Revealing Transient Strain in Geodetic Data with the Radial Basis Function Finite Difference Method

Trever T. Hines and Eric A. Hetland

December 2, 2016

1 Introduction

Crustal strain rates are fundamentally important quantities for assessing seismic hazard. This is because the locations where strain is rapidly accumulating are the locations where we can expect strain energy to be released seismically. It is then important to develop and improve upon methods for mapping strain in tectonically active regions because such maps could concievably feed into seismic hazard models such as UCERF3 (Field et al., 2014).

Methods for estimating strain from geodetic data fall in one of two catagories. There are model-based approaches which assume that strain is the result of loading on faults which have a known geometry, and there are data-based approaches which make no assumptions about the source of deformation. We will exclusively consider data-based approaches in this paper. The classic and simplest method for estimating strain is to assume that the strain rate is constant in time and spatially uniform within subnetworks of the geodetic data. Linear least squares is then used to find the components of the strain rate tensor for each subnetwork (e.g Frank, 1966; Prescott, 1976; Savage et al., 1986; Feigl et al., 1990; Murray and Lisowski, 2000). Several algorithms have been developed to improve upon this procedure. Shen et al. (1996) and Shen et al. (2015) discuss an algorithm where, instead of using the immediately adjacent stations to calculate strain at a position, the strain is computed with a weighted average over the entire network where the weighting is smaller for more distant stations. Another strategy is to fit a set of interpolating basis functions to the deformation field and then compute the strain from the analytical derivative of the interpolant (e.g. Beavan and Haines, 2001; Tape et al., 2009; Sandwell and Wessel, 2016).

The aforementioned studies have all been concerned with estimating long term strain rates. Time dependent strain would be useful for studying geophysical processes which occur over timescales of days to years such as slow slip events, postseismic relaxation, or volcanic deformation. Ohtani et al. (2010) identified transient strain events by fitting a set of spatial wavelet basis functions to the deformation field at discrete time epochs, and a Kalman filtering strategy was used to ensure that the coefficients for each basis function varied smoothly in time. Holt and Shcherbenko (2013) calculated time dependent strain by differentiating a bicubic interpolant that was fit to each epoch of a temporally smoothed deformation field.

Each of the methods described above are designed to overcome two complications that arise in estimating deformation gradients: (1) geodetic data are noisy and differentiation will only amplify the noise and (2) geodetic data are not observed on a regular grid, which prevents the use of standard finite difference methods for computing derivatives. In this paper we demonstrate that both of these complications can be elegantly handled with the Radial Basis Function-Finite Difference (RBF-FD) method.

The RBF-FD method was developed simultaneously and independently by Tolstykh and Shirobokov (2003), Shu et al. (2003), Cecil et al. (2004), and Wright and Fornberg (2006) as a computationally efficient way to solve large scale partial differential equations over irregular, multi-dimensional domains. The RBF-FD method can be thought of as a generalization of the traditional finite difference method, where the node layout is no longer restricted to regular grids. Indeed, the RBF-FD method can be used to estimate derivates of discrete data located at arbitrary scattered positions in multi-dimensional space. The RBF-FD method is particularly appealing because it is algorithmically simple, regardless of the domain shape or node layout, and also because the method has performed well in numerous benchmark tests (Fornberg and Flyer, 2015, and references therein).

In this paper, we do not use the RBF-FD method to solve a partial differential equation, but rather we use it to spatially smooth geodetic data and to compute deformation gradients. Our smoothing strategy can be viewed as a non-parametric, low-pass filter for scattered data where the degree of smoothness is controlled by a user specifies cutoff frequency. This can be contrasted with interpolation based methods where the resulting interpolant can be largely and perhaps unpredictably controlled by the choice of basis function. This process of smoothing and differentiating can be extended to estimate time dependent strain rates. In that case, we first temporally smooth and differentiate GPS displacement time series to get time dependent velocities. We

then spatially smooth and differentiate the resulting velocities for each time epoch to get time dependent strain rates.

The method proposed in this paper has numerous advantages which set it apart from other methods for computing strain rates. The method is computationally efficient and stable (there is no inversion of an ill-conditioned matrix). There are no hyper parameters or penalty parameters that need to be tuned for each application. As opposed to interpolation strategies such as Beavan and Haines (2001), Tape et al. (2009), or Ohtani et al. (2010), our method assumes that velocities are locally rather than globally continuous, which allows us to easily handle discontinuities resulting from, for example, a creeping fault.

We begin this paper by summarizing the RBF-FD scheme and explaining how we construct differentiation matrices for scattered data. We then introduce the RBF-FD filter, which is used to smooth the observed geodetic data prior to differentiation. We provide two real world demonstraitions of our method for calculating strain rates. First we calculate the long term strain rates in Southern California from the CMM3 velocity data set (Shen et al., 2011), and we verify that our results are consistent with other studies. We then calculate time dependent strain rates in Cascadia from the GPS data provided by UNAVCO. In Cascadia, we analyze strain resulting from slow slip events and compare it to the long term tectonic strain accumulation. Slow slip events are found to produce compression in the Olympic Peninsula, which is in addition to the compression resulting from tectonic loading. Further south in Oregon, the slow slip events tend to release the compressional strain that is accumulated tectonically. While similar conclusions have been drawn from fault slip inversions for slow slip events, it is important to recognize that slip inversion are the product of inverting an ill-conditioned matrix making it difficult to determine whether slip inferences are real or just an artifact of the inversion. The strain rates presented in this paper are more direct observations and can be interpretted with a higher degree of confidence.

2 Method

2.1 Differentiating Scattered Data

In this section we briefly summarizing the RBF-FD method and we refer the reader to Wright and Fornberg (2006) or Fornberg and Flyer (2015) for additional details. Consider a set of observation points $\mathbf{x} = \{x_i\}_{i=1}^{N}$ in $\mathbb{R}^d$ and a corresponding vector $\mathbf{u} = [u(x_1), ..., u(x_N)]^T$. Consider also a set of target points $\mathbf{y} = \{y_i\}_{i=1}^{M}$ in $\mathbb{R}^d$. We seek to find a differentiation matrix $\mathbf{L}$, such that $\mathbf{L}\mathbf{u} \approx [\mathcal{L}[u(x)] : x \in \mathbf{y}]^T$, where $\mathcal{L}$ is a differential operator. For each target point y_i, we approximate $\mathcal{L}[u(x)]\big|_{x=y_i}$ as a weighted sum of $\{u(x_j) : j \in \mathcal{S}^i\}$, where $\mathcal{S}^i$ consist of the indices for the n nearest points to y_i in $\mathbf{x}$. The approximation can be written as

$$\mathcal{L}[u(x)]\big|_{x=y_i} \approx \sum_{j \in \mathcal{S}^i} L_{ij} u(x_j). \tag{1}$$

We refer to the n nearest points to y_i in $\mathbf{x}$ as the stencil for y_i, and we denote the stencil as $\mathbf{x}^i = \{x_j : j \in \mathcal{S}^i\} = \{x_j^i\}_{j=1}^{n}$. The corresponding weights for each node in $\mathbf{x}^i$ are denoted as $\mathbf{w}^i = \{L_{ij} : j \in \mathcal{S}^i\} = \{w_j^i\}_{j=1}^{n}$. We can then equivalently write eq. (1) as

$$\mathcal{L}[u(x)]\big|_{x=y_i} \approx \sum_{j=1}^{n} w_j^i u(x_j^i). \tag{2}$$

