Reconstruction of Sea Level Around the Korean Peninsula Using Cyclostationary Empirical Orthogonal Functions

Se-Hyeon Cheon[1], Benjamin D. Hamlington[1], and Kyung-Duck Suh[2]

1Department of Ocean, Earth and Atmospheric Sciences, Old Dominion University, Norfolk, VA 23529, USA

2Institute of Construction and Environmental Engineering, Seoul National University, Seoul 08826, Republic of Korea

Correspondence to: Benjamin D. Hamlington (bhamling@odu.edu)

Abstract. Since the advent of the modern satellite altimeter era, the understanding of the sea level has increased dramatically. The satellite altimeter record, however, dates back only to the 1990s. The tide gauge record, on the other hand, extends through the 20th century, but with poor spatial coverage when compared to the satellites. Many studies have been conducted to create a dataset with the spatial coverage of the satellite datasets and the temporal length of the tide gauge records by finding novel ways to combine the satellite data and tide gauge data in what is known as sea level reconstructions. However, most of the reconstructions of sea level were conducted on a global scale, leading to reduced accuracy on regional levels, especially when there are relatively few tide gauges. The sea around the Korean Peninsula is one such area with few tide gauges before 1960. In this study, new methods are proposed to reconstruct the past sea level around the Korean Peninsula. Using spatial patterns obtained from a cyclostationary empirical orthogonal function decomposition of satellite data, we reconstruct sea level over the period from 1900 to 2014. Sea surface temperature data and altimeter data are used simultaneously in the reconstruction process, leading to an elimination of reliance on tide gauge data. Although we did not use the tide gauge data in the reconstruction process, the reconstructed results showed better agreement with the tide gauge observations in the region than previous studies that incorporated the tide gauge data. This study demonstrates a reconstruction technique that can be used at regional levels, with particular emphasis on areas with poor tide gauge coverage.

1 Introduction

Although sea level rise is a global phenomenon, the impacts are local and are happening now. Changes in sea level are affecting communities across the globe on an almost daily basis through increased erosion, greater saltwater intrusion, more frequent nuisance flooding, and higher storm surge causing severe damages on the coastal structures (e.g., Nicholls, 2011; Cheon and Suh, 2016; Suh et al., 2013). Planning for, adapting to, and mitigating current and future sea level has necessarily begun in many threatened areas. Expensive decisions - both in economic and societal terms - are already being made. Examples can be found throughout the world, with coastal communities making difficult decisions on how to address concerns associated with future sea level rise (e.g., Nicholls, 2011). The present and near-term threat of sea level rise across the globe highlight the immediate need for actionable regional sea level projections. In that respect, it is crucial to understand the long-term sea level changes to produce accurate projections in the future.

Before the satellite altimeter era, the only available sea level observations came from tide gauge (after this TG) records. The TGs provide the records of local sea level variations, covering nearly two hundred years in some locations around the globe. Using TG data, scientists have studied past sea level changes not only at specific locations but around the world. However, TGs provide a biased spatial coverage as they are necessarily only located at coastal sites and weighted towards the Northern Hemisphere. On the other hand, the satellite altimeters collecting data since 1992 have the near-global coverage of sea level but a relatively short observation period compared to TG observations, which is a severe handicap to analyzing long-term changes in sea level. This disadvantage is particularly true given the presence of sea level variability with decadal timescales.

Chambers et al. (2002) attempted to reconstruct sea level anomalies (SLA) by combining TG data and satellite altimeter data. In their research, they studied low-frequency variability in global mean sea level (GMSL) from 1950 to 2000. They interpolated sparse TG data into a global gridded SLA pattern applying EOFs (Empirical Orthogonal Functions) of SLA using data from the TOPEX/Poseidon satellite altimeter to capture the interannual-scale signals, e.g., ENSO (El Nino-Southern Oscillation) and PDO (Pacific Decadal Oscillation). Building on previous studies (Chambers et al., 2002; Kaplan et al., 1998, 2000), Church et al. (2004) created a reconstruction from 1950 to 2001 using EOFs of SLA data measured from satellite altimeter and a reduced space optimal interpolation scheme. This research was subsequently updated to increase temporal coverage from 1870 to the present (Church and White, 2006, 2011) and the reconstructions have been made available to the public through the website (http://www.cmar.csiro.au/sealevel/sl_data_cmar.html). In these studies, GMSL was found to rise approximately 210 mm from 1880 to 2009, with a linear trend from 1900 to 2009 of 1.7 ± 0.2 mm/yr. The resulting SLA is one of the most comprehensive and widely cited reconstructions. While these studies focused largely on the reconstruction of GMSL, Hamlington et al. (2011) applied cyclostationary empirical orthogonal functions (CSEOF) as basis functions for the reconstruction of SLA in an attempt to improve the representation of variability about the long-term trends. This approach provided an advantage for describing local variations such as ENSO and PDO. After that, Hamlington et al. (2012a) proposed an improved scheme of their reconstruction using sea surface temperature (hereafter SST). Given the limited TG data in the past, the reconstruction of SLA relying only on TGs were inaccurate, particularly before 1950. Leveraging other ocean observations (e.g., SST) led to an improved SLA reconstruction further into the past. In addition, this approach provides an advantage for describing local variations such as ENSO and PDO because the SST data gave an information on deep oceans where only few TGs are available.

While sea level is a global phenomenon, the extent of sea level change varies dramatically across the globe. During the 24-years satellite altimeter record, regional trends have been measured to be four times greater than the global average in some areas (Fig. 1). Therefore, the sea level assessment on a regional level is necessary to plan for future sea level accurately. Several studies focused on regional reconstructions targeting a particular area of interest. As an example, using an optimal interpolation method, Calafat and Gomis (2009) reconstructed the distribution of SLA in the Mediterranean Sea over 1945-2000. They used EOFs of satellite altimeter data spanning from 1993 to 2005 as basis functions and interpolated the TG data using these spatial patterns. A spatial distribution of sea level rise trends for the Mediterranean for the period of 1945-2000 indicated a positive trend in most areas. Calafat and Gomis (2009) estimated the linear trend of Mediterranean Sea's SLA from 1945 to 2000 at 0.7 ± 0.2 mm/yr. Hamlington et al. (2012b) performed a regional SLA reconstruction using CSEOFs as basis functions

(Hamlington et al., 2011) with a domain covering only the Pacific Ocean. They found that a choice of basis functions had a significant effect on the spatial pattern of the sea level rise and the ability to capture internal variability signals. A global basis functions, either CSEOFs or EOFs, are typically dominated by large-scale variability in the Pacific Ocean associated with ENSO or the PDO. As a result, global reconstructions are poorer in some ocean basins (e.g., Indian Ocean, Atlantic Ocean) than others (Pacific Ocean). This issue is likely exacerbated even further when looking at even smaller regions.

In this paper, we focus on one such region: the Korean Peninsula where over seventy-five million people live. In South Korea, over twenty-seven percent of people live in coastal city areas, and nearly thirty-six percent of GRDP (Gross Regional Domestic Product) is produced by coastal city regions (Choi and Jeong, 2015). As a result, policymakers have a keen interest in a sea level rise around the Korean Peninsula (hereafter KP; a suffix, '-KP' means the spatial domain of the data or variable is around the Korean Peninsula) to establish proper remedies to sea level rise. Studying SLA-KP, researchers have used globally reconstructed SLAs (Hamlington et al., 2011; Church and White, 2011) by trimming around the KP. However, subtracting SLA-KP (or more generally any small regions) from a grobally reconstructed SLA have some problems. First, global scale reconstructions use a limited number of basis functions to prevent interpolation from over-fitting and creating spurious sea level fluctuations. There is a difference between the dominant modes of global scale and local scale; e.g., there is a high possibility that the globally selected basis functions, which represent 90 % of the total variance in the global level will not represent 90 % of the total variance in local scale. Second, the temporal coverage of the TG around the KP (TG-KP) started around 1930 when the only TG had been available by 1950; it is too little to secure accuracy on these local scales. As mentioned above, TG-KP coverage is poor extending back into the 20th century, and relatively few TGs are available to analyze in some areas (Fig. 2). Hence, the goal of this study is proposing a new scheme that builds off of Hamlington et al. (2012b) that applies CSEOFs to reconstruct local SLA where the TG data is not enough to ensure a quality of reconstruction through the 20th century. We focus on the KP both due to its exposure to risk from impending sea level rise and also as a test case to demonstrate how this technique could be applied at other locations across the globe.

2 Data and Methods

2.1 Data

2.1.1 Sea level anomalies

The basis functions of this study's reconstructions are the CSEOFs of a gridded satellite data of SLA from 1993 to 2015. This monthly data has a $0.25^\circ \times 0.25^\circ$ resolution and AVISO (the Archiving, Validation, and Interpretation of Satellite Oceanographic) has provided this data since 2009; this satellite data opens in public (ftp://ftp.aviso.altimetry.fr/global/delayed-time/grids/climatology/monthly_mean/) and hereafter this dataset is written as AVISO. Before conducting the CSEOF decomposition, mean values for each grid point were removed to center the data. The annual signal has not been removed as it is accounted for by the CSEOF analysis (see more details in section 2.2.1). The data was trimmed to contain only the seas around the KP

(31°-46°N and 117°-142°E; hereafter AVISO-KP) and it was multiplied by the square root of the cosine of latitude to consider the actual area of each grid.

2.1.2 Sea surface temperature

In this study, two SST reconstruction datasets were used: ERSST (Extended Reconstructed Sea Surface Temperature; Huang et al., 2015, 2016; Liu et al., 2015) and COBESST2 (Centennial in situ Observation-Based Estimates; Ishii et al., 2005). The ERSST dataset is a global monthly SST dataset based on the observation of ICOADS (International Comprehensive Ocean-Atmosphere Dataset). This monthly analysis has a $2^\circ \times 2^\circ$ grid resolution and its time coverage is from 1854 to the present, and the included data are anomalies based on a monthly climatology computed from 1971-2000. The COBESST2 dataset is a monthly interpolated $1^\circ \times 1^\circ$ SST product from 1850 to the present. It integrates several SST observations: ICOADS 2.5, satellite SST, and satellite sea ice. The bucket correction process was applied to the data up to 1941. In addition to OI (Optimal Interpolation) scheme, this dataset used an EOF reconstruction.

Each data was trimmed as three different regions: a global domain (no trim), the Northwest Pacific (NWP) domain (25°-55°N and 110°-160°E), and around the KP area; to indicate the domains of dataset we put '-NWP' and '-KP' behind the name of dataset. Before conducting the CSEOF decomposition, these datasets were treated as follows. 1) The mean values for each grid point were removed to prevent those values to have a significant influence on CSEOFs. 2) The data were weighted by the square root of the cosine of latitude to consider the actual area of each grid. 3) Any grid points that were not continuous in time were removed. Like the satellite altimeter dataset, an annual signal was not removed.

