



**HAL**  
open science

## Steric sea level variability (1993-2010) in an ensemble of ocean reanalyses and objective analyses

Andrea Storto, Simona Masina, Magdalena Balmaseda, Stéphanie Guinehut, Yan Xue, Tanguy Szekely, Ichiro Fukumori, Gael Forget, You-Soon Chang, Simon A. Good, et al.

### ► To cite this version:

Andrea Storto, Simona Masina, Magdalena Balmaseda, Stéphanie Guinehut, Yan Xue, et al.. Steric sea level variability (1993-2010) in an ensemble of ocean reanalyses and objective analyses. *Climate Dynamics*, 2017, 49, pp.709-729. 10.1007/s00382-015-2554-9 . insu-03684928

**HAL Id: insu-03684928**

**<https://insu.hal.science/insu-03684928>**

Submitted on 5 Feb 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# *Steric sea level variability (1993-2010) in an ensemble of ocean reanalyses and objective analyses*

Article

Accepted Version

Storto, A., Masina, S., Balmaseda, M., Guinehut, S., Xue, Y., Szekely, T., Fukumori, I., Forget, G., Chang, Y.-S., Good, S. A., Kohl, A., Vernieres, G., Ferry, N., Peterson, K. A., Behringer, D., Ishii, M., Masuda, S., Fujii, Y., Toyoda, T., Yin, Y., Valdivieso, M. ORCID: <https://orcid.org/0000-0002-1738-7016>, Barnier, B., Boyer, T., Lee, T., Gourrion, J., Wang, O., Heimback, P., Rosati, A., Kovach, R., Hernandez, F., Martin, M. J., Kamachi, M., Kuragano, T., Mogensen, K., Alves, O., Haines, K. ORCID: <https://orcid.org/0000-0003-2768-2374> and Wang, X. (2017) Steric sea level variability (1993-2010) in an ensemble of ocean reanalyses and objective analyses. *Climate Dynamics*, 49 (3). pp. 709-729. ISSN 0930-7575 doi: <https://doi.org/10.1007/s00382-015-2554-9> Available at <https://centaur.reading.ac.uk/52358/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

Published version at: <http://link.springer.com/article/10.1007/s00382-015-2554-9>

To link to this article DOI: <http://dx.doi.org/10.1007/s00382-015-2554-9>

Publisher: Springer

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

[www.reading.ac.uk/centaur](http://www.reading.ac.uk/centaur)

## **CentAUR**

Central Archive at the University of Reading

Reading's research outputs online

1           **Steric sea level variability (1993-2010) in an ensemble of ocean**  
2                           **reanalyses and objective analyses**

3 Andrea Storto \* <sup>1</sup>, Simona Masina<sup>1</sup>, Magdalena Balmaseda<sup>2</sup>, Stéphanie Guinehut<sup>3</sup>,  
4 Yan Xue<sup>4</sup>, Tanguy Szekely<sup>5</sup>, Ichiro Fukumori<sup>6</sup>, Gael Forget<sup>7</sup>, You-Soon Chang<sup>8,24</sup>,  
5 Simon A. Good<sup>9</sup>, Armin Köhl<sup>10</sup>, Guillaume Vernieres<sup>11</sup>, Nicolas Ferry<sup>12</sup>, K.  
6 Andrew Peterson<sup>9</sup>, David Behringer<sup>4</sup>, Masayoshi Ishii<sup>15</sup>, Shuhei Masuda<sup>16</sup>, Yosuke  
7 Fujii<sup>15</sup>, Takahiro Toyoda<sup>15</sup>, Yonghong Yin<sup>18</sup>, Maria Valdivieso<sup>19</sup>, Bernard  
8 Barnier<sup>23</sup>, Tim Boyer<sup>14</sup>, Tony Lee<sup>6</sup>, Jérôme Gouillon<sup>20</sup>, Ou Wang<sup>6</sup>, Patrick  
9 Heimback<sup>7</sup>, Anthony Rosati<sup>8</sup>, Robin Kovach<sup>11</sup>, Fabrice Hernandez<sup>12,25</sup>, Matthew  
10 J. Martin<sup>13</sup>, Masafumi Kamachi<sup>15</sup>, Tsurane Kuragano<sup>15</sup>, Kristian Mogensen<sup>2</sup>,  
11 Oscar Alves<sup>18</sup>, Keith Haines<sup>19</sup>, and Xiaochun Wang<sup>22</sup>

12           <sup>1</sup>Centro Euro-Mediterraneo sui Cambiamenti Climatici, Bologna, Italy

13           <sup>2</sup>European Center for Medium-Range Weather Forecast, Reading, UK

14           <sup>3</sup>Collecte Localisation Satellites, Ramonville Saint-Agne, France

15           <sup>4</sup>National Center for Environmental Predictions, NOAA, College Park, MD, USA

16           <sup>5</sup>Institut Universitaire Européen de la Mer, CNRS, Brest, France

17           <sup>6</sup>Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA,  
18                           USA

19           <sup>7</sup>Massachusetts Institute of Technology, Cambridge, MA, CA, USA

20           <sup>8</sup>Geophysical Fluid Dynamics Laboratory, Princeton, NJ, USA

21           <sup>9</sup>Met Office Hadley Centre, Exeter, UK

22           <sup>10</sup>Institute of Oceanography, University of Hamburg, Hamburg, Germany

23           <sup>11</sup>Global Modeling and Assimilation Office, NASA Goddard Space Flight Center,  
24                           Greenbelt, MA, USA

25           <sup>12</sup>Mercator Océan, Ramonville-Saint-Agne, France

26           <sup>13</sup>Met Office, Exeter, UK

27           <sup>14</sup>National Oceanographic Data Center, Silver Spring, MA, USA

28           <sup>15</sup>Meteorological Research Institute, Japan Meteorological Agency, Tsukuba,  
29                           Japan

30           <sup>16</sup>Research and Development Center for Global Change, Japan Agency for  
31                           Marine-Earth Science and Technology, Yokohama, Japan

32           <sup>18</sup>Centre for Australian Weather and Climate Research, Bureau of Meteorology,  
33                           Melbourne, Australia

34           <sup>19</sup>Department of Meteorology, University of Reading, Reading, UK

35           <sup>20</sup>Centre national de la recherche scientifique, CORIOLIS, Brest, France

36           <sup>21</sup>Global Environment and Marine Department, Japan Meteorological Agency,  
37                           Tokio, Japan

38           <sup>22</sup>Joint Institute for Regional Earth System Science and Engineering, University of  
39                           California, Los Angeles, CA, USA

40           <sup>23</sup>Laboratoire des Ecoulements Géophysiques et Industriels, Grenoble, France

41           <sup>24</sup>Department of Earth Science, Kongju National University, Kongju, South Korea

42           <sup>25</sup>Institut de Recherche pour le Développement (IRD), Toulouse, France

---

## Abstract

Quantifying the effect of the seawater density changes on sea level variability is of crucial importance for climate change studies, as the sea level cumulative rise can be regarded as both an important climate change indicator and a possible danger for human activities in coastal areas. In this work, as part of the Ocean Reanalysis Intercomparison Project (ORA-IP), the global and regional steric sea level changes are estimated and compared from an ensemble of 16 ocean reanalyses and 4 objective analyses. These estimates are initially compared with a satellite-derived (altimetry minus gravimetry) dataset for a short period (2003-2010). The ensemble mean exhibits a significant high correlation at both global and regional scale, and the ensemble of ocean reanalyses outperforms that of objective analyses, in particular in the Southern Ocean. The reanalysis ensemble mean thus represents a valuable tool for further analyses, although large uncertainties remain for the inter-annual trends. Within the extended intercomparison period that spans the altimetry era (1993-2010), we find that the ensemble of reanalyses and objective analyses are in good agreement, and both detect a trend of the global steric sea level of 1.0 and  $1.1 \pm 0.05$  mm/yr, respectively. However, the spread among the products of the halosteric component trend exceeds the mean trend itself, questioning the reliability of its estimate. This is related to the scarcity of salinity observations before the Argo era. Furthermore, the impact of deep ocean layers is non-negligible on the steric sea level variability (22% and 12% for the layers below 700 and 1500 m of depth, respectively), although the small deep ocean trends are not significant with respect to the products spread.

---

**Keywords:** *ocean reanalysis evaluation, sea level variability, altimetry, gravimetry.*

## 1 Introduction

Sea level change is a key issue in contemporary climate change, as its global and regional variations are both fundamental indicators of climate change itself and may have a strong impact on human activities in coastal areas.

According to recent estimates (e.g. Cazenave and Llovel, 2010), the contribution of thermal expansion to global mean sea level is of the order of  $30\% \pm 12\%$  for the 1993-2007 period, the remaining contribution being mostly given by glaciers and ice sheet melting. Furthermore, future projections of sea level rise indicate that thermal expansion is likely to continue in the XXI century, and may account for the 32 to 36 % of the global mean sea level rise in 2100 on average, depending on the emission scenario, although large uncertainty is associated with the contribution of land-ice melting (IPCC, 2013).

While mean sea level rise may have different causes, regional sea level rise is generally dominated by the steric component (Fukumori and Wang, 2013) and, in particular, the thermosteric one in most areas (Stammer et al., 2013). Sea level projections tend to confirm this tendency for future scenarios (Sukuki and Ishii, 2011). It is also acknowledged that the sea level low-frequency and inter-annual variability is mostly dominated by the steric sea level variability (Piecuch and Ponte, 2011), although there exists several extra-tropical areas where this simplification does not hold (Piecuch et al., 2013). Attention therefore is being devoted to the monitoring of steric sea level variability and the understanding of the mechanisms causing its long-term rise.

---

\* *Corresponding author address:* Andrea Storto, Centro Euro-Mediterraneo sui Cambiamenti Climatici Viale Aldo Moro 44, I-40135 Bologna, Italy.  
E-mail: andrea.storto@cmcc.it

65 Traditional methods to investigate contemporary steric sea level change are the objec-  
66 tive analyses (Ishii et al., 2006; Levitus et al., 2012) that perform a statistical interpolation  
67 of the available in-situ observations, usually blended with a climatology, or a persistence  
68 background field. These methods, which do not make use of a dynamical ocean model,  
69 have the advantages of being relatively simple and computationally cheap. However, ob-  
70 jective analyses strongly depend on the model used for the correlation function and on  
71 the data availability. For data sparse regions (e.g. the Southern Ocean before the full  
72 deployment of the Argo floats observing network during the 2000s), the analysis can be  
73 unrealistically close to the climatology fields. Furthermore, recent comparisons of the  
74 statistical approaches implemented in the objective analyses highlight that the choice of  
75 mapping practices and climatological references is non-negligible for the resulting esti-  
76 mates (Abraham et al., 2013; Chang et al., 2014b). This issue has been proven even  
77 more important for steric sea level assessment studies that use a combination of in-situ  
78 and satellite observations (Chambers, 2006a; Llovel et al., 2010; Leuliette and Willis,  
79 2011).

80 An alternative methodology is offered by ocean retrospective analyses (or simply re-  
81 analyses or ocean syntheses) that typically combine a data assimilation system capable of  
82 ingesting most of the available ocean observations (in-situ and remotely sensed) with an  
83 ocean general circulation model (OGCM), with surface boundary conditions usually com-  
84 ing from atmospheric reanalysis or similar products. Ocean reanalyses use the same ocean  
85 model and analysis configuration (resolution, parametrizations, numerical schemes, data  
86 assimilation setup, etc.) throughout the reanalyzed period. Thus, they aim at building a  
87 dataset with quality as consistent and coherent over time as possible, in order to provide  
88 a robust tool for climate applications. Closely related to ocean reanalyses are ocean state  
89 estimates based on smoother, instead of filter methods, for assimilating data (e.g. Wunsch  
90 and Heimbach, 2013). For the present purpose we will not make that distinction, and will  
91 refer to them as "reanalyses".

92 Reanalysis approaches contrast with operational oceanographic analysis systems that  
93 aim at achieving the day-by-day best analysis and forecast, eventually upgrading the  
94 ocean model and analysis system as soon as improvements are available. Analyses for  
95 seasonal forecasting – including some of the analyses used here – fit somewhere in between.  
96 Consistency throughout the analysis is again paramount, but the need to transition to  
97 near real time with suitable continuity is also required in order that the forecast from the  
98 real time products can be properly informed by re-forecasts (or hindcasts).

99 Many efforts have been recently devoted to the assessment of ocean reanalyses, which  
100 have now reached some degree of maturity (Lee et al., 2009). This justifies the idea of  
101 promoting a coordinated multi-reanalysis intercomparison, the Ocean Reanalyses Inter-  
102 comparison Project (ORA-IP, Balmaseda et al., 2015), in the framework of which this  
103 work was performed. Provided that OGCMs are becoming reliable tools for climate in-  
104 vestigations, the use of reanalyses in contrast to objective analyses seems appealing for  
105 climate applications because the scarcity of in-situ observations in the pre-Argo ocean  
106 observing network may be partly overtaken by satellite data, information from the atmo-  
107 spheric forcing and the ocean model dynamics. In general, model-based products provide  
108 a full view of the ocean state, thus allowing process oriented studies. Even without data  
109 assimilation, OGCMs are able to capture the bulk of the inter-annual variability of sea  
110 level. For instance, Lombard et al. (2009) showed that non-assimilative hindcasts can  
111 reproduce steric sea level trends at regional scale, with a reasonable accuracy; similarly,  
112 Griffies et al. (2014), in the frame of the CORE-II simulations, found that OGCMs re-  
113 produce regional trends over the last two decades in general agreement with satellite

114 estimates, and dominated by the thermal component.

115 Recent studies (Purkey and Johnson, 2010; Ponte, 2012) pointed out that deep  
116 regions may have a non-negligible contribution to the total steric sea level. This makes the  
117 recourse to reanalyses appealing, since observations in the upper ocean may partly con-  
118 strain the deep ocean through the vertical background-error correlations and the physical  
119 balances implied by data assimilation systems and OGCMs, respectively, especially in  
120 the future when the accuracy and resolution of OGCMs and reanalyses are expected to  
121 increase. However, given the paucity of deep ocean observations, validating reanalyses in  
122 the deep ocean is an extremely challenging task, especially to evaluate the reliability of  
123 reanalyses in capturing the climate change signal in deep waters. The use of an ensemble  
124 of ocean reanalyses has also started to be used for key climate indexes (Stammer et al.,  
125 2010), as many ocean research groups are continuously producing ocean reanalyses. The  
126 underlying assumption is that the averaging operation over the ensemble members is able  
127 to reduce the systematic biases of the individual products.

128 In this paper, we analyze the steric sea level variability from an ensemble of ocean  
129 reanalyses and objective analyses, focusing on the performance of the ensemble means.  
130 First, data and methods used within the comparison are introduced in Section 2. Second,  
131 a comparison against reference steric sea level estimates is presented for a short 8-year  
132 period (2003-2010, Section 3). The comparison is then extended to the 1993-2010 period  
133 (Section 4). Finally, the main conclusions are given in Section 5.

## 134 2 Data and methods

### 135 2.1 Strategy for the comparison

136 The strategy used in this work consists of two separate phases of steric sea level intercom-  
137 parison, which span different periods: a first validation period (2003-2010), corresponding  
138 to the period when ocean syntheses, altimetric and gravimetric data overlap, and a second  
139 extended comparison period (1993-2010) that covers the altimetric missions. Within the  
140 validation period, reanalyses are compared with reference steric sea level estimates from  
141 altimetry minus gravimetry, as detailed later in this Section. For the extended comparison  
142 period, the products are compared to estimate their consistency and uncertainty, focusing  
143 on the significance of the mean climate signal.

### 144 2.2 Estimation of steric sea level

145 Following Gill and Niiler (1973), Landerer et al. (2007) and Griffies et al. (2014), the sea  
146 level anomaly  $\eta$  can be formulated as a function of the anomalies of three terms with  
147 respect to the time mean (the space and time dependence is dropped for simplicity):

$$\eta = \eta_a + \frac{p_b}{g\rho_0} + \eta_s \quad (1)$$

148 where  $\eta_a$  is the contribution from atmospheric pressure at sea level,  $\frac{p_b}{g\rho_0}$  accounts for the  
149 bottom pressure effects (change in water mass) and  $\eta_s$  is the steric sea level, which is in  
150 turn given by

$$\eta_s = - \int_{-H}^{\eta} \frac{\rho}{\rho_0} dz, \quad (2)$$

151  $z = H$  being the ocean depth and  $z = \eta$  the (time and space varying) ocean surface, the in-  
152 tegral covering the whole water column. This formulation corresponds to the “local steric

153 effect” defined by Griffies and Greatbatch (2012) (or “global steric effect” in case of global  
 154 average, see Section 2.4), to whom the reader is referred for a detailed discussion. Note  
 155 that the “non-Boussinesq steric effect” is not accounted for, because of the Boussinesq  
 156 approximation made in OGCMs. The integration of Equation (2) is then approximated  
 157 to extend from the ocean depth to the mean sea surface  $z = 0$ . The density anomaly can  
 158 be approximately decomposed into thermal, haline and pressure contributions by consid-  
 159 ering the variations of density induced by only temperature ( $T$ ), salinity ( $S$ ) or pressure  
 160 anomalies, respectively, keeping constant the other parameters. The pressure-steric sea  
 161 level contribution to  $\eta_s$  is negligible and is dropped from the equations (Griffies et al.,  
 162 2014). Thus, Equation (2) becomes:

$$\eta_s \simeq \eta_t + \eta_h = - \int_{-H}^0 \frac{\rho(T, \bar{S})}{\rho_0} dz - \int_{-H}^0 \frac{\rho(\bar{T}, S)}{\rho_0} dz, \quad (3)$$

163 where the over-bar denotes the time-averaged value. It was shown by Balmaseda et al.  
 164 (2012) that the inaccuracies introduced by the decomposition in Equation (3) with respect  
 165 to Equation (2) are negligible.

