JGR Oceans

RESEARCH ARTICLE
10.1029/2020JC016270

20th Century Multivariate Indian Ocean Regional Sea Level Reconstruction

Praveen Kumar[1], Benjamin Hamlington[2], Se-Hyeon Cheon[2], Weiqing Han[3], and Philip Thompson[4]

[1]Center for Coastal and Physical Oceanography, Old Dominion University, Norfolk, VA, USA, [2]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA, [3]Department of Atmospheric and Oceanic Sciences, University of Colorado, Boulder, CO, USA, [4]Department of Oceanography, University of Hawai'i at Mānoa, Honolulu, HI, USA

Key Points:
- We have developed a new multivariate sea level reconstruction technique designed to overcome sparse spatio-temporal tide gauge sampling
- Created a new sea level data set with improved 20th century interannual to decadal sea level variability estimates for the Indian Ocean
- This sea level reconstruction is used to give context to satellite measured sea level in the Indian Ocean

Supporting Information:
- Supporting Information S1

Correspondence to:
P. Kumar,
pkuma002@odu.edu

Citation:
Kumar, P., Hamlington, B., Cheon, S.-H., Han, W., & Thompson, P. (2020). 20th century multivariate Indian Ocean regional sea level reconstruction. *Journal of Geophysical Research: Oceans*, *125*, e2020JC016270. https://doi.org/10.1029/2020JC016270

Received 23 MAR 2020
Accepted 21 SEP 2020
Accepted article online 22 SEP 2020

Abstract Despite having some of the world's most densely populated and vulnerable coastlines, Indian Ocean sea level variability over the past century is poorly understood relative to other ocean basins primarily, due to the short and sparse observational records. In an attempt to overcome the limitations imposed by the lack of adequate observations, we have produced a 20th century Indian Ocean sea level reconstruction product using a new multivariate reconstruction technique. This technique uses sea level pressure and sea surface temperature in addition to sea level data to help constrain basin-wide sea level variability by (1) the removal of large spurious signals caused as a result of insufficient tide gauge data specifically during the first half of the 20th century and (2) through its information on large-scale climate modes such as El Niño-Southern Oscillation and Indian Ocean Dipole. Basis functions generated by Cyclostationary Empirical Orthogonal Functions are used for the reconstruction. This new multivariate technique provides improved regional sea level variability estimates along with a longer record length in comparison to existing globally reconstructed sea level data. The biggest advantage of using this multivariate reconstruction technique lies in its ability to reconstruct Indian Ocean sea level for the first half of the 20th century, providing a long sea level record for the study of Indian Ocean internal climate variability. This will enable future studies to help improve the understanding of how sea level trends and variability can be modulated by internal climate variability in the Indian Ocean.

Plain Language Summary Densely populated coastal regions of the countries surrounding the Indian Ocean are becoming increasingly vulnerable to the effects of sea level rise. Accurate predictions of future sea level change will help minimize the social and economic damage posed to these coastal communities. The key to predicting future sea level lies in how well we understand past and present sea level change. However, due to the presence of only a few short and scattered sea level observational records, change in sea level over the Indian Ocean is not well understood. In this paper we have created a new sea level data set by developing new methodology which is particularly well suited for filling data gaps left by the observational record in the Indian Ocean. Additionally, this new technique allows us to reliably extend the observational record to now span the entire 20th century and has been shown to have improved sea level estimates when compared to an older existing sea level data product. Using this new longer sea level data set, we hope that future studies will help shed more light and improve the understanding of Indian Ocean sea level change over long time scales.

1. Introduction

The rise in global mean sea level (GMSL) associated with melting of land ice and thermal expansion of seawater over the last century threatens coastal infrastructure and ecosystems. For many coastal and island populations the effects of sea level (SL) rise are already being experienced (Church & White, 2011). Despite having some of the most densely populated and vulnerable coastal regions in the world (Swapna et al., 2020), along with considerable evidence that the impact of Indian Ocean (IO) variability on climate extends beyond the basin to global scales, IO SL has received considerably less attention compared to that of the Pacific Ocean (Han, Vialard, et al., 2014). One reason for this is the absence of a network of high-quality long tide gauge (TG) records, limiting the study of IO SL to use only short observational data (e.g., Lee & McPhaden, 2008) or model-based approaches (e.g., Han et al., 2010; Han, Meehl, et al., 2014; Nidheesh et al., 2013; Trenary & Han, 2013). With the TG stations sparsely distributed around the IO

coasts and islands, the IO basin interior is left with no long-term observational records. IO SL trend estimates made using the sparse TG record are significantly impacted by the influence of regional and local variability, such as local land movement (Hamlington & Thompson, 2015).

With the advent of TOPEX/Poseidon satellite data in October 1992, satellite altimetry has ushered in a new era of monitoring the GMSL over the ice-free oceans, providing sea surface height measurements that are continuous and near-global. Unlike TG data, the satellite altimeter derived data allow for basin-wide SL measurements with unprecedented accuracy and stability (Nerem et al., 2010). The short record of satellite altimetry, however, makes it challenging to distinguish the longer-term variability from secular trend. Yet separating and removing internal variability from the SL record is essential for estimating long-term trends (e.g., Cazenave et al., 2014; Hamlington et al., 2013, 2014). Therefore, long SL records are needed to achieve more accurate depictions of SL induced by internal climate variability and to estimate longer-term trend in the Indo-Pacific region (Frankcombe et al., 2015).

In order to better understand the uncertainties present in the estimates of decadal variability due to the limitations imposed by sparse spatial coverage of the TG stations and the relatively short temporal length of the satellite altimetry record, researchers have developed a method that combines the long time-series data provided by TG stations with the information about the spatial covariance of SL from satellite altimetry (and sometimes models) to reconstruct historic SL fields with global coverage (Chambers et al., 2002; Church et al., 2004; Church & White, 2006, 2011; Hamlington, Leben, Nerem, Han et al., 2011; Hamlington, Leben, & Kim, 2012; Hamlington, Leben, Wright, et al., 2012; Hay et al., 2015; Meyssignac et al., 2012; Ray & Douglas, 2011). SL reconstructions fill in the unsampled regions in the sparse historic network with SL information derived from spatial patterns during time periods when the observational network is dense. Although originally developed for the use with sea surface temperature (SST) (e.g., Kaplan et al., 1998, 2000; Smith et al., 1996), the method of empirical orthogonal function (EOF) reconstruction has since then been modified for SL studies by employing the weighted least squares fits of satellite altimetry-derived EOF to TG data. Following an EOF decomposition, a data record is split into a set of individual modes, where in each mode consists of a single temporal evolution and a single spatial structure. The work of Hamlington, Leben, Nerem, Han, and Kim (2011) highlights the advantages of using basis functions derived from cyclostationary empirical orthogonal function (CSEOF) over EOF derived basis functions for SL reconstructions. Unlike an EOF mode that consists of only one spatial pattern, making it strictly time-independent (stationary), a single CSEOF mode has a time-dependent variability associated with its spatial structure which allows it to have multiple spatial structures (for each mode) that are constrained to be periodic with a selected nested period. As an example, consider the SL annual cycle that has a physical oscillation with a 1-year period and changes in strength from year to year. The CSEOF spatial structures capture the physical oscillation and effectively separates it from its amplitude modulation which is in turn represented by the CSEOF temporal evolution. In this way the annual cycle can be captured using a single CSEOF mode thus allowing us to retain the seasonal cycle before doing the CSEOF decomposition. This cyclostationary nature of CSEOF spatial patterns is a significant advantage over an EOF spatial pattern when performing SL reconstructions, as the CSEOF basis functions allow for a better representation of the longer-time-scale fluctuations and time-varying spatial patterns present in geophysical signals (Hamlington et al., 2011).

While data products based on the SL reconstruction technique discussed above are a very good resource and are commonly employed for the study of SL variability, data products based on ocean reanalysis (e.g., Carton & Giese, 2008) and thermosteric SL data based on in situ observations (e.g., Levitus et al., 2012) are other sources of data that can be used for the study of decadal SL variability. Using different methods, researchers have developed a number of SL products to help better understand and quantify SL variability but in doing so are now confronted with a major problem of choosing one product from all the available SL products generated. In an effort to test the interproduct SL consistency and coherence of the decadal variability displayed by the available SL products for the Indo-Pacific region, Nidheesh et al. (2017) compared 10 gridded multidecadal SL products. Of these, six had been generated by the method of SL reconstruction, three were reanalysis products, and one was based on in situ data, all with a temporal length spanning their overlapping period from 1960 to 2010. This study reported that although the data show an interproduct consistency in the Pacific Ocean, there exists a significant disagreement in the IO, further underscoring the importance of improving SL estimates in the IO.

For the method of SL reconstruction, one way to improve regional SL estimates is to use basis functions that are computed regionally instead of globally, which are often limited in their ability to capture regional variability and trends. Accurate estimates of SL variability at a regional scale are vital to policymakers in order to make informed decisions about the present and future SL. The importance of regionally reconstructing SL data is evident from the study of Hamlington, Leben, and Kim (2012) that noted a significant discrepancy in the spatial variability of the secular trends computed for the Pacific Ocean from two reconstructions: one for global and the other for the Pacific Ocean. In the IO, to our knowledge, there are no previous attempts to regionally reconstruct 20th century SL, and this is a primary motivation for the current work. Due to the limited temporal coverage of TG data in the IO, a reconstruction technique that relies on using only historic TG data (e.g., Hamlington, Leben, Nerem, Han, & Kim, 2011) is inadequate for reconstructing 20th century SL, particularly during the first half of the 20th century when TG data are lacking. To overcome this limitation, Hamlington, Leben, Wright, and Kim (2012) introduced a reconstruction technique that used SST along with SL anomaly (SLA). This reconstruction relied on a multivariate linear regression technique to extract relationships and establish links between SLA and SST. Climate variables in the Earth climate system are often connected to each other (e.g., an increase in SST will most likely be associated with an increase in SL). Climate variables can also be influenced by a common physical process and large-scale climate modes (e.g., El Niño-Southern Oscillation). The inclusion of SST into the reconstruction helps to improve the reconstructed SL estimates by constraining large-scale climate modes that have both SL and SST signatures. In addition to this SST also adds information about variability over the open ocean (over which TG data is very sparse).