2.1.3 Tide gauge data

Monthly mean records of 47 TG-KP were obtained from the Permanent Service for Mean Sea Level (PSMSL, Fig. 2) from 1930 to 2013. The earliest data of TG-KP is traced back to 1930 at Wajima Station (Fig. 2). Before 1965, the number of available TG datasets is fewer than 10, with only one TG (Wajima Station) providing data before 1950. An ongoing GIA (Glacial Isostatic Adjustment) correction was applied to the TG data using ICE-5G VM2 model (Peltier, 2004). Since an IB (Inverted Barometer) correction was applied to the satellite altimetry data, the TG-KP data are IB-corrected based on the pressure fields from 20th Century Reanalysis V2c data (Compo et al., 2006, 2011; Hirahara et al., 2014). The TG-KP data in this study are modified with further editing criteria. The techniques for editing are similar to those of Hamlington et al. (2011), with TG-KP that have shorter record length than 5 years and unphysical trends (greater than 7 mm/yr) likely owing to uncorrectable vertical land motions being removed prior to analysis. After calculating a month-to-month change, jumps greater than 250 mm were also removed.

2.1.4 Reconstructed sea level anomalies of previous studies

Church and White (2011, 2006) created the reconstruction of a global SLA from 1870 to 2009 using EOFs of SLA from satellite altimeter over 1993-2009. They applied a reduced space optimal interpolation technique. According to their research,

the GMSL rose about 210 mm over 1880-2009, and the linear trend through 1900-2009 was 1.7 ± 0.2 mm/yr. The resulting SLA is one of the most comprehensive and widely cited reconstruction results. This dataset was employed for long-term background trend for this study (see section 2.2.3). The GMSL timeseries (Church and White, 2011, 2006) has been extended and made publicly available (http://www.cmar.csiro.au/sealevel/GMSL_SG_2011_up.html). To reconstruct the past SLA, Hamlington et al. (2011) combined the CSEOFs of the satellite altimetry and historical TG record. This weekly analysis has a $0.5^\circ \times 0.5^\circ$ grid resolution and its time coverage is over 1950-2009. This dataset was used for the comparison with the reconstruction of this study (see section 2.2.3). This reconstruction dataset (Hamlington et al., 2011) can be downloaded from a NASA JPL/PO.DAAC (ftp://podaac.jpl.nasa.gov/allData/recon_sea_level/preview/L4/tg_recon_sea_level/).

2.2 Methods

A modified reconstruction method is proposed for an area such as the KP having poor TG coverage. The approach is based on the CSEOF decomposition and multivariate regression while taking into account a time lag. This approach is a progression from the technique described in Hamlington et al. (2012a). In this section, the procedure of the proposed scheme and fundamental theories are shown.

2.2.1 Cyclostationary empirical orthogonal functions

To understand the complex response of a physical system, the decomposition of data into a set of basis functions is frequently applied. The decomposed basis functions have the potential to give a better understanding of complex variability of the fundamental phenomenon. The simplest and most common computational basis functions are EOFs, which have often served as the basis for climate reconstructions. When a reconstruction selects the EOFs as basis functions, one basis function is defined as a single spatial map accompanied by a time series representing the amplitude modulation of this spatial pattern over time. The EOF decomposition of the spatio-temporal system, $T(r,t)$, is defined by the Eq. (1):

$$T(r,t) = \sum\nolimits_i LV_i(r) PCT_i(t), \tag{1}$$

where $LV(r)$ is a physical process (or loading vector) modulated by a time series $PCT(t)$ (principal component time series or PC time series). Combining each LV and PCT pair, a signal of single EOF mode can be produced.

The assumption underlying EOF-based reconstruction is the stationarity of the spatial pattern represented by the EOF over the entire period. However, the fact that many geophysical phenomena are cyclostationary is well known Kim et al. (2015). That is, some processes are periodic over a certain inherent timescale, with corresponding amplitudes varying over time. Even though EOFs represent cyclostationary signals through a superposition of multiple modes, as stated in Dommenget and Latif (2002), representing the cyclostationary signals with stationary EOFs can lead to an erroneous and ambiguous interpretation of the data. It also requires many EOFs to explain a relatively small amount of variability in a dataset.

To remedy some of these issues, Hamlington et al. (2011) introduced CSEOFs as the basis for the global reconstruction of SLA instead of EOFs. The CSEOF analysis proposed to capture the cyclostationary patterns and longer scale fluctuations in geophysical data (Kim et al., 1996; Kim and North, 1997; Kim and Wu, 1999; Kim and Chung, 2001; Kim et al., 2015). The

CSEOF analysis can capture the time varying signals as a single mode by giving a time dependency to the loading vectors. The system is defined as Eq.(2) and (3).

$$T(r,t) = \sum_i CSLV_i(r,t) PCT_i(t) \tag{2}$$

$$CSLV(r,t) = CSLV(r,t+d) \tag{3}$$

where $CSLV(r,t)$ is a Cyclostationary Loading Vector (for convenience, we call this as LV) and it is time dependent and periodic with a particular period (called a "nested period"). Previous studies (Kim et al., 1996; Kim and North, 1997; Kim et al., 2015) provide more detailed walk-through for the CSEOF computation and properties. CSEOFs have a significant advantage over EOFs since CSEOFs can explain cyclostationary signals in one mode; that is, CSEOFs of periodic processes are much easier to interpret than EOFs (Kim et al., 1996; Kim and North, 1997; Kim and Wu, 1999; Kim et al., 2015). Hamlington et al. (2011, 2012a, b) demonstrated that CSEOFs provided significant benefits dealing with repeating signals such as ENSO and Modulated Annual Cycle signals.

2.2.2 Multivariate regression using CSEOFs

When considering the complete Earth climate system, one variable is often directly connected to another variable. In some cases, they are impacted by a common physical process, or in other cases, one variable may directly influence another. To take advantage of these relationships and establish links, we can perform a multivariate linear regression as following Eq. (4).

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \epsilon \tag{4}$$

where $\beta_0, \beta_1, \beta_2, \cdots, \beta_k$ are regression coefficients and the is random error. In this study, the response variables are each PCT of AVISO-KP's CSEOF and the predictor variables are all PCT of each SST dataset's CSEOF. Eq. (4) can be re-written as follows:

$$PCT^m_{SLA} = \beta^m_0 + \beta^m_1 PCT^1_{SST} + \beta^m_2 PCT^2_{SST} + \cdots \beta^m_k PCT^k_{SST} + \epsilon^m \tag{5}$$

where PCT^m_{SLA} is the *m*-th PCT of AVISO-KP's CSEOF and β^m_k are regression coefficients for the *m*-th target and *k*-th PCT of SST (*m* = 1, 2, ..., *M*; *M* is total number of target's modes), PCT^k_{SST} is the *k*-th PCT of SST's CSEOF. The matrix form of the Eq. (5) is:

$$\begin{bmatrix} T^m_1 \\ T^m_2 \\ \vdots \\ T^m_n \end{bmatrix} = \begin{bmatrix} 1 & P^1_1 & P^2_1 & \cdots & P^k_1 \\ 1 & P^1_2 & P^2_2 & \cdots & P^k_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & P^1_n & P^2_n & \cdots & P^k_n \end{bmatrix} \times \begin{bmatrix} \beta^m_0 \\ \beta^m_1 \\ \vdots \\ \beta^m_k \end{bmatrix} \tag{6}$$

where T^m_n is the *n*-th component of PCT^m_{SLA}, P^k_n is the n-th component of PCT^k_{SST}.

Additionally, many geophysical signals have lagged relations with other geophysical signals (Bojariu and Gimeno, 2003; Dettinger et al., 1998; Hamlet et al., 2005; Hendon et al., 2007; Kawamura et al., 2004; McPhaden et al., 2006; Redmond and Koch, 1991). Hence, by assuming that the each mode of CSEOF represents an independent physical event, we can conclude the PCTs which are mathematically independent of each other also can have a lagged relationship. If we consider the lagged relationships between the target and predictor variables and use the predictors having a higher correlation, we can reduce the number of predictors in the regression; generally, the more predictors applied for the regression, the more noise is likely to appear in the simulation. Before performing the multivariate linear regression system as in (5), we calculated the cross-correlation between the target PCT of AVISO-KP and predictor PCTs of SST. The predictors were selected based on their cross-correlation values. The threshold cross-correlation value did not have a sensitive effect on the regression if the value can select more than ten predictors; in this study, we used 0.3 as the threshold. By assuming the lag of the *i*-th mode having maximum cross-correlation at lag ρ_i, the *m*-th mode's PCT of AVISO-KP can be given as follow based on the Eq. (5).

$$PCT^m_{SLA} = \beta^m_0 + \sum\nolimits_{i=1}^{k} \beta^m_i PCT^i_{SST}(t - \rho_i) + \epsilon^m \tag{7}$$

2.2.3 Reconstruction of SLA-KP

In this section, we explain the overall procedures of the current reconstruction. For the convenience sake, we use 'Re' as a prefix to indicate a reconstructed data, for example, ReSLA-KP means a reconstructed SLA-KP. Every SST dataset was trimmed to have the time span of 1891-2014. The AVISO was trimmed around the KP and the southeast sea of the Japanese islands was removed. Every SST data was cut into three regions: around the KP, the Northwest Pacific Ocean, and global (no trimming). All grid points that were not continuous in time were removed for every dataset. In total, we tested six different SST combinations. GMSL and mean values were removed from AVISO-KP at each grid point (hereafter we add suffix '0' to show the data have no GMSL, e.g., AVISO-KP0). Each data was weighted by the square root of the cosine of latitude to consider the actual area of each grid. We conducted the CSEOF decomposition for all data (AVISO-KP0 and six SST combinations)

with twelve month nested period. The lagged relation between PCTs of AVISO-KP0 and PCTs of each SST dataset were estimated with two years maximum lagging boundary. Using the PCTs of each dataset's CSEOF, we built the multiple linear regression systems based on Eq. (7) over 1993-2014. In this regression, the target variables were each PCT of AVISO-KP0 and the predictors are PCTs of each SST combination. The regression coefficients and their confidence intervals were estimated to extend the target variables. Applying Monte Carlo simulation that used the confidence intervals of regression coefficients, we randomly generated a thousand sample-sets of each extended PCT of AVISO-KP0. By combining the extended PCTs to the LVs of AVISO-KP0, we produced a thousand ReSLA-KP0. By adding randomly generated GMSLs (Church and White, 2011) to the ReSLA-KP0, a thousand of ReSLA-KP were generated. Finally, by statistical analysis of each time step of the random samples, we estimated the mean variation and their confidence intervals of each reconstruction.

For comparison, in addition to the TG-KP, we used the reconstructed dataset of Hamlington et al. (2011). We trimmed the dataset to have same domain with AVISO-KP; hereafter ReSLA-H. The reconstruction results over 1970-2009 are quite reliable, because, after 1970, the number of available TG record around the world is enough to guarantee the reconstruction results. The correlation coefficient (ρ) and NRMSE (Normalized Root Mean Square Error; we obtain this value through dividing RMSE by the standard deviation of the reference dataset; Eq. (8)) values for the entire domain and each TG location were calculated. By using these two values, we decided the best reconstruction case among the six reconstructions which are introduced in section 3.2.

$$NRMSE = 1 - \frac{\|x_{ref}(i) - x(i)\|}{\|x_{ref}(i) - \mu_{x_{ref}}\|} \tag{8}$$

where $\|\blacksquare\|$ indicates the 2-norm of a vector, x_{ref} and x are reference data and tested data respectively.

