Revealing Transient Strain in Geodetic Data with Gaussian Process Regression

Trever T. Hines and Eric A. Hetland

1   Introduction

Crustal strain rates are fundamentally important quantities for assessing seismic hazard. Knowing where and how quickly strain is accumulating gives insight into where we can expect stored elastic energy to be released seismically. Consequently, secular crustal strain rates estimated from GNSS data have been used to constrain seismic hazard models such as UCERF3 (Field et al., 2014). Dense networks of continuous GNSS stations, such as the Plate Boundary Observatory (PBO), make it feasible to estimate transient crustal strain, which we consider to be any deviation from the secular strain rate. Transient crustal strain can potentially illuminate geophysical signal that may not be immediately apparent from the GNSS displacement data. The ability to detect transient geophysical phenomena is also relevant for assessing seismic hazard because there are instances of major earthquakes being triggered by slow slip events (Roeloffs, 2006) and postseismic deformation (Freed and Lin, 2001).

Most methods for estimating crustal strain rates from GNSS data assume some parametric form of the deformation signal. The simplest method for estimating secular crustal strain rates assumes that GNSS derived velocities can be described with a first order polynomial (i.e., having constant deformation gradients) over some subnetwork of the GNSS stations (e.g., Feigl et al., 1990; Murray and Lisowski, 2000). The components of the strain rate tensor for each subnetwork are then determined through a least squares fit to the observations. The assumption that deformation gradients are spatially uniform is not appropriate when subnetworks span too large of an area. To help overcome this deficiency, Shen et al. (1996, 2015) used an inverse distance weighting scheme, in which the estimated strain rate at some point is primarily controlled by observations at nearby stations. However, the methods described in Shen et al. (1996, 2015) are still formulated by assuming that the deformation gradients are uniform over the entire network. The errors in this assumption manifest as implausibly low formal uncertainties for the estimated strain rates. Other methods for estimating secular strain rates have parameterized GNSS derived velocities with bi-cubic splines (Beavan and Haines, 2001), spherical wavelets (Tape et al., 2009), and elastostatic Green's functions (Sandwell and Wessel, 2016). The type of basis functions and the number of degrees of freedom for a parameterization can be subjective. If there are too few degrees of freedom in the parameterization, then estimated strain rates will be biased and the uncertainties will be underestimated. On the other hand, if there are too many degrees of freedom, then there will not be any coherent features in the estimated strain rates. The methods described by Beavan and Haines (2001), Tape et al. (2009), and Sandwell and Wessel (2016) also require the user to specify penalty parameters that control a similar trade-off between bias and variance in the solution. One could parameterize deformation with a physically motivated model of interseismic deformation (e.g., Meade and Hager, 2005; McCaffrey et al., 2007). In such models the lithospheric rheology and fault geometries are assumed to be known. Any errors in the assumed physical model could similarly result in biased strain estimates and underestimated formal uncertainties.

The aforementioned studies are concerned with estimating secular strain rates. In recent years the Southern California Earthquake Center (SCEC) community has shown interest in developing methods for detecting transient strain. SCEC supported a transient detection exercise (Lohman and Murray, 2013), where several research groups tested their methods for detecting transient geophysical signals with a synthetic GNSS dataset. Among the methods tested were the Network Strain Filter (NSF) (Ohtani et al., 2010) and the Network Inversion Filter (NIF) (Segall and Mathews, 1997). The NSF uses a wavelet parameterization to describe the spatial component of geophysical signal. The NIF, which is intended for imaging slow fault slip from geodetic data, uses the elastic dislocation Green's functions from Okada (1992). For the NSF and NIF, the time dependence of the geophysical signal is modeled as integrated Brownian motion. The method described in Holt and Shcherbenko (2013) was also tested in the SCEC transient detection exercise, which calculates strain rates using a bi-cubic spatial parameterization of displacements between time epochs. Holt and Shcherbenko (2013) defined a detection threshold based on the

strain rate magnitude, and below we demonstrate that this is indeed an effective criterion for identifying geophysical signal. For the same reasons descibed above, the transient deformation and corresponding uncertainties estimated by these methods can be biased by the chosen spatial parameterization. It is then difficult to distinguish signal from noise with these methods, which limits their utility for transient detection.

Here we propose using Gaussian process regression (GPR) (Rasmussen and Williams, 2006) to estimate transient strain from GNSS data. GPR is a Bayesian, non-parametric method for inferring a continuous signal from scattered data. Since GNSS stations are irregularly spaced and observation times may differ between stations, GPR is an ideal tool for synthesizing GNSS data into a spatially and temporally continuous representation of surface deformation. GPR is closely related to kriging (Cressie, 1993) and least squares collocation (Kato et al., 1998), and it has recently gained popularity in the machine learning community. GPR is Bayesian in that we describe our prior understanding of the geophysical signal with a Gaussian process. A Gaussian process is a normally distributed, continuous, stochastic process that is fully defined in terms of a mean function and a positive-definite covariance function. For example, Brownian motion, $B(t)$, is a well known Gaussian process in $\mathbb{R}^1$ which has zero mean and covariance function $\mathrm{cov}(B(t), B(t')) = \min(t, t')$, where $t, t' \geq 0$. If no prior information is available for the geophysical signal, then maximum likelihood methods can be used to choose a prior Gaussian process that is most consistent with the observations. We incorporate GNSS observations with the prior to form a posterior estimate of transient strain. The posterior transient strain is also a Gaussian process, and we can use its distribution to confidently discern geophysical signal from noise. Here we use GPR to infer transient strain resulting from slow slip events (SSEs) in Cascadia. Our results demonstrate that GPR can be an effective tool for detecting SSEs and revealing the spatio-temporal evolution of crustal strain during SSEs.

2   Estimating Transient Strain Rates

We seek a spatially and temporally dependent estimate of transient crustal strain rates. Our attention is limited to horizontal strain rates in this study. We denote transient crustal strain rates as $\dot{\underline{\varepsilon}}(p)$, where p represents the ordered pair $(\vec{x}, t)$, $\vec{x}$ are spatial coordinates in $\mathbb{R}^2$, and t is time. We determine $\dot{\underline{\varepsilon}}(p)$ by spatially and temporally differentiating estimates of transient displacements, $\vec{u}(p)$. We make a prior assumption that each component of $\vec{u}$ is a Gaussian process,

$$u_i(p) \sim \mathcal{N}\left(0, C_{u_i}\right), \tag{1}$$

where $C_{u_i}(p, p')$ is a covariance function indicating how we expect $u_i(p)$ to covary with $u_i(p')$. For simplicity, we treat each component of displacement independently and ignore any potential covariance. We drop the component subscripts with the understanding that the same analysis is being repeated to estimate the easting and northing components of $\vec{u}$. We assume that C_u can be separated into positive definite spatial and temporal functions as

$$C_u\left((\vec{x}, t), (\vec{x}', t')\right) = X(\vec{x}, \vec{x}')T(t, t'). \tag{2}$$

The appropriate choice for X and T may vary depending on the geophysical signal we are trying to describe (e.g. postseismic deformation or deformation from slow slip events), and we discuss this matter in Section 4.2.

We constrain u with GNSS data, which records u as well as other physical and non-physical processes which we are not interested in. We describe GNSS observations at position $\vec{x}_i$ and time t_j as a realization of the random variable

$$\begin{aligned}
d_{ij} =&\, u(\vec{x}_i, t_j) + \eta(\vec{x}_i, t_j) + w_{ij} + a_i^{(1)} + a_i^{(2)} t_j + \\
&\, a_i^{(3)} \sin(2\pi t_j) + a_i^{(4)} \cos(2\pi t_j) + a_i^{(5)} \sin(4\pi t_j) + a_i^{(6)} \cos(4\pi t_j),
\end{aligned} \tag{3}$$

where $a_i^{(1)}$ is an offset that is unique for each GNSS monument, $a_i^{(2)}$ is the secular velocity at $\vec{x}_i$, and the sinusoids describe seasonal deformation (using units of years for t_j). We use w_{ij} to denote normally distributed, uncorrelated noise. Correlated noise which does not have a parametric representation is denoted by η. For example, η can consist of temporally correlated noise describing benchmark wobble (e.g., Wyatt, 1982, 1989), and/or spatially correlated noise describing common mode error (e.g., Wdowinski et al., 1997). For now, we will only assume that $\eta \sim \mathcal{N}(0, C_\eta)$. We consider the six coefficients in eq. (3) to be uncorrelated random variables distributed as $\mathcal{N}(0, \kappa^2)$ in the limit as $\kappa \to \infty$ (i.e., the coefficients have diffuse priors). Of course, the secular velocities, $a_i^{(2)}$, are spatially correlated and we could invoke a tectonic model to form a prior on $a_i^{(2)}$. However, in our application to Cascadia, we will be using displacement time series which are long enough to sufficiently constrain $a_i^{(2)}$ for each station, avoiding the need to

incorporate a prior on it. Likewise, seasonal deformation is spatially correlated (Dong et al., 2002; Langbein, 2008), and it may be worth exploring and exploiting such a correlation in a future study.

We now consider the column vector of n GNSS observations, $\boldsymbol{d}_*$. Let $\boldsymbol{P}$ be the set of $(\vec{x}_i, t_j)$ pairs describing where and when each of the GNSS observations have been made. Let $\boldsymbol{a}$ be the vector of coefficients from eq. (3) for each of the m GNSS stations. We use $\boldsymbol{G}$ to represent the $n \times 6m$ matrix of corresponding basis functions evaluated at each point in $\boldsymbol{P}$. We also denote the vector of uncorrelated noise for each observation as $\boldsymbol{w}$, whose standard deviations are given by the formal data uncertainty, $\boldsymbol{\sigma}$. The observations can then be viewed as a realization of the random vector

$$\boldsymbol{d} = u(\boldsymbol{P}) + \eta(\boldsymbol{P}) + \boldsymbol{w} + \boldsymbol{G}\boldsymbol{a}, \tag{4}$$

which is distributed as $\mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma} + \kappa^2 \boldsymbol{G}\boldsymbol{G}^T)$, where

$$\boldsymbol{\Sigma} = C_u(\boldsymbol{P}, \boldsymbol{P}) + C_\eta(\boldsymbol{P}, \boldsymbol{P}) + \mathrm{diag}\left(\boldsymbol{\sigma}^2\right). \tag{5}$$

It should be understood that notation such as $u(\boldsymbol{P})$ and $C_u(\boldsymbol{P}, \boldsymbol{P})$ represents the column vector $[u(P_i)]_{P_i \in \boldsymbol{P}}$ and the matrix $[C_u(P_i, P_j)]_{(P_i, P_j) \in \boldsymbol{P} \times \boldsymbol{P}}$, respectively.