The climate modes of El Niño-Southern Oscillation (ENSO) and Indian Ocean Dipole (IOD) are known to have a significant influence on the IO SL, SST, and SLP (Murtugudde et al., 2000; Saji et al., 1999; Webster et al., 1999). Both these climate modes impact large areas over the western tropical IO and eastern IO off the coast of Java/Sumatra. The IO SST shows a basin wide warming pattern referred to as the IO Basin Mode (e.g., Klein et al., 1999) as a direct response to ENSO. Additionally, during IOD and ENSO events, cooler (warmer) than average SST anomalies appear off the coast of Java/Sumatra (over large parts of the western tropical IO). The cooling off the coast of Java/Sumatra suppresses convection over the eastern tropical IO and enhances convection over the western IO basin altering SLP patterns over the IO. From this we can infer that the addition of SST and SLP into the reconstruction will help to better constrain large-scale climate modes in addition to providing information about variability over the open ocean and yield better estimates of IO SL variability. Building on the reconstruction methodology of Hamlington, Leben, Wright, and Kim (2012), we propose a new multivariate reconstruction (MV-R) technique that allows for the inclusion of multiple climate variables to help improve the reconstruction of SL that reflect coupled climate variability. For the current study, atmospheric sea level pressure (SLP) will be included as the third variable along with SST and SLA, which is shown to improve 20th century SLA variability estimates for the dynamically complex IO.

In this paper we summarize the results obtained using the MV-R technique to create a new regionally reconstructed SL product for the IO spanning the time period of 1900–2018. Section 2 of this paper provides a detailed discussion outlining the data used and a short summary of the CSEOF MV-R technique. A brief explanation of the procedure used to create the final reconstructed SLA data has also been included. A detailed description of the results obtained from comparisons made between the MV-R SLA with satellite altimetry and an existing global reconstruction by Hamlington et al. (2014) is presented in section 3. This is followed by results obtained from examining the ability of the reconstruction to recreate climate modes over the tropical IO. Section 4 provides a summary of the results and describes the benefit of using this newly proposed MV-R technique to reconstruct 20th century SLA in the IO.

2. Data and Methods

The work of Hamlington, Leben, and Kim (2012) and more recently Cheon et al. (2018) show that regionally computed basis functions are capable of capturing and representing regional SL variability more accurately compared to globally computed basis functions. We follow a similar approach and perform the MV-R over the IO region, defined as the region bound by 40°S to 30°N and 20°E to 120°E. Below is a brief description of the data used by the MV-R technique.

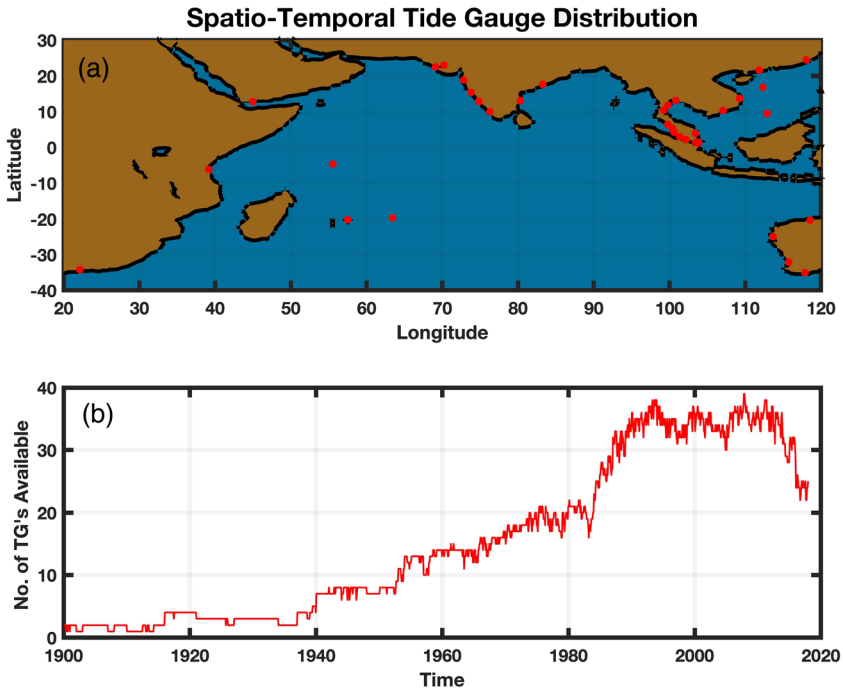

Figure 1. (a) Spatial location of tide gauge stations used for the multivariate reconstruction (MV-R). (b) Number of tide gauge stations available over time for use in the MV-R.

2.1. Data

2.1.1. Tide Gauge Data

Monthly mean TG data archived at the Permanent Service for Mean Sea Level (PSMSL; Woodworth & Player, 2003) subsampled for the IO were used as a source of in situ historical SL data. We only used the Revised Local Reference (RLR) TG data set as it makes use of a constant local datum over the full record of the SL measurement at each TG location. "Metric" data provided by PSMSL were not used for this study as it is often subjected to unknown datum shifts, which makes the SL data unreliable for time series analysis. With the multivariate nature of this reconstruction technique and a limited spatio-temporal coverage of TG records in the IO, we have opted to keep the selection and quality control criteria for TG data lenient. All the TG editing criteria used for this study are similar to those described in Church et al. (2004) and yields a similar TG set for the IO. Using rates derived from the ICE-5G VM2 model (Peltier, 2004) the TG records were corrected for the ongoing glacial isostatic adjustment (GIA). There has been no effort made to avoid regional clustering of TG stations. The robustness of this reconstruction technique toward data outliers allows us to avoid selectively editing TG records. However, we note that it will be appropriate for future studies to implement a more careful TG selection and editing procedure. The location and number of TG stations available through time used for the MV-R over the IO are shown in Figure 1. The list of TG stations used has been provided in the supporting information Figure S1.

To test the sensitivity of the MV-R to an individual TG record a TG cross-validation analysis was done (results not shown here). For this, the IO SLA was reconstructed leaving out a single TG resulting in 45 additional reconstructions referred to as MV-Rcv. The IO SL variability estimates obtained from MV-Rcv were compared to MV-R, and the results showed no significant differences between the estimates of SL variability. The MV-R and MV-Rcv SLA were then subsampled at the individual TG location which was left out and compared to the SLA of the respective TG that was left out. The correlation coefficient obtained from correlating subsampled MV-R and TG SLA was only marginally greater that the correlation coefficient of MV-Rcv versus TG SLA. These results suggest that the MV-R is not overly sensitive to a particular TG station. Similar results were obtained when using reconstruction created by leaving out all the TGs off the Australian coast (MV-RcvAus) and reconstructions created by leaving out all the South China Sea TGs (MV-RcvChina). Estimates of SL variability between the MV-RcvAus, MV-RcvChina, and MV-R were found to be very similar to each other.

2.1.2. Satellite Altimetry Data

Monthly mean SLA data used for generating SL basis functions required for the reconstruction were obtained from Archiving, Validation, and Interpretation of Satellite Oceanographic (AVISO) altimetry. This data set was produced by the multimission altimeter processing system DUACS/SSALTO (Data Unification and Altimeter Combination System/Segment Sol multimissions d'ALTimetrie, d'orbitographie et de Localisation precise) that processes data collected by TOPEX/POSEIDON, ERS-1 and 2, Geosat Follow-On, Envisat, Jason-1, and OSTM satellites with intercalibration and improved homogenous corrections applied to the entire data record. This is a delayed time, gridded data product with a spatial resolution of 0.25° × 0.25° on a Cartesian grid. Monthly mean SLA is computed by averaging weekly maps of SLA, month by month from January 1993 to produce one map per month since January 1993. Intercalibrated highly accurate along track data associated with a consistent mean is obtained using local inverse methods and global crossover minimizations. Finally, a global space-time objective mapping technique is used to merge along track data that corrects for correlated noise (Le Traon et al., 1998). In order to avoid an introduction of a low frequency power into each CSEOF mode, global mean SL trend of 3.4 mm/year was removed from the SLA data before computing the basis functions (1993–2018). Results from Tai (1989) suggest a RMS signal reduction of only 30% when a linear trend is removed from a 20 year record, and thus the ability of the reconstruction to capture decadal scale variations should not be impacted by the linear detrending of altimeter data before CSEOF decomposition. To account for latitudinal effects, the data were weighted using the square root of the cosine of latitude. From here on in this paper, the satellite altimeter data will be referred to as AVISO data.

2.1.3. Sea Surface Temperature Data

For historical SST data we used the Extended Reconstruction SST (ERSST) version 5 (Huang et al., 2017) data set, which is derived from the International Comprehensive Ocean-Atmosphere Data Set (ICOADS). With a temporal resolution from 1854 to present, this gridded SST data set was produced using enhanced statistical methods on a 2° latitude × 2° longitude grid with spatial completeness. No additional quality control was performed on the data since the use of averaged observations is known to curtail the influence of extreme outliers, and the technique of CSEOF reconstruction is known to be relatively insensitive to the presence of outliers. A map containing percentage of available ERSSTv5 data for 1900–1960 (Figure S2a) and 1960–2018 (Figure S2b) included in the supporting information shows 100% data availability over the IO for the full record length. The MV-R was also done using ERSSTv4 (Huang et al., 2015) and HadSST3 (Kennedy et al., 2011) to evaluate the sensitivity to the choice of historical SST data used. Our results (not shown) indicated no significant difference in the IO SL variability estimated using different historical SST data from which we concluded that the MV-R is relatively insensitive to the choice of historical SST data.

To generate the CSEOF SST basis functions for the period between 1993 and 2018 NOAA Optimum Interpolation (OI) SST v2 data (Reynolds et al., 2002) was used. This data set is comprised of monthly SST fields which have been obtained by averaging daily values that were linearly interpolated from weekly OI version 2 fields. The OI SST analysis is a blend of SST data obtained from in situ, satellite and SST simulated by sea-ice cover. The data have a spatial resolution of 1° × 1° and are available from 1981 to present. Before creating the CSEOF SST basis functions a linear trend was removed from each spatial grid point, and the data were weighted using the square root of the cosine of latitude to account for latitudinal effects. The seasonal cycle was retained prior to performing the CSEOF analysis as the analysis allows to capture the seasonal cycle in an individual mode.

2.1.4. Sea Level Pressure Data

SLP from International Comprehensive Ocean-Atmosphere Data Set (ICOADS) was used as a source of historic SLP data. The enhanced SLP product combines observations from ships and other in situ marine platform like drifting and moored buoys. This is a monthly mean gridded SLP data product and has a spatial resolution of 2° × 2° and temporal coverage from 1800 to present. For this study we use the ICOADS Release 3.0 (Freeman et al., 2017) which has new updated input data sources (more ship and buoy observations) with improvements made to the existing data sources along with removal of erroneous data. Data density maps shown in the supporting information indicate a better spatio-temporal data coverage for the period 1960–2018 (Figure S2d) compared to the data available for the period 1900–1960 (Figure S2c).

