Estimating ocean heat content from the ocean thermal expansion parameters using satellite data

Vijay Prakash Kondeti[1], Palanisamy Shanmugam[1]

[1]Ocean Optics and Imaging Laboratory, Department of Ocean Engineering, Indian Institute of Technology Madras, Chennai-600036, India

Correspondence to: Palanisamy Shanmugam (pshanmugam@iitm.ac.in)

Abstract. Ocean heat content (OHC) is a depth-integrated physical oceanographic variable used to precisely measure ocean warming. Because of the limitations associated with in-situ CTD data and Ocean Reanalysis system products, satellite-based approaches have gained importance in estimating the daily to decadal variability of OHC over the vast oceanic region. Efforts to minimize the biases in satellite-based OHC estimates are needed to realize the actual response of the ocean to the brunt of climate change. In the current study, an attempt has been made to better implement the satellite-based ocean thermal expansion method to estimate OHC at 17 depth extents ranging from the surface to 700m. To achieve this objective, artificial neural network (ANN) models were developed to derive thermosteric sea level (TSL) from a given dataset of sea surface temperature, sea surface salinity, geographical coordinates, and climatological TSL. The model-derived TSL data were further used to estimate OHC changes based on the thermal expansion efficiency of heat. Statistical analysis showed high correlation coefficients and low errors in validation of satellite-derived TSL / OHC of 700 m modeling depth (N 388469, R 0.9926 / 0.9922, RMSE 1.16 m / 1.56 GJ m^{-2}, MBE -0.19 m / -0.24 GJ m^{-2}, MBPE -0.46% / -0.03%, MAE 0.76 m / 1.03 GJ m^{-2}, and MAPE 2.34% / 0.13%) and nearly similar results at the remaining modeling depths. These results suggest that the proposed ANN models are capable of generating satellite-based daily OHC maps by covering both shallower and deeper oceanic regions of varying bathymetry levels (≥ 20 m). In addition, the first-ever attempt to estimate the ocean thermal expansion component (*i.e.*, TSL) from satellite data was successful, and the model-derived TSL can be used to obtain high-end sea-level rise products in the global ocean.

1. Introduction

Owing to the vast spatial coverage and high heat capacity, oceans balance the planet's temperatures by absorbing 89% of the excess atmospheric heat caused by the greenhouse effect and global warming (Abraham et al., 2013; IPCC, 2014; Roemmich et al., 2015; Riser et al., 2016; Trenberth et al., 2016; Meyssignac et al., 2019; Von Schuckmann et al., 2023). A precise understanding of the depth-wise penetration of this heat and its accumulation in the upper oceanic layers is of great importance to the scientific community (Liang et al., 2015; Baxter, 2016; IPCC, 2022). Ocean heat content (OHC), a depth-integrated physical oceanographic variable that refers to the amount of heat energy accumulated between any two depths, has gained attention in various studies of the Earth's Energy Imbalance (Von Schuckmann et al., 2016; Trenberth et al., 2016; Cheng et al., 2017; Meyssignac et al., 2019; Cheng et al., 2022). Thus, accurate estimation of OHC changes at various depth extents is vital and is the motivation of the current study.

To obtain a complete picture of OHC changes at different depths, direct measurements of in-situ conductivity, temperature, and depth (CTD) profiles are necessary. These in-situ measurements of the ocean properties are limited in terms of depth and spatial coverage, leading to the biased global reconstruction of OHC estimates owing to the sparse measurement data and spatial coverage deficiencies (Jagadeesh et al., 2015; Meyssignac et al., 2019; Marti et al., 2022). However, in-situ CTD profile measurements have been used to develop and validate the different OHC models (Momin et al., 2011; Jagadeesh et al., 2015; Su et al., 2020; Vijay and Shanmugam, 2022). On the other hand, synthetic CTD profile data generated by Ocean Reanalysis systems (ORA) have been used to study OHC variability in spatial and temporal scales (Balmaseda et al., 2015; Palmer et al., 2017). More recently, satellite-based methods have become crucial to overcome the limitations associated with in-situ CTD data and ORA products, to ensure the OHC trend at a global scale, and to understand the evolution of the Earth's climate system (Meyssignac et al., 2019; Vijay and Shanmugam, 2022).

The existing satellite-based OHC algorithms can be broadly grouped into three approaches based on the principles/parametrizations employed: (i) internal tide oceanic tomography (ITOT), (ii) ocean net surface heat fluxes, and (iii) ocean thermal expansion. Apart from these approaches, research is exploring ways to make use of tidal magnetic satellite observations (Irrgang et al., 2019), electrical conductance (Trossman and Tyler, 2019), and atmospheric oxygen & carbon dioxide concentrations (Resplandy et al., 2018) to infer OHC changes. The ITOT technique involves correlating the satellite altimeter-derived internal tide phase changes with ocean warming to estimate OHC variability. This technique is still at the proof-of-concept level, and the associated challenges remain to be addressed (Zhao, 2016a, 2017; Meyssignac et al., 2019). The OHC estimation through ocean net surface heat fluxes employs several assumptions and approximations in deriving the input parameters to compute radiative and turbulent heat fluxes, which in turn leads to higher uncertainty in global OHC changes (Wild et al., 2015; L'Ecuyer et al., 2015; Meyssignac et al., 2019).

On the other hand, the ocean thermal expansion method is a promising technique for the estimation of OHC by considering the thermosteric sea level (TSL) and expansion efficiency of heat (EEH). Numerous satellite-based OHC models have been developed based on the sea surface height anomaly from altimeters, water mass change equivalent sea level anomaly from the Gravity Recovery and Climate Experiment mission (GRACE), sea surface temperature from the various radiometers onboard satellites, and wind speed/stress from scatterometers/numerical weather models. Pioneering works done by White and Tai (1995), Chambers et al. (1997), Polito et al. (2000), and Sato et al. (2000) have attempted to implement the ocean thermal expansion method based on a relationship between OHC and satellite altimeter-based sea surface height anomaly (SSHA). It should be mentioned that regardless of the source, the volume of seawater changes when it is subjected to heating/cooling, and it eventually reflects in sea surface topography. The SSHA data recorded by the satellite altimeters comprise the sea surface topography changes due to tides, atmospheric pressure, salinity (haline), and barotropic flows along with the thermal effects. The SSHA changes due to the tides and atmospheric pressure can be corrected, but the effects of salinity and barotropic flows remain unresolved with the OHC estimates produced by Wang and Tai (1995) and Chambers et al. (1997). Sato et al. (2000) have introduced a haline correction factor as the integral product of the haline contraction coefficient and salinity anomaly from in-situ CTD profile data. Owing to the limitations associated with in-situ data, the in-situ-based haline correction cannot be applied to satellite altimeter-based SSHA data while correlating with the space and time-varying OHC data. Jayne et al. (2003) have proposed the Alt-GRACE approach to resolve the effect of barotropic flows in sea surface topography by subtracting the satellite gravimetry-derived water mass change component from SSHA data. Though the Alt-GRACE approach has improved the accuracy of satellite-based OHC estimates compared to Wang and Tai (1995), Chambers et al. (1997), Polito et al. (2000), and Sato et al. (2000), the issues associated with the haline effects and other approximations on the ocean thermal expansion coefficient and seawater density data have led to significant uncertainties in satellite-based OHC estimates.

With the advancement of artificial intelligence, several researchers have attempted to model OHC by directly relating it with the satellite-based parameters of relevance by using deep-learning regression techniques (Jagadeesh and Ali, 2006; Momin et al., 2011; Chacko et al., 2015; Jagadeesh et al., 2015; Su et al., 2020, 2021; Marti et al., 2022; Lyman and Johnson, 2023). These deep-learning models have oversimplified the OHC problem by neglecting the effects of salinity and barotropic flows. In addition, no previous work has accounted for the space and time-varying nature of the ocean thermal expansion coefficient and seawater density in OHC computations. The other common drawbacks of the existing works are discussed in Sect. 4.3. Consequently, there is a need for developing a satellite-based model to accurately implement the ocean thermal expansion method to estimate OHC by resolving all the issues associated with salinity variation, barotropic flows, ocean thermal expansion, seawater density, choice of temperature and its scale.

Given the above background, we have made a major attempt to develop and implement the satellite-based ocean thermal expansion models for estimating OHC changes at various depth extents (such as 20 m, 30 m, 40 m, 50 m, 100 m, 150 m, 200 m, 250 m, 300 m, 350 m, 400 m, 450 m, 500 m, 550 m, 600 m, 650 m, and 700 m). It enables the research community to generate satellite-based OHC maps of varying bathymetry levels (≥ 20 m) by covering both shallower and deeper oceanic waters. For this, artificial neural network (ANN) architectures were developed to estimate TSL for the given sea surface temperature (SST), sea surface salinity (SSS), geographical coordinates, and climatological TSL. The model-derived TSL estimates were then used to model OHC changes by accounting for the expansion efficiency of heat. The proposed models are capable of estimating TSL and OHC at multiple depth extents accurately. The robustness of the new models was tested by comparison of model-derived TSL and OHC with in-situ data.

2. Data

2.1. In-situ data for model development and in-situ validation

For this study, in-situ CTD profile data (collected by Argo floats) were obtained from the World Ocean Database-2018 (WOD) at the NOAA's National Centers for Environmental Information data archive for the period of 2005-2020 (Boyer et al., 2018a). WOD has been extensively used by the research community for various ocean applications (Levitus et al., 2009; Momin et al., 2011; Levitus et al., 2012; Cheng et al., 2014; Roemmich et al., 2015; Jagadeesh et al., 2015; Su et al., 2020). WOD comprises the oceanographic data of diverse biogeochemical parameters that have been collected by various institutions, agencies, individual researchers, and data recovery initiatives. The quality-controlled CTD profile data (*accepted_value* flag) of standard depth levels recommended by the International Association of Physical Oceanography (1936) were considered in this study to compute the TSL_d and OHC_d parameters and to obtain the SST and SSS data. The standard depth levels considered for deriving the TSL and OHC are given as 20 m, 30 m, 40 m, 50 m, 100 m, 150 m, 200 m, 250 m, 300 m, 350 m, 400 m, 450 m, 500 m, 550 m, 600 m, 650 m, and 700 m. The in-situ TSL_d and OHC_d parameters were computed by applying the integration formulae (Eqs. 1 & 2) on the CTD profile data of depth range from the ocean surface to the respective standard depth (d), and the corresponding SST and SSS values were extracted. Similarly, the climatological parameters such as $TSL_{clim,d}$ and $OHC_{clim,d}$ were computed from the monthly climatological temperature and salinity data of 41 vertical levels obtained from the World Ocean Atlas-2018 (WOA) (Boyer et al., 2018b). The theoretical considerations for computing OHC change at a particular depth can be found in Vijay and Shanmugam (2022) (Vijay and Shanmugam, 2022), and the same was adopted in this study. The Gibbs-SeaWater (GSW) Oceanographic Toolbox of TEOS-10 (IOC et al., 2010) was used to compute the in-situ-based OHC and TSL.

$$OHC_d = \int_0^d \rho C_P \Theta \, dz \tag{1}$$

$$TSL_d = \int_0^d \alpha \Theta \, dz \tag{2}$$

where OHC_d refers to the heat energy accumulated in an oceanic layer of depth range from the surface to a stipulated depth (d) and is given in the units of joules per unit area (J m^{-2}). Similarly, TSL_d (in meters) refers to the thermosteric sea level integrated from the surface to a stipulated depth (d). And, Θ is the conservative temperature in K (derived from in-situ temperature, absolute salinity, and pressure), ρ is the seawater density in kg m^{-3} (derived from the conservative temperature, absolute salinity, and pressure), C_P is the specific heat capacity (= 3991.87 J kg^{-1} K^{-1}), and α is the thermal expansion coefficient in K^{-1} (derived from the conservative temperature, absolute salinity, and pressure).

Python programming was used to prepare the individual databases for all the standard depth levels by extracting CTD profile data from the WOD and WOA NetCDF files with the help of NetCDF4, NumPy, Pandas, and GSW libraries. Each database (in-situ OHC, in-situ TSL, in-situ SST, in-situ SSS, climatological OHC, climatological TSL, and WOA geographical coordinates) was divided into two datasets, one for the model development spanning from 2005-2016 and the other for validating the model spanning from 2017-2020, by ensuring a well distribution in spatiotemporal scales over the global open ocean. The spatial distribution of data points used to model TSL_{700} and OHC_{700} is shown in Fig. A1. The in-situ CTD profiles of depth coverage shallower than 700 m are also included in this process of deriving the TSL and OHC of remaining depth extents. Indeed, the number of CTD profiles and their distribution in global oceans is higher than the CTD profile density of modeling depth 700 m.

2.2. Satellite-based validation

For the validation period 2017-2020, the NOAA Advanced Very High-Resolution Radiometer (AVHRR) Optimum Interpolation Sea Surface Temperature product (OISST v2.1) was used for daily SST data at 0.25° spatial resolution (Huang et al., 2021). Daily SSS data of the same spatial resolution were obtained from the ORAS5 Ocean reanalysis system of the European Centre for Medium-Range Weather Forecasts at the CMEMS portal (Product ID: GLOBAL_REANALYSIS_PHY_001_031) (Zuo et al., 2017). The NetCDF4 and NumPy Python libraries were used to read and resample satellite data to the WOA-18 grid and to collocate with the corresponding Argo in-situ data points. The accuracy of the satellite-based SST and ORA-based SSS was verified by comparing with Argo-measured SST and SSS data (N = 244722). The observed R, RMSE, MBE, and MAE values in SST & SSS validations are 0.99 & 0.99, 0.51°C & 0.26 PSU, -0.05°C & -0.006 PSU, and 0.33°C & 0.12 PSU, respectively. High correlation coefficients and low errors indicate the minimal deviation of satellite-based data from the actual (in-situ) data and ensure the reliability of satellite data in accurately representing the physical oceanographic conditions.

