Rebuttal – reviewer 1

We would like to thank the anonymous reviewer for the thorough review of our manuscript. We update the manuscript such that it answers the questions and implements the recommendations of the reviewer. Below we wrote a point-by-point response to the reviewer comments.

General comments:
Accurately determining vertical land motion at tide gauges is an important scientific issue with crucial societal implications associated with future relative sea levels at the coast. The study by Kleinherenbrink et al builds upon the most recent estimates of vertical land motion from GNSS data analyses and the combination of satellite altimetry and tide gauge data. The authors perform a detailed and honest critical review of the estimates available from the literature, while they provide ways to overcome some of the limitations. For instance, wherever there is no permanent GNSS antenna at the very top of the tide gauge (co-location), but multiple GNSS receivers are in the vicinity, they explore different methods to deal with this situation. In addition, they delve into the details of the best possible way of deriving estimates from the combination of satellite altimetry and tide gauge data with insightful outcomes too.

The manuscript reflects a sound scientific approach. The methods applied are clearly outlined. Some minor technical details are missing, however, and require clarification (see below). The results are discussed in detail, and overall the results and discussion provide a substantial contribution to the area of research on determining vertical land motion at tide gauges. In addition, the manuscript is well structured, clear and concise, and the conclusions are supported by the data. A somewhat negative note is that I miss that the authors are not providing their best estimates on vertical land motion (with the error bars) in a supplemental material. Similar to the studies they build upon, they should provide their estimates for future investigation. Perhaps this can be considered by the authors for the final version. In conclusion, my suggestion is a minor revision before publication.
We were already planning to make the data publicly available. The vertical land motion estimates for all altimetry-tide gauge correlation settings and the median GNSS approach are now provided in the supplementary material. The aforementioned sentence is adjusted accordingly.
The technical details and other comments are discussed below.

Other (minor or technical) comments:

p.1, Title: The term “weighting” does not correspond to several of the approaches examined in this study. See also 1st and 10th lines in the abstract). In addition, I would change “derive” to “estimate” to underline that behind the scenes the results from these methods are based on an estimation procedure, not directly observed.
We changed the title to: “A comparison of methods to estimate vertical land motion trends from GNSS and altimetry at tide gauge stations.” The text is adjusted as well, so ‘approach’ is used.

p.1, Lines 2-3: It should be clarified that these methods are considered to deal with the situation of multiple GNSS stations nearby a TG.

Two sentences are added at the beginning of the abstract and the existing text is adjusted accordingly. "Global Navigation Satellite System (GNSS) are usually not co-located with Tide Gauges (TGs). Therefore trends from neighbouring GNSS stations are combined to estimate a VLM trend at the TG. This study compares eight methods to estimate Vertical Land Motion (VLM) trends at 570 TG stations using GNSS."

p.1, Line 20: Ostanciaux et al. did not established the magnitude that can reach the GIA effect. I suggest to quote an original early reference such as Gutenberg, in Bull. geol. Soc. Am. (1941).
The reference to Ostanciaux is replaced with the one to Gutenberg.

p. 1, Lines 21-22: The statement that trends at TGs are affected by erosion is not obvious to me. Please, quote a reference that demonstrates this relationship.
We changed the sentence. Erosion and gas extraction are removed. So now the sentence reads: "including water storage, postseismic deformation and anthropogenic activities (references)."

p. 3, Line 26: For the sake of consistency, I wonder why Hector is not applied for the GNSS trends too. Can you develop the §with your arguments, please?
We wanted to have the best possible GNSS trends for which we do not have to apply any screening. The trends are strongly affected by jumps in the time series. Based on the Blewitt et al. (2016) their MIDAS method has the smallest equivalent step size detection. Therefore we selected this method. We cannot apply the same method to the altimetry-tide gauge time series, since the time series cover only ~200 months, which is rather short for the MIDAS approach. We rephrased several sentences in Sect. 2.1.1 to clarify the reasoning.

p.3, Line 31: The issue is primarily that the differential land motion between the GNSS antenna and the tide gauge is not monitored locally, for instance via repeated levelling campaigns. Thus, a lack of information.
A sentence is added that addresses this issue.

p4. Line 7: I guess "However" is not correct here. Considering revisiting this since the decrease in accuracy is not associated with the use of the software and its advantages.
We rephrased the sentence and split it up in two separate sentences. "The software is also able to estimate and detect discontinuities that occur due to earthquakes and equipment changes. Even though a large fraction of the trend estimates have formal accuracies better than 1 mm/yr, undetected discontinuities might significantly bias the estimated trends (Gazeaux et al., 2013)."

p4. Line 11: the term measurements is not appropriate here, the positioning time series are outcomes (estimates) of the measurements analysis.
The term measurements is replaced by estimates.

p.4, Line 12: Please, develop how the scaling is performed (what is its origin).
We added an equation and a reference to Wilcox (2005).

p.4, Line 12: typo in "devations", should be "deviations"
Updated.

Section 2.1.1: did you screen the GNSS time series for apparent transient processes that would impact (question the validity of) the linear trend estimation?
No, we did not screen the time series, because we use the pre-computed trends from MIDAS. Any non-linear behavior might bias the trend, but it will also inflate the error bars as described in the section. We now clearly state that we do not apply any screening.

Section 2.1.2: See above my comment on the term "weighting". Within this section you use the term "approach" which is definitely more appropriate.
The term weighting is either replaced by 'method' or by 'approach' throughout the manuscript.

p.4, Line 17: at some point (here or later in the manuscript) you should discuss this vague statement "a record long enough".
We changed the sentence, such that: "currently only a few have a record that ensures a trend accuracy of 1 mm/yr or less".

p.5, Lines 3-5: You detailed the "obvious" relationship of method [7], you should detail that of method [8], which is less obvious to me.
An equation is added for method [8].

p.5, Line 7: Holgate is published in 2013 (not 2012). See also reference list (p.22, Line 33).
The reference is updated.

p.5, Lines 22-25: Please, rephrase. I had to read the sentence several times. Consider splitting it into two sentences.
The sentence is rephrased.

p.5, Line 30: Please, develop the rationale for 250 km (why not 200 km, or 270 km, or ...).
Outside of the equatorial regions and the continental shelves, ocean correlation scales are below 250 km (Ducet et al. 2000; Roemmich et al. 2009), so we do not expect significant improvements if observations outside of the 250 km range are included. We could probably find some long-shore correlation along the shelves over longer distances, but it would not be appropriate to take those observations into account, since long-term trends do not have to resemble anymore, i.e. large-scale signals like GIA trends are not equal to the TG location anymore. On top of that, at least one track of the altimeters is always passing through the 250 km region. Making the radius smaller, reduces the number of observations substantially, especially at lower latitudes.

p.7, Table 2: The information conveyed by this table is too technical. Consider moving it to an Appendix or Supplemental material. Clarify what are these differences (related to J1? TP-J1, then J2-J1?). In addition, add error bars to the parameter estimates, and/or say if all these parameters are statistically significant at the 95% level.
We moved the table to the appendix.
The caption is extended, to clarify what the differences mean.
*Since we do not apply a full error propagation on these values, it is difficult to determine whether they are statistically significant or not. If we use the variances (sigma*2) of the residuals to compute the, then the errors for the coefficient c=sigma^2*(A^T*A)^-1, several of coefficients are not statistically significant (primarily in the equatorial regions for TP-J1). This is primarily, because we average*

the altimetry differences per latitude band (1/8 degree wide) first and then compute the polynomials. The degrees of freedom (~10 dof) is therefore rather small to estimate proper statistics for the equatorial regions. We therefore stick to the polynomials as used in the Ablain's paper, which is referenced to in the text.

p.8, Line 4: "are computed" should be "is computed".
Updated.

p.9, Lines 6-7: The sentence has a problem. I don't understand, please rephrase.
The sentence is rephrased.

p.9, Line 15: What is the rationale for the 50km radius. Please, develop.
Most studies involving sea level include observations within radii of 10-100 km. We took the radius right in the middle, but we could have increased or descreased the radius. A radius of 100 km would include observations with errors due to local VLM of more than 0.5 mm/yr on average (Santamaria-Gomez et al., 2017), while taken a small range reduces the number of trends substantially. Tests, however, demonstrated that similar results are obtained for 30 and 70 km, but with slightly less or more trends estimates, respectively. Several sentences are added in the methodolgy section.

p.10, Table 3: Consider adding a mnemonic keyword (after the number) to designate the approach, for instance "closest", "longest", etc.
We added keywords in the table.

p.13, Lines 7-8: Can you quantify the amount of reduction using equation (4)?
Yes, we can. The median of the spectral indices (for the same stations) is closer to zero for higher correlation settings. We added a line with the statistics.

p.20, Line 2: Strictly speaking, "observations" is not appropriate (estimates? Data?)
We guess this should be line 7. The term observations is replaced with data.

Rebuttal – reviewer 1

We would like to thank reviewer Alvaro Santamaria-Gomez for his comments that helped to improve and clarify the paper. We addressed the comments below and changed the manuscript accordingly.

General comments:
This paper addresses the methods of estimating the linear trend of the vertical land motion (VLM) at tide gauge (TG) stations using GPS and satellite altimetry minus TG observations (ALT-TG). Since the satellite altimetry and most of the GPS data are not provided at the TG location itself, the paper focuses on the different choices to extrapolate these datasets to the TG location. Some of the investigated choices have been used in past sea-level studies and are relevant for comparison purposes, whereas others (especially the treatment of satellite altimetry observations) are new in this paper. The comparison of the different choices provides valuable information to other scientists working on this subject and the preferred choices of the authors lead to a reduction of the VLM differences between GPS and ALT-TG techniques compared to previous studies.

The writing is clear globally, but some sentences (details given below) need clarification or correction. Parts of the methodology need also clarification (for instance, concerning the pole tide or the use of errorbars). The authors focus on describing the results of their analysis without going in depth with their discussion and implications, which undermines the conclusions to some extent.

Specific comments:

The title: I would suggest changing "data weighting methods" by "methods" or "approaches".
The title has been changed to: "A comparison of methods to estimate vertical land motion trends from GNSS and altimetry at tide gauge stations."

The abstract needs to be improved to make it more clear and self-contained. As it is now, it looks like a compressed listing of the results so it may be hard for the readers to understand without a minimal background and way out (recommendations or take-home messages).
We extended the abstract with several sentences where we focus on the recommendations that are given in the conclusions.

P1L18: several VLM processes are modelled. This is very ambiguous. It may be true (we could model anything), but only GIA models are actually being used. Later on, it's said that local VLM processes cannot be captured by models.
We removed the sentence. We added: "The large scale VLM processes, such as Glacial Isostatic Adjustment (GIA) and the elastic response of the Earth due to present-day mass redistribution can be modelled to accuracies close to 1 mm/yr. However, TG are often only corrected for the GIA signal, which typically … ".

P2L11: Actually, Santamaría-Gómez et al. 2017 did not conclude on the accuracy, but they show bigger differences between ULR and NGL than ULR and the other solutions being compared. Compared to the other solutions, NGL velocities also had larger errorbars to accommodate these differences.

We changed the sentence to: "They concluded that the number of stations in the NGL database was larger, but that the differences between neighbouring stations was significantly larger than the Jet Propulsion Laboratory (JPL) and ULR6 solutions."

P3L16-20: In addition to the ocean signal, the ALT-TG correlation can be used to infer the correlation between the TG record and the VLM *of* the TG itself, especially with low-pass filtered series as you did, for instance if the VLM at the TG is not linear during the altimetry period. This is inseparable from the ocean signal you mention (see discussion in Santamaría-Gómez et al. 2014, JoGE).

While the largest interannual signals come from the ocean, time series affected by for example discontinuities caused by earthquakes or other non-linear VLM behavior are also removed, because they will not have a high correlation. We tried to address this point on page 17, but we added a part to the sentence on page 3 to make sure this is stated.

P4L5-7: Note that all the ULR solutions have been computed using the CATREF software (Altamimi et al. 2016). CATS has been used to re-estimate the trend uncertainty, but the estimated trend does not change statistically. The "slight change in trend" comes together with the increased uncertainty and it is just the consequence of inverting a more complex covariance matrix in time with probably a small contribution also from the different use of spatial covariance between CATREF and CATS.

We agree that "The slight change in trend" is not significant and it leads to confusion. Therefore this part of the sentence is removed.

P4L29: I assume the approaches using the longest time series (5) and the smallest error (6) are also using the closest GPS station, but it would be better to clarify.

No, they do not. They take the longest time series, or the time series with the smallest error within the 50 km radius. We changed the sentences to clarify this.

P5L3: It would be better to add here the equation of the approach (8) to see how the approaches (4 distances) and (7 weighted mean) are combined. To me, this is in theory the best approach since it uses more information available that the other approaches. However, the way the distance and uncertainty are combined may still be very important. Also, the propagation of the VLM uncertainty from the GPS to the TG should be commented on as it varies for each approach.