Reanalysis SLP data produced by a cooperative project between the National Centers for Environmental Prediction (NCEP) and the National Center for Atmospheric Research (NCAR), NCEP-NCAR Reanalysis

1 (Kalnay et al., 1996; Kistler et al., 2001), were used to generate the CSEOF SLP basis functions for the period between 1993 and 2018. This is a gridded, monthly mean SLP data set with a temporal resolution covering the period from 1948 to present day and a spatial resolution of 2.5° × 2.5°. Data that are assimilated for this product includes land surface, ship, rawinsonde, pibal, aircraft, and satellite data. The SLP CSEOF basis functions were computed over the IO after the removal of a linear trend from each spatial grid point with the seasonal cycle retained. To account for latitudinal effects, the data were weighted using the square root of the cosine of latitude before obtaining the CSEOF basis functions.

2.1.5. Other Gridded Sea Level Products

Globally reconstructed gridded SL data by Hamlington, Leben, et al. (2014) available from 1950 to 2009 with a spatial resolution of 0.5° × 0.5° were downloaded from NASA PODAAC and will be referred to as H-R for this study. H-R uses globally computed basis functions and reconstructs SLA using data from TG's only. A detailed description of these data can be found in Hamlington, Leben, Nerem, Han, and Kim (2011). The data have a temporal resolution of 7 days which has been averaged to monthly mean values.

Gridded monthly SL data from the Simple Ocean Data Assimilation (SODA) version 2.2.4 (Giese & Ray, 2011) were downloaded from the IRI/LDEO Climate Data Library. This version represents their first assimilation run of over 100 years and is available for the time period between 1871 and 2010. The data have a global coverage with a spatial resolution of 0.5° × 0.5°. These data will be referred to as SODA for here on.

The ECMWF Ocean ReAnalysis System 5 (ORAS 5) SL data were downloaded from Integrated Climate Data Center. ORAS 5 contains five ensemble members (see Zuo et al., 2019), and for the current study we have used the control member without perturbation. The data have a spatial resolution of 1° × 1° with a temporal resolution that spans the period of 1979 to present and a backward extension that spans 1958–1978. We have combined the ORAS 5 and ORAS 5 backward extension data, so the data now have a temporal resolution of 1958 to present and will hence forth be referred to it as ORAS5 in this paper.

2.2. Methods

Basis functions for the MV-R technique were generated using CSEOF in an attempt to represent the cyclostationary variability present in geophysical data. To give the reader some perspective, we begin by first introducing EOFs and describe common terminology associated with this method. This is followed by a brief summary of the concept of CSEOF and a discussion on the advantages of using CSEOF over EOF for generating basis functions. Finally, a detailed explanation of the newly proposed MV-R technique is provided.

2.2.1. Empirical Orthogonal Function

In the study of climate variability, linear estimation problems often involve the use of EOFs to estimate the contributions from individual statistical modes by exploiting the ability of EOFs to act as a natural basis set. The capability of EOFs lies in their ability to split data into orthogonal modes, each consisting of a single spatial pattern referred to as the loading vector (LV) and a corresponding amplitude time series called the principal component time series (PCTS) (Kim et al., 1996). An example of the mathematical representation of this system is described in the equation below:

$$D(x,\ t) = \sum_{n=1}^{m} S_n(x)\ T_n(t) \tag{1}$$

In Equation 1, $D(x, t)$ is the input spatio-temporal data, m is the total number of modes, $S(x)$ is the spatial component (LV) associated with a physical process, and $T(t)$ is a time series (PCTS) that modulates the spatial pattern. Although $S(x)$ and $T(t)$ are associated with amplitudes that are arbitrary, their product is always constrained by $D(x, t)$. In order to represent the amplitude of $S(x)$ in usual physical units a common choice is to normalize $T(t)$ to a standard deviation of one. The EOF analysis also facilitates the decomposition of the total variance into a sum of variances, thereby assigning each mode with a fraction of the total variance. This greatly helps the investigator in identifying and removing known dominant modes of variability (e.g., the annual cycle) from the data record, to help focus on the lesser known modes of variability thus making it easier to study the response of a complicated physical system. Since each EOF mode is associated with a single spatial pattern, by definition, the EOF LVs are orthogonal and considered time-independent or stationary while the amplitude of the PCTS varies in time. This stationary nature of EOF basis functions makes them

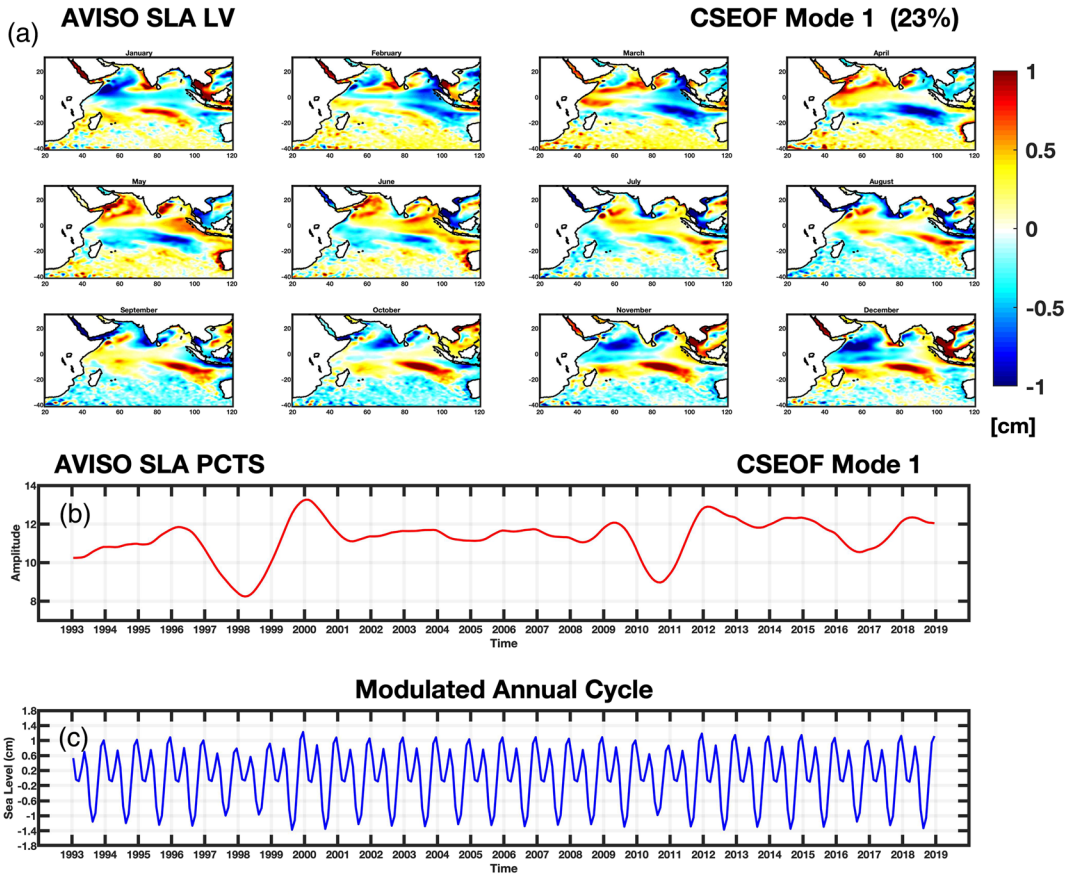

Figure 2. Mode 1 computed from the CSEOF decomposition of satellite altimetry data which represents the Modulated Annual Cycle. (a) The spatial plots show time-dependent monthly LVs and (b) associated PCTS. (c) Time series of the Modulated Annual Cycle obtained by reforming (LV*PCTS) the CSEOF mode.

suboptimal for the study of climate variability as most of the variability is associated with time-varying patterns that occurs on multiple time scales (Hamlington, Leben, Nerem, Han, & Kim, 2011; Hamlington, Leben, Nerem, & Kim, 2011; Kim et al., 1996).

2.2.2. Cyclostationary Empirical Orthogonal Function

In an attempt to better represent the cyclostationary variability present in geophysical signals Kim and Chung (2001), Kim and North (1997), Kim and Wu (1999), and Kim et al. (1996) introduced the concept of CSEOF. The primary difference between an EOF and a CSEOF mode is the time dependence of the CSEOF LVs. Unlike in an EOF mode, each CSEOF mode is associated with spatial patterns periodically varying in time determined by a defined nested period (Hamlington, Leben, Nerem, Han, & Kim, 2011). Equation 2 is the mathematical expression of this system:

$$D(x, t) = \sum_{n=1}^{m} S_n(x, t)\, T_n(t) \tag{2}$$

$$S(x, t) = S(x, t + z) \tag{3}$$

In Equation 2, the spatial pattern $S(x, t)$ is time-dependent (as defined in Equation 3) and periodic with a nested period specified by (z). As an example, consider the modulated annual cycle (MAC), created by the change in amplitude of the annual cycle oscillating in addition to its native one-year periodicity. This can be seen in Figure 2 which shows the first mode associated with the CSEOF decomposition of AVISO monthly data specifying a nested period of 1 year. There are 12 spatial patterns (LV, Figure 2a) associated with this CSEOF mode as we have specified a nested period of 1 year. The 12-month period of the annual

cycle is described by the 1-year periodicity of the LVs. The longer time-scale fluctuations of the annual cycle are described by the PCTS (Figure 2b) which is normalized to a standard deviation of one. Combining the LVs of each mode with their associated PCTS allows us to reconstruct and evaluate the contribution of each mode to the data. Figure 2c shows the time series of the MAC obtained by reforming (LV*PCTS) CSEOF mode and represents the MAC contribution to the IO regional mean sea level (MSL) explaining 23% of the total variance. It is important to note that the characteristic 6-month semiannual cycle is well represented in this reconstructed mode. For the purposes of this paper we have only used a 1-year nested period for all the CSEOF decompositions performed. Although Equations 1 and 2 look similar, the computation of CSEOF differ significantly from that of EOF. The reader is directed to the papers by Kim and Chung (2001), Kim and North (1997), Kim and Wu (1999), and Kim et al. (1996) which provide additional details and description about the computation and explanation of CSEOF.

2.2.3. Choice of Basis Function: EOF Versus CSEOF

The advantages of using CSEOF derived basis functions over EOF derived basis functions for SL reconstructions have already been discussed in detail by Hamlington, Leben, Nerem, Han, and Kim (2011) and Hamlington, Leben, Wright, and Kim (2012) and will only be briefly summarized in this subsection. The time dependence of the spatial pattern associated with CSEOF LVs allows for optimally capturing low frequency oscillations originating from signals containing nested oscillations. The ability of CSEOF to capture low- and high-frequency oscillations allows for the retention of the annual cycle in the data prior to performing a CSEOF decomposition. By limiting the variability that is lost as a result of mode mixing, CSEOF helps reconstruct individual signals (e.g., MAC) more accurately. Finally, the use of CSEOF basis functions for the reconstruction of SL makes it possible to fit a larger window of data, thus reducing the number of historical data points to derive meaningful results along with decreasing sensitivity toward inaccurate TG measurements. These are some of the main reasons that strongly favor the use of CSEOF basis functions over EOF derived basis functions for SL reconstructions.

