Geophys. J. Int. (2009) **179,** 945–971 doi: 10.1111/j.1365-246X.2009.04337.x

Multiscale estimation of GPS velocity fields

Carl Tape,[1] Pablo Musé,[1]* Mark Simons,[1] Danan Dong[2] and Frank Webb[2]

[1]*Seismological Laboratory, California Institute of Technology, Pasadena, CA, USA. E-mail: carltape@gps.caltech.edu*
[2]*Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA*

Accepted 2009 July 11. Received 2009 July 10; in original form 2009 March 9

SUMMARY

We present a spherical wavelet-based multiscale approach for estimating a spatial velocity field on the sphere from a set of irregularly spaced geodetic displacement observations. Because the adopted spherical wavelets are analytically differentiable, spatial gradient tensor quantities such as dilatation rate, strain rate and rotation rate can be directly computed using the same coefficients. In a series of synthetic and real examples, we illustrate the benefit of the multiscale approach, in particular, the inherent ability of the method to localize a given deformation field in space and scale as well as to detect outliers in the set of observations. This approach has the added benefit of being able to locally match the smallest resolved process to the local spatial density of observations, thereby both maximizing the amount of derived information while also allowing the comparison of derived quantities at the same scale but in different regions. We also consider the vertical component of the velocity field in our synthetic and real examples, showing that in some cases the spatial gradients of the vertical velocity field may constitute a significant part of the deformation. This formulation may be easily applied either regionally or globally and is ideally suited as the spatial parametrization used in any automatic time-dependent geodetic transient detector.

Key words: Wavelet transform; Satellite geodesy; Seismic cycle; Transient deformation; Kinematics of crustal and mantle deformation.

1 INTRODUCTION

Deformation of Earth's crust occurs at multiple length scales. For example, we might associate a scale of 5000 km with the signature of postglacial isostatic rebound of the crust (e.g. Walcott 1973; Simons & Hager 1997), a scale of 100 km might be appropriate for a given plate boundary, a scale of 10 km might be appropriate for local deformation of a basin due to hydrological variations (e.g. Argus *et al.* 2005), and a scale of 1 km might distinguish the signature of a shallowly locked fault from an aseismically slipping fault (e.g. Savage & Prescott 1976). The primary objective of this methodological study is to present a wavelet-based multiscale representation of three-component surface velocities, as a tool to facilitate analysis of geodetic observations from dense GPS networks.

We may categorize GPS-based studies of continental deformation based on 'physical' and 'non-physical' model parametrizations (Table 1). By 'physical,' we mean that the study imposes a physical description of the system, frequently in the form of elastic block models with specific sets of faults. 'Non-physical' parametrizations, such as the one in this paper, use convenient mathematical functions to estimate the velocity fields. Both approaches have their strengths. Our focus on the non-physical approach is driven by our desire to eventually detect time-dependent subtle signals of unknown origin in addition to 'steady-state' (or 'secular') plate motion.

Previous studies that have directly or indirectly estimated a velocity model from GPS data can be characterized as adopting a parametrization that is either variable in spatial scale or alternatively quasi-uniform in spatial scale, both of which differ from a multiscale approach. In a variable scale approach, the resolution of the model parametrization varies over the area of observations. Two examples of the variable scale approach include: (1) the classic estimation of strain from a triangulated network (e.g. Feigl *et al.* 1990), whereby the scale varies according to the distance between adjacent stations and (2) the estimation of slip rates on faults that separate rigid or elastic blocks (e.g. Meade *et al.* 2002), whereby the scale is tied to the geometry of the blocks adopted in the model. In a quasi-uniform scale approach, a single scale parameter is used, either based on the spacing within a nearly uniform grid of estimation points (e.g. Haines & Holt 1993; Beavan & Haines 2001) or based on explicit use of a parameter that weights observations based on distance from the estimation point (e.g. Shen *et al.* 1996; Ward 1998b; Hsu

*Now at: Department of Signal and Image Processing, Universidad de la República, Montevideo, Uruguay.

Table 1. Some examples of geodetic-based studies of crustal deformation.

	'Physical parametrization'	'Non-physical parametrization'
Time-independent	Meade *et al.* (2002) Meade & Hager (2005) Becker *et al.* (2005) McCaffrey *et al.* (2007)	Ward (1998b,a) Beavan & Haines (2001) Spakman & Nyst (2002) This study
Time-dependent	Heki (1997) Dragert *et al.* (2001) McGuire & Segall (2003)	Future study

et al. 2009). We use the term 'quasi-uniform' since the resolution of the velocity field (or strain-rate field) is still variable depending on the spatially variable density of observations. In contrast to all these approaches, in a multiscale parametrization the actual value of a given field at a given point in space is a superposition of values at different spatial scales. The multiscale parametrization naturally allows separation of processes that occur at different spatial scales.

From a set of discrete observations on the surface of Earth, we estimate a continuous multiscale vector velocity field using spherical wavelet functions. These functions have previously been used in characterizing other geophysical signals, such as gravitational and magnetic fields (e.g. Freeden & Windheuser 1997; Bayer *et al.* 2001; Holschneider *et al.* 2003; Chambodut *et al.* 2005). In comparison to previous approaches to the surface velocity or strain problem, novel aspects of our approach include: (1) an explicit and consistent decomposition of the velocity field into multiple scales at all locations; (2) a minimum scale at which we estimate the velocity field at a particular location that is controlled by the local station coverage; (3) inclusion of the vertical velocity observations if they are available and (4) use of spherical wavelets in representing the velocity field.

We are motivated to develop this multiscale representation of surface velocities in order to eventually incorporate it within a framework of a time-varying 'geodetic event detector.' Several studies have used GPS observations to detect instances of transient crustal deformation: for example, the subduction zones of Japan (Heki 1997; Ozawa *et al.* 2002; Miyazaki *et al.* 2003, 2004), Andes (Melbourne *et al.* 2002; Pritchard & Simons 2006) and Cascadia (Dragert *et al.* 2001; Rogers & Dragert 2003). These are the type of events we would like to be able to detect in an automated fashion, whereby we would have a characterization of both the temporal and spatial scale of the event. The decomposition of the velocity field into wavelets automatically focuses on locally correlated behaviour by relying on quantities inferred from multiple neighbouring sites and de-emphasizes the importance of individual stations. We anticipate that in the context of a transient event detector, our built-in emphasis on locally correlated behaviour will provide more sensitivity when looking for anomalous events, rather than using individual time-series for detection. As a corollary, some sources of noise in GPS data sets, such as reference frame error, will naturally fall into specific scales, thereby leaving other scales cleaner and more sensitive to transients.

In Section 2, we demonstrate the multiscale approach using spherical wavelets to decompose the velocity fields. The multiscale aspects are achieved by using wavelets from progressively finer meshes, but only where justifiable based on the local site density. From each estimated velocity field we can automatically compute maps of scalar quantities that are useful in interpreting crustal deformation, such as dilatation rate, strain rate, and rotation rate (Section 3). In Section 4, we present three synthetic examples to validate the approach, and then illustrate the method using the three-component NASA REASoN velocity field for southern California. Finally, in Section 5 we discuss implications for the technique and areas for possible extension and improvement.

2 ESTIMATION OF GPS VELOCITY FIELDS

We begin by introducing our choice of frame functions (defined below) for representing each component of a three-component velocity field on the sphere. A 'frame' is a set of functions that spans the whole space (here, of square-integrable functions), but are not necessarily linearly independent. The notion of frame is more general and hence less constrained than that of basis, thus providing more flexibility to represent signals. The redundancy of frames can be particularly useful when the goal is to approximate a signal with a small number of template functions. Our choice of frame functions is spherical wavelets. Any velocity field on the sphere can be represented as a linear combination of spherical wavelets. We describe the inverse problem of estimating the coefficients of the spherical wavelets describing the velocity field based on an irregularly distributed set of point observations. We adopt a least-squares approach that assumes Gaussian distributed data errors, but this assumption is not critical and is only chosen for convenience.

2.1 Fitting scattered data on the sphere

Fitting scattered data on the sphere is a long-standing problem that has been approached using local regression (Hastie & Loader 1993; Shen *et al.* 1996; Teza *et al.* 2008), kernel-based methods (Silverman 1986), statistics-based methods such as kriging (Matheron 1963; Reguzzoni *et al.* 2005), spline-based methods (Wahba 1981) or by estimating a smooth field using an expansion of basis functions (such as spherical harmonics), with a smoothing constraint applied to stabilize the inversion (Beavan & Haines 2001). Our approach falls into the last category, with the basis functions being a particular class of wavelets defined on the sphere.

We denote the time-dependent position of the ith GPS station in a network by $\mathbf{x}_i(t)$, whose origin is that of an external reference frame (e.g. ITRF2005 or fixed North America). Over an interval of time, Δt, the station moves $\Delta\mathbf{x}_i$, and thus the velocity of the station over this interval is $\mathbf{v}_i = \Delta\mathbf{x}_i/\Delta t$.

A velocity field defined on a sphere of radius R can be written as

$$\mathbf{v}(\theta,\phi) \;=\; v_r(\theta,\phi)\,\hat{\mathbf{r}} \;+\; v_\theta(\theta,\phi)\,\hat{\boldsymbol{\theta}} \;+\; v_\phi(\theta,\phi)\,\hat{\boldsymbol{\phi}}, \tag{1}$$

where $\hat{\mathbf{r}}$ is the vertical direction, $\hat{\boldsymbol{\theta}}$ is the south direction and $\hat{\boldsymbol{\phi}}$ is the east direction. Our goal is to estimate the field $\mathbf{v}(\theta,\phi)$ from a set of measurements taken at N irregularly distributed stations, with coordinates (θ_i, ϕ_i), $i = 1, \ldots, N$. These measurements, denoted by

$$\mathbf{v}^i = \left(v_r^i,\; v_\theta^i,\; v_\phi^i\right) \tag{2}$$

$$= [v_r(\theta_i,\,\phi_i),\; v_\theta(\theta_i,\,\phi_i),\; v_\phi(\theta_i,\,\phi_i)] \tag{3}$$

are assumed to be contaminated by noise. We are interested in functional descriptions of the field $\mathbf{v}(\theta,\phi)$ having the capability to adapt the scale locally in relation to the local density of the GPS stations. These descriptions should also allow decoupling of the velocity field at different spatial scales, since they may be associated with different physical phenomena or different sources of noise.

2.1.1 Wavelets on the sphere

Our desire to efficiently distinguish between multiple phenomena and noise sources that may be present in a given data set motivates the use of a spatio-spectral (or multiscale) representation (e.g. Simons & Hager 1997; Simons *et al.* 1997). Wavelets provide a natural representation for such multiscale signals. The formal extension of wavelet analysis to a spherical manifold has been an active research field in the last 15 yr (Dahlke & Maass 1996; Freeden & Windheuser 1996, 1997; Holschneider 1996; Simons & Hager 1997; Antoine & Vandergheynst 1999; Bogdanova *et al.* 2005; Wiaux *et al.* 2005). Applications in geophysics include analyses of Earth's magnetic field and gravity field (Holschneider *et al.* 2003; Chambodut *et al.* 2005). An application to the global temperature field is presented in Oh & Li (2004).

In order to define a wavelet analysis in the space $L^2(S^2)$ of square-integrable functions on the unit sphere, two basic operations are required: rotations (which play the role of translations in the Euclidean case), and dilatations. The major difficulty in deriving wavelets on the sphere lies in the definition of a proper dilatation operator, directly linked in wavelet theory to the notion of scale. Here we follow the approach proposed by Antoine & Vandergheynst (1999), whereby continuous spherical wavelets are derived from Euclidean wavelets by inverse stereographic projection. An alternative would be to use spherical wavelets constructed from spherical harmonics (e.g. Holschneider *et al.* 2003; Simons *et al.* 2006; Guilloux *et al.* 2009).

Bogdanova *et al.* (2005) proved that the inverse stereographic projection of any admissible 2-D wavelet in $\mathbb{R}^2$ yields an admissible wavelet on the sphere. This important result leads to a large variety of choices for spherical wavelets. In particular, in this work we consider the 'Difference of Gaussians' (DOG) spherical wavelet, which is an axisymmetric function. This spherical wavelet is obtained as the difference of the inverse stereographic projection of an isotropic Gaussian and of a dilated version of it. An example of a DOG spherical wavelet and its spectrum is shown in Fig. 1.

Rotation with Euler angles (ϕ, θ, α) of an axisymmetric function corresponds to centring the function at pole (θ, ϕ). Hence, by applying dilatations and rotations to a mother axisymmetric spherical wavelet ψ centred at, say the North Pole, the infinite set $\{\psi_{\mathbf{x},a}, \mathbf{x} \in S^2, a > 0\}$ of spherical wavelets with scale a, centred at pole $\mathbf{x}$, can be generated.

The decomposition of the velocity field on the sphere into a discrete set of spherical wavelets has to span the whole space $L^2(S^2)$, but does not need to be a linearly independent set. As described earlier, such a discrete set of not necessarily linearly independent functions, having the property of spanning the whole function space, is called 'frame'. We resolve the inherent non-uniqueness of the frame using a form of regularization. We refer the reader to Kovacevic & Chebira (2007a,b) for a detailed exposition of frame theory. Bogdanova *et al.* (2005) have shown that a discrete frame can be obtained by properly sampling the set of DOG spherical wavelets. This sampling requires discretization of both position, $\mathbf{x}$, and scale, a. The discretization of position amounts to defining a set of grids $\mathcal{G}_q$ on the sphere, for a set of resolutions q. We obtain each grid by successive subdivisions of an initial icosahedron $\mathcal{G}_0$, whereby each triangle is subdivided into four new equilateral triangles, and the new vertices are projected into the sphere. Three such meshes ($\mathcal{G}_2$, $\mathcal{G}_3$ and $\mathcal{G}_4$) are shown in Fig. 1(a). The discretization of the scale parameter depends on the set of resolutions chosen for the spherical grid. We provide details on the construction of the discrete frame of DOG spherical wavelets in Appendix A.

A frame of band-limited velocity fields on the sphere will then be a discrete and finite set of DOG spherical wavelets, of different scales, centred at gridpoints,

$$\mathcal{F}_{q_{\max}} = \left\{\psi_{\mathbf{x}_{(q,j)},a_q}(\mathbf{x}),\;\; \mathbf{x}_{(q,j)} \in \mathcal{G}_q,\;\; q = 0, \ldots, q_{\max}\right\}, \tag{4}$$

where $a_q = 2^{-q}$ denotes the discretized scale. Reasonable values of $q_{\max}$ for the fields we want to approximate range from 7 to 9, and correspond to a scale of 55 to 14 km, respectively, on the surface of Earth (see values of $\bar{L}_q$ in Table 2).

