The Global Mesoscale Eddy Available Potential Energy Field in Models and Observations

C. A. Luecke,[1] B. K. Arbic,[1] S. L. Bassette,[2] J. G. Richman,[3] J. F. Shriver,[4] M. H. Alford,[5] O. M. Smedstad,[6] P. G. Timko,[1,7] D. S. Trossman,[8,9] A. J. Wallcraft[4]

Corresponding author: C. A. Luecke, Department of Earth and Environmental Sciences, University of Michigan, 2534 C.C. Little Building, Ann Arbor MI 48109, USA. (cluecke@umich.edu)

[1]Department of Earth and Environmental

Abstract. Global maps of the mesoscale Eddy Available Potential Energy (EAPE) field at a depth of 500m are created using potential density anomalies in a high-resolution $\frac{1}{12.5}^{\circ}$ global ocean model. Maps made from both a free-running simulation and a data-assimilative reanalysis of the HYbrid Coordinate Ocean Model (HYCOM) are compared with maps made by other researchers from density anomalies in Argo profiles. The HYCOM and Argo maps display similar features, especially in the dominance of western boundary currents. The reanalysis maps match the Argo maps more closely, demonstrating the added value of data assimilation. Global averages of the simulation, reanalysis, and Argo EAPE all agree to within about 10%.

The model and Argo EAPE fields are compared to EAPE computed from temperature anomalies in a dataset of "Moored Historical Observations" (MHO) in conjunction with buoyancy frequencies computed from a global climatology. The MHO dataset allows for an estimate of the EAPE in high-frequency motions that is aliased into the Argo EAPE values. At MHO locations, 15-32% of the EAPE in the Argo estimates is due to aliased motions having periods of 10 days or less. Spatial averages of EAPE in HYCOM, Argo, and MHO data, agree to within 50% at MHO locations, with both model estimates lying within error bars observations.

Analysis of the EAPE field in an idealized model, in conjunction with published theory, suggests that much of the scatter seen in comparisons of dif-

Sciences, University of Michigan, Ann

ferent EAPE estimates is to be expected given the chaotic, unpredictable nature of mesoscale eddies.

Arbor, Michigan, USA.

1. Introduction

Low-frequency flow in the ocean is dominated by mesoscale eddies. These eddies have time scales on the order of 30-200 days, have length scales on the order of 50-200 km, and have been studied through many observational platforms including moored current meters and thermistors [e.g., *Richman et al.*, 1977; *Schmitz*, 1988], satellite altimetry [e.g., *Chelton et al.*, 2007], and surface drifters [e.g., *Lumpkin and Pazos*, 2007]. In recent years, increased computer power has led to high-resolution, three-dimensional ocean models that are able to simulate and forecast mesoscale eddies on a global scale [e.g., *Hecht and Hasumi*, 2008; *Chassignet et al.*, 2009]. Because global eddying models are used for forecasting oceanic flows and for dynamical process studies, the question arises as to how accurate these models are at representing energetic phenomena in the ocean. A number of recent studies have focused on comparisons of high-resolution three-dimensional ocean models with observational data. For example, model kinetic energy in both low-frequency [*Penduff et al.*, 2006; *Scott et al.*, 2010; *Thoppil et al.*, 2011] and tidal [*Timko et al.*, 2012, 2013] bands have been compared with kinetic energy measured from moored ocean current meters and surface drifters.

This paper presents global maps of mesoscale Eddy Available Potential Energy (EAPE) from both a free-running simulation and a reanalysis of the HYbrid Coordinate Ocean Model [HYCOM: *Chassignet et al.*, 2009], run on a $\frac{1}{12.5}^{\circ}$ global grid. The global maps of low-frequency EAPE computed from HYCOM are compared to global EAPE maps

---

[2]Department of Physics, University of

enabled by, and recently computed from, the Argo float array [*Roullet et al.*, 2014]. To our knowledge, no truly global model-data comparison of EAPE has been attempted until

Michigan, Ann Arbor, Michigan, USA.

[3]Center for Ocean-Atmospheric Prediction Studies, Florida State University, Tallahassee, FL

[4]Oceanographic Division, Naval Research Laboratory, Stennis Space Center, Mississippi, USA

[5]Scripps Institution of Oceanography, UC San Diego, La Jolla, USA

[6] QinetiQ North America, Stennis Space Center, MS, USA.

[7]Welsh Local Centre, Royal Meteorological Society, UK

[8]Goddard Earth Science Technology and Research, NASA GSFC, Greenbelt, Maryland, USA

[9]Department of Earth and Planetary Sciences, Johns Hopkins University, Baltimore, Maryland, USA

now.

The model-data comparison of EAPE in this work tests the ability of models to represent mesoscale eddy dynamics in the ocean interior. Available Potential Energy (APE) is defined as the amount of potential energy in a stratified fluid that is available for mixing and conversion into kinetic energy [*Munk and Wunsch*, 1998; *Huang*, 1998]. Although much literature has been focused on the subject of the time-mean APE [*Winters et al.*, 1995; *Saenz et al.*, 2012; *Tailleux*, 2013], there is little in the way of validating the realism of EAPE in realistic, high-resolution ocean models. The computation of EAPE differs from the computation of time-mean APE in the works cited above in that it requires no complex reference state or background potential energy, and instead relies upon a locally calculated mean-isopycnal state. For the purpose of this paper, a low-frequency "eddy" is defined as a departure from a long-term temporal mean, with a period of greater than two days. EAPE –the energy of fluctuations in density around a time-mean– is a function of both background stratification and isopycnal fluctuations, and is, therefore, a critical component of the structure of the pycnocline [*Gnanadesikan*, 1999]. Validating the ability of HYCOM to predict EAPE lays the groundwork for further study of energetics in high-resolution ocean models.

We compare the HYCOM simulation, HYCOM reanalysis, and Argo EAPE to independent estimates computed from anomaly time series at 1,057 individual moored historical instrument observations, in conjunction with buoyancy frequencies taken from a global climatology. Hereafter, the Moored Historic Observations will be referred to as MHO. An advantage of using MHO instruments as an observational dataset for EAPE is that the relatively high temporal resolution of many MHO instruments permits us to separate the

EAPE due to low-frequency mesoscale eddy motions from EAPE due to high-frequency motions such as internal gravity waves and tides. High-frequency motions are aliased into Argo records which have a $\sim$ 10 day cycle time. Disadvantages of the MHO dataset include its sparse spatial coverage, a likely seasonal bias due to the fact that it is simpler to deploy moorings in summer than in winter, and the lack of salinity data at the majority of MHO locations. Due to the lack of salinity data, the MHO EAPE are calculated using temperature as a proxy for density, a procedure that has a long historical precedent in oceanography [*Dantzler*, 1977; *Wunsch*, 1999]. The use of temperature as a proxy for density is less accurate in the upper ocean due to density compensated motions [*Rudnick and Ferrari*, 1999]. For this reason, we compute EAPE only at MHO instruments that are 60m or more below the surface. We also test the differences between EAPE computed from density variations versus EAPE computed using temperature as a proxy for density, using both HYCOM output and output from modern McLane in-situ profiler instruments [*Doherty et al.*, 1999].

An additional goal of this paper is to test whether modeled EAPE is improved with data assimilation. Recent advances in both remotely sensed and in-situ oceanic observations have dramatically increased the amount of data available for assimilation in a reanalysis. In this paper we will demonstrate that the data assimilation in the reanalysis improves the background buoyancy frequency, a critical constituent of the EAPE calculation. We will also show that the spatial structures of EAPE in the global HYCOM reanalysis more closely resemble the spatial structures seen in the Argo maps than do the spatial structures in the HYCOM simulation map. Finally, we employ the *Murphy* [1988] skill score to further quantify improvement in the reanalysis relative to the simulation. In the case

of global EAPE, the skill score is computed between HYCOM and Argo EAPE, while at MHO locations, the skill score is computed using the square of temperature anomalies in HYCOM and MHO.

We will show that there is considerable scatter whenever averages of the estimates used here – HYCOM simulation, HYCOM reanalysis, MHO, and Argo– are compared against each other. A brief analysis of the EAPE fields in a horizontally homogeneous quasi-geostrophic (QG) turbulence model demonstrates that statistical scatter is inherent in EAPE comparisons made from a chaotic mesoscale eddy field, even without the complicating factors present in the actual ocean (laterally inhomogeneous environments, varying topography etc.). Established theory [*Flierl and McWilliams*, 1977] in conjunction with our analysis of the idealized model quantifies the expected scatter as a function of record length, which is typically on the order of several months to a year for MHO records.

2. Models, Observational Data, and Methods

We use a free running global HYCOM simulation run for 20 years from 1993-2012, and a data-assimilative reanalysis run over the same time period. For our global comparison with Argo, all HYCOM EAPE is computed using density anomalies. Due to the size of the 20 year model output, we chose output from one model year (2003) for our global comparisons. Our global comparisons are made at a depth of 500m, close to the depth of the EAPE maximum presented in *Roullet et al.* [2014]. Model EAPE is calculated using the native $\frac{1}{12.5}^\circ$ resolution, and then decimated to $\frac{1}{4}^\circ$ degree resolution for mapping purposes. For our MHO-HYCOM comparison, we choose an EAPE computation method that is straightforward and consistent with the limitations inherent in the MHO data. Salinity data are generally not available alongside the historical temperature time series

records. With this in consideration, we take temperature anomalies – low-frequency departures from a temporal mean – as a proxy for density anomalies. Where Argo EAPE is compared with our MHO estimates, we still use values drawn from *Roullet et al.* [2014], which are computed from density anomalies. The MHO-Argo comparisons are thus inconsistent in this way. The errors introduced in using temperature as a proxy for density are examined using in-situ data from McLane profilers [*Doherty et al.*, 1999]. Because MHO temperature time-series data is generally too sparse in the vertical direction to allow for the quantification of oceanic stratification at mooring locations, we use the World Ocean Atlas 2009 [WOA: *Locarnini et al.*, 2010; *Antonov et al.*, 2010] records to compute background Brunt-Väisälä buoyancy frequency at the mooring sites. For the sake of clarity, EAPE calculated from a combination of MHO data and WOA climatology will be referred to as "MHO EAPE."

2.1. Models

EAPE is computed from runs of the Naval Research Laboratory (NRL) Global Ocean Forecast System [GOFS: *Metzger et al.*, 2014]. The free-running simulation (hereafter, "simulation") and reanalysis both employ the HYbrid Coordinate Ocean Model (HYCOM) configured with 32 layers in the vertical direction, and an equatorial horizontal resolution of 0.08° ($\frac{1}{12.5^\circ}$ or 9 km) on a tripolar grid. The model is spun up for twenty years with the atmospheric forcing from an annual climatology of the National Center for Environmental Prediction Climate Forecast System (CFS) and is then run with hourly forcing from the CFS Reanalysis [*Saha et al.*, 2010] from 1993 to 2010 and from CFSv2 [*Saha et al.*, 2013] for the remaining two years (2011 and 2012). During the 20 year

spin-up, the simulation stratification drifts away from climatology as will be seen later. Neither the HYCOM simulation nor the reanalysis contain tides. For the simulation, designated internally at NRL as GLBb0.08_expt10.2, daily means are archived on the Navy Department of Defense (DoD) Shared Resource Center (DSRC) at Stennis Space Center.

For the reanalysis, designated at NRL as GLBb0.08_expt19.0/19.1, an analysis using the Navy Coupled Ocean Data Assimilation system [NCODA: *Cummings and Smedstad*, 2013] is performed daily. Daily means are archived at the DSRC at Stennis Space Center and are available at the HYCOM consortium server (https://hycom.org/dataserver/glb-reanalysis). While the model code and configuration are fixed for the reanalysis, the observing network changes significantly over the twenty years. Both satellite altimeters and Argo floats are important sources of observations to be assimilated in the reanalysis. Altimetric sea surface height anomalies are converted to synthetic profiles of temperature and salinity in NCODA. The number of altimeters available for assimilation varies between two and four over the twenty years of the run. During the last decade of the reanalysis, the Argo floats provide approximately 370 profiles of temperature and salinity per day.

While global maps of HYCOM EAPE are computed from one model year (2003), the full 20 year model output is used for our comparison to MHO EAPE. Because the model runs are 20 years in duration, we will refer to "20 year runs" and "20 years of output" although reanalysis output for year 2001 has been discarded due to data corruption in a large portion of the model output for that year. The majority of MHO records are of order one year in length, much shorter than the 20 year model outputs. To ensure a

consistent comparison between the long model runs and the shorter in-situ MHO data sets, we analyze individual model years as well as 20 year means.

2.2. Observations

Argo global EAPE fields are obtained from *Roullet et al.* [2014] on a $\frac{1}{2}^\circ$ grid. *Roullet et al.* [2014] can be consulted for a discussion of the methods employed to extract EAPE from Argo floats. Discussions of the Argo array are given in numerous sources including *Roemmich and Owens* [2000], www.argo.net, and www.argo.ucsd.edu. For our later comparisons involving MHO data, Argo EAPE from the closest location and nearest depth to the MHO instrument in question is used.

Time-series of temperature from moored historical observations (MHO) are obtained from the Global Multi-Archive Current Meter Database [GMACMD: *Scott et al.*, 2010; *Timko et al.*, 2012, 2013], and can be found at http://stockage.univ-brest.fr/~scott/GMACMD/gmacmd.html. These records span from 1974 to 2008, and are generally not contemporaneous with the HYCOM runs or Argo data. To avoid problems with abyssal temperature records, in which the small magnitude of the fluctuating temperature variations are not well resolved by the measurements, we use only instruments in water depths of 1,500 meters or less. Only locations with seafloor depths greater than 500 meters are included, owing to the lack of climatological data for computing buoyancy frequency on the shelf. Data within the mixed layer shallower than 60 meters are excluded; at such locations the buoyancy frequency is often locally very small, causing the numerical EAPE estimates (see Equation 1) to become unphysically large. Additionally, within the mixed layer, temperature anomaly correlates less strongly with density fluctuations as a result of density compensation [*Rudnick and Ferrari*, 1999]. Finally, we use

only instruments between $65°N$ and $65°S$ in order to eliminate locations where salinity fluctuations play a larger role in density anomalies; the sparsity of Argo in polar regions is another reason to exclude them.

We select MHO temperature time-series that are longer than 180 days, and exclude records containing gaps in the time-series. The remaining records are then visually inspected and quality controlled for instrument errors and other problems such as severe discretization of temperature anomalies, thermistor calibration drift, and non-stationarity in the variance of temperature signals. The total number of instruments excluded for these reasons is relatively small ($< 10\%$). Our selection criteria yields a total of 1,057 instruments distributed globally. The horizontal locations of the MHO instruments are given in the left panel of Figure 1. The spatial coverage is sparse and uneven. Some basins (e.g. the North Atlantic and North Pacific) are relatively well sampled, while others (e.g. the South Pacific) are sampled very little. The vertical coverage of the MHO dataset is given in the right panel of Figure 1. Generally, the number of records is larger in shallower depths (upper 500 or so meters). To remove high-frequency motions such as internal tides and internal gravity waves, we low-pass filter the MHO records using a second-order Butterworth filter with a 2 day cut-off period. We also remove linear trends from the time series in order to remove seasonal trends not fully resolved by records shorter than a year. In order to validate our use of temperature time series in this paper, we compare temperature and density variance using McLane in-situ profile data [*Doherty et al.*, 1999] at 10 locations marked by a red "X" in the left panel of Figure 1. At these 10 locations, time series of temperature and salinity were taken at selected depths between 200 and 1300 meters. This yielded a total of 31 distinct temperature and salinity time series where

density variance could be compared with density variance using temperature as a proxy for density.

2.3. Calculation of EAPE

Consistent with the method of calculation used in the global Argo estimates [*Roullet et al.*, 2014], we use the "APE_3" term in *Kang and Fringer* [2010] to calculate EAPE. Note that, as is standard, the potential densities ρ and density anomalies ρ' are computed with respect to the local vertical position. To first order, the "APE_3" term, adjusted by a factor of ρ_0, is given by:

$$EAPE = \frac{g^2 \rho'^2}{2\rho_0^2 N^2}, \tag{1}$$

where g is the acceleration due to gravity, ρ' is a seawater potential density anomaly defined as a departure from a time-mean of density, ρ_0 is the average density of seawater, and the square of the Brunt-Väisälä buoyancy frequency $N^2 = -\frac{g}{\rho_0}\frac{d\rho}{dz}$, where z represents the vertical coordinate. We have divided the *Kang and Fringer* [2010] EAPE by an additional factor of ρ_0, in order to obtain units consistent with the units given by *Roullet et al.* [2014].

For our HYCOM global map calculations, model density output at a depth of 500m is used to calculate density anomalies, and N^2 is calculated with a centered difference derivative using model output at 550m and 450m. For the computation of global averages from HYCOM maps, model output is interpolated to the Argo native resolution of $\frac{1}{2}^\circ$. Spatial averages are computed only over locations where both model output and Argo data is available. In our analysis of the MHO, the buoyancy frequency is determined using WOA annual mean temperature and salinity climatology and the TEOS-10 Gibbs

Sea Water equation of state package [*McDougall and Barker*, 2011]. In the calculation of HYCOM EAPE at MHO locations, the modeled buoyancy frequency is computed from annually averaged model output interpolated to a depth level grid, in analogy to the WOA climatology used in conjunction with the MHO data.

For the calculation of MHO and HYCOM EAPE at MHO locations, we estimate a density anomaly ρ' from the temperature anomaly T', using a linearized equation of state for seawater:

$$\rho' \approx \rho_0[-\alpha T' + \beta S'] \approx -\rho_0 \alpha T', \tag{2}$$

where α and β are the thermal expansion and haline contraction coefficients of seawater respectively. The prime notation again denotes a departure from the time-averaged value, and the salinity term S' is dropped due to its absence in most historical in-situ measurements. For consistency, the salinity term is also dropped in the calculation of HYCOM EAPE at MHO locations. The coefficient α is calculated locally at instrument locations and depths from either the World Ocean Atlas climatology (for MHO data) or from annual averages (for HYCOM output.)