Following Fornberg and Flyer (2015), we find the components of $\mathbf{w}^i$, and thus also of $\mathbf{L}$, by solving the linear system of equations

$$
\begin{bmatrix}
\phi(||x_1^i - x_1^i||) & \cdots & \phi(||x_n^i - x_1^i||) & \vdots & \psi_1(x_1^i) & \cdots & \psi_m(x_1^i) \\
\vdots & & \vdots & \vdots & \vdots & & \vdots \\
\phi(||x_1^i - x_n^i||) & \cdots & \phi(||x_n^i - x_n^i||) & \vdots & \psi_1(x_n^i) & \cdots & \psi_m(x_n^i) \\
\hdashline
\psi_1(x_1^i) & \cdots & \psi_1(x_n^i) & \vdots & 0 & \cdots & 0 \\
\vdots & & \vdots & \vdots & \vdots & & \vdots \\
\psi_m(x_1^i) & \cdots & \psi_m(x_n^i) & \vdots & 0 & \cdots & 0
\end{bmatrix}
\begin{bmatrix}
w_1^i \\
\vdots \\
w_n^i \\
\hdashline
\lambda_1 \\
\vdots \\
\lambda_m
\end{bmatrix}
=
\begin{bmatrix}
\mathcal{L}\left[\phi(||x - x_1^i||)\right]\big|_{x=y_i} \\
\vdots \\
\mathcal{L}\left[\phi(||x - x_n^i||)\right]\big|_{x=y_i} \\
\hdashline
\mathcal{L}\left[\psi_1(x)\right]\big|_{x=y_i} \\
\vdots \\
\mathcal{L}\left[\psi_m(x)\right]\big|_{x=y_i}
\end{bmatrix}
\tag{3}
$$

for each stencil. In eq. (3), ϕ is a radial basis function (RBF) which we describe below, $||\bullet||$ indicates the L_2 norm, $\{\psi_i\}_{i=1}^{m}$ are monomial basis functions that span the space of all d-dimensional polynomials with a
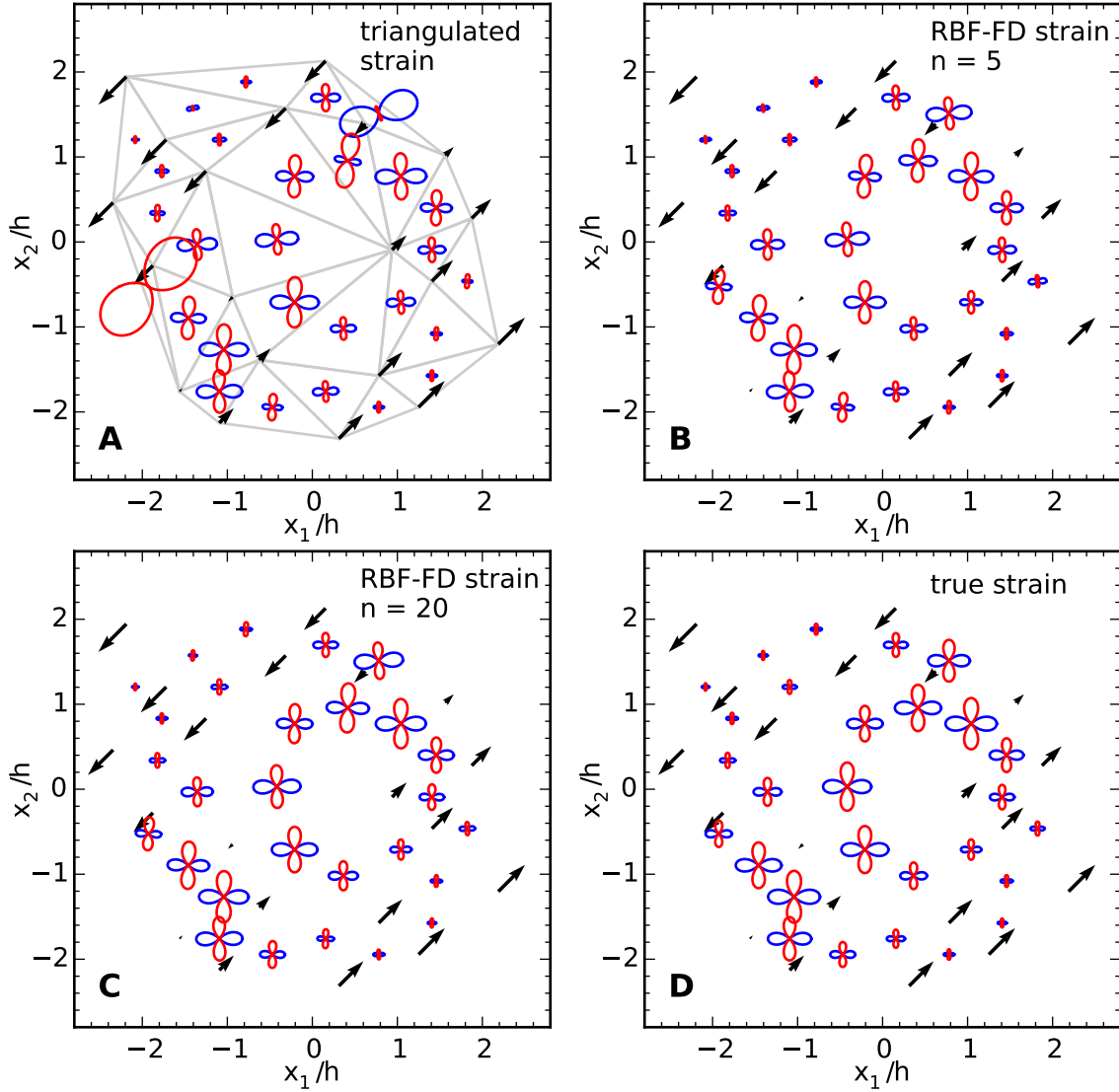
Figure 1: estimated synthetic strain

specified degree p (e.g. $\{1, x, y\}$ for $d = 2$ and $p = 1$), and $\{\lambda_i\}_{i=1}^{m}$ are parameters that are estimated along with $\mathbf{w}^i$ when eq. (3) is inverted but they serve no purpose and can be discarded.

Throughout this paper we use a cubic RBF for ϕ,

$$\phi(r) = r^3. \tag{4}$$

The cubic RBF is an odd degree polyharmonic spline which has the benefit of being scale invariant and thus there is no scaling parameter that needs to be optimized, unlike for many other common choices of RBFs (e.g. Larsson and Fornberg, 2003). We note that the results presented in this paper remain virtually unchanged when we use other polyharmonic splines for ϕ, which is consistent with the findings of Flyer et al. (2016).

We now elaborate on the stencil size, n, and the polynomial degree, p. We choose p to be equal to the degree of the derivative which we are approximating. This choice is based on the analysis of Flyer et al. (2016) and it ensures that eq. (1) will converge to the true derivative as the distance between nodes decreases. The accuracy of eq. (1) also generally improves with larger values of n, but at the expense of computational costs. We then choose n to be large enough for eq. (1) to converge. For the demonstrations in this paper, we find that $n = 30$ is an appropriate choice. It is worth noting that eq. (3) cannot be inverted when the number of nodes is less than the number of monomial basis functions, m. We then have a lower bound on n which is $n \geq m = \binom{p+d}{d}$.

In Figure 1, we demonstrate using the RBF-FD scheme to calculate strain from displacements. Our synthetic displacements are from an elastic, strike-slip, interseismic strain accumulation model (Savage and Burford, 1973). We use the RBF-FD scheme to compute the deformation gradients which are then used to obtain estimated strains (panel B and C). We also calculate strain with the classic method where we assume strain is uniform within each triangle of a Delaunay triangulation of the stations (e.g. Calais et al., 2002) (panel A). For ease

of comparison, strain is shown at the same positions for each panel in Figures 1. The strain estimated with the RBF-FD scheme is notably more accurate than the strain estimated by triangulation. This should be intuitivate because the RBF-FD scheme uses more data for each estimate. The strains shown in panel B and C are estimated with a stencil size of 5 and 20, respectively. As noted above, we can see that there is a point of diminishing returns when increasing the stencil size does not substantially improve the accuracy.

