A Prototype for Remote Monitoring of Ocean Heat Content Anomalies

DAVID S. TROSSMAN[a,b,c] AND ROBERT H. TYLER[d,e]

[a] *Department of Oceanography and Coastal Sciences, Louisiana State University, Baton Rouge, Louisiana*
[b] *Center for Computation and Technology, Louisiana State University, Baton Rouge, Louisiana*
[c] *Global Science and Technology, NOAA/NESDIS/STAR, College Park, Maryland*
[d] *Geodesy and Geophysics Laboratory, NASA Goddard Space Flight Center, Greenbelt, Maryland*
[e] *Joint Center for Earth Systems Technology, University of Maryland, Baltimore County, Baltimore, Maryland*

(Manuscript received 2 April 2021, in final form 4 November 2021)

ABSTRACT: To overcome challenges with observing ocean heat content (OHC) over the entire ocean, we propose a novel approach that exploits the abundance of satellite data, including data from modern satellite geomagnetic surveys such as Swarm. The method considers a novel combination of conventional in situ (temperature and pressure) as well as satellite (altimetry and gravimetry) data with estimates of ocean electrical conductance (depth-integrated conductivity), which can potentially be obtained from magnetic observations (by satellite, land, seafloor, ocean, and airborne magnetometers). To demonstrate the potential benefit of the proposed method, we sample model output of an ocean state estimate to reflect existing observations and train a machine learning algorithm [Generalized Additive Model (GAM)] on these samples. We then calculate OHC everywhere using information potentially derivable from various global satellite coverage—including magnetic observations—to gauge the GAM's goodness of fit on a global scale. Inclusion of in situ observations of OHC in the upper 2000 m from Argo-like floats and conductance data each reduce the root-mean-square error by an order of magnitude. Retraining the GAM with recent ship-based hydrographic data attains a smaller RMSE in polar oceans than training the GAM only once on all available historical ship-based hydrographic data; the opposite is true elsewhere. The GAM more accurately calculates OHC anomalies throughout the water column than below 2000 m and can detect global OHC anomalies over multiyear time scales, even when considering hypothetical measurement errors. Our method could complement existing methods and its accuracy could be improved through careful ship-based campaign planning.

SIGNIFICANCE STATEMENT: The purpose of this manuscript is to demonstrate the potential for practical implementation of a remote monitoring method for ocean heat content (OHC) anomalies. To do this, we sample data from a reanalysis product primarily because of the dearth of observations below 2000 m depth that can be used for validation and the fact that full-depth-integrated electrical seawater conductivity data products derived from satellite magnetometry are not yet available. We evaluate multiple factors related to the accuracy of OHC anomaly estimation and find that, even with hypothetical measurement errors, our method can be used to monitor OHC anomalies on multiyear time scales.

KEYWORDS: Energy budget/balance; Temperature; In situ oceanic observations; Satellite observations; Ocean models

1. Introduction

In this paper, we propose a remote monitoring technique for ocean heat content (OHC) anomalies and evaluate the method using the output of an ocean state estimate. With the ocean absorbing more than 93% of the excess heat accumulating on our planet due to the presence of greenhouse gases humans have emitted (e.g., Levitus et al. 2001; Trenberth et al. 2014, 2016; von Schuckmann et al. 2016), OHC—the specific heat- and density-weighted depth integral of temperature—has been considered a proxy for Earth's energy imbalance and therefore a critical climate variable to monitor. Further, the associated ocean temperature increase has accounted for roughly half of the observed global mean sea level rise from 1972 to 2008 (Church et al. 2011, 2013; Gregory et al. 2013) and about one-third of the observed global mean sea level rise since 2005 (Chambers et al. 2017; WCRP Global Sea Level Budget Group 2018).

Since the 1970s, large-scale warming of the upper 700 m of the ocean has been well studied (Domingues et al. 2008; Ishii and Kimoto 2009; Durack and Wijffels 2010; Levitus et al. 2012; Abraham et al. 2013; Balmaseda et al. 2013; Lyman and Johnson 2014; Roemmich et al. 2015; Gleckler et al. 2016; Boyer et al. 2016; Ishii et al. 2017; Cheng et al. 2017; Zanna et al. 2019). Our ability to monitor the temperature of the upper 2000 m in regions that are not covered by sea ice has improved considerably since the 2000s due to the deployment of Argo floats (e.g., Riser et al. 2016). However, below 2000 m depth, we must rely on hydrography measured along ship tracks, which has been shown to provide insufficient sampling of the ocean's temperature at such deep depths (Garry et al. 2019). Over the past several decades, observations of hydrographic transects have been the primary source of information about the full-depth OHC; only very recently has the deep Argo program begun to supplement these data (e.g., Johnson et al. 2019). Thus, ocean heat content anomalies have been a challenge to accurately monitor.

Corresponding author: David Trossman, david.s.trossman@gmail.com

DOI: 10.1175/JTECH-D-21-0037.1

Widely accepted approaches to monitoring full-depth OHC include constructing objective analyses of observations, producing reanalyses with data assimilation–based modeling systems, or performing large model ensembles of hindcasts (e.g., Trenberth et al. 2014, 2016), but these approaches substitute the problem of incomplete observations with either interpolation or imperfect modeling systems. Across many different ocean reanalysis products, the ensemble mean can be smaller than the ensemble spread in depth-integrated temperature anomalies over the upper 700 m, particularly in coastal and high-latitude regions (Palmer et al. 2017). Ocean reanalysis products tend to use a combination of in situ observations of temperature (e.g., Kouketsu et al. 2011; Abraham et al. 2013; Balmaseda et al. 2013; Roemmich et al. 2015; Gleckler et al. 2016; Boyer et al. 2016; Ishii et al. 2017; Cheng et al. 2017; Meyssignac et al. 2019) and fill in the spatial gaps with reanalysis data. These spatial gaps are significant because ~13% of the OHC resides in regions, such as those covered by sea ice and at depths below 2000 m (Purkey and Johnson 2010), which are not well sampled by observations (Desbruyères et al. 2016). However, there is considerable bias (Garry et al. 2019) and uncertainty (Llovel et al. 2014) in the extent of the unobserved OHC and warming of such regions has been increasing with time (Gleckler et al. 2016). The hindcast approach relies not only on imperfect ocean models, but also imperfect reanalysis forcing fields. For these reasons, other methods to monitor OHC have been developed.

A number of independent approaches for monitoring OHC have been proposed. These include acoustic time travel measurements (Munk and Wunsch 1979; ATOC Consortium 1998; Dushaw et al. 2009; Wu et al. 2020), satellite altimetry observations of internal tides phase speed changes along their ray paths (Zhao 2016), atmospheric measurements of oxygen and carbon dioxide concentrations (Resplandy et al. 2019), theoretical and model-derived relationships between sea surface heights (from satellite altimetry) and bottom pressure (via satellite gravimetry) with ocean heat content (Jayne et al. 2003; Fasullo and Gent 2017), ocean net surface fluxes (radiative and turbulent adjusted for mass transfer) from space to get the net ocean heating rate (L'Ecuyer et al. 2015), thermal expansion as a residual inferred from space-based observations and the sea level budget (Chambers et al. 2017; WCRP Global Sea Level Budget Group 2018; Hamlington et al. 2020), and depth-integrated electrical conductivity ("conductance") and depth-integrated conductivity-weighted velocity ("conductivity transport") measurements from in situ observations and inferred from satellite magnetometry (Tyler and Sabaka 2016; Irrgang et al. 2017, 2019; Trossman and Tyler 2019). Recently, Tyler (2021) extracted a century of tidal variability from hourly magnetic data taken at the Honolulu Geomagnetic Observatory and found the modulations and trend to be correlated with that seen in tide-gauge data and proposed by Colosi and Munk (2006) to be due to the effect of warming on the internal-tide wave speed. But no single method is known to be capable of both monitoring changes in global OHC accurately enough to resolve annual variations—relevant for quantifying the energy imbalance of our planet—and distinguishing full-depth OHC changes at each horizontal location of the ocean—relevant for hurricane forecasting.

A combination of many of the above methods could be used in conjunction with a machine learning method to monitor OHC. Machine learning methods have a history rooted in statistical regression techniques (e.g., Hastie et al. 2001). Their framework is useful for the purpose of calculating OHC because of established associations with sea surface height, bottom pressure, conductance, and seafloor depth (Jayne et al. 2003; Fasullo and Gent 2017; Trossman and Tyler 2019). One primary difference between a general linear regression technique and a machine learning method is that the latter needs to find a balance between the bias and variance of its predictions through a regularization term. This term prevents the machine learning method from overfitting to a particular training dataset, so that the approach can be applied to other datasets for prediction purposes. To guarantee that the machine learning model does not overfit to the training data, a type of cross-validation method is typically applied by leaving out some of the training data, predicting those data, and repeating for different combinations of the training dataset. Two examples of machine learning methods that are cast in a regression-like framework include the Generalized Additive Model (GAM; Wood 2006; Trossman et al. 2011) and the artificial neural network (e.g., Hsieh and Tang 1998; Wahle et al. 2015; Lary et al. 2016; Irrgang et al. 2019). In this study, we use a GAM to establish whether remotely monitoring OHC is possible using quantities that can be inferred using in situ and/or satellite data—in particular, sea surface height plus seafloor depth (or depth of the water column), bottom pressure, conductance, and OHC in the upper 2000 m.

Because some of the data—whose information potential for remotely monitoring OHC we aim to probe here—are not currently available on a global scale—but are likely to be in the future—we use the output from an ocean state estimate: the Estimating the Circulation and Climate of the Ocean (ECCO) framework. The specific advantages of using ECCO for this study include the following: 1) conductance from satellite magnetometry and its uncertainty are not currently available data products, 2) sea surface height from satellite altimetry and its uncertainty in sea ice–covered regions are not currently available data products, 3) OHC in the upper 2000 m is sparsely and/or intermittently observed in regions without Argo floats, 4) OHC throughout the full water column is observed only where there are historical hydrographic transects and now in select additional locations because of other instruments, 5) we can sample ECCO in any way we want to probe the information potential of any variable on a global scale, 6) we have inserted a conductivity subroutine into ECCO (Trossman and Tyler 2019) so that we retain thermodynamic consistency instead of approximating the conductivity using monthly averaged temperature and salinity fields like other models do, and 7) ECCO has been validated by independent observations from the ones it assimilates and is considered realistic as a data product (Heimbach et al. 2019). For at least some of these reasons, both Trossman and Tyler (2019) and Fournier et al. (2020) used ECCO output to investigate the proxy potential of model variables that are

monitored using satellites. With the help of ECCO in the present study, we aim to establish whether it is possible to remotely monitor OHC anomalies. We do this by treating some ECCO output variables that can be remotely monitored—sea surface heights, bottom pressures, conductance, and Argo-derived OHC in the upper 2000 m—like they are retrievals that can be run through an algorithm—the GAM—to calculate OHC (anomalies). We aim to subsample each of these ECCO output variables according to existing—historical hydrographic transect—observations as training data for the GAM and predict OHC (anomalies) elsewhere—with satellite retrievals. Note that, in this study, we do not use sea surface temperature (SST) or sea surface salinity (SSS), each of which can also be monitored via satellites. Satellite-based microwave measurements, for example, detect the top micrometer and centimeter, respectively, instead of bulk temperature and bulk salinity over the top 10 m, which are ECCO's SST and SSS. Importantly, OHC (anomalies) in the upper 2000 m—which we use here—contains information about SST and is a better approximation of full-depth OHC (anomalies); also, conductance (anomalies)—which we use here as well—contains information about both SST and SSS and is a better proxy of full-depth OHC (anomalies). In the absence of conductance observations, we could use SST and SSS observations but we aim to demonstrate the importance of conductance in the present study.

The structure of this manuscript is as follows. First, we describe the GAM that is used to calculate OHC (anomalies) from several observables and outline the data that the GAM is trained on. We utilize ECCO output in order to examine which potentially satellite-derivable quantities are most associated with OHC (anomalies). We specifically examine the added information from each potential data source and their potential to complement existing methods to monitor OHC anomalies. We do this by evaluating multiple factors related to the accuracy of OHC (anomaly) estimation, including the time period length, inclusion of particular quantities, and sampling strategy of the training data. We last assess the feasibility of our OHC anomaly monitoring strategy and emphasize that there is a balance between retraining the GAM using data over relatively short time scales with the amount of training data that grows over longer time scales in monitoring OHC changes.

2. Modeling system and observations

To accomplish our goal of establishing how accurate an observational network can monitor OHC anomalies—even before consideration of measurement uncertainty—we require a global realistic set of data that can be sampled for several variables. An ocean state estimate that is in excellent agreement with historical observations and their changes relative to ocean reanalyses (Heimbach et al. 2019) is utilized here. We use version 4, revision 3, of ECCO (ECCOv4r3; Fukumori et al. 2017) framework for the ocean state estimate, which is based on the Massachusetts Institute of Technology General Circulation Model (MITgcm) from 1992 to 2015. The fields used for this analysis were generated by a rerun of ECCOv4r3 using the MITgcm. The model's output is sampled along historical hydrographic transects in order to train our statistical model and then the model's globally complete output is compared to the resulting statistical model's estimates of OHC.

a. Ocean state estimation framework

The MITgcm is used to generate the data analyzed for the purpose of this study and is briefly described here. The MITgcm uses the so-called LLC90 grid, which is at a nominal $1° \times 1°$ ($0.5° \times 0.5°$ at the equator) resolution with 50 vertical levels. The shallowest model layer is centered at 5 m depth (5 m thick) and the deepest model layer is centered at 5906.2 m depth (nearly 500 m thick). The model features curvilinear Cartesian coordinates (Forget et al. 2015, see their Figs. 1–3), rescaled height coordinates (Adcroft and Campin 2004), and a partial cell representation of bottom topography (Adcroft et al. 1997). The MITgcm uses a dynamic/thermodynamic sea ice component (Menemenlis et al. 2005; Losch et al. 2010; Heimbach et al. 2010) and a nonlinear free surface with freshwater flux boundary conditions (Campin et al. 2004). The wind speed and wind stress are specified as 6-hourly varying input fields over a 24-yr period (1992–2015). There are 14-day adjustments to the wind stress, wind speed, specific humidity, shortwave downwelling radiation, and surface air temperature. These adjustments are based on estimated prior uncertainties for the chosen atmospheric reanalysis (Chaudhuri et al. 2013), which is ERA-Interim (Dee et al. 2011). The net heat flux is then computed via a bulk formula.

The ECCO framework reconstructs the history of the ocean over the recent satellite era by filling in the gaps of incomplete observations in a dynamically and kinematically consistent manner (Stammer et al. 2016) using the MITgcm and its adjoint-based data assimilation capabilities. The reason why the ECCO state estimate can be performed in a dynamically and kinematically consistent manner is that its initial conditions and model parameters are estimated via least squares using assimilated observational constraints. The least squares problem solved by the ECCO framework utilizes the method of Lagrange multipliers through iterative improvement, which relies upon a quasi-Newton gradient search (Nocedal 1980; Gilbert and Lemarechal 1989). The tangent linear model (Jacobian) and its transpose (the adjoint) are needed to solve for the Lagrange multipliers. Algorithmic (or automatic) differentiation tools (Griewank 1992; Giering and Kaminski 1998) have allowed for the practical use of Lagrange multipliers in a time-varying nonlinear inverse problem such as the one for the ocean because the discretized adjoint equations no longer need to be explicitly hand coded. Each of the data points in the time interval of 1992–2015 is weighted by a best-available estimate of its error variance. The observational data assimilated into the ECCO framework to arrive at the model's objective—to reconstruct the ocean's historical conditions—are discussed in Wunsch and Heimbach (2013). These data include satellite-derived ocean bottom pressures, sea ice concentrations, sea surface temperatures, sea surface salinities, sea surface height anomalies, and mean dynamic topography, as well as profiler- and mooring-derived temperatures and salinities (Fukumori et al. 2017). The control variables that are solved for by ECCO include the initial

condition of the velocities, sea surface heights, temperatures, and salinities; time-mean three-dimensional Redi (Redi 1982) coefficients, Gent–McWilliams (Gent and McWilliams 1990) coefficients, and vertical diffusivities (Gaspar et al. 1990); and time-varying two-dimensional surface forcing fields. The Redi and Gent–McWilliams coefficients are needed because our configuration of ECCO is not eddy resolving. If ECCO were optimized in an eddy-resolving configuration, minimizing misfits between the model and observations could result in overfitting the ocean parameters and state to each eddy over such a long time period (decades), so the ECCO state estimate has been produced at non-eddy-resolving configurations. Note that the effects of (sub)mesoscale eddies could alter the relationships between, for example, sea surface height/bottom pressure anomalies and OHC anomalies relevant to this study, but because satellites typically cannot resolve eddies either, using a non-eddy-resolving model is ideal for the purpose of the present study. Fifty-nine iterations in the optimization run of ECCO were performed to arrive at the solution we start from. The initial conditions and model parameters calculated in this final iteration are used for the reruns performed here. We perform a rerun of the ECCOv4r3 solution, which goes from 1992 through 2015 (Fukumori et al. 2017). Schemes for calculating the conductivity and specific heat at each time step as the model runs are taken from the TEOS-10 package (McDougall and Barker 2011). The relationship between the conductivity and the electromagnetic fields measured by satellite magnetometry is described below.

b. *Calculation of electromagnetic fields*

Ohm's law for a moving conductor,

$$\mathbf{J} = \sigma(\mathbf{E} + \mathbf{u} \times \mathbf{F}), \tag{1}$$

is a vector equation describing the electric current density $\mathbf{J}$ generated by an electric field $\mathbf{E}$ and/or the relative velocity of seawater $\mathbf{u}$ of the conducting fluid as it moves through the magnetic field $\mathbf{F}$, which we take to be the prescribed background main magnetic field; the total magnetic field $\mathbf{B} = \mathbf{F} + \mathbf{b}$ includes a component $\mathbf{b}$ associated with $\mathbf{J}$. When the flow velocity is not considered, Eq. (1) reduces to $\mathbf{J} = \sigma\mathbf{E}$, and the electrical conductivity σ can be observed as simply the ratio of $\mathbf{J}$ and $\mathbf{E}$, as might be obtained from in situ measurements, for example. Alternatively, in experiments where σ is observed, one may infer instead the flow velocity components $\mathbf{u}$. Hence, it is fairly direct to see how in situ electromagnetic (EM) observations can be used to infer or constrain ocean conductivity and/or velocity.