2.4. Statistical Analysis

We use several metrics to quantify differences between EAPE in HYCOM and observations. For our global comparisons of HYCOM and Argo, we present global area-weighted averages of EAPE. For our MHO location comparison, we present HYCOM, MHO and Argo EAPE, as well as the constituent terms in the calculation of the EAPE, T'^2 and N^2. A linear regression coefficient A for EAPE is calculated using standard methods. The

ratio γ of the means of the model and observations is defined as:

$$\gamma = \frac{\sum_{i=1}^{n} EAPE_{model}}{\sum_{i=1}^{n} EAPE_{observed}}, \tag{3}$$

where i is a location index and n is the total number of MHO instruments used in the calculation. Additionally, a correlation coefficient R is calculated between pairs of estimates (model, MHO, Argo) across the MHO locations. The ideal values expressing a perfect comparison are equal to one for all of the metrics A, γ, and R. Following the above methodology, means and correlations for the constituent terms T'^2 and N^2 are also calculated. Comparison statistics denoted as "20 year mean" are calculated on output that has been binned yearly, then averaged over 20 years, and then spatially averaged, prior to the calculation of statistics.

In order to quantify improvement between the model simulation and reanalysis, we employ a skill score used in *Murphy* [1988]. The skill score (SS) is defined as:

$$SS(R, S, Obs) = 1 - [MSE(R, Obs)/MSE(S, Obs)], \tag{4}$$

Where the mean square error (MSE) is defined as:

$$MSE(R, Obs) = \frac{1}{n}\sum_{i=1}^{n}(R_i + Obs_i)^2, \tag{5}$$

Where R denotes model reanalysis predictions, S denotes model simulation predictions, and Obs denotes the observational predictions. In the case of our global comparisons with Argo, we use EAPE predictions, and in the case of our Model vs. WOA comparisons we use temperature variance. The skill score is positive when the accuracy of the reanalysis is greater than that of the simulation. $SS = 1$ represents a reanalysis that perfectly matches observations ($MSE(R, Obs) = 0$), while $SS = 0$ when $MSE(R, Obs) = MSE(S, Obs)$, representing no improvement in the reanalysis. Multiplying by 100 translates SS into a

measure of percent improvement.

Because we present spatial means as a metric for comparison between our model runs and observational data, it is convenient to include estimates of the error of these means. We employ bootstrap methods to estimate 95th percentile confidence intervals on our means. Bootstrapping is performed with N=1000 bootstrap re-samples of our original data. In the case of our global area-weighted integrals, where the global integral and non-weighted average differ slightly, we estimate the percent error using the global average, and apply it to the global integral.

3. Results

We first present global maps of EAPE from HYCOM and Argo. We then examine HYCOM, Argo, and MHO EAPE at the MHO locations. We compare HYCOM and MHO values of the EAPE constituent terms N^2, the square of the Brunt-Väisälä buoyancy frequency, and T'^2, the square of the low-passed temperature anomalies.

3.1. Global EAPE Maps

Global maps of EAPE at 500m computed from the HYCOM simulation and reanalysis (top and middle panels respectively, of Figure 2) show spatial structures consistent with the Argo map (bottom panel), such as increased EAPE in western boundary currents and in the Southern Ocean. Because Argo, as well as other observations including satellite altimetry, are used as a source of assimilative observations for the reanalysis, one would expect that the reanalysis EAPE maps would more closely reproduce Argo estimates, and this is indeed the case. The spatial structure in the reanalysis more closely resembles the Argo structure in several respects, confirming the added value of data assimilation.

Perhaps the most apparent improvement is the lack of an artificial South Atlantic "eddy train" in the reanalysis. The simulation contains a distinctive "eddy train," resulting from eddies escaping the Agulhas Current into the South Atlantic. We note that this "eddy train" is not a unique feature of the HYCOM simulation, and can be seen in other high-resolution simulations [*Maltrud and McClean*, 2005], where the train was diagnosed using sea surface hight variance and was not seen in altimeter observations. One possible cause of this "eddy train" is illustrated in *McClean et al.* [2011], where introduction of ocean-atmospheric fluxes in a coupled model is shown to improve the realism of these eddies. It is also possible that improvements in the ocean-atmosphere wind shear implemented in HYCOM could improve the dynamics in the region. The train results in a large local over-estimation of EAPE in the simulation when compared to Argo, while much of the rest of the simulation South Atlantic EAPE is below the Argo values. In the reanalysis however, the eddy train is no longer apparent, and the rest of the South Atlantic is more energetic, in line with the Argo maps.

The reanalysis also matches the spatial structure of the Argo maps in the Kuroshio and Gulf Stream regions more closely than the simulation does. For instance, in both Argo and the HYCOM reanalysis, the Gulf Stream hooks northward at about 45°W, while the HYCOM simulation does not. There are a number of factors that may cause a model to differ from observation. *Chassignet and Xu* [2017], for instance, show that resolution plays a role in the realism of the Gulf Stream. While the model runs used in this paper contain an energetic eddy field, they lack sufficient resolution to accurately portray all featured of western boundary currents [*Thoppil et al.*, 2011]. While the simulation over-estimates EAPE in the Indian Ocean between 10°S and 30°S, the reanalysis predicts EAPE

values closer to that of Argo. On the other hand, the simulation arguably recreates more accurately the 500m EAPE fields in the near equatorial Pacific, and parts of the Southern Ocean.

Figure 3 displays point to point comparisons of the 500m EAPE values between (left) the simulation and Argo, and (right) the reanalysis and Argo. Model output was decimated to the Argo native resolution of $\frac{1}{2}^\circ$ for both global point-to-point scatter-plots. The simulation exhibits more scatter, and lower correlation with respect to Argo, with $R = 0.52$ and $R = 0.65$ for the simulation and reanalysis respectively. To further quantify differences between the simulation and reanalysis, we compute a skill score (SS). Globally referenced to Argo EAPE, the skill score for HYCOM is $SS(\mathrm{R}_{EAPE}, \mathrm{S}_{EAPE}, \mathrm{Argo}_{EAPE}) = 0.50$ implying that the modeled EAPE is improved by close to 50 percent in the reanalysis compared to the simulation through the introduction of data assimilation.

The visual impression from Figure 2 is that the HYCOM simulation and reanalysis are both more energetic than the Argo maps. We confirm this by computing global averages (left column of Table 1). The model simulation provides the highest estimate with a global average EAPE of 185 ± 6 cm$^2 s^{-2}$, while the reanalysis is slightly lower at 183 ± 4 cm$^2 s^{-2}$. EAPE estimated using Argo provides the lowest global estimate at 168 ± 4 cm$^2 s^{-2}$, about 10% lower than the HYCOM estimates. Zonal averages of EAPE (Figure 4) demonstrate that both the simulation and reanalysis reproduce the qualitative structure of observed Argo EAPE between about $60^\circ N$ and $55^\circ S$. However both model runs generally predict slightly higher values of EAPE than does Argo between these latitudes. In the Southern Hemisphere between $35^\circ S$ and $55^\circ S$, the reanalysis over predicts EAPE. However the peaks are more closely positioned in latitude in the Argo observations

when compared to the simulation, which has a peak EAPE that is slightly shifted to the North. The simulation also has local EAPE maxima just poleward of $20°N$ and $20°S$ that do not appear in either the reanalisys or in Argo. From $30°N$ - $60°N$, both model runs agree closely in both latitudinal dependence and magnitude. However, once again, the simulation predicts a maximum EAPE slightly shifted towards the equator than either the reanalysis or Argo. Poleward of $55°S$ and $60°N$, there is a marked disconnect between Argo and HYCOM, with HYCOM exhibiting much higher zonal EAPE in the north, and much lower zonal EAPE in the south. In polar regions, weak stratification causes issues with our expression for EAPE. Additionally, south of $55°S$ Argo observations generally becomes sparse. These factors are most likely one cause of the somewhat poor model-data agreement in these regions. Prompted by the differences between the HYCOM and Argo EAPE maps, we present another observational EAPE estimate for comparison to HYCOM and Argo values. In the next 3 sections we compute EAPE from HYCOM, Argo, and MHO at the individual locations of the MHO instruments. We also compare the constituent terms used in our calculation of EAPE, using HYCOM output and MHO/WOA data.

3.2. Stratification

We compare N^2 in HYCOM versus the WOA observational climatology at the MHO locations shown in Figure 1. Both the simulation and reanalysis stratifications are initialized from the same 20 year spin-up period with climatological forcing. However, during the spin-up period, the simulation drifts away from the WOA climatology. We might expect that once inter-annual forcing is applied during the 1993-2012 analysis period, the simulation will remain relatively far from climatology, whereas the data assimilation employed in the reanalysis should result in more accurate model stratification. Consistent with

this expectation, the reanalysis outperforms the simulation with respect to climatological stratification over the duration of the model runs. A comparison of the stratification in the HYCOM runs versus WOA is shown in Figure 5. The scatter in Figure 5a (20 year simulation average vs. WOA annual average) is visually greater than in Figure 5b (20 year reanalysis average vs. WOA annual average). Both the reanalysis and the simulation have linear regression coefficients A and ratios of the means γ that are very close to one (Table 2), suggesting that when averaged over MHO locations, the model shows fairly good skill with respect to reproducing accurate stratification. When viewed in the context of point to point correlation, the reanalysis shows notable improvement, with correlation being higher between the reanalysis and WOA ($R = 0.97$) than between the simulation and WOA ($R = 0.78$). In a free-running simulation, the forcing that produces the stratification in the model is independent of the climatological stratification that was used to initialize the run. The dynamic stratification produced in the model is a product of the mixing occurring within the model. It is possible that the simulation drifts from the climatology because of inaccuracies in model mixing dynamics.

To illustrate the temporal drift of model stratification, we display in Figures 6a-c the vertical profiles of N^2, spatially averaged over the locations shown in Figure 1 for model years 1993, 2002, and 2012. The N^2 profile in the simulation has a maximum that is slightly deeper than seen in WOA. The maximum N^2 values in the reanalysis, while slightly larger than those in WOA, occur at depths that are noticeably closer to the depths seen in WOA. It is also evident that the temporal drifts in the stratification are greater in the simulation than in the reanalysis, with the pycnocline both broadening and deepening over time. The temporal drift of the stratifications in the reanalysis and the simulation

is also seen in Figure 6d, which displays the yearly spatially averaged N^2 profiles in the upper 300 meters. From this it is clear that the HYCOM reanalysis more closely recreates the depth of maximum stratification than does the simulation. It is important to note that the stratification drift of the model, along with the background stratification used in our calculation of MHO EAPE is biased toward the summer seasonal pycnocline. As many of our MHO observations are located shallower than 250m, this most likely contributes to a bias in both our model and MHO EAPE estimates at MHO locations presented in Section 3.4, as well as in the stratification temporal drift discussed here. Furthermore, as we are averaging over a globally distributed data-set of MHO locations, there is certainly introduced bias from global-merging of stratification. We believe however, that despite this, spatial and yearly means of N^2 still provide a useful metric to diagnose model performance.

3.3. Temperature and Density Anomaly

In this section we compare T'^2 between HYCOM and MHO. However, before we discuss the results of the MHO analysis, we display some typical model results at selected locations. Frequency spectra of the low-passed temperature anomalies at individual locations for MHO records, simulation, and reanalysis are shown in Figure 7. The simulation and reanalysis results are computed from model year 1993. The corresponding annually averaged Brunt-Väisälä stratification frequency profiles from the WOA climatology, simulation, and reanalysis output are also shown. Note that inter-annual EAPE variability in the model output does not appear to have significant impact on the agreement between model and data at most locations.

At the locations shown in Figures 7a and 7c, the simulation temperature spectra fall off more steeply at periods shorter than $\sim$ 20 days (0.05 cpd), while the reanalysis spectra lie closer to the MHO spectra. In Figures 7b and 7d, both the simulation and reanalysis temperature frequency spectra are in fairly close agreement with the MHO spectra. While HYCOM displays some skill at many locations, records exist where much of the variance occurring at periods of less than $\sim$ 20 days is not captured by the dynamics of the simulation alone. It is possible that assimilation in the reanalysis introduces some of this unresolved low-frequency mesoscale energy. At the same time, data assimilation increments can introduce artificially high levels of gravity waves through geostrophic adjustment. However, any such excess high-frequency energy would be reduced by the daily averages employed here. The better agreement of the reanalysis EAPE maps with Argo EAPE maps, as is seen in the global model vs. Argo comparison of Figure 2, and the close agreement of the globally averaged EAPE in the reanalysis versus the simulation, suggest that the data assimilation is not introducing gross inaccuracies in the reanalysis, at least in the daily averaged fields. Further evidence of this is seen in the frequency spectra of the reanalysis shown in Figure 7; no artificial peaks are seen as one approaches the high frequencies characteristic of gravity waves. At all four locations shown in Figure 7, buoyancy frequency profiles reveal a greater accuracy in the reanalysis over the simulation, consistent with the discussion in Section 3.2.

A point-to-point comparison of T'^2 in HYCOM vs. MHO, shown in Figure 8, reveals substantial scatter but a small bias. The scatter in the reanalysis plot (Figure 8b) is marginally tighter than that in the simulation plot (Figure 8a). However, the reduction in the scatter between the reanalysis and simulation comparisons with MHO for T'^2 is

not as visually striking as in the comparison with WOA of N^2 (Figure 5). While Table 2 shows that the correlation between the model and MHO T'^2 is improved in the reanalysis ($R = 0.76$) vs. the simulation ($R = 0.54$), the reanalysis overestimates T'^2 slightly more in a spatial average than does the simulation, as can be seen in the ratio of the means ($\gamma = 1.12$) vs. ($\gamma = 1.10$) for the reanalysis and simulation respectively. We compute the skill score (SS) of HYCOM with respect to MHO T'^2 as a metric for examining the impact of data assimilation on HYCOM. We find the skill score $SS(\mathrm{R}_{T'^2}, \mathrm{S}_{T'^2}, \mathrm{MHO}_{T'^2}) = 0.47$, again implying a close to 50 percent improvement between the reanalysis and simulation. The large spread in the comparisons of T'^2 dominates the scatter seen in our EAPE comparison in the next section, and will be further addressed in Section 4.

To test the accuracy of the approximate linearized equation of state discussed in Section 2.3, we compare the square of the inferred density anomaly $(-\rho_0 \alpha T')^2$ against the square of the density anomaly ρ'^2 calculated from the full non-linear equation of state [*McDougall and Barker*, 2011]. The left panel of Figure 9 displays a scatterplot of inferred verses actual density values taken from the HYCOM reanalysis at MHO locations. The right panel of Figure 9 displays the same comparison at the McLane profile locations described in section 2.2, where observations of both temperature and salinity are available. In both cases, the majority of locations lie close to the 1:1 line. In both the HYCOM and McLane exercises, the correlation between the actual and inferred density anomaly is $R = 0.93$. The reasonably good comparison seen in Figure 9 suggests that the error in EAPE introduced via this approximation (Equation 2) is smaller than other sources of scatter (discussed below, and in Section 4).

3.4. MHO EAPE

Both the simulation and reanalysis EAPE, averaged over all MHO locations, are about 16% lower than the MHO EAPE (Table 1). The HYCOM bias to lower energies evident in Table 1 can also be seen in the scatterplots of EAPE in Figure 10, where the bulk of the HYCOM points compared fall slightly to the right of the 1:1 comparison line. The reanalysis vs. MHO EAPE scatterplot (Figure 10b) does show a marginally tighter clustering than the scatterplot of EAPE in the simulation vs. MHO (Figure 10a). The statistical comparison metrics outlined in Section 2.4 display similar trends for the HYCOM comparisons to both MHO and Argo (Table 3). The value of γ effectively remains constant from the simulation to the reanalysis, while the linear regression coefficient A shows a slight improvement in the reanalysis when compared to the simulation. It is clear that R, the model correlation with MHO on a point to point basis, is improved in the reanalysis ($R = 0.84$) versus the simulation ($R = 0.56$), and this trend is repeated when the model EAPE is compared to Argo at the same locations ($R = 0.62$ vs. $R = 0.47$ respectively). Both model runs tend to slightly under estimate the EAPE relative to MHO estimates, although all mean EAPE values are comparable, especially when one considers the wide scatter shown in Figure 10.

Lastly, we compare our EAPE results at MHO locations with the EAPE calculated from Argo floats [*Roullet et al.*, 2014]. Because of the relatively frequent temporal sampling of the MHO records, we are able to provide estimates of the amount of aliasing that may be present in the Argo EAPE estimates. As a proxy for Argo sampling, we subsample the MHO records at 10 day intervals, and compute the variance. As in *Roullet et al.* [2014], any motions that occur at periods less than ten days will be aliased into the Argo-like

estimate of low-frequency variance. We also compute variance from a 10 day low-pass of the MHO temperature time series. The variance computed from the 10 day low-pass does not contain aliased high-frequency motions. We find that 32% of the variance in the Argo-like estimates is due to aliased contributions from motions with periods of 10 days or less. Because the MHO EAPE estimates presented here employ a two day low-pass filter, some but not all of the aliased energy that is in Argo is included in our MHO EAPE estimates. By comparing the variance in the Argo-like sampled MHO observations with the two day low-passed MHO data used in this paper, we conclude that Argo should over estimate EAPE by about 15% when compared to our MHO predictions, suggesting that in the locations of the MHO comparison, the true observed low-frequency energy should be slightly less than the the low-frequency Argo EAPE estimates. It is worth noting however, that the Argo model still predicts the lowest EAPE estimates when averaged over MHO locations (Table 1).