The prior for transient displacements is then conditioned with $\boldsymbol{d}_*$ to form a posterior estimate of transient displacements, $\hat{u} = u|\boldsymbol{d}_*$. We will assume that appropriate covariance functions and corresponding hyperparameters for X, T, and C_η have already been chosen. In Section 4.1 and 4.2, we discuss how the covariance functions are chosen for our application to GNSS data from Cascadia. If κ is kept finite then, following Rasmussen and Williams (2006), we find that $\hat{u}$ is distributed as $\mathcal{N}(\mu_{\hat{u}}, C_{\hat{u}})$, where

$$\mu_{\hat{u}}(p) = C_u(p, \boldsymbol{P}) \left(\boldsymbol{\Sigma} + \kappa^2 \boldsymbol{G}\boldsymbol{G}^T\right)^{-1} \boldsymbol{d}_* \tag{6}$$

and

$$C_{\hat{u}}(p, p') = C_u(p, p') - C_u(p, \boldsymbol{P}) \left(\boldsymbol{\Sigma} + \kappa^2 \boldsymbol{G}\boldsymbol{G}^T\right)^{-1} C_u(\boldsymbol{P}, p'). \tag{7}$$

However, we are interested in the limit as $\kappa \to \infty$, and the form for eq. (6) and eq. (7) is not suitable for evaluating this limit. We use the partitioned matrix inversion identity (e.g., Press et al., 2007) to rewrite eq. (6) and eq. (7) as

$$\mu_{\hat{u}}(p) = \left[\begin{array}{cc} C_u(p, \boldsymbol{P}) & \boldsymbol{0} \end{array}\right] \left[\begin{array}{cc} \boldsymbol{\Sigma} & \boldsymbol{G} \\ \boldsymbol{G}^T & -\kappa^{-2}\boldsymbol{I} \end{array}\right]^{-1} \left[\begin{array}{c} \boldsymbol{d}_* \\ \boldsymbol{0} \end{array}\right] \tag{8}$$

and

$$C_{\hat{u}}(p, p') = C_u(p, p') - \left[\begin{array}{cc} C_u(p, \boldsymbol{P}) & \boldsymbol{0} \end{array}\right] \left[\begin{array}{cc} \boldsymbol{\Sigma} & \boldsymbol{G} \\ \boldsymbol{G}^T & -\kappa^{-2}\boldsymbol{I} \end{array}\right]^{-1} \left[\begin{array}{c} C_u(\boldsymbol{P}, p') \\ \boldsymbol{0} \end{array}\right]. \tag{9}$$

Taking the limit as $\kappa \to \infty$, we get the solution for the mean and covariance of $\hat{u}$,

$$\mu_{\hat{u}}(p) = \left[\begin{array}{cc} C_u(p, \boldsymbol{P}) & \boldsymbol{0} \end{array}\right] \left[\begin{array}{cc} \boldsymbol{\Sigma} & \boldsymbol{G} \\ \boldsymbol{G}^T & \boldsymbol{0} \end{array}\right]^{-1} \left[\begin{array}{c} \boldsymbol{d}_* \\ \boldsymbol{0} \end{array}\right] \tag{10}$$

and

$$C_{\hat{u}}(p, p') = C_u(p, p') - \left[\begin{array}{cc} C_u(p, \boldsymbol{P}) & \boldsymbol{0} \end{array}\right] \left[\begin{array}{cc} \boldsymbol{\Sigma} & \boldsymbol{G} \\ \boldsymbol{G}^T & \boldsymbol{0} \end{array}\right]^{-1} \left[\begin{array}{c} C_u(\boldsymbol{P}, p') \\ \boldsymbol{0} \end{array}\right]. \tag{11}$$

We use eq. (10) and (11) to find the posterior easting and northing components of transient displacements. Using $\hat{u}_i$ to denote the posterior transient displacements along direction i and x_i to represent the components of $\vec{x}$, we can write the components of $\underline{\dot{\varepsilon}}$ as

$$\dot{\varepsilon}_{ij}(p) = \frac{1}{2} \frac{\partial}{\partial t} \left(\frac{\partial \hat{u}_i(p)}{\partial x_j} + \frac{\partial \hat{u}_j(p)}{\partial x_i}\right). \tag{12}$$

The transient strain rate components are Gaussian processes with mean functions

$$\mu_{\dot{\varepsilon}_{ij}}(p) = \frac{1}{2} \frac{\partial}{\partial t} \left(\frac{\partial \mu_{\hat{u}_i}(p)}{\partial x_j} + \frac{\partial \mu_{\hat{u}_j}(p)}{\partial x_i}\right) \tag{13}$$

and covariance functions

$$C_{\dot{\varepsilon}_{ij}}(p, p') = \frac{1}{4} \frac{\partial^2}{\partial t \, \partial t'} \left(\frac{\partial^2 C_{\hat{u}_i}(p, p')}{\partial x_j \, \partial x'_j} + \frac{\partial^2 C_{\hat{u}_j}(p, p')}{\partial x_i \, \partial x'_i}\right). \tag{14}$$

Our motivation for estimating transient strain rates is, in part, to detect geophysical phenomena. We propose using a signal-to-noise ratio, SNR, based on $\underline{\dot{\varepsilon}}$ as our detection criterion. The Frobenius norm of $\underline{\dot{\varepsilon}}$, $||\underline{\dot{\varepsilon}}||_\mathrm{F}$, which is also referred to as the second invariant of strain rate in the geodetic literature, is often used as a metric for the strain rate "magnitude". Noting that $||\underline{\dot{\varepsilon}}||_\mathrm{F}$ is a random variable, we choose SNR to be the ratio of the estimated mean and standard deviation of $||\underline{\dot{\varepsilon}}||_\mathrm{F}$. Using nonlinear uncertainty propagation, we find SNR to be

$$\mathrm{SNR}(p) = \frac{\mu_{\dot{\varepsilon}_\mathrm{nn}}(p)^2 + \mu_{\dot{\varepsilon}_\mathrm{ee}}(p)^2 + 2\mu_{\dot{\varepsilon}_\mathrm{en}}(p)^2}{\left(C_{\dot{\varepsilon}_\mathrm{nn}}(p,p)\mu_{\dot{\varepsilon}_\mathrm{nn}}(p)^2 + C_{\dot{\varepsilon}_\mathrm{ee}}(p,p)\mu_{\dot{\varepsilon}_\mathrm{ee}}(p)^2 + 4C_{\dot{\varepsilon}_\mathrm{en}}(p,p)\mu_{\dot{\varepsilon}_\mathrm{en}}(p)^2\right)^{\frac{1}{2}}}, \tag{15}$$

where the subscripts "n" and "e" denote north and east, respectively. For simplicity, we have ignored covariances between the strain rate components in eq. (15), even though they are non-zero.

3 Outlier detection

In our formulation for estimating transient strain rates, we have assumed that noise in the data vector is normally distributed. This is not an appropriate assumption for GNSS data which are prone to more outliers than would be predicted for normally distributed noise. It follows that proposed methods for analyzing GNSS data should be robust against outliers (e.g., Blewitt et al., 2016). In order to make our estimates of transient strain more robust, we automatically identify and remove outliers in the GNSS data as a pre-processing step.

Our method for detecting outliers is based on the data editing algorithm described in Gibbs (2011). We calculate the residuals between the observations and a best fitting model. Data with residuals that are anomalously large are then identified as outliers. We treat $\boldsymbol{d}_*$ as a sample of $\boldsymbol{d}$ and assume that there is no correlated noise (i.e., $\eta(p) = 0$). The best fitting model for $\boldsymbol{d}_*$ is considered to be the expected value of the random vector $u(\boldsymbol{P}) + \boldsymbol{Ga}$ after conditioning it with non-outlier observations. We still consider u to have a separable covariance function as in eq. (2), and the choice for X and T does not need to be the same as that used in Section 2. Since outliers are determined based on how well a spatially and temporally dependent model fits the data, we are able to identify anomalous observations which may not be immediately apparent based on inspection of individual time series.

To begin the algorithm, we let $\boldsymbol{\Omega}$ be the index set of non-outliers in $\boldsymbol{d}_*$ and initiate it with all n indices. This algorithm is iterative, and for each iteration we calculate the residual vector

$$\boldsymbol{r} = \frac{\boldsymbol{d}_* - \mathrm{E}\left[(u(\boldsymbol{P}) + \boldsymbol{Ga})|\tilde{\boldsymbol{d}}_*\right]}{\boldsymbol{\sigma}} \tag{16}$$

$$= \frac{1}{\boldsymbol{\sigma}}\left(\boldsymbol{d}_* - \begin{bmatrix} C_u(\boldsymbol{P},\tilde{\boldsymbol{P}}) & \boldsymbol{G} \end{bmatrix} \begin{bmatrix} C_u(\tilde{\boldsymbol{P}},\tilde{\boldsymbol{P}}) + \mathrm{diag}(\tilde{\boldsymbol{\sigma}}^2) & \tilde{\boldsymbol{G}} \\ \tilde{\boldsymbol{G}}^T & \boldsymbol{0} \end{bmatrix}^{-1} \begin{bmatrix} \tilde{\boldsymbol{d}}_* \\ \boldsymbol{0} \end{bmatrix}\right), \tag{17}$$

where the tilde indicates that only elements corresponding to indices in $\boldsymbol{\Omega}$ are retained (e.g., $\tilde{\boldsymbol{P}} = \{P_i\}_{i\in\boldsymbol{\Omega}}$). We then update $\boldsymbol{\Omega}$ to be

$$\boldsymbol{\Omega} = \{i : |r_i| < \lambda \cdot \mathrm{RMS}\}, \quad r_i \in \boldsymbol{r}, \tag{18}$$

where RMS is the root-mean-square of $\tilde{\boldsymbol{r}}$ and λ is an outlier tolerance. We use $\lambda = 4$ in this study, which in our experience accurately identifies outliers without unnecessarily decimating the data. Iterations continue until the new $\boldsymbol{\Omega}$ is equal to the previous $\boldsymbol{\Omega}$.