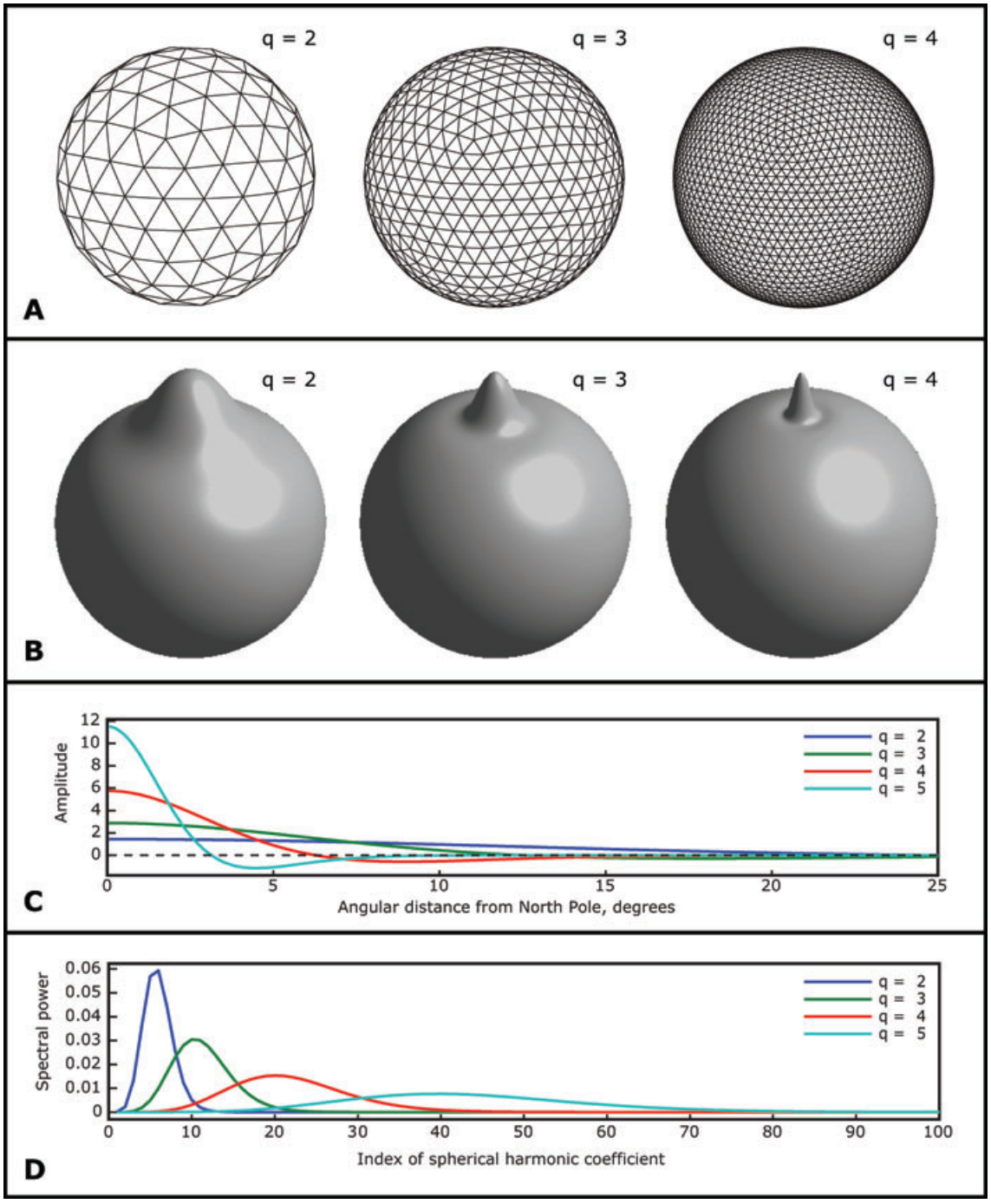

Figure 1. Spherical wavelet frame functions. (a) Triangulated spherical grids used for determining the locations for the centres of the spherical wavelet frame functions. From left- to right-hand side are grids for orders $q = 2$ (162 vertices), $q = 3$ (642 vertices) and $q = 4$ (2562 vertices) (Table 2). (b) Three different scales of a DOG (Difference of Gaussian) spherical wavelet centred at the North Pole. (c) Corresponding profiles of wavelets in (b), for a fixed longitude ϕ. (d) Corresponding spectra of wavelets in (b).

2.2 Decomposition of the velocity field in spherical wavelets

For the sake of clarity, we assume from now on that an ordering on the elements of $\mathcal{F}_{q_{\max}}$ has been established, and we write $\mathcal{F}_{q_{\max}} = \{g_k(\mathbf{x}), k = 1, \ldots, M\}$. Any scalar function $f \in L^2(S^2)$ whose bandwidth does not exceed the one associated to the finest scale $q_{\max}$ can be written as

$$f(\mathbf{x}) = \sum_{k=1}^{M} m_k \, g_k(\mathbf{x}) = \mathbf{g}^T(\mathbf{x})\,\mathbf{m}. \tag{5}$$

Table 2. Geometric properties of the spherical-triangular grids, after Wang & Dahlen (1995).

Order, q	Faces, F_q	Vertices, V_q	Average side arc-length,[a] $\bar{L}_q$		Max. representable bandwidth, $\bar{l}_q$
0	20	12	63.435°	7053.64 km	3
1	80	42	31.718°	3526.82 km	6
2	320	162	15.859°	1763.41 km	11
3	1 280	642	7.929°	881.71 km	23
4	5 120	2 562	3.965°	440.85 km	45
5	20 480	10 242	1.982°	220.43 km	91
6	81 920	40 962	0.991°	110.21 km	182
7	327 680	163 842	0.496°	55.11 km	363
8	1310 720	655 362	0.248°	27.55 km	726
9	5242 880	2621 442	0.124°	13.78 km	1453
10	20 971 520	10 485 762	0.062°	6.89 km	2906
11	83 886 080	41 943 042	0.031°	3.44 km	5811
12	335 544 320	167 772 162	0.016°	1.72 km	11 623
q	$F_q = 20 \times 4^q$	$V_q = 10 \times 4^q + 2$	$\bar{L}_q = 2^{-q}\,\bar{L}_0$		$\bar{l}_q = \text{int}(\pi/\bar{L}_q)$

[a] $\bar{L}_0 = \cos^{-1}\{\cos(72^\circ)/[1-\cos(72^\circ)]\} \approx 1.1071 \approx 63.43^\circ$.

Since the frame $\mathcal{F}_{q_{\max}}$ is not an orthogonal basis, the problem of finding optimal coefficients m_k is underdetermined and the solution is not unique. However, by including some form of regularization, a unique optimal representation can be found. We discuss the issue of regularization in more detail in the next section. The velocity vector field approximation problem can be reduced to three scalar estimation problems. If each of the three components is square-integrable, with bandwidth smaller than $\bar{l}_{q_{\max}}$, we can then write

$$\mathbf{v}(\theta,\phi) \;=\; \sum_{k=1}^{M}[a_k\, g_k(\theta,\phi)\hat{\mathbf{r}}(\theta,\phi) \;+\; b_k\, g_k(\theta,\phi)\,\hat{\boldsymbol{\theta}}(\theta,\phi) \;+\; c_k\, g_k(\theta,\phi)\,\hat{\boldsymbol{\phi}}(\theta,\phi)]. \tag{6}$$

The evaluation of (6) at the observation sites (θ_i,ϕ_i), $i = 1,\ldots,N$ (to which the noisy measurements $(v_r^i, v_\theta^i, v_\phi^i)$ have been associated), leads to

$$v_r^i = \sum_{k=1}^{M} a_k\, g_k(\theta_i,\phi_i) + n_r^i, \quad v_\theta^i = \sum_{k=1}^{M} b_k\, g_k(\theta_i,\phi_i) + n_\theta^i, \quad v_\phi^i = \sum_{k=1}^{M} c_k\, g_k(\theta_i,\phi_i) + n_\phi^i\ , \tag{7}$$

where n_r^i, n_θ^i and n_ϕ^i denote the measurement noise. Consequently, in the following we consider the estimation of a generic spherical scalar field $f(\mathbf{x})$ from N irregularly distributed sample points $\mathbf{x}_i$, $i = 1,\ldots,N$

$$f_i = f(\mathbf{x}_i) \;=\; \mathbf{g}^T(\mathbf{x}_i)\,\mathbf{m} \;+\; n_i. \tag{8}$$

Eq. (8) can be written in matrix form as

$$\mathbf{f} = \mathbf{G}\,\mathbf{m} \;+\; \mathbf{n}, \tag{9}$$

where $\mathbf{f} \in \mathbb{R}^N$ is the vector of measurements, $\mathbf{G} \in \mathbb{R}^{N\times M}$ is the design matrix with $G_{ik} = g_k(\mathbf{x}_i)$ and $\mathbf{m} \in \mathbb{R}^M$ is the model parameter vector.

2.3 The inverse problem, including uncertainties and regularization

We estimate the model vector, $\mathbf{m}$, by minimizing the regularized least-squares functional

$$F(\mathbf{m}) \;=\; \frac{1}{2}\,(\mathbf{G}\mathbf{m}-\mathbf{d})^T\,\mathbf{C}_{\mathrm{D}}^{-1}\,(\mathbf{G}\mathbf{m}-\mathbf{d}) + \frac{1}{2}\lambda^2\mathbf{m}^T\mathbf{S}\mathbf{m}. \tag{10}$$

The observations are assumed to be contaminated by a Gaussian noise of zero mean and data covariance matrix $\mathbf{C}_{\mathrm{D}}$. Each diagonal element of the data covariance matrix is the variance associated with each observation. Thus, the first term in (10) is the data misfit, and the second term is the model regularization, with regularization parameter λ.

Typical choices of regularization for when model $\mathbf{m}$ represents a spatial quantity include the norm of the model, the norm of the model gradient, or the Laplacian of the model. In our case, we regularize according to the norm of the gradient of the estimated model. The

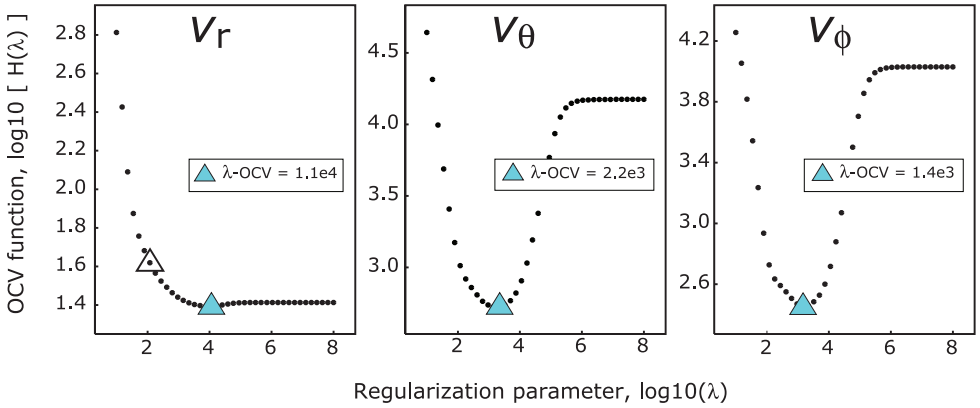

Figure 2. Selection of regularization parameter (Appendix B), illustrated for the example in Section 4.2. Each curve is the ordinary cross-validation (OCV) function, $H(\lambda)$ (eq. B11), for a component of the velocity field (v_r, v_θ, v_ϕ). The minimum of the curve (solid triangle) provides the regularization parameter. The $H(\lambda)$ curve for v_r does exhibit a minimum, though it is barely apparent with the y-axis limits shown here. For this particular example, we chose $\lambda = 100$ (open triangle) for estimating v_r (Section 4.2.2).

regularization matrix $\mathbf{S} \in \mathbb{R}^{M\times M}$ is then derived as follows:

$$\begin{aligned}
\|\nabla f(\theta,\phi)\|_2^2 &= \int_{S^2} \nabla f(\theta,\phi)^T\, \nabla f(\theta,\phi)\, \mathrm{d}\Omega \\
&= \int_{S^2} \nabla \left[\sum_{k=1}^{M} m_k\, g_k(\theta,\phi)\right]^T \nabla \left[\sum_{k'=1}^{M} m_{k'}\, g_{k'}(\theta,\phi)\right] \mathrm{d}\Omega \\
&= \sum_{k=1}^{M} m_k \sum_{k'=1}^{M} m_{k'} \int_{S^2} \nabla g_k(\theta,\phi)^T\, \nabla g_{k'}(\theta,\phi)\, \mathrm{d}\Omega \\
&= \sum_{k=1}^{M} m_k \sum_{k'=1}^{M} m_{k'} S_{kk'} \\
&= \mathbf{m}^T \mathbf{S} \mathbf{m},
\end{aligned} \tag{11}$$

where we have set

$$S_{kk'} = \int_{S^2} \nabla g_k(\theta,\phi)^T\, \nabla g_{k'}(\theta,\phi)\, \mathrm{d}\Omega. \tag{12}$$

The parameter λ controls the trade-off between the data fit and smoothness of the solution. The effect of minimizing the regularization term is to force energy into the longer-scale wavelet coefficients, since a smaller wavelet scale will have a larger gradient norm and thus a larger corresponding matrix element $S_{kk'}$. Regularization also helps resolve the lack of complete orthogonality of the spherical wavelets within and between scales (they form a frame, not a basis).

The solution to (10) is given by (e.g. Menke 1989)

$$\mathbf{m} = \left(\mathbf{G}^T \mathbf{C}_\mathrm{D}^{-1} \mathbf{G} + \lambda^2\, \mathbf{S}\right)^{-1} \mathbf{G}^T \mathbf{C}_\mathrm{D}^{-1} \mathbf{d}. \tag{13}$$

We choose λ by ordinary cross-validation, which tends to pick models that are rougher than would be selected by a typical L-curve trade-off approach (Fig. 2 and Appendix B).

The posterior data covariance is given by (e.g. Tarantola 2005)

$$\widetilde{\mathbf{C}}_\mathrm{D} = \mathbf{G} \left(\mathbf{G}^T \mathbf{C}_\mathrm{D}^{-1} \mathbf{G} + \lambda^2\, \mathbf{S}\right)^{-1} \mathbf{G}^T, \tag{14}$$

with the correspondence between the model covariance matrix and regularization matrix given by $\mathbf{C}_\mathrm{M} = (\lambda^2\, \mathbf{S})^{-1}$. We use $\widetilde{\mathbf{C}}_\mathrm{D}$ to compute a plotting mask for the estimated velocity field (Section 2.4).

2.4 Selection of frame functions

So far we have not specified any upper or lower limits to the scales of spherical wavelets that we will use in the estimation problem. In Table 2, we list spherical grids of orders $q = 0$–12. For each grid there are $10 \times 4^q + 2$ possible frame functions. Thus, the total number of possible frame functions in terms of q_max scales by $4^{q_\mathrm{max}+1}$ and is given by

$$m(q_\mathrm{max}) = \sum_{q=0}^{q_\mathrm{max}} [10 \times 4^q + 2] = \frac{1}{3}\left(4^{q_\mathrm{max}+1} - 1\right) + 2q_\mathrm{max}, \tag{15}$$

Table 3. Spectral and spatial support of spherical wavelets.

Scale		Spatial support		Spectral support
q	a_q			$0.990\|\psi_a\|^2$
0	2^0	82.442°	9167.1 km	4
1	2^{-1}	47.310°	5260.7 km	6
2	2^{-2}	24.707°	2747.3 km	10
3	2^{-3}	12.500°	1389.9 km	20
4	2^{-4}	6.268°	697.0 km	41
5	2^{-5}	3.136°	348.8 km	83
6	2^{-6}	1.569°	174.4 km	166
7	2^{-7}	0.784°	87.2 km	332
8	2^{-8}	0.392°	43.6 km	665
9	2^{-9}	0.196°	21.8 km	1329

Notes: Spatial support is defined as the distance to the first zero-crossing (see Fig. 1c). Spectral support is computed as the spherical harmonic degrees at which the energy reaches 99.0 per cent of the total energy (see Fig. 1d).