As the first departure from these truly in situ observations, one can describe configurations where EM observations on the seafloor, for example, can be used to estimate bulk integrated ocean parameters. Consider a controlled electric current source on the seafloor and assume the cable/antenna length is of a scale exceeding that of the ocean thickness. The electrical currents return throughout the water column and their amplitude will be modulated by any changes in the conductance. We see then a potential observational advantage as this seafloor system can be used to monitor depth integrated ocean parameters. Where the conductivity fluctuates due to change in water temperature, for example, this system could be regarded as a bulk thermometer of ocean temperature. Using an alternating current source to remove problems such as electrode drift, very high accuracy could be achieved. One would likely operate this system at frequencies low enough such that the ocean appears "electrically thin," meaning that the electromagnetic wavelengths in the ocean are much larger than the ocean thickness such that the return electric currents reach through the water column as described. The associated period increases with conductance, and therefore typically also with ocean thickness, but does not exceed 10 min even in the thickest ocean regions (Tyler 2017).

One need not, however, have in situ or seafloor observations of $\mathbf{J}$ and $\mathbf{E}$ in order to make parameter estimates. Maxwell's equations can be combined with Eq. (1) into a governing electromagnetic induction equation:

$$\partial_t \mathbf{B} = \nabla \times \left(\mathbf{u} \times \mathbf{B} - \frac{1}{\mu_0 \sigma} \nabla \times \mathbf{B}\right), \tag{2}$$

where μ_0 is the vacuum permeability constant. Note that the three (of the four) underlying Maxwell's equations, from which the induction equation is derived, remain approximately invariant when adopting the rotating frame of the planet and considering global scales and periods similar to a day (Tyler and Mysak 1995). Here the opportunity for inferring/constraining the ocean parameters σ, $\mathbf{u}$ from remote observations of $\mathbf{B}$ is expressed. Specialized forms of the induction equation appropriate for large scales near Earth's surface are described in detail in Tyler (2017). Even when due to electric currents within the ocean, the magnetic fields pass through sea ice and can reach satellite altitudes. But because of geometric attenuation away from the sources, the fields associated with features having length scales much smaller than the satellite altitude will be reduced. Hence, the remote magnetic fields mostly describe depth-integrated, large-scale ocean features. One can see in the specialized forms of the induction equation (Tyler 2017) that the ocean parameters that are potentially inferred are the conductance $\Sigma = \int_h \sigma \, dr$ and the conductivity transport $\mathbf{T}_\sigma = \int_h \sigma \mathbf{u} \, dr$.

One can regard the electric currents in the ocean (and their associated magnetic fields, which reach beyond the ocean) as generated by either a time-dependent component of the magnetic field incident on the ocean surface, or as due to the ocean flow whereby a small part of the flow's kinetic energy is spent driving these currents. The first process is referred to as electromagnetic "induction" and a very common application involves magnetic fields incident on the ocean due to electric currents in the ionosphere and magnetosphere. One can regard the induction process as one where electric currents at one location (e.g., the ionosphere) entrain electric currents in another conductor (e.g., the ocean) through the connection of their Coulomb fields which can reach over great distance and even through insulators. The second process is referred to as "motional induction" and can be loosely regarded as due to the tendency of a moving electrical conductor to entrain a permeating

magnetic field. In the case of a perfect conductor, the magnetic field is regarded as "frozen in" and moves with the conductor. The frozen in scenario is not typically achieved in ocean applications as the conductivity is not high enough (or the flow time scales short enough) to reduce the importance of the magnetic diffusion term [the last term in Eq. (2)].

Finally, for the purposes of this paper it should be noted that while Σ is a parameter potentially recoverable from either induction or motional induction processes, $\mathbf{T}_\sigma$ can be recovered/constrained only in processes of motional induction. Because Σ and $\mathbf{T}_\sigma$ might be inferred in different and varying conditions that also involve a range of expected errors, in this study we shall consider the addition of Σ and $\mathbf{T}_\sigma$ to the GAM separately. Then we use prescribed reference error levels with the most important satellite-derivable quantities.

c. *Observed hydrographic transects*

The hydrographic transects used in this study are taken from the World Ocean Circulation Experiment (WOCE) and the Climate Variability and Predictability (CLIVAR) programs (Fig. 1). Specifically, we use transects that have adequate information about both temperature and salinity to calculate a density and/or electrical conductivity (σ), as in previous studies that require one or more of these three quantities (e.g., Kunze 2017). Because we will never measure the full-depth OHC at every horizontal location in the ocean, it is not very practical to use observations at every point in the ocean to train the GAM and then estimate the full-depth OHC at each location. Instead, we train the GAM along particular transects that have been sampled by ships and apply the GAM.

Along these transects, we sample the following variables from the ECCO output: sea surface height (SSH), bottom pressure (p_b), conductance (Σ), seafloor depth (H_0), and conductivity transport vector ($\mathbf{T}_\sigma$) and its divergence. From the ship, the SSH can be measured. The SSH anomaly is related to the sea level anomaly, which is primarily a function of OHC and added mass. The deepest measurement taken from the ship can be used to infer p_b. The p_b anomalies are important to account for the added mass contributions to sea level as a correction to using SSH anomalies as a proxy for OHC anomalies (Jayne et al. 2003). Seafloor cables (e.g., Schnepf et al. 2021), such as the Florida Cable, measure voltage differences, which are converted to an estimate of the conductivity-weighted depth-averaged flow velocity crossing the cable (Sanford 1971) or what we refer to here as the cross-cable component of $\mathbf{T}_\sigma$. Mooring arrays (e.g., Lozier et al. 2019), such as OSNAP, provide both conductivities and velocities such that Σ and $\mathbf{T}_\sigma$ can be calculated. Σ could also be delivered by seafloor and surface-towed observatories, which could monitor the full-depth conductance quite accurately but would realistically provide spatial coverage similar to ship-based hydrographic data. Σ is expected to be spatially correlated with OHC over much of the ocean because both Σ and OHC are depth integrated. Because $\mathbf{T}_\sigma$ is not necessarily measured—in situ—along the same hydrographic transects as the other variables, there may be logistical difficulties with using all of these data to train a GAM that calculates OHC. This is why we examine the importance of including $\mathbf{T}_\sigma$ in the GAM in the present study. Last, H_0 can be inferred from ship-based measurements. H_0 is important to account for because a deeper ocean has the capability to hold more heat at a given location. H_0 can change over time due to processes such as bottom deformation (Vishwakarma et al. 2020), but is assumed to be temporally constant here because ECCO, whose output we sample, makes this assumption. A GAM can update the relationship between OHC and a static field such as H_0 when all other (above-listed) potentially satellite-derivable predictors are included over a relevant time scale; this is why the GAM could require updated training data for each time period over which the OHC is monitored.

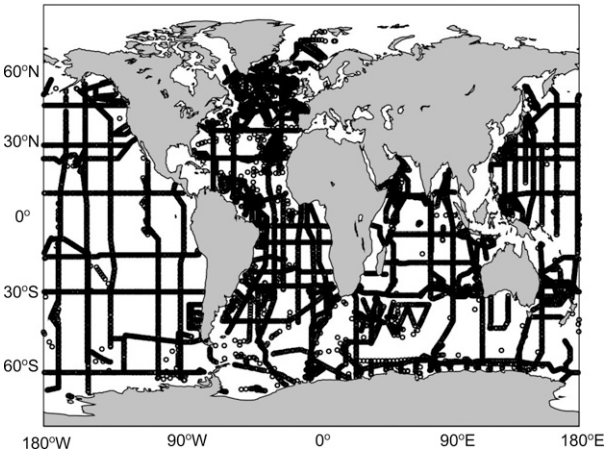

FIG. 1. Locations (black dots) where hydrographic transect observations were taken by ship and used in the present analysis.

SSH anomalies are routinely monitored by satellite altimetry over the global ocean, whereas p_b, Σ, and $\mathbf{T}_\sigma$ can potentially be derived from satellite data. In particular, p_b—in water column thickness equivalent—can be derived from satellite gravimetry (Ponte et al. 2007). The values of Σ and $\mathbf{T}_\sigma$ may be derived or constrained using electric and magnetic field observations (including observations by satellite magnetometers) and this paper therefore considers their use in a GAM. The value of H_0 has been derived from ship-based and satellite measurements and is essentially time invariant, requiring that the GAM be retrained at each time we want to estimate OHC. After being trained on transects of in situ measurements of SSH, p_b, Σ, H_0, and $\mathbf{T}_\sigma$, OHC can then be estimated using bathymetry and time-dependent satellite observations with a GAM of the form given in the following subsection. Using hypothetical measurement errors (see below), we further examine what the practical limitations are to using SSH, p_b, Σ, H_0, $\mathbf{T}_\sigma$, or some subset of these variables, as predictors of OHC.

d. *Generalized Additive Model specifications*

We present scatterplots of the 1992–2015 averages of the OHC versus each of its potential predictors in Fig. 2, including the aforementioned electromagnetic variables. The strongest

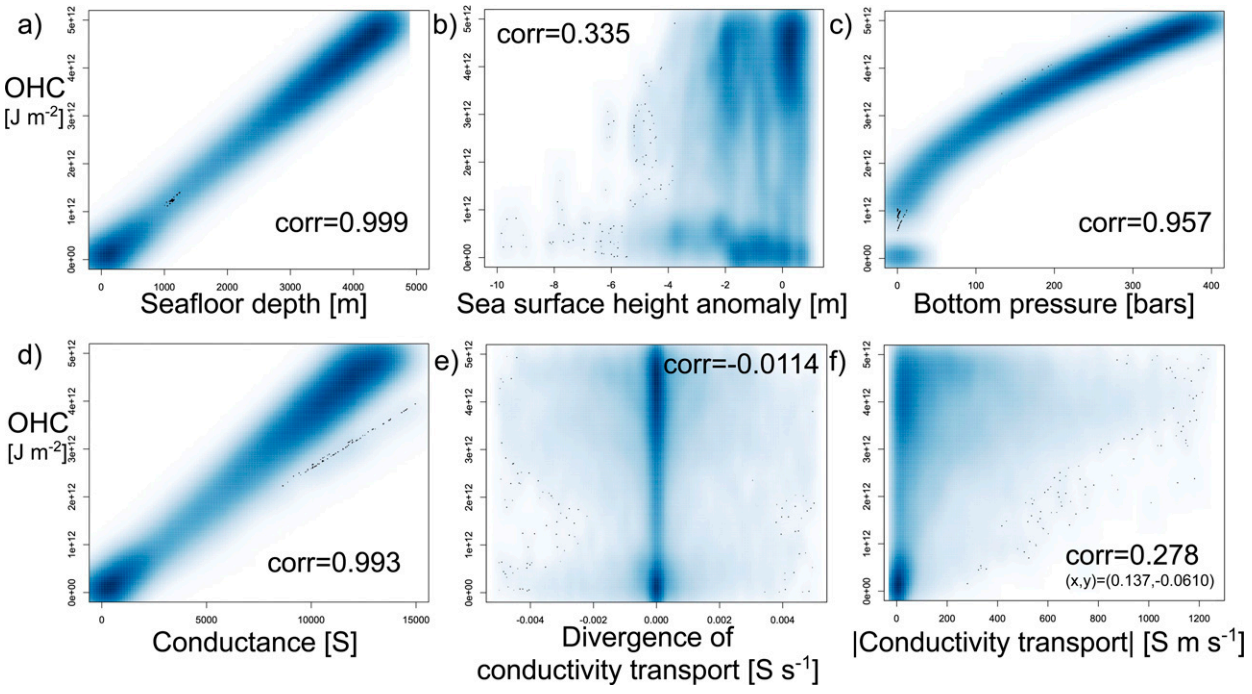

FIG. 2. Scatterplots between ocean heat content (OHC; units: J m^{-2}) and (a) the seafloor depths (units: m), (b) the sea surface height anomalies (units: m), (c) the bottom pressures (units: bars), (d) the conductances (units: S), (e) the divergences of the conductivity transports (units: S s^{-1}), and (f) the magnitudes of the conductivity transports (units: S m s^{-1}), each averaged over 1992–2000. The darker blue colors indicate there is a greater density of dots. The black dots are individual datum; these points are not great enough in number to correspond to a shade of blue. Also listed are the correlations between each of the quantities plotted (corr).

correlation is between OHC and H_0 (Fig. 2a). There is a fairly good correlation between OHC and SSH anomaly, as has been noted in previous studies (Fig. 2b). There is a strong correlation between OHC and both p_b (Fig. 2c) and Σ (Fig. 2d). The spatial distributions of each of OHC and each of the most highly correlated variables (H_0, p_b, and Σ) are shown in Fig. 3. This figure suggests that these variables are most strongly correlated in deeper parts of the ocean and more weakly correlated elsewhere (e.g., the Arctic Ocean). The dominant mechanisms explaining ocean heat uptake variability—namely, Ekman heat transport convergences, air–sea fluxes, and other aspects of ocean dynamics—are different over various regions of the ocean (e.g., the subtropical versus subpolar North Atlantic; Buckley et al. 2015). Because conductivity primarily varies with temperature in most regions of the ocean, we expect $\mathbf{T}_\sigma$ to vary with heat transport, and the divergence of the conductivity transport, $\nabla \cdot \mathbf{T}_\sigma$, would be related to heat transport convergence. However, in the ECCO output, $\nabla \cdot \mathbf{T}_\sigma$ is poorly correlated with OHC (Fig. 2e), and is therefore excluded from the rest of our analysis. Reasons for the poor correlation include the boundary (e.g., air–sea) fluxes that can lead to nonconservation of temperature and our depth integration over the full water column instead of a fraction of the upper ocean. Also, $\nabla \cdot \mathbf{T}_\sigma$ is related to the time rate of change in OHC, not OHC itself, such that a running time integral of $\nabla \cdot \mathbf{T}_\sigma$—which should be $-\partial\Sigma/\partial t$ if conductivity is a conservative tracer—could be used. However, this would present additional difficulties—associated with sustaining an observing system for $\nabla \cdot \mathbf{T}_\sigma$ over longer periods of time—in addition to the challenges with calculating divergences along the transects used to train our GAM. Despite how the sign of each individual component of $\mathbf{T}_\sigma$ could matter, the individual components of $\mathbf{T}_\sigma$ are poorly correlated with OHC. However, there is a marginally fair correlation between OHC and $|\mathbf{T}_\sigma|$ (Fig. 2f), which is likely due to greater heat/conductivity transport magnitudes found where there is more heat/conductivity and their common depth integrals. The scatterplots shown in Fig. 2 look virtually identical when either monthly or annual averages of each quantity are considered, and their correlations are qualitatively the same.

