Geophysical Research Letters

Supporting Information for

Robustness of observation-based decadal sea-level variability in the Indo-Pacific Ocean

A.G. Nidheesh[1,3]

, M. Lengaigne[1,2], J. Vialard[1], T. Izumo[1,2], A.S. Unnikrishnan[3], B. Meyssignac[4], B. Hamlington[5], C. de Boyer Montegut[6]

[1]*LOCEAN-IPSL, Sorbonne Univ. (UPMC, Univ Paris 06)-CNRS-IRD-MNHN, Paris, France*

[2]*Indo-French Cell for Water Sciences, IISc-NIO-IITM-IRD Joint International Laboratory, CSIR-NIO, Goa, India*

[3]*CSIR-National Institute of Oceanography, Goa, India*

[4]*LEGOS, CNES, CNRS, IRD, UPS, Université de Toulouse, Toulouse, France*

[5]*Old Dominion University, Department of Ocean, Earth and Atmospheric Sciences, Norfolk, VA, USA*

[6]*Laboratoire d'Océanographie Physique et Spatiale, UBO/Ifremer/CNRS/IRD, Brest, France*

Contents of this file

Additional Supporting Information (Files uploaded separately)

- None

A. Supplementary information for data and method

1. Supplementary information on the datasets used in the current study

Gridded sea-level data examined in this paper consist of three ocean reanalyses, six reconstructions and one thermosteric (in-situ-derived) dataset.

Re-analyses: Re-analyses data are basically derived from ocean general circulation models in which observations are assimilated. These ocean models generally assimilate diverse oceanic data (e.g. temperature, salinity), including sea-level measurements (from altimetry). ORA-S4 (Ocean Reanalysis System 4 from the ECMWF) reanalysis [*Balmaseda et al.,* 2013] is based on the NEMO ocean model forced by ECMWF atmospheric reanalysis data (ERA-40 and ERA-interim). The simulated sea-level is further corrected by assimilating along-track altimeter SLA and global mean sea-level trend estimated from altimeter data since 1993. Similarly, SODA2.2.4 (Simple Ocean Data Assimilation) reanalysis [*Carton and Giese,* 2008] is based on the Parallel Ocean Programme model forced by 20^{th} Century Re-analysis (20CR) wind forcing. It assimilates temperature and salinity data from the World Ocean Data and SST from ICOADS. Finally, we also use GECCO (German ECCO - Consortium for Estimating the Circulation and Climate of the Ocean; hereafter GE) reanalysis [*Kohl and Stammer,* 2008], which is based on the ECCO/MIT ocean model forced by NCEP1 winds and assimilates various hydrographic and satellite data. It must hence be noticed that these reanalyses assimilate different observational data and are forced by different wind products [*McGregor et al.* 2012].

Sea-level reconstructions: The most reliable and directly available sea-level data are provided by tide gauges and satellite altimetry. The modern satellite altimetry offers sea-level measurements with a near-global coverage but only spans about twenty-five years (not sufficient for analyzing sea-level changes at decadal time scales). On the other hand, tide gauge sea-level measurements are mostly confined to coastal oceans, preventing a thorough assessment of open ocean variability, except in the western tropical Pacific where many islands have long-lasting tide gauge measurements (Fig. 1 of *Church et al.* [2004]). The lack of "long and spatially-dense" observational data in the

investigation of long-term sea-level changes hence prompted the scientific community to reconstruct past sea-level by combining data from both satellite altimetry and tide gauges by using sophisticated statistical techniques; that eventually lead to the development of sea-level reconstructions [e.g. *Church et al.,* 2004, *Church and White* 2011; *Meyssignac et al.,* 2012a].

Sea-level reconstructions spatially extrapolate sea-level from long tide-gauge records, available publically from the archives of Permanent Service for Mean Sea-Level (PSMSL) database [*Woodworth and Player* 2003], through a time-varying linear combination of several spatial sea-level fields (or *basis functions*) to "reconstruct" past sea-level *spatio-temporal* variations. Clearly, the quality of the reconstructed sea level for a given region primarily depends on the availability of long tide-gauge record (s) in that region with minimum missing data. A careful selection and edit on all the available records is usually done prior to final use to eliminate those "unqualified" records (see *Church et al.* [2004] for a detailed description of these editing procedure). More than 400 tide gauge records are used in both CW and HA final reconstructions. However, this must be noted that, while the gauge distribution is dense over certain regions (for example in the western Pacific Ocean), the distribution is very sparse over regions like south IO (see Fig. 1 of *Hamlington et al.* [2011]).