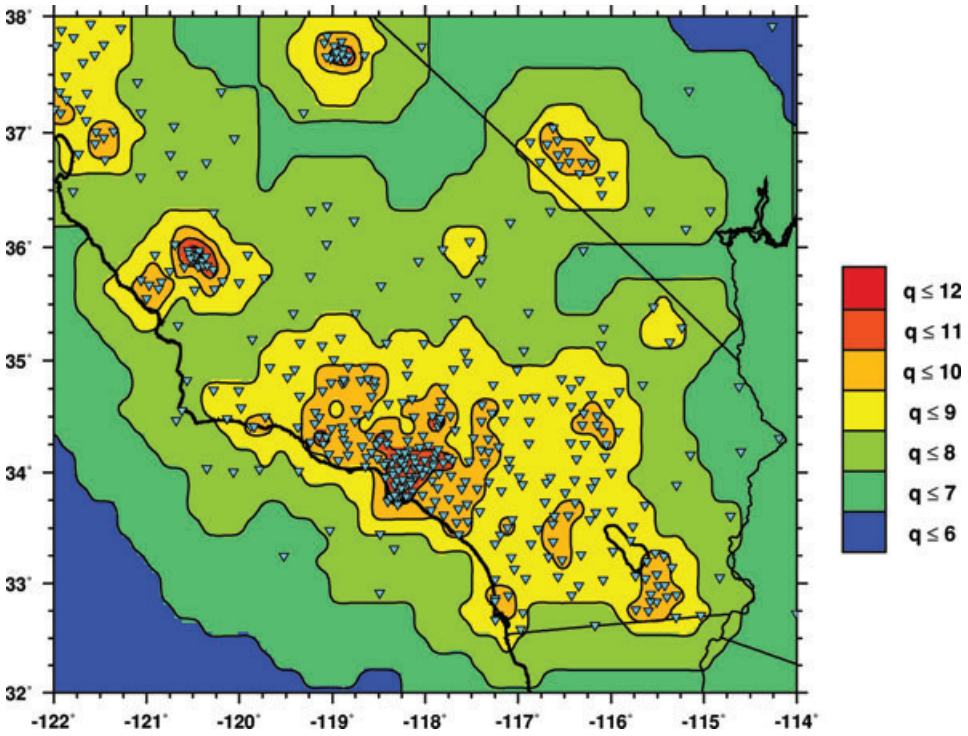

Figure 3. Selection of spherical wavelets based on the REASoN observation points, using scales $q = 3$–12. The density of the observation points, plotted as inverted triangles, controls the selection of spherical wavelets, whose centre points are not shown. The colour map shows the maximum-q scale wavelet that 'covers' each area, where the coverage is determined by the length scale for each spherical wavelet (Table 3). Where stations are dense, wavelets with all scales $q = 3$–12 (red) are available; where stations are sparse, only wavelets with long length scales $q = 3$–6 (blue) are available.

where we have used the formula for summing a geometric series. With no computational restrictions and careful choice of regularization, it is possible to use all available frame functions for every grid. However, for computational and practical purposes, it is desirable to use a reduced set of frame functions. Our algorithm first eliminates frame functions that have fewer than three stations within their 'footprint,' defined (somewhat arbitrarily) as the region inside the first zero-crossing of the spherical wavelet (Table 3). This procedure thereby determines the range of allowable q-values based on the local density of stations. An example is shown in Fig. 3 based on the REASoN stations in southern California; the density of stations allows for $q_{\max} = 12$. One may also choose to choose a $q_{\max}$ that is less than what is permitted by the station density; this choice will stabilize the estimation problem and reduce the possibility of overfitting the observations. We have chosen $q_{\max} \leq 9$ for each example in Section 4.

We also restrict $q_{\min}$ based on the approximate outer length scale, L, of the network. We select $q_{\min}$ such that the support of the $q_{\min}$ wavelets is less than $2L$. This restriction is not required, since proper regularization should leave little power in wavelets that are much larger than the network, and we already remove a pure rotational field, as described next.

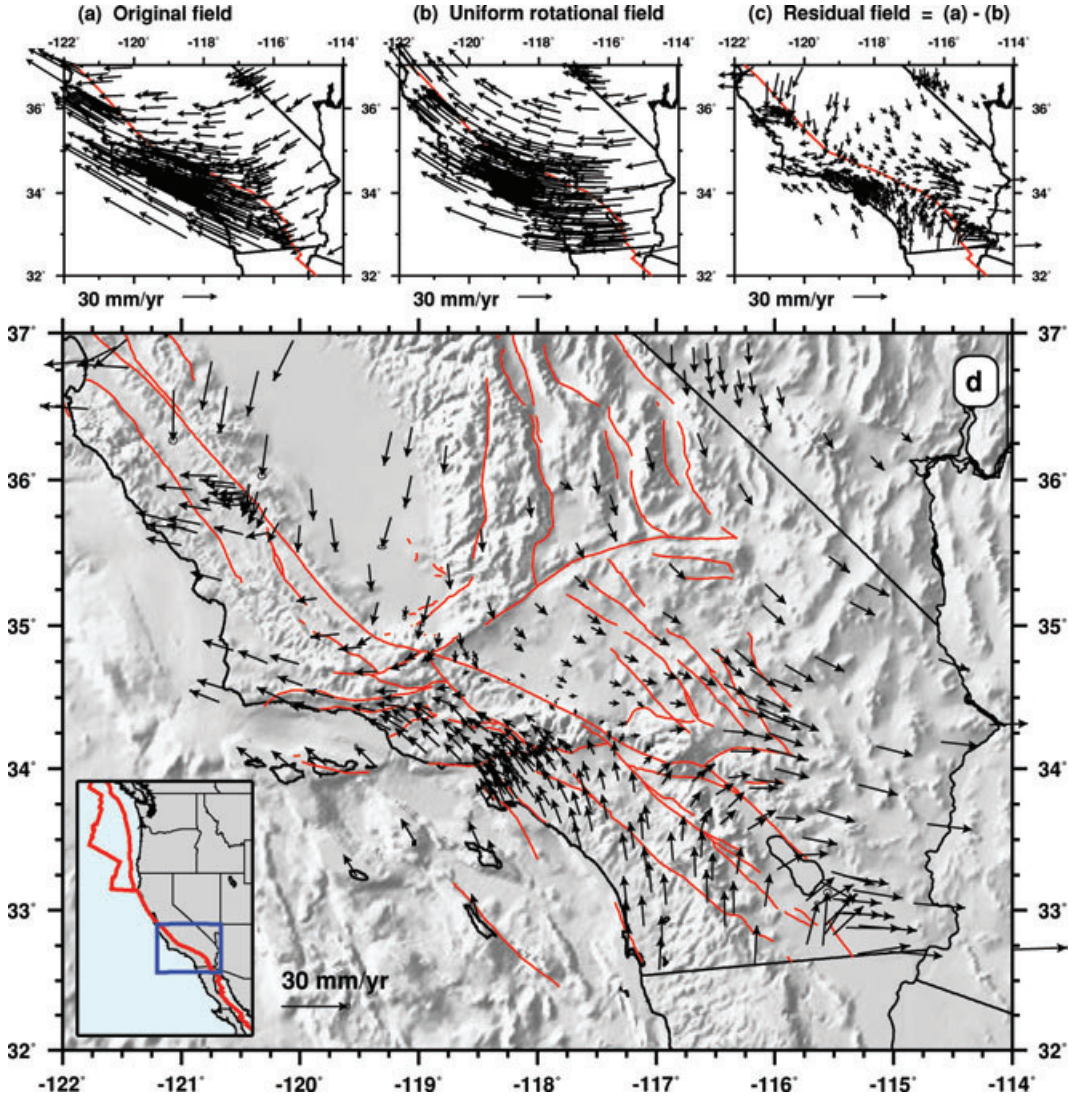

Figure 4. Removing a pure rotational field from the REASoN horizontal velocity field. (a) REASoN horizontal velocity field in southern California. (b) Best-fitting rotational field to the field in (a). (c) Residual horizontal velocity field: (a) minus (b). (d) Expanded view of (c), superimposed on shaded relief from Amante & Eakins (2008). See Fig. 9 for principal faults and features. See Fig. 10 for the REASoN vertical velocity field.

We now consider the longest-scale features in a given velocity field. We begin by considering a uniformly rotating sphere. Locally the velocity field appears as a simple translation, which, on the sphere, corresponds to a uniform rotational field. A rotational field on the sphere is described by a three-parameter Euler vector. In our approach, we optionally estimate an Euler vector that best fits the original velocity field, remove the associated pure rotational field, and proceed with the residual field. It is challenging to obtain accurate uncertainty bounds on the three Euler vector parameters (e.g. latitude, longitude and angular rotation rate); however, we are mainly interested in using the rotational field as a means of removing the largest-scale features of the velocity field, rather than relying on several large-scale spherical wavelets to account for such motions. An example of this procedure is shown in Fig. 4. Note that removal of a uniform rotational field does not affect estimates of strain-rate.

Finally, we mask regions of the estimated fields based on the diagonal terms of the posterior data covariance (14). We use the spatial variation of $\widetilde{\mathbf{C}}_{\mathrm{D}}$, together with a user-specified threshold value, to compute the masks used in the plots. Use of this mask means that all regions of the scalar fields that are shown are resolved to the indicated minimum scale (see Figs 12d, g and j). The use and choice of a specific binary threshold is admittedly *ad hoc*. Future analysis will consider different approaches to quantitatively conveying the utility of the derived quantities as a function of scale and position.

2.5 Representation and interpretation of multiscale results

A multiscale representation provides great flexibility in analysing pointwise position data. In our approach, the maximum resolution (or, equivalently, the finest adopted scale) of the derived field depends on the local density of observations. One could choose to only evaluate the

velocity field at scales that are allowable everywhere, in other words, the maximum value of the smallest scale allowed by a given network configuration. This choice would unnecessarily limit the detail of the resultant field. An alternative approach is to construct a single velocity field where the scale of information represented by the field varies as a function position in a given spatially heterogeneous network of observations. This approach may be useful for generating velocity field maps that are spatially uninterrupted, but is dangerous because it encourages the comparisons of derived quantities (such as strain rates) evaluated at different scales, thereby allowing one to spuriously infer spatial variability where none exists. In this study, we resist the temptation to produce a single map but insist on a perspective that considers scales of deformation separately, or at most, sums up all resolvable scales, but through error-based spatial masks, only allows for comparison of regions that are comparable at a given spatial resolution.

In the future, one might compare our estimates of strain-rate-like quantities with those from physical models (e.g. Meade & Hager 2005). Given that our estimates are essentially filtered versions of a true underlying field, such comparisons should be made by taking model predictions of velocities evaluated at the same points as the observed data and then estimating any derived quantities using the same approach (including the same value of regularization) that was used when treating the observed data (e.g. Simons *et al.* 1997).

3 STRAIN RATE MAPS AND OTHER SCALAR QUANTITIES

A vector and tensor are physical quantities. However, in practice we choose a basis (or coordinate system, e.g. $\hat{\mathbf{r}}$–$\hat{\boldsymbol{\theta}}$–$\hat{\boldsymbol{\phi}}$) to describe them. Once a basis is chosen, we can discuss the *components* of a vector (e.g. v_r, v_θ, v_ϕ) or tensor. The norm of a vector or tensor does not depend on the choice of basis, but it does depend on the choice of reference frame (for example, what is taken to be fixed). The norm of the gradient of the velocity field does not depend on the choice of basis or the reference frame. In this section, we examine what kinds of scalar quantities are sensible to derive from tensor quantities. Maps of these scalar quantities illuminate regions of deformation that may be difficult to detect from a visual inspection of the estimated velocity field.

Many different scalar quantities have been used in previous studies. Here we list several of these, assuming that the strain-rate tensor, $\mathbf{D}$ (e.g. eq. 16), is 3×3 and has eigenvalues λ_1, λ_2 and λ_3, with $\lambda_1 \geq \lambda_2 \geq \lambda_3$. For simplicity, we leave out multiplicative constants.

(i) $\lambda_1 + \lambda_2 + \lambda_3$ (e.g. Ward 1998b; Beavan & Haines 2001), equal to the trace of $\mathbf{D}$.
(ii) $\max[\,|\lambda_1|\ |\lambda_2|\ |\lambda_3|\,]$ (Ward 1998b).
(iii) $\sqrt{\mathbf{D} : \mathbf{D}}$ (see eq. D5) (Bos *et al.* 2003).
(iv) $\sqrt{\mathbf{D} : \mathbf{D}}$, but ignoring components of $\mathbf{D}$ associated with r: $\sqrt{D_{\phi\phi}^2 + D_{\theta\theta}^2 + 2D_{\phi\theta}^2}$ (Kreemer *et al.* 2003).
(v) $|\,\lambda_1 - \lambda_3\,|$ (Beavan & Haines 2001).
(vi) $D_{r\theta}$, $D_{r\phi}$, or $D_{\theta\phi}$ (Flesch *et al.* 2005).
(vii) $W_{\theta\phi}$ (Allmendinger *et al.* 2007), an element of the rotation tensor $\mathbf{W}$ (e.g. eq. 17).

Other more complicated expressions have also been used (e.g. Holt *et al.* 2000, eq. 14). We prefer to use scalar quantities that are invariant under coordinate transformations (Section 3.2). Because the eigenvalues do not depend on the choice of coordinate system, any scalar function in terms of eigenvalues is also an invariant quantity.

3.1 The velocity gradient tensor

Once we have computed the estimated velocity field, $\mathbf{v}(\theta, \phi)$, we can readily compute its surface derivatives and other tensor and scalar quantities. The spatial velocity gradient, $\mathbf{L}$, is defined by

$$\mathbf{L} \equiv (\boldsymbol{\nabla}\mathbf{v})^T = \mathbf{v}\boldsymbol{\nabla},$$

and can be decomposed as $\mathbf{L} = \mathbf{D} + \mathbf{W}$, where $\mathbf{D}$ is symmetric and $\mathbf{W}$ is antisymmetric. For a 3-D velocity field, these expressions are[†]

$$\mathbf{D} = \frac{1}{2}(\mathbf{L} + \mathbf{L}^T) = \begin{bmatrix} \frac{\partial v_r}{\partial r} & \frac{1}{2r}\left(-v_\theta + \frac{\partial v_r}{\partial \theta} + r\frac{\partial v_\theta}{\partial r}\right) & \frac{1}{2r}\left(-v_\phi + \frac{1}{\sin\theta}\frac{\partial v_r}{\partial \phi} + r\frac{\partial v_\phi}{\partial r}\right) \\ D_{12} & \frac{1}{r}\left(v_r + \frac{\partial v_\theta}{\partial \theta}\right) & \frac{1}{2r}\left(-v_\phi \cot\theta + \frac{1}{\sin\theta}\frac{\partial v_\theta}{\partial \phi} + \frac{\partial v_\phi}{\partial \theta}\right) \\ D_{13} & D_{23} & \frac{1}{r}\left(v_r + v_\theta \cot\theta + \frac{1}{\sin\theta}\frac{\partial v_\phi}{\partial \phi}\right) \end{bmatrix}, \tag{16}$$

$$\mathbf{W} = \frac{1}{2}\left(\mathbf{L} - \mathbf{L}^T\right) = \begin{bmatrix} 0 & \frac{1}{2r}\left(-v_\theta + \frac{\partial v_r}{\partial \theta} - r\frac{\partial v_\theta}{\partial r}\right) & \frac{1}{2r}\left(-v_\phi + \frac{1}{\sin\theta}\frac{\partial v_r}{\partial \phi} - r\frac{\partial v_\phi}{\partial r}\right) \\ -W_{12} & 0 & \frac{1}{2r}\left(-v_\phi \cot\theta + \frac{1}{\sin\theta}\frac{\partial v_\theta}{\partial \phi} - \frac{\partial v_\phi}{\partial \theta}\right) \\ -W_{13} & -W_{23} & 0 \end{bmatrix}. \tag{17}$$

Physically speaking, $\mathbf{D}$ is the strain-rate tensor and $\mathbf{W}$ is the rotation-rate tensor, though typically this terminology is reserved for the condition of 'small' deformation, where $D_{ij} \ll 1$ and $W_{ij} \ll 1$ (Malvern 1969, pp. 148–150). For the examples in Section 4, D_{ij} and W_{ij} are typically less than 10^{-7}.

