Ocean Modelling 159 (2021) 101760

Contents lists available at ScienceDirect

Ocean Modelling

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

Observation and model resolution implications to ocean prediction

Gregg Jacobs [a,*], Joseph M. D'Addezio [a], Hans Ngodock [a], Innocent Souopgui [b]

[a] Naval Research Laboratory, Stennis Space Center MS, USA
[b] University of New Orleans, Department of Physics, LA, USA

ARTICLE INFO

Keywords:
Data assimilation
OSSE
Ensemble
Uncertainty

ABSTRACT

We address ocean modeling capability that has grown exponentially while ocean observation growth has not maintained pace, a situation leading to seemingly degraded forecast skill when model resolution is increased. Skill in predicting ocean instabilities such as mesoscale eddies requires satellite and in situ observations continually correcting numerical model conditions. Observations constrain positions of larger ocean model features, while smaller features are unconstrained. By means of an Observation System Simulation Experiment (OSSE), we show that time–space observation coverage controls the separation of constrained and unconstrained feature scales. Using 1000 independent surface drifters, we show constrained scales have deterministic prediction skill and unconstrained scales predict areas of higher expected errors. The results are shown to be consistent with ensemble forecasts. Separating constrained and unconstrained features, and using information within each appropriately, allows us to manage the present gap between observation and model resolution.

1. Introduction

Ocean forecasting has progressed in recent decades since the beginning of the Global Ocean Data Assimilation Experiment (GODAE) (Smith, 2000). Global and multiple nested area predictions occur in operational centers around the world (Lellouche et al., 2018; Lemieux et al., 2016; Metzger et al., 2014; Rowley and Mask, 2014; Schiller et al., 2020). These systems regularly correct a prior forecast with recent satellite and in situ observations. The satellite sea surface height (SSH) observations are particularly critical for inferring ocean interior thermohaline variability within mesoscale eddies (Le Traon et al., 2017), and the Argo program provides regular in situ observations of the ocean interior (Roemmich et al., 2009). In situ data has increased with deployment of autonomous systems such as gliders (Rudnick, 2016) and development of surface drifter swarms (Novelli et al., 2017). While computational capability, and thus numerical model resolution and complexity, has grown exponentially with Moore's Law, observational sampling capability has not maintained pace. Thus, a significant gap now exists between scales resolved by regular observations and scales numerical models can physically represent. Sandery and Sakov (2017) showed decreased ocean predictive skill as model horizontal resolution increases, and we find the gap between observation and model resolution is a key contributor to this outcome. This can have consequences in emergency response situations such as the Deep Water Horizon incident that used forecasts from high-resolution models to predict surface oil trajectory. The associated ocean-oil modeling effort did not account for the significant model and observation resolution gap, and the result was wide discrepancy between forecasts and observations (Liu et al., 2011).

To illustrate the situation, we examine trajectories of more than 1000 surface drifters from the Lagrangian Submesoscale Experiment (LASER) in the Gulf of Mexico (Özgökmen et al., 2018) relative to a 1 km data assimilative numerical model (Fig. 1a). The cycling assimilation began several months prior to the drifter deployment, and every day the system assimilated all regular observations (satellite altimeter sea surface height, satellite sea surface temperature, near real time in situ profiles, and ship observations). The regular satellite SSH observations over 5 days are sparse relative to the finest scales of features represented by the model. The drifters are independent from the model. At larger scales, there is qualitative correspondence between model and drifters in the prominent Loop Current Eddy (LCE) in the southern area, the southward flow across the Florida shelf, and the southward flow in the range of 24–28°N, 89–90°W. Within smaller scale features, the correspondence degrades as the satellite observations do not resolve these small scales, and thus corrections are not made to the model to constrain the positions of the features. The illustration in Fig. 1a using the 1 km resolution model is not extreme. Prior model results at 250 m resolution in this area contain strong instabilities along the shelf break front and sharpening of fronts that the 1 km model does not represent well (Jacobs et al., 2016). Thus, present model resolution

* Corresponding author.
E-mail addresses: Gregg.Jacobs@nrlssc.navy.mil (G. Jacobs), joseph.daddezio@nrlssc.navy.mil (J.M. D'Addezio), Hans.Ngodock@nrlssc.navy.mil (H. Ngodock), isouopgu@uno.edu (I. Souopgui).

https://doi.org/10.1016/j.ocemod.2021.101760
Received 1 June 2020; Received in revised form 1 December 2020; Accepted 14 January 2021
Available online 29 January 2021

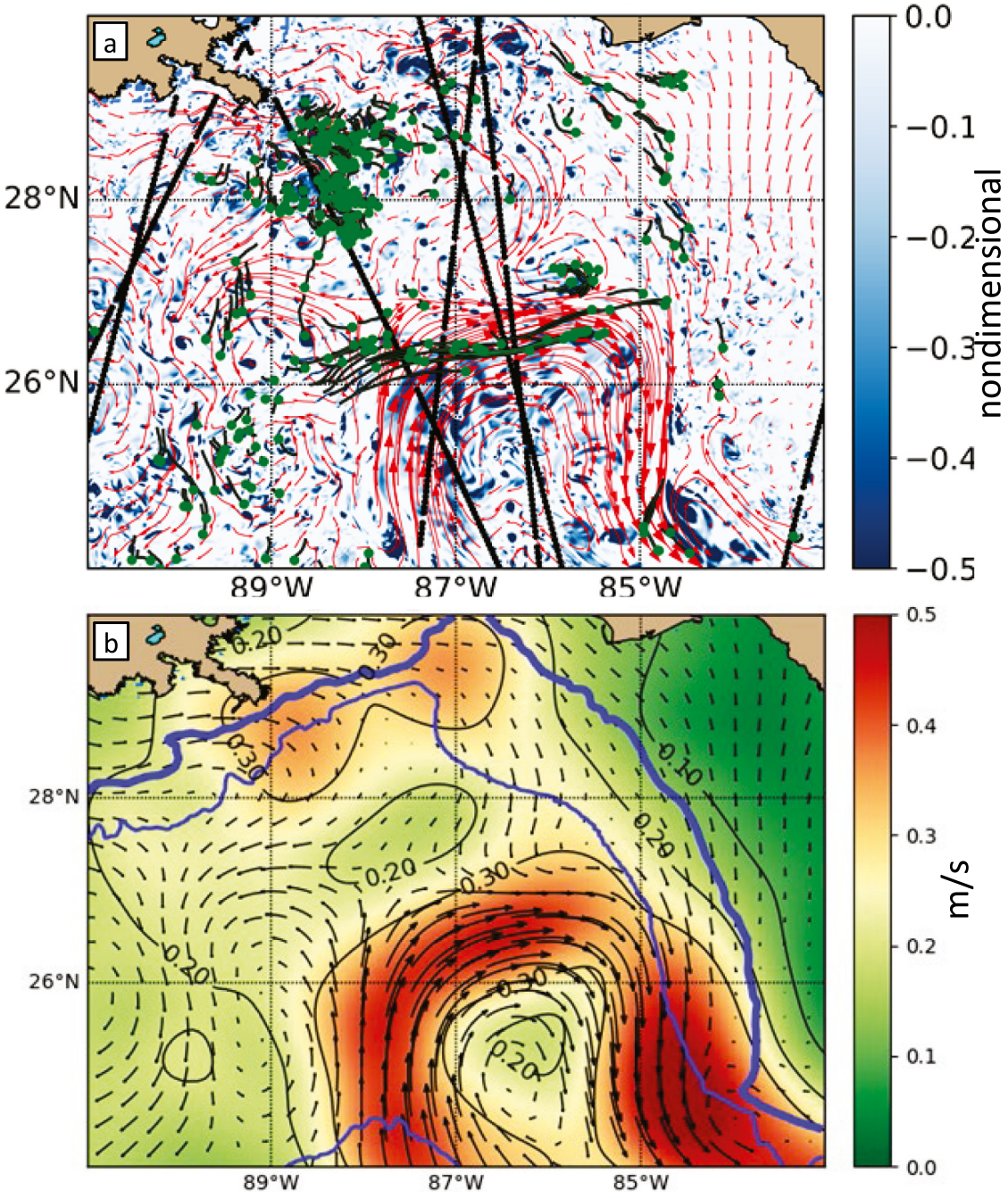

Fig. 1. (a) A 1 km ocean model snapshot on February 26, 2016 shows the large-scale features apparent in the surface currents (red vectors indicating 24 h streamlines), and smaller scale features are apparent in the Okubo–Weiss parameter normalized by the snapshot standard deviation that highlights vortices (background color). Observed 24-hour surface drifter trajectories (black lines with green dots indicating latest position) have correspondence with the model large-scale features, but small-scale features are not consistent with the trajectories. Satellite altimeter observations over the prior 5 days (black lines) correct the model. This satellite coverage is typical, and positions of the observation tracks change day to day. (b) The model currents filtered to retain the constrained scales (black vectors) eliminate the small-scale features. The thick blue line is the 100 m isobath, and the thin blue line is the 1000 m isobath. The small-scale local RMS variability (background color) indicates areas that contain mispositioned strong fronts, submesoscale eddies (as in the Loop Current interior), and shelf break instabilities (in the northern areas).

capability can produce forecasts at scales far smaller than resolved by regular observations.

The situation creates a contradiction within operational centers conducting ocean forecasts. If there is skill only in larger scales, why should a center dedicate oftentimes-significant computer resources to produce forecasts at smaller scales? In this paper, we address the problem by understanding how model and observation resolution control skill across the scale spectrum and what useful information is contained within partitions of said scale spectrum. There are three main aspects, and we place them in a framework aligned with wavenumber spectra (Fig. 2). (1) Larger scales are constrained by observations. Noting wavenumber as k, we define constrained scales as having power spectral density of errors $\epsilon(k)$ less than the power spectral density of the true world $\gamma(k)$, or scales at which $\epsilon(k)/\gamma(k) < 1$. (2) Observations and the data assimilation system determine features that are resolved and corrected, and these control the boundary separating constrained and unconstrained scales. The boundary is defined as $\epsilon(k)/\gamma(k) = 1$. (3) Features within unconstrained scales are misplaced in model forecasts and therefore uncorrelated with the true world features (Fig. 1a).

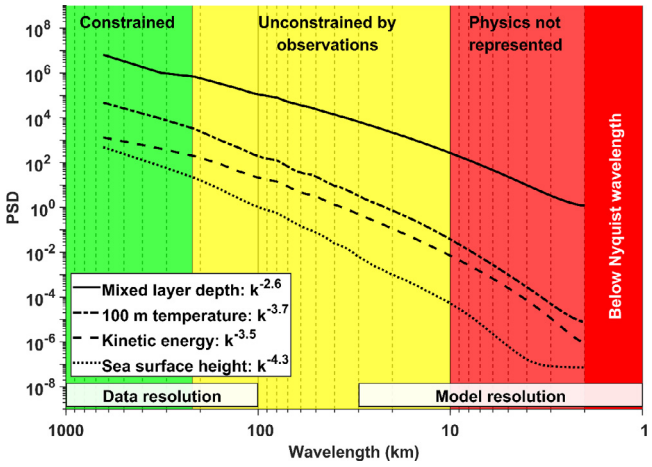

Fig. 2. The power spectral density (PSD) of mixed layer depth (m^2/cpkm), 100 m depth temperature ($°C^2$/cpkm), surface kinetic energy (m^2 s^{-2}/cpkm), and sea surface height (m^2/cpkm) from a 1-year model at 1 km resolution. The respective power-law slopes in the legend were fit over 10 to 200 km wavelengths. Data resolution controls the boundary between constrained and unconstrained, and the boundary value for surface trajectories is used here as an example (220 km) (Jacobs et al., 2020). The model resolution controls the Nyquist wavelength and the separation of the unconstrained band from the band not represented by model physics.