3. Methodology

3.1. Theoretical formulations

Ocean thermal expansion is the best proxy to model the heat content accumulated in an oceanic layer. Unlike freshwater, seawater expands when it warms and contracts when it cools to temperatures above its freezing point. The volumetric expansion of seawater is non-isotropic in nature due to the differences in the degree of constraint in different directions. In a vertical direction, atmospheric pressure exerts a normal force on the seawater parcel at the surface. The magnitude of this normal/vertical force is less compared to the horizontal forces exerted by physical barriers such as continental boundaries and geographic features on the ocean floor. It allows the ocean thermal expansion of seawater in the vertical direction rather than the horizontal direction, as the seawater is less constrained in the vertical direction compared to the horizontal direction. The amount of change in seawater volume in response to the net warming/cooling depends on the absolute conservative temperature and ocean thermal expansion coefficient (Eq. 2). Following are the GSW functions (Eqs. 3-5) (IOC et al., 2010) involved in the calculation of TSL (Eq. 2) for the given set of measured temperature (T), practical salinity (SP), pressure (P), longitude (x), and latitude (y).

$$Absolute\ salinity\ (SA) = gsw.SA\_from\_SP\ (SP, P, x, y) \tag{3}$$

$$\Theta = gsw.CT\_from\_T(SA, T, P) \tag{4}$$

$$\alpha = gsw.Alpha(SA, \Theta, P) \tag{5}$$

Hence, an attempt has been made in this study to model TSL as a function of SST, SSS, and geographical coordinates. The existing correlations between the proposed input parameters and the targeted output parameter were explored by employing in-situ-based data used in the model development process (Fig. 1).

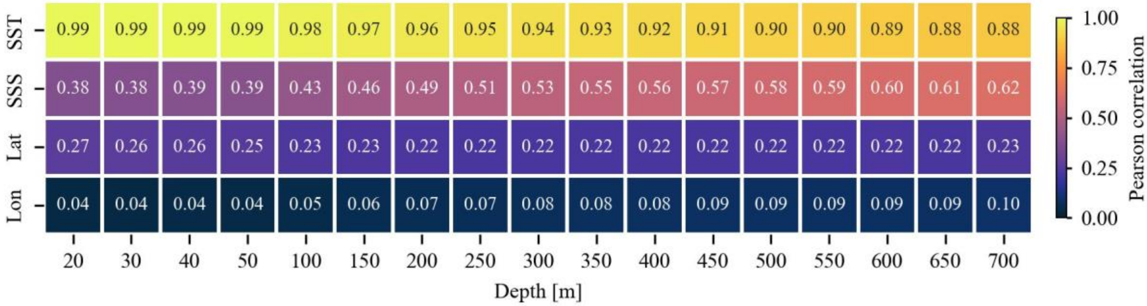

Figure 1. Heatmap showing the Pearson correlation coefficients between the input parameters (*i.e.*, SST, SSS, and geographical coordinates) and the output parameter (TSL) of various depth extents.

It is observed that SST has an almost one-to-one correlation with TSL at shallower depth extents, and can be solely used to model the thermal expansion of upper oceanic layers. Despite a decreasing trend in correlation strength when moving towards deeper depths, SST plays a primary role in accounting for TSL variations at deeper depths, because of its strong correlations with TSL. Observed weaker correlations between SSS and TSL which are plausible owing to the salinity's secondary role in TSL variations as compared to the temperature. However, an increasing trend in correlation coefficients between SSS and TSL is observed towards the deeper depth extents. Hence, SST and SSS are complementary to each other in resolving the TSL variations, and their combination plays a major role in modeling TSL of all depth extents considered in this study. Apart from these physical parameters, absolute salinity used in the computation of seawater density, conservative temperature, and ocean thermal expansion coefficient is a function of geographical coordinates along with practical salinity and pressure (Eq. 3). By considering all these theoretical considerations and observed correlations, an attempt has been made to model TSL of various depth extents by employing SST, SSS, and geographical coordinates as the input parameters along with the climatological TSL (Fig. 2). Here, TSL_d is an external manifestation of OHC_d stored in an oceanic layer based on EEH_d (Eq. 6). The model-derived TSL is further used to estimate OHC changes (as shown in Fig. 2 along with climatological OHC) as follows,

$$OHC_d = \frac{TSL_d}{EEH_d} \qquad (6)$$

where *EEH* is a conversion factor that explains the relationship between the relative changes in OHC and the corresponding TSL. EEH is not a constant value over the global ocean as it varies as a function of temperature, salinity, and pressure. Hence, ANN modeling is employed in this study to estimate OHC from TSL by accounting the complex variations in EEH.

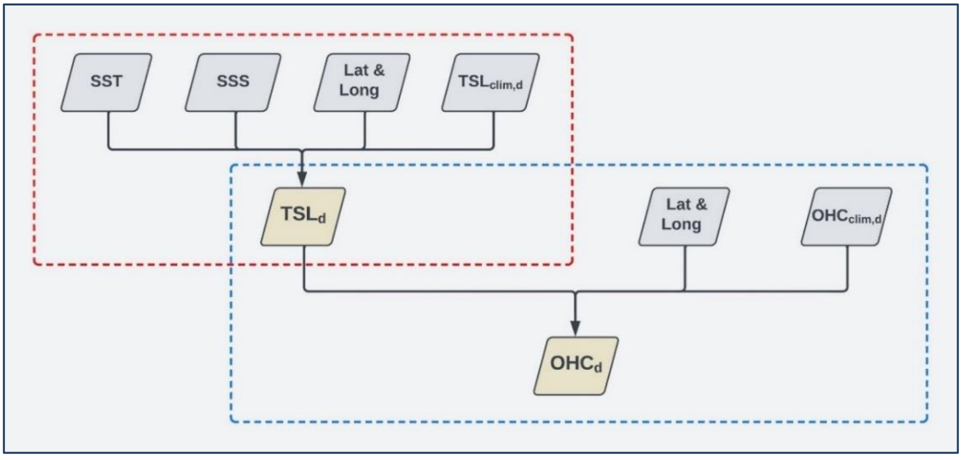

Figure 2. Flow chart representing the parameters involved in TSL and OHC modeling. The red and blue dashed boxes represent the TSL and OHC frameworks employed in ANNs, respectively.

3.2. ANN model description

This section explains the various steps and architectures involved in the ANN modeling of TSL and OHC. The multilayer perceptron regressor algorithm of deep neural networks was used to model both TSL and OHC (Pedregosa et al., 2011). It is observed that the input data of geophysical parameters have different units and scales. The range and order of SST, SSS, latitude, and longitude data are -1.8 °C to 34.15 °C & *O(10^1)*, 2.53 PSU to 40.45 PSU & *O(10^1)*, -76° to 80° & *O(10^1)*, and -180° to 180° & *O(10^2)*, respectively. In addition, the range and order of TSL$_{clim,d}$ and OHC$_{clim,d}$ are also distinct and vary with water depth. Hence, the input data were normalized using the StandardScaler class of Scikit-Learn and feed-forwarded through the neural networks. This StandardScaler normalizes the raw data to ensure the mean and standard deviation of each input parameter as 0 and 1, respectively. It allows the ANN model to focus on the relative importance and relationships between the input parameters rather than their magnitude. The standardized input data were injected into the corresponding neurons in the input layer and forward propagated through the hidden layers and then the output layer by applying the random weights and rectified linear unit (ReLU) activation function at each neuron (Fig. 3). The model outputs were compared with the actual data and computed mean squared error (MSE) using a loss function (Eq. 7). In addition, L2 regularization (α_{L2}) was employed to add a penalty term to the loss value to prevent overfitting. The observed error was then backpropagated through the network to update weights and biases using the Adam optimizer based on the learning rate and gradient of the error (see Eq. 8 in Vijay and Shanmugam, 2022). This process is repeated until the validation score improves more than 0.0001.

$$MSE = \frac{1}{N}\sum(Y_{pred,i} - Y_{act,i})^2 \qquad (7)$$

where N is the number of samples, $Y_{pred,i}$ is the predicted data, and $Y_{act,i}$ is the actual data. The model development work was carried out by employing both the input and output parameters from the in-situ sources. It enables the ANN models to implement the input data of any remote sensing sources to produce OHC estimates subject to the reliability and accuracy of those data sources. The particle swarm optimization technique (Kennedy and Eberhart, 1995; Shi and Eberhart, 1998) was employed for hyperparameter tuning, and the hyperparameters' combinations corresponding to each modeling depth are presented in Table 1. The Joblib module of the Scikit-Learn library was used to save all the TSL and OHC models of various depths considered in this study, and the same module was used to load the TSL and OHC models of desired depth with the help of a unified Python script.

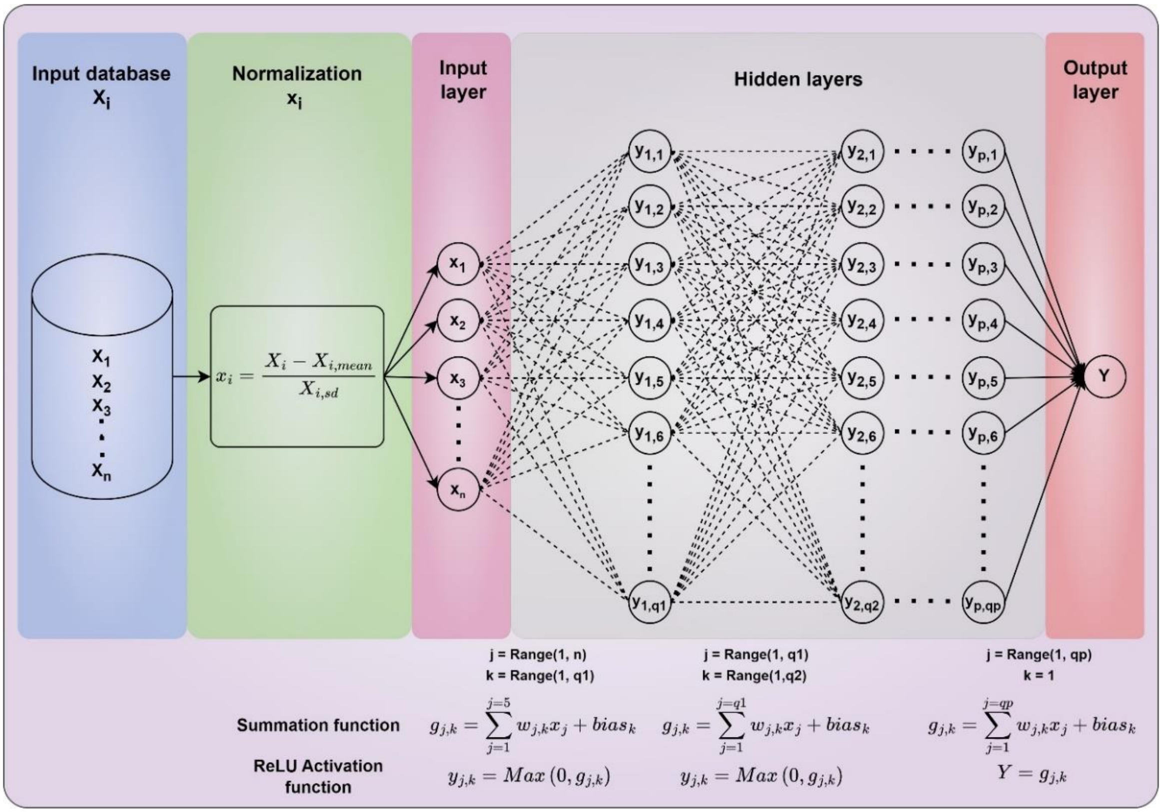

Figure 3. Schematic of the ANN architecture employed in the modeling of TSL and OHC parameters. The flow of the modeling and the associated mathematical transformations/formulations are given by considering a typical ANN architecture with *n* input parameters, one output parameter, *p* hidden layers, and *q1* to *qp* neurons in each hidden layer.

Table 1. ANN model hyperparameters employed in TSL (regular font) and OHC (bold font) modeling of various depth extents.

Depth (m)	Hidden layers	Batch size	α_{L2}	Learning rate	No. of iterations
20	38, 10, 55	178	0.00422	0.0004	14
	49, 12, 34	**183**	**0.09023**	**0.0001**	**26**
30	100, 97, 36	165	0.00001	0.0001	14
	11, 50, 55	**58**	**0.00079**	**0.0001**	**16**
40	64, 71, 5	106	0.00001	0.0001	16
	57, 89, 46	**148**	**0.09691**	**0.0001**	**19**
50	64, 99, 30	241	0.01478	0.0001	17
	56, 59, 10	**139**	**0.07188**	**0.0001**	**22**
100	70, 100, 100	256	0.00001	0.0009	30
	25, 36, 63	**256**	**0.03556**	**0.0016**	**44**
150	47, 83, 92	60	0.00001	0.0005	34
	49, 77, 28	**69**	**0.05176**	**0.0318**	**16**
200	100, 100, 16	256	0.00315	0.0022	33
	27, 48, 67	**202**	**0.05638**	**0.0367**	**18**
250	56, 82, 67	174	0.00001	0.0019	39
	2, 100, 77	**73**	**0.00001**	**0.0037**	**22**
300	83, 28, 74	128	0.00001	0.0028	36
	48, 92, 10	**87**	**0.01364**	**0.0459**	**12**
350	85, 25, 67	128	0.04606	0.0013	20
	27, 53, 48	**141**	**0.08585**	**0.0851**	**14**
400	89, 75, 96	64	0.04859	0.0007	26
	49, 1, 80	**138**	**0.00001**	**0.0031**	**20**
450	51, 83, 95	128	0.08582	0.0005	42
	47, 27, 52	**32**	**0.00263**	**0.0055**	**24**
500	71, 100, 62	128	0.00001	0.0012	27
	45, 100, 63	**126**	**0.05162**	**0.0607**	**15**
550	47, 89, 91	256	0.00843	0.0011	44
	64, 75, 78	**114**	**0.05176**	**0.0634**	**15**
600	98, 65, 6	16	0.00001	0.0001	48
	63, 17, 10	**180**	**0.04654**	**0.0538**	**23**
650	100, 69, 75	16	0.00001	0.0001	18
	53, 74, 40	**176**	**0.07072**	**0.0048**	**20**
700	98, 37, 37	164	0.04262	0.0015	32
	83, 63, 79	**216**	**0.01217**	**0.0742**	**19**

4. Results and discussion

The performance of TSL and OHC models on unseen data from the in-situ and satellite sources was assessed using density scattergrams and statistical metrics. These metrics include mean bias error (MBE), mean bias percentage error (MBPE), mean absolute error (MAE), mean absolute percentage error (MAPE), root mean square error (RMSE), Pearson correlation coefficient (R), slope, and intercept (also referred and presented in Vijay and Shanmugam, 2022). To better understand the model performance, mean values of in-situ data were computed for the validation period and used to compute the weighted average of validation metrics across all the modeling depths.

