Revisiting the 2003–18 Deep Ocean Warming through Multiplatform Analysis of the Global Energy Budget

ANDREA STORTO,a LIJING CHENG,b,c AND CHUNXUE YANGa
a Institute of Marine Sciences, National Research Council, Rome, Italy
b International Center for Climate and Environment Sciences, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China
c Center for Ocean Mega-Science, Chinese Academy of Sciences, Qingdao, China

(Manuscript received 17 September 2021, in final form 27 April 2022)

ABSTRACT: Recent estimates of the global warming rates suggest that approximately 9% of Earth’s excess heat has been cumulated in the deep and abyssal oceans (below 2000-m depth) during the last two decades. Such estimates assume stationary trends deducted as long-term rates. To reassess the deep ocean warming and potentially shed light on its interannual variability, we formulate the balance between Earth’s energy imbalance (EEI), the steric sea level, and the ocean heat content (OHC), at yearly time scales during the 2003–18 period, as a variational problem. The solution is achieved through variational minimization, merging observational data from top-of-atmosphere EEI, inferred from Clouds and the Earth’s Radiant Energy System (CERES), steric sea level estimates from altimetry minus gravimetry, and upper-ocean heat content estimates from in situ platforms (mostly Argo floats). Global ocean reanalyses provide background-error covariances for the OHC analysis. The analysis indicates a 2000-m-bottom warming of 0.08 ± 0.04 W m⁻² for the period 2003–18, equal to 13% of the total ocean warming (0.62 ± 0.08 W m⁻²), slightly larger than previous estimates but consistent within the error bars. The analysis provides a fully consistent optimized solution also for the steric sea level and EEI. Moreover, the simultaneous use of the different heat budget observing networks is able to decrease the analysis uncertainty with respect to the observational one, for all observation types and especially for the 0–700-m OHC and steric sea level (more than 12% reduction). The sensitivity of the analysis to the choice of the background time series proved insignificant.

SIGNIFICANCE STATEMENT: Several observing networks provide complementary information about the temporal evolution of the global energy budget. Here, satellite observations of Earth’s energy imbalance (EEI) and steric sea level and in situ–derived estimates of ocean heat content anomalies are combined in a variational analysis framework, with the goal of assessing the deep ocean warming. The optimized solution accounts for the uncertainty of the different observing networks. Furthermore, it provides fully consistent analyses of global ocean heat content, steric sea level, and EEI, which show smaller uncertainty than the original observed time series. The deep ocean (below 2000-m depth) exhibits a significant warming of 0.08 ± 0.04 W m⁻² for the period 2003–18, equal to the 13% of the total ocean warming.

KEYWORDS: Energy budget/balance; Heat budgets/fluxes; In situ oceanic observations; Inverse methods; Radiation budgets; Sea level

1. Introduction

The recent multidecadal ocean heat content (OHC) increase is widely considered a direct proxy of human-induced climate change because of the role of the oceans in accumulating the majority of the excess Earth’s system energy (von Schuckmann et al. 2016). According to the most recent estimate, compiled from a variety of observational sources, the oceans have gained 89% of Earth’s heat during the period 1971–2018 (von Schuckmann et al. 2020). About 9% of such heat is estimated to be cumulated in the deep and abyssal oceans (below 2000 m of depth; von Schuckmann et al. 2020), which accounts for about 50% of the total ocean volume. This estimate is obtained from repeated World Ocean Circulation Experiment (WOCE) transects and equal to 0.07 ± 0.04 W m⁻² (Purkey and Johnson 2010), confirmed by updated observation-based estimates making use of both Argo float measurements and CTD transects (0.07 ± 0.06 W m⁻², Desbruyères et al. 2017) or estimates by means of Green’s function ocean heat content reconstructions (0.06 W m⁻², Zanna et al. 2019). A recent update confirmed the estimate to be 0.06 ± 0.03 W m⁻² for the 1992–2011 period (Johnson et al. 2018, 2020), while using autoregressive artificial neural networks, Bagnell and DeVries (2021) found a deep ocean warming rate of 0.10 ± 0.04 W m⁻² (16% of the total warming) for the 1990–2019 period.

Because of the sparseness of measurements below 2000 m of depth, the previous estimates present smaller signal-to-noise ratios than upper-ocean warming rates (e.g., Cheng et al. 2017), and are provided as long-term mean warming. The interannual and decadal variability of the deep ocean heat...
content is indeed not possible to estimate directly due to the limited observational sampling. For instance, it is not clear whether ocean ventilation mechanisms responsible for deep ocean heat penetration may have steady (Wunsch and Heimbach 2014) or nonstationary components (Balmaseda et al. 2013; Desbruyères et al. 2017). Furthermore, the extensively debated “warming hiatus” appearing in surface and upper-ocean warming estimates during the 2000s (e.g., Fyfe et al. 2016) is claimed in several studies to be due to the vertical (Balmaseda et al. 2013; Trenberth 2015; Yan et al. 2016) and poleward (Drijfhout et al. 2014) redistribution of heat, eventually concealed by the climate system natural variability (Watanabe et al. 2014; Cheng et al. 2015). It is therefore evident that mapping the interannual heat redistribution in the oceans has crucial importance to understanding and attributing the present climate change. There are also several societal, biological, and economic reasons to improve our knowledge of the deep ocean interannual variability, as this portion of the global ocean does not only provide crucial climate regulation but also encases a large variety of mineral, energetic, and biological resources and constitutes a fundamental repository of biological diversity (Levin et al. 2019).