The variance of errors is larger than the variance of the features in unconstrained scales, or $\varepsilon(k)/\gamma(k) > 1$.

The meteorology community has examined this problem in high-resolution models intended to represent convection and rain. Gilleland et al. (2009) overview a range of methods developed to show the skill, and therefore benefit, of high-resolution models. Individual small-scale rain areas may not be accurately predicted, but when measured by different metrics, the character and coverage of events have skillful information. The approach we use here is similar to the 'neighborhood' or 'scale-separation' categories (Gilleland et al., 2010). Contingency diagrams for metrics such as rainfall threshold suffer a double penalty. Small-scale misplaced rain events in a high-resolution model lead to false positives, and small-scale non-rain events lead to false negatives as the events in the model do not align with observed small-scale events. In our case, this effect manifests as error variance at unconstrained scales reaching peak values of twice the true variance, or $\varepsilon(k)/\gamma(k) = 2$. Developing different metrics has avoided the double penalty (Mittermaier and Roberts, 2010). Using only the constrained scales aids in removing the double penalty, and there is valuable information in the unconstrained scales if used appropriately. Just as in the case of high-resolution atmospheric models predicting the general areas of rainfall events, such as along atmospheric fronts, the information of variability at small scales has value in predicting areas in which the small ocean features occur and lead to larger forecast errors. The available ocean observations are not persistent high-resolution gridded data such as radar-observed rainfall rates used in meteorological studies. The LASER data is a very dense data set for oceanography, but it is comparatively sparse and not uniform. This limits our ability to develop metrics, and we restrict ourselves to using RMS error statistics at observation locations. Thus, we proceed with a different approach to show the value of high-resolution model forecasts.

Model skill assessments typically use the full model result as a deterministic forecast. Separating the constrained and unconstrained portions of the forecast allows us to use the constrained portion as a deterministic forecast and the unconstrained portion as an estimate of forecast error. There is confidence in the prediction at constrained scales and uncertainty in unconstrained scales. To substantiate the framework considered here, we provide evidence for several points:

(1) We show that removing the model unconstrained features results in lower errors relative to observations, and we build on prior work that estimated the constrained scales (Jacobs et al., 2020). Errors of velocity relative to the LASER drifters indicate the majority of the error variance is due to unconstrained features rather than errors in the constrained features.

(2) The results demonstrated in (1) are consistent with ensemble forecast methodologies historically used to quantify uncertainty. Computing the mean across the ensemble removes features not correlated between members. The power spectra of the mean across the ensemble contain much less energy at unconstrained scales when compared to the spectra of individual ensemble members.

(3) Observation ability to resolve features controls the boundary between constrained and unconstrained scales. An Observation System Simulation Experiment (OSSE) using simulated Surface Water/Ocean Topography (SWOT) observations provides this demonstration.

(4) The unconstrained portion of the model solution has skill in predicting forecast errors. This is demonstrated by using the unconstrained portion of the model forecast in combination with model errors with respect to the independent 1000 LASER drifters over 3 months. The unconstrained variability is also shown to be consistent with ensemble variance estimates.

The paper is organized as follows. Section 2 provides the setup of model experiments, which include the cycling data assimilative 1 km model, the 3 km ensemble, and the OSSE setup. Section 3 discusses the results separated into the constrained solution (Section 3.1), the boundary controlled by observations (Section 3.2), and the unconstrained solution (Section 3.3). Section 4 discusses the applicability and factors that influence the results. Finally, Section 5 provides overall conclusions and final thoughts.

2. Model experiment setup

2.1. 1 km assimilative model

The Navy Coastal Ocean Model (NCOM) produced our simulated three-dimensional ocean fields throughout the study. NCOM is based on the hydrostatic and Boussinesq approximations, and explicitly represents time-evolving temperature, salinity, velocity, and sea surface height. The model integrates the primitive equations in time using an explicit leapfrog scheme with a split barotropic/baroclinic mode (Barron et al., 2006). The model covers the entire Gulf of Mexico with a horizontal resolution set to 1 km and using 50 sigma/z levels in the vertical (D'Addezio et al., 2019a,b; Jacobs et al., 2020, 2016; Souopgui et al., 2020). The Global Ocean Forecast System (GOFS) based on the HYbrid Coordinate Ocean Model (HYCOM) (Metzger et al., 2014) delivered boundary conditions at 3-hour frequency that were interpolated in time to the NCOM time step. GOFS did not contain tides. A global tide solution (Egbert and Erofeeva, 2002) was the source of tidal velocities and surface heights in addition to the time-interpolated GOFS boundary conditions. A separate atmospheric model within the Coupled Ocean Atmosphere Mesoscale Prediction System (COAMPS) (Hodur, 1997) provided data for computing atmospheric momentum flux, air temperature, humidity, and solar radiation at hourly frequency for surface forcing to the ocean model. Daily freshwater inflow of rivers was represented as a velocity transport into the model domain at land boundary points with a salinity of 0 and velocity distributed over depth. The transports were from the United States Geological Survey (USGS) river monitoring stations using the instruments closest to the model land boundary. For rivers without observations, historical monthly transports were used.

The NCOM model connects to the Navy Coupled Ocean Data Assimilation (NCODA) system that utilizes a 3-dimensional variational (3DVar) approach (Smith et al., 2011). Each day of the model run, recent observations minus the prior model forecast provide innovations, and the 3DVar computes an analysis to correct the model fields. There are several important aspects of the data assimilation process that have been examined and optimized in previous work.

The error autocovariances of the prior forecast used as the background have been examined closely. Details of the error covariance specification are in Jacobs et al. (2014a). The background error autocovariance is cast as a variance amplitude and autocorrelation separable in horizontal, vertical, and flow-dependent functions. The horizontal autocorrelation is a second order autoregressive (SOAR) function with length scale based on the Rossby radius of deformation multiplied by a scaling factor. The scaling factor is examined closely in Jacobs et al. (2020), and we use the optimal value from that study resulting in a length scale of the SOAR of approximately 36 km averaged over the domain. The vertical correlation is also a SOAR function with a length scale based on the background vertical density gradient so that positions in the water column with high vertical gradient have short length scales in order to maintain sharp vertical features. The flow dependent correlation decreases correlations in the direction of the horizontal SSH gradient, which prevents data from affecting the analysis across mesoscale fronts. Prior work has shown the two dominant factors in the autocovariance are the evolution of the variance amplitude and the time period prior to the analysis over which observations are used. The time evolving variance formulation has been shown to be a good representation of forecast minus observation variances, and including SSH observations over a 5 day window also advances the forecast skill (Jacobs et al., 2014a).

2.2. Observations and observation errors

The regular observations include the satellite SSH (Jason-2, CryoSat −2, SARAL/AltiKa), the satellite SST observations separated into day and night observations, profile data from ship of opportunity XBT, Argo floats, ocean gliders, and surface ship data. These are the observations that are regularly publicly available in near real time and processed operationally. We refer to these as the regular observations.

The satellite SST temporal–spatial sampling is relatively high, and the spacing between observations is typically much less than the horizontal decorrelation scale of the assimilation. The nadir SSH observations from Jason-2, CryoSat −2, SARAL/AltiKa are typically very close in the direction along the satellite ground track (approximately 6.5 km), but on any given day, the ground tracks are typically far apart in the across-track direction (Fig. 1a). In both these cases, there is substantial redundant information that greatly slows the computational solution process of the data assimilation without adding significant value. Very close observations can also result in a 3DVar analysis with amplitude greater than any of the observation innovations. For these reasons, observations are combined into super-observations (super-obs) by averaging spatially close observations or thinning that selects one data point from a set of spatially close observations. For the altimeter SSH, the super-obbing is conducted by averaging SSH values that are close as measured by the horizontal decorrelation scale of the background error. The super-obbed SSH is then converted to a synthetic vertical temperature and salinity profile using correlations based on historical in situ observations (Helber et al., 2013). Satellite SST data are more uniform, and super-obbing is also done by spatial averaging of values. In situ profiles from certain platforms such as gliders or towed CTD may be very dense in one direction. For these data types, thinning is used to prevent averaging from smoothing sharp features. The average temperature profile may be unrealistic, and therefore data is thinned to maintain individual observed profiles. Data thinning and super-obbing are conducted in the data preparation stage.

Observation errors are a sum of instrument and representativeness errors. Representativeness errors are intended to be the variance of true features that the model is not expected to represent, and representativeness errors are much larger than instrument errors. The value of the representativeness error in the assimilation system is a function of the background gradients in the horizontal and vertical. SST representativeness errors are higher in areas of sharp fronts as the fronts in the true ocean can move more rapidly than in the model. The average SST total error is about 0.6 °C with minimum errors of 0.1 °C and maximum errors reaching 1.8 °C. Profile representativeness errors are higher at depths of high vertical gradients as internal waves can create large variations that are not represented in the model. Using one day of observations in the 1 km system, Fig. 3 shows the effects of including the vertical gradient dependence in the total errors ascribed to Argo and the synthetic profiles derived from SSH observations. The assimilation system computes error estimates for each profile individually. The minimum, mean, and maximum values across all profiles are shown in Fig. 3. In Fig. 2, the representativeness error is the variance at scales of physics not represented by the model and scales below the Nyquist. We will see in the discussion that when the model has unconstrained features, these features should contribute additional errors to observations. This consideration was not taken into account in the experiments, and the discussion considers how aspects of the assimilation process may change the results.

This study uses the model and assimilation system throughout. Results in Fig. 1a are from the model and assimilation system using all regular observations to correct the model state on a daily basis, and the system was initialized in 2012. The drifter trajectories do not affect the results and therefore represent independent observations.

