

Dear Dr. Maraun,

The following are duplicates of the responses I previously posted on the interactive discussion in response to the three reviewers. In addition, I am adding a copy of the revised text showing all the changes.

In the following document, the reviewers original comments are in italics, my response is in plain text, and any changes made to the revised text is in quotation marks. If the change is within previous text, I indicate it is new by enclosing in [].

I hope you find this revised manuscript is suitable for publication.

Cheers,

Don Chambers

Response to Reviewer # 1

1. Pag 1836, line 12

The signal on which you run the EMD is built only trough noise. Explain better this, please. You should also explain why this experiment is interesting for the target of your paper.

I tried to explain the motivation in the following paragraph, but I can see now that could be confusing. Thus, I have substantially revised this section to better explain my motivation by bringing that discussion into this paragraph and expanding it. I also believe this will answer several of your other questions, notable Comments 6, 7, and 8.

“Moreover, Wu and Huang (2004) have shown that EMD behaves as a low-pass filter. If one runs random noise with a normal Gaussian distribution through the process, low-frequency signals will be seen in the resulting IMFs. They found there is roughly a doubling of the average period with each subsequent mode. This is a significant issue. It means that any random (or near-random, high-frequency) signal will propagate into low-frequency signals in the recovered IMFs. Wu and Huang (2009) proposed a method to quantify the uncertainty caused by this behavior by computing an ensemble mean of IMFs, starting from the same time-series but with different amounts of added random noise.

However, this method ignores that all geophysical time-series have an underlying real signal that has high variance and little serial correlation; i.e., a high-frequency, near-random signal. This “signal” will also be filtered by the EMD process and will likely appear as a quasi-stationary oscillation in higher order IMFs that is not real. Although adding multiple realizations of random noise to a time-series will account for uncertainty in the IMFs from random error in the measurement, it will not account for the shifting of high-frequency signal to low-frequency signal in the recovered IMFs from the high-frequency signal. One of the assumptions in EMD processing is this is captured in the lowest IMFs. Our testing indicates it is not.

To demonstrate the potential size of this problem, we ran EMD on a monthly-resolution time-series that is 150-years long with randomly-correlated values that have a standard deviation of 60 mm. We used 60 mm because this is the standard deviation of residual monthly sea level at San Francisco after fitting and removing a quadratic function plus annual and semiannual sinusoids, so is representative of high-frequency sea level at a typical site, although some sites can have significantly higher variability. We ran another case using a colored noise model that exactly reproduces the autocovariance of the San Francisco tide gauge residuals. The results were nearly identical to the ones shown with the random residuals, so we choose to use random values as they are faster to compute for the several thousand simulations we plan to run. The EMD finds IMFs that have quasi-period fluctuations of nearly 60-years and amplitudes as large as 10 mm (Figure 1); fluctuations at quasi-30-year periods are the same magnitude.”

2. *Pag 1839, line 7*

I suggest to insert the value of the correlation of SOI and PDO also before you have worked on them.

We have added that information in the revision:

“Secondly, because the two indices are slightly correlated (-0.21 , $p < 0.001$) due to similar interannual (< 10 year) variations...”

3. *Pag 1840, line 11*

Where will you note that none accurately captures the input seasonal variation?

We note it right there. We do not show it, because it is beyond the scope of the paper and most authors use EMD to extract low-frequency variations, not seasonal. This is discussed in the introduction. Thus, we feel showing this is irrelevant.

4. *When you compute the correlation between the “best IMF” and the simulated oscillation for the 1000 simulations, it should be interesting for the reader seeing an histogram (for the case 1) to have a better idea of the distribution of this parameter (with also a mean value with error).*

We have added a new figure (Figure 4 in revision) and added the statistics. The new text reads:

“Figure 4 shows the histogram of computed correlations. Note that the correlations were not the same in every simulation. The 13-year oscillation had a mean correlation of 0.66 (standard deviation = 0.09), the 55-year had a mean of 0.52 (standard deviation = 0.11), while the mean correlation of the 80-year signal was 0.74 (standard deviation = 0.09).”

5.

Pag 1842, line 12

Please, explain better the following part, it is not clear: “We isolated this signal by looking at the autocorrelation of the remaining IMFs uncorrelated with either PDO or

ENSO. The IMF with an autocorrelation greater than 0.9 at a lag of 1 year was selected.“

I can see this was a little confusing. I've modified the text thusly:

“In addition, we found in nearly every case (99%) the EMD computed one to two IMFs with a periodic signal that did not correlate highly with either PDO or ENSO, but had a low-frequency. Because this was not always contained in a single IMF between the two prescribed periodic fluctuations, we had to adapt a method to search for it. We isolated this signal by looking at the autocorrelation of the IMFs after removing those correlated with PDO or ENSO, as well as the last mode. To find the mode with the longest-period fluctuation, we examined the autocorrelation at a 1-year lag. Only IMFs with an autocorrelation greater than 0.9 at a lag of 1-year were examined, and if two existed, the one with the higher autocorrelation was selected.”

6. Pag 1836, from line 20

You cite (Wu et al., 2004) saying what they do in their paper, but I know that they do an other thing. I know that they propose a test useful when you analyze a signal in which is present some noise. The test is useful to identify the IMFs due to noise (“non significative IMFs”), in such a way to not consider them for a physic discussion about the intrinsic oscillations present in the signal. Perhaps do you talk about the work present in the other reference you cite in line 20, or about (Wu et al., 2007)? This confused me because you propose your approach as alternative also to their works, but actually I have a comment exactly on the test of (Wu et al., 2004), in particular I don't understand why you don't apply the test (and so I write what follow in 8).

Thanks for pointing out the citation was wrong. You are correct; it should be Wu and Huang (2009), for their Ensemble EMD. I have corrected this in the text (see reply for Point 1). As far as why I don't compute the confidence tests of modes described in Wu and Huang (2004), I don't need to, since I know what the signal should be and can compare the IMFs to it. This is a much more stringent test than the statistical testing of Wu and Huang (2004), which assumes high-frequency signal is properly captured in low IMFs and does not distort higher IMFs. I will discuss this more in my response to Comment 8.

7. You decide to use the random noise to represent high-frequency variability. You chose a noise with a variance to match the variance of the difference between the original data and the model. This signal is actually due both to noise part and some other signal with appreciable characteristic frequencies (one way to appreciate these is, for example, applying the EMD on this). So using the noise to represent this “high frequency variability”, you actually represent only the noise of this (and you should say this).

Here is where we begin to disagree. Using the random values is not just representing noise in this simulation. It is representing the magnitude of high-frequency variability that is real signal in a simple, repeatable simulation. Real signals will of course have serial correlations in the data, but for tide gauge measurements they are not large after a

few months. I actually did test with a colored-noise model that reproduces the autocovariance of the San Francisco tide gauge residuals quite well. The results were no different than the random values, so I chose to only discuss those as they are more easily reproduced.

I added a comment regarding this in the section I added to your Comment 1:

“We ran another case using a colored noise model that exactly reproduces the autocovariance of the San Francisco tide gauge residuals. The results were nearly identical to the ones shown with the random residuals, so we choose to use random values as they are faster to compute for the thousands of simulations we plan to run.”

Moreover, one might expect that by averaging 1000 different IMFs computed from different randomly simulated residuals added to the base signal, that the mean would be zero. This is NOT the case, and what this simulation clearly demonstrates. This is exactly the type of error that can still remain in the Ensemble EMD method, which is what motivated this paper.

8. Performing your experiment, in any of 1000 run, I don't understand why you don't apply the noise test (Wu et al., 2004), that give you the possibility to isolate, and not consider, the part of signal due to noise (“non significative signal”). I know that clearly in the assumption that you represent the “high frequency signal” with noise, all the noise is significative (because you insert it!), nevertheless in this way you discuss also about IMFs due to noise. The crucial point of this, is that performing the EMD on a generic real signal you can apply the test and so avoid to consider the part of signal due to noise. The problem of non apply the test could be that, if you find a “problematic” IMF, you are finding a “problem” in a IMF that could be not actually significative (i.e. due to noise), so in a IMF that is actually due to a part of signal that you can avoid to consider. I observe that you don't discuss about the first IMFs, and usually applying the test you discover that IMFs due to noise are the first but it's not absolutely a rule; so in any of the 1000 simulations, if you find a “problem” in one IMF, before say that this is a real “problem” you should ascertain that is not due to noise, applying the test.

There is no reason to conduct the significance testing of Wu and Huang (2004) in this experiment, as I know what the answer should be and can compare the IMFs to it. This is a much more stringent test than the statistical testing of Wu and Huang (2004), which assumes high-frequency signal is properly captured in low IMFs and does not distort higher IMFs. This is the whole point of the experiment – to determine if the EMD method can find input, known signals in the presence of high-frequency variability with realistic variance. Although I have not performed the significance testing, I suspect it will say the lowest high-frequency modes are not significant, but the higher, low-frequency ones are. Just based on the spread of the ensemble, they appear significant. But they are wrong!

I've added a paragraph after the discussion of the results for Case 1 to try to highlight the probably misinterpretation of an analysis of the EMD results for Case 1. Note, Figure 5 is the old Figure 4.

“More importantly, consider the interpretation of the results from this simplistic simulation in terms of longer-term climate change if only the EMD results (Figure 5) were analyzed. Based only on the returned IMFs, one could easily argue that there was no significant low-frequency variation in the sea level before 1950, then a rather dramatic rise in the 1970s, followed by a return to normal condition. In fact, there were equally large sea level shifts in the early part of the simulated record that were lost due to the way the EMD method partitions the real signal. “

9. Comment on “Case 2” You study if it is possible to reproduce each simulated signal through one IMF. It should be observed a conceptual difference that exist between case 1 and 2. Actually, in principle, you can reproduce each simulated signal through one IMF (for each signal) only in the case 1, because sinusoids respond to the definition of a IMF (Huang et al., 1998). In the case 2, instead, because of ENSO/PDO doesn't respond to IMF's definition, you know already in principle that you can't capture this signal through a single IMF. In principle, you should need at least two IMF (the sum of two IMF doesn't have to respond to the definition of IMF) to reproduce that signal.

So a part of the signal of ENSO/PDO is diffused (necessarily) in other IMFs and we can expect this before performing the EMD.

I don't disagree with anything you say here. I point out what I wrote in the introduction:

“However, there are some potential pitfalls that we believe have not been fully addressed in previous papers utilizing the method. First and foremost, EMD is a purely mathematical deconstruction of the data, with no regard to intrinsic covariance of the signals or physics. Second, it assumes that IMFs are comprised of fluctuating signals where the magnitude of nearby peaks and troughs are balanced to create a zero mean – an assumption not based on any physical requirement, as real observations can have quite large ranges in magnitudes, especially sea level data affected by climate signals and synoptic storm events.”