4.1 In-situ validations with unseen data

The main objective of the in-situ-based validation with unseen data is to evaluate the generalization ability and overall accuracy of TSL and OHC-ANN models on unseen data. For this purpose, the in-situ measured data of SST, SSS, latitude, and longitude were used to predict the model-derived TSL and OHC values, which were then compared with in-situ TSL and OHC data. The number of validation data points and their spatial distribution are presented in Table 2 and Fig. A1(b). The performance of the TSL models is exceptionally good on unseen data of all the modeling depths without any overfitting (Table 2 and Fig. 4). Similar model performance can also be observed in the case of OHC estimates as it primarily depends on the TSL estimates (Table 2 and Fig. 5). The high values of R indicate a strong positive correlation between the predicted and in-situ OHC (TSL) values. This suggests that the models are generally capable of capturing OHC (TSL) patterns in the data. The slope and intercept of the regression line between predicted and actual values are close to 1 and 0, respectively. This suggests that the model-derived estimates have good agreement with the actual data with a minimal bias. The RMSE values are notably small implying that the predicted OHC values have a little random error when compared to the actual data. The MBE and MBPE values are close to zero, indicating that the model-derived estimates have a negligible systematic error when compared to the actual values. The low MAE and MAPE values are also indicating a high accuracy with the model-predicted OHC values. These results clearly demonstrate that the proposed ANN models succeeded in generalizing and accurately predicting the OHC (TSL) data with a high accuracy.

The spatial distribution of mean percentage error (MPE) over the global open oceanic regions was computed by averaging the observed percentage errors of all modeling depths available at each pixel (Fig. A2). It is observed that the models' performance is comparatively low over the north-western parts of the North Atlantic gyre, southwestern parts of the South Atlantic gyre, Kuroshio extension, and Antarctic circumpolar current regions due to the high eddy kinetic energy (Beech et al., 2022; Ni et al., 2023). An elaborate note on the potential sources of the observed MPE values is given in Sect. 4.4. Further, the entire validation dataset was divided into two parts in terms of the observed overestimation and underestimation of data. In the cases of overestimation (underestimation), 95% of the data points have MPE of less than or equal to 0.47% (0.44%). The lower values of MPE indicate that the proposed ANN models are capable of capturing OHC patterns in all major oceanic basins and can be used to produce accurate OHC products upon implemention on real-time data.

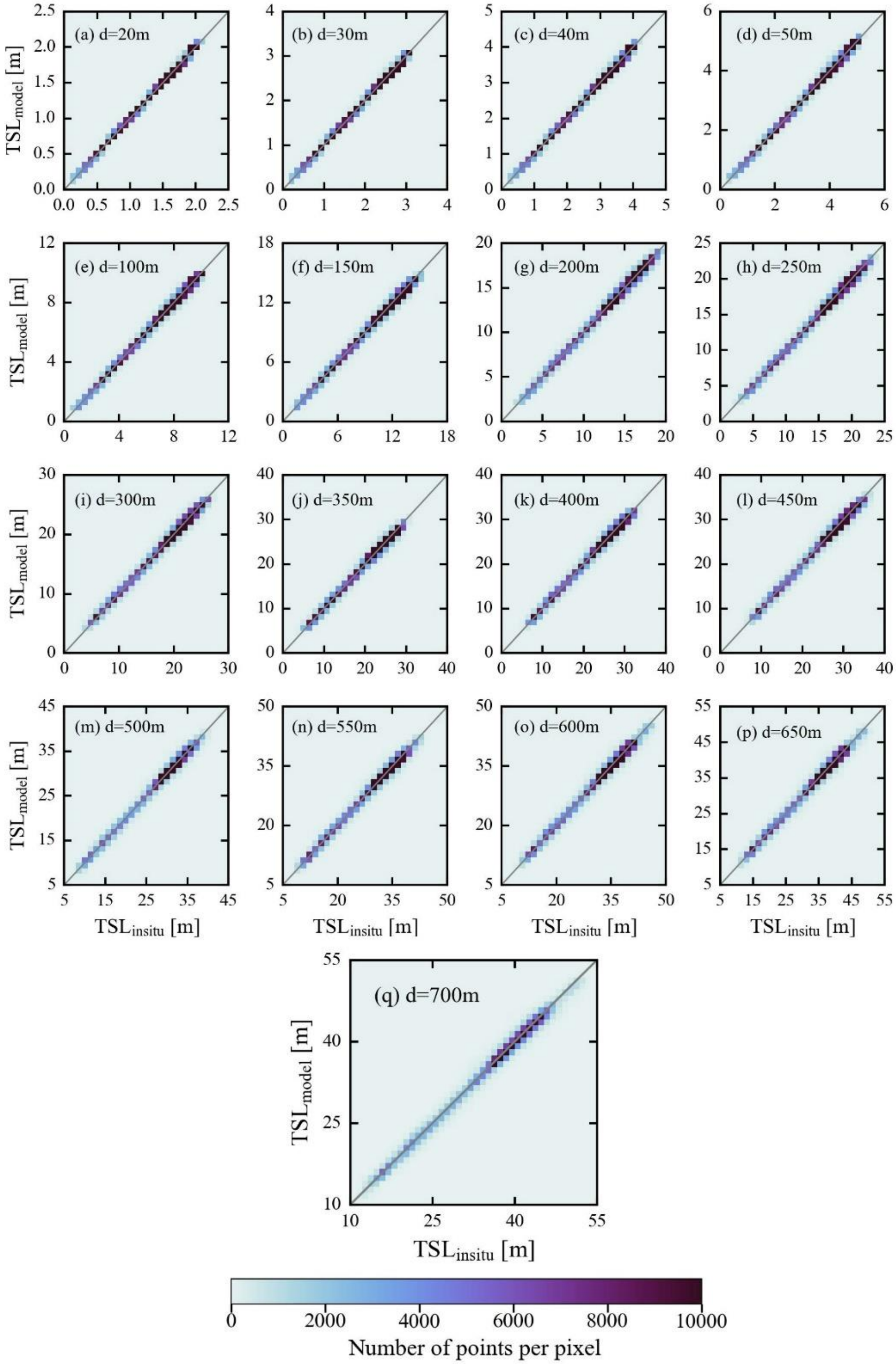

Figure 4. Density scattergrams showing the observed agreement between model-derived TSL values and in-situ measured TSL values during in-situ-based validations.

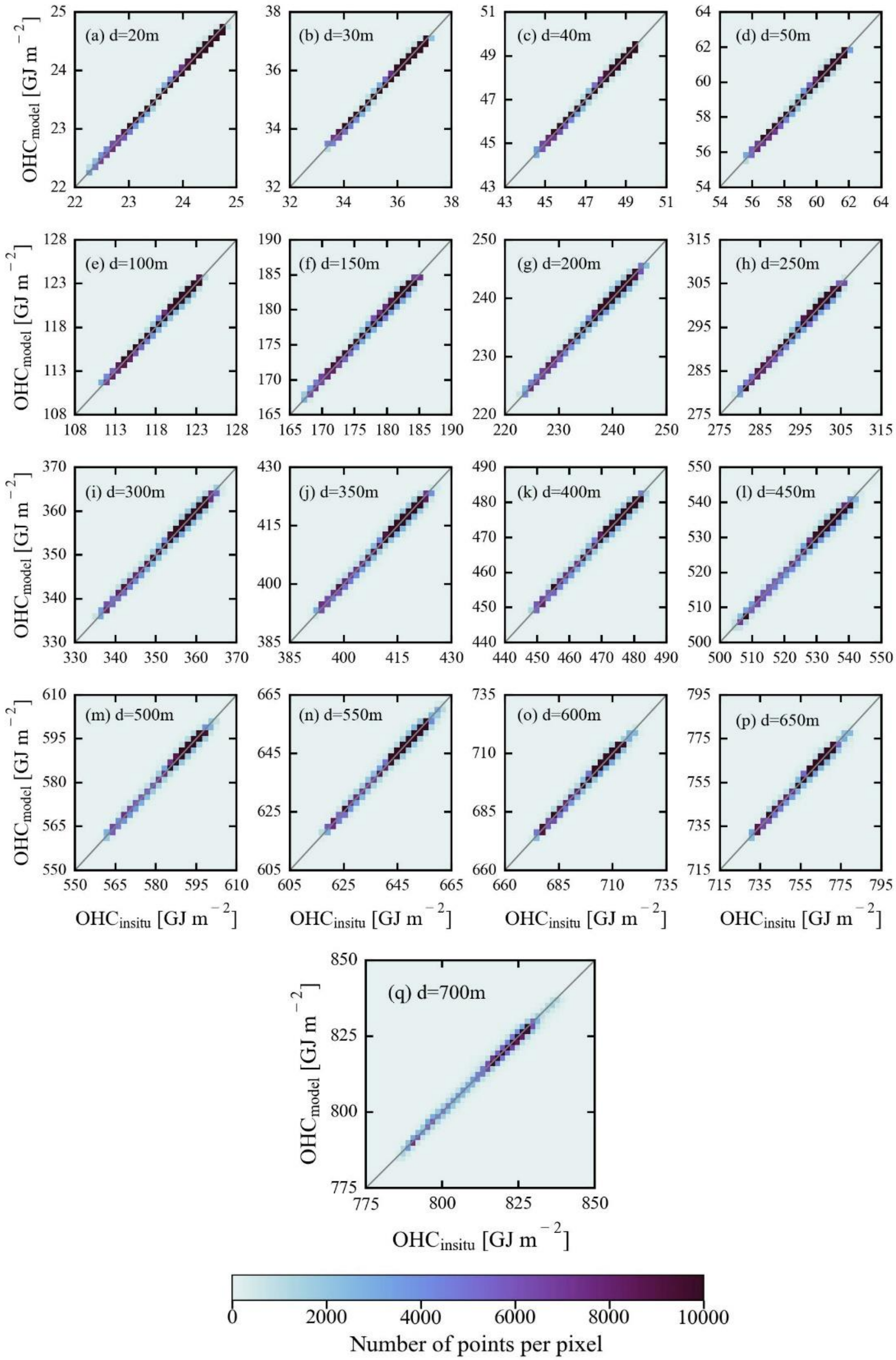

Figure 5. Density scatterplots showing the observed agreement between model-derived OHC values and in-situ measured OHC values during in-situ-based validations.

Table 2. Statistical results from the in-situ-based validation of TSL (regular font) and OHC (bold font) against in-situ data. The units for the various metrics used in TSL & OHC validations are given as follows: Mean (m & GJ m^{-2}), RMSE (m & GJ m^{-2}), MBE (m & GJ m^{-2}), MBPE (%), MAE (m & GJ m^{-2}), MAPE (%), and intercept (m & GJ m^{-2}).

Depth (m)	**N**		**Mean**	**R**	**RMSE**	**MBE**	**MBPE**	**MAE**	**MAPE**	**Slope**	**Intercept**
	Model development	**Model validation**									
20	801303	536719	1.44	0.9997	0.01	-0.0007	0.0575	0.006	0.60	0.9981	0.002
			23.91	**0.9997**	**0.02**	**-0.0011**	**-0.0047**	**0.009**	**0.04**	**0.9987**	**0.030**
30	794166	532149	2.15	0.9993	0.03	0.0029	0.3764	0.015	0.99	0.9982	0.007
			32.85	**0.9992**	**0.04**	**0.0010**	**0.0027**	**0.021**	**0.06**	**0.9992**	**0.030**
40	787074	526571	2.85	0.9988	0.05	-0.0009	0.1325	0.027	1.28	0.9988	0.002
			47.78	**0.9988**	**0.07**	**-0.0008**	**-0.0014**	**0.038**	**0.08**	**0.9978**	**0.103**
50	779134	520102	3.54	0.9984	0.07	-0.0008	0.0861	0.042	1.47	0.9975	0.008
			59.70	**0.9984**	**0.10**	**0.0015**	**0.0028**	**0.057**	**0.10**	**0.9972**	**0.169**
100	731065	476709	6.80	0.9974	0.18	-0.0129	-0.1725	0.120	2.09	0.9960	0.015
			119.00	**0.9973**	**0.25**	**-0.0279**	**-0.0233**	**0.169**	**0.14**	**0.9981**	**0.196**
150	712120	460278	9.83	0.9967	0.29	-0.0407	-0.3419	0.205	2.41	0.9905	0.053
			177.97	**0.9965**	**0.40**	**-0.0369**	**-0.0198**	**0.279**	**0.16**	**0.9867**	**2.331**
200	697314	446979	12.64	0.9961	0.38	-0.0001	0.0571	0.272	2.51	0.9960	0.050
			236.62	**0.9959**	**0.53**	**-0.0076**	**-0.0029**	**0.372**	**0.16**	**0.9939**	**1.426**
250	686378	436906	15.28	0.9959	0.46	-0.0361	-0.1803	0.332	2.49	0.9943	0.051
			295.04	**0.9957**	**0.63**	**-0.0242**	**-0.0078**	**0.450**	**0.15**	**0.9918**	**2.392**
300	678526	429501	17.80	0.9956	0.55	-0.0471	-0.0023	0.392	2.53	0.9851	0.218
			353.29	**0.9954**	**0.74**	**-0.0155**	**-0.0039**	**0.525**	**0.15**	**0.9889**	**3.902**
350	672148	423688	20.23	0.9949	0.65	-0.1035	-0.3383	0.462	2.59	0.9860	0.179
			411.40	**0.9947**	**0.87**	**-0.0357**	**-0.0081**	**0.613**	**0.15**	**0.9861**	**5.676**
400	666605	418686	22.57	0.9947	0.72	-0.0425	-0.0526	0.505	2.52	0.9887	0.213
			469.39	**0.9945**	**0.97**	**-0.0067**	**-0.0010**	**0.676**	**0.14**	**0.9879**	**5.683**
450	661336	413987	24.83	0.9946	0.78	-0.1227	-0.4726	0.547	2.47	0.9916	0.087
			527.25	**0.9943**	**1.06**	**-0.1681**	**-0.0315**	**0.741**	**0.14**	**0.9872**	**6.588**
500	654880	408240	27.03	0.9949	0.80	-0.0604	-0.1866	0.558	2.29	0.9945	0.089
			585.03	**0.9947**	**1.07**	**-0.0761**	**-0.0127**	**0.747**	**0.13**	**0.9894**	**6.105**
550	649850	403357	29.14	0.9948	0.85	-0.0462	-0.0937	0.586	2.19	0.9911	0.213
			642.69	**0.9945**	**1.15**	**0.0347**	**0.0057**	**0.787**	**0.12**	**0.9900**	**6.479**
600	645150	398855	31.21	0.9945	0.91	-0.0390	-0.0205	0.623	2.18	0.9883	0.327
			700.28	**0.9942**	**1.23**	**0.0298**	**0.0046**	**0.838**	**0.12**	**0.9873**	**8.937**
650	640479	392921	33.18	0.9941	0.99	0.0185	0.0903	0.670	2.19	0.9949	0.189
			757.74	**0.9939**	**1.33**	**0.0086**	**0.0014**	**0.892**	**0.12**	**0.9904**	**7.296**
700	633004	388469	35.13	0.9941	1.04	-0.1928	-0.4791	0.711	2.17	0.9858	0.307
			815.15	**0.9938**	**1.41**	**-0.2413**	**-0.0292**	**0.960**	**0.12**	**0.9836**	**13.134**
Weighted average				0.9961	0.74	-0.0620	-0.1591	0.513	2.29	0.9927	0.177
				0.9960	**1.03**	**-0.0515**	**-0.0087**	**0.708**	**0.13**	**0.9914**	**6.648**