2.2 RBF-FD Filter

We calculate strain rates by spatially differentiation geodetically observed velocities. In order to differentiate noisy velocity data with the method described in Section 2.1, we must first smooth the data. Existing strategies for smoothing scattered data can be classified as parametric and non-parametric approaches. Parametric approaches involves fitting a set of basis functions to the data, perhaps using least-squares or regularized least-squares (e.g. Fasshauer, 2007). Non-parametric approaches include kernel smoothing (e.g. Hastie and Tibshirani, 1990) and Gaussian process regression (e.g. Rasmussen and Williams, 2006). Kriging is among the better known examples of Gaussian process regression (Matheron, 1963). Many of the wide variety of the existing strategies could surely produce a sufficiently smooth velocity field for us to calculate a coherent strain map. But since the intent of this paper is, in part, to demonstrate the utility of the RBF-FD method in geodesy, we present a smoothing strategy which we refer to as the RBF-FD filter. This method is non-parametric and it offers several features which we find valuable. The RBF-FD filter is computationally efficient and stable, it can be viewed as a low-pass filter with a well defined cutoff frequency, and we can easily specify discontinuities which we do not want to smooth across. In the following discussion of the RBF-FD filter, we seek to find a smoothed solution, $\mathbf{u}_{\mathrm{post}}$, from observed data, $\mathbf{u}_{\mathrm{obs}}$. We constrain $\mathbf{u}_{\mathrm{post}}$ with the observation equation

$$\mathbf{u}_{\mathrm{post}} = \mathbf{u}_{\mathrm{obs}} + \epsilon, \quad \epsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_{\mathrm{obs}}), \tag{5}$$

and the prior model

$$\mathbf{u}_{\mathrm{prior}} \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_{\mathrm{prior}}), \tag{6}$$

where ϵ and $\mathbf{u}_{\mathrm{prior}}$ are considered to be Gaussian processes with zero mean and covariances $\mathbf{C}_{\mathrm{obs}}$ and $\mathbf{C}_{\mathrm{prior}}$ respectively. The solution for $\mathbf{u}_{\mathrm{post}}$ minimizes the objective function

$$||\mathbf{u}_{\mathrm{post}} - \mathbf{u}_{\mathrm{obs}}||^2_{\mathbf{C}_{\mathrm{obs}}} + ||\mathbf{u}_{\mathrm{post}}||^2_{\mathbf{C}_{\mathrm{prior}}} \tag{7}$$

and is itself a Gaussian process with a distribution described by

$$\mathbf{u}_{\mathrm{post}} \sim \mathcal{N}(\bar{\mathbf{u}}_{\mathrm{post}}, \mathbf{C}_{\mathrm{post}}). \tag{8}$$

We use $\bar{\mathbf{u}}_{\mathrm{post}}$ and $\mathbf{C}_{\mathrm{post}}$ to denote the mean and covariance of $\mathbf{u}_{\mathrm{post}}$ respectively. Using Bayesian linear regression (Tarantola, 2005) these values are found to be

$$\begin{aligned} \bar{\mathbf{u}}_{\mathrm{post}} &= (\mathbf{C}_{\mathrm{obs}}^{-1} + \mathbf{C}_{\mathrm{prior}}^{-1})^{-1} \mathbf{C}_{\mathrm{obs}}^{-1} \mathbf{u}_{\mathrm{obs}} \\ \mathbf{C}_{\mathrm{post}} &= (\mathbf{C}_{\mathrm{obs}}^{-1} + \mathbf{C}_{\mathrm{prior}}^{-1})^{-1}. \end{aligned} \tag{9}$$

$\mathbf{C}_{\mathrm{obs}}$ is presumably well known, while $\mathbf{C}_{\mathrm{prior}}$ needs to be chosen based on an understanding of the underlying signal which we are trying to estimate. In Section 2.2.1 we discuss our choice for $\mathbf{C}_{\mathrm{prior}}$ and provide demonstrations with one-dimensional data. The natural extension for $\mathbf{C}_{\mathrm{prior}}$ when dealing with d-dimensional data is discussed in Section 2.2.2.

2.2.1 Filtering in One Dimension

For one-dimensional data we consider a prior which can be stated implicitly as

$$\mathbf{D}_n \mathbf{u}_{\mathrm{prior}} = \mathbf{q}, \quad \mathbf{q} \sim \mathcal{N}(0, \lambda^2), \tag{10}$$

where $\mathbf{D}_n$ is an n'th order differentiation matrix, and $\mathbf{q}$ is a vector of white noise with constant variance λ^2. We leave the details of $\mathbf{D}_n$ unspecified for now. The inverse of the prior covariance matrix is

$$\mathbf{C}_{\mathrm{prior}}^{-1} = \lambda^2 \mathbf{D}_n^T \mathbf{D}_n, \tag{11}$$

and the mean and covariance for the posterior are then

$$\begin{aligned} \bar{\mathbf{u}}_{\mathrm{post}} &= (\mathbf{C}_{\mathrm{obs}}^{-1} + \frac{1}{\lambda^2}\mathbf{D}_n^T\mathbf{D}_n)^{-1}\mathbf{C}_{\mathrm{obs}}^{-1}\mathbf{u}_{\mathrm{obs}} \\ \mathbf{C}_{\mathrm{post}} &= (\mathbf{C}_{\mathrm{obs}}^{-1} + \frac{1}{\lambda^2}\mathbf{D}_n^T\mathbf{D}_n)^{-1}. \end{aligned} \tag{12}$$

This solution is closely tied to several well established methods of smoothing. For example, one can immediately recognize eq. (12) as an example of Tikhonov regularization (Tikhonov and Arsenin, 1978). We also note a similarity between eq. (12) and smoothing splines (Wahba, 1990). To see this similarity, recall that a one-dimensional smoothing spline is defined as the function, $f(t)$, which minimizes,

$$\sum_{i=1}^{N} \left((\mathbf{u}_{\text{obs}})_i - f(t_i) \right)^2 + \alpha \int_{t_1}^{t_N} \left(f^{(n)}(t) \right)^2 dt, \tag{13}$$

where $(\mathbf{u}_{\text{obs}})_i$ is an observation at time t_i, N is the number of observations, α is a smoothing parameter, and $f^{(n)}$ denotes the n'th time derivative of f. In comparison, if we ignore data uncertainties (i.e. $\mathbf{C}_{\text{obs}} = \mathbf{I}$), $\bar{\mathbf{u}}_{\text{post}}$ is the discrete function which minimizes

$$||\mathbf{u}_{\text{obs}} - \bar{\mathbf{u}}_{\text{post}}||_2^2 + \frac{1}{\lambda^2} ||\mathbf{D}_n \bar{\mathbf{u}}_{\text{post}}||_2^2. \tag{14}$$

If the sampling rate for $\mathbf{u}_{\text{obs}}$ is constant and the penalty parameters are appropriately chosen then eq. (14) can be recognized as a discretized form of (13). We would thus expect $f(t)$ and $\bar{\mathbf{u}}_{\text{post}}$ to be effectively identical. The key difference between smoothing splines and the method presented here is that the former produces a globally smooth interpolant while we are only requiring local smoothness in our solution. The benefits of imposing local smoothness are discussed in Section 2.3.

We discuss the penalty parameter, λ. One common method for choosing an appropriate penalty parameter is generalized cross-validation (Craven and Wahba, 1979), which yields a smoothed solution with the maximum predictive power. There is merit to using an entirely objective approach such as cross-validation, and this would be appropriate if there is no prior knowledge of the signal's characteristic period. Otherwise, it may be better to chose a penalty parameter that damps out all the high frequency oscillations which are known to be noise.