As shown in Figure 11, the vertical profiles of spatial mean EAPE between the simulation, reanalysis, MHO, and Argo at MHO locations are in qualitative agreement, with all EAPE estimates having a local minimum at about 200m, and a local interior maximum between 300 and 700m. For many of the depths between 200 and 600m, it is worth noting that both model EAPE predictions are "bounded" by observations, with MHO serving as an upper limit, and Argo EAPE representing a lower limit. The globally- and depth-averaged EAPE values given in Table 1 display similar trends; Argo is the lowest estimate, MHO is the highest estimate, and the models fall in between. The vertical profile of spatially averaged EAPE shown in Figure 11 is presented as a useful and interesting metric. However, due to the sparse sampling in the vertical, as well as the spatial sampling bias

depicted in Figure 1, it is not representative of the global vertical structure of EAPE.

Scatterplots of HYCOM and MHO EAPE values against Argo EAPE values, at the MHO locations shown in Figure 1, are displayed in Figure 12. Although on a point-to-point basis, the subplots in Figure 12 show large scatter, it is also true that EAPE calculated from the MHO, simulation and reanalysis are close to values in *Roullet et al.* [2014] when spatially averaged. As shown in Table 1, the spatial means of both the simulation and reanalysis EAPE (600 $\pm$90 and 598 $\pm$87 $cm^{-2}s^{-2}$ respectively) lie between the mean values inferred from MHO and Argo (709 $\pm$143 and 462 $\pm$55 $cm^{-2}s^{-2}$ respectively). The large errors of the EAPE calculated at MHO locations in comparison to our global EAPE estimates are due to the decreased sample size of our MHO database. Both the simulation and reanalysis EAPE are less than 16% lower than MHO and less than 30% higher than Argo estimates, and all four mean EAPE values are within a factor of 1.5 of one another Both the simulation and reanalysis EAPE means lie within the error bars of MHO and Argo, however the estimated error in Argo EAPE and MHO EAPE at MHO locations do not overlap. Interestingly, the scatter seen in MHO versus Argo EAPE estimates ($R = 0.60$) is comparable to the scatter in the comparisons of the HYCOM simulation and reanalysis versus Argo given in Table 3. The fact that two different observationally-based EAPE estimates yield a similar scatter to that seen in the model-Argo comparisons reinforces the notion that such model-data comparisons of mesoscale eddy fields are prone to large scatter. In the following section, we use an idealized model to show the influence of sampling times on this scatter.

4. Estimates of Inherent Scatter in Eddy Statistics

Mesoscale eddies are, by their nature, chaotic. In model-data comparisons such as the ones presented in this paper, the question arises as to how much of the "scatter" seen in the model-data scatterplots is due simply to the unpredictable nature of the underlying EAPE fluctuations. HYCOM and other realistic-domain high-resolution ocean models exhibit dynamical variability due to complex and varying topography, atmospheric forcing, and horizontal inhomogeneities arising from basin geometries. The scatter plots in Figures 3, 8, 10, and 12 are made from model output and observations that are impacted by all of these factors. However, some of the scatter is due simply to the fact that we are sampling a chaotic field, in both model output and observations, irrespective of the complexities introduced by the horizontally inhomogeneous factors described above.

Estimates of the temporal sampling requirements for chaotic systems have been previously made, giving predictions for the degree of spread expected in the variance of temperature time-series data of a given length. *Flierl and McWilliams* [1977] estimate the error in temperature variance (a quantity proportional to EAPE assuming a given buoyancy frequency) to be on the order of 20-30% for records of 700 days in length, and 10% or lower only for records longer than 15 years. Because many of the MHO records used in this data are on the order of a year in length, and because the MHO, HYCOM, and Argo EAPE estimates used here are not contemporaneous, we expect a substantial amount of spread in even the best model-data comparisons.

To illustrate the intrinsic spread expected in our model-data comparison, we compare EAPE between different grid points in a simulation of an idealized model that is horizontally homogeneous. As our idealized model is horizontally homogeneous, the confounding

spatially varying factors mentioned above –topography, atmospheric forcing, and basin geometry– are not present. The idealized model is quasi-geostrophic (QG), containing two vertical layers on an f-plane domain. The forcing for the QG model consists of a horizontally homogeneous mean flow that is vertically sheared to induce baroclinic instability. The model is damped by linear bottom Ekman friction with a decay coefficient R_2. The nondimensional bottom friction strength parameter is $\kappa = [\frac{R_2 L_d}{\overline{u_1}-\overline{u_2}}]$, where L_d is the first baroclinic mode Rossby radius of deformation, overbars denote an imposed time mean, and $\overline{u_1} - \overline{u_2}$ is the difference in the imposed (zonal) mean flow in the upper (1) and lower (2) layers. The value of κ in the simulation used here is 0.4. Snapshots of the model output are saved at every unit of non-dimensional time $[\frac{L_d}{\overline{u_1}-\overline{u_2}}]$. The correlation time is about 16.5 snapshots. Further details about the simulation used here can be found in *Arbic et al.* [2012, 2014] and references therein. Because we are using the QG simulation as an analogue for the mid-latitude mesoscale eddy field, we equate the 16.5 snapshot correlation time for the simulation with 40 days, a typical correlation time for mid-latitude mesoscale eddies [*Kuragano and Kamachi*, 2000; *Jacobs et al.*, 2001]. EAPE in the QG model is given by:

$$EAPE = \frac{1}{2}\frac{\psi_{bc}^2}{L_d^2} = \frac{1}{2}\frac{\delta(\psi_1 - \psi_2)^2}{(1+\delta)^2 L_d^2}, \tag{6}$$

where ψ_{bc} is the baroclinic streamfunction, ψ_1 and ψ_2 are the upper and lower layer streamfunctions, and δ is the ratio between the top and bottom layer thickness. In the simulation used here, $\delta = 0.2$.

Because our HYCOM-MHO comparisons involved about 1100 instruments, we compare EAPE at 1100 randomly selected unique locations in the QG model versus EAPE at 1100 different randomly selected locations. EAPE is averaged over 8 and 182 model correlation

time periods (approximately equivalent to 320 days and 20 years for mid-latitude oceanic eddies respectively). As predicted by *Flierl and McWilliams* [1977], the longer period of time averaging dramatically reduces the scatter between the point to point comparisons as seen in the difference between Figures 13a and 13b.

We are able to validate our QG model runs against the quantitative predictions *Flierl and McWilliams* [1977] make on the amount of error expected in temperature variance for various record durations. To make these comparisons, we use the QG analogue of T', the temperature anomaly. It can be shown using the thermal wind relation that the temperature anomaly $T' \propto (\psi_1 - \psi_2)$. The temporal variance of T' is then calculated for several different lengths of time. We use 100, 320, and 700 days, which correspond to temporal-averaging lengths assessed in *Flierl and McWilliams* [1977]. In order to quantify statistical errors in temperature variance in the model, we calculate the ratio of the standard deviation of the time-mean temperature variance computed over all 1100 points, to the magnitude of the time-mean temperature variance averaged over the same model grid points. For 100, 320, and 700 days of sampling time, we estimate the temperature variance error to be 63%, 36%, and 23% respectively, which agree well with the 60%, 40%, and 20% estimates made by *Flierl and McWilliams* [1977]. Additionally it takes 12 years of time averaging for our QG temperature variance error to drop below 10%, which lies fairly close to the 15 years predicted by *Flierl and McWilliams* [1977]. This result suggests that given the duration of sampling common in the MHO records, an order 30-40% error in EAPE estimates is to be expected. It is reasonable then, that the discrepancies in EAPE estimates displayed in Table 1 fall roughly within this range. The large error in temperature variance due to the chaotic nature of mesoscale eddies must certainly account

for some of the spread seen in the scatterplots of T'^2 and EAPE shown in Figures 3, 8, 10, and 12.

5. Conclusions

In this paper, we have assessed the ability of both a simulation and reanalysis of a three-dimensional global ocean model (HYCOM) to reproduce the statistics of the low-frequency eddy available potential energy (EAPE) field in global maps made from Argo floats [*Roullet et al.*, 2014] and in local moored historical observations (MHO). EAPE plays an important role in the vertical structure and mixing in the ocean, as well as in the overall oceanic energy budget. It is therefore essential that high-resolution ocean models, which are increasingly being used for ocean forecasting and dynamical process studies, be evaluated for the accuracy of their EAPE fields. As far as we know, this study is the first to compare, on a global scale, the EAPE fields in high-resolution ocean models with EAPE fields computed from observations. Both the HYCOM simulation and reanalysis predict global area averaged EAPE estimates that are within 10% of Argo global estimates. At MHO locations, the spatially averaged EAPE falls within 16% of MHO estimates, and within 30% of Argo estimates. At the MHO locations, both the highest EAPE estimate (MHO) and lowest estimate (Argo) only differ about 50%, and effectively bound the estimates from the models. Both model EAPE estimates fall within the error of our observations. If account is taken of the fact that Argo estimates include aliased high-frequency motions, then the Argo EAPE values spatially averaged over the MHO locations are lower than the model estimates by about 50%, and in the globally averaged EAPE estimates made from maps, Argo becomes lower than the model estimates by 30%. Point-to-point comparisons of Argo, simulation, reanalysis and MHO EAPE at

MHO locations exhibit considerable scatter. However we show improvement in the local point-to-point correlation of EAPE from the simulation to the reanalysis. As we have shown in an idealized quasi-geostrophic model, and as discussed in *Flierl and McWilliams* [1977], some amount of scatter is to be expected given the chaotic nature of the mesoscale eddies underlying the EAPE fields. While both the HYCOM simulation and reanalysis stratification profiles agree reasonably well with climatological estimates, it is clear that the reanalysis stratification stays closer to climatology, and exhibits less temporal drift than does the simulation stratification. Data assimilation in the reanalysis also improves the spatial structure of the global EAPE with respect to the Argo maps. We show using the skill score that in both our Argo and MHO comparisons, model perfomance (EAPE and T'^2 respectively) is increased through data assimilation. The results presented in this paper show that HYCOM recreates the global low-frequency EAPE field reasonably well. This suggests that it would be reasonable to use HYCOM to quantify global– and basin–scale EAPE reservoirs.

Acknowledgments. We thank Rob Scott and Darran Furnival for access to, and help with, the GMACMD. We thank Rui Xin Huang for advice on estimating ρ' from the historical data. We thank Effie Fine for the entertaining discussions that led to improvements in our estimates of error. We thank two anonymous reviewers of this manuscript, and the two anonymous reviewers of an earlier version, for their helpful comments, which have greatly improved the presentation of our work. CAL, PGT, and BKA acknowledge support from Office of Naval Research (ONR) grants N00014-11-1-0487 and N00014-15-1-2288. SLB, DST, and BKA acknowledge support from the National Science Foundation grants (NSF) OCE-0968783, OCE-1351837, OCE-0960820, and a Research Experience

for Undergraduates (REU) supplement for SLB. JGR, JFS, OMS, and AJW acknowledge support from the Office of Naval Research (ONR) through the "6.2 Eddy Resolving Global Ocean Prediction Including Tides" supported by ONR program element 0602435N. This is contribution NRL/JA/7320-2015-2721 and has been approved for public release. Distribution is unlimited. This work was supported in part by a grant of computer time from the Department of Defense (DoD) High Performance Computing Modernization Program at the Navy DSRC. Daily output files for the two model runs analyzed in this paper are archived at the Department of the Navy Shared Resource Center (DSRC) at the Stennis Space Center. The files stored there can be accessed after obtaining an account at the facility. The corresponding author can be contacted for information to access the archived data once an account has been established. The reanalysis output files are also archived at http://www.hycom.org and are available to any user at that site.

References

Antonov, J. I., D. Seidov, T. P. Boyer, R. A. Locarnini, A. V. Mishonov, H. E. Garcia, O. K. Baranova, M. M. Zweng, and D. R. Johnson (2010), *World Ocean Atlas 2009, Volume 2: Salinity*, NOAA Atlas NESDIS 68, U.S. Government Printing Office, Washington, D.C.

Arbic, B. K., R. B. Scott, G. R. Flierl, A. J. Morten, J. G. Richman, and J. F. Shriver (2012), Nonlinear cascades of surface oceanic geostrophic kinetic energy in the frequency domain, *J. Phys. Oceanogr.*, *42*(9), 1577–1600, doi:10.1175/JPO-D-11-0151.1.

Arbic, B. K., M. Müller, J. G. Richman, J. F. Shriver, A. J. Morten, R. B. Scott, G. Sérazin, and T. Penduff (2014), Geostrophic turbulence in the frequency-wavenumber

domain: Eddy-driven low-frequency variability, *J. Phys. Oceanogr.*, *44*(8), 2050–2069, doi:10.1175/JPO-D-13-054.1.

Chassignet, E. P., and X. Xu (2017), Impact of horizontal resolution (1/12 to 1/50) on gulf stream separation, penetration, and variability, *Journal of Physical Oceanography*, *47*(8), 1999–2021, doi:10.1175/JPO-D-17-0031.1.

Chassignet, E. P., H. E. Hurlburt, E. J. Metzger, O. M. Smedstad, J. A. Cummings, G. R. Halliwell, R. Bleck, R. Baraille, A. J. Wallcraft, C. Lozano, H. L. Tolman, A. Srinivasan, S. Hankin, P. Cornillon, R. Weisberg, A. Barth, R. He, F. Werner, and J. Wilkin (2009), US GODAE: Global ocean prediction with the HYbrid Coordinate Ocean Model (HYCOM)., *Oceanography.*, *22(2)*, 64–75, doi:dx.doi.org/10.5670/oceanog.2009.39.

Chelton, D. B., M. G. Schlax, R. M. Samelson, and R. A. de Szoeke (2007), Global observations of large oceanic eddies, *Geophysical Research Letters*, *34*(15), n/a–n/a, doi:10.1029/2007GL030812, l15606.

Cummings, J. A., and O. M. Smedstad (2013), Variational data assimilation for the global ocean, in *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications (Vol. II)*, edited by S. K. Park and L. Xu, pp. 303–343, Springer Berlin Heidelberg, doi:10.1007/978-3-642-35088-7_13.

Dantzler, H. K. (1977), Potential energy maxima in the tropical and subtropical north atlantic, *Journal of Physical Oceanography*, *7*(4), 512–519, doi:10.1175/1520-0485(1977)007¡0512:PEMITT¿2.0.CO;2.

Doherty, K. W., D. . E. Frye, S. P. Lberatore, and J. M. Toole (1999), A moored profiling instrument, *Journal of Atmospheric and Oceanic Technology*, *16*, 1816–1829.

Flierl, G. R., and J. C. McWilliams (1977), Sampling requirements for measuring moments of eddy variability, *Journal of Marine Research*, *35*, 797–820.

Gnanadesikan, A. (1999), A simple predictive model for the structure of the oceanic pycnocline, *Science*, *283*, 2077–2079.

Hecht, W. M., and H. Hasumi (2008), *Ocean Modeling in an Eddying Regime*, vol. Geophysical Monograph; 177, American Geophysical Union, 2000 Florida Avenue N. W., Washington, D.C.

Huang, R. X. (1998), Mixing and available potential energy in a Boussinesq ocean, *J. Phys. Oceanog.*, *28*, 669–678.

Jacobs, G. A., C. N. Barron, and R. C. Rhodes (2001), Mesoscale characteristics, *Journal of Geophysical Research: Oceans*, *106*(C9), 19,581–19,595, doi:10.1029/2000JC000669.

Kang, D., and O. Fringer (2010), On the calculation of available potential energy in internal wave fields, *J. Phys. Oceanogr.*, *40*, 2539–2545.

Kuragano, T., and M. Kamachi (2000), Global statistical space-time scales of oceanic variability estimated from the topex/poseidon altimeter data, *Journal of Geophysical Research: Oceans*, *105*(C1), 955–974, doi:10.1029/1999JC900247.

Locarnini, R. A., A. V. Mishonov, J. I. Antonov, T. P. Boyer, H. E. Garcia, O. K. Baranova, M. M. Zweng, and D. R. Johnson (2010), *World Ocean Atlas 2009, Volume 1: Temperature*, NOAA Atlas NESDIS 68, U.S. Government Printing Office, Washington, D.C.

Lumpkin, R., and M. Pazos (2007), *Lagrangian Analysis and Prediction of Coastal and Ocean Dynamics:*, Cambridge University Press, Cambridge, doi: 10.1017/CBO9780511535901.

Maltrud, M. E., and J. L. McClean (2005), An eddy resolving global 1/10 ocean simulation, *Ocean Modelling*, *8*(12), 31 – 54, doi: https://doi.org/10.1016/j.ocemod.2003.12.001.

McClean, J. L., D. C. Bader, F. O. Bryan, M. E. Maltrud, J. M. Dennis, A. A. Mirin, P. W. Jones, Y. Y. Kim, D. P. Ivanova, M. Vertenstein, J. S. Boyle, R. L. Jacob, N. Norton, A. Craig, and P. H. Worley (2011), A prototype two-decade fully-coupled fine-resolution ccsm simulation, *Ocean Modelling*, *39*(1), 10 – 30, doi: https://doi.org/10.1016/j.ocemod.2011.02.011, modelling and Understanding the Ocean Mesoscale and Submesoscale.

McDougall, T. J., and P. M. Barker (2011), *Getting Started with TEOS-10 and the Gibbs Seawater (GSW) Oceanographic Toolbox*, SCOR/IAPSO WG127.

Metzger, E. J., O. M. Smedstad, P. G. Thoppil, H. E. Hurlburt, J. A. Cummings, A. J. Wallcraft, L. Zamudio, D. S. Franklin, P. G. Posey, M. W. Phelps, P. J. Hogan, F. L. Bub, and C. J. DeHaan. (2014), US Navy Operational Global Ocean and Arctic Ice Prediction Systems., *Oceanography*, *27*(3), 32–43, doi: http://dx.doi.org/10.5670/oceanog.2014.66.