166 The global sea level – the spatial average over the whole ocean of the terms of Equation  
 167 (1) – is therefore given by a barostatic term that accounts for variations on the ocean water  
 168 mass (from time-varying contributions of evaporation, precipitation, terrestrial runoff and  
 169 land ice melting) and a steric term that accounts for the expansion and contraction of  
 170 water. Note that the term “barostatic” denominates the contribution of changes in the  
 171 ocean mass to the global mean sea level, and is preferred to the term “eustatic”, in  
 172 accordance with Gregory et al. (2013). The steric component, at global scale, is primarily  
 173 driven by changes in the seawater temperature, the haline contribution being small since,  
 174 to first approximation, the global water cycle does not change sufficiently to affect the  
 175 global salinity budget. At basin and local scales, the bottom pressure term corresponds  
 176 to the redistribution of the seawater masses and the haline contribution of the steric term  
 177 may be as important as the thermal contribution (e.g. Ivchenko et al., 2008; Lombard  
 178 et al., 2009), even for long-term basin-scale effects (Durack et al., 2014). However, the  
 179 barotropic response of the sea level to a local change of ocean water mass occurs on time  
 180 scales generally of the order of a few weeks to few months because of its fast adjustment  
 181 (Ponte, 2006); therefore, within studies encompassing monthly to inter-annual time scales  
 182 the bottom pressure term, which is often dominated by barotropic fluctuations, acts as  
 183 a globally uniform term, with few notable exceptions (Landerer et al., 2007; Yin et al.,  
 184 2010; Piecuch et al., 2013; Griffies et al., 2014).

### 185 2.3 Reanalyses and objective analyses

186 The products participating in the comparison are summarized in Table 1. There are  
 187 20 products, of which 16 are reanalyses (REA) and 4 are objective analyses (OA). We  
 188 consider as objective analyses here the products that do not include any dynamical balance  
 189 through an ocean general circulation model, but make only a statistical use of observations  
 190 to estimate the three-dimensional state of the ocean. They are ARMOR, CORA, EN3  
 191 and IK09. Another fundamental difference is that objective analyses do not use any  
 192 information about the air-sea fluxes, unlike ocean reanalyses.

193 The resolution of the products ranges from eddy-permitting (about 1/4 degree) to  
 194 much coarser resolution, the majority of products being in the range 1/2 degree to 1 de-  
 195 gree of horizontal resolution. A few products (CFSR, ECDA and MOVEC) are the ocean  
 196 components of assimilative coupled atmosphere ocean general circulation models. Further-



197 more, 13 reanalyses assimilate sea level anomalies, while one objective analysis (ARMOR)  
 198 projects onto vertical profiles of temperature and salinity the information from satellite  
 199 altimetry. All the reanalyses except three products assimilate salinity measurements from  
 200 in-situ observations. The data assimilation methods include optimal interpolation (OI),  
 201 three- and four- dimensional variational assimilation (3DVAR, 4DVAR), Kalman filter  
 202 or related sequential smoother (KF, SS), ensemble optimal interpolation and ensemble  
 203 Kalman filter (EnOI, EnKF). All reanalyses except three (ECCOV4, ECDA and UR025.4)  
 204 include an additional constraint to avoid model biases and drifts, either implementing a  
 205 bias correction scheme or restoring to climatological surface or subsurface fields, or a  
 206 combination of them. Note however that the restoring time scales, when implemented,  
 207 may vary notably among the products – from a few days to several years . The different  
 208 characteristics summarized in Table 1 suggest that the collection of products includes a  
 209 large diversity of model and data assimilation configurations.

210 The steric sea level anomaly was calculated by means of Equation (3) for all products,  
 211 using monthly means of temperature and salinity. The steric sea level fields were then  
 212 interpolated onto a regular grid of 1x1 degree horizontal resolution, regardless of their  
 213 native resolution. The same computation was also carried out for a number of depth  
 214 levels: i) 0-100 m; ii) 100-300 m; iii) 300-700 m; iv) 700-1500 m; v) 1500-3000 m; vi)  
 215 3000-4000 m; vii) 4000-bottom. This allows us to have an insight into the contributions  
 216 of different vertical levels. For each of the vertical levels and the total column, not  
 217 only the steric sea level but also the thermo- and halo- steric components were provided  
 218 separately. For the sake of comparison with the verifying dataset (see next Section), only  
 219 the interior ocean (distance from coast greater than 100 km) and the ocean between 70S  
 220 and 70N contribute to the global mean steric sea level. We introduce also three derived  
 221 datasets, ALLENS, REAENS and OAENS, which are the ensemble means calculated from  
 222 all products, from the reanalyses only, and from the objective analyses only, respectively.

## 223 2.4 Reference steric sea level estimates

224 Previous studies have combined gravimetric data with altimetric data in order to estimate  
 225 the steric sea level. Lombard et al. (2007) and Willis et al. (2008) compared steric sea  
 226 level inferred from altimetry minus gravimetry with steric sea level objectively analyzed  
 227 from in-situ observations. They were not able to close the sea level global budget, namely  
 228 to match the total sea level from altimetry with the the sum of gravimetry and Argo-  
 229 derived steric sea level, within the error bars of the observational networks, and suggested  
 230 that the discrepancies may be related to the sampling of the in-situ observing network  
 231 or inaccuracies in the processing of one or more observing systems. Later, an improved  
 232 processing of the observational datasets allowed Leuliette and Miller (2009) to close the  
 233 budget, demonstrating the complementarity of gravimetry, altimetry and Argo network  
 234 in estimating the sea level budget. Since then, the methodology of combining the different  
 235 observing networks has largely been explored, also for studies at the basin scale (Garcia-  
 236 Garcia et al., 2010; Chang et al., 2014a).

237 Steric sea level estimates for use in the comparison have been calculated from Equation  
 238 (1) by formulating steric sea level as a difference of total sea level minus bottom pressure  
 239 and atmospheric pressure. For both terms, we decompose the time-varying sea level  
 240 anomaly terms into globally averaged and spatially varying sea level variations:

$$\eta(x, y, t) = \langle \eta(t) \rangle_D + \tilde{\eta}(x, y, t), \quad (4)$$

241 where  $\langle \dots \rangle_D$  is the spatial averaging operator and  $\tilde{\eta}(x, y, t)$  can be computed as the

242 difference between the previous two terms and has spatial average equal to zero. This  
 243 allowed us to optimize the processing of data depending on whether the analysis focuses  
 244 on global mean or regional sea level. For the global mean sea level  $\langle \eta(t) \rangle_D$ , we used  
 245 the dataset of Nerem et al. (2010), which represents a seamless dataset cross-calibrating  
 246 the TOPEX and JASON altimeter missions. The spatially varying term of Equation  
 247 (4) for the total component ( $\tilde{\eta}(x, y, t)$ ) is provided by the AVISO delayed-time monthly  
 248 gridded altimetric products, with the time-varying global mean removed. Altimetric  
 249 observations were subject to the usual geophysical removals and multi-satellite cross-  
 250 correction (Le Traon et al., 1998). For the bottom pressure term (Chambers and Schröter,  
 251 2011), the global mean value  $\langle p_b(t) \rangle_D$  was taken from Johnson and Chambers (2013),  
 252 which uses the Gravity Recovery and Climate Experiment (GRACE) RL05 data only in  
 253 the ocean interior to avoid possible land and ice contamination. The spatially-varying  
 254 term ( $\tilde{p}_b(x, y, t)$ ) was taken from release RL05 of GRACE gravimetric data (Chambers  
 255 and Bonin, 2012), which is provided with the area-weighted global mean set to zero.  
 256 The release RL05 disseminates data after the application of a destriping procedure and  
 257 a 500 km wide Gaussian filter to remove the meridional stripes typical of gravimetric  
 258 data. The attenuation of systematic biases in the processing chain of GRACE data  
 259 was attempted using the ensemble mean of the three RL05 releases from CSR (Center  
 260 for Space Research, University of Texas), GFZ (GeoForschungsZentrum, the German  
 261 Research Centre for Geosciences) and JPL (Jet Propulsion Laboratory), mostly differing  
 262 in the data pre-processing. Note that we do not estimate steric sea level errors from the  
 263 altimetric and gravimetric data errors, as it is beyond the scope of this work.

264 Results from the construction of the verifying dataset are presented in Figure 1, which  
 265 reproduces the zonal averages of the 2003-2010 monthly means for the three components  
 266 (total, bottom pressure, and steric sea level inferred from the previous two). The total  
 267 sea level seasonality is primarily affected by the heat content seasonal cycle, peaking in  
 268 September (March) for the Northern (Southern) Hemisphere. Consequently, it is dom-  
 269 inated by a strong hemispheric separation during all months. On the other hand, the  
 270 bottom pressure signal does not exhibit a hemispheric separation. This is due to the fact  
 271 that at monthly time-scales bottom pressure is rather uniform and driven by the Northern  
 272 Hemisphere seasonal cycle of water stored inland (snow, ice) and in rivers, reservoirs and  
 273 underground (Leuliette and Willis, 2011). The resulting steric sea level (bottom panel) re-  
 274 sembles the seasonal cycle of the total sea level, but with amplitudes and phases modified  
 275 according to the bottom pressure component.

276 Although many studies have suggested the low signal-to-noise ratio of GRACE data  
 277 over the oceans (e.g. Chambers, 2006b) and the critical role of the data processing in  
 278 estimating inter-annual trends and variability from gravimetry (Quinn and Ponte, 2010),  
 279 the use of such data to infer steric sea level allows us to build a validation dataset for the  
 280 ocean syntheses and further test the consistency between the altimetric and gravimetric  
 281 datasets and the ocean reanalyses. In other words, the validation allows us to identify  
 282 the products that better close the sea level budget, given the altimetric and gravimetric  
 283 missions. The validation dataset is not strictly independent, as most of the products  
 284 assimilate altimetric data. However, without going into detail, there is a tremendous  
 285 diversity in the methods that altimetry is assimilated with, ranging from hydrostatic  
 286 adjustments (e.g. Storto et al., 2011), to simplified barotropic and baroclinic adjustments  
 287 (e.g. Fukumori et al., 1999), to combined analytic and statistical balances (e.g. Weaver  
 288 et al., 2005). This suggests that steric sea level estimates from altimetry and gravimetry  
 289 represent a useful validation dataset, despite the assimilation of altimetric data in most  
 290 products.

291 Hereafter, ALT-GRV will denote the steric sea level estimates presented in this Section.

## 292 **3 Validation period (2003-2010)**

### 293 **3.1 Global steric sea level comparison**

294 In this Section, globally averaged values of the steric sea level of reanalyses are compared  
295 to that obtained by altimetry minus gravimetry for the 2003-2010 period. The globally  
296 averaged estimates from all products and the validation dataset are shown in Figure 2. For  
297 comparison with the verifying dataset (see Section 2.4), global averages of the reanalyses  
298 and objective analyses include only the ocean between 70S and 70N. In the figure legend,  
299 we report the correlations of the full, seasonal and inter-annual signals, with respect to  
300 the verifying dataset. The way through which the signal is decomposed is detailed in the  
301 Appendix. Generally, all products except three exhibit a significant correlation (0.21 is  
302 the minimum significant correlation according to a two-sided t-test with 95% confidence  
303 level). The most skillful product is ARMOR, which shows very high correlations for all  
304 the signals.

305 The performance of the ensemble means are also very satisfactory, with values of (full  
306 signal) correlations of 0.79, 0.80 and 0.71 for the ensemble of all products, the ensemble of  
307 reanalyses and the ensemble of objective analyses, indicating that the ensemble mean of  
308 the reanalyses outperforms the ensemble mean of the objective analyses. The correlation  
309 difference between REAENS and OAENS is significant, according to a Steiger's Z test  
310 (Steiger, 1980), with 95% confidence level, for two dependent overlapping correlations.  
311 Note that not only the correlation of REAENS is greater than OAENS, but is also greater  
312 than that of ALLENS, indicating that increased ensemble sizes do not ensure better skill  
313 scores. This result is due to the fact that two objective analyses out of four present a  
314 negative correlation with the verifying dataset for the inter-annual signal, namely they are  
315 not able to capture the year-by-year variations seen by the altimetry minus gravimetry  
316 dataset.

317 The decomposition of the correlation coefficient for the seasonal and inter-annual  
318 signal (in the legend of Figure 2) suggests also that the correlation of the seasonal signal is  
319 almost always greater than that of the inter-annual signal. Because of the definition of the  
320 seasonal signal (corresponding to the detrended signal, see the Appendix), this indicates  
321 that linear trends of most of the products are not in agreement with the verifying dataset  
322 and are likely affected by biases and drifts. Conversely, the seasonality of the steric sea  
323 level is generally well-captured.

324 To refine the analysis of the seasonal and inter-annual components of the global steric  
325 sea level, Figure 3 shows two components (annual and inter-annual), in which we have  
326 divided the signal. The semi-annual component is not shown, for sake of simplicity. The  
327 first panel (annual component) shows the amplitude and phase in polar coordinates. The  
328 majority of the products under-estimate the annual amplitude with respect to ALT-GRV  
329 (5.3 mm for ALT-GRV against 3.9 mm for REAENS). OAENS, on the contrary, shows  
330 an amplitude of 5.1 mm, very close to the one observed. Almost all products capture  
331 the phase of the annual global steric sea level, whose peak occurs in mid-April. The  
332 semi-annual amplitude (not shown) shows less consistency among the products, with an  
333 amplitude generally under-estimated.

334 The second panel of Figure 3 shows the linear trends of the global steric sea level  
335 for all the products, with the 95% confidence level estimated by means of the bootstrap  
336 methodology. There is a large diversity in the trend values, provided that ALT-GRV

337 exhibits a value of  $1.0 \pm 0.4 \text{ mm yr}^{-1}$ , while the products' trends range from about -2  
338 to  $2 \text{ mm yr}^{-1}$ , with 9 products (along with OAENS) showing a negative trend. REAENS  
339 shows a weakly positive trend ( $0.2 \pm 0.3 \text{ mm yr}^{-1}$ ). Although 2003-2010 is a short  
340 time period to have a climatologically significant trend, here the purpose is to evaluate  
341 the consistency or the diversity of the products in estimating trends in comparison with  
342 the verifying dataset. For this well observed period, there is no clear consensus on the  
343 inter-annual trend, i.e. the trend of the global mean steric sea level from the individual  
344 products has large variability. There may be several reasons for such a result, including  
345 the sensitivity of the trend calculation over a short period, the impact of the abrupt change  
346 of the observing network (Argo network deployment) in data assimilation systems, or the  
347 global mean imbalances in the atmospheric reanalyses used to force the models.

### 348 3.2 Regional steric sea level comparison

349 The regional steric sea level has been evaluated by correlating the point-by-point time-  
350 series of the individual products with the verifying dataset. This is summarized in Table 2,  
351 which reports for each product and the ensemble means the area averages of the point-by-  
352 point correlation. Regional correlations for the ensemble means resemble the correlation  
353 coefficients of the global steric sea level (by comparison of Table 2 with Figure 2). However,  
354 11 products exhibit spatially averaged regional correlations greater than that of global  
355 steric sea level while 9 products present the opposite behavior, suggesting that there is  
356 no clear evidence whether the products capture the regional signal better than the global  
357 one. However, all the correlation scores in the Table 2 are significant. For all products  
358 the correlation in the tropics is greater than the global average and the one averaged in  
359 the southern extra-tropics is smaller, indicating the strong latitudinal dependence of the  
360 correlation coefficient.

361 The top panel of Figure 4 shows the correlation map for REAENS with the verifying  
362 dataset, indicating the very high correlation globally (the global mean correlation is 0.8,  
363 see Table 2), except at high latitudes. Only south of about 60S, the reanalysis ensemble  
364 mean does not show significant correlation (less than 0.21).

365 The correlation is close to 1 in the tropical band, where reanalyses are most successful  
366 at capturing the evolution of density anomalies. Decreases in correlation can be seen in  
367 regions around the Kuroshio Extension, the Gulf Stream, the Falkland current, and off  
368 the coast of Peru. However, in these regions the ensemble spread among the products is  
369 high (not shown), a consequence of the high eddy variability, which most models do not  
370 resolve.

371 The bottom panel shows the map of correlation difference with the verifying dataset  
372 between REAENS and OAENS. Where the differences are positive (negative), the ensem-  
373 ble of reanalyses outperforms (underperforms) the ensemble of objective analyses. The  
374 picture thus suggests that while in the tropics the differences are negligible due to the  
375 high correlations of both datasets, at mid and high latitudes the differences are evident  
376 and reveal the better fit of REAENS to the verifying estimates of steric sea level. This is  
377 particularly evident for the Southern Ocean, especially in the Indian and Atlantic sectors,  
378 where differences reach values up to 0.4. The correlation difference between REAENS and  
379 ALLENS (not shown) verifies that ensemble size does not compromise the comparison.