It should be noted that this algorithm does not identify jumps in GNSS time series, which are another common issue. Some, but not all, jumps can be automatically removed by looking up the dates of equipment changes and earthquakes. However, it is still necessary to manually find and remove jumps of unknown origin. That being said, this outlier detection algorithm significantly reduces the effort needed to manually clean GNSS data.

4 Application to Cascadia Slow Slip Events

We demonstrate our method for estimating transient strain rates with GNSS data from Cascadia. We are interested in identifying transient strain resulting from SSEs (e.g., Dragert et al., 2001). Our results illuminate where elastic strain energy, which can eventually be released seismically, is accumulating during SSEs. We demonstrate that transient strain rates estimated with the method described above are an effective tool for detecting SSEs and potentially other geophysical signal in GNSS data. In Cascadia, SSEs can be detected by monitoring for associated
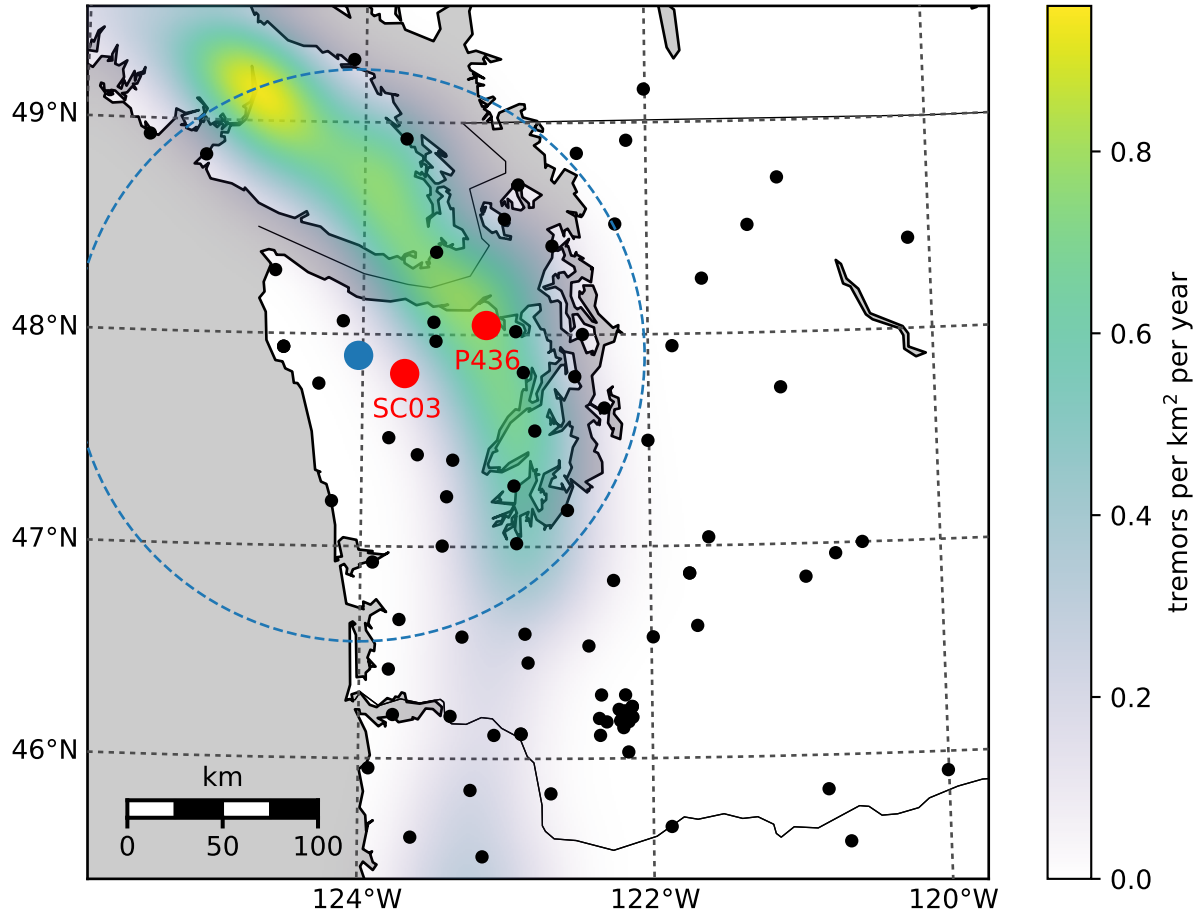
Figure 1: Positions of continuous GNSS stations used to estimate transient strain rates. The colored regions indicate the distribution of seismic tremor as determined by Wech (2010). The red dots show the positions of GNSS stations mentioned in this paper. The blue dot indicates the location of the strain rates shown in Figure 7 and 8. The blue dashed circle demarcates the spatial extent of the tremors shown in Figure 8.

seismic tremor (Rogers and Dragert, 2003), which is actively being done by the Pacific Northwest Seismic Network (PNSN) (Wech, 2010). We verify that transient strain rates estimated with the method described above are indeed identifying strain from SSEs by comparing $\dot{\underline{\underline{\varepsilon}}}$ to the tremor records. We envision that GNSS derived transient strain rates could be useful for detecting SSEs at subduction zones where there is no associated tremor (Schwartz and Rokosky, 2007).

We use publicly available continuous GNSS data from the University Navstar Consortium (UNAVCO) (Herring et al., 2016), which can be found at www.unavco.org. We limit the dataset to the stations and time ranges which are pertinent to the seven most recent SSEs in the Puget Sound region. The earliest SSE considered in this study began in August 2010, and the most recent SSE began in February 2017. We use these most recent SSEs because the station coverage is sufficiently dense for us to use maximum likelihood methods to constrain prior models. The positions of GNSS stations used to estimate transient strain rates are shown in Figure 1.

4.1 Noise model

Before we determine the transient strain rates, we must establish a prior for the transient displacements, u, and the noise, η. In this section we discuss our choice for the noise covariance function C_η. There have been numerous studies on temporally correlated noise in GNSS data (e.g., Zhang et al., 1997; Mao et al., 1999; Williams et al., 2004;

Langbein, 2008). In these studies, temporally correlated noise was described with some combination of Brownian motion, a first-order Gauss-Markov (FOGM) process, and/or flicker noise. There is some physical justification for using Brownian motion as a noise model because it accurately describes the power spectrum of motion resulting from instability in geodetic monuments (e.g., Wyatt, 1982, 1989). Here we describe the time dependence of η as a FOGM process and consider η to be spatially uncorrelated. A FOGM process is a solution to the stochastic differential equation

$$\dot{\eta}(t) + \alpha\eta(t) = \beta w(t), \tag{19}$$

where $w(t)$ is white noise with unit variance. The FOGM process degenerates to the commonly used Brownian motion noise model under the condition that $\alpha = 0$ and $\eta(0) = 0$. Our noise model that satisfies eq. (19) is a Gaussian process with zero mean and the covariance function

$$C_\eta\left((\vec{x}, t), (\vec{x}', t')\right) = \frac{\beta^2}{2\alpha} \exp\left(-\alpha|t - t'|\right) \delta(\|\vec{x} - \vec{x}'\|_2). \tag{20}$$

We constrain the hyperparameters for η, α and β, with a set of 38 continuous GNSS stations in the Pacific Northwest which are east of 121°W. These stations are sufficiently far from the subduction zone that they are unlikely to contain transient signal associated with the Cascadia SSEs. We clean the data for these stations by removing jumps at times of equipment changes, and we remove outliers that have been detected with the algorithm described in Section 3. We then find α and β for each station time series with the Restricted Maximum Likelihood (REML) method (e.g., Harville, 1974; Cressie, 1993; Hines and Hetland, 2017). The REML method finds the hyperparameters, which we collectively refer to as $\boldsymbol{\theta}$, that maximize the likelihood function

$$\mathcal{L}(\boldsymbol{\theta}) = \left(\frac{\left|\boldsymbol{G}^T\boldsymbol{G}\right|}{(2\pi)^{n-6m}\left|\boldsymbol{\Sigma}(\boldsymbol{\theta})\right|\left|\boldsymbol{G}^T\boldsymbol{\Sigma}\boldsymbol{G}\right|}\right)^{\frac{1}{2}} e^{-\frac{1}{2}\boldsymbol{d}_*^T\boldsymbol{K}(\boldsymbol{\theta})\boldsymbol{d}_*}, \tag{21}$$

where

$$\boldsymbol{K}(\boldsymbol{\theta}) = \boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1} - \boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\boldsymbol{G}\left(\boldsymbol{G}^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\boldsymbol{G}\right)^{-1}\boldsymbol{G}^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}. \tag{22}$$

Harville (1974) showed that choosing the hyperparameters which maximize eq. (21) is equivalent to choosing the hyperparameters which maximize the probability of drawing $\boldsymbol{d}_*$ from $\boldsymbol{d}$. We independently estimate $\boldsymbol{\theta}$ for each station, and so $\boldsymbol{d}_*$ consists of displacements for an individual station. We are assuming $u(p) = 0$ when estimating the noise hyperparameters for this section. We use the REML method over the maximum likelihood (ML) method (e.g., Langbein and Johnson, 1997) because the REML method accounts for the improper prior that we assigned to $\boldsymbol{a}$ (Hines and Hetland, 2017). The distribution of inferred α and β are shown in Figure 2. The amplitude of FOGM noise, β, for the easting and northing components is notable low and are clustered around 0.5 mm/yr$^{0.5}$. The corresponding estimates of α tend to cluster around 0 yr^{-1}, suggesting Brownian motion. We also estimate hyperparameters for the vertical component of displacements, even though verticals are not used for estimating strain. Our hope is that the vertical deformation gradients could reveal some geophysical signal. The amplitude of FOGM noise for the vertical component is relatively large with a median value of 13.5 mm/yr$^{0.5}$. The inferred values for α are also higher for the vertical component with a median value of 8.21 yr^{-1}. In Figure 3, we use the median values of α and β to generate two random samples of FOGM noise for each component. The samples span five years and over these five years the easting and northing samples drift by about 1 mm. In the context of detecting SSEs, which produce several mm's of surface displacement on the time-scale of weeks, the estimated FOGM noise for the easting and northing component is negligible. In contrast, the estimated FOGM noise for the vertical component is larger than the signal we would expect from SSEs. We suspect that the higher amplitude for the FOGM noise in the vertical component is accommodating for deficiencies in our rather simple seasonal model. Based on this analysis, we henceforth ignore temporally correlated noise in the easting and northing component because of its low amplitude. We also do not use vertical displacements because of the presumably low signal-to-noise ratio.