2.3. Power spectral density

The power spectral density (PSD) in Fig. 2 represents a space and time average. An approximately 6 degree square subdomain in deep water was chosen for two-dimensional PSD calculations (D'Addezio et al., 2019b). Four variables were considered: mixed layer depth (MLD) diagnosed from the model temperature, 100 m depth temperature, surface kinetic energy (KE), and sea surface height (SSH). Model data with 3-hour frequency were saved for PSD computation. At each 3-hour interval, a two-dimensional fast Fourier transform computed the two-dimensional PSD. A time-average of the two-dimensional PSD provided a single mean two-dimensional PSD estimate. To obtain a one-dimensional spectrum, the time mean two-dimensional PSD was averaged in direction along concentric circles of constant wavenumber k (D'Addezio et al., 2019b; Richman et al., 2012). The two-dimensional PSD is defined as:

$$G\left(k_x, k_y\right) = \left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g\left(x, y\right) e^{ik_x x} e^{ik_y y} dx dy\right)(*) \tag{1}$$

where $\left(k_x, k_y\right)$ is the wavenumber vector, the first term on the right side is the Fourier transform of the function $g\left(x, y\right)$, and $(*)$ is the complex conjugate of the Fourier transform. The directional average is then

$$G_R\left(k\right) = \frac{1}{2\pi k}\int_{-\pi}^{\pi} G\left(k_x, k_y\right) k d\theta \tag{2}$$

where $k = \sqrt{k_x^2 + k_y^2}$, θ is direction, and within the integral there is an implicit conversion from $G\left(k_x, k_y\right)$ to $G\left(k, \theta\right)$. The plots in Fig. 2 are $G_R\left(k\right)$. To obtain variance, $G_R\left(k\right)$ must be multiplied by $2\pi k$ and then integrated over k. The total variance V is:

$$\begin{aligned} V &= \int_0^{\infty} 2\pi k G_R\left(k\right) dk \\ &= \int_0^{\infty} 2\pi k \int_{-\pi}^{\pi} \frac{1}{2\pi k} G\left(k_x, k_y\right) k d\theta dk \\ &= \int_0^{\infty} \int_{-\pi}^{\pi} G\left(k_x, k_y\right) k d\theta dk \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G\left(k_x, k_y\right) dk_x dk_y \end{aligned} \tag{3}$$

Because Fig. 2 is plotted in log space, multiplying by k is equivalent to adding 1 to the spectral slopes. For surface kinetic energy, we applied the processing described above separately to surface zonal and meridional velocity and then averaged the two resulting spectra (Richman et al., 2012).

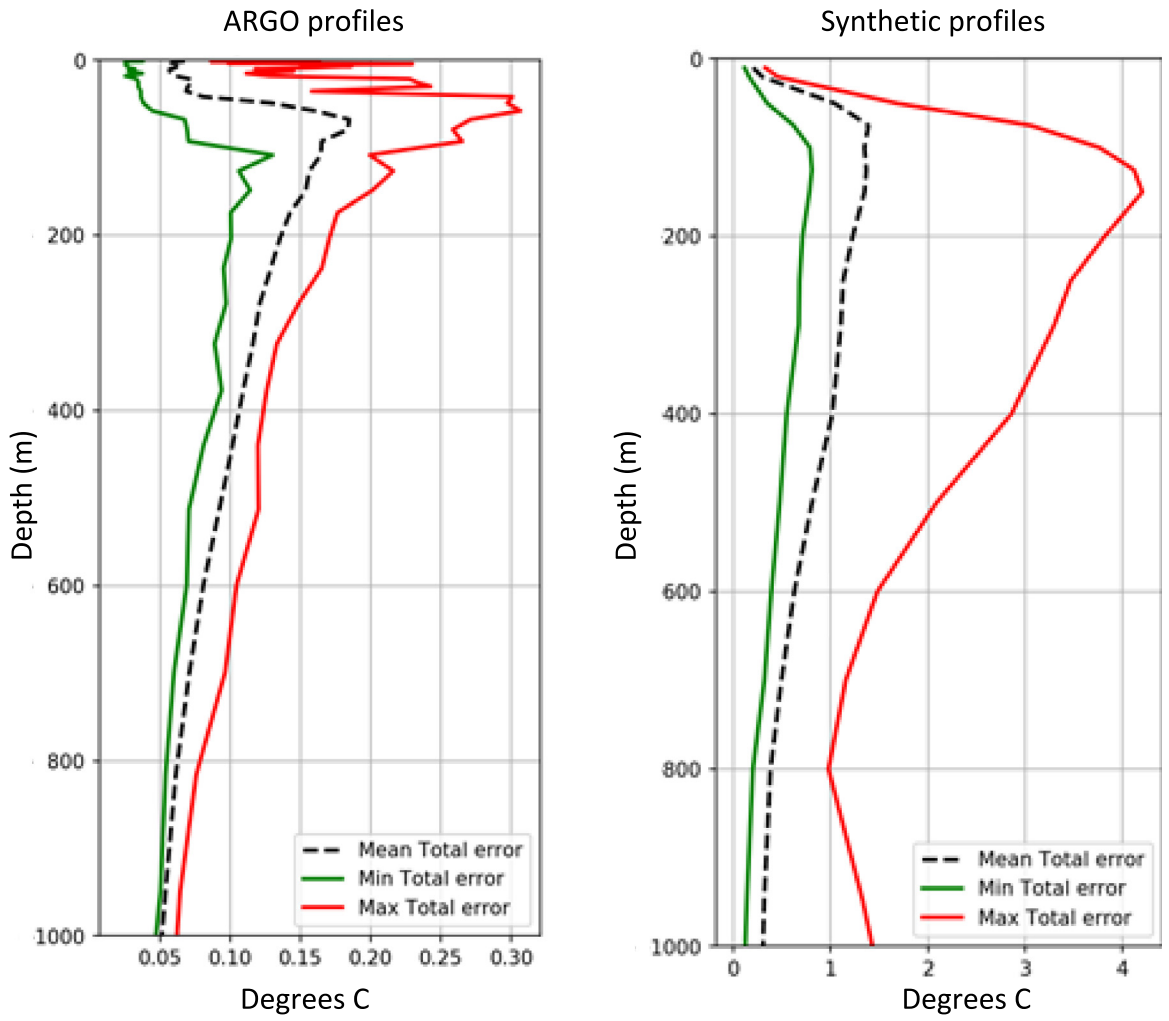

Fig. 3. The total errors of Argo temperature profiles (left) and synthetic temperature profiles (right) from the assimilation system indicate the higher errors at the thermocline depth in the Gulf of Mexico due to the higher vertical temperature gradients. The solid line indicates the peak errors, and the dashed line is the mean error over all the observations.

2.4. 3 km ensemble model

Ensemble ocean predictions are composed of a set of numerical forecast models. The system used here was set up at 3 km resolution using the NCOM model. In this study, a 32-member ensemble used an Ensemble-Transform (ET) methodology. One member was the control run, which conducted a daily data assimilation cycle. After the data assimilation, the perturbations of the remaining members from the control run were scaled so that the ensemble variance matched the analysis variance estimated during the data assimilation process (McLay et al., 2008). All the members were then integrated forward in time until the next assimilation cycle. Thus, the analysis variance controls the spread of the ensemble at the analysis time, and the ensemble growth during the forecast is dependent on the nonlinear model development. The ensemble prediction system was initialized in 2012 and conducted daily cycles. At any location and for any variable, the ensemble set provides a sample intended to represent the probability density function of the state (Evensen, 1994). This ensemble system has previously shown skill in forecasting errors in the Gulf of Mexico domain (Wei et al., 2014). The variance across the ensemble increases during the forecast period. The spread-reliability is an evaluation of the variance across the ensemble to the error variance with observations. The spread-reliability shows the ensemble is able to predict areas of larger and smaller errors consistent with observations. The ensemble under-predicts error in that the variance across the ensemble is typically lower than the observed error variance. This is also reflected in the Talagrand histograms (Wei et al., 2013).

Part of the error of any ensemble member is due to unconstrained features. At any time, the mean over ensemble members reduces the influence of the unconstrained features if the features are uncorrelated across members. This phenomenon will be explored more thoroughly in Section 3.1. We analyze results of the ensemble ocean system covering January through June 2016 using the power spectral density analysis described in Section 2.3. The order of ensemble-averaging, time-averaging, and computing PSD is important as the computation of PSD is a non-linear operation. Define an ensemble member variable v_{ij} as the ith ensemble member at the jth time. Define the mean PSD of a variable as:

$$\frac{1}{N_i N_j} \sum_{i=1}^{N_i} \sum_{j=1}^{N_j} PSD\left(v_{ij}\right) \tag{4}$$

where $PSD()$ is the squared amplitude of the Fourier Transform averaged over direction, N_i is the number of ensemble members, and N_j is the number of time steps. The mean PSD represents the energy distribution of the variable that would occur on average within any snapshot of any ensemble member. Define the ensemble-mean PSD

as:

$$\frac{1}{N_j}\sum_{j=1}^{N_j} PSD\left(\frac{1}{N_i}\sum_{i=1}^{N_i} v_{ij}\right). \quad (5)$$

This averages the ensemble members, which reduces the effects due to unconstrained features, before computing the PSD and averaging realizations over time. The ratio of ensemble-mean PSD to mean PSD should be less than 1 and indicates the fraction of energy remaining after averaging across the ensemble (Fig. 4). At the largest scales, the ratio is near 1, which indicates the energy of features is similar across all the ensemble members at the larger scales. At smaller scales, features across the ensembles are uncorrelated, and averaging reduces the energy in the ensemble-mean PSD. The characteristic of importance is the wavenumber range over which the ratio decreases substantially from a value of 1. In Section 3.1, we consider this wavenumber range to show consistency with estimated constrained scales.

2.5. 1 km OSSE setup

In the OSSE experiment, the term 'nature run' is the system with all realistic boundary conditions and surface forcing but without any data assimilation. The OSSE setup uses a 1 km nature run (D'Addezio et al., 2019a,b). The nature run allows the numerical model physics to represent instability processes that cascade energy across the wavenumber spectrum (Fig. 2). The nature run also provides the data sampled by simulated observing systems. We sampled the nature run at all regular satellite and in situ observation locations and by the SWOT simulator (version 2.0.0) (Gaultier et al., 2016). A second 1 km OSSE model assimilates the simulated observations.

The GOFS state on November 1, 2015 provides the initial condition to the nature run. This initial condition was 2 months prior to the OSSE experiment to allow spin-up. After spin-up to January 1, 2016, a 1-year integration of the nature run saved 3-hourly values of 3-dimensional temperature, salinity, and velocity and 2-dimensional sea surface height. The OSSE initial condition on December 1, 2015 was set to the nature state on December 1, 2016. This provided a seasonally realistic initial condition with ocean mesoscale features that were out of place relative to the nature run. The OSSE is spun-up for one month, assimilating daily simulated nature run observations, and is evaluated relative to the nature run starting January 1, 2016 (D'Addezio et al., 2019b). This OSSE is a fraternal twin experiment that does not represent errors in the physics of the numerical model. We sampled the nature run at real observation locations and times for the regular observing network (in situ, satellite sea surface temperature, and satellite altimetry), and the SWOT simulator sampled the nature run using a 2 km grid in the along- and across-track directions. The SWOT simulator includes estimates of the errors from a range of sources. In the experiments here, we include only an estimate of the sensor noise from the KaRIn instrument (Gaultier et al., 2016). The expected sampling for SWOT is 1 km in the along- and across-track directions. In Section 2.2, we discussed the super-obbing and thinning that takes place with respect to the horizontal decorrelation scale. Even at 2 km, the SWOT observations are significantly reduced. A 1 km sampling would result in the same data quantity used in the assimilation. Examples of the SWOT super-obbing are provided in D'Addezio et al. (2019b) (their Fig. 9).