380 Provided that the correlation coefficients are computed over monthly means and that  
381 the annual amplitude of steric sea level is greater than the inter-annual trend and vari-  
382 ability (see Figure 3), the general high correlation resides in the high seasonality of the  
383 steric signal that may be easily captured by the products. To appreciate the effect of

384 the seasonality on the skill scores previously presented, we show in Figure 5 the same  
385 correlation maps of Figure 4 but with the seasonal signal removed. Note that the cor-  
386 relation for the case with inter-annual signal removed behaves closely to that of the full  
387 signal, and is not shown. The correlation of REAENS with the ALT-GRV dataset (top  
388 panel) shows in this case a generally lower correlation, confirming that the steric sea level  
389 seasonality importantly contributes to the correlation scores. However, correlations are  
390 still significant all over the global ocean, except south of 60S and over a few other areas  
391 (Arabian Sea and Angola Basin) and still quite high, close to 1, in the Tropical Pacific  
392 except over the areas corresponding to the Intertropical and South Pacific Convergence  
393 Zones. However, within the latter areas the inter-annual variability of the steric sea level  
394 is generally small (not shown), contributing to the local decrease of the correlation score.  
395 Note also that the dominance of the seasonal signal on the correlation of the full signal  
396 may also explain, for some products, the low correlation scores, due to a change of phase  
397 in the seasonal signal.

398 The correlation difference map (bottom panel) exhibits similar patterns to those of  
399 Figure 4. However, these differences are slightly larger, confirming that reanalyses capture  
400 the steric sea level inter-annual variability better than objective analyses, as for the global  
401 mean steric sea level.

402 There are several interpretations for the ensemble reanalyses to have higher correla-  
403 tions especially in the Southern Ocean. A first interpretation is given by the fact that  
404 in this region the in-situ observing network is poor, even for the Argo floats (see e.g.  
405 Figure 1d in Storto et al. (2013)). In this case, the background used by objective analy-  
406 ses, which is usually either a climatology or a previous analysis, is not impacted by the  
407 analysis step because of the scarcity of observations and does not change, producing no  
408 inter-annual variability. On the other hand, reanalyses can still take advantage of space-  
409 borne measurements (SLA and SST) over ice-free areas – and SIC over ice-contaminated  
410 areas. They also use the information from the atmospheric forcing that, despite its large  
411 uncertainty at high latitudes, still provides a time-varying forcing.

412 Another concurrent explanation is that the Southern Ocean is a current system where  
413 deep variations have large contributions to steric sea level variability. A way to show  
414 this is given by Figure 6, which reproduces the REAENS explained variance of the 700  
415 m to bottom steric sea level contribution to the total steric sea level. Percentage values  
416 of the explained variance are generally low (less than 10%), suggesting that the steric  
417 sea level variability can be generally explained by the contribution of the first 700 m. An  
418 exception is the Southern Ocean, where percentage values are in the range 20% to 50%, in  
419 agreement with a generally deeper mixed layer depth, on the average, in this region (see  
420 e.g. de Boyer Montégut et al., 2004). Similar results are found when the linear trend is  
421 removed or the explained variance is calculated from OAENS and do not seem affected by  
422 possible drifts of individual members (not shown). The sea level variability is characterized  
423 by deep steric contributions in the Southern Ocean. This may be better simulated by  
424 reanalyses than by objective analyses through the inclusion of the atmospheric forcing,  
425 satellite data and the dynamical ocean balances.

## 426 4 Extended intercomparison period (1993-2010)

427 In this Section, we compare the products for the extended period 1993-2010. Within  
428 this period, we do not consider any validation dataset and we focus on the statistical  
429 significance of the ensemble mean with respect to the ensemble spread (defined hereafter  
430 as the standard deviation over the ensemble members). Different regions and vertical

431 levels are considered.

#### 432 4.1 Global steric sea level comparison

433 The comparison for global steric sea level is summarized in Figure 7. The figure shows the  
434 globally averaged values of steric sea level and its thermo- and halo- steric components for  
435 the individual products (gray lines), the ensemble mean of the reanalyses and objective  
436 analyses (cyan and blue lines, respectively) and their ensemble spreads (red and orange  
437 lines, respectively). Note that y-axes are different for the ensemble means and spreads.  
438 The analysis of total steric sea level suggests that although discrepancies and outliers  
439 are visible among the individual products, the two ensemble means are reasonably close.  
440 This is confirmed by the correlation between the two ensemble means, equal to 0.93. The  
441 spread is also generally comparable, although the size of the two ensembles is different.  
442 The shape of the spread stems from the definition of the steric sea level data as anomalies  
443 with respect to the whole 1993-2010 period. Consequently, different linear trends of the  
444 individual products lead to a convex parabola-like shape for the spread.

445 Qualitatively similar results were found for the thermosteric component, for which  
446 the correlation between the two ensemble means is even higher (0.97), and the spread  
447 has a very similar magnitude and behavior. We observe a more pronounced seasonality  
448 for the reanalysis ensemble spread, while an abrupt increase of spread for the objective  
449 analyses in 2003 is likely to be related to the increase of Argo observations. The match  
450 of the two ensemble means for halosteric sea level (bottom panel of Figure 7) is lower,  
451 the correlation dropping to 0.55. In particular, the seasonal variability is different, with  
452 a seasonal amplitude in REAENS of 0.7mm being considerably less than the seasonal  
453 amplitude of 2.7mm in OAENS. This may be related to inaccurate representation of the  
454 global freshwater budget in many of the reanalyses, and in particular to the fact that  
455 some reanalyses unrealistically show an inter-annual variability larger than the seasonal  
456 variability.

457 Although the spread of the halosteric component is smaller than that of the ther-  
458 mosteric one, their values are higher than the variability of the signal. This is illustrated  
459 in Table 3, where we report the standard deviation of the ensemble mean of the global  
460 time series, along with their time-averaged spread and the normalized spread (the ratio  
461 between the spread and the ensemble mean standard deviation), for both reanalyses and  
462 objective analyses. While for the steric and thermo-steric sea level the ratios between the  
463 spread and the signal variability are smaller than unity, both reanalyses and objective  
464 analyses show a value greater than 1 (1.61 and 1.71) for the normalized spread in case of  
465 the halosteric component. This means that the uncertainty is greater than the variability  
466 of the signal, suggesting the low reliability of the halosteric time-series. The same exercise  
467 of computing the normalized spread is repeated for four different reanalysis subgroups,  
468 categorized according to possible additional constraints on salinity, i.e. no constraint,  
469 bias correction, restoring to climatological sea surface salinity or restoring to subsurface  
470 climatological salinity. Results are reported in Table 3, whose caption details the pro-  
471 cedure to compute these values accounting for different group sizes. Products with no  
472 constraints have small spread also associated with small variability, i.e. the ratio is equal  
473 to 1.35. The use of bias correction increases both spread and variability, leading to a  
474 ratio comparable with the one from the reanalysis group with no salinity constraints. By  
475 using the restoring to sea surface salinity or subsurface salinity fields, reanalyses have an  
476 intermediate spread and variability, leading to a smaller normalized spread (1.20 and 1.12,  
477 respectively), still greater than 1. This suggests that, based on our ensemble of reanalyses,

478 while the restoring helps in slightly decreasing the normalized spread, namely increasing  
479 the signal-to-noise of the halosteric component, all subgroups of reanalyses have a spread  
480 exceeding the variability, suggesting that other issues such as the global freshwater budget  
481 uncertainty linked to the uncertainty in the modeling of the hydrological cycle together  
482 with the scarcity of salinity observations lead to this result, regardless of the bias and  
483 drift correction possibly implemented in the reanalyses.

484 By looking at the geographical patterns of the normalized spread (not shown), the  
485 steric and thermosteric components have values less than 1 everywhere except in the  
486 ACC region, with minimum values in the Equatorial region (high signal-to-noise ratio).  
487 On the contrary, the halosteric component shows values greater than unity everywhere  
488 (low signal-to-noise ratio), with high values especially in the Atlantic Ocean, indicating  
489 the large uncertainty of the salinity content in this basin.

490 The reanalyses spread also exhibits local maxima at the beginning and in the middle  
491 of the period, with a parabola-like shape (minimum or maximum of the halosteric sea  
492 level at the middle of the period) due to a few outlier products.

493 Global steric sea level trends are reported in Figure 8 in mm/yr. The total steric sea  
494 level trends range from 0.1 to 3.1 mm/yr, and are all positive. For the ensemble means,  
495 the trends have values of  $1.02 \pm 0.05$  mm/yr and  $1.11 \pm 0.08$  mm/yr for the reanalyses  
496 and objective analyses, respectively, with a standard deviation of the trends (red bars)  
497 of the order of 0.5 mm/yr for both. Similar results apply to the thermosteric sea level,  
498 although two products exhibit negative trends. For the ensemble means, the trends equal  
499  $1.03 \pm 0.05$  mm/yr and  $1.17 \pm 0.05$  mm/yr, for the reanalyses and objective analyses,  
500 respectively, the standard deviation of trends being equal to 0.6 mm/yr for both. The  
501 halosteric trend exhibits no clear consistency between the products, ranging from about  
502 -0.8 to 0.9 mm/yr, most products showing a negative trend. The reanalyses ensemble  
503 mean filters out these discrepancies, exhibiting almost no trend ( $-0.01 \pm 0.01$  mm/yr ).  
504 OAENS shows a slightly negative trend ( $-0.07 \pm 0.06$  mm/yr).

## 505 4.2 Regional steric sea level comparison

### 506 4.2.1 Steric sea level trends and their significance

507 Maps of linear trends (ensemble mean of all products) are presented in Figure 9 for the to-  
508 tal steric sea level and the two components separately. The total steric and thermosteric  
509 sea level look very similar, i.e. the local trends are in general dominated by the ther-  
510 mosteric component. Well-known maxima of the trends are found in the western tropical  
511 Pacific (up to 8 mm/yr). Also shown are signal-to-spread ratios (SSR, see the Appendix),  
512 with solid (dashed) contours referring to significant positive (negative) trends, i.e. with  
513 absolute SSR values greater than 1. Areas of significant positive trend are found in the  
514 western tropical Pacific, the central north Pacific, the Indonesian Archipelago and the  
515 southern Indian Ocean, a few areas part of the ACC and some areas in the Atlantic  
516 Ocean, in particular within the tropics and in the North Atlantic subpolar gyre and the  
517 Labrador Sea. Significant negative trends (up to -4 mm/yr) are found only in the eastern  
518 Pacific Ocean and in the Alaskan gyre. The patterns of thermosteric trends are simi-  
519 lar, with a more pronounced positive trend in the Labrador and North Atlantic subpolar  
520 gyre. (up to 7 mm/yr). The halosteric component trend map shows only few regions  
521 with SSR values greater than 1. While there are several regions with positive significant  
522 trends within the ACC and in the Western Tropical Pacific, the only significant negative  
523 trends are located within the North Pacific subpolar gyre. Negative trends characterize  
524 the Atlantic Ocean, showing the typical compensating effect between the halosteric and

525 thermosteric components (Lowe and Gregory, 2006).

526 To better understand why halosteric trends are non-significant almost everywhere,  
527 we report in Figure 10 a graphic showing the percentage of the global ocean area with  
528 significant trend, as a function of the starting and ending year for the trend computation.  
529 We consider 5 years as minimum period for calculating the trend. Trends are considered  
530 significant if their ensemble mean exceeds the spread of the trends. We also report the  
531 number of in-situ observations for clarity. The figure provides insights into the significance  
532 of the trends, i.e. the capability of capturing the steric trends in the altimetric period.  
533 Generally, thermosteric trends are more significant: the percentage of area with significant  
534 trends is almost always above 20%, peaking at 37%. For the halosteric component, this  
535 percentage is below 7% except for the last 5-year period, i.e. for trends computed starting  
536 at least in 2002 and ending at least in 2008, peaking in the 2006-2010 period with a value of  
537 about 15%. For periods starting before 2002, there are small differences in the significant  
538 area for the halosteric trends, suggesting a close relation with the number of observations,  
539 below 2 million per year before 2001. In other words, before the full deployment of the  
540 Argo floats, only a small percentage of the ocean show significant trends, questioning  
541 the reliability of the halosteric trend estimates. In contrast, the number of temperature  
542 observations is larger (between 3 and 6 million before 2002, generally three to four times  
543 larger than the amount of salinity observations). Thermosteric trends are also particularly  
544 significant in the last period (for trends computed starting from 2003). However, there  
545 are periods of increasing (for trends including the 2000 and 2001 years or the 2004 to 2006  
546 period) and decreasing (for trends starting in 1999 or 2000) percentage, which suggests  
547 that the change in observation coverage (i.e. Argo floats deployment) is not the only cause  
548 for these variations. In particular, these anomalous increases are primarily related to the  
549 ENSO variability. For instance, by comparing the maps of significance for thermosteric  
550 trends ending in 2010 but starting either in 1998 or in 1999 (not shown), it turns out that  
551 the loss of significance is located in the Western Tropical Pacific, related to a La Nina  
552 event. On the contrary, the increase of significance when the years 2001 or 2004-2005 fall  
553 within the trend computation results from the inclusion of El Niño events. To summarize,  
554 halosteric trends are very dependent on the observational coverage and the assimilation  
555 method. Thermosteric trends are, however, more robust. This finding complements the  
556 previous result about the low signal-to-noise ratio of the global halosteric component due  
557 to the freshwater budget uncertainty.

#### 558 4.2.2 Thermo- and halo- steric contributions

559 A secondary objective of the intercomparison is the quantification of the thermal and  
560 haline steric sea level contributions to the total steric sea level. This is summarized in  
561 Figure 11 in terms of linear trends and explained variance of the full and inter-annual  
562 signals for the main ocean basins (see also the Appendix). The figure also reports the  
563 standard deviation of the contribution among all products in order to evaluate the signif-  
564 icance of the results. All basins exhibit positive trends for the thermosteric contribution.  
565 This positive trend is significant with respect to the ensemble spread (signal greater than  
566 the spread) except in the Southern Ocean. The Atlantic Ocean exhibits the largest trend  
567 (1.7 mm/yr). The halosteric trend never appears significant except in the Atlantic Ocean,  
568 which presents a salinification corresponding to a negative trend of the steric sea level  
569 equal to -0.5 mm/yr. While non-significant, the Southern Ocean exhibits a positive trend  
570 in halosteric sea level, consistent with the recent freshening (Böning et al., 2008). In the  
571 Southern Ocean the two contributions are comparable (about 0.3 mm/yr), in agreement



572 with Purkey and Johnson (2010, 2013).

573 All basins report a value for the halosteric sea level explained variance between 15%  
574 and 25%, except for the Southern Ocean where the value reaches 36% (Figure 11, middle  
575 and bottom panels). Similar values are found for the inter-annual only signal, except for  
576 the Southern Ocean. The latter presents a value of about 45% for the halosteric contri-  
577 bution, indicating that the inter-annual variability of halosteric sea level is comparable to  
578 the thermosteric component.

### 579 4.2.3 Contributions from vertical levels

580 In this Section we examine the contributions of the different depth levels. Figure 12 reports  
581 the same diagnostics of Figure 11 for the 7 vertical levels analyzed. The global steric sea  
582 level exhibits a positive and significant trend for the top 700 m of depth, peaking between  
583 100 and 300 m with around 0.39 mm/yr. Depths below 700 m show no significant trends,  
584 with abyssal waters (4000 m to bottom) contributing with a slightly negative trend (-0.08  
585 mm/yr). Similar qualitative results apply for all the basins, with some notable exceptions:  
586 i) in the Atlantic Ocean, the largest contribution comes from the 300-700 m level (0.37  
587 mm/yr); ii) the 100-300 m Pacific waters exhibit the largest positive contribution (0.58  
588 mm/yr) along with the Indian waters in the 700-1500 m level. The Pacific abyssal waters  
589 also showed the largest negative trend (-0.12 mm/yr); iii) the Southern Ocean presents  
590 small values for the vertical contributions, all of them non-significant except in the 100-  
591 300 m level ( $0.19 \text{ mm/yr} \pm 0.12 \text{ mm/yr}$ ). Note also that the spread of trends in deep  
592 ocean layers are of the same magnitude of upper ocean trends, indicating that while there  
593 is decrease of contribution to linear trends from top to the bottom of the ocean, the  
594 uncertainty remains unchanged, leading to decreased SSR and suggesting that reanalysis  
595 trends below 700 m are not robust (except for the Indian Ocean).

596 The explained variance shows a predominant role of the 0-100 m level, because of  
597 the seasonality of the the air-sea fluxes. This is particularly notable in the Indian and  
598 Southern Oceans, where the explained variance is equal to 80 and 65%, respectively,  
599 against 43% at global scale. Below 100 m, the values decrease and, in particular below  
600 1500 m, the fluctuation among the products is larger than the value itself, indicating the  
601 different behavior of the products in representing the deep ocean variability.

