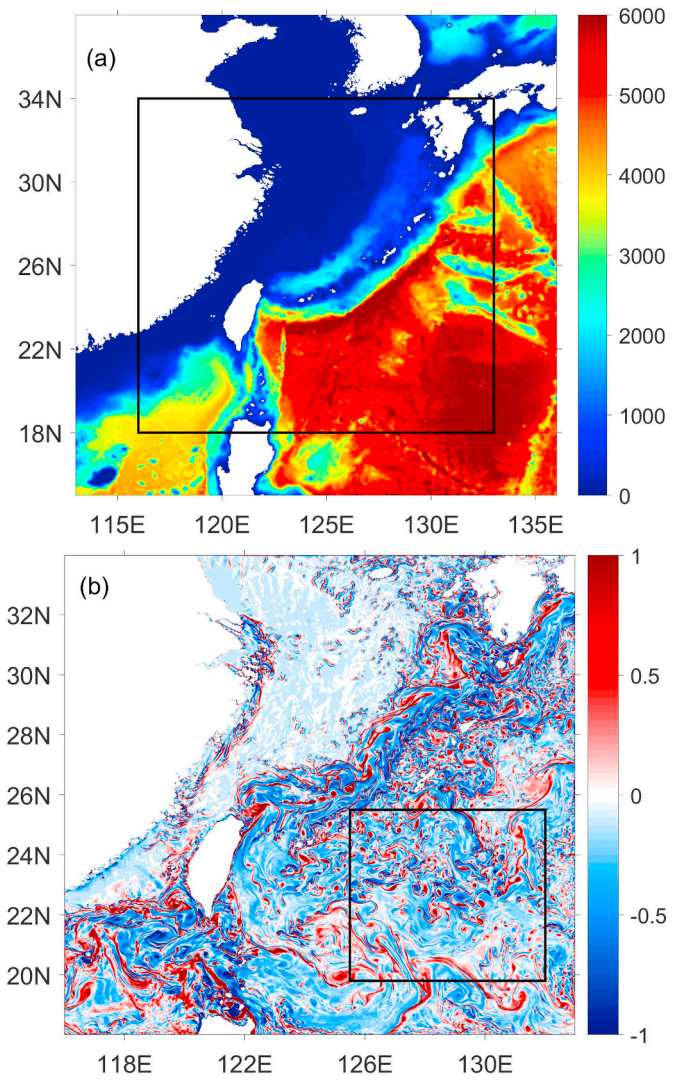

Fig. 1. (a) Bathymetry (m) of the western Pacific Ocean domain used for the larger 3 km nest (113°E–136°E; 15°N–38°N). The enclosed black box shows the domain for the final 1 km nest (116°E–133°E; 18°N–34°N). (b) Surface relative vorticity normalized by the Coriolis parameter from the 1 km NATURE run on January 1, 2016. The black box denotes the square subregion over which wavenumber spectra analyses were performed.

model resolution was set to 1 km, allowing NCOM to produce ocean features with wavelengths down to O (10 km), which is a minimal resolution for representing submesoscale processes (Capet et al., 2008). Using a model grid that can explicitly resolve submesoscale features provides the ability to conduct an important hypothesis test: can assimilating simulated SWOT observations into the high-resolution system constrain features at submesoscale wavelengths? Using a model grid any coarser than that chosen would preclude the rejection or acceptance of any such hypothesis.

Lateral ocean boundary conditions were provided by a double, one-way nesting procedure. First, boundary conditions from the operational global 1/12° HYbrid Coordinate Ocean Model (HYCOM) were provided to the larger nest (113°E–136°E; 15°N–38°N) (Fig. 1a): a regional NCOM simulation with a horizontal resolution of 3 km. Boundary conditions were generated from the 3 km simulation and provided to the 1 km nest (116°E–133°E; 18°N–34°N) (Fig. 1b). This approach was taken to mitigate numerical instabilities that can arise from interpolating boundary conditions from a model at least 3 times coarser than the nested simulation. Atmospheric forcing was provided by the NAVy Global Environmental Model (NAVGEM) (Hogan et al., 2014). The forcing

derived from this model has a horizontal resolution of 37 km and a frequency of 3 h. Surface wind stress and precipitation were extracted directly from the NAVGEM output. Latent and sensible heat flux were computed by Coupled Ocean-Atmosphere Response Experiment (COARE) version 3.0 (Fairall et al., 2003). Solar radiation provided by NAVGEM was absorbed through the water column using a Jerlov water type II (Jerlov, 1976). River in flow was introduced at fixed points approximating the geographic location of large river mouths using the NCOM river database (Barron and Smedstad, 2002). The final 1 km nest included tidal forcing from the global Oregon Tidal Inverse Solution (OTIS) (Egbert and Erofeeva, 2002). The tidal forcing was not included in global HYCOM or the 3 km outer nest. Any analyses featuring either SSH or velocity have had the barotropic tidal signal removed from each variable through a post processing procedure. Nonetheless, this harmonic analysis will inevitably miss incoherent baroclinic tides that have non-tidal frequencies. This phenomenon is present in temperature and salinity below the mixed layer, SSH, as well as surface and subsurface velocities and will therefore be included in our error analyses. This problem deserves attention as these waves are expected to "containment" balanced mesoscale and submesoscale SWOT observations (Chavanne et al., 2010), but the subject is outside the scope of this study. In all, we endeavored to produce a highly realistic system that includes short time and small spatial scale variability representative of the physical ocean and will therefore be present in the SWOT observations.

2.2. NATURE setup and validation

This section describes NATURE and a validation of NATURE statistics in comparison with climatological observations of the real ocean and an independent assimilative ocean forecasting system. The 3 km NCOM nest was initialized on November 1, 2015 and spun up until the end of the month. The 3 km December 1, 2015 forecast was then interpolated onto the final 1 km grid and used as the initial condition for NATURE. NATURE was then spun up for the entirety of December 2015 to ensure that both mesoscale and submesoscale features had sufficient time to properly develop. The simulation was then integrated forward for an entire year with the complete time period being: January 1, 2016 to December 31, 2016. No data assimilation was applied to NATURE.

The completely unconstrained nature of this simulation could introduce problems in the OSSE whereby the statistics of NATURE deviate so far from the real ocean that conclusions gleaned from the experiment may not be applicable to the physical world. For this reason, it is common practice to validate the NATURE simulation to ensure that it is sufficiently realistic (Atlas, 1997). The purpose of this procedure is not to test the error of the solution in terms of how it exactly compares with the real ocean at equivalent times. At any given time step, NATURE will be quite different from the real ocean. Instead, we are testing to determine if NATURE statistics are consistent with the statistics of the real ocean.

To begin, surface temperature, mixed layer, and thermocline means are analyzed. Qualitatively, the mean NATURE SST compares favorably with both the observation-based Generalized Digital Environmental Model version 4 (GDEM4) (Teague et al., 1990; Carnes et al., 2010) SST climatology and the assimilative HYCOM 2016 annual SST mean (Fig. 2). All three means have similar magnitude and comparable locations of large scale fronts, such as the Kuroshio western boundary current that meets the coastal waters along the continental shelf. By taking the area-averaged value at each time step of each dataset, the seasonal cycle of NATURE can be validated (Fig. 2d). Both NATURE and HYCOM deviate from the area-averaged GDEM4 climatology within each month, with the model deviations from climatology being similar. The seasonal cycle amplitude is consistent between the models and climatology. The higher frequency events in the model time series are similar and are responses to atmospheric events. Using a 0.2 °C temperature criteria to derive mixed layer depth (MLD) produces a similar mean spatial distribution from all three of the examined datasets (Fig. 3). In the deep ocean, GDEM4 has slightly shallower MLD than both NATURE and HYCOM. NATURE produces the shallow MLD along the Kuroshio pathway evident in the GDEM4 data much better than HYCOM. Across the annual cycle, NATURE and HYCOM track the GDEM4 climatology closely, although mixed layers are generally deeper in both simulations between January and the end of March (Fig. 3d). Below the deepest regional mean MLD (~90 m; Fig. 3d), 100 m temperature provides a representation of the mean thermocline structure (Fig. 4). The location of the Kuroshio front is in good agreement between the three datasets and deep ocean magnitudes are highly comparable. Through time, the NATURE spatial mean shows deviations from both GDEM4 and HYCOM on the order of approximately 1 °C between February and November (Fig. 4d). The seasonal cycle of three datasets, however, are in good agreement. In all, these analyses suggest that NATURE SST, MLD, and 100 m temperature means have similar magnitudes, spatial distributions, and seasonal cycles to that of the climatology of the real ocean.

Next, NATURE energy is evaluated. The mean surface eddy kinetic energy (EKE) from NATURE compares favorably with drifter observations and assimilative HYCOM (Fig. 5). All drifter data collected between 1983 and 2009 were binned into 1° × 1° grid boxes. Only grid points with at least 100 drifter observations were considered in the temporal mean (Thoppil et al., 2011). The spatial distribution of mean EKE is similar in all three datasets, but NATURE magnitudes are more similar to that of the observations than HYCOM. This is most likely due to model differences in horizontal resolution. Prior experiments have shown that a model horizontal resolution of at least 3 km is required to completely represent mesoscale dynamics (e.g. Hogan and Hurlburt, 2000). Spatially, the three datasets are highly consistent with higher EKE in the deep ocean and along the pathway of the Kuroshio western boundary current as well as lower EKE north of the current along the wide continental shelf. Strong EKE seasonality is not observed and despite differences in magnitude the two time-series have strong correlation through time (Fig. 5d). 2016 annual NATURE SSH root mean square (RMS) is evaluated against nadir altimetry observations and assimilative HYCOM over the same time period (Fig. 6). All three datasets feature comparable magnitudes. The points of maximum RMS in each dataset are located primarily along the pathway of the Kuroshio and to the east of the boundary current. South of approximately 20°N, RMS values are lower in each. Overall, HYCOM compares most favorably with the observations in terms of feature location as it assimilated the altimeter data throughout 2016. The non-assimilative NATURE does not produce local maxima in exactly the same locations as observed, but the overall magnitude and general spatial pattern is consistent with the observations. Finally, annual subsurface temperature standard deviation along 21°N is examined (Fig. 7). All three datasets produce depth maximum temperature standard deviation within the top 150 m. Lower amplitude thermocline variability is present below this level and the 1 °C standard deviation contour extends down to approximately 500 m in NATURE and GDEM4. The same contour in HYCOM is shallower, with a maximum depth of approximately 250 m. In all, the NATURE temperature depth variance structure is more comparable with the climatological data than HYCOM.

Overall, NATURE statistics are in good agreement with observations and an assimilative ocean model. With this confirmation, NATURE was used as a reasonable approximation of the real ocean and sampled to provide simulated observations to the OSSE experiments. The NATURE sampling, OSSE configuration, and data assimilation methodology are described in the remainder of Section 2.

2.3. OSSE experiments

If the same model, horizontal resolution, initial condition, boundary conditions, and surface forcing were used to generate the OSSE experiments, an exact replica of NATURE would be created. Formally,

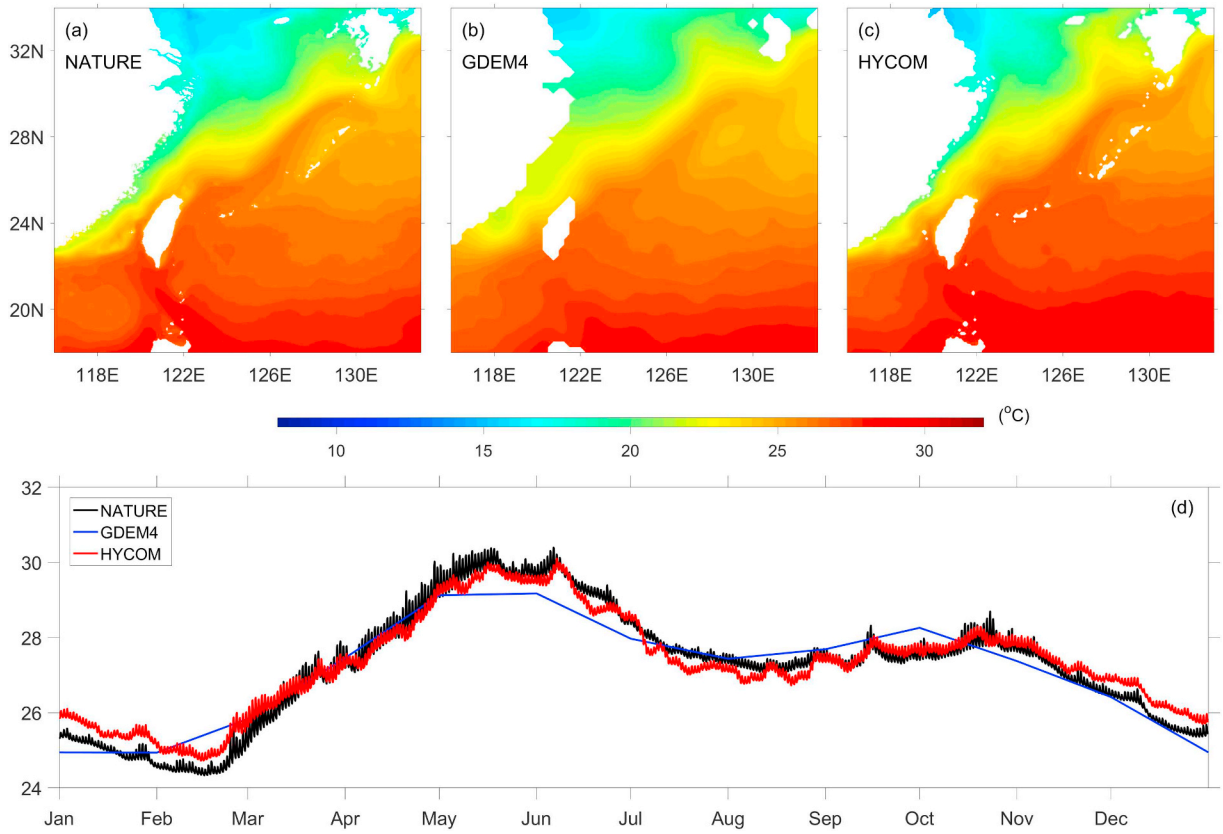

Fig. 2. Mean sea surface temperature (SST; ˚C) of (a) NATURE (2016), (b) GDEM4 climatology, and (c) global HYCOM (2016). (d) Area-averaged NATURE, GDEM4, and HYCOM SST for each of their respective time steps.