The global ocean observing system is being complemented with deep Argo floats, capable to sense the ocean up to 5000 m and more of depth (Zilberman 2017). Such new instruments will fill in the future the deep ocean sampling gap, although many years of instrumented records will be necessary to estimate a robust long-term climate signal, also in consideration of the very slow deep ocean heat uptake (Johnson et al. 2019; Silvy et al. 2020; G. Thomas et al. 2020). For the time being, it is necessary to assess the deep ocean warming through indirect estimates and reconstructions, which is the primary goal of the present study. While ocean reanalyses may provide complementary information on the deep ocean variability (Storto et al. 2019a), their disagreement below 2000 m of depth is significant (Palmer et al. 2017) and likely caused by suboptimal vertical correlations used in data assimilation and unbalanced analyses, which often lead to spurious variability (Storto et al. 2017a), thus making these products hard to use at present in the context of deep ocean heat content estimation. Opposed to these methodologies, there exist also observation-only estimates, where a few deep ocean sampling CTD are used to build relationships between the upper and the deep ocean, for subsequent use with Argo floats (Chang et al. 2019).

While the aforementioned observational scariness holds for the deep ocean heat content below 2000 m of depth, there exist several indirect means for estimating the total (full column) ocean heat content. In particular, Clouds and the Earth’s Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) data (Loeb et al. 2009) provide estimates of the global and regional Earth’s energy imbalance at the top of the atmosphere (TOA). An important portion (~90%) of it approximates well the long-term global ocean heat content evolution (Meyssignac et al. 2019; Hakuba et al. 2021; Loeb et al. 2021) over the observed period (since 2000). Additionally, steric sea level estimates, inferred from the difference of total sea level (from altimetry) minus barystatic sea level (from gravimetry) (Storto et al. 2017a, 2019b) are in direct proportionality with the ocean heat content changes (e.g., Chambers et al. 1997), considering that the halosteric sea level contribution is negligible at the global scale (Storto et al. 2019b). All of these data sources, together with upper-ocean heat content estimates, may be merged and analyzed together to provide insights into the deep ocean variability.

Closing budgets through the use of multiple and heterogeneous observing networks is a relatively new area, emerging from the continuous improvement and diversification of the global observing network (e.g., Johnson et al. 2016). For instance, about 10 years ago, the first study combining altimetry and gravimetry satellite with Argo float measurements allowed closing the sea level budget (Leuliette and Miller 2009), namely, the sum of the independent components agrees, within the respective error bars, with altimetry-derived measurements of total sea level (Leuliette and Willis 2011). Such initial studies aimed at closing the budget in a long-term mean sense, and within the uncertainty estimates of each observing network. Applying the budget closure (viz., using the residuals of the sea level budget equation) to infer deep ocean heat content provided too large uncertainties for obtaining robust deep ocean warming estimates (Llovel et al. 2014; Dieng et al. 2015), calling in turn for more sophisticated methodologies. Later on, Volkov et al. (2017) improved this approach providing realistic variability for the deep ocean thermosteric sea level rise. Since these initial studies, some progress has been made in the methodological formulation of the budget closure problem. Notably, L’Ecuyer et al. (2015) introduced a cost function framework to close the global energy budget through weighted analyses of the surface energy fluxes and their errors. The methodology has been recently extended by C. M. Thomas et al. (2020), who confirm the feasibility of adopting a variational approach for budget studies, although both works did not focus on estimating the deep ocean variability.

The goal of the present study is to formulate and assess a variational analysis to estimate deep ocean heat content yearly variations from several in situ and satellite observations. The paper is structured as follows: after this introduction, section 2 presents data and methods used to estimate the deep ocean heat content yearly means. Section 3 summarizes the results of our analysis, and section 4 discusses and concludes.

2. Data and methods

2a. Methodological aspects

1) THE ANALYSIS SCHEME AND ANALYSIS DIAGNOSTICS

The problem of analyzing the deep ocean heat content is here treated as a variational data assimilation problem. Using a variational approach allows us to combine multiple and heterogeneous observing networks through defining appropriate observation operators, even in the case when the problem is overdetermined, and provides an analysis that is the optimal statistical merging of the different datasets, weighed by their respective errors, with the observational information spatially and temporally spread by means of the error covariances.
Thus, our problem is different from a classical budget study, where the unknown component is found by simple addition [in case of a total component; see, e.g., Leuliette and Miller (2009) for the sea level budget] or as residual [see, e.g., Llovel et al. (2014) for the deep ocean heat content].

The analysis state variables are the ocean heat content (anomalies) in different layers (0–300 m, 300–700 m, 700–2000 m, and 2000 m–bottom)—the global ocean being represented as a 4-layer box—as yearly means. Such temporal frequency is chosen to improve the analysis accuracy (use of yearly mean observed data, namely, without the seasonal cycle, are more accurate than values of ocean heat content anomaly, based on the available dilution to analyze simultaneously all multiannual and multilayer from 1 to (number of years). The goal is here to use data assimilation to calculate the vertically varying thermal anomalies whose terms minimize the distance between the analysis and the background, and the analysis and the observations, associated uncertainties; second, observations not measuring directly OHC can be easily ingested through the formulation of the observation operators that maps ocean heat content increments into the observations [e.g., steric sea level, Earth’s energy imbalance (EEI)]. In particular, we have developed and used the variational formalism for possible inclusion of nonlinear operators in the future (e.g., air–sea fluxes, lateral transports); using variational schemes is well known, indeed, to be much easier in case of complicated observation operators.

Another advantage of applying a variational analysis to a relatively small-sized problem is that several diagnostics can be readily calculated. The analysis error covariance matrix $A$ can be computed explicitly [see, e.g., Boutrier and Courtier (2002) and Moore et al. (2012) for the derivation and different formulations], as

$$A = (I - KH)B,$$

with $K$ being the Kalman gain matrix, given by $K = BH(IBH^T + R)^{-1} [\text{see, e.g., Boutrier and Courtier (2002) and Katzfuss et al. (2016) for the derivation}].$ Equation (3) can be applied directly, once the error covariance matrices are defined, to calculate the analysis error covariances. It is also useful to calculate the analysis and the analysis error in observation space so as to compare the optimized solution with the observation estimates and errors. To this end, we consider the following expressions as the optimized solution and its error variance in observation space, using the tangent-linear approximation to efficiently project error variances in observation space, given respectively by

$$y_a = H(x_a) \equiv H\delta x_a + H(x_b) \quad \text{and}$$

$$\sigma^2 = \text{diag}(HAIH^T);$$

namely, Eqs. (4) and (5) diagnose the analysis and its error, respectively, in the space of the observations (e.g., EEI) using the tangent-linear observation operators (defined in detail in section 2b).