2.2.4. Multivariate Reconstruction Technique

The reconstruction framework used for this paper was originally introduced by Hamlington, Leben, Nerem, Han, and Kim (2011). However, their reconstruction technique relied only on SL obtained from TG records, resulting in the reconstructed SL to have poor estimates for the periods where TG data were limited. This technique was later updated by Hamlington, Leben, Wright, and Kim (2012) by adding a second variable SST to help improve the reconstructed SL estimates especially during periods where the TG coverage was poor. Building on this technique we have created a new MV-R framework that now allows the input of multiple climate variables. For the current study SLP is included as the third variable which along with SL and SST allows us to reconstruct reliable estimates of 20th century IO SL.

The MV-R technique used to create gridded space-time SL data for this paper can be divided into three sections. First, the CSEOF Decomposition wherein the data were decomposed using the CSEOF analysis over the IO. For this we decomposed SLA (AVISO), SST (OISST), and SLP (NCEP-NCAR) data spanning the period between 1993 and 2018. The seasonal cycle was not removed prior to CSEOF decomposition as CSEOF has the ability to reconstruct individual signals more accurately by limiting the variability that is lost as a result of mode mixing. From the CSEOF decomposition, the first 21 modes for SLA, SST, and SLP were retained which explained at least 90% of the total variance in each data set. Since we are not interested in reconstructing the SL seasonal cycle, we have opted to leave out the mode associated with the seasonal cycle before continuing with the following steps. The second step is the Regression which is based on the understanding that it is possible to find physical and dynamic consistency between two separate CSEOF sets computed from two different variables. The regression technique introduced by Lim and Kim (2007) was used in this step of the reconstruction. The SST and SLP PCTS are regressed onto each individual SLA PCTS generating a set of regression coefficients. These regression coefficients were then used along with the original SST and SLP LVs to create a new set of SST and SLP LVs which now have the amplitude fluctuations as described by the SLA PCTS. The final step is the Estimation of PCTS which was done by simultaneously fitting the new SST and SLP LVs along with the SLA LVs to historical measurements of SST (ERSST), SLP (ICOADS), and SLA (TG) thus, reconstructing a SLA PCTS with a temporal resolution that covers the period between 1900 and 2018.

Like any other EOF-based reconstruction technique, an important assumption made while performing a reconstruction using this technique is that the basis functions and regression relationship both, calculated

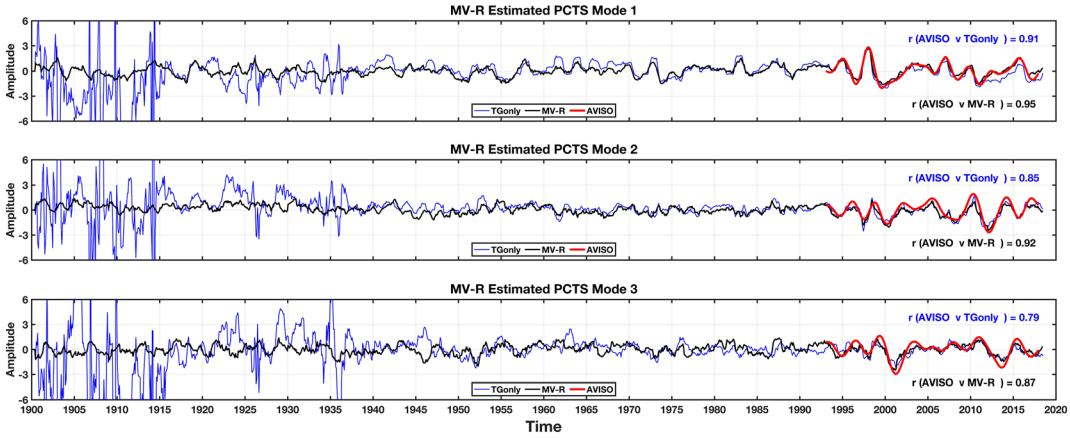

Figure 3. Estimated PCTS for the first three modes (seasonal cycle not included) for a tide gauge only reconstruction (blue line) and multivariate reconstruction (MV-R, black line). PCTS for the AVISO SLA (red line) is in good agreement with the PCTS estimated from the MV-R. For the first half of the 20th century, PCTS is observed to be overestimated by the TG only reconstruction.

over the time period 1993–2018, hold good over the entire time period of reconstruction. Due to the short record length of the AVISO data a single CSEOF mode is not capable of capturing the secular trend pattern and in the event that it did capture it, we cannot assume the secular trend pattern to be stationary over the reconstructed time period. We choose to take a similar approach used by Hamlington, Leben, Nerem, Han, and Kim (2011) wherein the computation of the secular trend was separated from the reconstruction procedure. Furthermore, since the primary goal of this study is to improve the estimates of regional IO SL variability using a regional reconstruction technique, we made no attempt to reproduce trends in GMSL. To remove the effect of GMSL, before performing a CSEOF decomposition a trend of 3.4 mm/year for the period between 1993 and 2018 was subtracted from the AVISO SLA. It is important to note that all the trends presented in this paper are relative to a background trend in GMSL over the period for which it is computed. Estimation of GMSL is not an easy task and for a global MV-R, GMSL can be accounted for by using an approach similar to the one presented in Hamlington, Leben, Nerem, Han, and Kim (2011).

From previously published research it has been noted that errors emerging from truncating the basis functions estimated from the AVISO data set in combination with instrument errors do not significantly influence the results presented in this paper. As error estimation for a SL reconstruction procedure is not a trivial task, errors are not explicitly addressed in this paper following a similar approach taken by previous studies (e.g., Church et al., 2004; Hamlington, Leben, Nerem, Han, & Kim, 2011). Reconstruction techniques that involve estimating the GMSL generally include a detailed error analysis, but as we do not include the GMSL such a discussion is not presented.

3. Results

3.1. Estimated PCTS

To emphasize the importance of using this new MV-R technique specifically in a region characterized by a sparse spatio-temporal TG coverage, we have compared PCTS estimated from the MV-R technique to PCTS estimated from a "TG only" reconstruction (Figure 3). While the PCTS of MV-R was estimated using historical SST and SLP in addition to SLA, we estimated the PCTS of the TG only reconstruction by using TG data alone as a source of historical data. For this kind of reconstruction technique, the reliability and quality of the estimated PCTS depend largely on the spatio-temporal resolution of the input TG data. For this reason, a larger number of longer TG records spread evenly across the study area is ideal. However, for a region like the IO that has an extremely poor spatio-temporal resolution of TG data, estimating historic PCTS from a TG only reconstruction is not ideal. This is illustrated in Figure 3, which shows estimated PCTS of the first three modes (not including the seasonal cycle) for a TG only reconstruction (blue line) and MV-R (black line). The

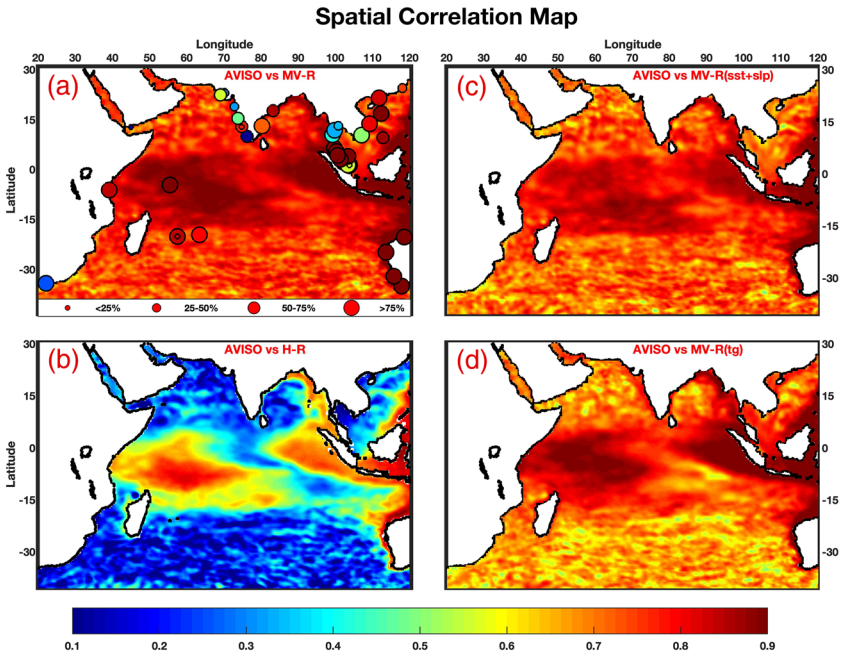

Figure 4. (a) Spatial correlation map between the AVISO SLA and multivariate reconstruction (MV-R) SLA computed for the time period between 1993 and 2008. The spatial correlation map for the period 1993–2018 was computed and looks similar to panel (a). The bubbles represent correlation between subsampled MV-R SLA and tide gauge data and the size of each bubble corresponds to the percentage of available tide gauge data over the period 1993–2018. (b) Spatial correlation map between the AVISO SLA and Hamlington, Leben, et al. (2014) (H-R) SLA created for the period between 1993 and 2008. (c) Spatial correlation map (1993–2008) between AVISO SLA and reconstructed SLA created using both SST and SLP without the use of TG data referred to as MV-R (sst + slp). (d) Spatial correlation map (1993–2008) between AVISO SLA and reconstructed SLA created using only TG data referred to as MV-R (tg). It should be noted that MV-R (tg) is same as "TG only" reconstruction that is discussed in Figure 3. Prior to calculating the correlation coefficients, the seasonal cycle was removed and SLA was detrended. The color bar represents the correlation coefficient value.

spatial pattern associated with these three modes have been shown in the supporting information (Figure S3). Although both the reconstructions seem to agree well with the AVISO PCTS (red line), it is important to note that PCTS estimated from the MV-R shows a better agreement to AVISO PCTS as indicated by the correlation values listed in Figure 3.

Differences between the estimated PCTS begin to grow as we go back in time notably for the first half of the 20th century. The TG only reconstruction is observed to largely overestimate the PCTS before 1940, thus limiting its use to only the second half of the 20th century. The reason for this large variability observed in the estimated PCTS of the TG only reconstruction can be explained based on the drastic decline in the number of TG stations (see Figure 1), decreasing from 20 in 1980 to only three available TG stations before 1940. Although we use the same TG set for the MV-R technique, the inclusion of historic SST and SLP to supplement the lack of spatio-temporal TG data results in estimating a more consistent PCTS for the entire 20th century. This is a major advantage of using the MV-R technique over a TG only reconstruction technique for reconstructing 20th century SL variability in the IO.