I stand by all of that, which is in close agreement to what you state. EMD is a mathematical deconvolution of the data, and no individual mode has any physical meaning. But many scientists are trying to analyze individual IMFs for climatic signals. I give a list of some I know about in the beginning section. Thus, I believe my analysis of whether EMD can extract physically meaningful signals in a single mode is justified. I also point out that although you state that in Case 1 you should be able to extract the physical modes in a single EMD, I demonstrate that you cannot.

You say (pag 1843, lines 1-4): “We know of none that find multiple modes that add up to correlate with an ENSO index. Thus, we argue it is more relevant to quantify if EMD can

extract physically meaningful climate modes than whether it can extract modes with interannual and multi-decadal variability". Performing the decomposition with other techniques you obtain different results, clearly we know that each techniques work in a different way. I agree that it's very important that a technique give you modes that have physical meaning. But with EMD, actually mathematically you already know before performing the analysis that you can't obtain this mode in a unique IMF (the same for PDO). After a decomposition, for sure if you retain for some reason that physically this signal is a "unique signal" you have to sum the IMF that give you the signal (clearly if you know already what you want to build, after performing EMD), but EMD can't say this to us (see for example Alberti et al., 2014. NOTE: the citation of this reference should be interesting to give the reader the awareness, although this "critical" paper, that EMD is a delicate tool but useful when used in the right way).

Thanks for sending me that reference. I had not seen it in my literature review. In that case, though, you and your colleagues were using the EMD as a type of low-pass filter by adding up the higher modes that did not pass the significance testing. I have no qualm with that. I have added a comment on low-pass filtering with EMD in the conclusions:

"EMD is a quick and relatively easy tool to identify possible multidecadal fluctuations in a sea level record. However, it should not be used solely to quantify magnitude and phase, nor should analysis of climatic signals be based on a single IMF. One should also be cautious in interpreting acceleration computed from the final IMF, especially in light of the significant errors found in the early and later parts of the low-frequency IMFs (Figures 5, 6, and 7). Where EMD has shown to be useful has been in low-pass filtering data to reduce the impact of high-frequency variability and noise (e.g., Alberti et al., 2014). In that case, the sum of the higher IMFs are used as the low-pass filter."

...

Besides, looking very crudely at fig. 5 seems that the sum of "unsimulated low frequency" IMF and the "PDO IMF" give a good approximation of the total PDO signal, except for first years (regarding this, however, I already said in you that it wasn't clear what you said in pag 1842, line 12).

But the change in the early part of the "PDO" IMF could lead to erroneous climatic interpretations. I added a small comment to the section where I discussed this, noted here in bold.

"As with the ENSO-mode, the mean PDO-mode IMF tracks the general periodicity of the PDO, although the amplitudes are on average too small. Again, the standard deviation suggests any single simulation would give a considerable range of amplitudes. We note that as with the Case 1 results, there is a tendency for an increasing amplitude in time for the mean IMF, **which could be misinterpreted as a sign of climate change; this is inconsistent with the true signal**, where the first two peaks in the given PDO signal are roughly the same magnitude."

According to me, it should be interesting to perform the same experiment using, instead of ENSO and PDO signal like simulated signal, some IMFs ("simulated IMFs") obtained performing EMD on an other signal. I suggest to do it. You could use also the ENSO and PDO to extract and define the "simulated IMFs". I think this procedure should be interesting because in this case, like in case 1, the EMD could actually extract the "simulated IMFs" in a new IMFs from a theoretical point of view.

I don't see how this would provide any more insight than the experiments I have already conducted and decline to add them.

10. Pag 1844, line 1 (About the case 3)

You say: "By enforcing an unrealistic balance of equal highs and lows, the method creates a low-frequency oscillation that does not exist." However I think that should be necessary comment the result of EMD's application to "case 3" comparing this with "case 3 without add the extreme event". I say this because also in "case 3 without add the extreme event" I expect that you will obtain some oscillation that "does not exist" (no prescribed oscillations), and this should be clarified.

I showed the result of the random-only case in Figure 1, but I did not reference it here. I've fixed that in the revision. The addition is shown in bold:

"Because the EMD method implicitly assumes local highs are balanced perfectly by nearby lows, it cannot handle an extreme event like this. By enforcing an unrealistic balance of equal highs and lows, the method creates a low-frequency oscillation that does not exist. **Although the random-only case (Figure 1) also produces low-frequency erroneous oscillations, the amplitudes are significantly less for the longer-period IMFs.** With a larger pulse, the magnitude of the error is even higher."

11. Period IMFs How do you obtain the periods of IMFs? From instantaneous frequency, from values peak-peak or?

I assume this refers to the statement:

"Notice the large, non-stationary oscillation with a period of about 10-years in IMF6."

Since I don't dwell on the period and just use it to describe a rough period, I don't feel it is necessary to describe the details as it would clutter the text. It is based on dividing the length of the time-series by the number of peaks.

Technical Comments

1. Pag 1837, line 7

Before introduce the cases, you should add that you will analyze three cases. After this talk about them.

Two lines above that we state:

“Thus, we propose to test the EMD process not on real observations where one does not know the underlying modes, but on three simulated data sets where the modes are prescribed.”

2. *"Data and methods"*

I suggest to present the three cases in a more schematic way, to give a more immediate vision to the reader ... (You could use the same division in "Results and analysis"). I suggest also to insert the analytic expression for the third case using a Dirac's delta to underline that is only one the point in which you insert the extreme value.

This is more a matter of writing styles than a technical problem. We prefer the text the way it is written.

3. *Pag 1842, line 11*

The sentence:

“In addition, we found in nearly every case (99 %) the EMD computed an IMF with a periodic signal between the ENSO and the PDO signal.”

should be: “In addition, we found in nearly every case (99 %) the EMD computed an IMF with a periodicity between the periodicity of the IMFs designed to describe ENSO and the PDO.”

That sentence has already been revised according to Comment 5 above:

““In addition, we found in nearly every case (99%) the EMD computed one to two IMFs with a periodic signal that did not correlate highly with either PDO or ENSO, but had a low-frequency.”

4. *Figures In figures in which the average periods of IMFs are missed, I suggest to insert them.*

I'm sorry, but I do not understand this request and so do not know how to respond.

5.

Pag 1847, line 10

The title of the follow reference is not correct.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, E. H., Zheng, Q., Tung, C. C., and Liu, H. H.: The Empirical Mode Decomposition and the Hilbert spectrum for non stationary time series analysis, P. R. Soc. London, 454, 903–995, 1998.

Thanks for catching that. One of the pitfalls of typing in reference is missing a word or two in the title.

Response to Reviewer # 2

Comment # 1

=====

However, as I understand it, as a simply mathematical procedure the identified signals do not necessarily match a “real” signal. Like in a classical EOF analysis, it could happen that a “real” signal is actually accounted for by the combination of two or more modes. Is this what is happening when more than one IMF is correlated with a climate index in case 2? If so, my feeling is that the evaluation of the ability of EMD is not fair.

Reply

=====

I agree with the reviewer that the “real” signals are likely spread over multiple modes, but I’m sure the reviewer also knows that many users of these techniques (EOFs and EMDs both) tend to focus on just a single mode and say it’s the ENSO, or PDO mode, for example. That is the motivation for this. I have attempted to make this clearer in the revision by adding more verbiage in the introduction and conclusions. Here is the new text. Updated sentences are in [].

Intro

“However, there are some potential pitfalls that we believe have not been fully addressed in previous papers utilizing the method. First and foremost, EMD is a purely mathematical deconstruction of the data, with no regard to intrinsic covariance of the signals or physics. Second, it assumes that IMFs are comprised of fluctuating signals where the magnitude of nearby peaks and troughs are balanced to create a zero mean – an assumption not based on any physical requirement, as real observations can have quite large ranges in magnitudes, especially sea level data affected by climate signals and synoptic storm events. [Thus, it is unlikely that a single IMF from the EMD analysis can represent a real, physical climate variation. Because of the assumption in the method, it is more likely that multiple modes will be needed to quantify the physical climate mode. However, without some a priori knowledge of this mode, how can one know which IMFs to add together? In the worst case, the climate signal could be spread among a large number of modes. Already, several authors have performed an EMD on ENSO indices and argued they have extracted distinct modes of interannual to multi-decadal variability (Wu and Huang, 2004; Franzke, 2009; Pecaí et al., 2010); their argument based solely on the fact such modes are extracted during the EMD process, but with no physical explanation for them.]”

And in the Conclusions:

“EMD is a quick and relatively easy tool to identify possible multidecadal fluctuations in a sea level record. [However, we have demonstrated that real climatic non-stationary signals are generally spread among multiple modes. Analyzing a single IMF for climate variability will likely lead to significantly biased interpretations. Thus, we feel that EMD

analysis should not be used solely to quantify magnitude and phase of non-stationary climate variations, nor should analysis of climatic signals be based on a single IMF.]”

With these additions to the scope of the paper, we feel our analysis for Case 2 is perfectly justified.

Comment # 2

=====

What I find most important is the significant difference in the acceleration computed from the highest order IMF. In the introduction, the author states that, due to way the last IMF is computed, it is equivalent to a direct quadratic fitting of the time series. However, his results show that the two fittings differ. The author should explain this apparent inconsistency.

Reply

=====

The EMD “acceleration” curve is based on a quadratic fit, but only to the final oscillatory IMF as discussed in the Introduction. I’ve modified it to make this more clear:

“There is also a subtlety in finding the last IMF that is not discussed in the literature. Since the EMD process requires fitting of cubic splines, the last IMF mode that can be calculated has more than one local minima and more than one local maxima, but fewer than four. The only way to get the final IMF shown in most studies, which shows a continuously increasing sea level mode, is to fit a quadratic to the final IMF from the EMD process, and plot the resulting fit. [This is conceptually no different than fitting a quadratic to the original time series, other than the fact it is done to the final mode, which has significant lower variance than the original data. This should] improve the estimate – provided there are no systematic errors or biases in the final IMF that would bias the result.”

I have also revised the conclusions to explain what must be happening to result in different, biased accelerations:

“Finally, authors have asserted that the acceleration that comes out of an EMD process is more accurate, as they believe the IMFs better separate the high- and low-frequency fluctuations than linear least squares. [Their argument assumes that the high-frequency variations and shorter-period non-stationary signals in the original time-series are biasing a quadratic fit to the original data. By eliminating these signals in the EMD process in specific IMFs, they believe the final IMF contains the “true” acceleration plus residual low-frequency variability.] Our experiments, however, show the opposite. The quadratic fit to the last IMF is either no more accurate than one fit with least squares to the full, unfiltered data set, or, in some cases, is significantly biased. In the experiment with ENSO- and PDO-like oscillations, the acceleration estimated from the final IMF was nearly 100% too large on average. In individual experiments, the error was even more.