3 Results and Discussions

3.1 Sea Level Anomalies around the KP

Using AVISO-KP over 1993-2015, a linear trend map was estimated as shown in Fig. 3. The mean trend was found to be 3.1 $\pm$ 0.5 mm/yr. The linear trend of mean SLA-KP (hereafter MSLA-KP; we use 'M' to show the spatially averaged data) agrees closely with the GMSL trend, 3.0 $\pm$ 0.0 mm/yr (Fig. 4). Due to the similarity between the long-term trends of MSLA-KP and GMSL (Fig 4), it is reasonable that we assume the SLA-KP can be described as the combination between background signals (GMSL) and the residuals which contain local characteristics of SLA-KP. Most of the SLA-KP trends were close to the mean, but some parts of the East/Japan Sea, and of the Yellow Sea close to land, exhibited extreme patterns. Some areas showed trends over 7 mm/yr, while in other regions there were trends less than 1 mm/yr of the linear trend (Fig. 3). To check whether the extreme trends patterns had a significant influence on MSLA-KP, we compared the MSLA of the area having the extreme trends and the other area. We calculated the mean correlation (hereafter $\bar{\rho}$) of each grid point of AVISO-KP to saperate the two areas. For example, $\bar{\rho}$ at a single grid point *P* was calculated by taking mean of ρ values that had been estimated between *P* and

all other points. By repeating these calculations at all the points of AVISO-KP, we obtained Fig. 5. We deemed that the SLAs of the regions having relatively high $\bar{\rho}$ fluctuate together, on the other hand, the SLAs of the low $\bar{\rho}$ regions oscillate separately. The regions that had the relatively low correlation coefficient agreed with the regions that had the extreme linear trends (Fig. 3 and 5). We divided the SLA-KP into two regions according to the mean correlation coefficient; we roughly selected the threshold value as 0.5, which can separate the area having extreme trend and the remaining area. The MSLA of each region shows a good agreement each other (Fig. 6). This demonstrates that the small-scale extreme features tend to cancel out and do not significantly impact on MSLA-KP. This also suggests that the entire region can be treated as local variability fluctuating about some background long-term mean, an important feature for this reconstruction procedure.

The linear trend at each TG location was estimated and it was compared with the nearest point in AVISO-KP; using the same data, the ρ values were estimated and the mean value of the ρ was about 0.72 (Fig. 7). The comparison showed that only five TGs showed acceptable accuracies having less than 30% of difference with the AVISO-KP's linear trend. Eleven TGs showed more than 30% of underestimation and twenty-one TGs had more than 30% of over estimation. To figure out the effect of these disagreements, the MSLA of AVISO-KP was compared with TG-KP's MSLA, and they showed $\bar{\rho}$ = 0.89 and NRMSE = 0.52 (Fig. 8). The linear trend of MSLA of TG-KP was estimated as 4.31 mm/yr and this value is about 40% higher than the MSLA of AVISO-KP. This disagreement originated from the short time period and a lack of TGs. Unresolved vertical land motion at the TG-KP could also lead to such disagreements.

CSEOF decomposition was conducted to investigate periodic orthogonal behaviors for SLA-KP with twelve month nested period after removing mean values at each grid point. The first mode represents an annual variation considering the spatial patterns and PCT of the CSEOF (Fig. 10). Nearly 60% of SLA-KP variations can be presented by the first mode (Fig. 9). The second mode shows similar spatial patterns having positive value for all months, and the PCT shows clear positive trend (Fig. 11). This means the second mode is a sea level rising mode, and it represents 10% of variations of SLA-KP roughly. The third and fourth modes were not able to relate to specific phenomenon, and the modes explain variability in the SLA-KP at about 5% and 3%, respectively. Using the four modes, we can explain about 70% of SLA-KP. The first and second modes have the linear trend, but the linear trend in the first mode is negligibly small compared with the signal itself (Fig. 12 and 13). Hence, we can say that the second mode is the most important key to estimating SLA-KP.

3.2 Sea Level Reconstruction around the KP

We calculated MSLA using ReSLA-KP, ReSLA-H, and TG-KP; to calculate the spatial means of the reconstructed data, we considered the each grid's area; for TG-KP, we estimated the average of sea level differences between each time steps at all TGs, and integrated the averaged differences to have the MSLA-KP.

Six reconstructions were conducted and the MSLAs of six reconstructions showed a reasonable agreement with the MSLA of TG-KP over 1965-2014. For the period prior to 1965, however, the results showed considerable diversity (Fig. 15). The six MSLAs were compared with ReSLA-H and TG-KP over 1970-2008 because there were a few TG data available before 1970. Both a correlation coefficient and a NRMSE were applied for the quantified comparison (Fig. 14). Considering the NRMSE, the SSTs of NWP and KP provided better reconstructions than ReSLA-H. However, in terms of correlation coefficient, only

SSTs of NWP showed better results than ReSLA-H. Finally, we selected the reconstruction using COBESST2-NWP as the best reconstruction considering both NRMSE and correlation coefficient.

Most of reconstructions show better agreement than ReSLA-H in terms of correlation coefficients despite we did not use TG data during the reconstruction process. The mean ReSLA-KP shows good agreement with the mean ReSLA-H, but poor agreement with the MSLA-TG (Fig. 15). This disagreement, however, is likely caused by lack of high-quality TGs before 1970. We further calculated correlation coefficients and linear trends using TG-KP and ReSLA-KP, ReSLA-H at the each TG location; for the reconstructed data, we calculated the linear trends at the nearest grid points. We made two correlation comparisons: one between ReSLA-KP and TG-KP, and the other between ReSLA-H and TG-KP to check if ReSLA-KP showed better representation of each TG-KP. The ReSLA-KP showed higher correlation coefficient values than ReSLA-H (Fig. 16a) demonstrating the better agreement between the current reconstruction and TG-KP. The linear trends of TG-KP, ReSLA-KP, and ReSLA-H were estimated at the TG location over 1970 to the present; for the calculation, each time series was edited to have the same time span. The estimated linear trends are given in Fig. 16b. The ReSLA-KP has similar linear trends with ReSLA-H at the TG location, and the variance of the trends are smaller than TG-KP (Fig. 16b); we conducted t-test to check statistical significances of the trend values, and the p-values were shown in Fig. 16c. ReSLA-KP shows better agreement with the AVISO-KP than ReSLA-H over satellite era (Fig. 17). It also has more fluctuations (Fig. 15) and these detailed fluctuations are closer to the AVISO-KP, and this is likely a result of the applied number of modes for the reconstruction process. Hamlington et al. (2011) used a limited number (< 90% of total variance) of CSEOF modes to avoid over-fitting issues, but in this study, nineteen CSEOF modes are used which explain 98% of total variance of SLA-KP.

Using Monte Carlo simulation, the means and standard devations of ReSLA were estimated for the best reconstruction case (COBESST2-NWP). By applying the means and standard deviations of regression coefficients (Eq. 8), each mode's PCT was randomly extended into the past, and the process was repeated by a thousand times. The extended PCTs were combined with corresponding LVs of AVISO-KP. Through this process, a thousand of ReSLA-KP were generated, and the mean and standard deviation were estimated at each time step and grid point. The resulting MSLA-KP and 95% confidence interval are shown in Fig. 18.

The linear trend in ReSLA-KP over 1900-2014 is estimated as 1.71 ± 0.04 mm/yr, and this value is similar to the linear trend of Church and White (2011) as 1.70 ± 0.02 mm/yr. A linear trend map of ReSLA-KP was calculated, and the maximum and minimum linear trends are about 2.1 mm/yr and 1.4 mm/yr, respectively (Fig. 19). The difference on two extreme values of the ReSLA-KP is much less than the AVISO-KP over 1993-2015, particularly in the Yellow Sea, (Fig. 3 and 19). This alleviation means that the extended reconstruction period can reduce the effect of the internal variability having a large amplitude.

4 Summary

There were two primary goals of the work presented in this study: 1) Improve the understanding of the sea level around the KP both in the past and present and 2) Present a new reconstruction scheme for local areas with insufficient tide gauge coverage. To meet these goals, we used the satellite altimeter data from AVISO and the TG data from PSMSL to investigate

the characteristics of SLA-KP. The linear trend of MSLA-KP was estimated as 3.1 $\pm$ 0.5 mm/yr from the satellite altimeter data (Fig. 4). However, when we looked into the trend map, some areas (such as near the river mouth in the Yellow Sea and in the middle of the East/Japan Sea) showed significant departures from the mean trend (Fig. 3).

To investigate this further, the reconstruction was performed using AVISO-KP and two SST reanalysis datasets. Each SST dataset was divided into three cases (global, NWP, KP). The six datasets were decomposed by CSEOF analysis; the AVISO-KP was decomposed into CSEOF modes after removing the GMSL. The decomposed LVs played a role of basis functions for the reconstruction, and the main process of reconstruction was extending the PCTs of each mode into the past. The six reconstructed SLA-KPs were generated by this study over 1900-2014. Using the correlation coefficient and the NRMSE, the best reconstruction was selected. The best reconstruction was produced by COBESST2-NWP. Through the best reconstruction results, the linear trend of SLA-KP was estimated as 1.71 $\pm$ 0.04 mm/yr over 1900-2014 (Fig. 18). The extreme linear trends shown in Fig. 3 did not appear in the ReSLA-KP (Fig. 3 and 19).

While we focus here on a specific example (the KP), this study can be used to inform other efforts in studying past and present sea level in areas with poor tide gauge coverage. Our interest was on the KP, specifically, but it was found that including information from the NWP improved the localized representation of sea level. Consequently, considering large-scale ocean variability and teleconnections between different parts of the ocean is important when selecting the reconstruction domain. This study also demonstrates that TG data may not even be necessary to understand sea level in the past. Using only satellite-based sea level information and SST, we found dramatic improvements between the current reconstruction and past efforts, particularly when comparing to the TG variability. Many TGs are influenced by vertical land motion that cannot easily be corrected for. Relying on SST alleviates concerns associated with non-ocean related trends. It should be noted that this reconstruction may not work as well in other parts of the ocean, especially those with a less pronounced agreement between sea level and SST. This study does, however, demonstrate the extended efforts that must be made to obtain accurate information about past sea level. As planning efforts get underway in more parts of the world, such comparisons between past and present sea level will become more important, and alternative approaches to simply using TG information are going to be needed.

Acknowledgements. S.H.C. and K.D.S. were supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT and Future Planning (NRF-2014R1A2A2A01007921). B.D.H. acknowledges support from NASA PO NNX15AG45G and NNX16AH56G.

References

Bojariu, R. and Gimeno, L.: The role of snow cover fluctuations in multiannual NAO persistence, Geophysical Research Letters, 30, 2003.

Calafat, F. M. and Gomis, D.: Reconstruction of Mediterranean sea level fields for the period 1945–2000, Global and Planetary Change, 66, 225–234, 2009.

Chambers, D. P., Mehlhaff, C. A., Urban, T. J., Fujii, D., and Nerem, R. S.: Low-frequency variations in global mean sea level: 1950–2000, Journal of Geophysical Research: Oceans, 107, 2002.

Cheon, S.-H. and Suh, K.-D.: Effect of sea level rise on nearshore significant waves and coastal structures, Ocean Engineering, 114, 280–289, 2016.