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. When not accounted for, common mode error manifests as spatially uniform undulations in estimated transient displacements. However, estimated transient strain rates are insensitive to common mode error. We therefore do not include common mode error in our noise model. We then make the simplifying assumption that $\eta(p) = 0$ for the easting and northing component of GNSS data.
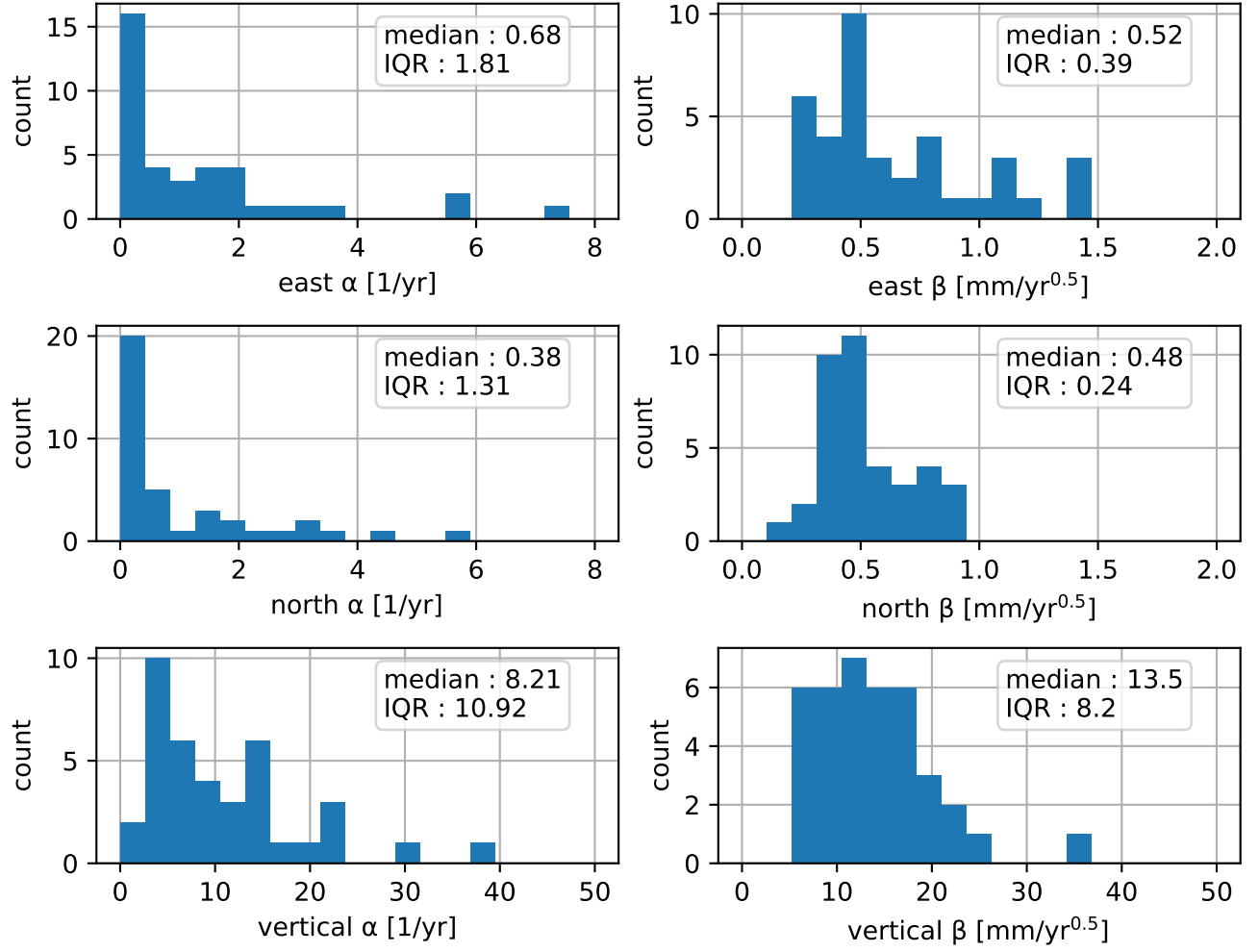
Figure 2: Distribution of estimated FOGM hyperparameters (eq. 20). Hyperparameters are estimated with the REML method for 38 stations in the Pacific Northwest that are east of 121°W. "IQR" is the inter-quartile range.
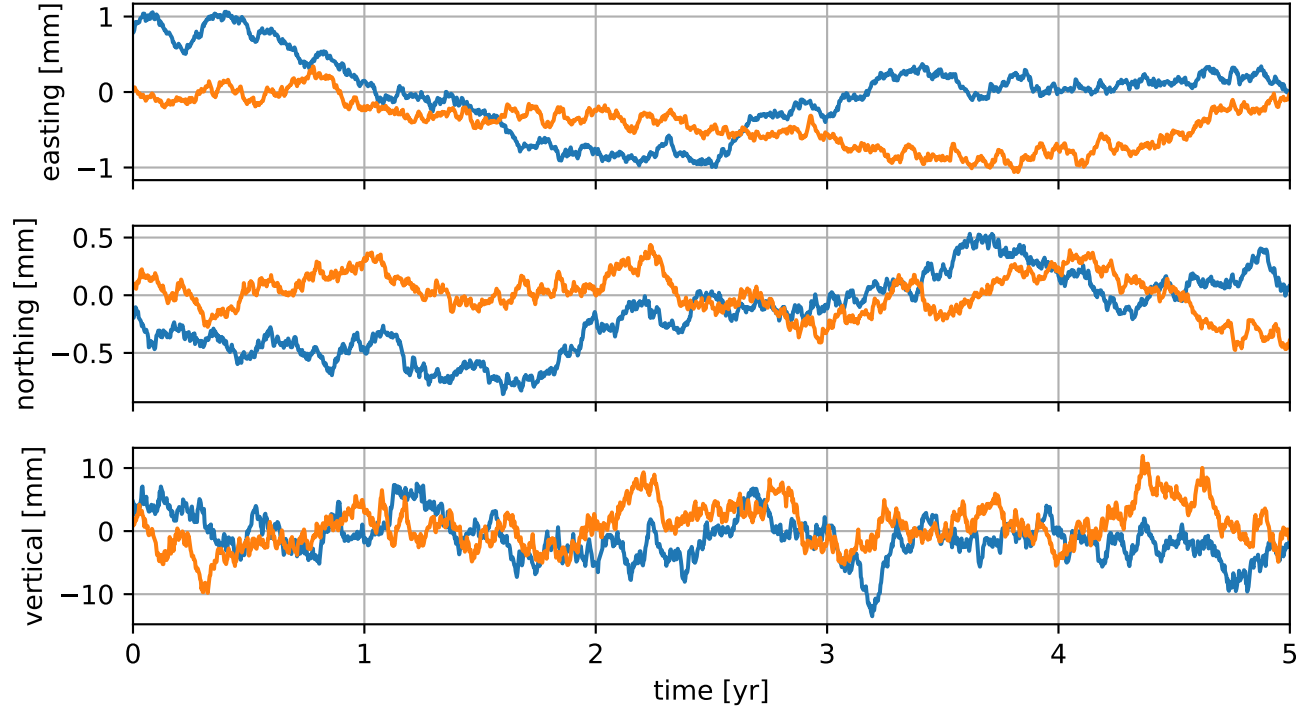
Figure 3: Two FOGM noise samples for each component. The FOGM hyperparameters have been set to the median values from Figure 2.

4.2 Transient displacement model

We next establish our prior model for transient displacements. Specifically, we discuss our choice for the covariance functions $X(\vec{x}, \vec{x}')$ and $T(t, t')$. For X, we use the squared exponential (SE) covariance function,

$$X(\vec{x}, \vec{x}') = \exp\left(\frac{-||\vec{x} - \vec{x}'||_2^2}{2\ell^2}\right). \tag{23}$$

The SE covariance function is commonly used in kriging (e.g, Cressie, 1993) and Gaussian process regression (e.g., Rasmussen and Williams, 2006). The SE is a positive definite covariance function for any number of spatial dimensions. A Gaussian process with an SE covariance function is isotropic and has realizations that are infinitely differentiable. In terms of geodetic applications, Kato et al. (1998) and El-Fiky and Kato (1999) demonstrated that the SE accurately describes the covariance of secular GNSS derived velocities in Japan.

We consider three potential models for the temporal covariance of u. First, we consider the one-dimensional SE covariance function,

$$T(t, t') = \phi^2 \exp\left(\frac{-|t - t'|^2}{\tau^2}\right). \tag{24}$$

Note that T includes the hyperparameter ϕ, which serves to scale the covariance function C_u. Second, we consider integrated Brownian motion (IBM). IBM has zero mean and its covariance function can be found by integrating the covariance function for Brownian motion as

$$T(t, t') = \int_0^t \int_0^{t'} \phi^2 \min(s, s') \, ds' \, ds \tag{25}$$

$$= \frac{\phi^2}{2} \min(t, t')^2 \left(\max(t, t') - \frac{1}{3} \min(t, t')\right), \quad t, t' \geq 0. \tag{26}$$

IBM has been used in the context of Kalman filtering as a non-parametric model for the time dependence of geophysical signals (e.g., Segall and Mathews, 1997; McGuire and Segall, 2003; Ohtani et al., 2010; Hines and Hetland, 2016). It should be emphasized $t = 0$ is a reference time at which the Gaussian process is exactly zero.