[†] These can be found in Malvern (1969, p. 671), though note the $\cos\phi$ typo (instead of $\cos\theta$) in his expression for $D_{\theta\phi}$.

3.1.1 Some practical reductions for simple rheologies

With observations only at the surface, we do not have access to vertical derivatives of the velocity field, $\partial v_r/\partial r$, $\partial v_r/\partial\theta$, and $\partial v_r/\partial\phi$. Furthermore, in many scenarios the vertical component of the velocity field (v_r) and its horizontal derivatives ($\partial v_r/\partial\theta$, $\partial v_r/\partial\phi$) are not available. However, these observational limitations do not mean that **L**, **D** and **W** reduce to 2×2 tensors, even if one assumes that the missing quantities are equal to zero, because $D_{r\theta}$ contains v_θ, and $D_{r\phi}$ contains v_ϕ. To eliminate some of the unknown terms, we can use the fact that we are observing at the free-surface (i.e. zero tractions) in combination with simple assumptions regarding the rheological behaviour that is dominant at the surface on the timescale of the observations. We first assume a spherical Earth, ignoring both ellipticity and topography, that is, ignoring any deviations of the free surface from $r = R$. The free surface condition then becomes

$$\boldsymbol{\sigma}\,\hat{\mathbf{r}} = \mathbf{0}, \tag{18}$$

where $\boldsymbol{\sigma}$ is the second-order symmetric stress tensor, and $\hat{\mathbf{r}} = (1, 0, 0)$ is the upward-pointing unit normal in $r{-}\theta{-}\phi$ local coordinates. Eq. (18) provides the constraints

$$\sigma_{11} = 0, \qquad \sigma_{21} = \sigma_{12} = 0, \qquad \sigma_{31} = \sigma_{13} = 0. \tag{19}$$

For both simple isotropic viscous and isotropic elastic rheologies, then $D_{21} = D_{12} = 0$ and $D_{31} = D_{13} = 0$. If we assume a linear elastic rheology, then for observations at the surface, eqs (16) and (17) reduce to (see Appendix C)

$$\mathbf{D}_0^{\text{elastic}} = \begin{bmatrix} F\,(D_{22} + D_{33}) & 0 & 0 \\ 0 & \frac{1}{R}\left(v_r + \frac{\partial v_\theta}{\partial\theta}\right) & \frac{1}{2R}\left(-v_\phi\cot\theta + \frac{1}{\sin\theta}\frac{\partial v_\theta}{\partial\phi} + \frac{\partial v_\phi}{\partial\theta}\right) \\ 0 & D_{23} & \frac{1}{R}\left(v_r + v_\theta\cot\theta + \frac{1}{\sin\theta}\frac{\partial v_\phi}{\partial\phi}\right) \end{bmatrix}, \tag{20}$$

$$\mathbf{W}_0^{\text{elastic}} = \begin{bmatrix} 0 & \frac{1}{R}\left(-v_\theta + \frac{\partial v_r}{\partial\theta}\right) & \frac{1}{R}\left(-v_\phi + \frac{1}{\sin\theta}\frac{\partial v_r}{\partial\phi}\right) \\ -W_{12} & 0 & \frac{1}{2R}\left(-v_\phi\cot\theta + \frac{1}{\sin\theta}\frac{\partial v_\theta}{\partial\phi} - \frac{\partial v_\phi}{\partial\theta}\right) \\ -W_{13} & -W_{23} & 0 \end{bmatrix}, \tag{21}$$

where

$$F = -\lambda\,/\,(\lambda + 2\mu) \tag{22}$$

is a constant, and λ and μ are Lamé parameters. For a Poisson solid, $\lambda = \mu$, and $F = -1/3$.

We now assume a linear viscous rheology of the form (Malvern 1969, p. 298)

$$\boldsymbol{\sigma} = \text{Tr}(\boldsymbol{\sigma})\mathbf{I} + \Lambda\,\text{Tr}(\mathbf{D})\mathbf{I} + 2\eta\mathbf{D}\,, \tag{23}$$

where η and Λ characterize the viscosity. With the assumption of incompressibility, $\boldsymbol{\nabla}\cdot\mathbf{v} = \text{Tr}\,(\mathbf{L}) = \text{Tr}\,(\mathbf{D}) = 0$, we obtain the constraint $D_{11} = -(D_{22} + D_{33})$. In comparison with eq. (20), this corresponds to $F = -1$. Applying the free-surface constraints in eq. (19), we obtain the additional constraints, $D_{12} = D_{13} = 0$, and the algebra follows the elastic case, but with $F = -1$.

3.2 Scalar quantities derived from tensors

In our examples, we consider three invariant quantities, related to dilatation rate (Θ), strain rate (Σ) and rotation rate (Ω). In functional form, these quantities are written as $\Theta(\theta, \phi)$, $\Sigma(\theta, \phi)$ and $\Omega(\theta, \phi)$, since our observations are restricted to the surface of the sphere.

3.2.1 Dilatation rate

The dilatation rate, or volumetric strain rate, is the divergence of the velocity field, or, equivalently, the first invariant of the velocity gradient tensor

$$\begin{aligned}\Theta &\equiv \boldsymbol{\nabla}\cdot\mathbf{v} = \text{Tr}(\mathbf{L}) = \text{Tr}(\mathbf{D}) \\ &= r^{-1}\left[2\,v_r + v_\theta\cot\theta + r\frac{\partial v_r}{\partial r} + \frac{\partial v_\theta}{\partial\theta} + \frac{1}{\sin\theta}\frac{\partial v_\phi}{\partial\phi}\right],\end{aligned} \tag{24}$$

listed in Malvern (1969, p. 670). The dilatation rate, after applying the free-surface condition, is obtained from eq. (20)

$$\Theta_0^{\text{elastic}} = \frac{1}{R}\,(F + 1)\left(2\,v_r + v_\theta\cot\theta + \frac{\partial v_\theta}{\partial\theta} + \frac{1}{\sin\theta}\frac{\partial v_\phi}{\partial\phi}\right), \tag{25}$$

which is used in the examples in Section 4.

3.2.2 Strain rate

We define a strain-rate quantity, Σ, to be the Frobenius norm (Appendix D) of the strain-rate tensor $\mathbf{D}$. Mathematically, Σ can be expressed as (Appendix D)

$$\Sigma = \|\mathbf{D}\|_{\mathrm{F}} = \sqrt{\mathbf{D} : \mathbf{D}} = \sqrt{tr\,(\mathbf{DD})}\,, \tag{26}$$

which is used in Bos *et al.* (2003). For the remainder of this paper, we will abbreviate 'Frobenius norm of the strain-rate tensor' as 'strain rate.'

3.2.3 Rotation rate

The rotation vector, $\mathbf{w} = \frac{1}{2}\,\nabla \times \mathbf{v}$, is determined from the components of the rotation-rate tensor, $\mathbf{W}$, as (Malvern 1969, p. 147)

$$\mathbf{w} = -W_{\theta\phi}\hat{\mathbf{r}} + W_{r\phi}\hat{\boldsymbol{\theta}} - W_{r\theta}\hat{\boldsymbol{\phi}}. \tag{27}$$

Substituting the terms in eq. (21), we obtain

$$\mathbf{w} = \frac{1}{2R}\left(v_\phi \cot\theta - \frac{1}{\sin\theta}\frac{\partial v_\theta}{\partial \phi} + \frac{\partial v_\phi}{\partial \theta}\right)\hat{\mathbf{r}} + \frac{1}{R}\left(-v_\phi + \frac{1}{\sin\theta}\frac{\partial v_r}{\partial \phi}\right)\hat{\boldsymbol{\theta}} + \frac{1}{R}\left(v_\theta - \frac{\partial v_r}{\partial \theta}\right)\hat{\boldsymbol{\phi}}, \tag{28}$$

which is the rotation vector associated with a point on the surface of the sphere, (R, θ, ϕ). The magnitude of rotation is

$$\Omega \equiv |\mathbf{w}| = \left(W_{\theta\phi}^2 + W_{\phi r}^2 + W_{r\theta}^2\right)^{1/2}. \tag{29}$$

The units of Ω should be radians per unit time. This quantity is invariant under coordinate transformations; it can also be expressed as (Appendix D)

$$\Omega = \frac{1}{\sqrt{2}}\,\|\mathbf{W}\|_{\mathrm{F}} = \sqrt{-\frac{1}{2}\,\mathrm{Tr}(\mathbf{WW})} = \sqrt{\frac{1}{2}\,\mathbf{W} : \mathbf{W}}. \tag{30}$$

The location of the rotation pole is the (longitude, latitude) where the unit vector $\mathbf{w}/|\mathbf{w}|$ intersects the sphere.

For physical interpretations, we partition the rotation vector (eq. 28) into two vectors, $\mathbf{w} = \mathbf{w}_{\mathrm{s}} + \mathbf{w}_{\mathrm{t}}$, where

$$\mathbf{w}_{\mathrm{s}} = \frac{1}{2R}\left(v_\phi \cot\theta - \frac{1}{\sin\theta}\frac{\partial v_\theta}{\partial \phi} + \frac{\partial v_\phi}{\partial \theta}\right)\hat{\mathbf{r}} - \frac{v_\phi}{R}\,\hat{\boldsymbol{\theta}} + \frac{v_\theta}{R}\,\hat{\boldsymbol{\phi}}, \tag{31}$$

$$\mathbf{w}_{\mathrm{t}} = \frac{1}{R}\left(\frac{1}{\sin\theta}\frac{\partial v_r}{\partial \phi}\right)\hat{\boldsymbol{\theta}} + \frac{1}{R}\left(\frac{\partial v_r}{\partial \theta}\right)\hat{\boldsymbol{\phi}}. \tag{32}$$

The subscript 's' is applied to denote 'sphere,' in that all velocity terms are tangential to the surface of the sphere (v_θ and v_ϕ); the subscript 't' is applied to denote 'tilt,' in that all terms are related to local tilt-like motion.

3.3 Linear scalar quantities

The tensors $\mathbf{L}$, $\mathbf{D}$ and $\mathbf{W}$—and the rotation vector $\mathbf{w}$—are all linear with respect to decompositions of $\mathbf{v}$ into different scales. However, of the three scalar quantities discussed above—dilatation rate (Θ), strain rate (Σ) and rotation rate (Ω)—only dilatation rate is linear with respect to decompositions of $\mathbf{v}$. For example, if we decompose $\mathbf{v}$ into five subfields, then the sum of the dilatation rates of the subfields is equal to the dilatation of $\mathbf{v}$. The strain rate quantity is not linear because of the squared terms in eq. (26). The rotation rate quantity is not linear because of the squared terms in eq. (29).

4 EXAMPLES

Here we present a series of examples illustrating the multiscale estimation technique described in Section 2. For the computations of the velocity gradient, we employ the free-surface condition and assume a linear elastic rheology, as summarized in Appendix C, with the key formulas eqs (20) and (21). The scales used in each example are summarized in Table 2: low q refers to long length scales, while high q refers to short length scales.

As previously described, spherical wavelets are automatically selected based on station density, and optionally, a user-defined $q_{\max}$ (Section 2.4 and Fig. 3). In estimating the coefficients for each wavelet, a regularization parameter is obtained using ordinary cross-validation (Appendix B and Fig. 2).

4.1 Synthetic examples

Results from three synthetic examples are shown in Figs 5–8. Most of the synthetic fields do not have any direct relevance to southern California; we use them only to highlight different styles of deformation that could be observed in a tectonic setting.

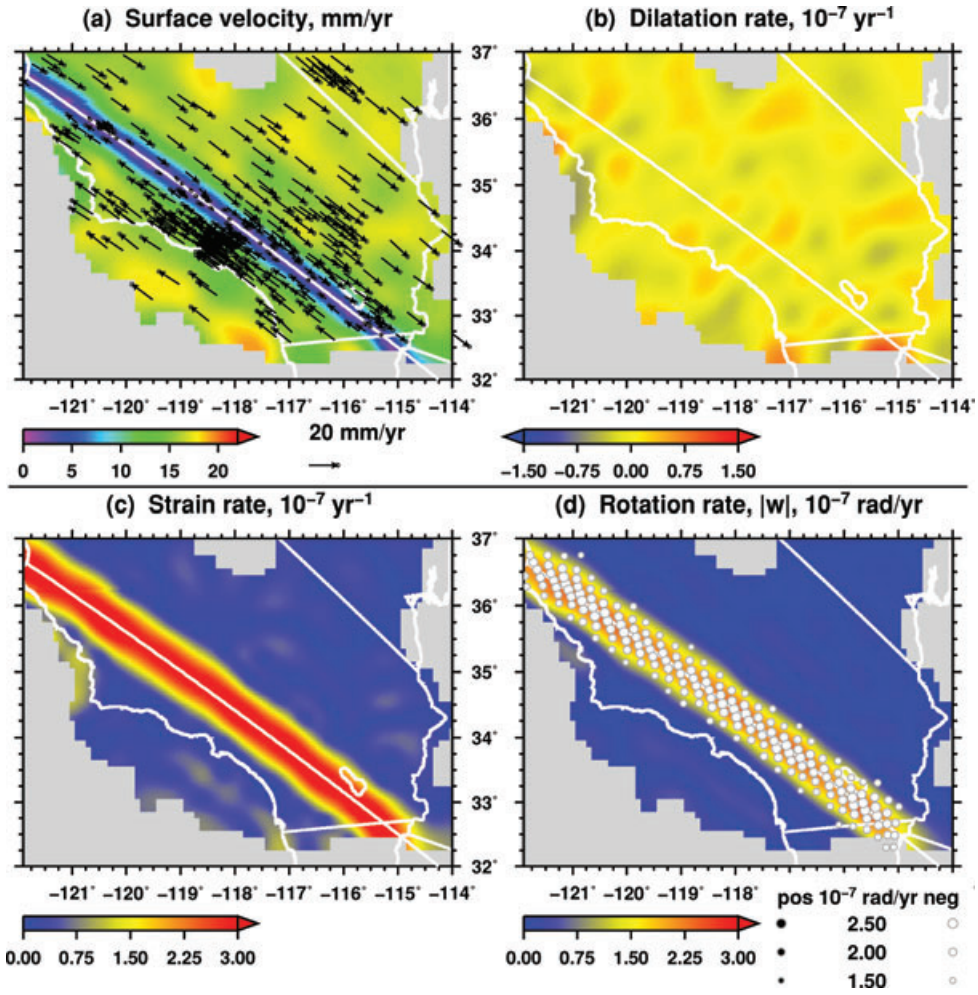

Figure 5. Interseismic field for a locked strike-slip fault (Section 4.1.1). (a) Black arrows denote a synthetic horizontal velocity field computed at the actual GPS observation points. Background colour is the magnitude of the multiscale estimated horizontal velocity field, using scales $q = 3$–7. (b)–(d) Dilatation rate (b), strain rate (c) and rotation rate (d), each computed directly from the estimated velocity field. Superimposed on the rotation-rate field (d) are Euler poles and antipoles for which $|\mathbf{w}| \geq 1.5 \times 10^{-7}$ rad yr^{-1}. White circles denote antipoles, that is, poles associated with a negative (i.e. clockwise) rotation direction.