We demonstrate how λ can be chosen so that frequencies greater than ω_c are attenuated. We make the simplifying assumption that eq. (12) is a linear time-invariant (LTI) filter. In doing so, we assume that $\mathbf{u}_{\text{obs}}$ has a constant sampling rate, the data uncertainty is constant and uncorrelated ($\mathbf{C}_{\text{obs}} = \sigma \mathbf{I}$), and $\mathbf{u}_{\text{prior}}$ is periodic. We enforce periodicity in $\mathbf{u}_{\text{prior}}$ by making $\mathbf{D}_n$ the periodic spectral differentiation matrix (e.g. Trefethen, 2000). Under the discrete Fourier transform, $\mathbf{D}_n$ then has the properties

$$\mathcal{F}[\mathbf{D}_n \mathbf{g}]_k = (2\pi i \omega_k)^n \mathcal{F}[\mathbf{g}]_k \tag{15}$$

and

$$\mathcal{F}[\mathbf{D}_n^T \mathbf{g}]_k = (-2\pi i \omega_k)^n \mathcal{F}[\mathbf{g}]_k, \tag{16}$$

where ω_k is the frequency domain variable and $\mathbf{g}$ is an arbitrary vector. With the LTI assumptions, the discrete Fourier transform of $\bar{\mathbf{u}}_{\text{post}}$ is

$$\mathcal{F}[\bar{\mathbf{u}}_{\text{post}}]_k = \frac{\frac{1}{\sigma^2}}{\frac{1}{\sigma^2} + \frac{(2\pi\omega_k)^{2n}}{\lambda^2}} \mathcal{F}[\mathbf{u}_{\text{obs}}]_k. \tag{17}$$

We make the change of variables

$$\lambda^2 = (2\pi\omega_c)^{2n} \sigma^2 \tag{18}$$

which simplifies eq. (17) to

$$\mathcal{F}[\bar{\mathbf{u}}_{\text{post}}]_k = \frac{1}{1 + \left(\frac{\omega_k}{\omega_c}\right)^{2n}} \mathcal{F}[\mathbf{u}_{\text{obs}}]_k. \tag{19}$$

Based on eq. (19), we can recognize the frequency response of eq. (12) to be qualitatively similar to the frequency response for an n'th order low-pass Butterworth filter with cut-off frequency ω_c. In particular, the frequency response for eq. (12) is flat in the passband and decays log-linearly in the stopband (Figure 2). In the limit as $n \to \infty$, eq. (12) becomes an ideal low-pass filter which removes all frequencies above ω_c and leaves lower frequencies unaltered.

The above Fourier analysis reveals how eq. (12) will behave under LTI assumptions. For the GPS data considered in this paper, this assumption is not appropriate because neither the data uncertainties nor the sampling rates are constant. When the LTI conditions are not satisfied, it is no longer obvious how eq. (12) behaves or whether it can still be viewed as a low-pass filter with cutoff frequency ω_c.

We explore the behavior of eq. (12) under more general conditions. We continue to substitute λ in eq. (12) with eq. (18), which changes our free parameter to ω_c. Since we are no longer assuming a constant variance, we replace σ^2 in eq. (18) with a characteristic variance, $\bar{\sigma}^2$, which we define as
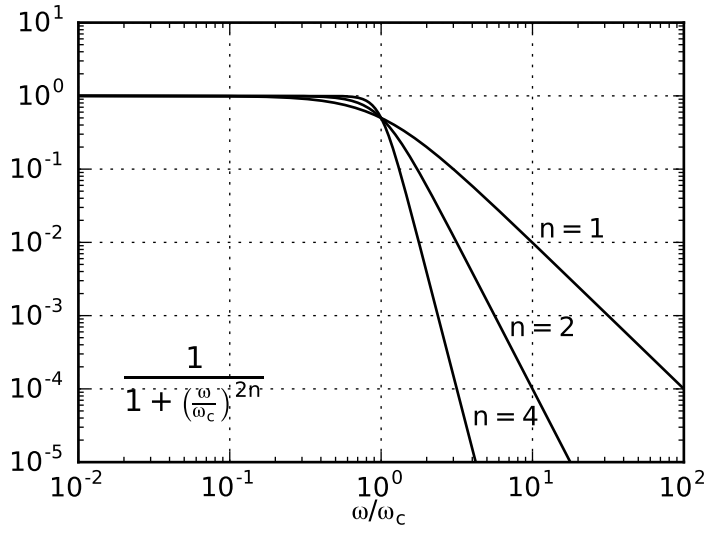
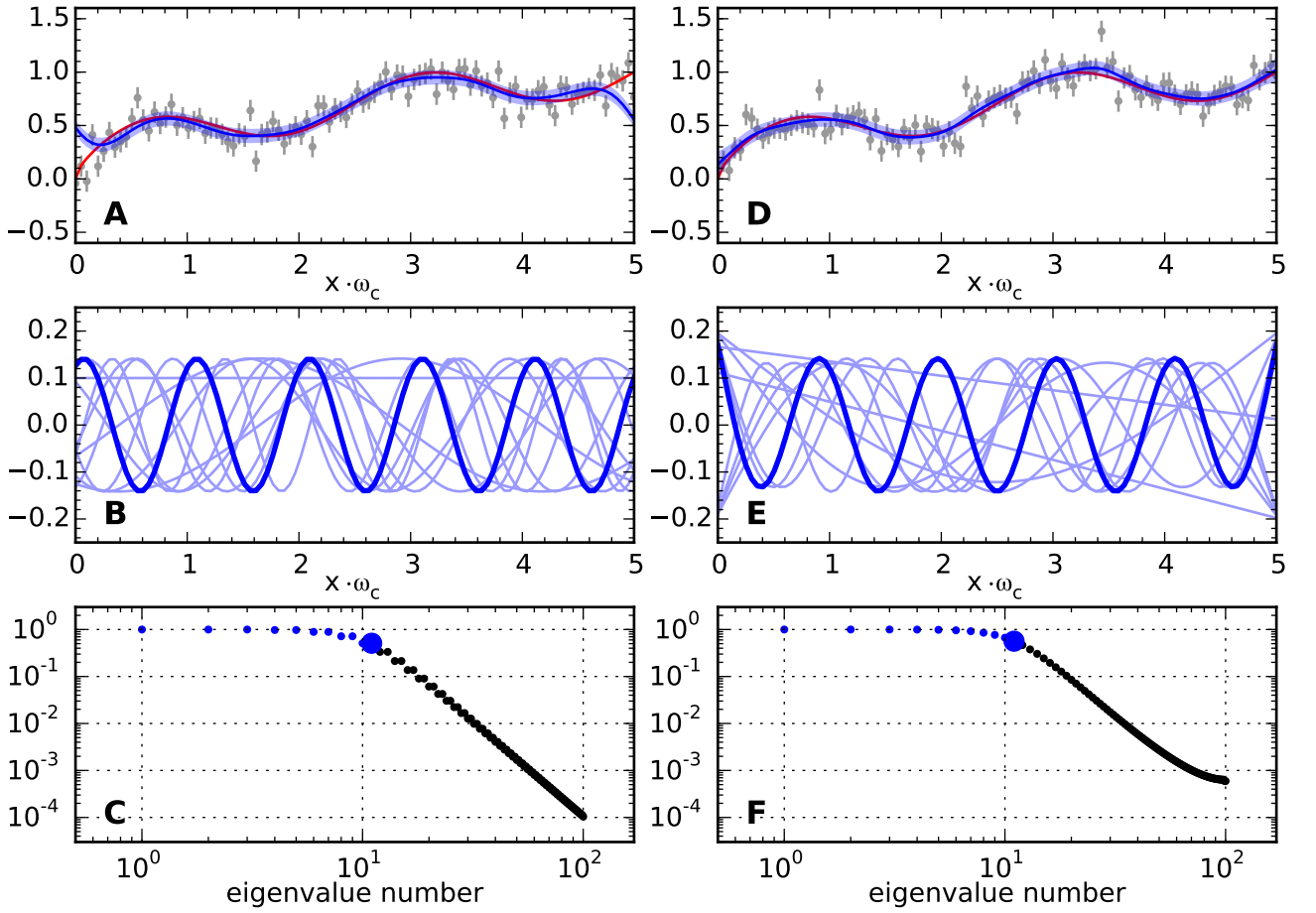Figure 2: Frequency response of the RBF-FD filter from eq. (19) for different values of n.