We added the equation for approach 8 and we agree that in theory approach 8 would be the best. A possible reason that it does not perform best is probably the limited number of tide gauge for which we have multiple GNSS stations. The statistics are therefore obscured by several 'exceptional cases'. We can find multiple examples in the Netherlands where a close distance is not a good indicator for the representativeness of the VLM at the tide gauge, mainly due to anthropogenic effects, for example local gas extraction or groundwater fluctuations. Uncertainties are not provided directly, because it would in several variants not respresent a proper estimation of the real uncertainty. For example, the weighting used in method 5 (longest) would just give you the GNSS trend uncertainty, but it does not include any relative local VLM information in the uncertainty. It is therefore definitely underestimated. We also do not use any uncertainty information in the rest of the manuscript. In the additional material we provide the trends and uncertainties and distance to the GNSS stations from the tide gauges. On top of that we provide the best (median) solutions with error bars. We added a line about this in the *conclusions. If somebody wishes to compute the uncertainties for the other methods, it can easily be done by propagating the uncertainties of the individual GNSS trends.*

Section 2.3: I am confused about which altimetry series did you use and when. You say that an "additional" filtered set was used to test interannual correlation (P6L10-12) and that before estimating the correlation, you removed residual seasonal cycles (P6L15). So, where do these residual seasonal cycles come from if the series were filtered? Is the yearly moving-average not enough to remove unmodeled Sa tides from altimetry or the low-pass filter allows for annual variations at the TG? Finally, the filtered series were kept for the analysis (P7L1), but Figures 1 and 5 show both filtered and unfiltered series.
The time series are filtered with a moving-average low-pass filter of a year. A moving-average low-pass filter does not completely remove semi-(annual) cycles, which you can see by the response in the frequency domain. An alternative would be to use a different filter, but this leaves larger transient zones at the beginning and the end of the time series. We added a sentence to clarify this.
It is slightly confusing indeed. Basically we create two low-pass filtered time series: one monthly, one yearly. The yearly one is used to determine interannual correlation. If the yearly ones have a correlation with the TG time series higher than the correlation threshold, the corresponding monthly ones are kept. The other ones are removed. In the figure we averaged the remaining monthly ones, and then low-pass filtered them to show that the interannual signals are reduced. We have rewritten several sentences and updated the caption to clarify this.

P6L13-15: Is it necessary to remove the ocean pole tide from the ALT and TG records? Are they significantly different? Concerning the solid Earth pole tide, I would suggest adding that the RADS solid Earth pole tide model is consistent with the Desai's model concerning a linear mean pole trajectory, so that the interannual vertical deformation is preserved in the TGs when subtracting one and adding the other (I assume this was the purpose, but it could be said explicitly). However, what is the rationale for adding the IERS solid Earth pole tide to the TG records after removing the RADS model (P8L1-3)? Contrary to the RADS or Desai's models, the IERS solid Earth pole tide model does not correct the interannual deformation (see King and Watson, 2014). The interannual deformation was removed by the RADS model and is not restored by the IERS model. In doing so, the ALT-TG VLM will not be consistent with the GPS VLM that is still affected by this interannual deformation from the IERS model. If I understood your treatment, I think you should add the Desai's model in both cases.
The ocean pole tide can be as large as 2 cm and will therefore affect the correlation parameters. The RADS pole tide includes contributions from the solid Earth, loading and ocean tide tides. The altimeter is affected by all three. The tide gauge is only affected by the ocean tide. We subtract however the full RADS pole tide from both of the time series. To be consistent, we have to add the loading and the solid Earth tide. The loading tide is very small (typically 10 % of the ocean tide), so it barely has an effect on the correlation and therefore we ignore this term (especially after filtering). The solid Earth tide amounts to almost a centimeter and even though it will probably not affect the correlation, we add it back to be safe. In both cases Desai's model is used. The interannual deformation, or the non-linear part of the mean pole, is therefore corrected for in both cases.

P8L5-8: note that the IERS conventions were updated about this issue in June 2015, and even though the issue still persists, most of the GPS VLM estimates are based on the old IERS implementation, at least the ULR and

NGL solutions you used. The 0.1 mm/yr error arises in a regional sea-level reconstruction using GPS-corrected TG records with old IERS model. The VLM effect at individual GPS sites may be 3 times larger (King and Watson, 2014). Explain how this error is corrected using the mass redistribution fingerprints. Section 2.4 could be integrated into the 2.3.
We therefore apply the old solution to get the ALT-TG trends in the same system as the NGL solutions. This means that the trends can differ as mentioned in King and Watson, 2014) due to the non-linear drift of the pole. The non-linear drift of the pole is primarily caused by the melting of the ice sheets at locations away from the rotation axis of the Earth (mostly Greenland), which is captured by the sea level equation (it includes rotation changes). Several sentences are added to section 2.3 to clarify this.
We keep section 2.4 and 2.3 separate, because in the results we discuss the solutions without the present-day mass redistribution correction first separately.

P9L10: change ULR by ULR5, which is the solution used by Wöppelmann and Marcos, 2016
Updated.

Section 3.1 and elsewhere: direct/indirect are ambiguous terms. I would suggest using GNSS and ALT-TG for consistency.
We changed the terms to GNSS and ALT-TG.

Figure 3 and elsewhere: change spread by range
The term range is used instead of spread.

P10L7: change solutions by weighting methods for consistency or even to approaches, which may be more appropriate.
The term weighting method is not used anymore in combination with GNSS trends throughout the manuscript.

P11L1-4: The range values are driven by the extremes, which are obtained from the "mean", "median" and "inverse distance" approaches. None of these approaches is using the information provided by the VLM errorbars, which can be as large as 1 mm/yr, and only the "median" approach is less affected by outlier VLM values (but only if we have a large sample and we assume the VLM estimates in 50 km follow a Gaussian distribution, which may not). I would suggest using the interquantile range instead of the range to evaluate the dispersion of the different approaches.
Instead of providing the mean and the median, we now give the mean and the 25-75 % percentiles.

P11L6-7: In relation to my comment before, these global estimates of spatial variations of VLM were given as 1 sigma standard deviations. You would have to multiply them by 5 or more to obtain something close to the range of the extremes (for instance, by 10 in areas with strong GIA gradient). On top of that, a global figure will never fit all locations which will be underestimated or overestimated.
We adjusted the sentence and removed the word underestimation.

P11L11-16: Table 3 shows the VLM differences at 70 TGs between using the closest ULR5 value and 8 different approaches with the NGL velocities. It is surprising that the RMS of the differences is the highest for the closest NGL value (approach 3), which will use the same GPS station as in ULR5 for many TGs, whereas it is minimum for the median of the NGL values 50 km

around the TG (approach 2). The WRMS of the differences between ULR6 and NGL is about 0.7 mm/yr. You are using ULR5 and not ULR6 here, but the RMS for the closest NGL station is two times larger and appears unreasonable to me. It may be due to the VLM errorbars not being used. Also the ranking of the methods in this table and that in Figure 7 matches exactly as if the ULR5 velocities were providing the same benchmarking information as the ALT-TG trends. Is this coincidental?
We remove all GNSS trends with uncertainties larger than 1 mm/yr. This removes GNSS stations that might be co-located with the tide gauge. If then the closest station is used from NGL, it might be that it is not the same station as used by ULR5. In that case the closest station method depends on a single station not co-located with the tide gauge and therefore it is likely that some outliers are present. From the 70 stations, we find three trend differences larger than 3 mm/yr for the closest station method, while only one for the median method. We added a short discussion on this matter.

Figure 4: Change "reduction" by "change" or invert the sign of the scale for consistency (positive reduction is good, otherwise is bad).
We changed the word reduction to change.

Figure 7: It would be easier to read the legend if the mean RMS of each line, with fairly constant values, is added on the right of the figure, for instance.
We added the mean RMS in the figure.

P16L6-8:
[A] Please explain how the median takes into account the standard deviation of the GNSS trends as in a weighted mean (approach 7).
Also, any approach using more than one GNSS trend in 50 km around the TG is filtering the spatial variations in VLM, including the variance weighting (weighted mean) approach.
[B] From these lines on, it is decided that the median approach is the best candidate, but I'm not fully convinced and I would suggest adding more discussion on these results. For instance, the fact that a simple median provides better results than the more complex approach of including distance and uncertainty information needs better discussion. The combination of the distance and errorbar information is not trivial and may depend on the TG location, so this may have flawed this approach.
[C] However, even the weighted mean is using additional relevant information, but it is ranked after the median and the mean. This makes me think whether the evaluation using the ALT-TG trends is the best benchmark. For instance, the ALT-TG VLM uncertainties are probably large as well, with important variations among the TGs (correlation, length of the series, etc), and it seems to me that they were not used for the benchmarking either.
[D] On the other hand, the alternative explanation would be that the trend uncertainties of the NGL solution are not providing a useful value of their precision. For instance, it is known that there are trend biases not explained by their formal uncertainty and caused by a combination of the time series length and non-linear effects like seasonal signals, discontinuities, interannual deformation, transients, etc. Different processes would also bias the ALT-TG trends (orbital error, altimeter bias drift, etc.).
[A] Suppose that we have no relative vlm movement in the area. Then the expected value between the tide gauge vlm trend and the observed GNSS trends is zero. The

GNSS trends with larger uncertainties are likely to have a deviation further from zero. Therefore it is likely that the median value is closer to the GNSS trends with smaller uncertainties. Besides, any 'outlying' values do not affect the median. Now suppose that the have relative vlm movement in the area. When variance weighting is applied, and the GNSS station with the large relative difference has by coincidence the lowest variance, it will get the highest weight, while actually it is the worst proxy. In the case of median weighting, the median will not be affected by this outlier. We rephrased the sentence, such that: "The median method is less sensitive to large values caused by GNSS trends with larger uncertainties (for which the mean method is sensitive) and also less to outliers caused by large local VLM differences (for which the variance weighting method is sensitive)."
[B] The distance-variance weighting approach does is more sensitive to the distance than to the variance, especially because the maximum uncertainty for the GNSS trends is set to a maximum of 1 mm/yr. For example, a trend found at 10 kilometer distance is already 10 times weaker than one at 1 km, while the uncertainties have often similar values, mostly 0.7-1 mm/yr. Therefore it effectly reduces the number of GNSS trends being used. We added a note in the methods section that the method is strongly depending on the distance. As a recommendation, we mention that using another distance weighting method might be better, but that it would require information of VLM correlation distances to find an optimum.
[C] The VLM uncertainties where indeed not used for benchmarking, because the uncertainties for the GNSS methods are not trivial and probably do not properly respresent the true uncertainty, because the uncertainty information due to relative VLM between GNSS and the TG is not present. As mentioned in [A], the weighted mean can be strongly affected by outliers due to local VLM differences. Just to give an indication of the uncertainties of the ALT-TG time series, we added some statistics to the ALT-TG results. Of the 663 trends for no correlation threshold, 293 have an uncertainty smaller than 1 mm/yr. Of the 344 trends for a correlation threshold of 0.7, 284 trends have an uncertainty smaller than 1 mm/yr.
[D] The MIDAS method takes annual differences in vertical location, so the seasonal signals are reduced to the minimum. Interannual deformation, transients, etc. will widen the distribution of the annual differences. Therefore the uncertainties increase. Since we use a maximum of 1 mm/yr on the uncertainty, it will therefore remove all trends computed from time series with substantial earthquake activity (see for example Japan) or interannual signals, like groundwater storage. We added a line in the GNSS methods section to state this.
Altimeter stability is guaranteed up to 0.4 mm/yr. If the altimeter would really be drifting with 0.4 mm/yr, this would increase the mean of the ALT-TG vs GNSS differences, but this has only a 0.06 mm/yr effect on the RMS. Temporally varying orbital errors would show up in the ALT-TG time series, so their contribution is captured in the uncertainty estimates produced by Hector.

P2L20 and elsewhere: the correct reference for the ULR5 solution is Santamaría-Gómez et al 2012 Glob. Planet Change.
The reference is added.

I fully agree with the last sentence and I would add that, whenever possible, one should always inspect the data being used. A much extended (and faster) practice is always using the trend uncertainties together with the trends, because they (should) carry relevant information on the linearity of the observed series.
We added a final sentence that the error bars carry relevant information about the non-linearity of the time series, but when a single station is present it does not contain any information about local VLM variability.

A comparison of ~~data weighting~~ methods to ~~derive~~ estimate vertical land motion trends from GNSS and altimetry at tide gauge stations

Marcel Kleinherenbrink[1], Riccardo Riva[1], and Thomas Frederikse[1]

[1]Department of Geoscience and Remote Sensing, Delft University of Technology, P.O. Box 5048, 2600 GA Delft, The Netherlands

Correspondence to: Marcel Kleinherenbrink (m.kleinherenbrink@tudelft.nl)

Abstract. ~~This study compares eight weighting techniques for~~ Tide-gauge (TG) records are affected by Vertical Land Motion (VLM), causing them to observe relative instead of geocentric sea level. VLM can be estimated from Global Navigation Satellite System (GNSS) ~~derived~~ time series, but only a few TGs are equipped with a GNSS receiver. Hence, (multiple) neighouring GNSS stations can be used to estimate VLM at the TG. This study compares eight approaches to estimate Vertical Land Motion (VLM) trends at 570 ~~tide gauge (TG) stations . The spread~~ TG stations using GNSS, by taking into account all GNSS trends with an uncertainty smaller than 1 mm yr^{-1} within 50 km. The range between the methods has a comparable size as the formal uncertainties of the GNSS trends. Taking the median of the surrounding GNSS trends shows the best agreement with differenced altimetry - tide gauge (ALT-TG) trends. An attempt is also made to improve VLM trends from ALT-TG time series. Only using highly correlated along-track altimetry and TG time series, reduces the standard deviation of ALT-TG time series up to 10%. As a result, there are spatially coherent changes in the trends, but the reduction in the RMS of differences between ALT-TG and GNSS trends is insignificant. However, setting correlation thresholds also acts like a filter to remove problematic TG ~~stations~~time series. This results in sets of ALT-TG VLM trends at 344-663 TG locations, depending on the correlation threshold. Compared to other studies, we decrease the RMS of differences between GNSS and ALT-TG trends (from 1.47 to 1.22 mm yr^{-1}), while we increase the number of locations (from 109 to 155), Depending on the ~~weighting~~ methods the mean of differences between ALT-TG and GNSS trends varies between 0.1-0.2 mm yr^{-1}. We reduce the mean of differences by taking into account the effect of elastic deformation due to present-day mass redistribution into account. At varying ALT-TG correlation thresholds, we provide new sets of trends for 759 to 939 different TG stations. If both GNSS and ALT-TG solutions are available, we recommend to use the GNSS solutions, because redisual ocean signals might correlate over long distances. However, if large discrepancies ($>$ 3 mm yr^{-1}) between both methods are present, local VLM differences between the TG and the GNSS station are likely the culprit and therefore is is better to take the ALT-TG solution. Especially GNSS estimates where only a single GNSS station and no ALT-TG solution is available, might still require some inspection before they are used in sea level studies.

