Improving sea level reconstructions using non-sea level measurements

B. D. Hamlington,¹ R. R. Leben,¹ and K.-Y. Kim²

Received 12 June 2012; revised 11 September 2012; accepted 13 September 2012; published 31 October 2012.

¹Colorado Center for Astrodynamics Research, Department of Aerospace Engineering Sciences, University of Colorado Boulder, Boulder, Colorado, USA.
²School of Earth and Environmental Science, Seoul National University, Seoul, South Korea.

Corresponding author: B. D. Hamlington, Colorado Center for Astrodynamics Research, Department of Aerospace Engineering Sciences, University of Colorado Boulder, ECNT 320, 431 UCB, Boulder, CO 80309-0431, USA. (hamlingt@colorado.edu)

©2012. American Geophysical Union. All Rights Reserved.
0148-0227/12/2012JC008277

1. Introduction

[1] We present a new method for reconstructing sea level involving cyclostationary empirical orthogonal functions (CSEOFs). While we show results from a CSEOF reconstruction using basis functions computed from satellite altimetry and subsequently fit to tide gauge data, our focus is on how other ocean observations such as sea surface temperature can be leveraged to create an improved reconstructed sea level data set spanning the time period from 1900 to present. Basis functions are computed using satellite measurements of sea surface temperature, and using a simple regression technique, these basis functions are transformed to represent a similar temporal evolution to corresponding satellite altimeter-derived sea level basis functions. The resulting sea level and sea surface temperature basis functions are fit to tide gauge data and historical sea surface temperature data, respectively, to produce a reconstructed sea level data set spanning the period from 1900 to present. We demonstrate the use of this reconstructed data set for climate monitoring, focusing primarily on climate signals in the Pacific Ocean. The CSEOF reconstruction technique can be used to create indices computed solely from sea level measurements for monitoring signals such as the eastern Pacific (EP) El Niño–Southern Oscillation (ENSO), central Pacific (CP) ENSO, and Pacific Decadal Oscillation (PDO). The EP ENSO, CP ENSO, and PDO signals are all well represented in the CSEOF reconstruction relying solely on sea level measurements from 1950 to present; however, significant improvement can be made in reconstructing these signals during the first half of the twentieth century by including sea surface temperature measurements in the sea level reconstruction procedure.

an active research area [e.g., Smith, 2000; Chambers et al., 2002; Church et al., 2004; Church and White, 2006, 2011; Berge-Nguyen et al., 2008; Hamlington et al., 2011b; Ray and Douglas, 2011; Meyssignac et al., 2012a, 2012b]. To date, sea level reconstructions have relied solely on tide gauge measurements to extend the sea level data record beyond the beginning of the satellite altimeter era. Prior to 1950, however, the number of available tide gauges decreases rapidly. This makes it exceedingly difficult to reconstruct climate signals like El Niño–Southern Oscillation (ENSO) and the Pacific Decadal Oscillation (PDO), where there are few or no tide gauges in the respective regions of high variability. As a result of this, previous reconstructions have primarily been used to reconstruct GMSL prior to 1950, and published papers have contained only a limited discussion of the ability of reconstructions to capture actual climate signals. This deficiency in sea level reconstructions severely limits the possibilities for using sea level as a variable for climate monitoring and raises questions about conducting studies on interannual to decadal-scale variations in sea level prior to 1950.

[4] By finding a way to leverage other ocean variables such as sea surface temperature (SST), it is possible to create an improved sea level reconstruction from the 19th century to the present that better represents climate variability in the ocean. While the number of available tide gauges decreases rapidly before 1950, there are many in situ measurements of SST with reasonably good coverage of the ocean. Not only does this improve the reconstruction of large-scale climate signals, incorporating SST measurements allows for more basis functions to be used in the reconstruction and thereby more sea level variability to be captured by the reconstruction. Here, we show initial results combining sea level and SST measurements to reconstruct sea level from 1900 to 2010 for the Pacific Ocean region. This reconstruction will be referred to as a bivariate sea level reconstruction throughout the paper. For comparison, we also compute a sea level reconstruction relying only on sea level anomaly (SLA) measurements from 1900 to 2010 and a sea level reconstruction using only in situ SST measurements over the same time period. While one could argue that reconstructing sea level in the Pacific Ocean is relatively easy as a result of the large-scale signals like ENSO and the PDO that dominate the variability, we believe this is a good region to begin the discussion of the bivariate technique and demonstrate the reconstruction of climate signals in sea level back to 1900, something that has not been done previously. This reconstruction technique relies on the use of cyclostationary empirical orthogonal functions (CSEOFs) as basis functions for the reconstruction and the regression analysis introduced by Kim and North [1997] and Lim and Kim [2007]. Additional climate variables could be included to improve and extend the reconstructions, however, we will only use sea level and SST data in the reconstruction presented here. Section 2 describes the CSEOF technique, and section 3 outlines the sea level and SST data that is used to perform the reconstruction. In section 4 we provide details on CSEOFs, the regression technique, and the reconstruction procedure. Section 5 contains results from the initial bivariate sea level reconstruction from 1900 to 2012. Finally, in section 6, we summarize our results and discuss how this technique represents an improvement over previous reconstructions and how it can be used to further improve sea level reconstructions in the future.