4.2. Satellite-based validations with unseen data

The performance of the proposed ANN models in satellite-based applications has been assessed by injecting daily SST and SSS data from the satellite sources (refer to Sect. 2.2) in place of in-situ sources. The choice of satellite sources for SST and

SSS data is completely subjective to the intended application and their compatibility in terms of spatial and temporal resolutions, whereas geographical coordinates data can be employed from WOA corresponding to the climatological TSL and OHC data. It is recommended to resample SST and SSS data to the WOA grid to eliminate the discrepancies arising from the non-uniform spatial references among the input data. The satellite-based SST, ORA-based SSS, latitude, and longitude data were then given as the inputs to the ANN models for producing TSL and OHC estimates of all the modelling depths considered in this study. Consequently, the model-derived TSL and OHC estimates were compared with Argo-measured in-situ data, and the observed validation results are presented in this section (Table 3 and Figs. 6 and 7).

The performance of the proposed ANN models on satellite-based validation data (Table 3, Figs. 6 and 7) is almost similar to their performance on in-situ-based validation data (Table 2, Figs. 4 and 5). However, the models' performance on satellite-based validation data was marginally low as compared to the in-situ-based validation, likely due to the errors associated with the satellite-based products. According to the statistical results, the R values were observed to be slightly lower by an average percentage decrease of 0.11% across all the modeling depths. Similarly, the RMSE, MBE, MBPE, MAE, and MAPE values were slightly larger than those observed during the in-situ-based validation. This relatively lower performance of the proposed models on the satellite-based validation datasets can be observed by comparing the spatial maps and the distribution of MPE (Figs. A2 and A3). And, 95% of the data have MPE of less than or equal to 0.56% (0.5%) in the cases of overestimation (underestimation), which is higher than those reported in Sect. 4.1. Though the performance of the proposed models on satellite-based data is comparatively lower than the in-situ-based validation, the observed difference in various validation metrics is insignificant. It indicates the efficiency of the proposed models in estimating OHC from satellite data at various depths over the major oceanic basins. However, it should be noted that the validation results presented in this section are subject to vary with the other sources of satellite-based SST and SSS data.

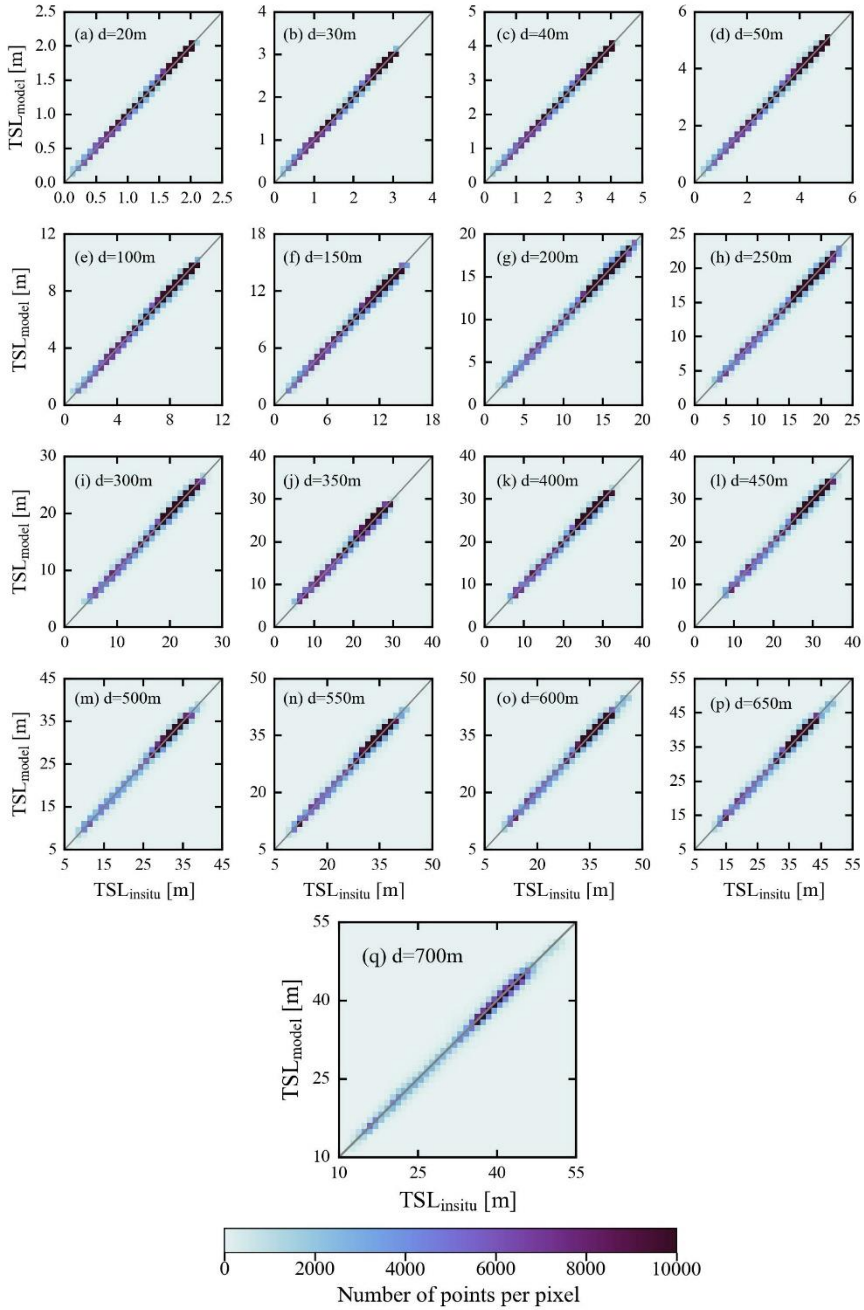

Figure 6. Density scatterplots showing the observed agreement between model-predicted TSL values and in-situ measured TSL values during satellite-based validation.

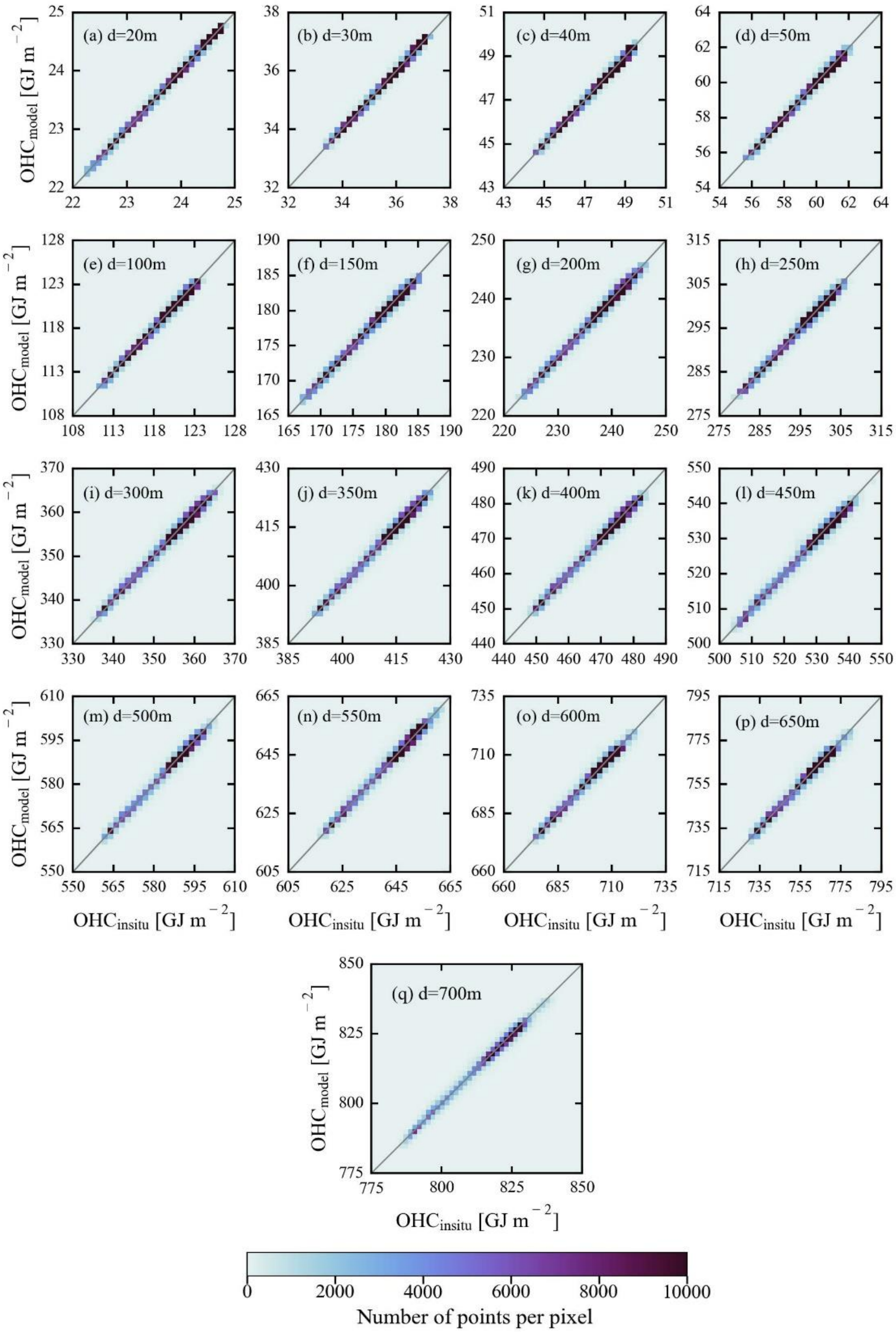

Figure 7. Density scatterplots showing the observed agreement between model-predicted OHC values and in-situ measured OHC values during satellite-based validation.

Table 3. Statistical results from satellite-based validation data of TSL (regular font) and OHC (bold font) against unseen Argo measured in-situ data. The units for the various metrics used in TSL & OHC validations are given as follows: Mean (m & GJ m^{-2}), RMSE (m & GJ m^{-2}), MBE (m & GJ m^{-2}), MBPE (%), MAE (m & GJ m^{-2}), MAPE (%), and intercept (m & GJ m^{-2}).