602 When the seasonal signal is removed (bottom panel of Figure 12), the contribution  
603 of the top 100 m decreases. The 100-300 m level gives the main contribution at global  
604 scale (30%). The Atlantic Ocean shows a similar fraction of explained variance (around  
605 25%) provided by the 0-100, 300-700 and 700-1500 m levels. The Pacific Ocean shows a  
606 dominant contribution of the 0-300 m level, due to the dominant role of the near-surface  
607 tropical waters. Finally, the Southern Ocean provides higher values for deeper levels. In  
608 particular, the contribution of the waters below 1500 m depth sums up to 22% of the  
609 inter-annual variability, in agreement with Ponte (2012), reaching 47% for the waters  
610 below 700 m.

611 Trend maps for thermosteric and halosteric contributions to sea level from the 0-700  
612 m, 700-1500 m, 1500-4000 m and the 4000-bottom levels are shown in Figure 13 and  
613 14 respectively, along with the contour lines corresponding to a SSR equal to 1. The  
614 0-700 m thermosteric trend is found very similar to the 0-bottom trend (middle panel  
615 of Figure 9), with all significant patterns being located in the same areas. Intermediate  
616 waters (between 700 and 1500 m) show significant positive trends (up to 2.5 mm/yr) in  
617 the Labrador Sea and in a some areas of the ACC, while non-significant elsewhere. The  
618 evident warming of the intermediate waters in the Labrador Sea is in agreement with

619 in-situ measurements (Avsic et al., 2006), and may be related to a reduction of deep  
620 convection in this region (Schott and Brandt, 2007).

621 Deep waters (between 1500 and 4000 m) generally show positive trends in the Atlantic  
622 Ocean (up to 4 mm/yr in the Labrador Sea), which is however smaller than the spread  
623 of the ensemble trends.

624 Abyssal waters (bottom right panel) show smaller trends, generally negative in the  
625 Indo-Pacific Ocean and positive in the Atlantic Ocean, except in the Gulf Stream region.  
626 The latter feature agrees with observation-based studies (e.g. Purkey and Johnson, 2010),  
627 which highlight the abyssal warming in the western Atlantic basin coming from the warm-  
628 ing of the Antarctic Bottom Water. Only a few significant areas are visible in the Western  
629 Pacific Ocean. The Pacific Ocean cooling appears to be related to the outlying behavior  
630 of a few individual members that exhibit a large negative in temperature for that area.  
631 Our current approach is to treat all products equally, but in the future it may be better  
632 to discard outlying members.

633 For the halosteric trends, the maps confirm small trends everywhere. For the top 700  
634 m, there exist significant positive values in the ACC, in the region of the South Pacific  
635 Convergence Zone and in the Western North Pacific zone, while significant negative trends  
636 are located only around the Alaskan gyre and in some Atlantic areas. This map agrees well  
637 with the total halosteric map, indicating the significance of the Southern Ocean freshening.  
638 The maps of the intermediate and deep waters suggest a negative (significant) trend in  
639 the Indonesian Throughflow region, up to -4 mm/yr, and in the Alaskan gyre, other  
640 areas exhibiting a smaller and non-significant trend. Finally, although non-significant,  
641 the ensemble mean of the abyssal water halosteric trend shows negative contributions  
642 in the North Atlantic and North and South Pacific gyres, with peaks of -0.6 mm/yr,  
643 while the Gulf Stream region is characterized by a positive trend (up to 0.5 mm/yr). In  
644 contrast, the Indian Ocean shows a uniform positive trend (0.2 mm/yr).

## 645 5 Summary and discussion

646 We have analyzed steric sea level variability from an ensemble of global ocean reanalyses  
647 and objective analyses during the period 1993-2010, in the framework of the ORA-IP  
648 project.

649 The relatively large number of global ocean products included in the comparison al-  
650 lowed us to follow a multi-system ensemble approach and exploit the statistical properties  
651 of the ensemble to detect the consistencies among the different products. It should be  
652 noted that this is in contrast to the atmospheric reanalysis community, where the number  
653 of the state-of-the-art reanalyses is in general of the order of 4 to 6, rendering an ensemble  
654 approach more difficult. We believe that the large number and variety of ocean reanaly-  
655 ses should be extensively exploited in the future. For instance, by objectively discarding  
656 ocean reanalyses with an outlier behavior or weighting the reanalyses in a super-ensemble  
657 context (e.g. Krishnamurti et al., 2000), it will be possible to construct an optimal steric  
658 sea level dataset, which would also contain uncertainty estimates. At this stage, however,  
659 the goal of this comparison is not to build such a dataset, but rather to evaluate the  
660 performance of the reanalyses and their ensemble mean.

661 As a preliminary step towards the assessment of the products included in the com-  
662 parison, we have constructed a validation dataset for the period 2003-2010 by combining  
663 altimetric and gravimetric satellite data. This allowed us to evaluate the performance of  
664 the individual products and that of the ensemble mean, separated into ensemble means  
665 of reanalyses and objective analyses. Within the 8-year validation period, the individual

666 products are satisfactorily able to capture the global steric sea level seasonality, while  
667 they show large discrepancies in the inter-annual trends. The ensemble means agree well  
668 at both the global and regional scale with the altimetry minus gravimetry dataset, and  
669 prove a valuable tool for potential use in studies encompassing longer time periods. Fur-  
670 thermore, the use of the ensemble spread to evaluate the significance of important climate  
671 signals such as inter-annual trends seems an appealing strategy, and reinforces the impor-  
672 tance of sustaining the development and production of multiple reanalyses and adopting  
673 a multi-reanalysis approach to ocean variability studies at inter-decadal scale.

674 The steric sea level variability in the tropics is particularly well represented in the  
675 ensemble of products with a correlation close to 1 with the validating data set. In general,  
676 we found that the ensemble mean of the reanalyses outperforms that of the objective  
677 analyses at both the global and regional scale. This is particularly evident for the Southern  
678 Ocean at the inter-annual scale, and we speculate that this may be caused by the small  
679 amount of in-situ observations and the strong contributions to the total steric sea level  
680 of deep ocean layers in this area, associated to the fact that reanalyses make use of  
681 satellite data and bear information about time-varying atmospheric forcing, affected by  
682 large uncertainty, though.

683 The comparison was then extended to the 1993-2010 period, showing close agreement  
684 of the ensemble of reanalyses and objective analyses in reproducing the steric and ther-  
685 mosteric sea level, with a high cross-correlation. Consequently, the estimates of global  
686 steric sea level trends are similar ( $1.02 \pm 0.05$  and  $1.11 \pm 0.06$  mm/yr, respectively) and  
687 in agreement with recent estimates (Hanna et al., 2013, and references therein). However,  
688 we did not find a significant consensus among the products for the halosteric trends, at  
689 both global and regional scale, although we showed that in a few specific regions (e.g. the  
690 Southern Ocean) its inter-annual variability has effects comparable to the thermosteric  
691 component.

692 Given the approximate linear relationship between global mean halosteric and barystatic  
693 sea level variations due to the land-ice melting affecting both (Munk, 2003), a global mean  
694 barystatic sea level rise of 2 mm/yr roughly corresponds to a global mean halosteric sea  
695 level rise of about 0.05 mm/yr. Our comparison thus indicates that many ocean synthe-  
696 ses have unrealistically large global mean halosteric changes, further to a large ensemble  
697 spread. The disagreement among the global mean halosteric sea level estimates reflects  
698 the uncertainty of the freshwater budget and the disparity among its different model-  
699 ing, intrinsic inaccuracies in the ocean model (e.g. deep ocean variability), as well as  
700 data assimilation assumptions, especially for the pre-Argo era when salinity is in fact  
701 unconstrained or constrained by temperature profiles or sea surface observations only or  
702 climatological assumptions. The global freshwater budget is poorly known, and most re-  
703 analyses adopt a climatological representation for the continental runoff, therefore unable  
704 to follow the variations of the continental ice melting. Along with the commonly adopted  
705 strategy of climatological restoring of the sea surface and/or subsurface salinity, this al-  
706 lows on one hand to mitigate biases arising from atmospheric forcing and ocean model  
707 inaccuracies; on the other hand, it forces the products towards a climatological state and  
708 neglects climate change signals (Griffies et al., 2014). In the future, more sophisticated  
709 methods to correct biases and drifts without compromising the climate change signal  
710 should be explored.

711 Only a very small region of the global ocean has a statistically significant trend in  
712 the halosteric component, although this has recently proven to be non-negligible even  
713 at basin scale (Durack et al., 2014). Our ability to measure the temporal variability of  
714 this component is hampered by the scarcity of salinity observations before the deploy-

715 ment of the Argo float network, with a clear implication regarding the optimal design  
716 and maintenance of the in-situ observing network. This questions the reliability of the  
717 estimates of the halosteric inter-annual variability before the 2000s. It will be essential  
718 in the future to evaluate whether the next generation of ocean reanalyses is able to re-  
719 duce the relative uncertainty of the halosteric component by making better use of the few  
720 observations available (for instance by improving the cross-parameter correlations, the  
721 air-sea coupling, the representation of the global freshwater budget, etc.) or this feature  
722 intrinsically resides in the characteristics of the pre-Argo observing network. This gives  
723 further validation to the idea of promoting and maintaining inter-comparison exercises  
724 in the future. However, provided that the halosteric component explains only a small  
725 portion of the total steric component at both global and regional scales, its uncertainty  
726 does not result in large uncertainties in the total steric sea level.

727 Qualitatively similar conclusions can be drawn when analyzing the separate impact of  
728 different depth levels. While we found a non-negligible effect of the deep waters on inter-  
729 annual variability – about 29% and 12% of explained variance for the waters below 700  
730 and 1500 m ( $\pm 9\%$  and  $\pm 7\%$ ), respectively, with a 1500 m to bottom explained variance  
731 over 20% in the Atlantic and Southern Oceans, steric sea level trends contributed by the  
732 waters below 700 m are generally non-significant with respect to the ensemble spread.  
733 Thus, while the ensemble mean reproduces some notable processes, such as the abyssal  
734 warming in the Western Atlantic basin (Purkey and Johnson, 2010), their quantification  
735 remains difficult, and the reanalyses are not able to provide robust results.

### 736 *Acknowledgments.*

737 During the preparation of this article, our co-author Nicolas Ferry passed away. He was an  
738 active and supportive member of the ORA-IP and CLIVAR-GSOP activities. This work  
739 has received funding from the Italian Ministry of Education, University and Research and  
740 the Italian Ministry of Environment, Land and Sea under the GEMINA project and from  
741 the European Commission Copernicus program, previously known as GMES program,  
742 under the MyOcean and MyOcean2 projects. The authors thank the CLIVAR Project  
743 Office for supporting the participation in conferences and workshops for presenting pre-  
744 liminary results about the steric sea level comparison. Simon Good was supported by the  
745 Joint UK DECC/Defra Met Office Hadley Centre Climate Programme (GA01101). K.A.  
746 Peterson was supported by the UK Public Weather Service research programme. The  
747 GLORY2V3 reanalysis has been developed at Mercator Océan with the support of the  
748 European Commission funded projects MyOcean (FP7-SPACE-2007-1) and MyOcean2  
749 (FP7-SPACE-2011-1). GRACE ocean data were processed by Don P. Chambers, sup-  
750 ported by the NASA MEaSUREs Program, and are available at <http://grace.jpl.nasa.gov>.  
751 The authors would like to thank Don Chambers (USF) for his help in the correct use of  
752 gravimetry data. Mean sea level data were provided by the Sea Level Research Group,  
753 University of Colorado. The altimeter products were produced by Ssalto/Duacs and dis-  
754 tributed by AVISO, with support from CNES (<http://www.aviso.oceanobs.com/duacs/>).  
755 The authors are also very grateful to three anonymous reviewers and the editor for their  
756 precious suggestions and valuable comments, and for the improvement of the quality of  
757 this paper.

758 **Appendix: Mathematical definitions**

759 We briefly introduce in this Appendix some mathematical definitions that are used in the  
760 paper.

761 **Signal-to-spread Ratio**

762 In order to evaluate how distinguishable is the climate signal of the reanalysis ensemble  
763 with respect to its uncertainty, we define the signal-to-spread ratio (*SSR*), or more  
764 generally, the signal-to-noise ratio of the ensemble for a generic parameter  $p$  as

$$SSR = \frac{EM}{ES} = \frac{\langle p \rangle}{\sqrt{1/N \sum_{i=1}^{i=N} (p_i - \langle p \rangle)^2}} \quad (5)$$

765 with  $EM$  and  $ES$  being the ensemble mean and spread, respectively, and  $N$  being the  
766 ensemble size, with

$$\langle p \rangle = 1/N \sum_{i=1}^{i=N} p_i. \quad (6)$$

767 Note that the ensemble spread is defined as the sample standard deviation. Values of  
768 *SSR* smaller (greater) than 1 indicate that the discrepancy of reanalyses is greater (less)  
769 than their mean signal.

770 **Annual and Seasonal Decomposition**

771 It is also useful to decompose the steric sea level signal onto the seasonal (annual and  
772 semi-annual) and linear trend (inter-annual) components. To do this, we assume that  
773 every time-series of the variable  $x$  be of the form:

$$x(t) = mt + c + A_a \cos\left(\frac{2\pi}{12}t - \varphi_a\right) + A_s \cos\left(\frac{2\pi}{6}t - \varphi_s\right) + \varepsilon(t) \quad (7)$$

774 where  $t$  is the time (in months),  $m$  is the linear trend,  $A_a$  and  $\varphi_a$  are the annual amplitude  
775 and angular phase, respectively, and  $A_s$  and  $\varphi_s$  are the semi-annual amplitude and angular  
776 phase and  $\varepsilon$  are the residuals. The decomposition is carried out by a least-squares fitting  
777 of Equation (7), i.e. by minimizing the sum of  $\varepsilon^2(t)$ .

778 The seasonal signal  $x_S$ , introduced in the text in Sections 3 and 4, is defined as the  
779 full signal minus the linear trend:

$$x_S(t) = x(t) - mt, \quad (8)$$

780 namely it corresponds to the detrended signal. Conversely, the inter-annual signal  $x_I$ , is  
781 defined as the full signal to which the fitted seasonal signal is subtracted:

$$x_I(t) = x(t) - A_a \cos\left(\frac{2\pi}{12}t - \varphi_a\right) - A_s \cos\left(\frac{2\pi}{6}t - \varphi_s\right). \quad (9)$$

782 In the above definitions, only the complementary fitted signal is removed, while the  
783 residuals are always kept. Note that the time-series corresponding to the inter-annual  
784 and seasonal signals have the same length of the time-series of the full signal, implying  
785 that the same minimum values for testing the significance of the correlations apply.

786 **Explained Variance**

787 The (percentage) explained variance of a component  $y$  with respect to the (total) com-  
788 ponent  $z$  is defined as

$$EV(y) = 100 \frac{VAR(z) - VAR(z - y)}{VAR(z)}, \quad (10)$$

789 with  $VAR(\dots)$  being the variance operator. When the explained variance of the inter-  
790 annual signal is introduced (e.g. in Figure 11), it means that the explained variance is  
791 calculated on the timeseries, after removal of the seasonal signal, for both components  $y$   
792 and  $z$ .

## 793 References

- 794 Abraham, J., et al., 2013: A review of global ocean temperature observations: Implica-  
795 tions for ocean heat content estimates and climate change. *Rev. Geophys.*, **51**, 450–483.
- 796 Avsic, T., J. Karstensen, U. Send, and J. Fischer, 2006: Interannual variability of newly  
797 formed Labrador Sea Water from 1994 to 2005. *Geophys. Res. Lett.*, **33**, L21S02.
- 798 Balmaseda, M., et al., 2015: The Ocean Reanalyses Intercomparison Project (ORA-IP).  
799 *J. Oper. Oceanogr.*, **7**, 31 pp, accepted for publication.
- 800 Balmaseda, M. A., K. Mogensen, and A. Weaver, 2012: Evaluation of the ECMWF ocean  
801 reanalysis system ORAS4. *Q. J. R. Meteorol. Soc.*, **139**, 1132–1161.
- 802 Behringer, D., 2007: The global ocean data assimilation system at NCEP. *11th Symp.*  
803 *on Integrated Observing and Assimilation Systems for Atmosphere, Oceans and Land*  
804 *Surface*, Amer. Meteor. Soc.
- 805 Blockley, E. W., et al., 2014: Recent development of the Met Office operational ocean  
806 forecasting system: an overview and assessment of the new Global FOAM forecasts.  
807 *Geoscientific Model Development Discussions*, **6** (4), 6219–6278, doi:10.5194/gmdd-6-  
808 6219-2013, URL <http://www.geosci-model-dev-discuss.net/6/6219/2013/>.
- 809 Böning, C., A. Dispert, M. Visbeck, S. Rintoul, and F. Schwarzkopf, 2008: The response  
810 of the Antarctic Circumpolar Current to recent climate change. *Nature Geoscience*, **1**,  
811 864–869.
- 812 Cabanes, C., et al., 2013: The CORA dataset: validation and diagnostics of in-situ ocean  
813 temperature and salinity measurements. *Ocean Sci.*, **9**, 1–18.
- 814 Cazenave, A. and W. Llovel, 2010: Contemporary Sea Level Rise. *Annu. Rev. of Marine*  
815 *Sci.*, **2**, 145–173.
- 816 Chambers, D., 2006a: Observing seasonal steric sea level variations with GRACE and  
817 satellite altimetry. *J. Geophys. Res.*, **111**, C03 010, doi:10.1029/2005JC004836.
- 818 Chambers, D. and J. Bonin, 2012: Evaluation of Release-05 GRACE time-variable gravity  
819 coefficients over the ocean. *Ocean Sci.*, **8**, 859–868.
- 820 Chambers, D. and B. Schröter, 2011: Measuring ocean mass variability from satellite  
821 gravimetry. *J. Geodyn.*, **52**, 333–343.
- 822 Chambers, D. P., 2006b: Evaluation of new GRACE timevariable gravity data over the  
823 ocean. *Geophys. Res. Lett.*, **33**, LI7603, doi:10.1029/2006GL027296.
- 824 Chang, Y., S. Zhang, A. Rosati, T. Delworth, and W. Stern, 2013: An assessment of  
825 oceanic variability for 1960-2010 from the GFDL ensemble coupled data assimilation.  
826 *Clim. Dynam.*, **40**, 775–803.
- 827 Chang, Y.-S., A. Rosati, and G. A. Vecchi, 2014a: Basin patterns of global sea level  
828 changes for 2004-2007. *J. Mar. Syst.*, **80**, 115–124.
- 829 Chang, Y.-S., G. A. Vecchi, A. Rosati, S. Zhang, and X. Yang, 2014b: Comparison of  
830 global objective analyzed T-S fields of the upper ocean for 2008-2011. *J. Mar. Syst.*,  
831 **137**, 13–20.