[This is most likely due to the aliasing behavior of EMD, where some of the high-frequency variance is aliased into the low-frequency modes, as we have demonstrated.]”

Response to Reviewer # 3

Comment # 1

=====

The first comment is related to the title. A major part of this manuscript deals with the identification of an acceleration in sea level. This should be included in the title.

Reply

=====

That's a good suggestion. I've changed the title to:

Evaluation of Empirical Mode Decomposition for Quantifying Multi-Decadal Variations and Acceleration in Sea Level Records

Comment # 2

=====

You should discuss your results with respect to those from figures 3 and 4 in Franzke (2009, <http://journals.ametsoc.org/doi/pdf/10.1175/JCLI-D-11-00293.1>). Franzke already showed, in a different simulation study, that EMD and wavelet methods perform worse compared to classical approaches such as OLS when searching for a known trend. Your results clearly underpin this finding. Franzke, however, argued that this is only the case if the real signal is known. If there is an exponential trend and you fit a linear, he suggests that EMD is the better choice, since the error bars of a linear OLS will increase exponentially. My personal opinion is a bit different to that. I prefer to apply different linear and nonlinear approaches to search for the real signal rather than using one individual model. I think that this issue still needs further independent investigations: of course not here, but you should discuss this point.

Reply

=====

Note: The paper mentioned is from 2011, not 2009.

Thanks for pointing me to that paper. I did comment on another Franzke paper (one from 2009), but that was on using EMD on climate indices. I had not found this second paper in my literature review, probably because it dealt with apply EMD to SST data, not sea level. But after reading it, I see your point. I've added some commentary on that paper in both the Introduction and the Conclusions. It's repeated below:

In the Introduction, after discussing fitting quadratic terms to the highest IMF.

“However, Franzke (2011) conducted an experiment of detecting non-linear trends (i.e., an acceleration) to a small suite of 100 simulated temperature time-series, using different methods including ordinary least squares and EMD. The results showed no statistically significant improvement in EMD. In fact, in most tests, ordinary least squares computed a non-linear trend closer to the input signal.”

In the conclusions, when discussing the recovery of acceleration:

“Finally, authors have asserted that the acceleration that comes out of an EMD process is more accurate, as they believe the IMFs better separate the high- and low-frequency fluctuations than linear least squares. Their argument assumes that the high-frequency variations and shorter-period non-stationary signals in the original time-series are biasing a quadratic fit to the original data. By eliminating these signals in the EMD process in specific IMFs, they believe the final IMF contains the “true” acceleration plus residual low-frequency variability. Even Fanzke (2011), who demonstrated that EMD was no better than this than ordinary least squares and a parametric model argued that EMD was still better if the trend was non-linear, especially exponential. Our experiments, however, show the opposite. The quadratic fit to the last IMF is either no more accurate than one fit with least squares to the full, unfiltered data set, or, in some cases, is significantly biased.”

Comment # 3

=====

You decided to use random noise for the residual signal. However, Dangendorf et al. (2014, <http://onlinelibrary.wiley.com/doi/10.1002/2014GL060538/abstract>) have shown that the residual signal (after accounting for ENSO variations) in San Francisco is long-term correlated. Did you test whether a different choice of residual noise (i.e. long-term correlated noise, for instance simulated with an ARFIMA model) affects your simulations?

Reply

=====

I did test a colored noise model (based on a AR(3) model) in an early experiment, and found the results were not significantly different than using random noise. I chose to use random noise for faster computations and to make it easier to reproduce my results. I make a note of that in the revised manuscript in the introduction.

“We ran another case using a colored noise model that exactly reproduces the autocovariance of the San Francisco tide gauge residuals. The results were nearly identical to the ones shown with the random residuals, so we choose to use random values as they are faster to compute for the several thousand simulations we plan to run.”

Comment # 4

=====

Case 3: You include an extreme event in terms of monthly means. This is not a storm surge in its classical expression, which is defined as a high frequency event with a duration of a few hours or days. Your extreme event is rather comparable to an anomalous ENSO event connected with larger scale ocean dynamics

Reply

=====

True, this is not a storm surge, but the reflection of a storm surge in a monthly average. For example, Tropical Storm Debby in June 2012 caused a storm surge of more than a meter at the tide gauge. This is reflected in the monthly mean for June as the highest June mean value in the record.

I have addressed your point with a slight revision of the sentence:

“Case 3 starts as the baseline model, adds random noise with a standard deviation of 60 mm (representative of the high-frequency variability in San Francisco sea level), then adds an extra 350 mm for January 1956 to represent the signal of a large anomalous high-water event, such as the effect of a large storm surge event on the monthly average, a large flooding event from sustained rainfall, or climatic variations in winds that can cause sustained high water levels.”

Evaluation of Empirical Mode Decomposition for Quantifying Multi-Decadal

Variations and Acceleration in Sea Level Records

Don P. Chambers¹

¹ College of Marine Science, University of South Florida, St. Petersburg, FL

Correspondence to: D. Chambers (donc@usf.edu)

Abstract. The ability of empirical mode decomposition (EMD) to extract multidecadal variability from sea level records is tested using three simulations: one based on a series of purely sinusoidal modes, one based on scaled climate indices of El Niño and the Pacific Decadal Oscillation (PDO), and the final one including a single month with an extreme sea level event. All simulations include random noise of similar variance to high-frequency variability in the San Francisco tide gauge record. The intrinsic mode functions (IMFs) computed using EMD were compared to the prescribed oscillations. In all cases, the longest-period modes are significantly distorted, with incorrect amplitudes and phases. This affects the estimated acceleration computed from the longest periodic IMF. In these simulations, the acceleration was underestimated in the case with purely sinusoidal modes, and overestimated by nearly 100% in the case with prescribed climate modes. Additionally, in all cases, extra low-frequency modes uncorrelated with the prescribed variability are found. These experiments suggest that using EMD to identify multidecadal variability and accelerations in sea level records should be used with caution.

1. INTRODUCTION

Over the last decade, several papers have used the method of empirical mode decomposition (EMD) (Huang et al., 1998; Huang and Wu, 2008) to evaluate non-stationary patterns in time series as disparate as electromyographic signals (Andrade et al., 2006) and sea level (Breaker and Ruzmaikin, 2011; Ezer and Corlett, 2012; Ezer et al., 2013; Lee, 2013; Chen et al., 2014). The use of EMD in sea level records has been motivated in large part by numerous papers discussing the appearance of decadal and longer period fluctuations in tide gauge records and global mean sea level estimates based on tide gauge records (e.g., Feng et al., 2004; Miller and Douglas, 2007; Woodworth et al., 2009; Bromirski et al., 2011; Sturges and Douglas, 2011; Chambers et al., 2012; Calafat and Chambers, 2013; Becker et al., 2014; Dangerdorf et al., 2014).

At first glance, EMD appears to be a useful tool to find non-stationary, low-frequency fluctuations in sea level, as it breaks the time-series into a set of Intrinsic Mode Functions (IMFs) that have progressively longer quasi-period fluctuations. IMFs extracted from various tide gauge records have been shown to be correlated with several climate indices (e.g., Ezer and Corlett, 2012; Ezer et al., 2013), which gives some credence to extracted signals. Moreover, authors have argued that the final IMF, representing the continuously increasing sea level mode, is a better representation of an acceleration than simply fitting a quadratic to the original data (Huang and Wu, 2008; Ezer and Corlett, 2012; Ezer et al., 2013).

However, there are some potential pitfalls that we believe have not been fully addressed in previous papers utilizing the method. First and foremost, EMD is a purely mathematical deconstruction of the data, with no regard to intrinsic covariance of the signals or physics. Second, it assumes that IMFs are comprised of fluctuating signals where the magnitude of nearby peaks and troughs are balanced to create a zero mean – an assumption not based on any

physical requirement, as real observations can have quite large ranges in magnitudes, especially sea level data affected by climate signals and synoptic storm events. Thus, it is unlikely that a single IMF from the EMD analysis can represent a real, physical climate variation. Because of the assumption in the method, it is more likely that multiple modes will be needed to quantify the physical climate mode. However, without some a priori knowledge of this mode, how can one know which IMFs to add together? In the worst case, the climate signal could be spread among a large number of modes. Already, several authors have performed an EMD on ENSO indices and argued they have extracted distinct modes of interannual to multi-decadal variability (Wu and Huang, 2004; Franzke, 2009; Pecai et al., 2010); their argument based solely on the fact such modes are extracted during the EMD process, but with no physical explanation for them.

The basic idea of EMD is to fit cubic splines to the local maxima and minima of a time-series separately, average the splines, then remove the average from the time-series. The process is iterated on the residual time-series until the average of the splines converges to have a standard deviation less than some pre-set tolerance. This is the first IMF. This is then subtracted from the original time-series and the process is repeated until only one minimum and one maximum remain. For details of the procedure, readers are referred to the original paper by Huang et al. (1998) or more recent applications (e.g., Huang and Wu, 2008; Ezer et al., 2013).

There is also a subtlety in finding the last IMF that is not discussed in the literature. Since the EMD process requires fitting of cubic splines, the last IMF mode that can be calculated has more than one local minima and more than one local maxima, but fewer than four. The only way to get the final IMF shown in most studies, which shows a continuously increasing sea level mode, is to fit a quadratic to the final IMF from the EMD process, and plot the resulting fit. This is conceptually no different than fitting a quadratic to the original time series, other than the fact it

is done to the final mode, which has significant lower variance, than the original data. This should improve the estimate – provided there are no systematic errors or biases in the final IMF that would bias the result.

However, Franzke (2011) conducted an experiment of detecting non-linear trends (i.e., an acceleration) to a small suite of 100 simulated temperature time-series, using different methods including ordinary least squares and EMD. The results showed no statistically significant improvement in EMD. In fact, in most tests, ordinary least squares computed a non-linear trend closer to the input signal.

Moreover, Wu and Huang (2004) have shown that EMD behaves as a low-pass filter. If one runs random noise with a normal Gaussian distribution through the process, low-frequency signals will be seen in the resulting IMFs. They found there is roughly a doubling of the average period with each subsequent mode. This is a significant issue. It means that any random (or near-random, high-frequency) signal will propagate into low-frequency signals in the recovered IMFs. Wu and Huang (2009) proposed a method to quantify the uncertainty caused by this behavior by computing an ensemble mean of IMFs, starting from the same time-series but with different amounts of added random noise.

However, this method ignores that most geophysical time-series have an underlying real signal that has high variance and little serial correlation; i.e., a high-frequency, near-random signal. This “signal” will also be filtered by the EMD process and will likely appear as a quasi-stationary oscillation in higher order IMFs that is not real. Although adding multiple realizations of random noise to a time-series will account for uncertainty in the IMFs from random error in the measurement, it will not account for the shifting of high-frequency signal to low-frequency

Don Chambers 1/16/15 10:37 AM

Deleted: although the variance of the final

Don Chambers 1/16/15 10:37 AM

Deleted: EMD mode is far lower

Don Chambers 1/16/15 10:38 AM

Deleted: , which

Don Chambers 1/16/15 10:38 AM

Deleted: should

signal in the recovered IMFs. One of the assumptions in EMD processing is this is captured in the lowest IMFs. Our testing indicates it is not.