3.2. Spatial Comparison With AVISO Data

The ability to reproduce spatial variability of SLA over the IO can be evaluated by checking linear dependence with AVISO SLA using a spatial correlation map. We compared the AVISO SLA to the SLA obtained from the MV-R and H-R. To decrease the chances of obtaining high spatial correlation coefficients, commonly generated by the presence of a seasonal cycle and trends being included in the data, before calculating the correlation coefficients the data sets were detrended, and the seasonal cycle removed. Figure 4a shows the spatial correlation between the AVISO SLA and the MV-R SLA for the period between 1993 and 2008. The spatial correlation between the AVISO SLA and H-R SLA for the same period is shown in Figure 4b.

It should be noted that the spatial correlation between the MV-R SLA and AVISO SLA for the period between 1993 and 2018 was checked and found to be similar to Figure 4a. For both the reconstructions, spatial correlations with AVISO SLA are observed to be higher in the equatorial IO (10°S to 10°N) and along the west Australian coast with a spatially averaged correlation coefficient of 0.79 for the MV-R and 0.37 for H-R over the entire IO.

For the H-R SLA (Figure 4b), extremely poor correlations with the AVISO SLA are observed in the southern IO and regional Arabian Sea and Bay of Bengal. The overall poor correlations observed between H-R and AVISO could be a result of a combination of factors. H-R uses globally computed basis functions which may not be able to resolve IO variability at a regional scale. The basis functions used for H-R were computed using a shorter altimeter record length, and H-R SLA was reconstructed using TGs only. By using IO regionally computed basis functions derived from a longer altimeter record coupled with the use of the MV-R technique, an overall good correlation with the AVISO SLA is observed in Figure 4a. This is an indication that the new regionally reconstructed MV-R data have improved regional SL variability estimates for the IO compared to H-R. The fill color of the bubbles in Figure 4a represents the correlation coefficient between the TG SLA and MV-R SLA. The size of each bubble corresponds to the percentage of TG data available between 1993 and 2018. Majority of the TG stations show a very good agreement to the MV-R SLA indicating that the MV-R technique is successful in representing the SL variability observed in TG data. It is observed that TG stations that have data records of lesser than 50% between 1993 and 2018 tend to show a lesser agreement with the MV-R and AVISO SLA. The TG station of Aden (Red Sea) has a very short and discontinuous SLA record resulting in a poor correlation coefficient. In contrast, although the TG stations along the coast of Thailand have a long SL record length, the SLA shows large variability which is not seen in the MV-R SLA resulting in lower correlation coefficients for these stations. Along the west coast of India, TG stations are characterized by SL records that have a large variability accompanied by long data gaps resulting in low correlation coefficients. Overall, the SLA estimated from the MV-R technique is seen to have a good estimate of regional SL of the IO over the altimeter time period.

To help highlight the benefits of reconstructing SLA using the MV-R (SST and SLP in addition to SLA), we created two additional reconstructions (1) that used both SST and SLP without the use of TG data referred to as MV-R (sst + slp) and (2) that used only TG data referred to as MV-R (tg). It should be noted that MV-R (tg) is same as "TG only" reconstruction that is discussed in section 3.1. Spatial correlations were then computed over the period 1993–2008 between MV-R (sst + slp) and AVISO SLA (Figure 4c) and MV-R (tg) and AVISO SLA (Figure 4d). Results of spatial correlations computed over 1993–2018 were found to be similar to Figures 4c and 4d. It can be observed that the correlation values in Figure 4c specifically over the southcentral IO, Arabian Sea and Bay of Bengal are marginally higher when compared to Figure 4d. This demonstrates the advantage of including SST and SLP to help improve SL variability over the open ocean (over which TG data are very sparse). The advantages of regionally reconstructing SL data are apparent from the comparison between Figure 4b (H-R was created using globally computed basis functions) and Figure 4d which is created using IO regionally computed basis functions. Although both H-R and MV-R (tg) use the similar set of TGs in the IO, H-R basis functions were created using a shorter altimeter data length which could also be why Figure 4b shows poor correlations. Overall, the higher spatial correlation observed in Figure 4a when compared to Figures 4c and 4d is evidence that shows the benefit of using SLA, SST, and SLP to reconstruct SLA. Although one can argue that there is only a marginal increase in correlation values (Figure 4a compared to Figures 4c and 4d) the real advantage of using MV-R (SST and SLP in addition to SLA) is in reconstructing SL variability over the first half of the 20th century.

3.3. IO Regional SL Trends

By making direct trend comparisons to AVISO SLA we attempt to validate the ability of the MV-R technique to accurately reproduce regional trends observed in the IO SL. For all the trend maps presented in this section it is important to note that regional trends are computed relative to a background GMSL trend over the same time period as we have chosen not to account for GMSL (see section 2.2.4). Prior to computing the trend map the seasonal cycle from the AVISO and MV-R SLA was removed. The spatial distribution of regional SLA trends over the IO for the period between 1993 and 2018 is shown in Figure 5 for the (a) AVISO and (b) MV-R SLA. Higher trends are observed in the eastern IO and parts of the southern IO. It is encouraging to

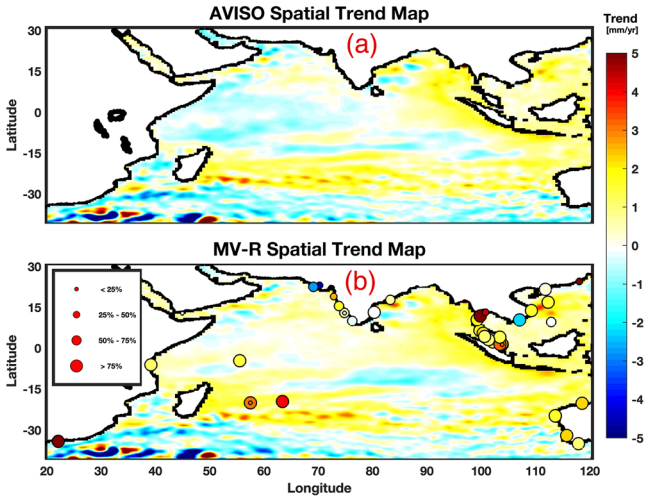

Figure 5. Spatial trend map for the (a) AVISO SLA and (b) multivariate reconstruction (MV-R) SLA computed from 1993 to 2018. Color of bubbles represent the tide gauge trend value and the size of each bubble corresponds to the percentage of available tide gauge data over the period 1993–2018. Before calculating the trend, seasonal cycle and a trend of 3.4 mm/year were removed for the period between 1993 and 2018. The color bar represents the trend in units of mm/year.

observe that the overall spatial trend patterns of the MV-R SLA closely resemble spatial trend patterns of regional IO as observed from the AVISO SLA.

In Figure 5b the color of the bubbles represents the trend value at each TG station. Prior to computing trends, the seasonal cycle and a GMSL trend have been removed from the TG data for the period between 1993 and 2018. In general, trends computed from TG stations that have more than 75% of continuous uniformly varying SLA data tend to show a very good agreement to MV-R SLA trends. A good example of some such TG stations are those along the Australian coast. On the other hand, TG stations with short SLA records, those containing long periods of missing data, and an overall noncoherent SLA variability are seen to be associated with trend estimates that disagree with the MV-R SLA trends, the west coast of India being one such example. Overall, majority of the TG stations show a reasonably good agreement between the trends computed at each TG station and trends computed from the subsampled MV-R SLA at the respective TG location.

Results from the study of Nidheesh et al. (2017) suggest that there exists disagreements in the interproduct SL consistency and coherence among various reconstructed and reanalysis products over the southwest tropical IO. This is observed in Figure 6 which shows a comparison of spatial trend maps computed from the (a) MV-R, (b) H-R, (c) SODA, and (d) ORAS5 for the common time period of 1958–2005 over which the data overlap. Before computing spatial trend maps the seasonal cycle and GMSL rise time series have been removed. Multidecadal SL trends over the tropical

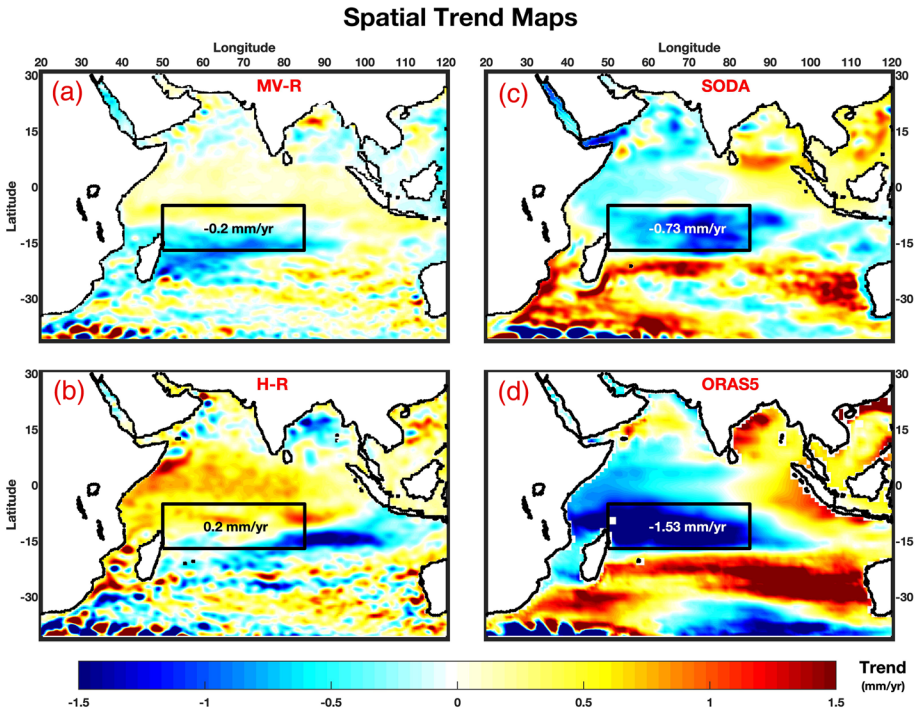

Figure 6. Spatial trend map computed from (a) multivariate reconstruction (MV-R), (b) Hamlington, Leben, et al. (2014) (H-R), (c) SODA and (d) ORAS5 for the period 1958–2005. The seasonal cycle and global mean sea level rise time series have been removed before computing the trend. The boxed region (17°S–5°S, 50°E–85°E) is situated in the thermocline ridge of the southwest tropical IO, and the value in each box represents the spatially averaged trend value over the box. The color bar represents the trend in units of mm/year.