We also show the scatterplots for OHC anomalies—computed monthly relative to the average over the entire 1992–2015 time interval—versus a subset of the variables that were found to be fairly well correlated to OHC in Fig. 4. (The anomalies in $\nabla \cdot \mathbf{T}_\sigma$ are poorly correlated with OHC anomalies—magnitudes less than about 0.05. The anomalies in $|\mathbf{T}_\sigma|$ have a range of correlations with OHC anomalies similar to seafloor depths and do not provide significant unique information about OHC anomalies—shown later, but for OHC magnitudes only.) One month (January 1992) is shown in Fig. 4, but we inspected other months and indicate a range of correlations with OHC anomalies, depending upon the month, in each panel. Σ anomalies are more strongly correlated with OHC anomalies than any other combination of variables (Fig. 4a). This highlights the importance of including Σ anomalies as a

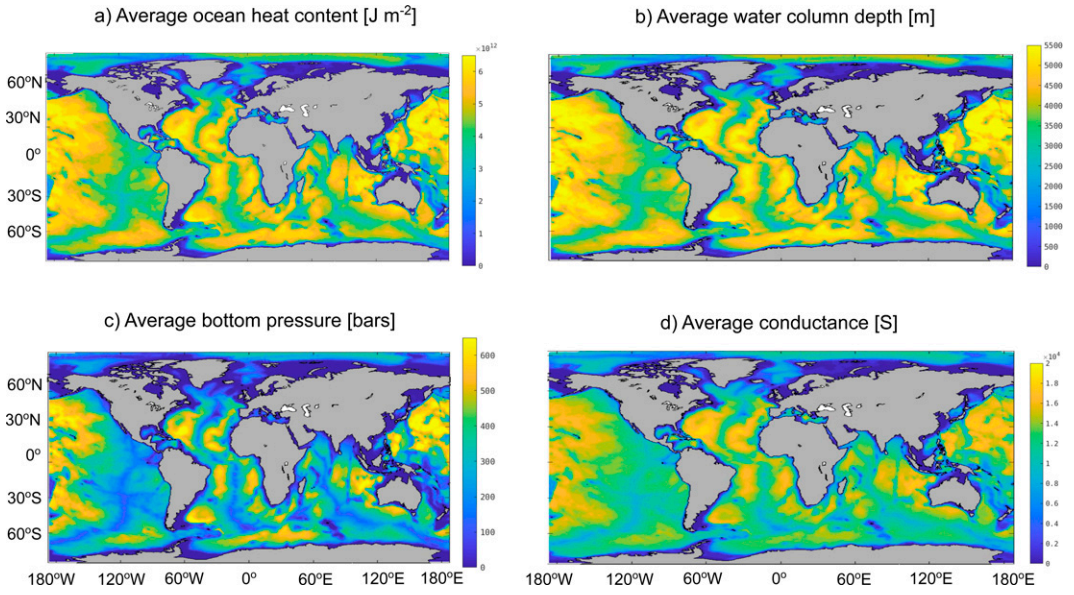

FIG. 3. Maps of (a) the OHC (units: J m^{-2}), (b) the water column depth (units: m), (c) the bottom pressures (units: bars), and (d) the conductances (units: S), each averaged over 1992–2015.

potential proxy for OHC anomalies. The next strongest correlation is between OHC anomalies and SSH anomalies (Fig. 4b), like Jayne et al. (2003) found. p_b anomalies and H_0 marginally correlate with OHC anomalies (Figs. 4c,d), but because Jayne et al. (2003) found that p_b anomalies were important to consider, we include them here in our GAM. We later include—as a predictor in the GAM—one last variable that can be relayed via satellites, but is derived from in situ measurements: the OHC in the upper 2000 m, or $\mathrm{OHC}_{2000\mathrm{m}}$. We choose 2000 m here because of the depth range that is covered nearly globally by Argo floats. Including $\mathrm{OHC}_{2000\mathrm{m}}$ makes it more challenging to remotely monitor OHC globally because of regions that are ice covered for part of each year, but when known, this variable is very accurately known.

This motivates our use of a GAM for OHC of the form

$$\begin{aligned}\widehat{\mathrm{OHC}} = {} & f_0 + f_1\left(\widehat{\mathrm{SSH}}' + \hat{H}_0\right) + f_2(\hat{p}_b) \\ & + f_3(\hat{\Sigma}) + f_4\left(\widehat{\mathrm{OHC}_{2000\mathrm{m}}}\right) \\ & + g\left(\widehat{\mathrm{SSH}}', \hat{p}_b, \hat{\Sigma}, \hat{H}_0, \widehat{\mathrm{OHC}_{2000\mathrm{m}}}\right), \end{aligned} \tag{3}$$

$$\widehat{\mathrm{SSH}}' = \mathrm{SSH}'(\eta_{\mathrm{fac}} + \epsilon_{\eta_{\mathrm{fac}}})\theta(\lambda - \lambda_z) + \epsilon_{\mathrm{SSH}'},$$

$$\hat{p}_b = p_b(m + \epsilon_m) + \epsilon_{p_b},$$

$$\hat{\Sigma} = \Sigma(\mathbf{b} + \boldsymbol{\epsilon}_\mathbf{b})\theta\left[\hat{G}_\Sigma(\mathbf{b} + \boldsymbol{\epsilon}_\mathbf{b})\right] + \epsilon_\Sigma,$$

$$\hat{H}_0 = H_0 + \epsilon_{H_0},$$

$$\widehat{\mathrm{OHC}_{2000\mathrm{m}}} = \mathrm{OHC}_{2000\mathrm{m}} + \epsilon_{\mathrm{OHC}_{2000\mathrm{m}}},$$

$$\theta(x) = \begin{cases} 1, & \text{if } x \geq 0 \\ 0, & \text{if } x < 0 \end{cases},$$

and for OHC anomalies (OHC') of the form

$$\begin{aligned}\widehat{\mathrm{OHC}'} = {} & f_0 + f_1\left(\widehat{\mathrm{SSH}'}\right) + f_2(\hat{p}'_b) \\ & + f_3(\hat{\Sigma}') + f_4\left(\widehat{\mathrm{OHC}'_{2000\mathrm{m}}}\right) \\ & + g\left(\widehat{\mathrm{SSH}'}, \hat{p}'_b, \hat{\Sigma}', \widehat{\mathrm{OHC}'_{2000\mathrm{m}}}\right), \end{aligned} \tag{4}$$

$$\widehat{\mathrm{SSH}'} = \mathrm{SSH}'(\eta_{\mathrm{fac}} + \epsilon_{\eta_{\mathrm{fac}}})\theta(\lambda - \lambda_z) + \epsilon_{\mathrm{SSH}'},$$

$$\hat{p}'_b = \hat{p}'_b(m + \epsilon_m) + \epsilon_{p'_b},$$

$$\hat{\Sigma}' = \Sigma'(\mathbf{b} + \boldsymbol{\epsilon}_\mathbf{b})\theta\left[\hat{\mathbf{G}}_{\Sigma'}(\mathbf{b} + \boldsymbol{\epsilon}_\mathbf{b})\right] + \epsilon_{\Sigma'},$$

$$\widehat{\mathrm{OHC}'_{2000\mathrm{m}}} = \mathrm{OHC}'_{2000\mathrm{m}} + \epsilon_{\mathrm{OHC}'_{2000\mathrm{m}}},$$

where $f_i(\cdot)$ for $i = 0, \ldots, 4$ are smoother functions, $g(\cdot)$ is the sum of tensor products of each cross-pairwise combination of arguments (i.e., squares of each variable are not included), the hat/circumflex ($\hat{\cdot}$) indicates a measurement (its absence indicates the truth), the variables with arguments and without a hat are derived from the quantities that a satellite measures (arguments being intermediate quantities that are inferred), and ϵ_X indicates measurement error for variable X. The GAM-based estimates of OHC using Eq. (3) are insensitive to the definition of OHC as a function of temperature anomaly or temperature because f_0 is a spatially varying but temporally constant offset. SSH′ is a function of all of the correction factors (η_{fac}) involved in the retrieval algorithm

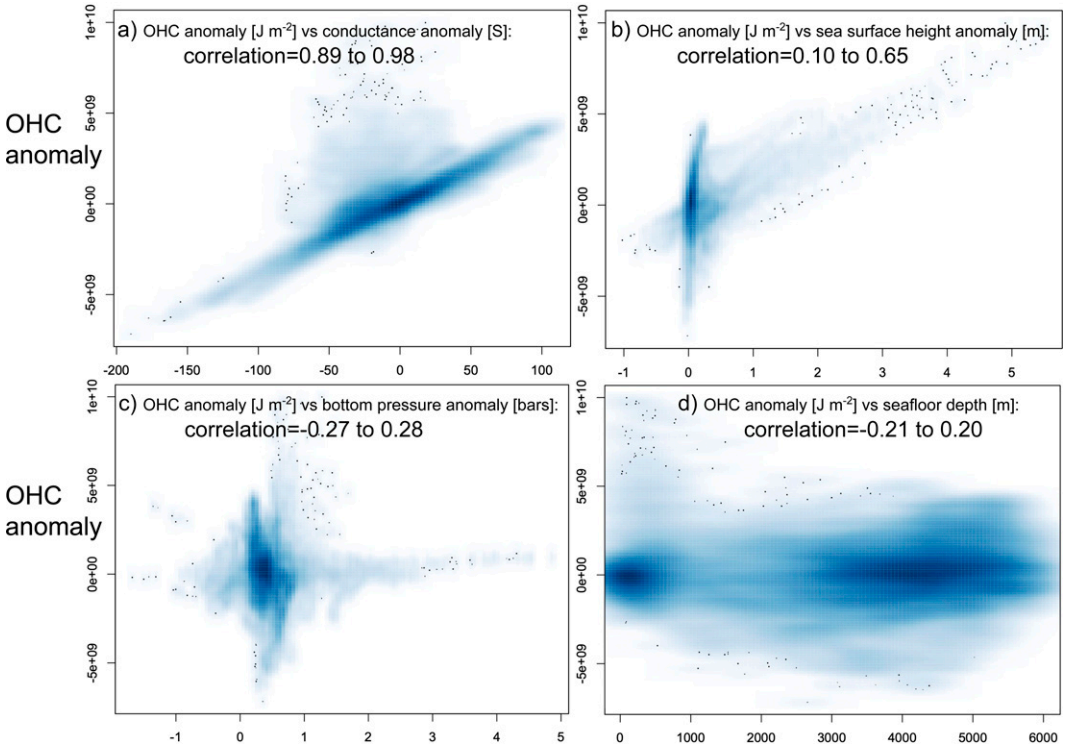

FIG. 4. Scatterplots between OHC (units: J m^{-2}) anomalies and (a) conductance (units: S), (b) the sea surface height anomalies (units: m), (c) the bottom pressure anomalies (units: bars), and (d) the seafloor depths (units: m) for the month of January 1993. The darker blue colors indicate there is a greater density of dots. The black dots are individual datum; these points are not great enough in number to correspond to a shade of blue. Also listed are the range of correlations between each of the quantities plotted (corr), depending upon the month selected.

and postprocessing from satellite altimetry (e.g., the tides). SSH has been observed over poleward latitudes (λ) of $\lambda_z = 66°$ for only a subset of the history of satellite altimetry. The bathymetry (H_0) is included in the GAM such that $H = H_0 + \mathrm{SSH}'$ is the water column depth. Alternatively, we can formulate the GAM in terms of the absolute dynamic topography, but the geoid—like H_0—could be separable into the constant function, f_0, for each time considered. If the nonlinear interaction terms in $g(\cdot)$ are significant, then the geoid and H_0 are not separable and either the geoid or H_0 need to be accounted for. For proof-of-concept purposes, we stick with using H_0 and not the geoid. The bottom pressure p_b is a function of the mass (m) inferred from the retrieval algorithm from satellite gravimetry. The conductance Σ and conductivity transport $\mathbf{T}_\sigma$ are functions of the magnetic field $\mathbf{b}$ inverted from the retrieval algorithm and postprocessing from satellite magnetometry; the functions that indicate whether these inversions are possible (when ≥ 0) are represented by $\hat{\mathbf{G}}_\Sigma$ and $\hat{\mathbf{G}}_{|\mathbf{T}_\sigma|}$, respectively. Although ECCOv4r3 does not include tides to demonstrate whether tidal amplitudes are useful information for the GAM in addition to the other predictors, the information these tidal amplitudes provide [e.g., M_2, like Irrgang et al. (2019) included] is likely similar to that provided by Σ and/or $\mathbf{T}_\sigma$.

Using many terms in the GAMs for OHC and OHC anomalies, we perform a preliminary sensitivity analysis to the uncertainties in the observations. The accuracy in which Σ and/or $\mathbf{T}_\sigma$ may be estimated from satellite magnetic data has not yet been established, so we only examine sensitivities of the RMSE to example values. To do this for each variable in Eqs. (3) and (4), random noise is selected from a normal distribution with mean zero and standard deviation equal to various levels [ϵ_X in Eqs. (3) and (4) for each variable X]. This noise is added to the predictors in Eqs. (3) and (4) because the satellite data carry the majority of the observational uncertainties. OHC is reestimated using the GAM approach with the added noise. Unless otherwise stated, the standard deviations (of the normal distributions of the randomly sampled measurement errors) are set to be $\epsilon_{\mathrm{SSH}'} = 1$ cm, $\epsilon_{p_b} = 2$ bars, $\epsilon_\Sigma = 3$ S, $\epsilon_H = 1$ m, and $\epsilon_{|\mathbf{T}_\sigma|} = 0.5$ S m s^{-1} for the sensitivity calculations and the same values are used for the standard deviations in each variable's corresponding anomaly. $\mathrm{OHC}_{2000\mathrm{m}}$ is excluded for the purpose of the preliminary error sensitivity analysis, but included in the final GAM because of its significant predictive information. The hypothetical measurement errors for ocean heat content over the top 2000 m $\mathrm{OHC}_{2000\mathrm{m}}$ are at most $\epsilon_{\mathrm{OHC}_{2000\mathrm{m}}} = 1029 \times 4850 \times 2000 \times 0.002 = 19\,962\,600$ J m^{-2}, compared to a typical value of OHC, $\sim 4 \times 10^{12}$ J m^{-2}. Because of this accuracy, we do not include $\mathrm{OHC}_{2000\mathrm{m}}$ in the preliminary error analysis, but we will include $\mathrm{OHC}_{2000\mathrm{m}}$ in our final estimates that demonstrate the feasibility of our remote monitoring method.

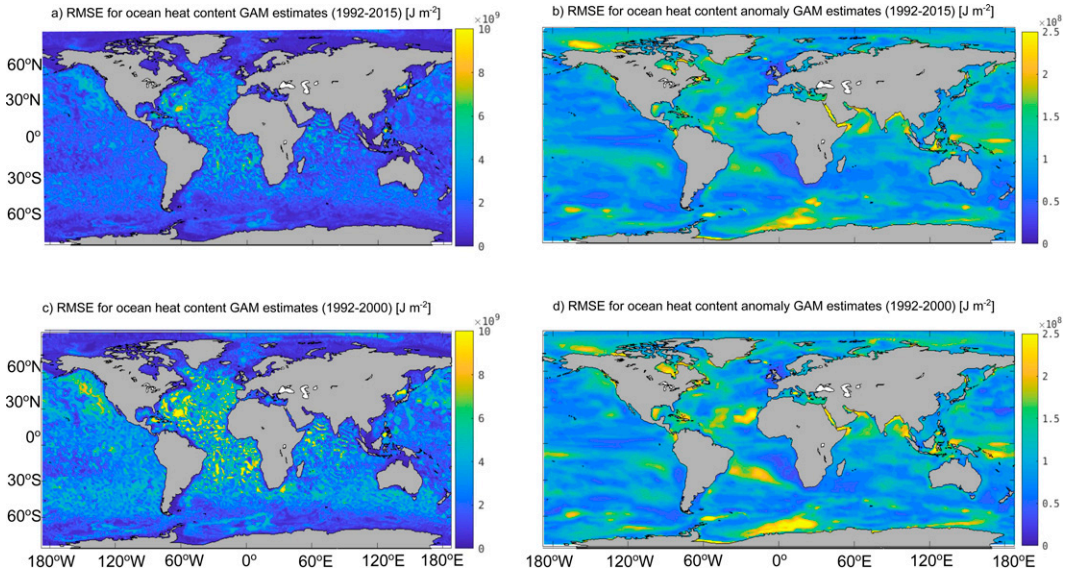

FIG. 5. Importance of time period length in estimating OHC: shown are the root-mean-square errors (RMSEs) (a) using the GAM, $\mathrm{OHC} \sim f_1(\mathrm{SSH}') + f_2(p_b) + f_3(\Sigma) + f_4(H_0) + g(\mathrm{SSH}', p_b, \Sigma, H_0)$, for the full length of ECCOv4r3 (1992–2015) and (c) using the same GAM specification, but for the first 9 years (1992–2000). Also shown is the same for OHC anomalies (OHC′) (b) using the GAM, $\mathrm{OHC}' \sim f_1(\mathrm{SSH}') + f_2(p_b') + f_3(\Sigma') + f_4(H_0) + g(\mathrm{SSH}', p_b', \Sigma', H_0)$ for the full length of ECCOv4r3 (1992–2015) and (d) using the same GAM specification, but for the first 9 years (1992–2000). The Black Sea and Caspian Sea are not included in ECCO.