Figure 3: eigenvectors and eigenvalues

$$\frac{1}{\overline{\sigma^2}} = \frac{1}{N}\mathrm{tr}\left(\mathbf{C}_{\mathrm{obs}}^{-1}\right),\tag{20}$$

where N is the number of observations. We now write our smoothed solution as

$$\bar{\mathbf{u}}_{\mathrm{post}} = (\mathbf{C}_{\mathrm{obs}}^{-1} + \frac{1}{(2\pi\omega_c)^{2n}\bar{\sigma}^2}\mathbf{D}_n^T\mathbf{D}_n)^{-1}\mathbf{C}_{\mathrm{obs}}^{-1}\mathbf{u}_{\mathrm{obs}}$$

$$\mathbf{C}_{\mathrm{post}} = (\mathbf{C}_{\mathrm{obs}}^{-1} + \frac{1}{(2\pi\omega_c)^{2n}\bar{\sigma}^2}\mathbf{D}_n^T\mathbf{D}_n)^{-1}.\tag{21}$$

The behavior of eq. (21) can be revealed by analyzing the eigen decomposition of the matrix mapping $\mathbf{u}_{\mathrm{obs}}$ to $\bar{\mathbf{u}}_{\mathrm{post}}$,

$$\mathbf{K} = (\mathbf{C}_{\mathrm{obs}}^{-1} + \frac{1}{(2\pi\omega_c)^{2n}\bar{\sigma}^2}\mathbf{D}_n^T\mathbf{D}_n)^{-1}\mathbf{C}_{\mathrm{obs}}^{-1}.\tag{22}$$

The eigenvalues of $\mathbf{K}$, $\{s_1,\ldots,s_N\}$, are real and bounded between 0 and 1. The eigenvectors, $\{\mathbf{v}_1,\ldots,\mathbf{v}_N\}$, are real and, when $\mathbf{K}$ is symmetric (e.g. when $\mathbf{C}_{\mathrm{obs}} = \sigma^2\mathbf{I}$), form an orthogonal basis set. Each eigenvalue, s_i, describes the amount that $\mathbf{v}_i$ will be shrunk under the mapping $\mathbf{K}$. The eigenvectors associated with eigenvalues close to 1 can then be interpretted as components which are retained in $\bar{\mathbf{u}}_{\mathrm{post}}$. When the LTI conditions are satisfied, the above Fourier analysis reveals that $\{\mathbf{v}_1,\ldots,\mathbf{v}_N\}$ are a set of orthogonal sinusoids and $\{s_1,\ldots s_N\}$ are the coresponding frequency responses from eq. (19) (Figure 3A-C). We show the eigen decomposition of $\mathbf{K}$ under non-LTI condtions where (1) $\mathbf{u}_{\mathrm{prior}}$ is aperiodic (Figure 3D-F), (2) $\mathbf{u}_{\mathrm{prior}}$ is aperiodic and the data uncertainties are not constant (Figure 4A-C), and (3) $\mathbf{u}_{\mathrm{prior}}$ is aperiodic and the data sampling rate is not constant (Figure 4D-F). For case (1) and (3), the eigenvectors resemble sinusoids, except at the boundaries, and the eigenvalues decay in a manner that is consistent with the frequency response function in eq. (19). Namely, eq. (19) tells us that the eigenvectors associated with eigenvalues $\geq 1/2$ are sinusoids with frequency $\leq \omega_c$. For case (2), we provide a demonstration where the data uncertainties are anomalously high from the time interval 0.2-0.5. When looking outside of this time interval, the eigenvalues and eigenvectors are again consistent with the frequency response from eq (19). Within the time interval 0.2-0.5, the eigenvectors are elongated, which is desireable behavior because $\mathbf{u}_{\mathrm{post}}$ will not contort to fit the less certain observations. Since the frequency content of $\mathbf{u}_{\mathrm{post}}$ varies locally depending on the data uncertainty, we can only loosely view ω_c as a cutoff frequency. Nonetheless, we find in application that $1/\omega_c$ is a good predictor of the minimum feature wavelength retained in $\mathbf{u}_{\mathrm{post}}$, regardless of the data uncertainty or spacing.

2.2.2 Filtering in Higher Dimensions

We expand our discussion to smoothing data which is observed in d-dimensional space. We now consider the prior model

$$\mathbf{L}_n\mathbf{u}_{\mathrm{prior}} = \mathbf{q},\quad \mathbf{q} \sim \mathcal{N}(0, (2\pi\omega_c)^{2n}\bar{\sigma}^2)\tag{23}$$

where $\mathbf{L}_n$ is a differentiation matrix which approximates the operation

$$\sum_{i=1}^d \frac{\partial^n}{\partial x_i^n}\tag{24}$$

and n is an even integer. The variance chosen for $\mathbf{q}$ was motivated by the analysis in Section 2.2.1. The mean and covariance of $\mathbf{u}_{\mathrm{post}}$ is now described by

$$\bar{\mathbf{u}}_{\mathrm{post}} = (\mathbf{C}_{\mathrm{obs}}^{-1} + \frac{1}{(2\pi\omega_c)^{2n}\bar{\sigma}^2}\mathbf{L}_n^T\mathbf{L}_n)^{-1}\mathbf{C}_{\mathrm{obs}}^{-1}\mathbf{u}_{\mathrm{obs}}$$

$$\mathbf{C}_{\mathrm{post}} = (\mathbf{C}_{\mathrm{obs}}^{-1} + \frac{1}{(2\pi\omega_c)^{2n}\bar{\sigma}^2}\mathbf{L}_n^T\mathbf{L}_n)^{-1}.\tag{25}$$

If we again assume that the observation are regularly spaced, have constant variance, and $\mathbf{L}_n$ is the corresponding spectral differentiation matrix, then the d-dimensional discrete Fourier transform of $\bar{\mathbf{u}}_{\mathrm{post}}$ is