8

T	direction	ℓ	ϕ	τ	diff. log(REML)
SE	east	92 ± 25 km	0.62 ± 0.11 mm	0.026 ± 0.011 yr	-
SE	north	91 ± 53 km	0.43 ± 0.05 mm	0.030 ± 0.017 yr	-
Wendland	east	95 ± 30 km	0.66 ± 0.15 mm	0.093 ± 0.044 yr	0.78 ± 0.87
Wendland	north	92 ± 57 km	0.46 ± 0.10 mm	0.116 ± 0.057 yr	0.08 ± 0.58
IBM	east	110 ± 130 km	290 ± 420 mm/yr$^{1.5}$	-	-16.4 ± 7.8
IBM	north	150 ± 560 km	110 ± 250 mm/yr$^{1.5}$	-	-10.1 ± 2.3

Table 1: Optimal hyperparameters for the prior on transient displacements determined with the REML method. The temporal covariance function is indicated by the "T" column. The SE, IBM, and Wendland covariance functions are defined in eqs. (24), (25), and (27), respectively. The spatial covariance function, X, is the squared exponential (eq. 23) in all cases. The hyperparameters are estimated for each of the seven SSEs considered in this study, and the tabulated values indicate the median and interquartile ranges of estimates. The "diff log(REML)" column compares the log REML likelihood to the log REML likelihood when using the SE covariance function for T. Positive values indicate that observations are more consistent with the SE covariance function.

For some geophysical signals, it is appropriate to have this reference time. For example, if we are trying to identify postseismic deformation then t should be zero at the time of the earthquake. However, if we are interesting in detecting transient events, where there is no known start time, then IBM may not be an appropriate prior and an isotropic Gaussian process should be preferred. In the following analysis, we make the quite arbitrary choice that t is zero on the first epoch of $\boldsymbol{d}_*$. Using an earlier reference time does not change the results discussed in this section. Our third option for T is the Wendland class of covariance functions (Wendland, 2005). Wendland covariance functions have compact support and hence their corresponding covariance matrices will be sparse. In our analysis, we exploit this sparsity with the CHOLMOD software package (Chen et al., 2008). We elaborate on the computational advantages of the Wendland functions in Section 5. Wendland functions are positive definite in $\mathbb{R}^d$, and they describes an isotropic Gaussian process with realizations that can be differentiated k times. The form of the covariance function depends on the choice of d and k. We use $d = 1$ since we are describing the temporal covariance of u. We use $k = 1$ because we need u to be at least once differentiable in time to determine strain rates. The corresponding Wendland covariance function is

$$T(t, t') = \phi^2 \left(1 - \frac{|t - t'|}{\tau}\right)_+^3 \left(\frac{3|t - t'|}{\tau} + 1\right), \tag{27}$$

where

$$(t)_+ = \begin{cases} t, & t > 0 \\ 0, & \text{otherwise.} \end{cases} \tag{28}$$

We next determine appropriate hyperparameters for X and each of the three candidate covariance functions for T. First, we clean the GNSS datasets by removing offsets at times of equipment changes and removing outliers with the method describe in Section 3. For the outlier detection algorithm, our prior model, u, is chosen to have a length-scale and time-scale which is able to approximately describe SSE displacements. We use the SE covariance function for X with length-scale $\ell = 100$ km, and we use the Wendland covariance function for T, due to its computational efficiency, with time-scale $\tau = 0.1$ yr and $\phi = 1$ mm. The outlier detection algorithm is particularly effective at removing outliers for stations at high elevation (Figure 4), which can be adversely affected by ice or snow during the winter (Lisowski et al., 2008). After cleaning the dataset, we divide it into seven subsets which are four months long and each centered on the time of an SSE. The times of the seven SSEs are determined with tremor records from Wech (2010). We use the REML method to find the optimal hyperparameters for T and X for each subset of data. We chose to make each data subsets four months long because it is long enough to encompass a SSE in Cascadia, while it is short enough to still be computationally tractable. However, four months is too short to resolve the sinusoids in $\boldsymbol{d}$, and they are omitted from $\boldsymbol{d}$ for the REML analysis. The estimated hyperparameters for u are summarized in Table 1. Based on the interquartile ranges (IQR), the estimated hyperparameters for the SE and Wendland covariance functions do not vary significantly between SSEs. This suggests that the median of estimated hyperparameters should be an appropriate prior model for all Cascadia SSEs. For the IBM model, there are several anomalously large values for ℓ and ϕ, which results in large IQRs.

Next we identify which covariance function for T best describes the SSEs. One approach is to compare the REML likelihood for each covariance function, similar to the analysis in Langbein (2004). In Table 1, we summarize how the log REML likelihoods for the Wendland and IBM covariance functions compare to the SE covariance function.
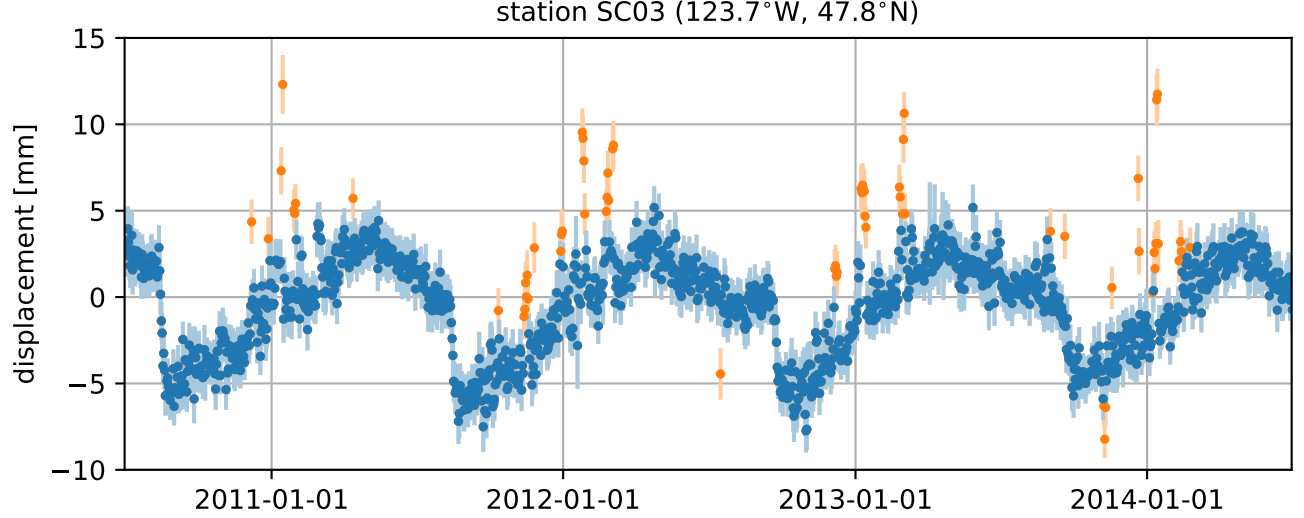
Figure 4: Detrended easting component of GNSS displacement data for station SC03, which is located on Mount Olympus in Washington. The orange markers indicate outliers which were automatically detected using the algorithm from Section 3. The error bars show one standard deviation uncertainties. Note that outliers tend be observed in the winter, suggesting that they were caused by snow or ice.

Based on the differences in log REML likelihoods, the data is substantially more likely to come from a Gaussian process with a SE or Wendland covariance function than an IBM covariance function. The REML likelihoods do not definitively indicate whether the SE or Wendland covariance function is preferable.

To further explore which covariance function for T best describes SSEs, we compare the observations to the predicted displacements for each covariance function. We consider the data prediction vector to be $\hat{\boldsymbol{d}} = (u(\boldsymbol{P}) + \boldsymbol{G}\boldsymbol{a})\,|\boldsymbol{d}_*$. It can be shown that $\hat{\boldsymbol{d}}$ is normally distributed with mean

$$\boldsymbol{\mu}_{\hat{d}} = \left[\begin{array}{cc} C_u(\boldsymbol{P},\boldsymbol{P}) & \boldsymbol{G} \end{array}\right] \left[\begin{array}{cc} \boldsymbol{\Sigma} & \boldsymbol{G} \\ \boldsymbol{G}^T & \boldsymbol{0} \end{array}\right]^{-1} \left[\begin{array}{c} \boldsymbol{d}_* \\ \boldsymbol{0} \end{array}\right] \tag{29}$$

and covariance

$$\boldsymbol{C}_{\hat{d}} = C_u(\boldsymbol{P},\boldsymbol{P}) - \left[\begin{array}{cc} C_u(\boldsymbol{P},\boldsymbol{P}) & \boldsymbol{G} \end{array}\right] \left[\begin{array}{cc} \boldsymbol{\Sigma} & \boldsymbol{G} \\ \boldsymbol{G}^T & \boldsymbol{0} \end{array}\right]^{-1} \left[\begin{array}{c} C_u(\boldsymbol{P},\boldsymbol{P}) \\ \boldsymbol{G}^T \end{array}\right]. \tag{30}$$

We compute $\hat{\boldsymbol{d}}$ using SE, Wendland, and IBM covariance functions for T and the median hyperparameters from Table 1. Figure 5 compares the easting component of $\boldsymbol{d}_*$ to $\hat{\boldsymbol{d}}$ for the Winter 2015-2016 SSE at Station P436. The data prediction vector appears to accurately describe displacements throughout the SSE, regardless of the choice of T. Since the SSE signal is strongest at Station P436, it can be assumed that T adequately describes the signal elsewhere. The prediction for the IBM covariance function contains slightly more high frequency, and perhaps spurious, features. The predictions for the Wendland and SE covariance functions are nearly indistinguishable. In our estimates of transient strain discussed in the next section, we ultimately settle on the Wendland covariance function for T and use the median values from Table 1 for the hyperparameters. We choose the Wendland covariance function over the SE covariance function because of its computational advantages.