Depth (m)	N: Data for model development	N: Data for model validation	Mean	R	RMSE	MBE	MBPE	MAE	MAPE	Slope	Intercept
20	801303	536719	1.44	0.9987	0.03	-0.0034	-0.0822	0.016	1.67	0.9960	0.002
			23.91	**0.9987**	**0.04**	**-0.0049**	**-0.0201**	**0.023**	**0.09**	**0.9965**	**0.080**
30	794166	532149	2.15	0.9984	0.04	-0.0008	0.2562	0.027	1.88	0.9961	0.008
			32.85	**0.9984**	**0.06**	**-0.0043**	**-0.0118**	**0.037**	**0.10**	**0.9969**	**0.108**
40	787074	526571	2.85	0.9980	0.07	-0.0054	0.0211	0.041	2.08	0.9969	0.003
			47.78	**0.9980**	**0.09**	**-0.0070**	**-0.0143**	**0.057**	**0.12**	**0.9959**	**0.191**
50	779134	520102	3.54	0.9977	0.09	-0.0060	-0.0262	0.057	2.17	0.9960	0.008
			59.70	**0.9976**	**0.12**	**-0.0056**	**-0.0090**	**0.077**	**0.13**	**0.9956**	**0.257**
100	731065	476709	6.80	0.9966	0.20	-0.0206	-0.2651	0.140	2.56	0.9951	0.013
			119.00	**0.9965**	**0.28**	**-0.0385**	**-0.0322**	**0.194**	**0.16**	**0.9971**	**0.301**
150	712120	460278	9.83	0.9958	0.32	-0.0496	-0.4165	0.229	2.81	0.9897	0.052
			177.97	**0.9956**	**0.44**	**-0.0491**	**-0.0266**	**0.311**	**0.17**	**0.9858**	**2.474**
200	697314	446979	12.64	0.9951	0.43	-0.0091	-0.0022	0.300	2.83	0.9951	0.053
			236.62	**0.9950**	**0.59**	**-0.0200**	**-0.0081**	**0.409**	**0.17**	**0.9929**	**1.653**
250	686378	436906	15.28	0.9948	0.52	-0.0450	-0.2117	0.364	2.79	0.9928	0.065
			295.04	**0.9946**	**0.71**	**-0.0365**	**-0.0119**	**0.492**	**0.17**	**0.9904**	**2.807**
300	678526	429501	17.80	0.9943	0.62	-0.0556	-0.0279	0.428	2.79	0.9837	0.235
			353.29	**0.9941**	**0.83**	**-0.0271**	**-0.0071**	**0.571**	**0.16**	**0.9875**	**4.398**
350	672148	423688	20.23	0.9939	0.71	-0.1052	-0.3291	0.494	2.80	0.9846	0.206
			411.40	**0.9936**	**0.95**	**-0.0381**	**-0.0086**	**0.655**	**0.16**	**0.9847**	**6.264**
400	666605	418686	22.57	0.9935	0.79	-0.0450	-0.0422	0.540	2.72	0.9869	0.252
			469.39	**0.9933**	**1.06**	**-0.0103**	**-0.0017**	**0.723**	**0.15**	**0.9860**	**6.557**
450	661336	413987	24.83	0.9934	0.87	-0.1234	-0.4559	0.586	2.67	0.9898	0.129
			527.25	**0.9931**	**1.17**	**-0.1694**	**-0.0316**	**0.792**	**0.15**	**0.9854**	**7.508**
500	654880	408240	27.03	0.9934	0.91	-0.0707	-0.2034	0.605	2.50	0.9924	0.134
			585.03	**0.9933**	**1.21**	**-0.0909**	**-0.0151**	**0.807**	**0.14**	**0.9874**	**7.293**
550	649850	403357	29.14	0.9932	0.97	-0.0484	-0.0768	0.636	2.40	0.9887	0.280
			642.69	**0.9929**	**1.30**	**0.0315**	**0.0053**	**0.851**	**0.13**	**0.9876**	**8.021**
600	645150	398855	31.21	0.9930	1.03	-0.0431	-0.0139	0.675	2.38	0.9861	0.392
			700.28	**0.9927**	**1.39**	**0.0242**	**0.0039**	**0.906**	**0.13**	**0.9850**	**10.52**
650	640479	392921	33.18	0.9926	1.11	0.0193	0.1132	0.719	2.37	0.9925	0.267
			757.74	**0.9924**	**1.48**	**0.0092**	**0.0015**	**0.957**	**0.13**	**0.9880**	**9.090**
700	633004	388469	35.13	0.9926	1.16	-0.1917	-0.4560	0.763	2.34	0.9835	0.387
			815.15	**0.9922**	**1.56**	**-0.2400**	**-0.0290**	**1.029**	**0.13**	**0.9813**	**14.982**
Weighted average				0.9950	0.83	-0.0657	-0.1645	0.554	2.54	0.9909	0.224
				0.9948	**1.15**	**-0.0566**	**-0.0104**	**0.763**	**0.14**	**0.9896**	**7.799**

4.3. Comparison with the contemporary satellite-based OHC models

Comparison of our ANN models with the existing models is crucial to determine the relative uncertainty in the OHC estimates. Previously, an ANN algorithm suite was developed by the National Remote Sensing Centre (NRSC) of ISRO to disseminate the daily OHC products over the North Indian Ocean (40°E-120°E, 0°-30°N) at a spatial resolution of 0.25 degree (Ali et al., 2012; Jagadeesh et al., 2015). This algorithm suite includes ANN models to estimate OHC at multiple depth extents such as 50 m, 100 m, 150 m, 200 m, 300 m, 500 m, and 700 m for the given input data of SSHA, SST, and $OHC_{clim,d}$. NRSC-ANN models estimate OHC changes by employing the satellite altimetry-based SSHA data from AVISO (Archiving, Validation, and Interpretation of Satellite Oceanographic data) data portal, SST from the Advanced Microwave Scanning Radiometer-2 onboard JAXA's Global Change Observation Mission - Water (GCOM-W1), and climatological OHC from the World Ocean Atlas-2009 monthly climatological CTD profiles. The multilayer perceptron regressor algorithm of neural networks with three hidden layers was used to estimate OHC of all seven depth extents. The number of data points used to develop and validate the NRSC-ANN algorithm were 11472 and 2479, respectively. To compute in-situ OHC at different depths, this algorithm employed the Celsius scale, in-situ temperature, and average density data instead of the Kelvin scale, conservative temperature, and instantaneous density, respectively (see Eq. 3 in Jagadeesh et al., 2015).

Validation datasets were prepared for the period of 2017-2020 by computing in-situ OHC in both Kelvin and Celsius scales for the depth extents of 50 m, 100 m, 150 m, 200 m, 300 m, 500 m, and 700 m from Argo program. Daily OHC data were downloaded from the NRSC's Bhuvan portal and collocated with the corresponding Celsius-scaled in-situ OHC data to evaluate the NRSC-ANN models. Similarly, satellite-based SST, ORA-based SSS data, geographical coordinates, and climatological TSL and OHC data were extracted by collocating with Kelvin-scaled in-situ OHC data for our ANN models to generate the OHC estimates. Evaluation of these two OHC estimates was done separately by means of the normalized metrics such as R, MBPE, and MAPE (Table 4).

Table 4. Statistical results observed during the validation of model-derived OHC estimates against in-situ OHC data.

Depth (m)	N	R		MBPE (%)		MAPE (%)	
		NRSC-ANN model	Proposed ANN model	NRSC-ANN model	Proposed ANN model	NRSC-ANN model	Proposed ANN model
50	15595	0.9223	0.9303	-0.0012	0.0227	1.4762	0.1104
100	14546	0.8575	0.8780	-0.3539	0.0303	2.5145	0.1732
150	14303	0.7678	0.8215	-0.6887	-0.0263	3.2401	0.2053
200	13513	0.7169	0.8152	-1.1048	0.0072	3.4667	0.1903
300	12833	0.7732	0.8690	-1.2656	0.0218	3.1671	0.1525
500	12410	0.8965	0.9346	-0.6996	-0.0052	2.3939	0.1073
700	11959	0.9447	0.9628	-0.6214	-0.0370	2.0035	0.0891

As expected, our ANN models produced relatively high accurate OHC estimates at all depth extents and hence yielded higher correlation coefficients and lower errors as compared to the NRSC-ANN models. The accuracy of OHC estimates produced by our ANN model also increased with depth in contrast to that of NRSC-ANN OHC estimates. Determination of key input parameters based on a precise theoretical basis, accurate computation of in-situ OHC, and use of suitable ANN architectures for each modeling depth enabled our ANN models to produce accurate OHC estimates.

It should be mentioned that SSHA is the combined outcome of temperature (thermosteric), salinity (halosteric), and water mass changes in the oceanic water column. The direct use of satellite altimeter-derived SSHA without eliminating halosteric and water mass change components results in weaker correlations with OHC. Moreover, the different time spans were used in the computation of the mean sea level at AVISO (1993-2012) and monthly climatology data at WOA09 (1955-2006). The combination of merged SSHA data from AVISO/CMEMS and climatological OHC data from WOA could lead to discrepancies in OHC estimates. Hence, the prime criterion followed in determining the input parameters in the current study is the theoretical relationship between the input and output parameters rather than the direct usage of all the relevant parameters. The one-to-one relationship between OHC and TSL is employed in the OHC modeling. To arrive at TSL, the theoretical dependency of TSL on temperature, salinity, and geographical coordinates is considered in TSL modeling work. However, SSHA and climatological OHC data of the same base period are desirable and can be used in OHC (TSL) modeling if available in the future.

Celsius scale can be used to compute in-situ OHC where the temperature is always on the positive side. The usage of the Celsius scale when the temperatures are less than zero and greater than the seawater freezing point is not appropriate because of the negative values. In addition, the conservative temperature is an accurate variable to calculate OHC compared to the measured in-situ temperature or potential temperature (IOC et al., 2010; Pawlowicz, 2013). Thus, conservative temperatures in the absolute scale (Kelvin scale) are used to compute in-situ OHC estimates in the current study. On the other hand, employing instantaneous density rather than average density value is important to account for the dynamic variations in seawater density.

The vertical distribution of conservative temperature follows a non-linear profile with a mixed layer at the top, a thermocline at the middle, and a deep ocean layer at the bottom. This suggests that it is appropriate to customize the ANN hyperparameters for each modeling depth. In this study, hyperparameter tuning was performed for each modeling depth and it resulted in a better understanding of OHC patterns at various depth extents. Though a clear improvement was achieved with the proposed OHC models, relatively lower correlations were observed for our ANN models in the depth range of 100-300 m over the North Indian Ocean (refer to Table 4). It indicates that the ANN models less generalized the OHC patterns at the intermediate depths over the North Indian Ocean and the corresponding underlying factors are discussed in the following section. Nevertheless, the observed results demonstrated that the proposed ANN models are capable of improving the accuracy and quality of OHC products through the ocean thermal expansion method.

4.4. Potential sources of uncertainty in OHC estimates

The relationship between the surficial parameters (SST and SSS) and depth-integrated parameters (TSL and OHC) is the prime factor determining the efficiency of the proposed OHC models of various depth extents (Klemas and Yan, 2014). This relationship is expected to account for a wide range of geophysical processes including ocean currents, vertical mixing (upwelling/downwelling), stratification, fronts, gyres, eddies, and air-sea interface processes. In addition, different climate modes and oscillations, solar radiation, sea ice, phytoplankton growth, freshwater inputs, and winds can also be considered in this context. Monthly climatological CTD profiles obtained from WOA-18 were objectively analyzed to calculate the mean SST and SSS fields over a period of 1955-2017. Hence, these climatological data along with real-time SST and SSS data enabled the ANN models to better generalize the prevailing geophysical processes and subsequent patterns in TSL & OHC of various depth extents. The same can be perceived from the improved accuracy levels observed during the validations carried out on unseen data (refer to Sects. 4.1 and 4.2) and the comparison with NRSC-OHC products (Sect. 4.3).

It should be noted that the established relationship between the input parameters (surficial and climatological) and output parameters (TSL & OHC patterns) may not hold great in the events of complex geophysical processes where the physical oceanographic conditions differ significantly from the prevailing conditions. Moreover, the relative contributions of these geophysical processes are subject to vary depending on the time and location of the water parcel in oceans. The slightly lower accuracy of the proposed ANN models can be attributed to the influence of these complex geophysical processes. The in-situ and satellite-based retrieval of all these atmospheric/surface/subsurface processes and their incorporation into the ANN models is difficult because of the scarcity/sparsity of the required datasets in different spatial, temporal, and vertical scales. The above factors constitute a potential source of uncertainty in OHC estimates and reduce the generalization ability of the model. Hence, it is advisable to carry out vicarious calibration with the help of contemporary in-situ CTD profiles before adopting the OHC estimates for further scientific analyses of specific interest in both regional and global scales. Further efforts are needed to better understand, quantify, and eliminate the different sources of identified uncertainties caused by the complex geophysical processes. More number of in-situ CTD profiles are required to be collected and analyzed in such oceanic regions to better account for the associated complex patterns and processes.

5. Spatiotemporal variability of OHC

Here, we present the long-term variability of model-derived OHC and its comparison with the existing global OHC products for the period 1993-2020. The time period (1993-2020) was chosen based on the availability of satellite-based input data to generate the model-derived OHC estimates and the existing OHC products considered. Thus, model-derived OHC estimates were generated from 1993 to 2020 at a spatial resolution of 0.25° and computed annual time series of model-derived OHC anomalies (OHCA) with reference to the 1993-2020 long-term mean. It is worth mentioning that the model-derived OHCA estimates presented in this section represent heat changes in both shallower and deep oceanic basins of bathymetry levels $\geq$ 20 m. The bathymetry values of each pixel were rounded off to the nearest and lowest modeling depth (d) with the help of

GEBCO-2020 bathymetry data, and the corresponding $OHCA_d$ values were considered for that pixel (GEBCO Compilation Group, 2020).

On the other hand, OHCA time series annual maps obtained from various global OHC products such as the National Centers for Environmental Information (NCEI), Institute of Atmospheric Physics (IAP), Pacific Marine Environmental Laboratory (PMEL), and OPEN-LSTM have been employed for comparison. NCEI employs the Objective analysis method on in-situ CTD profile data of World Ocean Database-2009 and estimates annual OHCA at a spatial resolution of 1° with reference to the 1955-2006 long-term mean (Levitus et al., 2012). Similarly, IAP employs the ensemble optimal interpolation with a dynamic ensemble approach on in-situ CTD profile data of World Ocean Database-2013 and distributes monthly OHC estimates at a spatial resolution of 1° (Cheng et al., 2017). Annual OHC means were computed from IAP monthly OHC data, and annual OHCA estimates were generated with reference to the 1993-2020 long-term mean. Recently, PMEL has developed a random forest regression model to predict OHCA of 0-40 m, 40-90 m, 90-190 m, 190-290 m, 290-450 m, 450-700 m, 700-950 m, 950-1450 m, etc with reference to the 1993-2022 long-term mean. This PMEL random forest regression model employs satellite-based SST, SSHA, latitude, longitude, and time data to predict weekly OHCA estimates at a spatial resolution of 0.25° (Lyman and Johnson, 2023). In the current study, PMEL layer-wise OHCA estimates from surface to 700 m have been summed up at each pixel to arrive at weekly OHCA spatial maps and subsequently computed corresponding annual OHCA estimates. Similarly, Su et al., (2021) have developed a long short-term memory neural network method to produce monthly OHC estimates (OPEN-LSTM) at a spatial resolution of 1°. OPEN-LSTM employs satellite-based SSHA, SST, zonal and meridional components of sea surface wind, latitude, longitude, and day of the year to predict monthly OHC. Annual OHC means were computed from OPEN-LSTM monthly OHC data, and annual OHCA estimates were generated with reference to the 1993-2020 long-term mean.