2) IMPACT OF THE OBSERVING NETWORKS

As an additional diagnostic, variational data assimilation may provide information on the impact of individual observing systems onto the analyzed OHC. To diagnose the sensitivity of the analysis to the observations and following Cardinali et al. (2004), we introduce the influence matrix $S$ of a (linear) data assimilation system as the derivative of the analysis in

$$x_{t+1} = x_t + \delta x_t,$$

where $T(x, t = j)$ is the vertically varying (function of $z$) global mean temperature anomaly during the $j$th year; $z^0$ and $z^5$ are the depth integration limits for the $i$th vertical layer; $c_p$ is the specific heat capacity of seawater (3991.9 J kg$^{-1}$ °C$^{-1}$); and $\rho_0$ is the reference density (1026 Kgm$^{-3}$). Anomalies are calculated with respect to the entire analysis period 2003–18.

Variational analysis techniques are widely used in meteorology and oceanography. They solve iteratively a cost function whose terms minimize the distance between the analysis and the background, and the analysis and the observations, weighted by the respective error covariances (see, e.g., Talagrand and Courtier 1987):

$$J(\delta x) = \frac{1}{2} \delta y^T(\delta y) - \frac{1}{2} \ddot{y}^T(\ddot{y}) + \frac{1}{2} (\ddot{y} - \ddot{y})^T R^{-1} (\ddot{y} - \ddot{y}),$$

where $\delta x = x - x_b$ is the increment with respect to the background $x_b$, $d = y - H(x_b)$ is the vector of innovations (or misfits), with $y$ being the vector of the observations, $H$ is the (nonlinear) observation operator, and $I$ is its tangent-linear version; in the case of linear observation operator, $H\delta x - d$ is simply given by $Hx - y$; $B$ and $R$ are the background- and observation-error covariance matrices, respectively. We assume the observation error covariance matrix $R$ to be diagonal, with the diagonal terms for the $k$th observation given by $R_k = \sigma_{k1}^2 + \sigma_{k2}^2$, with subscript $n$ and $r$ indicating the instrumental (or measurement) and representativeness (or representation) components of the observational error, respectively. See section 2b for the observation error quantification.

The analysis ($x_a = x_b + \delta x_a$) is defined with $\delta x$ at the minimum of the cost function $J$, and it is the statistically optimal solution (with minimum trace of the analysis error covariance matrix; see, e.g., Kalman and Bucy 1961; Daley 1994). The first term of Eq. (2) contains the temporal and vertical background-error correlations (see the appendix), which implicitly drive the temporal and spatial spreading of the observational information. The appendix also presents details of the variational solution.
observation space with respect to the observations, which can be in turn calculated as

\[ \mathbf{S} = \frac{\partial \mathbf{H}_x}{\partial y} = \mathbf{K}^T \mathbf{H}^T. \]  

(6)

Equation (6) diagnoses the sensitivity of the analysis (in observation space) to the observations, thus quantifying their impact. Similarly, it can be shown (Cardinali et al. 2004) that the analysis sensitivity to the background in observation space is equal to \( I - \mathbf{S} \), with \( I \) being the identity matrix of size equal to the number of observations. The so-called degrees of freedom for signal (DFS) are then defined as the trace of the influence matrix [Eq. (6)], eventually scaled by the total number of observations to provide the percentage impact.

3) TREND UNCERTAINTY

To estimate the OHC trends coming from the variational analysis, weighted least squares regression is applied to the analysis time series (e.g., Wunsch 1996, p. 121; von Schuckmann and Le Traon 2011). To estimate the trend uncertainty, it is suggested (e.g., MacInintosh et al. 2017) to consider two components: the statistical uncertainty in fitting a trend to a finite-length time series with respect to the signal variability (i.e., the trend representativeness error \( \varepsilon^2_d \)) and the accuracy of the time series itself propagated onto the trend (i.e., the trend measurement error \( \varepsilon^2_m \)). These two components sum up to the total trend uncertainty (\( \varepsilon^2 = \varepsilon^2_d + \varepsilon^2_m \)). To estimate the first term, we consider the diagnostic expression of Leroy et al. (2008) to assess the effect of the internal variation on the trend estimation, which takes into account the autocorrelation of the time series at all possible time scales, and is given by

\[ \varepsilon^2_d = \frac{12}{L} \sum_{i=0}^{\infty} \rho_i \sigma^2. \]  

(7)

where \( L \) is the length of the OHC time series, \( \sigma^2 \) is the standard deviation of the regression residuals, and \( \rho_i \) is the autocorrelation at lag \( i \), the sum extending to all possible lags. To estimate the time series accuracy propagation into the trend, we take the expression for the parameter uncertainty variance from the weighted least squares regression (e.g., Fu 2016), equal to

\[ \varepsilon^2_m = (\mathbf{D}^T \mathbf{A}^{-1} \mathbf{D})^{-1}. \]  

(8)

where \( \mathbf{A} \) is the OHC analysis error covariance matrix [Eq. (3)] and \( \mathbf{D} \) is the vector of the time coordinates \( \mathbf{D} = (t_1, t_2, \ldots, t_n)^T \).

b. Global datasets

In this section, we review the datasets used in our ocean heat content analysis, starting with the choice of the background time series, followed by the observational datasets.

1) OHC BACKGROUND-ERROR COVARIANCES

The background OHC field is provided by the ensemble mean of global ocean reanalyses, which also provide the ensemble anomalies to compute the OHC background-error covariances (see the appendix). Indeed, reanalyses are crucial in our technique as they provide ensemble anomalies from which the error covariances are estimated and embedded in the variational scheme.

