November 11, 2017

Dear Editor,

We are grateful for the comments from you and the reviewers. We are pleased that the comments are generally positive and that most of the concerns seem to be with the presentation of the material rather than the material itself. We agree that the previous draft required a substantial amount of clarification in order to be more accessible to geophysicists. Consequently, we have revised most of the methods section (Section 2) with the hopes of improving clarity. Details of our revisions are given below in our response to the reviewers' comments.

In the revised draft, we have addressed several issues that we found, which were not brought up by the reviewers:

1) The strain covariance function, eq. 14 of the previous draft, was incorrect. That covariance function was only correct for the shear strain. The normal strains differ by a factor of 2. This error was NOT in our software implementation of the method, and so our presented results have not changed.

2) There was an error in eq. 20 of the previous draft. We used a Dirac delta function where we should have used a Kronecker delta.

3) In the previous draft, Figure 5 was the only figure that showed how well our model fits the GNSS data. The data consists of seven years of daily displacements from 94 stations, so it is not possible to show all the data. However, we are now showing the fit at 3 stations, rather than 1. These stations were chosen to illustrate a slight over-smoothing in our fit to the data.

4) In addition to clarifying Section 2, we have significantly rewritten Sections 3 (application to the Pacific Northwest) and Section 4 (the discussion) so that they have a clearer focus and structure. However, the content being presented has not changed.

Each of the reviewers comments are written below and followed by our response. We hope that we have addressed their concerns sufficiently in our revision.

Reviewer #1:

Review of the paper : Revealing transient strain in geodetic data with Gaussian process regression

Overview of the work: The paper focuses on a new non-parametric Bayesian method to estimate transient strain rate from GNSS data. It is based on Gaussian Process regression. The method also includes an outlier removal processing step. This technique is then applied to the detection of slow slip events in Cascadia. The validity of the detected transient signals is accomplished via comparison with the seismic tremor recorded along the Cascadia range.

The article is well-written and your algorithm is very interesting. I think the use of Gaussian Process Regression is novel and open new roads in GNSS data analysis. I believe that you spend a lot of efforts in trying to justify most of the assumptions behind your model, together with the effects of using different kind of prior. Perhaps, that overshadows slightly your results. You could consider to reorganize the article to make it more readable for geophysicists.

I have taken several steps to make my revised draft more readable for geophysicists. For details, see my response to the editor's comment "Many of the reviewers' comments ... ".

I also require that a comparison between your outlier removal model and one other algorithm could help to support your claims about the efficiency of your method.

See my response to the comment "Also, the paper would be improved if the proposed outlier detection ..."

Otherwise, I have underlined a few clarifications needed about the stochastic model and some paragraphs within the text.

General Comments:

About the method:

The method is well established. The assumptions to justify the various hypothesis on the deterministic models of the GNSS time series are generally well supported (seasonal signal, tectonic

rate, offsets).

1- Now, the correlation of the stochastic processes is not completely clear. When the authors justify the model behind equation 3, wij is a normally distributed, uncorrelated noise. In statistic, it is equal to a WGN process. Then, the authors introduce a separate parameter (eta) which models the temporally correlated noise. This parameter is also following a Gaussian distribution, with a zero-mean and a covariance matrix C_η (i.e. line 10 p5). This approach is not common in modelling GNSS time series. The standard is based on Williams 2003, where the author established a stochastic model with a covariance matrix as sum of identity matrix for the white noise and another matrix which represents the coloured noise. The coloured noise covariance matrix is defined differently (see the literature – Bock and Melgar Reviews of Geophysics 2016, He et al., Journal of geodynamics 2017).

I have changed my description of the data (eq. 3 in the revised draft) so that it includes a single term, η, that encompasses correlated and uncorrelated noise. I then break η down into correlated and uncorrelated components in Section 3.1 (eq. 25). The covariance matrix for η in eq. 26 is more consistent with the literature.