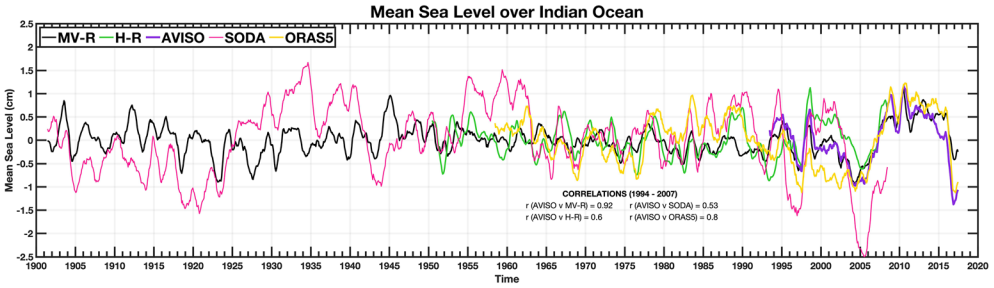

Figure 7. Spatially averaged mean sea level (MSL) computed from the multivariate reconstruction (MV-R), Hamlington, Leben, et al. (2014) (H-R), AVISO, SODA, and ORAS5 SLA over the Indian Ocean. The seasonal cycle and global mean sea level rise time series have been removed before computing the MSL. To focus on the ability of the MV-R to represent variability about a trend, the MSL has been detrended and smoothed using a running mean over the period of a year.

southwest IO show falling SL since the 1960s (Han et al., 2010). To make quantitative comparisons in Figure 6, spatially averaged trends are computed over the boxed region (17°S–5°S, 50°E–85°E) which is situated in the thermocline ridge of the southwest tropical IO. These trend values over the boxed region have been listed in Figure 6. A falling SL trend is observed for the MV-R, SODA, and ORAS5 data (Figures 6a, 6c, and 6d). However, the trend computed from H-R (Figure 6b) shows a rising SL trend value. Although magnitude of the trend in each box over southwest tropical IO is different, there is no simple way to ascertain which data product best represents the IO SL variability. Nonetheless, the ability of the MV-R to capture this falling SL trend over the southwest tropical IO represents a significant improvement in reconstructed regional IO SL variability estimates when compared to the reconstructed SL of H-R.

3.4. IO Mean Sea Level Variability

The robustness of this MV-R technique lies in its capability of capturing variability about a trend. This can be evaluated by the ability to reproduce the variability observed in the IO MSL. The spatially averaged IO MSL computed from the MV-R, H-R, AVISO, SODA, and ORAS5 data over the IO is presented in Figure 7. Before computing the MSL, the seasonal cycle and global mean SL rise time series have been removed from the data. To focus on the ability of the reconstruction to represent variability about a trend, MSL has been detrended and smoothed by computing running means over 1-year periods. In Figure 7, over the more recent altimeter time period (1994–2008) for which all the data sets overlap, MSL calculated from the MV-R has a very good agreement with the AVISO MSL and shows a similar pattern of variability when compared to the MSL calculated from ORAS5 data. MSL computed from H-R and SODA is seen to overestimate the amplitude of IO MSL variability when compared to AVISO MSL. Over the longer period for which the data overlap (1959–2005), the MV-R MSL is observed to have highest correlation with SODA MSL (0.44), followed by ORAS5 (0.34) and H-R (0.33). As we go back in time, the similarities in the variability of IO MSL between the SODA and MV-R begin to decrease and have the most difference for the first half of the 20th century. Based on the good agreement with AVISO MSL, the reduced variability of the MV-R IO MSL observed for the first half of the 20th century can be considered to be more realistic than the IO MSL computed from SODA data. A good agreement between the AVISO MSL and MV-R MSL along with similarities in the pattern of MSL variability between the MV-R and SODA MSL is evidence that the MV-R technique is robust in representing variability about a trend. These results further suggest that the MV-R technique shows an improvement in representing the regional IO SLA variability over the SLA variability estimates from H-R.

3.5. Climate Modes Over the Tropical IO

In its mean state, the thermocline in the tropical IO is deep in the eastern equatorial IO warm pool region due to the mean surface westerly winds in the equatorial IO. The resulting structure of the thermocline, which is deeper in the east as compared to the west, favors a more convective atmosphere over the eastern equatorial IO. However, during particular years unusually warm SSTs are observed over large parts of the western IO and unusually cold SSTs off Java/Sumatra which is accompanied by wind and rainfall anomalies. This anomalous departure from the mean state of the ocean-atmosphere system is referred to as the Indian Ocean Dipole (IOD) (Murtugudde et al., 2000; Saji et al., 1999; Webster et al., 1999) and is well recognized as

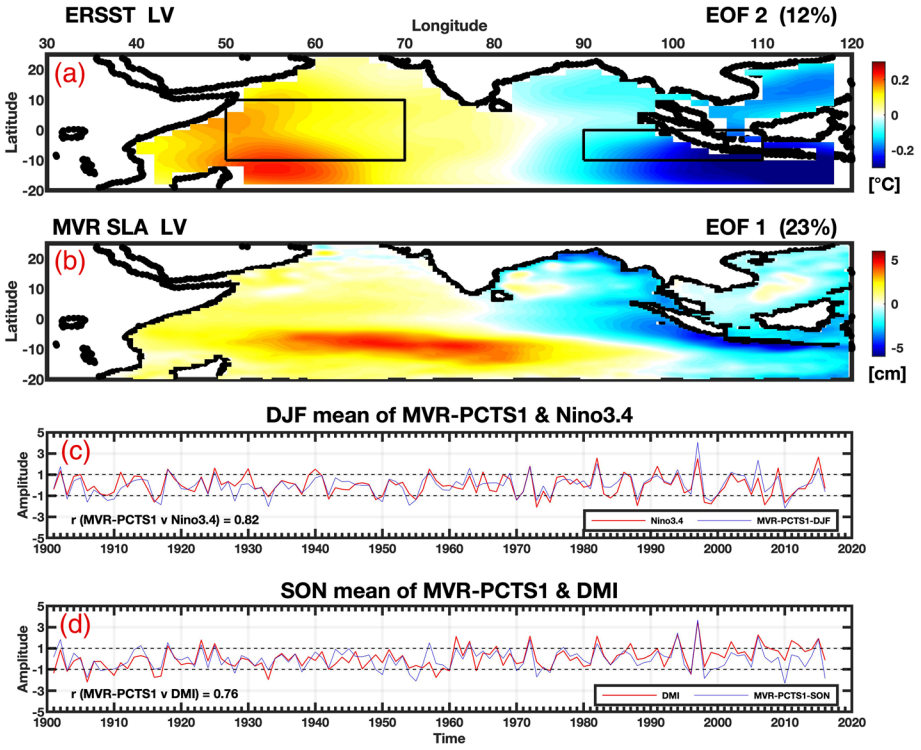

Figure 8. (a) Spatial structure associated with the second EOF mode of ERSST data recognized as the Indian Ocean Dipole. (b) Spatial structure of the first EOF mode of multivariate reconstruction (MV-R) SLA showing spatial characteristics of a dipole. (c) December, January, February (DJF) average value of MV-R PCTS 1 and Niño3.4 index. (d) September, October, November (SON) average value of MV-R PVTS 1 and DMI index. Black dashed line in (c) and (d) indicates one standard deviation.

the dominant signal of coupled interannual variability in the IO. The spatial pattern associated with the second EOF mode of SST data is known to be characteristic to the IOD spatial structure and can be seen from Figure 8a, which shows the second EOF mode of ERSST.

For SLA data, the spatial pattern associated with the first EOF mode shows a dipole structure with higher SLA over parts of the western and central equatorial IO and lower SLA in the eastern IO. Figure 8b is the spatial pattern of the first EOF mode of the MV-R SLA. The northwest-southeast slope of higher SLA has been shown to be driven by anomalous downwelling Rossby waves that are radiated from the southeastern IO, primarily forced by anticyclonic wind stress curl in the southern tropical IO (Rao & Behera, 2005; Xie et al., 2002). The lower SLA seen in the eastern equatorial IO is a result of Kelvin waves excited by equatorial easterly wind anomalies along with local southeasterly along shore winds that drive a coastal upwelling along the Java/Sumatra coast. These upwelling Kelvin waves are reflected as coastally trapped Kelvin waves and associated with lower SLA observed along the Bay of Bengal (Rao et al., 2002; Vinayachandran et al., 1999, 2002).

Results from the study of Zhang and Han (2020) suggest that the SLA pattern as observed in Figure 8b is associated with the climate modes of IOD and ENSO. Their results indicate that the eastern Pacific El Niño causes a reduced convection over the eastern tropical IO and an enhanced convection over the western IO basin driving easterly wind anomalies which in turn strongly influences the upwelling in the eastern tropical IO. Since the ENSO is highly seasonal and phase locked to boreal winter, we have computed the December, January, February (DJF) average value of MV-R PCTS 1 and Niño3.4 index (used for representing ENSO events). The impact of ENSO on IO SLA can be observed from Figure 8c which shows a high correlation of 0.82 with the MV-R PCTS 1. Saji et al. (1999) proposed an index for representing IOD events called the Dipole Mode Index (DMI) which relates to the PCTS of the second EOF mode of SST data over the IO. DMI is defined as the difference in the spatially averaged SST anomalies between the tropical western IO box and

tropical south-eastern IO box (Figure 8a). The IOD is also highly seasonal but phase locked to boreal fall. For this reason, we computed the September, October, November (SON) average value of MV-R PVTS 1 and DMI which is presented in Figure 8d. The high correlation of 0.76 highlights the dominance of IOD on IO SL. It should be noted that in Figures 8c and 8d, the data have been normalized by its standard deviation before plotting. From Figure 8d it can be observed that the positive IOD events of 1961, 1963, 1972, 1982, 1994, 1997, and 2006 are captured by the MV-R PCTS 1. Although some of the positive IOD events co-occur with El Niño, the ability of the reconstruction to capture independent IOD events of 1961, 1994, and 2006 reassures us that the MV-R technique is capable of representing the regional IO SL variability significantly well.

4. Discussion and Conclusion

The large uncertainty associated with the precise quantification of 20th century SL rise in the IO can be attributed to the presence of short and sparse historic IO TG records. This has also resulted in disagreements on the relationship between IO SL and internal variability. To capture and accurately represent the SL contribution of internal variability on interannual to decadal timescales, long SL data records are essential. SL reconstruction techniques provide a way to overcome these sampling challenges by combining historic TG data with the information about the spatial covariance of SL from satellite altimetry. However, due to the lack of historical SL data records coupled with poor spatial and temporal sampling by TG stations, it is extremely challenging to reconstruct SL data in the IO. With only three available TG records before 1940, using a TG only reconstruction technique in the IO limits reliable estimates SL to only the second half of the 20th century. In this paper we have proposed a new regional CSEOF MV-R framework that represents a significant advancement over traditional SL reconstruction techniques. The ability of CSEOF to capture the evolution of cyclostationary geophysical signals in a single mode make CSEOF a more appropriate set of basis functions for reconstructing SL than compared to the more traditional choice of using EOF. Building upon an existing CSEOF SL reconstruction technique, SLP has been included as the third variable which along with SLA and SST is used to reconstruct IO SLA back to 1900.