The simulated observations were assimilated by the OSSE using the NCODA 3DVar system (Smith et al., 2011). The OSSE used a two-step, multi-scale formulation of NCODA (Li et al., 2015; Souopgui et al., 2020). The first step of the analysis used same settings as discussed in the 1 km model of Section 2.1, which are mesoscale-oriented. The second step used the results of the first step as the background for computing innovations (observation minus background) and a smaller decorrelation length scale (Souopgui et al., 2020). This methodology improves analysis skill when the observing network is able to resolve a wide range of spatial scales (Li et al., 2015, 2019; Muscarella et al., 2014). The OSSE multi-scale 3DVAR assimilated observations every 24 h for a six-month period.

For the second analysis step, the horizontal scale of the SOAR function was about 10 km. The 10 km scale represents about 20 km features, which is near the lower limit of features SWOT expects to resolve. Each daily analysis used data over a prior period. The first step used prior data periods of 12-day, 5-day, and 1-day for in situ profiles, satellite altimetry, and satellite sea surface temperature, respectively. The second analysis used data over a 1-day period for all observation types as the smaller spatial scales have shorter time scales (Souopgui et al., 2020).

Temperature at 100 m depth was used for evaluation (Fig. 5). The OSSE assimilation experiment minus the nature run provided full time varying 3D error fields. A spectral analysis performed on the errors provided the ability to identify constrained and unconstrained scales.

3. Results

We use the wavenumber PSD to illustrate the relations between model resolution, observation resolution, and the gap between the two (Fig. 2). The data are from a 1 km resolution ocean model (Section 2.1), and the PSD lines are one-dimensional spectra representing time and direction averages (Section 2.3). Features smaller than 2 km are beyond the Nyquist wavelength of the model 1 km grid. Second order accurate numerical schemes require about 10 grid points within one wavelength to reasonably represent a first derivative (Roache, 1998). For the 1 km model considered, there is a limit at 10 km wavelength separating physics the model represents from those it does not. The level of desired numerical accuracy may change from one situation to another, so this limit may not be a clear and distinct line. Model resolution controls the limits at the smallest scales in the spectrum. As model resolution increases, both the Nyquist and representation limits move to smaller wavelengths to the right. Now we consider what information is contained in the larger constrained scales, what controls the boundary between the constrained and unconstrained, and what information lies within the unconstrained scales.

3.1. Constrained scales and confidence

By the definition introduced in Section 1, error variance in the constrained wavenumber band (Fig. 2) is lower than the variance of the features themselves (D'Addezio et al., 2019b). For the OSSE experiments, we can compute the PSD of forecast errors to determine the boundary between constrained and unconstrained. Because we do not have regularly gridded observations of the true ocean, we cannot directly compute PSD of errors. An approach to estimating the boundary separating constrained and unconstrained scales is to filter unconstrained features from the model to reduce error variance relative to observations (Jacobs et al., 2020). The filtering methodology is described in Appendix A, which uses a Gaussian convolution kernel. Initially removing small-scale features reduces errors to observations, and gradually increasing the size of the filter reduces errors until a minimum error is reached. Continuing to increase the filter size and removing larger scales increases errors. Errors in surface trajectories are a minimum at a filter length scale of 58 km defined by a Gaussian e-folding scale in a convolution kernel or a 220 km wavelength defined by the $\frac{1}{4}$ power point of the filter. The errors as a function of filtering scale in Jacobs et al. (2020) are a broad minimum because error PSD divided by true ocean PSD ($\varepsilon(k)/\gamma(k)$) gradually transitions from less than 1 to greater than 1. This is also apparent in the OSSE results in Section 3.2. In addition, the Gaussian kernel does not provide a sharp cutoff in wavenumber space. This boundary can change depending on the particular situation, and therefore this boundary is not a clear and distinct line. In Fig. 2, the unconstrained band lies between the boundary of skillful prediction of constrained features larger than 220 km wavelength and the model representation limit at 10 km wavelengths.

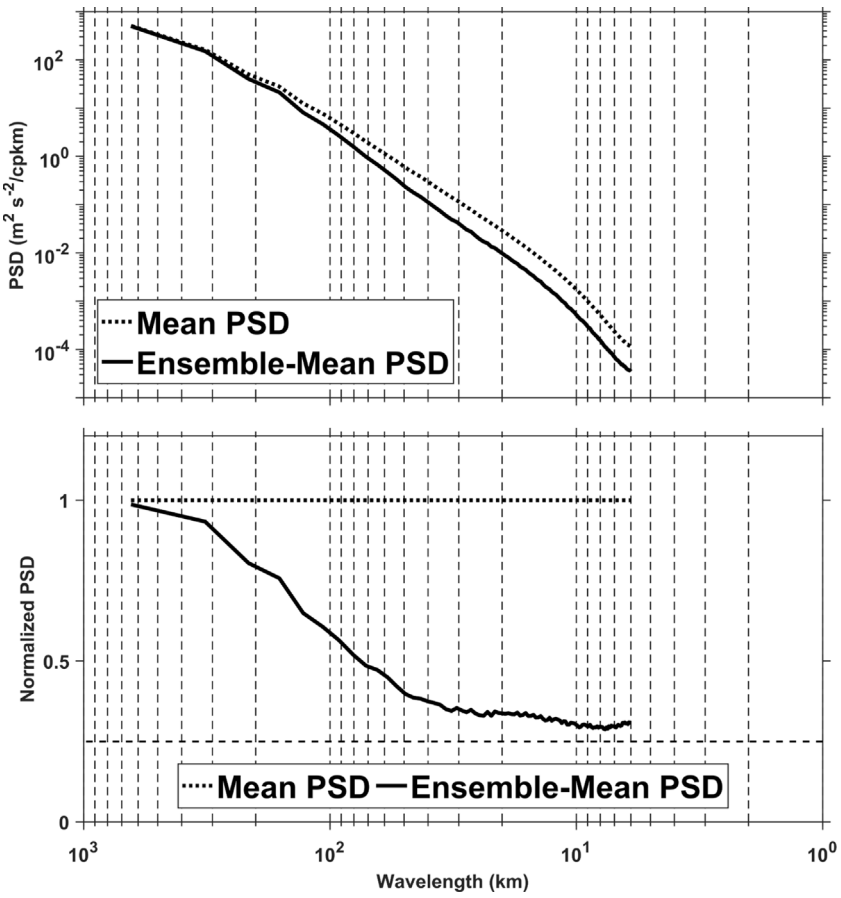

Fig. 4. The mean PSD (dotted line) and the ensemble-mean PSD (solid line) for surface eddy kinetic energy (m^2 s^{-2}/ cpkm) (top) and the ratio of the two (bottom). The drop in the energy ratio in the range of 200 km to 60 km is consistent with features not being correlated between ensemble members that ensemble-averaging removes. These are unconstrained features represented by the ensemble.

In tests varying the horizontal decorrelation length scale within the data assimilation, the error minimum occurred at the same filtering scale for the range of decorrelation scales tested except for the longest decorrelation scale that was unrealistically large (Jacobs et al., 2020).

Ocean forecasting experiments indicate that forecast skill degrades with increasing model resolution (Sandery and Sakov, 2017). Fig. 2 helps interpret this result. Assume the model resolution were such that the scale of represented features were equivalent to scales resolved by observations and allowed by the data assimilation. In this case, there would be no unconstrained wavenumber band. As model resolution increases, the unconstrained band between constrained scales and the model representation limit increases. All of the variance in the unconstrained band adds to the error level because, by definition, there is no skill in the unconstrained band. The error variance added from the unconstrained band is more than (up to twice) the variance of the features in the unconstrained band because the model and true feature positions are uncorrelated (Appendix B). As model resolution continues to increase, the unconstrained band becomes wider. The variance added from the unconstrained band continues to increase errors as model resolution increases.

Errors at small scales significantly affect ocean surface trajectory forecasts (Griffa et al., 2004), and we show the total error variance in the constrained band is less than the total error variance in the unconstrained band. Differencing a model field from observations, three sources contribute to error variance (Appendix B): (1) errors in the constrained band (EC), (2) the model unconstrained scale variance (MU), and (3) the true ocean variance at the unconstrained scale (TU). For the moment, we are neglecting the variance of the true ocean at scales not represented by the model. After filtering MU features from numerical model fields, RMS error decreases to 0.80 times the original value (Jacobs et al., 2020). Assuming a zero mean, the RMS squared is the variance, and therefore filtering the MU reduces variance to 0.64 of the original value. Assuming the three error contributions are statistically independent, the MU variance is 0.36 of the total error variance. The remaining 0.64 of the total variance is the sum of EC and TU. A conservative assumption is that TU variance is equal to MU variance. Then the contribution of the EC band is 0.28 of the total error variance, which is less than TU and MU errors. The assumption is conservative as we expect TU variance is greater than MU variance since model resolution limits features physically represented (Fig. 2). Thus, the contribution of the EC errors to the total error would be less than the estimated value of 0.28, and the conclusion remains valid. Initially we neglected errors due to variance at scales smaller than represented by the model. If we consider the contribution of errors due to variability in the true ocean at scales not represented in the model, the EC errors become smaller, and the conclusion remains valid. The contribution to errors due to variability in the unconstrained band is larger than the contribution to errors in the constrained band. The majority of uncertainty is in the unconstrained portion of the solution. The constrained portion of the solution can be described as the portion in which we have confidence. This result is a reflection of the definition that constrained scales have error variance smaller than the variance of the solution (Section 1).

This conclusion is for the observations of velocity obtained from the LASER drifters. The PSD slope can change the boundary position for different variables. Because the PSD curves represent an average over direction, integration to obtain variance in a wavenumber band requires multiplication by wavenumber (Section 2.3). When considering the contribution to variance in different wavenumber bands, a value of 1 should be added to the slopes of the lines in Fig. 2. The ratio of