Munk, W., and C. Wunsch (1998), Abyssal recipes ii: Energetics of tidal and wind mixing, *Deep-Sea Research I*, *45*.

Murphy, A. H. (1988), Skill scores based on the mean square error and their relationships to the correlation coefficient, *Monthly Weather Review*, *116*(12), 2417–2424, doi: 10.1175/1520-0493(1988)116¡2417:SSBOTM¿2.0.CO;2.

Penduff, T., B. Barnier, J.-M. Molines, and G. Madec (2006), On the use of current meter data to assess the realism of ocean model simulations, *Ocean Modelling*, *11*, 399–416.

Richman, J. G., C. Wunsch, and N. G. Hogg (1977), Space and time scales of mesoscale motion in the western north atlantic, *Reviews of Geophysics*, *15*(4), 385–420, doi: 10.1029/RG015i004p00385.

Roemmich, D., and B. W. Owens (2000), The argo project: Global ocean observations for understanding and prediction of climate variability, *Oceanography*, *13*.

Roullet, G., X. Capet, and G. Maze (2014), Global interior eddy available potential energy diagnosed from argo floats, *Geophysical Research Letters*, *41*(5), 1651–1656, doi: 10.1002/2013GL059004.

Rudnick, D. L., and R. Ferrari (1999), Compensation of horizontal temperature and salinity gradients in the ocean mixed layer, *Science*, *283*(5401), 526–529, doi: 10.1126/science.283.5401.526.

Saenz, J. A., A. M. Hogg, G. O. Hughes, and R. W. Griffiths (2012), Mechanical power input from buoyancy and wind to the circulation in an ocean model, *Geophysical Research Letters*, *39*(13), doi:10.1029/2012GL052035, l13605.

Saha, S., S. Moorthi, H.-L. Pan, X. Wu, J. Wang, S. Nadiga, P. Tripp, R. Kistler, J. Woollen, D. Behringer, H. Liu, D. Stokes, R. Grumbine, G. Gayno, J. Wang, Y.-T. Hou, H.-Y. Chuang, H.-M. H. Juang, J. Sela, M. Iredell, R. Treadon, D. Kleist, P. Van Delst, D. Keyser, J. Derber, M. Ek, J. Meng, H. Wei, R. Yang, S. Lord, H. Van Den Dool, A. Kumar, W. Wang, C. Long, M. Chelliah, Y. Xue, B. Huang, J.-K. Schemm, W. Ebisuzaki, R. Lin, P. Xie, M. Chen, S. Zhou, W. Higgins, C.-Z. Zou, Q. Liu, Y. Chen, Y. Han, L. Cucurull, R. W. Reynolds, G. Rutledge, and M. Goldberg (2010), The NCEP climate forecast system reanalysis, *Bull. Amer. Meteor. Soc.*, *91*(8), 1015–1057, doi:10.1175/2010BAMS3001.1.

Saha, S., S. Moorthi, X. Wu, J. Wang, S. Nadiga, P. Tripp, D. Behringer, Y.-T. Hou, H.-y. Chuang, M. Iredell, M. Ek, J. Meng, R. Yang, M. P. Mendez, H. van den Dool, Q. Zhang, W. Wang, M. Chen, and E. Becker (2013), The ncep climate forecast system version 2, *J. Climate*, *27*(6), 2185–2208, doi:10.1175/JCLI-D-12-00823.1.

Schmitz, W. J. (1988), Exploration of the eddy field in the midlatitude North Pacific, *J. Phys. Oceanogr.*, *18*(3), 459–468, doi:10.1175/1520-0485.

Scott, R. B., B. K. Arbic, E. P. Chassignet, A. C. Coward, M. Maltrud, W. J. Merryfield, A. Srinivasan, and A. Varghese (2010), Total kinetic energy in four global eddying ocean circulation models and over 5000 current meter records, *Ocean Modelling*, *32*, 157–169.

Tailleux, R. (2013), Available potential energy and exergy in stratified fluids, *Annual Review of Fluid Mechanics*, *45*(1), 35–58, doi:10.1146/annurev-fluid-011212-140620.

Thoppil, P. G., J. G. Richman, and P. J. Hogan (2011), Energetics of a global ocean circulation model compared to observations, *Geophysical Research Letters*, *38*(15), doi:10.1029/2011GL048347.

Timko, P. G., B. K. Arbic, J. G. Richman, R. B. Scott, E. J. Metzger, and A. J. Wallcraft (2012), Skill tests of three-dimensional tidal currents in a global ocean model: A look at the North Atlantic, *J. Geophys. Res: Oceans*, *117*(C8), doi:10.1029/2011JC007617.

Timko, P. G., B. K. Arbic, J. G. Richman, R. B. Scott, E. J. Metzger, and A. J. Wallcraft (2013), Skill testing a three-dimensional global tide model to historical current meter records, *J. Geophys. Res: Oceans*, *118*(12), 6914–6933, doi:10.1002/2013JC009071.

Winters, K. B., P. N. Lombard, J. J. Riley, and E. A. D'Asaro (1995), Available potential energy and mixing in density-stratified fluids, *J. Fluid Mech.*, *94*, 3187–3200.

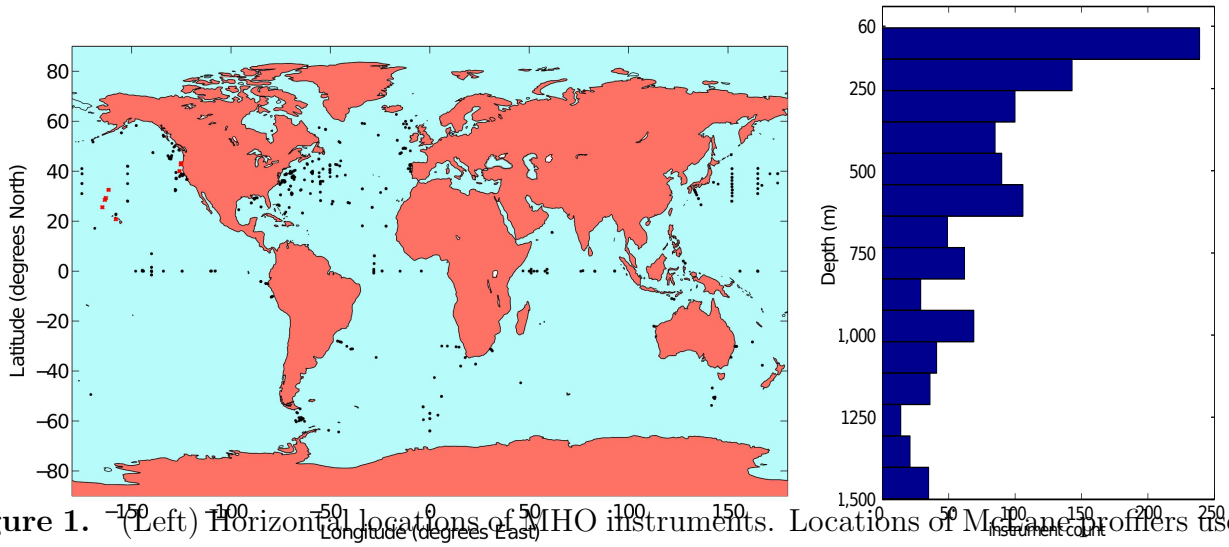


Figure 1. (Left) Horizontal locations of MHO instruments. Locations of Methane profilers used to test our use of temperature as a proxy for density are shown with a red X. (Right) Distribution of MHO instruments by depth, binned into 15 equally distributed depths from 60 to 1500m.

Wunsch, C. (1999), A summary of north atlantic baroclinic variability, *Journal of Physical Oceanography*, *29*(12), 3161–3166, doi:10.1175/1520-0485(1999)029¡3161:ASONAB¿2.0.CO;2.

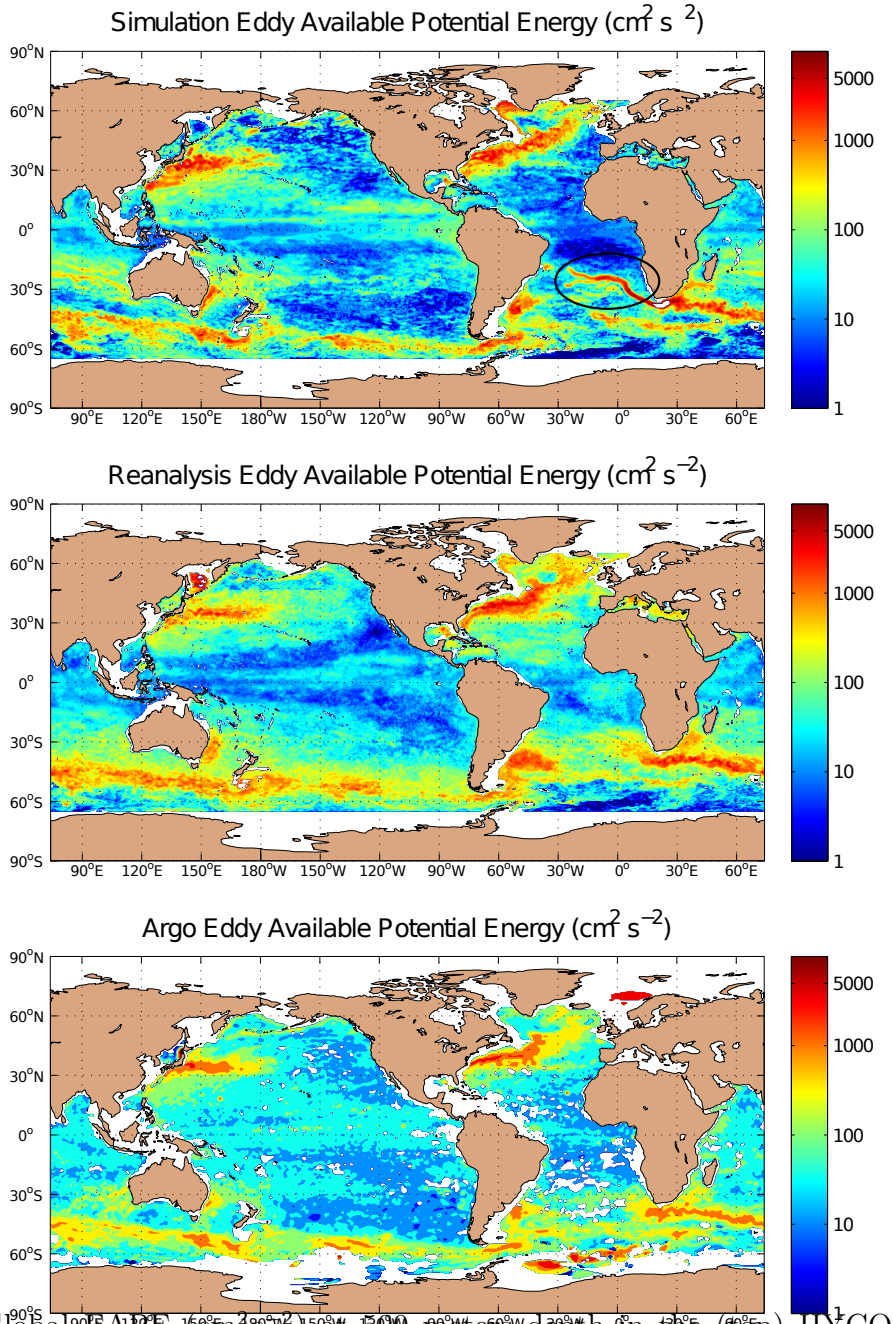


Figure 2. Global EAPE (cm^2 s^{-2}) at 300 meters depth in the (top) HYCOM simulation, (middle) HYCOM reanalysis, and (bottom) *Roullet et al.* [2014] Argo analysis. The eddy "train" discussed in the results section is encircled in the top sub-figure. In this figure, both HYCOM maps are given on a $\frac{1}{4}^\circ$ grid while the Argo map is shown on its native $\frac{1}{2}^\circ$ resolution.

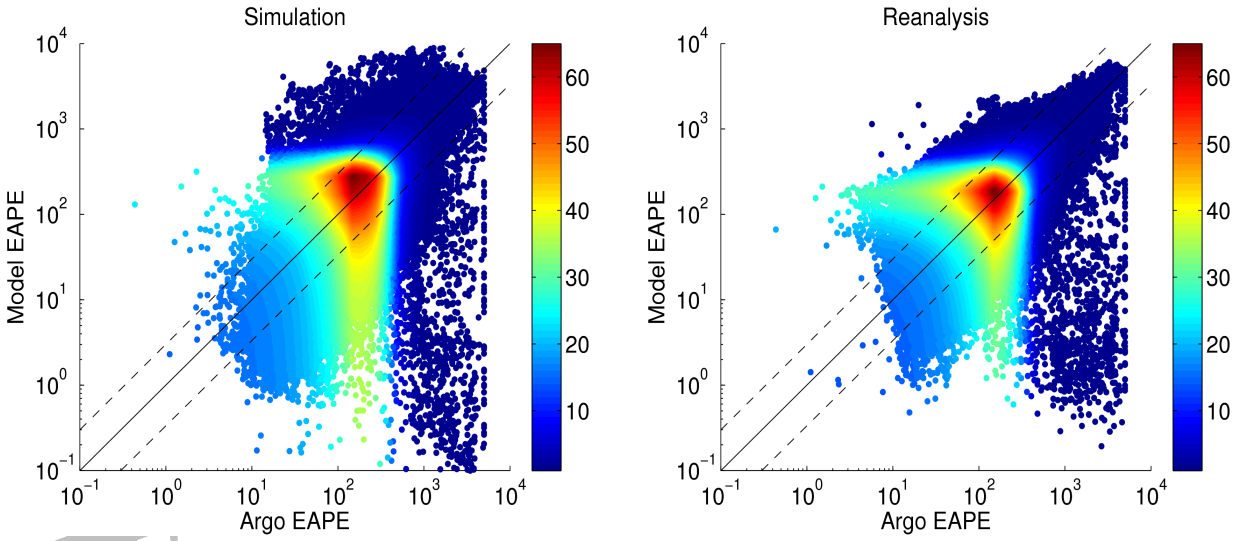


Figure 3. A point-to-point comparison of global EAPE ($cm^2 s^{-2}$) at 500m (a) between simulation and Argo, and (b) between reanalysis and Argo. Population density is given by color, with the most tightly grouped data shown in red, and the sparsest data in blue. The one-to-one line is shown in black along with bounding lines representing a factor of 3.

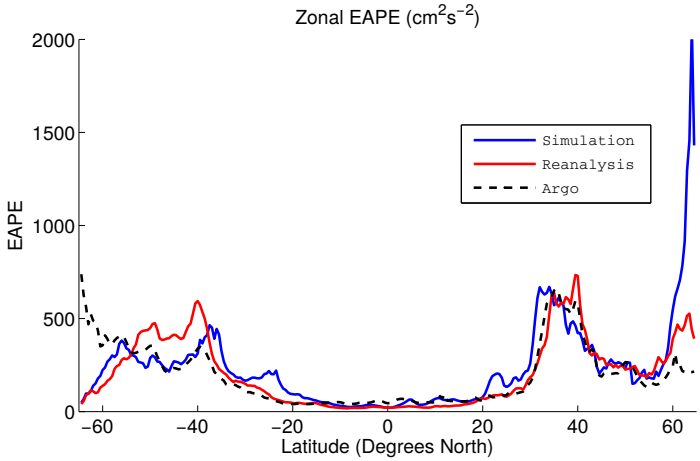


Figure 4. Zonal-mean distribution of EAPE ($cm^2 s^{-2}$) for simulation, reanalysis, and Argo.

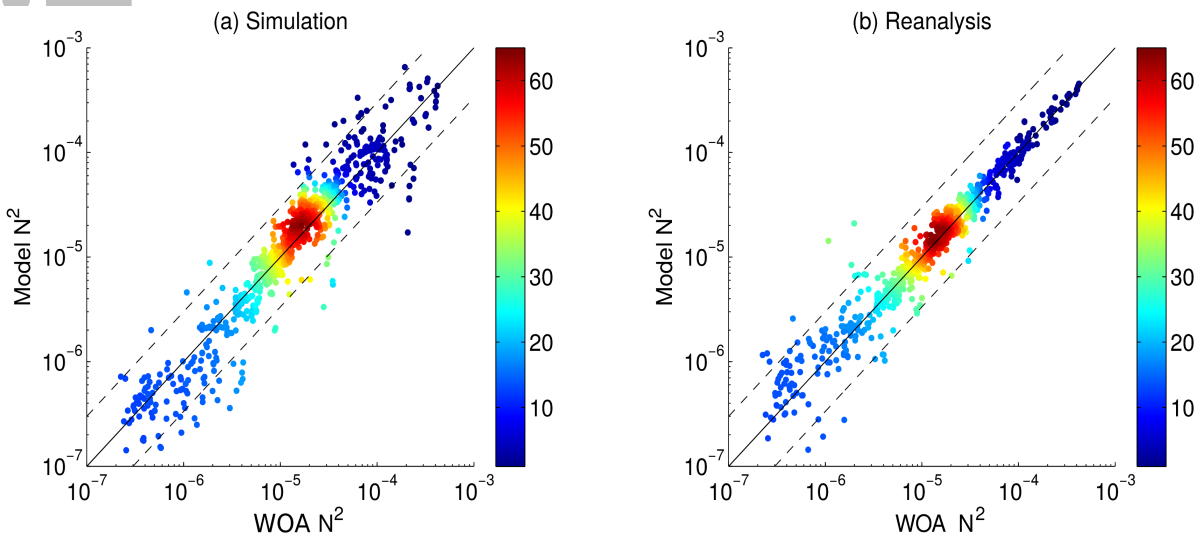


Figure 5. Global point-to-point comparison of squared Brunt-Väisälä buoyancy frequency (s^{-2}) at MHO locations, averaged over 20 years of model output (a) between simulation and WOA, and (b) between reanalysis and WOA. Population density is given by color, with the most tightly grouped data shown in red, and the sparsest data in blue. The one-to-one line is shown in black along with bounding lines representing a factor of 3 as described in Figure 3.