2. CSEOFs

[5] Previous reconstructions, both SST and sea level, have generally relied on empirical orthogonal functions (EOFs) to form the basis for the reconstruction. When compared to CSEOFs, however, EOFs have characteristics that make them suboptimal for use as basis functions for sea level reconstruction. EOFs enforce a stationarity on the spatial variability. A single spatial map defines the basis function, and the reconstruction procedure simply computes the amplitude modulation of this map through time. Given the evidence that many signals in geophysical data are cyclostationary, CSEOFs provide significant advantages over EOFs when dealing with signals such as modulated annual cycle and ENSO signals. Although EOFs are able to represent cyclostationary features through a superposition of multiple modes, CSEOFs are able to explain cyclostationary signals in a single mode, increasing the opportunity for interpretability.

[6] The decomposition of data in terms of a set of basis functions is often very useful in understanding the complicated response of a physical system. By decomposing into less complicated patterns, it may be easier to understand and shed light into the nature of the variability in a data set. While theoretical basis functions have been studied extensively, exact theoretical basis functions are very difficult to find and in general, computational basis functions are sought instead. Perhaps the simplest and most common computational basis functions are EOFs. Consider a simple system defined by

\[
T(r, t) = \sum_i LV_i(r) PCTS_i(t),
\]

where \(LV(r)\) is a physical process (termed to be the loading vector) modulated by a stochastic time series \(PCTS_i(t)\), which is called the principal component time series. Each loading vector and principal component time series pair represents a single EOF mode. As mentioned above, however, physical processes and the corresponding statistics are time-dependent. Representing the data with stationary EOFs can lead to erroneous and difficult interpretation of the data [Dommenger and Latif, 2002].

[7] Kim et al. [1996], Kim and North [1997], Kim and Wu [1999], and Kim and Chung [2001] introduced the concept of cyclostationary empirical orthogonal function (CSEOF) analysis to more compactly capture the time-varying spatial patterns and longer-timescale fluctuations present in geophysical signals. The significant difference between CSEOF and EOF analysis is the LVs’ time dependence, which allows the spatial pattern of each CSEOF mode to vary in time, with the temporal evolution of the spatial pattern of the CSEOF LVs constrained to be periodic with a selected “nested period.” In other words, the system is defined as

\[
T(r, t) = \sum_i LV_i(r, t) PCTS_i(t),
\]

\[
LV(r, t) = LV(r, t + d),
\]
CSEOF analysis minimizes mode mixing, which is a common problem with EOF decomposition. When mode mixing occurs, the annual cycle frequently spreads across several modes, which is one reason the signal is usually removed from the data by some other means. Recent studies, however, have demonstrated the efficacy of using CSEOFs to extract robust modes representing the modulated annual cycle (MAC) and ENSO variability [Trenberth et al., 2005; Hamlington et al., 2011a, 2011b]. This leads to robust estimates of the MAC or ENSO variability from satellite altimetry data without affecting signals associated with other ocean variability. For further details on CSEOFs and the procedure for computing CSEOFs, the reader should refer to Kim and North [1997] in which a detailed description of the computation of CSEOFs is provided. Additionally, while seemingly similar in purpose, CSEOFs and extended EOFs are quite different mathematically, and the reader should refer to Kim and Wu [1999] and Seo and Kim [2003] for a full discussion of the differences between the two techniques. 

By fitting CSEOFs in place of EOFs to tide gauges to reconstruct sea level for the 1950–2010 time period [Hamlington et al., 2011b], we have found that it is possible to create an alternate and, we believe, improved reconstruction to those based on EOFs. Our motivation for using CSEOFs, as described by Hamlington et al. [2011b], in place of EOFs is fourfold: (1) EOFs are not an optimal basis for nonstationary signals with nested oscillations that are undergoing low-frequency oscillation; (2) CSEOFs account for both the high- and low-frequency components of the annual cycle and do not require the removal of the annual signal from the satellite altimetry nor the tide gauge records before reconstruction; (3) specific signals, such as those relating to the MAC and ENSO can be reconstructed individually using CSEOFs with little mixing of variability between modes; and (4) the reconstruction procedure using CSEOFs spans data gaps smaller than the nested period and fits a larger window of data, allowing for the use of fewer historical data to obtain meaningful results. In the work by Hamlington et al. [2011b] the impact of tide gauge selection, editing, and weighting on the fidelity of the reconstructions was also investigated. Error analyses were performed using Monte Carlo simulations.

3. Data

The bivariate reconstruction we present here relies on observations from two ocean climate variables: sea level and sea surface temperature. Both in situ and satellite data are used in the reconstruction procedure. In situ data provide the historical measurements required for the reconstruction back in time, while the satellite data provides the basis functions fit to the historical data. The respective data sets are described in the following sections.

3.1. Sea Level Data

3.1.1. Tide Gauge Data

The central sea level data set we use for the period 1900 through 2010 is monthly mean sea level records gathered from the data archive at the Permanent Service for Mean Sea Level (PSMSL). We use only the Revised Local Reference (RLR) data, which are measured sea level at each site relative to a constant local datum over the complete record. At present, we have not included the metric data offered by PSMSL as they can have substantial unknown datum shifts and their use in time series analysis is generally not recommended.

Since the focus of this paper is on the bivariate reconstruction technique and the reconstruction of large-scale climate signals, we have selected tide gauges for use in our analysis with very lenient editing criteria. In short, we have used every available tide gauge in the Pacific Ocean region from 40°S to 60°N and 120°E to 75°W. We choose such lenient editing criteria due to the larger fitting window afforded by the CSEOF technique which makes the reconstruction relatively insensitive to a small number of poor tide gauges. We did edit and remove data found in tide gauge records undergoing a month-to-month change of greater than 250 mm. It is without doubt that some included tide gauges may warrant more careful consideration in the future, but we sought to avoid selectively editing tide gauges and instead show that this reconstruction method is robust to data outliers and can be used even without expert editing of individual tide gauges. While it may be appropriate in future studies, we also made no attempt to avoid regional clustering when selecting tide gauge records, resulting in a large cluster of tide gauges in Japan.