1 Introduction

Tide Gauges (TGs) measure local relative sea level, which means that they are affected by geocentric sea level, but also by Vertical Land Motion (VLM). Knowing VLM at TGs is essential to convert the observed sea level into a geocentric reference frame, in which among others satellite altimeters operate. TGs used in sea level reconstructions also require a correction for VLM. The mean of VLM at TGs is not equal to that of the basin, and therefore local VLM estimates are required to get an accurate estimate of ocean volume change. ~~Several VLM processes are modelled. On a global scale the largest VLM trend in TG records is~~ The models for large scale VLM processes, such as Glacial Isostatic Adjustment (GIA) ~~,~~ and the elastic response of the Earth due to present-day mass redistribution, are becoming more accurate. TGs are often only corrected for the GIA signal, which typically reaches values of 10 mm yr^{-1} in Canada and Scandinavia ~~(*Ostanciaux et al.*, 2012). On large scales~~ (*Gutenberg et al.*, 1941). The elastic deformation due to present-day mass redistribution ~~also affects the trends and~~ is often ignored. However, elastic deformation is becoming larger due to the ~~accelerating~~ increasing rate of Greenland's ice mass loss, and to a lesser extent other processes~~, it accelerates them~~. Trends at TGs are also affected by a large number of other local signals, including ~~erosion, ground water depletion~~water storage, postseismic deformation and ~~gas extraction~~ anthropogenic activities (*Hamlington et al.*, 2016; *Wöppelmann and Marcos*, 2016). Since the local VLM processes cannot be captured by models, and the large-scale processes contain large uncertainties, observations of VLM at TGs are essential.

One method to estimate VLM at TGs uses geodetic Global Positioning System (GPS) receivers at fixed stations or Doppler Orbitography and Radiopositioning Integrated by satellite (DORIS) observations. Since many other navigation satellites are currently providing range estimates as well, we will refer to the GPS stations as Global Navigation Satellite System (GNSS) stations. Most studies compute GNSS VLM at TG stations from one of the datasets by University of La Rochelle (ULR) (*Wöppelmann et al.*, 2007; *Pfeffer and Allemand*, 2016; *Wöppelmann et al.*, 2014; *Wöppelmann and Marcos*, 2016). Even though ULR contains several GNSS solution inland, its main focus is the coastal zone. Currently, 754 GNSS stations are processed in the ULR6 database. A more extensive database with approximately 14000 GNSS is processed by the Nevada Geodetic Laboratory (NGL). They use a different processing procedure to estimate trends from time series, which makes trends less vulnerable to jumps (*Blewitt et al.*, 2016). A statistical comparison between several GNSS solutions was recently made by *Santamaría-Gómez et al.* (2017). They concluded that the number of stations in the NGL database was larger, but that the ~~accuracy was on average significantly lower than the~~ differences between neighbouring stations was significantly larger than the Jet Propulsion Laboratory (JPL) and ULR6 ~~stations~~solutions. They also discussed systematic errors due to differences in the origin of the reference frames, which were in the order of 0.2 mm yr^{-1} globally. Furthermore, they found that the local VLM uncertainty at tide gauge was increased by 4×10^{-3} mm yr^{-1} per kilometer distance between the TG and the GNSS station (*Santamaría-Gómez et al.*, 2017). Most studies use the trends of either co-located GNSS stations or the closest GNSS station or the mean of all GNSS stations within a radius of several tens of kilometers (*Santamaría-Gómez et al.*, 2014; *Pfeffer and Allemand*, 2016). Only *Hamlington et al.* (2016) involved a more complex GNSS ~~weighting~~ post-processing procedure using NGL trends, based on a combination of spatial filtering, Delaunay triangulation and median weighting. One way to quantify the accuracy of GNSS-based VLM trends at TGs is to compute the spread of individual geocentric sea level estimates or the spread of geocentric sea level between regions (*Wöppelmann and Marcos*, 2016). The spread of regional trends reduced from 0.9 mm yr^{-1} in the ULR1 solution (*Wöppelmann et al.*, 2007) to 0.5 mm yr^{-1} in the ULR5 solution ~~(*Wöppelmann et al.*, 2014)~~(*Santamaría-Gómez et al.*, 2012; *Wöppelmann et al.*, 2014), which is approximately the expected residual climatic signal. Any further improvements in the GNSS trends require therefore another validation technique.

A second way to observe VLM at TGs, to overcome the limitations of sparsely distributed GNSS network, is differencing satellite altimetry and TG time series, which we will refer to as ALT-TG time series from here on. Initially, the ALT-TG time series were used to monitor the stability of satellite altimeters for the Global Mean Sea Level (GMSL) record, which is currently gauranteed up to 0.4 mm yr^{-1} (*Mitchum*, 1998, 2000). The first study to infer VLM trends from ALT-TG time series was *Cazenave et al.* (1999). Based on the method of *Mitchum* (1998) they compared ALT-TG to DORIS at six stations. Later, several studies were conducted on regional and global scale of which an overview is given by *Ostanciaux et al.* (2012). The first study to estimate more than 100 VLM trends (*Nerem and Mitchum*, 2002) obtained error bars for 60 of 114 TGs smaller than 2 mm yr^{-1}. However, they noted that the TGs should be inspected on a case-by-case basis to determine if the result was truly VLM. *Ostanciaux et al.* (2012) increased the number of ALT-TG VLM trend estimates sixfold to 641, but it included some outliers with trends above 20 mm yr^{-1}. They also made a comparison between their study and several earlier studies. The best agreement was found over a small set of 28 tide gauges, where the results of *Ostanciaux et al.* (2012) differed from (*Ray et al.*, 2010) by an RMS of 1.2 mm yr^{-1}.

Recently, several studies have compared the GNSS trends to those of ALT-TG globally (*Santamaría-Gómez et al.*, 2014; *Wöppelmann and Marcos*, 2016; *Pfeffer and Allemand*, 2016). Several other studies did an equivalent comparison with DORIS and ALT-TG for a limited number of stations (*Cazenave et al.*, 1999; *Nerem and Mitchum*, 2002; *Ray et al.*, 2010). While the older studies primarily used along-track data from the Jason (TOPEX/POSEIDON (TP), Jason-1 (J1) and Jason-2 (J2)) series of satellite altimeters, the latest studies used preprocessed grids and *Wöppelmann and Marcos* (2016) made a comparison between several gridded products and one along-track dataset. All recent studies used ULR5 GNSS trends for comparison. The best results were obtained with an interpolated altimetry grid provided by AVISO (*Pujol et al.*, 2016), yielding a median of differences of 0.25 mm yr^{-1} with an RMS of 1.47 mm yr^{-1} based on a comparison at 107 locations (*Wöppelmann and Marcos*, 2016). It is important to note that the time series for all sites were visually inspected, primarily to remove those with non-linear behaviour. Additonally, the corresponding correlation between altimetry and TG time series were found to be highest for AVISO. *Pfeffer and Allemand* (2016) did not apply visual inspection and obtained a comparable result for 113 stations (an RMS of 1.7 mm yr^{-1}), while only incorporating GNSS trends from stations within 10 km from the tide gauge.

This study aims to further reduce the discrepancies between GNSS and ALT-TG trends, while increasing the number of trend pairs. To do this, we will apply several steps to improve the VLM estimates at tide gauges. First of all, the number of reliable trend estimates are increased by using the GNSS trends from the larger NGL database. Most TGs will neighbour multiple GNSS stations for which several ~~weighting~~ methods are applied to determine the best procedure. Correlations between altimetry and TG time series are exploited to reduce residual ocean variability, which is often present in ALT-TG time series (*Vinogradov and Ponte*, 2011). The reduction in ocean variability should lead to more reliable ALT-TG VLM trends. Correlation thresholds additionally function as a filter, to remove time series that are uncorrelated due to differences in ocean signals~~and~~, possible (undocumented) jumps in the TG time series, or interannual VLM signals that cannot be separated from the ocean signal (*Santamaría-Gómez et al.*, 2014). Additionally, we address the problem of contemporary mass redistribution on trends over different time spans using a fingerprinting method.

2 Data and Methods

In this section, we describe the processing procedures for deriving ~~direct and differenced~~ GNSS and ALT-TG VLM trends for comparison at TG locations. First, we will address the estimation of GNSS trends at the TG locations. The estimation of ALT-TG differenced trends is discussed in several steps. We briefly discuss the selection of the tide gauges. After that we will discuss the altimetry processing procedures. We briefly review the Hector software (*Bos et al.*, 2013) for the estimation of trends from differenced ALT-TG time series. Eventually, trend corrections for contemporary mass redistribution using fingerprinting methods are described.

2.1 GNSS trends

The trend estimation at tide gauges primarily deals with two problems. First, a trend ~~has to be~~ is estimated from a GNSS time series, which contains an autocorrelated noise signal, and often undocumented jumps. ~~This is discussed~~ We use pre-computed trends, of which the procedure is briefly reviewed in Sect. 2.1.1. Second, since many GNSS stations are not directly co-located to the ~~tide-gauge station, the~~ TG station. Regular leveling campaigns, to monitor the relative VLM between the TG and the GNSS stations, after often absent. Therefore, the assumption is made that both locations are affected by the same VLM signal. When multiple GNSS receivers are present in the vicinity of the tide gauge, a ~~weighting~~ method is required to estimate a single VLM trend from multiple GNSS stations. This is discussed in Sect. 2.1.2.

2.1.1 ~~Trend~~ GNSS trend estimation

To obtain VLM trends at TGs, often the products of the Université de La Rochelle (ULR) are used. ULR ~~version~~ versions 5 and 6 ~~makes~~ make use of the Create and Analyze Time Series (CATS) software (*Williams*, 2008), which is able to estimate trends and errors from time series, taking into account temporally correlated noise. ~~Next to a slight change in trend, it~~ It has the advantage that it computes a more realistic trend uncertainty. ~~However, the frequently occuring jumps,~~ The software is also able to estimate and detect discontinuities that occur due to earthquakes ~~or equipment changes, decrease the accuracy, which is therefore often larger~~ and equipment changes. Even though a large fraction of the trend estimates have formal accuracies better than 1 mm~~yr^{-1}~~/yr, undetected discontinuities might bias the estimated trends (*Gazeaux et al.*, 2013).

In this study the results of NGL (*Blewitt et al.*, 2016) are used. *Blewitt et al.* (2016) proposed the Median Interannual Difference Adjusted for Skewness (MIDAS) approach, which is based on the Theil-Sen estimator. The procedure estimates trends from couples of daily data points separated by 365 days. It then removes all ~~measurements~~ estimates outside two standard deviations, which are computed by scaling the ~~median of absolute devations~~ Median of Absolute Devations (MAD) by 1.4826 (*Wilcox*, 2005), with respect to the median of the trend couples. Afterwards, a new median is computed, which serves as the trend estimate. *Blewitt et al.* (2016) demonstrated that MIDAS has a smaller equivalent step detection size than methods which included step detection, as those ~~used by ULR~~compute with CATS and used by ULR5. Besides the advantage of detecting smaller jumps, approximately 14000 GNSS time series are processed, which is almost 20 times more than ULR6. In this study no screening, like performed by (*Wöppelmann and Marcos*, 2016), is applied to the time series or trends.

2.1.2 ~~Station weighting and selection~~Trend estimation at tide gauges

Despite several recommendations to co-locate GNSS receivers with TGs, currently only a few have a record ~~long enough to estimate reliable trends~~that ensures a trend uncertainty of 1 mm yr^{-1} or better. Therefore we take all stations into account that are within 50 km from a TG, provided that the standard deviation on the trend is lower than 1 mm yr^{-1} as estimated from the MIDAS algorithm. The threshold on the standard deviation ensures that most records containing large non-linear effects, due to for example earthquakes and water storage changes, are removed from the analysis. Other studies used ranges from 10 km (*Pfeffer and Allemand*, 2016) up to 100 km (*Hamlington et al.*, 2016). At 100 km the error due relative VLM trends increases substantially, on average with more than 0.5 mm yr^{-1} (*Santamaría-Gómez et al.*, 2017) for the NGL estimates, while taking a range of 10 km reduces the number of trends substantially. Therefore the range is set to 50 km, but comparable results are found for 30 and 70 km yielding a different number of trends (not shown).

Most studies simply average all neighbouring TG trends or take the trend from the closest station. However, many other and possibly better, techniques are possible. We compare trends from several approaches in Sect. 3.1 and with the ALT-TG trends in Sect. 3.3.