2- First, the author should give a summary of stochastic noise modelling in GNSS. I also think that a discussion is needed to relate to previous models. Why using separated stochastic models? Also I would comment on the possibility that the (low) spatial correlation between the parameters in your deterministic model (i.e. secular velocities) could be absorbed in the estimation of the covariance matrix C_η.

I believe the reviewer is refering to page 4-5 of the original draft, where I first introduce η. I agree that a more thorough discussion on the noise model is necessary, but I do not think this is an appropriate place in the text for it. The paper is laid out so that Section 2 contains a general description of the method for estimating transient strain rates, and Section 3 contains details that are specific to the particular application. There is no universally appropriate noise model, so I decided to save a discussion on the noise model for Section 3. This is clarified in Section 2 of the revised draft when I say "The appropriate noise model may vary depending on the application, and we hold off on specifying the covariance matrix, C_{η_i}, until Section 3.1."

Now, I am confused When looking at p.8 Section 4.1, the description of the noise models should be given as a subsection of Section 2 p.4. I would also improve the literature review. Several other models have also been discussed (i.e. Montillet et al., 2014 uses a fractional Brownian motion model; ARMA, ARFIMA, GGM or Band pass noise – see Het et al., 2017).

Again, we think that the description of the noise model should remain in Section 3 because it is specific to the Pacific Northwest application.

We have added a few more references in Section 3.1 of the revised draft to emphasize that GNSS noise modeling is a well-worked area. We refer the reader to Bock and Melgar 2016 and He et al 2017 for a more complete review.

Outlier detection:

1- P7 : I would give a longer summary in the introduction about all the efforts in outlier detection in GNSS time series. The first paragraph p.7 is not enough.

In Section 2.2 of the revised draft, I have expanded my discussion on how previous studies have treated outliers in GNSS data.

2- Also, the paper would be improved if the proposed outlier detection method is compared with existing ones such as the Hector software package (Bos et al., 2013). For example, the comparison between this method and another one could be done when applying the algorithm in p. 13 (last paragraph below equation 18).

My outlier detection algorithm is quite similar to what is used in Bos et al 2013 (and many other papers), in that the outliers are determined based on the residuals for a best fitting model. The only difference is that I am fitting the data with a stochastic model rather than a parametric model consisting of a linear trend and sinusoids. When using a parametric model, I found that deformation from slow slip events would sometimes be erroneously identified as an outlier. The stochastic model used in my outlier detection method is more flexible and able to describe this transient deformation, and so only high frequency anomalous observations get identified as outliers. I have included Figure A1 to demonstrate this.

Of course, a parametric model that includes, for example, B-splines would be just as effective at detecting outliers without removing geophysical signal. So I am not suggesting that my outlier detection method is a

significant advancement or even a novel method. Instead, I am only describing my outlier detection method for the sake of completeness. The previous version of my paper seems to have put too much emphasis on the outlier detection method, when the main focus should remain on the method for estimating transient strain rates. To help keep my paper focused, I placed the description of the outlier detection method in the appendix.

Results:

The results are generally well explained. However, the use of different priors and the comparison in the text is sometimes confusing. Perhaps, a table could summary the different priors and the main results. It would ease the reading of Section 4.2, 4.3 and then the discussion in Section 5.

In the previous draft, I was discussing two different priors: a prior for the outlier detection algorithm, and a prior used in computing transient strain rates. I acknowledge that this was convoluted. In the revised draft, all the details of the outlier detection algorithm are confined to the appendix. I hope this clarifies any confusion.

Also, I have rewritten most of the results section to be more structured and hopefully have a clearer focus.

Minor issues:

P2 line 30 : "fidelity" replace with "reliability".

Done

P2 line 30-31 " Developing and improving upon methods for deriving secular...area of research." Need references.

The introduction has been revised and this sentence has been removed.

p.2 line 43 :"too large of an area" ... perhaps " a very large area"

Done

P5 line 32 : "formal data uncertainty": can you define it?