4.3   Transient Strain Rates

Having established a noise model and a prior for transient displacements, we use the cleaned GNSS dataset to calculate transient strain rates in the Puget Sound region. We calculate transient strain rates for each day from January 1, 2010 to March 15, 2017. The strain rates are estimates at a grid of points spanning the study area. In Figure 6 we show the transient strain rates on January 1, 2016, which coincides with the height of an SSE. We have included a video showing the map view of strain rates over time as supplementary material. The strain rates shown in Figure 6 are generally similar to the strain rates for the other six SSEs considered in this study. The SSEs cause compression in the Olympic Peninsula and extension east of Puget Sound. For comparison, estimated secular
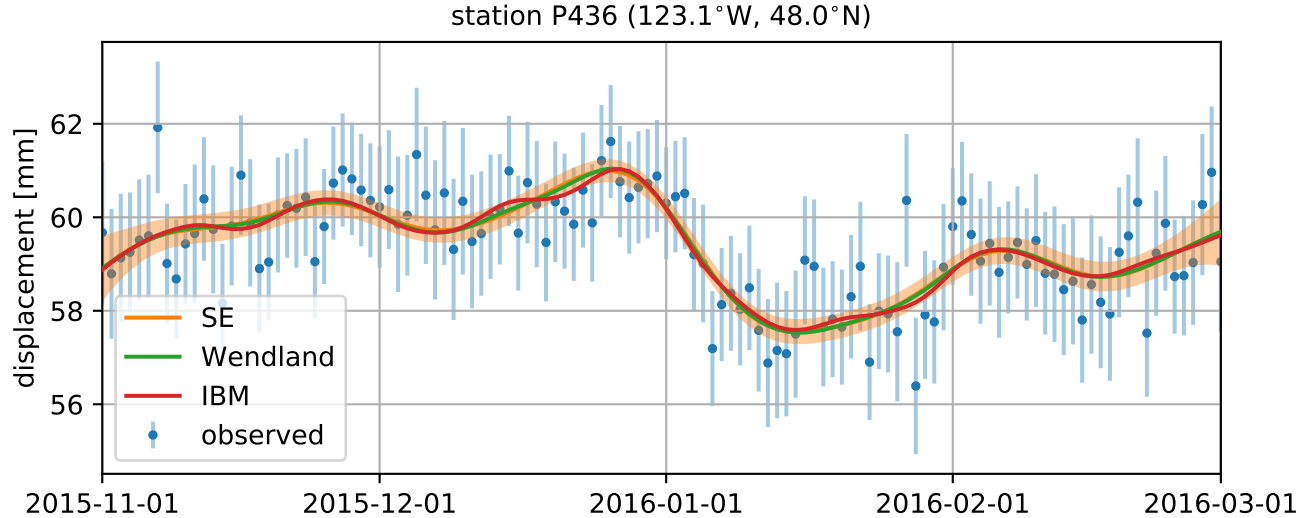
Figure 5: Observed easting component of displacements at station P436 and predicted displacements when using different covariance functions for T. The one standard deviation uncertainties are shown for the observations and the predicted displacements when using the SE covariance function. For clarity, uncertainties are not shown for the IBM and Wendland covariance functions, but they are nearly equivalent to the uncertainties for the SE covariance function.

strain rates indicate trench perpendicular compression throughout this study region (Murray and Lisowski, 2000; McCaffrey et al., 2007, 2013). The SSEs are thus concentrating tectonically accumulated strain energy trench-ward, and presumably pushing the subduction zone closer to failure. Similar conclusions have been drawn based on fault slip models (e.g., Dragert et al., 2001; Wech et al., 2009; Schmidt and Gao, 2010), which reveal that SSEs are occurring down-dip of the seismogenic zone and migrating stress upward. A key difference between the strain inferred here and strain that can be derived from fault slip models is that our estimated strain rates are not based on an assumed physical model. In contrast, fault slip models can be biased by errors in the assumed fault geometry or lithospheric rheology. Moreover, the degrees of freedom in fault slip models usually cannot be constrained by GNSS data alone, and it is necessary to impose regularization which further biases the results. Since our estimated strain rates lack such systematic errors, we can be more confident that our solution is unbiased and has meaningful uncertainties.

In Figure 7 we show the time dependence of estimated transient strain rates at a position on the Olympic Peninsula, where transient strain rates from SSEs are largest. To verify that the estimated transient strain rates are accurately identifying geophysical signal, we compare the signal-to-noise ratio from eq. (15), SNR, to the frequency of seismic tremor (Figure 8). An SNR greater than ~3 can be interpreted as a detected geophysical signal. For each detected event there is a corresponding peak in seismic tremor. We are also able to clearly identify transient strain associated with a more subtle inter-SSE event in August 2014. In between SSEs, the SNR is consistently between 0 and 2, indicating that no anomalous non-SSE events are being detected, at least at this location.