3. Results

We first present results to identify some important factors to consider when estimating OHC and/or OHC anomalies with a machine learning algorithm, such as a GAM. These factors include the time period length, the inclusion of particular satellite-derived quantities, and the spatiotemporal sampling of any in situ observations that can help reduce the root-mean-square error (RMSE) of the GAM estimates. For each of these results, we train our GAM on (spatially sampled) monthly averaged output of ECCOv4r3 and estimate OHC everywhere, iterating this procedure for each month over specified time intervals. We present RMSE values with respect to OHC and OHC anomalies when demonstrating the relative importance of each predictor. Last, we present additional factors to consider when practically implementing our proposed remote monitoring method for OHC. We investigate whether we can detect OHC anomalies using the GAM-based estimates. We determine this in two ways on variable time scales: 1) GAM-based estimates are within a 95% confidence interval of the ECCOv4r3 values or 2) RMSE $< 100\%$. To determine whether our method can be practically implemented with observations in a future study with the second criterion, we present normalized RMSE values (percent RMSE) when considering all factors. We show that conclusions from each of the two criteria are often similar.

a. *OHC versus OHC anomaly estimation: Important factors*

We first consider the importance of the time period length over which we estimate OHC and OHC anomalies using monthly data everywhere over the full 1992–2015 time interval of ECCOv4r3 versus the first 9 years (1992–2000). The GAM-based estimation procedure's RMSE for OHC is smaller when performed over the entire time period of ECCOv4r3 than over the first 9 years (Figs. 5a,c). A similar result is shown in Figs. 5b and 5d for OHC anomalies, but the difference between the RMSEs over the two different time periods is not as obvious as it is for the OHC estimates. The spatial patterns shown in Figs. 5a–d are qualitatively different. The globally averaged OHC anomalies estimated by the GAM gradually go from larger to smaller than the globally averaged

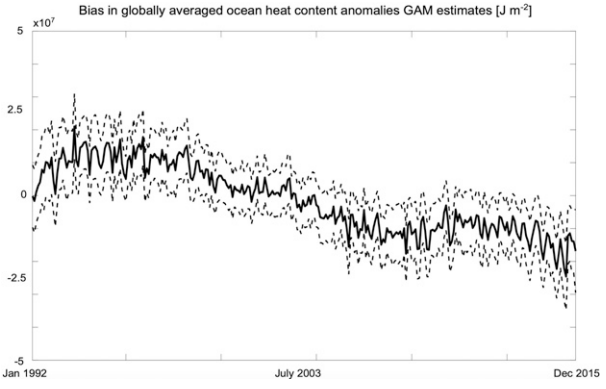

FIG. 6. Accuracy in estimating OHC anomalies: shown with a black curve are the OHC anomalies using the GAM, $\mathrm{OHC}' \sim f_1(\mathrm{SSH}') + f_2(p_b') + f_3(\Sigma') + f_4(H_0) + g(\mathrm{SSH}', p_b', \Sigma', H_0)$, minus the OHC anomalies for the full length of ECCOv4r3 (1992–2015) and the confidence intervals to the 95% level based on the standard errors (dashed curves) from the GAM (assuming they are normally distributed).

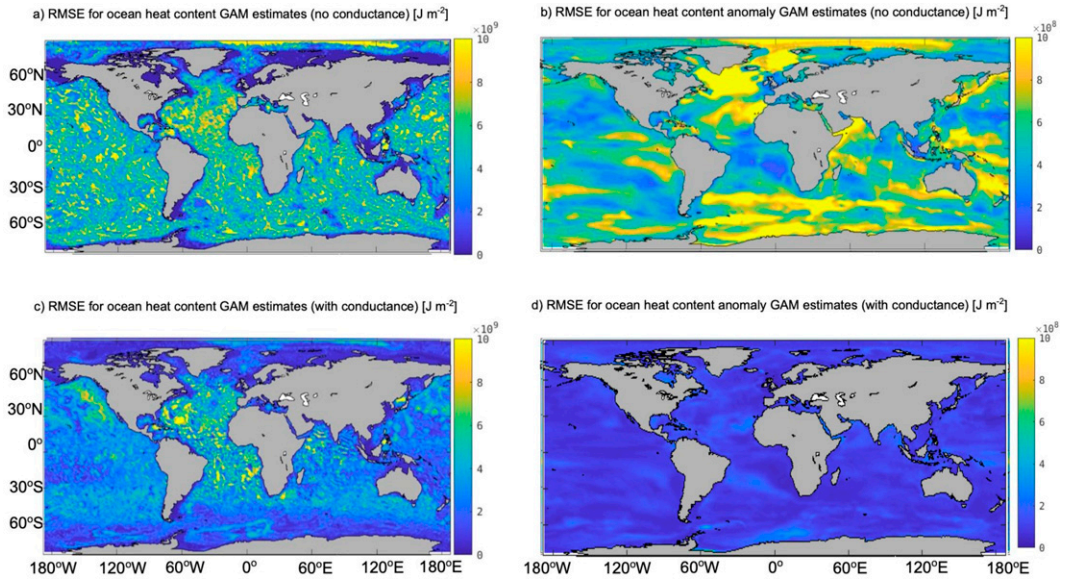

FIG. 7. Importance of including conductance in estimating OHC: shown are the RMSEs (a) using the GAM, OHC ~ $f_1(\mathrm{SSH}') + f_2(p_b) + f_3(H_0) + g(\mathrm{SSH}', p_b, H_0)$, over the first 9 years of ECCOv4r3 (1992–2000) and (c) for the GAM, OHC ~ $f_1(\mathrm{SSH}') + f_2(p_b) + f_3(\Sigma) + f_4(H_0) + g(\mathrm{SSH}', p_b, \Sigma, H_0)$, over the first 9 years (1992–2000). Also shown is the same for OHC′ (b) using the GAM, OHC′ ~ $f_1(\mathrm{SSH}') + f_2(p_b') + f_3(H_0) + g(\mathrm{SSH}', p_b', H_0)$, over the first 9 years of ECCOv4r3 (1992–2000) and (d) for the GAM, OHC′ ~ $f_1(\mathrm{SSH}') + f_2(p_b') + f_3(\Sigma') + f_4(H_0) + g(\mathrm{SSH}', p_b', \Sigma', H_0)$, over the first 9 years (1992–2000). Note that the color-bar scale is different in (b) and (d) from Figs. 5b and 5d. The Black Sea and Caspian Sea are not included in ECCO.

ECCOv4r3 OHC anomalies (Fig. 6). The errors in the globally averaged OHC anomalies are typically insignificant according to the confidence intervals at the 95% level and are almost always within 10% of the globally averaged ECCOv4r3 OHC anomalies (not shown). It is tempting to take time differences across GAM-based OHC estimates to get OHC anomalies that are more accurate than the GAM-based OHC anomaly estimates in some locations, but it remains to be seen in our analysis whether this would be useful. These findings further suggest that the remainder of our analysis will be sensitive to the time periods over which we perform the GAM-based estimation procedures; we focus on the 1992–2000 time period for worst case scenario RMSEs.

We next examine the information that conductance provides to OHC and OHC anomaly estimates using monthly data everywhere over the time period with the larger RMSE (1992–2000). Recall that we provide information from the variables shown in Fig. 3 for the OHC calculations and exclude seafloor depths for the OHC anomaly calculations because seafloor depths provide little information about OHC anomalies (Fig. 4d). The RMSE of the OHC estimates are an order of magnitude larger when conductance is excluded from the GAM than when it is included (Figs. 7a,c). This is true even in regions where temperature does not explain the majority of the variability in conductivity (Trossman and Tyler 2019, see gray circles in their Figs. 5a,b). The RMSE of the OHC anomaly estimates are much larger when conductance is excluded from the GAM compared to when conductance is included (Figs. 7b,d; note the different scale in Fig. 7d). For both OHC and OHC anomalies, the RMSE of the GAM-based estimates is larger everywhere without conductance in the GAM, not only in general. Later, we will assess the importance of each predictor in the GAM. We expect that the conductance will stand out as one of the most important predictors to include in the GAM and the conductivity transport and its divergence will be the least important.

Another important factor to consider including in the GAM that could be practically relevant when implementing a remote monitoring system for OHC is the use of OHC observed over the entire water column in the GAM's training data. Of course, OHC over the top 2000 m is included within these training data so we examine the information that OHC over the top 2000 m provides to the GAM. The OHC in the top 2000 m is highly correlated is the OHC over the full water column (Fig. 3a; Pearson correlation coefficient of 0.95). The OHC anomalies in the top 2000 m is even more highly correlated with OHC anomalies throughout the water column (Pearson correlation coefficient of 0.97–0.99). The percentage of the full-column OHC that is in the top 2000 m varies in space correlates with the underlying bathymetry, which explains the high correlations that are close to constant over time. We now assume that we have OHC over the top 2000 m everywhere to demonstrate that it provides unique information, in addition to that of water column depth, ocean bottom pressure, and conductance. The RMSE of the GAM-based OHC estimates is reduced by another order of magnitude when the OHC in the upper 2000 m is included in the GAM (Figs. 8a,c). The RMSE of the GAM-based OHC anomaly estimates experiences an even more dramatic reduction upon

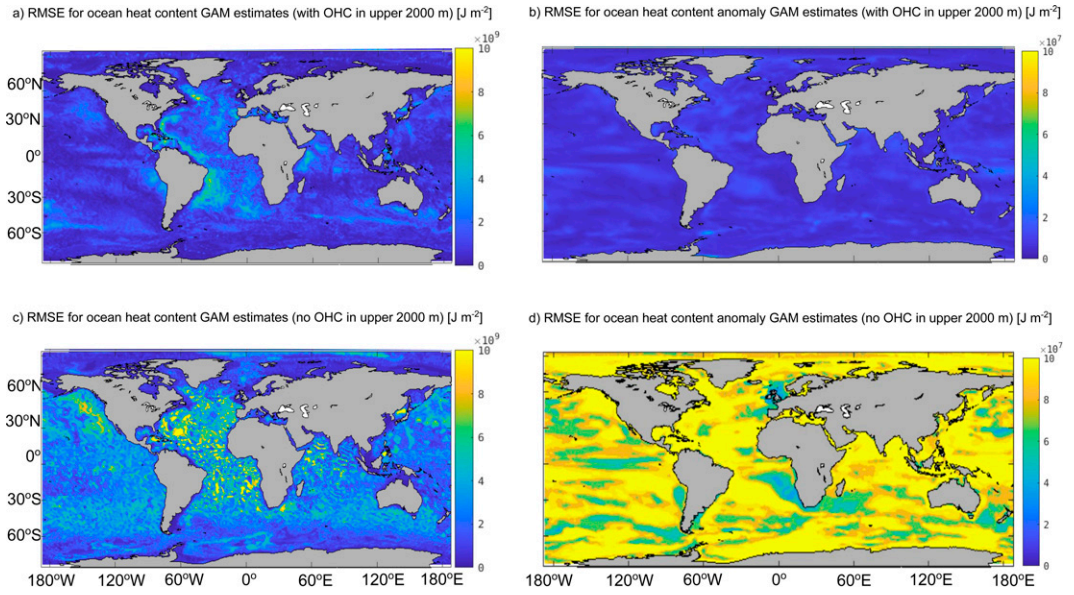

FIG. 8. Importance of including OHC over the top 2000 m in estimating OHC throughout the water column: shown are the RMSEs (a) using the GAM, $\mathrm{OHC} \sim f_1(\mathrm{SSH}') + f_2(p_b) + f_3(\Sigma) + f_4(H_0) + f_5(\mathrm{OHC}_{2000\,\mathrm{m}}) + g(\mathrm{SSH}', p_b, \Sigma, H_0, \mathrm{OHC}_{2000\mathrm{m}})$, over the first 9 years of ECCOv4r3 (1992–2000) and (c) for the GAM, $\mathrm{OHC} \sim f_1(\mathrm{SSH}') + f_2(p_b) + f_3(\Sigma) + f_4(H_0) + g(\mathrm{SSH}', p_b, \Sigma, H_0)$, over the first 9 years (1992–2000). Also shown is the same for OHC′ (b) using the GAM, $\mathrm{OHC}' \sim f_1(\mathrm{SSH}') + f_2(p_b') + f_3(\Sigma') + f_4(H_0) + f_5(\mathrm{OHC}'_{2000\,\mathrm{m}}) + g(\mathrm{SSH}', p_b', \Sigma', H_0, \mathrm{OHC}'_{2000\,\mathrm{m}})$, over the first 9 years of ECCOv4r3 (1992–2000) and (d) for the GAM, $\mathrm{OHC}' \sim f_1(\mathrm{SSH}') + f_2(p_b') + f_3(\Sigma') + f_4(H_0) + g(\mathrm{SSH}', p_b', \Sigma', H_0)$, over the first 9 years (1992–2000). Note that the color-bar scale is different in (b) and (d) from Figs. 5b, 5d, and 7b, 7d. The Black Sea and Caspian Sea are not included in ECCO.

inclusion of the OHC anomalies in the upper 2000 m (Figs. 8b,d). If full-depth conductance from in situ observations is incomplete or satellite-derived conductance is unavailable, it would be possible to use conductance over the upper 2000 m (e.g., from Argo) instead of conductance over the full water column. If conductance over the upper 2000 m is used instead of conductance over the full water column in the calculation for Figs. 8b and 8d, then the RMSE is globally 12.1% greater. These findings indicate that temperature and salinity data from the upper 2000 m from available measurements (e.g., Argo) should be included in the GAM's training data, but if full-depth conductance is available (e.g., from satellite, land, seafloor, ocean, and airborne measurements), then it is better to include both OHC in the upper 2000 m and full-depth conductance.

One practical consideration for our remote monitoring system for OHC is that we will not have global coverage of OHC observations through the water column, let alone in the upper 2000 m, with which to train our GAM. No assessment has been made or implied concerning the potential replacement of more conventional observations by magnetometric methods thus far; instead, we have performed analyses to examine what conductance data can add to other ongoing satellite and in situ data collections. There are other observational programs that can potentially add information about OHC magnitudes and anomalies. One such program, which can provide OHC over the entire water column, is the deep Argo program. We next sample the ECCOv4r3 output and train our GAM on about 400 (full-depth) deep Argo floats distributed throughout the ocean. We show Fig. 9, which frames the RMSE in terms of percentage of the 24-yr average OHC over 1992–2015 using two different spatiotemporal sampling schemes for the GAM's training data. As expected, having a certain number (~400) of horizontal locations sampled for the GAM's training data (Fig. 9a) results in a smaller-percentage RMSE than training the GAM on a small (~30) but increasing number of sampled horizontal locations up to the same number (~400) over time (Fig. 9). The RMSE of the OHC in the subpolar North Atlantic, region east of the Weddell Sea, and continental shelf regions each heavily depends upon the spatial sampling for a given time. Because our samples are only at locations where hypothetical deep Argo floats exist, these results suggest that a remote monitoring system for OHC will require more full-depth training data for the GAM than 400 locations. This motivates our use of ship-based hydrographic transect data.

b. *Remote monitoring prototype*

We next find combinations of transects that have been historically observed by ships (Fig. 1) that can train the GAM to have a minimal spatial RMSE in estimating global OHC. By iterating the GAM training and estimation steps with single transects, all combinations of pairs of transects, all combinations of triplets of transects, and so on until all historical transects are included in the training step, we find several optimal combinations of transects that can be used to minimize the global RMSE in estimating OHC at each location with zero

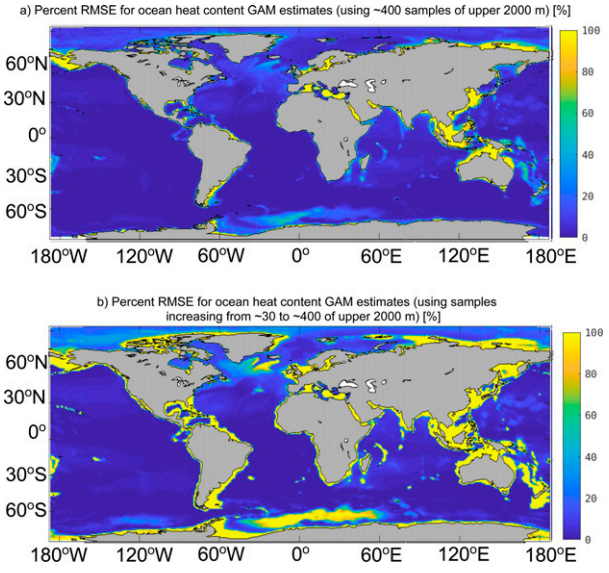

FIG. 9. Importance of spatial sampling of OHC over the top 2000 m: the percent RMSEs using the GAM, OHC ~ f_1 (SSH′) + f_2 (p_b) + f_3 (Σ) + $f_4(H_0)$ + $f_5(\mathrm{OHC}_{2000\ \mathrm{m}})$ + g(SSH′, p_b, Σ, H_0, $\mathrm{OHC}_{2000\mathrm{m}}$), over the final 15 years of ECCOv4r3 (2001–15), (a) trained on nearly 400 horizontal grid locations each year and (b) trained on an increasing number of horizontal grid locations over time (increasing from about 30 to nearly 400). The Black Sea and Caspian Sea are not included in ECCO.

measurement errors for one example month (April 1992). These combinations of transects can yield a spatial RMSE in estimating OHC of about 0.15%–0.25% with all of the predictors (top right in Fig. 10), given zero or very small measurement errors (an order of magnitude smaller than the ones we stated in the last paragraph of section 2d). The combination of transects mapped in Fig. 10 (bottom left) is an example of one that leads to minimal RMSE. Other combinations with transects in the Pacific Ocean can yield similar RMSE values, but those are not shown. The example shown in Fig. 10 includes most historical transect data in the Indian Ocean, several select hydrographic transects in the Atlantic Ocean (including the long-running AR07/OSNAP-West line), and nothing in the Pacific Ocean. Table 1 tabulates the RMSE using this example combination of transects, but with different combinations of predictors. Table 1 demonstrates that a predictor that has a smaller correlation with OHC reduces the RMSE by less than a predictor that has a higher correlation with OHC when added to the GAM.