In general, the *basis functions* used in reconstructions are a few leading Empirical Orthogonal Functions (EOFs) derived from satellite altimeter record [e.g. *Church et al.,* 2004, *Church and white,* 2011]. The resulting reconstructed sea-level data have the spatial resolution of these basis functions and the record length of tide gauges. EOF-based basis functions (used in CW) are time-independent (i.e. they do not vary in time) and do not contain information on temporal modulation of variability (i.e. modulation of annual cycle and ENSO). Hence, HA reconstruction implemented *cyclostationary* EOFs as basis functions to combine the satellite data with in situ tide gauge measurements [*Hamlington et al.,* 2011]. This technique is expected to reduce the impacts of poor tide gauge sampling and sampling errors in a given location, as the fitting of in situ data involves multiple windows of basis functions (see *Hamlington et al.* [2011] for more details). On the other hand, the use of longer ocean reanalysis to define basis functions (instead of altimetry which is used in CW, HA and M1) is motivated by the fact that the

altimeter record is too short to capture variability over decadal time scales [*Meyssignac et al.,* 2012a]. Hence, we use three more reconstruction datasets whose basis functions are constructed using the three ocean reanalyses discussed in this paper: ORA-S4, SODA and GECCO [*Meyssignac personal communication*], abbreviated M2, M3 and M4 respectively in the paper. These later datasets are an update of the one described in *Meyssignac et al.* [2012a] but with different ocean reanalyses.

WO thermosteric sea-level data. We also use thermosteric sea-level computed from the World Ocean Data [*Levitus et al.,* 2012; hereafter WO], using all available ocean temperature observational data world-wide to build a gridded thermosteric sea-level dataset through optimal interpolation [see *Levitus et al.,* 2012 for details].

Altimeter data. The combined altimeter data from TOPEX/Poseidon, Jason 1 and 2 (TPJ) distributed by the CSIRO [Commonwealth Scientific and Industrial Research Organization, http://www.cmar.csiro.au/sealevel/sl_data_cmar.html] over the 1993-2013 period are also considered in this study, as a point of reference. Since 21 years of altimeter data is not sufficient to draw conclusions on decadal scale variability, we do not consider this data in the ensemble mean analyses presented in this paper, but included for a comparison in the supplementary figures.

A note on sea level in gridded datasets. The OGCMs used in the three reanalyses above conserve volume (under Boussinesq approximation) and do not include a spatially uniform sea-level rise due to thermal expansion or mass addition. Those models do not incorporate addition of mass from continental ice-storage changes and hence also lack ocean mass-induced global mean sea-level rise. This is evident in Global Mean Sea Level (GMSL) time series estimated for these three re-analyses products (Fig. S1), as there is no global mean sea-level change observed. However, ORA-S4 assimilates SLA and GMSL from Altimetry since 1993, and hence includes a GMSL rise from 1993 onward. On the other hand, sea-level datasets from reconstructions are basically an interpolation of "observed sea level" that includes *all* factors that induce a change in sea level. Consequently, the six analyzed reconstructions exhibit GMSL rise over the entire period (1960-2010), that is consistent with the trend observed in satellite altimetry after 1993 (Fig. S1).

WO thermosteric sea level reflects the ocean heat content variations over the upper 2000m layer but lacks sea-level changes associated with deep ocean warming, halosteric sea level changes and *ocean mass* variations. For decadal time scales discussed in this paper, the tropical ocean sea-level changes are mainly driven by wind-driven thermocline variability, i.e. steric sea-level variations are close to *actual sea-level* changes. *Nidheesh et al.* [2013] showed that the decadal steric sea-level variability in the tropical Indo-Pacific Ocean is mostly driven by ocean thermal variations (thermosteric changes). WO thermosteric sea- level is hence suitable to our objectives. We also remove GMSL time series from each dataset before all the analyses in this paper in order to focus on regional decadal sea-level variability.

2. Mathematical formulation of the "agreement ratio" between products

For each grid-point in space, we compute a metric named *agreement ratio*, which is basically derived from the ensemble variability over ten products as shown below. At a given grid-point, we note $P_{(n,\,t)}$ the value of the product n at time t, where, n varies from 1 to N (total number of products) and t varies from 1 to T (total number of time steps)

The ensemble average at a given time t is computed as:

$$Ens_t = \frac{1}{N} \sum_n P_{n,t}$$

The temporal standard deviation of this ensemble average is an indicator of the mean amplitude of decadal variability across products. It is computed as:

$$\sigma = \sqrt{\frac{1}{T-1} \sum_t [Ens_t - \overline{Ens}]^2}$$

where

$$\overline{Ens} = \frac{1}{T} \sum_t Ens_t$$

Finally, the spread (*s*) in variability among products over the entire period is estimated as the square root of the mean difference (i.e. averaged over products and time) between the product-ensemble and individual variability:

$$s = \sqrt{\frac{1}{NT-1} \sum_t dif_t}$$

With the *summed difference* between ensemble variability and variability among individual products (dif_t) is computed as:

$$dif_t = \sum_n [P_{n,t} - Ens_t]^2$$

The *agreement ratio* is computed as spread (*s*) divided by the mean amplitude of decadal variability across products (σ), i.e

$$Agreement\ ratio = s / \sigma$$

Hence a value of agreement ratio below 1 indicates that the spread between products is smaller than the amplitude of ensemble variability in those products.