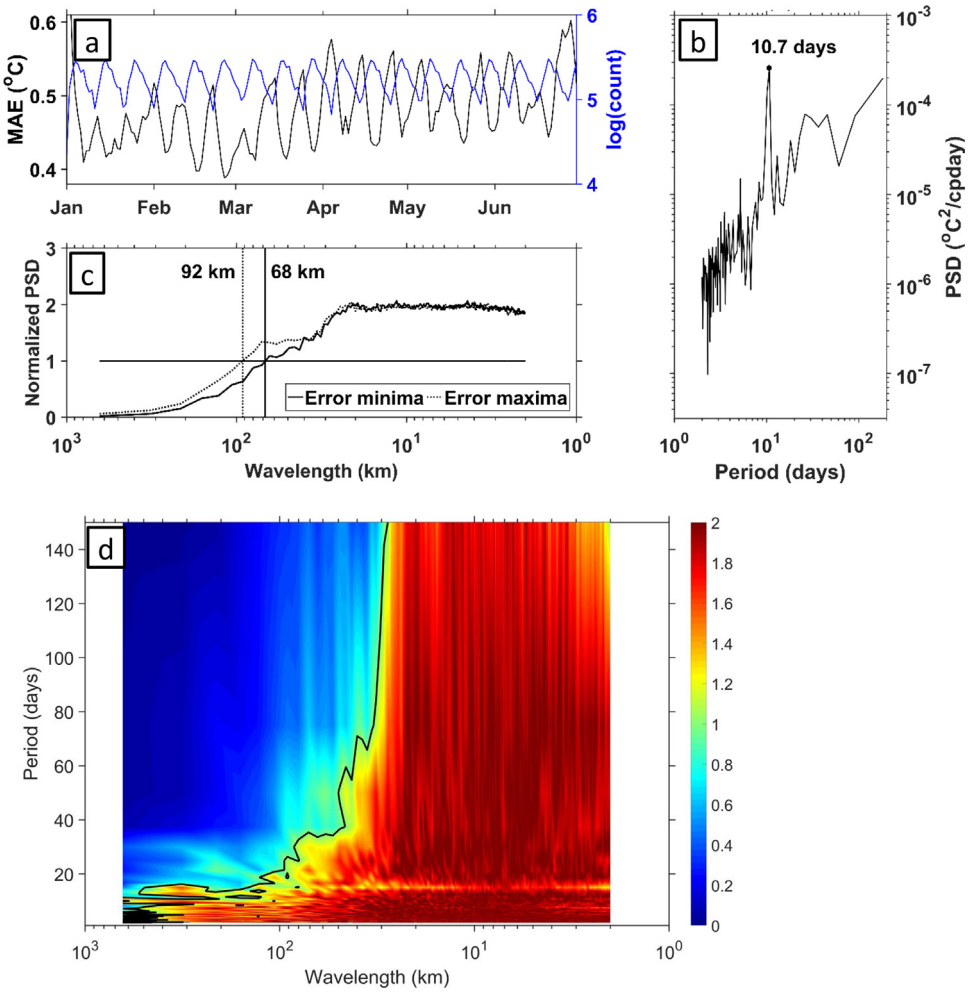

Fig. 5. Results from a 6 month Observation System Simulation Experiment (OSSE) indicate (a) a periodic signal in the mean of the absolute value of the error (MAE) in 100 m depth temperature (black line) and the number of observations (blue line) available in each daily data assimilation cycle. The MAE frequency power spectral density (PSD) (b) contains an energy peak at a 10.7-day period, which is a result of the SWOT sampling pattern. The ratio of error PSD to the average of nature and OSSE PSDs is computed separately for data at the local maxima and at the local minima of the MAE time series (c). The constrained scales indicated in (c) change from 92 km when SWOT swaths are not present to 68 km when SWOT swaths are present. The ratio of error PSD to the average of nature and OSSE PSDs over both wavenumber and frequency uses the same definition of skill with a value of 1 plotted as a black contour (d). Areas less than 1 indicate skill, which transition from shorter time periods at large scales to only longer time periods at small scales.

unconstrained to constrained variance (U/C variance) is the ratio of the PSD integrated over the unconstrained band to the PSD integrated over the constrained band. The SSH slope magnitude is largest and the mixed layer depth (MLD) slope magnitude is smallest. Therefore, SSH U/C variance is smaller than MLD U/C variance. This has implications for evaluating the impact of added observations. OSSEs suggest that when increasing SSH observations, SSH skill advancements are small, while skill advancements in MLD are larger (D'Addezio et al., 2019b; Jacobs et al., 2014a). When judging observation system impact, it is important to consider a relevant variable that is sensitive to the observations and contains requisite smaller scale structure.

Finally, we evaluate consistency of the estimated constrained and unconstrained bands with an ensemble representation. Results from the 32-member ocean model ensemble indicate smaller scales being unconstrained. The mean PSD contains more variance than the ensemble-mean PSD, and the ratio of the two shows larger energy reductions at smaller scales (Fig. 4). Averaging over the ensemble realizations reduces the signal of unconstrained features not correlated between ensemble members. Larger scale features are similar amongst the ensemble members, and averaging across the ensembles does not greatly reduce the constrained energy. Smaller scales contain a more pronounced reduction of energy when averaging over the ensemble set with the change in relative energy occurring in the range of 100 to 200 km wavelengths. This is consistent with the estimated constrained scale of 220 km obtained by filtering the 1 km model result to minimize errors to the LASER data (Jacobs et al., 2020). Thus, our conclusions are consistent with ensemble approaches, and we consider this further in the discussion (Section 4).

3.2. The boundary separating constrained and unconstrained scales

Ocean features resolved by observations and corrected by the assimilation control the boundary between constrained and unconstrained scales. Prior theory and experiments have shown computing model corrections at scales smaller than resolved by observations leads to degradation of results (Li et al., 2015). Thus, limited independent observations to resolve features represent a wall to predictive skill. Demonstrations of high-resolution in situ data (Özgökmen et al., 2018) and future high-resolution satellite ocean observing systems, such as SWOT (Fu and Ubelmann, 2014), can reduce errors. One such example is the application of dense surface drifter deployments to predicting trajectories. Inferred velocities from 300 drifters in the Grand Lagrangian Deployment (GLAD) (Olascoaga et al., 2013) correct model forecasts, and forecast drift trajectories improve significantly as the data constrain smaller scale features through a 4DVar assimilation (Carrier et al., 2014). An explanation of the effect is that the additional data moves the

boundary to smaller scales to the right in Fig. 2. The constrained band increases in size, the unconstrained band reduces in size, and the total variance contributing to errors in the unconstrained band decreases. We use an OSSE to demonstrate and quantify this more clearly.

An OSSE of the SWOT observing system (Section 2.5) demonstrates this in terms of the PSD of errors. Satellite orbits are predictable far in advance. Therefore, we know the data distribution that will occur. However, the sampled points in an area do not always form a similar regular pattern every day. From the SWOT OSSE results, the mean of the absolute value of the error (MAE) of 100 m depth temperature (Fig. 5a) contains a periodicity, which correlates negatively to the number of observations within each daily assimilation. We use the 100 m temperature for evaluation rather than the SSH because the SSH PSD slope is steeper, which leads to 100 m temperature being a more sensitive metric. In addition, the SSH is the main observed variable, and evaluation of a less observed variable provides a more independent test. The PSD of the MAE time series contains a local peak at 10.7 days (Fig. 5b). This is roughly half the 21-day exact repeat period of the SWOT orbit, and it is due to the manner in which the satellite samples the area during the orbit cycle. While SWOT data are dense along satellite observation swaths, distribution of swaths in an area is irregular in time. During the 21-day SWOT repeat period, a set of swaths first covers the area leading to increased data, then there is a time gap of no SWOT swaths leading to data decrease, and later a second set of swaths provides another dense sampling set followed by another data gap. This leads to the 10.7 day periodicity in 100 m temperature errors.

Because we have full time-varying fields from the nature and OSSE models, we can compute full error fields. We compute constrained scales in 100 m depth temperature using data at times of the local minima in the MAE and separately at the times of local maxima using wavenumber spectral analysis. To compute the normalized error PSD for the OSSE experiment, we used (D'Addezio et al., 2019b)

$$\frac{\epsilon_{OSSE}}{\langle \gamma_{NATURE}, \gamma_{OSSE} \rangle}, \tag{6}$$

where ϵ_{OSSE} is the OSSE error PSD (PSD of the nature run minus OSSE), γ_{NATURE} is the nature run PSD, γ_{OSSE} is the OSSE PSD, and brackets denote the mean of the two PSD. Values range from 0 to 2 (Fig. 5). A value of 0 indicates the OSSE has no error at the particular wavelength. A value of 2 indicates the OSSE and the nature run are uncorrelated, and the error variance is the sum of the OSSE variance and the nature run variance. We define the boundary between constrained and unconstrained scales as the wavelength where the normalized error PSD has a value of 1 as in Section 1. The constrained scales have errors smaller than the variability of the features. The boundaries between constrained and unconstrained scales are 68 km during times of error minima and data maxima versus 92 km during times of error maxima and data minima (Fig. 5c). Thus, we see the boundary separating constrained and unconstrained scales oscillating throughout the SWOT observing cycle.

With the full time-varying temperature field, we can examine the details of the boundary both in spatial scale and temporal scale. The wavenumber-frequency spectra of errors, the nature run, and the OSSE experiment were computed from the 00Z value of 100 m temperature on each day of the 6 months. This produces a three dimensional matrix with dimensions [f, k_x, k_y]. The [k_x, k_y] PSD were directionally-averaged as described in Section 2.3, generating a final [f,k] PSD matrix. Normalizing the error PSD according to Eq. (6) and using the definition of the boundary between constrained and unconstrained as the normalized PSD having a value of 1, we can see how the time scale of the boundary changes as a function of spatial scale (Fig. 5d). In this analysis, we have not separated data from times of error minima and maxima. At the largest scales, there is skill at time periods longer than approximately 15 days. The boundary time scale increases as spatial scales become shorter. For example, at 50 km scales only time periods longer than 40 days have skill. The shortest scales at which there is skill are about 30 km at which the time scales move to 180 days.

The ability of the observation data to resolve smaller features moves the boundary. While observations move the boundary to smaller scales and advance prediction skill, we must properly utilize the constrained and unconstrained information from forecasts regardless of the boundary position.

3.3. Unconstrained scales and uncertainty

Though there is not skill in predicting the exact location of features in the unconstrained band, we can still recover valuable information from this band. In the constrained scales, there is skill in predicting the position of features, which is deterministic forecast skill. Assuming the model physics are realistic, the unconstrained solution contains statistical prediction skill. While the position of any unconstrained feature is not predictable, the model can predict areas in which the unconstrained features develop. Filtering can separate scales, and the unconstrained features are the residual of the full model solution minus the constrained scales (Appendix A). At any time, define the local unconstrained variability at a point (x_o, y_o) as the spatially weighted RMS of the local unconstrained features: $\hat{v}_U(x_o, y_o)$ (Appendix A). In areas containing unconstrained features with large amplitude, $\hat{v}_U$ is large, and areas containing small amplitude unconstrained features, $\hat{v}_U$ is small. The vectors in Fig. 1a are the full model solution, and vectors in Fig. 1b are the constrained currents that remain after filtering the results shown in Fig. 1a. The background color in Fig. 1b is the local unconstrained variability of speed (square root of eastward and northward $\hat{v}_U^2$). Larger scale constrained ocean features such as mesoscale eddies have strong currents around the periphery, which form sharp fronts. The small-scale nature of the fronts results in errors due to mispositioning with high values of $\hat{v}_U$. Bathymetric features such as the shelf break along Louisiana to the Florida panhandle are areas of active generation of small features (Jacobs et al., 2016). Mesoscale eddies modulate stratification that becomes conducive or inhibitive to the formation of submesoscale eddies (Capet et al., 2008; D'Asaro et al., 2011; Jacobs et al., 2014b; McWilliams, 2016; Qiu et al., 2014; Shcherbina et al., 2013; Thomas et al., 2013; Zhong and Bracco, 2013). The anticyclonic Gulf of Mexico Loop Current Eddy (LCE) contains a much deeper thermocline than the surrounding area. This effect decreases stratification in the upper water column and leads to a much deeper mixed layer that is conducive to the formation of submesoscale eddies. The LCE is typically an area of high submesoscale eddy activity. Forecast errors or uncertainty are not expected to be distributed uniformly but should be related to the underlying processes.