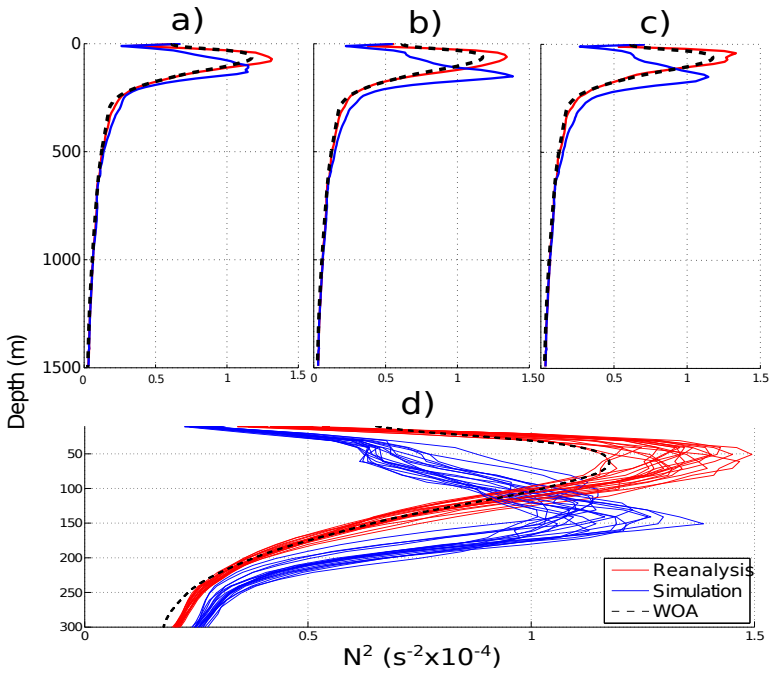


Figure 6. Vertical profile of spatially averaged buoyancy frequency N^2 (s^{-2}) in HYCOM and WOA taken over comparison points shown in Figure 1. Profiles are shown of yearly averaged snapshots of model output for a) 1993, b) 2002, c) 2012, and, d) profiles for all years of model output over the upper 300m of water column. In all subplots, we display the WOA annual profiles averaged over the same locations.

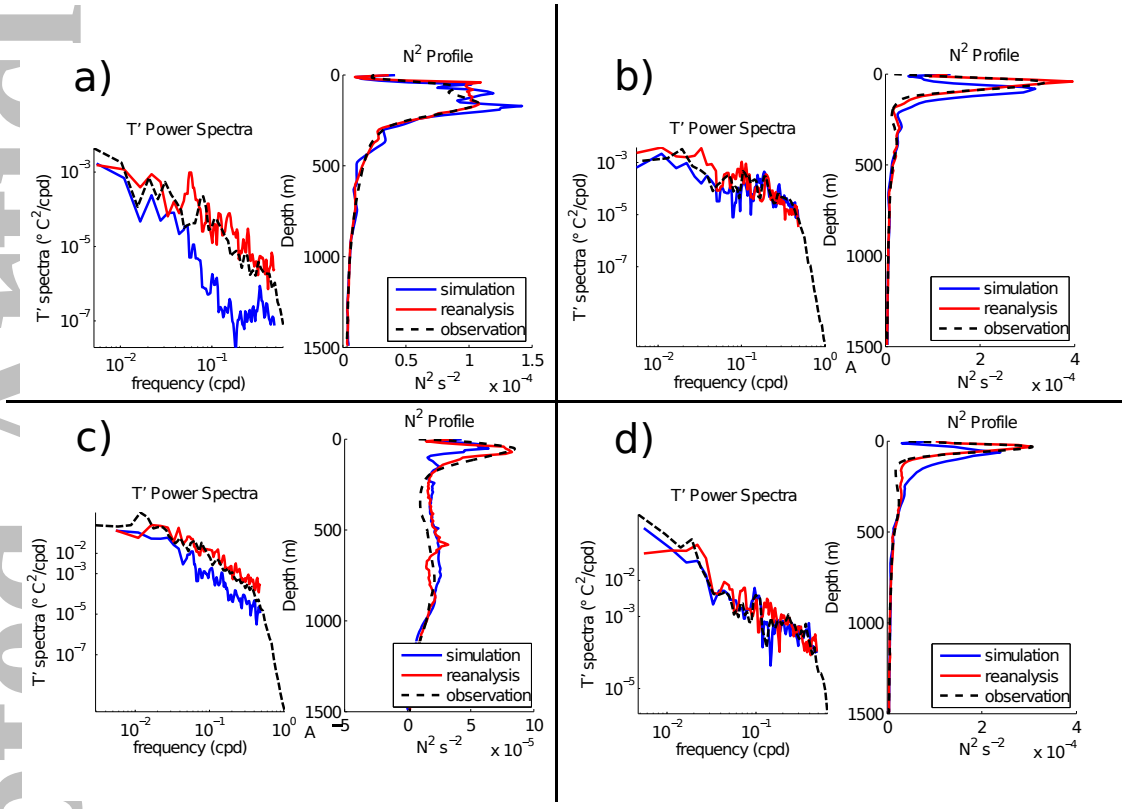


Figure 7. Temperature anomaly frequency spectra (at depths given below) and annual average N^2 profiles, for 1993 HYCOM simulation and reanalysis output, showing different levels of agreement with spectra computed from MHO observations (left plots) and N^2 computed from WOA observational climatology (right plots). Locations are: a) 32.44°N, 127.769°W (454m), in the North-Eastern Pacific, b) 0.02°S, 110.21°W (927m) in the Eastern Equatorial Pacific, c) 37.8°N, 55.7°W (497m) in the Gulf Stream, and d) 5.96°S, 82.50°W (100m) off the coast of Peru.

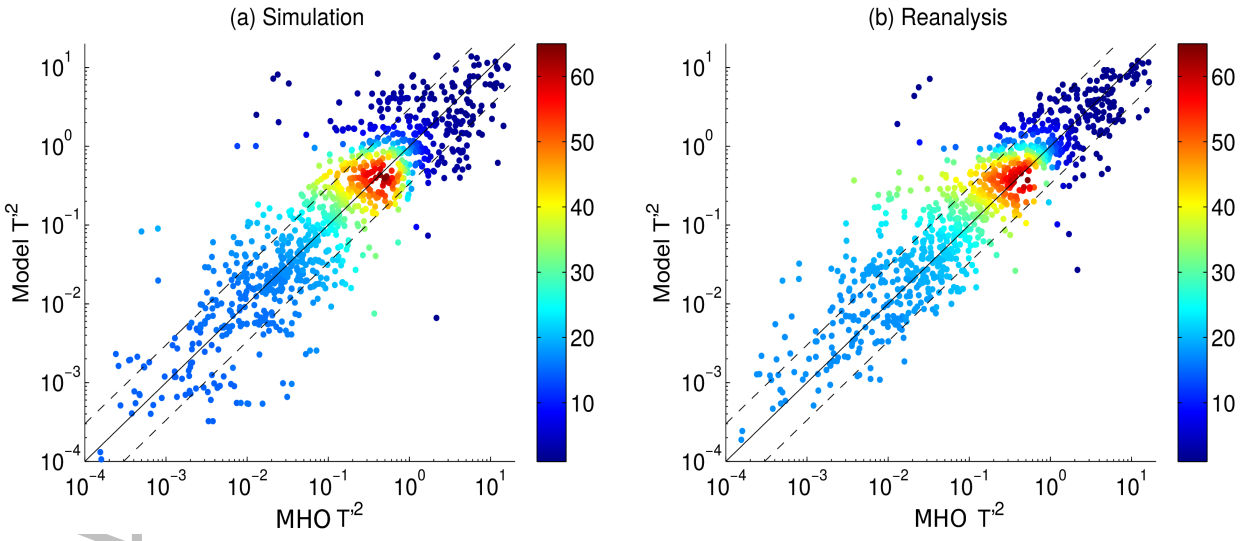


Figure 8. As in Figure 3, but for a point-to-point comparison of 20 year averaged T'^2 ($^\circ C^2$) (a) between simulation and MHO, and (b) between reanalysis and MHO.

Table 1. Means of EAPE, computed over the entire globe (left column), and over all available MHO locations (right column), for MHO, Argo, HYCOM simulation and reanalysis. For the global calculations, we use model year 2003 for the HYCOM simulation and reanalysis, and for the MHO comparison, HYCOM 20-year means are used.

Mean EAPE ($cm^2 s^{-2}$)		
	global	at MHO locations
Simulation	185 ±6	600 ±90
Reanalysis	183 ±4	598 ±87
MHO	N/A	709 ±143
Argo	168 ±4	462 ±55

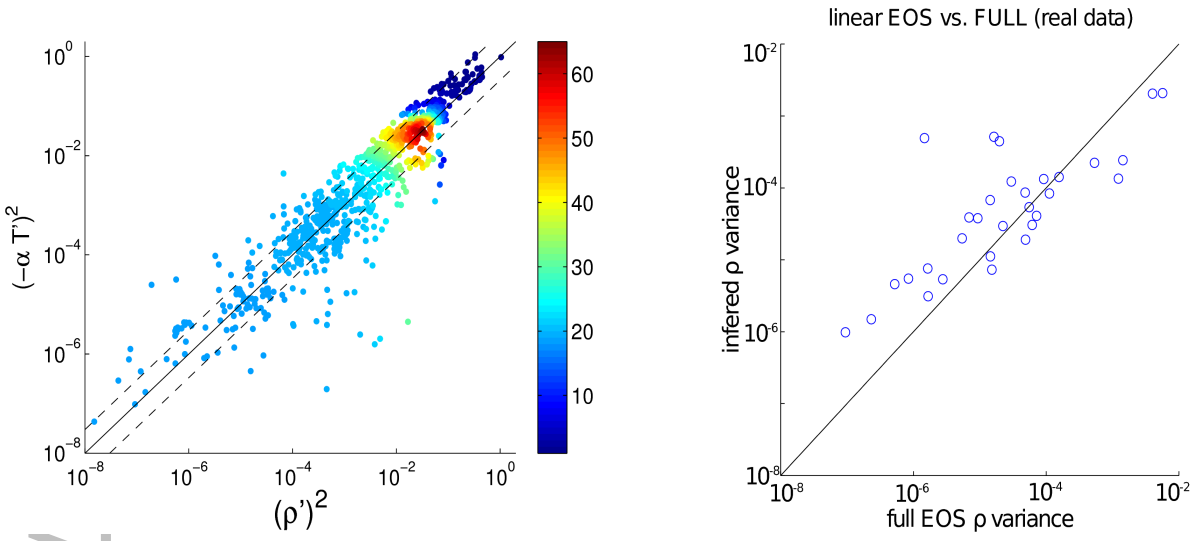


Figure 9. (Left) As in Figure 3 but for a comparison of the squared inferred density anomaly $(-\rho_0\alpha T')^2$ against the actual squared density anomaly ρ'^2 calculated using the full non-linear equation of state [*McDougall and Barker*, 2011], with HYCOM reanalysis salinity and temperature fields as inputs at MHO locations. (Right) A similar comparison performed at 31 depths over 10 locations using in-situ McLane profiler temperature and salinity data. Units are $(\frac{kg}{m^3})^2$.

Table 2. Statistical metrics and 20 year means for EAPE constituent terms, the square of the temperature anomaly and the square of the buoyancy frequency at MHO locations, denoted by terms in$\langle\rangle$.

	Temperature Variance			
	$\langle T'^2\rangle_{MHO}{}^\circ C^2$	$A_{T'^2}$	$\gamma_{T'^2}$	$R_{T'^2}$
Data	0.95 (MHO)			
Simulation	1.06	0.68	1.10	0.54
Reanalysis	1.08	0.83	1.12	0.76

Stratification			
$\langle N^2\rangle_{WOA}10^{-4}s^{-2}$	A_{N^2}	γ_{N^2}	R_{N^2}
0.50 (WOA)			
0.56	1.00	1.15	0.78
0.50	1.00	1.02	0.97

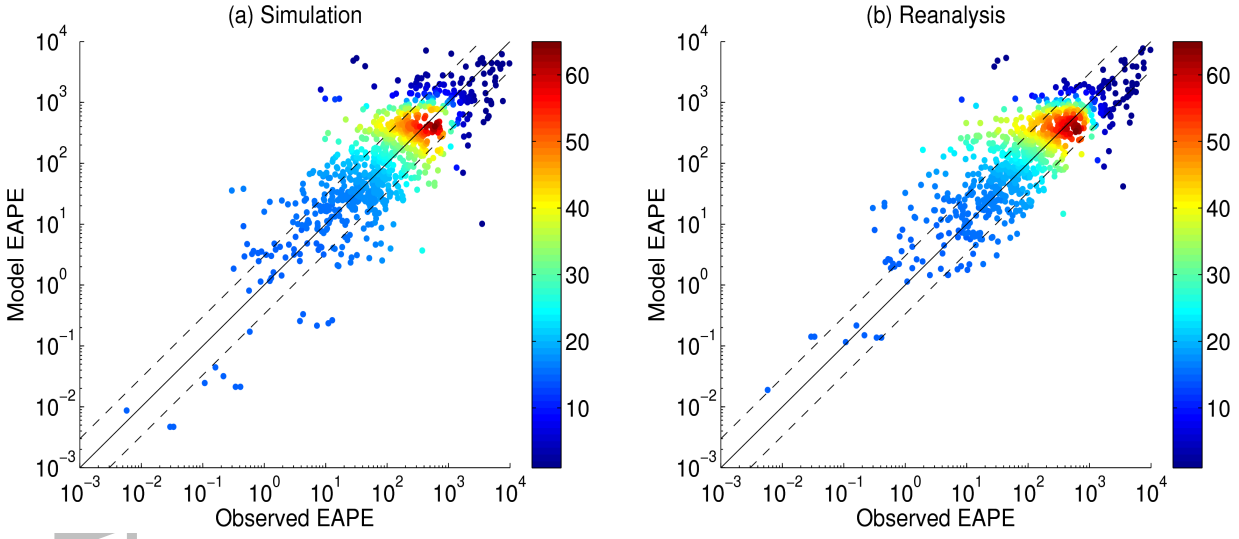


Figure 10. As in Figure 3, but for a point-to-point comparison of EAPE ($cm^2 s^{-2}$) (a) between a 20 year average of the simulation and MHO, and (b) between a 20 year average of the reanalysis and MHO.

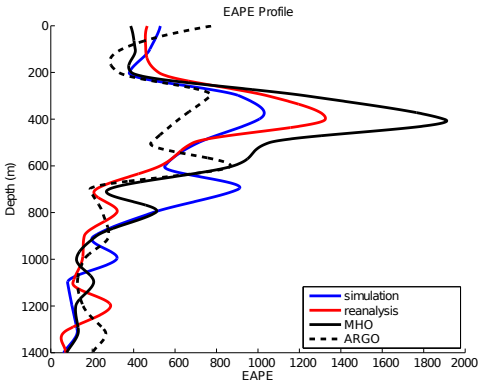


Figure 11. Vertical distribution of EAPE ($cm^2 s^{-2}$) for simulation, reanalysis, MHO, and Argo over MHO locations. EAPE is binned into 15 evenly distributed depth bins in the vertical, and then splined for continuity.

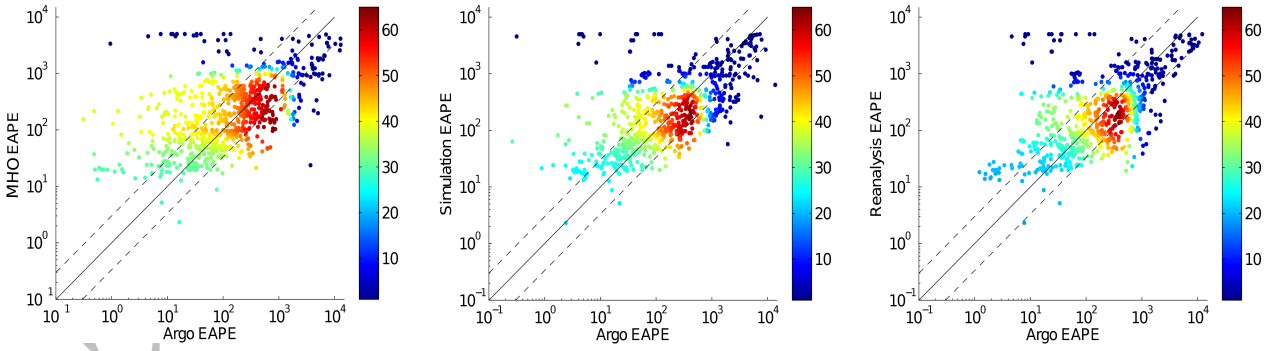


Figure 12. As in Figure 3, but for a point-to-point comparison EAPE ($log_{10}(\mathrm{cm}^2 s^{-2})$) in (left) reanalysis, (middle) simulation, and (right) MHO, versus Argo EAPE [*Roullet et al.*, 2014] at MHO locations.

Table 3. Statistical comparison metrics (see text for definitions) for EAPE between HYCOM (computed using 20 years of output) and observations (MHO and Argo).