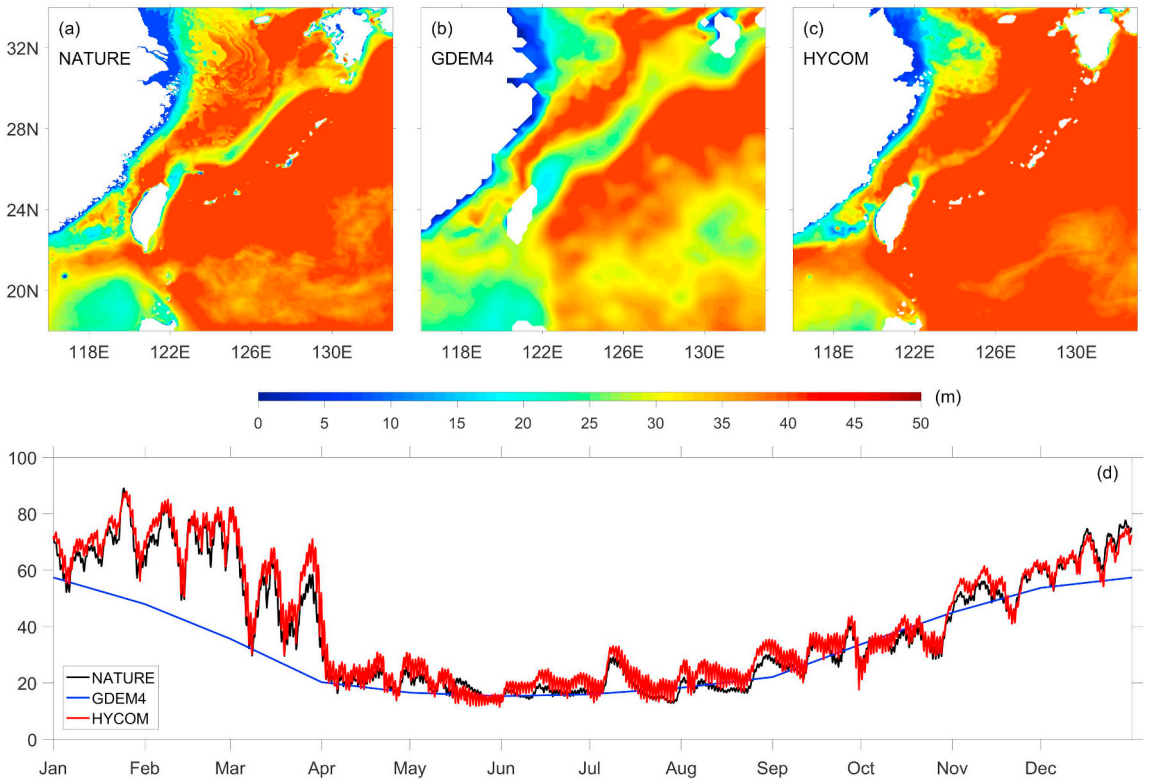

Fig. 3. Mean mixed layer depth (MLD; m) of (a) NATURE (2016), (b) GDEM4 climatology, and (c) global HYCOM (2016). (d) Area-averaged NATURE, GDEM4, and HYCOM MLD for each of their respective time steps.

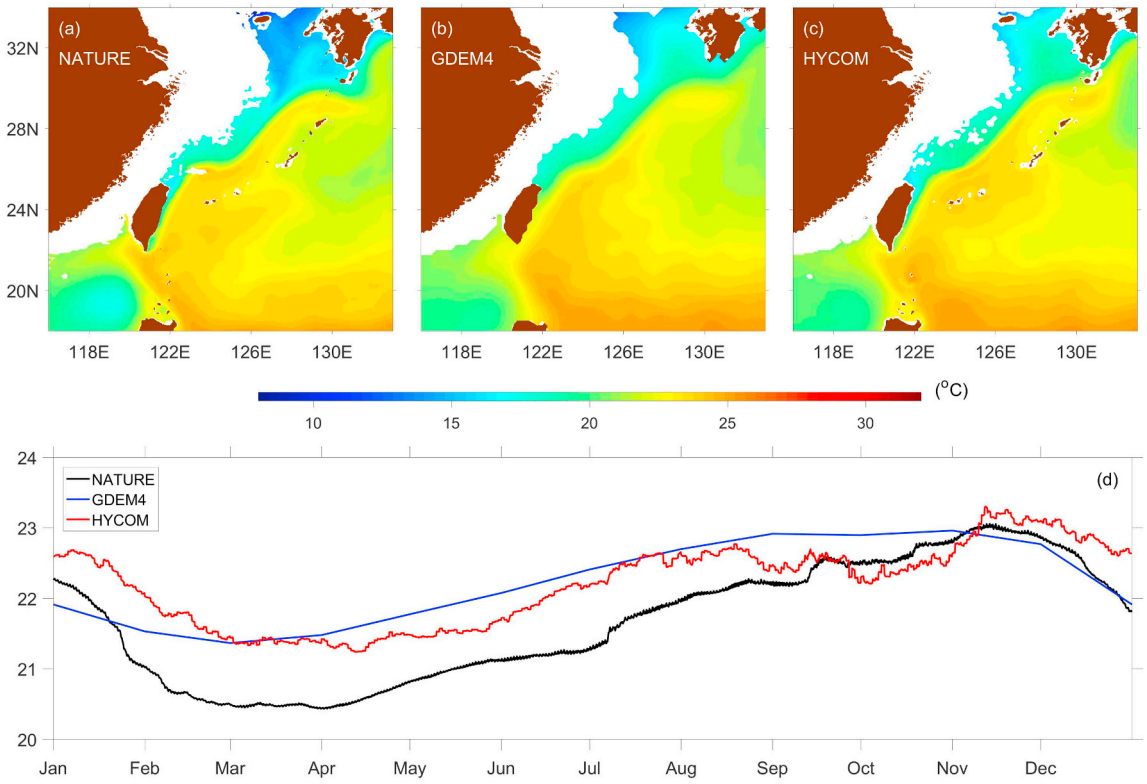

Fig. 4. Mean 100 m temperature (°C) of (a) NATURE (2016), (b) GDEM4 climatology, and (c) global HYCOM (2016). (d) Area-averaged NATURE, GDEM4, and HYCOM 100 m temperature for each of their respective time steps.

OSSEs should utilize a completely different model run at a coarser resolution to represent errors in model physics and resolution limitations (Atlas, 1997). The approach taken in this study, however, is to use the same dynamical model and resolution and a realistically different initial condition for the OSSEs. The hypothesis is that the new initial condition is sufficiently different and that the processes inside the 1 km numerical

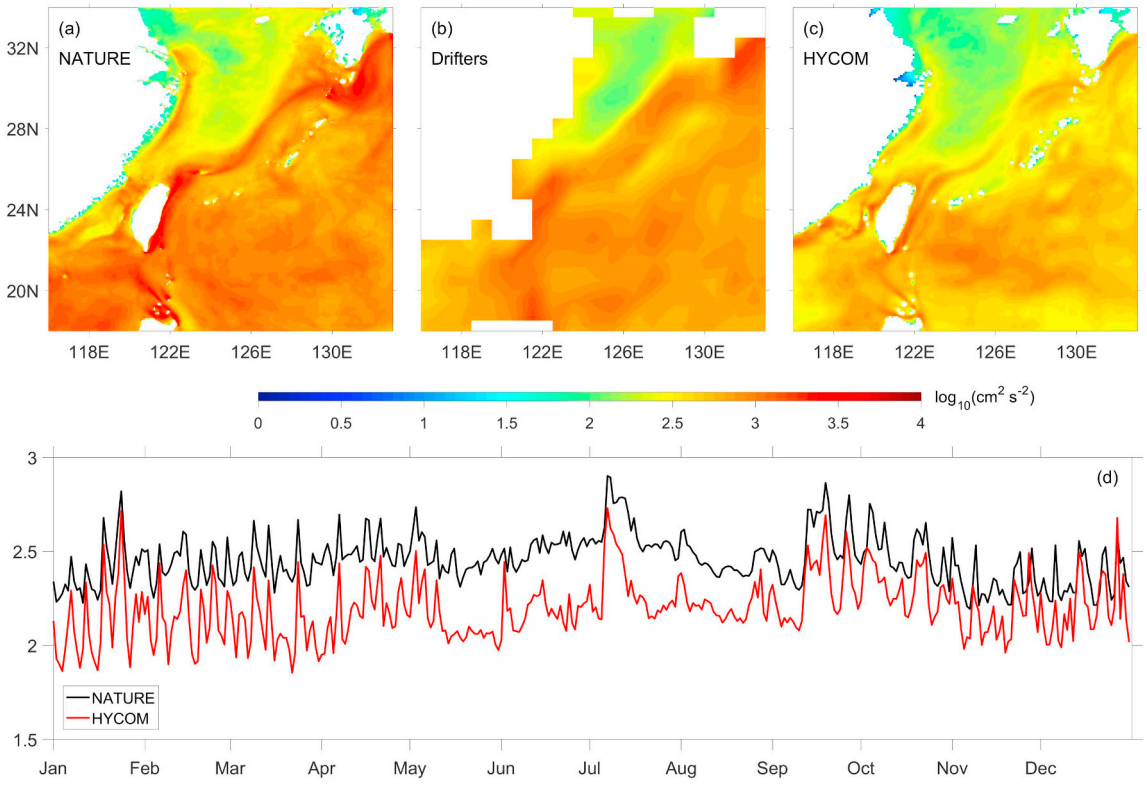

Fig. 5. Mean surface eddy kinetic energy (EKE; $cm^2\,s^{-2}$) of (a) NATURE (2016), (b) surface drifter observations (1983–2009), and (c) global HYCOM (2016). (d) Area-averaged NATURE and HYCOM EKE for each of their respective time steps.