Choi, H. and Jeong, M.: Demographic Change in the Coastal Zones, Proceeding of Korean Society for Marine Environment and Energy, p. 115, 2015.

Church, J. A. and White, N. J.: A 20th century acceleration in global sea-level rise, Geophysical research letters, 33, 2006.

Church, J. A. and White, N. J.: Sea-level rise from the late 19th to the early 21st century, Surveys in Geophysics, 32, 585–602, 2011.

Church, J. A., White, N. J., Coleman, R., Lambeck, K., and Mitrovica, J. X.: Estimates of the regional distribution of sea level rise over the 1950–2000 period, Journal of climate, 17, 2609–2625, 2004.

Compo, G. P., Whitaker, J. S., and Sardeshmukh, P. D.: Feasibility of a 100-year reanalysis using only surface pressure data, Bulletin of the American Meteorological Society, 87, 175–190, 2006.

Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., Gleason, B. E., Vose, R. S., Rutledge, G., Bessemoulin, P., et al.: The twentieth century reanalysis project, Quarterly Journal of the Royal Meteorological Society, 137, 1–28, 2011.

Dettinger, M. D., Cayan, D. R., Diaz, H. F., and Meko, D. M.: North–south precipitation patterns in western North America on interannual-to-decadal timescales, Journal of Climate, 11, 3095–3111, 1998.

Dommenget, D. and Latif, M.: A cautionary note on the interpretation of EOFs, Journal of climate, 15, 216–225, 2002.

Hamlet, A. F., Mote, P. W., Clark, M. P., and Lettenmaier, D. P.: Effects of temperature and precipitation variability on snowpack trends in the western United States, Journal of Climate, 18, 4545–4561, 2005.

Hamlington, B., Leben, R., Nerem, R., Han, W., and Kim, K.-Y.: Reconstructing sea level using cyclostationary empirical orthogonal functions, Journal of Geophysical Research: Oceans, 116, 2011.

Hamlington, B., Leben, R., and Kim, K.-Y.: Improving sea level reconstructions using non-sea level measurements, Journal of Geophysical Research: Oceans, 117, 2012a.

Hamlington, B., Leben, R., Wright, L., and Kim, K.-Y.: Regional sea level reconstruction in the Pacific ocean, Marine Geodesy, 35, 98–117, 2012b.

Hendon, H. H., Wheeler, M. C., and Zhang, C.: Seasonal dependence of the MJO–ENSO relationship, Journal of Climate, 20, 531–543, 2007.

Hirahara, S., Ishii, M., and Fukuda, Y.: Centennial-scale sea surface temperature analysis and its uncertainty, Journal of Climate, 27, 57–75, 2014.

Huang, B., Banzon, V. F., Freeman, E., Lawrimore, J., Liu, W., Peterson, T. C., Smith, T. M., Thorne, P. W., Woodruff, S. D., and Zhang, H.-M.: Extended reconstructed sea surface temperature version 4 (ERSST. v4). Part I: Upgrades and intercomparisons, Journal of climate, 28, 911–930, 2015.

Huang, B., Thorne, P. W., Smith, T. M., Liu, W., Lawrimore, J., Banzon, V. F., Zhang, H.-M., Peterson, T. C., and Menne, M.: Further exploring and quantifying uncertainties for extended reconstructed sea surface temperature (ERSST) version 4 (v4), Journal of Climate, 29, 3119–3142, 2016.

Ishii, M., Shouji, A., Sugimoto, S., and Matsumoto, T.: Objective analyses of sea-surface temperature and marine meteorological variables for the 20th century using ICOADS and the Kobe collection, International Journal of Climatology, 25, 865–879, 2005.

Kaplan, A., Cane, M. A., Kushnir, Y., Clement, A. C., Blumenthal, M. B., and Rajagopalan, B.: Analyses of global sea surface temperature 1856–1991, Journal of Geophysical Research: Oceans, 103, 18 567–18 589, 1998.

Kaplan, A., Kushnir, Y., and Cane, M. A.: Reduced space optimal interpolation of historical marine sea level pressure: 1854–1992, Journal of Climate, 13, 2987–3002, 2000.

Kawamura, R., Suppiah, R., Collier, M. A., and Gordon, H. B.: Lagged relationships between ENSO and the Asian Summer Monsoon in the CSIRO coupled model, Geophysical Research Letters, 31, 2004.

Kim, K.-Y. and Chung, C.: On the evolution of the annual cycle in the tropical Pacific, Journal of Climate, 14, 991–994, 2001.

Kim, K.-Y. and North, G. R.: EOFs of harmonizable cyclostationary processes, Journal of the Atmospheric Sciences, 54, 2416–2427, 1997.

Kim, K.-Y. and Wu, Q.: A comparison study of EOF techniques: Analysis of nonstationary data with periodic statistics, Journal of Climate, 12, 185–199, 1999.

Kim, K.-Y., North, G. R., and Huang, J.: EOFs of one-dimensional cyclostationary time series: Computations, examples, and stochastic modeling, Journal of the atmospheric sciences, 53, 1007–1017, 1996.

Kim, K.-Y., Hamlington, B., and Na, H.: Theoretical foundation of cyclostationary EOF analysis for geophysical and climatic variables: concepts and examples, Earth-Science Reviews, 150, 201–218, 2015.

Liu, W., Huang, B., Thorne, P. W., Banzon, V. F., Zhang, H.-M., Freeman, E., Lawrimore, J., Peterson, T. C., Smith, T. M., and Woodruff, S. D.: Extended reconstructed sea surface temperature version 4 (ERSST. v4): Part II. Parametric and structural uncertainty estimations, Journal of Climate, 28, 931–951, 2015.

McPhaden, M. J., Zhang, X., Hendon, H. H., and Wheeler, M. C.: Large scale dynamics and MJO forcing of ENSO variability, Geophysical research letters, 33, 2006.

Nicholls, R. J.: Planning for the impacts of sea level rise, Oceanography, 2011.

Peltier, W.: Global glacial isostasy and the surface of the ice-age Earth: the ICE-5G (VM2) model and GRACE, Annu. Rev. Earth Planet. Sci., 32, 111–149, 2004.

Redmond, K. T. and Koch, R. W.: Surface climate and streamflow variability in the western United States and their relationship to large-scale circulation indices, Water Resources Research, 27, 2381–2399, 1991.

Suh, K.-D., Kim, S.-W., Kim, S., and Cheon, S.: Effects of climate change on stability of caisson breakwaters in different water depths, Ocean Engineering, 71, 103–112, 2013.

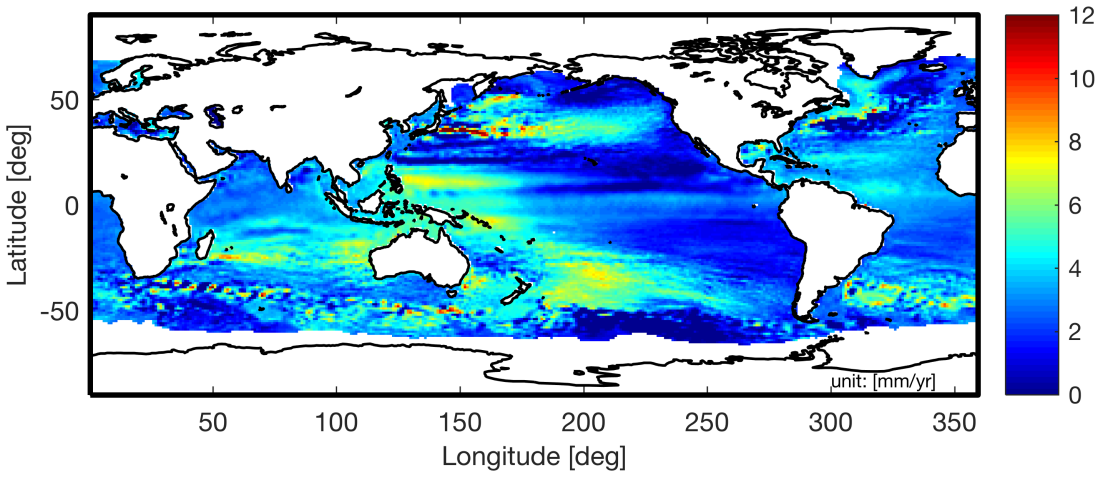

Figure 1. Global linear trend map of sea level anomalies using AVISO from 1993 to 2015

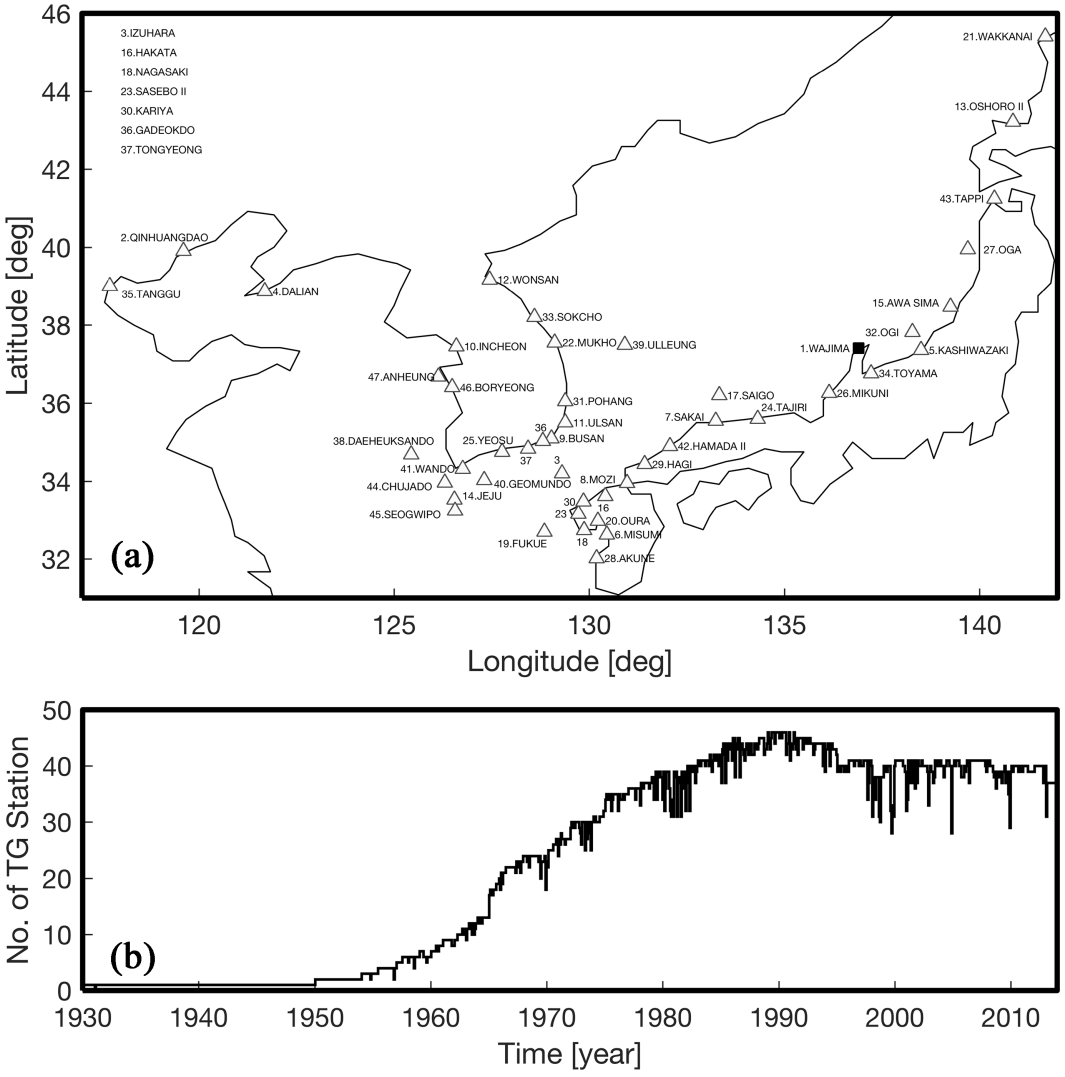

Figure 2. (a) The locations of tide gague station used in this study around the Korean Peninsula. The black square is Wajima station which has the longest record length (1930-present); (b) The number of tide gague stations provided by PSMSL around the Korean Peninsula