By "formal data uncertainties" I meant "the derived uncertainties that accompany the GNSS displacement solutions". The revised text has been clarified.

P 6 : Need reference to show Equation 15.

I have added a few more steps in the derivation of eq. 15.

P 7 line 54-60: This paragraph can be further enhanced with references (Bock and Melgar, 2016, Gazeaux et al., 2013).

I have added a reference to Gazeaux et al. 2013.

P 8 : first paragraph. add references on Cascadia range (Aguiar, Melbourne, Scriver 2009), (Melbourne and Szeliga 2005), (Melbourne and Webb, 2003, 2002).

We agree that we need to acknowledge the extensive amount of research on Cascadia SSEs. In the revised draft, we refer the reader to the review by Schartz and Rokosky 2007.

P.8 line 20 delete www.unavco.org (only in Acknowledgment)

Done

p. 11 line 45 "common mode error" you need to add some references and define it properly.

In Section 3.1 of the revised draft, where we elaborate on the noise model, we describe common mode error by saying "Another significant source of noise in GNSS data is common mode error (e.g., Wdowinski et al. 1997; Dong et al. 2006), which is noise that is highly spatially correlated." We believe this definition and these references are sufficient.

p.13 : "Wendland functions have compact support and hence their corresponding covariance.. sparse". I would just mention that the associated covariance matrices are sparse.

I recognize that it may be a bit redundant to say that the covariance function is compact and thus the covariance matrices are sparse. However, Wendland covariance functions are typically described as "compact" in the literature, and I think it is a key adjective that should not be dropped out. I have left the wording as is.

p.13 line 22 what is an "isotropic Gaussian process"? please add a definition.

An Gaussian process is isotropic if its covariance function, $C(x, x')$, can be written as a function of $||x - x'||_2^2$. An isotropic Gaussian process is also stationary, meaning that its statistical properties are invariant to translations. We have replace "isotropic" with "stationary" in the text because "stationary" is a more appropriate, and perhaps more familiar, term.

P.20 : " Using a compact covariance function .." Again I would just say that the covariance matrix is sparse and that s why you have the mathematical simplification.

My personal preference is to leave the wording as is.

p.20 line 33 " computational burden" I think that refers to p.20 line 60 "prohibitive when using sveral years of daily GNSS". It would interesting to have some number. How do you quantify the computation time to process longer and longer time series?

I agree that numerical values would be useful; however, I am reluctant to quantify the computational cost of evaluation transient strain rates. There are many factors that influence the computational cost (whether sparse matrices are used, the type of sparse matrices that are used, the linear solver algorithm, etc.) and I am not convinced that my implementation is the most efficient. So I do not give details on the computation time, but I note in the discussion that this is an area for further research.

Reviewer #2:

Comments to the Author(s) Hines/Hetland GJI Review 08/2017

This is a well written paper and I have just a few minor clarifications and suggestions below. One point that might be of interest to address is would it be possible to "approximate" this approach so that it can run much faster? I gather from the text that the method is intensive for CPU and memory use.

Our approach is indeed computationally intensive and approximation methods for GPR do exist. In the discussion section for the revised draft, we suggest exploring these approximation methods in future research.

Page 3: Last line: min(t,t') needs as scaling factor to make units for covariance e.g., m^2/yr

We added the scale factor ϕ to the covariance function for Brownian motion.

Page 4: Stochastic model: It would seem that the assumption of no covariance between the north and east components is not a very good one since there is a tectonic framework in which the transients occur and its unlikely this framework aligns along the Cardinal directions. Maybe add a statement of impact of the neglect and possibly the idea of reorienting the "axes" to align with the tectonic framework e.g., perpendicular to the subduction zone interface in this case?

This is a good point that we neglected to mention in the previous draft. We have added a paragraph to Section 2 that explains the reasoning and implications for this assumption. The paragraph starts with "It should be noted that we have ignored any covariances ...".