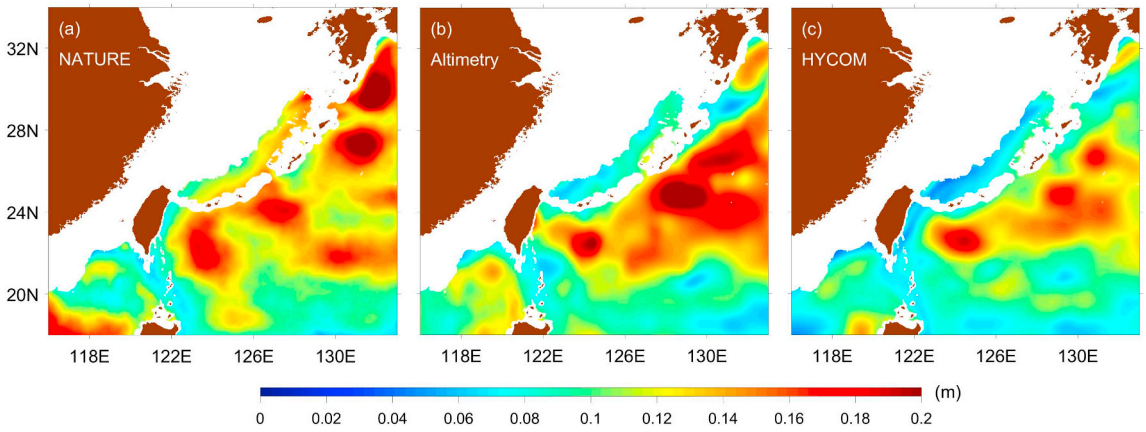

Fig. 6. Root mean square (RMS) of sea surface height (SSH; m) anomalies from (a) NATURE (2016), (b) gridded nadir altimetry (2016), and (c) global HYCOM (2016). NATURE and HYCOM SSH anomalies were derived by removing each model 2016 mean SSH from every 3 h snapshot.

simulation are non-deterministic so the NATURE and OSSE simulations are realistically different from one another over the experiment time period. There are many examples of successful experiments using this methodology (e.g. Jacobs et al., 2014a; Li et al., 2015a; Carrier et al., 2016). Ultimately, the study makes the strong hypothesis of a perfect model, whereby having a perfect set of observations at all model grid points would provide the perfect forecast. This removes an independent variable from the OSSE error analysis (i.e. model error), and allows us to focus on the impact of the different observation types. For the experiments presented here, the OSSE initial condition was perturbed by offsetting the initial condition by one year. The NATURE state on December 1, 2016 was used as the initial condition of the OSSEs starting December 1, 2015. The validity of the initial condition perturbation approach was tested by generating a parallel, non-assimilative experiment using this new initial condition called the Free Run. The Free Run was spun up for the entirety of December 2015 beginning with the altered initial condition and then run out for all of 2016, which is the same time frame as NATURE. The Free Run was then compared with NATURE during 2016 to demonstrate that the two solutions do indeed differ due to non-deterministic processes.

Fig. 8 shows mean absolute error (MAE) between NATURE and the Free Run every 3 h for SSH, surface speed, and MLD at all grid points within the simulation domain with water depth of at least 1000 m. Shelf dynamics are largely deterministic, so we focus on deep water variability where mesoscale and submesoscale nonlinearities should cause the two simulations to differ substantially over time. SSH errors oscillate around a RMS of 9 cm with a minimum of 5.5 cm and maximum of 12 cm (Fig. 8a). For reference, the RMS of NATURE SSH anomalies for all of 2016 was calculated and found to be 13 cm. This suggests that Free Run SSH errors with respect to NATURE are on average more than half of the signal. Surface speed MAE features less variability around a RMS of 20 cm s^{-1} (Fig. 8b). The RMS of NATURE speed anomalies are found to be 25 cm s^{-1}, which is very close to the Free Run error amplitude. Initially, MLD MAE is large (~30 m), but then a sharp decline

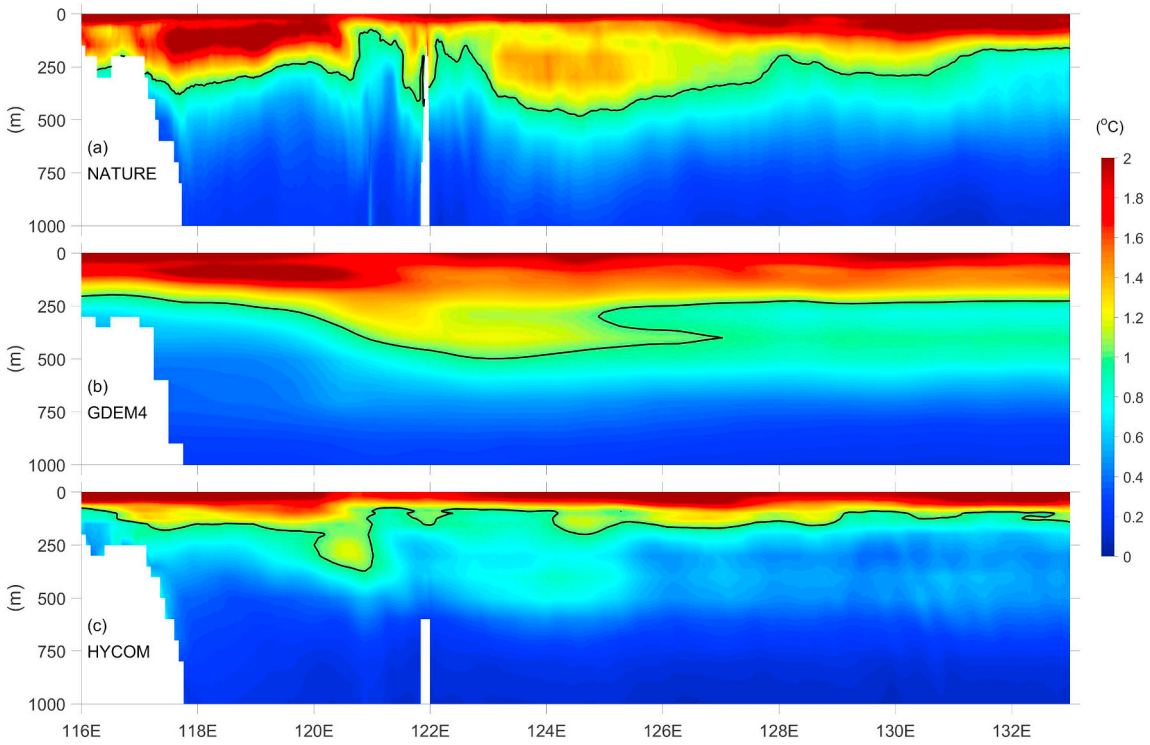

Fig. 7. Longitude-depth transect at 21°N of annual temperature standard deviation (°C) from (a) NATURE (2016), (b) GDEM4 climatology, and (c) global HYCOM (2016). The solid-black line denotes the 1 °C standard deviation contour.

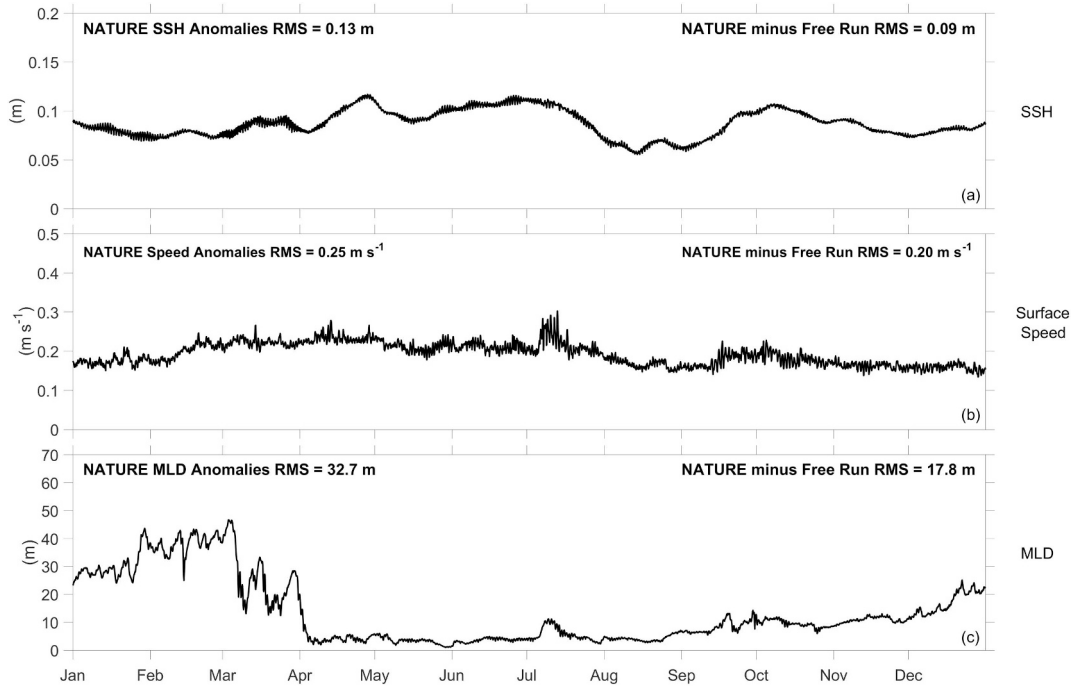

Fig. 8. Mean absolute error (MAE) between NATURE and the Free Run at each 3 h time step in water depths > 1000 m for (a) sea surface height (SSH; m), (b) surface speed (m s^{-1}), and (c) mixed layer depth (MLD; m). NATURE anomaly root mean square (RMS) for all time steps in 2016 at all grid points with at least 1000 m water depth is shown in the top left corner of each subplot. NATURE anomalies were derived by removing the NATURE 2016 mean from each 3-h snapshot. The RMS of the NATURE minus Free Run time series is shown in the top right corner of each subplot.

in error is observed starting in March (Fig. 8c). This is not due to the Free Run dramatically converging towards NATURE, but instead due to seasonal variations in the mixed layer. Boreal winter is characterized by deep mixed layers and therefore larger errors across the domain. As the region transitions towards boreal spring and summer, stratification increases in the near surface due to increased solar radiation and MLD becomes shallower. This causes a natural reduction of error until December when the transition towards boreal winter causes mixed layers to deepen again. The RMS of NATURE MLD anomalies was found to be 17.8 m and Free Run error RMS was found to be 32.7 m. The analyses performed demonstrate that the use of a perturbed initial condition for the Free Run causes it to differ from NATURE throughout the 2016 forward integration period with expected magnitudes. Thus, error reduction in the assimilative runs is attributed to assimilation and not a steady decline in error due to the identical numerical model used in all runs.

With confidence that the methodology outlined above does produce a different solution, the next step was to sample NATURE to generate simulated observations. For observational datasets currently in operation (SST, in situ, and altimetry), NATURE was sampled at real observation locations and times throughout 2016. Simulated SST and in situ observations were assumed to be without error (i.e. no noise was added to the sampled observations). Errors were not added to these data because the impact of the altimetry data is desired. NATURE was sampled at Jason-2, AltiKa, and CryoSat-2 mission observation locations and times, and error was added to the simulated observations based on a random Gaussian distribution with a standard deviation of 3 cm. To generate simulated SWOT observations, we employed the Jet Propulsion Laboratory's (JPL) SWOT simulator (Gaultier et al., 2016). Version 2.0.0 of the simulator was used to sample NATURE on a 2 km grid in both along- and across-track directions. NATURE was output in netcdf format in 3 hourly time intervals. The SWOT simulator interpolated the model fields in space and time to obtain the simulated observations on the SWOT grid along the 21-day repeat tracks. The simulator includes functionality to add random realizations of simulated error to the SWOT observations based on an estimated SWOT error spectrum. These errors include KaRIN noise, roll error, phase error, timing error, baseline error, and wet troposphere error (Gaultier et al., 2016). These simulated errors are known to currently be conservative (i.e. expecting worst-case errors), especially for errors with long wavelengths (i.e. roll, phase, and baseline errors) (Ubelmann et al., 2018; https://spark.adobe.com/page/OKWKAikjWvm0E/). 5000 random realizations of simulated SWOT errors using version 2.0.0 of the SWOT simulator were generated and found to have an error standard deviation of 6.5 cm (not shown). A proposed solution is to reduce the magnitude of these strongly correlated errors by performing crossover calibration (Ubelmann et al., 2018). With such large simulated errors and with ongoing work being performed to optimize the simulator, we chose to apply the same random Gaussian errors that were applied to the simulated nadir altimetry data to the simulated SWOT observations. Therefore, this experiment quantifies differences in analysis/forecast skill due only to differences between altimetry and SWOT sampling with a prescribed white noise level (i.e. resolution, swath width, and sampling frequency).