B. Supplementary figures

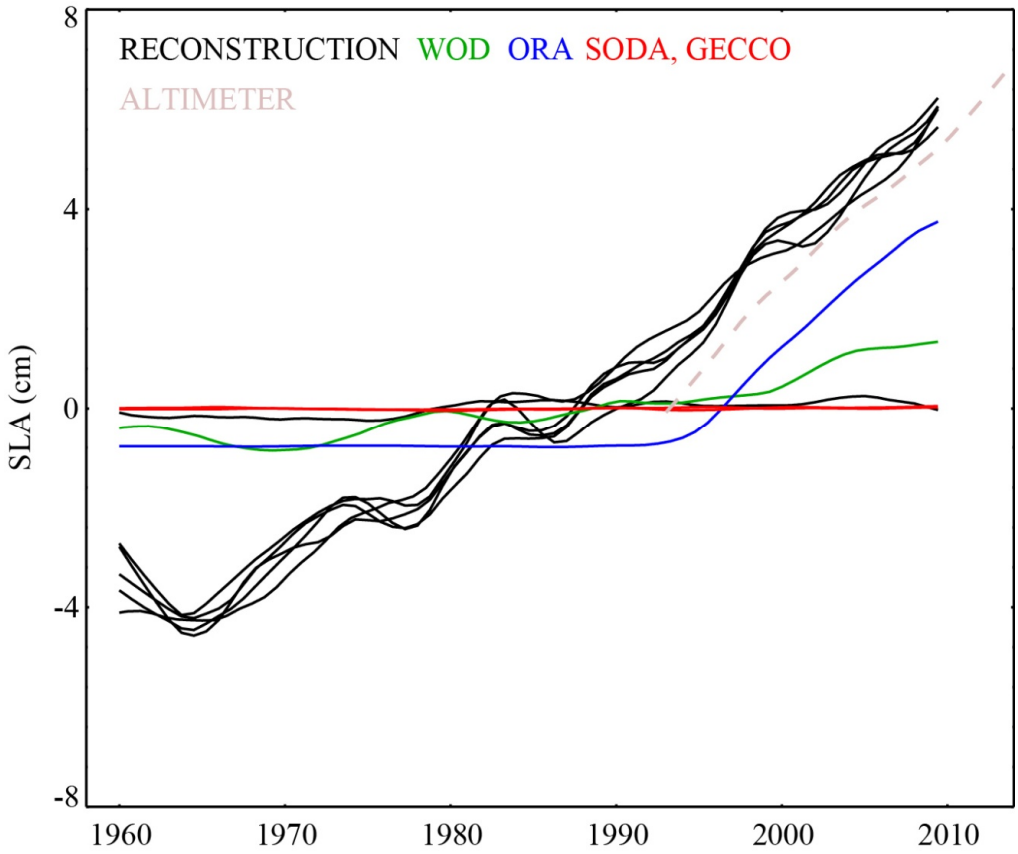

Figure S1: Global mean sea-level (GMSL) time series from 11 gridded products (including Altimetry – TP/J) analysed in this study: WO (green), OR (blue), SO & GE (red), HA, CW, M1, M2, M3, M4 (black) and TP/J (dashed grey). This GMSL is obtained as the spatial average over the 65°N-65°S latitudinal band of the decadal sea-level component of each gridded product.