To demonstrate the potential size of this problem, we ran EMD on a monthly-resolution time-series that is 150-years long with randomly-correlated values that have a standard deviation of 60 mm. We used 60 mm because this is the standard deviation of residual monthly sea level at San Francisco after fitting and removing a quadratic function plus annual and semiannual sinusoids, so is representative of high-frequency sea level at a typical site, although some sites can have significantly higher variability. We ran another case using a colored noise model that exactly reproduces the autocovariance of the San Francisco tide gauge residuals. The results were nearly identical to the ones shown with the random residuals, so we choose to use random values as they are faster to compute for the several thousand simulations we plan to run. The EMD finds IMFs that have quasi-period fluctuations of nearly 60-years and amplitudes as large as 10 mm (Figure 1); fluctuations at quasi-30-year periods are the same magnitude.

Thus, we propose to test the EMD process not on real observations where one does not know the underlying modes, but on three simulated data sets where the modes are prescribed. We believe the differences between the recovered IMFs and given signals will be a better measure of the accuracy of the EMD method than what has previously been discussed in the literature. Two different simulations will be examined with fluctuating signals and differing random noise to represent high-frequency variability: one using purely sinusoidal oscillations with multiple periods ranging from 13-years to 80-years, the second with variations based on band-pass filtered and scaled El Niño-Southern Oscillation (ENSO) and Pacific Decadal Oscillation (PDO) indices, both with additional random noise applied. The third case will examine a situation with only seasonal fluctuations, random noise, and a single month with a variation larger than 3-standard

Don Chambers 1/15/15 11:30 AM

Moved (insertion) [2]

Don Chambers 1/15/15 2:12 PM

Deleted: Moreover, Wu and Huang (2004) have shown that EMD behaves as a low-pass filter. If one runs random noise with a normal Gaussian distribution through the process, low-frequency signals will be seen in the resulting IMFs. They found there is roughly a doubling of the average period with each subsequent mode. attempted to account for this behavior in the error statistics of their recovered IMFs. This is done by creating several pseudo-time-series by adding a small amount of random noise to the original time-series, running a large number of EMDs, and considering the average. The standard deviation of the differences represents the uncertainty. Hit does not consider the effect of the real signal in the tide gauge data that has high variance and little serial correlation. This "signal" will also be filtered by the EMD process and will likely appear as a quasi-stationary oscillation in higher order IMFs that is not real (i.e., similar to the EMD filtering of purely random noise shown in Figure 1). ... [1]

Don Chambers 1/15/15 11:32 AM

Moved (insertion) [3]

Don Chambers 1/15/15 11:24 AM

Deleted: noise that has

Don Chambers 1/15/15 11:22 AM

Moved (insertion) [1]

Don Chambers 1/15/15 11:22 AM

Deleted: Note that we

Don Chambers 1/15/15 11:22 AM

Deleted: ,

Don Chambers 1/15/15 11:25 AM

Deleted: one

Don Chambers 1/15/15 11:25 AM

Deleted: clearly observes

Don Chambers 1/15/15 11:22 AM

Moved up [1]: Note that we used 60 mm because this is the standard deviation of residual monthly sea level at San Franc ... [2]

Don Chambers 1/15/15 11:36 AM

Deleted: Wu and Huang (2004) and others (Ezer and Corlett, 2012; Ezer et al., 2013) have attempted to account for this behavior i ... [3]

Don Chambers 1/15/15 11:30 AM

Moved up [2]: attempted to account for this error statistics of their recovered IMFs. This is done by creating several pseudo-time-s ... [4]

Don Chambers 1/15/15 11:32 AM

Moved up [3]: it does not consider the effect of the real signal in the tide gauge data that has high variance and little serial ... [5]

deviations. This represents an extreme event, typically caused by major storm surge, which is a common feature in many sea level records. We will demonstrate that the EMD method leads to spurious IMFs in all cases, and where the IMFs are correlated with the input signal, their amplitudes and phases are significantly biased in many periods of the record.

2. DATA AND METHODS

We use the monthly sea level record from the San Francisco tide gauge for our reference sea level record. It was downloaded from the Permanent Service for Mean Sea Level (PSMSL) (Woodworth and Player, 2003; PSMSL, 2012). We fit annual and semi-annual sinusoids along with a trend and an acceleration term to the data using ordinary least squares to obtain the baseline sea level variability for our model (y_{base}), where

$$y_{base}(t) = -78.3 + 0.92*dt + 0.0081*dt**2. + 4.2*cos(2\pi*dt) - 31.8*sin(2\pi*dt) + 20.3*cos(4\pi*dt) + 17.7*sin(4\pi*dt), \quad (1)$$

and t is the time in years, with $dt = t - 1900.0$. This baseline model is the same for all experiments.

For Case 1, we add three long-period sinusoids (13-years, 55-years, and 80-years) to the baseline model along with random noise ($\epsilon(t)$):

$$y_{case1}(t) = y_{base}(t) - 9.8*cos((2\pi/13)*dt) + 12.5*sin((2\pi/13)*dt) - 6.3*cos((2\pi/55)*dt) + 12.3*sin((2\pi/55)*dt) + 9.6*cos((2\pi/80)*dt) - 15.2*sin((2\pi/80)*dt) + \epsilon(t). \quad (2)$$

The random noise has a variance to match the variance of the residuals of the real tide gauge data minus the model. 1000 different random noise models were applied to create 1000 different versions of Case 1 for statistical testing. The periods and amplitudes of the long-period sinusoids

were chosen arbitrarily to approximate the level of multidecadal fluctuations in the San Francisco sea level record (Figure 2a).

Case 2 starts from the same baseline model, but instead of prescribing sinusoidal oscillations, we use non-stationary climate indices for El Niño-Southern Oscillation (ENSO) variations and the Pacific Decadal Oscillation (PDO). We use the Southern Oscillation Index based on the pressure differences between Tahiti and Darwin, Australia to represent ENSO variability (Trenberth, 1984; Ropelewski and Jones, 1987; downloaded from <http://www.cgd.ucar.edu/cas/catalog/climind/soi.html> on 5 March 2014), and the PDO index based on the leading principal component of sea surface temperature in the North Pacific (Zhang et al., 1997; Mantua et al., 1997; downloaded from <http://jisao.washington.edu/pdo/PDO.latest> on 5 March 2014).

Several additional processing steps are required before using these indices for our experiment. First, neither index covers the same period as the tide gauge (January 1856 to December 2010). The SOI index starts in January 1866 while the PDO index begins in January 1900. In order to have a simulated record as long as possible, we start in January 1866 and use values from the end of PDO record to fill in the missing data before January 1900. Recall we are not worried about the “true” ENSO or PDO variability, only a simulation of the type of variability and how well EMD can recover it.

Secondly, because the two indices are slightly correlated (-0.21 , $p < 0.001$) due to similar interannual (< 10 year) variations, we low-pass the PDO index with a 5-year Gaussian, and band-pass filter the SOI by first removing the 5-year Gaussian of the SOI, and then filtering the residuals with a 0.5-year Gaussian. After doing this, the correlation between the two filtered indices $PDO_{LP}(t)$ and $SOI_{BP}(t)$ is insignificant (-0.003 , $p < 0.01$).

The final step is to determine the scaling factor to apply to both the $PDO_{LP}(t)$ and $SOI_{BP}(t)$ variations. This is done by first normalizing both time-series by their standard deviation. Then, after removing the estimated trend, acceleration and seasonal variations, the sea level data are low-pass and band-pass filtered as the climate indices were, and the standard deviation of the filtered residuals is calculated. The scaling factor applied to $PDO_{LP}(t)$ is the standard deviation of the low-pass filtered sea level residuals (20.8 mm); the scaling factor applied to $SOI_{BP}(t)$ is the standard deviation of the band-pass filtered sea level residuals scaled by -1 to account for the fact El Niño sea level variations at San Francisco are positive when the SOI is negative (-28.7 mm).

The final time-series for Case 2 is assembled as for Case 1, including the random noise term based on the standard deviation of the residuals and the model:

$$y_{case2}(t) = y_{base}(t) - 28.7 * SOI_{BP}(t) + 20.8 * PDO_{LP}(t) + \epsilon(t), \quad (3)$$

noting $PDO_{LP}(t)$ and $SOI_{BP}(t)$ are normalized as described previously. One time-series is shown in Figure 2b to show the model does a reasonable job of simulating the San Francisco tide gauge record.

Case 3 starts as the baseline model, adds random noise with a standard deviation of 60 mm (representative of the high-frequency variability in San Francisco sea level), then adds an extra 350 mm for January 1956 to represent the signal of a large anomalous high-water event, such as the effect of a large storm surge event on the monthly average, a large flooding event from sustained rainfall, or climatic variations in winds that can cause sustained high water levels. Such a value is possible in sea level records, depending on the size and duration of the storm (e.g., the maximum deviation of monthly sea level residuals after removing a trend for the San Francisco tide gauge record is 4.9 times higher than the standard deviation). Moreover, most tide gauge records have numerous events instead of just one; San Francisco has six monthly residuals

exceeding 200 mm and two exceeding 300 mm. For this study, however, we consider just one to demonstrate the effect on EMD if authors do not consider this possibility in their analysis.

We perform EMD using the EMD Toolkit for SciLab (<http://www.scilab.org>), based on code documented in Rilling et al. (2003). The specific function utilized was *emdc*, which stops iterating when a tolerance is reached. We used a tolerance value of 0.05.

3. RESULTS AND ANALYSIS

Figure 3 shows the low-frequency IMFs for a single simulation of Case 1, along with the input signals. IMFs 1-5 are all much higher frequency and so are not considered, although we will note that none accurately captures the input seasonal variation. However, we point out that some of the artifacts shown in Figure 3 for the low-frequency IMFs are a direct result of correcting for errors in the higher frequency IMFs not shown, so that the sum of all matches the original data.

The correlation of IMF6 with the prescribed 13-year sinusoid is significant (> 0.5), but not high. It is clear there are several periods where the EMD method would suggest no variability at a 13-year period (1870-1890, 1950-1970) and other periods (~ 1910) where the variation is significantly faster. Moreover, the amplitude of the recovered IMF is steadily increasing after 1980, although the phase is about correct. The next IMF is an artifact of the method, with no significant correlation with any input signal, yet showing a periodicity of ~ 20 -years with amplitudes as high as 20 mm.