In particular, we have chosen three state-of-the-art global ocean reanalyses, which are reported in Table 1. These reanalyses have been selected among several products (see, e.g., Balmaseda et al. 2015) because they reach near–real time, and their deep ocean warming trend is found within the uncertainty bar of the latest literature value (von Schuckmann et al. 2020), at 99% confidence level. The selection procedure ensures that the background-error covariance matrix does not include spurious correlations due to, for example, deep ocean OHC drifts, provided that many reanalyses, in the deep ocean, are known to be affected by spurious variability and drifts (Palmer et al. 2017; Storto et al. 2017a).

Two reanalyses are part of the Copernicus Marine Environment Monitoring Service (CMEMS) catalogue of delayed time products, while the third one is the ECCO consortium reanalysis. These reanalyses differ for the ocean model, observational dataset, and assimilation schemes along with differences in the reanalysis initial state, atmospheric forcing, sea ice dynamics, parameterizations, which all concur to the dispersion of the reanalysis ensemble (Storto et al. 2019b).

Table 1 reports the OHC time series from the three reanalyses and their ensemble mean. The top 300 m OHC exhibits significant consistency of the three time series (except during 1993–2000). Below 300 m, in all layers there exists large spread between the reanalyses, although the long-term tendency appears reasonably similar between each other. The three reanalyses also show different interannual variability, some of them with rather smoothed OHC increase; in general, the CMCC Global Ocean Reanalysis System (C-GLORS) exhibits more pronounced interannual variations. In some cases, an initial cooling is visible, which might be ascribed to the climatic signature of the Pinatubo eruption in 1991.

Estimation of the background-error covariance matrix is performed from the ensemble anomalies of these three

<table>
<thead>
<tr>
<th>Reanalysis</th>
<th>Ocean and sea ice model</th>
<th>Assimilation scheme (time window)</th>
<th>Assimilated observations</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMCC-CGLORS v7</td>
<td>NEMO3.6 + LIM2</td>
<td>3DVAR + large-scale bias correction (7 days)</td>
<td>SLA, SST, and in situ profiles</td>
<td>Storto and Masina (2016)</td>
</tr>
<tr>
<td>ECCO-KFS</td>
<td>MITgcm</td>
<td>Kalman filter/smoother (7 days)</td>
<td>SLA, SST, and in situ profiles</td>
<td>Fukumori (2002)</td>
</tr>
<tr>
<td>ECMWF ORAS5</td>
<td>NEMO3.4 + LIM2</td>
<td>3DVAR + bias correction (5 days)</td>
<td>SLA, SST, and in situ profiles</td>
<td>Zuo et al. (2019)</td>
</tr>
</tbody>
</table>
The vertical background-error covariance matrix, which is found positive definite, is computed from these 78 anomalies (26 years × 3 ensemble members), where the anomalies are given with respect to the time-varying ensemble mean. Temporally subsampling the dataset to ensure ergodicity of the ensemble anomalies did not provide any significant difference nor in the background-error covariance matrix nor in the resulting analysis.

Vertical variances (top-left panel of Fig. 2) exhibit the largest (absolute) value within the deepest layer 2000 m–bottom (around 20,000 ZJ; 1 ZJ = 1.0 × 10^{21} J). Cross covariances imply positive correlations between adjacent layers, while negative correlations are visible between the top layer 0–300 m and the two bottom layers 700–2000 m and 2000 m–bottom, suggesting mechanisms of compensation and vertical heat redistribution. The temporal autocorrelation (Fig. 2, bottom-left panel) decreases almost linearly with the time lag, crossing the 0.6 and the 0.0 correlation value after about 2 and 8 years, respectively. The full background-error covariance matrix (Fig. 2, right panel) thus merges these features in its block-diagonal structure.

2) OHC

Ocean heat content estimates are provided by the recent Earth heat inventory performed by von Schuckmann et al. (2020) and conducted under the aegis of the Global Climate Observing System (GCOS). Such assessment has released the ocean heat content yearly time series and their uncertainty for a number of vertical regions as ensemble mean and standard deviation compiled from many independent assessments using in situ oceanic data. The data are publicly available (https://doi.org/10.26050/WDCC/GCOS_EHI_EXP_v2). We used the estimates and accuracy taken for the layers 0–700 m and 700–2000 m as ocean heat content anomalies. The accuracy (ensemble standard deviation) is considered to be the measurement error $\sigma_m$, and the representation error $\sigma_r$ is assumed to be related to the partial coverage of the ocean heat content analyses (see Table 2 for the list of measurement and representation errors used in this study). In particular, we estimated $\sigma_r$ from the standard deviation of the differences of the global ocean heat content minus that in the 70°S–70°N region, both taken from the ensemble mean reanalysis, and equal to 2.0 and 0.3 ZJ for the vertical regions 0–700 m and 700–2000 m, respectively. This error represents the inaccuracy of the global ocean heat content estimates due to the lack of temperature observations over ice-covered areas. Other possible geographical representativeness errors (e.g., that caused by the poor sampling in coastal and/or shallow-water areas or semienclosed basins) are assumed to be negligible, also for consistency with the satellite datasets.

3) EEI

Total-column EEI estimates are derived from CERES EBAF (edition 4.1; Loeb et al. 2016). We use cumulative values.
(i.e., temporally integrated from the beginning of the analysis period) to convert the EEI in ocean heat content-equivalent measurements. Indeed, most of the top-of-atmosphere EEI is accumulated in the oceans, and thus is strictly related to the variations of ocean heat content (e.g., Storto et al. 2017b).

The observation operator is given by

\[ H_{\text{eei}} := c_p \rho_b \sum_i \int_{z_i}^{z_{i+1}} T \, dz, \tag{9} \]

where \( \beta \) is the amount of EEI uptaken by the ocean and is currently set to 0.90 according to the latest estimate (von Schuckmann et al. 2020) for the period 2010–18, which slightly scales down previous estimates (e.g., 0.93; Allan et al. 2014).