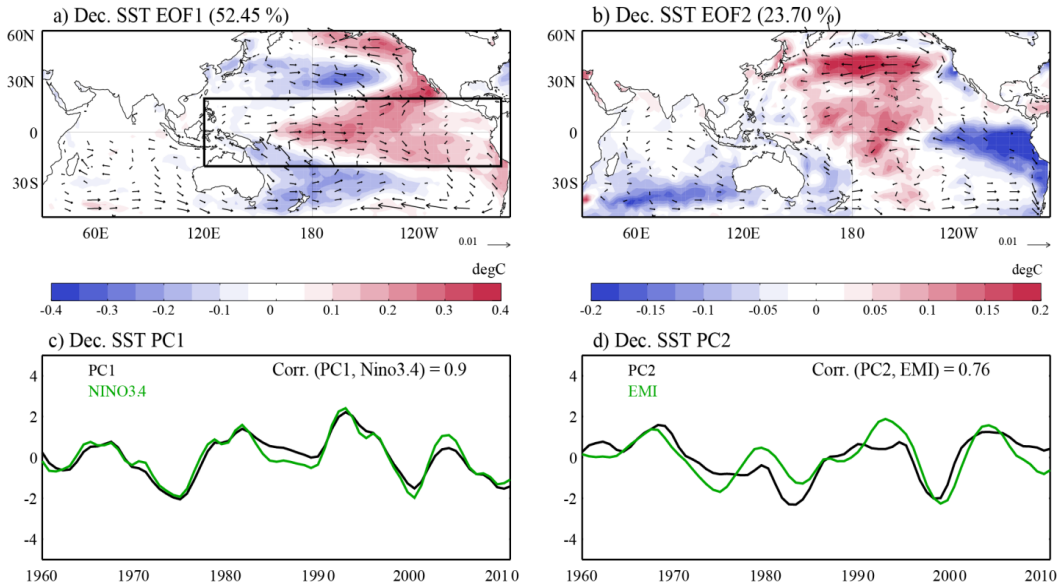

Figure S2: First two EOFs of decadal SST variability in the tropical Pacific (120°E-70°W, 20°S-20°N; black rectangle on panel a) over the 1960-2010 period. (a, b) SST (color) and wind-stress (arrow) patterns associated with EOF1 and EOF2. (c, d) Corresponding principal components (PCs). Decadal components of Niño3.4 SST index and EMI (El Nino Modoki Index; Ashok et al. 2007) are overlaid on panels c and d respectively with corresponding correlation coefficient. Even though the EOFs are computed for the tropical Pacific, the SST and wind are regressed onto the corresponding normalized PCs over the entire Indo-Pacific region to assess their Indo-Pacific signature. PCs shown in panels c and d are used in this study as indices for decadal ENSO and Modoki variability respectively.

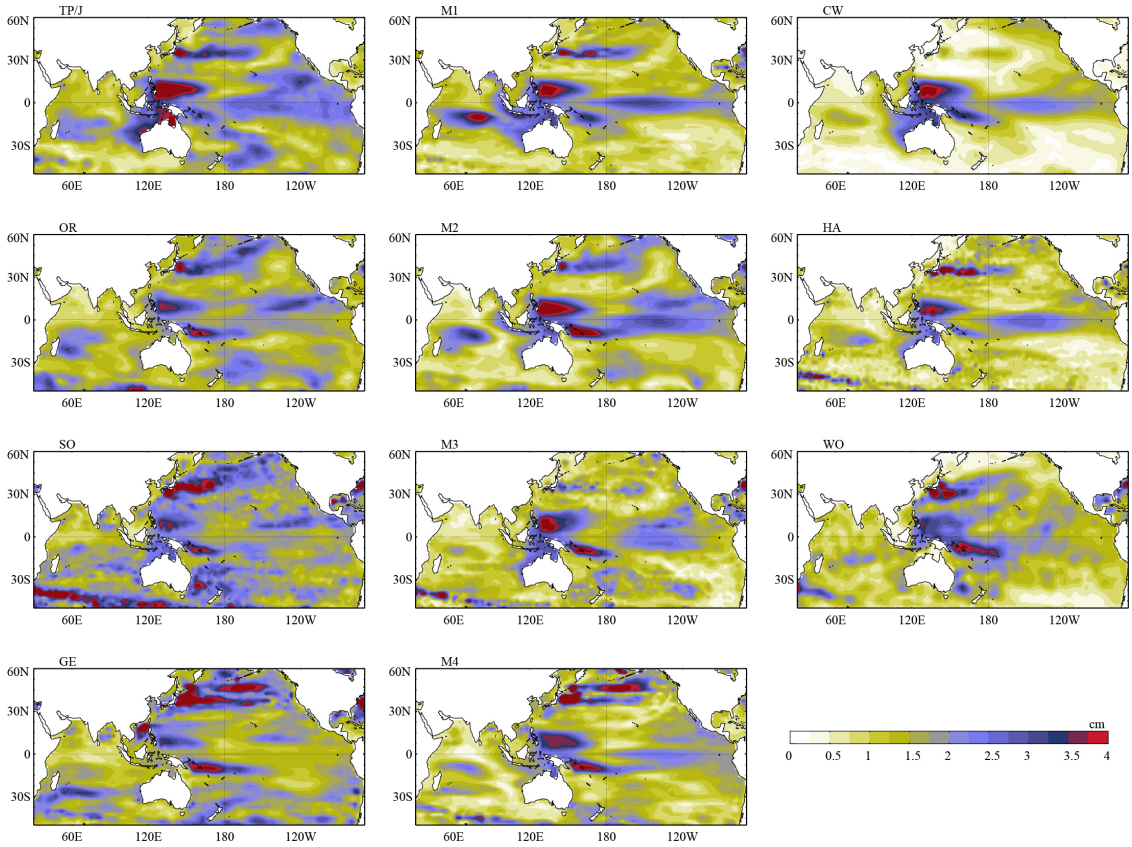

Figure S3: Standard deviation of decadal SLA from the 10 gridded products analyzed in this study and for the shorter altimeter data (TP/J). The M1, CW & HA reconstructions use basis functions from TP/J data. The M2, M3 and M4 datasets respectively use basis functions from OR, SO and GE, and are displayed side-by-side to facilitate comparison between each reconstruction and the reanalysis on which it is based.