The longer period IMFs also have problems (Figure 3). The one best correlated with the 55-year sinusoid (IMF8) is out of phase with the real signal until about 1940, and the amplitude is increasing in time. The 80-year IMF exhibits a similar behavior of increasing amplitude (Figure

3). Finally, the estimated quadratic term to the longest oscillatory IMF (IMF9 in this case), significantly underestimates the prescribed acceleration.

This is just a single example, however. To see if averaging of multiple versions of random noise will better match the prescribed signals, we examine the 1000 different IMF clusters from the simulations. One cannot simply separate the corresponding low-frequency modes based on the IMF number, however, as the total number of IMFs changed from 9 to 11 in the 1000 different simulations. We found that the 13-year signal was found in IMFs ranging from number 4 to 7, while the 55-year mode ranged from IMF7-9. The last mode found ranged from IMF9-11. Thus, we had to rely on correlation with the known oscillation to identify the relevant IMF. This was done by computing the correlation of each recovered IMF from each simulation with the prescribed sinusoids. Figure 4 shows the histogram of computed correlations. Note that the correlations were not the same in every simulation. The 13-year oscillation had a mean correlation of 0.66 (standard deviation = 0.09), the 55-year had a mean of 0.52 (standard deviation = 0.11), while the mean correlation of the 80-year signal was 0.74 (standard deviation = 0.09).

So that the lower correlation in the 55-year test did not bias our results, a minimum bound was set to 0.5. If two or more modes had correlations > 0.5 with one of the input signals, the one with the highest correlation was chosen. Figure 5 summarizes the results, showing the mean IMF with the standard deviation as a shaded uncertainty band. Not every simulation found an IMF that had a correlation > 0.5 with all the prescribed sinusoids. The 13-year oscillation had 848 matches, the 55-year only 550, and the 80-year 990. It appears that the extra mode or two that often pops up between 13-years and 55-years in the EMD distorts the recovery of the 55-year signal (e.g., Figure 3).

Don Chambers 1/15/15 1:40 PM

Deleted: A

Don Chambers 1/15/15 1:41 PM

Deleted: -

Don Chambers 1/15/15 1:41 PM

Deleted: 4

Don Chambers 1/15/15 1:41 PM

Deleted: 890

Although the phase of the mean 13-year IMF is consistent with the prescribed signal, the mean amplitude is too small (Figure 5). The standard deviation is also quite high relative to the amplitude (80%), suggesting the actual recovered IMF could be nearly zero for any realization, or two times too large, depending on how the high-frequency variability affects it.

For the longer-period oscillations, there is a systematic error in the mean IMF. It is roughly the same in both the 55-year and 80-year signal: the phase is only correct at the end of the record, and the amplitude is unrealistically increasing in time (Figure 5), from almost no fluctuation at the beginning to larger variations than were prescribed at the end. The scatter is again relatively large compared to the largest amplitude (60-80%). Finally, the acceleration estimated from the final IMF mode is systematically too small (Figure 5).

We acknowledge this test has its limitations. The final peaks of the 55-year 80-year sinusoids are very close to each other. However, a rather simplistic harmonic analysis using least squares over ranges of given periods found all three sinusoids precisely with small errors (< 5 mm). If EMD can purportedly uncover non-stationary oscillations in a data set accurately, then it should be able to compute perfectly stationary ones just as well. The level that EMD cannot do this is a reasonable estimate of its accuracy.

More importantly, consider the interpretation of the results from this simplistic simulation in terms of longer-term climate change if only the EMD results (Figure 5) were analyzed. Based only on the returned IMFs, one could easily argue that there was no significant low-frequency variation in the sea level before 1950, then a rather dramatic rise in the 1970s, followed by a return to normal condition. In fact, there were equally large sea level shifts in the early part of the simulated record that were lost due to the way the EMD method partitions the real signal.

Don Chambers 1/15/15 1:45 PM

Deleted: 4

Don Chambers 1/15/15 1:45 PM

Deleted: 4

Don Chambers 1/15/15 1:47 PM

Deleted: 4

Figure 6 summarizes the results of Case 2, the simulation based on the ENSO and PDO indices. As with the experiment in Case 1, 1000 different simulations were run, differing only by the random noise. The IMF with the highest correlation greater than 0.5 with both the prescribed ENSO and PDO index was averaged. In addition, we found in nearly every case (99%) the EMD computed one to two IMFs with a periodic signal that did not correlate highly with either PDO or ENSO, but had a low-frequency. Because this was not always contained in a single IMF between the two prescribed periodic fluctuations, we had to adapt a method to search for it. We isolated this signal by looking at the autocorrelation of the IMFs after removing those correlated with PDO or ENSO, as well as the last mode. To find the mode with the longest-period fluctuation, we examined the autocorrelation at a 1-year lag. Only IMFs with an autocorrelation greater than 0.9 at a lag of 1-year were examined, and if two existed, the one with the higher autocorrelation was selected.

We should note that typically there were several IMFs that correlated significantly with the ENSO index. For the statistics shown in Figure 6, only the one with the highest correlation was chosen. Although we found that by adding the 1-2 additional IMFs to the most significant ENSO mode resulted in a better correlation, we felt this was not a fair evaluation of the EMD process. ENSO is a physical process, and the relationship between the climate indices and the physics of the strength and timing of an ENSO event related to the index has been well established, (e.g., Philander, 1990; 2006). Although some authors have run EMD on ENSO indices and argued they have extracted distinct modes of interannual to multi-decadal variability (Wu and Huang, 2004; Franzke, 2009; Pecaí et al., 2010), their conclusions are based solely on the fact such modes are extracted during the EMD process; they have offered no physical explanation for them. We note that other statistical based methods (such as principal component analysis) run on

Don Chambers 1/15/15 3:10 PM
Deleted: 5

Don Chambers 1/15/15 3:32 PM
Deleted: In addition, we found in nearly every case (99%) the EMD computed an IMF with a periodic signal between the ENSO and the PDO signal.

Don Chambers 1/15/15 1:54 PM
Deleted: remaining

Don Chambers 1/15/15 1:54 PM
Deleted: uncorrelated with either

Don Chambers 1/15/15 1:55 PM
Deleted: The

Don Chambers 1/15/15 1:56 PM
Deleted: was selected.

Don Chambers 1/15/15 3:10 PM
Deleted: 5

Don Chambers 1/15/15 3:10 PM
Formatted: Font:Not Italic

environmental data like sea surface temperature, precipitation, sea level, winds, etc. find modes highly correlated with ENSO and PDO indices (e.g., Mantua et al., 1997; Wolter and Timlin, 1998; Chambers et al., 1999; Bond et al., 2003). We know of none that find multiple modes that add up to correlate with an ENSO index. Thus, we argue it is more relevant to quantify if EMD can extract physically meaningful climate modes than whether it can extract modes with interannual and multi-decadal variability.

The ENSO-mode IMF on average matches the timing of the input ENSO variability (Figure 6), although the amplitude is smaller; on average it underestimates the size of the El Niño and La Niña events by a factor of 2 to 3. Moreover, the standard deviation is large, ranging from 50% to 250% of the estimated peak values. This means that no single decomposition exactly matches the simulated ENSO variability. Some may catch a peak or two properly, but other El Niño or La Niña events are not captured at all.

The non-simulated low-frequency IMF has a period of between 25-30 years (Figure 5), with an average amplitude ranging from 10 mm to 20 mm. This is the same magnitude of variability as the PDO-related variability, although IMFs extracted from a single simulation could have an amplitude nearly 3 to 4 times higher, based on the standard deviation. Without knowing *a priori* what variations were in the data, this mode would be interpreted as a real, physical oscillation in sea level, when in fact it is a bogus artifact of the analysis.

As with the ENSO-mode, the mean PDO-mode IMF tracks the general periodicity of the PDO, although the amplitudes are on average too small. Again, the standard deviation suggests any single simulation would give a considerable range of amplitudes. We note that as with the Case 1 results, there is a tendency for an increasing amplitude in time for the mean IMF,

Don Chambers 1/15/15 3:10 PM

Deleted: 5

inconsistent with the true signal, which could be misinterpreted as a sign of climate change; the first two peaks in the given PDO signal are roughly the same magnitude.

Finally, the average long-term rise computed from the last IMF is wrong (Figure 6). The trend at 1900 is 36% lower than prescribed, and the overall acceleration is 83% higher.

Figure 7 shows the results from the EMD of Case 3, with the single extreme event. Notice the large, non-stationary oscillation with a period of about 10-years in IMF6. The amplitude reaches 25 mm around 1956. Recall that this experiment only included seasonal variations, random noise, and this single large event. Because the EMD method implicitly assumes local highs are balanced perfectly by nearby lows, it cannot handle an extreme event like this. By enforcing an unrealistic balance of equal highs and lows, the method creates a low-frequency oscillation that does not exist. Although the random-only case (Figure 1) also produces low-frequency erroneous oscillations, the amplitudes are significantly less for the longer-period IMFs. With a larger pulse, the magnitude of the error is even higher. It does not affect just this mode. It also shows up in IMF7 and IMF8, especially distorting the end of the record (Figure 7). We have not tested by adding more extreme events, but we would assume it would cause even more spurious signals like these.

4. CONCLUSIONS

While at first glance empirical mode decomposition appears to be a useful tool for quantifying non-stationary multidecadal oscillations in sea level records, the results of our experiments suggest there are several issues. Probably the biggest one is the fact the EMD process applied to random noise consistent with high-frequency sea level variability and single extreme events will cause relatively large and systematic multidecadal oscillations that are not

Don Chambers 1/15/15 3:19 PM
Deleted: 5

Don Chambers 1/15/15 3:19 PM
Deleted: 6

Don Chambers 1/15/15 3:21 PM
Deleted: 6

real. This will distort any underlying true signal. Our results suggest this is especially a problem for the longest period fluctuations; the IMFs are systematically biased away from the true signal, both in amplitude and phase. In some cases the amplitude increases in time, which could lead to incorrect interpretations regarding acceleration.

Moreover, there always appears to be one or more IMFs that are completely spurious fluctuations. These are needed to correct the errors in the other IMFs so they all sum to the original data. With no knowledge of the underlying physical modes, how is one to know which of the signals is spurious? In the articles that have applied EMD to sea level, all long-period IMFs have been assumed real and analyzed in regards to climatic or dynamical fluctuations in sea level. Based on the results of our experiments, we cannot believe that all the analyzed modes are true.