We show skill in $\hat{v}_U^2$ being a predictor for error variance by using trajectories observed by more than 1000 surface drifters from LASER in the Gulf of Mexico (Özgökmen et al., 2018). The value of $\hat{v}_U^2$ is an estimate of the expected errors due to unconstrained ocean features. If the estimator $\hat{v}_U^2$ has skill in predicting areas of small-scale activity, we expect a relation with observed error variance. This relation changes if we are computing errors to the full model solution or errors to only the constrained portion of the model solution (Appendix B). The expected relation between $\hat{v}_U^2$ and the constrained model solution error variance is a constant value due to errors in the constrained field plus the value of $\hat{v}_U^2$ due to the true ocean variance at unconstrained scales. The expected relation between $\hat{v}_U^2$ and the full model solution error variance is the same constant due to errors in the constrained field plus twice the value of $\hat{v}_U^2$ due to the unconstrained variance in the model solution and the true ocean variance at unconstrained scales.

To demonstrate $\hat{v}_U^2$ has predictive skill, for every drifter velocity at a location and time, we first compute the squared observed velocity error. The squared error adds to the value in a bin determined by the coincident value of $\hat{v}_U^2$. Within each bin of $\hat{v}_U^2$, we compute the mean of all squared velocity errors. Velocity squared errors are accumulated separately from both the full model field and the constrained model

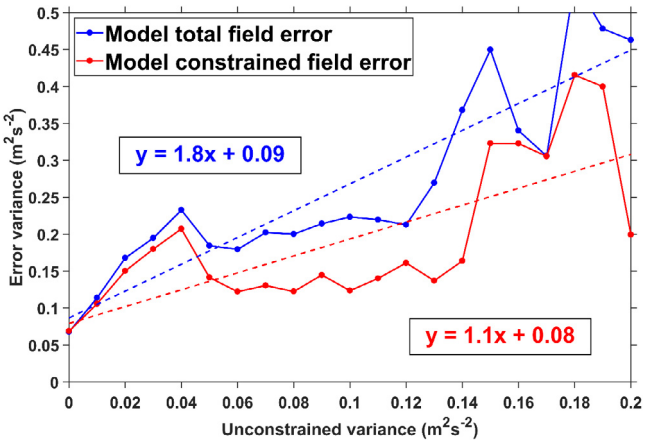

Fig. 6. The mean of the squared errors of drifter-inferred velocity minus model velocity as a function of the predicted unconstrained local variance. The least linear squared regressions to the data (dashed lines) are consistent with expected values. The expected errors in the total model velocity have a 2:1 relation, the expected errors in the confident portion of the model velocity have a 1:1 relation, and the vertical intercept is expected to be the same between the two.

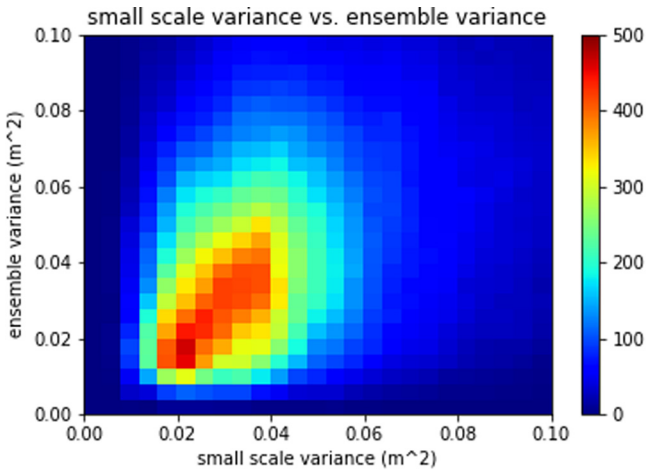

Fig. 7. The probability density function of the small-scale variance of surface velocity from the 1 km model vs. the variance across the ensemble of surface velocity during the LASER time period. The results include data only in locations where the water depth is greater than 1000 m.

field. The results (Fig. 6) are consistent with expected relations. The error variance of both the full model solution and the constrained solution have about the same constant value. The constrained solution error variance increases with a slope of about 1, and the total solution error variance increases with of slope of about 2. This supports two conclusions. First, using only the constrained scales in deterministic forecasts produces lower errors in model variables in an RMS sense. Second, the unconstrained local variance $\hat{v}_U^2$ provides statistical prediction of error variance. We can accomplish this only if forecasts have sufficient resolution to represent the small-scale features.

The ensemble results also aid in verifying the predictive skill in $\hat{v}_U^2$. We compute the small-scale variance of surface velocity by summing the eastward and the northward small-scale variance of velocity components from the 1 km model. From the ensemble, at each time we sum the eastward and the northward variance across the ensemble. We restrict the comparisons of the information to water depths that are greater than 1000 m where mesoscale and submesoscale instabilities are the dominant process. The probability density function of the small-scale surface velocity variance and the variance of surface velocity across the ensemble indicate a high correlation (Fig. 7). This further advances the utility of the scale separation method, as the result correlates well with the ensemble methodology.

We also note that the forecast time period considered in the evaluations here is only the 0 to 24 h forecast after the data assimilation cycle each day. It would be expected that as longer forecast times are considered, the boundary between constrained and unconstrained energy will move to larger scales, or to the left in Fig. 2. The rate at which the boundary moves to the left would be expected to be a function of the instability growth time across the wavenumber spectrum. The SWOT OSSE results (Fig. 5) give some indication of the speed at which errors increase over time as the MAE increases during periods of no SWOT coverage from the minimum value. The rate at which the constrained scales change during the forecast time is an important aspect left for future examination.

4. Discussion

Results in Fig. 6 are consistent with the concept that the forecast solution contains skill in the constrained scales, the unconstrained scales add error, the small-scale variability is predictor of the error variance, and thus the high-resolution forecasts have skill in predicting areas in which the small-scale variability will occur. This is similar to the meteorological problem of showing skill in high-resolution model rainfall (Gilleland et al., 2010). A range of metrics have been developed in meteorology such as Fractional Skill Score, which is skill in predicting the fraction of an area over which rainfall exceeds a threshold (Mittermaier and Roberts, 2010). Such metrics are enabled by large area gridded data of rainfall rates from radar systems. Even with the extensive drifter deployment during the LASER experiment, the ocean sampling was relatively sparse and non-uniform. Constructing regular gridded fields for evaluation was not possible, so our metrics are simple RMS errors at point locations.

Improved metrics could advance the demonstration of value in high-resolution ocean models. It is also important to consider the variables used in the metrics, and the spectral slopes in Fig. 2 are important considerations. The steep slope of SSH indicates this is not a sensitive variable for demonstrating advances in predictive skill. The shallow slope of the EKE and MLD shows the influence of smaller scale features. Previous Observation System Experiments (OSEs) (Jacobs et al., 2014b) indicate the improvement when adding an additional altimeter data stream is much smaller in SSH than in frontogenesis forcing and surface divergence skill. Frontogenesis forcing is based on the flow field strain and buoyancy gradients, and surface divergence is a gradient of the velocity field. The slope of the frontogenesis forcing and divergence PSD would be shallower by a value of 1 than EKE, and the slope of EKE is about 1 shallower than the slope of SSH that would be expected if currents were in geostrophic balance (Fig. 2). Because frontal areas typically contain stronger currents, they are important for many applications including fisheries, commercial endeavors, and emergency response. The surface divergence driven by frontogenesis may be sufficiently observable in the future by airborne instruments (Rodríguez et al., 2018). As we look to developing metrics more closely related to human activity, these will tend toward variables that have a much flatter slope and thus much more variability at small scales.

Part of the forecast process is the assimilation and associated settings including length scales of background errors, flow dependent correlations, and observation errors. Prior results indicate the boundary separating constrained and unconstrained scales is more sensitive to some parts of the data assimilation than others. In experiments varying the horizontal decorrelation scale (Jacobs et al., 2020), the boundary did not change except for the largest, unrealistic decorrelation scale. In OSSE experiments (Souopgui et al., 2020), adding SWOT data increased skill in both single-scale and multi-scale assimilation. The multi-scale assimilation consistently provided increased forecast skill relative to the single-scale assimilation, and the boundary scale decreased slightly.

Thus, changes in the data assimilation process can have impact on the boundary between constrained and unconstrained scales. It should be noted that the OSSE experiments were able to compute the boundary scale from the PSD of full time-evolving 3D fields, while the estimates of the boundary scale through comparison to non-uniformly distributed drifters used a filtering approach. The OSSE experiments are a much more sensitive tool than what can be measured by direct observations. Together, these results imply that advancing the data assimilation process and settings increases forecast skill and can have impacts on the boundary of constrained scales. From the OSSE results in this paper, we find additional observations that resolve smaller scale ocean features reduce the constrained scales.

The features leading to the unconstrained energy are due primarily to flow instabilities. These include small mesoscale eddies on the scale of 100 km, submesoscale eddies with high Rossby number, and frontogenesis processes that are driven by straining of the mesoscale and submesoscale flow and buoyancy fields (Hoskins, 1982). The motivation for conducting the experiments at 1 km resolution was to resolve the processes associated with submesoscale eddies (Capet et al., 2008). Prediction of instability processes requires regular corrections to model initial conditions, and high-resolution models with insufficient observations leads to unconstrained scales. Other dynamical regimes may not contain processes leading to the separate unconstrained features. In particular, coastal environments are dominated by wind-driven events and fresh water fluxes from land rather than instabilities. Thus, caution should be exercised prior to applying these concepts to other dynamical regimes.

In light of the unconstrained variability, we must reconsider observation errors in high-resolution models. Errors in instrumentation are often the smallest source of discrepancy between observations and model forecasts. Representativeness errors are regarded as signal due to small-scale features in the true world that the numerical model does not represent. In Fig. 2, this is the variance in the band at scales smaller than represented by the model or smaller than 10 km in the case of the 1 km resolution model used here. Considering Fig. 2 again, there is large variance in the unconstrained band that creates deviations of the model from observations, and this is an additional error source. The unconstrained errors in Fig. 2 are larger than the representativeness errors because the model resolution represents features down to 10 km, and observations constrain only much larger features. In addition, because the model unconstrained features are uncorrelated with the true ocean features, the total unconstrained error variance is twice the variance of the true ocean in the unconstrained band.

If, as in Appendix B, we separate the model and true world into variability in the constrained scales and unconstrained scales, and if we assume the model and true fields are uncorrelated in the unconstrained band, we find additional errors that are twice the variance in the unconstrained scales represented by the model when comparing to the LASER drifters. To the representativeness errors we must add twice the unconstrained variance. This is analogous to the double penalty errors pointed out by Gilleland et al. (2009). In the assimilation experiments here, the unconstrained variance error was not considered. Including the error would be expected to reduce the influence of observations, and appropriate errors could increase forecast skill and move the boundary to smaller scales. While the quantitative value of the boundary position may change, the concepts of separating constrained and unconstrained scales would not.