- 832 de Boyer Montégut, C., G. Madec, A. S. Fischer, A. Lazar, and D. Iudicone, 2004: Mixed  
833 layer depth over the global ocean: an examination of profile data and a profile-based  
834 climatology. *J. Geophys. Res.*, **109**, c12003.
- 835 Durack, P., S. Wijffels, and P. Gleckler, 2014: Long-term sea-level change revisited: the  
836 role of salinity. *Environ. Res. Lett.*, **9**, 114 017–1–11.
- 837 Forget, G. and R. Ponte, submitted: The Partition of Regional Sea Level Variability.  
838 *Progress in Oceanography*.
- 839 Fujii, Y., N. Nakaegawa, S. Matsumoto, T. Yasuda, G. Yamanaka, and M. Kamachi, 2009:  
840 Coupled climate simulation by constraining ocean fields in a coupled model with ocean  
841 data. *J. Climate*, **22**, 5541–5557.
- 842 Fukumori, I., 2002: A partitioned Kalman filter and smoother. *Mon. Wea. Rev.*, **130**,  
843 1370–1383.
- 844 Fukumori, I., R. Raghunath, L. Fu, and Y. Chao, 1999: Assimilation of  
845 TOPEX/POSEIDON data into a global ocean circulation model: How good are the  
846 results? *J. Geophys. Res.*, **104**, 25 647–25 665.
- 847 Fukumori, I. and O. Wang, 2013: Origins of heat and freshwater anomalies underlying  
848 regional decadal sea level trends. *Geophys. Res. Lett.*, **40**, 563–567.
- 849 Garcia-Garcia, D., B. Chao, and J.-P. Boy, 2010: Steric and mass-induced sea level  
850 variations in the Mediterranean Sea revisited. *J. Geophys. Res.*, **115**, C12 016,  
851 doi:10.1029/2009JC005928.
- 852 Gill, A. and P. P. Niiler, 1973: The theory of the seasonal variability in the ocean. *Deep-*  
853 *Sea Res.*, **20**, 141–177.
- 854 Gregory, J., et al., 2013: Twentieth-Century Global-Mean Sea Level Rise: Is the Whole  
855 Greater than the Sum of the Parts? *J. Climate*, **26**, 4476–4499.
- 856 Griffies, S. and R. Greatbatch, 2012: Physical processes that impact the evolution of  
857 global mean sea level in ocean climate models. *Ocean Model.*, **51**, 37–72.
- 858 Griffies, S., et al., 2014: An assessment of global and regional sea level for years 1993-2007  
859 in a suite of interannual CORE-II simulations. *Ocean Model.*, **78**, 35–89.
- 860 Guinehut, S., A. Dhomp, G. Larnicol, and P.-Y. Le Traon, 2012: High resolution 3D  
861 temperature and salinity fields derived from in situ and satellite observations. *Ocean*  
862 *Sci.*, **8**, 845–857.
- 863 Haines, K., M. Valdivieso, H. Zuo, and V. Stepanov, 2012: Transports and budgets in a  
864 1/4 global ocean reanalysis 1989-2010. *Ocean Sci.*, **8**, 333–344.
- 865 Hanna, E., et al., 2013: Ice-sheet mass balance and climate change. *Nature*, **498**, 51–59.
- 866 Ingleby, B. and M. Huddleston, 2007: Quality control of ocean temperature and salinity  
867 profiles - Historical and real-time data. *J. Mar. Syst.*, **65**, 158–175.
- 868 IPCC, 2013: *Climate Change 2013: The Physical Science Basis. Contribution of Working*  
869 *Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate*  
870 *Change*. Cambridge University Press, Cambridge, United Kingdom and New York,  
871 NY, USA, 1535 pp.



- 872 Ishii, M., M. Kimoto, K. Sakamoto, and S.-I. Iwasaki, 2006: Steric sea level changes esti-  
873 mated from historical ocean subsurface temperature and salinity analyses. *J. Oceanogr.*,  
874 **62**, 155–170.
- 875 Ivchenko, V., S. Danilov, D. Sidorenko, J. Schröter, M. Wenzel, and L. Aleynik, 2008:  
876 Steric height variability in the Northern Atlantic on seasonal and interannual scales. *J.*  
877 *Geophys. Res.*, **113**, C11 007, doi:10.1029/2008JC004836.
- 878 Johnson, G. and D. Chambers, 2013: Ocean bottom pressure seasonal cycles and decadal  
879 trends from GRACE Release-05: Ocean circulation implications. *J. Geophys. Res.*, **118**,  
880 4228–4240.
- 881 Köhl, A., 2014: Evaluation of the GECCO2 Ocean Synthesis: Transports of Volume, Heat  
882 and Freshwater in the Atlantic. *Q. J. R. Meteorol. Soc.*, doi:10.1002/qj.2347.
- 883 Krishnamurti, T., C. Kishtawal, Z. Zhang, T. Larow, D. Bachiochi, and E. Williford,  
884 2000: Multimodel ensemble forecasts for weather and seasonal climate. *J. Climate*, **13**,  
885 4196–4216.
- 886 Landerer, F., J. Jungclauss, and J. Marotze, 2007: Regional Dynamic and Steric Sea Level  
887 Change in Response to the IPCC-A1B Scenario. *J. Phys. Oceanogr.*, **37**, 296–312.
- 888 Le Traon, P., F. Nadal, and N. Ducet, 1998: An improved mapping method of multi-  
889 satellite altimeter data. *J. Atmos. Ocean. Technol.*, **15**, 522–534.
- 890 Lee, T., T. Awaji, M. Balmaseda, E. Grenier, and D. Stammer, 2009: Ocean State  
891 Estimation for Climate Research. *Oceanography*, **22**, 160–167.
- 892 Leuliette, E. and W. Miller, 2009: Closing the sea level rise budget with altimetry, Argo,  
893 and GRACE. *Geophys. Res. Lett.*, **36**, L04 608, doi:10.1029/2008GL036010.
- 894 Leuliette, E. and J. Willis, 2011: Balancing the sea level budget. *Oceanography*, **24**,  
895 122–129.
- 896 Levitus, S., et al., 2012: World ocean heat content and thermosteric sea level change  
897 (0–2000 m), 1955–2010. *Geophys. Res. Lett.*, **39**, L10 603.
- 898 Llovel, W., S. Guinehut, and A. Cazenave, 2010: Regional and interannual variability in  
899 sea level over 2002–2009 based on satellite altimetry, Argo float data and GRACE ocean  
900 mass. *Ocean Dynam.*, **60**, 1193–1204.
- 901 Lombard, A., G. Garric, and T. Penduff, 2009: Regional patterns of observed sea level  
902 change: insights from a 1/4 global ocean/sea-ice hindcast. *Ocean Dyn.*, **59**, 433–449.
- 903 Lombard, A., et al., 2007: Estimation of steric sea level variations from combined GRACE  
904 and Jason-1 data. *Earth Planet. Sci. Lett.*, **254**, 194–202.
- 905 Lowe, J. and J. Gregory, 2006: Understanding projections of sea level rise in a Hadley  
906 Centre coupled climate model. *J. Geophys. Res.*, **111**, C11 014.
- 907 Masuda, S., et al., 2010: Simulated Rapid Warming of Abyssal North Pacific Waters.  
908 *Science*, **329**, 319–322.
- 909 Munk, W., 2003: Ocean Freshening, Sea Level Rising. *Science*, **300**, 2041–2043.

- 910 Nerem, R., D. Chambers, C. Choe, and G. Mitchum, 2010: Estimating Mean Sea Level  
911 Change from the TOPEX and Jason Altimeter Missions. *Mar. Geodesy*, **33**, 435–446.
- 912 Piecuch, C. and R. Ponte, 2011: Mechanisms of interannual steric sea level variability.  
913 *Geophys. Res. Lett.*, **38**, L15 605.
- 914 Piecuch, C., K. Quinn, and R. Ponte, 2013: Satellite-derived interannual ocean bottom  
915 pressure variability and its relation to sea level. *Geophys. Res. Lett.*, **40**, 3106–3110.
- 916 Ponte, R. M., 2006: Oceanic response to surface loading effects neglected in volume-  
917 conserving models. *J. Phys. Oceanogr.*, **36**, 426–434.
- 918 Ponte, R. M., 2012: An assessment of deep steric height variability over the global ocean.  
919 *Geophys. Res. Lett.*, **39**, L04 601, doi:10.1029/2011GL050681.
- 920 Purkey, S. and G. Johnson, 2010: Warming of Global Abyssal and Deep Southern Ocean  
921 Waters between the 1990s and 2000s: Contributions to Global Heat and Sea Level Rise  
922 Budgets. *J. Climate*, **23**, 6336–6351.
- 923 Purkey, S. and G. Johnson, 2013: Antarctic Bottom Water Warming and Freshening:  
924 Contributions to Sea Level Rise, Ocean Freshwater Budgets, and Global Heat Gain. *J.*  
925 *Climate*, **26**, 6105–6122.
- 926 Quinn, K. J. and R. M. Ponte, 2010: Uncertainty in ocean mass trends from GRACE.  
927 *Geophys. J. Intern.*, **181**, 762–768, doi:10.1111/j.1365-246X.2010.04508.x.
- 928 Schott, F. and P. Brandt, 2007: *Ocean Circulation: Mechanisms and Impacts - Past and*  
929 *Future Changes of Meridional Overturning*, chap. Circulation and Deep Water Export  
930 of the Subpolar North Atlantic During the 1990’s, 91–118. Edited by A. Schmittner, J.  
931 C. H. Chiang and S. R. Hemming. American Geophysical Union, Washington, D. C.
- 932 Stammer, D., A. Cazenave, R. Ponte, and M. Tamisiea, 2013: Causes for contemporary  
933 regional sea level changes. *Annu. Rev. of Marine Sci.*, **5**, 21–46.
- 934 Stammer, D., et al., 2010: Ocean Information Provided through Ensemble Ocean Syn-  
935 theses. *Sustained Ocean Observations and Information for Society*, OceanObs’09, 21-25  
936 September 2009, J. Hall, D. Harrison, D. Stammer Eds., ESA Publication WPP-306.
- 937 Steiger, J., 1980: Tests for comparing elements of a correlation matrix. *Psychological*  
938 *Bull.*, **87**, 245–251.
- 939 Storto, A., S. Dobricic, S. Masina, and P. Di Pietro, 2011: Assimilating along-track alti-  
940 metric observations through local hydrostatic adjustments in a global ocean reanalysis  
941 system. *Mon. Wea. Rev.*, **139**, 738–754.
- 942 Storto, A., S. Masina, and S. Dobricic, 2013: Ensemble spread-based assessment of obser-  
943 vation impact: application to a global ocean analysis system. *Q. J. R. Meteorol. Soc.*,  
944 **139**, 1842–1862.
- 945 Storto, A., S. Masina, and S. Dobricic, 2014: Estimation and Impact of Nonuniform  
946 Horizontal Correlation Length Scales for Global Ocean Physical Analyses. *J. Atmos.*  
947 *Ocean. Technol.*, **31**, 2330–2349.

- 948 Sukuki, T. and M. Ishii, 2011: Regional distribution of sea level changes resulting from  
949 enhanced greenhouse warming in the Model for Interdisciplinary Research on Climate  
950 version 3.2 version. *Geophys. Res. Lett.*, **38**, L02601.
- 951 Toyoda, T., Y. Fujii, T. Yasuda, N. Usui, T. Iwao, T. Kuragano, and M. Kamachi, 2013:  
952 Improved analysis of the seasonal-interannual fields by a global ocean data assimilation  
953 system. *Theoret. and Appl. Mech. Japan*, **61**, 31–48.
- 954 Weaver, A. T., C. Deltel, E. Machu, S. Ricci, and N. Daget, 2005: A multivariate balance  
955 operator for variational ocean data assimilation. *Q. J. R. Meteorol. Soc.*, **131**, 605–625.
- 956 Willis, J., D. Chambers, and R. Nerem, 2008: Assessing the globally averaged sea level  
957 budget on seasonal to interannual timescales. *J. Geophys. Res.*, **113**, C06015.
- 958 Wunsch, C. and P. Heimbach, 2013: *Ocean Circulation and Climate - A 21st century*  
959 *perspective, International Geophysics Series, Vol.103*, chap. Chapter 21: Dynamically  
960 and Kinematically Consistent Global Ocean Circulation and Ice State Estimates, 553–  
961 579. Edited by G. Sielder, J. Church, S. Griffes, J. Gould, and J. Church. Academic  
962 Press, Elsevier.
- 963 Xue, Y., B. Huang, Z. Hu, A. Kumar, C. Wen, D. Behringer, and S. Nadiga, 2011: An  
964 Assessment of Oceanic Variability in the NCEP Climate Forecast System Reanalysis.  
965 *Clim. Dynam.*, **37**, 2511–2539.
- 966 Yin, J., S. Griffies, and R. Stouffer, 2010: Spatial Variability of Sea Level Rise in Twenty-  
967 First Century Projections. *J. Climate*, **23**, 4585–4607.
- 968 Yin, Y., O. Alves, and P. Oke, 2011: An ensemble ocean data assimilation system for  
969 seasonal prediction. *Mon. Wea. Rev.*, **139**, 786–808.

Table 1: List of products participating the steric sea level comparison, with details about a bibliographic reference, the product type, the spatial resolution and the data assimilation configuration. The symbol <sup>+</sup> indicates that the product is the ocean component of a coupled ocean-atmosphere system. For the resolution, the symbol \* indicates that the computational grid has a resolution refinement in the Equatorial region. The penultimate column reports the data assimilation method used and the observations assimilated (T: in-situ observations of temperature; S: in-situ observations of salinity; SST: satellite observations of sea surface temperature; SLA: altimetric observations). The last column reports additional constraints included in the product (SSS: restoring to climatological sea surface salinity; 3D: three-dimensional restoring to climatological fields ; BC: bias-correction ). The product GLORYS2V3 is called GLORYS2 hereafter.

Product Name	Producing Institute	Reference	Type of Product	Resolution	Assimilation Method (Obs)	Additional Constraint
ARMOR	CLS	Guinehut et al. (2012)	OA	1/3 x 1/3	OI(T, S, SST, SLA)	NO
CFSR <sup>+</sup>	NOAA/NCEP	Xue et al. (2011)	REA	1/2 x 1/2*	3DVAR(T, SST)	SSS
C-GLORS	CMCC	Storto et al. (2014)	REA	1/2 x 1/2	3DVAR(T, S, SST, SLA)	SSS + BC
CORA	Ifremer	Cabanes et al. (2013)	OA	1/3 x 1/3	OI(T, S)	NO
ECCO-NRT	JPL/NASA	Fukumori (2002)	REA	1 x 1*	KF-SS(T, SLA)	SSS
ECCOV4	MIT/AER/JPL	Forget and Ponte (submitted)	REA	1 x 1*	4DVAR(T, S, SST, SLA)	NO
ECDA <sup>+</sup>	NOAA/GFDL	Chang et al. (2013)	REA	1 x 1*	EnKF(T, S, SST)	NO
EN3	UK MetOffice	Ingleby and Huddleston (2007)	OA	1 x 1	OI(T, S)	3D
GECCO2	Univ. Hamburg	Köhl (2014)	REA	1 x 1*	4DVAR(T, S, SST, SLA)	SSS + 3D
GEOS5	NASA/GMAO		REA	1/2 x 1/2	EnOI(T, S, SST, SLA)	SSS
GLORYS2V3	MERCATOR		REA	1/4 x 1/4	KF(T, S, SST, SLA)	BC
GLOSEA5	UK MetOffice	Blockley et al. (2014)	REA	1/4 x 1/4	3DVAR(T, S, SST, SLA)	SSS + 3D
GODAS	NOAA/NCEP	Behringer (2007)	REA	1 x 1*	3DVAR(T, SST, SLA)	SSS
IK09	JAMSTEC	Ishii et al. (2006)	OA	1 x 1	OI(T, S)	NO
K7OC (ESTOC)	JAMSTEC	Masuda et al. (2010)	REA	1 x 1	4DVAR(T, S, SST, SLA)	BC
MOVEC <sup>+</sup>	MRI/JMA	Fujii et al. (2009)	REA	1 x 1*	3DVAR(T, S, SST, SLA)	SSS + 3D + BC
MOVEG2	MRI/JMA	Toyoda et al. (2013)	REA	1 x 1/2*	3DVAR(T, S, SST, SLA)	3D
ORAS4	ECMWF	Balmaseda et al. (2012)	REA	1 x 1*	3DVAR(T, S, SST, SLA)	SSS + 3D + BC
PEODAS	BoM/CAWCR	Yin et al. (2011)	REA	2 x 1*	EnKF(T, S, SST)	SSS + 3D
UR025.4	Univ. Reading	Haines et al. (2012)	REA	1/4 x 1/4	OI(T, S, SST, SLA)	NO

Table 2: Spatially area averages of the point-by-point temporal correlation (2003-2010) of the steric sea level product with the verifying dataset (ALT-GRV). The four regions are defined with respect to latitudinal bands: between 60S and 60N for the global average, and between 20N and 60N, 20S and 20N and 60S and 20S for the Northern Extra-Tropics, the Tropics and the Southern Extra-Tropics, respectively.