We upsampled the monthly tide gauge data to weekly sampling using linear interpolation to match the 1 week temporal resolution of the altimeter-derived CSEOFs, which will be discussed more in the next section. We found the nearest grid point for each tide gauge as the basis functions obtained from the satellite altimetry are on a half-degree resolution spatial grid. Finally, the available tide gauge records were averaged to produce a single time series if there were multiple tide gauges associated with a single spatial grid point. After the editing procedure, 427 tide gauges remained for use in our reconstruction.

The PSMSL sea level data are relative sea level; therefore the records must be corrected for the ongoing glacial isostatic adjustment (GIA). We use the ICE-5G VM2 model [Peltier, 2004] to perform the GIA correction. Since an inverted barometer correction is applied to the satellite altimetry data, the sea level measurements from the tide gauges were corrected using an inverted barometer response of sea level to atmospheric loading based on the pressure fields from the NCEP/NCAR reanalysis. It should be noted, however, that neither of these corrections were found to have a noticeable effect on the resulting sea level reconstruction.

3.1.2. Satellite Altimetry Data

CSEOF sea level basis functions for our reconstruction were estimated from the AVISO quarter-degree resolution multiple altimeter product based on satellite altimeter measurements spanning 1993–2011 collected by the TOPEX/POSEIDON, ERS-1 and 2, Geosat Follow-On, Envisat, Jason-1 and OSTM satellites. This updated and reprocessed gridded data product, which was released in December 2011, was created using the delayed time Ssalto/DUACS multimission altimeter data processing system with improved homogenous corrections and intercalibration applied to the entire data record. Global crossover minizations and local inverse methods are used to derive intercalibrated highly accurate along track data that are...
referenced to a consistent mean. The along track data were then merged through a global space-time objective mapping technique that takes into account correlated noise \cite{LeTraon1998}.

[15] We applied very little additional processing other than removing the mean and a linear least squares fit from the time series at each spatial grid point. A CSEOF decomposition of the satellite altimetry data is not able to extract the change in mean sea level into a single mode. It is therefore necessary to remove mean sea level from the satellite altimetry data before computing the basis functions to avoid putting low-frequency power into each CSEOF mode. Although our technique does result in the removal of the spatial pattern of sea level trends, it is unlikely that the regional distribution of sea level trends over the past two decades is the same as that since 1900 and it is, therefore, unwise to force this stationary pattern upon the reconstruction. Removing the trend from each grid point does not significantly affect the ability of the reconstruction to capture decadal timescale signals. Using the study of Tai \cite{Tai1989}, the percentage of signal reduction at various periods caused by removing the linear trend from a 17 year record can be computed. Signals with periods of approximately 10 years undergo an RMS signal reduction of less than 5% and even signals at 20 year periods are reduced by only 30%. Linear detrending of the altimeter data before computing the CSEOFs should therefore have little effect on decadal-scale variations and our ability to capture them in the reconstruction. The seasonal signal was not removed from the time series prior to computing the CSEOFs. We do not use the CSEOF mode representing the seasonal signal in the reconstruction; we find that there is little difference in the reconstruction regardless of whether or not the annual cycle is removed from both the historical data and satellite altimetry data. Before CSEOF decomposition, the data was weighted using the square root of the cosine of latitude to yield area-weighted variance decomposition.

3.2. SST Data

3.2.1. Historical Measurements

[16] The historical SST measurements come from the International Comprehensive Ocean-atmosphere Data Set (ICOADS) covering the time period from 1900 to present. We use a gridded monthly data set of superobservations \cite{Smith1996} formed by averaging observations on to a 2° latitude \times 2° longitude. A monthly climatology was computed from 1960 to 1980 and removed from the data covering the period from 1900 to present in order to remove the seasonal signal from the observations. In future reconstructions, the seasonal signal will likely be retained, as the CSEOFs are capable of capturing and explaining the seasonal variability. As with the tide gauges, we only use the observations taken in the region from 40°S to 60°N and 120°E to 75°W. The historical bias adjustment for the data collected before 1942 outlined by Smith and Reynolds \cite{Smith2002} and updated by Smith and Reynolds \cite{Smith2004} is also applied to the data. Other bias adjustments are available, but for the purposes of this paper, we only apply the one provided by Smith and Reynolds \cite{Smith2004}. We do not perform any additional quality control beyond that already completed prior to releasing ICOADS. Additional quality control may be appropriate in the future, but the CSEOF reconstruction technique is relatively insensitive to outliers and the use of averaged observations should limit the influence of extreme outliers.

3.2.2. Satellite Measurements

[17] CSEOF SST basis functions were computed from the weekly 1° grid optimal interpolation SST (OISST) data set covering the period from 1993 to 2011. The analysis uses in situ and satellite-measured SST data in addition to SST simulated by sea ice cover. Since the CSEOF analysis is able to separate the seasonal signal into a single mode, we do not remove it from the data prior to analysis. As for the satellite altimetry-derived CSEOF basis functions, we have removed the linear trend from each spatial grid point prior to performing the CSEOF analysis. Again, only the region from 40°S to 60°N and 120°E to 75°W is used for computation of the CSEOF basis functions.