In total eight different approaches are considered. The first two involve all of the trends neighbouring GNSS stations by computing their mean [1] and median [2]. Method [1] is among others applied by (*Frederikse et al.*, 2016) for regional sea level reconstructions. One of the most frequently applied approach uses the trend at the closest station [3]. It is used in two recent studies by *Santamaría-Gómez et al.* (2012) and *Pfeffer and Allemand* (2016). We also investigate inverse distance weighting [4] in which the trend $\frac{dh_{TG}}{dt}$ is estimated as:

$$\frac{dh_{TG}}{dt} = \frac{\sum \frac{1}{d_i}\frac{dh_i}{dt}}{\sum \frac{1}{d_i}}, \tag{1}$$

where d_i and $\frac{dh_i}{dt}$ respresent the distance to the tide gauge station and the trend at GNSS station i. We also use the GNSS trends based on the longest time series [5] and smallest error [6] from stations within the 50 km radius. The seventh approach involves weighting with the variances σ_i^2 of the trends [7], such that:

$$\frac{dh_{TG}}{dt} = \frac{\sum \frac{1}{\sigma_i^2}\frac{dh_i}{dt}}{\sum \frac{1}{\sigma_i^2}}. \tag{2}$$

And the last method [8] takes into account spatial dependency and trend uncertainty by combining methods [4] and [7], i.e. by weighting with the variance and with the distance~~.~~, so that:

$$\frac{dh_{TG}}{dt} = \frac{\sum \frac{1}{\sigma_i^2 d_i}\frac{dh_i}{dt}}{\sum \frac{1}{\sigma_i^2 d_i}} \tag{3}$$

Method [8] is a variant to the technique used in the altimeter calibration study of *Watson et al.* (2015). Note that ~~in~~ the uncertainties range mostly between 0.7-1 mm yr^{-1} and therefore method [8] is more sensitive to the distance from the TG than to the variance of the GNSS trends. The distance weights used in methods [4] and [8] quickly reduce, effectively lowering the number of GNSS trends involved. In several studies the ~~weighting method for GNSS trends~~ method to estimate VLM trends at tide gauges from GNSS is not documented.

2.2 Tide gauge time series

Monthly TG data are obtained from the PSMSL database (*Holgate et al.*, 2013). All time series flagged after 1993 are removed. Any observations that are outside of 1 meter from the mean are considered outliers and removed from the data. This number is similar to our altimetry sea level treshold and based on the criterion used by NOAA for their global mean sea level estimates (*Masters et al.*, 2012). To be consistent with the altimetry observations, we apply a Dynamic Atmosphere Correction (DAC) consisting of a low-frequency inverse barometer correction and short-term wind and pressure effects *Carrère and Lyard* (2003). Initially, we consider all TGs with at least 10 years of valid data.

2.3 Differenced ALT-TG time series

Wöppelmann and Marcos (2016) obtained the smallest standard deviation in the differenced time series by averaging grid cells within 1 degree from the TG using the AVISO interpolated product. The results obtained by taking the most correlated grid point from AVISO within 4 degrees around the TG increased the standard deviation. *Wöppelmann and Marcos* (2016) obtained lower correlations by averaging Goddard Space Flight Center (GSFC) along-track altimetry measurements within a radius of 1 degree from the TG. Note that the AVISO grid is constructed using correlation radii of 50-300 km (*Ducet et al.*, 2000) and it includes measurements from all altimetry satellites, not only the Jason series. The AVISO grid therefore effectively averages over a much larger radius around the TG and it includes observation from more satellites. The larger uncorrelated noise using GSFC compared to AVISO, as shown by the combination of the increased RMS and the spectral index (*Wöppelmann and Marcos*, 2016), is therefore likely an effect of the limited number of GSFC altimetry measurements. However, using the large effective radius of AVISO~~might increase the correlated noise, due to the inclusion sea level observations~~, data far away from the TG ~~, which do~~ is included, which might not correlate with the sea level ~~observed at the tide gauge~~signal at the TG. This can result in a remaining ocean signal in ALT-TG time series, which contaminates the VLM trend estimates.

~~Values for the parameters of the latitudinal intermission bias correction. These numbers are added to the sea surface height anomalies of the respective satellites. TP asc. TP desc. Jason-2 Parameter Lat(deg) Value Lat(deg) Value Lat(deg) Value c_0(mm) (-66.2,-1.5) 80.3 (-66.2,-1.5) 77.3 (-66.2,66.2) 98.1 c_1(mm deg^{-1}) -2.3·10^{-1} -1.7·10^{-1} -9.3·10^{-2} c_2(mm deg^{-2}) -1.1·10^{-2} 1.2·10^{-3} 3.8·10^{-3} c_3(mm deg^{-3}) -3.0·10^{-4} 2.9·10^{-4} 8.4·10^{-7} c_4(mm deg^{-4}) -2.4·10^{-6} 3.8·10^{-6} 7.6·10^{-7} c_0(mm) (-1.5,0.2) 83.8 (-1.5,1.3) 79.9 c_1(mm deg^{-1}) 1.3 2.4 c_2(mm deg^{-2}) -1.3 5.2·10^{-1} c_3(mm deg^{-3}) -5.3·10^{-1} c_4(mm deg^{-4}) c_0(mm) (0.2,4) 84.9 (1.3,4) 73.3 c_1(mm deg^{-1}) -8.0·10^{-1} 13.7 c_2(mm deg^{-2}) -8.6·10^{-1} -5.1 c_3(mm deg^{-3}) 1.5·10^{-1} 4.9·10^{-1} c_4(mm deg^{-4}) c_0(mm) (4,66.2) 72.9 (4,66.2) 75.8 c_1(mm deg^{-1}) 8.1·10^{-1} 7.9·10^{-1} c_2(mm deg^{-2}) -2.8·10^{-2} -3.3·10^{-2} c_3(mm deg^{-3}) 3.4·10^{-4} 6.4·10^{-4} c_4(mm deg^{-4}) -1.1·10^{-6} 3.9·10^{-6}~~

Table 1. List of geophysical corrections and orbits applied in this study.

Satellite	T/P	Jason-1&2
Orbits	CCI	GDR-E
Ionosphere	Smoothed dual-frequency	
Wet troposphere	Radiometer	
Dry troposphere	ECMWF	
Ocean tide	GOT4.10	
Loading tide	GOT4.10	
Solid Earth tide	Cartwright	
Sea state bias	CLS	
Mean sea surface	DTU15	
Dynamic atmosphere	MOG2D	

To overcome the limitations of gridded products, we work with along-track data and exploit the correlations between sea level at the satellite measurement location and at the TG on interannual and decadal scales by using a low-pass filter. We start by creating sea level time series every 6.2 km along-track using the measurements from TP, J1 and J2 from the RADS database (*Scharroo et al.*, 2012) between 1993-2015. In order to get a consistent set of altimetry observations, the same geophysical correction are used for all satellites, which are given in Table 1. All time series within 250 km from the TG are taken into account. This radius is larger than the open ocean correlation distances used by *Ducet et al.* (2000) and *Roemmich and Gilson* (2009), except for the equatorial region where the correlation scales become much larger. At distances larger than 250 km, one will still find some highly correlated signals, but the trends caused by large scale processes like GIA and present-day mass redistribution will differ from that at the TGs. It also ensures that at least one ground track of the altimeters is within the range of the tide gauge at the equator. Reducing the 250 km radius leads to a decreased number of trends.

Additionally, intermission biases between TP-J1 and J1-J2 are removed. *Ablain et al.* (2015) revealed a large dependence of the intermission biases on the latitude. For the J1-J2 differences, a single polynomial is estimated through the differences between the sea level observations of both instrument, such that the correction $\Delta h_{sla,ib}(\lambda)$ becomes:

$$\Delta h_{sla,ib}(\lambda) = c_0 + c_1 \cdot \lambda + c_2 \cdot \lambda^2 + c_3 \cdot \lambda^3 + c_4 \cdot \lambda^4, \tag{4}$$

with λ the latitude of the altimetry observations. For the TP-J1 differences, separate polynomials are estimated for four latitude regions and the ascending/descending tracks (*Ablain et al.*, 2015). The values for the parameters c_n are given in Table 6. More details on the computation procedure are found in Appendix A.

The Jason satellite series samples sea level every ten days, hence we average monthly 3-4 measurements in order to make a first set of time series that is compatible with the monthly TG observations. As for the case of the TG monthly solutions, observations more than 1 m from the mean sea surface are removed and the time series should have at least 10 years of valid observations. Additionally, a second set of ~~monthly~~ time series at each satellite measurement location is created, by applying a yearly moving-average filter. This second set of altimetry time series is correlated with a yearly low-pass filtered version of the TG series, in order to test whether their signals match on interannual and longer time scales. The yearly moving-average filter allows to suppress the noise present in individual altimetry measurements. The full pole tide from RADS (which contains a solid Earth, loading and ocean tide as in *Desai et al.* (2015)) is subtracted from both time series before correlation, whereas for the TG time series we restore the solid Earth pole tide as computed in *Desai et al.* (2015). The loading tide is at its maximum only a few millimeters, which has no significant effect on the interannual correlation, and is therefore not restored. We also remove residual ~~seasonal~~ annual and semi-annual cycles and a linear trend before correlation~~.~~, because the yearly moving-average filter has side-lobes causing these seasonal signals to be partly retained. Other longer filters are considered to reduce the side lobes, but they would introduce larger transient zones. An iterative procedure removes sea surface heights outside of 3 RMS up to a maximum of 10 % of the observations. The outlier removal is primarily implemented to remove any spurious data present in the RADS database. It is unlikely that more than 10 % of the observations contain processing problems or outliers due to extreme events. If more observations would be discarded, high correlations might not represent corresponding ocean signal anymore. The result is a set of correlations that indicate which altimetry sea level time series resemble the TG time series on interannual time scales and longer.

The monthly low-pass filtered altimetry time series are kept, if the corresponding correlation from yearly low-pass filtered time series are above a certain threshold. We combine the remaining monthly altimetry time series, to get one averaged altimetry time series per TG. Alternatively, we also use the correlations as weights, to get one correlation-weighted altimetry time series per tide gauge. In this case the monthly low-pass filtered time series are weighted by their corresponding correlation, then added together and accordingly normalized, so that the weights sum up to one. The resulting time series are subtracted from the TG time series if there are at least ten altimetry time series with a correlation above the threshold. The resulting differenced ALT-TG time series with less than 15 years of valid observations are further discarded. This last requirement is due to the fact that remaining ocean signals can still affect the estimated trends significantly. An example of the reduction of variability due to correlation thresholds and weighting is shown in Fig. 1. The white noise in the unfiltered time series is reduced in the red curve, however the opposite might happen if the number of altimetry time series decreases. Most important is to note that there is a strong reduction in the variance of temporally correlated residuals, represented here by the low-pass filtered time series. Correlated residual signal can strongly affect the estimated trend, especially in areas with large ~~interannual~~ variability due to interannual event like ENSO. Note that for the differentiation of the time series only the solid Earth part of the pole tide is added to the TGs, but this time as is done in the IERS2010 conventions (*Petit and Luzum*, 2010), such that the trends are consistent with those of the GNSS data. The main difference is that the altimetry pole tide correction of *Desai et al.* (2015) ~~are~~ is computed with respect to linearly drifting mean pole, while in the IERS conventions the mean pole location is modelled as a third order polynomial. If the pole tide is not taken into account consistently, it can introduce biases of 0.1 mm yr~~−1~~ $^{-1}$ (*Santamaría-Gómez et al.*, 2017). Since the change rate of the mean pole is non-linear, this will introduce trend biases if the time spans between GNSS and altimetry do not match. The drift of the mean pole is caused by redistribution of mass in the Earth system. This is corrected for using the ~~mass redistribution~~ mass-redistribution fingerprints discussed in Sect. 2.5~~.~~, which are computed using a model that includes elastic responses and rotation changes. The drifting mean pole is primarily captured by the C_{21} and S_{21} spherical harmonic coefficients (*Wahr et al.*, 2015).

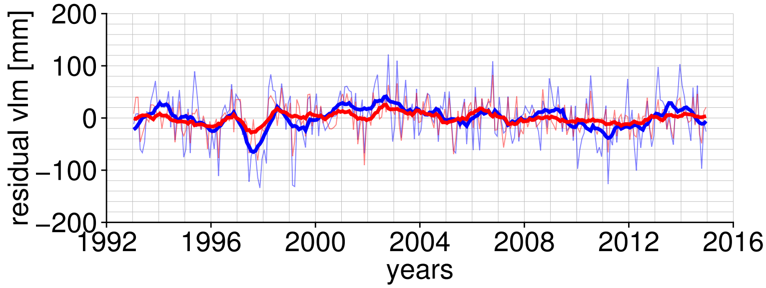

Figure 1. ~~Low-pass filtered time~~ Time series of ALT-TG differenced VLM at Winter Harbour. After averaging or weighting with the correlation a moving-average filter is applied to visualize the remaining interannual variability. In blue: without a threshold on the correlation and without correlation weighting. In red: with a threshold of 0.7 for the correlation and with correlation weighting. In the background the time series without ~~a low-pass~~ the moving-average filter applied.

2.4 Differenced ALT-TG trends

The ALT-TG time series have a monthly resolution, so they contain less observations, and they exhibit substantial interannual variability. These time series are therefore less suitable to be processed with the MIDAS algorithm used to compute GNSS trends. For the computation of the ALT-TG trends and the corresponding standard deviation, we fit a power-law in combination with a white noise model by using the Hector software (*Bos et al.*, 2013). The spectrum of the white noise is flat, while the spectrum of power-law noise, $P(f)$, decays with frequency and is given by (*Bos et al.*, 2013):

$$P(f) = \frac{1}{f_s^2} \frac{\sigma^2}{(2\sin(\pi f/f_s))^{2d}}, \tag{5}$$

where f_s is the sampling frequency, σ the power-law noise scaling factor and d links to the spectral index κ in *Wöppelmann and Marcos* (2016) by $\kappa = -2d$. The value of d affects the effective number of autoregressive parameters (*Bos et al.*, 2013). This is required to capture the temporal correlation in the ALT-TG time series as shown by Fig. 2 in which the low-pass filtered time series give an idea of the memory in the system. In order to handle several weakly non-stationary ALT-TG time series we use the function 'PowerlawApprox', which uses a Toeplitz approximation for power-law noise (*Bos et al.*, 2013).