Page 4/5 Equation (3): Is it worth discussing at this point whether the estimates of u and n (eta) can be separated without explicit prior knowledge (which we may not have). For example, orbit modeling errors on one satellite (e.g., due to an unknown yaw problem in the spacecraft) wilt generate a spatially and temporally correlated error in d_{ij}. How does this (stochastically non-modeled) error not project in the u estimates?

At this point in the text, it is assumed that the statistical properties of u and eta are known, although not yet specified. So I do not think this would be an appropriate place to discuss a scenario in which we do not know u or eta. I also think that it should go with saying that the validity of our results are contingent on having a good model for u and eta.

In Section 3.1 and 3.2, we discuss how we choose a model for u and eta. The key piece of "prior knowledge" that allows us to discern u from eta is that u is negligible for inland stations. This allowed us to islolate and constrain eta first, and then we constrained u. If we allowed eta to be spatially correlated (which we did not), then we would have conceivably been able to describe the spatially and temporally correlated noise from, for example, orbit modeling errors.

Page 6: Equation 10: Maybe it is discussed later in the paper but it is probably worthwhile stating at this point how the inverse is performed with a block 0 in the matrix?

The matrix being inverted in eq. 10 and 11 of the revised draft is still invertible with the block 0, provided that

the columns of G are linearly dependent. This is explained in the revised draft.

Page 6: Equation 12: Maybe some additional explanation is needed here. How are these partial derivatives formed. My understanding the u estimates at this point correspond one-to-one with the positions and times of the GNSS position determinations. Did I miss something?

The posterior displacements, $\hat{u}$, are actually a spatially and temporally continuous stochastic process. We can evaluated $\hat{u}$ at any position and time that we may be interested in (not just the position and times of the data). We are also able to compute the spatial and temporal derivatives analytically. In Section 2 of the revised draft, we explain how the derivatives are computed in greater detail. We hope this clarifies any confusion.

Page 8: Partly addresses issue raised about eqn 3. It probably should be noted that wether conditions east of -121 deg longitude are different to the coast so there could be possible problems with this noise model.

This is also a good point that we neglected to mention in the first draft. In Section 3.1 of the revised draft, we acknowledge the potentially dubious assumption that noise east of -121 is representative of the entire region. We added the disclaimer saying "We assume the noise at these inland stations is representative of the noise at all the stations considered in this study, which is probably a poor assumption since the inland stations are subject to distincly different climatic conditions".

Editor:

Editor's Review of GJI MS 17.0593 "Revealing transient strain..." by Hines and Hetland This is an novel algorithm applied to a problem of great current interest. Based on the reviewer's comments and my own reading, I am recommending moderate revision. The authors' response should include a description of their changes, and also a version of the paper on which these can be seen (just as highlights in the text).

Many of the reviewers' comments address places where the paper could be more clearly written, and my overall comment would be the same. The current version perhaps spends a little too much time on the background and not enough on explaining the method and its derivation. As shown by the references below (these include some the authors cite), and the Dermanis paper, which I have also attached, strain estimation from scattered data is an area in which there has been a lot more prior work than the authors mention. I am including these references not because the authors need to cite many or even most of them, but only to make the point that the problem of estimating strain, and of looking for transients, is a familiar one: for many readers, myself among them, much more familiar than the statistical methods applied here. So a slightly fuller discussion of these methods (perhaps a little more background and some reminders of, eg, what a hyperparameter is) would be welcome to many readers; and increased clarity of presentation usually means that the paper will be used and cited more.

I appreciate this comment, and I have extensively revised the paper in order to make it more readable. To list some of the major changes:

1) I properly define a Gaussian process in Section 2. I elaborate on its definition to give it a more tangible meaning, and I use Brownian motion as an example of a Gaussian process.

2) I give more references, including section numbers for books, to make it easier for the reader to get additional information.

3) In section 2 of the revised draft, I give an example of a prior for transient displacements. I include a figure showing what the covariance function looks like and what a corresponding realization looks like.