Finally, authors have asserted that the acceleration that comes out of an EMD process is more accurate, as they believe the IMFs better separate the high- and low-frequency fluctuations than linear least squares. Their argument assumes that the high-frequency variations and shorter-period non-stationary signals in the original time-series are biasing a quadratic fit to the original data. By eliminating these signals in the EMD process in specific IMFs, they believe the final IMF contains the “true” acceleration plus residual low-frequency variability. Even Fanzke (2011), who demonstrated that EMD was no better than this than ordinary least squares and a parametric model, argued that EMD was still better if the trend was non-linear, especially exponential. Our experiments, however, show the opposite. The quadratic fit to the last IMF is either no more accurate than one fit with least squares to the full, unfiltered data set, or, in some cases, is significantly biased. In the experiment with ENSO- and PDO-like oscillations, the acceleration estimated from the final IMF was nearly 100% too large on average. In individual

Don Chambers 1/16/15 10:44 AM

Deleted: However, our

experiments, the error was even more. This is most likely due to the aliasing behavior of EMD where some of the high-frequency variance is aliased into the low-frequency modes, as we have demonstrated.

EMD is a quick and relatively easy tool to identify possible multidecadal fluctuations in a sea level record. However, we have demonstrated that real climatic non-stationary signals are generally spread among multiple modes. Analyzing a single IMF for climate variability will likely lead to significantly biased interpretations. Thus, we feel that EMD analysis should not be used solely to quantify magnitude and phase of non-stationary climate variations, nor should analysis of climatic signals be based on a single IMF. One should also be cautious in interpreting acceleration computed from the final IMF, especially in light of the significant errors found in the early and later parts of the low-frequency IMFs (Figures 5, 6, and 7). Where EMD has shown to be useful has been in low-pass filtering data to reduce the impact of high-frequency variability and noise (e.g., Alberti et al., 2014). In that case, the sum of the higher IMFs are used as the low-pass filter.

Instead, we believe other more traditional methods, such as harmonic analysis (e.g., Chambers et al., 2012), linear regression against climatic indices or physical parameters (e.g., Calafat and Chambers, 2013), running means of linear trends evaluated over discrete window-lengths (e.g., Holgate, 2007; Merrifield et al., 2009), or simply low-pass filtering on different time-scales should also be utilized along with EMD to study low-frequency climatic variability. This is in order to find possible spurious signals in the IMFs arising from the way the EMD process filters random noise and extreme events. At the very least, authors should carefully remove extreme events from the sea level records before performing EMD to reduce biasing

Don Chambers 1/16/15 10:31 AM

Deleted: However,

Don Chambers 1/16/15 10:31 AM

Deleted: it

Don Chambers 1/15/15 3:03 PM

Deleted: and one

Don Chambers 1/16/15 9:37 AM

Deleted: 4

Don Chambers 1/16/15 9:37 AM

Deleted: 5

Don Chambers 1/16/15 9:37 AM

Deleted: 6

low-frequency IMFs. Unless other methods are utilized and shown to agree with the EMD results, we remain skeptical of many interpretations of EMD processed sea level data.

Acknowledgements

The San Francisco tide gauge data are furnished by the Permanent Service for Mean Sea Level (<http://www.psmsl.org/>). We thank J. Bonin and T. Wahl for helpful comments on an early draft of this paper. This research was carried out in part under a grant from the NASA Interdisciplinary Science Program.

REFERENCES

- Alberti, T., Lepreti, F., Vecchio, A., Bevacqua, E., Capparelli, V., and Carbone, V.: Natural periodicities and Northern-Hemisphere-Southern Hemisphere connection of fast temperature changes during the last glacial period: EPICA and NGRIP revisited, *Clim. Past*, 10, 1751-1762, doi:10.5194/cp-10-1751-2014, 2014.
- Andrade, A. O., Nasuto, S., Kyberd, P., Sweeney-Reed, C. M., and Van Kanijn, F. R.: Biomedical Signal Processing and Control, 1, 44-55, doi:10.1016/j.bspc.2006.03.003, 2006.
- Becker, M., Karpytchev, M., and Lennartz-Sassinek, S.: Long-term sea level trends: Natural or anthropogenic?, *Geophys. Res. Lett.*, 41, 5571–5580, doi:10.1002/2014GL061027, 2014.
- Bond, N. A., Overland, J. E., Spillane, M., and Stabeno, P.: Recent shifts in the state of the North Pacific, *Geophys. Res. Lett.*, 30, 2183, doi:10.1029/2003GL018597, 2003.
- Breaker, L. C., and Ruzmaikin, A.: The 154-year record of sea level at San Francisco: extracting the long-term trend, recent changes, and other tidbits, *Clim. Dyn.*, 36, 545-559, doi:10.1007/s00382-010-0865-4, 2011.
- Bromirski, P. D., Miller, A. J., Flick, R. E., and Auad, G.: Dynamical suppression of sea level rise along the Pacific coast of North America: Indications for imminent acceleration, *J. Geophys. Res.*, 116, C07005, doi:10.1029/2010JC006759, 2011.
- Calafat, F. M., and Chambers, D. P.: Quantifying recent acceleration in sea level unrelated to internal climate variability, *Geophys. Res. Lett.*, 40, doi:10.1002/grl.50731, 2013.
- Chambers D. P., Tapley, B. D., and Stewart, R. H.: Anomalous warming in the Indian Ocean coincident with El Niño, *J. Geophys. Res.*, 104, 3035-3047, 1999.
- Chambers, D. P., Merrifield, M. A., and Nerem, R. S.: Is there a 60-year oscillation in global mean sea level?, *Geophys. Res. Lett.*, 39, L18607, doi:10.1029/2012GL052885, 2012.
- Chen, X., Feng, Y., and Huang, N. E.: Global sea level trend during 1993-2012, *Glob. Plan. Change*, 112, 26-32, doi:10.1016/j.gloplacha.2013.11.001, 2014.
- Chen, X., Zhang, Y., Zhang, M., Feng, Y., Wu, Z., Qiao, F., and Huang, N. E.: Intercomparison between observed and simulated variability in global ocean heat content using empirical mode decomposition. Part I: Modulated annual cycle, *Clim. Dyn.*, doi:10.1007/s00382-012-

1554-2, 2012.

Dangendorf, S., Rybski, D., Mudersbach, C., Müller, A., Kaufmann, E., Zorita, E., and Jensen, J.: Evidence for long-term memory in sea level, *Geophys. Res. Lett.*, 41, 5564–5571, doi:[10.1002/2014GL060538](https://doi.org/10.1002/2014GL060538), 2014.

Ezer, T., and Corlett, W. B.: Is sea level rise accelerating in the Chesapeake Bay? A demonstration of a novel new approach for analyzing sea level data, *Geophys. Res. Lett.*, 39, L19605, doi:[10.1029/2012GL053435](https://doi.org/10.1029/2012GL053435), 2012.

Ezer, T., Atkinson, L. P., Corlett, W. B., and Blanco, J. L.: Gulf Stream's induced sea level rise and variability along the U.S. mid-Atlantic coast, *J. Geophys. Res. Oceans*, 118, 685–697, doi:[10.1002/jgrc.20091](https://doi.org/10.1002/jgrc.20091), 2013.

Feng, M., Li, Y., and Meyers, G.: Multidecadal variations of Fremantle sea level: Footprint of climate variability in the tropical Pacific, *Geophys. Res. Lett.*, 31, L16302, doi:[10.1029/2004GL019947](https://doi.org/10.1029/2004GL019947), 2004.

Franzke, C.: Multi-scale analysis of teleconnection indices: climate noise and nonlinear trend analysis, *Nonlin. Processes Geophys.*, 16, 65–76, 2009.

[Franzke, C.: Nonlinear trends, long-range dependence, and climate noise properties on surface temperature, *J. Clim.*, 25, 4172–4183, doi: 10.1175/JCLI-D-11-00293.1, 2011.](#)

Holgate, S. J.: On the decadal rates of sea level change during the twentieth century, *Geophys. Res. Lett.*, 34, L01602, doi:[10.1029/2006GL028492](https://doi.org/10.1029/2006GL028492), 2007.

Huang, N. E., and Wu, Z.: A review on Hilbert-Huang transform: the method and its applications on geophysical studies, *Rev. Geophys.*, 46, RG2006, doi:[10.1029/2007RG000228](https://doi.org/10.1029/2007RG000228), 2008.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, E. H., Zheng, Q., Tung, C. C., and Liu, H. H.: The Empirical Mode Decomposition and the Hilbert spectrum for [nonlinear and non-stationary](#) time series analysis, *Proc. R. Soc. London.*, 454, 903–995, 1998.

Lee, H. S.: Estimation of extreme sea levels along the Bangladesh coast due to storm surge and sea level rise using EEMD and EVA, *J. Geophys. Res. Oceans*, 118, 4273–4285, doi:[10.1002/jgrc.20310](https://doi.org/10.1002/jgrc.20310), 2013.

Mantua, N. J., Hare, S. R., Zhang, Y., Wallace, J. M., and Francis, R. C.: A Pacific interdecadal climate oscillation with impacts on salmon production, *Bulletin of the American Meteorological Society*, 78, 1069–1079, 1997.

Merrifield, M. A., Merrifield, S. T., and Mitchum, G. T.: An anomalous recent acceleration of global sea level rise. *Journal of Climate* 22:5772–5781, 2009.

Miller, L., and Douglas, B. C.: Gyre-scale atmospheric pressure variations and their relation to 19th and 20th century sea level rise, *Geophys. Res. Lett.*, 34, L16602, doi:[10.1029/2007GL030862](https://doi.org/10.1029/2007GL030862), 2007.

Molla, M. K. I., Rahman, M. S., Sumi, A., and Banik, P.: Empirical mode decomposition analysis of climate changes with special reference to rainfall data, *Discrete Dynamics in Nature and Society*, 2006, 1–17, doi:[10.1155/DDNS/206/45348](https://doi.org/10.1155/DDNS/206/45348), 2006.

Philander, S. G.: *El Niño, La Niña and the Southern Oscillation*. Academic Press, San Diego, 293 pp, 1990.

Philander, S. G.: *Our Affair With El Nino: How We Transformed an Enchanting Peruvian Current Into a Global Climate Hazard*, Princeton University Press, Princeton, NJ, 275 pp., 2006.

Rilling, G., Flandrin, P., and Gonçalves, P.: On empirical mode decomposition and its algorithms. *IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing NSIP-03*, Grado, Italy, 2003.