2.4. Data assimilation

A 3DVAR data assimilation was used in the experiments in a manner similar to that used in the operational global HYCOM system. While the more advanced 4DVAR scheme is becoming more widely used in oceanography, particularly in regional nests (e.g. Ngodock and Carrier, 2014), the 3DVAR provides a lower computational cost to enable us to perform the multiple experiments run at such high resolutions. 4DVAR systems are expected to provide greater skill and the magnitude of this increased skill remains an important research problem for the future. The Navy Coupled Ocean Data Assimilation (NCODA) system (Cummings, 2005) was used for these experiments. Vertical correlations of temperature, salinity, geopotential, and velocity were provided by the Improved Synthetic Ocean Profile (ISOP; Helber et al., 2013). The ISOP system is described more fully in the following paragraph. Data

assimilation was performed daily at 00z, using the previous 24 h forecast as the background to be updated. Available in situ, altimetric, and SST observations 12-days, 5-days, and 12 h prior were collected for the analysis, respectively. The temporal and spatial scarcity of the in situ observations (e.g. Argo; http://www.argo.ucsd.edu/) dictate the longer window when compared to the other observation types. Conversely, the short 12-hour window used for SST is due to the relative ubiquity of these observations. Using a 5-day observation window for the altimetric observations might be too short to properly represent the mesoscale field and too long to correctly fit transient submesoscale features. This choice is a compromise based off prior results that tested different sets of observation windows (Jacobs et al., 2014b). Section 4 elaborates on how future work might search for an optimal window to accommodate the different time scales of the mesoscale and submesoscale phenomena. The 12-day and 5-day windows cause individual observations to be assimilated on successive days, a practice that violates optimal estimation theory. Using observations repeatedly, however, improves assimilation performance and is a technique employed by a number of different Global Ocean Data Assimilation Experiment (GODAE) systems (Martin et al., 2007; Cummings et al., 2009; Jacobs et al., 2014a, 2014b). Innovations, or the difference between the background and observations, were derived using the First Guess at Appropriate Time (FGAT) method and used to minimize the cost function to achieve an optimal increment (Cummings, 2005). A 6 h hindcast was performed to incrementally insert the analysis correction. The increment was divided by the total number of time steps in the 6-hour hindcast prior to the 00z time of the analysis, and the divided increment was added to the model tendency at each time step. Inserting the entire increment directly into the full background at 00z can generate spurious gravity waves and inertial oscillations that take time for the forward solution to damp out. The incremental insertion method attempts to mitigate the development of such motions and therefore provide an initial state with lower energy in these extraneous features. Nonetheless, we did observe analysis steps that still featured these transient features (not shown), which suggests some potential influence on our error analysis. Finally, background error variances are a function of the state variable, geographic position, and depth and are based on the GDEM4 climatology. The background error variances are not scaled by the decorrelation length scale that is discussed in Section 2.4.1.

SST and in situ observations were used directly to generate innovations. Nadir altimetry and SWOT observations were treated in a different manner. Using ISOP, altimetric observations were converted into subsurface temperature and salinity anomalies using a one-dimensional variational analysis employing vertical correlations based on a relatively coarse set of in situ observations of mesoscale processes. Climatological temperature and salinity provide additional constraint and allow for full temperature and salinity profiles. The ISOP climatological temperature and salinity were derived from in situ observations collected by the Navy's Master Oceanographic Observation Data Set (MOODS), World Ocean Database (WOD) 2005, and delayed-mode Argo (Helber et al., 2013). The synthetic temperature and salinity were then used to generate innovations, and ultimately increments. The most important ISOP variational constraint is to steric height anomaly, not SSH anomaly. Generally, SSH and steric height are strongly correlated, but many processes cause discrepancies (e.g. wind-driven, high-frequency barotropic flows). Because our experiments were conducted in a controlled environment, we optimized the process by sampling NATURE steric height anomalies instead of SSH anomalies. NATURE steric height referenced to 1000 m was calculated and the GDEM4 mean steric height field was removed to generate anomalies.

2.4.1. Analysis decorrelation length scale and super-observations

An important element of the 3DVAR assimilation procedure is the analysis decorrelation length scale (i.e. decorrelation length scale of the background error). The parameter controls the length scale over which analysis increments are spread. A few factors motivate the magnitude of the decorrelation length scale. A numerical grid represents physical features that are approximately 10 grid points in size. Hence, the system should not attempt to correct features that the model cannot reasonably represent, and this dictates a lower bound on the scale. Additionally, the assimilation system should correct features that are believed to be the primary error sources. Historically, these have been mesoscale eddies (Cummings et al., 2009). The Rossby radius of deformation describes the length scale for which planetary vorticity is primarily balanced by pressure (i.e. geostrophy) and therefore provides a convenient, physically relevant constraint for fitting mesoscale features. Thus, the decorrelation length scale is often taken to be related to the deformation radius (Cummings, 2005). These considerations have worked in concert in the past.

The SWOT observations provide information on scales below the deformation radius and therefore considerations must be made to deal with the high density of these data with respect to the decorrelation length scale. As more observations cluster within a decorrelation length scale, the computational cost of the observation preparation and deriving the minimization increases and in an extreme case could cause the solution to never converge due to practical limitations (e.g. round off errors). Additionally, the nadir altimeter and SWOT surface observations are converted into synthetic temperature and salinity profiles by the ISOP system. Vertical representation error becomes problematic as more small-scale information is introduced into the assimilation by the high-resolution observations that SWOT provides. ISOP is based off a mesoscale climatology, which does not accurately represent smaller-scale phenomena such as balanced submesoscales and internal gravity waves that are explicitly represented in our modeling framework (and thus also within the observations sampled from NATURE). These issues provide motivation to thin the observations.

There are two approaches to reducing the number of observations. One is thinning that simply selects a subset of the available observations. The second approach is through the creation of a 'super-observation', whereby a weighted average is derived over the local analysis decorrelation length scale and treated as the true observation to be assimilated. The latter methodology is used in this experiment. The weighting is Gaussian using distance from the centroid of the data going into the super-observation with an e-folding scale of the local analysis decorrelation length scale. A simple observation operator maps the control variables from the nearest model grid point to the centroid over which the super-observation was created to calculate innovations. The super-observation reduces observation error and maintains features larger than the decorrelation scale. Therefore, the super-observation produces an observation that is consistent with the analysis scales and the types of physics that ISOP is able to reproduce.

For SST, a set of super-observations was derived and passed directly to NCODA for assimilation. For nadir altimetry and SWOT data, a set of super-observations was derived and then passed to ISOP to generate synthetic temperature and salinity profiles at the centroid of the data over which the super-observation was taken. This drastically reduces the number of observations that are assimilated into the system (Fig. 9). The full observation set (Fig. 9a–b) includes SWOT observations on a 2 km along- and across-track grid. Fig. 9c–d shows how using a mean 15 km decorrelation length scale affects the volume of observations after super-observations are constructed. Using a larger average analysis decorrelation length scale (30 km; Fig. 9e–f) results in a smaller number of super-observations.

The errors incurred by inferring synthetic temperature and salinity using the default 30 km decorrelation length scale super-observations (Fig. 9e–f) of steric height anomalies between January and March 2016 are shown in Fig. 10. Error in the observed steric height anomalies were taken into account when the ISOP synthetics were created. Thus, the ISOP errors shown in Fig. 10 are a function of both measurment error and fitting imperfect temperature and salinity profiles. For both temperature (Fig. 10a) and salinity (Fig. 10b), the comparison between NATURE and ISOP synthetic data fall primarily along the 1:1 line.

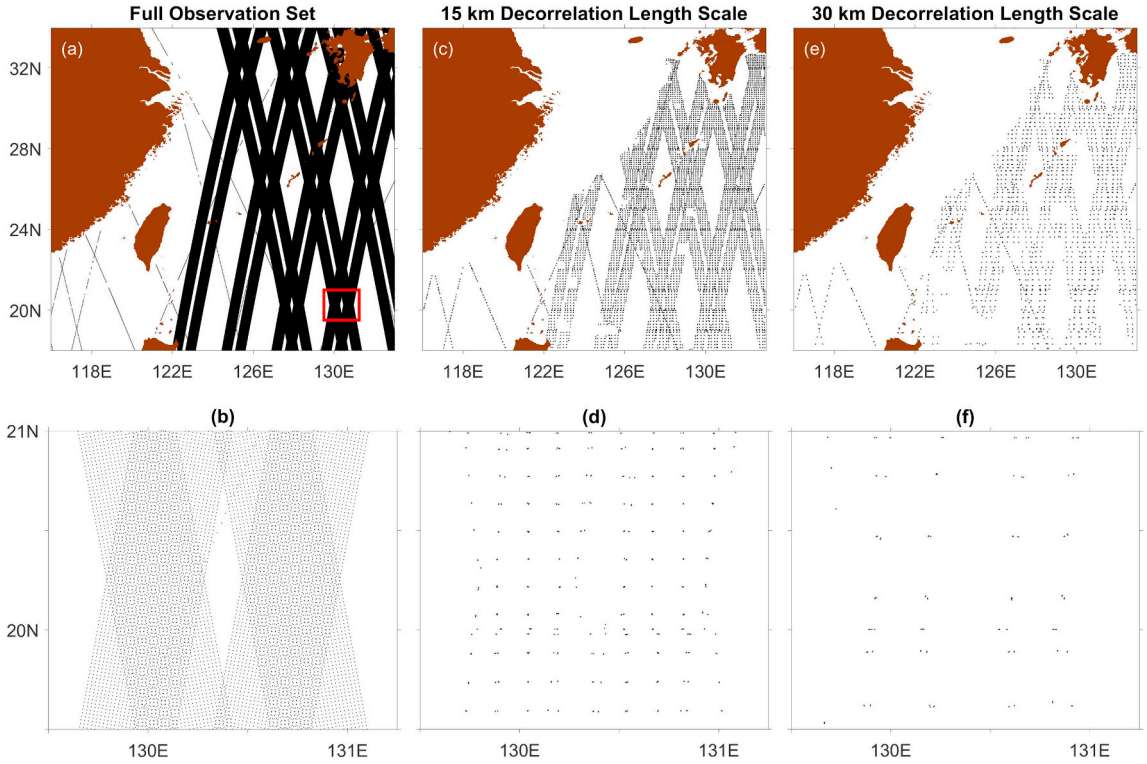

Fig. 9. (a) Full set of simulated nadir altimeter and SWOT observations over an example 5-day observation window. (b) Zoom in of the simulated nadir altimeter and SWOT observations over the subregion highlighted by red in (a). The full set of data was thinned using the super-observation technique over a mean analysis decorrelation length scale of (c–d) 15 km and (e–f) 30 km. Synthetic temperature and salinity profiles were not created for water depths < 150 m causing those surface observations to be absent in (c–d) and (e–f).

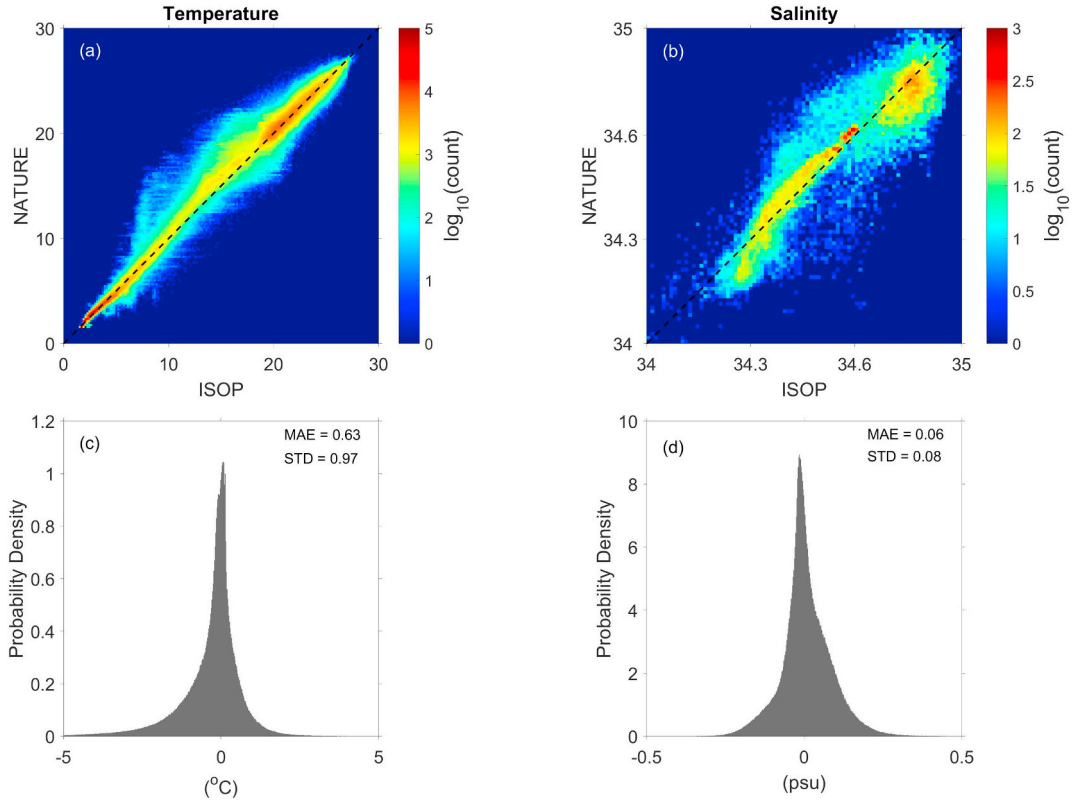

Fig. 10. Log-scale scatter diagram of (a) NATURE temperature vs. ISOP synthetic temperature and (b) NATURE salinity vs. ISOP synthetic salinity. Probability density of (c) ISOP temperature errors and (d) ISOP salinity errors.

Table 1
Description of OSSE experiments differentiated by the types of data that were assimilated.