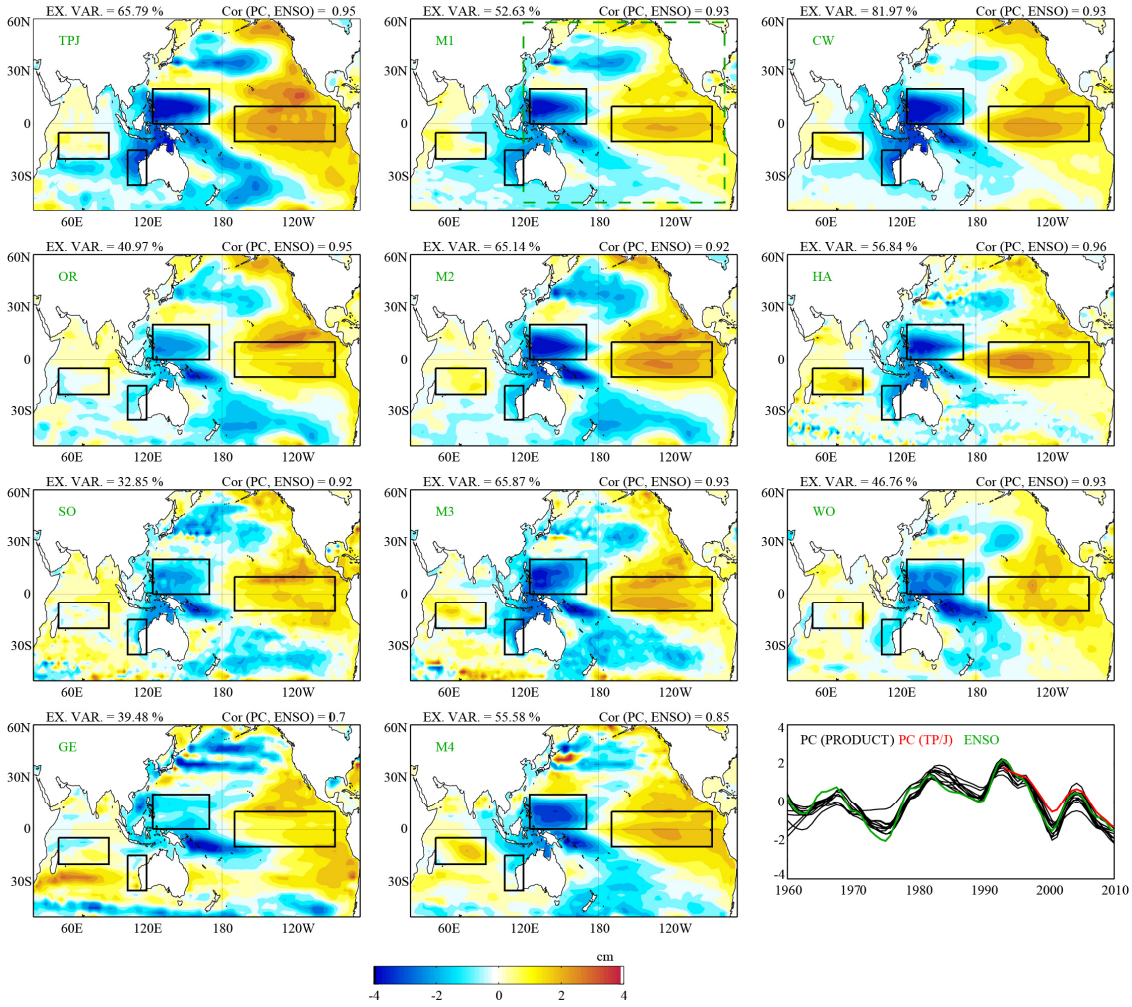

Figure S4: Indo-Pacific SLA signature of Pacific decadal SLA EOF1 (spatial maps) from 10 sea-level data used in this study and for the shorter merged altimeter dataset. PC1 from 10 products (black) and TP/J (red) are shown in the last panel with decadal ENSO time series (green). The M1, CW & HA reconstructions all use basis functions from the TP/J dataset. The M2, M3 and M4 datasets respectively use basis functions from OR, SO and GE, and are displayed side-by-side to facilitate comparison between each reconstruction and the reanalysis on which it is based.