This MV-R technique is seen to be particularly effective in estimating PCTS for the first half of the 20th century (see Figure 3), which is characterized by an extremely poor spatio-temporal TG coverage in the IO. For the more recent altimeter time period (1993–2008), the MV-R SLA shows a significantly better agreement to AVISO SLA than compared to the agreement shown between H-R SLA and AVISO SLA (Figure 4). Similarities in the spatial trend patterns calculated from the MV-R SLA and AVISO SLA seen from Figure 5 highlight the strength of the MV-R technique in representing regional IO SLA trends. TG stations that have a long continuous SLA record show a good correlation and have trend estimates that are similar to the MV-R SLA subsampled at the respective TG station. These results show that the MV-R technique is capable of providing reliable estimates of the IO SL over the altimeter time period, which in turn implies that this technique can be reliably used to generate realistic IO SL trends and variability for the 20th century. The ability of the MV-R technique to capture the falling SL trend over the southwest tropical IO (Figure 6a) indicates a significant improvement in regional IO SL estimates when compared to H-R. Consistent with previous literature, the first EOF mode of the MV-R SLA was observed to be associated with a dipole structure and the PCTS showed a good agreement to both the DMI and Niño3.4 index. The IOD events that occurred independent of El Niño were successfully captured indicating the capability of the MV-R technique in representing internal climate variability observed in the IO.

Through this work it has been shown that the MV-R technique is robust and particularly well suited to overcome the reconstruction challenges imposed by the sparse spatio-temporal TG sampling in the IO. Using this MV-R technique, we have successfully reconstructed regional IO SLA for the 20th century. This new technique provides improved SL variability estimates and a longer temporal coverage in comparison to the reconstructed product of Hamlington, Leben, et al. (2014). Perhaps the biggest advantage of using this reconstruction technique is the ability to reliably estimate SLA back to 1900 which will enable future studies to shed more light on the IO internal climate variability on interannual to decadal timescales. Further, by removing the fraction of variability associated with internal climate variability the underlying anthropogenic trend in the IO can be isolated. This will help gain valuable insights into the past, present, and

future SL changes occurring as a result of anthropogenic influences. An improved representation of internal climate variability and the anthropogenic trend in the IO will ultimately benefit all the adaptation and mitigation strategies developed for the densely populated IO rim countries that are becoming increasingly vulnerable to the effects of higher SL.

Data Availability Statement

The new reconstructed multivariate IO SL data have been made publicly available and can be downloaded online (https://doi.org/10.6084/m9.figshare.12609908). Tide gauge data are available from the Permanent Service for Mean Sea Level (PSMSL; https://www.psmsl.org/), and the satellite altimetry data are available through the Archiving, Validation and Interpretation of Satellite Oceanographic (AVISO; https://www.aviso.altimetry.fr/en/data/products/sea-surface-height-products/global.html). SST and SLP data are provided by NOAA/OAR/ESRL Physical Sciences Division, Boulder, Colorado, USA. The Extended Reconstruction SST (ERSST) version 5 SST data were downloaded from https://psl.noaa.gov/data/gridded/data.noaa.ersst.v5.html, and the NOAA Optimum Interpolation (OI) SST v2 data were downloaded online (https://www.esrl.noaa.gov/psd/data/gridded/data.noaa.oisst.v2.html). SLP data from the International Comprehensive Ocean-Atmosphere Data Set (ICOADS) were download at https://www.esrl.noaa.gov/psd/data/gridded/data.coads.2deg.html and the NCEP-NCAR Reanalysis 1 SLP data were download online (https://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis.surface.html). The globally reconstructed SLA data by Hamlington, Leben, et al. (2014) can be downloaded at https://podaac-opendap.jpl.nasa.gov/opendap/allData/recon_sea_level/preview/L4/tg_recon_sea_level/, Simple Ocean Data Assimilation (SODA) version 2.2.4 were downloaded from https://iridl.ldeo.columbia.edu/SOURCES/.CARTON-GIESE/.SODA/.v2p2p4/index.html?Set-Language=en, and ECMWF Ocean ReAnalysis System 5 (ORAS 5) SL data and backward extension were downloaded online (http://icdc.cen.uni-hamburg.de/projekte/easy-init/easy-init-ocean.html).

Acknowledgments

This research was carried out at the Center for Coastal and Physical Oceanography, Old Dominion University, Norfolk, VA, USA. Two anonymous reviewers provided many useful suggestions that helped to improve the manuscript. P. K., B. D. H, S. C., W. H., and P .R. T. acknowledge support from the National Science Foundation, award PO1558741. The authors would like to thank the data providers that make this work possible.

References

Carton, J. A., & Giese, B. S. (2008). A reanalysis of ocean climate using Simple Ocean Data Assimilation (SODA). *Monthly Weather Review*, *136*(8), 2999–3017. https://doi.org/10.1175/2007MWR1978.1

Cazenave, A., Dieng, H.-B., Meyssignac, B., Von Schuckmann, K., Decharme, B., & Berthier, E. (2014). The rate of sea-level rise. *Nature Climate Change*, *4*(5), 358–361. https://doi.org/10.1038/nclimate2159

Chambers, D. P., Mehlhaff, C. A., Urban, T. J., Fujii, D., & Nerem, R. S. (2002). Low-frequency variations in global mean sea level: 1950–2000. *Journal of Geophysical Research*, *107*(C4), 3026. https://doi.org/10.1029/2001JC001089

Cheon, S.-H., Hamlington, B. D., & Suh, K.-D. (2018). Reconstruction of sea level around the Korean Peninsula using cyclostationary empirical orthogonal functions. *Ocean Science*, *14*(5), 959–970. https://doi.org/10.5194/os-14-959-2018

Church, J. A., & White, N. J. (2006). A 20th century acceleration in global sea-level rise. *Geophysical Research Letters*, *33*, L01602. https://doi.org/10.1029/2005GL024826

Church, J. A., & White, N. J. (2011). Sea-level rise from the late 19th to the early 21st century. *Surveys in Geophysics*, *32*(4–5), 585–602. https://doi.org/10.1007/s10712-011-9119-1

Church, J. A., White, N. J., Coleman, R., Lambeck, K., & Mitrovica, J. X. (2004). Estimates of the regional distribution of sea level rise over the 1950–2000 period. *Journal of Climate*, *17*(13), 2609–2625. https://doi.org/10.1175/1520-0442(2004)017<2609:EOTRDO>2.0.CO;2

Frankcombe, L. M., McGregor, S., & England, M. H. (2015). Robustness of the modes of Indo-Pacific sea level variability. *Climate Dynamics*, *45*(5–6), 1281–1298. https://doi.org/10.1007/s00382-014-2377-0

Freeman, E., Woodruff, S. D., Worley, S. J., Lubker, S. J., Kent, E. C., Angel, W. E., et al. (2017). ICOADS release 3.0: A major update to the historical marine climate record. *International Journal of Climatology*, *37*(5), 2211–2232. https://doi.org/10.1002/joc.4775

Giese, B. S., & Ray, S. (2011). El Niño variability in simple ocean data assimilation (SODA), 1871–2008. *Journal of Geophysical Research*, *116*, C02024. https://doi.org/10.1029/2010JC006695

Hamlington, B. D., Leben, R. R., & Kim, K.-Y. (2012). Improving sea level reconstructions using non-sea level measurements. *Journal of Geophysical Research*, *117*, C10025. https://doi.org/10.1029/2012JC008277

Hamlington, B. D., Leben, R. R., Nerem, R. S., Han, W., & Kim, K.-Y. (2011). Reconstructing sea level using cyclostationary empirical orthogonal functions. *Journal of Geophysical Research*, *116*, C12015. https://doi.org/10.1029/2011JC007529

Hamlington, B. D., Leben, R. R., Nerem, R. S., & Kim, K.-Y. (2011). The effect of signal-to-noise ratio on the study of sea level trends. *Journal of Climate*, *24*(5), 1396–1408. https://doi.org/10.1175/2010JCLI3531.1

Hamlington, B. D., Leben, R. R., Strassburg, M. W., & Kim, K.-Y. (2014). Cyclostationary empirical orthogonal function sea-level reconstruction. *Geoscience Data Journal*, *1*, 13–19. https://doi.org/10.1002/gdj3.6

Hamlington, B. D., Leben, R. R., Strassburg, M. W., Nerem, R. S., & Kim, K.-Y. (2013). Contribution of the Pacific Decadal Oscillation to global mean sea level trends. *Geophysical Research Letters*, *40*, 5171–5175. https://doi.org/10.1002/grl.50950

Hamlington, B. D., Leben, R. R., Wright, L. A., & Kim, K.-Y. (2012). Regional sea level reconstruction in the Pacific ocean. *Marine Geodesy*, *35*(sup1), 98–117.

Hamlington, B. D., Strassburg, M. W., Leben, R. R., Han, W., Nerem, R. S., & Kim, K. Y. (2014). Uncovering an anthropogenic sea-level rise signal in the Pacific Ocean. *Nature Climate Change*, *4*, 782–785. https://doi.org/10.1038/nclimate2307

Hamlington, B. D., & Thompson, P. R. (2015). Considerations for estimating the 20th century trend in global mean sea level. *Geophysical Research Letters*, *42*, 4102–4109. https://doi.org/10.1002/2015GL064177

Han, W., Meehl, G. A., Hu, A., Alexander, M. A., Yamagata, T., Yuan, D., et al. (2014). Intensification of decadal and multi-decadal sea level variability in the western tropical Pacific during recent decades. *Climate Dynamics*, *43*, 1357–1379. https://doi.org/10.1007/s00382-013-1951-1

Han, W., Meehl, G. A., Rajagopalan, B., Fasullo, J. T., Hu, A., Lin, J., et al. (2010). Patterns of Indian Ocean sea-level change in a warming climate. *Nature Geoscience*, *3*(8), 546–550. https://doi.org/10.1038/ngeo901

Han, W., Vialard, J., McPhaden, M. J., Lee, T., Masumoto, Y., Feng, M., & De Ruijter, W. P. M. (2014). Indian Ocean decadal variability: A review. *Bulletin of the American Meteorological Society*, *95*(11), 1679–1703. https://doi.org/10.1175/BAMS-D-13-00028.1

Hay, C. C., Morrow, E., Kopp, R. E., & Mitrovica, J. X. (2015). Probabilistic reanalysis of twentieth-century sea-level rise. *Nature*, *517*(7535), 481–484. https://doi.org/10.1038/nature14093

Huang, B., Banzon, V. F., Freeman, E., Lawrimore, J., Liu, W., Peterson, T. C., et al. (2015). Extended reconstructed sea surface temperature version 4 (ERSST. v4). Part I: Upgrades and intercomparisons. *Journal of Climate*, *28*(3), 911–930. https://doi.org/10.1175/JCLI-D-14-00006.1