Model Run	Comparison Metric		
	A	γ	R
Simulation EAPE (MHO)	0.40	0.84	0.56
Reanalysis EAPE (MHO)	0.55	0.84	0.84
Simulation EAPE (Argo)	0.91	1.30	0.47
Reanalysis EAPE (Argo)	1.07	1.30	0.62

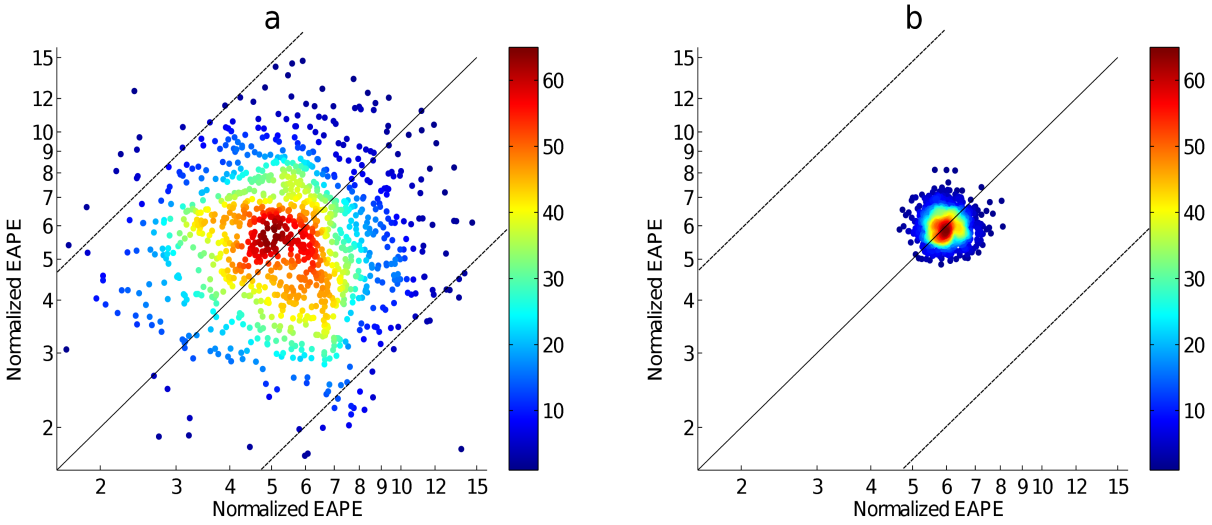
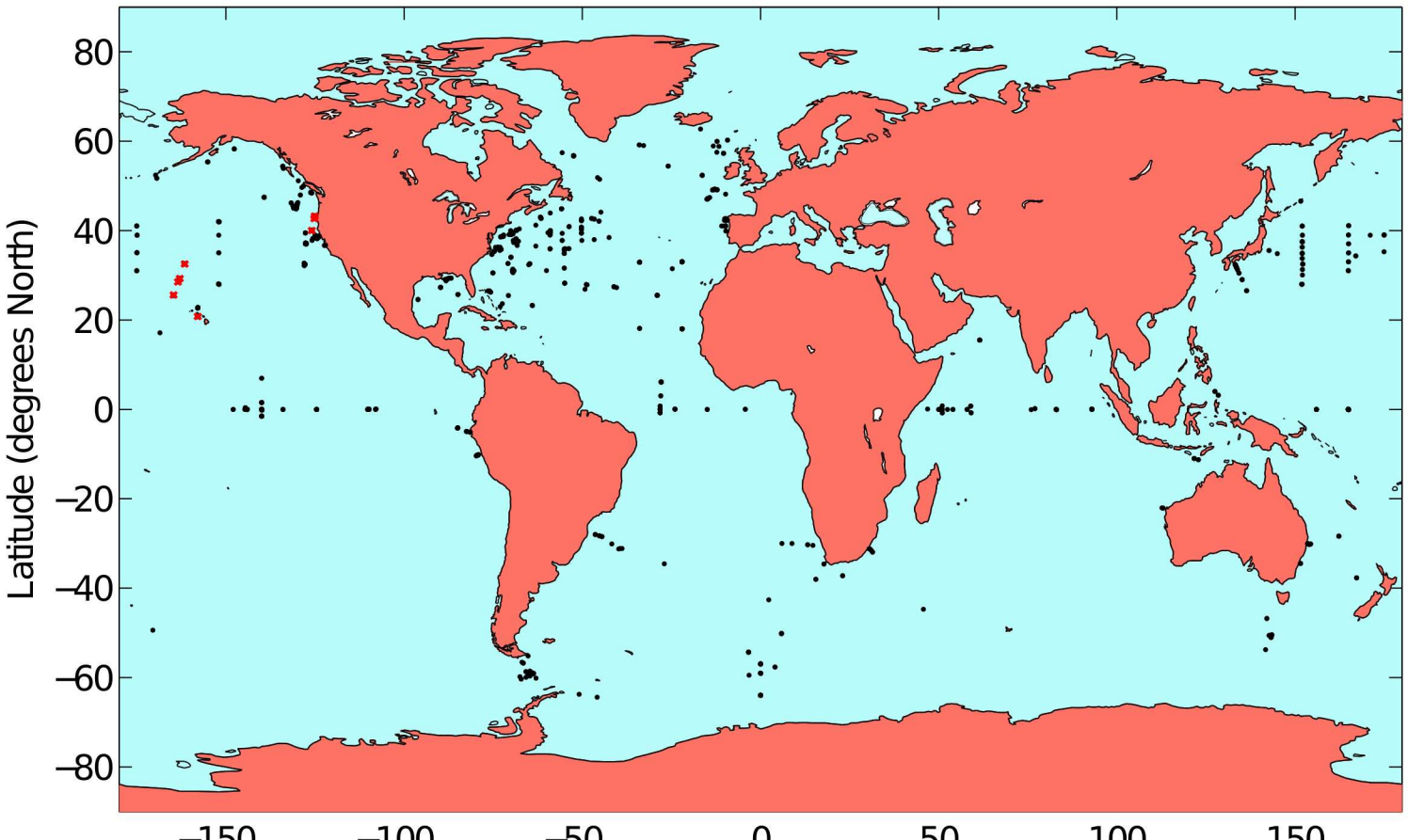


Figure 13. Scatterplots of EAPE taken from 1100 random points in an idealized horizontally homogeneous QG turbulence model verses EAPE taken from 1100 different random points within the same QG model, used to illustrate of the effect of record duration on scatter. The QG EAPE [normalized by $\frac{1}{2}(\overline{u_1} - \overline{u_2})^2$] is averaged over 8 model decorrelation time periods in (a), and over 182 model decorrelation time periods in (b) (approximately equivalent to 320 days and 20 years for mid-latitude oceanic eddies respectively.) As in previous plots, bounding lines represent a factor of 3, and colorbars represent population density.

Figure 1.



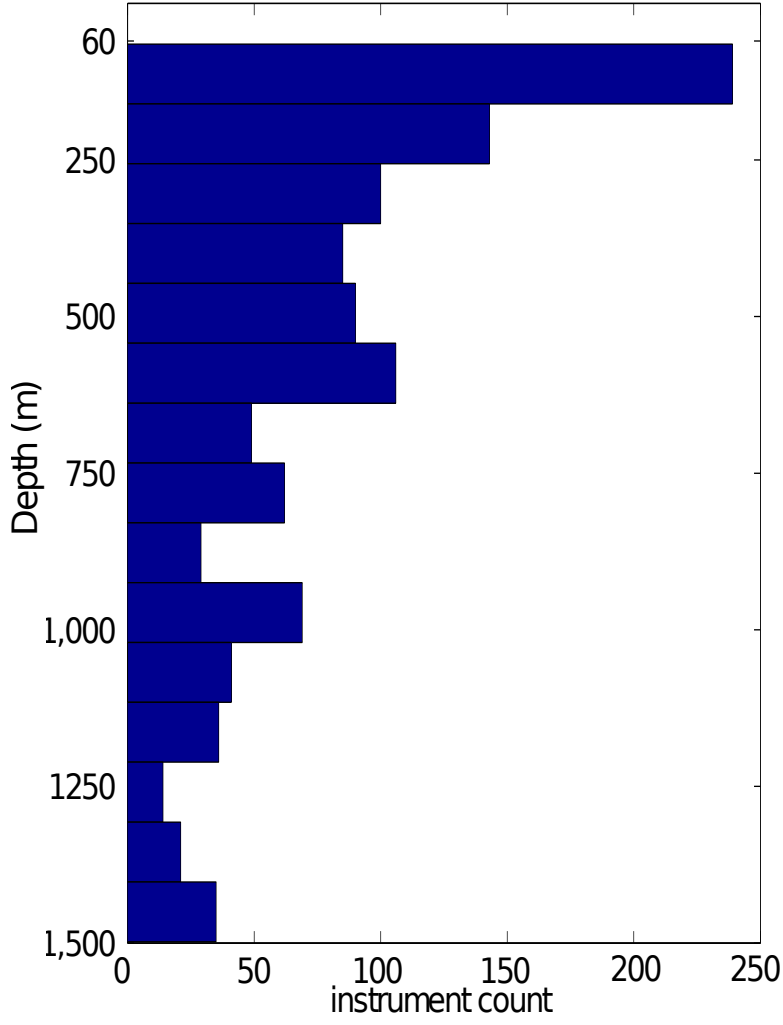


Figure 2.

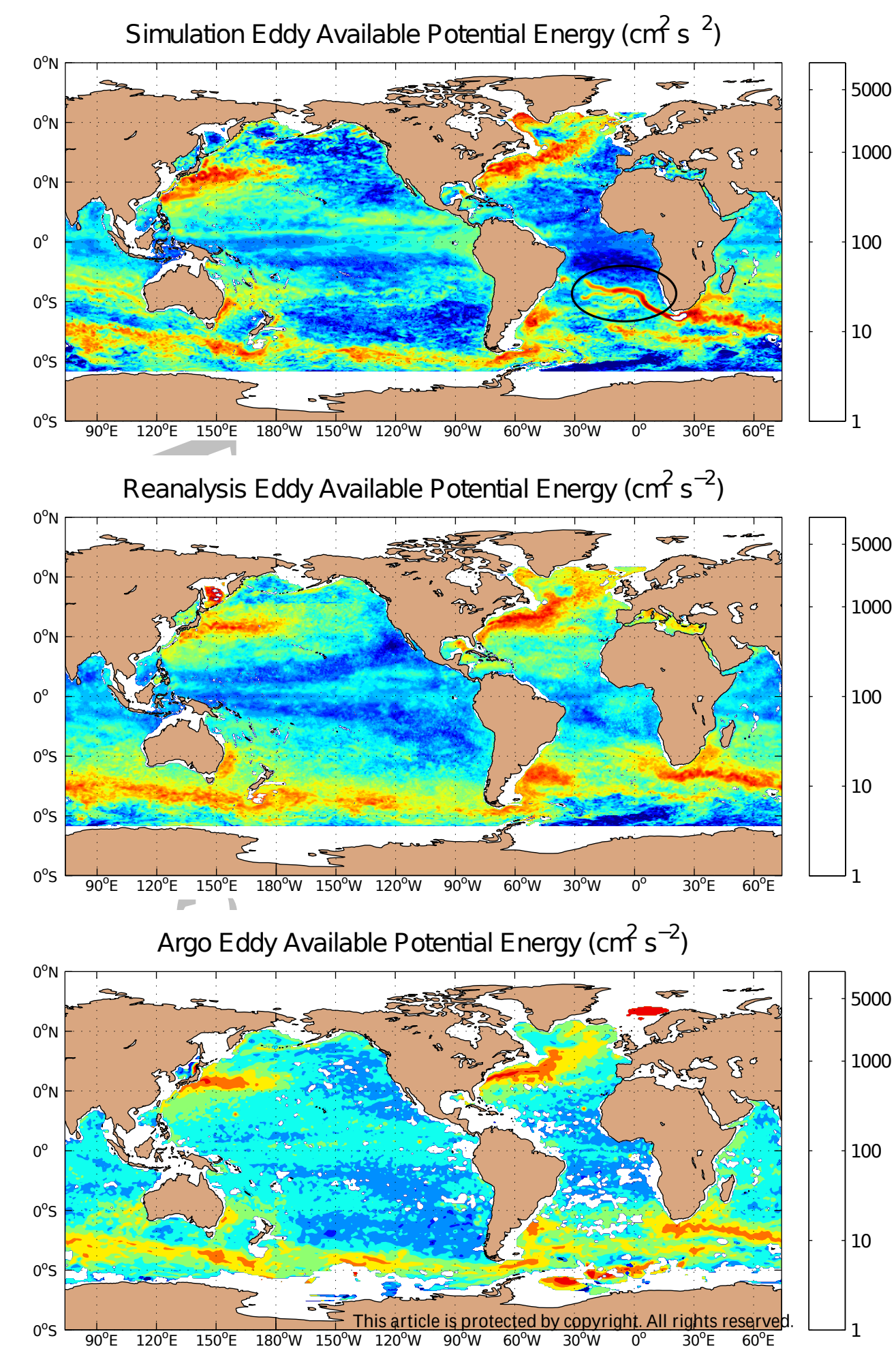


Figure 3.

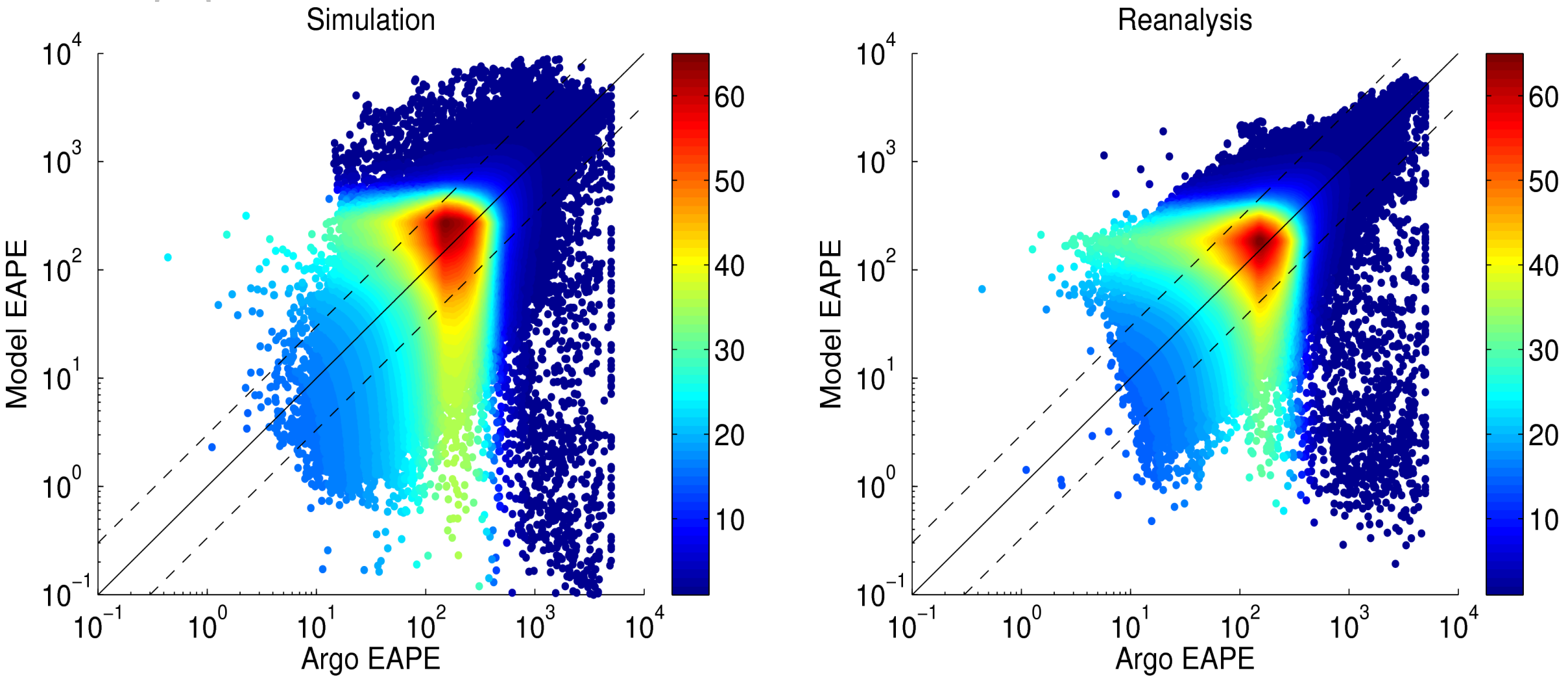


Figure 4.