The most important factors that determine the spatial RMSE are the variables included in the GAM (Table 1) and the variance in OHC used to train the GAM. The minimization of spatial RMSE using a GAM often requires training data that sufficiently span the range and domain of the statistical model (e.g., Trossman et al. 2011). The example shown in Fig. 10 satisfies that criterion, as there is an inverse relationship between the percent spatial RMSE of the GAM and the standard deviation of the OHC in the training (hydrographic transect) data per number of transects, regardless of whether errors in the satellite data are accounted for. This inverse relationship still holds when $\mathbf{T}_\sigma$ and other variables are excluded from the GAM (not shown), and with similar spatial RMSEs in estimated OHC (Table 1). The number of observations used to train the GAM as a function of time throughout the water column and only in the upper 2000 m are shown in Fig. 10 as reference for the accuracy of our method as a function of time (see later).

1) PRELIMINARY ERROR ANALYSIS

Next, we evaluate how the spatial RMSE for OHC estimates can be impacted by the presence of measurement errors in the satellite data. The percentage change in spatial RMSE due to a change in the ratio of the standard deviation of the added random noise to the mean value of the variable (i.e., percentage change in spatial RMSE times the signal-to-noise ratio) is quantified in Fig. 11. These sensitivities reflect how informative each variable is to the GAM. Here, we separate out the SSH anomalies and seafloor depth to highlight the sensitivity to the measurement errors of individual variables. We also include $\mathbf{T}_\sigma$ to demonstrate the relatively uninformative role that $\mathbf{T}_\sigma$ plays in the GAM.

Accounting for only one variable's measurement error, the sensitivity of the spatial RMSE to the level of noise is shown in Fig. 11a. This figure suggests that the spatial RMSE in OHC is most sensitive to conductance measurement errors when the other measurement errors are negligible and similarly sensitive to seafloor depth measurement errors when all other measurement errors are ignored. Accounting for all measurement errors simultaneously, the spatial RMSE in OHC is most sensitive to seafloor depth measurement errors (Fig. 11b). This is because the strongest correlation between OHC and any variable is between OHC and seafloor depth (Fig. 2a). However, because not all of the standard deviation levels we chose for measurement error magnitudes are well known, we further examine the sensitivity of spatial RMSE to measurement errors in the partial models included in Table 1. The spatial RMSE in OHC is most sensitive to bottom pressure—and similarly sensitive to conductance—in each of the partial models that excludes seafloor depth (Fig. 11c). When seafloor depth is included, the measurement errors associated with seafloor depth dominate the sensitivity in spatial RMSE to measurement errors. The measurement errors in each of the variables used in the GAM will be important to quantify if our technique is going to be practically applied to monitor OHC, but the variable that needs to be most accurately known is the one that changes the least and could be most well observed: the seafloor depths.

We perform the same hypothetical measurement error sensitivity analysis for Fig. 11b, except for OHC anomalies in Fig. 12. Here, we include all of the variables included in our final GAM specification: SSH anomalies, bottom pressure anomalies, conductance anomalies, and OHC anomalies in the upper 2000 m. We repeat the calculations by training the GAM in two different ways: one by only training on data that are available for the year in which a GAM estimate is being made (Fig. 12a) and one by training on all available data from

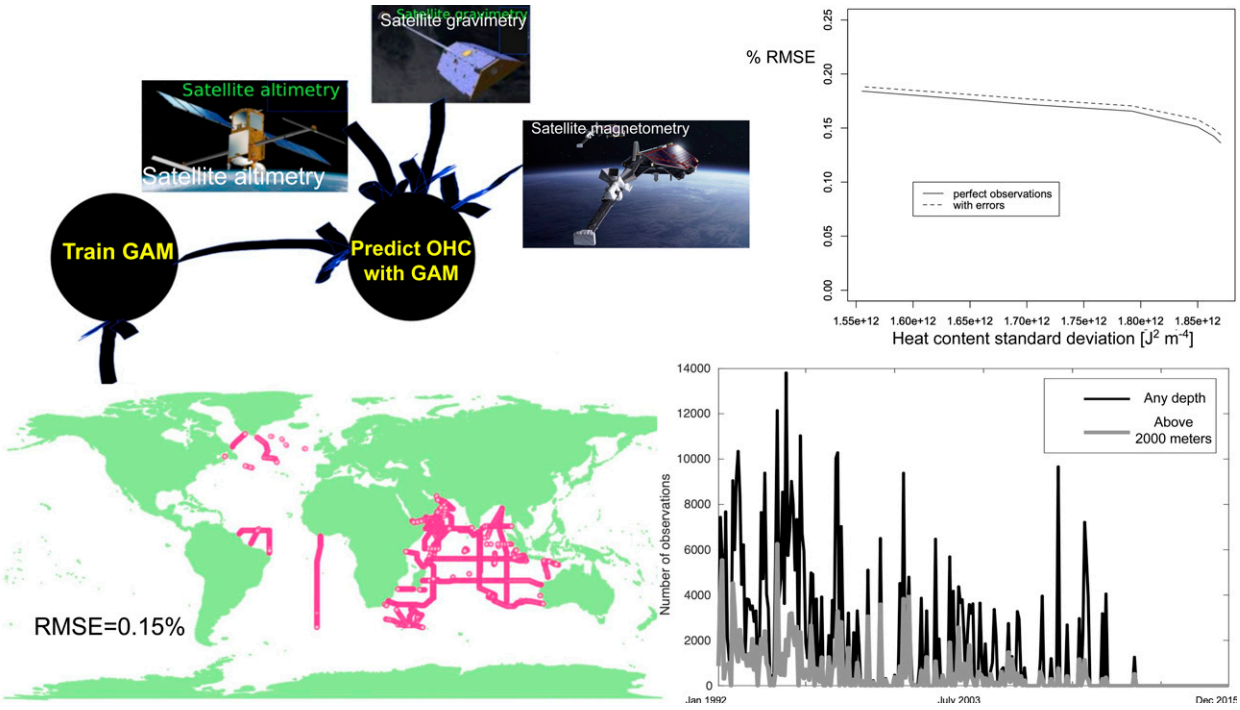

FIG. 10. (top left) Flowchart for how the remote monitoring system for OHC would work. First, a Generalized Additive Model (GAM) is trained using hydrographic transect observations of ocean heat content, sea surface heights, bottom pressure, depth-integrated conductivity (conductance), and seafloor depth at hydrographic transect locations. Then the GAM is used at every wet point of the World Ocean where satellite altimetry (sea surface heights), gravimetry (bottom pressure), and magnetometry (conductance) observations exist to estimate the OHC. (top right) Example relationships between the standard deviation of OHC from all training hydrographic transects per transect and the RMSE of the resulting GAM with and without considering errors in the satellite observations; these errors are an order of magnitude smaller than the ones we stated in the last paragraph of section 2d. (bottom right) The number of observations from all hydrographic transects shown in Fig. 1 at any depth (black curve) and above 2000 m depth (gray curve). (bottom left) An example combination of hydrographic transect locations that determines one of the smallest RMSEs in estimated OHC, as determined by random sampling of every combination of hydrographic transects, which includes A01W, A14, AR04, AR07E, I01E, I01W, I02E, I02W, I03, I04, I05E, I07N, I09N, IR01W, IR03, IR04, IR06, ISS1, and ISS2.

any time (Fig. 12b). The former results in smaller sensitivities to measurement errors than the latter, but the relative sensitivities to individual variables are about the same, whether the GAM is trained using one set of data or another. The sensitivities to bottom pressure anomaly measurement errors are largest, followed by sensitivities to measurement errors in OHC anomalies in the upper 2000 m and then the sensitivities to conductance anomaly measurement errors. The sensitivities to SSH anomaly measurement errors are negligible compared to the others. Because of the large sensitivity to bottom pressure anomalies, it should be noted that Schindelegger et al. (2021) estimated the errors in bottom pressure from satellite gravimetry are typically about 2-cm-equivalent sea level, which is about two orders of magnitude smaller than the amount we perturbed the bottom pressure anomaly measurement errors. So while the sensitivities are largest to bottom pressure anomaly measurement errors, there is only a percent or two RMSE increase due to the measurement errors of bottom pressure anomalies from satellite gravimetry. More concerning is the sensitivity to measurement errors in OHC anomalies in the upper 2000 m and whether the actual conductance anomaly measurement errors from satellite magnetometry are larger than our hypothetical values.

2) TRADE-OFF BETWEEN OHC VARIANCE IN TRAINING DATA AND FREQUENCY OF GAM TRAINING UPDATES

Training the GAM using the transects shown in Fig. 10 and applying the GAM with measurement errors in the predictors to estimate OHC as above—but for each month as a function of time—we can accurately estimate OHC relative to the ECCOv4r3 output. Here, we exclude $\mathbf{T}_\sigma$ due to potential challenges with observing the velocities along hydrographic transects concurrently with the other variables and its relatively small impact on RMSE. Figure 13a shows the temporally averaged residuals of the GAM estimates at each location of the ocean over 1992–2015. The GAM-based OHC estimates are too small for each month over 1992–2015 primarily because the Arctic Ocean has not been sampled in the training data and because the global relationships between each of the predictors and OHC are different in the Arctic compared to the rest of the world. These temporal residuals are fairly constant over time in the shelf regions, but vary dramatically

TABLE 1. The globally area-averaged percent root-mean-square errors (RMSE) between the ECCO-derived ocean heat content (OHC) and the Generalized Additive Model (GAM)-derived OHC for many different GAMs. The percent RMSE in OHC is computed by calculating the root-mean-square error between the ECCO-derived OHC and the GAM-derived OHC and dividing by the area-averaged ECCO-derived OHC ($\approx 4.1 \times 10^{12}$ J m^{-2}). No measurement errors were accounted for in these calculations so perfect information along each of the randomly chosen hydrographic transects and inferred from the satellites is assumed. This example uses data to predict OHC during April 1992. The smoother functions, $f_n(\cdot)$, are different in each row and for different $n = 1, \ldots, 5$. The tensor product functions, $g(\cdot)$, are also different in each row.

Terms included in GAM: OHC = f_0 + …	Percent RMSE in OHC
$f_1(\mathrm{SSH}')$	43.6%
$f_1(p_b)$	5.51%
$f_1(\Sigma)$	5.92%
$f_1(H_0)$	0.60%
$f_1(\lvert\mathbf{T}_\sigma\rvert)$	41.4%
$f_1(\mathrm{SSH}') + f_2(p_b)$	6.12%
$f_1(\mathrm{SSH}') + f_2(p_b) + g(\mathrm{SSH}', p_b)$	6.10%
$f_1(\mathrm{SSH}') + f_2(p_b) + f_3(\Sigma)$	1.92%
$f_1(\mathrm{SSH}') + f_2(p_b) + f_3(\Sigma) + g\ (\mathrm{SSH}', p_b, \Sigma)$	0.93%
$f_1(\mathrm{SSH}') + f_2(p_b) + f_3(\Sigma) + f_4(H_0)$	0.21%
$f_1(\mathrm{SSH}') + f_2(p_b) + f_3(\Sigma) + f_4(H_0) + g(\mathrm{SSH}', p_b, \Sigma, H_0)$	0.15%
$f_1(\mathrm{SSH}') + f_2(p_b) + f_3(\Sigma) + f_4(H_0) + f_5(\lvert\mathbf{T}_\sigma\rvert)$	0.21%
$f_1(\mathrm{SSH}') + f_2(p_b) + f_3(\Sigma) + f_4(H_0) + f_5(\lvert\mathbf{T}_\sigma\rvert) + g(\mathrm{SSH}', p_b, \Sigma, H_0, \lvert\mathbf{T}_\sigma\rvert)$	0.15%

over time in the Arctic Ocean, as indicated by the temporal standard deviations of the residuals (Fig. 13b). The temporal RMSE becomes strongly correlated (0.9998) with the seafloor depth over long (>10 years) time periods (Fig. 13c), suggesting that OHC could be remotely monitored over decadal time scales with a predictable RMSE. However, the biases in the global OHC estimates with the GAM are not highly predictable for each month, as evidenced by how the temporal standard deviation of the residuals (Fig. 13b) dominate the bias contribution to the temporal RMSE (Fig. 13c) and by the fair correlation between the GAM-based global OHC estimates and the ECCOv4r3 global OHC estimates (0.5). The correlation between the GAM-based global OHC estimates and the ECCOv4r3 global OHC estimates increases to nearly 0.6 when the Arctic Ocean is excluded. Only coastal regions have statistically significant differences between the GAM-based estimates and the ECCOv4r3 estimates of OHC (magenta crosses in Fig. 13c) and these regions have the smallest OHC.

There is an optimal balance between the variance of OHC data used to train the GAM and the frequency over which the GAM is trained. Figures 13d–f demonstrate that using hydrographic transects only for the year over which the GAM-based estimates are being applied does not necessarily reduce the temporal residuals, standard deviation of the residuals, or RMSE. The residuals are largest in the same locations, whether we use all of the transects shown in Fig. 10 or only the transects for the year over which the GAM-based estimates are being applied (typically 1–10 transects per year). However, the (relatively small) differences between the residuals between use of these two training datasets are incoherent in their spatial patterns (Fig. 13d, which should be noted is an order of magnitude smaller in scale than Fig. 13a). The standard deviations of the residuals and the temporal RMSE also look similar, regardless of which training dataset is used, but both the standard deviations of the residuals (Fig. 13e) and the temporal RMSE (Fig. 13f) are larger in open ocean regions when the transects for the year over which the GAM-based estimates are being applied are used. This is an example of how the number of transects used to train the GAM can be more important for accuracy of GAM-based estimates than use of the relevant times to train the GAM, but the opposite can also happen (e.g., if less transects were used than shown in Fig. 10).

A similar exercise to the one shown in Fig. 13 is repeated for OHC anomalies with all significant predictors (SSH′, Σ', p_b', and $\mathrm{OHC}'_{2000\mathrm{m}}$). We account for possible measurement errors here by randomly sampling values from normal distributions with standard deviations equal to their hypothetical measurement errors. We present the (percent) RMSE over time (Fig. 14) in order to highlight the potential accuracy to monitor OHC anomalies using our method. When all of the transects collected over any year (Fig. 1) are used to train a GAM that is never retrained, the RMSE is typically of order $\sim 10^9$–10^{10} J m^{-2} (Fig. 14a), which can be less than 100% of the OHC anomaly magnitudes outside of high-latitude regions (blue colors in Fig. 14b). The polar oceans have greater than 1000% RMSE. This makes OHC anomalies potentially detectable using our method on a long-term (multiyearly or decadal) basis. Outside of the polar regions, the GAM could be trained on the ship-based full-column hydrographic transect observations that have already been collected over many decades as well as available altimetry, gravimetry, magnetometry, and in situ upper 2000 m fine structure data.