Model-derived annual OHCA estimates were regridded to 1° spatial resolution to maintain uniform spatial reference among all the OHC products considered. As the proposed models are built for open oceanic regions, the regions covered by sea ice are masked in both the north and south poles by verifying the corresponding sea ice concentration data obtained from the National Snow and Ice Data Center (Meier et al., 2021). Subsequently, long-term variability maps (Fig. 8) and time series plots (Fig. 9) were produced to compare model-derived OHC estimates with the existing global OHC products. Further, the information on percentage variance explained (PVE) by the observed long-term trend values is provided to realise the short-term trends or periodic signals in OHC variability (Fig. A4). Higher PVE values indicate the persistent increase or decrease in OHC throughout the study period, and vice versa.

Lower magnitudes of long-term warming/cooling trends (± 0.05 GJ m^{-2} $Year^{-1}$) are observed throughout the global ocean (Fig. 8a). The corresponding PVE values are observed to be very low (≤ 30%) which infer the intermittent trends in majority of the global ocean rather than persistent warming/cooling (Fig. A4a). The same can be observed from the non-linear distribution of OHCA time series indicating short-term periods of alternate warming and cooling during the study period (Fig. 9). However, the oceanic regions linked with Kusoshio current, Gulf stream, Antarctic circumpolar current, North Atlantic cold blob, southeastern Pacific are experiencing relatively higher magnitudes of persistent warming/cooling (± 0.1 to 0.15 GJ m^{-2} $Year^{-1}$, PVE 50-90%).

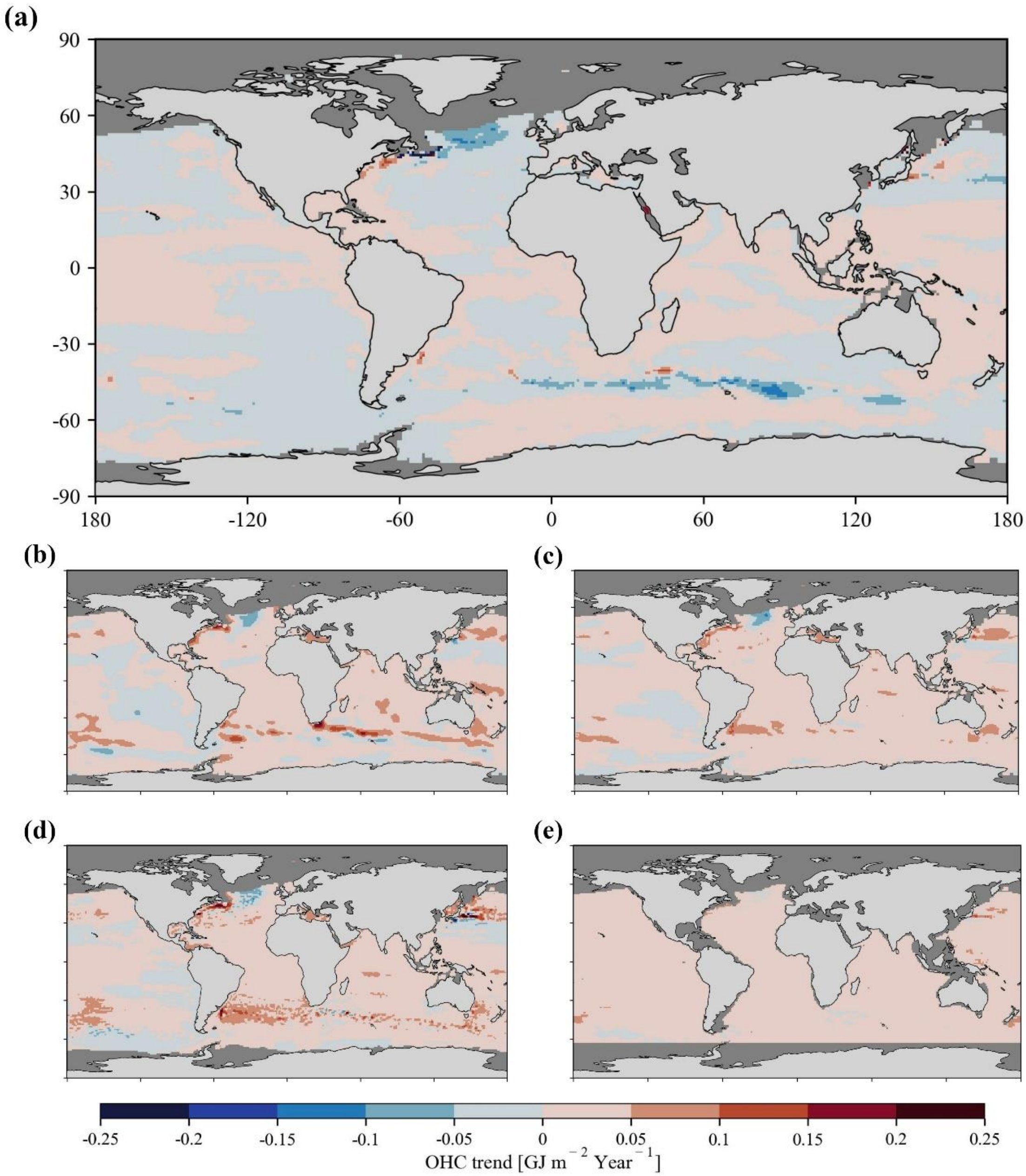

Figure 8. Spatial maps showing the long-term trends of OHC obtained from (a) the current model, (b) NCEI, (c) IAP, (d) PMEL, and (e) OPEN-LSTM products. Note that the oceanic regions shallower than 20 m depth and/or covered with sea ice are masked with a dark gray color.

The spatial patterns of OHC trends observed from NCEI (Fig. 8b), IAP (Fig. 8c), and PMEL (Fig. 8d) products are almost similar and relatively more warming regions compared to the model-derived OHC estimates (Fig. 8a). NCEI, IAP, and PMEL products indicating persistent warming conditions over the vast oceanic regions of the Pacific, Atlantic, and Indian Oceans with higher PVE values (Figs. A4b-A4d). The same can be observed from the persistent long-term warming throughout the study period (Fig. 9). On the other hand, OPEN-LSTM OHC estimates indicate lower warming patterns all over the globe except the North Atlantic cold blob and some parts of the Antarctic circumpolar current (Fig. 8e) with higher PVE values over vast oceanic regions of Pacific, Atlantic, and Indian Oceans (Fig. A4e). As a result, OPEN-LSTM also showed persistent long-term warming throughout the study period (Fig. 9).

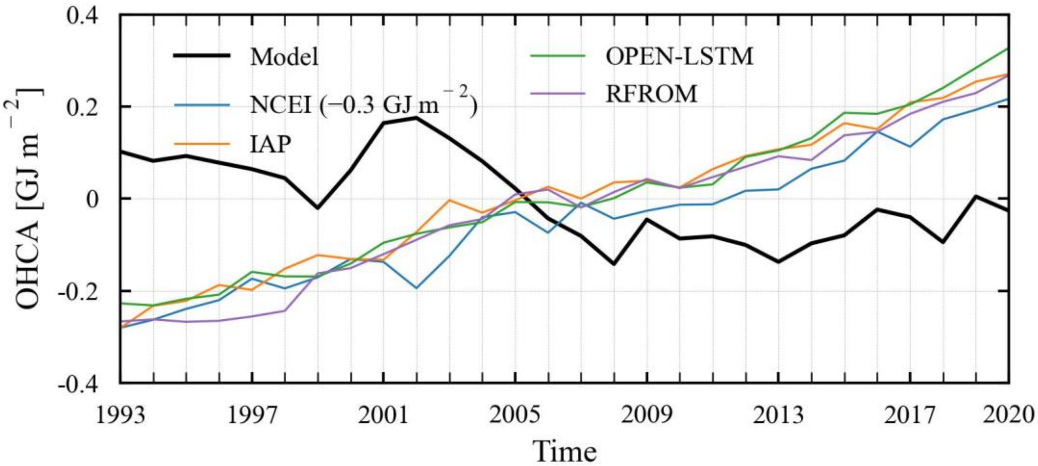

Figure 9. Time series distribution of global mean OHCA obtained from the current model and the existing OHC products observed over the period 1993-2020. Note that the NCEI time series has been shifted by subtracting 0.3 GJ m^{-2} to better compare with the remaining OHC time series plots.

The observed time series plots have indicated contrasting trends between the current OHC model and the existing products. The time series plot of model-derived OHCA has indicated alternate periods of short-term cooling and warming during the study period. Global open oceans have witnessed a cooling trend of −0.017 GJ m^{-2} Year^{-1} (PVE 76.99%) during 1993-1999, a warming trend of +0.069 GJ m^{-2} Year^{-1} (PVE 92.73%) during 2000-2002, a cooling trend of −0.054 GJ m^{-2} Year^{-1} (PVE 99.71%) during 2003-2008, and a warming trend of +0.007 GJ m^{-2} Year^{-1} (PVE 36.50%) during 2009-2020. The observed results indicate the efficiency of the current model by capturing the ocean cooling during 2003-2006 (Loehle, 2009; Lyman et al., 2006) and the global warming hiatus during 1998-2013 (Trenberth, 2015). Whereas the observed time series plots of NCEI, IAP, PMEL, and OPEN-LSTM products indicated persistent warming trends of +0.017 GJ m^{-2} Year^{-1} (PVE 95.75%), +0.019 GJ m^{-2} Year^{-1} (PVE 97.94%), +0.0198 GJ m^{-2} Year^{-1} (PVE 97.19%), and +0.0195 GJ m^{-2} Year^{-1} (PVE 97.48%), respectively. However, full-depth pan-global OHCA estimates by including OHC estimates over ice-covered oceanic regions are required to substantiate these global ocean cooling and global warming hiatus signatures, and to realize the role of excess heat added by anthropogenic climate change.

6. Conclusion

Accurate reconstruction of OHC and analysis of its regional patterns and long-term global records are critical for estimating the Earth's energy imbalance and understanding the evolution of climate change. Owing to the lack of instrumentation to cover geographic and depth ranges, OHC estimates from the in-situ measured temperatures are temporally limited and insufficiently widespread to capture its spatiotemporal changes and structures. OHC estimates from either different mapping methods or Ocean reanalyses (ORAs) have yielded large uncertainties in past studies. Thus, improving OHC estimates through a novel satellite-based method is the major step forward in overcoming sparse observations and reducing the uncertainty in OHC trends. In this study, we proposed an artificial network model to estimate OHC changes in global oceans. The proposed ANN model incorporates the ocean thermal expansion method as a promising tool to estimate OHC changes from satellite data. Accurate implementation of the ocean thermal expansion method was challenging due to the inability of the present-day satellite systems to directly measure the ocean thermal expansion/contraction component. In this study, we proposed a satellite-based novel approach to better implement the ocean thermal expansion method by establishing a relationship between the surficial parameters such as SST & SSS and subsurface T-S profiles. This model predicts the depth-integrated TSL component by making use of SST & SSS data and then utilizes the predicted TSL to estimate OHC changes. For this application, we developed ANN models for TSL and OHC of various depth extents such as 20 m, 30 m, 40 m, 50 m, 100 m, 150 m, 200 m, 250 m, 300 m, 350 m, 400 m, 450 m, 500 m, 550 m, 600 m, 650 m, and 700 m. The performance of these TSL & OHC models was assessed by carrying out in-situ-based and satellite-based validations by using the unseen in-situ CTD profiles from the Argo program. Observed high correlations and low errors indicated that the proposed ANN models performed exceptionally good on unseen data of all modeling depths without any overfitting and can be used in conjunction with the sea ice thermodynamics-based OHC models of the ice-covered oceans (Vijay and Shanmugam, 2022) to better study the pan-global OHC changes by covering both open and ice-covered oceanic regions of varying bathymetry levels (≥ 20 m).

The model development and validation databases were prepared by using in-situ CTD profiles obtained from the Argo program and collocated with the corresponding satellite-based daily data of SST (AVHRR v2.1) and SSS (ORAS5). The multilayer perceptron regressor algorithm of deep neural networks was used and its architecture was optimized by evaluating different combinations of hyperparameters for each modeling depth using the particle swarm optimization technique. Precise consideration of theoretical aspects in the selection of input parameters, accurate computation of in-situ OHC, and customized ANN architectures enabled the proposed models to establish the accurate relationships between the surficial parameters and depth-integrated OHC (TSL) of various depths extents. The overall performance of the proposed models on satellite data was good, suggesting that these models can be used for a variety of applications subjected to the accuracy requirements and can produce accurate satellite-based OHC (TSL) estimates at various depth extents than previously possible. However, the influence of complex geophysical processes on the generalization ability of ANN models is discussed, and realized that the proposed models relatively less generalized the data in the events of complex geophysical processes. Further research should focus on implementation of these models over the oceanic regions with complex geophysical processes. More number of in-situ CTD profiles need to be collected and analyzed in such oceanic regions to better account the associated complex patterns. However, the scope of the current research includes minimizing the observed marginal gap by exploring new methods/parametrizations in satellite-based OHC modeling approaches.

CRediT authorship contribution statement

Vijay Prakash Kondeti: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Software, Validation, Visualization, and Writing - original draft. **Palanisamy Shanmugam:** Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, and Writing - review & editing.

Code and Data availability

Data will be made available on request.

Declaration of competing interest

The authors declare no known competing financial or personal interests in this paper.

Acknowledgement

This research work was supported by The Prime Minister's Research Fellows (PMRF) Scheme and in part by the National Geospatial Programme (NGP) of Department of Science and Technology of Government of India (Grant No: OEC1819150DSTXPSHA). The authors are thankful to the Argo program for providing in-situ CTD profiles. They are grateful to NOAA for WOD-18, WOA-18, and SST data; CMEMS for SSS data; NCEI, IAP, PMEL, and Science Data Bank for OHC estimates; NSIDC for sea ice concentration data; and GEBCO for bathymetry data. The authors are thankful to the two anonymous researchers for their constructive comments and recommendations.

References

Abraham, J. P., Baringer, M., Bindoff, N. L., Boyer, T., Cheng, L. J., Church, J. A., Conroy, J. L., Domingues, C. M., Fasullo, J. T., Gilson, J., Goni, G., Good, S. A., Gorman, J. M., Gouretski, V., Ishii, M., Johnson, G. C., Kizu, S., Lyman, J. M., Macdonald, A. M., Minkowycz, W. J., Moffitt, S. E., Palmer, M. D., Piola, A. R., Reseghetti, F., Schuckmann, K., Trenberth, K. E., Velicogna, I., and Willis, J. K.: 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, 2013.