	SST	In situ	Altimeter	SWOT
NATURE	None	None	None	None
Free Run	None	None	None	None
Altim	On	On	On	None
SWOT	On	On	None	On
Altim + SWOT	On	On	On	On

Temperature statistics show that ISOP synthetic profiles have a MAE of 0.63 °C and an error standard deviation of approximately 1 °C (Fig. 10c). ISOP synthetic salinity profiles have a MAE of 0.06 psu and an error standard deviation of 0.08 psu (Fig. 10d). These errors are consistent with prior validation comparing ISOP derived synthetic temperature and salinity with in situ observations (Helber et al., 2013). The non-Gaussian distribution of both temperature and salinity errors is an undesirable result, but further investigation of this issue is outside the scope of this work.

3. Results

Table 1 provides a brief summary of each of the different experiments conducted, differentiated by the types of data incorporated into the daily analysis cycle. Four OSSE experiments were performed and evaluated: the Free Run, Altim, SWOT, and Altim + SWOT. All four OSSEs started with the same initial condition on December 1, 2015 by using the state from NATURE on December 1, 2016. All experiments were integrated forward without assimilation to January 1, 2016. Starting January 1, 2016, cycling assimilation was performed every 24 h, with the exception of the Free Run which did not assimilate any information. The Altim run assimilated SST, in situ profiles, and traditional nadir altimetry. The SWOT experiment is similar to Altim except that it assimilated SWOT data instead of nadir altimetry data. Finally, Altim + SWOT assimilated all possible observation types used in this study. The assimilative OSSEs (Altim, SWOT, and Altim + SWOT) used the default mean decorrelation length scale of 30 km for the initial set of results shown in Sections 3.1 and 3.2. Each OSSE was integrated forward until March 31, 2016 and then evaluated relative to NATURE.

Error statistics were computed for three variables: SSH, MLD, and 100 m temperature. SSH defines the surface pressure field and is a depth-integrated quantity approximating the underlying baroclinic structure that is primarily controlled by mesoscale processes, but is also influenced by smaller-scale variability such as balanced submesoscales and internal gravity waves. MLD is a complicated variable that is a function of both large-scale forcing and local frontogenesis. Model skill with respect to this variable is essential for accurately representing oceanic variability and air-sea exchange. Finally, 100 m temperature provides an estimate of internal ocean variability. The maximum mean depth of the mixed layer in this region is < 100 m (Fig. 3), which allows the use of 100 m temperature to quantify model skill with respect to thermocline variability. Deeper variables (e.g. 1000 m temperature) could also be considered. However, balanced submesoscales are primarily confined within the mixed layer and upper thermocline (Bachman et al., 2017), and therefore any deep-water analysis would feature only mesoscale and internal gravity wave variability. Using these surface and near surface variables allows us to focus on the widest continuum of ocean physics present in this modeling framework. The comparison of results begins with area-averaged errors to gain a broad view of model skill, and then the evaluation focuses on wavenumber spectra to determine constrained spatial scales.

3.1. Area-averaged errors: 30 km decorrelation length scale

Fig. 11 shows MAE between NATURE and each OSSE for 100 m temperature, SSH, and MLD. The MAE includes information at a location only if the water depth is at least 1000 m. The Free Run errors are consistently higher than any of the assimilative experiments for all three variables tested. As the assimilation process begins on January 1, a sharp decrease in errors is observed for both 100 m temperature and SSH. The initial rate of error decrease correlates with the quantity of observations used in the assimilation process. In the SSH MAE, the Altim + SWOT errors decrease most rapidly followed by the SWOT experiment. Because SSH is a direct observation, this provides verification that the quantity of data is constraining the system more rapidly. By January 11, the Altim experiment error levels are roughly equivalent to those of the SWOT and Altim + SWOT experiments. All three assimilative experiment error levels then oscillate around a MAE of 0.65 °C for 100 m temperature and 5 cm for SSH. With respect to MLD, the assimilative experiment errors are generally lower compared to the Free Run, but the difference between the Free Run and the assimilative run MAE is smaller than the other variables.

3.2. Wavenumber spectra: 30 km decorrelation length scale

The bulk statistics do not finely differentiate the observation system impacts. The spectra of ocean variables tend to be red and it is not possible to determine if smaller scales are more accurately represented by the area-averaged errors. Therefore, we quantify the wavelengths that are constrained within the OSSEs through wavenumber spectral analysis. Two dimensional power spectral density (PSD) was calculated for each 3 hour time step in the square subdomain shown in Fig. 1b. The time series of two dimensional PSD was then averaged into a single two dimensional field in k_x, k_y wavenumber space. The time-averaged PSD was then averaged radially to produce a one dimensional spectrum. The time and space averaged 100 m temperature, SSH, and MLD NATURE spectra for 2016 are shown in Fig. 12. All three variables have an unbroken cascade of PSD from the largest observed scale (640 km) down to the Nyquist wavelength (2 km). Each variable has a different spectral slope between the 200 km and 10 km wavelength range. SSH produces the steepest slope, followed by 100 m temperature, and then MLD. These slopes are consistent with previous studies that examined the same variables at comparable model resolutions (Capet et al., 2008; Richman et al., 2012).

Spatial gradients accentuate small-scale variability. Temperature is heavily influenced by horizontal stirring caused by mesoscale and submesoscale velocities; the horizontal gradient of SSH being a reasonable approximation of the velocity vector. Therefore, 100 m temperature PSD has a flatter slope as much of its spatial variability is tied to intermediate to small-scale horizontal motions. MLD is driven by mesoscale and submesoscale processes controlling stratification, stirring, and frontogenesis; the latter being responsible for convergence/divergence along eddy fronts. The frontal processes are a derivative quantify of the horizontal velocity and buoyancy fields. MLD is therefore characterized by even greater small-scale variability and subsequently has a flatter slope than both SSH and 100 m temperature. These dynamics will be important considerations later when discussing the ramifications of 3DVAR data assimilation in this high-resolution modeling environment. The spectra of many other oceanic variables could be analyzed as well. For example, kinetic energy, an important metric, could be analyzed, but due to the nature of our assimilation method of fitting temperature and salinity synthetics, velocity is indirectly altered by corrections in surface and subsurface pressure (i.e. through geostrophic correlations). Therefore, any error trends in the pressure field will simply be extrapolated onto the resulting velocity. In the subsurface, temperature and salinity primarily act as passive tracers and therefore have similar spectral slopes (not shown). For these reasons, redundancy in the results is reduced by focusing on 100 m temperature, SSH, and MLD, which are each controlled by a unique set of physical processes.

Conceptually, we separate the wavenumber space into two domains:

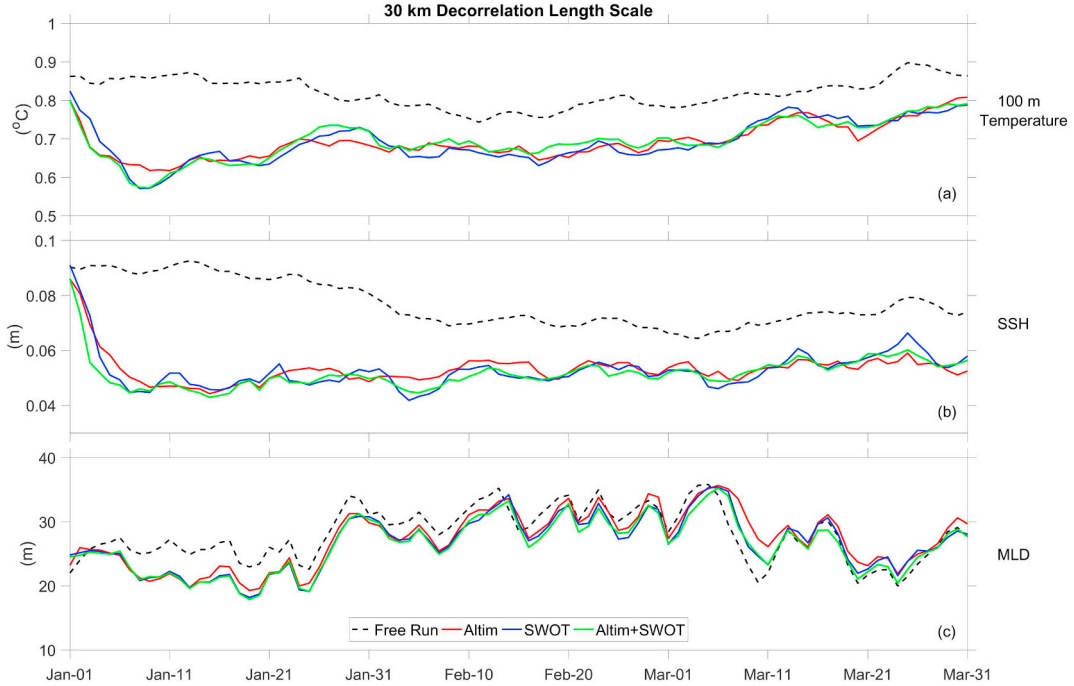

Fig. 11. Mean absolute error (MAE) between NATURE and each experiment at all grid points with at least 1000 m water depth at each 00z time step for (a) 100 m temperature (°C), (b) sea surface height (SSH; m), and (c) mixed layer depth (MLD; m). See Table 1 for descriptions of each experiment.

larger scales (small wavenumbers) that are constrained by the available observations and smaller scales (larger wavenumbers) which are unconstrained. An appropriate assimilation should place constrained OSSE features in locations and times that are similar to those in NATURE. The spatial scales constrained by each OSSE may be determined more clearly by differencing fields of each OSSE from NATURE, which provides the error in the OSSE. The amplitude of the errors in the constrained features should be reduced below the amplitude of the features, and the spectral energy should also be removed from the constrained scales in the PSD of the errors. At small scales, unconstrained features in the OSSE experiments will not be at the same locations and times as in NATURE. The result is that amplitudes of errors at unconstrained scales become greater than the amplitude of the features and spectral energy is increased in the unconstrained scales of the PSD of the errors. We can observe this effect most clearly by normalizing the error spectrum:

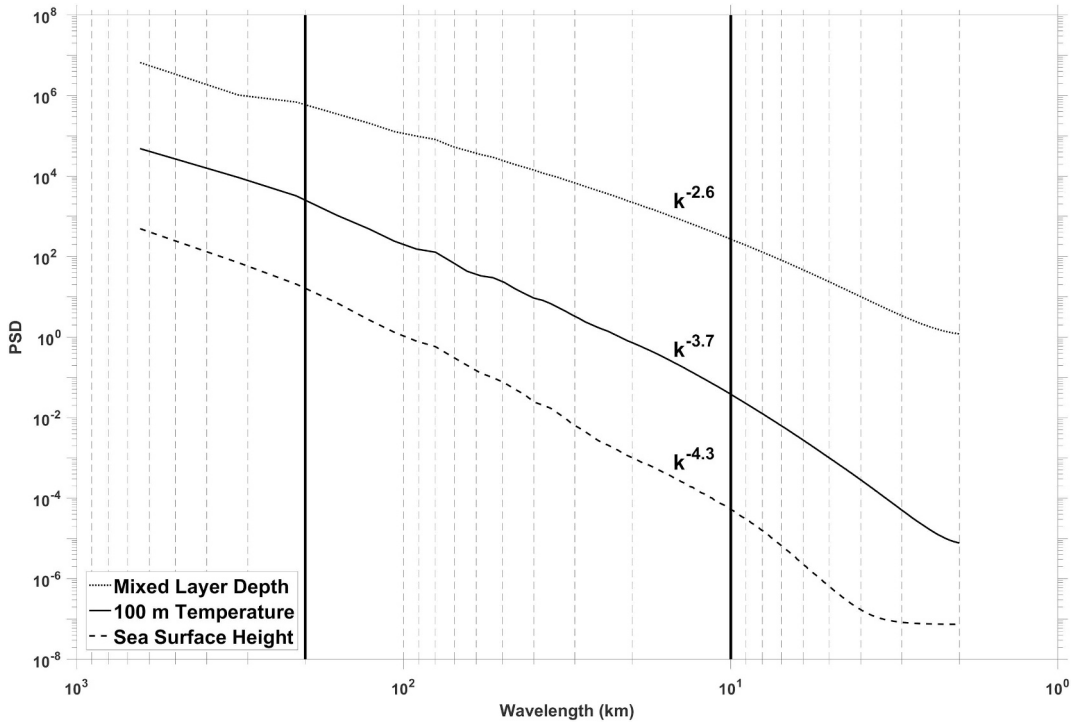

Fig. 12. NATURE wavenumber spectra averaged over all time steps in 2016 for mixed layer depth (MLD; m^2/cpkm), 100 m temperature ($°C^2$/cpkm), and sea surface height (SSH; m^2/cpkm). The indicated spectral slopes were derived over the 200 km–10 km wavelength range.

