

Associate Editor

Both reviewers expressed that the proposed MODWT method for slow-slip detection shows potential, but were concerned about whether the method could be applied successfully to regions without existing high-quality slow-slip or tremor catalogs. An existing catalog appears to be necessary to select the detection threshold from the ROC diagram, and to identify false detections. One way to address this concern would be, as Reviewer #1 suggests, by applying the method to another subduction zone. Alternatively, as Reviewer #2 suggests, the manuscript could be revised to reframe the proposed method. Successfully applying the method to another subduction zone would make the strongest case for the MODWT method, although it would be more work. Either of these options for revision are acceptable to me, so I leave the choice to the authors.

We added a section about the analysis of GPS data from New Zealand to show how the method could still be applied if tremor data are not available.

Reviewer 1

In "Detection of slow slip events using wavelet analysis of GNSS recordings", the authors propose a technique to detect slow slip events using the Maximal Overlap Discrete Wavelet Transform (MODWT). While previous authors have used or advocated the use of wavelets for the detection of slow slip in geodetic time series, this manuscript provides a rubric for detection and provides a connection to slow slip duration. Applying this rubric to slow slip events in northern Cascadia, the authors provide a method for using the MODWT on both geodetic and tremor duration time series to demonstrate the general applicability of the technique. The authors advocate using this technique in situations where the connection between slow slip and tectonic tremor is less obvious.

This manuscript provide a useful and much needed exploration and explanation of the relationship between wavelet levels and offset duration. In particular, advocating for the use of the MODWT over the standard discrete wavelet transform provides a nice way to avoid boundary effects at the edge of each wavelet octave and provides better temporal location for identification of wavelet peaks. The applicability to the case of Cascadia presented in the manuscript is clear and concise as well. My one big wish for this manuscript is a demonstration of its power in the situations described in the Introduction (lines 69-74), in particular, for those subduction zones where the connection between tremor and slow slip is not straightforward. To that end, I think the biggest test missing from this manuscript is a test in one of the problematic subduction zones, such as New Zealand or Alaska. Much of the demonstration of the power of this technique lies in selecting the threshold. For instance, how applicable are the thresholds determined from northern Cascadia data to other regions (say NZ or AK)? If those subduction zones were examined, how would one go about selecting a threshold? In short, is it possible to select a threshold in the absence of a catalog or does threshold selection depend on slow slip events already being know from other methods?

We added a section about the analysis of GPS data from New Zealand to show how the method could still be applied if tremor data are not available.

To select the threshold in the absence of a tremor catalog, a first possibility is to fit the value of the threshold using data from a subduction zone where tremor data are available. That is what

we did in the manuscript: we use for New Zealand the same threshold as we use for northern Cascadia. Another possibility is to try different thresholds, and keep the one that give the most plausible results, with some slow slip events detected but without too much transients that look like noise.

While Figure 10 does show that nearly all the thresholds examined demonstrate a higher true positive rate than false positive, not all threshold value have equal identification power. Additionally, the actual value of the threshold is missing from the diagram (it would be a third dimension in the diagram). It is therefore difficult to ascertain whether the threshold levels explored provide complete coverage of reasonable values.

We modified Figure 10 to better show the results corresponding to the different values of the threshold, by using a different color for each point depending on the values taken by the two thresholds.

Thus, while I feel that this manuscript highlights a potentially useful mathematical technique for identifying slow slip in geodetic time series, its broader applicability is not demonstrated. This manuscript would be improved through a demonstration of the proposed method, along with a comparison to an existing slow slip catalog, in a separate slow slip area.

We added a comparison of slow slip events detected with this method and slow slip events detected by Todd and Schwartz (2016) in New Zealand.

In addition, I note the following items:

(Line 172): I'm pretty sure the reference frame for the PANGA GPS time series is ITRF14. Orbits from JPL in the IGB08 reference frame end in early-mid 2018 and the time series for PGC5 shown in Figure 3 extend at least half-way through 2020. Therefore, the orbits used to process PGC5 can't be in IGB08.

Thank you for pointing out this mistake. We corrected the sentence.

(Line 391) If the time series is categorized by the values -1, 0, and 1, wouldn't that be a trinary time series, not a binary time series?

We replaced binary by trinary.

(Lines 458-460) These lines discuss a small displacement in the time series at 2007.5, but I don't see it. Should I be looking at Figure 9? If so, which latitude band should I look at and which time series are they composed of? That would be the place to look for causal offsets.