Ali, M. M., Jagadeesh, P. S. V., Lin, I. I., and Hsu, J. Y.: A neural network approach to estimate tropical cyclone heat potential in the Indian Ocean, IEEE Geosci. Remote Sens. Lett., 9, 1114–1117, https://doi.org/10.1109/LGRS.2012.2190491, 2012.

Balmaseda, M. A., Hernandez, F., Storto, A., Palmer, M. D., Alves, O., Shi, L., Smith, G. C., Toyoda, T., Valdivieso, M., Barnier, B., Behringer, D., Boyer, T., Chang, Y. S., Chepurin, G. A., Ferry, N., Forget, G., Fujii, Y., Good, S., Guinehut, S., Haines, K., Ishikawa, Y., Keeley, S., Köhl, A., Lee, T., Martin, M. J., Masina, S., Masuda, S., Meyssignac, B., Mogensen, K., Parent, L., Peterson, K. A., Tang, Y. M., Yin, Y., Vernieres, G., Wang, X., Waters, J., Wedd, R., Wang, O., Xue, Y., Chevallier, M., Lemieux, J. F., Dupont, F., Kuragano, T., Kamachi, M., Awaji, T., Caltabiano, A., Wilmer-Becker, K., and Gaillard, F.: The ocean reanalyses intercomparison project (ORA-IP), J. Oper. Oceanogr., 8, s80–s97, https://doi.org/10.1080/1755876X.2015.1022329, 2015.

Baxter, J. M.: Explaining Ocean Warming: Causes, scale, effects and consequences, edited by: Laffoley, D. and Baxter, J. M., IUCN, International Union for Conservation of Nature, https://doi.org/10.2305/IUCN.CH.2016.08.en, 2016.

Beech, N., Rackow, T., Semmler, T., Danilov, S., Wang, Q., and Jung, T.: Long-term evolution of ocean eddy activity in a warming world, Nat. Clim. Chang., 12, 910–917, https://doi.org/10.1038/s41558-022-01478-3, 2022.

Boyer, T. P., Baranova, O. K., Coleman, C., Garcia, H. E., Grodsky, A., Locarnini, R. A., Mishonov, A. V, Paver, C. R., Reagan, J. R., Seidov, D., Smolyar, I. V, Weathers, K. W., and Zweng, M. M.: NOAA Atlas NESDIS 87. World Ocean Database 2018, 1–207, 2018.

Chacko, N., Dutta, D., Ali, M. M., Sharma, J. R., and Dadhwa, V. K.: Near-real-time availability of ocean heat content over the north indian ocean, IEEE Geosci. Remote Sens. Lett., 12, 1033–1036, https://doi.org/10.1109/LGRS.2014.2375196, 2015.

Chambers, D. P., Tapley, B. D., and Stewart, R. H.: Long-period ocean heat storage rates and basin-scale heat fluxes from TOPEX, J. Geophys. Res. Ocean., 102, 10525–10533, https://doi.org/10.1029/96JC03644, 1997.

Cheng, L., Zhu, J., and Sriver, R. L.: Global representation of tropical cyclone-induced ocean thermal changes using Argo data – Part 2 : Estimating air – sea heat fluxes and ocean heat content changes, Ocean Sci. Discuss., 11, 2907–2937, https://doi.org/10.5194/osd-11-2907-2014, 2014.

Cheng, L., Trenberth, K. E., Fasullo, J., Boyer, T., Abraham, J., and Zhu, J.: Improved estimates of ocean heat content from 1960 to 2015, Sci. Adv., 3, 1–10, https://doi.org/10.1126/sciadv.1601545, 2017.

Cheng, L., Foster, G., Hausfather, Z., Trenberth, K. E., and Abraham, J.: Improved Quantification of the Rate of Ocean Warming, J. Clim., 35, 4827–4840, https://doi.org/10.1175/jcli-d-21-0895.1, 2022.

Huang, B., Liu, C., Banzon, V., Freeman, E., Graham, G., Hankins, B., Smith, T., and Zhang, H. M.: Improvements of the Daily Optimum Interpolation Sea Surface Temperature (DOISST) Version 2.1, J. Clim., 34, 2923–2939, https://doi.org/10.1175/JCLI-D-20-0166.1, 2021.

IOC, SCOR, and IAPSO: The international thermodynamic equation of seawater-2010: Calculation and use of thermodynamic properties Intergovernmental Oceanographic Commission, 2010.

IPCC: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, J. Cryst. Growth, 2014.

IPCC: Changing Ocean, Marine Ecosystems, and Dependent Communities, 447–588 pp., https://doi.org/10.1017/9781009157964.013, 2022.

Irrgang, C., Saynisch, J., and Thomas, M.: Estimating global ocean heat content from tidal magnetic satellite observations, Sci. Rep., 9, 1–8, https://doi.org/10.1038/s41598-019-44397-8, 2019.

Jagadeesh, P. S. V. and Ali, M. M.: Estimation of upper ocean heat content from remote sensing observations in the Arabian Sea, Remote Sens. Model. Atmos. Ocean. Interact., 6404, 64041C, https://doi.org/10.1117/12.699319, 2006.

Jagadeesh, P. S. V., Suresh Kumar, M., and Ali, M. M.: Estimation of Heat Content and Mean Temperature of Different Ocean Layers, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., 8, 1251–1255, https://doi.org/10.1109/JSTARS.2015.2403877, 2015.

Kennedy, J., & Eberhart, R.: Particle Swarm Optimization, in: In Proceedings of ICNN’95-international conference on neural networks, IEEE, 1942–1948, https://doi.org/10.1007/978-3-319-46173-1_2, 1995.

Klemas, V. and Yan, X. H.: Subsurface and deeper ocean remote sensing from satellites: An overview and new results, Prog. Oceanogr., 122, 1–9, https://doi.org/10.1016/j.pocean.2013.11.010, 2014.

L’Ecuyer, T. S., Beaudoing, H. K., Rodell, M., Olson, W., Lin, B., Kato, S., Clayson, C. A., Wood, E., Sheffield, J., Adler, R., Huffman, G., Bosilovich, M., Gu, G., Robertson, F., Houser, P. R., Chambers, D., Famiglietti, J. S., Fetzer, E., Liu, W. T., Gao, X., Schlosser, C. A., Clark, E., Lettenmaier, D. P., and Hilburn, K.: The observed state of the energy budget in the early twenty-first century, J. Clim., 28, 8319–8346, https://doi.org/10.1175/JCLI-D-14-00556.1, 2015.

Levitus, S., Antonov, J. I., Boyer, T. P., Locarnini, R. A., Garcia, H. E., and Mishonov, A. V.: Global ocean heat content 1955-2008 in light of recently revealed instrumentation problems, Geophys. Res. Lett., 36, 1–5, https://doi.org/10.1029/2008GL037155, 2009.

Levitus, S., Antonov, J. I., Boyer, T. P., Baranova, O. K., Garcia, H. E., Locarnini, R. A., Mishonov, A. V, Reagan, J. R., Seidov, D., Yarosh, E. S., and Zweng, M. M.: World ocean heat content and thermosteric sea level change (0–2000 m), 1955–2010, Geophys. Res. Lett., 39, 1–5, https://doi.org/10.1029/2012GL051106, 2012.

Liang, X., Wunsch, C., Heimbach, P., and Forget, G.: Vertical redistribution of oceanic heat content, J. Clim., 28, 3821–3833, https://doi.org/10.1175/JCLI-D-14-00550.1, 2015.

Loehle, C.: Cooling of the global ocean since 2003, Energy Environ., 20, 101–104, https://doi.org/10.1260/095830509787689141, 2009.

Lyman, J. M. and Johnson, G. C.: Global High-Resolution Random Forest Regression Maps of Ocean Heat Content Anomalies Using In Situ and Satellite Data, J. Atmos. Ocean. Technol., 40, 575–586, https://doi.org/10.1175/JTECH-D-22-0058.1, 2023.

Lyman, J. M., Willis, J. K., and Johnson, G. C.: Recent cooling of the upper ocean, Geophys. Res. Lett., 33, 1–5, https://doi.org/10.1029/2006GL027033, 2006.

Marti, F., Blazquez, A., Meyssignac, B., Ablain, M., Barnoud, A., Fraudeau, R., Jugier, R., Chenal, J., Larnicol, G., Pfeffer, J., Restano, M., and Benveniste, J.: Monitoring the ocean heat content change and the Earth energy imbalance from space altimetry and space gravimetry, Earth Syst. Sci. Data, 14, 229–249, https://doi.org/10.5194/essd-14-229-2022, 2022.

Meyssignac, B., Boyer, T., Zhao, Z., Hakuba, M. Z., Landerer, F. W., Stammer, D., Köhl, A., Kato, S., L’Ecuyer, T., Ablain, M., Abraham, J. P., Blazquez, A., Cazenave, A., Church, J. A., Cowley, R., Cheng, L., Domingues, C., Giglio, D., Gouretski, V., Ishii, M., Johnson, G. C., Killick, R. E., Legler, D., Llovel, W., Lyman, J., Palmer, M. D., Piotrowicz, S., Purkey, S., Roemmich, D., Roca, R., Savita, A., Schuckmann, K. von, Speich, S., Stephens, G., Wang, G. G., Wijffels, S. E., and Zilberman, N.: Measuring global ocean heat content to estimate the earth energy imbalance, Front. Mar. Sci., 6, 1–31, https://doi.org/10.3389/fmars.2019.00432, 2019.

Momin, I. M., Sharma, R., and Basu, S.: Satellite-derived heat content in the tropical Indian Ocean, Remote Sens. Lett., 2, 269–277, https://doi.org/10.1080/01431161.2010.519001, 2011.

Ni, Q., Zhai, X., LaCasce, J. H., Chen, D., and Marshall, D. P.: Full-Depth Eddy Kinetic Energy in the Global Ocean Estimated From Altimeter and Argo Observations, Geophys. Res. Lett., 50, https://doi.org/10.1029/2023GL103114, 2023.

Palmer, M. D., Roberts, C. D., Balmaseda, M., Chang, Y. S., Chepurin, G., Ferry, N., Fujii, Y., Good, S. A., Guinehut, S., Haines, K., Hernandez, F., Köhl, A., Lee, T., Martin, M. J., Masina, S., Masuda, S., Peterson, K. A., Storto, A., Toyoda, T., Valdivieso, M., Vernieres, G., Wang, O., and Xue, Y.: Ocean heat content variability and change in an ensemble of ocean reanalyses, Clim. Dyn., 49, 909–930, https://doi.org/10.1007/s00382-015-2801-0, 2017.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., and Vanderplas, J.: Scikit-learn: Machine learning in Python, J. Mach. Learn. Res., 12, 2825–2830, 2011.

Polito, P. S., Sato, O. T., and Liu, W. T.: Characterization and validation of the heat storage variability from TOPEX/Poseidon at four oceanographic sites, J. Geophys. Res. Ocean., 105, 16911–16921, https://doi.org/10.1029/1999JC000048, 2000.

Resplandy, L., Keeling, R. F., Eddebbar, Y., Brooks, M. K., Wang, R., Bopp, L., Long, M. C., Dunne, J. P., Koeve, W., and Oschlies, A.: Quantification of ocean heat uptake from changes in atmospheric O2 and CO2 composition, Nature, 563, 105–108, https://doi.org/10.1038/s41586-018-0651-8, 2018.

Riser, S. C., Freeland, H. J., Roemmich, D., Wijffels, S., Troisi, A., Belbéoch, M., Gilbert, D., Xu, J., Pouliquen, S., Thresher, A., Le Traon, P. Y., Maze, G., Klein, B., Ravichandran, M., Grant, F., Poulain, P. M., Suga, T., Lim, B., Sterl, A., Sutton, P., Mork, K. A., Vélez-Belchí, P. J., Ansorge, I., King, B., Turton, J., Baringer, M., and Jayne, S. R.: Fifteen years of ocean observations with the global Argo array, Nat. Clim. Chang., 6, 145–153, https://doi.org/10.1038/nclimate2872, 2016.

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

Sato, O. T., Polito, P. S., and Liu, W. T.: Importance of salinity measurements in the heat storage estimation from TOPEX/POSEIDON, Geophys. Res. Lett., 27, 549–551, https://doi.org/10.1029/1999GL011003, 2000.

Von Schuckmann, K., Palmer, M. D., Trenberth, K. E., Cazenave, A., Chambers, D., Champollion, N., Hansen, J., Josey, S. A., Loeb, N., Mathieu, P. P., Meyssignac, B., and Wild, M.: An imperative to monitor Earth's energy imbalance, Nat. Clim. Chang., 6, 138–144, https://doi.org/10.1038/nclimate2876, 2016.

Shi, Y., & Eberhart, R.: A Modified Particle Swarm Optimizer Algorithm, in: IEEE international conference on evolutionary computation proceedings, 69–73, https://doi.org/10.1109/ICEMI.2007.4350772, 1998.

Su, H., Zhang, H., Geng, X., Qin, T., Lu, W., and Yan, X. H.: OPEN: A new estimation of global ocean heat content for upper 2000 meters from remote sensing data, Remote Sens., 12, https://doi.org/10.3390/rs12142294, 2020.

Su, H., Qin, T., Wang, A., and Lu, W.: Reconstructing ocean heat content for revisiting global ocean warming from remote sensing perspectives, Remote Sens., 13, https://doi.org/10.3390/rs13193799, 2021.

Trenberth, K. E.: Has there been a hiatus?, Science (80-.)., 349, 691–692, https://doi.org/10.1126/science.aac9225, 2015.

Trenberth, K. E., Fasullo, J. T., von Schuckmann, K., and Cheng, L.: Insights into Earth's energy imbalance from multiple sources, J. Clim., 29, 7495–7505, https://doi.org/10.1175/JCLI-D-16-0339.1, 2016.