Product	Globally Averaged Correlation	Northern Extra-Tropics Correlation	Tropics Correlation	Southern Extra-Tropics Correlation
ARMOR	0.762	0.760	0.853	0.673
CFSR	0.527	0.511	0.701	0.364
C-GLORS	0.841	0.852	0.917	0.762
CORA	0.531	0.509	0.730	0.346
ECCO-NRT	0.618	0.564	0.828	0.438
ECCOV4	0.631	0.580	0.835	0.455
ECDA	0.526	0.502	0.759	0.308
EN3	0.512	0.484	0.726	0.315
GECCO2	0.616	0.501	0.801	0.488
GEOS5	0.555	0.490	0.816	0.329
GLORYS2	0.865	0.876	0.906	0.820
GLOSEA5	0.831	0.896	0.902	0.731
GODAS	0.514	0.493	0.729	0.313
IK09	0.541	0.535	0.754	0.335
K7OC	0.449	0.465	0.533	0.359
MOVEC	0.647	0.687	0.784	0.494
MOVEG2	0.707	0.698	0.869	0.552
ORAS4	0.627	0.579	0.861	0.419
PEODAS	0.565	0.522	0.778	0.375
UR025.4	0.757	0.806	0.895	0.599
REAENS	0.799	0.787	0.901	0.704
OAENS	0.647	0.640	0.818	0.482
ALLENS	0.780	0.766	0.894	0.675

Table 3: Standard deviation, spread and normalized spread of the two ensemble means (REA: reanalyses; OA: objective analyses) for the global steric, thermo- and halo- steric sea level during the period 1993-2010. The table reports in mm the standard deviation of the ensemble means, along with the time-averaged ensemble spread and the ratio between the latter and the standard deviation. The same computation is reported also for the halosteric sea level, by categorizing the reanalyses in four different groups, depending on constraints on salinity (NOREST: no restoring to SSS, nor to subsurface fields; BCORR: bias correction scheme implemented; SREST : restoring to climatological SSS; 3DREST: restoring to climatological subsurface fields). For the latter, products were subsampled by computing the spread and variability of all combinations of two members included in each group, and then averaging over the combinations, in order to avoid the influence of the group size on the results. Note that each product may belong to more than one group.

Products	Steric Component	Monthly Variability	Mean Spread	Normalized Spread
REA	Steric	8.12	6.40	0.79
	Thermosteric	7.17	6.78	0.95
	Halosteric	2.94	5.02	1.71
OA	Steric	6.27	5.22	0.83
	Thermosteric	6.23	5.35	0.86
	Halosteric	0.75	1.21	1.61
NOREST	Halosteric	2.03	2.75	1.35
BCORR	Halosteric	3.85	4.91	1.28
SREST	Halosteric	3.37	4.05	1.20
3DREST	Halosteric	3.79	4.26	1.12

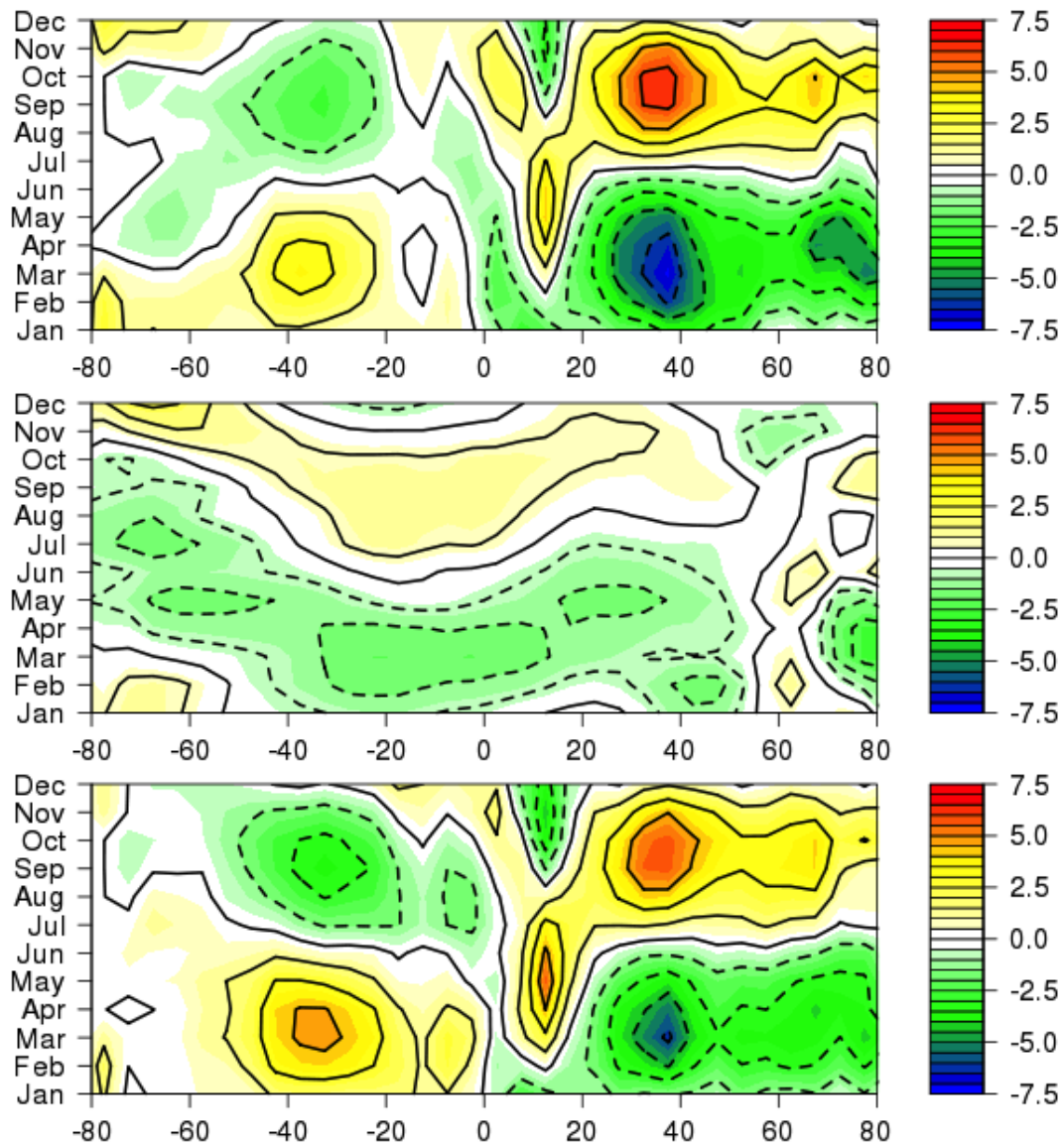


Figure 1: Seasonal cycle (2003-2010 mean) as a function of latitude for the total sea level (top panel) derived by satellite altimetry, mass component (middle panel), derived by satellite gravimetry and steric sea level ALT-GRV (bottom panel), derived by subtracting the barystatic sea level from the total sea level. Data processing is explained in the text. Units are cm.

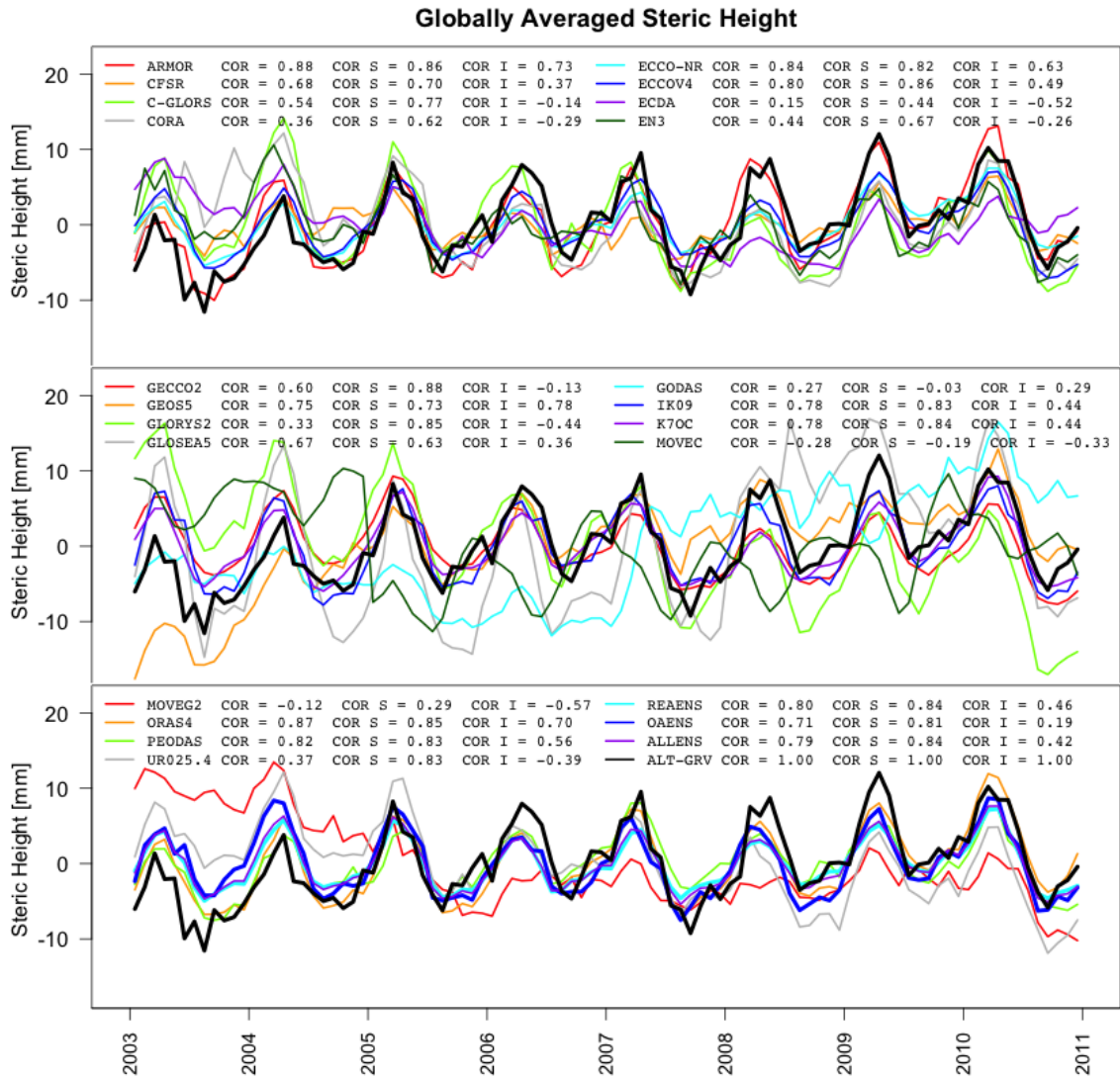


Figure 2: Monthly time-series of global steric sea level for the period 2003-2010 for the different products and the verifying dataset (black lines). The temporal correlation of each product with the verifying dataset for the full signal (COR), the seasonal signal (COR S, i.e. interannual signal removed) and the inter-annual signal (COR I, i.e. seasonal signal removed) is also shown.



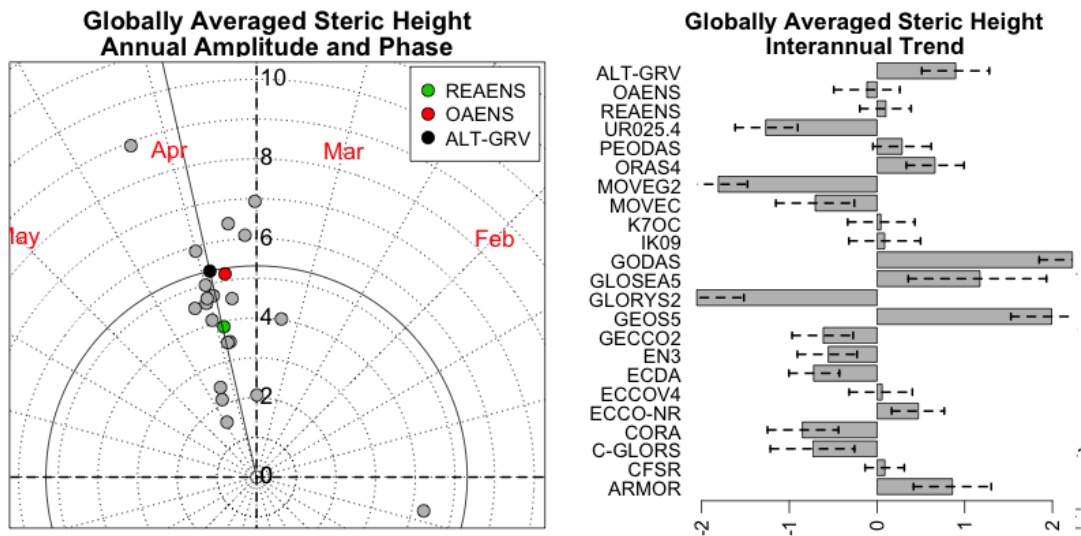


Figure 3: Decomposition of the 2003-2010 steric sea level in annual and inter-annual (linear trend) components for all the products (see Appendix for the definitions). For the annual component, the plot shows the amplitude and phase in polar coordinates, with the radius corresponding to the amplitude and the angle with respect to the x-axis to the phase (corresponding to the maximum reached in the annual cycle), for the individual products (gray circles), the ensemble of the reanalyses (green circle) and the objective analyses (red circle) and the verifying dataset (black circle). The y-axis reports the values for the amplitude in mm, while the phase is given in months from the beginning of the year, respectively, and reported in red. Note that the month labels are located at the middle of the month, and the radial graduation is by half month. For the inter-annual linear trend, units are  $\text{mm yr}^{-1}$ .

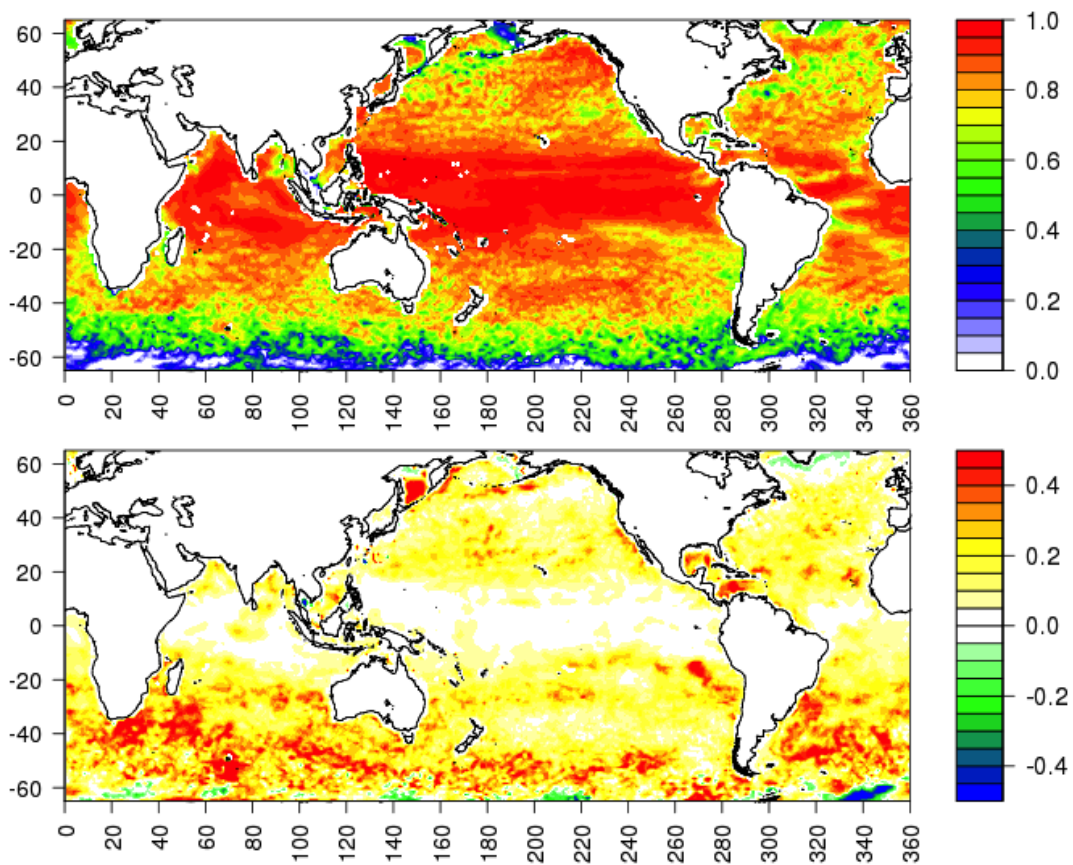


Figure 4: Correlation map of the reanalysis ensemble with the verifying dataset (top panel) and difference of correlation with the verifying dataset between REAENS and OAENS (bottom panel). Note that the color palettes are different.

