

Response to the major comments

We thank the reviewer for the comments and insight into the paper. We have made many adjustments to the paper and have added clarifications where necessary. We think the paper is much improved as a result.

(a) The sea level change may be associated with many factors such as ocean temperature (including SST), salinity, currents, and surface winds etc. Therefore, the multi-regression between SST and SL PCTs may not include all aspects of SL changes. I am wondering whether the reconstruction could further be improved if more physical variables are considered.

To apply multi-variables to current reconstruction scheme, there are several problems. First, when we applied multi-variable's PCTs as predictor, the over-fitting occurred because as the mode goes higher, the possibility of overfitting increases. Second, the reanalysis of other variables does not appear to be of sufficiently high quality back through time, based on our tests. Actually, we tried to reconstruct SLA using wind and SST data simultaneously, but the result is poorer than each individual reconstruction case. To solve these problems, we it will require significant effort beyond the scope of this first study with the goal of demonstrating the basic technique. However, we do agree the reconstruction applying multi-variables is valuable topic for the future study.

(b) The SL reconstruction does not include TG observations, but have a clear improvement over a similar reconstruction that includes TG observations. I am wondering whether the SL reconstruction could further be improved if all available TG observations are included.

To include TG data, there are two problems. First, using TG data we cannot conduct CSEOF analysis because they have lots of discontinuous points and their spatial coverage are too poor. So we cannot establish the proper regression relationship between TG and SST. Second, the TG data's quality is not good. The vertical land motions cannot be explained and accounted for and many of the Japanese TG are impacted by earthquakes or volcanic activities. To use TG data, the additional researches are necessary to correct the vertical land motions.

(c) How to validate the SL reconstruction in the early period over 1900-30 when no TG observations are available. It might be a little risky to include the reconstruction in this period.

This is indeed an issue with any reconstruction study. As we can see in Fig. 11, even we have TG data for vilification, prior to 1970, the agreement between MSLA from TG and ReSLA-KP is very poor. But we thought that it does not mean our reconstruction is not good because the TG data is not enough to verify. Even we cannot verify the reconstruction results, we think the result is still valuable. However, we have tried to make clear in the paper that the quality of the reconstruction is difficult to assess into the past.

(d) Writing and presentation may need improving. There are too many abbreviations such as SL, MSL, GMSL, SL-KP. For example, MSL and GMSL could be explained in figure captions. KP is unnecessary because the study focuses on KP region only.

We have modified the abbreviations. We have, however, left in the KP in some locations, to distinguish from the global terms.

(e) Figure captions should identify the data source and average region etc.

We have made this change to the figure captions.

Detailed comments

P1L11, revise: extend the spatial resolution ..into the past

Original	Many studies have been conducted to extend the spatial resolution of the satellite data into the past by finding novel ways to combine the satellite data and tide gauge data in what are known as sea level reconstructions.
Revised	Many studies have been conducted to create a dataset with the spatial coverage of the satellite datasets and 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 a sea level reconstruction.

P3L5, CSEOF is not defined

We have now defined CSEOF in P2L22.

P3L14, “KP” could be deleted throughout the manuscript since the study has been limited over the KP region anyway, which will greatly improve the readability. “KP” could be noted in the figure caption when necessary.

We have deleted ‘KP’ and add explanation about the default domain.

P3L21, revise: looking at the regional level will lead to

Original	As mentioned above, TG coverage is poor extending back into the 20th century, and looking at the regional level will lead to relatively few gauges to analyse in most areas.
Revised	As mentioned above, TG-KP coverage is poor extending back into the 20th century, with many areas completely lacking tide gauge records.

P4L11, annual signal=> seasonal signal?

The terms, annual signal and seasonal signal, are having same meaning. But to prevent confusions, we change seasonal signal to annual signal.

P4L20, include data => included data?

We have fixed it.

P4L21, over => from?

We have fixed it.

P6L5-7, revise the sentence

Original	If previous reconstruction schemes are applied that rely only on sea level, then it is likely only possible to obtain reliable results after 1970.
Revised	If a reconstruction scheme for SLA-KP relies only on TG data, then the results are only reliable after 1970 when TG coverage improves.