$$\mathcal{F}\left[\bar{\mathbf{u}}_{\mathrm{post}}\right]_k = \frac{1}{1 + \left(\sum_{i=1}^d \left(\frac{\omega_{i,k}}{\omega_c}\right)^n\right)^2}\mathcal{F}\left[\mathbf{u}_{\mathrm{obs}}\right]_k.\tag{26}$$
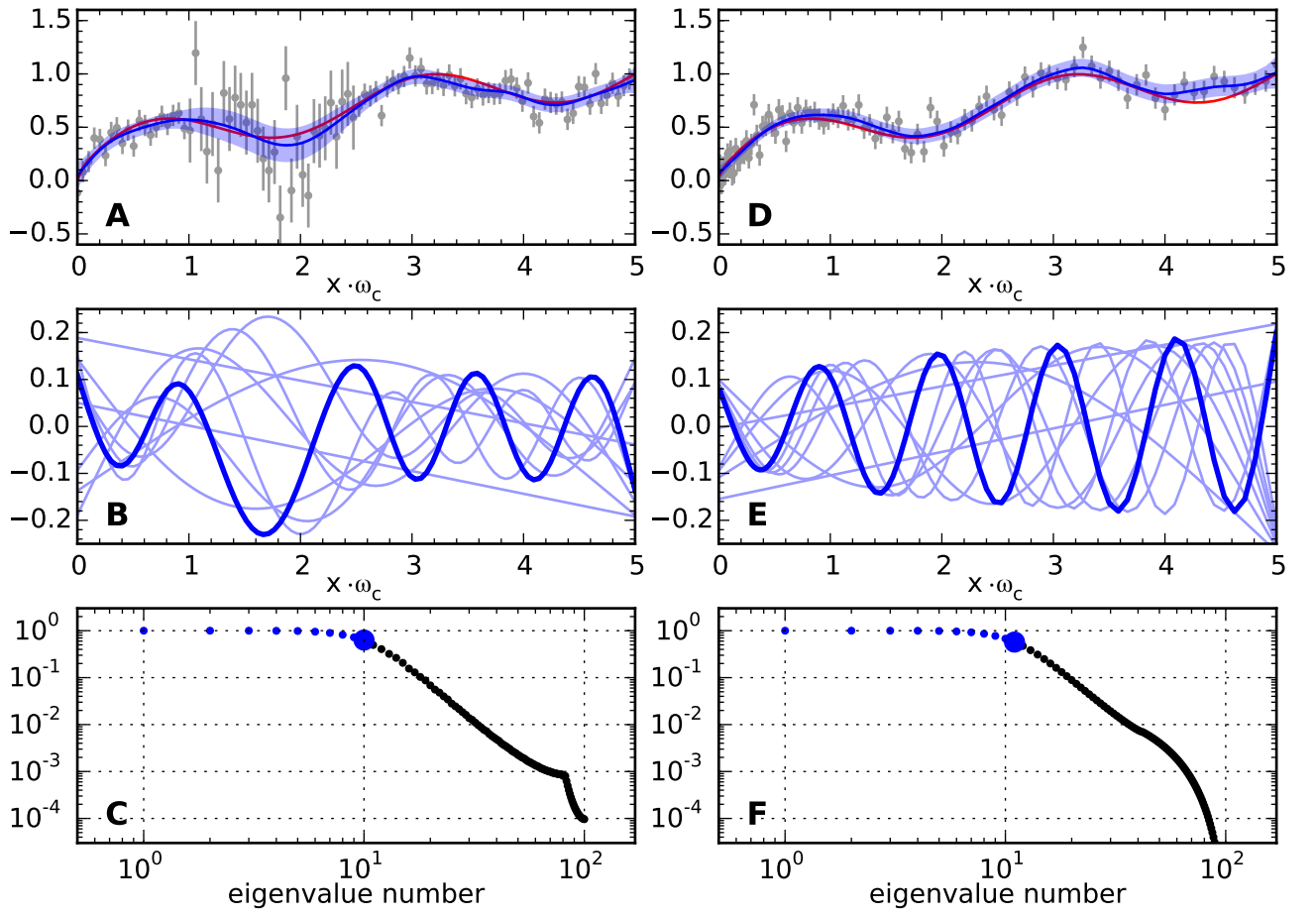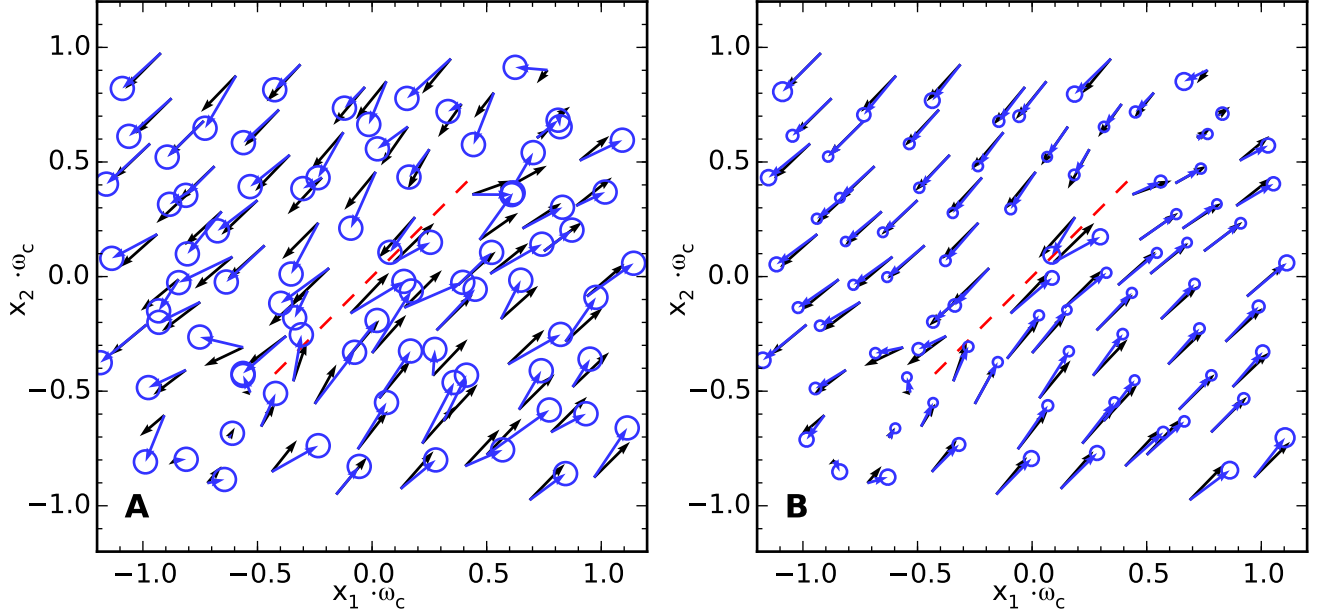
Figure 4: eigenvectors and eigenvalues

Figure 5: eigenvectors and eigenvalues

The frequency response for eq. (25) can once again be recognized as a low-pass filter with cutoff frequency ω_c. In the limit as $n \to \infty$, the frequency response becomes a d-dimensional box which is zero for all the frequency tuples $(\omega_{1,k}, \ldots \omega_{d,k})$ which have at least one component whos magnitude is greater than ω_c.

In general, $\mathbf{L}_n$ can be constructed with the RBF-FD scheme described in Section 2.1, which allows us to use eq. (25) to smooth irregularly spaced data. When that data are irregularly spaced, or if any of the above mentioned idealized conditions are not satisfied, eq. (25) exhibits the same behavior noted in Section 2.2.1. In particular, it can be verified through eigen decomposition that the frequency content of $\bar{\mathbf{u}}_{\mathrm{post}}$ is consistent with ω_c, regardless of the spacing or uncertainty of $\mathbf{u}_{\mathrm{obs}}$.

2.3   Accounting for Known Discontinuities

The RBF-FD filter from Section 2.2 and the RBF-FD differentiation scheme from Section 2.1 can both account for known discontinuities from, for example, a creeping fault segment. When using the RBF-FD scheme to estimate the strain rate from a velocity field, it is important not to use stencils that includes stations which are on opposite sides of a creeping fault segment. Doing so could result in estimated strain rates which are erroneously high. The stencils should thus be constructed so that no two stations within a stencil form an edge which intersects the trace of the creeping segment. We also construct stencils in this manner when forming the differentiation matrix $\mathbf{L}_n$ from eq. (25). By doing so, the RBF-FD filter will not produce a solution that is incorrectly tapered across a known discontinuity.

We elaborate on the statistical implications for how we handle discontinuities for the RBF-FD filter. When $\mathbf{L}_n$ is constructed with the RBF-FD scheme, we can interpret $\mathbf{u}_{\mathrm{prior}}$ as a Gaussian Markov random field (e.g. Rue and Held, 2005). That is to say, if $\mathbf{u}_{\mathrm{prior}}$ is known at stations within a neighborhood of x_i then $\mathbf{u}_{\mathrm{prior}}$ at station x_i is conditionally independent of $\mathbf{u}_{\mathrm{prior}}$ outside of that neighborhood. The neighborhood consists of all stations which share a stencil with x_i. Our strategy for handling discontinuities thus imposes conditional independence between $\mathbf{u}_{\mathrm{prior}}$ at any two stations which form an edge that intersects the trace of a creeping fault segment.

We demonstrate smoothing synthetic, scattered, two-dimensional vector data where there is a known discontinuity which we do not want to smooth across (Figure 5). The synthetic data is intended to simulates GPS velocities near a creeping fault segment. We generate the synthetic data with the elastic dislocation solution provided by Okada (1992) and obscure the underlying signal with white noise. We smooth both directional components of the synthetic data independently with eq. (25). Figure 5A and 5B show the observed, $\mathbf{u}_{\mathrm{obs}}$, and smoothed, $\mathbf{u}_{\mathrm{post}}$, data compared to the true signal. As noted in Section 2.2.1, $\mathbf{u}_{\mathrm{post}}$ tends to have the highest uncertainty at the domain edges, and that is where $\mathbf{u}_{\mathrm{post}}$ does the poorest job at recovering the original signal. Nonetheless, almost every smoothed datum in $\mathbf{u}_{\mathrm{post}}$ is an improvement over $\mathbf{u}_{\mathrm{obs}}$ at representing the underlying signal.

9

2.4 Numerical Notes

Do not compute the inverse for upost!

This is well conditioned and lends itself well to iterative matrix solvers

The posterior covariance matrix is generally dense. If memory does not permit the covariance matrix to be computed and one is only interested in the the data uncertainties then one can find the uncertainties by...