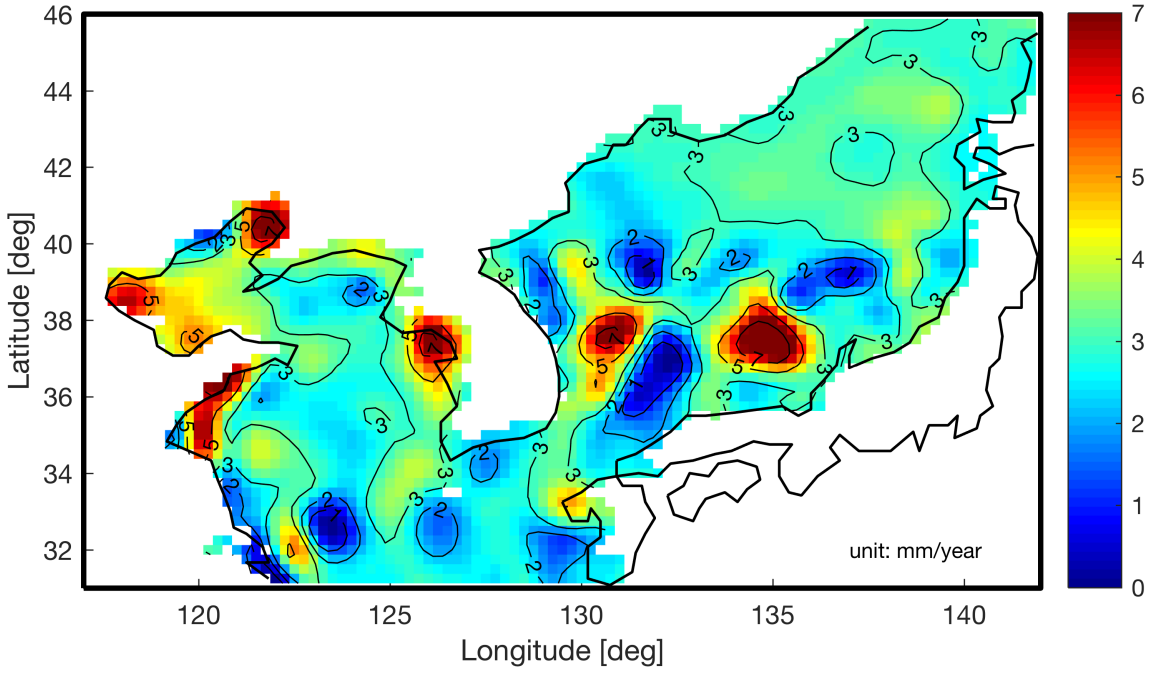

Figure 3. Linear trend map of sea level anomalies around the Korean Peninsula from AVISO w/o annual signal from 1993 to 2015

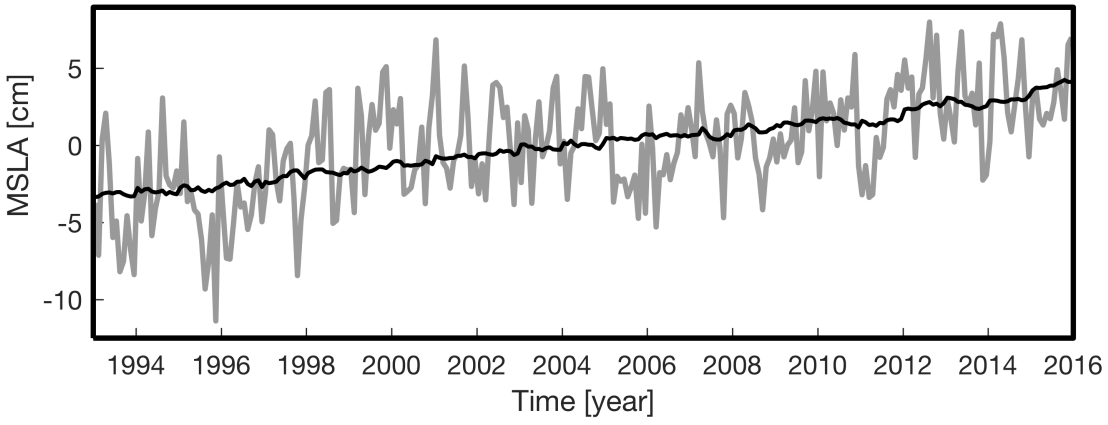

Figure 4. Spatial mean time series of sea level anomalies around the Korean Peninsula (gray) and global (black) from AVISO w/o annual signal

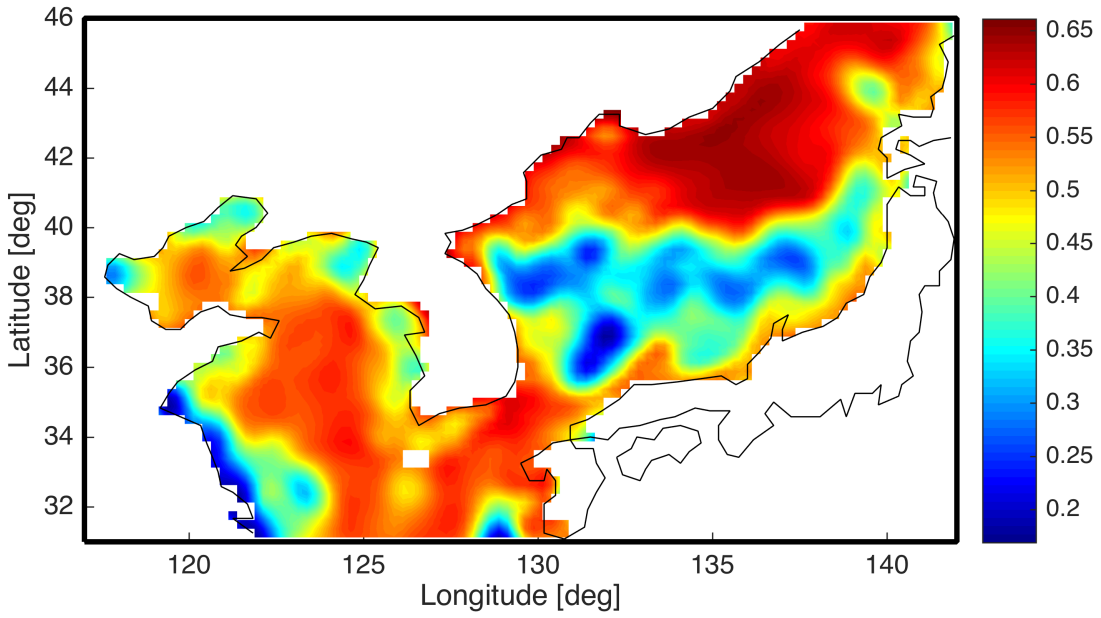

Figure 5. Mean correlation coefficient map of sea level anomalies around the Korean Peninsula from AVISO w/o annual signal from 1993 to 2015

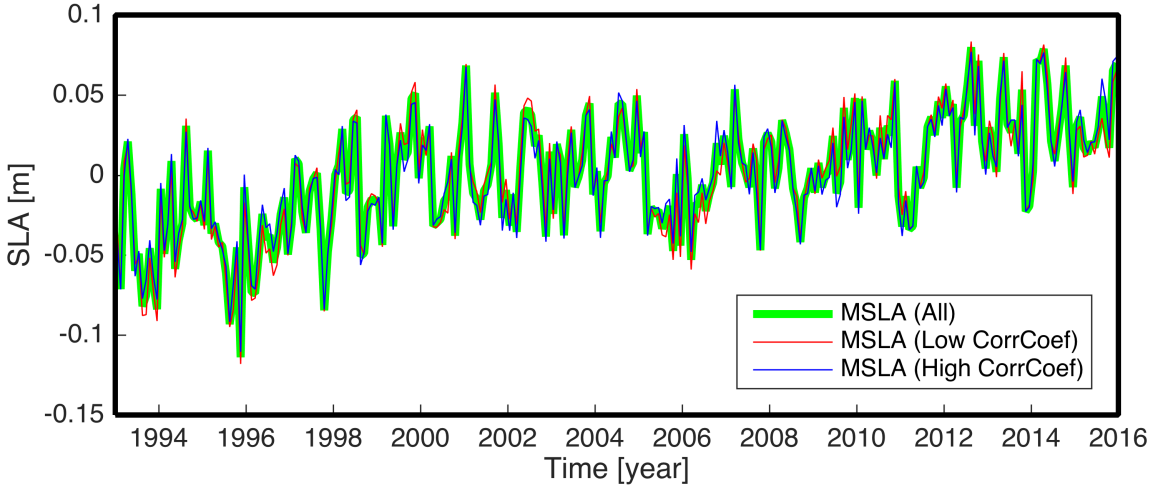

Figure 6. Spatial mean time series of sea level anomalies from two regions based on the correlation coefficients in Fig. 5

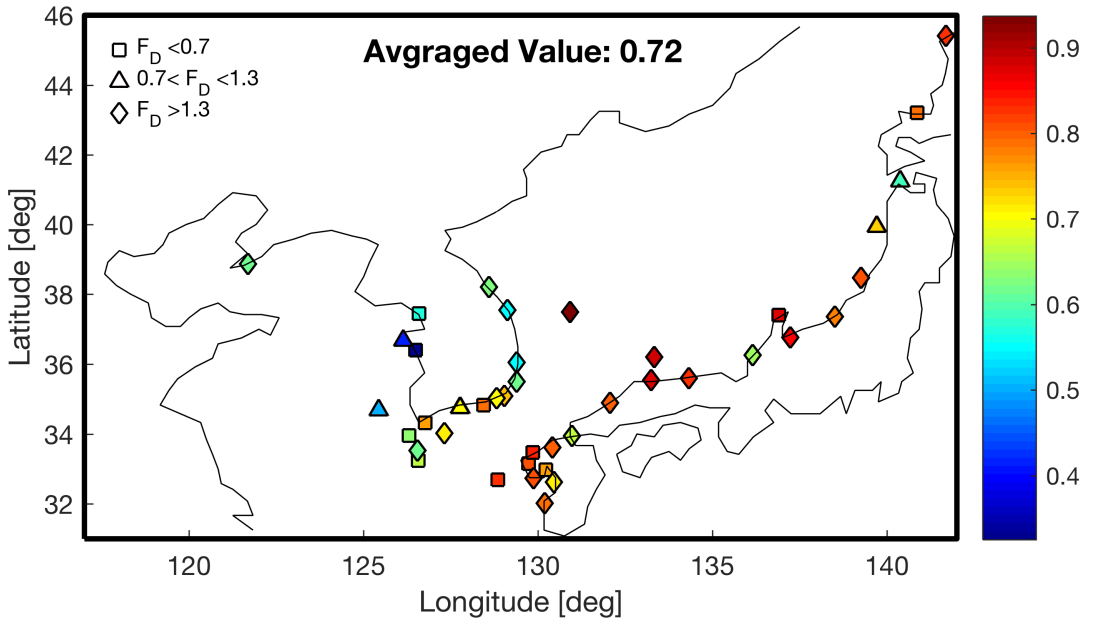

Figure 7. Linear Trends comparison (shapes) and correlation coefficients (colors) between tide gauge and the closest AVISO grid point (< 12 km) from 1993 to 2014, where $F_D = SLR_{TG} / SLR_{AVISO}$ (w/o annual signals)