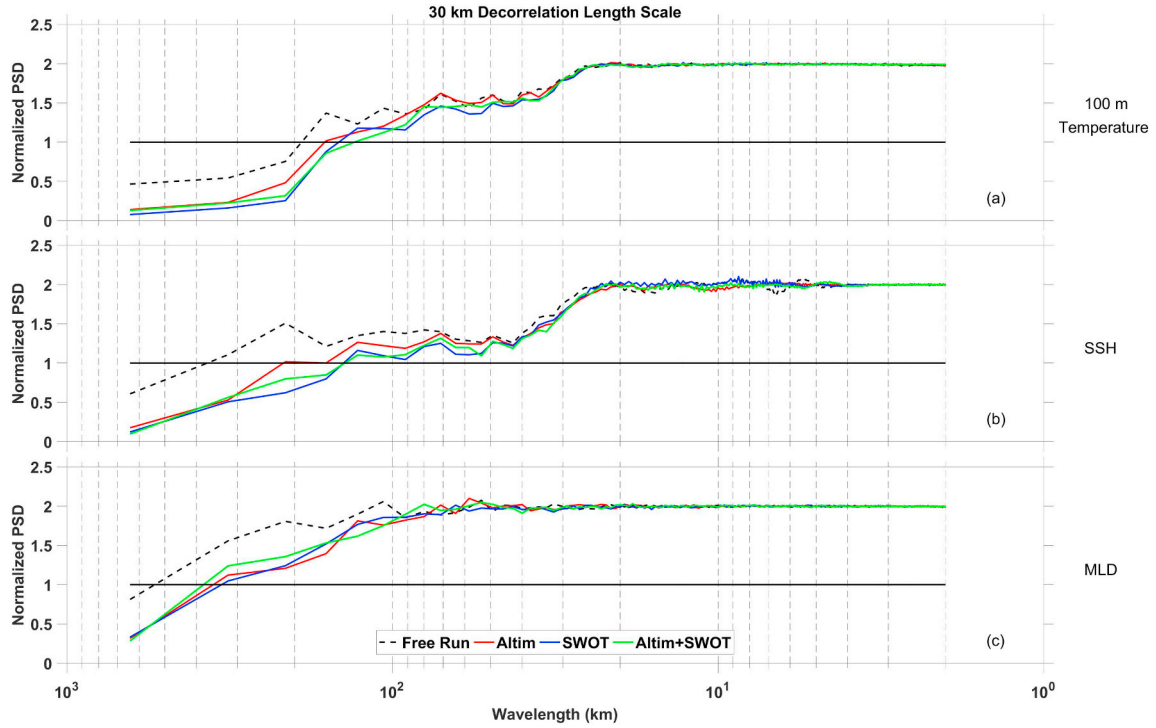

Fig. 13. Normalized wavenumber spectra of (a) 100 m temperature, (b) sea surface height (SSH), and (c) mixed layer depth (MLD) differences from NATURE. See Eq. (1) for more information on the normalization scheme.

$$\frac{\varepsilon_{OSSE}}{\langle \gamma_{NATURE}, \gamma_{OSSE} \rangle} \tag{1}$$

where ε_{OSSE} is the PSD of the error (NATURE minus OSSE), γ_{NATURE} is the PSD of NATURE, γ_{OSSE} is the PSD of the OSSE, and brackets denote the mean of the two spectra. This normalization creates a convenient metric for measuring the error of each experiment as a function of spatial scales. At a particular wavenumber, if the normalized value is zero, NATURE and the respective experiment share exactly the same features in that spatial scale. If features are completely unconstrained, OSSE and NATURE features are uncorrelated random variables and the PSD of the error is the sum of the PSD of the OSSE and NATURE. Thus, a normalized PSD value of 2.0 implies that the assimilation has not impacted features at that particular spatial scale. A normalized PSD value of 1.0 indicates that the correlation between the OSSE and NATURE is 0.5, and we define this value as the separation point between constrained and unconstrained spatial scales.

The spectral analyses were derived for the four OSSE experiments during the month of February (N = 232 time steps). The results contain several trends (Fig. 13). The Free Run performs the worst with respect to all three variables (100 m temperature, SSH, and MLD) with a normalized PSD below one at only the largest of spatial scales. For 100 m temperature, the Altim experiment produces the next highest errors (Fig. 13a). Both SWOT assimilating experiments have lower errors and cross the normalized PSD threshold of 1.0 at lower wavelengths than the other OSSE experiments. We quantify the performance of each OSSE by evaluating the ratios of integrated spectra:

$$\frac{\int_{k_{min}}^{k_{Nyquist}} \varepsilon_{OSSE}\, dk}{\int_{k_{min}}^{k_{Nyquist}} \langle \gamma_{NATURE}, \gamma_{OSSE} \rangle\, dk} \tag{2}$$

where k denotes wavenumbers, k_{min} is the minimum resolved wavenumber ($1/640\,km^{-1}$), and $k_{Nyquist}$ is the Nyquist wavenumber ($1/2\,km^{-1}$). As the assimilation more accurately accounts for more of the features in NATURE, this value should trend towards zero. Table 2 provides a summary of the ratios, and is consistent with the prior analysis. For 100 m temperature, experiments that assimilated simulated SWOT observations contain less error than both the Free Run and Altim experiments with the SWOT experiment producing the smallest value, which is a 23% decrease in error over the Altim experiment.

Analyses of SSH results in similar conclusions. Visual inspection shows that both experiments using the simulated SWOT observations outperform the Free Run and Altim experiments (Fig. 13b). SSH ratios of integrated spectra, per Eq. (2), demonstrate that the SWOT experiment again performs best with a 19% decrease in error over the Altim experiment (Table 2). Finally, the same analyses are performed for MLD. This variable is the least constrained of the three tested with the Altim, SWOT, and Altim + SWOT producing similar normalized spectra (Fig. 13c). The ratio of integrated spectra shows that the reduction in error in the two SWOT experiments is only 2% when compared to the result from the Altim experiment (Table 2).

Finally, the smallest wavelength at which the normalized spectra cross the 1.0 threshold is determined for each OSSE and for each variable (Fig. 14). For 100 m temperature (Fig. 14a), the Altim experiment constrains wavelengths down to 161 km, a 28 km improvement over the Free Run. The SWOT experiment constrains wavelengths down to 145 km and the Altim + SWOT experiment down to 130 km. This result shows that the experiment using both current nadir altimetry and simulated SWOT observations constrains an additional 31 km over the experiment utilizing only nadir altimetry. This is significant, as it approaches the deformation radius for this geographical region (~50 km) (Chelton et al., 1998). Again, the analysis of SSH produces similar results (Fig. 14b). The Altim + SWOT experiment produces the lowest minimum constrained wavelength of 139 km, a value slightly higher than that produced when analyzing 100 m temperature. MLD produces the largest minimum constrained wavelength of 323 km (Fig. 14c). In each case, the OSSEs that include simulated SWOT observations produce the smallest minimum constrained wavelength. Therefore, the results indicate that including SWOT observations in a highly realistic cycling forecast/analysis system does produce an advancement in forecast skill.

Table 2
Ratio of integrated spectra. See Eq. (2) for more detail.

	30 km decorrelation length scale			15 km decorrelation length scale		
100 m temperature	0.81	0.48	0.37	0.43	0.4	0.42
Sea surface height	0.96	0.59	0.48	0.51	0.53	0.7
Mixed layer depth	1	0.84	0.82	0.82	0.76	0.79
	Free Run	Altim	SWOT	Altim + SWOT	SWOT	Altim + SWOT

3.3. Area-averaged errors and wavenumber spectra: 15 km decorrelation length scale

Gaultier et al. (2016) and Ubelmann et al. (2015) found that two-dimensional fields of SWOT SSH constructed using optimal and dynamic interpolation respectively can resolve scales of 100 km and less. These interpolation schemes are difficult to compare to the combined variational assimilation/primitive equation forecasting performed here. The approach we use relies on dynamic extrapolation into the future from an initial condition. The interpolation schemes, however, benefit from a set of observations prior and posterior to the mapping time. Regardless, SWOT observations clearly contain information on scales smaller than those shown to be constrained by the analysis/forecast system in our first set of experiments (Fig. 14). This prompts a second set of experiments in which the assimilation settings were altered in an attempt to further utilize the high-density SWOT observations.

In our experiments, the horizontal resolution of the forward model is 1 km, and the model can therefore represent scales smaller than the regional Rossby radius of deformation (~50 km; Chelton et al., 1998). In addition, the SWOT observations resolve much smaller scales in two dimensions than the nadir altimeters. This transition to both higher resolution models and observations motivates that the long held mesoscale assumptions built into the current analysis/forecast systems be challenged (e.g. the decorrelation length scale as a function of the deformation radius). In this section, the SWOT and Altim + SWOT experiments were repeated with a mean analysis decorrelation length scale of 15 km. The effects of changing the decorrelation length scale on thinning show more of the data are retained within the analysis (Fig. 9c–d). The less severely thinned data were used to obtain different ISOP temperature and salinity synthetics over a larger set of geographic locations. This should ultimately bias the analysis towards the smaller scale features present in the high-density SWOT data and potentially allow the system to constrain a larger set of wavelengths.

Area-averaged errors and wavenumber spectra were recalculated for the updated analysis/forecast fields to quantify improvement in skill. The Altim experiment was not altered and therefore serves as a useful reference between the previous results using a larger analysis decorrelation length scale (30 km) and those shown in this section using a shorter analysis decorrelation length scale (15 km). Compared with the experiments run with the larger 30 km decorrelation length scale, we observe a slight increase in area-averaged errors for 100 m temperature and SSH and a decrease in errors for MLD (Fig. 15). Wavenumber spectral analysis again provides a finer differentiation between the OSSEs (Fig. 16). In comparison to the experiments using the larger decorrelation length scale, visual inspection suggests little difference or a slight decrease in skill for 100 m temperature and SSH and an increase in skill with respect to MLD. This observation is quantified by the ratio of integrated spectra (Table 2). For 100 m temperature, the ratios are similar between the experiments run with a larger or shorter decorrelation length scale. For SSH, a slight decrease in skill is noted for the SWOT experiment run with the shorter decorrelation length scale, but a substantial decrease in skill occurs for the Altim + SWOT experiment. As observed in the area-averaged errors, the experiments utilizing the shorter decorrelation length scale also produced greater skill in wavenumber space with respect to MLD. Finally, these observations are mirrored in the estimations of the minimum constrained wavelengths

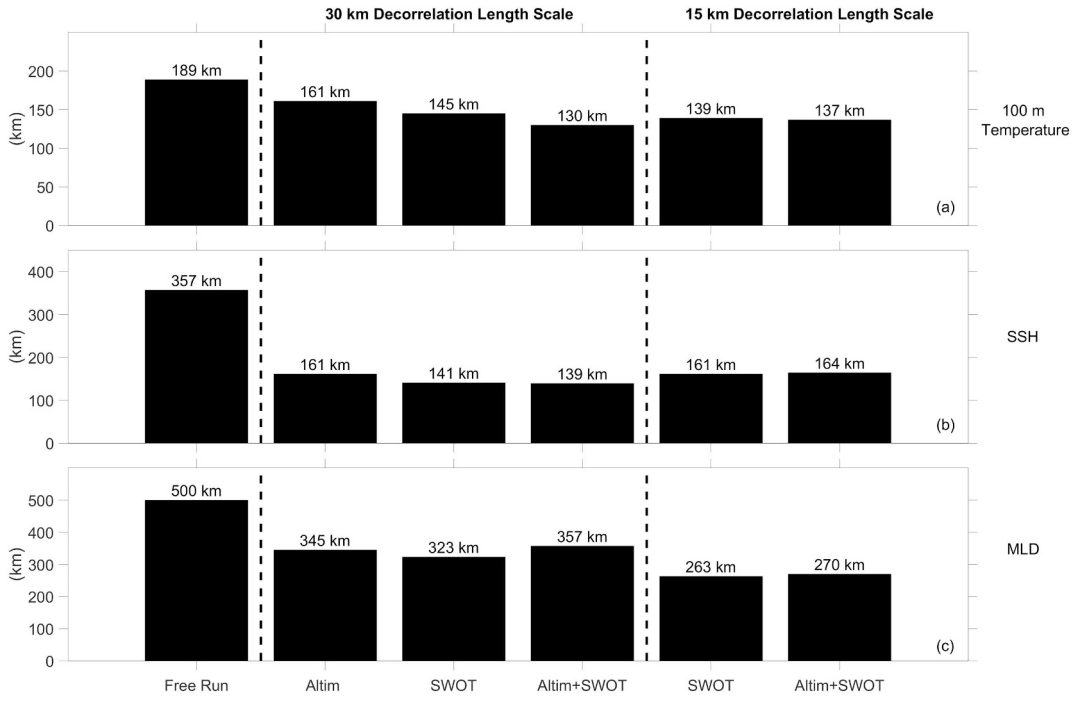

Fig. 14. Minimum constrained wavelength (correlation 0.5) for each experiment when analyzing (a) 100 m temperature, (b) sea surface height (SSH), and (c) mixed layer depth (MLD).