Don Chambers 1/15/15 3:40 PM

Deleted:

- Ropelewski, C.F., and Jones, P.D.: An Extension of the Tahiti-Darwin Southern Oscillation Index, *Monthly Weather Review*, 115, 2161-2165, 1987.
- Sturges, W., and Douglas, B. C.: Wind effects on estimates of sea level rise, *J. Geophys. Res.*, 116, C06008, doi:10.1029/2010JC006492, 2011
- Trenberth, K. E.: Signal versus Noise in the Southern Oscillation, *Monthly Weather Review*, 112, 326-332, 1984.
- Yang, P. C., Wang, C. L., Bian, J. C., and Zhou, X. J.: The prediction of non-stationary climate series based on empirical mode decomposition, *Adv Atmos Sci*, 27, 845–854, doi:10.1007/s00376-009-9128-x, 2010.
- Wolter, K., Timlin M.S.: Measuring the strength of ENSO events - how does 1997/98 rank? *Weather*, 53, 315-324, 1998.
- Woodworth, P. L., White, N. J., Jevrejeva, S., Holgate, S. J., Church, J. A., and Gehrels, W. R.: Evidence for the accelerations of sea level on multi-decade and century timescales. *Int. J. Climatol.*, 29, DOI: 10.1002/joc.1771, 777-789, 2009.
- Wu, Z., and Huang, N. E.: A study of the characteristics of white noise using the empirical mode decomposition method, *Proc. R. Soc. London A*, 460, 1597–1611, doi:10.1098/rspa.2003.1221, 2004.
- Wu, Z., and Huang, N. E.: Ensemble empirical mode decomposition: a noise-assisted data analysis method. *Advances in Adaptive Data Analysis*, 1(01), 1-41.
- Wu, Z., Huang, N. E., Long, S. R., and Peng, C.-K.: On the trend, detrending and variability of nonlinear and nonstationary time series. *Proc. Natl. Acad. Sci. U.S.A.*, 104, 14889–14894, 2007.
- Zhang, Y., Wallace, J.M., and Battisti, D.S.: ENSO-like interdecadal variability: 1900-93, *J. Climate*, 10, 1004-1020, 1997.

Figure Captions

Figure 1. Low frequency IMFs resulting from EMD of random noise with a standard deviation of 60 mm.

Figure 2. True monthly sea level recorded at San Francisco tide gauge (blue), and a) simulated by a trend + acceleration + seasonal + 13-year, 55-year, and 80-year sinusoidal functions with additional random noise (Case 1), and b) simulated by a trend + acceleration + seasonal + ENSO + PDO (Case 2). See text for details.

Figure 3. True oscillations and long-term trend + acceleration (red) for simulation shown in Figure 2a, along with the closest correlating IMF (blue).

Figure 4. Histogram of correlation values for IMFs in Case 1 correlated with the (a) 13-year, (b) 55-year, and (c) 55-year signals.

Figure 5. Mean (solid blue line) and standard deviation (light blue envelope) of IMFs calculated from the 1000 different Case 1 simulations, along with the true signal (red).

Figure 6. Mean (solid blue line) and standard deviation (light blue envelope) of IMFs calculated from the 1000 different Case 2 simulations, along with the true signal (red).

Figure 7. Low frequency IMFs resulting from EMD of a simulated signal with a trend + acceleration + seasonal variations + random noise + a single large anomalous event in January 1956.

Don Chambers 1/15/15 3:46 PM
Formatted: Font:Bold

Don Chambers 1/15/15 3:45 PM
Deleted: 4

Don Chambers 1/15/15 3:45 PM
Deleted: 5

Don Chambers 1/15/15 3:45 PM
Deleted: 6

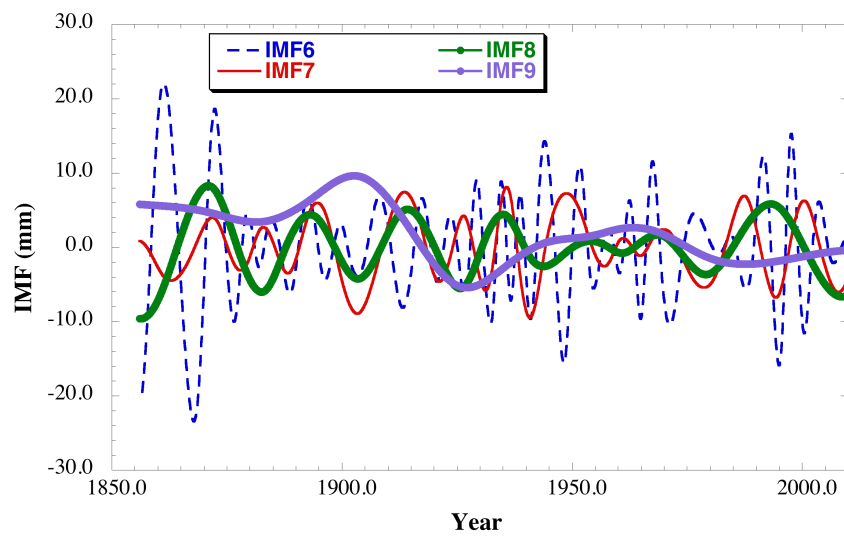


Figure 1. Low frequency IMFs resulting from EMD of random noise with a standard deviation of 60 mm.

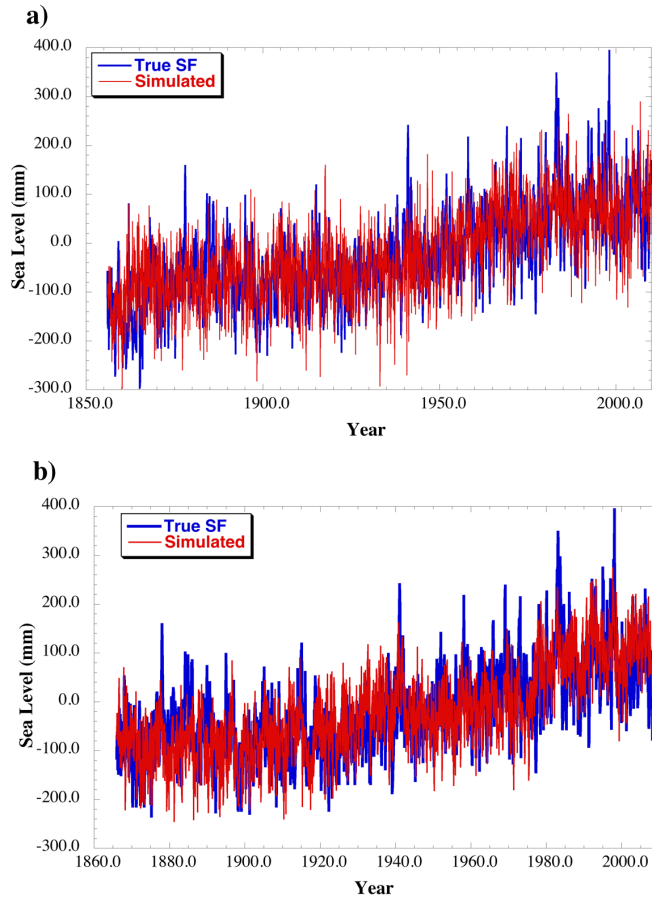


Figure 2. True monthly sea level recorded at San Francisco tide gauge (blue), and a) simulated by a trend + acceleration + seasonal + 13-year, 55-year, and 80-year sinusoidal functions with additional random noise (Case 1), and b) simulated by a trend + acceleration + seasonal + ENSO + PDO (Case 2). See text for details.

|

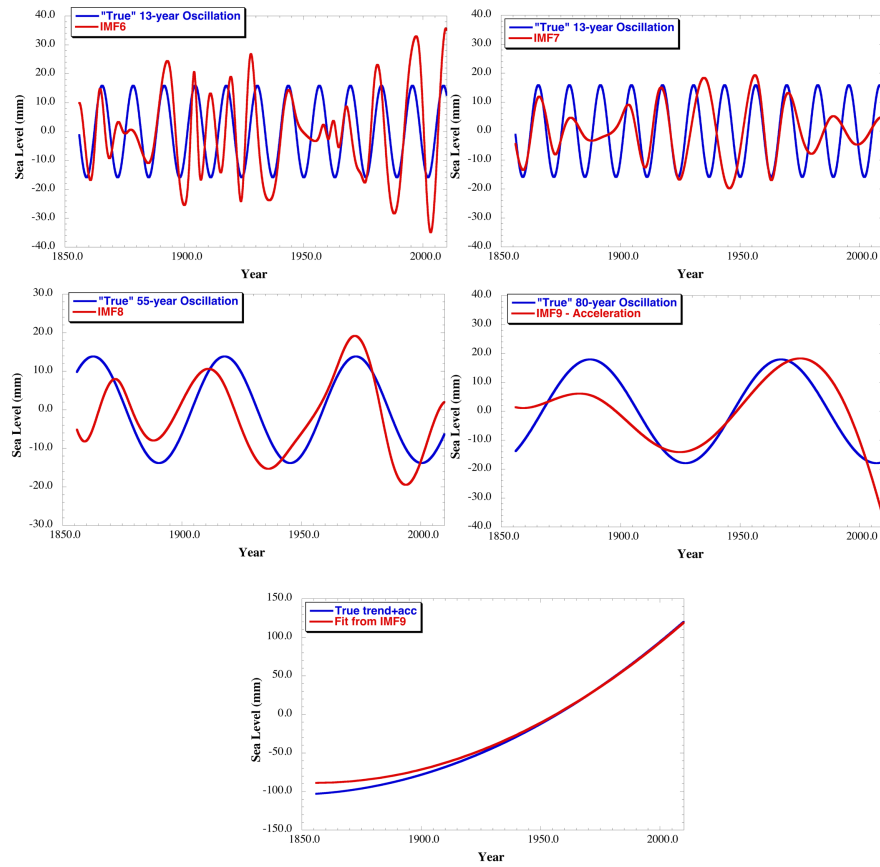


Figure 3. True oscillations and long-term trend + acceleration (red) for simulation shown in Figure 2a, along with the closest correlating IMF (blue).

|

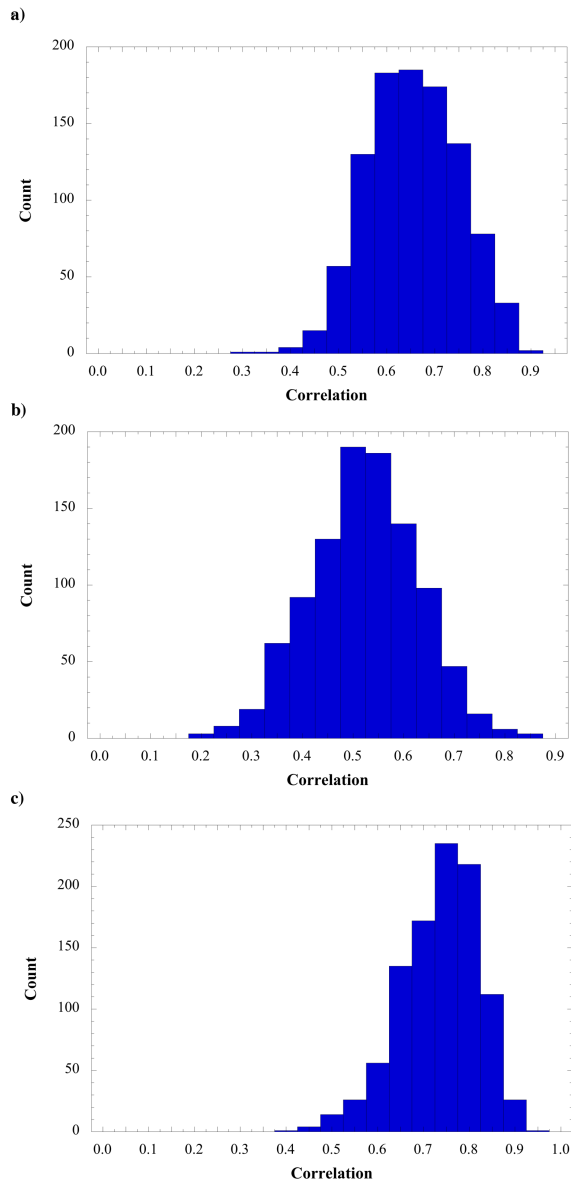


Figure 4. Histogram of correlation values for IMFs in Case 1 correlated with the (a) 13-year, (b) 55-year, and (c) 55-year signals.