The observational error variance is set equal to the squared sum of 1) the instrumental error (0.43 W m\(^{-2}\)), calculated as standard deviation of the difference of yearly mean EEI between the CERES sensors onboard \( \text{Terra} \) and \( \text{Aqua} \) (edition 4.1) or \( \text{Terra} \) and \( \text{NPP} \) (edition 1A) over the overlapping period (2012–17); 2) the representativeness error associated with the temporal variability of \( \beta \) at the yearly time scale, and set equal to 0.30 W m\(^{-2}\), and estimated as standard deviation of differences between yearly OHC and OHC filtered with a 3-yr running mean. The 3-yr time scale is chosen after cross-wavelet analysis of OHC and EEI time series (Fig. 3), which shows that their maximum coherence is found for periods equal to 3 years.

The total observational error is then equal to 0.56 W m\(^{-2}\). The use of a constant \( \beta \) coefficient, although unavoidable due to the very limited knowledge of the OHC–EEI relationship over short time scales, is accounted for in the definition of the representativeness error; however, including the representation error of \( \beta \) is not the same as varying \( \beta \) from year to year. Thus, care should be taken when analyzing the resulting OHC interannual changes that might be erroneously affected by the use of this constant coefficient.

In addition, note that Argo observations are used in the adjustment scheme of Loeb et al. (2009) in order for

<table>
<thead>
<tr>
<th>Observing system</th>
<th>Representation error</th>
<th>Measurement error</th>
<th>Total error</th>
</tr>
</thead>
<tbody>
<tr>
<td>OHC 0–700 m</td>
<td>2.0 ZJ</td>
<td>3.1–16.8 ZJ</td>
<td>3.7–16.9 ZJ</td>
</tr>
<tr>
<td>OHC 700–2000 m</td>
<td>0.3 ZJ</td>
<td>1.7–4.8 ZJ</td>
<td>1.7–4.8 ZJ</td>
</tr>
<tr>
<td>Steric sea level</td>
<td>((2.1^2 + 1.3^2)^{1/2}) mm</td>
<td>((1.4^2 + 0.15^2)^{1/2}) mm</td>
<td>2.84 mm</td>
</tr>
<tr>
<td>EEI</td>
<td>((0.30^2 + 0.21^2)^{1/2}) W m(^{-2})</td>
<td>0.43 W m(^{-2})</td>
<td>0.56 W m(^{-2})</td>
</tr>
</tbody>
</table>
CERES EBAF EEI to match the EEI measured by the in situ oceanic network; namely, CERES EBAF and OHC data are not fully independent, in strict terms. However, Argo data are used only to constrain the climatological monthly means of the TOA fluxes, while interannual variations remain unchanged.

4) STERIC SEA LEVEL

Steric sea level estimates are derived from the difference of altimetry minus gravimetry, as in several previous studies (Leuliette and Miller 2009; Storto et al. 2017a; Storto 2019). In particular, global sea level from altimetry is computed as the ensemble mean of three independent global sea level assessment: from the University of Colorado Sea Level Research Group (Veng and Andersen 2020), the CMEMS (Taburet et al. 2019), and the National Aeronautics and Space Administration (Beckley et al. 2017). The global barystatic sea level (mass component) is calculated from the GRACE and GRACE-FO solutions of equivalent ocean sea level (Chambers and Bonin 2012; Johnson and Chambers 2013). We have used in particular the latest release RL06Mv2 (Watkins et al. 2015; Wiese et al. 2019), which embeds correction for the glacial isostatic adjustment from ICE6G-D (Roy and Peltier 2015).

The observation operator is given by

$$H_{st} := \sum_i \alpha_i \int_{z_0}^{z_i} T \, dz,$$

where $\alpha_i$ is the thermal expansion coefficient associated with the $i$th vertical layer (see, e.g., MacIntosh et al. 2017). Note that $\alpha$ can be computed either analytically or statistically (see, e.g., WCRP Global Sea Level Budget Group 2018). Here, we have preferred to fit the reanalysis OHC data with the steric data from the reanalyses to provide estimates of $\alpha$ varying in the four vertical layers and fully consistent with the background fields. Testing different ways of regressing the steric sea level to the thermal expansion coefficients did not lead to any significant difference in the resulting analyses, the differences being much smaller than the analysis confidence limits.

In the case of steric sea level, the observational error variance is set equal to the squared sum of the 1) instrumental error of the altimetry [provided in turn as a sum of systematic and random errors at yearly time scales, equal to 1.4 mm according to Ablain et al. (2019)]. We also calculated the ensemble spread of the three global sea level time series, equal to 1.0 mm, but we prefer to use the more conservative estimate from Ablain et al. (2019), considering the limited number of global sea level time series used; 2) instrumental error of gravimetry-based barystatic sea level estimates, calculated projecting the GRACE RL06 uncertainty, which also accounts for leakage errors (Wiese et al. 2016), onto global yearly means, and equal to 0.15 mm; 3) observation operator representativeness error, as variance of the residual of the linear fit to calculate $\alpha$, and equal to 2.1 mm, which represents how good the linear regression parameters can capture the steric sea level variability (note that the latter implicitly accounts also for the neglected contribution of the halosteric variations onto the global steric sea level obtained by the difference of altimetry minus gravimetry global sea level estimates); 4) geographical representativeness error, which considers the nonglobal coverage of altimetry data. Such errors are estimated from the ensemble mean reanalysis and set equal to the standard deviation of the differences between fully globally averaged sea surface height and averaged fields without the grid points where the mapped altimetry data from CMEMS (Taburet et al. 2019) do not contain valid data. This
was found equal to 1.3 mm. The total observational error for
the steric estimates is thus equal to 2.84 mm.

3. Results

a. OHC analysis

The analyzed ocean heat content anomalies \( \mathbf{x}_a \), from the
minimization of the cost function in Eq. (2) are shown in
Fig. 4, in comparison with the ensemble mean of the global
ocean reanalyses \( \mathbf{x}_b \). Time series are also given in the online supplemental material as text files.
Peaks of heat content are clearly different: the background
shows extreme OHC values at the beginning and end of the
0–300-m and 700–2000-m depth time series, while the analy-
sis mostly in correspondence of the 0–300-m layer. There is
also a clear attenuation of the 700–2000-m peaks. The analy-
sis shows indeed a more pronounced interannual variabil-
ity (simply defined here as standard deviation of the
detrended yearly means) than the background at all depths
(Fig. 5). The interannual variability of the deep ocean
increases significantly from the background, although the
interannual variability in the top layer remains the largest
ones. Such increased deep ocean variability may be due to
the underestimated background variability—following the
ensemble-average operation—and/or spurious variability
borne by the heterogeneous observing networks used in
our analysis; therefore, it will need to be addressed in detail
when deep ocean observations, for example, from the deep
Argo network, will be available.