4) I provide a clearer derivation of the posterior transient displacements (eq. 6 and 7 of the revised draft).

5) In the revised draft, eq. 19 and 20 were added to help the reader understand how the transient strain rates are computed.

6) I provide a clearer derivation of the signal-to-noise ratio.

7) I moved details of the outlier detection method to the appendix because it is not essential to the text.

8) I expanded the explanation of the REML method, which is used to constrain u and eta in Section 3.1 and 3.2.

9) I attempt to avoid esoteric jargon such as "hyperparameter". In the revised draft, "hyperparameter" has

been replaced with "parameter", which does not degrade the precision of our description of the method.

10) Most of the paper was reworded / restructed to improve clarity. Despite the fact that very little of the original text is in the revised daft, there are almost no differences in the content being presented.

I feel the same way about the section on detecting outliers: again, a well-worked area. Iterative fitting and outlier detection, as an overall strategy, hav e been used for a long time. I appreciate that by imposing spatial and temporal conditions you can identify outliers not otherwise obvious, but it would help to show an example of this, rather than the current Figure 4, in which most of the identified cases are ones that any method could easily find.

I recognize that outlier detection is a well-worked area. I included the description of the outlier detection method for completeness, and not because I thought it was an innovative or superior method. In order to keep the focus of my paper on my method for computing transient strain, I move the outlier detection method to the appendix.

In my experimentation with the outlier detection algorithm, it turns out that the prior spatial covariance has virtually no effect on the detected outliers. However, the temporal covariance is important. The prior stochastic model needs to be flexible enough to describe deformation from, for example, SSEs. This is illustrated in Figure A1 of the revised draft.

A few specifics (numbers are page+line):

2+26: possibility not risk: the latter is a technical term in seismic hazard.

Done

2+47-48: I do not think this description of Shen's method (which I've used and programmed) is right, since it allows the deformation gradients to vary in space. I've always thought of it as a kind of adaptive kernel smoother, somewhat like applying 2-D loess; so it is only depen- dent on nearby points, and has no assumption of uniformity anywhere. (Actually this raises a more general question: is the method given here use local or global support?)

I have updated my description of Shen's method in the revised draft.

3+7: could also

Done

3+32: redo the sentence; as it stands, you are saying that the SCEC exercise calculates strain rates.

The text has been clarified say "The method described in Holt 2013 was also tested in the SCEC transient detection exercise. Holt 2013 detects transient signal from an estimate of transient strain rates, and they estimate transient strain rates using a ..."

3+37: a geophysical signal

Done

3+47: function not signal

Done

3+56: by a Gaussian, not with

This sentence has been removed in the revised draft

4+37: displacement (not plural)

This sentence has been removed in the revised draft

6+56: I suspect that equation (15) took some effort to derive; could some details go into an Appendix?

I added a clearer derivation in the revised draft.

Figure 6: I like that the strain uncertainties can be shown, but is there a reason not to plot principle strains using arrows, showing errors by an ellipse around the tips?

I have put a lot of thought into representing strain, and I recognize that strain is conventionally represented by

showing the principle strains axes. However, there is no good way of showing the uncertainties for the principle strain axes. For example, consider a perfectly dilational strain that is perturbed by a small amount of noise. The orientation of principle strains will be highly uncertain, but the magnitudes of the principle strains will be well constrained. If we were to draw some sort of confidence interval around the tips of the principle strains, then they would have a crescent shape (not an ellipse). It is unclear how I would write a program to draw these confidence intervals.

Instead of showing the principle strains, I am showing the normal strains as a function of azimuth. For a given azimuth, the uncertainty in the normal strain is computed by linear propagation of the uncertainties in the strain. In my opinion, it is much easier to show the uncertainties in an azimuthal representation of strain.

I did not include this justification in the revised draft because I think it would be too much of a digression.

Thank you for your continued
consideration,

Trever T. Hines