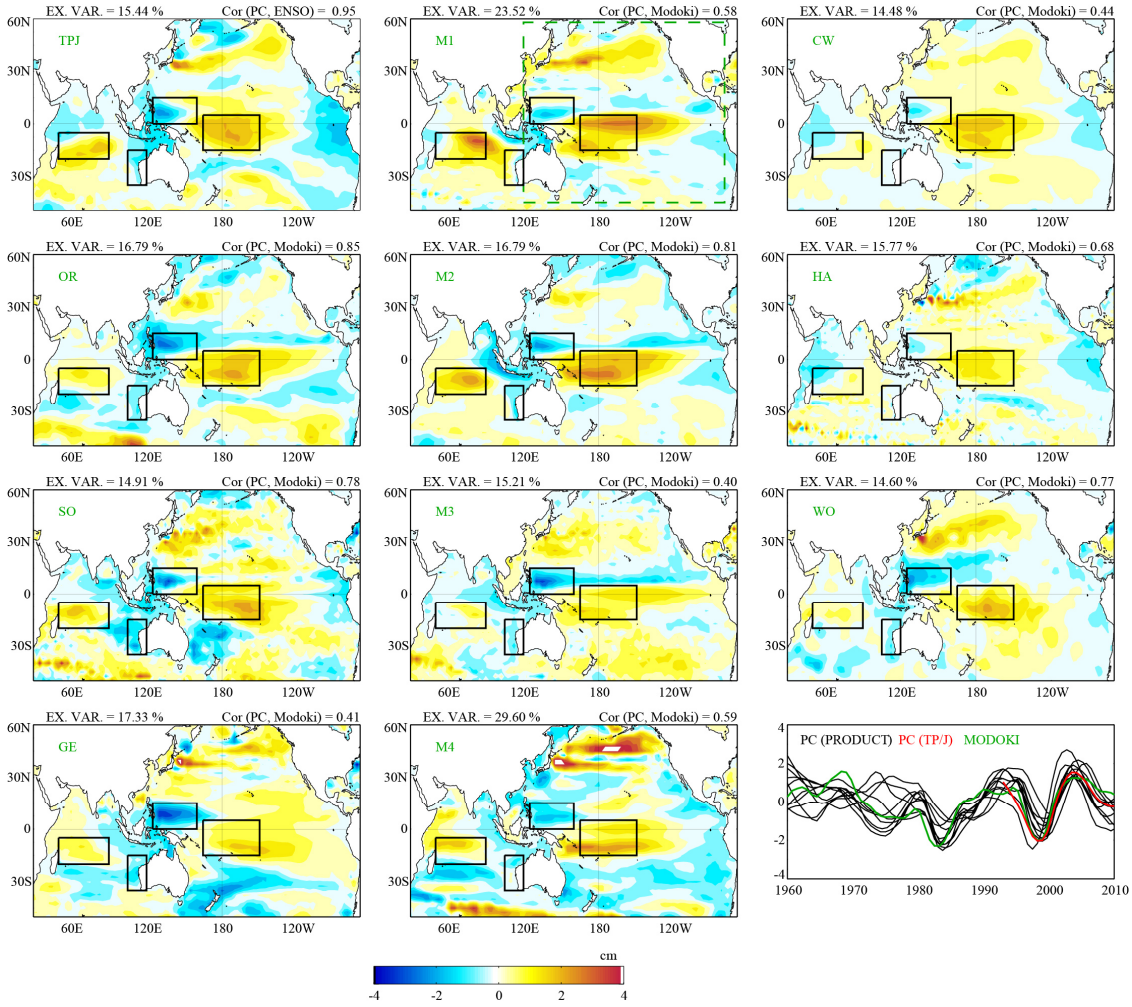

Figure S5: As Figure S4, but for Indo-Pacific SLA signature of Pacific decadal SLA EOF2. In the last panel, decadal MODOKI time series (green) is shown with sea-level PCs. The M1, CW & HA reconstructions all use basis functions from the TP/J dataset. The M2, M3 and M4 datasets respectively use basis functions from OR, SO and GE, and are displayed side-by-side to facilitate comparison between each reconstruction and the reanalysis on which it is based.