The deep ocean heat content anomalies (Fig. 6) are char-
acterized by significantly smaller uncertainty than the
background. Apart from the interannual fluctuations, there
is qualitative agreement with the trend estimate from
von Schuckmann et al. (2020) in terms of both the analysis and
its uncertainty. Maxima of the deep ocean OHC occur during
2012, 2016 and 2018. There occur drops of heat content during
2015 and 2017. Note, however, that the interannual variations
of deep ocean heat content are not significant with respect to
the analysis uncertainty (Fig. 6), thus its temporal modulation
must be taken with care, considering also the intrinsic limita-
tions of the constant \( \beta \) coefficient in the assimilation of the
EEI observations [section 2b(3)].

The analyzed ocean heat content, steric sea level and cumu-
lativ e EEI from the analysis [in observation space, Eq. (4)]
are shown in Fig. 7, together with the background-equivalent
values and the observations. The plot indicates the very good
adherence between the observations and the analysis, which
is able to account simultaneously for all observational con-
strains. The OHC 0–700 m at the beginning of the time series
(2003–04) differs from that of the observations, due to the
large observational errors (see later); for the rest of the time
series and also within the layer 700–2000 m, values of analysis
and observations are very close. The analyzed steric sea level
time series shows only slight adjustments made by our estima-
tion method with respect to the original values from satellite
observations. The cumulative EEI time series shows almost
identical behaviors in the satellite-derived and analyzed time
series.

Another important result is about the increase of accuracy of
the analyzed time series with respect to the observational error
estimates, in the comparison between observational errors
(diagonal elements of \( \mathbf{R} \) and analysis errors in observation
space [Eq. (5)]. This is shown in Fig. 8 for the four
observational datasets used. Because of temporal correlations, steric sea level and CERES errors are slightly larger than their average at the beginning and end of the time series. OHC presents instead a time-varying definition of the errors in both observations (von Schuckmann et al. 2020) and analysis, which tend to decrease with time in the 0–700-m layer because of the increased observational sampling of the Argo floats. In the 700–2000-m layer, the behavior of the observational error has a less obvious variability. Significant error reductions occur for the steric sea level data (12% error reduction) and for the 0–700-m ocean heat content (14%) but are nonnegligible also for the CERES (6%) and OHC 700–2000 m (4%). This implies that the variational formulation of the OHC analysis is not only able to provide an optimized solution for the global energy budget but also is able to decrease the uncertainty associated with the observational data.

The DFS (Fig. 9) summarize the impact of the different components of the variational formulation onto the resulting analysis [Eq. (6)]. While the background has the smallest impact (14.6%), the total observations impact is equal to 85.4%. CERES and the two OHC time series exhibit the largest impacts, all of them between 20% and 24%, the OHC in the 700–2000-m layer leading the ranking (23.5%). Steric sea level data have smaller impact in the analysis (19.4%).

Figure 10 shows the ocean warming trends by vertical levels (left) and cumulated from the surface (right), together with the error bars as described in section 2 [Eqs. (7) and (8)], while Table 3 reports the trend values and associated uncertainty. The total column warming is equal to 0.62 ± 0.08 W m⁻² for the period 2003–18. The individual layers account for 0.33 ± 0.07 W m⁻² (0–700 m; 53% of the total), 0.22 ± 0.02 W m⁻²
(700–2000 m; 35%), and 0.08 ± 0.04 W m\(^{-2}\) (2000 m–bottom; 13%). In particular, the deep ocean warming is slightly larger than previous works—both in absolute value and percentage with respect to the global ocean warming—but within the error bars of all previous estimates. The variational analysis also provides an optimized solution for the steric sea level trend (1.60 ± 0.14 mm yr\(^{-1}\)) and the EEI trend (0.69 ± 0.03 W m\(^{-2}\)). Note that the latter assumes by construction that 90% of the EEI is cumulated in the oceans.

By comparing the background and the analysis trends in Fig. 10, it is also possible to conclude that the ensemble of reanalyses is in good agreement, within the error bars, with the analysis at all levels, except for the overestimated trend within the 0–300-m layer. However, the combination of all the observational data leads to a significant reduction of the warming rate uncertainties. For instance, the uncertainty is reduced from 0.17 to 0.04 W m\(^{-2}\) in the deep and abyssal oceans and from 0.22 to 0.08 W m\(^{-2}\) for the total column. Note that the 2000-m–bottom trend from the background time series is not statistically significant.

b. Robustness of the assessment

We tested the robustness of the analysis by performing a random combination of the background time series taken from the three reanalyses instead of their ensemble mean and permuting their use in each vertical level. The aim of such procedure is to assess the impact of the choice of the background on the resulting optimized solution. The number of permutations performed is 22. Results are summarized in Fig. 11 and Table 3. Figure 11 shows the mean and standard deviation of the ensemble of the analyses. Small or no impact of the background is found for the layer 300–700 m and 700–2000 m, respectively. For the latter, this is due to background-error variances largely exceeding the observational error variances. The top layer 0–300 m is the most affected by the perturbations, although there are no significant changes in the time series behavior. The deep ocean is only slightly affected toward the middle of the time series, while the background perturbation has small or no impact toward the beginning or the end of the time series, respectively. In terms of the resulting trends (Table 3), the percentage ratio of the ensemble standard deviation and the ensemble mean is used to identify the level of noise associated with the background perturbation. Values are small, all below 5% (4.7% for the OHC 0–300 m). The deep ocean warming exhibits a value of 2.4%, while values for steric sea level and EEI are below 0.3%. All these values are significantly below the uncertainty bars identified by the analysis, meaning that the choice of the background has no significant impact on the optimized solution.