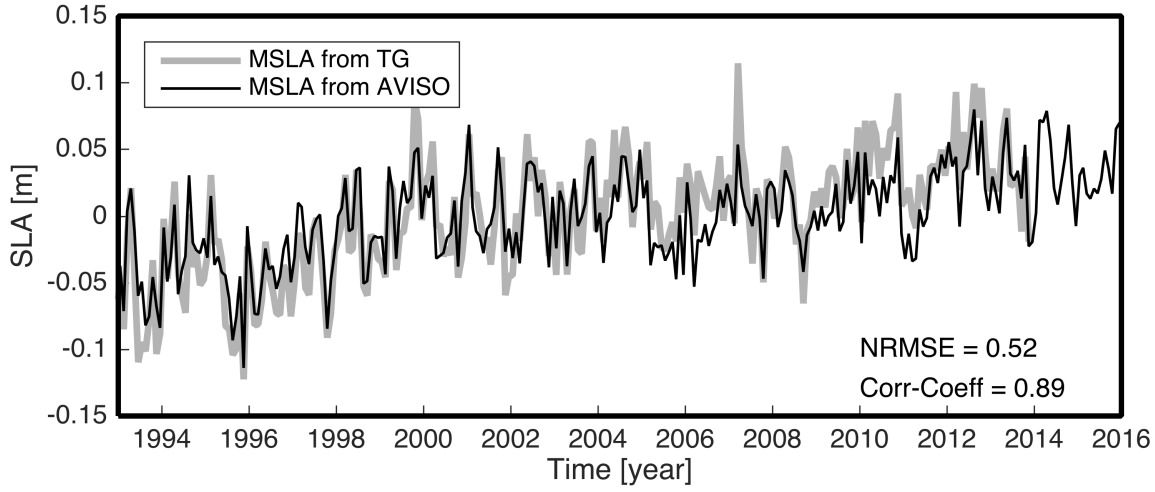

Figure 8. Spatial mean time series of sea level anomalies of tide gauge and AVISO around the Korean Peninsula w/o annual signal