2.5 Contemporary mass redistribution

The trends estimated from GNSS time series are computed over different time spans than the ALT-TG trends and will be affected by non-linear VLM induced by elastic deformation due to present-day ice melt and changes in land hydrology storage (*Riva et al.*, 2017). To quantify those non-linear VLM signals, the response to mass redistribution is computed using a fingerprinting method at yearly resolution. We take into account the loads of Greenland, Antarctica and glacier mass loss, the

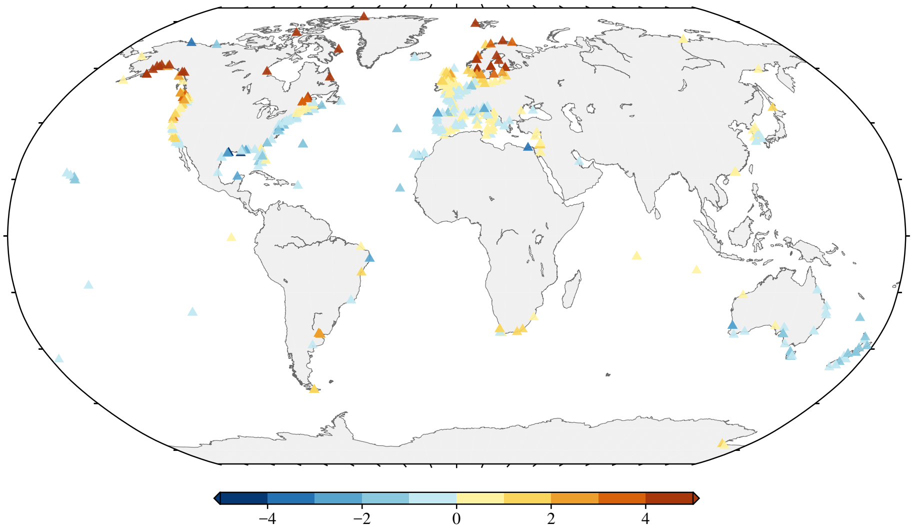

Figure 2. VLM (mm/yr) at TGs using the median of the neighbouring trends.

effects of dam retention and hydrological loads. A detailed description of the input loads is given in (*Frederikse et al.*, 2016). To estimate the fingerprints of VLM, the sea level equation is solved, including the rotational feedback (*Farrell and Clark*, 1976; *Milne and Mitrovica*, 1998). Since not all load information for 2015 and 2016 is available yet, we will limit the time series of ALT-TG up to 2015. ~~The~~ Some GNSS trends are ~~pre-computed and therefore~~ estimated from time series that span beyond 2015. Therefore we linearly extrapolate the fingerprint data, ~~is~~ if necessary, to 2015 and 2016 ~~by taking~~ based on the difference between years 2013 and 2014.

3 Results

This section first addresses the trends obtained from GNSS stations. The averaging methods are discussed and the NGL trends are compared to those of ~~ULR~~ULR5. Then the results of the correlation-weighted ALT-TG trends are discussed. These are compared to those from *Wöppelmann and Marcos* (2016). After that, the GNSS and ALT-TG trends are compared and optimal settings are discussed. For the comparison we take into account that both trends are not computed from time series covering the same period by correcting for non-linear VLM trends estimated from fingerprints.

3.1 Direct GNSS trends

For 570 TGs at least one GNSS station is found within a 50 km radius with an uncertainty on the trend that is below 1 mm yr^{-1}. The VLM for these TGs is shown in Fig. 2 using the median of the surrounding GNSS stations in case there are multiple trends available. The signature of GIA dominates the signal on large scales, which is primarily visible in Scandinavia and Canada.

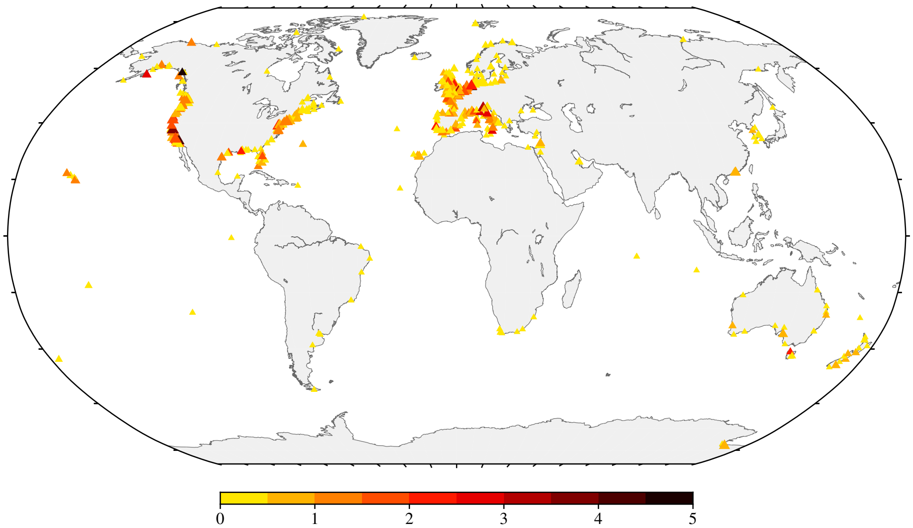

Figure 3. ~~Spread~~ Range (mm/yr) ~~in~~ of VLM estimates at TGs using eight different approaches. The size of the symbols indicates the number of GNSS trends available (with a maximum of ten).

Table 2. Statistics of trend differences between NGL and ~~ULR~~ ULR5 at 70 stations for ~~various weighting techniques~~the eight approaches.

		RMS	Mean	Median
Approach	~~mm yr−1~~ Keyword	mm yr^{-1}	mm yr~~−1~~$^{-1}$	mm yr~~−1~~$^{-1}$
1	mean	1.11	0.07	0.05
2	median	1.05	0.12	0.03
3	closest	1.36	0.02	0.02
4	dist. weight.	1.21	0.00	0.03
5	longest	1.29	0.32	0.20
6	smallest error	1.15	0.24	0.17
7	error weight.	1.11	0.08	0.02
8	dist./error weight.	1.23	0.01	0.05

In Alaska there might be a significant contribution of present-day ice mass loss. If GIA is removed the VLM signals typically range between -3 and 3 mm yr^{-1} (*Wöppelmann and Marcos*, 2016), with a few exceptions.

Even though the large-scale GIA process appears to be captured properly, regional VLM has a large effect on the GNSS trends. In Fig. 3 the differences between the lowest and highest VLM estimate from the eight methods discussed in Sect. 2.1.2 are shown. The extreme values primarily resulted from the 'mean', 'median' and 'inverse distance' ~~weighting~~ methods (not shown). The figure shows that the ~~spread~~ range is generally higher, where more GNSS trends are available. In particular the seismically active zones like the US West Coast show a larger ~~spread. The spread between~~ range. The range of solutions, when considering all TGs with at least two GNSS trends, has a mean of 0.92 mm yr^{-1} with 25 and 75 percentiles of 0.38

Table 3. Number of TGs at which trends are estimated from differenced ALT-TG time series. The '-1.0' indicates no correlation threshold is set.

Threshold	Number of TGs
-1.0	663
0.0	660
0.1	658
0.2	655
0.3	638
0.4	602
0.5	549
0.6	470
0.7	344

and ~~a median of 0.71~~ 1.20 mm yr^{-1}. In case at least three available GNSS trends are considered, the ~~differences rise~~ mean of the differences rises to 1.09 mm yr^{-1} and ~~0.85~~ the 25 and 75 percentiles to 0.56 and 1.34 mm yr^{-1}~~, respectively~~. Since we only considered GNSS trends with a maximum standard deviation of 1 mm yr^{-1}, this implies that a significant contribution of kilometer-scale VLM variations is present along the West Coast of the US, where the difference between methods is often larger than 1 mm yr^{-1}. Note that the ~~spread between~~ range of individual GNSS trends is on average even larger than the ~~spread~~ range between methods. *Santamaría-Gómez et al.* (2017) estimated the global numbers for the impact of spatial variations in VLM at 30 km and 100 km separation to be 0.2 mm yr^{-}1 and 0.5 mm yr^{-1}. ~~However, at several coastal locations, especially in~~ At coasts of Europe and North America~~this appears to be an underestimation.~~, where most tide gauges are located, these numbers are substantially larger, i.e. even the range between methods is on average larger than 1 mm yr^{-1}. The differences between methods is often comparable in size as the VLM signal, especially after the GIA is removed.

Wöppelmann and Marcos (2016) show that a comparison between their ALT-TG trends and their GNSS trends yields an RMS of 1.47 mm yr^{-1}. They use visual inspection to remove tide gauges where clear non-linear effects or discontinuities were present. In Table 2 a comparison is made between the eight different ~~weighting methods~~ approaches and the GNSS trends of *Wöppelmann and Marcos* (2016) that were used in the aforementioned comparison with ALT-TG trends at 70 locations. The values show that a substantial fraction of the RMS between ~~direct and differenced~~ GNSS and ALT-TG trends can already be explained by different GNSS averaging and processing methods. Using the closest station (approach 3) an RMS of 1.36 mm yr^{-1}, which is comparable in magnitude to the RMS between GNSS and ALT-TG trends found by *Wöppelmann and Marcos* (2016). ~~The~~ Note that we remove all NGL GNSS trends with an uncertainty larger than 1 mm yr^{-1} and therefore co-located stations are sometimes removed. The closest GNSS station in our selection is therefore not always the same as the one used by *Wöppelmann and Marcos* (2016). The best comparison is found with the median (approach 2), even though the RMS of differences is still above 1 mm yr^{-1}. Since the closest station method depends on a single station, there is larger chance some

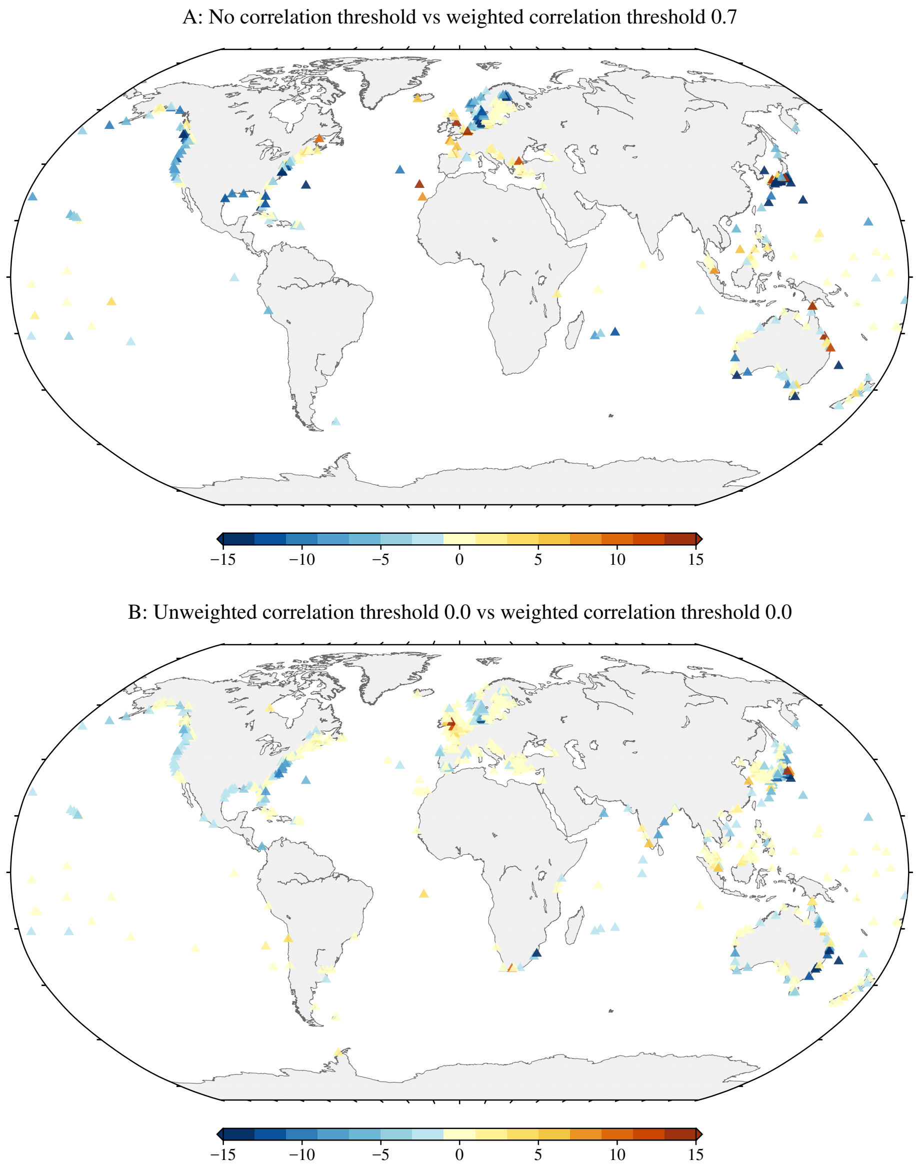

Figure 4. ~~Reduction~~ Change in standard deviation (mm) of the differenced time series using correlation thresholds and weighting. Note that a correlation threshold of 0.0 indicates positive correlations only.

outliers are present, which substantially increases the RMS of differences. For the closest station method three trend differences larger than 3 mm yr^{-1} are found, whereas only one is found for the median method.

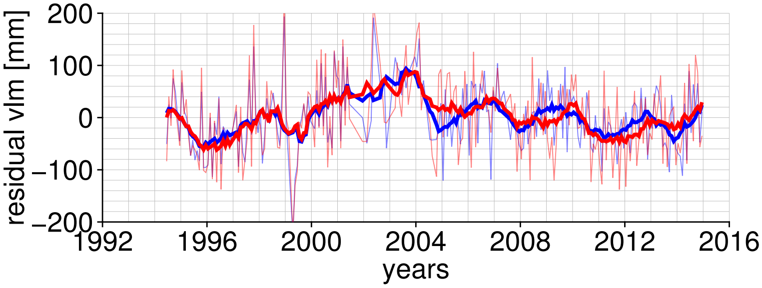

Figure 5. ~~Low-pass filtered time~~ Time series of ALT-TG differenced VLM at the Llandudno (UK) TG. A moving-average filter is applied to visualize the interannual variability. In blue: with a threshold of 0.0 for the correlation, but without correlation weighting. In red: with a threshold of 0.0 for the correlation and with correlation weighting. In the background the time series without a ~~low-pass~~ moving-average filter applied.