4. Methods

[18] We have developed a method for performing a bivariate reconstruction of sea level by using CSEOFs and a simple regression computation to relate the SST patterns to the SLA basis functions. In this section, we will describe these methods and outline how the reconstruction is computed over the period from 1900 to 2010.

4.1. Regression

[19] As mentioned above, CSEOF LVs have a temporal dependence, which allows the spatial pattern of each CSEOF mode to vary in time, with the temporal evolution of the spatial pattern of the CSEOF LVs constrained to be periodic with a selected “nested period.” The PCTS describes the change in amplitude of the LV over time. Using a regression technique similar to that introduced by Lim and Kim \cite{Lim2007}, it is possible to create SST LVs that have the same PCTS (or alternatively, the same temporal evolution) as corresponding SLA LVs. Two sets of CSEOFs computed independently for two variables do not necessarily behave similarly, but this simple regression allows for the possibility of finding physical and dynamical consistency. The goal of the reconstruction procedure is to estimate the PCTS back through time by simultaneously fitting SLA and SST patterns with identical PCTS over the satellite record, to historical in situ measurements so that the SLA LVs and estimated PCTS can be used to reconstruct time-space gridded maps of the sea level record.

[20] The following four steps outline the procedure for performing a bivariate reconstruction of sea level using SLA and SST data.

[21] 1. Perform a CSEOF decomposition on satellite-measured SLA and SST data over a common time period, in this case 1993 to present.

[22] 2. Regress all SST PCTS on each individual SLA PCTS to compute a set of regression coefficients, \( \alpha \):

\[
PCTS_{SLA}(t) = \sum_i \alpha_i \PCTS_{SST,i}(t) + \epsilon(t). \tag{3}
\]

[23] 3. Use the regression coefficients computed in step 2 along with original SST LV, \( \text{LV}_{SST} \), to form SST spatial patterns, \( \text{LVR}_{SST}(r,t) \), with amplitude fluctuations described by \( \text{PCTS}_{SLA}(t) \):

\[
\text{LVR}_{SST}(r,t) = \sum_i \alpha_i \text{LV}_{SST,i}(r,t). \tag{4}
\]
[24] 4. Simultaneously fit LVR$_{SST}$ and LVR$_{SLA}$ to historical SST and tide gauge measurements, respectively, to reconstruct PCT$_{SST}$ and PCT$_{SLA}$ back through time (further details in section 4.2).

[25] Assuming the regression performed in step 3 is of sufficiently high quality, the result will be two sets of LVs, one for SLA and one for SST, that have the same PCTS over the calibration period (1993–2011, in this case). By using both sets of LVs simultaneously to reconstruct sea level, we assume that the regression relationship computed from 1993 to 2011 holds true over the entire span of the reconstruction. Additionally, the regression relationship between SST and SSH is only as good as the relationship between the SST and the vertical temperature profile. This could be remedied by using data from the full water column in our reconstruction. This is difficult, however, given the lack of data sets to both compute basis functions and to subsequently fit these basis functions to. These assumptions about the regression are not dissimilar from the assumption made when performing any reconstruction that the basis functions computed during the calibration period are applicable and relevant over the entire reconstruction. It should be noted that this will not be the case at certain times and locations during the reconstruction period, but in general, the inclusion of SST greatly improves the reconstruction of sea level, as seen through the results presented in this paper.

4.2. Reconstruction

[26] Once the regression has been completed for all CSEOF modes, sea level is reconstructed by simultaneously fitting the SLA LVs and the regressed SST LVs to historical measurements. The SLA LVs are fit to the tide gauge data spanning the period from 1900 to 2010. The SST regressed LVs are fit simultaneously to the ICOADS 3° resolution SST anomalies, again spanning the time period from 1900 to 2010. We follow a procedure very similar to that described by Kaplan et al. [2000], Church et al. [2004] and Hamlington et al. [2011b] to solve for the reconstructed PCTS from 1990 to 2010. With the bivariate nature of this reconstruction, however, the weighting used in the reconstruction requires careful consideration. Previous SST and sea level reconstructions have generally relied on a combination of truncation and measurement errors for weighting. While we take a similar approach, the SLA and SST components of the reconstruction must also be weighted relative to each other. In general, there are many more historical SST measurements than tide gauge measurements that are being fit at each point of time. After analyzing the effect of this relative weighting on the reconstruction, we have opted to equally weight the SLA and SST portions of the reconstruction. In other words, the sum of the weights for the SLA component is set to be equal to the sum of the weights for the SST component at each point in time. This generally results in a single tide gauge being weighted more heavily than a single historical SST measurement. We conduct historical grid tests as will be discussed in section 5.3 to test different weighting schemes, and find that the equal weighting chosen here is suitable for the overall time period of the reconstruction.