P6L11, delete “in this case”, “really” P6L12, independent of => independent from?
We have erased this sentence.

P9L3, How does “summing” actually do, arithmetic or squareroot?
We mean Root Sum Square. This part has been deleted.

P10L10, this is an indication that SL is not merely dependent on SST.
We have deleted this part. And we gave up to explain the physical reasons for the extreme SLR values.

P11L22, delete “then”
We have deleted it.

P11L25, delete “cases of”
We have deleted them.

P11L28-29, delete “considering the available number of TG data”
We have deleted them.

P12L4, It is not clear how MSLA-KP is defined (assuming every ocean grid in reconstruction). How MSLA-KP can be compared with TG-KP (only in TG grids).
We have added more explanations as follow.

To check the reconstruction results, we calculated MSLA of TG-KP, ReSLA-H, and ReSLA-KP. Spatial mean was calculated for the two grid datasets. For TG-KPs, we calculated mean differences between each time steps and we integrated the differences. The integrated mean differences became the MSLA of TG-KP.

P12L12, revise “was edited to have the same time span data gaps”
We have erased ‘data gaps’.

P12L14-15, revise the sentence: ReSLA-KP show a better agreement of AVISO-KP than ReSLAH.
We have revised.

P12L17-18, how many modes are used in Hamlington?
We have added detailed number.
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.

P12L23, thousand => a thousand
We have corrected.

P13L17-18, authors should extend the conclusion of a better current SL reconstruction. there is no way from Figures 16-17 to tell the current study is better.

We have deleted this part.

It is not clear in Figure 13 either. It may be necessary to point to Figure 14a. A better way is to calculate the RMSE.

We have added more figures.

Fig. 1, digital quality should be improved.

I think it has a high resolution, 600 ppi. We will ensure that it is of sufficiently high resolution in any future submission.

Fig. 2, coastal line should be consistent with those in other figures.

We have changed the figure.

Fig. 3, I could caption the figure as “Mean SLA in KP (gray) and global (black) regions from AVISO” so that I can get rid of some abbreviations.

We have changed the caption and figure.

Fig. 4, add “AVISO” in caption

We have modified.

Fig. 5, add “AVISO” in caption

We have modified caption.

Fig. 6, revise: trends (shapes) and correlation (color), change the red color of triangle into black so that the color will not be confused with correlation.

We have modified the caption and figure.

Fig. 7, NRMSE, I don't know the advantage of using normalized RMSE instead of RMSE.

‘NRMSE’ and ‘RMSE’ very similar, but when NRMSE has ‘zero’ value this means the regression is same with some constant value cases and if the value are negative that the compared data is less agreed its mean value. So, we believe that NRMSE gives some intuitive interpretation.

Figs. 8-9, I am confused how the 3-month averaged mode is plotted. I assume there is only one CSEOF associated with one PCT for a particular mode.

Yes, you are right. But the evolution is small through the months, therefore we represent the results as seasonal mean values to save some space.

Fig. 10, I assume this is for KP region

We have modified the caption.

Fig. 11, which region, KP region?

We have modified the caption.

Fig. 12, Why does Hamlington have a constant Corr and NRMSE?

Because we have 6 cases but ReSLA-H is just one case.

Fig. 14, “yellow” is barely identifiable. Why the correlation is over 1993-2008 while trend is over 1970-2008?

I think the color problem is related to the resolution, I provide 600 ppi image and that figure has no problem to recognize yellows.

And the time period is 1970-2008 for the both cases.

Fig. 15, The figure look great but there is a question: Since the study uses the CSEOF derived from AVISO, therefore validation against AVISO is considered to be not independent. One may argue that if authors use Hamlington deriving CSEOF, the performance reconstruction may be close to Hamlington.

Yes, that’s right. Nevertheless, ReSLA-H has very poor agreement. We just want to show the limit of global reconstruction, as you can see the below figure. Over 1993-2015, the correlation coefficient is pretty high, this means if we applied Hamlington et al. (2011)’s method in local scale, the correlation coefficients must be higher than current Fig.