3.2 Differenced ALT-TG trends

Using correlation thresholds, we try to minimize the residual ocean signal in ALT-TG time series Additionally, it will filter problematic stations, where no correlation between TG and altimetry observations is found. A higher threshold reduces therefore the number of ALT-TG trends. Table 3 shows the reduction of the differenced VLM trends, when the correlation threshold increases. After a correlation threshold of 0.4, the number of observations drops substantially. At a threshold of 0.7, the number of TGs for which a trend is computed, is only half of that without a threshold. The remaining trends are generally more reliable, because of two reasons: VLM time series that exhibit relatively large residual ocean signals are removed; and secondly, TG time series that contain large jumps due to unidentified reasons (e.g. earthquakes or equipment changes) are removed.

In order to show that the method decreases the oceanic signal, we compare the standard deviation reduction by using correlation thresholds and weighting (Fig. 4). The plot in the top panel shows the comparison between the standard deviation of the differenced time series using no correlation threshold and the time series using a threshold of 0.7 together with a correlation weighting. The mean reduction in standard deviation is 3.9 mm, whereas the mean standard deviation is 37 mm. The change in standard ~~devations~~ deviations at several locations are coherent, which is expected because the sea level fluctuations along continental slopes are coherent (*Hughes and Meridith*, 2006). Substantial reductions in standard deviation are apparent at both North American coasts, in Japan and in Northern Europe. *Vinogradov and Ponte* (2011) had already observed large discrepancies in interannual ocean signals between TGs and altimetry in North America and in Japan. It suggests that our technique is capable to reduce these ocean signals. ~~The~~ This is confirmed by the change in the median of the spectral indices, κ, as discussed in Sect. 2.4. The median of the spectral indices changes from -0.63 to -0.57, which indicates that the autocorrelation in the residuals decreased. The Winter Harbour (Canada) VLM time series (Fig. 1) shows a typical example in which especially the correlated noise is reduced. However, there are several locations where the standard deviation increases substantially. Most of them are sporadic, but in a few locations, like in the UK and France there is coherent increase. Though reduced in magnitude, similar, patterns are observed for the not-weighted against weighted VLM time series with a correlation threshold of 0.0 (bot-

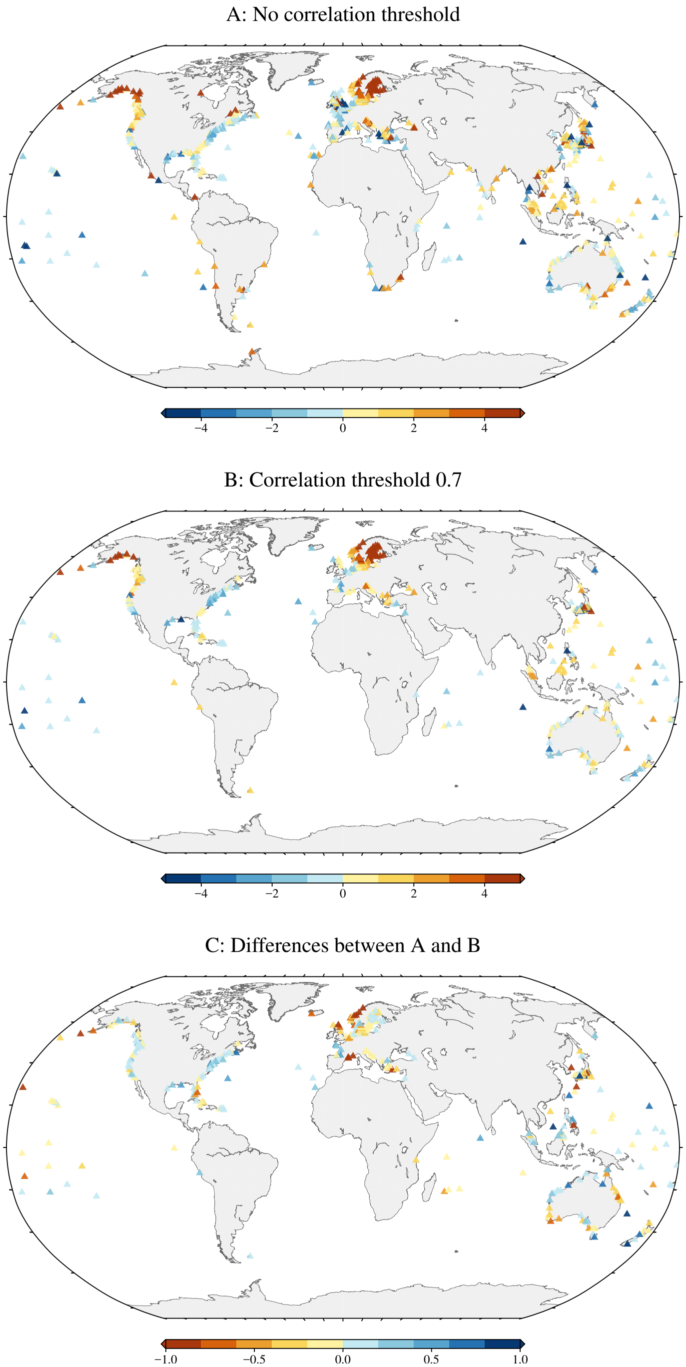

Figure 6. ALT-TG trends (mm yr^{-1}) estimated using no threshold (A), with a correlation threshold and correlation weighting (B) and the difference between them (C).

tom of Fig. 4), i.e. when only positively correlated altimetry time series are taken into account. Instead of 344 VLM trends, as for the comparison discussed above, 660 trends are compared. The mean reduction of the standard deviation is 1.4 mm, whereas the mean standard deviation is 38 mm. Remarkable is the strong reduction of the standard deviation at the southeast side of Australia. In the UK and France an increase in standard deviation is present again. In most cases an increase in white noise, likely due to the decreased effective number of altimetry measurements, is responsible for the higher standard deviation, as demonstrated in Fig. 5 for a VLM time series at Llandudno, UK. In most cases of an increasing standard deviation, the correlated ocean signals are still reduced or remain approximately equal.

Fig. 6 shows the VLM trends estimated from the ALT-TG time series using no correlation threshold and a threshold of 0.7. A comparison of Fig. 2 and Fig. 6 reveals that especially the Indian Ocean and the southern Pacific Ocean are sampled better using ALT-TG instead of GNSS trends. If the correlation threshold is set to 0.7, the number of trends decreases, which has particularly an impact on the number of trends in South America and Africa. Hence, for regional reconstructions, a careful choice should be made for the correlation threshold.

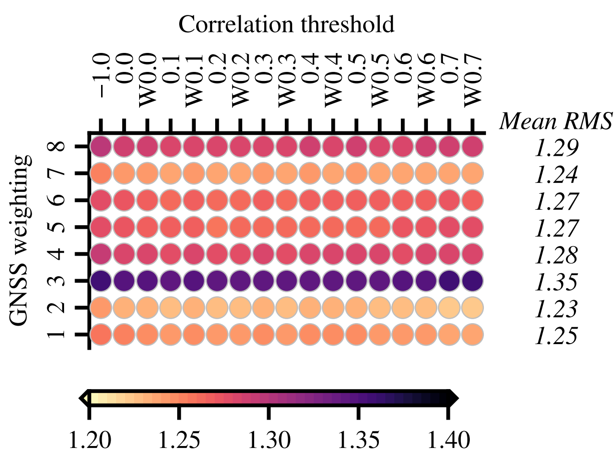

Figure 7. RMS (mm/yr) of differences between ~~direct~~ GNSS and ~~indirect inferred~~ ALT-TG VLM trends. The 'W' indicates weighting by correlation. The '-1.0' indicates no correlation threshold is set. The numbers of the y-axis refer to the ~~weighting methods for~~ approaches used to combine the GNSS trends as described in Sect. 2.1.2.

Compared to the GNSS trends, the neighbouring ALTG-TG trends show more variation, which is especially true for the UK and Japan. It is difficult to say whether this is a true VLM signal, but it is important to note that many GNSS stations are placed on bedrock, which exhibits more stable trends than the coastal locations of tide gauges. Secondly, the GNSS trends with an ~~error~~ uncertainty larger than 1 mm yr^{-1} are removed, which reduces the variability. ~~However, the larger~~ Of the 663 ALT-TG trends, 293 (44 %) have a trend uncertainty smaller than 1 mm yr^{-1}. Therefore larger spatial trend variability can also be induced by remaining ocean ~~signal~~ signals in the VLM time series. In the Fig. 6B, showing the 0.7 treshold trends, the number of trends is reduced due to the correlation threshold. It removes most tide gauges in the highly variable regions mentioned before and the neighbouring differences are therefore less erratic. 284 out of 344 trends (83 %) have a trend uncertainty smaller than 1 mm yr^{-1} using the 0.7 correlation threshold.

Table 4. Statistics of the differences between the median of the GNSS trends ~~using median weighting~~ (approach 3) and the ALT-TG trends for various correlations thresholds. The 'W' indicates that the altimetry time series are weighted by the correlation. The row 'W&M' shows the comparison with *Wöppelmann and Marcos* (2016) trends. The column 'NoT' indicates the number of trends. On the left side of the table all stations are taken into account, on the right side only stations are taken into account for which a solution exist for all correlations thresholds (and including those from W&M).

	All				Same			
Correlation	RMS	Mean	Median	NoT	RMS	Mean	Median	NoT
	mm yr^{-1}	mm yr^{-1}	mm yr^{-1}		mm yr^{-1}	mm yr^{-1}	mm yr^{-1}	
-1.0	2.141	-0.241	-0.107	294	1.234	-0.167	-0.099	137
0.0	2.108	-0.248	-0.101	294	1.226	-0.175	-0.068	137
0.0W	2.103	-0.250	-0.036	294	1.219	-0.172	-0.056	137
0.1	2.113	-0.258	-0.096	293	1.219	-0.174	-0.074	137
0.1W	2.108	-0.260	-0.043	292	1.218	-0.170	-0.045	137
0.2	2.082	-0.233	-0.073	292	1.217	-0.163	-0.074	137
0.2W	2.080	-0.234	-0.015	292	1.216	-0.168	-0.042	137
0.3	1.986	-0.152	0.047	283	1.221	-0.157	-0.066	137
0.3W	1.991	-0.157	0.056	283	1.217	-0.165	-0.044	137
0.4	1.695	-0.106	0.065	264	1.223	-0.152	-0.050	137
0.4W	1.696	-0.112	0.071	264	1.218	-0.158	-0.041	137
0.5	1.554	-0.086	0.044	239	1.220	-0.153	-0.058	137
0.5W	1.552	-0.087	0.056	239	1.217	-0.155	-0.067	137
0.6	1.417	-0.093	-0.065	204	1.209	-0.155	-0.087	137
0.6W	1.416	-0.093	-0.083	204	1.208	-0.156	-0.094	137
0.7	1.220	-0.142	-0.123	155	1.206	-0.140	-0.060	137
0.7W	1.220	-0.144	-0.124	155	1.206	-0.142	-0.074	137
W&M	1.658	-0.177	-0.050	211	1.328	-0.101	0.020	137

The results of applying correlation weighting and thresholding are shown Fig. 6C. Two spots of coherent changes in the trends can be clearly identified: in Norway the trends increased by approximately 1 mm yr^{-1}, while in the East Coast of the United states the opposite happens. These spots exhibit longshore coherent sea level signals that are not found in the open ocean (*Calafat et al.*, 2013; *Andres et al.*, 2013). Note that both locations also exhibit a strong reduction in standard devation (Fig. 4). Coherent changes are also present around Denmark. Other regions, where substantial reductions in the standard deviation are found, do not experience coherent changes in trends.

3.3 GNSS vs ALT-TG trends

In this section the VLM trends from GNSS using the eight ~~weighting solutions~~ approaches as described in Sect. 2.1.2 are compared with the differenced ALT-TG VLM trends using various correlation thresholds. Based on the intercomparison we determine the best solution for the GNSS ~~weighting~~ approach and the correlation thresholds for altimetry. Additionally, a comparison is made with *Wöppelmann and Marcos* (2016). We also investigate the effect of present-day mass redistribution on the difference in trends due to varying time spans of the ~~direct and the differenced~~ GNSS and the ALT-TG method.

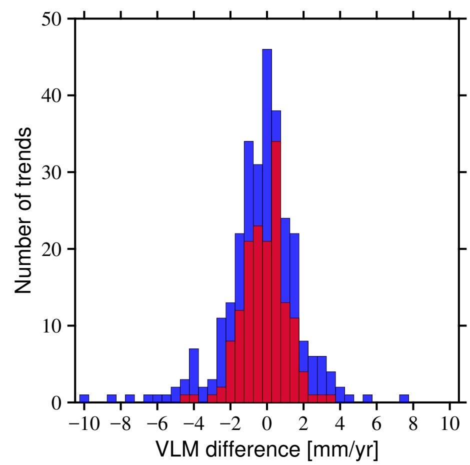

Figure 8. Histogram of GNSS and ALT-TG trend differences. In blue the results without any correlation threshold and in red with a correlation threshold of 0.7 and correlation weighting.