Trossman, D. S. and Tyler, R. H.: Predictability of Ocean Heat Content From Electrical Conductance, J. Geophys. Res. Ocean., 124, 667–679, https://doi.org/10.1029/2018JC014740, 2019.

Vijay, P. K. and Shanmugam, P.: Artificial Neural Network Model for Estimating Ocean Heat Content in the Sea Ice-Covered Arctic Regions Using Satellite Data, IEEE Access, 10, 109544–109557, https://doi.org/10.1109/ACCESS.2022.3213942, 2022.

White, W. B. and Tai, C.: Inferring interannual changes in global upper ocean heat storage from TOPEX altimetry, J. Geophys. Res. Ocean., 100, 24943–24954, https://doi.org/10.1029/95JC02332, 1995.

Wild, M., Folini, D., Hakuba, M. Z., Schär, C., Seneviratne, S. I., Kato, S., Rutan, D., Ammann, C., Wood, E. F., and König-Langlo, G.: The energy balance over land and oceans: an assessment based on direct observations and CMIP5 climate models, Clim. Dyn., 44, 3393–3429, https://doi.org/10.1007/s00382-014-2430-z, 2015.

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

Zhao, Z.: Propagation of the Semidiurnal Internal Tide: Phase Velocity Versus Group Velocity, Geophys. Res. Lett., 44, 11,942-11,950, https://doi.org/10.1002/2017GL076008, 2017.

Zuo, H., Balmaseda, M. A., and Mogensen, K.: The new eddy-permitting ORAP5 ocean reanalysis: description, evaluation and uncertainties in climate signals, Clim. Dyn., 49, 791–811, https://doi.org/10.1007/s00382-015-2675-1, 2017.

Appendix A

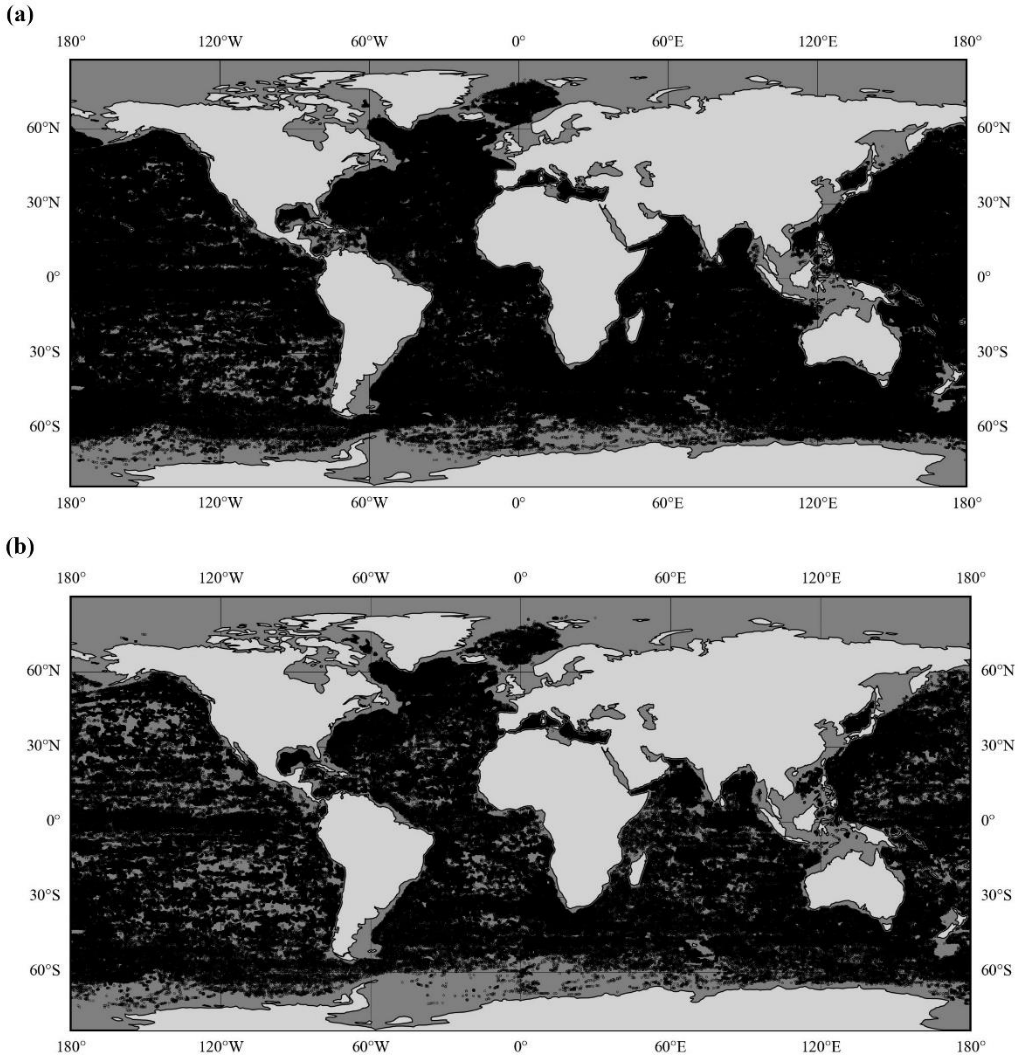

Figure A1. The spatial distribution of in-situ data points used for (a) model development (N=633004 Argo CTD profiles) and (b) validation (N=388469 unseen Argo CTD profiles) in the case of TSL_{700} and OHC_{700}.

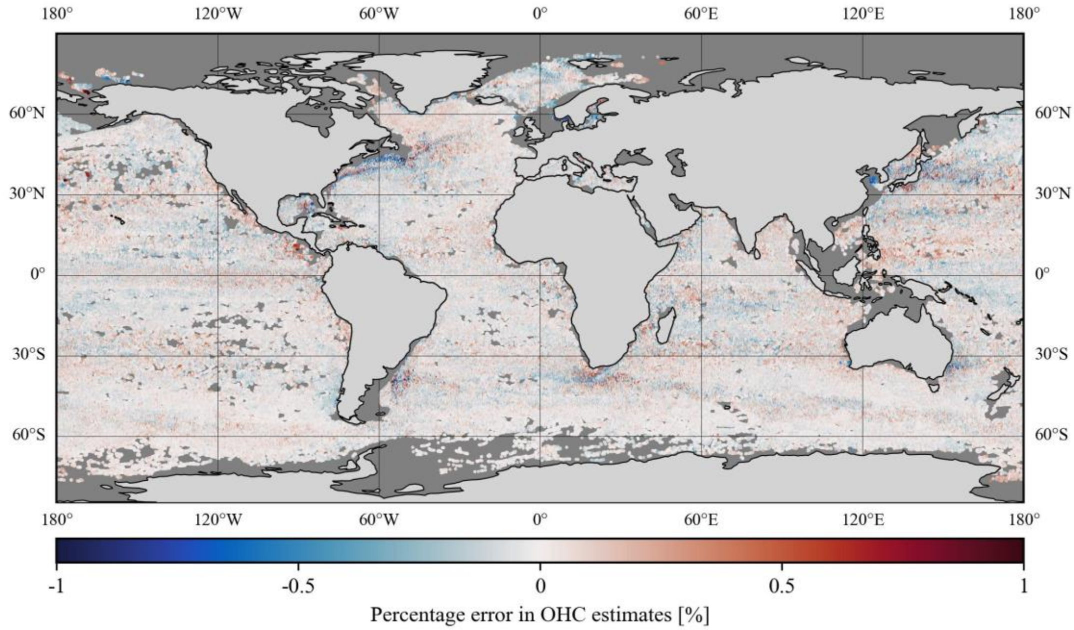

Figure A2. Spatial distribution of mean percentage errors observed during the in-situ-based validation of OHC models. The oceanic regions shallower than 20 m and/or covered with sea ice are marked with a dark gray color.

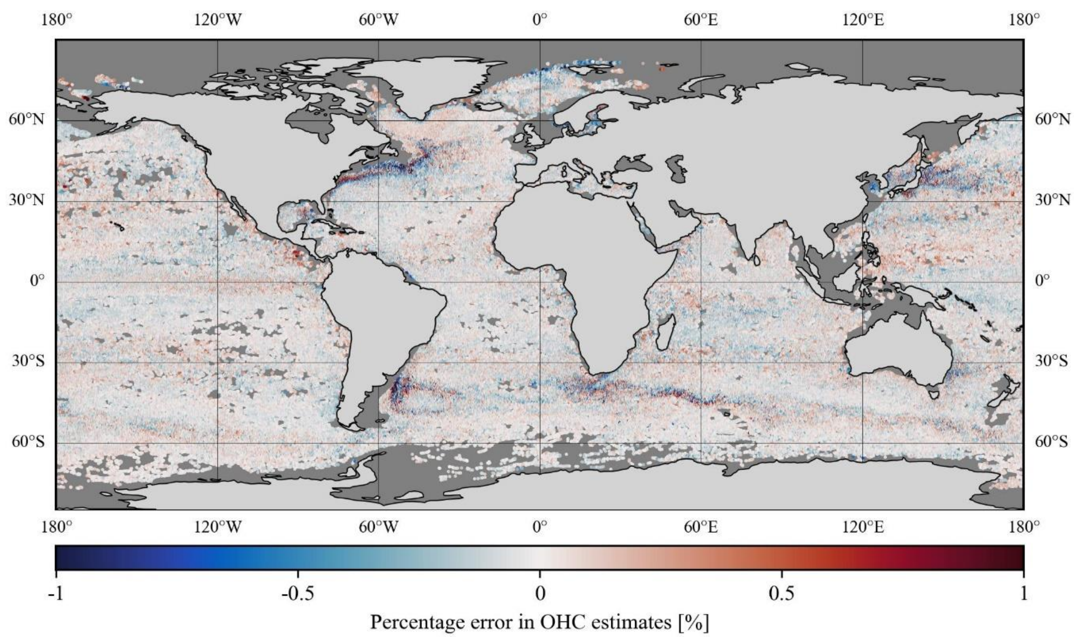

Figure A3. Spatial distribution of mean percentage errors observed during the satellite-based validation of OHC. The oceanic regions shallower than 20 m and/or covered with sea ice are marked with a dark gray color.

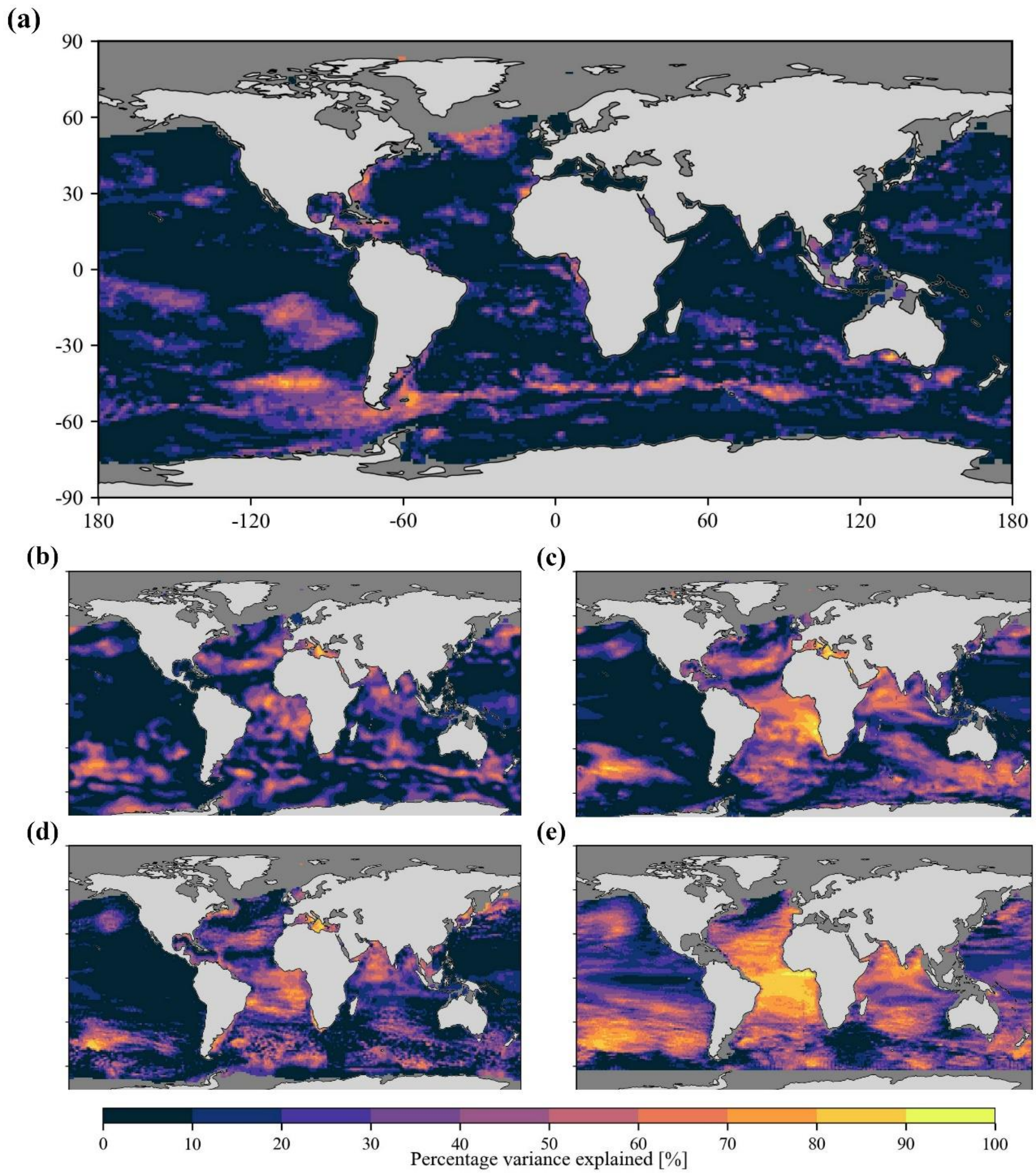

Figure A4. Spatial maps showing the percentage variance explained by the OHC trends obtained from (a) the current model, (b) NCEI, (c) IAP, (d) PMEL, and (d) OPEN-LSTM products. Note that the oceanic regions shallower than 20 m depth and/or covered with sea ice are masked with a dark gray color.