We next show that there is also a trade-off between the training data's OHC variance and GAM training frequency when predicting OHC anomalies. We do this by comparing the OHC anomalies from the GAM trained only once on all available ship-based hydrographic transect data with the

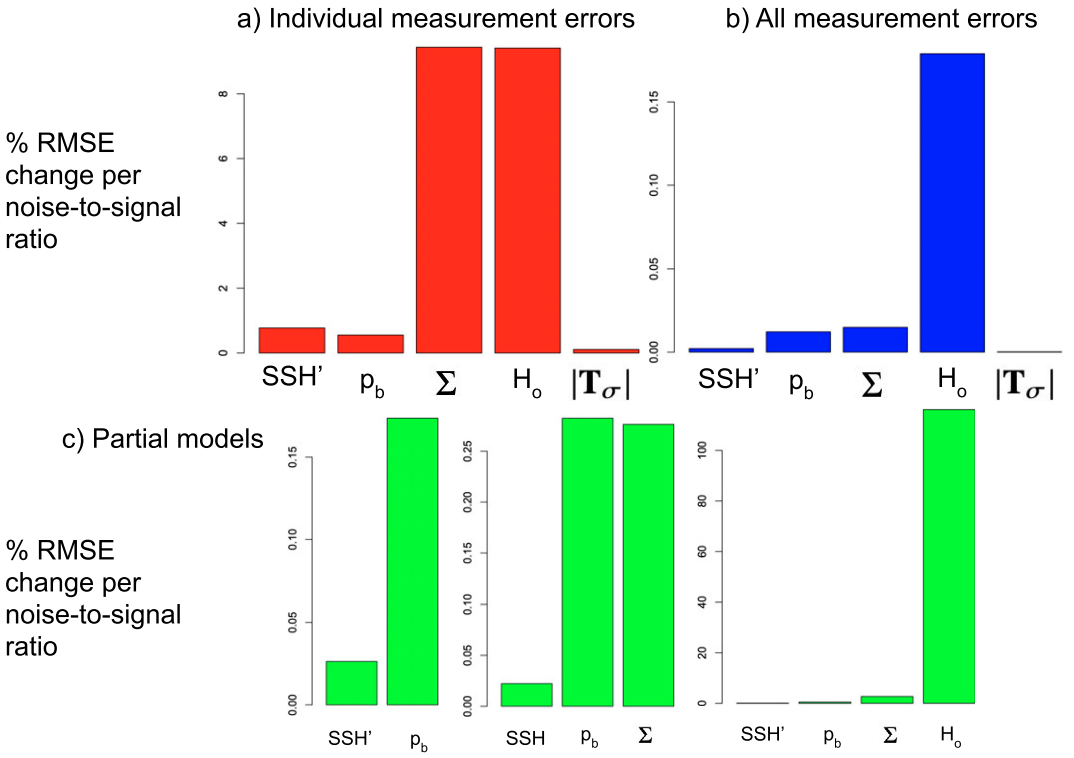

FIG. 11. Some sensitivities of the OHC RMSE due to measurement errors. Shown are (a) the percent RMSE increase per noise-to-signal ratio due to measurement error for each individual variable's contribution (one at a time with no other errors), and (b) as in (a), but accounting for all errors at once. (c) As in (b), but for reduced models (those specified in Table 1): $\mathrm{OHC} \sim f_1(\mathrm{SSH}') + f_2(p_b) + g(\mathrm{SSH}', p_b)$, $\mathrm{OHC} \sim f_1(\mathrm{SSH}') + f_2(p_b) + f_3(\Sigma) + g(\mathrm{SSH}', p_b, \Sigma)$, $\mathrm{OHC} \sim f_1(\mathrm{SSH}') + f_2(p_b) + f_3(\Sigma) + f_4(H_0) + g(\mathrm{SSH}', p_b, \Sigma, H_0)$, and $\mathrm{OHC} \sim f_1(\mathrm{SSH}') + f_2(p_b) + f_3(\Sigma) + f_4(|\mathbf{T}_\sigma|) + g(\mathrm{SSH}', p_b, \Sigma, |\mathbf{T}_\sigma|)$ from left to right. The units of SSH′ are meters, of p_b are bars, of Σ are S, of $|\mathbf{T}_\sigma|$ are S m s^{-1}, and of H_0 are meters.

OHC anomalies from GAMs that are retrained based on only the data available for a given year in which a prediction is made. If we retrain our GAM for each year where ship-based hydrographic transect data exist and supplement those with only the concomitant satellite-derivable data in the training dataset, then the RMSE is generally a similar order of magnitude (Fig. 14c). In high-latitude regions, this yearly retraining procedure allows us to detect OHC anomalies on a multiyear basis because the RMSE is less than 100% (Fig. 14d). However, outside of the polar oceans, OHC anomalies using the yearly retraining procedure have larger RMSE than using a GAM that is trained only once on all available historical data. The physical relationship between our predictors and full-column OHC anomalies is likely the most important information to provide the GAM, but the GAM needs to be retrained on short-term (e.g., yearly) bases to account for the relatively quickly changing polar oceans. The RMSE is about 7.7% larger, on global average, when the strategic set of transects (Fig. 10) are sampled in January of each year and used to train the GAM only once instead of training the GAM only once (Fig. 14a) on all available historically sampled transects (Fig. 1). The RMSE is about an order of magnitude larger when the GAM is retrained each year on the strategic set of transects that are sampled in January, for example, of each year (not shown) instead of retraining the GAM on any available historically sampled transects from a given year (Fig. 14c). Due to the trade-off highlighted here between the training data's OHC variance and GAM training frequency, we suggest that OHC anomalies can be monitored outside of the polar oceans by training the GAM only once on all available ship-based coverage and in the polar oceans by retraining the GAM over shorter periods of time (e.g., annual time scales).

The accuracy in estimating OHC anomalies throughout the water column and the accuracy of the same below 2000 m are highlighted with time series of the globally averaged percentage differences between the GAM-based estimates and the ECCOv4r3 values in Fig. 15. When the GAM is trained on all available transects, at least once per year the percentage differences are less than 100% and within the confidence intervals of the globally averaged estimates to the 95% level for the estimates throughout the water column (Fig. 15a) and below 2000 m depth (Fig. 15b). Each of these time series has a seasonal cycle amplitude of tens of percent error because of the times when the transects were sampled within particular years. The percentage errors reverse in sign and drift toward negative values in the 2000s (Fig. 15a), when there are less

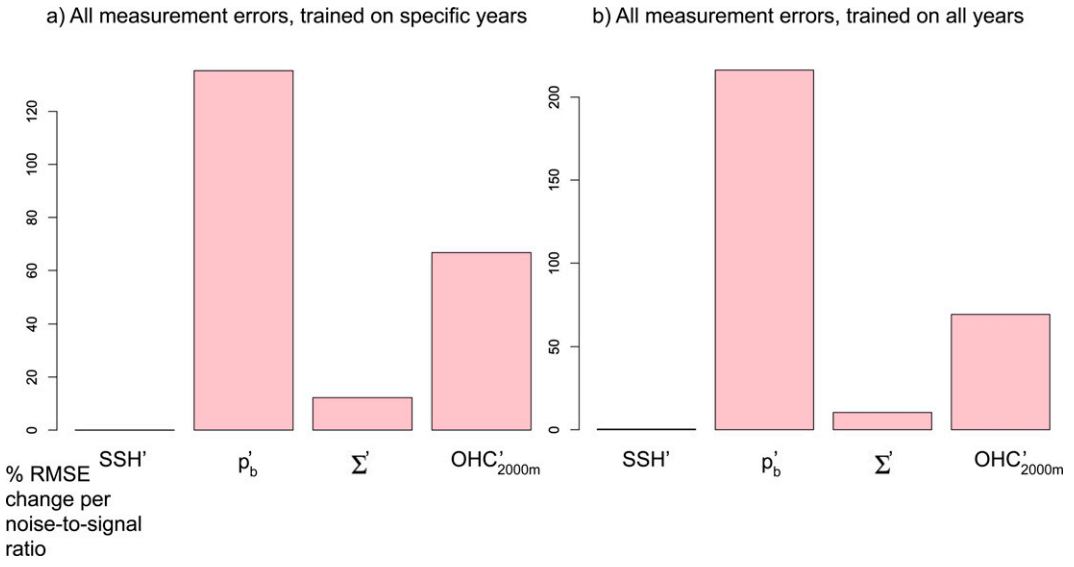

FIG. 12. Some sensitivities of the OHC anomalies RMSE due to measurement errors. Shown are the percent RMSE increase per noise-to-signal ratio due to measurement error for each individual variable's contribution accounting for all errors at once, like the calculations for Fig. 11b but for OHC anomalies. (a) The sensitivities for each variable's measurement error when the GAM for OHC anomalies is trained only on data for the year when the estimate is being made. (b) As in (a), but the GAM for OHC anomalies is trained on all available data from any time. The units of SSH′ are meters, of p_b' are bars, of Σ′ are S, and of $\mathrm{OHC}'_{2000\mathrm{m}}$ are J m^{-2}.

hydrographic transect data (Fig. 10). When the number of observations used to retrain the GAM for a given year is sufficient, the percentage differences are less than 100% and the 95% confidence intervals always overlap 0% error in the GAM-based estimates of globally averaged OHC anomalies, whether the OHC anomalies are throughout the water column (Fig. 15c) or only below 2000 m (Fig. 15d). In 1996 and 2000, for example, there were too few data to retrain an accurate GAM at the beginning of those years (Figs. 15c,d). In 1999, there were too few data below 2000 m to retrain an accurate GAM at the beginning of that year (Fig. 15d). The correlation between the absolute value of the errors in the globally averaged OHC anomalies in Figs. 15c and 15d and the number of observations used to retrain the GAM each year over the full water column (below 2000 m) is −0.13 (−0.11) over 1992–2015. Each correlation is closer to −0.2 over 1992–2000 because of the relative dearth of training data post-2000 (Fig. 10). The time series shown in Fig. 15 corroborate how the accuracy of our method is sensitive to the spatiotemporal sampling of ship-based campaigns.

4. Conclusions

Using the output of an ocean state estimate (ECCO), we trained a statistical model (GAM) on variables that can be monitored remotely (SSH and p_b), can be potentially monitored remotely (Σ), have time-invariant estimates (H_0), and in situ data that are routinely relayed via satellites ($\mathrm{OHC}_{2000\mathrm{m}}$) to compute root-mean-square errors (RMSEs) in OHC. We repeated these exercises for OHC anomalies using SSH, p_b, Σ, and $\mathrm{OHC}_{2000\mathrm{m}}$ anomalies. Practical considerations require that the GAM is trained on available ship-based hydrographic transect data as well as potentially deep Argo measurements because these are our only full-depth measurements of OHC. Our focus was on the 1992–2000 time period because the RMSE is largest for this period; if the RMSE is acceptable for this time period, then we can use our method anytime between 1992 and 2015.

We assessed the importance of particular predictors, observational coverage, and measurement errors. Our interest in Σ comes from the fact that it is a full-depth integrated parameter and can potentially be remotely observed over the globe (even under ice) using satellite magnetometry. We highlighted the importance of Σ and $\mathrm{OHC}_{2000\mathrm{m}}$ as predictors when trying to reduce the RMSE, particularly to the level of temporal variations over short time intervals. We also showed that increasing the number of deep Argo-like floats sampling the entire water column's OHC will reduce the RMSE in OHC. We then trained the GAM on hydrographic transects, and demonstrated that a GAM can be used to accurately monitor global OHC to within about 0.15% RMSE on yearly time scales, assuming perfect information (i.e., no measurement errors and no sampling/aliasing/retrieval problems). This is not sufficiently accurate to monitor the changes in OHC over short time intervals, particularly when measurement errors are accounted for, even when global satellite observational coverage is attainable. For OHC, measurement errors associated with seafloor depths were shown to dominate all others for the variable the machine learning algorithm is most sensitive to. However, we were able to detect OHC anomalies using our method to within 100% RMSE almost everywhere using a combination of approaches to training the GAM, even when we accounted for hypothetical

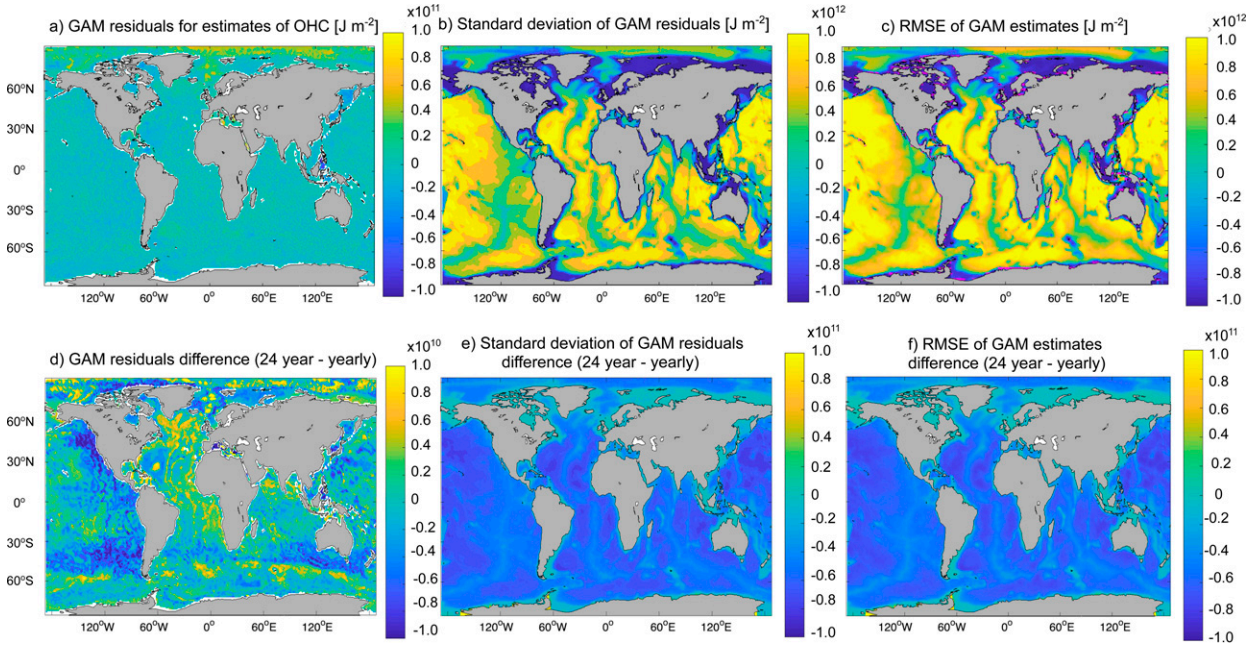

FIG. 13. Including measurement errors in the data used to plug into the GAM and the transects shown in Fig. 10 for training data, shown are maps of (a) the temporally averaged residuals from the GAM-estimated OHC (units: J m^{-2}), (b) the temporal standard deviations of the residuals from the GAM-estimated OHC (units: J m^{-2}), and (c) the temporal RMSEs of the GAM-estimated OHC (units: J m^{-2}). Magenta crosses are shown in (c) wherever the GAM-estimated OHC is statistically significantly different from the ECCO-estimated OHC to the 95% confidence level, using 1.96 times the standard errors computed by the GAM as the half-width of the 95% confidence intervals. Also shown are the differences between the (d) temporally averaged residuals, (e) temporal standard deviations of the residuals, and (f) temporal RMSEs of the GAM-estimated OHC when trained on all transects shown in Fig. 10 and when trained only on transects from a given year for which the estimates are made ("yearly"). Yellow colors in (d)–(f) mean that values are greater using all transects shown in Fig. 10 and blue colors in (d)–(f) mean that values are greater using only transects from the given year for which the estimates are made. Note that the color-bar scale in (d) maxes out at an order of magnitude less than that in (a) to emphasize how small the bias is relative to temporal standard deviation of the errors. The Black Sea and Caspian Sea are not included in ECCO.

measurement errors (see Figs. 14b,d). The measurement errors associated with bottom pressure anomalies were shown to dominate all others. Given the small uncertainties in bottom pressures derived from observations (e.g., Schindelegger et al. 2021), these results suggest that OHC anomalies could potentially be monitored—over multiyear time scales—using our proposed remote monitoring method.