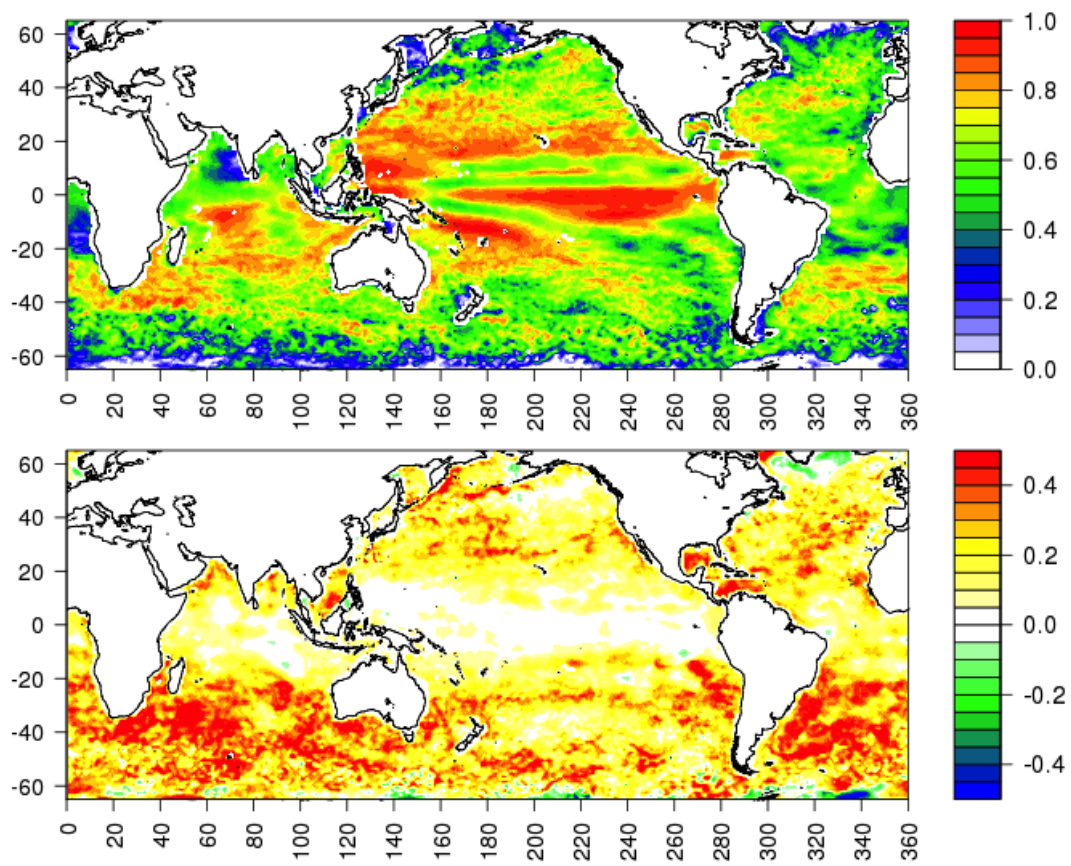


Figure 5: As in Figure 4, but for the inter-annual signal only (seasonal signal removed).

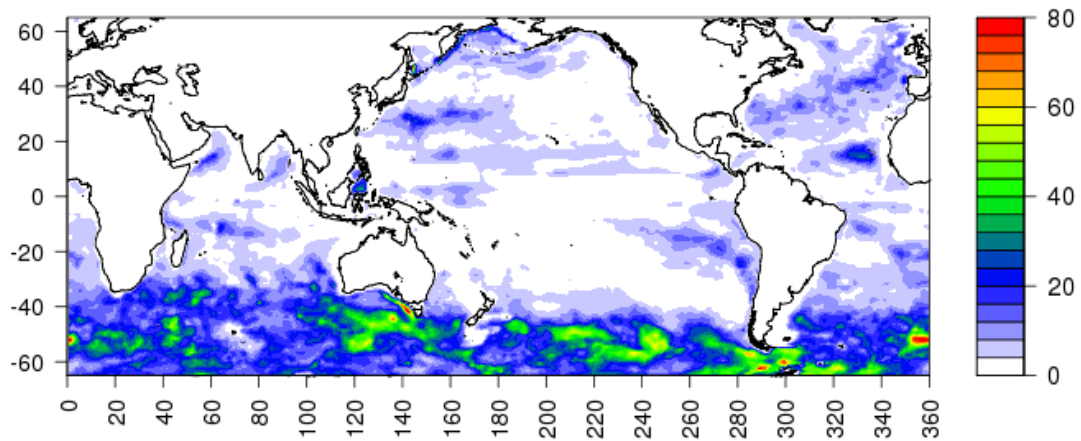


Figure 6: Explained variance of the steric sea level in the layer 700 m to bottom with respect to the surface to bottom variance, calculated on the reanalysis ensemble mean during the period 2003-2010. Units are %. The figure provides a quantitative evaluation of the percentage impact of the relatively deep ocean (below 700 m of depth) on the steric sea level variability.

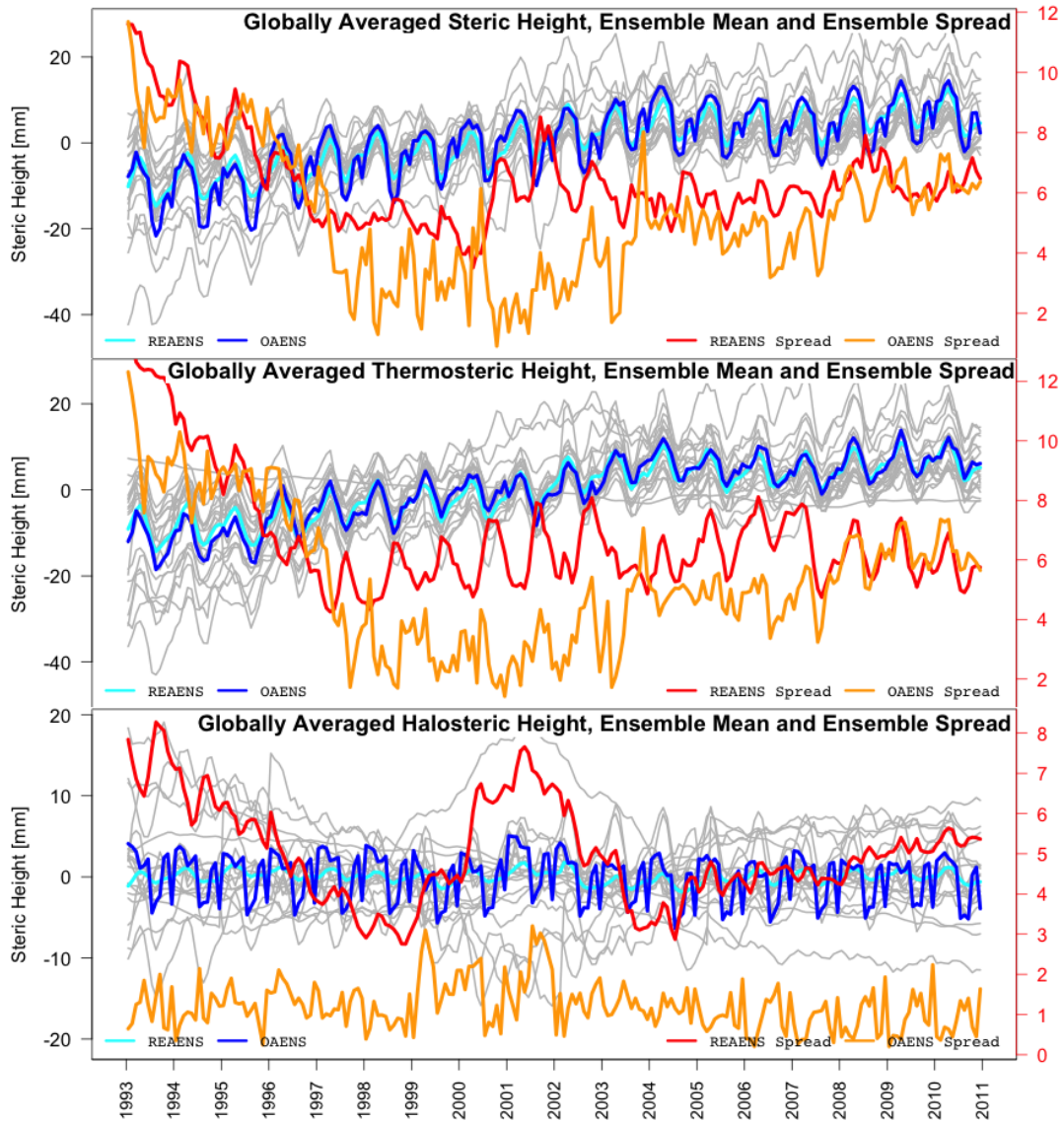


Figure 7: Monthly time-series of global steric (top panel), thermosteric (middle panel) and halosteric (bottom panel) sea level for the period 1993-2010 (Extended Intercomparison Period). Gray lines correspond to the individual products, while the cyan (blue) line corresponds to the ensemble mean of reanalyses (objective analyses). Also shown in red (orange) is the time-series of the ensemble spread from the reanalyses (objective analyses). For the ensemble spreads, the y-axis is in red on the right-side of the panels and the unit is mm.

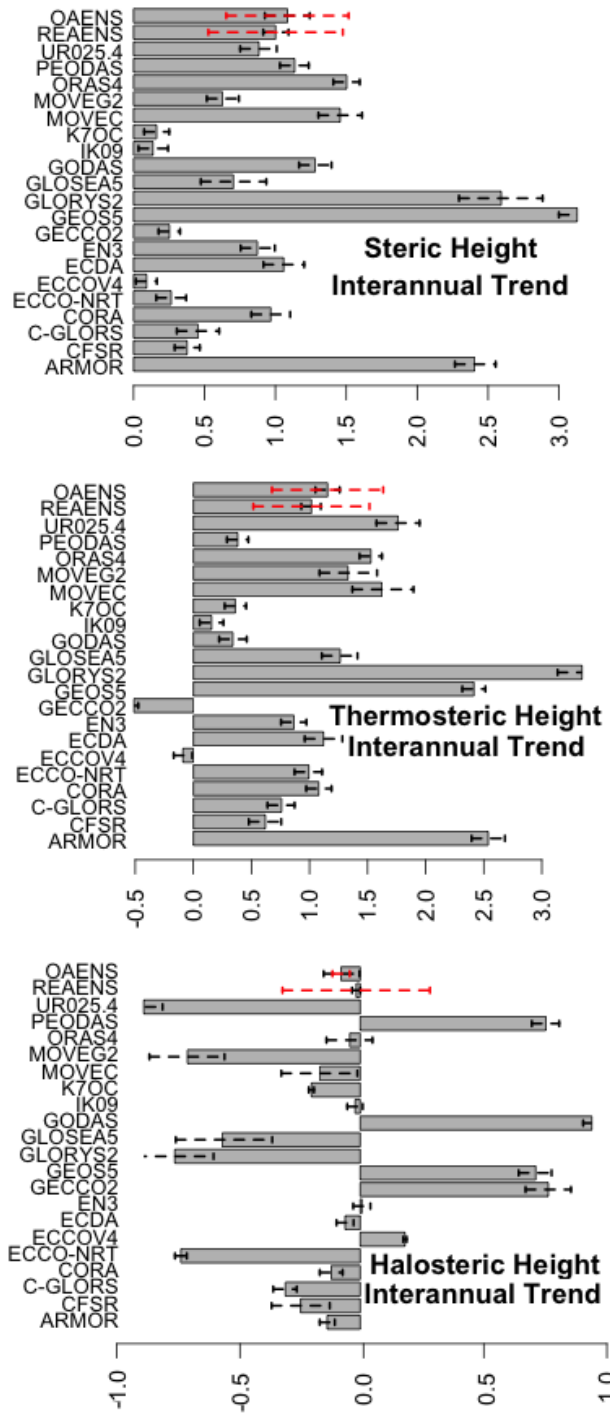


Figure 8: Global sea level linear trends (1993-2010) from all the products, the ensemble mean of reanalyses (REAENS) and objective analyses (OAENS) for the steric (top panel), thermosteric (middle panel) and halosteric (bottom panel) sea level, with the 95% confidence level calculated using a bootstrap algorithm. Units are mm yr<sup>-1</sup>. For REAENS and OAENS, red bars correspond to the spread (standard deviation) of the trends from the individual products.

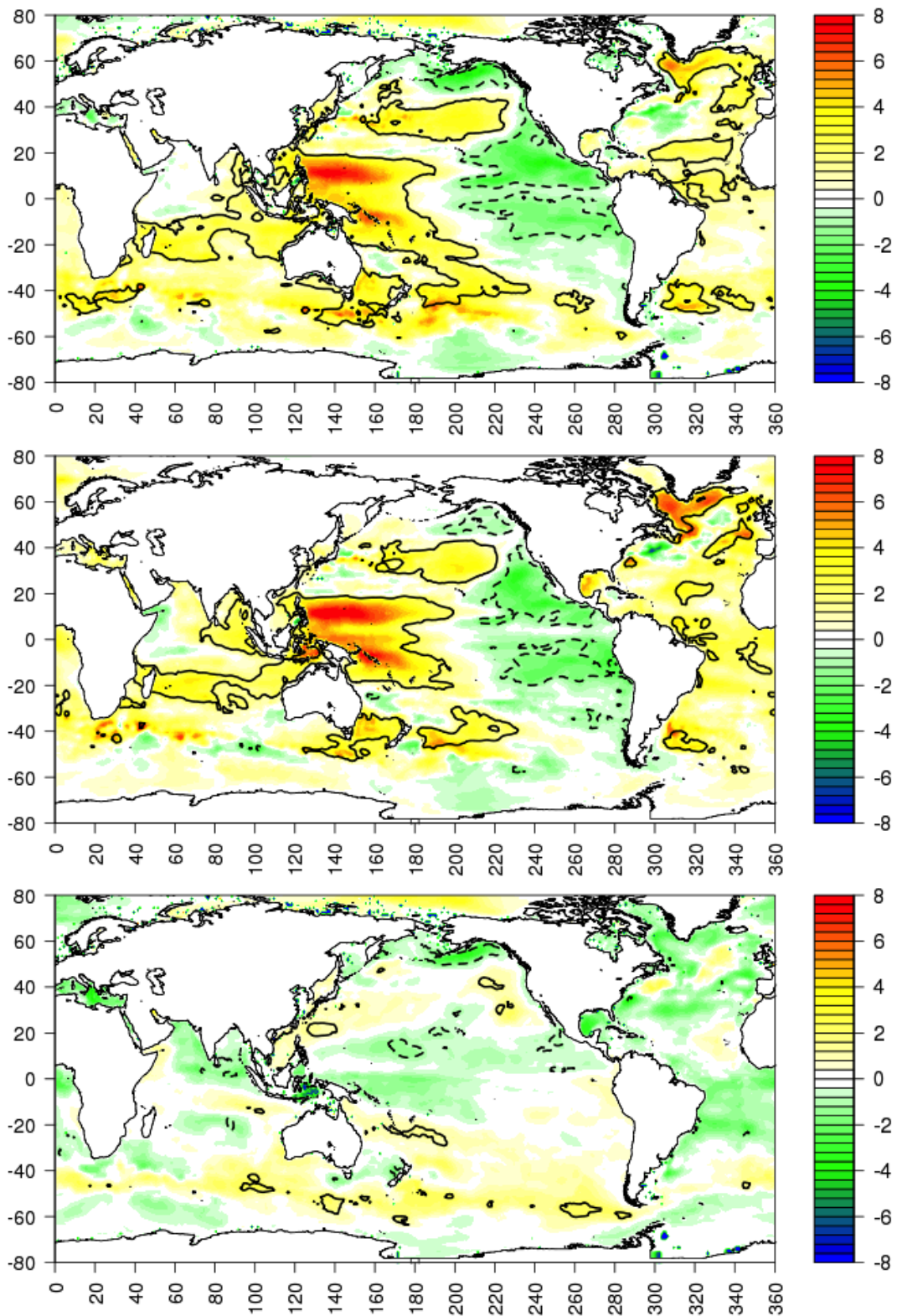


Figure 9: Map of 1993-2010 linear trends ensemble mean (including all the products) for the steric (top panel), thermosteric (middle panel) and halosteric (bottom panel) sea level. Units are  $\text{mm yr}^{-1}$ . Solid (dashed) contour lines denote regions of signal-to-noise ratio (ensemble mean divided by ensemble spread) equal to 1 (-1).