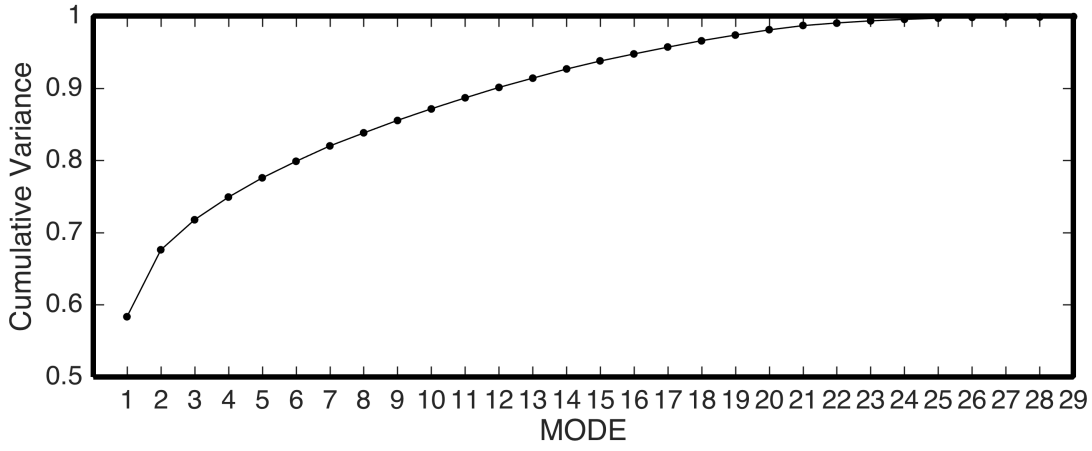

Figure 9. Cummulative variance of each CSEOF mode of the AVISO-KP

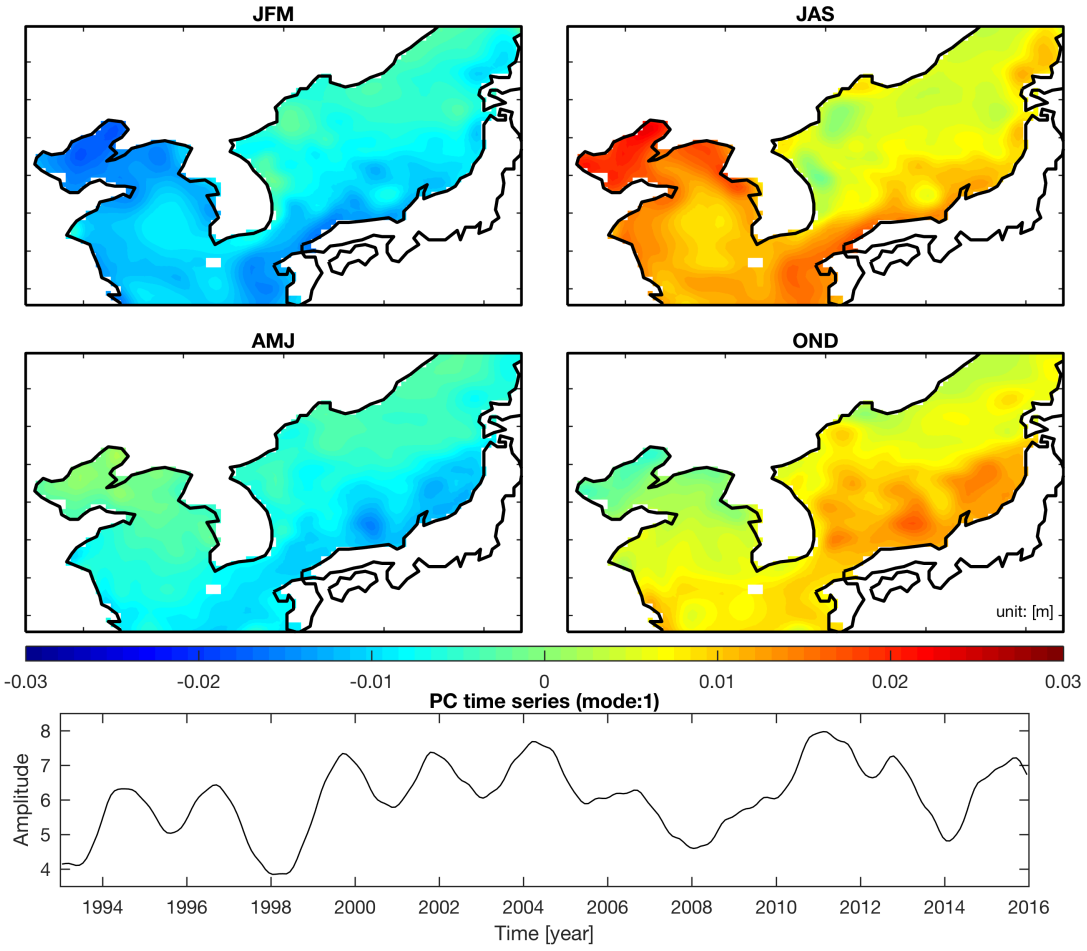

Figure 10. The first CSEOF mode of AVISO-KP

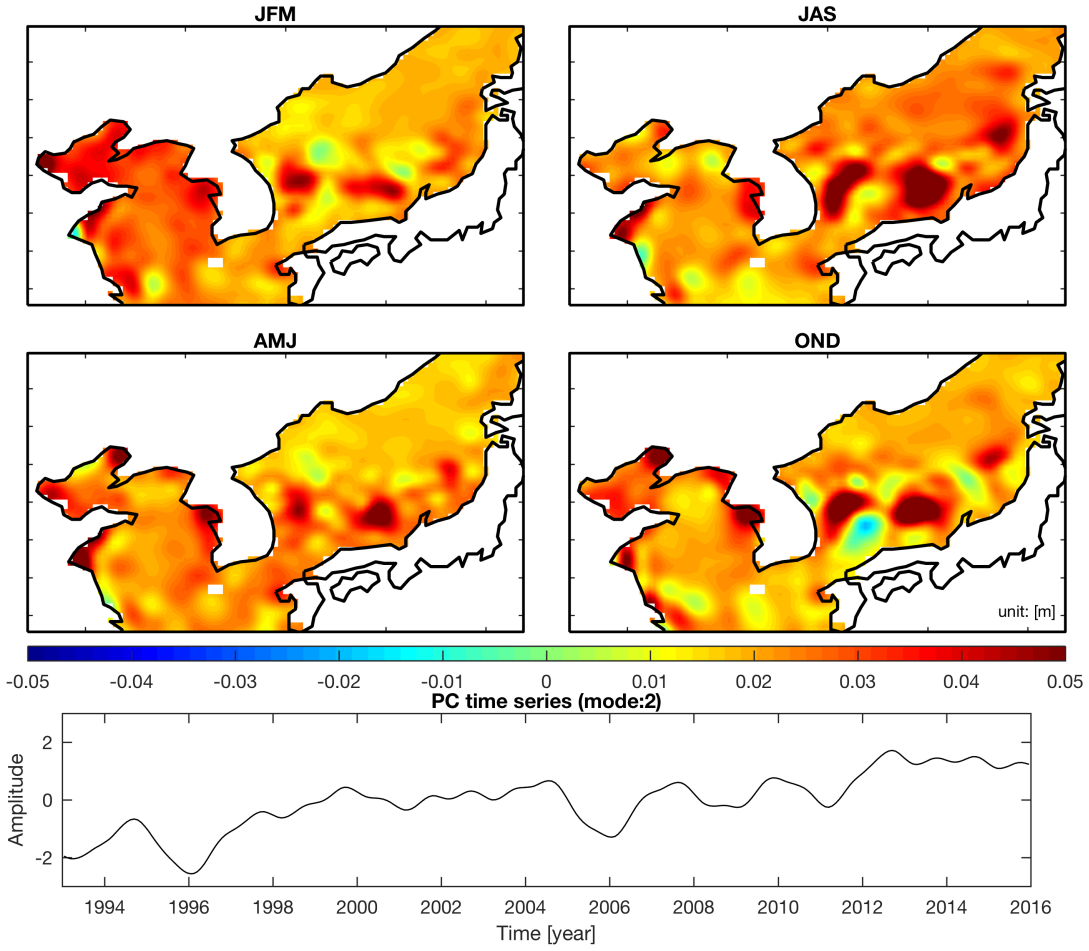

Figure 11. The second CSEOF mode of AVISO-KP

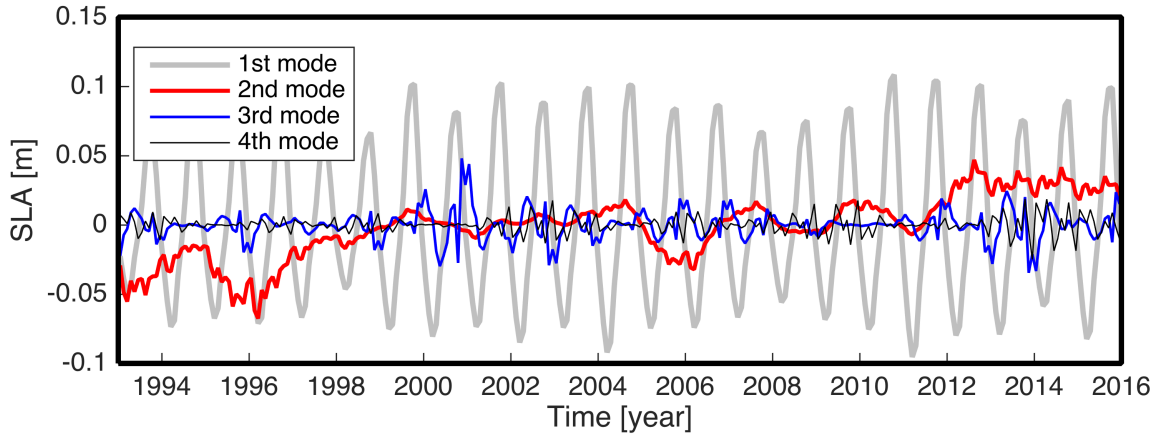

Figure 12. MSLA of the four biggest modes of CSEOF decomposition of AVISO-KP

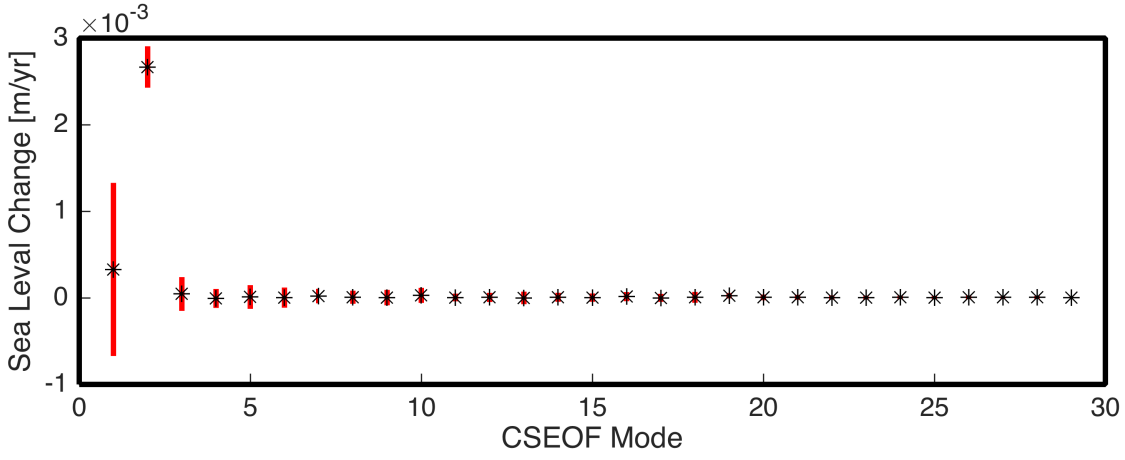

Figure 13. Linear trends (black '*') and 95% confidence intervals (red line) of the spatially averaged CSEOF mode of AVISO-KP

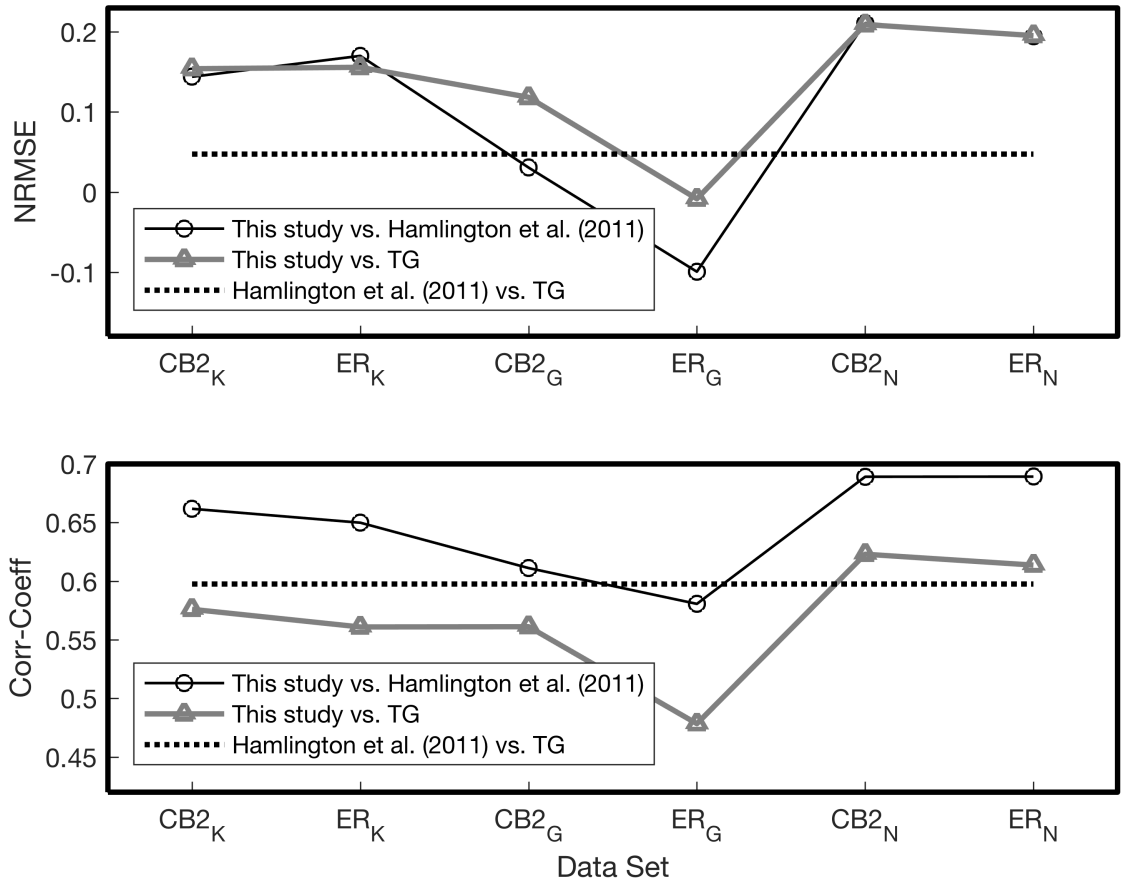

Figure 14. Results of goodness of fit test for Reconstructed MSLA according to Hamlington et al. (2011) and TG MSLA; the top figure include normalized root mean squared error and the other include the correlation coefficients; here subscripts K, G, and N represent around the Korean Peninsula, Global, and the North-West Pacific, respectively and CB2 and ER represent COBESST2 and ERSST

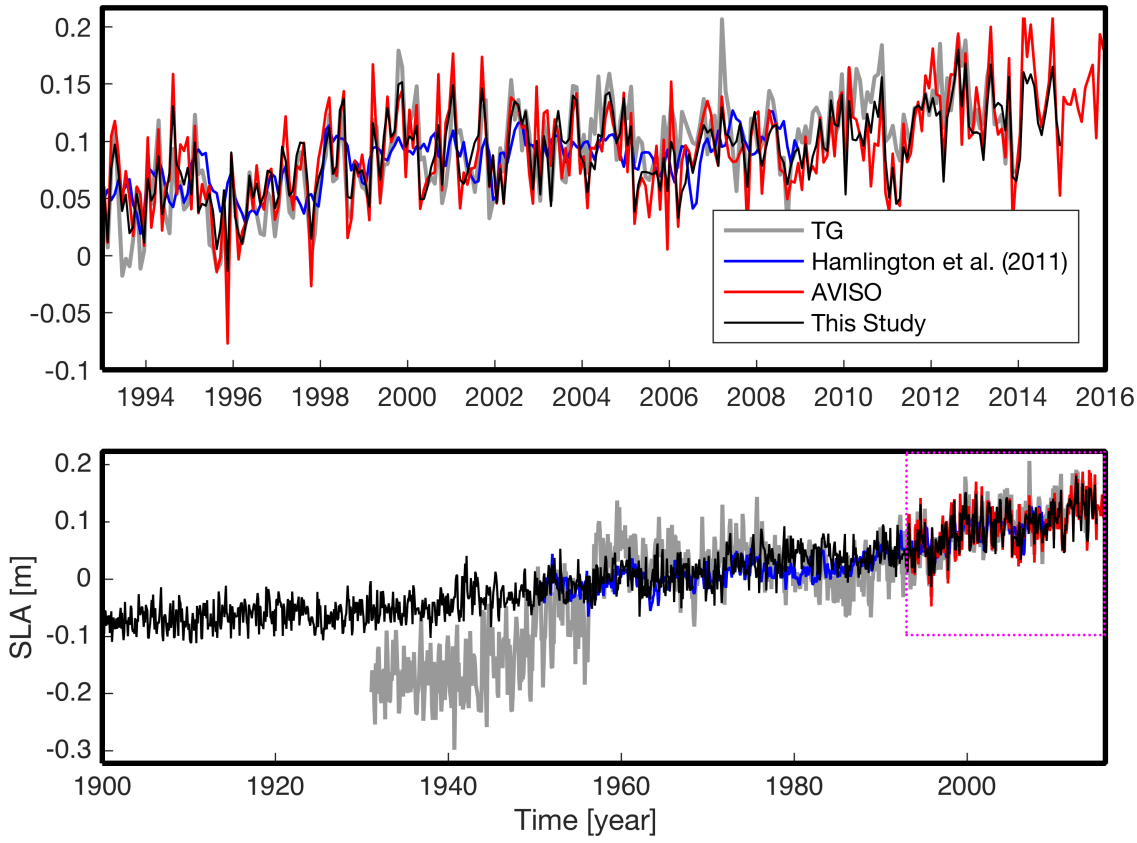

Figure 15. Comparison of spatial mean time series of sea level anomalies (tide gauge, AVISO, ReSLA-H (Hamlington et al, 2011), and this study) around the KP w/o annual signal