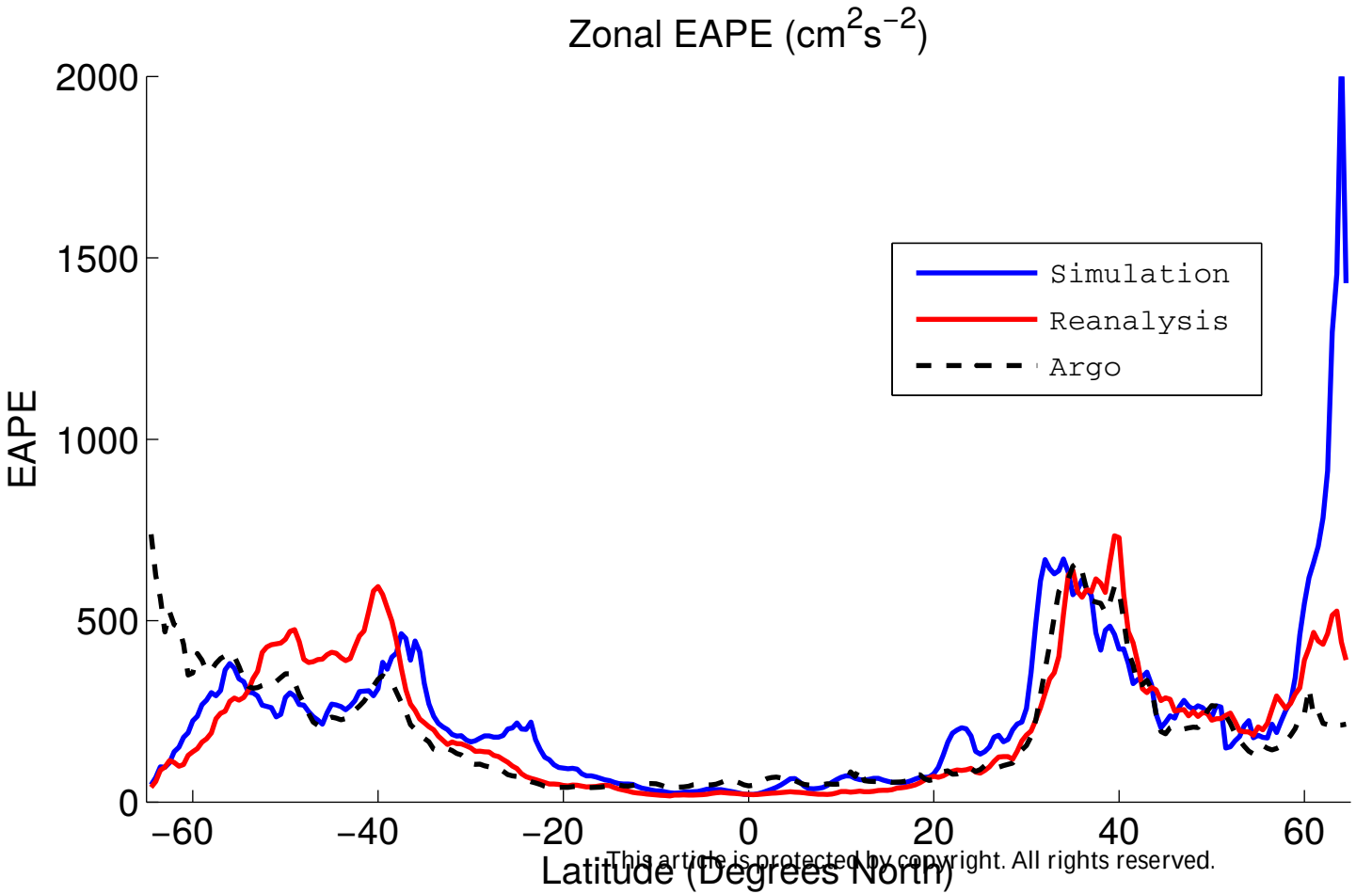


Figure 5.

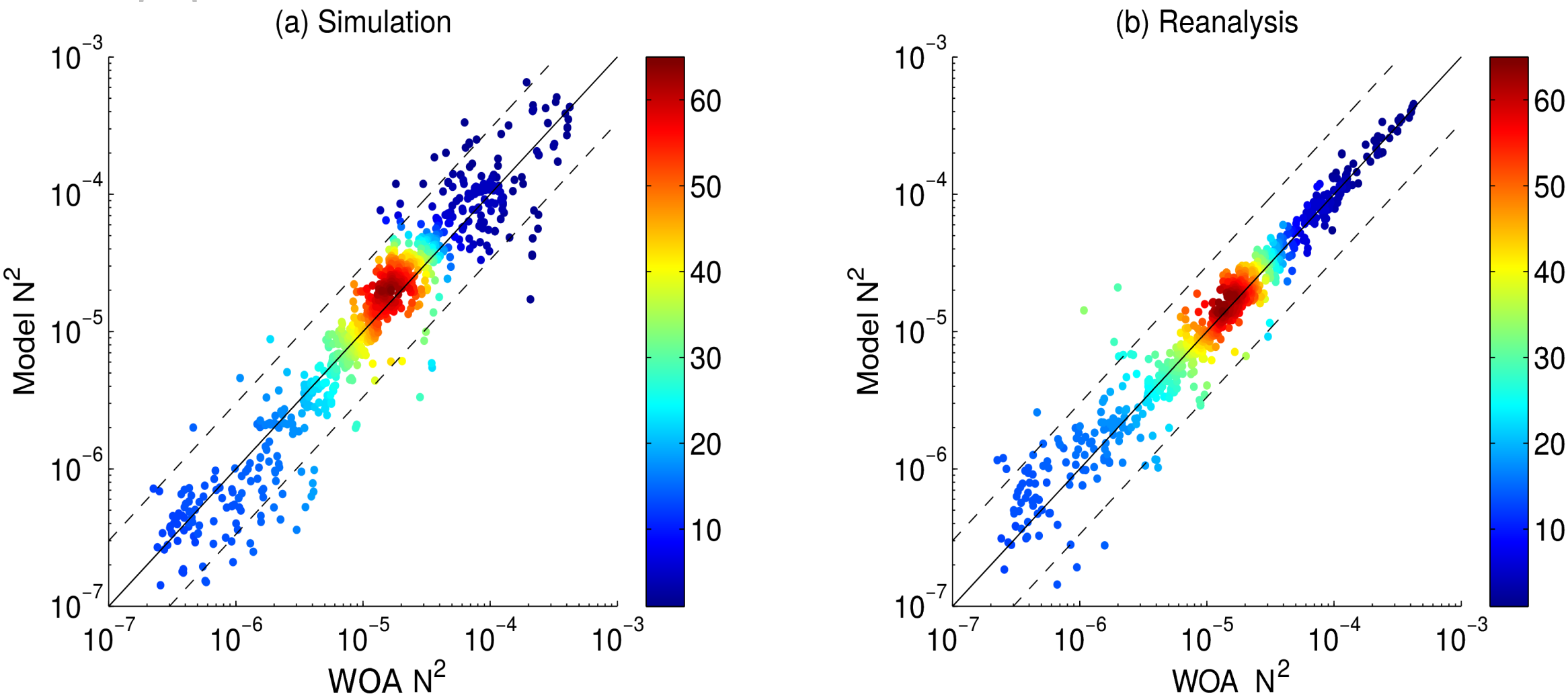


Figure 6.

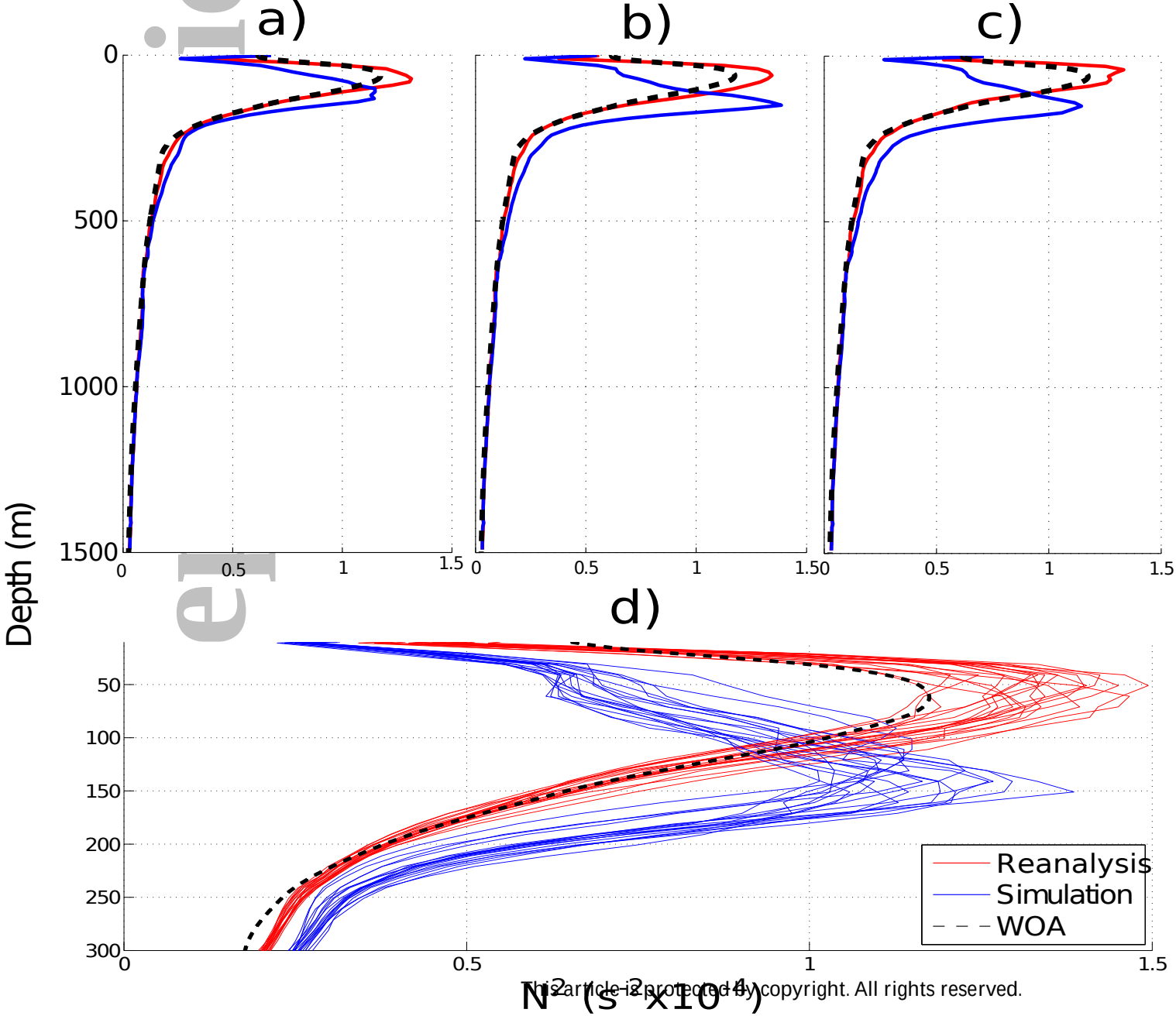


Figure 7.

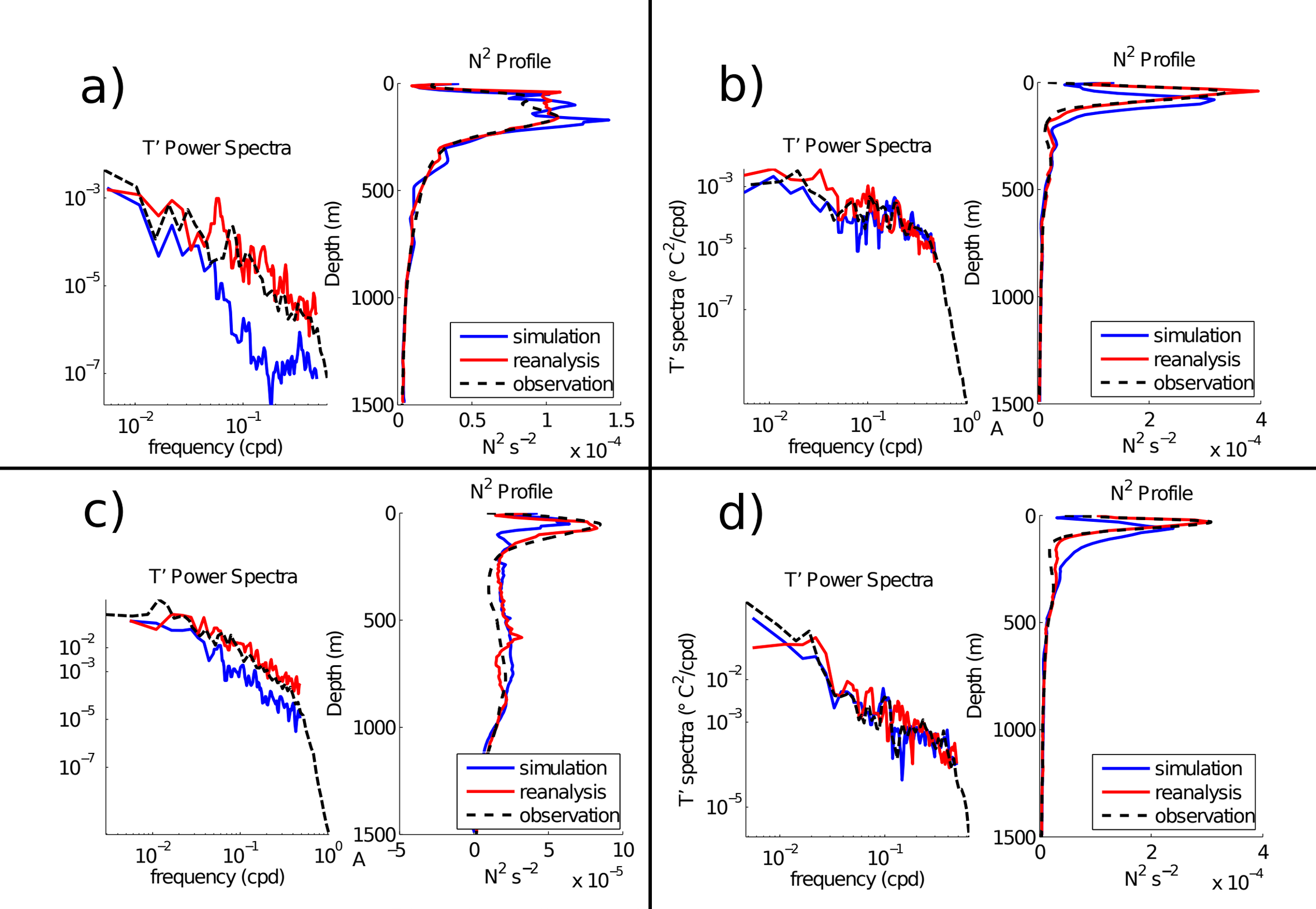


Figure 8.

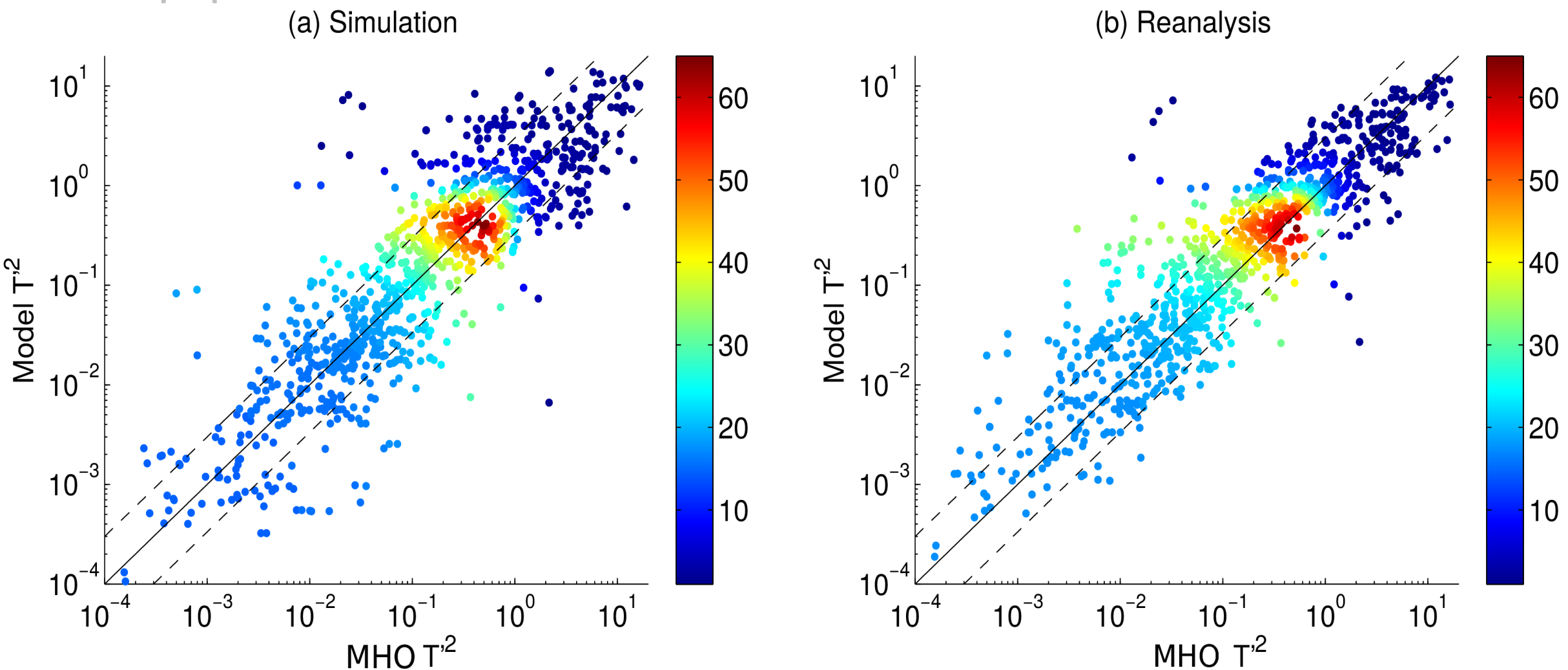
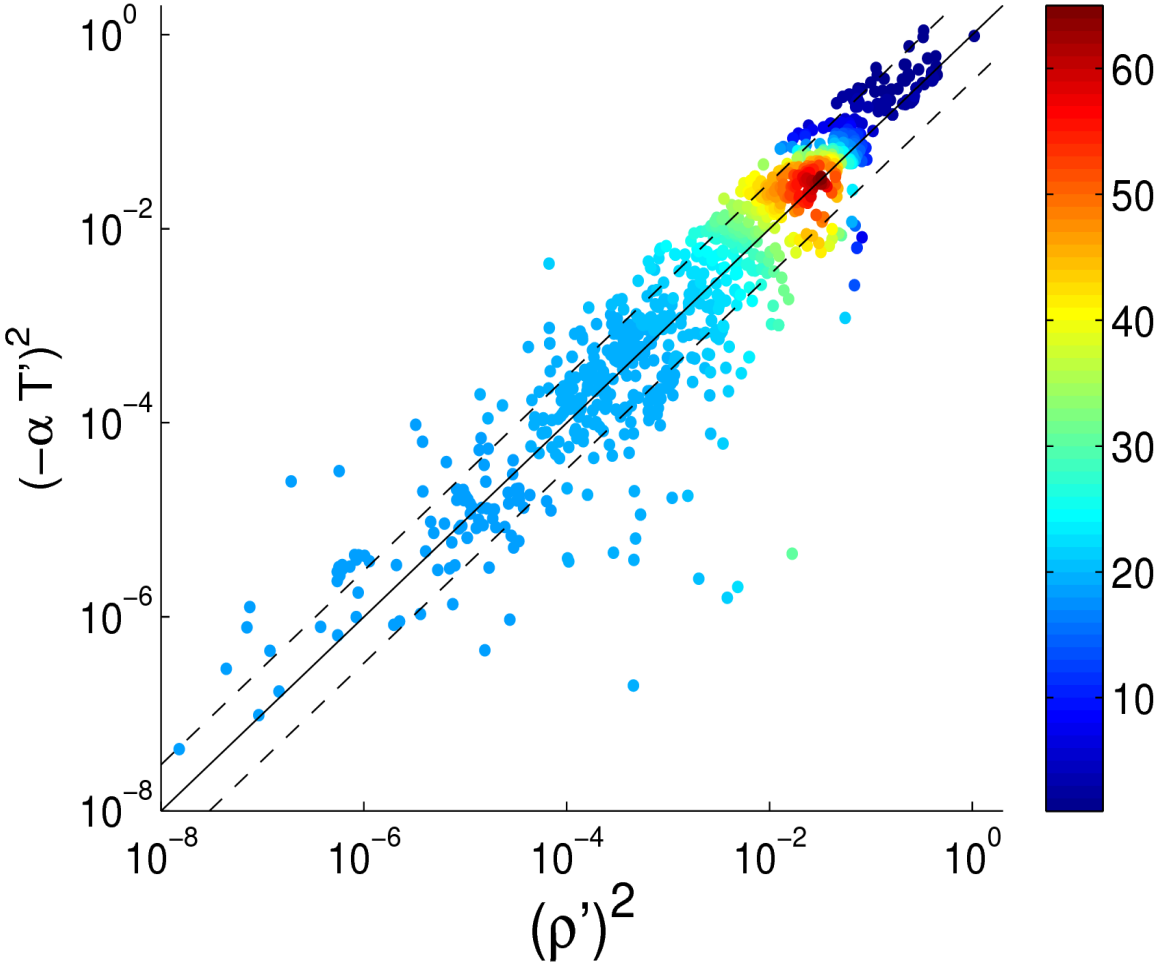


Figure 9.