For each synthetic velocity field, we consider both uniform and irregular grids of observations, both with and without errors added. For the irregular grids, the reference observation points are the GPS locations for the REASoN data set in southern California (Fig. 4). We add Gaussian random errors to the synthetic velocity fields. In most synthetic examples, we only show the results for the case of irregularly spaced observations with errors added, which is the closest to a real scenario. All observation points for the velocity fields are on the surface of a sphere, and the components are given in terms of local directions of up ($\hat{\mathbf{r}}$), south ($\hat{\boldsymbol{\theta}}$) and east ($\hat{\boldsymbol{\phi}}$) (eq. 1).

Fig. 5 shows our standard presentation for each example. In (a) we show the 'target' data as vectors (all with errors added), as well as the magnitude of the estimated horizontal velocity field (as described in Section 2) as the underlying coloured scalar field. Each velocity field is estimated using spherical wavelets with scales from $q_{\min}$ to $q_{\max}$ (in Fig. 5, $q_{\max} = 7$), with the local q values determined directly from the local station density (e.g. Fig. 3). The selection of the minimum scale in the estimation is determined from the length scale of the network; for the southern California region we use $q_{\min} = 3$ (see Section 2.4). From the estimated velocity fields, we can directly compute the estimated fields for dilatation rate (Fig. 5b), strain rate (Fig. 5c) and rotation rate (Fig. 5d).

4.1.1 An interseismic strike-slip velocity field

We first consider a velocity field at the surface due to a strike-slip fault locked at depth (e.g. Chinnery 1961, Fig. 5). Two parameters are needed to specify the model: the locking depth and the slip rate. We approximate the interseismic field as a 'great-circle fault,' one that is defined on the entire sphere. First, given two latitude–longitude points on the fault, we compute the pole to the fault. We then compute the velocity field for a spinning hemisphere whose points on the fault move with half the slip rate. For the other hemisphere we compute the velocity field for a spinning hemisphere in the opposite direction with half the slip rate. These oppositely spinning hemispheres produce the

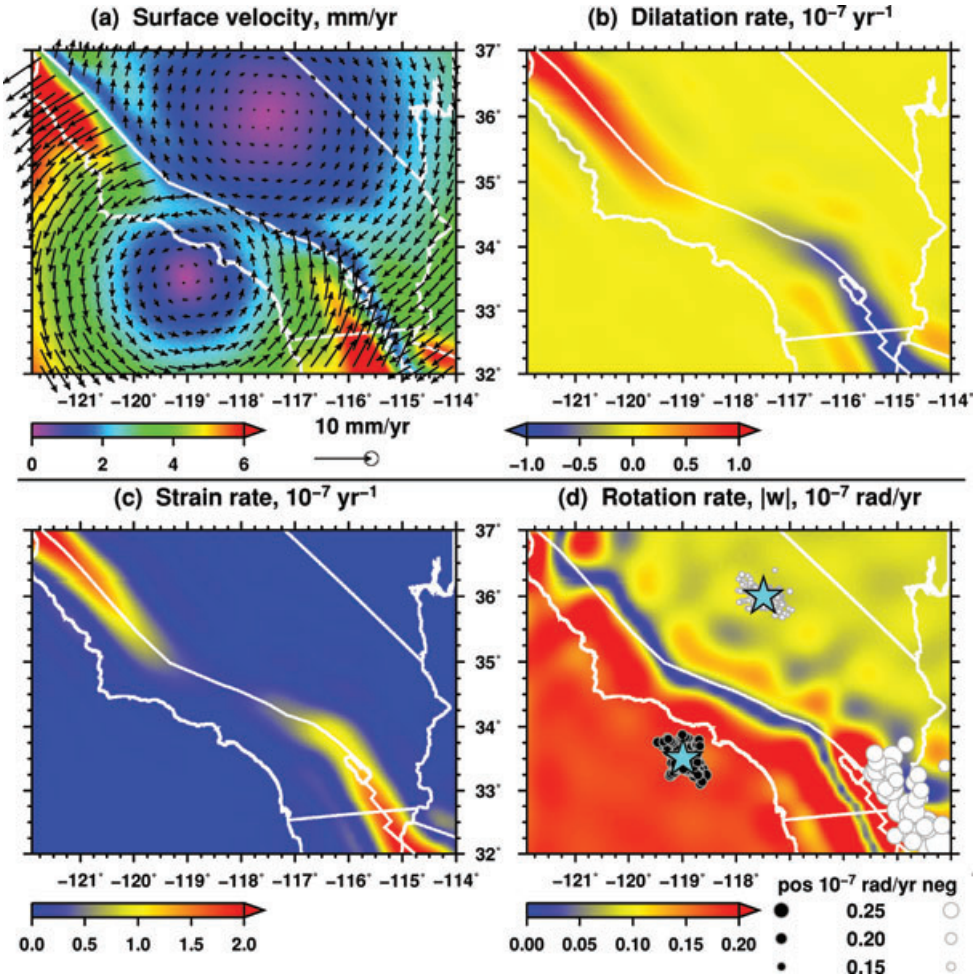

Figure 6. Example of two-microplate rotational field (Section 4.1.2). (a) Black arrows denote a synthetic horizontal velocity field computed at a regular set of gridpoints. Background colour is the magnitude of the multiscale estimated horizontal velocity field, using scales $q = 3$–7. (b)–(d) Dilatation rate (b), strain rate (c) and rotation rate (d), each computed directly from the estimated velocity field. Superimposed on the rotation-rate field (d) are the locations of Euler poles (black circles) and antipoles (white circles). The three clusters represent 60 per cent of the Euler vectors corresponding to the rotational field; the remaining poles (and antipoles) are outside the region or are not plotted. For each of the semi-concentric clusters, we have computed the mean direction, plotted as a blue star.

'unlocked' fault velocity field, which is then modulated by the locking depth with a function of the form

$$F(\Delta, D) = \frac{2}{\pi} |\tan^{-1}[R(\Delta - \pi)/D]|, \tag{33}$$

where R is the radius of Earth, and Δ is the angular distance from an observation point to the pole of the great-circle fault. The resultant synthetic velocity fields contain no vertical component, and the corresponding dilatation-rate field is zero.

Fig. 5 shows the results for a synthetic velocity field computed for the strike-slip fault, locked down to a depth of 15 km, beneath which the relative slip is 35 mm yr^{-1}, representing the plate motion. The strain-rate field (Fig. 5c) exhibits the expected interseismic pattern for a locked fault, with a maximum along the fault and symmetric to either side. The maximum strain rate along the fault is approximately 35×10^{-8} yr^{-1}.

As expected for this example, the estimated dilatation rate (Fig. 5b) is nearly zero. The extent to which it is non-zero results from a combination of (1) errors added to the synthetic data and (2) artefacts from the estimation. The mean and standard deviation of the unmasked points are $(-0.054 \pm 1.8) \times 10^{-8}$; in other words, the spurious variation in dilatation rate is approximately 30 times less than the maximum strain rate.

The rotation rate (Fig. 5d) has a maximum of approximately 15×10^{-8} rad yr^{-1}, indicating that rotation can partly account for the deformation in the vicinity of the fault. The presence of only antipoles (white circles) are indicative of the clockwise motion associated with the right-lateral fault.

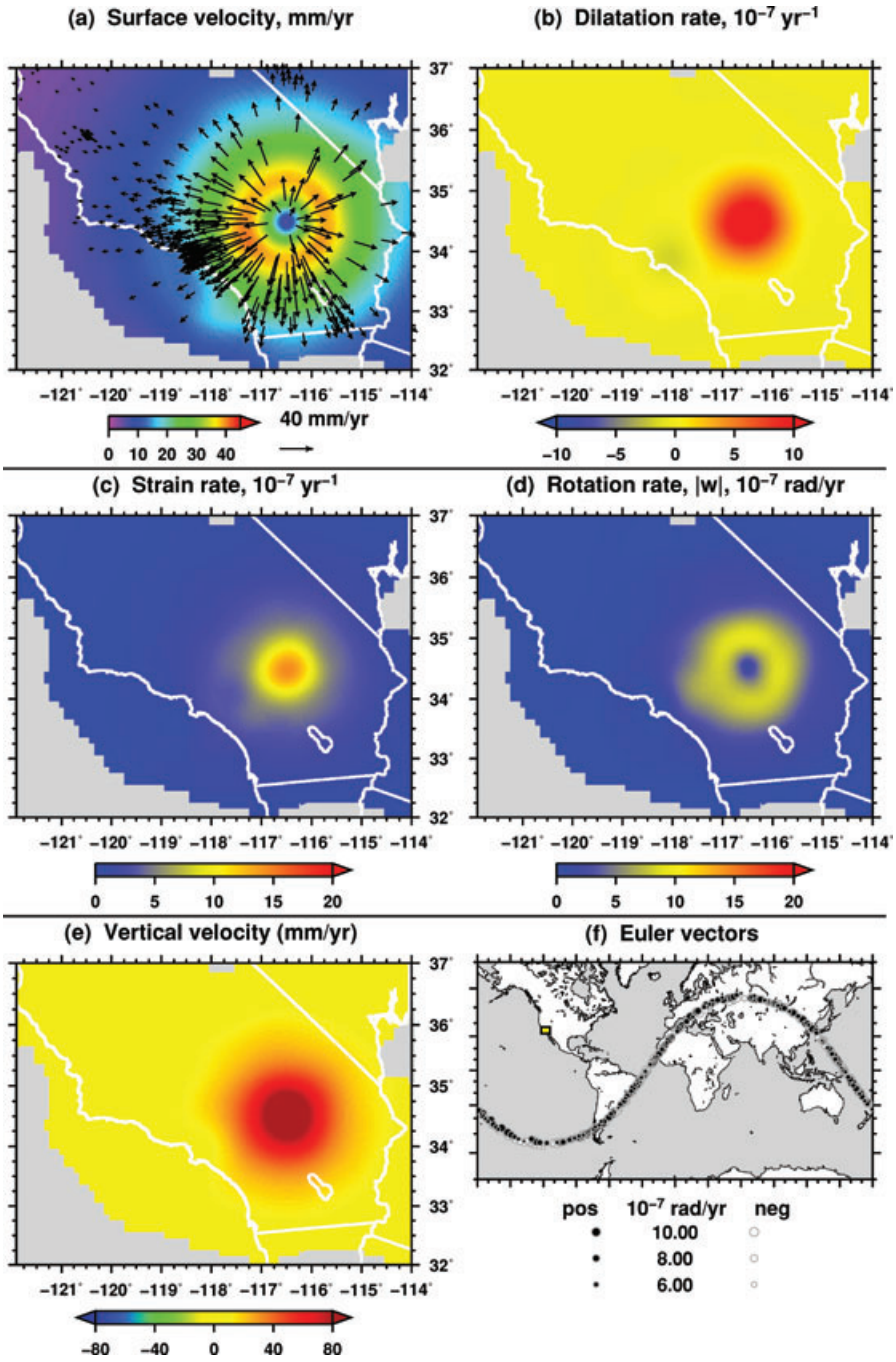

Figure 7. Example of two dilatational sources (Section 4.1.3). (a) Black arrows denote a synthetic horizontal velocity field computed at the GPS observation points of the REASoN data set (Dong *et al.* 2009). Background colour is the magnitude of the multiscale estimated horizontal velocity field, using scales $q = 3$ to $q = 7$. Vertical velocity field is shown in Fig. 8(b). (b)–(d) Dilatation rate (b), strain rate (c) and rotation rate (d), each computed directly from the estimated velocity field. In the rotation-rate field, $|\mathbf{w}|$, the ring-like pattern is due to the tilt terms in $|\mathbf{w}_t|$ (eq. 32). (e) Estimated vertical velocity field. (f) Euler vectors with magnitudes $|\mathbf{w}| \geq 4 \times 10^{-7}$ rad yr^{-1}. Antipoles (white) are plotted beneath the poles (black); both sets of poles lie along a great circle. Small box denotes the region shown in (a)–(e).

4.1.2 A rotational field with two microplates

We now consider another two-plate example with three differences compared with Fig. 5: (1) the two rotation poles are closer to the fault (rather than at the fault normal); (2) the boundary is irregular (rather than a great circle) and (3) the observation points are regularly spaced. Our observation points are divided by the approximate Pacific–North America Plate boundary (Fig. 6). The northeast 'microplate' rotates in a negative direction about $(-117.5^\circ, 36^\circ)$, and the southwest 'microplate' rotates in a positive direction about $(-119^\circ, 33.5^\circ)$.

Dilatation rate is negative along the southern segment of the plate boundary, where the velocities are strongly convergent. Along the central segment, dilatation and strain rates are almost absent because the velocities across the boundary are similar in direction and magnitude.

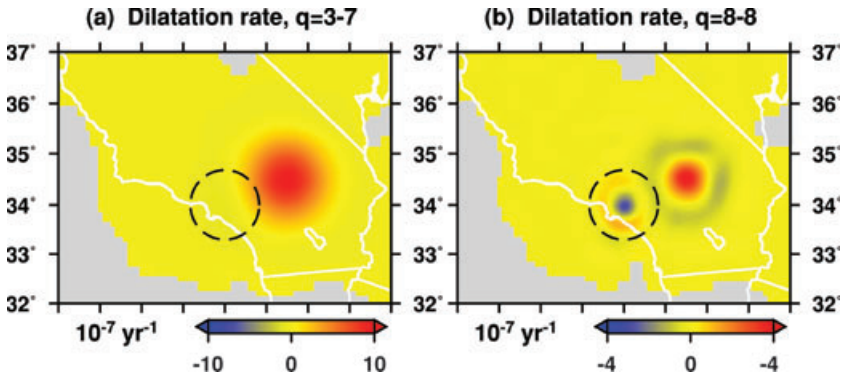

Figure 8. Multiscale estimation of the dilatational field for the synthetic velocity field shown in Fig. 7(a) (Section 4.1.3), using scales $q = 3$–8. Dashed circle corresponds to the region of the smaller dilatational source. (a) Dilatational field derived from the $q = 3$–7 scales of the $q = 3$–8 estimated velocity field. (b) Dilatational field derived from the $q = 8$ scale of the $q = 3$–8 estimated velocity field. Only at this scale is the smaller, negative dilatational source apparent.

Away from the boundary, values of the rotation-rate field are approximately 1×10^{-8} rad yr^{-1} in the northeast microplate and 2×10^{-8} rad yr^{-1} in the southwest microplate.

Each longitude–latitude point of the rotation-rate field, $|\mathbf{w}|$, has a corresponding Euler vector, $\mathbf{w}$, and Euler pole, $\mathbf{w}/|\mathbf{w}|$. In this example, most of the Euler poles fall into three clusters (Fig. 6d). The southeast cluster of relatively large negative rotations is associated with strong local rotations near the boundary, analogous to the poles plotted in Fig. 5(d). The two semi-concentric clusters of poles reveal the Euler vectors for each microplate (Fig. 6d). Note that the sign of rotation vectors is negative for the northeast microplate but positive for the southwest microplate, as expected from the velocity field in Fig. 6(a).