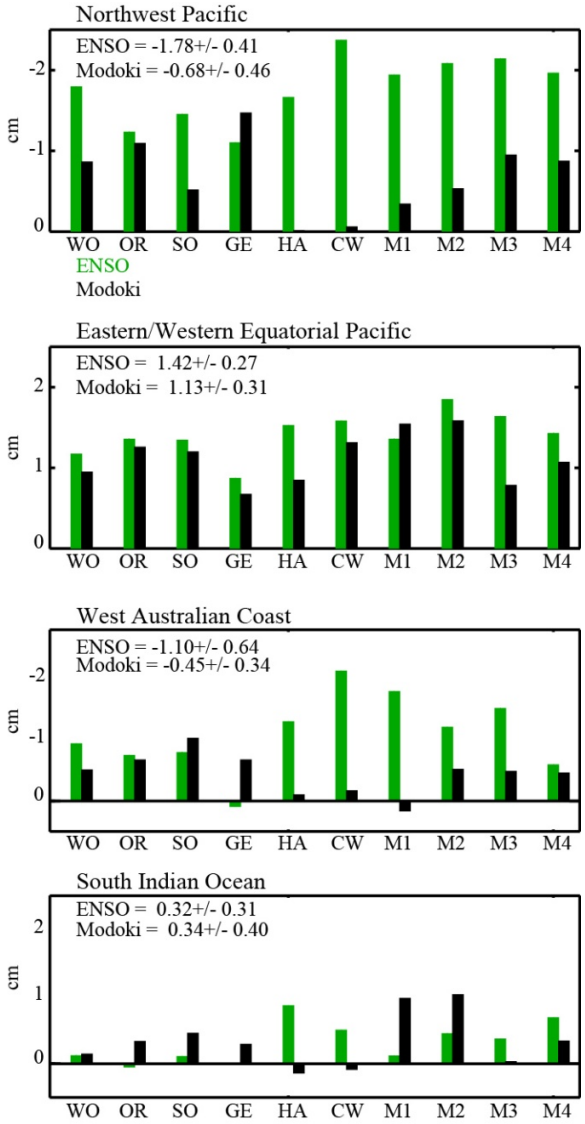

Figure S6: Decadal ENSO- (green) and Modoki- (black) related SLA amplitude at four regions shown in Figures S4 and S5 (four boxes), computed as the box-averaged value of EOF from each product. The mean amplitude and it's inter-product standard deviation are also given on each panel for both ENSO and Modoki. Note that in S4 and S5, the regions selected to highlight equatorial Pacific variability for ENSO and Modoki are different (eastern equatorial Pacific for ENSO and western equatorial Pacific for Modoki) based on their maximum centers of action.