[27] Estimating GMSL using reconstruction techniques is not a trivial task. Christiansen et al. [2010] discussed the difficulties of estimating GMSL using EOF reconstruction techniques, albeit using model data instead of satellite altimetry and tide gauge data. Although CSEOF basis functions describe the cyclostationary variability in sea level well, no single CSEOF mode captures the secular trend pattern because of the short length of the satellite altimetry record to date. EOF analysis of the satellite altimetry record has similar difficulties in extracting the secular trend from such a short record. Nevertheless, even if the CSEOF analysis was capable of extracting the secular trend from the AVISO data, we would not assume that the resulting spatial pattern determined for the period from 1993 to present is stationary over the entire 110 year time period. Church et al. [2004] approximate the trend in their reconstruction by introducing a constant basis function that is fit along with the other EOF basis functions. In the work by Hamlington et al. [2011b], we take a different approach by separating the estimation of GMSL from the reconstruction of the individual modes. In this paper, we make no attempt to account for mean sea level, and simply note that an estimate of mean sea level could be made using the procedure outlined by Hamlington et al. [2011b]. The focus of this paper is primarily on reconstructing the regional sea level variability, and since we are only reconstructing the Pacific Ocean region, a discussion of trends in GMSL as provided in other reconstruction papers is not applicable. Any trends shown in this paper are relative to the background trend over the Pacific Ocean region. For a global bivariate reconstruction, we would adopt an approach similar to Hamlington et al. [2011b] to account for GMSL in our reconstruction.

[28] We also do not explicitly address the errors in our reconstruction in this paper, having elected to treat the reconstruction error in a manner similar to previous studies [e.g., Church et al., 2004; Hamlington et al., 2011b; Meyssignac et al., 2012a, 2012b]. Errors can be estimated as a combination of instrument errors and the error resulting from truncating the basis functions estimated from the AVISO data set. Similar to previously published papers, we do not provide values for the error estimated in this way as such a discussion does not significantly impact the results of this study, but emphasize that such an analysis is possible and available. More exhaustive error analysis of the estimate of GMSL is generally included in published studies on sea level reconstructions, but since GMSL is not addressed here, such a discussion is omitted. The reader is referred to Kaplan et al. [1998, 2000], Smith et al. [1996, 2008] and Church et al. [2004] for a discussion on how errors are generally handled for sea level and SST reconstructions.

[29] After computing the reconstruction over the period from 1900 through 2010, the reconstructed PCTS can be combined with the satellite altimetry-derived LVs to create a sea level data set with the temporal span of the tide gauges and historical SST data and spatial coverage of the satellite altimetry in the Pacific Ocean. In this paper, we provide the initial results from a bivariate reconstruction incorporating SLA and SST measurements, and for the first time demonstrate efficacy in capturing climate variability in a sea level reconstruction prior to 1950.

5. Results

5.1. Regression Analysis

[30] An initial combined SLA-SST reconstruction has been produced for the Pacific Ocean region over the period...
from 1900 to 2010. For comparison, we have also computed a reconstruction relying only on tide gauges measurements of sea level for historical observations, and an additional reconstruction using only historical SST measurements.

[31] Using the methods described in sections 2 and 4, a CSEOF decomposition of the Pacific Ocean region using a 1 year nested period was completed for the weekly 1/4° resolution multiple altimeter 1993–2011 AVISO merged sea level anomaly data set. In the future, it may be appropriate to use other nested periods to specifically target a lower-frequency climate signal. However, for the purposes of this study a 1 year nested period is sufficient to capture the variability associated with the large-scale signals in the Pacific Ocean [Kim and Chung, 2001; Hamlington et al., 2011b]. A similar CSEOF analysis was performed on the weekly 1° resolution NOAA OISST data set spanning 1993–2011. This resulted in CSEOF modes having 52 spatial maps (one for each week of the year) and a PCTS spanning 1993–2011.

[32] The regression analysis outlined in section 4.1 was then applied to obtain SST LVs with the same PCTS as corresponding SLA LVs. The first 11 SST CSEOF modes were used to perform the regression with each SLA CSEOF PCTS (in other words, i = 11 in equations (3) and (4)) Figure 1 shows the result of the regression for the CSEOF mode that contains the variability associated with the eastern Pacific ENSO signal. This is the second CSEOF mode for both the SLA and SST decompositions (the seasonal signal is the dominant mode in both cases). As can be seen from Figure 1c, the regressed SST PCTS agrees very well in terms of both correlation and scale with the corresponding SLA PCTS. The information regarding the physical units is contained within the LVs shown in Figures 1a and 1b. Note, only one of the 52 weekly LVs (here, the first week of January) are shown in the interest of space. Similarly, in Figure 2c, the regression for third CSEOF mode is shown to have excellent agreement in the PCTS once the regression is computed. This same regression analysis was performed for the first 11 SLA CSEOF modes, which explain 80% of the variance once the modes explaining the seasonal signal are removed. While in previous CSEOF reconstructions additional modes explaining more of the variance were used, we limit the variance explained to 80% to avoid overfitting. Correlations between the SLA PCTS and regressed SST PCTS are greater than 0.8 for each of the 11 modes, with the first 5 modes having correlations greater than 0.95. It should be noted that historical grid testing (see section 5.3 for details) was conducted to determine if using 11 predictors in the regression would lead to overfitting and a subsequent degradation in the reconstruction. Given the low level of noise on the CSEOF PCTS (both SLA and SST), overfitting should not be expected, and this was confirmed by the historical grid sampling tests. After the regression is completed, the regressed SST LVs are fit simultaneously with the SLA LVs to produce PCTS that explain sea level variability over the time period from 1900 to 2010.