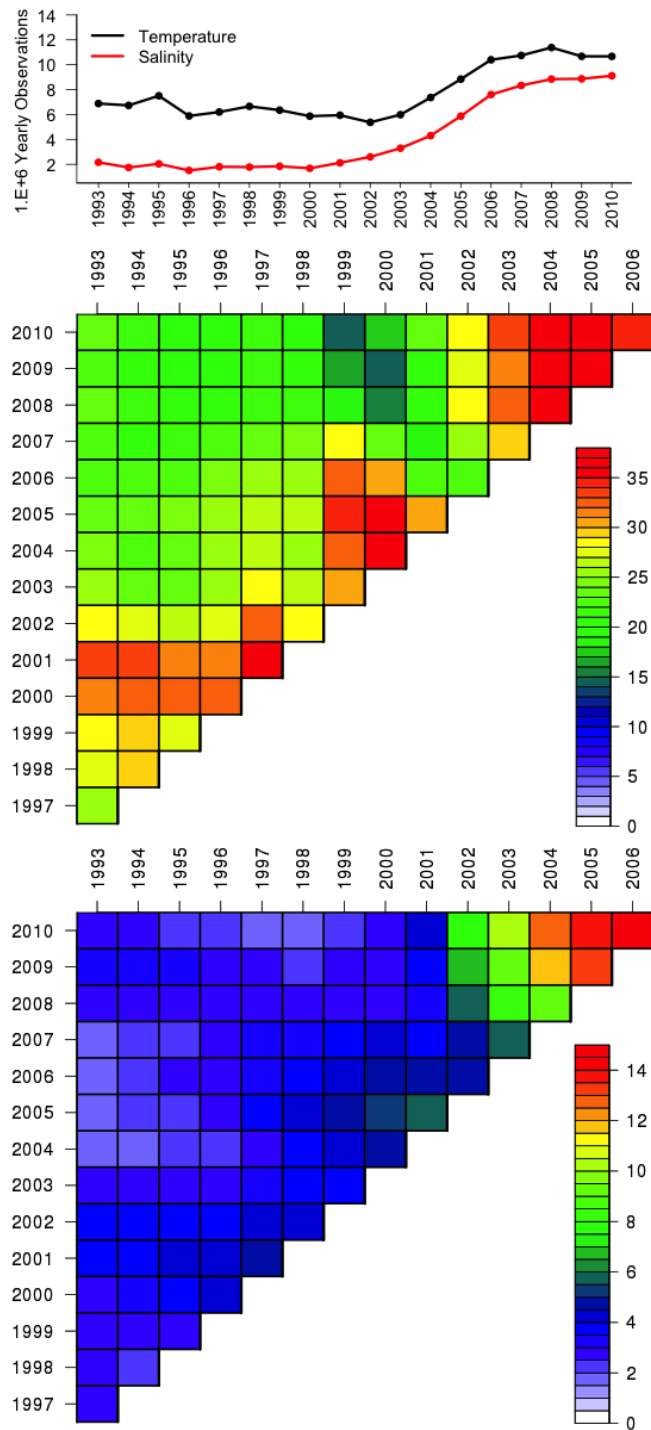


Figure 10: Top: Number of in-situ observations per year from the EN3 dataset (Ingleby and Huddleston, 2007). Units are millions of observations per year. Bottom: Area percentage of the global ocean exhibiting a signal-to-spread ratio greater than the unit for the linear trend as function of starting and ending year, for the thermosteric (middle panel) and the halosteric (bottom panel) sea level. The x-axis above the triangular plots refers to the starting year for the trend computation, while the y-axis refers to the ending year. A minimum period of 5 years is imposed for the linear trend computation. The SSR is computed as ratio between the linear trend ensemble mean and the linear trend ensemble standard deviation, using all the available products, regardless whether they are reanalyses or objective analyses.



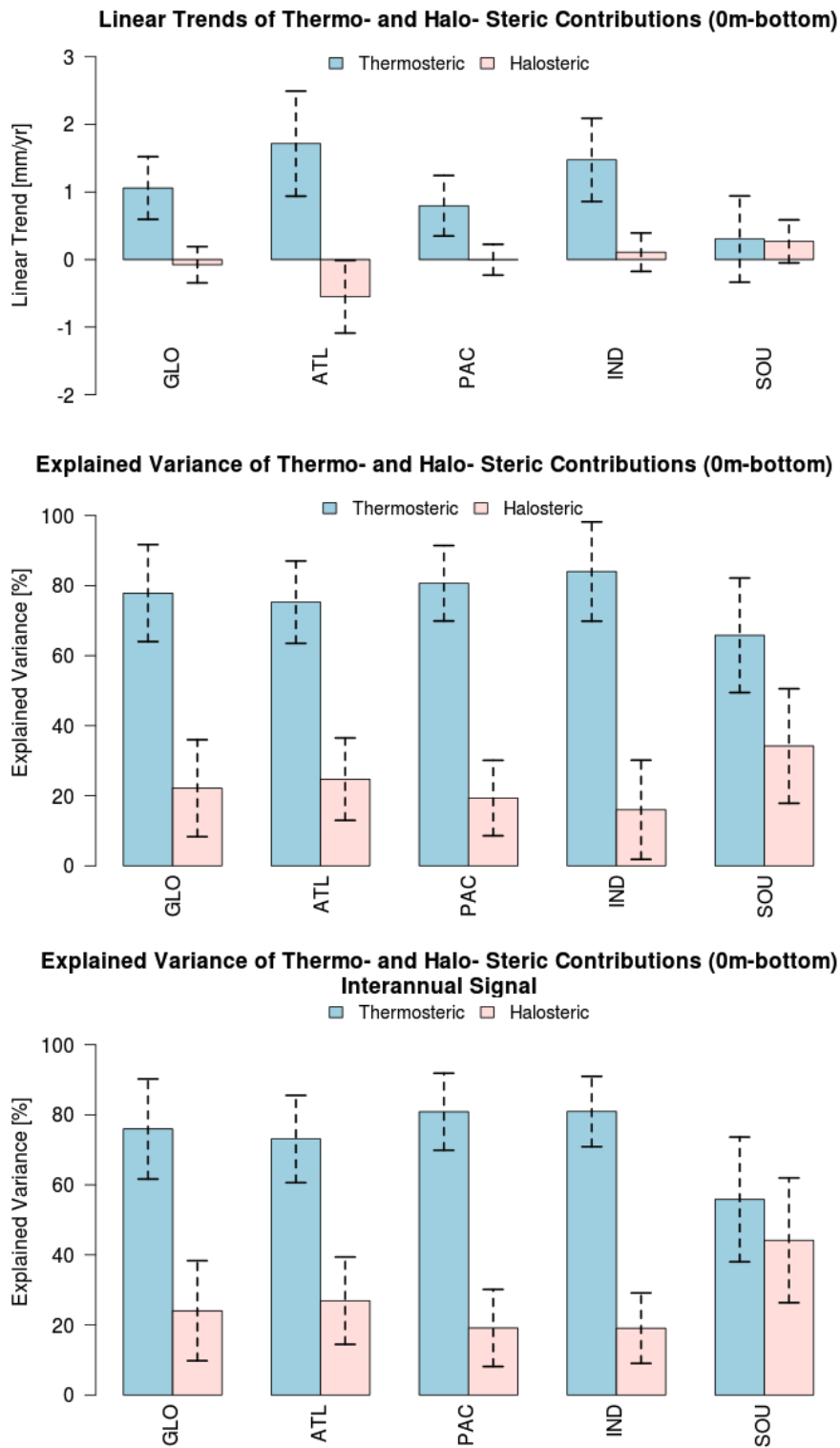


Figure 11: Contribution of the 1993-2010 linear trends from the thermo-steric and halo-steric component (top panel), and explained variance of the full (middle panel) and inter-annual (bottom panel) signal, for the main Ocean basins (GLO: Global Ocean; ATL: Atlantic Ocean; PAC: Pacific Ocean; IND: Indian Ocean; SOU: Southern Ocean). The Southern Ocean is defined as the part of Ocean south of 50S. The box-plot shows the mean and the standard deviation of the values from the different products.

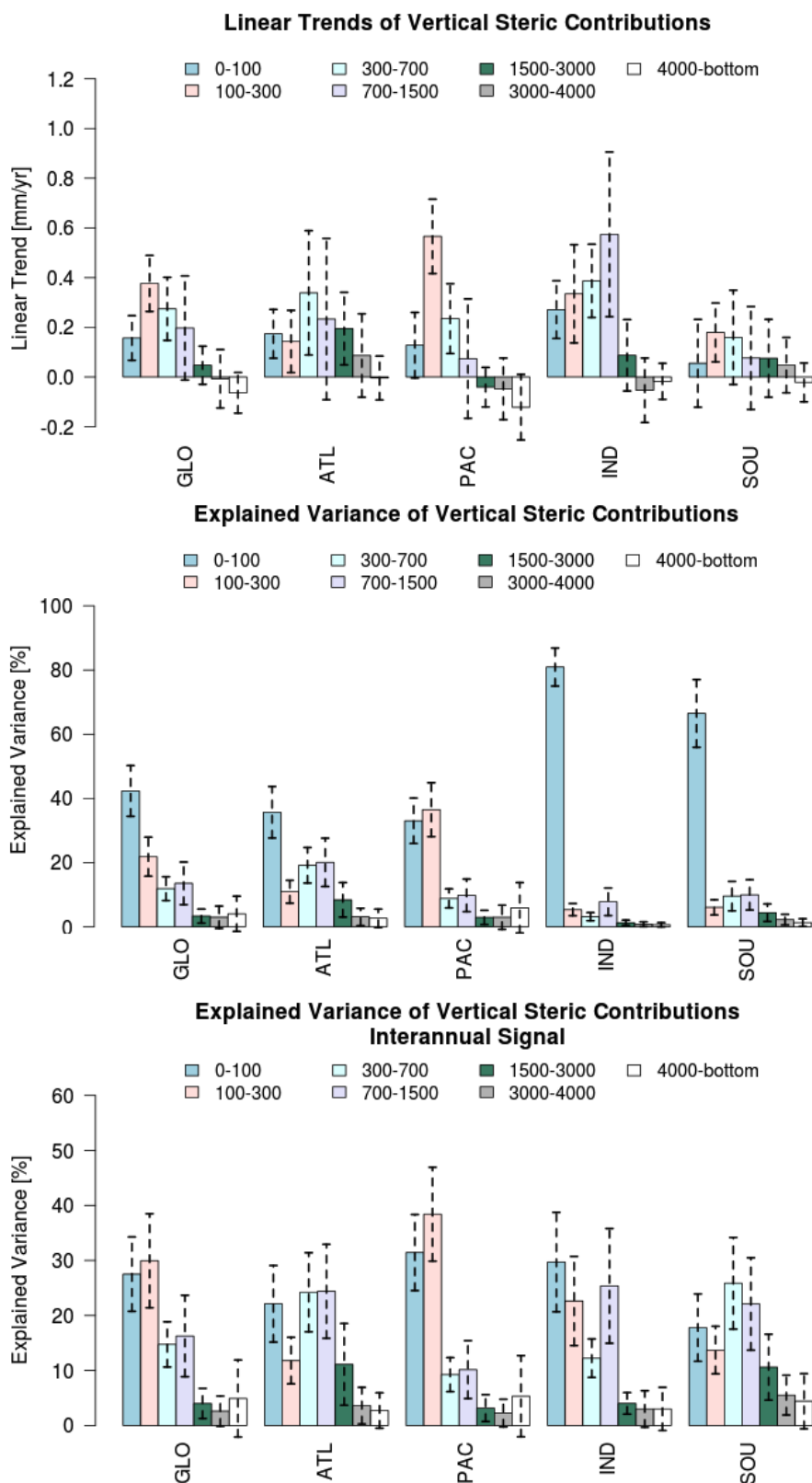


Figure 12: Contribution of the 1993-2010 linear trends from the vertical layers (top panel), and explained variance of the vertical layers for the full (middle panel) and inter-annual (bottom panel) signal, for the main Ocean basins (as in Figure 11).

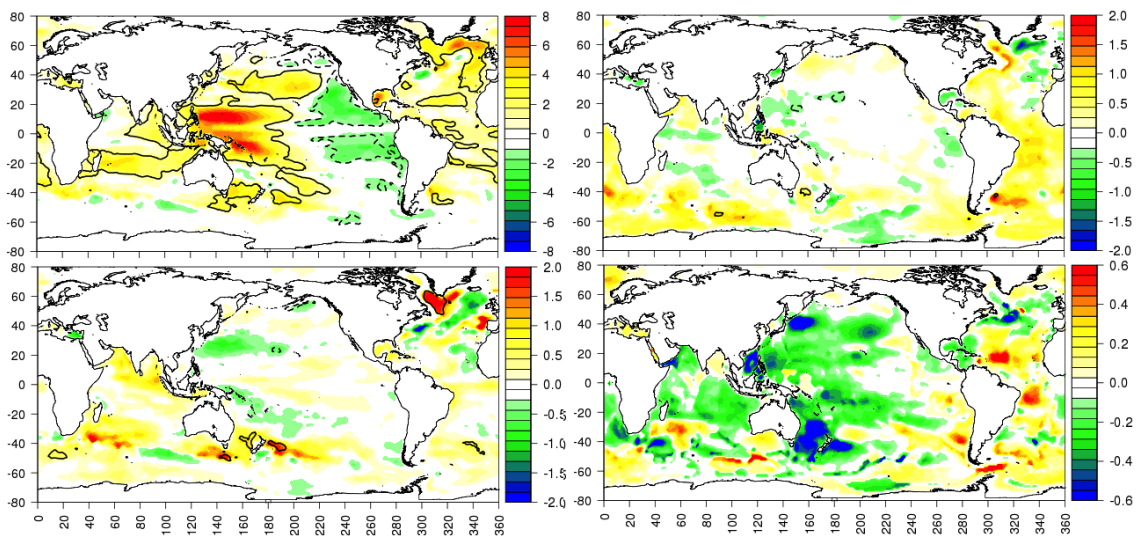


Figure 13: Map of the 1993-2010 ensemble mean of linear trends for the thermosteric sea level in the upper waters (0-700 m of depth, top left panel), intermediate waters (700-1500 m of depth, bottom left panel), deep waters (1500-4000 m of depth, top right panel) and abyssal waters (4000 m to bottom, bottom right panel). Units are  $\text{mm yr}^{-1}$ . Note that the color palettes are different for the four panels. Solid (dashed) contour lines denote regions of signal-to-noise ratio (ensemble mean divided by ensemble spread) equal to 1 (-1).

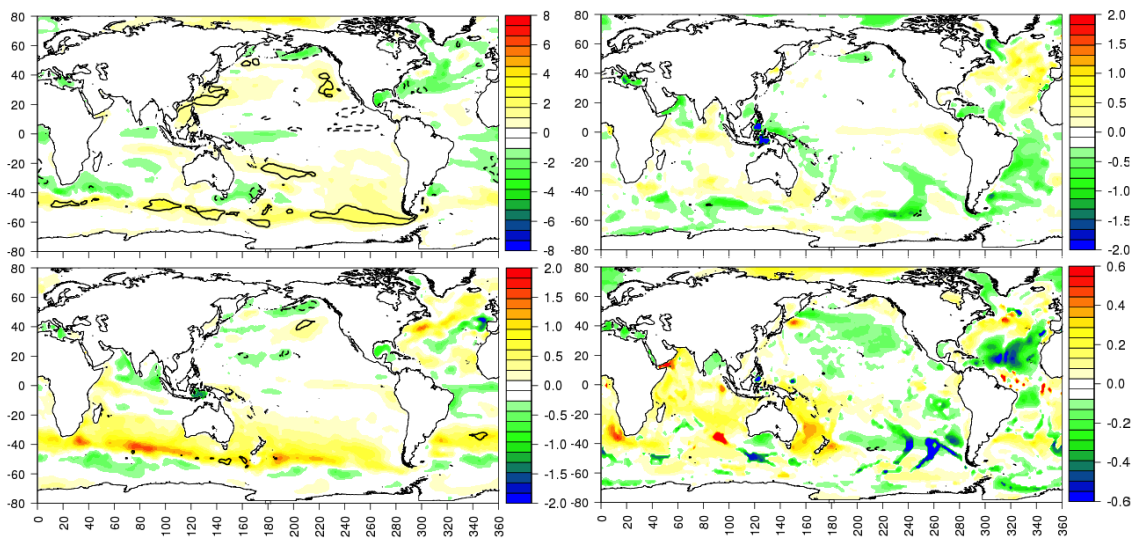


Figure 14: As in Figure 13, but for the haline component of the steric sea level. Note that the color palettes are different for the four panels.