The posterior uncertainty does not account for error in L

3 Applications

3.1 Strain Rate in Southern California

We provide a demonstration in which we estimate the long term strain rates in southern California from the SCEC crustal motion map (CMM4)(Shen et al., 2011). The crustal motion map provides an estimate of time-averaged horizontal velocities at over one thousand irregularly spaced points in sothern California. The CMM4 is constrained by three decades of geodetic data and velocities have been corrected for coseismic and postseismic motions. We independently smooth easting and northing components of the CMM4 velocities with the RBF-FD filter using a cutoff frequency $\omega_c = 0.01$ km^{-1}. We do not smooth the velocities across the creeping segment of the San Andreas Fault (SAF), which extends from Parkfield to Hollister. We show a comparison between the CMM4 velocities and smoothed velocities as supplementary material. We then differentiate the smoothed velocities with the RBF-FD scheme to find the velocity gradient and strain rates. The estimated strain rates, complete with propagated uncertainties, are shown in Figure 6, and it accurately illuminates well known features. The strain map highlights the SAF and shows that the orientation of maximum shear strain runs parallel with the SAF strike (Lisowski et al., 1991). Shear strain rates along the SAF are particularly high $(0.3 - 0.4$ micro-strain per year) in the Salton Trough and at the ends of the creeping segment. We can observe N-S contraction in the Los Angeles basin at a rate of $0.1 - 0.2$ micro-strain per year (Shen et al., 1996; Argus et al., 1999, 2005). We also observe a diffuse band of shear strain along the Eastern California shear zone (cite). These results are also consistent with the strain map by Shen et al. (2015), which was derived from the same data set.

The uncertainties on our strain estimates are shown as semi-transparent fields around the strain glyphs, and it may appear that our inferred strain rates are overly confident. In particular, the uncertainties on the estimated strain rates in Nevada are surprisingly low, given the sparse station coverage. There are two sources of error which could potentially result in underestimated uncertainties: (1) we could have chosen a cutoff frequency that is too low for the geophysical signal that we are interested in and (2) the data coverage may be too sparse for the differentiation matrices to be accurate numerical approximations of the corresponding differential operators. The former is a well recognized bias inherent in any smoothing method (Hastie and Tibshirani, 1990), and we address the latter here. Ideally, we would want the average distance between adjacent stations to be $<1/\omega_c$. Such a dense coverage would ensure that the differentiation matrix used to estimate the velocity gradient is sufficiently accurate. In this demonstration, the data spacing in Nevada is comparable to $1/\omega_c = 100$ km. One remedy for the sparse coverage is to use the RBF-FD filter to interpolate the smoothed solution onto a finer grid of points and then find the velocity gradient of the denser smoothed data set. As suggested in Section 2.2, the RBF-FD filter can be used for interpolation by treating the interpolation points as data with infinite uncertainty. Doing so will increase the accuracy of velocity gradient estimates, but at the expense of increased computational cost. Figure 7A shows a map of estimated strain rates where we improved the accuracy of velocity gradient estimate by interpolating the smoothed velocities on a grid with a ~5 km spacing. The inferred strain rates at points along the SAF, where the coverage is sufficiently dense, do not appear to be significantly improved by the added interpolation points. The uncertainties on the strain rates in Nevada are much higher, and likely more appropriate, when we included the interpolation points.

In Figure 7B, we highligh the creeping segment of the SAF, which is a notably problematic area when trying to estimate strain rates in southern California (e.g. Sandwell and Wessel, 2016). We find that the strain rate is greatest at the two ends of the creeping segment, near Parkfield and Hollister. We are also able to resolve a quadrant pattern of extension and compression at the ends of the creeping segment. This strain pattern is consistent with what would be expected from an elastic dislocation model, giving us confidence that the inferred strain rates are accurate.

3.2 Time Depenent Strain Rate in Cascadia

Dragert et al. (2001) first discovered slow slip events

Slow slip event depths from Dragert et al. (2001), Wech et al. (2009), Schmidt and Gao (2010), Bartlow et al. (2011) show slip concentrated at depths from 30 to 50 km detph.
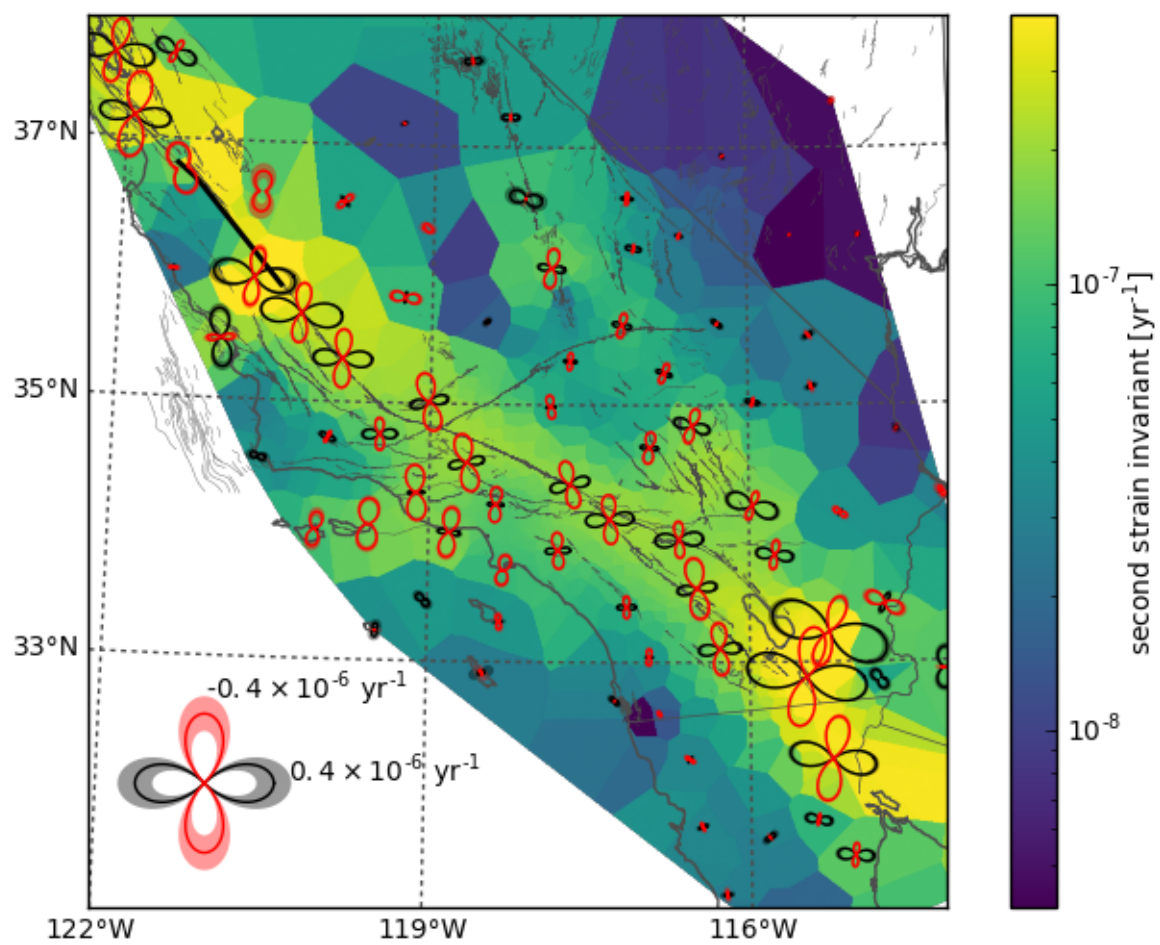
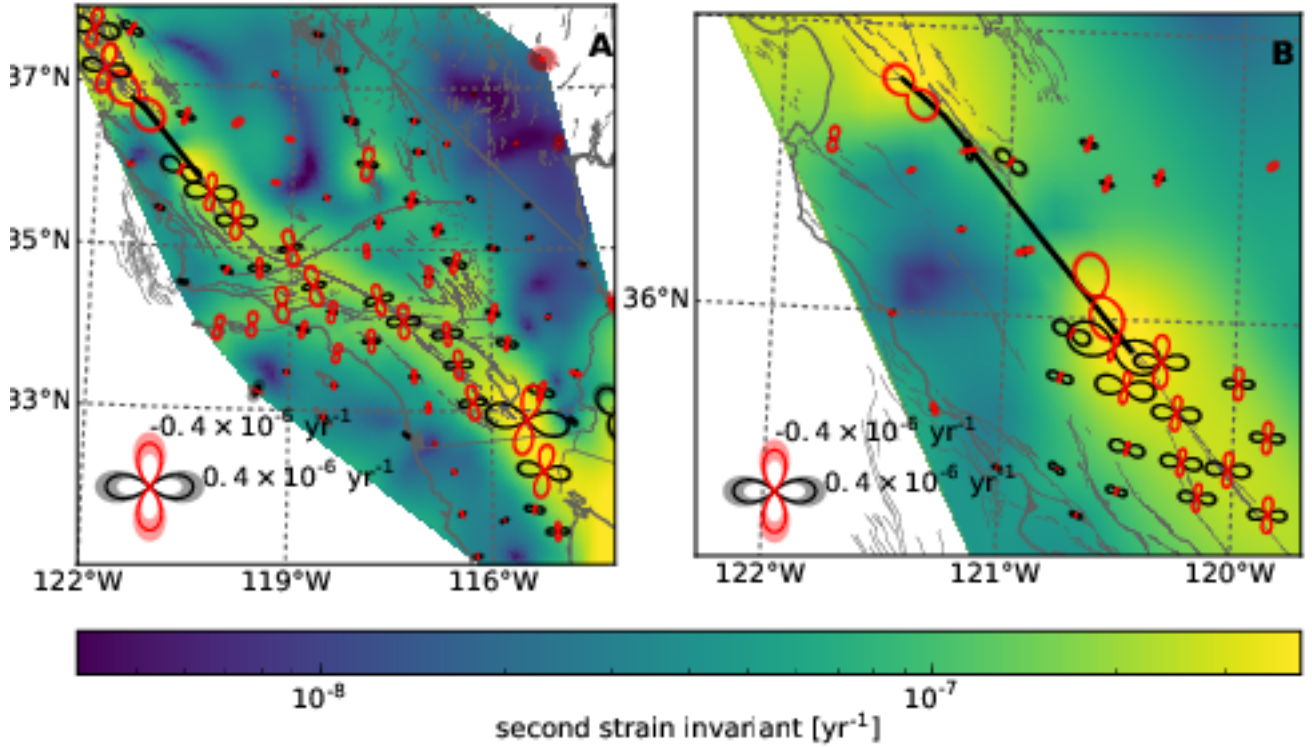Figure 6: Southern California strain map