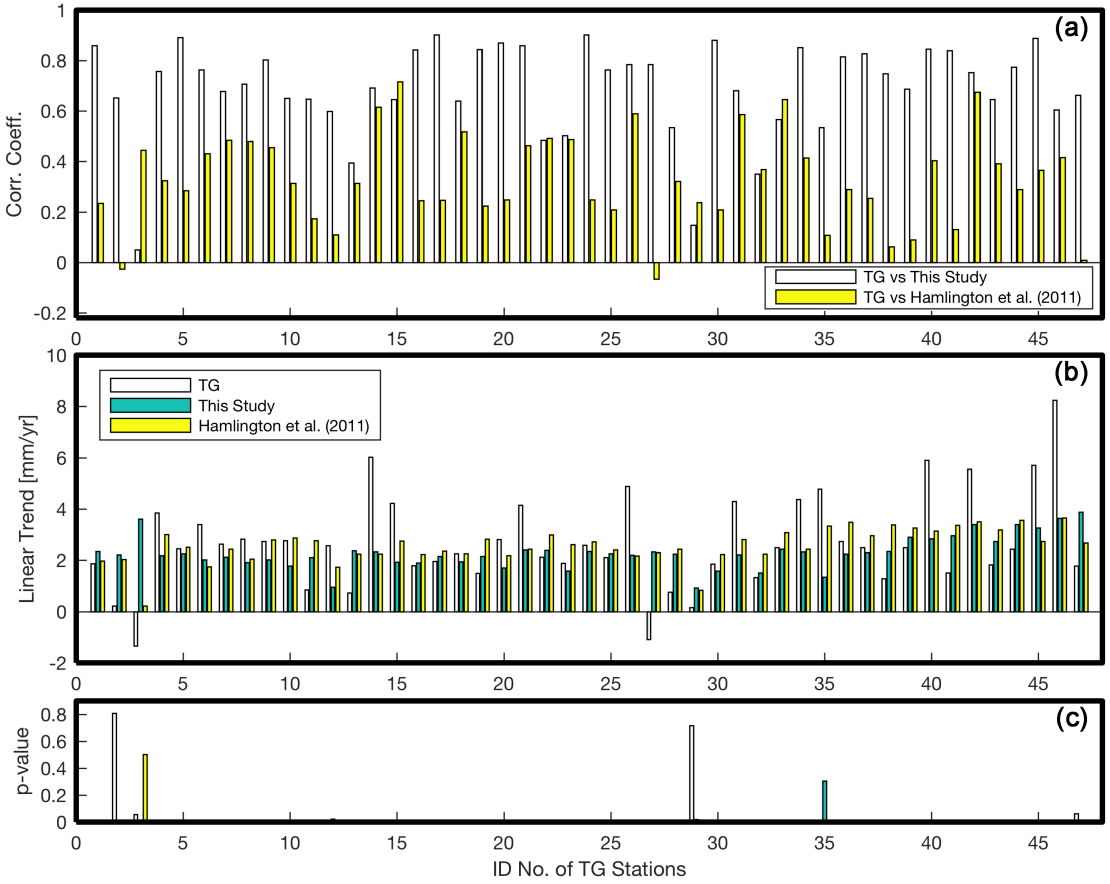

Figure 16. (a) Comparison of correlation coefficients between TG-KP and the reconstruction results over 1970-2008; (b) Comparison of linear trends over 1970-2008; (c) t-test result of (b)

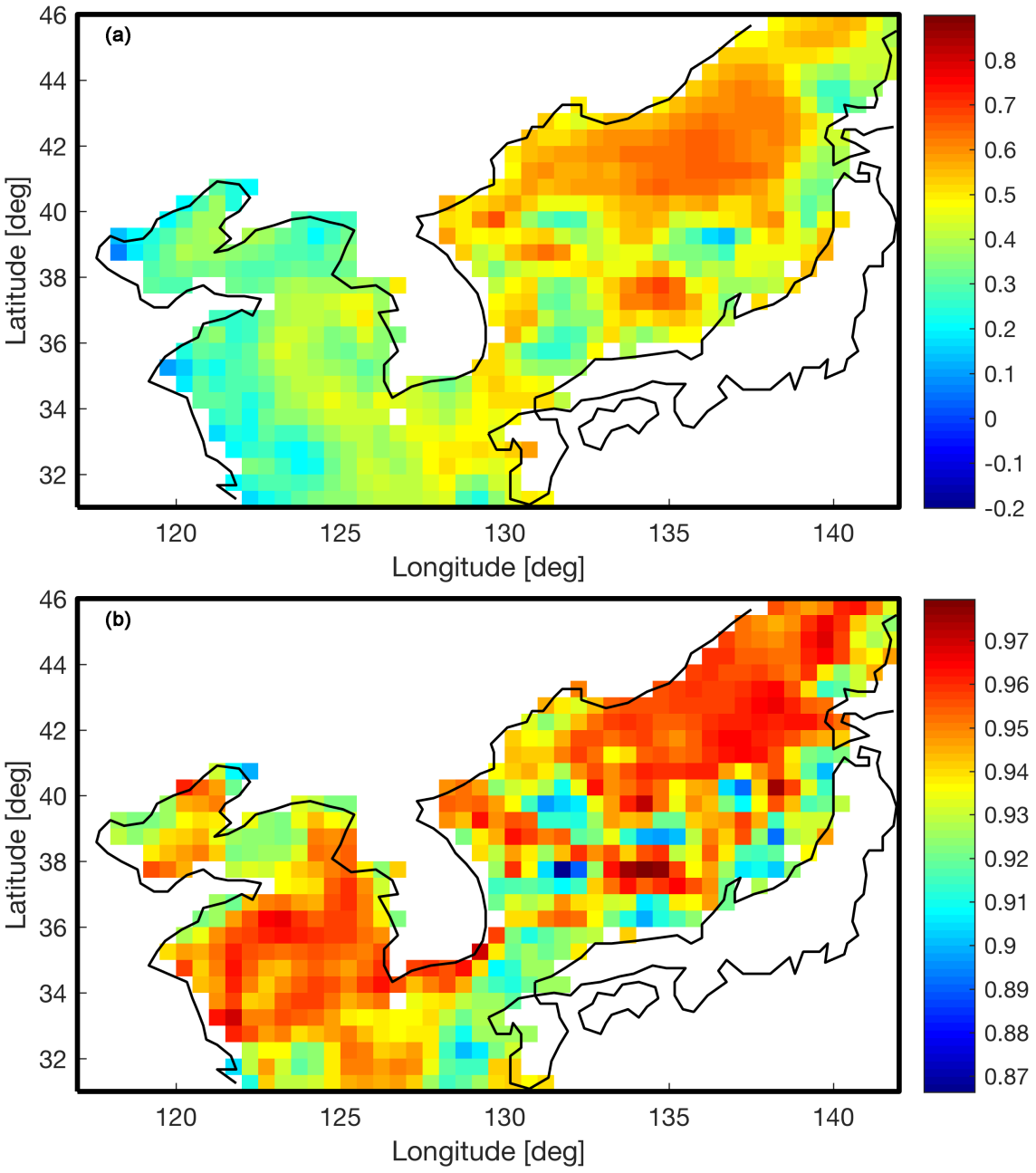

Figure 17. (a) Correlation coefficient map between Hamlington et al. (2011) and AVISO-KP over 1993-2008; (b) Correlation coefficient map between this study and AVISO-KP over 1993-2008

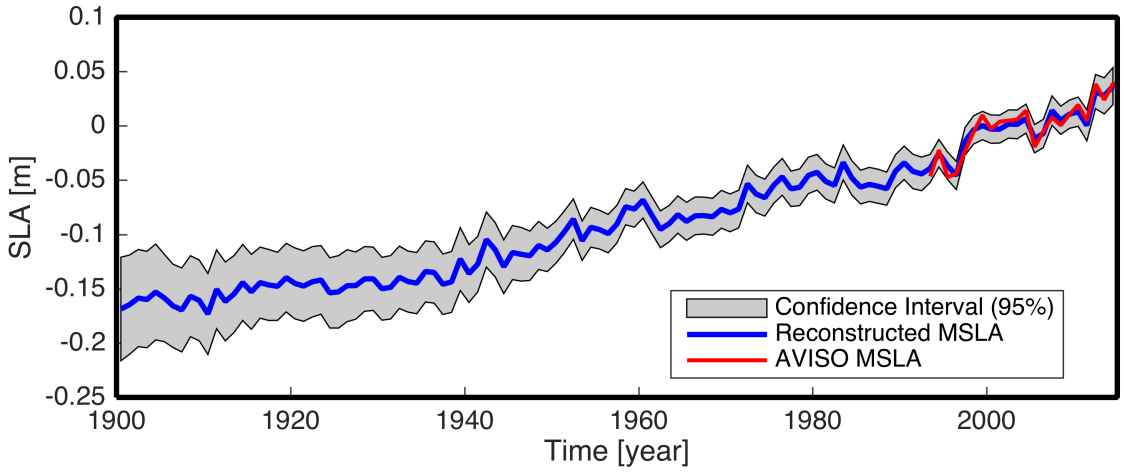

Figure 18. Spatial mean time series of sea level anomalies of the best recostruction case (COBESST2-NWP) and 95% confidence interval.

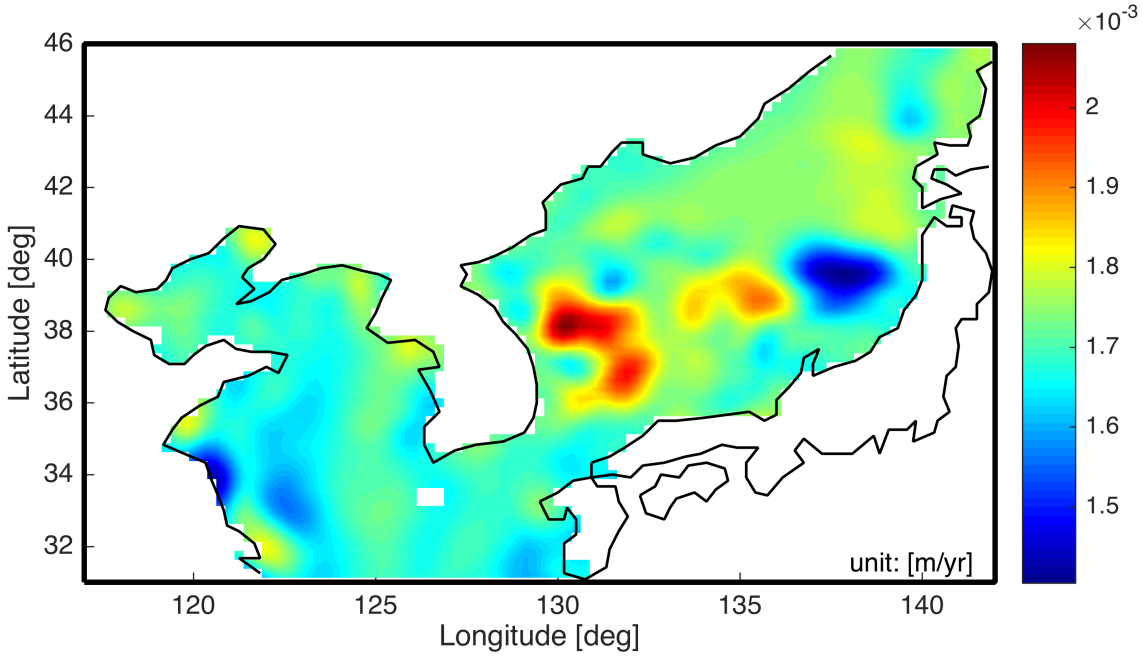

Figure 19. Linear trend map of the best reconstruction of current study from 1900 to 2014