We have added a Figure in the Supplement where we highlight the four possible small displacement that we see on the wavelet details.

(Figure 2) The caption text is not clear, in particular the phrase "for two locations of the missing values" is vague. The text (lines 290-295) does provide a better explanation of this figure however.

Rewording the caption to provide a more self-contained explanation would improve the caption. Also, the values on the time axis for the right hand column are difficult to read as formatted.

We modified the caption and the values on the time axis.

(Figure 5) The ordering of the data and the detail levels in this figure are opposite that of the Figures 1-3. For the sake of comparison, these should all be in the same order.

We reordered the data and the detail levels.

(Figure 6) When producing the stacked time series, is there overlap/redundancy? Looking at the map in Figure 4 suggests that most triangles will be made with an overlapping number of stations. How will this affect the mapping of strange station behavior? In particular, could the displacement discussed on Lines 458-460 all be caused by just one misbehaving station that shows up in multiple bins?

We added a table summarizing the number of stations and the number of overlapping stations for each location. There is no evidence for a single station misbehaving at time 2007.5 and several GPS stations show a small increase in the displacement in the eastern direction at that time. However, there are many missing values around time 2007.5, so it is difficult to conclude.

(Figure 7) Why do the obvious red-to-blue transitions in the tremor data seem to disappear from the detail level 7 figure, but are present in the detail level 8 (Figure 6) and detail level 6 (Figure 8) data? Can anything be interpreted from that pattern?

The number of red-to-blue transitions depends on the threshold chosen. For detail level 7, the false positive rate is smaller than for details levels 6 and 8, so all the transitions may not be seen clearly.

(Figure 9) When stacking levels 6, 7 and 8, won't the lower numbered levels dominate because of their (generally) larger amplitudes (as seen in Figure 3)? Is it valuable to normalize the levels before summing, or will this cause other issues?

The amplitude of level 6 is not much higher than the amplitude of level 8, so we do not think that it would dominate the stacking. A more complex way of stacking the coefficients would be to apply a wavelet-based denoising method, where each wavelet vector \tilde{W}_j is modified before reconstructing the initial time series by applying an inverse wavelet transform to the wavelet vectors. We could then stacked the denoised GPS time series instead of stacking the wavelet details.

(Figure 10) Is it possible to select a threshold in the absence of a catalog or does threshold selection depend on slow slip events already being know from other methods?

It is possible to choose a threshold without a tremor catalog. However, without the comparison with the tremor, we cannot decide which slow slip events are real and which slow slip events are false detections and we cannot produce the ROC curve.

(Figure 11) What are the dates selected by this algorithm? How do they compare with the dates given in Michel et al., 2019? Most look spot on, but others, in particular the 2009 event look like they are quite distant in time.

The timing of the events are taken from the plot of moment rate as a function of time from Figure 8 of Michel et al. (2019). For each event, we choose the middle of the time window as the time of the slow slip event.

There are some awkward sentences as well:

(Line 190-191): "where n is some integer greater than or equal to" sounds better than "higher or equal".

We modified this sentence.

(Line 199): The terminology "details" and "smooths" sounds awkward, is that the standard terminology in the wavelet literature?

Yes, this is the standard terminology in the wavelet literature.

(Line 287-288) The sentence "The straight line starts at and ends at." is incomplete.

We removed the incomplete sentence.

Reviewer 2

The authors present a wavelet-based method to identify slow slip events within the GNSS position time series. Rather than looking at the geodetic record of slow slip as a sum of periodic basis functions, the wavelet method allows for the isolation of slow slip signals across different time scales at different times during the period of study. The approach is of interest to the community and I think that this manuscript will eventually be publishable.

That said, my major concern is that the method as presented will be very difficult to apply to another region of study, especially one without a systematically generated catalog of tectonic tremor or low-frequency earthquakes. Here the authors validate the detections (i.e. establish a detection threshold) within the geodetic data with detections made in the seismic data. Looking at figures 6-8, there are many wiggles within the time series, and it's not obvious to me how to interpret these signals without the well-established catalogs of A. Wech and the PNSN (tremor) and Michel et al. (2019) (slow slip); such detailed and long-duration catalogs are not available outside of Japan to validate the results of the presented method. Aggravating my concern is the premise that the authors themselves set up in the introduction of the manuscript, that a method is needed to detect slow slip when tremor is not available as an independent control; the presented method does not reach that goal. I would also highlight that slow slip in northern Cascadia is some of the most well behaved slow slip across all subduction zones, with short-ish recurrence intervals and consistently clear geodetic signals captured by a relatively dense GNSS network, that will likely show up more reliably than elsewhere with this sort of approach.