5 Discussion

We have demonstrated that transient strain rates estimated with the method described in Section 2 can be used to detect SSEs and will be robust to detect other transient geophysical phenomena. Another potential application would be to use GNSS derived transient strain rates to develop noise models for borehole strain meters (BSMs). The Plate Boundary Observatory maintains 43 BSMs in the Pacific Northwest, and it has been demonstrated that BSMs are able to record transient geophysical events such as SSEs (e.g., Dragert and Wang, 2011). However, noise in BSM data is not well understood, which complicates the use of BSM data for detecting geophysical events. The noise consists, in part, of a long-term decay resulting from the instrument equilibrating with the surrounding rock (Gladwin et al., 1987). Typically, this noise is dealt with in an ad-hoc manner by fitting and removing exponentials and low-order polynomials. A more rigorous quantification of BSM noise could be performed by analyzing the
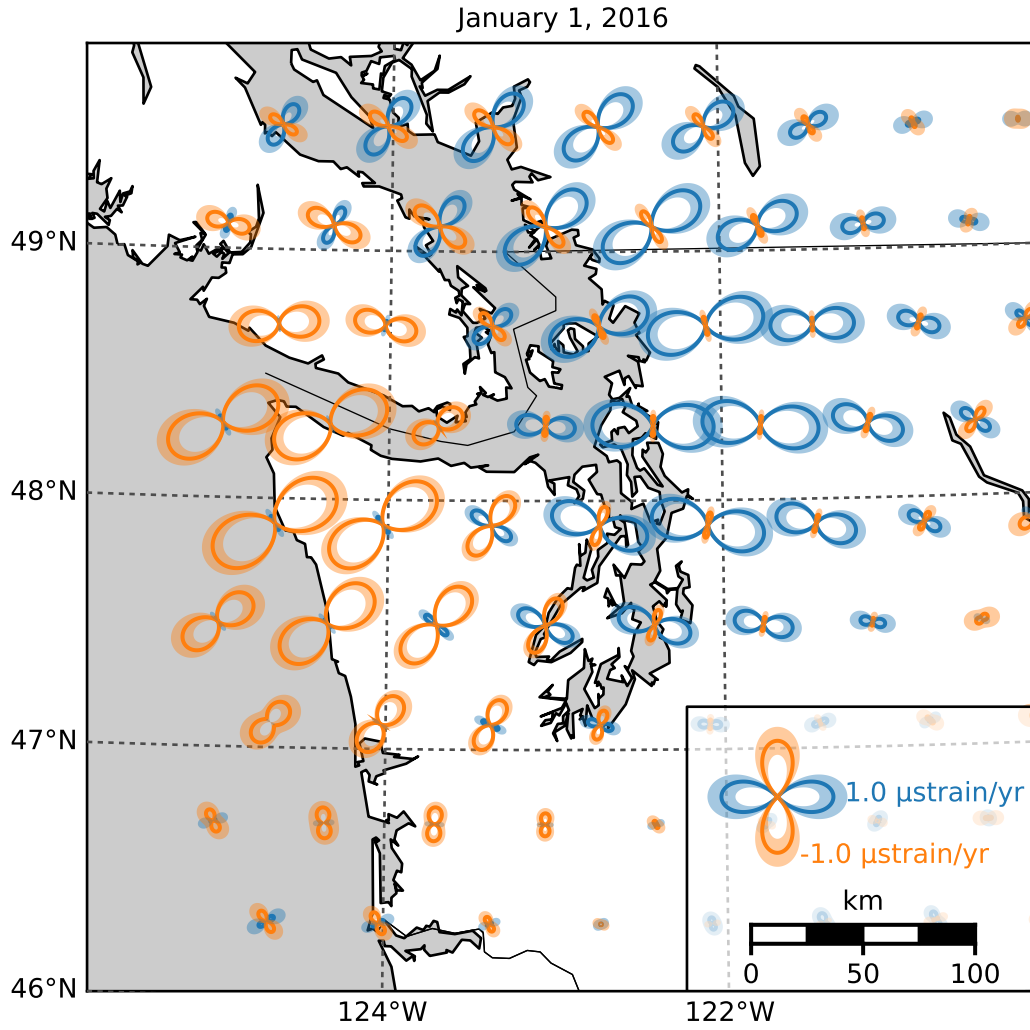
Figure 6: Estimated transient strain rates during the Winter 2015-2016 SSE. Strain glyphs show the normal strain rate along each azimuth, where orange indicates compression and blue indicates extension. The shaded regions indicate one standard deviation uncertainties in the normal strain rates.
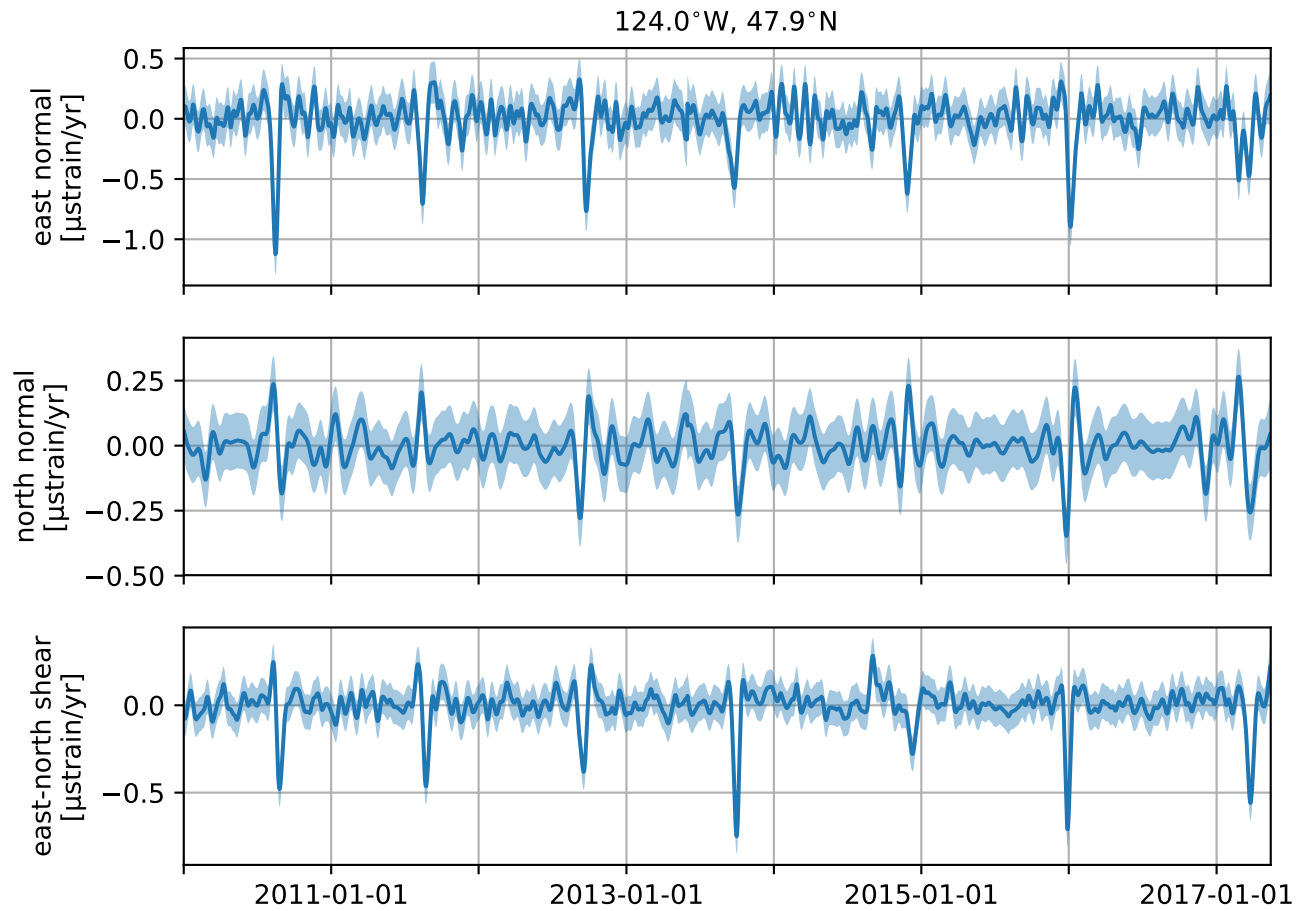
Figure 7: Three components of the transient horizontal strain rate tensor estimated at the position shown in Figure 1. The shaded regions indicate one standard deviation uncertainty.
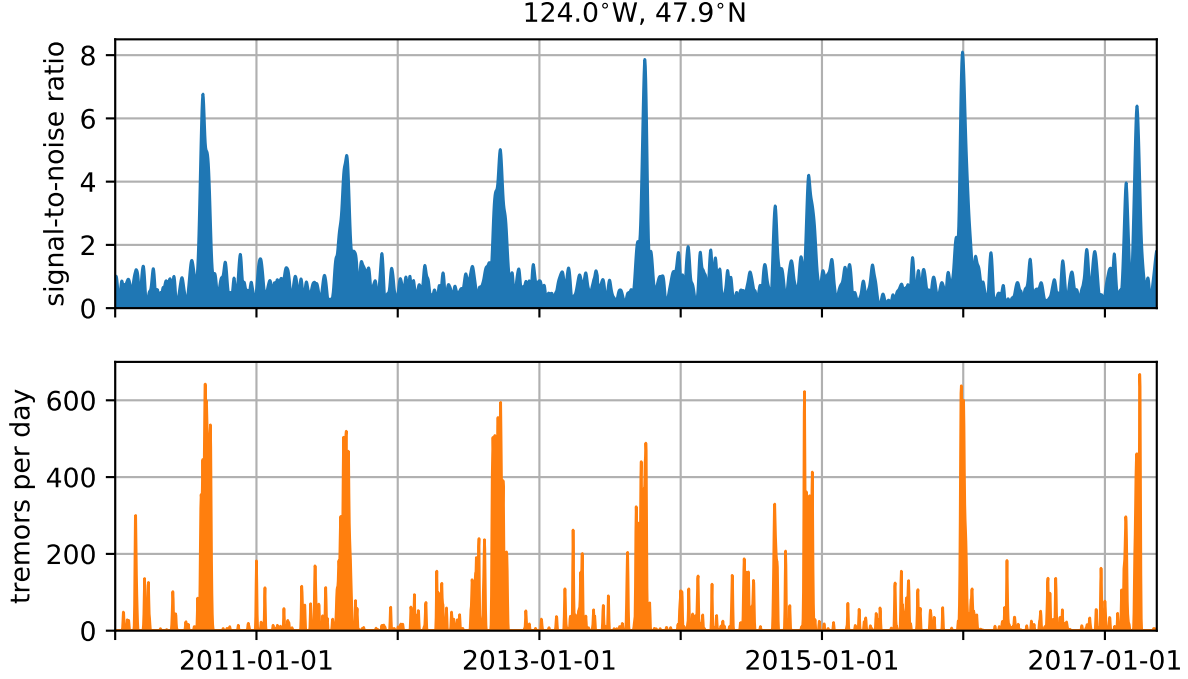
Figure 8: (top) Signal-to-noise ratio (eq. 15) at the position shown in Figure 1. (bottom) Frequency of tremors in the region shown in Figure 1.

residuals between the the BSM data and GNSS derived strain rates.

There is potential for a more thorough analysis of the spatio-temporal noise in GNSS data, η, than what was performed in Section 4.1. In this study, we did not explored the spatial covariance of η. Spatially correlated common mode noise, which results from factors such as reference frame error, is a non-negligible component of GNSS data. We ignore common mode error in this study because transients strain rates are insensitive to it. For other geophysical studies based on GNSS data, such as fault slip inversions, it may be necessary to incorporate a spatially covarying noise model (e.g., Miyazaki et al., 2003). We can also improve upon the seasonal model used in this study, which consists of four spatially uncorrelated sinusoids for each station. We did not explore the spatial covariance of seasonal deformation or the temporal "roughness" (i.e., the number of sinusoids needed to describe the observations). The periodic Gaussian process (Mackay, 1998) is an alternative model for seasonal deformation and is well suited for exploring the roughness of seasonal deformation. The periodic Gaussian process has zero mean and the covariance function

$$T(t,t') = \phi \exp\left(\frac{-\sin(\pi|t - t'|)^2}{2\tau^2}\right). \tag{31}$$

Realizations have annual periodicity and the roughness is controlled by τ. Decreasing τ has the same effect as including higher frequency sinusoids in the seasonal model. The optimal value for τ can be found with the REML method as described in Section 4.1.

The transient strain rates estimated in this study are constrained by about seven years of daily displacement observations from 94 GNSS stations. It can be computationally intensive to evaluate eqs. (10) and (11) for a dataset with this size. We significantly reduce the amount of memory needed to estimate transient strain rates by describing the temporal covariance of displacements with a compact Wendland covariance function. Using a compact covariance function for our prior turns eqs. (10) and (11) into sparse systems of equations, which we then solve with CHOLMOD. CHOLMOD is designed for solving sparse, positive definite systems of equations. The matrix being inverted in eqs. (10) and (11) is not positive definite; however, we can use another partitioned matrix inversion identity from Press et al. (2007) to partition it into positive definite submatrices to be inverted. Even when using a compact covariance function, it may still be necessary to reduce the computational burden by dividing the data into subsets and evaluating transient strain rates for each subset.

6 Conclusion

In this paper we propose using Gaussian process regression (GPR) to estimate transient strain rates from GNSS data. Most other methods for estimating strain rates assume a parametric representation of deformation, which can bias the results if the parameterization is not chosen carefully. Here we assume a stochastic, rather than parametric, prior model for displacements. Our prior model describes how much we expect transient displacements to covary spatially and temporally. If we know nothing about the underlying signal that we are trying to recover, then the prior model can be chosen objectively with maximum likelihood methods. Because GPR is a Bayesian method, the uncertainties on our estimated transient strain rates are well quantified, allowing one to discern geophysical signal from noise. We demonstrate that GPR is an effective tool for detecting geophysical phenomena, such as slow slip events, in our application to GNSS data from Cascadia. One limitation with GPR is that it is not robust against outliers. To overcome this limitation, we have introduced an effective pre-processing method for identifying and removing outliers from GNSS datasets. Another complication with GPR is that it usually involves inverting a dense matrix where the number of rows and columns is equal to the number of observations. This is prohibitive when using several years of daily GNSS observations from a network of several hundred stations. We significantly reduce the computational burden of GPR by using compact Wendland covariance function to describe our prior model. While this paper just focuses on estimating transient strain rates, we believe that GPR is a powerful tool that can be applied to a wide range of geophysical problems.

7 Acknowledgements

References

Beavan, J. and Haines, J. (2001). Contemporary horizontal velocity and strain rate fields of the Pacific-Australian plate boundary zone through New Zealand. Journal of Geophysical Research, 106(B1):741–770.

Blewitt, G., Kreemer, C., Hammond, W. C., and Gazeaux, J. (2016). MIDAS robust trend estimator for accurate GPS station velocities without step detection. Journal of Geophysical Research : Solid Earth, 121:2054–2068.

Chen, Y., Davis, T. a., and Hager, W. W. (2008). Algorithm 887 : CHOLMOD , Supernodal Sparse Cholesky Factorization and Update/downdate. ACM Transactions on Mathematical Software, 35(3):1–12.

Cressie, N. (1993). Statistics for Spatial Data. John Wiley & Sons, New York, rev. edition.