Huang, B., Thorne, P. W., Banzon, V. F., Boyer, T., Chepurin, G., Lawrimore, J. H., et al. (2017). Extended reconstructed sea surface temperature version 5 (ERSSTv5), upgrades, validations, and intercomparisons. *Journal of Climate*, *30*, 8179–8205. https://doi.org/10.1175/JCLI-D-16-0836.1

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., et al. (1996). The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society*, *77*(3), 437–471. https://doi.org/10.1175/1520-0477(1996)077<0437:tnyrp>2.0.co;2

Kaplan, A., Cane, M. A., Kushnir, Y., Clement, A. C., Blumenthal, M. B., & Rajagopalan, B. (1998). Analyses of global sea surface temperature 1856–1991. *Journal of Geophysical Research*, *103*(C9), 18,567–18,589. https://doi.org/10.1029/97JC01736

Kaplan, A., Kushnir, Y., & Cane, M. A. (2000). Reduced space optimal interpolation of historical marine sea level pressure: 1854–1992. *Journal of Climate*, *13*(16), 2987–3002. https://doi.org/10.1175/1520-0442(2000)013<2987:RSOIOH>2.0.CO;2

Kennedy, J. J., Rayner, N. A., Smith, R. O., Saunby, M., & Parker, D. E. (2011). Reassessing biases and other uncertainties in sea surface temperature observations since 1850 part 1: Measurement and sampling errors. *Journal of Geophysical Research*, *116*, D14103. https://doi.org/10.1029/2010JD015218

Kim, K.-Y., & Chung, C. (2001). On the evolution of the annual cycle in the tropical Pacific. *Journal of Climate*, *14*(5), 991–994. https://doi.org/10.1175/1520-0442(2001)014<0991:OTEOTA>2.0.CO;2

Kim, K.-Y., & North, G. R. (1997). EOFs of harmonizable cyclostationary processes. *Journal of the Atmospheric Sciences*, *54*(19), 2416–2427. https://doi.org/10.1175/1520-0469(1997)054<2416:EOHCP>2.0.CO;2

Kim, K.-Y., North, G. R., & Huang, J. (1996). EOFs of one-dimensional cyclostationary time series: Computations, examples, and stochastic modeling. *Journal of the Atmospheric Sciences*, *53*(7), 1007–1017. https://doi.org/10.1175/1520-0469(1996)053<1007:EOODCT>2.0.CO;2

Kim, K.-Y., & Wu, Q. (1999). A comparison study of EOF techniques: Analysis of nonstationary data with periodic statistics. *Journal of Climate*, *12*(1), 185–199. https://doi.org/10.1175/1520-0442-12.1.185

Kistler, R., Collins, W., Saha, S., White, G., Woollen, J., Kalnay, E., et al. (2001). The NCEP–NCAR 50-year reanalysis: Monthly means CD-ROM and documentation. *Bulletin of the American Meteorological Society*, *82*(2), 247–267. https://doi.org/10.1175/1520-0477(2001)082<0247:tnnyrm>2.3.co;2

Klein, S. A., Soden, B. J., & Lau, N.-C. (1999). Remote sea surface temperature variations during ENSO: Evidence for a tropical atmospheric bridge. *Journal of Climate*, *12*(4), 917–932. https://doi.org/10.1175/1520-0442(1999)012<0917:RSSTVD>2.0.CO;2

Le Traon, P. Y., Nadal, F., & Ducet, N. (1998). An improved mapping method of multisatellite altimeter data. *Journal of Atmospheric and Oceanic Technology*, *15*(2), 522–534. https://doi.org/10.1175/1520-0426(1998)015<0522:AIMMOM>2.0.CO;2

Lee, T., & McPhaden, M. J. (2008). Decadal phase change in large-scale sea level and winds in the Indo-Pacific region at the end of the 20th century. *Geophysical Research Letters*, *35*, L01605. https://doi.org/10.1029/2007GL032419

Levitus, S., Antonov, J. I., Boyer, T. P., Baranova, O. K., Garcia, H. E., Locarnini, R. A., et al. (2012). World ocean heat content and thermosteric sea level change (0–2000 m), 1955–2010. *Geophysical Research Letters*, *39*, L10603. https://doi.org/10.1029/2012GL051106

Lim, Y.-K., & Kim, K.-Y. (2007). ENSO impact on the space-time evolution of the regional Asian summer monsoons. *Journal of Climate*, *20*(11), 2397–2415. https://doi.org/10.1175/JCLI4120.1

Meyssignac, B., Becker, M., Llovel, W., & Cazenave, A. (2012). An assessment of two-dimensional past sea level reconstructions over 1950–2009 based on tide-gauge data and different input sea level grids. *Surveys in Geophysics*, *33*(5), 945–972. https://doi.org/10.1007/s10712-011-9171-x

Murtugudde, R., McCreary, J. P. Jr., & Busalacchi, A. J. (2000). Oceanic processes associated with anomalous events in the Indian Ocean with relevance to 1997–1998. *Journal of Geophysical Research*, *105*(C2), 3295–3306. https://doi.org/10.1029/1999JC900294

Nerem, R. S., Chambers, D. P., Choe, C., & Mitchum, G. T. (2010). Estimating mean sea level change from the TOPEX and Jason altimeter missions. *Marine Geodesy*, *33*(sup1), 435–446. https://doi.org/10.1080/01490419.2010.491031

Nidheesh, A. G., Lengaigne, M., Vialard, J., Izumo, T., Unnikrishnan, A. S., Meyssignac, B., et al. (2017). Robustness of observation-based decadal sea level variability in the Indo-Pacific Ocean. *Geophysical Research Letters*, *44*, 7391–7400. https://doi.org/10.1002/2017GL073955

Nidheesh, A. G., Lengaigne, M., Vialard, J., Unnikrishnan, A. S., & Dayan, H. (2013). Decadal and long-term sea level variability in the tropical Indo-Pacific Ocean. *Climate Dynamics*, *41*(2), 381–402. https://doi.org/10.1007/s00382-012-1463-4

Peltier, W. R. (2004). Global glacial isostasy and the surface of the ice-age Earth: The ICE-5G (VM2) model and GRACE. *Annual Review of Earth and Planetary Sciences*, *32*(1), 111–149. https://doi.org/10.1146/annurev.earth.32.082503.144359

Rao, S. A., & Behera, S. K. (2005). Subsurface influence on SST in the tropical Indian Ocean: Structure and interannual variability. *Dynamics of Atmospheres and Oceans*, *39*(1–2), 103–135. https://doi.org/10.1016/j.dynatmoce.2004.10.014

Rao, S. A., Behera, S. K., Masumoto, Y., & Yamagata, T. (2002). Interannual subsurface variability in the tropical Indian Ocean with a special emphasis on the Indian Ocean dipole. *Deep Sea Research Part II: Topical Studies in Oceanography*, *49*(7–8), 1549–1572. https://doi.org/10.1016/S0967-0645(01)00158-8

Ray, R. D., & Douglas, B. C. (2011). Experiments in reconstructing twentieth-century sea levels. *Progress in Oceanography*, *91*(4), 496–515. https://doi.org/10.1016/j.pocean.2011.07.021

Reynolds, R. W., Rayner, N. A., Smith, T. M., Stokes, D. C., & Wang, W. (2002). An improved in situ and satellite SST analysis for climate. *Journal of Climate*, *15*(13), 1609–1625. https://doi.org/10.1175/1520-0442(2002)015<1609:AIISAS>2.0.CO;2

Saji, N. H., Goswami, B. N., Vinayachandran, P. N., & Yamagata, T. (1999). A dipole mode in the tropical Indian Ocean. *Nature*, *401*(6751), 360–363. https://doi.org/10.1038/43854

Smith, T. M., Reynolds, R. W., Livezey, R. E., & Stokes, D. C. (1996). Reconstruction of historical sea surface temperatures using empirical orthogonal functions. *Journal of Climate*, *9*, 1403–1420. https://doi.org/10.1175/1520-0442(1996)009<1403:ROHSST>2.0.CO;2

Swapna, P., Ravichandran, M., Nidheesh, G., Jyoti, J., Sandeep, N., Deepa, J. S., & Unnikrishnan, A. S. (2020). Sea-level rise. In R. Krishnan, J. Sanjay, C. Gnanaseelan, M. Mujumdar, A. Kulkarni, S. Chakraborty (Eds.), *Assessment of climate change over the Indian region* (Chap. 9, pp. 175–189). Singapore: Springer. https://doi.org/10.1007/978-981-15-4327-2_9

Tai, C.-K. (1989). Accuracy assessment of widely used orbit error approximations in satellite altimetry. *Journal of Atmospheric and Oceanic Technology*, *6*(1), 147–150. https://doi.org/10.1175/1520-0426(1989)006<0147:AAOWUO>2.0.CO;2

Trenary, L. L., & Han, W. (2013). Local and remote forcing of decadal sea level and thermocline depth variability in the South Indian Ocean. *Journal of Geophysical Research: Oceans*, *118*, 381–398. https://doi.org/10.1029/2012JC008317

Vinayachandran, P. N., Iizuka, S., & Yamagata, T. (2002). Indian Ocean dipole mode events in an ocean general circulation model. *Deep Sea Research Part II: Topical Studies in Oceanography*, *49*(7–8), 1573–1596. https://doi.org/10.1016/S0967-0645(01)00157-6

Vinayachandran, P. N., Saji, N. H., & Yamagata, T. (1999). Response of the equatorial Indian Ocean to an unusual wind event during 1994. *Geophysical Research Letters*, *26*(11), 1613–1616. https://doi.org/10.1029/1999GL900179

Webster, P. J., Moore, A. M., Loschnigg, J. P., & Leben, R. R. (1999). Coupled ocean–atmosphere dynamics in the Indian Ocean during 1997–98. *Nature*, *401*(6751), 356–360. https://doi.org/10.1038/43848

Woodworth, P. L., & Player, R. (2003). The permanent service for mean sea level: An update to the 21stCentury. *Journal of Coastal Research*, 287–295.

Xie, S.-P., Annamalai, H., Schott, F. A., & McCreary, J. P. (2002). Structure and mechanisms of South Indian Ocean climate variability*. *Journal of Climate*, *15*(8), 864–878. https://doi.org/10.1175/1520-0442(2002)015<0864:samosi>2.0.co;2

Zhang, X., & Han, W. (2020). Effects of climate modes on interannual variability of upwelling in the tropical Indian Ocean. *Journal of Climate*, *33*(4), 1547–1573. https://doi.org/10.1175/JCLI-D-19-0386.1

Zuo, H., Balmaseda, M. A., Tietsche, S., Mogensen, K., & Mayer, M. (2019). The ECMWF operational ensemble reanalysis-analysis system for ocean and sea-ice: A description of the system and assessment. *Ocean Science Discussions*, *15*(3), 779–808. https://doi.org/10.5194/os-2018-154