Another specific sensitivity test was performed to assess the impact of CERES EBAF data on the resulting analysis. This is summarized in Fig. 12. Withholding the TOA radiative fluxes, the resulting analysis provides a smaller deep ocean warming trend (0.03 ± 0.09 W m\(^{-2}\)), a larger interannual variability (about 30% larger than the case with TOA radiative fluxes), and a smaller accuracy (average analysis error around 3 times as large). This indicates that CERES data have an important constraining role of the total water column and contribute significantly to the final OHC tendency, through increase of the accuracy and rectification of the long-term trend.

c. Heat redistribution

Figure 13 shows the time series of the implied heat fluxes, after a 3-yr running mean is applied to the time series. The 2000-m–bottom ocean heat content tendency corresponds to
the vertical heat flux (VHF) at 2000 m, whereas the EEI optimized solution indicates the net downward air–sea flux. Also shown are the ocean heat content tendencies within the layers 0–700 and 700–2000 m, which summed together perfectly match the difference between the global ocean heat content tendency (GOHCT) minus the deep ocean contribution (not shown). The figure also shows the time-averaged values for the ocean heat content tendencies. Interestingly, the analysis shows periods of opposite anomalous tendency between the deep ocean and the top layers. In particular, during 2011–15 the upper-ocean heat content tendency exceeds its time-averaged value, while the deep ocean tendency is below, whereas the GOHCT has a rather constant value during 2012–16. The large deep ocean OHC tendency during 2017 counteracts the anomalous cooling of the 700–2000 m layer.

Higher-than-average deep ocean warming during the first years of the time series (2006–08) counteracts the slight to neutral upper-ocean warming. Our time series is limited in time, mostly because of the availability of TOA radiative fluxes from CERES after 2000 and gravimetry data from GRACE after 2003, and caution should be taken in analyzing the interannual signal of our OHC analysis [e.g., section 2b(3)]. Nevertheless, the vertical heat redistribution indicates that minimizing energy budget terms’ error (Hedermann et al. 2017) might effectively resolve the dilemma about the apparent warming hiatus during the first years of the 2000s (see, e.g., Trenberth 2015) by reconciling integrated OHC estimates (from e.g., EEI and steric sea level) and upper-ocean variability from in situ observing networks. However, the downward VHF at 2000 m shows large year-to-year fluctuations that respond by construction to the balance between air–sea fluxes and upper-ocean warming. While its behavior comes from the statistically optimal formulation described in section 2, the reliability of its variations from the oceanographic point of view shall be validated in the future with independent estimates, which are not available at this time.

4. Summary and discussion

During the last decade, the climate community has highlighted the impact of climate change on the deep and abyssal waters, which have been found to significantly warm in studies that make use of repeated transects (Purkey and Johnson 2010), regional observing programs (Meinen et al. 2020), or indirect derivations (Volkov et al. 2017). At the same time, Durack et al. (2018), Palmer et al. (2017), and Storto et al. (2017a) have shown that coupled model simulations and ocean
Table 3. Trends and uncertainty of the global energy budget analysis are shown in the first row. The percentages indicate the amount with respect to the 0-m-bottom ocean warming. The second row reports the value of the ensemble spread (standard deviation) of the background perturbation, in terms of absolute values and, in parentheses, percentage with respect to the mean value (noise level).

<table>
<thead>
<tr>
<th>Observing system</th>
<th>OHC0–300 (W m⁻²)</th>
<th>OHC300–700 (W m⁻²)</th>
<th>OHC0–700 (W m⁻²)</th>
<th>OHC700–2000 (W m⁻²)</th>
<th>OHC0–2000 bottom (W m⁻²)</th>
<th>OHC2000–bottom (W m⁻²)</th>
<th>Steric sea level (mm yr⁻¹)</th>
<th>EEI (W m⁻²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trend 2003–18</td>
<td>0.23 ± 0.05 (37%)</td>
<td>0.10 ± 0.04 (16%)</td>
<td>0.33 ± 0.07 (53%)</td>
<td>0.22 ± 0.02 (35%)</td>
<td>0.54 ± 0.07 (87%)</td>
<td>0.08 ± 0.04 (13%)</td>
<td>0.62 ± 0.08</td>
<td>1.60 ± 0.14</td>
</tr>
<tr>
<td>Background perturbation</td>
<td>0.004 (4.7%)</td>
<td>0.003 (1.3%)</td>
<td>0.002 (0.5%)</td>
<td>0.0002 (0.1%)</td>
<td>0.002 (0.4%)</td>
<td>0.002 (2.4%)</td>
<td>0.001 (0.2%)</td>
<td>0.005 (0.3%)</td>
</tr>
</tbody>
</table>

The substantial limitations of the current global observing network to sample the deep and abyssal ocean (deeper than 2000 m) at the interannual time scale calls, however, to improving methods for indirect estimation of the deep ocean warming, which is the main goal of the present work. Several studies suggested that the deep ocean warming is mainly originated in the Antarctic Ocean and ventilated by Antarctic Bottom Waters (e.g., Purkey and Johnson 2010). In the present work, we have exploited the varioabale assimilation formalism to build an ocean heat content time series from a large drifts (e.g., Hobbs et al., 2018, 2020). The key feature of our methodology is that the optimized solution found through decreasing their uncertainty (e.g., Muetterties et al. 2018, 2020) but is consistent within the error bars. Quantifying the heat redistribution with an optimized solution, accounts for all observing networks and their uncertainties simultaneously may also help to reconcile inconsistent datasets about the apparent warming hiatus.

Quantifying the heat redistribution with an optimized solution, accounts for all observing networks and their uncertainties simultaneously may also help to reconcile inconsistent datasets about the apparent warming hiatus.
observed in upper-ocean heat content time series but not in TOA radiative fluxes. Our analysis indicates that during periods of smaller-than-average upper-ocean warming, the ocean below 700 m of depth is characterized by an increase of heat. However, longer time series are needed to validate and better understand the mechanisms that lead to such interannual variations of the deep vertical heat transports (e.g., at 2000 m).