5. Conclusions

We provide a framework in wavenumber space to understand concepts involved in ocean forecast systems that have model resolution representing features smaller than observations resolve and the assimilation system corrects. Constrained scales are defined as having error variance smaller than the true world variance. Model smaller scales are unconstrained and have errors larger than the true world variance. We have confidence in the forecasts at the constrained scales, and the unconstrained scales contain uncertainty since there is no predictive skill. Error variance in the unconstrained band is up to twice the variance of the true world features. Results from ensemble forecasts are consistent with the scales estimated to be constrained. The ratio of ensemble-mean PSD to the mean of the ensemble PSD shows the large reduction of energy in the unconstrained scales indicating the uncorrelated features between ensemble members at the smaller scales. The framework explains results in which increased resolution leads to degraded skill (Sandery and Sakov, 2017) since increased resolution increases the variance in the unconstrained scales.

The boundary between constrained and unconstrained bands is controlled by the capability of observations to resolve features and the assimilation to correct the model initial condition. Moving the boundary to smaller scales increases the constrained band and reduces forecast error. The assimilation system must use data appropriately to ensure smaller features are corrected as shown by Souopgui et al. (2020). Variables and metrics that are most sensitive to observation and assimilation impacts are those with larger energy at smaller scales, and thus have shallow PSD slopes. Ocean variables such as MLD and frontal forcing are better indicators of advanced forecasts than SSH.

Adding observations through the SWOT OSSE shows regular periodic intervals when SWOT swaths cover the domain with the boundary moving to smaller scales and intervals when SWOT swaths are absent from the domain with the boundary moving to larger scales. The OSSE also shows the separation of time and space scales of constrained features. The largest constrained features cover time scales of 15 days and longer, while the smaller constrained features occur only at time scales of 100 days and longer.

In the 1 km setup, the unconstrained local variance $\hat{v}_U^2$ has skill in predicting the forecast errors since the majority of the errors are due to features at scales below the constrained. In order for $\hat{v}_U^2$ to be a predictor, model resolution must be sufficient to represent unconstrained processes of importance. Twice the unconstrained variance must be included in observation errors in addition to representativeness errors. Based on the PSD of the 1 km model (Fig. 2), the unconstrained errors are larger than representativeness errors, which are typically larger than instrument errors.

The equivalent results of the unconstrained variance versus ensemble variance in Fig. 7 require several critical factors to be met. The locations of unconstrained variability must have predictability, and the areas of unconstrained variability are controlled by the larger scale constrained flow field. We must know the boundary position between the constrained and unconstrained portions of the fields, and the LASER observations play a key role in estimating the boundary scale. The forecast model must have sufficient resolution to represent the majority of variance at unconstrained scales. Finally, the majority of forecast error must be in the unconstrained band, and the PSD slope is an important factor in determining this (Fig. 2).

Knowing the separation between constrained and unconstrained energy allows high-resolution forecasts to be used most effectively. The constrained portion of the solution has lower errors. Information in the unconstrained local variance could be exploited in statistical forecasts (Piterbarg, 2001). We must continue to extend numerical representation of smaller scale dynamics, as these affect dispersion at scales well below model resolutions (Poje et al., 2014). Theory to exploit unconstrained statistical information in dispersion prediction has been utilized in the ocean (Haza et al., 2007). Ultimately, by knowing and exploiting both constrained and unconstrained information, we can justify continued increase in ocean model resolution beyond scales resolved by present observing networks.

CRediT authorship contribution statement

Gregg Jacobs: Coordinated numerical model experiments, Analyses of LASER drifter observations, Development, Writing - original draft.

Joseph M. D'Addezio: Set up and produced the OSSE model experiment results, Computed the spectral analyses, Produced figures, Writing - original draft. **Hans Ngodock:** Coordinated assimilation experiments, Analyses of GLAD drifters, Writing - original draft. **Innocent Souopgui:** Produced the OSSE experiment results, Writing - original draft.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research is funded by the NRL work unit Submesoscale Prediction of Eddies from Altimeter Retrieval (SPEAR), USA and a grant from BP/The Gulf of Mexico Research Initiative to the Consortium for Advanced Research on the Transport of Hydrocarbon in the Environment (CARTHE), USA. This paper is contribution NRL/JA-7320-20-4940 and has been approved for public release. The LASER drifter data analyzed during the current study are available in the Gulf of Mexico Research Initiative repository, https://doi.org/10.7266/N7W0940J and https://doi.org/10.7266/N7QN656H.

Appendix A. Separating large and small scales and the definition of local unconstrained local variability

Define the boundary scale as L_b, and the Gaussian convolution kernel at a grid point (x_o, y_o):

$$k\left(x_i, y_j\right) = mask\left(x_i, y_j\right) \exp\left(\left(-\left(x_i - x_o\right)^2 - \left(y_i - y_o\right)^2\right)/L_b^2\right). \tag{7}$$

Because land boundaries create an irregular ocean domain, the definition of the kernel includes the value $mask\left(x_i, y_j\right)$ defined as 1 if $\left(x_i, y_j\right)$ is water and 0 if land. In addition, $mask\left(x_i, y_j\right)$ is set to 0 if the point $\left(x_i, y_j\right)$ extends beyond the boundary of the model domain.

The constrained field of variable v at a grid point $\left(x_o, y_o\right)$ is defined as:

$$v_C\left(x_o, y_o\right) = \frac{1}{No} \sum_{i=i0-3L_b, j=j0-3L_b}^{i0+3L_b, j0+3L_b} v\left(x_i, y_j\right) k\left(x_i, y_j\right), \text{ where} \tag{8}$$

$$No = \sum_{i=i0-3L_b, J=j0-3L_b}^{i0+3L_b, j0+3L_b} k\left(x_i, y_j\right) \tag{9}$$

The unconstrained field of v is

$$v_U\left(x_o, y_o\right) = v\left(x_o, y_o\right) - v_C\left(x_o, y_o\right). \tag{10}$$

The local unconstrained variability at a grid point $\left(x_o, y_o\right)$ is

$$\hat{v}_U\left(x_o, y_o\right) = \sqrt{\frac{1}{No} \sum_{i=i0-3L_b, j=j0-3L_b}^{i0+3L_b, j0+3L_b} v_U^2\left(x_i, y_j\right) k\left(x_i, y_j\right)} \tag{11}$$

Appendix B. Expected errors in constrained and unconstrained portions of the solution

Separate the true ocean variable v based on the boundary scale so that it is composed of constrained and unconstrained components:

$$v_t = v_{tC} + v_{tU} \tag{12}$$

Similarly, separate the model variable

$$v_m = v_{mC} + v_{mU} \tag{13}$$

The model constrained field is equal to the true constrained field plus an error:

$$v_{mC} = v_{tC} + v_{mC\varepsilon} \tag{14}$$

The squared error of the model constrained field is:

$$\begin{aligned} \varepsilon_{mC}^2 &= \left(v_{mC} - v_t\right)^2 \\ &= \left(v_{tC} + v_{mC\varepsilon} - \left(v_{tC} + v_{tU}\right)\right)^2 \\ &= \left(v_{mC\varepsilon} - v_{tU}\right)^2 \end{aligned} \tag{15}$$

The squared error of the total model field is:

$$\begin{aligned} \varepsilon_m^2 &= \left(v_m - v_t\right)^2 \\ &= \left(v_{tC} + v_{mC\varepsilon} + v_{mU} - \left(v_{tC} + v_{tU}\right)\right)^2 \\ &= \left(v_{mC\varepsilon} + v_{mU} - v_{tU}\right)^2 \end{aligned} \tag{16}$$

Assume that the errors of the model constrained field are not correlated to other fields (either model or true) and that the model and true unconstrained fields are not correlated to each other:

$$\begin{aligned} &\langle v_{mC\varepsilon}, v_{tU} \rangle = 0 \\ &\langle v_{mC\varepsilon}, v_{mU} \rangle = 0 \\ &\langle v_{mU}, v_{tU} \rangle = 0 \end{aligned} \tag{17}$$

where the brackets indicate a statistical expected value. Then the expected values of the squared errors are:

$$\left\langle \varepsilon_{mC}^2 \right\rangle = \left\langle v_{mC\varepsilon}^2 \right\rangle + \left\langle v_{tU}^2 \right\rangle \tag{18}$$

$$\left\langle \varepsilon_m^2 \right\rangle = \left\langle v_{mC\varepsilon}^2 \right\rangle + \left\langle v_{mU}^2 \right\rangle + \left\langle v_{tU}^2 \right\rangle \tag{19}$$

Here we make two points. First, the errors of the full model field relative to observations are composed of the errors in the constrained field, the model variance in the unconstrained band, and the true world variance in the unconstrained band. We also conclude that the full model field errors are larger than the errors of only the constrained model field due to the additional model unconstrained variance $\left\langle v_{mU}^2 \right\rangle$. Assume that the model unconstrained variance and the true world unconstrained variance are both equal to the same value $\left\langle v_U^2 \right\rangle$. Then the expected squared errors of the constrained and total model fields are

$$\left\langle \varepsilon_{mC}^2 \right\rangle = \left\langle v_{mC\varepsilon}^2 \right\rangle + \left\langle v_U^2 \right\rangle \tag{20}$$

$$\left\langle \varepsilon_m^2 \right\rangle = \left\langle v_{mC\varepsilon}^2 \right\rangle + 2\left\langle v_U^2 \right\rangle \tag{21}$$

The model constrained field error is a constant due to errors in the constrained variable plus the variance in the unconstrained band. The error in the total model field is higher than the variance of the unconstrained. This is a result of the model and true unconstrained variability being uncorrelated. Considering errors within the unconstrained band, the model error variance is twice the unconstrained variance $\left\langle \hat{v}_U^2 \right\rangle$.

If the local spatial variance of the unconstrained scales in the model $\hat{v}_U^2$ is an estimate of the unconstrained variance, then the expected squared error in the model constrained field at a location is a constant due to the error in the model constrained field plus the estimated unconstrained variance $\hat{v}_U^2$. The expected squared error in the model total field is the same constant plus twice the estimated unconstrained variance $\hat{v}_U^2$.

References

Barron, C.N., Kara, A.B., Martin, P.J., Rhodes, R.C., Smedstad, L.F., 2006. Formulation, implementation and examination of vertical coordinate choices in the Global Navy Coastal Ocean Model (NCOM). Ocean Model. 11, 347–375.

Capet, X., Mcwilliams, J.C., Mokemaker, 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.

Carrier, M.J., Ngodock, H., Smith, S., Jacobs, G., Muscarella, P., Ozgokmen, T., Haus, B., Lipphardt, B., 2014. Impact of assimilating ocean velocity observations inferred from Lagrangian drifter data using the NCOM-4DVAR*. Mon. Weather Rev. 142, 1509–1524.

D'Addezio, J.M., Bingham, F.M., Jacobs, G.A., 2019a. Sea surface salinity subfootprint variability estimates from regional high-resolution model simulations. Remote Sens. Environ. 233, 111365.

D'Addezio, J.M., Smith, S., Jacobs, G.A., Helber, R.W., Rowley, C., Souopgui, I., Carrier, M.J., 2019b. Quantifying wavelengths constrained by simulated SWOT observations in a submesoscale resolving ocean analysis/forecasting system. Ocean Model. 135, 40–55.

D'Asaro, E., Lee, C., Rainville, L., Harcourt, R., Thomas, L., 2011. Enhanced turbulence and energy dissipation at ocean fronts. Science 332, 318–322.