Figure 1. Eastern Pacific ENSO (first basis function) (a) CSEOF SLA LV and (b) the regressed SST LV from the first week in January with (c) the respective PCTS shown.
5.2. Reconstructed Sea Level PCTS 1900–2010

Three different sea level reconstructions were computed for the period from 1900 to 2010: (1) using only SLA measurements, with satellite altimeter-derived basis functions fit to tide gauge data; (2) using only SST measurements, with regressed SST basis functions fit to historical SST measurements; and (3) using the bivariate approach, simultaneously fitting SLA and SST basis functions to their respective historical measurements. We do not include the seasonal signal in the reconstruction, but emphasize that it is possible and would improve the reconstructions given the qualities of the CSEOF analysis. Figure 3 shows the reconstructed PCTS for the first four modes from 1900 to 2010. Mode 1 represents the eastern Pacific ENSO signal. Comparison of the PCTS from the bivariate reconstruction with the reconstruction made using only the tide gauge data shows excellent agreement after 1950; however, the tide gauge-only sea level reconstruction has much larger amplitude variations in the PCTS in the early twentieth century when the number of available tide gauges decreases dramatically. This is also prevalent especially in the higher-order modes that contain smaller spatial scales indicating that the oscillations are likely an artifact of both overfitting the limited amount of available tide gauge data and not having sufficient tide gauge coverage to accurately estimate the amplitude of the variability. The inclusion of the SST measurements limits this effect prior to 1950 without significantly impacting the results after 1950.

5.3. Historical Grid Tests

The results shown in the previous section underscores the importance of including other types of data into a sea level reconstruction prior to 1950, and may be the reason why previous publications on sea level reconstructions have very limited discussion of any resolved sea level climate signals when extending the reconstruction period back through time. To test whether or not the historical sampling is sufficient to capture sea level variability, we perform tests where the distribution of historical sea level and SST measurements in 1920, 1940, 1960 and 1980 is used during the altimeter time period from 1993 to 2010 to reconstruct sea level from coincident tide gauge sea level and in situ SST measurements. In other words, we only use the handful of tide gauges and SST measurements at locations that were also available during past decades to perform the 1993–2010 reconstruction and subsequently compare to the AVISO data. This allows us to validate the reconstructed PCTS obtained from very limited historical sampling by comparing to the PCTS computed directly from the CSEOF decomposition of the satellite altimeter data. As seen in Table 1, the reconstruction relying only on tide gauges performs poorly in 1920 and 1940, with significant improvement seen in 1960 and 1980. Furthermore, Figure 4 shows the results of the reconstruction of the first four modes for the 1920 sampling. The reconstructed PCTS for the first four modes using the bivariate approach with both tide gauge and historical SST data compares very
favorably to the altimeter-derived PCTS. Using only tide
gauge measurements, however, there is very poor agreement
between the reconstructed PCTS and the altimeter-derived
PCTS. In every sampling test the reconstruction is improved
by including SST measurements. These results suggest
that before 1950, it is difficult to capture sea level variability
with tide gauges alone, and including SST measurements
improves the reconstruction regardless of the sampling pro-
vided by the tide gauges. There is likely a sampling density of
tide gauges at which including SST measurements would
likely degrade the reconstruction of sea level, but the avail-
able tide gauges do not provide such a sampling density over
the past 110 years.

It should be noted that we have not conducted the
cross-validation testing described by Smith [2000] that
incorporates the exclusion of data for the time being recon-
structed. The computational burden and complexity of
CSEOFs make full cross-validation testing challenging.
The historical grid tests performed here may tend to over-
estimate the performance of the reconstruction and the
amount of variability that can be used in the reconstruction.
In the future, cross-validation tests as defined by Smith
[2000] would provide a better understanding of how the
reconstruction is doing in the past. For the purposes of
this paper, however, we believe the historical grid tests are
sufficient to demonstrate the improvement seen by using the
bivariate approach.

### Table 1. Percentage Difference in RMSE Relative to the
Full Bivariate Sea Level Reconstruction for Several Different
Reconstructions

<table>
<thead>
<tr>
<th>Reconstruction Type</th>
<th>Year Specifying Spatial Sampling</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1920</td>
</tr>
<tr>
<td>CSEOF TG-only reconstruction</td>
<td>58.8</td>
</tr>
<tr>
<td>CSEOF SST-only reconstruction</td>
<td>3.15</td>
</tr>
<tr>
<td>Combined TG and SST reconstruction</td>
<td>2.22</td>
</tr>
</tbody>
</table>

*Percent equals \([\text{RMSE}_{\text{sub}} - \text{RMSE}_{\text{full}}]/\text{RMSE}_{\text{full}}\). RMSE is computed by
performing a sea level reconstruction over the period from 1993 to 2010 with
subsampled data from the current observational record and comparing the
resulting Pacific sea level reconstruction to the AVISO satellite altimetry
data. The 1993–2010 “historical” data used in the reconstruction has the
spatial coverage of the historical observations from the specified year. Using
the 1920 case as an example, only data from modern tide gauge and SST
observations from the same the tide gauge and SST locations sampled in
1920 are used to reconstruct sea level over the satellite altimetry time period.
We have also conducted tests to determine the best weighting of the SLA and SST components of the reconstruction. As seen in Table 2, we have tested two different time periods (1920 and 1980), and three different weighting schemes. For the 1920 sampling, the best result is found when weighting SST twice as heavily as SLA. For the 1980 sampling, the best result is found when using equal weighting. This suggests that as the tide gauge sampling gets worse back through time, the SST should be weighted more heavily. The overall effect of the weighting, however, appears to be minor.