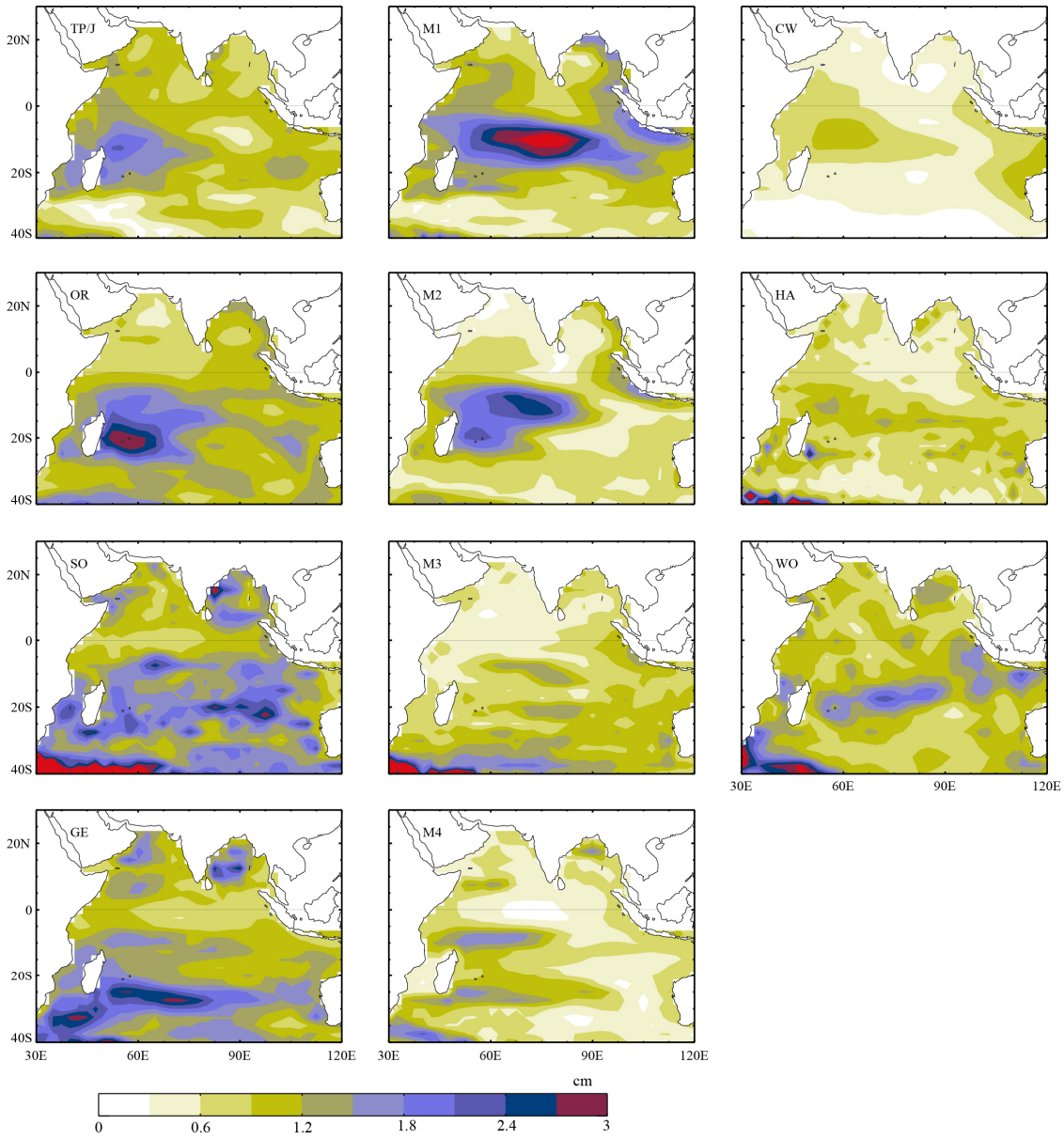

Figure S7: Standard deviation of Pacific-independent decadal SLA in the IO from individual products. The M1, CW & HA reconstructions all use basis functions from the TP/J data. The M2, M3 and M4 datasets respectively use basis functions from OR, SO and GE, and are displayed side-by-side to facilitate comparison between each reconstruction and the reanalysis on which it is based. Note that the result from altimeter displayed for information, but that the dataset is probably too short for performing this analysis.

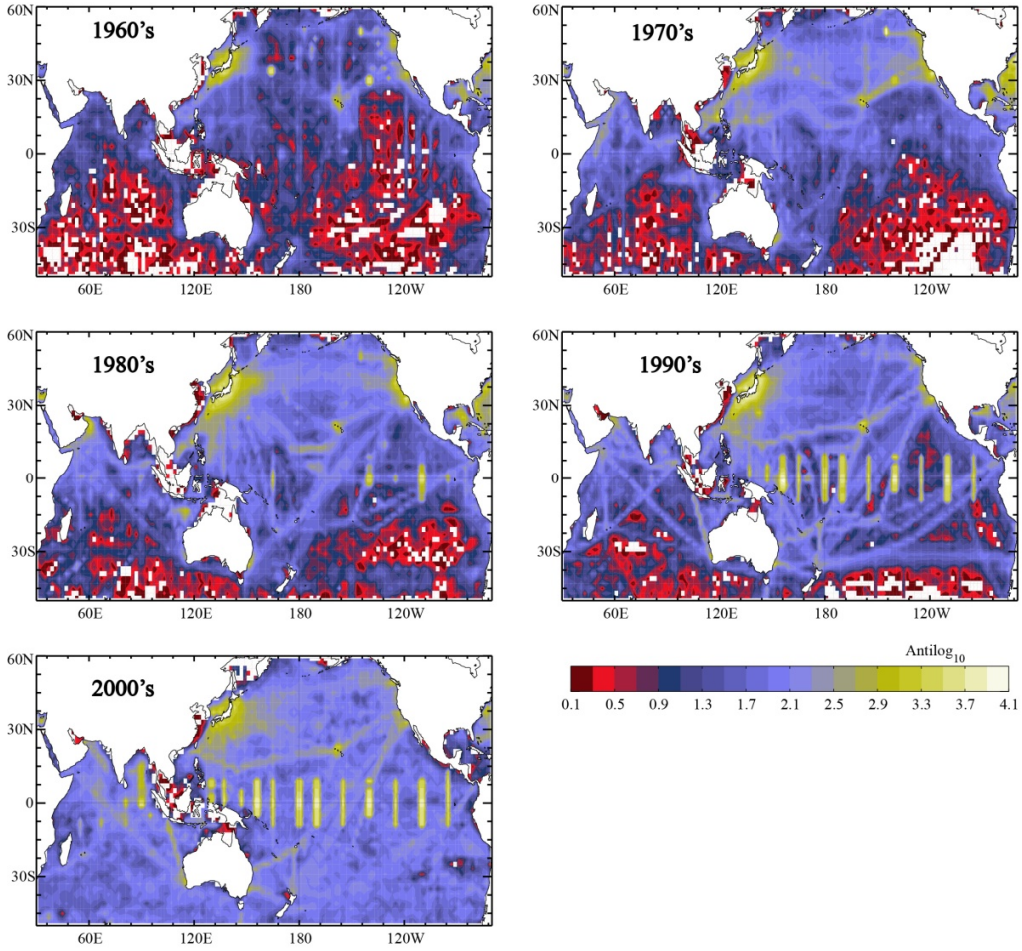

Figure S8: Decimal logarithm of the number of profiles per 2° x 2° box and per decade, with enough levels to estimate the heat content of the top 300 meters of the Indo-Pacific Ocean (i.e. that resolve vertical movements of the thermocline associated with natural decadal climate variability), during 1960-2010.