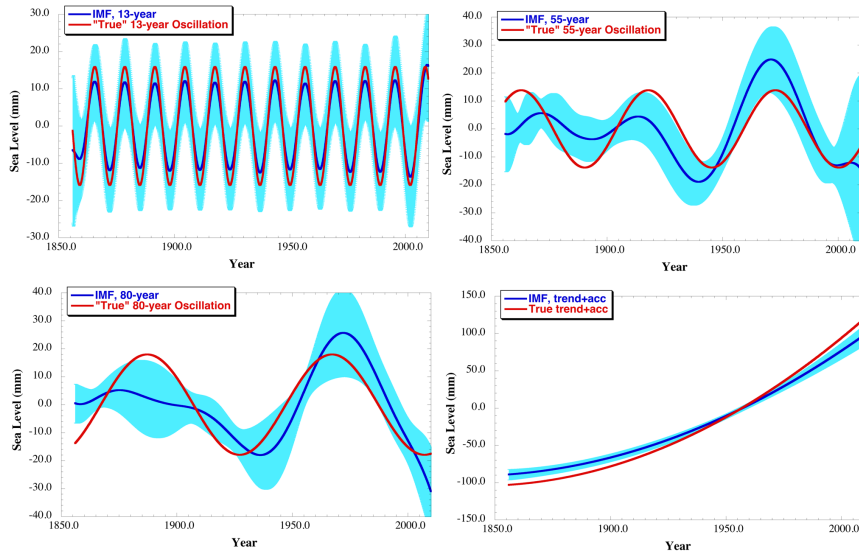


Figure 5. Mean (solid blue line) and standard deviation (light blue envelope) of IMFs

calculated from the 1000 different Case 1 simulations, along with the true signal (red).

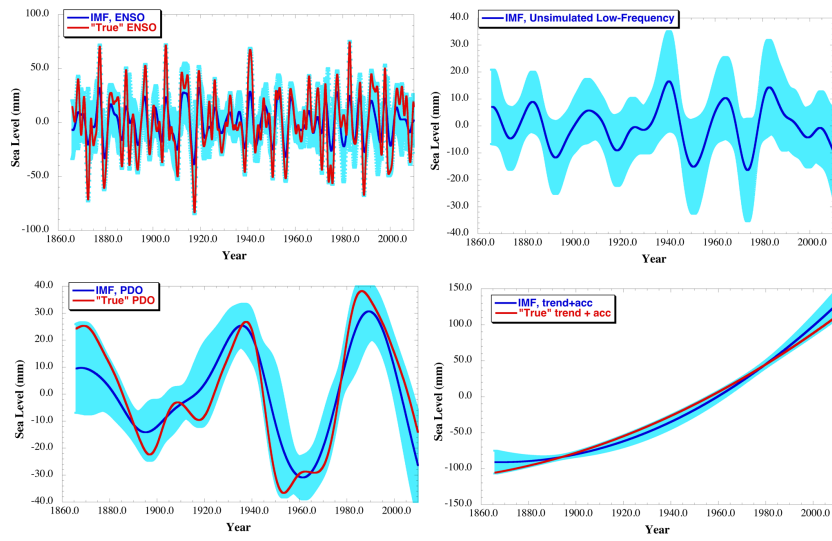


Figure 6. Mean (solid blue line) and standard deviation (light blue envelope) of IMFs calculated from the 1000 different Case 2 simulations, along with the true signal (red).

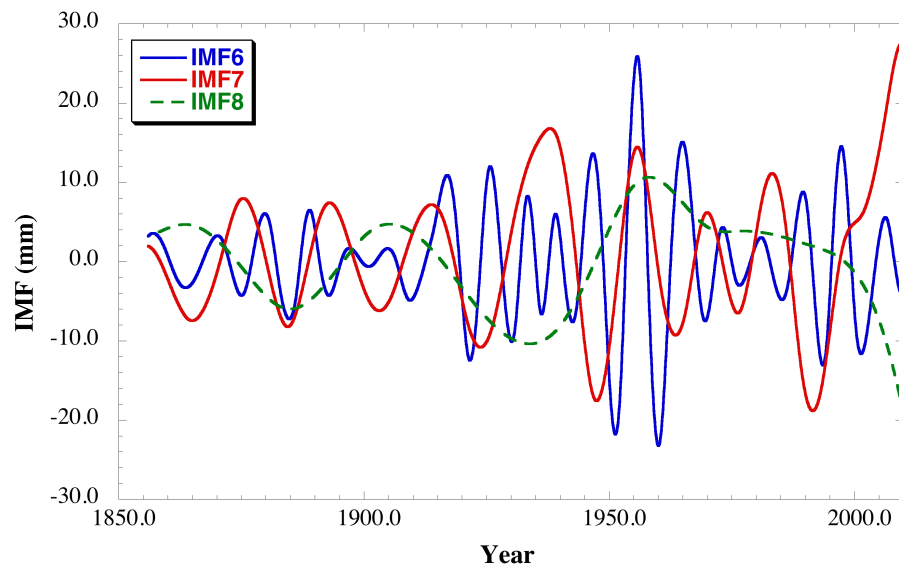
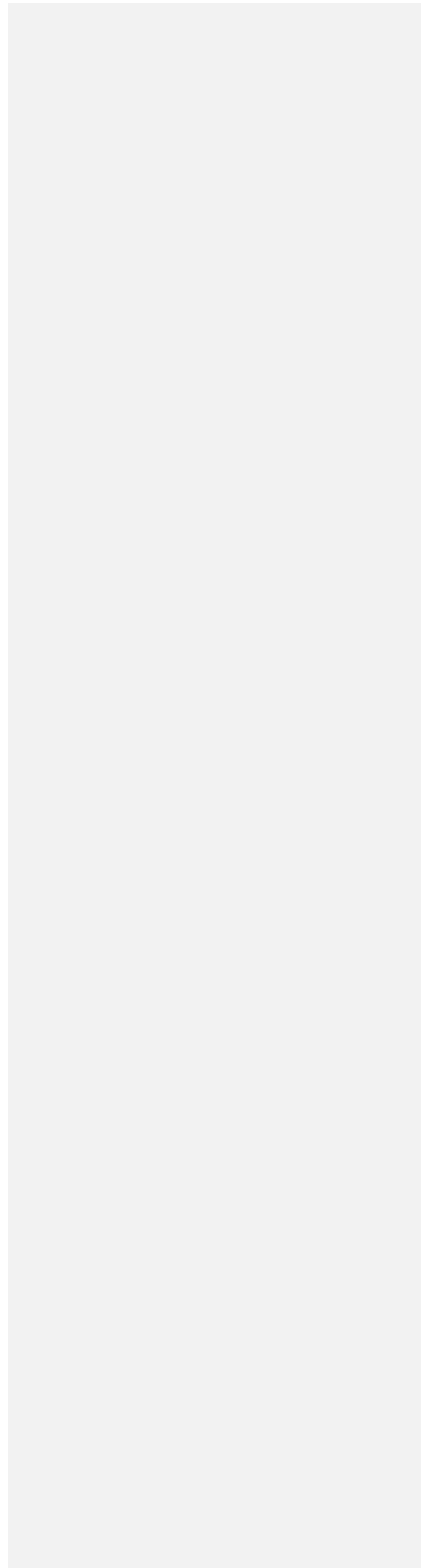


Figure 6. Low frequency IMFs resulting from EMD of a simulated signal with a trend + acceleration + seasonal variations + random noise + a single large anomalous event in January 1956.



Moreover, Wu and Huang (2004) have shown that EMD behaves as a low-pass filter. If one runs random noise with a normal Gaussian distribution through the process, low-frequency signals will be seen in the resulting IMFs. They found there is roughly a doubling of the average period with each subsequent mode. attempted to account for this behavior in the error statistics of their recovered IMFs. This is done by creating several pseudo-time-series by adding a small amount of random noise to the original time-series, running a large number of EMDs, and considering the average. The standard deviation of the differences represents the uncertainty. Hit does not consider the effect of the real signal in the tide gauge data that has high variance and little serial correlation. This “signal” will also be filtered by the EMD process and will likely appear as a quasi-stationary oscillation in higher order IMFs that is not real (i.e., similar to the EMD filtering of purely random noise shown in Figure 1).

In a simple example of running

Note that we used 60 mm because this is the standard deviation of residual monthly sea level at San Francisco after fitting and removing a quadratic function plus annual and semiannual sinusoids, so is representative of high-frequency sea level at a typical site, although some sites can have significantly higher variability.

Wu and Huang (2004) and others (Ezer and Corlett, 2012; Ezer et al., 2013) have attempted to account for this behavior in the error statistics of their recovered IMFs. This is done by creating several pseudo-time-series by adding a small amount of random noise to the original time-series, running a large number of EMDs, and considering the average.

The standard deviation of the differences represents the uncertainty. However, we see a problem with this method. Although adding random noise will account for uncertainty in the IMFs from uncertainty in the measurement, it does not consider the effect of the real signal in the tide gauge data that has high variance and little serial correlation. This “signal” will also be filtered by the EMD process and will likely appear as a quasi-stationary oscillation in higher order IMFs that is not real (i.e., similar to the EMD filtering of purely random noise shown in Figure 1).

Page 5: [4] Moved to page 5 (Move #2) Don Chambers

1/15/15 11:30 AM

attempted to account for this behavior in the error statistics of their recovered IMFs. This is done by creating several pseudo-time-series by adding a small amount of random noise to the original time-series, running a large number of EMDs, and considering the average. The standard deviation of the differences represents the uncertainty. H

Page 5: [5] Moved to page 5 (Move #3) Don Chambers

1/15/15 11:32 AM

it does not consider the effect of the real signal in the tide gauge data that has high variance and little serial correlation. This “signal” will also be filtered by the EMD process and will likely appear as a quasi-stationary oscillation in higher order IMFs that is not real (i.e., similar to the EMD filtering of purely random noise shown in Figure 1).