Our study indicates that sustaining in situ programs such as the Argo floats (Roemmich et al. 2019), both the upper ocean only and the deep floats (Zilberman 2017; Johnson et al. 2019), is crucial to quantify the global ocean warming, together with the design of future satellite missions to monitor with accuracy the Earth energy imbalance (e.g., Palchetti et al. 2020). This work also proposes a simple methodology to enhance the consistency of several independent observational datasets, which is among the priority of many space agencies and environmental monitoring programs (Popp et al. 2020). Although the quantification of the observational uncertainties has been achieved here with rather simple considerations, this can be certainly improved in the future, leading in turn to more accurate analysis estimates. In particular, deep Argo floats will offer in the future complementary and constraining observations to both increase the accuracy of the deep ocean warming estimates and improve the characterization of the background-error covariances. The correct separation of random and systematic errors (Wunsch 2018) is another important point that needs extensive investigations. Similarly, the weak dependence between CERES EBAF data and Argo-derived OHC estimates has been neglected in this study. Depending upon the availability of a detailed statistical characterization of the error correlations between these two datasets, the variational formulation may be updated in the future to account for nondiagonal terms of the error covariance matrix $\mathbf{R}$.

Backward extension of our analysis method relies on accurate estimates of ocean mass variability, the TOA radiative fluxes and their relationship with the ocean heat content on an interannual basis, which will be evaluated in the future. The variational analysis formalism may be also applied in the future to regional assessment (through studying the balance within ocean “boxes”), although in this case much care shall be devoted to the use of the observation data, as (i) the local halosteric component of sea level is not negligible, in general; (ii) TOA net heat fluxes need to account for precise atmospheric transports; and (iii) oceanic heat transports, taken, for example, from models or reanalyses, are characterized by rather low accuracy.
Acknowledgments. Author Cheng is supported by National Natural Science Foundation of China (Grants 42076202 and 42122046). The authors declare no conflicts of interest.

Data availability statement. Reanalysis data that form the background values for the analysis presented in this study are taken from the Copernicus Marine Environment Monitoring Service (https://marine.copernicus.eu) and ECCO consortium (https://www.ecco-group.org/products.htm). Ocean heat content time series are published by GCOS (https://doi.org/10.26050/WDCC/GCOS_EHI_EXP_v2). Ocean heat content anomalies and uncertainties within the different vertical regions. The time series underwent a 3-yr running filter, for consistency with the matched time scale between EEI and OHC (see Fig. 3).

Eq. (2) is made only once so as to analyze all layers and years simultaneously. A limited-memory quasi-Newton algorithm (Byrd et al. 1995) is used to achieve the convergence of cost function, which requires provision of the gradient of $J(\delta x)$:

$$\nabla_{\delta x} J(\delta x) = (B)^{-1} \delta x + H^T R^{-1} (H \delta x - d), \quad (A1)$$

which in turn requires us to code the adjoint ($H^T$) of the observation operators that will be introduced in section 2b. Note that when the observation operators used in Eq. (2) are all linear, using the variational cost function is equivalent to the solution of the best linear unbiased estimator (BLUE) that could be achieved analytically without the variational minimization (see, e.g., Daley 1994).

The background OHC field $x_b$ is provided by the ensemble mean of global ocean reanalysis realizations (see section 2b). Accordingly, the background-error covariance matrix, for use in the variational cost function of Eq. (2), is calculated from the ensemble anomalies collected during the analysis period, using the following decomposition:

$$B = \Sigma_u C \Sigma_v = B_u \circ C_v, \quad (A2)$$

where the matrix $\Sigma_u$ is the square-root block-diagonal vertical standard deviation matrix (a 64 $\times$ 64 matrix, with each of the 16 diagonal 4 $\times$ 4 blocks repeating the 4 $\times$ 4 diagonal matrix whose diagonal elements contain the vertical background-error standard deviations), and $C$ is the correlation matrix, containing both the vertical and the temporal correlations. Indeed, we assume that there exist temporal correlations induced by the inertia of the climate system, similarly to other statistical analyses of ocean observations (e.g., Gasparin et al. 2015; Kuusela and Stein 2018), and we thus introduce temporal correlations.

In Eq. (A2), for sake of visualization, the covariance matrix can be equivalently given as the Schur product between the $B_u$ total vertical covariance matrix, namely, the 64 $\times$ 64 matrix.

APPENDIX

Details of the Variational Formulation

Here, we detail some specificities of the variational formulation utilized in this work. The minimization implied by

![Figure 12](image_url)  
**Figure 12.** As in Fig. 6, but with the black line showing the analysis without the ingestion of CERES EBAF data. Accordingly, the gray shade represents the associated uncertainty.

![Figure 13](image_url)  
**Figure 13.** OHC tendencies and corresponding heat fluxes, for the different vertical regions. The time series underwent a 3-yr running filter, for consistency with the matched time scale between EEI and OHC (see Fig. 3).
with each of its $4 \times 4$ blocks containing the vertical covariance matrix $D$, between the four layers,

$$
B = \begin{pmatrix}
D_v & \cdots & D_v \\
\vdots & \ddots & \vdots \\
D_v & \cdots & D_v
\end{pmatrix},
$$

and the temporal correlation matrix $C$.

In practice, to set up $B$, each element $B_{i,j}$ is given by the vertical background-error covariance between the vertical level corresponding to the element $i$ and that of the element $j$, and the temporal correlation $c_t$ corresponding to the temporal distance between the elements $i$ and $j$. Formally:

$$
B_{i,j} = D_v(z_i,z_j) c_t(|t_i - t_j|),$$

(A3)

where both $D_v$ and $c_t$ are calculated from the dataset of the ensemble anomalies, as sample covariance and sample autocorrelation, respectively. The expression of Eq. (A2) built through Eq. (A3) is the one actually used to specify background-error covariances for use in the analysis of Eq. (2).

REFERENCES