4.1.3 Two dilatational 'volcanic' sources

We now consider two volcanic-like dilatational sources. The sources are simplified 'Mogi sources,' parametrized in terms of the location of the source (x_0, y_0, z_0) and the strength, c, which can be interpreted in terms of the rate of (de)pressurization or equivalently the rate of magma chamber volume change. The three-component displacement field is given by (Mogi 1958)

$$\mathbf{U}(x, y, z) = \frac{c\,(x - x_0)}{d(x, y, z)^3}\,\hat{\mathbf{x}} + \frac{c\,(y - y_0)}{d(x, y, z)^3}\,\hat{\mathbf{y}} + \frac{c\,(z - z_0)}{d(x, y, z)^3}\,\hat{\mathbf{z}}, \tag{34}$$

where

$$d(x, y, z) = \left[(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2\right]^{1/2}. \tag{35}$$

We implement this by converting our longitude–latitude (ϕ, θ) points to UTM (x, y) and assuming that all observation points are at the surface $(z = 0)$.

The first, and dominant, source for the synthetic field is located at $(-116.5^\circ, 34.5^\circ, 100\,\text{km})$ with amplitude $c_1 = 10^9$. The second, much smaller, source is located at $(-118^\circ, 34^\circ, 30\,\text{km})$ with amplitude $c_2 = -10^7$ (note c_2 is chosen to be of opposite sign of c_1 in this example). The horizontal components of the displacement field are plotted in Fig. 7(a).

As expected, the effect of these sources is clearest in the dilatation rate (Fig. 7b). The rotation-rate field (Fig. 7d) forms a ring-like pattern due to the 'tilt terms' in $|\mathbf{w}_t|$ (eq. 32). The maximum value of $|\mathbf{w}_t|$ is 7.7×10^{-7} rad yr^{-1}, while the maximum value of $|\mathbf{w}_s|$ is only 0.25×10^{-7} rad yr^{-1}. The corresponding Euler vectors lie approximately along a great circle (Fig. 7f).

The multiscale decomposition of the velocity field reveals the smaller dilatational source only at scale $q = 8$. The estimated dilatational field up to $q = 7$ does not show the signature of the smaller source (Fig. 8a), but the dilatation rate for scale $q = 8$ reveals it well (Fig. 8b). Thus, with the multiscale estimation we are able to isolate the location, spatial extent, magnitude and sign of the smaller dilatational source.

We also consider an example where we set the vertical component of the velocity field and its spatial derivatives to zero. As expected based on the discussion above, the rotation signal is then absent. Although not shown in figures, the dilatation for the with- and without-vertical cases are imperceptibly different: the dilatation in the case that includes the vertical velocities is overwhelmed by the spatial derivative terms $\partial v_\theta/\partial\theta$ and $\partial v_\phi/\partial\phi$.

4.2 Three-component southern California velocity field

4.2.1 Data selection: spatial coverage and time window of analysis

For example purposes, we will use a subset of the velocity solution from the NASA Earth Science Research, Education and Applications Solution Network (REASoN) project (Dong *et al.* 2009). This data set consists of three-component velocity field measurements derived from only continuously operating GPS sites over an 11.6-yr time period from 1997 to 2008 (Fig. 4). The velocity field is described in Dong *et al.* (2009). The version we use has had both seasonal variations and post-seismic signals removed. The post-seismic signals are from three

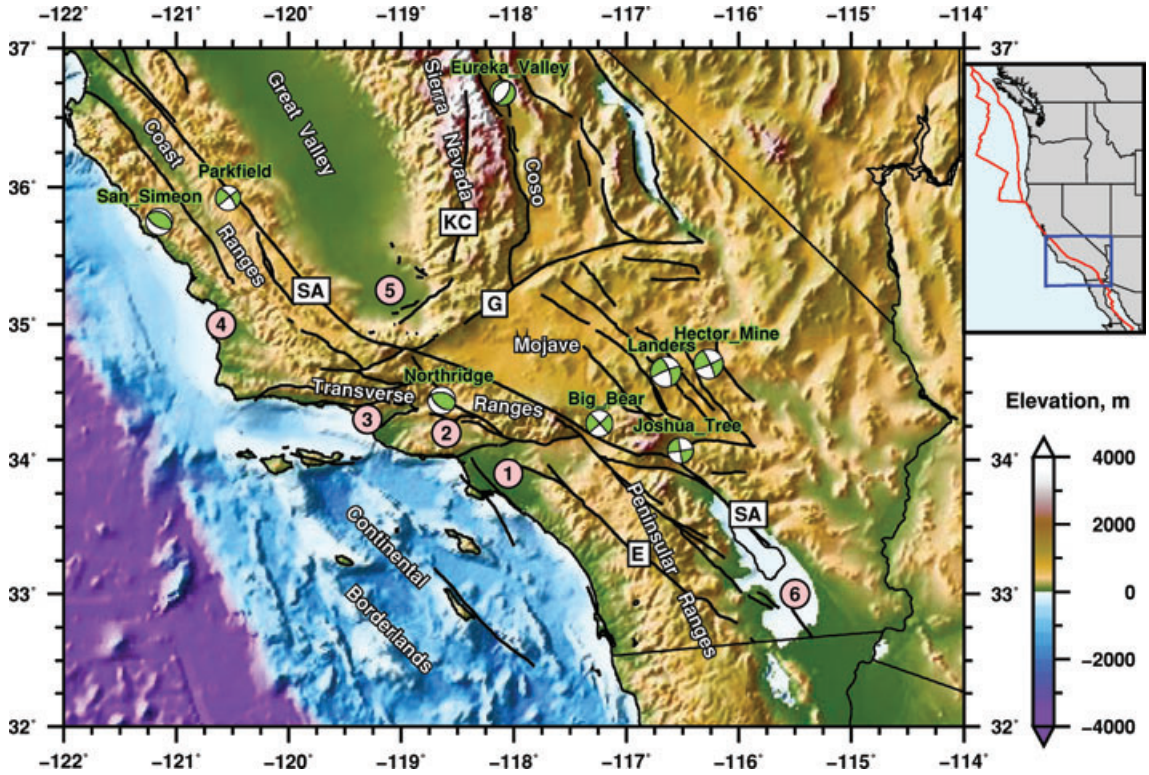

Figure 9. Base map for GPS velocity field study of southern California. Map shows topography and bathymetry (Amante & Eakins 2008), as well as active faults (Jennings 1994), including the Kern Canyon fault (Nadin & Saleeby 2009). Labels 1–6 denote the sedimentary basins of Los Angeles (1), San Fernando (2), Ventura–Santa Barbara (3), Santa Maria (4), southern San Joaquin (5) and the Salton trough (6), all of which have been active during the Neogene. Faults labelled for reference are: San Andreas (SA), Kern Canyon (KC), Garlock (G), and Elsinore (E). Global CMT fault-plane solutions (Dziewonski & Woodhouse 1983) for large ($M_w \geq 6.0$) earthquakes from 1990 to 2008 are: Landers (1992 June 28 M_w 7.3), Hector Mine (1999 October 16 M_w 7.1), Northridge (1994 January 17 M_w 6.7), San Simeon (2003 December 22 M_w 6.6), Joshua Tree (1992 April 23 M_w 6.2), Eureka Valley (1993 May 17 M_w 6.1) and Parkfield (2004 September 28 M_w 6.0). Inset map shows the plate boundary setting for western North America (Bird 2003).

earthquakes (Fig. 9): 1999 October 16 M_w 7.1 Hector Mine (e.g. Pollitz *et al.* 2001; Hudnut *et al.* 2002), 2003 December 22 M_w 6.5 San Simeon (e.g. McLaren *et al.* 2008) and 2004 September 28 M_w 6.0 Parkfield (Savage & Langbein 2008).

A denser set of geodetic measurements for California can be found in the California Crustal Motion Map, version 1.0 (CCMM1; Shen *et al.* 2006), which was preceded by the SCEC Crustal Motion Map, v3.0 (Shen *et al.* 2003). (In the southern California region, CCMM1 contains 1093 observations, while REASoN contains 408.) However, these data sets contain many campaign GPS sites, which are not as useful for examining temporal variations, which is our ultimate objective.

In Fig. 4, we show our subset of the REASoN data set, including uncertainty estimates. The reference frame for the velocity field is ITRF2005 (Altamimi *et al.* 2007), which leads to a significant component of rotation, when compared to a fixed-North-America reference frame, for example. Thus, as described in Section 2.4, we estimate a single Euler vector for the field, and remove the least-squares best-fitting purely rotational field (Figs 4a–c). The difference between the residual velocity field (c) and the original one (a) is a uniform rotation, which has no effect on quantities such as strain rate or dilatation rate.

4.2.2 Vertical velocity field

The vertical velocity field of our subset of the REASoN data set is plotted in Fig. 10(a). Since we are primarily interested in spatial gradients of the vertical field, we could remove a reference vertical rate to all points, similar to removing the effect of a uniform rotation on the horizontal field. However, because the vertical velocities are distributed about 0 mm yr^{-1}, we do not remove a reference value.

We estimate the vertical velocity field using scales $q = 3$–7 (Fig. 10b). In order to better fit the magnitude of the vertical field, we use a regularization parameter that is less than the minimum of the ordinary cross-validation function (Fig. 9). This underdamping leads to reduced residuals in Fig. 10(c), but at the expense of more 'oscillations' from one scale to the next in the estimation (Figs 12g and j).

The estimated vertical field considers the uncertainty in the observations and helps reveal the primary features of the velocity field. In particular, observation points above the large sedimentary basins (San Joaquin, Ventura, Los Angeles, Salton) are going down, while regions moving up include Sierra Nevada (Fay *et al.* 2008), Parkfield, Transverse Ranges and the eastern Mojave. The residual field reveals a spatial pattern dominated by the strongest subsidence signals in the San Joaquin basin and the Salton trough (Fig. 10c). By allowing for shorter

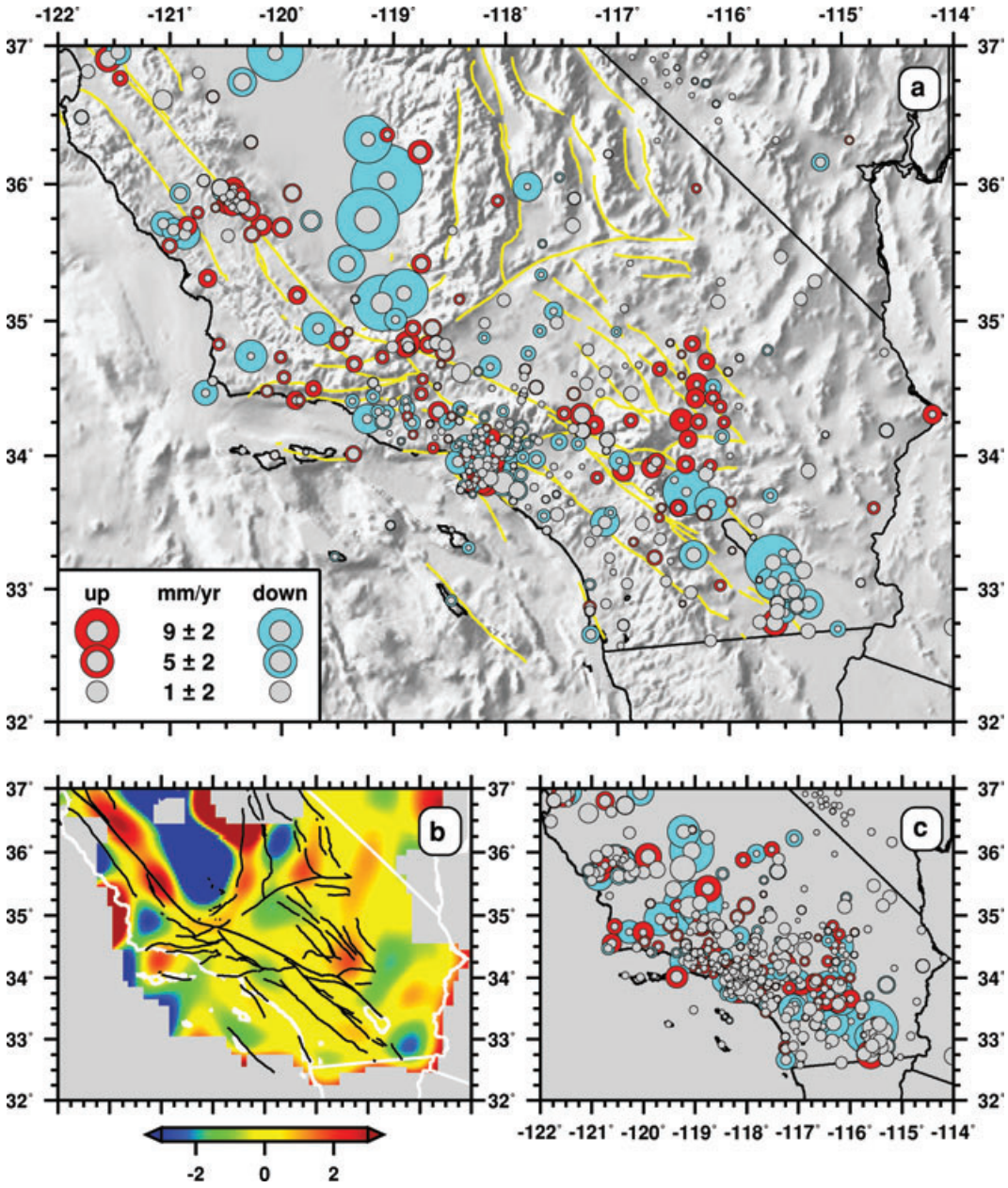

Figure 10. (a) Vertical component of the REASoN continuous GPS velocity field, southern California. Gray circles denote the estimated standard deviation; the legend shows uncertainty estimates of ± 2 mm yr^{-1}. The area of each plotted circle is proportional to the measurement value. See Fig. 9 for principal faults and features. (b) Estimated vertical velocity field, using scales $q = 3$–7. (c) Residual vertical velocity field: (a) minus (b). The legend is the same as in (a); note that the size of each grey circle remains the same.

scales in the estimation (such as $q = 8$) we could reduce the residuals. Undoubtedly, several of the observations contain non-tectonic signals (e.g. sediment compaction) or have anthropogenic causes (Dong *et al.* 2009).

4.2.3 Multiscale estimation and decomposition

We estimate the 3-D velocity field for $q = 3$–7 (Fig. 11). We also estimate the 2-D horizontal velocity field for the same scales (not shown). Comparing the scalar quantities estimated using the 2-D and 3-D fields, only the rotation-rate fields are perceptibly different. Thus our interpretations of strain rate and dilatation rate apply to both the 2-D and 3-D cases.