We showed that there is a trade-off between the variance in OHC (anomalies) in the training data for the GAM and the frequency of training the GAM, whether the OHC (anomalies) are throughout the water column or below 2000 m depth. To demonstrate this, we compared the RMSE from GAM-based predictions of OHC and OHC anomalies in two configurations: 1) the GAM is only trained once based on all available data and 2) the GAM is trained on only available data of a given year when a prediction is made. OHC anomalies could be monitored using the GAM with all available historical ship-based training data outside of the polar oceans. However, retraining the GAM over short (e.g., yearly) time scales could reduce the RMSE in the OHC anomalies in some regions, despite the sparseness of training data (with low OHC variance). While physical principles primarily dictate the relationships between each of the predictors and OHC anomalies, updates to the GAM are needed in regions of the ocean (e.g., polar oceans) where there are relatively quick changes in the spatial distribution of OHC anomalies. Observational campaigns could be designed to specifically minimize the RMSE in GAM-based predictions of OHC anomalies by sampling a sufficient amount of OHC anomaly variance over short time scales. These campaigns could similarly provide in situ conductance data for validation/calibration of remote conductance estimates.

The RMSE for OHC estimates can further increase due to difficulties with constraining target variables from satellite data, incomplete sampling, and aliasing. For example, there is currently no observational data product for Σ from satellite magnetometry or other magnetic data. Because of this, we discussed the value of Σ for estimating OHC separately and did not attempt to estimate OHC (anomalies) from observations using our method. It would also be useful for full-depth ship-based conductivity–temperature–depth (CTD) measurements to report their measured conductivity because full-depth conductance should be included in the training data for the approach we proposed in the current study. Also, salinity data have at least as large normalized measurement errors as conductivity because of error propagation of both temperature and conductivity measurements and because temperature

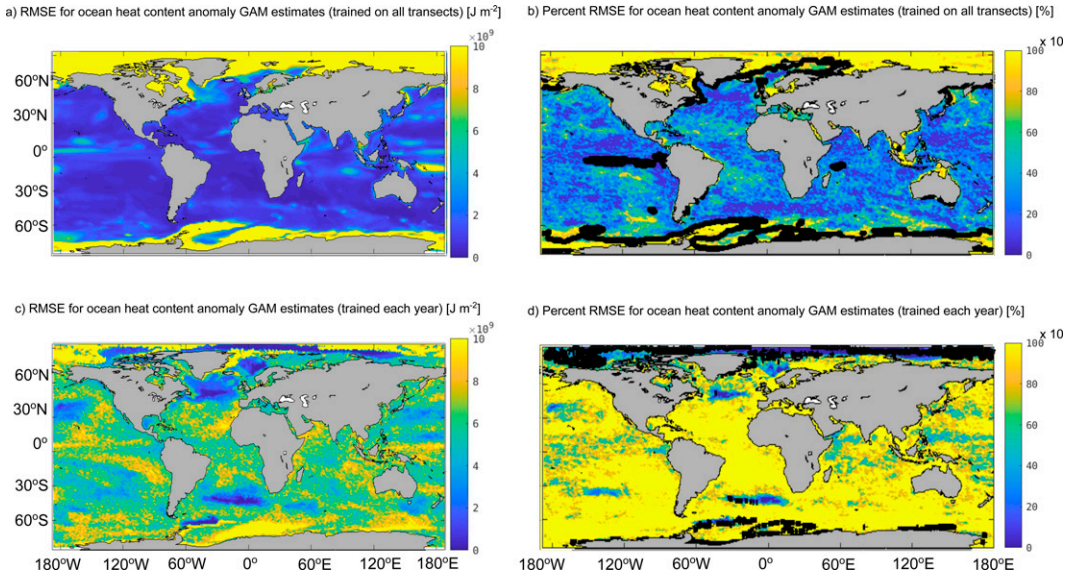

FIG. 14. Trade-off between OHC anomaly variance in training data and time period length: the (percent) RMSEs using the GAM, $\mathrm{OHC}' \sim f_1(\mathrm{SSH}') + f_2(p_b') + f_3(\Sigma') + f_4(\mathrm{OHC}'_{2000\mathrm{m}}) + g(\mathrm{SSH}', \mathrm{OHC}'_{2000\mathrm{m}})$, over the entire time period of ECCOv4r3 (1992–2015), (a),(b) trained on all available transects, regardless of time and (c),(d) trained on only the transects sampled for the current year in which a prediction is made. (a),(c) RMSEs and (b),(d) percent RMSEs. Areas where the GAM-based estimates of the OHC anomalies are statistically insignificant to the 95% level from the ECCOv4r3 OHC anomalies are outlined in thick black contours in (b) and (d). Hypothetical measurement errors specified in the main text are randomly sampled and added to the predictors' values. The Black Sea and Caspian Sea are not included in ECCO.

and conductivity measurements may not correspond to the same water parcel as a result of unequal sensor response times. Because Σ anomalies over the upper 2000 m provides some useful information about OHC anomalies, we suggest it would be useful for Argo to also report conductivity as measured rather than calculated from temperature and salinity—after being calculated from the measured conductivity. However, we calculated a >10% increase in RMSE in the OHC anomalies when Σ is approximated with Argo float–like data over the upper 2000 m alone, further motivating a method to constrain Σ from satellite magnetometry and other magnetic data. Further, some satellite data have experienced time periods with less-than-global coverage. For instance, throughout much of the 1990s, SSH was observed using satellite altimetry between 66°S and 66°N and not in polar regions. Only including SSH between these latitudes in the training data increases RMSE of OHC estimate by less than 0.1%. These factors will need to be accounted for if our proposed technique is going to be used to monitor OHC.

Some future research directions could refine our proposed remote monitoring system, which could be useful for either monitoring Earth's energy imbalance or (hurricane) forecasting purposes. First, measurement errors for each of the variables included in the GAM need to be refined. Second, additional training observations could improve the accuracy of the OHC estimates. Deep Argo (e.g., Gasparin et al. 2020), Deepglider (Osse and Eriksen 2007), SMART cables (Howe et al. 2019), and Arctic hydrographic transects could make a valuable additions to the hydrographic training data used here, which can be explored in a follow-up implementation study. Supplementing ship-based observations with hydrographic data from deep Argo could increase the variance in OHC in the training data and therefore reduce the RMSE of the OHC estimates. Third, there exists an optimal combination of factors to arrive at the most realistic monitoring method for OHC anomalies. For example, the GAM-training frequency given the existing hydrographic transects and other full-depth data is likely somewhere between once per 24 years and every year. We may need to more explicitly account for a seasonal cycle in the algorithm to estimate OHC anomalies on sub-multiyear time scales. This will not be feasible regionally by simply including time stamps in the GAM because of aliasing in many regions (e.g., the Arctic), but our method could be improved by supplementing it with the seasonal cycles of OHC anomalies from other methods, such as data assimilation—or some proxy-based estimates. There is also an optimal blend of our proposed method with other existing methods, which future studies will determine. The most accurate use of the method proposed here to remotely monitor OHC is to apply it to locations where observations do not exist over the full water column and data assimilation models do not accurately calculate the OHC. Finally, the opportunity for extracting Σ is currently being explored by several research groups. Tidal magnetic fields from Swarm could potentially be used as a substitute for Σ, given that electrical conductivity of the ocean allows

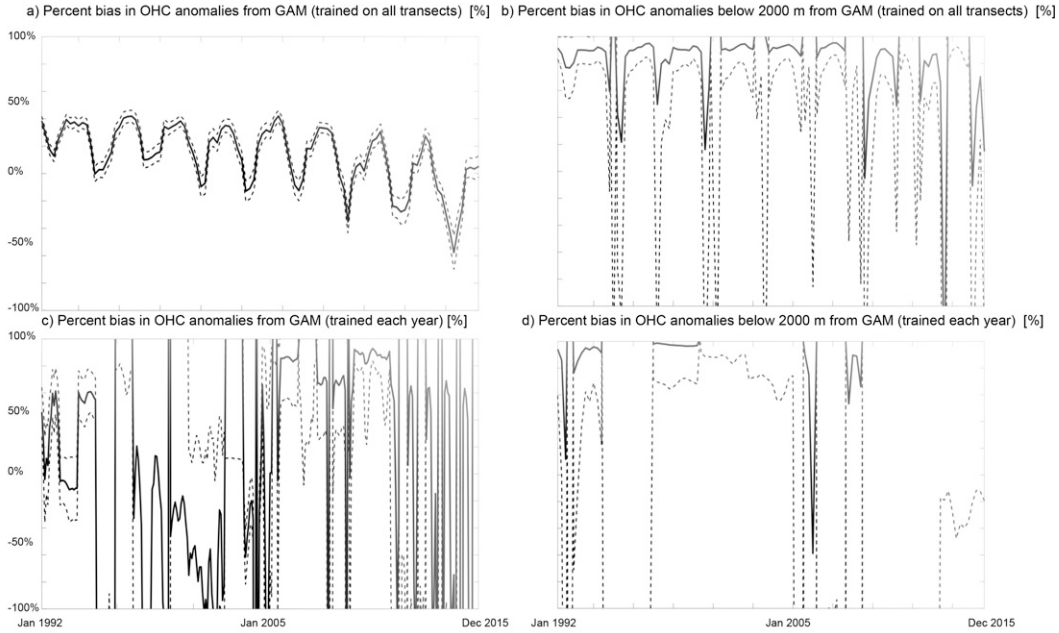

FIG. 15. Accuracy of GAM-estimated OHC anomalies throughout the water column vs below 2000 m: black curves show percentage differences between the area-weighted, globally averaged OHC anomalies using the GAM, $OHC' \sim f_1(SSH') + f_2(p'_b) + f_3(\Sigma') + f_4(OHC'_{2000m}) + g(SSH', OHC'_{2000m})$, minus the area-weighted, globally averaged OHC anomalies for the full length of ECCOv4r3 (1992–2015) and the confidence intervals to the 95% level based on the standard errors (dashed curves) from the GAM (assuming they are normally distributed). (a),(c) OHC anomalies throughout the water column and (b),(d) OHC anomalies below the top 2000 m. Plots use (a),(b) GAMs trained on all available transects, regardless of time, and (c),(d) GAMs trained on only the transects sampled for the current year in which a prediction is made.

for the generation of ocean tidal magnetic fields, there is a strong relationship between Σ (anomalies) and OHC (anomalies), and M_2 tidal magnetic fields can potentially predict global OHC in at least the upper 2000 m (Irrgang et al. 2019). A future study will ultimately make use of all available observations from at least as far back in time as 2002 to derive a time series and map of OHC with uncertainties and compare with other existing methods to estimate OHC. Furthermore, future investigations could inspect the potential to monitor freshwater fluxes into the ocean, infer ocean heat transport, and/or provide directional information about tsunami flow (e.g., Tyler 2005; Manoj et al. 2010; Minami et al. 2015; Minami 2017) using a similar approach to the one used in the present study.

Acknowledgments. The authors appreciate the comments provided by the anonymous reviewers. David Trossman acknowledges support by the NASA SLCT Grant 80NSSC17K0558 and NASA PO Pilot Grant 80NSSC19K1664 at The University of Texas at Austin as well as computing resources provided by the Texas Advanced Computing Center (TACC) at The University of Texas at Austin for the ECCO simulations. Robert Tyler acknowledges support by NASA PO Pilot funding. The model data used in this study are available at https://zenodo.org/deposit/3897466.

REFERENCES

Abraham, J. P., and Coauthors, 2013: A review of global ocean temperature observations: Implications for ocean heat content estimates and climate change. *Rev. Geophys.*, **51**, 450–483, https://doi.org/10.1002/rog.20022.

Adcroft, A., and J.-M. Campin, 2004: Rescaled height coordinates for accurate representation of free-surface flows in ocean circulation models. *Ocean Modell.*, **7**, 269–284, https://doi.org/10.1016/j.ocemod.2003.09.003.

——, C. Hill, and J. Marshall, 1997: The representation of topography by shaved cells in a height coordinate model. *Mon. Wea. Rev.*, **125**, 2293–2315, https://doi.org/10.1175/1520-0493(1997)125<2293:ROTBSC>2.0.CO;2.

ATOC Consortium, 1998: Ocean climate change: Comparison of acoustic tomography, satellite altimetry, and modeling. *Science*, **281**, 1327–1332, https://doi.org/10.1126/science.281.5381.1327.

Balmaseda, M. A., K. E. Trenberth, and E. Källén, 2013: Distinctive climate signals in reanalysis of global ocean heat content. *Geophys. Res. Lett.*, **40**, 1754–1759, https://doi.org/10.1002/grl.50382.

Boyer, T., and Coauthors, 2016: Sensitivity of global upper-ocean heat content estimates to mapping methods, XBT bias corrections, and baseline climatologies. *J. Climate*, **29**, 4817–4842, https://doi.org/10.1175/JCLI-D-15-0801.1.

Buckley, M. W., R. M. Ponte, G. Forget, and P. Heimbach, 2015: Determining the origins of advective heat transport

convergence variability in the North Atlantic. *J. Climate*, **28**, 3943–3956, https://doi.org/10.1175/JCLI-D-14-00579.1.

Campin, J.-M., A. Adcroft, C. Hill, and J. Marshall, 2004: Conservation of properties in a free surface model. *Ocean Modell.*, **6**, 221–244, https://doi.org/10.1016/S1463-5003(03)00009-X.

Chambers, D. P. A., N. Cazenave, H. Champollion, W. Dieng, R. Llovel, and R. Forsberg, 2017: Evaluation of the global mean sea level budget between 1993 and 2014. *Surv. Geophys.*, **38**, 309–327, https://doi.org/10.1007/s10712-016-9381-3.

Chaudhuri, A. H., R. M. Ponte, G. Forget, and P. Heimbach, 2013: A comparison of atmospheric reanalysis surface products over the ocean and implications for uncertainties in air–sea boundary forcing. *J. Climate*, **26**, 153–170, https://doi.org/10.1175/JCLI-D-12-00090.1.

Cheng, L., K. E. Trenberth, J. Fasullo, T. Boyer, J. Abraham, and J. Zhu, 2017: Improved estimates of ocean heat content from 1960 to 2015. *Sci. Adv.*, **3**, e1601545, https://doi.org/10.1126/sciadv.1601545.

Church, J. A., and Coauthors, 2011: Revisiting the Earth's sea-level and energy budgets from 1961 to 2008. *Geophys. Res. Lett.*, **38**, L18601, https://doi.org/10.1029/2011GL048794.

——, and Coauthors, 2013: Sea level change. *Climate Change 2013, The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 1137–1216.

Colosi, J. A., and W. Munk, 2006: Tales of the venerable Honolulu tide gauge. *J. Phys. Oceanogr.*, **36**, 967–996, https://doi.org/10.1175/JPO2876.1.

Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, https://doi.org/10.1002/qj.828.

Desbruyères, D. G., S. G. Purkey, E. L. McDonagh, G. C. Johnson, and B. A. King, 2016: Deep and abyssal ocean warming from 35 years of repeat hydrography. *Geophys. Res. Lett.*, **43**, 10 356–10 365, https://doi.org/10.1002/2016GL070413.

Domingues, C. M., J. A. Church, N. J. White, P. J. Gleckler, S. E. Wijffels, P. M. Barker, and J. R. Dunn, 2008: Improved estimates of upper-ocean warming and multi-decadal sea level rise. *Nature*, **453**, 1090–1093, https://doi.org/10.1038/nature07080.

Durack, P. J., and S. E. Wijffels, 2010: Fifty-year trends in global ocean salinities and their relationship to broad-scale warming. *J. Climate*, **23**, 4342–4362, https://doi.org/10.1175/2010JCLI3377.1.

Dushaw, B. D., and Coauthors, 2009: A decade of acoustic thermometry in the North Pacific Ocean. *J. Geophys. Res.*, **114**, C07021, https://doi.org/10.1029/2008JC005124.

Fasullo, J. T., and P. R. Gent, 2017: On the relationship between regional ocean heat content and sea surface height. *J. Climate*, **30**, 9195–9211, https://doi.org/10.1175/JCLI-D-16-0920.1.

Forget, G., J.-M. Campin, P. Heimbach, C. N. Hill, R. M. Ponte, and C. Wunsch, 2015: ECCO version 4: An integrated framework for nonlinear inverse modeling and global ocean state estimation. *Geosci. Model Dev.*, **8**, 3071–3104, https://doi.org/10.5194/gmd-8-3071-2015.

Fournier, S., T. Lee, X. Wang, T. W. K. Armitage, O. Wang, I. Fukumori, and R. Kwok, 2020: Sea surface salinity as a proxy for Arctic Ocean freshwater changes. *J. Geophys. Res. Oceans*, **125**, e2020JC016110, https://doi.org/10.1029/2020JC016110.

Fukumori, I., O. Wang, I. Fenty, G. Forget, P. Heimbach, and R. M. Ponte, 2017: ECCO version 4 release 3. JPL Rep., 10 pp., http://hdl.handle.net/1721.1/110380.