Figure 7: Southern California strain map

Interseismic locking depths from Flück et al. (1997), Murray and Lisowski (2000), McCaffrey et al. (2007) and McCaffrey et al. (2013), Burgette et al. (2009), Schmalzle et al. (2014) are consistent with a full locking down to about 20 km.

Studies which simultaneously modeled interseismic and ets are Holtkamp and Brudzinski (2010) and Schmalzle et al. (2014)

4 Discussion and Conclusion

References

Argus, D. F., Heflin, M. B., Donnellan, A., Webb, F. H., Dong, D., Hurst, K. J., Jefferson, D. C., Lyzenga, G. A., Watkins, M. M., and Zumberge, J. F. (1999). Shortening and thickening of metropolitan Los Angeles measured and inferred by using geodesy. *Geology*, 27(8):703–706.

Argus, D. F., Heflin, M. B., Peltzer, G., Crampé, F., and Webb, F. H. (2005). Interseismic strain accumulation and anthropogenic motion in metropolitan Los Angeles. *Journal of Geophysical Research B: Solid Earth*, 110(B04401):1–26.

Bartlow, N. M., Miyazaki, S., Bradley, A. M., and Segall, P. (2011). Spacetime correlation of slip and tremor during the 2009 Cascadia slow slip event. 38(L18309):1–6.

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.

Burgette, R. J., Weldon, R. J., and Schmidt, D. A. (2009). Interseismic uplift rates for western Oregon and along-strike variation in locking on the Cascadia subduction zone. *Journal of Geophysical Research: Solid Earth*, 114(1):1–24.

Calais, E., Nocquet, J. M., Jouanne, F., and Tardy, M. (2002). Current strain regime in the Western Alps from continuous Global Positioning System measurements, 1996-2001. *Geology*, 30(7):651–654.

Cecil, T., Qian, J., and Osher, S. (2004). Numerical methods for high dimensional Hamilton-Jacobi equations using radial basis functions. *Journal of Computational Physics*, 196(1):327–347.

Craven, P. and Wahba, G. (1979). Smoothing noisy data with spline functions: estimating the correct degree of smoothing by the method of generalized cross-validation. *Numerische Mathematik*, 403:377–403.

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

Fasshauer, G. E. (2007). *Meshfree Approximation Method with Matlab*. World Scientific Publishing Co., Singapore.

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.

Flück, P., Hyndman, R. D., and Wang, K. (1997). 3-D dislocation model for great earthquakes of the Cascadia subduction zone. *Journal of Geophysical Research*, 102(B9):20539–20550.

Flyer, N., Fornberg, B., Bayona, V., and Barnett, G. A. (2016). On the role of polynomials in RBF-FD approximations: I. Interpolation and accuracy. *Journal of Computational Physics*, 321:21–38.

Fornberg, B. and Flyer, N. (2015). *A Primer on Radial Basis Functions with Applications to the Geosciences*. Society for Industrial and Applied Mathematics, Philadelphia.

Frank, C. F. (1966). Deduction of earth strains from survey data. *Bulletin of the Seismological Society of America*, 56(1):35–42.

Hastie, T. J. and Tibshirani, R. J. (1990). *Generalized Additive Models*, volume 1. Chapman and Hall, London.

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.

Holtkamp, S. and Brudzinski, M. R. (2010). Determination of slow slip episodes and strain accumulation along the Cascadia margin. *Journal of Geophysical Research: Solid Earth*, 115(4):1–21.

Larsson, E. and Fornberg, B. (2003). A numerical study of some radial basis function based solution methods for elliptic PDEs. *Computers and Mathematics with Applications*, 46:891–902.

Lisowski, M., Savage, J. C., and Prescott, W. H. (1991). The velocity field along the San Andreas Fault in central and southern California. *Journal of Geophysical Research*, 96(B5):8369.

Matheron, G. (1963). Principles of geostatistics. *Economic Geology*, 58(8):1246–1266.

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.

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.

Prescott, W. H. (1976). An extension of Frank's method for obtaining crustal shear strains from survey data. *Bulletin of the Seismological Society of America*, 66(6):1847–1853.

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

Rue, H. and Held, L. (2005). *Gaussian Markov Random Fields: Theory and Applications.* Chapman and Hall, Boca Raton.

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

Savage, J. C. and Burford, R. O. (1973). Geodetic determination of relative plate motion in central California. *Journal of Geophysical Research*, 78(5):832–845.

Savage, J. C., Prescott, W. H., and Gu, G. (1986). Strain accumulation in southern California, 19731984. *Journal of Geophysical Research*, 91(B7):7455–7473.

Schmalzle, G. M., McCaffrey, R., and Creager, K. C. (2014). Central Cascadia subduction zone creep. *Geochemistry, Geophysics, Geosystems*, pages 1515–1532.

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.

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.

Shen, Z. K., King, R. W., Agnew, D. C., Wang, M., Herring, T. A., Dong, D., and Fang, P. (2011). A unified analysis of crustal motion in Southern California, 1970-2004: The SCEC crustal motion map. *Journal of Geophysical Research: Solid Earth*, 116(11):1–19.

Shu, C., Ding, H., and Yeo, K. (2003). Local radial basis function-based differential quadrature method and its application to solve two-dimensional incompressible NavierStokes equations. *Computer Methods in Applied Mechanics and Engineering*, 192(7-8):941–954.

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.

Tarantola, A. (2005). *Inverse problem theory and methods for model parameter estimation.* SIAM.

Tikhonov, A. N. and Arsenin, V. Y. (1978). Solutions of Ill-Posed Problems . by A . N . Tikhonov ; V . Y . Arsenin Review by : John B . Bell Stable URL : http://www.jstor.org/stable/2006360 . *American Mathematical Society*, 32(144):1320–1322.

Tolstykh, A. I. and Shirobokov, D. A. (2003). On using radial basis functions in a "finite difference mode" with applications to elasticity problems. *Computational Mechanics*, 33(1):68–79.

Trefethen, L. N. (2000). *Spectral Methods in Matlab.* SIAM.

Wahba, G. (1990). *Spline Models for Observational Data.* Philadelphia.

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.

Wright, G. B. and Fornberg, B. (2006). Scattered node compact finite difference-type formulas generated from radial basis functions. *Journal of Computational Physics*, 212(1):99–123.