Although a comparison may be appropriate in the future, we have not conducted tests to compare an EOF and CSEOF sea level reconstructions from 1900 to 2010. Recent publications on the topic of EOF reconstructions focus on GMSL and have not extensively discussed the reconstruction of actual climate signals before 1950. Furthermore, Ray and Douglas [2011] describe the inability of their reconstruction to capture the ENSO variability in 1940 with their EOF technique and edited tide gauge data set. Since our focus here

Table 2. Percentage Difference in RMSE Relative to the Full Bivariate Sea Level Reconstruction for Several Different Reconstructions Using Different Weighting for SLA and SST Components

<table>
<thead>
<tr>
<th>Subsampling Period</th>
<th>0.5</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1920</td>
<td>2.06</td>
<td>2.22</td>
<td>2.25</td>
</tr>
<tr>
<td>1980</td>
<td>0.98</td>
<td>0.38</td>
<td>0.76</td>
</tr>
</tbody>
</table>

*C*The historical data available for the reconstruction is subsampled to have the spatial coverage of the listed year.
is on improving the CSEOF reconstruction of climate signals over the past century and not on GMSL, we do not provide a detailed comparison of the two methods. The set of tide gauges used in our study would also not be suitable for an EOF reconstruction, particularly before 1950, as the EOF reconstruction is more sensitive to outliers and inaccurate historical data. Additionally, a bivariate approach with EOFs would be difficult given the mixing of modes and the increased number of basis functions required to perform the analysis. For all these reasons, it is difficult to make a direct comparison of the CSEOF bivariate sea level reconstruction to an EOF sea level reconstruction. The analysis described by Ray and Douglas [2011] and Hamlington et al. [2011b] has more information regarding the advantages and disadvantages of EOF and CSEOF reconstruction methods. In a future study, EOF and CSEOF reconstruction techniques should likely be studied and compared directly in an idealized setting using climate model output.

5.4. Climate Signals in Sea Level From 1900 to 2010

[39] To illustrate the ability to reconstruct ocean climate variability during the entire record, we compare sea level based climate indices to existing ocean climate indices. Figures 5, 6, and 7 compare indices derived from the combined tide gauge–SST sea level reconstruction with widely cited indices used for monitoring the eastern Pacific (EP) ENSO, the central Pacific (CP) ENSO, and the Pacific Decadal Oscillation (PDO). In Figure 5, the Extended Multivariate ENSO Index (MEI.ext) from Wolter and Timlin, 2011 is plotted relative to the reconstructed PCTS for the CSEOF mode shown in Figure 2. The correlation over the period from 1900 to 2010 is found to be 0.86. Figure 6 shows

Figure 5. Comparison of reconstructed PCTS for the eastern Pacific ENSO mode for a tide gauge–only reconstruction (blue) and a bivariate reconstruction (red) using both tide gauge and historical SST measurements. The Extended Multivariate ENSO Index (MEI.ext) is shown for comparison.

Figure 6. Comparison of an El Niño Modoki Index (EMI) computed for a tide gauge–only reconstruction (blue) and a bivariate reconstruction (red) using both tide gauge and historical SST measurements. The EMI computed from the ERSST reconstructed SST data set (black) is shown for comparison.
the El Niño Modoki Index (EMI), used for monitoring the El Niño Modoki or CP ENSO [Ashok et al., 2007], computed from the ERSST reconstructed SST data set and from our tide gauge–SST sea level reconstruction. Again, the correlation from 1900 to 2010 is very high with a value of 0.64. Finally, as seen in Figure 7, a sea level based PDO index (as defined by Cummins et al. [2005]) is computed from the multivariate sea level reconstruction introduced here and compared to the SST PDO index from Mantua et al. [1997] available at http://jisao.washington.edu/pdo/PDO.latest. The correlation between the two is found to be 0.58. These high correlations highlight the ability of reconstructions to capture large-scale ocean variability and climate signals, on timescales ranging from interannual to decadal, not only from 1950 onward, but also over the full record from 1900 to 2010. This degree of efficacy has not been demonstrated by any published sea level reconstruction to date.

Finally, to summarize and to further demonstrate the added benefit of incorporating SST into the sea level reconstruction, we compute correlations for the EP ENSO, CP ENSO and PDO using the three different reconstructions. We compute correlations for the period from 1900 to 2010 and also for the period starting in 1950. As seen in Table 3, in every case the combined reconstruction provides an improvement over a single-variable reconstruction of sea level. This suggests that not only does SST help improve sea level reconstructions, but also that tide gauge measurements could potentially be used to improve SST reconstructions.

### 5.5. Regional Sea Level Trend Patterns

In addition to computing and comparing climate indices, we can also use sea level reconstructions to investigate the spatial distribution of sea level trends for the Pacific Ocean. For direct comparison to the AVISO data, we compute trend maps for a tide gauge–only reconstruction, an SST-only reconstruction, and a bivariate reconstruction over the period from 1993 to 2010. The spatial patterns for each reconstruction are shown in Figure 8. Note, since we are not accounting for mean sea level (see section 4.2), each of these maps are relative to the background sea level trend over the Pacific Ocean region shown. The trends computed from each reconstruction show good agreement with the trends obtained directly from the altimetry-derived AVISO maps. Encouragingly, as seen in Figure 8c, we are able to reproduce much of the spatial variation in the sea level trends from 1993 to 2010 without any sea level measurements, using only the regressed SST LVs fit to the historical SST measurements. This strengthens the argument that sea level can still be accurately reconstructed prior to 1950 when SST measurements heavily outnumbered tide gauge measurements.