The dominant feature in the strain-rate field (Fig. 11c) is the high values along the San Andreas fault, in particular near the Parkfield region and the Salton Sea region. The dilatation (Fig. 11b) is correlated with the vertical velocity field (Fig. 10b): where dilatation is positive (Sierra Nevada, Parkfield), the vertical velocities point up; where the dilatation is negative (Ventura basin, L.A. basin, San Joaquin basin), the vertical velocities point down. As crustal material diverges, the observation sites experience subsidence; as material converges, the observation

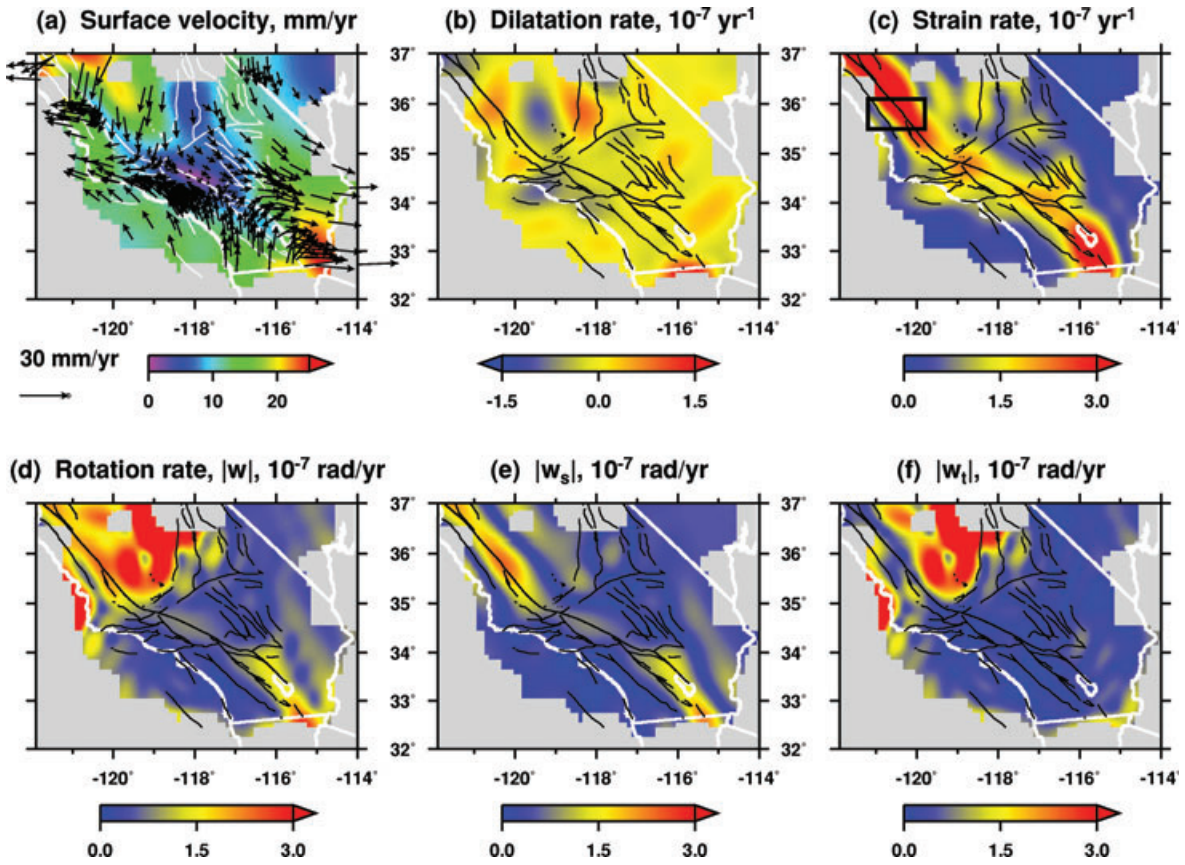

Figure 11. REASoN velocity field. (a) Black arrows denote the horizontal REASoN velocity field with estimated uncertainty ellipses (Fig. 4d). Background colour is the magnitude of the multiscale estimated horizontal velocity field, using scales $q = 3$ to $q = 7$. (b)–(d) Dilatation rate (b), strain rate (c) and rotation rate (d), each computed directly from the estimated velocity field, including the vertical component (Fig. 10). The box in (c) is the area shown in Fig. 15. (e) Component of rotation rate with velocity terms tangential to the sphere, $|\mathbf{w}_s|$ (eq. 31). (f) Component of rotation rate containing 'tilt terms,' $|\mathbf{w}_t|$ (eq. 32). See Section 4.2 for details.

sites experience uplift. One exception is the western Transverse Ranges, where the dilatation is negative and the vertical field is positive. Note that artefacts in the dilatation rate in the synthetic examples (for example, Fig. 5b) raise caution in interpretation.

The rotation rate (Fig. 11d) is strongly influenced by the vertical velocity field (Fig. 10). We partition the rotation vector as $\mathbf{w} = \mathbf{w}_s + \mathbf{w}_t$ (Section 3.2.3), and the magnitudes of these vectors are plotted in Figs 11(d)–(f). We use the previous synthetic examples to interpret these patterns. For the $|\mathbf{w}_s|$ rotation-rate field (Fig. 10e), the maxima in the vicinity of the San Andreas fault are associated with apparent local (negative) rotations such as those demonstrated in Fig. 5(d). For the $|\mathbf{w}_t|$ rotation-rate field (Fig. 10f), the maxima in the vicinity of the San Joaquin basin and southern Sierra Nevada are associated with tilt-like rotations such as those demonstrated in Fig. 7. In other words, the rotation rate derived from the 3-D REASoN velocity field (Fig. 11d) captures signatures of both a right-lateral strike-slip fault as well as a contractional source.

We now examine the multiscale decomposition of the estimated 3-D velocity field. Figs 12(a)–(c) shows overall the estimated velocity field, separated into the up (r), south (θ), and east (ϕ) components. Figs 12(d)–(f) shows the estimated field using scales $q = 3$–5; this corresponds to the longest length-scale features. Figs 12(g)–(i) and (j)–(l) show the respective scales $q = 6$ and 7; the colour scales for the horizontal fields are reduced in order to see the details in the fields. Each scalar field in the top row of Fig. 12 is a sum of the four scalar fields beneath it, revealing the multiscale nature of the estimated field. The mask used in the figure is controlled by the posterior covariance.

It is clear from Fig. 12 that the vertical velocity field is dominated by shorter scale features. Most of the estimated vertical field in (a) is present in $q = 7$ (j) (note same colour scale), while most of the horizontal velocity field (b, c) is in scales $q = 3$–6 (e, f, h, i). This difference arises because the horizontal field captures plate motion that occurs over a scale of a few hundred kilometres, while the vertical field captures shorter-scale processes, including those that might be non-tectonic. At scale $q = 7$ (j, k, l), each component has a comparable amplitude, with the colour scale ranging as ± 5 mm yr^{-1}.

Once we have estimated the multiscale velocity field, we can readily compute other scalar quantities, such as dilatation, strain and rotation. There are two obvious possibilities for plotting the multiscale decomposition, one by cumulative scale (Fig. 13) and one by incremental scale (not shown). In both cases, the derivative maps illuminate gradients in the velocity field that are not immediately apparent from the velocity field maps themselves. For example, the addition of scale $q = 8$ in the estimation reveals several features that are not associated with the

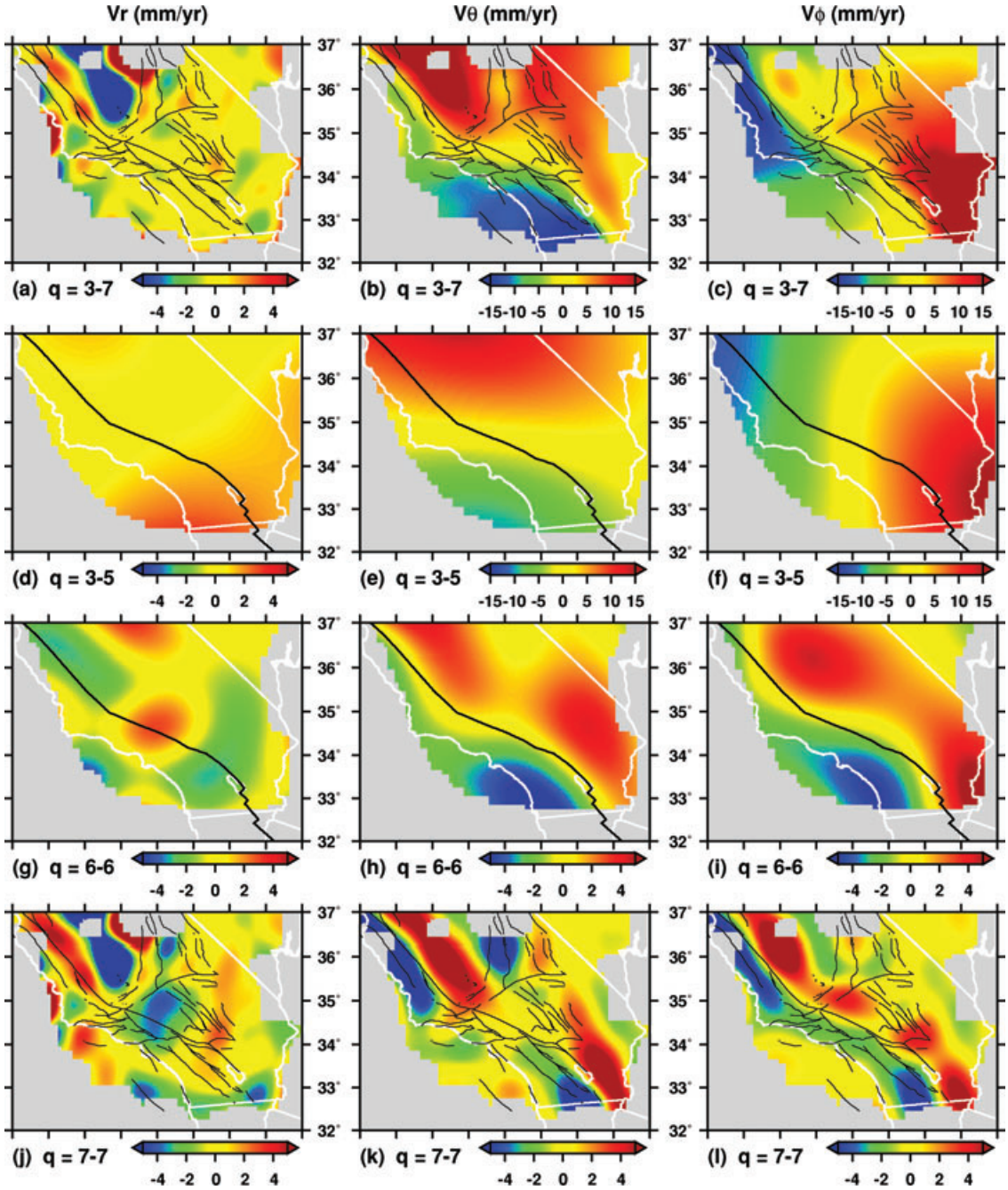

Figure 12. Multiscale estimation of REASoN velocity field in southern California. The columns correspond to the up (r), south (θ) and east (ϕ) components of the velocity field. Each figure in the top row (a–c) is the estimated field for scales $q = 3$–7, and each is the sum of the four fields plotted beneath it. The mask for each scale depends on the density of data points and thus increases for increasing scale level (i.e. decreasing length scale). See Section 4.2.3 for details.

San Andreas fault (Figs 13h and k). Furthermore, the rotation-rate field is concentrated only at the shortest scales ($q = 7$–8), whereas the strain-rate field illuminates the plate boundary at all scales.

Fig. 14 shows another way of viewing the multiscale estimation results, in this case for the horizontal field. As expected, as we allow shorter length-scale frame functions to be used in estimating the velocity field, the residual field vectors decrease in magnitude. The smallest residuals occur at observation points that are dense enough to fall within the support of the smallest length-scale frame functions, and also that have the smallest estimated uncertainties in the data. The largest residuals overall (Fig. 14d) are associated with the largest uncertainties in the data.

From the perspective of monitoring a GPS network, the residual map (Fig. 14d) may be helpful in detecting spurious behaviour of single stations. For example, a single observation near the Coso region (Fig. 9) is apparent in Figs 14(a)–(c), but disappears with the inclusion of scale $q = 8$. This is manifested by the strong dilatation-rate and strain-rate signals observed in the same region in Figs 13(j) and (k). Such an observation would warrant additional analysis of the GPS time-series and error estimate.

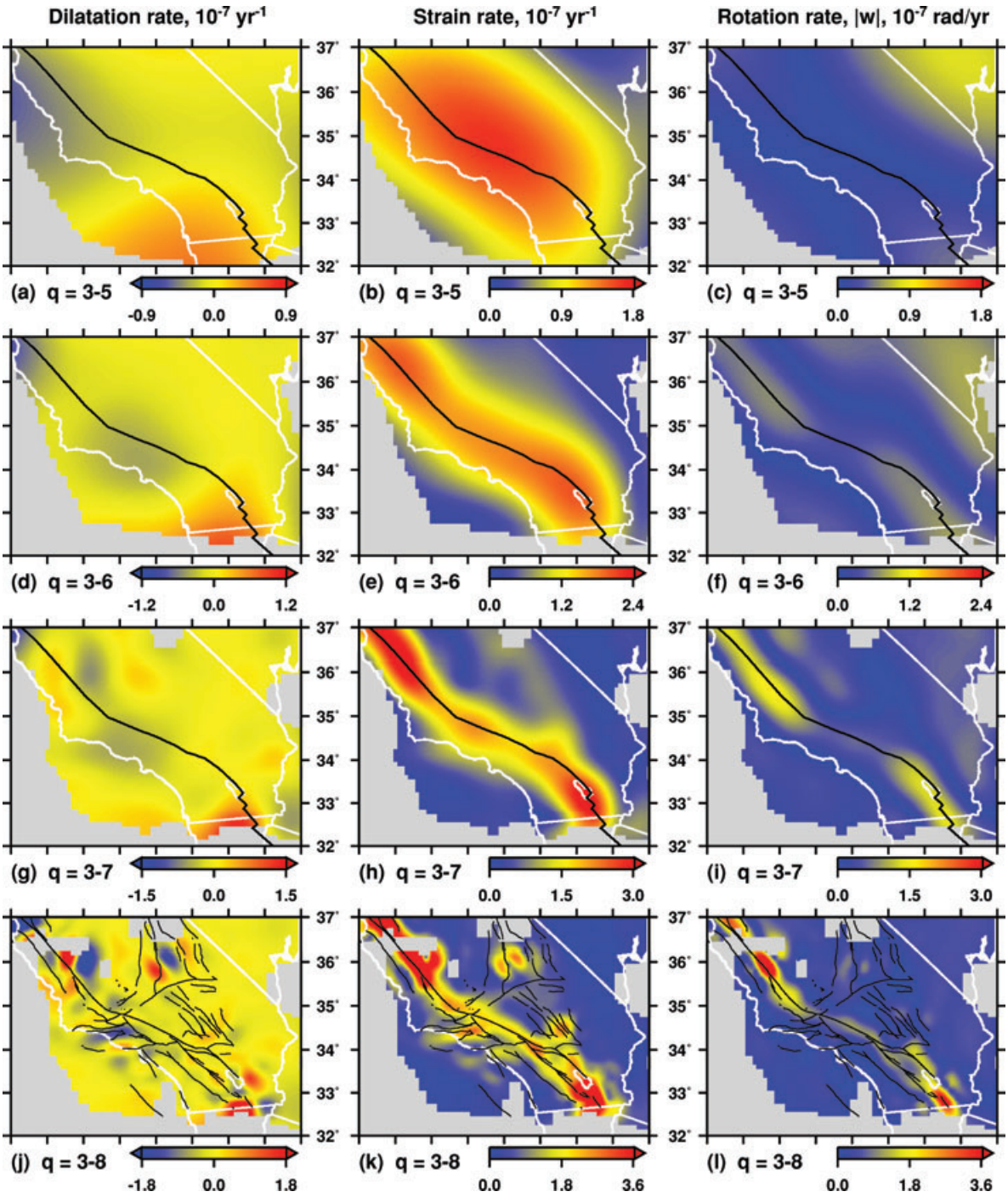

Figure 13. Multiscale scalar quantities derived from the velocity gradient for the estimated REASoN horizontal velocity field (Fig. 9), using scales $q = 3$–8. In each row, from top to bottom, we add an additional finer scale of spherical wavelets in the estimation of the velocity field. Note that the range of the colour scale increases in a linear manner. See Section 4.2.3 for details.