We added an application to GPS data in New Zealand, where tremor is not well spatially and temporally correlated, to show the broader applicability of the method.

The figures for the most part are clear, but there are some improvements to be made (see comments below). I also find that the introduction is lacking details and there are way too few articles cited for background reading for interested readers (cf. first paragraph with only two unique citations that discusses observations made across many subduction zones).

We added several references in the introduction.

I think a revised version of this manuscript that has reframed the purpose of the approach will be an interesting contribution that is appropriate to be published in BSSA. I list other minor comments below:

general - There is no discussion of common mode signals that will surely stack constructively across GNSS stations. The removal of sinusoidal seasonal signals will remove some common mode signal, but there can still exist some non-tectonic signals that will stack constructively within the bandwidth of interest here. There is a reasonable chance that this will have little impact on the results, but it should at least be verified/discussed given the proposed processing.

We added a short discussion on common modes.

general - I would avoid using the term ETS as it only adds confusion, as lines 69-74 highlight. I would suggest using "slow slip" to describe geodetic observations of aseismic slip and "tectonic tremor" for the accompanying seismic signals. The following paragraph further underlines this potential confusion: discussing the moment release of an "ETS event" is ambiguous. As described, it is not obvious to those unfamiliar with slow slip and tremor that the aseismic moment of slow slip dominates the seismic moment of tremor, because ETS conflates the geodetic observations with the seismic ones. I would simply use the term slow slip when referring to the dominant aseismic transient.

We replaced "ETS" by "slow slip" wherever it seemed more appropriate.

general - I'm not sure that the "true detection" terminology used here is quite appropriate. This implies that we have an absolute knowledge of what's happening on the plate interface, and that there can't be new detections made here without already be corroborated by other catalogs. Perhaps "robust detection" might be more appropriate? In any case, this underlines my above major comment that this method is difficult to interpret without other catalogs.

"true detection" is the standard term usually used in the literature on machine learning methods for classification. We agree that it would be valid only if there was perfect correlation between slow slip and tremor. As it may not be the case, we replace the term by "robust detection".

line 42 - "feature" is a strange word choice and does not tell the reader much at all

We replaced "feature" by "phenomenon".

line 60 - "it is more abundant in some places" doesn't tell us much

We modified this sentence.

line 64 - "for which" -> whose

We modified this sentence.

line 67 - Obara (2002) did not discuss slow slip, only tectonic tremor. I would suggest something like Obara et al., GRL, 2004

We replaced the reference by Obara et al., 2004.

line 126 - "levels" does not mean very much without context

We replaced "levels" by "time scales".

lines 148 - "details" same comment as above; this general term doesn't mean anything without context

We replaced "details" by "components at different time scales".

line 151-152 - why is there supposed to be a difference between the two tidal gauges and what does that mean? Not clear without the reader being forced to go and read the cited article.

We clarified this sentence.

line 156 - I assume longitudinal displacements is supposed to mean east-west motion?

Yes, we modified this sentence.

line 287-288 - missing text here

We removed this sentence.

figure 2 - I would say figure 2 could be moved to the appendix or supplementary material. It's not key to the message of the manuscript

We moved Figure 2 to the Supplementary Material.

figures 6-8 - I think this figure would be more easily interpretable as well as much effective if only one panel was used, with alternating geodetic and seismic data for a given location (red triangle from Figure 4) in two different colors (black and something else). The red/blue rectangles that indicate detections could be four similar colors for geodetic/seismic detections (blue/purple for negative peaks and red/orange for positive peaks for example). I'd also suggest that figures 7 and 8 could be made into supplementary figures as they show similar information to figure 6

To better compare between events detected with GPS data and events detected with tremor data, we added figures where both set events are added on top on the wavelet details. We believe the figures are thus more easily readable and interpretable.

figure 7 - I'm not sure I understand why the threshold selected here is less effective (there are less events identified) than for levels levels 6 and 8. What could explain this?

This is probably due to the value we have chosen for the threshold. In all cases, we tried to maximize the true positive rate and minimize the false positive rate. The true positive rate is higher for level 7 than for level 8 and the false positive rate is lower for level than for level 6 and 8. The downside is that there are more false negative (missed detections) for the level 7.

figure 10 - I think this figure would be a bit clearer if the 6/7/8 detail levels were in different colors

In Figure 10, we sum details 6, 7 and 8, we did not plot them separately.