Dong, D., Fang, P., Bock, Y., Cheng, M. K., and Miyazaki, S. (2002). Anatomy of apparent seasonal variations from GPS-derived site position time series. J. Geophys. Res., 107(B4):2075.

Dong, D., Fang, P., Bock, Y., Webb, F., Prawirodirdjo, L., Kedar, S., and Jamason, P. (2006). Spatiotemporal filtering using principal component analysis and Karhunen-Loeve expansion approaches for regional GPS network analysis. Journal of Geophysical Research: Solid Earth, 111(3):1–16.

Dragert, G., Wang, K., and James, T. S. (2001). A silent slip event on the deeper Cascadia subduction interface. Science, 292:1525–1528.

Dragert, H. and Wang, K. (2011). Temporal evolution of an episodic tremor and slip event along the northern Cascadia margin. Journal of Geophysical Research: Solid Earth, 116(12):1–12.

El-Fiky, G. S. and Kato, T. (1999). Continuous distribution of the horizontal strain in the Tohoku district, Japan, predicted by least-squares collocation. Journal of Geodynamics, 27(2):213–236.

Feigl, K. L., King, R. W., and Jordan, T. H. (1990). Geodetic measurement of tectonic deformation in the Santa Maria Fold and Thrust Belt, California. Journal of Geophysical Research: Solid Earth, 95(B3):2679–2699.

Field, E. H., Arrowsmith, R. J., Biasi, G. P., Bird, P., Dawson, T. E., Felzer, K. R., Jackson, D. D., Johnson, K. M., Jordan, T. H., Madden, C., Michael, A. J., Milner, K. R., Page, M. T., Parsons, T., Powers, P. M., Shaw, B. E., Thatcher, W. R., Weldon, R. J., and Zeng, Y. (2014). Uniform California Earthquake Rupture Forecast, version 3 (UCERF3) -The time-independent model. Bulletin of the Seismological Society of America, 104(3):1122–1180.

Freed, a. M. and Lin, J. (2001). Delayed triggering of the 1999 Hector Mine earthquake by viscoelastic stress transfer. Nature.

Gibbs, B. P. (2011). Advanced Kalman Filtering, Least-Squares and Modeling: A Practical Handbook. John Wiley & Sons, Hoboken, NJ.

Gladwin, M. T., Gwyther, R. L., Hart, R., Francis, M., and Johnston, M. J. S. (1987). Borehole tensor strain measurements in California. Journal of Geophysical Research: Solid Earth, 92(B8):7981–7988.

Harville, D. A. (1974). Bayesian Inference for Variance Components Using Only Error Contrasts. Biometrika, 61(2):383–385.

Herring, T. A., Melbourne, T. I., Murray, M. H., Floyd, M. A., Szeliga, W. M., King, R. W., Phillips, D. A., Puskas, C. M., Santillan, M., and Wang, L. (2016). Plate Boundary Observatory and related networks: GPS data analysis methods and geodetic product. Reviews of Geophysics, pages 1–50.

Hines, T. T. and Hetland, E. A. (2016). Rapid and simultaneous estimation of fault slip and heterogeneous lithospheric viscosity from post-seismic deformation. Geophysical Journal International, 204(1):569–582.

Hines, T. T. and Hetland, E. A. (2017). Unbiased characterization of noise in geodetic data. submitted to Journal of Geodesy.

Holt, W. E. and Shcherbenko, G. (2013). Toward a Continuous Monitoring of the Horizontal Displacement Gradient Tensor Field in Southern California Using cGPS Observations from Plate Boundary Observatory (PBO). Seismological Research Letters, 84(3):455–467.

Kato, T., El-Fiky, G. S., Oware, E. N., and Miyazaki, S. (1998). Crustal strains in the Japanese islands as deduced from dense GPS array. Geophysical Research Letters, 25(18):3445–3448.

Langbein, J. (2004). Noise in two-color electronic distance meter measurements revisited. Journal of Geophysical Research: Solid Earth, 109(4):1–16.

Langbein, J. (2008). Noise in GPS displacement measurements from Southern California and Southern Nevada. Journal of Geophysical Research: Solid Earth, 113(5):1–12.

Langbein, J. and Johnson, H. (1997). Correlated errors in geodetic time series: Implications for time-dependent deformation. Journal of Geophysical Research, 102(B1):591–603.

Lisowski, M., Dzurisin, D., Denlinger, R. P., and Iwatsubo, E. Y. (2008). Analysis of GPS-Measured Deformation Associated with the 2004 2006 Dome-Building Eruption of Mout St. Helens, Washington. Technical Report September 1984.

Lohman, R. B. and Murray, J. R. (2013). The SCEC Geodetic Transient-Detection Validation Exercise. Seismological Research Letters, 84(3):419–425.

Mackay, D. J. C. (1998). Introduction to Gaussian processes. Neural Networks and Machine Learning, 168(1996):133–165.

Mao, A., Harrison, G. A., and Dixon, H. (1999). Noise in GPS coordinate time series. Journal of Geophysical Research, 104(B2):2797–2816.

McCaffrey, R., King, R. W., Payne, S. J., and Lancaster, M. (2013). Active tectonics of northwestern U.S. inferred from GPS-derived surface velocities. Journal of Geophysical Research: Solid Earth, 118:709–723.

McCaffrey, R., Qamar, A. I., King, R. W., Wells, R., Khazaradze, G., Williams, C. A., Stevens, C. W., Vollick, J. J., and Zwick, P. C. (2007). Fault locking, block rotation and crustal deformation in the Pacific Northwest. Geophysical Journal International, 169(3):1315–1340.

McGuire, J. J. and Segall, P. (2003). Imaging of aseismic fault slip transients recorded by dense geodetic networks. Geophysical Journal International, 155:778–788.

Meade, B. J. and Hager, B. H. (2005). Block models of crustal motion in southern California constrained by GPS measurements. Journal of Geophysical Research: Solid Earth, 110:1–19.

Miyazaki, S., McGuire, J. J., and Segall, P. (2003). A transient subduction zone slip episode in southwest Japan observed by the nationwide GPS array. Journal of Geophysical Research, 108(B2):1–15.

Murray, M. H. and Lisowski, M. (2000). Strain accumulation along the Cascadia subduction zone in western Washington. Geophysical Research Letters, 27(22):3631–3634.

Ohtani, R., McGuire, J. J., and Segall, P. (2010). Network strain filter: A new tool for monitoring and detecting transient deformation signals in GPS arrays. Journal of Geophysical Research: Solid Earth, 115(12):1–17.

Okada, Y. (1992). Internal deformation due to shear and tensile faults in a half space. Bulletin of the Seismological Society of America, 82(2):1018–1040.

Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T. (2007). Numerical Recipes: The Art of Scientific Computing. Cambridge University Press, Cambridge, 3 edition.

Rasmussen, C. E. and Williams, C. K. I. (2006). Gaussian processes for machine learning. The MIT Press.

Roeloffs, E. A. (2006). Evidence for Aseismic Deformation Rate Changes Prior To Earthquakes. Annual Review of Earth and Planetary Sciences, 34(1):591–627.

Rogers, G. and Dragert, H. (2003). Episodic tremor and slip on the Cascadia subduction zone: the chatter of silent slip. Science, 300:1942–1943.

Sandwell, D. T. and Wessel, P. (2016). Interpolation of 2-D vector data using constraints from elasticity. Geophysical Research Letters, pages 1–7.

Schmidt, D. A. and Gao, H. (2010). Source parameters and time-dependent slip distributions of slow slip events on the Cascadia subduction zone from 1998 to 2008. Journal of Geophysical Research: Solid Earth, 115(4):1–13.

Schwartz, S. Y. and Rokosky, J. M. (2007). Slow slip events and seismic tremor at circum-Pacific subduction zones. Reviews of Geophysics, 45:1–32.

Segall, P. and Mathews, M. (1997). Time dependent inversion of geodetic data. Journal of Geophysical Research, 102(B10):22391–22409.

Shen, Z., Wang, M., Zeng, Y., and Wang, F. (2015). Optimal Interpolation of Spatially Discretized Geodetic Data. Bulletin of the Seismological Society of America, 105(4):2117–2127.

Shen, Z. K., Jackson, D. D., Ge, B. X., and Bob, X. G. (1996). Crustal deformation across and beyond the Los Angeles basin from geodetic measurements. Journal of Geophysical Research, 101(B12):27927–27957.

Tape, C., Musé, P., Simons, M., Dong, D., and Webb, F. (2009). Multiscale estimation of GPS velocity fields. Geophysical Journal International, 179(2):945–971.

Wdowinski, S., Zhang, J., Fang, P., and Genrich, J. (1997). Southern California Permanent GPS Geodetic Array: Spatial filtering of daily positions for estimating coseismic and postseismic displacements induced by the 1992 Landers earthquake. 102(97):57–70.

Wech, A. G. (2010). Interactive Tremor Monitoring. Seismological Research Letters, 81(4):664 – 669.

Wech, A. G., Creager, K. C., and Melbourne, T. I. (2009). Seismic and geodetic constraints on Cascadia slow slip. Journal of Geophysical Research: Solid Earth, 114(10):1–9.

Wendland, H. (2005). Scattered data approximation.

Williams, S. D. P., Bock, Y., Fang, P., Jamason, P., Nikolaidis, R. M., Prawirodirdjo, L., Miller, M., and Johnson, D. J. (2004). Error analysis of continuous GPS position time series. Journal of Geophysical Research: Solid Earth, 109(B3).

Wyatt, F. (1982). Displacement of Surface Monuments: Horizontal Motion. <u>Journal of Geophysical Research</u>, 87(B2):979–989.

Wyatt, F. K. (1989). Displacement of surface monuments: Vertical motion. <u>Journal of Geophysical Research</u>, 94(B2):1655–1664.

Zhang, J., Bock, Y., Johnson, H., Fang, P., Williams, S., Genrich, J., Wdowinski, S., and Behr, J. (1997). Southern California Permanent GPS Geodetic Array: Error analysis of daily position estimates and site velocities. <u>Journal of Geophysical Research</u>, 102(B8):18035–18055.