### Table 3. Comparison of the Correlation of the Multivariate ENSO Index (MEI), ENSO Modoki Index (EMI), and the Pacific Decadal Oscillation (PDO) to Similar Indices Computed From Sea Level Reconstructions Based on Fitting to Sea Level Measurements From Tide Gauge Data Only (TG-Only), In Situ SST Measurements Only (SST-Only), and Both Measurements (TG and SST)*

<table>
<thead>
<tr>
<th>Climate Index</th>
<th>MEI</th>
<th>EMI</th>
<th>PDO</th>
</tr>
</thead>
<tbody>
<tr>
<td>TG-only</td>
<td>0.77 (0.79)</td>
<td>0.92 (0.95)</td>
<td>0.34 (0.45)</td>
</tr>
<tr>
<td>SST-only</td>
<td>0.84 (0.89)</td>
<td>0.92 (0.94)</td>
<td>0.58 (0.76)</td>
</tr>
<tr>
<td>TG and SST</td>
<td>0.86 (0.91)</td>
<td>0.95 (0.97)</td>
<td>0.64 (0.80)</td>
</tr>
</tbody>
</table>

*In all cases, the indices based on the combined multivariate sea level reconstruction using TG and SST measurements exhibited the best correlations. Values in parentheses are the correlation coefficients after 1 year smoothing is applied. For the (MEI), see K. Wolter, Multivariate ENSO index (MEI), 2010, Earth System Research Laboratory, NOAA, http://www.esrl.noaa.gov/psd/people/klaus.wolter/MEI/.)
Additionally, with a reconstruction that can better capture and explain the sea level variability in the Pacific Ocean, we can compute a trend map over the period from 1900 to 2010, as seen in Figure 9. Shown here are the results for the tide gauge–only sea level reconstruction and the bivariate reconstruction. The two spatial patterns are in good agreement, although the magnitude of the trends for the tide gauge–only reconstruction are generally larger, perhaps suggesting some degree of overfitting.

6. Discussion and Conclusion

Sea level is an essential variable for climate monitoring. Finding a way to more accurately estimate sea level variability in the past will improve our ability to understand...

Figure 8. Regional sea level trends from 1993 to 2010 from (a) the AVISO data set, (b) the tide gauge–only reconstruction, (c) the SST-only reconstruction, and (d) the bivariate reconstruction.

Figure 9. Regional sea level trends from 1900 to 2010 from (a) the tide gauge–only reconstruction and (b) the bivariate reconstruction.
how the current state of the ocean compares to the state of the ocean in the past. Previous sea level reconstructions have focused primarily on GMSL prior to 1950 as a result of the limited tide gauge availability toward the start of the twentieth century. Through our historical grid testing and as seen in Figure 4, even capturing a signal as large as eastern Pacific ENSO is a challenge given the tide gauges available in 1920. Ray and Douglas [2011] confirmed this result, in part, through the poor representation of the 1940/1941 ENSO in their reconstruction. Using the bivariate approach to reconstruct sea level presented here, however, we have the opportunity to study sea level variability back to the beginning of the twentieth century and beyond, something that has not been possible to date. This results from the use of CSEOFs, which better capture sea level variability and form a more appropriate set of basis functions for a sea level reconstruction when compared to EOFs.

[43] With the application of a simple regression technique, SST LVs with the same temporal evolution as corresponding SLA LVs can be created and fit simultaneously to improve the reconstruction when tide gauge data is particularly sparse. We have demonstrated the use of such a bivariate reconstruction through the comparison to common climate indices. The bivariate reconstruction has correlations of 0.86, 0.64 and 0.58 over the period from 1900 to 2010 when compared to commonly used climate indices for the EP ENSO, CP ENSO and PDO, respectively. These correlations represent a significant improvement over similarly computed correlations for a sea level reconstruction relying on only tide gauge measurements for historical reference. Additionally, sea level trends can be estimated over the period from 1900 to 2010 (Figure 9), something that has generally been avoided in previous literature as a result of the poor tide gauge coverage.

[44] The sea level reconstruction for the Pacific Ocean region presented here represents the initial attempt at using the bivariate technique for improving reconstructions. In the future, we will extend this work to a global scale to try to improve the representation of climate signals in sea level for other regions like the Atlantic and Indian Oceans. Furthermore, while we have only included SLA and SST measurements in the reconstruction presented here, it is possible to incorporate additional observations such as sea level pressure to obtain further improvement. Assuming an adequate regression relationship with SLA, it is possible to reconstruct sea level using several different types of observations in a more expansive multivariate approach. As a by-product of this technique, we can also form reconstructions for the other variables included in the analysis, allowing for the improvement of SST reconstructions using sea level measurements, for instance.

[45] In summary, the focus of this paper is on the introduction of a new reconstruction technique that improves sea level reconstructions and makes them more useful for climate monitoring over a longer period of time. While GMSL is an important measure of the changing sea level in the ocean and a convenient way to inform the public about sea level rise, understanding how the current state of the ocean relates to the previous state of the ocean can potentially provide additional information on sea level changes over the world’s oceans in the future. Extending the discussion to actual climate signals and beyond GMSL represents an important and necessary step for improving the utility of sea level reconstructions for climate monitoring.

[46] Acknowledgments. This work was supported by NASA Ocean Surface Topography Mission Science Team grant NNX08AR60G and NASA ROSES grant NNX11AE26G. Ocean Topography Project Scientist support for R.R.L. by the NASA/JPL PO.DAAC is gratefully acknowledged. K.-Y.K. acknowledges support by the Ministry of Land, Transport, and Maritime Affairs (Ocean Climate Variability Program). The altimeter products used in this study were produced by Salto/Duacs and distributed by AVISO, with support from the Centre National d’Etudes Spatiales. We would also like to thank Richard W. Reynolds for his comments and insight into this work.

References