Fig. 7 shows the RMS of trends differences between various GNSS ~~weighting~~ combination methods and correlation thresholds for ALT-TG. The RMS of trend differences is computed at 155 TG stations for which all solutions are available. The colors exhibit small differences horizontally and large differences vertically, indicating that the GNSS ~~weighting~~ method is more important in reducing the RMS. The difference between the method with the lowest RMS of differences, which is obtained by taking the median of the GNSS trends (2), and the method with the highest RMS, which uses the closest GNSS station (3), is approximately ~~0.14~~ 0.12 mm yr^{-1}. *Hamlington et al.* (2016) computed VLM trends at TG locations by using a complex filtering procedure that also implicitly takes into account the median of the GNSS trends. Next to taking the median of the GNSS trends, taking the mean (1) within the 50 km radius and using variance weighting (7) also yield substantially lower RMS differences than the other five methods. However, the median method performs slightly better. Besides, ~~it inherently takes into account the standard deviation of the GNSS trends (which is not done by taking the mean~~ the median method is less sensitive to large values caused by GNSS trends with larger uncertainties (for which the mean method is sensitive) and also less to outliers caused by large local VLM differences (for which the variance weighting method is sensitive)~~and it filters the spatial variations in VLM (which is not done by variance weighting)~~.

In Table 4 we analyze the results for different correlation thresholds in more detail by comparing them to the GNSS trends based on the median method. On the left side of the table the RMS, mean and median are shown for all VLM estimates available for each correlation threshold. Setting no correlation thresholds yields 294 trends for comparison, while setting a threshold at 0.7 leaves only 155. While the number of trends decreases, the RMS decreases as well, indicating that the correlation thresholds can serve as a selection procedure, which filters out outliers. This is confirmed by Fig. 8, in which we see the decrease of the number of available trends, but also the removal of the outliers. If the threshold is set to 0.7 only three discrepancies in trends of larger than 3 mm yr^{-1} are found. Note that the reduction in RMS is not only caused by the removal of problematic ALT-TG time series. Large earthquakes for example might induces jumps or non-linear behaviour in both the TG and GNSS time series, so the larger ~~spread~~ range in Fig. 8 for no correlation threshold may be partly attributed problematic GNSS trends. In the last row the *Wöppelmann and Marcos* (2016) trends are compared with our GNSS trends. It has a similar RMS with the 0.4-0.5 correlation threshold trends, but it is computed with a substantially smaller number of trends.

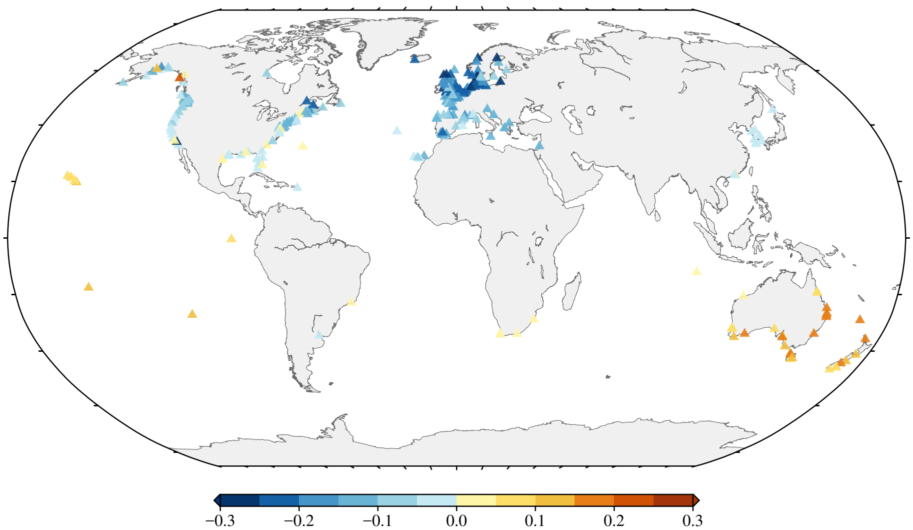

Figure 9. Trend differences (mm yr^{-1}) between the GNSS and ALT-TG time spans induced by non-linear VLM due to present-day mass redistribution.

On the right side of the table, we only included TGs for which all solutions are available, which reduces the number from 155 to 137, because W&M trends are also considered for comparison. The RMS of differences for 155 stations is only slightly larger as will be shown below in Table 5. Note that the RMS of the residuals using ALT-TG from W&M, is already 0.14 mm yr^{-1} lower than those in the study of *Wöppelmann and Marcos* (2016) and about 0.4 mm yr^{-1} less than in *Pfeffer and Allemand* (2016), who incorporated only 109 and 113 stations, respectively. This is a consequence of the combined use of the ~~NGL trends with median weighting~~ median of the NGL trends and the selection based on correlation. Our altimetry solutions further decrease the RMS by another 0.1 mm yr^{-1} compared to W&M, even when no threshold on the correlation is set. In the study of *Wöppelmann and Marcos* (2016), the along-track altimetry ALT-TG trends performed worse than the AVISO results.

The reason for this discrepancy could be the latitudinal intermission bias, or the small radius around the TG used in that study for including altimetry measurements.

Increasing the correlation threshold only slightly reduces the RMS between GNSS and ALT-TG trends and the additional weighting has a neglectable effect on the RMS. As mentioned before, the threshold increase and correlation weighting generally reduced the standard deviation (Fig. 4) of the ALT-TG time series and Fig. 6 showed coherent changes in trend. Additionally, the NGL and ULR trends showed an RMS of differences and ~~the weighting methods spreads~~ range between the GNSS approaches of more than a millimeter. We argue that the absence of a clear improvement or a change in RMS due to correlation thresholds is a result of the relatively large noise in the GNSS trends. The histogram in Fig. 8 shows that for 155 stations, only three discrepancies are larger than 3 mm yr^{-1}. These TGs (which are located at Galveston (US), Eureka (US) and the Cocos Islands (Australia)~~are all~~) are inspected and we find that the neighbouring GNSS stations are located at the other side of lagunas or on different islands. ~~The GNSS is~~ Local VLM differences between the GNSS stations and the TG are therefore the likely cause for the largest discrepancies and not the ALT-TG trend.

The third column of Table 4 shows that the mean is in all cases negative, i.e. the GNSS trends are larger than those of ALT-TG. Trends obtained with correlations -1.0, 0.0, 0.1 and 0.2 are barely statistically different from zero based on a 95% confidence level, while the others are not. The 95 % confidence level is taken as two times the standard deviation of the mean of the residual trends ($\frac{\sigma_n}{\sqrt{N}}$, where N is the number of trends and σ_n the standard deviation of the residual trends). In the right 'mean' column for the 137 stations, the means are statistically insignificantly different from zero at the 95%-confidence level, wheras at a 90%-confidence level several are not. The medians in both columns are closer to zero and deviate up to 0.2 mm yr^{-1} from the mean, which indicates a slightly skewed distribution.

There is a non-linear VLM signal due to present-day mass loss in both GNSS and ALT-TG trends and since they cover different time spans this causes small systematic differences between trends. Due to the inhomogeneous distribution of the TGs and the spatial signal of non-linear VLM, this does not only affect the mean, but also the skewness of the distribution. In Fig. 9 the trend difference between the GNSS and ALT-TG methods are visualized for all 294 stations. Most of the negative differences in trends are observed in Europe and parts of North-America, while positive differences in trends are observed in Australia. In Europe there is an uplift due to present-day mass loss, which increases over the last few years. Since the GNSS time series are generally shorter, they measure a larger uplift signal. By subtracting the present-day VLM that GNSS observes from altimetry observations, we obtain negative signals in Europe.

We applied a correction for the effect of present-day mass loss to the trends for the 155 stations for which a trend is found with all methods in Table 5. Similarly, this is done for the 137 stations, so that the results are comparable with Table 4. There is no significant reduction in RMS. The maximal deviation of the median from zero is 0.06 mm yr^{-1} for the 155 stations and maximally 0.07 mm yr^{-1} for the 137 stations, which is a reduction with respect to the values listed in Table 4. The mean is also reduced to approximately -0.1 mm yr^{-1}, which is statistically equal to zero. This result is at the level of the noise in the determination of the ITRF origin (*Santamaría-Gómez et al.*, 2017) and it is smaller than the 0.4 mm yr^{-1} to which global mean sea level trends from altimetry are gauranteed (*Mitchum*, 2000). Unless it is proven that the altimeters are more stable and the uncertainties in the ITRF origin are reduced, a mean of trend differences closer to zero cannot be expected.

Table 5. Statistics of trend differences with GPS for various solutions after applying a correction for non-linear VLM.

Correlation	NoT: 155			NoT: 137		
	RMS mm yr^{-1}	Mean mm yr^{-1}	Median mm yr^{-1}	RMS mm yr^{-1}	Mean mm yr^{-1}	Median mm yr^{-1}
-1.0	1.231	-0.102	-0.039	1.223	-0.100	0.030
0.0	1.225	-0.109	-0.027	1.215	-0.108	0.031
0.0	1.223	-0.106	0.016	1.209	-0.105	0.048
0.1	1.220	-0.107	-0.014	1.208	-0.107	0.034
0.1	1.222	-0.104	0.003	1.208	-0.104	0.072
0.2	1.220	-0.099	0.016	1.207	-0.096	0.027
0.2	1.221	-0.101	-0.001	1.206	-0.101	0.059
0.3	1.223	-0.091	0.011	1.211	-0.090	0.018
0.3	1.221	-0.098	-0.001	1.207	-0.098	0.036
0.4	1.226	-0.087	0.011	1.214	-0.085	0.021
0.4	1.223	-0.092	0.008	1.209	-0.091	0.037
0.5	1.225	-0.088	0.020	1.212	-0.086	0.042
0.5	1.222	-0.090	0.027	1.208	-0.088	0.045
0.6	1.222	-0.087	-0.007	1.202	-0.088	0.018
0.6	1.222	-0.087	-0.006	1.201	-0.089	0.028
0.7	1.220	-0.071	0.021	1.202	-0.073	0.037
0.7	1.219	-0.074	0.012	1.201	-0.075	0.036

4 Conclusions

We presented new ways to estimate VLM at TGs from GNSS ~~observations~~ and differenced ALT-TG time series. A comparison is made between eight different ~~weighting~~ methods to obtain VLM at the TG from NGL GNSS trends. The ~~spread~~ range of the trends between the ~~weighting methods~~ approaches is at the same level as the standard deviations of the GNSS trends, with a mean of 0.92 mm yr^{-1} and a median of 0.71 mm yr^{-1}. A comparison with the estimates of ULR5 (*Wöppelmann and Marcos*, 2016) at 70 stations yielded an RMS of at least 1.05 mm yr^{-1}. A comparison with ALT-TG showed that using the median of all neighbouring GNSS provided the best results.

For the ALT-TG trends we used along-track ~~observations~~ data from the Jason series of altimeters. At every 6 km along-track ~~observations~~ data were stacked, to create time series. The time series were low-pass filtered with a moving-average filter of one year and correlated with low-pass filtered TG time series. An average or weighted monthly time series for altimetry was created taking only into account the time series corresponding to correlations above a threshold. The TG time series were subtracted from the average of monthly low-pass filtered altimetry time series to create a ALT-TG time series. Using the Hector software between 344 and 663 trends were computed from the ALT-TG time series, depending on the correlation threshold set.

The standard deviation of the ALT-TG time series was reduced on average by approximately 10 % when a correlation threshold of 0.7 was used. Spatially coherent differences in trends between various tresholds are observed at the east coast of the US and in Norway. We argue that residual interannual ocean variability in ALT-TG time series can locally induce VLM trend biases, especially when time series are short. For 155 stations globally distributed, increasing the correlation threshold does not significantly affect the RMS of differences between GNSS and ALT-TG trends. However, the correlation threshold also works as a selection procedure. When considering 294 VLM estimates from GNSS and ALT-TG at the same TGs for comparison, with no threshold yielded an RMS of differences of 2.14 mm yr^{-1}, whereas an RMS of 1.22 mm yr^{-1} was reached using 155 stations and a threshold of 0.7. This is a substantial improvement with respect to the 1.47 mm yr^{-1} RMS of *Wöppelmann and Marcos* (2016) at 109 TGs, the best result so far. Note that increasing the threshold considerably reduces the number of time series in the southern hemisphere and therefore other thresholds might be better depending on the purpose.

The comparison with tide gauges also reveals that the trends from ALT-TG are biased low (similar to *Wöppelmann and Marcos* (2016)), even though this is barely significant. Using mass redistribution fingerprints, a correction is applied for trend differences caused by non-linear behaviour of present-day mass changes. The RMS of differences is barely affected, but the mean of differences is changed from about -0.2 to -0.1 mm yr^{-1}, which is now statistically insignificant.

The trends in this publication (median GNSS and ALT-TG for all correlations) ~~will be made publicly available.~~ are provided in the supplementary material. The ALT-TG trends are accompanied by errors bars computed using the Hector software. The provided uncertainties for the GNSS using the MAD from the median of the trends within 50 km scaled by 1.4826 (*Wilcox*, 2005), similar to the MIDAS algorithm. If only a single GNSS station is present, the MIDAS uncertainty is provided. If two GNSS stations are present and both trends are statistically equal, it takes the square-root of the mean of the GNSS variances to avoid very small error bars. When no correlation threshold is used ~~this implies that~~ 663 ALT-TG and 570 GNSS trends are available at 939 different TGs. Setting the correlation threshold to 0.7, the number of trends at different TGs decreases to 759. Depending on the application, the value of the treshold can be varied to find an optimum between the reliability and the number of trends. If both GNSS and ALT-TG trends are available, we recommend to use GNSS trends, because of correlated residual ocean signals between various ALT-TG time series. However, if a large discrepancy (> 3 mm yr^{-1}) is found between the GNSS and ALT-TG trends, we recommend to use the ALT-TG trend, because the culprit is likely local VLM differences between the TG and the GNSS stations. The GNSS - ALT-TG histogram for no correlation threshold reveals large discrepancies between the two methods up to 10 mm yr^{-1}. While the problem with ALT-TG trends are mostly resolved by setting a higher threshold, the GNSS trends might still require some inspection before they are used in sea level studies. A faster practise is to use trend uncertainties, that carry information about the linearity of the trends, and when the MAD is used as described above, also information about local VLM variability. However, when only one GNSS station is present the information about local VLM variations is absent.