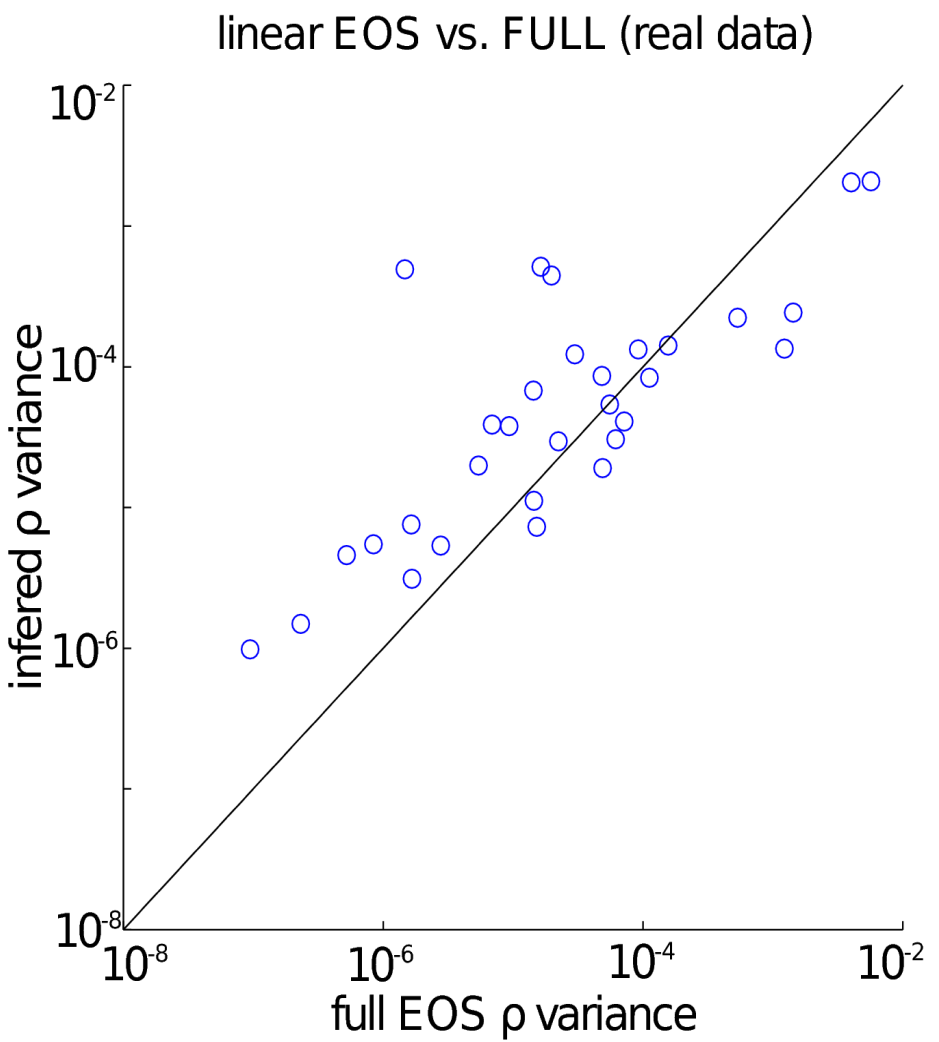


Figure 10.

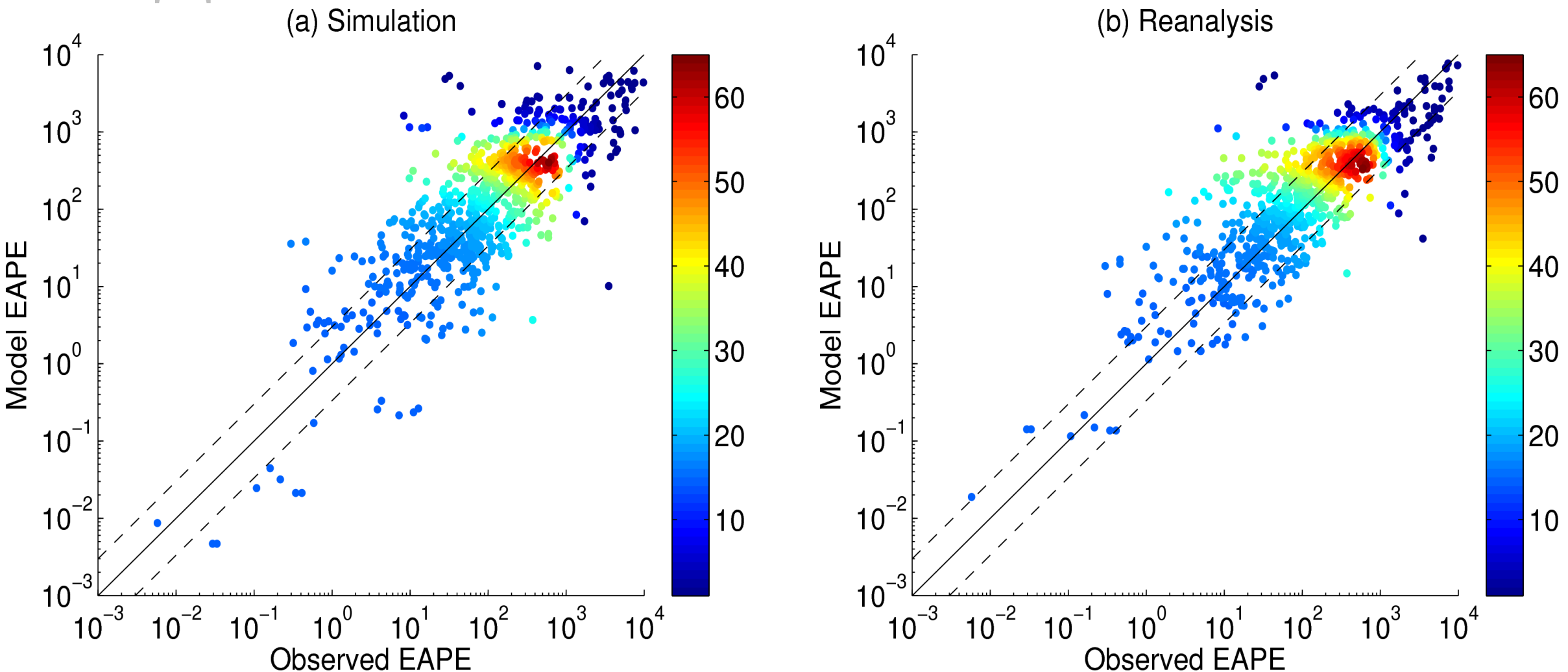


Figure 11.

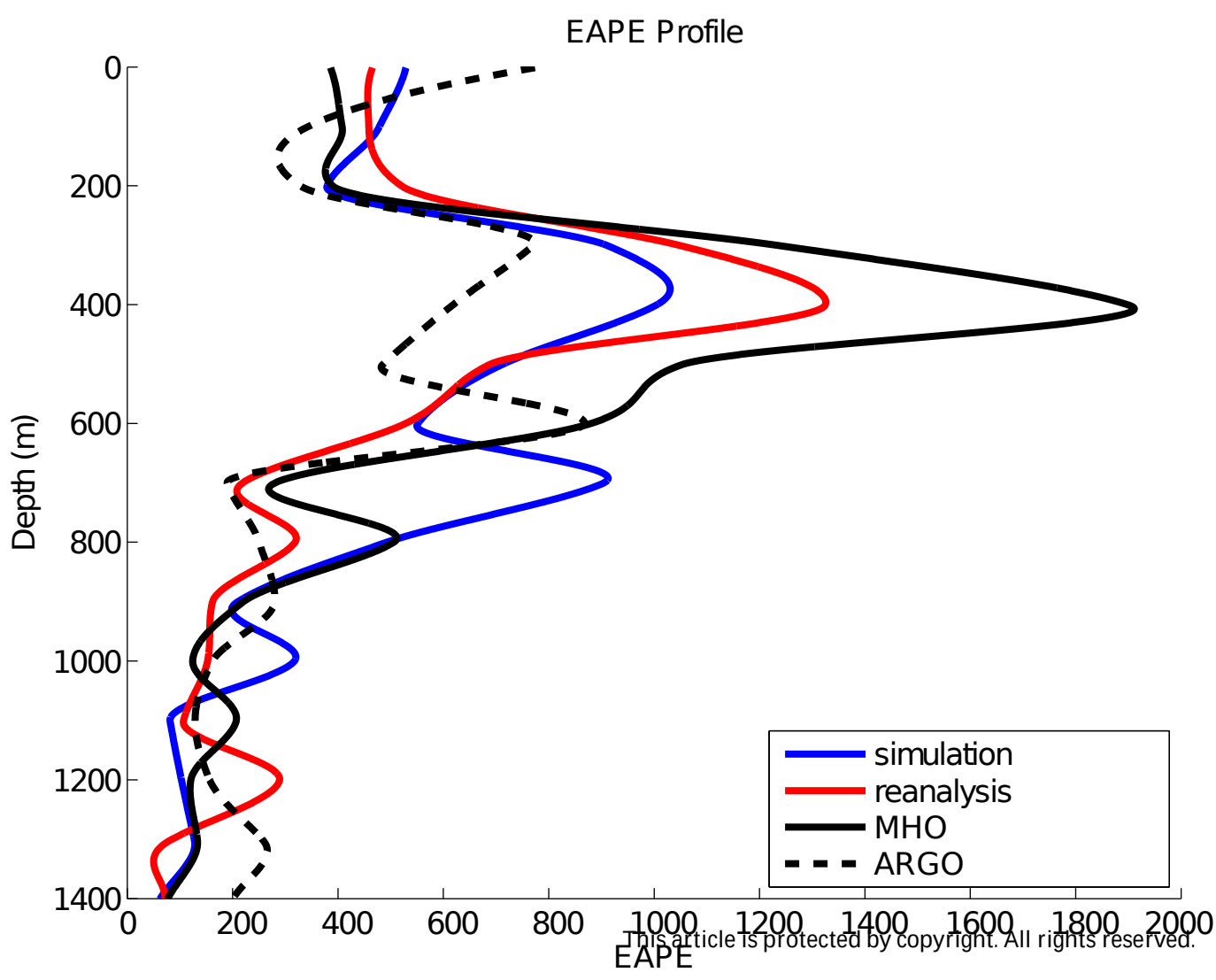
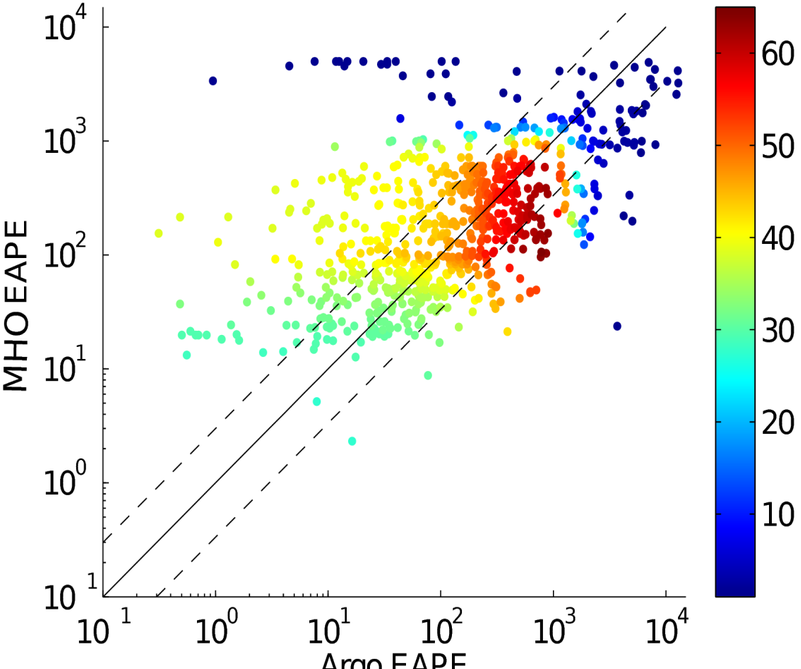
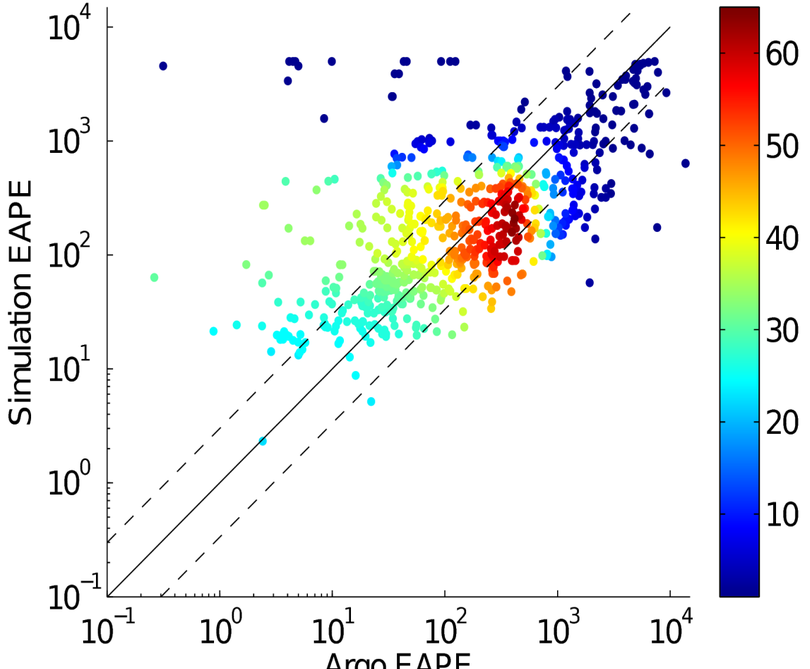


Figure 12.





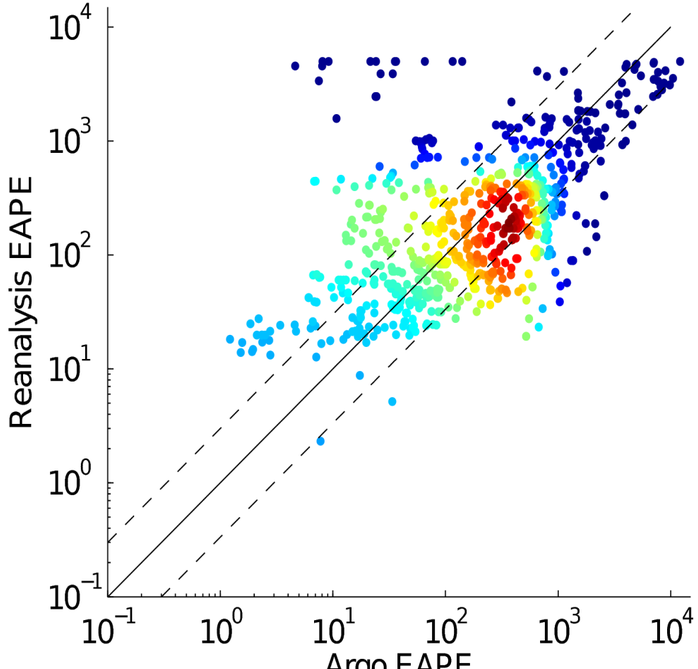


Figure 13.