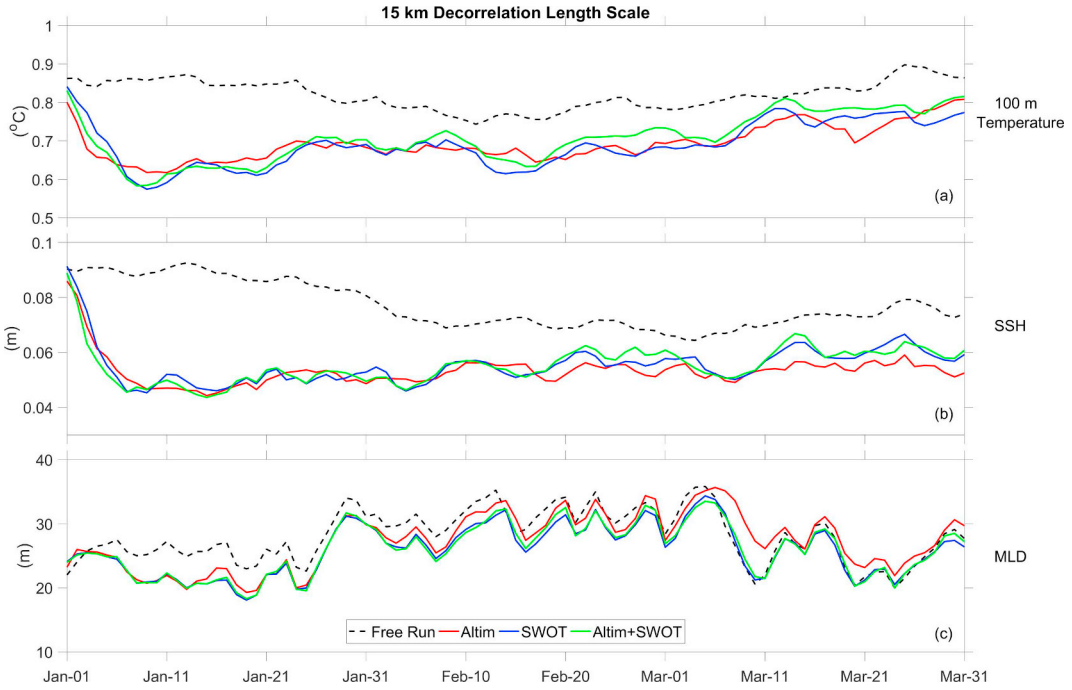

Fig. 15. Same as Fig. 11, except experiments SWOT and Altim + SWOT used a decorrelation length scale reduced by half.

(Fig. 14). The minimum constrained wavelength for 100 m temperature is slightly lower for the SWOT experiment and slightly higher for the Altim + SWOT experiment when rerun with the shorter decorrelation length scale. For SSH, both experiments using the shorter decorrelation length scale produce a larger minimum constrained wavelength when compared to the same experiments run with the larger decorrelation length scale. For MLD, the experiments using the shorter decorrelation length scale produces a minimum constrained wavelength up to 87 km smaller than the same experiments run with a larger decorrelation length scale. Overall, we observe that decreasing the decorrelation length scale reduces skill with respect to 100 m temperature and SSH, while increasing skill for MLD substantially. The apparent dichotomy of these results is explored in Section 4.

4. Discussion

Reducing the analysis decorrelation length scale in an attempt to further utilize the small-scale information within the high density SWOT observations resulted in conflicting results. A reduction of skill occurs for 100 m temperature and SSH, while an increase in skill occurs for MLD. Fig. 12 demonstrates that 100 m temperature and SSH have steeper spectral slopes and therefore a relatively higher concentration

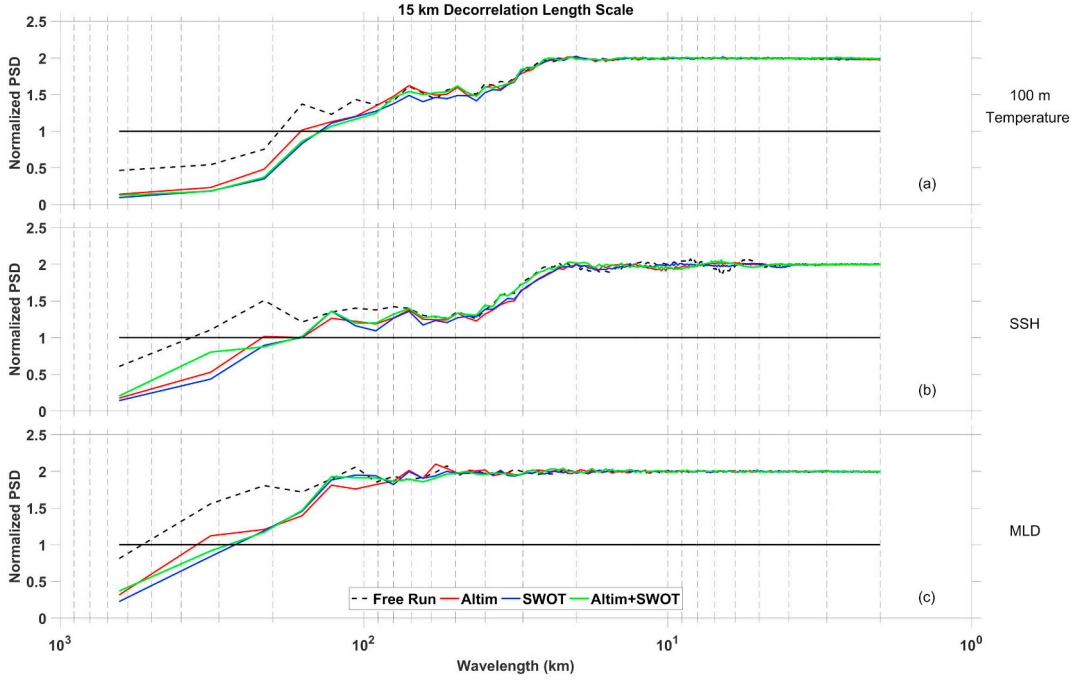

Fig. 16. Same as Fig. 13, except experiments SWOT and Altim + SWOT used a decorrelation length scale reduced by half.

of energy at larger wavelengths. MLD, however, has a flatter slope and therefore a relatively higher concentration of energy at smaller wavelengths. The results demonstrated in Section 3.3 therefore suggest that reducing the decorrelation length scale preferentially sets small-scale features at the expense of the large-scale phenomena. The reverse is true when using a relatively large analysis decorrelation length scale.

This situation has been addressed by implementing a multi-scale assimilation process for 3DVAR (Muscarella et al., 2014; Li et al., 2015a, 2015b; Miyazawa et al., 2017). These multi-scale systems use a two-step assimilation method whereby large-scale information from observations are fit to the background using a larger decorrelation length scale in the first pass. This analysis is then used as the background in the second step in which small scale information is assimilated using a smaller decorrelation length scale. This allows the analysis to compute a correction to both the large-scale and small-scale without compromising skill in one or the other. The initial results reported here indicate an advancement in skill when using SWOT observations and are consistent with the expected effects of a single-scale analysis. Examination of results derived using SWOT observations in a 3DVAR multi-scale analysis system would be an appropriate next step. This work should include investigations into the optimal decorrelation length scales used in each of the analysis steps. Our results suggest that larger, intermediary, and smaller decorrelation length scales affect ocean variables differently and some balance may be required depending on the needs of the user. Additionally, more than two assimilation steps might be considered so that no compromise is necessary. The drawback would be the additional computation cost of running the extra analyses. The selection of the observation window may also be critical. Here, a 5-day window for the nadir altimetry and SWOT observations was used, but both longer and shorter windows have merit. A longer window may allow a much more accurate representation of the mesoscale field especially when considering the relatively long 21-day repeat period SWOT will have. This long window, however, will severely misfit small-scale features that generally have shorter time scales and will no longer be at their previously observed location at the analysis time. A shorter time window would mitigate this issue, but would then generate an inferior representation of the mesoscale field. Finally, mesoscale and submesoscale physics are fundamentally different (McWilliams, 2016). The two-step assimilation should also alter the dynamical balances applied to the analysis increments to account for the physics of the features that each analysis step is trying to optimally fit.

A more sophisticated assimilation system could also be used, such as the 4DVAR. By inserting observations along a time-evolving trajectory, the action of the adjoint and tangent linear model dynamically spreads high-resolution information along all available wavelengths. In this framework, a longer observation window is always advantageous, but comes with a much greater computational cost. The 4DVAR, however, still uses an error covariance that is based on a preset decorrelation length scale that will generate an analysis increment that is too smooth to effectively constrain small-scale features. For this reason, there is ongoing work to generate multi-scale 4DVAR solutions as well, which also set each scale independently by partitioning observations and decorrelation length scales (Carrier et al., 2018). Finally, the topic of scale separation has also been approached using ensemble-variational data assimilation schemes (Buehner and Shlyaeva, 2015). These considerations should be a continued area of focus to extend the influence of high density observations such as those from SWOT.

5. Summary and conclusions

A set of OSSEs tested the utility of forthcoming SWOT data in a realistic, submesoscale resolving ocean assimilation and forecasting system. The usefulness of this new data type was first evaluated by calculating regional errors with respect to NATURE from January 1 to March 31, 2016. Experiments including simulated SWOT data converged towards NATURE at a faster rate. Area-averaged errors indicated experiments with and without SWOT observations reached a similar error level after approximately 11 days. A finer differentiation of the OSSEs was obtained using wavenumber spectral analysis of differences between each experiment and NATURE. The OSSEs utilizing the simulated SWOT observations consistently produced lower error in wavenumber space when compared to the experiment that assimilated only nadir altimetry data. With respect to 100 m temperature, the minimum constrained wavelength was found to be 161 km for the experiment using only nadir altimetry as opposed to 130 km for the experiment utilizing both nadir altimetry and SWOT observations. This 31 km improvement in skill is substantial, and suggests that operational ocean analysis/forecast systems can expect notable increases in predictive skill when SWOT data become available.

The smallest constrained wavelength found for all tested variables and all OSSE experiments, 130 km, is greater than the regional Rossby radius of deformation (~50 km), which serves as an estimate of scale separation between mesoscales and submesoscales. It was hypothesized that it may be necessary to reduce the analysis decorrelation length scale in order to fully utilize the high density SWOT observations. A degradation of skill was noted for 100 m temperature and SSH for the SWOT experiments using this reduced decorrelation length scale. An increase in skill for MLD was observed for the same experiments. Based off derived spectral slopes, SSH and 100 m temperature feature a relatively higher concentrations of PSD at longer wavelengths than MLD. This lead us to conclude that reducing the decorrelation length scale improved analysis/forecast skill for variables that feature more small-scale variability at the expense of variables that feature more large-scale variability. The reverse was shown to be true when using a relatively large decorrelation length scale. These findings further highlight the need for multi-scale assimilation solutions when utilizing a suite of observations which include both large-scale and small-scale information.

We have shown that SWOT observations will improve ocean analysis/forecast skill to a substantial degree when they become available. To fully constrain all the wavelengths that SWOT can observe, multi-scale assimilation solutions will need to be implemented. Work is currently underway to build this functionality into the analysis/forecast system used in this study with the intent of fully constraining the mesoscale field while also beginning to generate predictive skill into the submesoscale regime.

Acknowledgements

Joseph M. D'Addezio is supported by the Naval Research Laboratory Cooperative Agreement BAA-N00173-03-73-13-01 awarded to The University of Southern Mississippi. This research is funded by the Naval Research Laboratory base program Submesoscale Prediction of Eddies by Altimeter Retrieval (SPEAR). The authors appreciate constructive comments provided by the anonymous reviews and the editor, which improved the manuscript. The authors thank Prasad G. Thoppil for helping process the surface drifter observations. The authors are also grateful to numerous NRL colleagues for instructive discussions had throughout experimentation.

References

Ananda, P., Yannice, F., Gilles, L., Le Traon, P.-Y., 2006. Improved description of the ocean mesoscale variability by combining four satellite altimeters. Geophys. Res. Lett. 33, 13–16.

Andreadis, K.A., Clark, E.A., Lettenmaier, D.P., Alsdorf, D.E., 2007. Prospects for river discharge and depth estimation through assimilation of swath-altimetry into raster-based hydrodynamics model. Geophys. Res. Lett. 34, L10403. https://doi.org/10.1029/2007GL029721.

Atlas, R., 1997. Atmospheric observations and experiments to assess their usefulness in data assimilation. J. Meteorol. Soc. Jpn 75, 111–130.