Egbert, G.D., Erofeeva, S.Y., 2002. Efficient inverse modeling of barotropic ocean tides. J. Atmos. Ocean. Technol. 19, 183–204.

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

Fu, L.-L., Ubelmann, C., 2014. On the transition from profile altimeter to swath altimeter for observing global ocean surface topography. J. Atmos. Ocean. Technol. 31, 560–568.

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.

Gilleland, E., Ahijevych, D., Brown, B.G., Casati, B., Ebert, E.E., 2009. Intercomparison of spatial forecast verification methods. Weather Forecast. 24, 1416–1430.

Gilleland, E., Ahijevych, D.A., Brown, B.G., Ebert, E.E., 2010. Verifying forecasts spatially. Bull. Am. Meteorol. Soc. 91, 1365–1376.

Griffa, A., Piterbarg, L.I., Özgökmen, T., 2004. Predictability of Lagrangian particle trajectories: effects of smoothing of the underlying Eulerian flow. J. Mar. Res. 62, 1–35.

Haza, A.C., Piterbarg, L.I., Martin, P., Özgökmen, T.M., Griffa, A., 2007. A Lagrangian subgridscale model for particle transport improvement and application in the Adriatic Sea using the Navy Coastal Ocean Model. Ocean Model. 17, 68–91.

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. Naval Research Lab Stennis Detachment Stennis Space Center Ms Oceanography Div.

Hodur, R.M., 1997. The Naval Research Laboratory's coupled ocean/atmosphere mesoscale prediction system (COAMPS). Mon. Weather Rev. 125, 1414–1430.

Hoskins, B.J., 1982. The mathematical theory of frontogenesis. Annu. Rev. Fluid Mech. 14, 131–151.

Jacobs, G.A., Bartels, B.P., Bogucki, D.J., Beron-Vera, F.J., Chen, S.S., Coelho, E.F., Curcic, M., Griffa, A., Gough, M., Haus, B.K., 2014a. Data assimilation considerations for improved ocean predictability during the Gulf of Mexico Grand Lagrangian deployment (GLAD). Ocean Model. 83, 98–117.

Jacobs, G.A., D'Addezio, J.M., Bartels, B., Spence, P.L., 2020. Constrained scales in ocean forecasting. Adv. Space Res. http://dx.doi.org/10.1016/j.asr.2019.09.018.

Jacobs, G.A., Huntley, H.S., Kirwan Jr., A., Lipphardt Jr., B.L., Campbell, T., Smith, T., Edwards, K., Bartels, B., 2016. Ocean processes underlying surface clustering. J. Geophys. Res.: Oceans 121, 180–197.

Jacobs, G.A., Richman, J.G., Doyle, J.D., Spence, P.L., Bartels, B.P., Barron, C.N., Helber, R.W., Bub, F.L., 2014b. Simulating conditional deterministic predictability within ocean frontogenesis. Ocean Model. 78, 1–16.

Le Traon, P.-Y., Dibarboure, G., Jacobs, G., Martin, M., Rémy, E., Schiller, A., 2017. Use of Satellite Altimetry for Operational Oceanography, Satellite Altimetry over Oceans and Land Surfaces. CRC Press, Boca Raton, pp. 581–608.

Lellouche, J.-M., Greiner, E., Le Galloudec, O., Garric, G., Regnier, C., Drevillon, M., Benkiran, M., Testut, C.-E., Bourdalle-Badie, R., Gasparin, F., 2018. Recent updates to the copernicus marine service global ocean monitoring and forecasting real-time 1/ 12° high-resolution system. Ocean Sci. 14, 1093–1126.

Lemieux, J.F., Beaudoin, C., Dupont, F., Roy, F., Smith, G.C., Shlyaeva, A., Buehner, M., Caya, A., Chen, J., Carrieres, T., 2016. The Regional Ice Prediction System (RIPS): verification of forecast sea ice concentration. Q. J. R. Meteorol. Soc. 142, 632–643.

Li, Z., McWilliams, J.C., Ide, K., Farrara, J.D., 2015. A multiscale variational data assimilation scheme: formulation and illustration. Mon. Weather Rev. 143, 3804–3822.

Li, Z., Wang, J., Fu, L.L., 2019. An observing system simulation experiment for ocean state estimation to assess the performance of the SWOT mission: Part 1—A twin experiment. J. Geophys. Res.: Oceans 124, 4838–4855.

Liu, Y., Weisberg, R.H., Hu, C., Zheng, L., 2011. Trajectory forecast as a rapid response to the Deepwater Horizon oil spill. Monitoring and Modeling the Deepwater Horizon Oil Spill: A Record-Breaking Enterprise. Geophys. Monogr. Ser. 195, 153–165.

McLay, J.G., Bishop, C.H., Reynolds, C.A., 2008. Evaluation of the ensemble transform analysis perturbation scheme at NRL. Mon. Weather Rev. 136, 1093–1108.

McWilliams, J.C., 2016. Submesoscale currents in the ocean. Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 472, 20160117.

Metzger, E.J., Smedstad, O.M., Thoppil, P.G., Hurlburt, H.E., Cummings, J.A., Wallcraft, A.J., Zamudio, L., Franklin, D.S., Posey, P.G., Phelps, M.W., 2014. US Navy operational global ocean and Arctic ice prediction systems. Oceanography 27, 32–43.

Mittermaier, M., Roberts, N., 2010. Intercomparison of spatial forecast verification methods: Identifying skillful spatial scales using the fractions skill score. Weather Forecast 25, 343–354.

Muscarella, P.A., Carrier, M., Ngodock, H., 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.

Novelli, G., Guigand, C.M., Cousin, C., Ryan, E.H., Laxague, N.J., Dai, H., Haus, B.K., Özgökmen, T.M., 2017. A biodegradable surface drifter for ocean sampling on a massive scale. J. Atmos. Ocean. Technol. 34, 2509–2532.

Olascoaga, M., Beron-Vera, F., Haller, G., Triñanes, J., Iskandarani, M., Coelho, E., Haus, B., Huntley, H., Jacobs, G., Kirwan, A., 2013. Drifter motion in the Gulf of Mexico constrained by altimetric Lagrangian coherent structures. Geophys. Res. Lett. 40, 6171–6175.

Özgökmen, T.M., Boufadel, M., Carlson, D.F., Cousin, C., Guigand, C., Haus, B.K., Horstmann, J., Lund, B., Molemaker, J., Novelli, G., 2018. Technological advances for ocean surface measurements by the consortium for advanced research on transport of hydrocarbons in the environment (CARTHE). Mar. Technol. Soc. J. 52, 71–76.

Piterbarg, L.I., 2001. Short-term prediction of Lagrangian trajectories. J. Atmos. Ocean. Technol. 18, 1398–1410.

Poje, A.C., Özgökmen, T.M., Lipphardt, B.L., Haus, B.K., Ryan, E.H., Haza, A.C., Jacobs, G.A., Reniers, A., Olascoaga, M.J., Novelli, G., 2014. Submesoscale dispersion in the vicinity of the Deepwater Horizon spill. Proc. Natl. Acad. Sci. 111, 12693–12698.

Qiu, B., Chen, S., Klein, P., Sasaki, H., Sasai, Y., 2014. Seasonal mesoscale and submesoscale eddy variability along the North Pacific Subtropical Countercurrent. J. Phys. Oceanogr. 44, 3079–3098.

Richman, J.G., Arbic, B.K., Shriver, J.F., Metzger, E.J., Wallcraft, A.J., 2012. Inferring dynamics from the wavenumber spectra of an eddying global ocean model with embedded tides. J. Geophys. Res.: Oceans 117.

Roache, P.J., 1998. Fundamentals of Computational Fluid Dynamics. Hermosa Publishers, Albuquerque, NM.

Rodríguez, E., Wineteer, A., Perkovic-Martin, D., Gál, T., Stiles, B.W., Niamsuwan, N., Rodriguez Monje, R., 2018. Estimating ocean vector winds and currents using a Ka-band pencil-beam doppler scatterometer. Remote Sens. 10, 576.

Roemmich, D., Johnson, G.C., Riser, S., Davis, R., Gilson, J., Owens, W.B., Garzoli, S.L., Schmid, C., Ignaszewski, M., 2009. The Argo Program: Observing the global ocean with profiling floats. Oceanography 22, 34–43.

Rowley, C., Mask, A., 2014. Regional and coastal prediction with the relocatable ocean nowcast/forecast system. Oceanography 27, 44–55.

Rudnick, D.L., 2016. Ocean research enabled by underwater gliders. Annu. Rev. Mar. Sci. 8, 519–541.

Sandery, P.A., Sakov, P., 2017. Ocean forecasting of mesoscale features can deteriorate by increasing model resolution towards the submesoscale. Nature Commun. 8, 1–8.

Schiller, A., Brassington, G.B., Oke, P., Cahill, M., Divakaran, P., Entel, M., Freeman, J., Griffin, D., Herzfeld, M., Hoeke, R., 2020. Bluelink ocean forecasting Australia: 15 years of operational ocean service delivery with societal, economic and environmental benefits. J. Oper. Oceanogr. 13, 1–18.

Shcherbina, A.Y., D'Asaro, E.A., Lee, C.M., Klymak, J.M., Molemaker, M.J., McWilliams, J.C., 2013. Statistics of vertical vorticity, divergence, and strain in a developed submesoscale turbulence field. Geophys. Res. Lett. 40, 4706–4711.

Smith, N.R., 2000. The global ocean data assimilation experiment. Adv. Space Res. 25, 1089–1098.

Smith, S., Cummings, J., Rowley, C., Chu, P., Shriver, J., Helber, R., Spence, P., Carroll, S., Smedstad, O., Lunde, B., 2011. Validation Test Report for the Navy Coupled Ocean Data Assimilation 3D Variational Analysis (NCODA-VAR) System, Version 3.43. NRL Report NRL/MR/7320-11-9363, Naval Research Laboratory, Stennis Space Center, MS.

Souopgui, I., D'Addezio, J.M., Rowley, C.D., Smith, S.R., Jacobs, G.A., Helber, R.W., Yaremchuk, M., Osborne, J.J., 2020. Multi-scale assimilation of simulated SWOT observations. Ocean Model. 154, 101683.

Thomas, L.N., Taylor, J.R., Ferrari, R., Joyce, T.M., 2013. Symmetric instability in the Gulf Stream. Deep Sea Res. II 91, 96–110.

Wei, M., Jacobs, G., Rowley, C., Barron, C.N., Hogan, P., Spence, P., Smedstad, O.M., Martin, P., Muscarella, P., Coelho, E., 2013. The performance of the US Navy's RELO ensemble, NCOM, HYCOM during the period of GLAD at-sea experiment in the Gulf of Mexico. Deep Sea Res. II.

Wei, M.Z., Rowley, C., Martin, P., Barron, C.N., Jacobs, G., 2014. The US Navy's RELO ensemble prediction system and its performance in the Gulf of Mexico. Q. J. R. Meteorol. Soc. 140, 1129–1149.

Zhong, Y., Bracco, A., 2013. Submesoscale impacts on horizontal and vertical transport in the Gulf of Mexico. J. Geophys. Res.: Oceans 118, 5651–5668.