Garry, F. K., E. L. McDonagh, A. T. Blaker, C. D. Roberts, D. G. Desbruyères, E. Frajka-Williams, and B. A. King, 2019: Model-derived uncertainties in deep ocean temperature trends between 1990 and 2010. *J. Geophys. Res. Oceans*, **124**, 1155–1169, https://doi.org/10.1029/2018JC014225.

Gaspar, P., Y. Grégoris, and J.-M. LeFevre, 1990: A simple eddy kinetic energy model for simulations of the oceanic vertical mixing: Tests at Station Papa and long-term upper ocean study site. *J. Geophys. Res.*, **95**, 16 179–16 193, https://doi.org/10.1029/JC095iC09p16179.

Gasparin, F., M. Hamon, E. Rémy, and P.-Y. Le Traon, 2020: How deep Argo will improve the deep ocean in an ocean reanalysis. *J. Climate*, **33**, 77–94, https://doi.org/10.1175/JCLI-D-19-0208.1.

Gent, P. R., and J. C. McWilliams, 1990: Isopycnal mixing in ocean circulation models. *J. Phys. Oceanogr.*, **20**, 150–155, https://doi.org/10.1175/1520-0485(1990)020<0150:IMIOCM>2.0.CO;2.

Giering, R., and T. Kaminski, 1998: Recipes for adjoint code construction. *ACM Trans. Math. Software*, **24**, 437–474, https://doi.org/10.1145/293686.293695.

Gilbert, J. C., and C. Lemarechal, 1989: Some numerical experiments with variable-storage quasi-Newton algorithms. *Math. Program.*, **45**, 407–435, https://doi.org/10.1007/BF01589113.

Gleckler, P. J., P. J. Durack, R. J. Stouffer, G. C. Johnson, and C. E. Forest, 2016: Industrial-era global ocean heat uptake doubles in recent decades. *Nat. Climate Change*, **6**, 394–398, https://doi.org/10.1038/nclimate2915.

Gregory, J. M., and Coauthors, 2013: Twentieth-century global-mean sea level rise: Is the whole greater than the sum of the parts? *J. Climate*, **26**, 4476–4499, https://doi.org/10.1175/JCLI-D-12-00319.1.

Griewank, A., 1992: Achieving logarithmic growth of temporal and spatial complexity in reverse automatic differentiation. *Optim. Methods Software*, **1**, 35–54, https://doi.org/10.1080/10556789208805505.

Hamlington, B. D., and Coauthors, 2020: Understanding of contemporary regional sea-level change and the implications for the future. *Rev. Geophys.*, **58**, e2019RG000672, https://doi.org/10.1029/2019RG000672.

Hastie, T., R. Tibshirani, and J. Friedman, 2001: *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* Springer Verlag, 533 pp.

Heimbach, P., D. Menemenlis, M. Losch, J. M. Campin, and C. Hill, 2010: On the formulation of sea-ice models. Part 2: Lessons from multi-year adjoint sea ice export sensitivities through the Canadian Arctic Archipelago. *Ocean Modell.*, **33**, 145–158, https://doi.org/10.1016/j.ocemod.2010.02.002.

——, and Coauthors, 2019: Putting it all together: Adding value to the global ocean and climate observing systems with complete self-consistent ocean state and parameter estimates. *Front. Mar. Sci.*, **6**, 55, https://doi.org/10.3389/fmars.2019.00055.

Howe, B. M., and Coauthors, 2019: SMART cables for observing the global ocean: Science and implementation. *Front. Mar. Sci.*, **6**, 424, https://doi.org/10.3389/fmars.2019.00424.

Hsieh, W. W., and B. Tang, 1998: Applying neural network models to prediction and data analysis in meteorology and oceanography. *Bull. Amer. Meteor. Soc.*, **79**, 1855–1870, https://doi.org/10.1175/1520-0477(1998)079<1855:ANNMTP>2.0.CO;2.

Irrgang, C., J. Saynisch, and M. Thomas, 2017: Utilizing oceanic electromagnetic induction to constrain an ocean general circulation model: A data assimilation twin experiment. *J. Adv. Model. Earth Syst.*, **9**, 1703–1720, https://doi.org/10.1002/2017MS000951.

——, ——, and ——, 2019: Estimating global ocean heat content from tidal magnetic satellite observations. *Sci. Rep.*, **9**, 7893, https://doi.org/10.1038/s41598-019-44397-8.

Ishii, M., and M. Kimoto, 2009: Reevaluation of historical ocean heat content variations with time-varying XBT and MBT depth bias corrections. *J. Oceanogr.*, **65**, 287–299, https://doi.org/10.1007/s10872-009-0027-7.

——, and Coauthors, 2017: Accuracy of global upper ocean heat content estimation expected from present observational data sets. *SOLA*, **13**, 163–167, https://doi.org/10.2151/sola.2017-030.

Jayne, S. R., J. M. Wahr, and F. O. Bryan, 2003: Observing ocean heat content using satellite gravity and altimetry. *J. Geophys. Res.*, **108**, 3031, https://doi.org/10.1029/2002JC001619.

Johnson, G. C., S. G. Purkey, N. V. Zilberman, and D. Roemmich, 2019: Deep Argo quantifies bottom water warming rates in the southwest Pacific basin. *Geophys. Res. Lett.*, **46**, 2662–2669, https://doi.org/10.1029/2018GL081685.

Kouketsu, S., and Coauthors, 2011: Deep ocean heat content changes estimated from observation and reanalysis product and their influence on sea level change. *J. Geophys. Res.*, **116**, C03012, https://doi.org/10.1029/2010JC006464.

Kunze, E., 2017: Internal-wave-driven mixing: Global geography and budgets. *J. Phys. Oceanogr.*, **47**, 1325–1345, https://doi.org/10.1175/JPO-D-16-0141.1.

Lary, D. J., A. H. Alavi, A. H. Gandomi, and A. L. Walker, 2016: Machine learning in geosciences and remote sensing. *Geosci. Front.*, **7**, 3–10, https://doi.org/10.1016/j.gsf.2015.07.003.

L'Ecuyer, T., H. K. Beaudoing, M. Rodell, W. Olson, B. Lin, and S. Kato, 2015: The observed state of the energy budget in the early 21st century. *J. Climate*, **28**, 8319–8346, https://doi.org/10.1175/JCLI-D-14-00556.1.

Levitus, S., J. I. Antonov, J. Wang, T. L. Delworth, K. W. Dixon, and A. J. Broccoli, 2001: Anthropogenic warming of Earth's climate system. *Science*, **292**, 267–270, https://doi.org/10.1126/science.1058154.

——, and Coauthors, 2012: World Ocean heat content and thermosteric sea level change (0–2000 m), 1955–2010. *Geophys. Res. Lett.*, **39**, L10603, https://doi.org/10.1029/2012GL051106.

Llovel, W., J. K. Willis, F. W. Landerer, and I. Fukumori, 2014: Deep-ocean contribution to sea level and energy budget not detectable over the past decade. *Nat. Climate Change*, **4**, 1031–1035, https://doi.org/10.1038/nclimate2387.

Losch, M., D. Menemenlis, J. M. Campin, P. Heimbach, and C. Hill, 2010: On the formulation of sea-ice models. Part 1: Effects of different solver implementations and parameterizations. *Ocean Modell.*, **33**, 129–144, https://doi.org/10.1016/j.ocemod.2009.12.008.

Lozier, M. S., and Coauthors, 2019: A sea change in our view of overturning in the subpolar North Atlantic. *Science*, **363**, 516–521, https://doi.org/10.1126/science.aau6592.

Lyman, J. M., and G. C. Johnson, 2014: Estimating global ocean heat content changes in the upper 1800 m since 1950 and the influence of climatology choice. *J. Climate*, **27**, 1945–1957, https://doi.org/10.1175/JCLI-D-12-00752.1.

Manoj, C., A. Kuvshinov, S. Neetu, and T. Harinarayana, 2010: Can undersea voltage measurements detect tsunamis? *Earth Planets Space*, **62**, 353–358, https://doi.org/10.5047/eps.2009.10.001.

McDougall, T. J., and P. M. Barker, 2011: Getting started with TEOS-10 and the Gibbs Seawater (GSW) Oceanographic Toolbox. SCOR/IAPSO Rep. WG127, 28 pp.

Menemenlis, D., and Coauthors, 2005: NASA supercomputer improves prospects for ocean climate research. *Eos, Trans. Amer. Geophys. Union*, **86**, 89–96, https://doi.org/10.1029/2005EO090002.

Meyssignac, B., and Coauthors, 2019: Measuring global ocean heat content to estimate the Earth energy imbalance. *Front. Mar. Sci.*, **6**, 432, https://doi.org/10.3389/fmars.2019.00432.

Minami, T., 2017: Motional induction by tsunamis and ocean tides: 10 years of progress. *Surv. Geophys.*, **38**, 1097–1132, https://doi.org/10.1007/s10712-017-9417-3.

——, H. Toh, and R. H. Tyler, 2015: Properties of electromagnetic fields generated by tsunami first arrivals: Classification based on the ocean depth. *Geophys. Res. Lett.*, **42**, 2171–2178, https://doi.org/10.1002/2015GL063055.

Munk, W., and C. Wunsch, 1979: Ocean acoustic tomography: A scheme for large scale monitoring. *Deep-Sea Res.*, **26A**, 123–161, https://doi.org/10.1016/0198-0149(79)90073-6.

Nocedal, J., 1980: Updating quasi-Newton matrices with limited storage. *Math. Comput.*, **35**, 773–782, https://doi.org/10.1090/S0025-5718-1980-0572855-7.

Osse, T. J., and C. C. Eriksen, 2007: The Deepglider: A full ocean depth glider for oceanographic research. *Oceans 2007*, Vancouver, BC, Canada, IEEE, https://doi.org/10.1109/OCEANS.2007.4449125.

Palmer, M. D., and Coauthors, 2017: Ocean heat content variability and change in an ensemble of ocean reanalyses. *Climate Dyn.*, **49**, 909–930, https://doi.org/10.1007/s00382-015-2801-0.

Ponte, R. M., K. J. Quinn, C. Wunsch, and P. Heimbach, 2007: A comparison of model and GRACE estimates of the large-scale seasonal cycle in ocean bottom pressure. *Geophys. Res. Lett.*, **34**, L09603, https://doi.org/10.1029/2007GL029599.

Purkey, S. G., and G. C. Johnson, 2010: Warming of global abyssal and deep Southern Ocean between the 1990s and the 2000s: Contributions to global heat and sea level rise budgets. *J. Climate*, **23**, 6336–6351, https://doi.org/10.1175/2010JCLI3682.1.

Redi, M. H., 1982: Oceanic isopycnal mixing by coordinate rotation. *J. Phys. Oceanogr.*, **12**, 1154–1158, https://doi.org/10.1175/1520-0485(1982)012<1154:OIMBCR>2.0.CO;2.

Resplandy, L., and Coauthors, 2019: Quantification of ocean heat uptake from changes in atmospheric O_2 and CO_2 composition. *Sci. Rep.*, **9**, 20244, https://doi.org/10.1038/s41598-019-56490-z.

Riser, S. C., and Coauthors, 2016: Fifteen years of ocean observations with the global Argo array. *Nat. Climate Change*, **6**, 145–153, https://doi.org/10.1038/nclimate2872.

Roemmich, D., J. Church, J. Gilson, D. Monselesan, P. Sutton, and S. Wijffels, 2015: Unabated planetary warming and its ocean structure since 2006. *Nat. Climate Change*, **5**, 240–245, https://doi.org/10.1038/nclimate2513.

Sanford, T. B., 1971: Motionally induced electric and magnetic fields in the sea. *J. Geophys. Res.*, **76**, 3476–3492, https://doi.org/10.1029/JC076i015p03476.

Schindelegger, M., A. A. Harker, R. M. Ponte, H. Dobslaw, D. A. Salstein, 2021: Convergence of daily GRACE solutions and models of submonthly ocean bottom pressure variability. *J. Geophys. Res. Oceans*, **126**, e2020JC017031, https://doi.org/10.1029/2020JC017031.

Schnepf, N. R., M. C. Nair, J. Velímský, and N. P. Thomas, 2021: Can seafloor voltage cables be used to study large-scale circulation? An investigation in the Pacific Ocean. *Ocean Sci.*, **17**, 383–392, https://doi.org/10.5194/os-2019-129.

Stammer, D., M. Balmaseda, P. Heimbach, A. Köhl, and A. Weaver, 2016: Ocean data assimilation in support of climate applications: Status and perspectives. *Annu. Rev. Mar. Sci.*, **8**, 491–518, https://doi.org/10.1146/annurev-marine-122414-034113.

Trenberth, K. E., J. T. Fasullo, and M. A. Balmaseda, 2014: Earth's energy imbalance. *J. Climate*, **27**, 3129–3144, https://doi.org/10.1175/JCLI-D-13-00294.1.

——, ——, K. von Schuckmann, and L. Cheng, 2016: Insights into Earth's energy imbalance from multiple sources. *J. Climate*, **29**, 7495–7505, https://doi.org/10.1175/JCLI-D-16-0339.1.

Trossman, D. S., and R. H. Tyler, 2019: Predictability of ocean heat content from electrical conductance. *J. Geophys. Res. Oceans*, **124**, 667–679, https://doi.org/10.1029/2018JC014740.

——, L. Thompson, and S. L. Hautala, 2011: Application of thin-plate splines in two-dimensions to oceanographic tracer data. *J. Atmos. Oceanic Technol.*, **28**, 1522–1538, https://doi.org/10.1175/JTECH-D-10-05024.1.

Tyler, R. H., 2005: A simple formula for estimating the magnetic fields generated by tsunami flow. *Geophys. Res. Lett.*, **32**, L09608, https://doi.org/10.1029/2005GL022429.

——, 2017: Mathematical modeling of electrodynamics near the surface of Earth and planetary water worlds. NASA Tech. Rep. TM-2017-219022, 62 pp., https://ntrs.nasa.gov/citations/20170011279.

——, 2021: A century of tidal variability in the North Pacific extracted from hourly geomagnetic observatory measurements at Honolulu. *Geophys. Res. Lett.*, **48**, e2021GL094435, https://doi.org/10.1029/2021GL094435.

——, and L. Mysak, 1995: Electrodynamics in a rotating frame of reference with application to global ocean circulation. *Can. J. Phys.*, **73**, 393–402, https://doi.org/10.1139/p95-055.

——, and T. J. Sabaka, 2016: Magnetic remote sensing of ocean heat content. *2016 Fall Meeting*, San Francisco, CA, Amer. Geophys. Union, Abstract GC31D-1142, https://agu.confex.com/agu/fm16/meetingapp.cgi/Paper/191894.

Vishwakarma, B. D., S. Royston, R. E. M. Riva, R. M. Westaway, and J. L. Bamber, 2020: Sea level budgets should account for ocean bottom deformation. *Geophys. Res. Lett.*, **47**, e2019GL086492, https://doi.org/10.1029/2019GL086492.

von Schuckmann, K., and Coauthors, 2016: An imperative to monitor Earth's energy imbalance. *Nat. Climate Change*, **6**, 138–144, https://doi.org/10.1038/nclimate2876.

Wahle, K., J. Staneva, and H. Guenther, 2015: Data assimilation of ocean wind waves using neural networks. A case study for the German Bight. *Ocean Modell.*, **96**, 117–125, https://doi.org/10.1016/j.ocemod.2015.07.007.

WCRP Global Sea Level Budget Group, 2018: Global sea-level budget 1993–present. *Earth Syst. Sci. Data*, **10**, 1551–1590, https://doi.org/10.5194/essd-10-1551-2018.

Wood, S. N., 2006: *Generalized Additive Models: An Introduction with R.* Chapman and Hall/CRC Press, 410 pp.

Wu, W., Z. Zhan, S. Peng, S. Ni, and J. Callies, 2020: Seismic ocean thermometry. *Science*, **369**, 1510–1515, https://doi.org/10.1126/science.abb9519.

Wunsch, C., and P. Heimbach, 2013: Dynamically and kinematically consistent global ocean circulation and ice state estimates. *Ocean Circulation and Climate*, 2nd ed. G. Siedler et al., Eds., Elsevier, 553–580.

Zanna, L., S. Khatiwala, J. M. Gregory, J. Ison, and P. Heimbach, 2019: Global reconstruction of historical ocean heat storage and transport. *Proc. Natl. Acad. Sci. USA*, **116**, 1126–1131, https://doi.org/10.1073/pnas.1808838115.

Zhao, Z., 2016: Internal tide oceanic tomography. *Geophys. Res. Lett.*, **43**, 9157–9164, https://doi.org/10.1002/2016GL070567.