Bachman, S.D., Taylor, J.R., Adams, K.A., Hosegood, P.J., 2017. Mesoscale and submesoscale effects on mixed layer depth in the Southern Ocean. J. Phys. Oceanogr. 47,

2173–2188. https://doi.org/10.1175/JPO-D-17-0034.1.
Barron, C.N., Smedstad, L.F., 2002. Global river inflow within the Navy Coastal Ocean Model. In: Proceedings MTS/IEEE Oceans 2002 Conference, pp. 781–786.
Barron, C.N., Kara, A.B., Martin, P.J., Rhodes, R.C., Smedstad, L.F., 2006. Formulation, implementation and examination of vertical coordinates choices in the Global Navy Coastal Ocean Model (NCOM). Ocean Model 11, 347–375. https://doi.org/10.1016/j.ocemod.2005.01.004.
Biancamaria, S., Durand, M., Andreadis, K.M., Bates, P.D., Boone, A., Mognard, N.M., Rodriguez, E., Alsdorf, D.E., Lettenmaier, D.P., Clark, E.A., 2010. Assimilation of virtual wide swath altimetry to improve arctic river modeling. Remote Sens. Environ. 115 (2), 373–381. https://doi.org/10.1016/j.rse.2010.09.008.
Buehner, M., Shlyaeva, A., 2015. Scale-dependent background-error covariance localisation. Tellus A 67, 1–17. https://doi.org/10.3402/tellusa.v67.28027.
Capet, X., McWilliams, J.C., Molemaker, M.J., Shchepetkin, A.F., 2008. Mesoscale to submesoscale transition in the California Current System part I: flow structure, eddy flux, and observational tests. J. Phys. Oceanogr. 38, 29–43. https://doi.org/10.1175/2007JPO3671.1.
Carnes, M., Helber, R.W., Barron, C.N., Dastugue, J.M., 2010. Validation Test Report for GDEM4. NRL Tech Report NRL/MR/7330–10-9271. (Available from NRL, code 7330, Bldg. 1009, Stennis Space Center, MS 39529-5004).
Carrier, M.J., Ngodock, H.E., Smith, S.R., Souopgui, I., Bartels, B., 2016. Examining the potential impact of SWOT observations in an ocean analysis-forecasting system. Mon. Weather Rev. 144, 3767–3782. https://doi.org/10.1175/MWR-D-15-0361.1.
Carrier, M.J., Osborne, J.J., Ngodock, H.E., Smith, S.R., Souopgui, I., D'Addezio, J.M., 2018. A multiscale approach to high resolution profile observations within a 4DVAR analysis system. Mon. Weather Rev. https://doi.org/10.1175/MWR-D-17-0300.1. in press.
Chavanne, C., Flament, P., Luther, D., Gurgel, K.-W., 2010. The surface expression of semidiurnal internal tides near a strong source at Hawaii. Part II: interactions with mesoscale currents. J. Phys. Oceanogr. 40, 1180–1200. https://doi.org/10.1175/2010JPO4223.1.
Chelton, D.B., DeSzoeke, R.A., Schlax, G., Naggar, K.E., Siwertz, N., 1998. Geographical variability of the first baroclinic Rossby radius of deformation. J. Phys. Oceanogr. 28, 433–460.
Cummings, J.A., 2005. Operational multivariate ocean data assimilation. Q. J. R. Meteorol. Soc. 131, 3583–3604.
Cummings, J., Bertino, L., Brasseur, P., Fukumori, I., Kamachi, M., Martin, M.J., Mogensen, K., Oke, P., Testut, C.E., Verron, J., Weaver, A., 2009. Ocean data assimilation systems for GODAE. Oceanography 22 (3), 96–109. https://doi.org/10.5670/oceanog.2009.69.
Ducet, N., Le Traon, P.Y., Reverdin, G., 2000. Global high-resolution mapping of ocean circulation from the combination of TOPEX/Poseidon and ERS-1 and -2. J. Geophys. Res. 105 (8), 19477–19498.
Durand, D., Andreadis, K.M., Alsdorf, D.E., Lettenmaier, D.P., Moller, D., Wilson, M., 2008. Estimation of bathymetric depth and slope from data assimilation of swath altimetry into a hydrodynamic model. Geophys. Res. Lett. 35, L20401. https://doi.org/10.1029/2008GL034150.
Egbert, G.D., Erofeeva, S.Y., 2002. Efficient inverse modeling of barotropic ocean tides. J. Atmos. Ocean. Technol. 19, 183–204.
Fairall, C.W., Bradley, E.F., Hare, J.E., Grachev, A.A., Edson, J.B., 2003. Bulk parameterization of air-sea fluxes: Updates and verification for the COARE algorithm. J. Clim. 16. https://doi.org/10.1175/1520-0442(2003)016%3C0571:BPOASF%3E2.0.CO;2.
Fu, L.L., Ubelmann, C., 2014. On the transition from profile altimetry to swath altimeter for observing global ocean surface topography. J. Atmos. Ocean. Technol. 31, 560–568. https://doi.org/10.1175/JTECH-D-13-00109.1.
Gaultier, L., Ubelmann, C., Fu, L.-L., 2016. The challenge of using future SWOT data for oceanic field reconstruction. J. Atmos. Ocean. Technol. 33, 119–126. https://doi.org/10.1175/JTECH-D-15-0160.1.
Halliwell, G.R.Jr., Srinivasan, A., Kourafalou, V., Yang, H., Willey, D., Le Henaff, M., Atlas, R., 2014. Rigorous evaluation of a fraternal twin ocean OSSE system for the open Gulf of Mexico. J. Atmos. Ocean. Technol. 31, 105–129. https://doi.org/10.1175/JTECH-D-13-00011.1.
Helber, R.W., Townsend, T.L., Barron, C.N., Dastugue, J.M., Carnes, M.R., 2013. Validation Test Report for the Improved Synthetic Ocean Profile (ISOP) System, Part I: Synthetic Profile Methods and Algorithm. NRL Tech Report NRL/MR/7320–13-9364. (Available from NRL, code 7320, Bldg. 1009, Stennis Space Center, MS 39529-5004).
Hogan, P.J., Hurlburt, H.E., 2000. Impact of upper ocean-topographical coupling and isopycnal outcropping in Japan/East Sea models with 1/8° to 1/64° resolution. J. Phys. Oceanogr. 30, 2535–2561. https://doi.org/10.1175/1520-0485.
Hogan, T.F., et al., 2014. The navy global environmental model. Oceanography 27, 116–125. https://doi.org/10.5670/oceanog.2014.73.
Jacobs, G.A., Richman, J.G., Doyle, J.D., Spence, P.L., Bartels, B.P., Barron, C.N., Helber, R.W., Bub, F.L., 2014a. Simulating conditional deterministic predictability within ocean frontogenesis. Ocean Model 78, 1–16. https://doi.org/10.1016/j.ocemod.2014.02.004.
Jacobs, G.A., et al., 2014b. Data assimilation considerations for improved ocean predictability during the Gulf of Mexico Grand Lagrangian Deployment (GLAD). Ocean Model 83, 98–117. https://doi.org/10.1016/j.ocemod.2014.09.003.
Jerlov, N.G., 1976. Marine Optics. Elsevier, pp. 231.
Le Traon, P.-Y., et al., 2015. Use of satellite observations for operational oceanography: recent achievements and future prospects. J. Oper. Oceanogr. 8, s12–s27.
Levy, M., Ferrari, R., Franks, P.J.S., Martin, A.P., Riviere, P., 2012. Bringing physics to life at the submesoscale. Geophys. Res. Lett. 39, L14602. https://doi.org/10.1029/2012GL052756.
Li, Z., McWilliams, J.C., Ide, K., Farrar, J.D., 2015a. Coastal ocean data assimilation using a multi-scale three-dimensional variational scheme. Ocean Dyn. 65, 1001–1015. https://doi.org/10.1007/s10236-015-0850-x.
Li, Z., McWilliams, J.C., Ide, K., Farrara, J.D., 2015b. A multiscale variational data assimilation scheme: formulation and illustration. Mon. Weather Rev. 143, 3804–3822. https://doi.org/10.1175/MWR-D-14-00384.1.
Martin, A.J., Hines, A., Bell, M.J., 2007. Data assimilation in the FOAM operational short-range ocean forecasting system: a description of the scheme and its impact. Q. J. R. Meteorol. Soc. 133 (625), 981–995.
McWilliams, J.C., 2016. Submesoscale currents in the ocean. Proc. R. Soc. A 472, 20160117. https://doi.org/10.1089/rspa.2016.0117.
Miyazawa, Y., Varlamov, S.M., Miyama, T., Guo, X., Hihara, T., Kiyomatsu, K., Kachi, M., Kurihara, Y., Muakami, H., 2017. Assimilation of high-resolution sea surface temperature data into an operational nowcast/forecast system around Japan using a multi-scale three-dimensional variational scheme. Ocean Dyn. 67, 713–728. https://doi.org/10.1007/s10236-017-1056-1.
Munier, S., Polebistki, A., Brown, G., Belaud, G., Lettenmaier, D., 2015. SWOT data assimilation for operational reservoir management on the Upper Niger River Basin. Water Resour. Res. 51, 554–575. https://doi.org/10.1002/2014WR016157.
Muscarella, P.A., Carrier, M.J., Ngodock, H.E., 2014. An examination of a multi-scale three-dimensional variational data assimilation scheme in the Kuroshio Extension using the Naval Coastal Ocean Model. Cont. Shelf Res. 73, 41–48.
Ngodock, H., Carrier, M., 2014. A 4D-Var system for the Navy coastal ocean model part I: system description and assimilation of synthetic observations in the Monterey Bay. Mon. Weather Rev. 142, 2085–2107. https://doi.org/10.1175/MWR-D-13-00221.1.
Oubanas, H., Gejadze, I., Malaterre, P.-O., Mercier, F., 2018a. River discharge estimation from SWOT-type observations using variational data assimilation and the Full Saint-Venant hydraulic model. J. Hydrol. 559, 638–647. https://doi.org/10.1016/j.jhydrol.2018.02.004.
Oubanas, J., Gejadze, I., Malaterre, P.-O., Durand, M., Wei, R., Frasson, R.P.M., Domeneghetti, A., 2018b. Discharge estimation in ungauged basins through variational data assimilation: the potential of the SWOT mission. Water Resour. Res. 54. https://doi.org/10.1002/2017WR021735.
Pedinotti, V., Boone, A., Ricci, S., Biancamaria, S., Mognard, N., 2014. Assimilation of satellite data to optimize large scale hydrological model parameters: a case study for the SWOT mission. Hydrol. Earth Syst. Sci. 18 (11), 4485–4507. https://doi.org/10.5194/hess-18-4485-2014.
Richman, J.G., Arbic, B.K., Shriver, J.F., Metzger, E.J., Wallcraft, A.J., 2012. Inferring dynamics from wavenumber spectra of an eddying global ocean model with embedded tides. J. Geophys. Res. 117, C12012. https://doi.org/10.1029/2012JC008364.
Smedstad, O.M., Hurlburt, H.E., Metzger, E.J., Rhodes, R.C., Shriver, J.F., Wallcraft, A.J., Kara, A.B., 2003. An operational eddy resolving 1/16° global ocean nowcast/forecast system. J. Mar. Syst. 40–41, 341–361.
Su, Z., Wang, J., Klein, P., Thompson, A.F., Menemenlis, D., 2018. Ocean submesoscales as a key component of the global heat budget. Nat. Comm. 9, 775. https://doi.org/10.1038/s41467-018-02983-w.
Teague, W.J., Carron, M.J., Hogan, P.J., 1990. A comparison between the generalized digital environment model and levitus climatologies. J. Geophys. Res. 95 (C5), 7167–7183.
Thoppil, P.G., Richman, J.G., Hogan, P.J., 2011. Energetics of a global ocean circulation model compared to observations. Geophys. Res. Lett. 38, L15607. https://doi.org/10.1029/2011GL048247.
Ubelmann, C., Klein, P., Fu, L.-L., 2015. Dynamic interpolation of sea surface height and potential applications for future high-resolution altimetry mapping. J. Atmos. Ocean. Technol. 32, 177–184.
Ubelmann, C., Gaultier, L., Fu, L.-L., Wang, J., Steunou, N., Picot, N., Dibarboure, G., 2018. Ocean simulator for science applications. In: 2018 Science Team Meeting, . https://spark.adobe.com/page/OKWKAikjWvm0E.
Yoon, Y., Durand, M., Merry, C., Clark, E., Andreadis, K., Asldorg, D., 2012. Estimating river bathymetry from data assimilation of synthetic SWOT measurements. J. Hydrol. 464, 363–375. https://doi.org/10.1016/j.jhydrol.2012.07.028.