Appendix A: Intermission biases

The latitude-dependent intermission biases are computed from 1/8 degree latitudinally averaged sea surface height differences between TOPEXPOSEIDON and Jason-1 (TP-J1) and Jason-1 and Jason-2 (J1-J2). For the TP-J1 bias four separate polygons are estimated for ascending tracks and four for the descending tracks, while for J1-J2 a single polygon is estimated. Depending on the geophysical corrections and the processing of the altimetry data, not all parameters are statistically different from zero based on variances of the residuals. However, to be consistent with the study of *Ablain et al.* (2015), we maintain the polygons as such.

Table 6. Values for the parameters of the latitudinal intermission bias correction. These numbers are added to the sea surface height anomalies of the respective satellites. TP asc. and TP desc. indicates the function variables that should be added to the ascending and descending tracks of TOPEX/POSEIDON using Eq. (4), respectively. J2 indicates the function variables to be used for Jason-2.

	TP asc.		TP desc.		Jason-2	
Parameter	Lat(deg)	Value	Lat(deg)	Value	Lat(deg)	Value
c_0(mm)	(-66.2,-1.5)	80.3	(-66.2,-1.5)	77.3	(-66.2,66.2)	98.1
c_1(mm deg^{-1})		$-2.3\cdot10^{-1}$		$-1.7\cdot10^{-1}$		$-9.3\cdot10^{-2}$
c_2(mm deg^{-2})		$-1.1\cdot10^{-2}$		$1.2\cdot10^{-3}$		$3.8\cdot10^{-3}$
c_3(mm deg^{-3})		$-3.0\cdot10^{-4}$		$2.9\cdot10^{-4}$		$8.4\cdot10^{-7}$
c_4(mm deg^{-4})		$-2.4\cdot10^{-6}$		$3.8\cdot10^{-6}$		$-7.6\cdot10^{-7}$
c_0(mm)	(-1.5,0.2)	83.8	(-1.5,1.3)	79.9		
c_1(mm deg^{-1})		1.3		2.4		
c_2(mm deg^{-2})		-1.3		$5.2\cdot10^{-1}$		
c_3(mm deg^{-3})		$-5.3\cdot10^{-1}$				
c_4(mm deg^{-4})						
c_0(mm)	(0.2,4)	84.9	(1.3,4)	73.3		
c_1(mm deg^{-1})		$-8.0\cdot10^{-1}$		13.7		
c_2(mm deg^{-2})		$-8.6\cdot10^{-1}$		-5.1		
c_3(mm deg^{-3})		$1.5\cdot10^{-1}$		$4.9\cdot10^{-1}$		
c_4(mm deg^{-4})						
c_0(mm)	(4,66.2)	72.9	(4,66.2)	75.8		
c_1(mm deg^{-1})		$8.1\cdot10^{-1}$		$7.9\cdot10^{-1}$		
c_2(mm deg^{-2})		$-2.8\cdot10^{-2}$		$-3.3\cdot10^{-2}$		
c_3(mm deg^{-3})		$3.4\cdot10^{-4}$		$6.4\cdot10^{-4}$		
c_4(mm deg^{-4})		$-1.1\cdot10^{-6}$		$3.9\cdot10^{-6}$		

Acknowledgements. This study is funded by the Netherlands Organisation for Scientific Research (NWO) through VIDI grant 864.12.012 (Multi-Scale Sea Level (MuSSeL)).

The MIDAS GNSS trends are obtained from the Nevada Geodetic Laboratory (NGL).

http://geodesy.unr.edu/

The altimetry data are obtained from the Radar Altimetry Database System (RADS).

http://rads.tudelft.nl/rads/rads.shtml

Permanent Service for Mean Sea Level (PSMSL), 2017, "Tide Gauge Data", Retrieved 1 November 2016.

http://www.psmsl.org/data/obtaining/

We would like to thank Marta Marcos and Guy Wöppelmann for sharing their trend estimates.

References

Ablain, M., Cazenave, A., Larnicol, G., Balmaseda, M., Cipollini, P., Faugère, Y., ... & Benveniste, J. (2015). Improved sea level record over the satellite altimetry era (1993-2010) from the Climate Change Initiative Project. Ocean Science, 11, 67-82.

Andres, M., Gawarkiewicz, G. G., & Toole, J. M. (2013). Interannual sea level variability in the western North Atlantic: Regional forcing and remote response. Geophysical Research Letters, 40(22), 5915-5919.

Blewitt, G., Kreemer, C., Hammond, W. C., & Gazeaux, J. (2016). MIDAS robust trend estimator for accurate GPS station velocities without step detection, Journal of Geophysical Research: Solid Earth, 121, 2054-2068, doi:10.1002/2015JB012552.

Bos, M. S., Fernandes, R. M. S., Williams, S. D. P., & Bastos, L. (2013). Fast error analysis of continuous GNSS observations with missing data. Journal of Geodesy, 87(4), 351-360.

Bos, M. S., Williams, S. D. P., Araújo, I. B., & Bastos, L. (2013). The effect of temporal correlated noise on the sea level rate and acceleration uncertainty. Geophysical Journal International, 196(3), 1423-1430.

Bouin, M. N., & Wöppelmann, G. (2010). Land motion estimates from GPS at tide gauges: a geophysical evaluation. Geophysical Journal International, 180(1), 193-209.

Calafat, F. M., Chambers, D. P., & Tsimplis, M. N. (2013). Inter-annual to decadal sea-level variability in the coastal zones of the Norwegian and Siberian Seas: The role of atmospheric forcing. Journal of Geophysical Research: Oceans, 118(3), 1287-1301.

Carrère, L., & Lyard, F. (2003). Modelling the barotropic response of the global ocean to atmospheric wind and pressure forcing - comparisons with observations, Geophysical Research Letters, 30(6), 1275, 2003.

Cazenave, A., Dominh, K., Ponchaut, F., Soudarin, L., Cretaux, J. F., & Le Provost, C. (1999). Sea level changes from Topex-Poseidon altimetry and tide gauges, and vertical crustal motions from DORIS. Geophysical Research Letters, 26(14), 2077-2080.

Desai, S., Wahr, J., & Beckley, B. (2015). Revisiting the pole tide for and from satellite altimetry. Journal of Geodesy, 89(12), 1233-1243.

Ducet, N., Le Traon, P. Y., & Reverdin, G. (2000). Global high-resolution mapping of ocean circulation from TOPEX/Poseidon and ERS-1 and -2. Journal of Geophysical Research: Oceans, 105(C8), 19477-19498.

Farrell, W. E., & Clark, J. A. (1976). On postglacial sea level. Geophysical Journal International, 46(3), 647-667.

Frederikse, T., Riva, R., Kleinherenbrink, M., Wada, Y., Broeke, M., & Marzeion, B. (2016). Closing the sea level budget on a regional scale: Trends and variability on the Northwestern European continental shelf. Geophysical Research Letters, 43(20).

Gazeaux, J., Williams, S., King, M., Bos, M., Dach, R., Deo, M., Moore, A. W., Ostini, L., Petrie E., Roggero, M., Teferle, F. N., Olivares, G., & Webb, F.H. (2013). Detecting offsets in GPS time series: First results from the detection of offsets in GPS experiment. Journal of Geophysical Research: Solid Earth, 118(5), 2397-2407.

Gutenberg, B. (1941). Changes in sea level, postglacial uplift, and mobility of the Earth's interior, Bull. Geol. Soc. Am., 52, 721-772.

Hamlington, B. D., Thompson, P., Hammond, W. C., Blewitt, G., & Ray, R. D. (2016). Assessing the impact of vertical land motion on twentieth century global mean sea level estimates. Journal of Geophysical Research: Oceans, 121(7), 4980-4993.

Hughes, C.W., & Meredith, M.P. (2006). Coherent sea-level fluctuations along the global continental slope. Phil. Trans. R. Soc., 364, 885-901.

Holgate, S. J., Matthews, A., Woodworth, P. L., Rickards, L. J., Tamisiea, M. E., Bradshaw, E., Foden, P. R., Gordon, K. M., Jevrejeva, S., & Pugh, J. (2012). New data systems and products at the permanent service for mean sea level. Journal of Coastal Research, 29(3), 493-504.

Masters, D., Nerem, R. S., Choe, C., Leuliette, E., Beckley, B., White, N., & Ablain, M. (2012). Comparison of global mean sea level time series from TOPEX/Poseidon, Jason-1, and Jason-2. Marine Geodesy, 35(sup1), 20-41.

Milne, G. A., & Mitrovica, J. X. (1998). Postglacial sea-level change on a rotating Earth. Geophysical Journal International, 133(1), 1-19.

Mitchum, G. T. (1998). Monitoring the stability of satellite altimeters with tide gauges. Journal of Atmospheric and Oceanic Technology, 15(3), 721-730.

Mitchum, G. T. (2000). An improved calibration of satellite altimetric heights using tide gauge sea levels with adjustment for land motion. Marine Geodesy, 23(3), 145-166.

Nerem, R. S., & Mitchum, G. T. (2002). Estimates of vertical crustal motion derived from differences of TOPEX/POSEIDON and tide gauge sea level measurements. Geophysical Research Letters, 29(19).

Ostanciaux, É., Husson, L., Choblet, G., Robin, C., & Pedoja, K. (2012). Present-day trends of vertical ground motion along the coast lines. Earth-Science Reviews, 110(1), 74-92.

Petit, G., & Luzum, B. (2010). IERS conventions (2010) (No. IERS-TN-36). BUREAU INTERNATIONAL DES POIDS ET MESURES SEVRES (FRANCE).

Pfeffer, J., & Allemand, P. (2016). The key role of vertical land motions in coastal sea level variations: a global synthesis of multisatellite altimetry, tide gauge data and GPS measurements. Earth and Planetary Science Letters, 439, 39-47.

Pujol, M. I., Faugére, Y., Taburet, G., Dupuy, S., Pelloquin, C., Ablain, M., & Picot, N. (2016). DUACS DT2014: the new multi-mission altimeter data set reprocessed over 20 years. Ocean Science, 12(5).

Ray, R. D., Beckley, B. D., & Lemoine, F. G. (2010). Vertical crustal motion derived from satellite altimetry and tide gauges, and comparisons with DORIS measurements. Advances in Space Research, 45(12), 1510-1522.

Riva, R. E., Frederikse, T., King, M. A., Marzeion, B., & van den Broeke, M. R. (2017). Brief communication: The global signature of post-1900 land ice wastage on vertical land motion. The Cryosphere, 11(3), 1327.

Roemmich, D., & Gilson, J. (2009). The 2004-2008 mean and annual cycle of temperature, salinity, and steric height in the global ocean from the Argo Program. Progress in Oceanography, 82(2), 81-100.

Santamaría-Gómez, A., Gravelle, M., Collilieux, X., Guichard, M., Míguez, B. M., Tiphaneau, P., & Wöppelmann, G. (2012). Mitigating the effects of vertical land motion in tide gauge records using a state-of-the-art GPS velocity field. Global and Planetary Change, 98, 6-17.

Santamaría-Gómez, A., Gravelle, M., Dangendorf, S., Marcos, M., Spada, G., & Wöppelmann, G. (2017). Uncertainty of the 20th century sea-level rise due to vertical land motion errors. Earth and Planetary Science Letters, 473, 24-32.

Santamaría-Gómez, A., Gravelle, M., & Wöppelmann, G. (2014). Long-term vertical land motion from double-differenced tide gauge and satellite altimetry data. Journal of Geodesy, 88(3), 207-222.

Scharroo, R., Leuliette, E. W., Lillibridge, J. L., Byrne, D., Naeije, M. C., & Mitchum, G. T. (2012). RADS: Consistent multi-mission products. In Proceedings of Symposium on 20 Years of Progress in Radar Altimetry (Vol. 20).

Vinogradov, S. V., & Ponte, R. M. (2011). Low-frequency variability in coastal sea level from tide gauges and altimetry. Journal of Geophysical Research: Oceans, 116(C7).

Wahr, J., Nerem, R. S., & Bettadpur, S. V. (2015). The pole tide and its effect on GRACE time-variable gravity measurements: Implications for estimates of surface mass variations. Journal of Geophysical Research: Solid Earth, 120(6), 4597-4615.

Watson, C. S., White, N. J., Church, J. A., King, M. A., Burgette, R. J., & Legresy, B. (2015). Unabated global mean sea-level rise over the satellite altimeter era. Nature Climate Change, 5(6), 565-568.

Wilcox, R. R. (2005), Introduction to Robust Estimation and Hypothesis Testing, Elsevier Academic Press, Burlington, Mass.

Williams, S. D. P. (2008). CATS: GPS coordinate time series analysis software. GPS Solutions, 12(2), 147-153, doi:10.1007/s10291-007-0086-4.

Wöppelmann, G., Miguez, B. M., Bouin, M. N., & Altamimi, Z. (2007). Geocentric sea-level trend estimates from GPS analyses at relevant tide gauges world-wide. Global and Planetary Change, 57(3), 396-406.

Wöppelmann, G., Marcos, M., Santamaría-Gómez, A., Martín-Míguez, B., Bouin, M. N., & Gravelle, M. (2014). Evidence for a differential sea level rise between hemispheres over the twentieth century. Geophysical Research Letters, 41(5), 1639-1643.

Wöppelmann, G., & Marcos, M. (2016). Vertical land motion as a key to understanding sea level change and variability. Reviews of Geophysics, 54(1), 64-92.