4.2.4 Parkfield region

We use a subset of the REASoN data set to illustrate the simple point that estimates of strain rate are scale dependent. In the Parkfield region, the density of stations allows us to include $q = 9$ in the estimation. We remove a rotational field from the subset to obtain the velocity field plotted in Fig. 15. We then estimate the horizontal velocity field, from which we compute the strain-rate field, which is shown in Fig. 15. The high density of stations near the fault is able to capture the spatial gradient in the velocity field, and gives rise to an estimated strain-rate of 1.5×10^{-6} yr^{-1}, which is in between the values reported in Ward (1998b, 0.1×10^{-6} yr^{-1}) and Shen-Tu *et al.* (1999, 3×10^{-6} yr^{-1}). The spatial gradients of the vertical velocity field are small compared with the gradients of the horizontal velocity field and do not influence the pattern in Fig. 15.

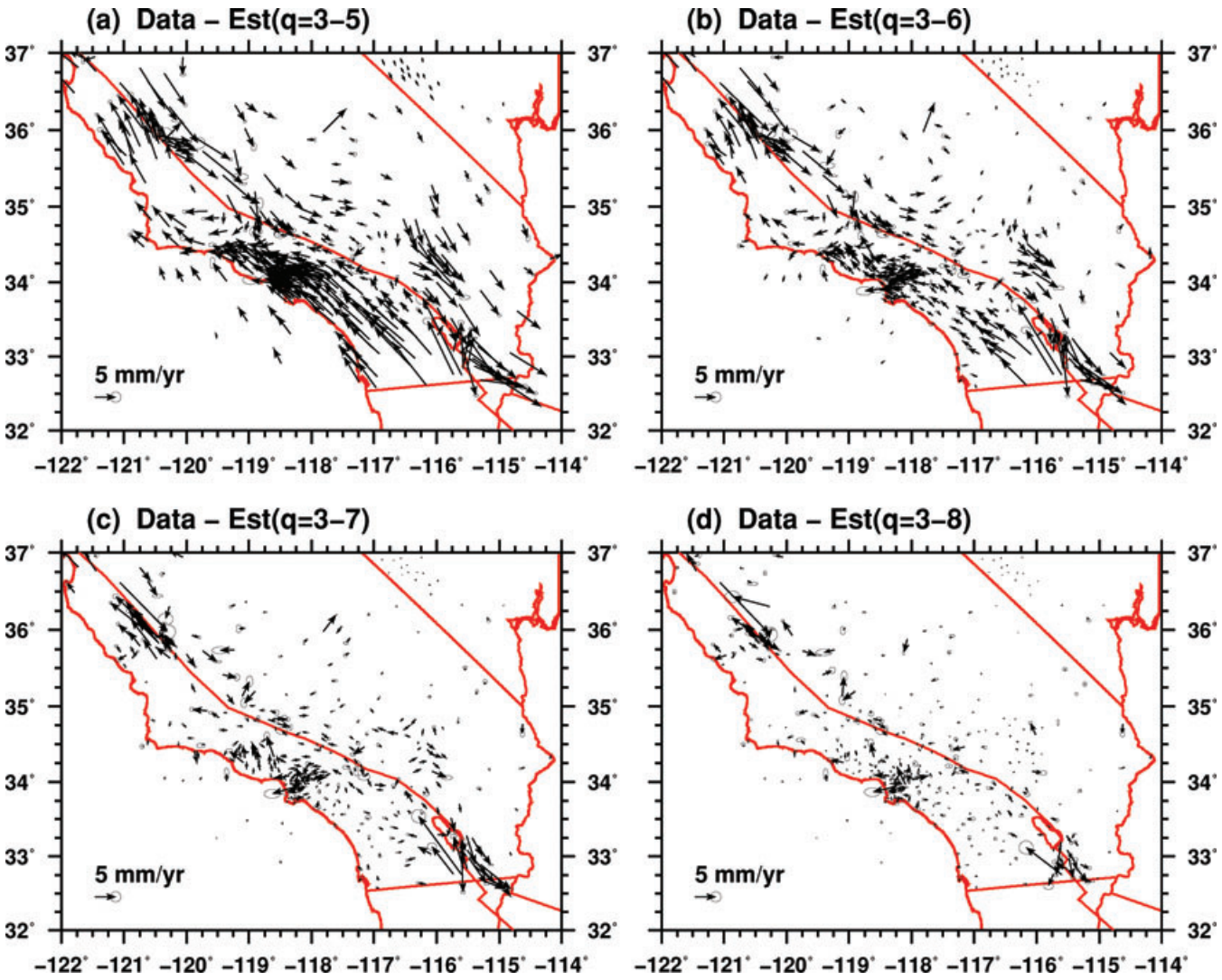

Figure 14. Residual horizontal velocity fields for the REASoN data set. The ellipses at the arrowheads do not change in each subplot. From (a) to (d), the minimum length scale in the estimated velocity field decreases as we increase the order of the grids from $q = 5$ to 8. As shorter length scales are included in the estimated field, the residuals decrease. See Fig. 12 and Section 4.2.3 for details.

5 DISCUSSION

5.1 Summary

We have presented both synthetic and real examples of multiscale estimations of velocity fields on the sphere. Our technique is well suited for dense geodetic networks and is straightforward to implement. First, we estimate a three-component velocity field using spherical wavelet frame functions of multiple scales. We then compute gradient quantities directly from the estimated field to identify potential deformation signals. The primary control on the estimation is the spacing among the stations. Where stations are dense, we allow for short-scale spherical wavelets in the estimation; where stations are sparse, we allow only long-scale spherical wavelets in the estimation. In some areas of our southern California example, such as the northeastern Mojave desert, we have the requisite station density to resolve small scale deformation signals, but we find nothing significant. In other areas, we do find such signals at small scales. For example, in the Parkfield area, we infer a strain-rate of 1.5×10^{-6} yr^{-1} across the fault (Fig. 15). Of course, without the concentration of stations in Parkfield, we would estimate a much lower strain rate and would not be able to distinguish among: (1) the fault creeping; (2) the fault locked at, say, 15 km depth and (3) distributed deformation that might be explained by a viscous model (e.g. Thatcher 1995).

In our multiscale estimation, the residual field between the original field and estimated field may reveal two key features (see Fig. 14). First, if there are systematic residuals in a particular region, then it is probable that one needs to include shorter-scale wavelets in the estimation. Second, if there is a strong residual at a single station, then the station is 'anomalous' and is either malfunctioning or is capturing a signal that is not spatially resolved.

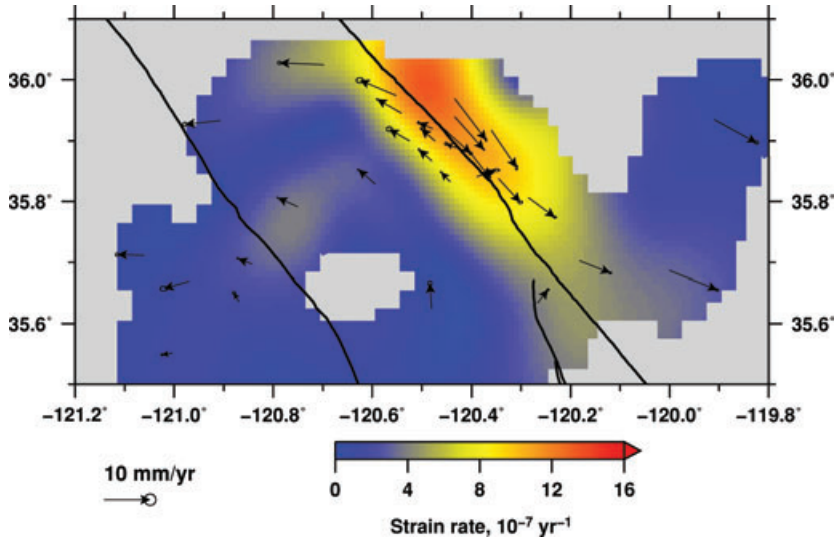

Figure 15. Strain-rate field derived from the multiscale estimation of the REASoN velocity field in the Parkfield region, using only scales $q = 6$–9. The estimated strain rate across the San Andreas fault is about 1.5×10^{-6} yr^{-1}. The superimposed velocity field observations, plotted as vectors, are obtained by first removing a uniform rotational field from the REASoN velocity field (e.g. Figs 4a–c). See Fig. 11(c) for context.

The synthetic examples (Section 4) illustrate several basic points. In the example with two dilatational sources, one much larger than the other, we saw the advantage of the separation of scales in identifying subtle signals. This example also suggests that derivative quantities based on observations from multiple sites will perform better in detecting signals than the individual components of the velocity field and that the vertical component, if available, should be used in estimating the velocity field, since deformation may not be predominantly in the horizontal directions.

With the real observations (Section 4.2), we show the usefulness of separating the three-component velocity field into constituent scales. The vertical velocity field contains spatially coherent features with length scales typically less than those observed in the horizontal velocity field, which reveal deformation across a broad plate boundary zone. The shortest length scales due to the horizontal deformation are found in the regions of Parkfield and Salton Sea, as previously established (e.g. Ward 1998b). Although some of the vertical velocity observations are certainly influenced by non-tectonic effects, they should still be included—along with proper uncertainty estimates—in the monitoring of signals in dense geodetic networks, especially in areas with larger expected vertical motions, such as subduction zones.

5.2 Regularization

Regularization is required to obtain a smooth estimated velocity field from the discrete observations. This is achieved through two possible actions. First, one can cull the set of possible spherical wavelets based on the coverage of observations. If each spherical wavelet has a sufficient number of observations constraining its coefficient, then no regularization is needed ($\lambda = 0$). Second, if *all* spherical wavelets are used for the inverse problem (eq. 15), then extensive regularization will be needed, since most wavelets will have zero observations constraining their corresponding coefficients. In our approach, we have chosen something in between these two 'end-members,' where we at the outset eliminate many candidate spherical wavelets based on data coverage (Fig. 3), but we still require a moderate amount of explicit regularization in the inversion (Fig. 2).

5.3 Towards multiscale time-dependent event detection

Our ultimate objective is to monitor time-dependent signals in dense GPS networks. Because our emphasis is on continuously recording GPS networks, our station coverage is not as dense as many published velocity fields, which include campaign GPS measurements (e.g. Shen *et al.* 2006; McCaffrey *et al.* 2007). In this study, we have only dealt with the spatial part of the problem, showing that the multiscale representation is well suited to identifying and characterizing geophysical signals of all scales. It also has the potential capability of removing scale-specific noise. This approach is a step towards global multiscale monitoring of time-dependent GPS displacement fields, in hopes of efficient and accurate characterization of Earth's surface deformation and the detection of geophysically important phenomena.

ACKNOWLEDGMENTS

We are grateful to John Haines, an anonymous reviewer, and editor John Beavan for comments that improved this manuscript. We thank Jean-Philippe Avouac for helpful discussions. We acknowledge the Southern California Integrated GPS Network and its sponsors, the W.M. Keck Foundation, NASA, NSF, USGS and SCEC, for providing data used in this study. This research was supported in part by the Gordon and Betty Moore Foundation. This is Caltech Tectonic Observatory Contribution 112.

REFERENCES

Allmendinger, R.W., Reilinger, R. & Loveless, J., 2007. Strain and rotation rate from GPS in Tibet, Anatolia, and the Altiplano, *Tectonics,* **26,** TC3013, doi:10.1029/2006TC002030.

Altamimi, Z., Collilieux, X., Legrand, J., Garayt, B. & Boucher, C., 2007. ITRF2005: A new release of the International Terrestrial Reference Frame based on time series of station positions and Earth Orientation Parameters, *J. geophys. Res.,* **112,** B09401, doi:10.1029/2007JB004949.

Amante, C. & Eakins, B.W., 2008. ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis, National Geophysical Data Center, NESDIS, NOAA, U.S. Department of Commerce, Boulder, CO, USA.

Antoine, J.P. & Vandergheynst, P., 1999. Wavelets on the 2-sphere: a group-theoretical approach, *Appl. Comput. Harmonic Anal.,* **7**(3), 262–291.

Argus, D.F., Heflin, M.B., Peltzer, G., Crampé, F. & Webb, F.H., 2005. Interseismic strain accumulation and anthropogenic motion in metropolitan Los Angeles, *J. geophys. Res.,* **110,** B04401, doi:10.1029/2003JB002934.

Bayer, M., Freeden, W. & Maier, T., 2001. A vector wavelet approach to iono- and magnetospheric geomagnetic satellite data, *J. Atmos. Solar-Terr. Phys.,* **63,** 581–597.

Beavan, J. & Haines, J., 2001. Contemporary horizontal velocity and strain rate fields of the Pacific-Australian plate boundary zone through New Zealand, *J. geophys. Res.,* **106**(B1), 741–770.

Becker, T.W., Hardebeck, J.L. & Anderson, G., 2005. Constraints on fault slip rates of the southern California plate boundary from GPS velocity and stress inversions, *Geophys. J. Int.,* **160,** 634–650.

Bird, P., 2003. An updated digital model of plate boundaries, *Geochem. Geophys. Geosyst.,* **4,** 1–52.

Bogdanova, I., Vandergheynst, P., Antoine, J.R., Jacques, L. & Morvidone, M., 2005. Stereographic wavelet frames on the sphere, *Appl. Comput. Harmonic Anal.,* **19**(2), 223–252.

Bos, A.G., Spakman, W. & Nyst, M.C.J., 2003. Surface deformation and tectonic setting of Taiwan inferred from a GPS velocity field, *J. geophys. Res.,* **108**(B10), 2458, doi:10.1029/2002JB002336.

Chambodut, A., Panet, I., Mandea, M., Diament, M., Holschneider, M. & Jamet, O., 2005. Wavelet frames: an alternative to spherical harmonic representation of potential fields, *Geophys. J. Int.,* **163**(3), 875–899.

Chinnery, M.A., 1961. The deformation of the ground around surface faults, *Bull. seism. Soc. Am.,* **51**(3), 355–372.

Dahlke, S. & Maass, P., 1996. Continuous wavelet transforms with applications to analyzing functions on spheres, *J. Fourier Anal. Appl.,* **2**(4), 379–396.

Dong, D., Fang, P., Bock, Y., Prawirodirdjo, L., Webb, F., Kedar, S. & Lundgren, P., 2009. Secular vertical crustal deformation field in California and Nevada regions from GPS data analysis: 1. Hydrological and anthropogenic signals, *J. geophys. Res,* in preparation.

Dragert, H., Wang, K. & James, T.S., 2001. A silent slip event on the deeper Cascadia subduction interface, *Science,* **292,** 1525–1528.

Dziewonski, A. & Woodhouse, J.H., 1983. Studies of the seismic source using normal-mode theory, in *Earthquakes: Observation, Theory and Interpretation: Notes from the International School of Physics 'Enrico Fermi' (1982, Varenna, Italy),* Vol. LXXXV, pp. 45–137, eds Kanamori, H. & Boschi, E., North-Holland Pub., Amsterdam.

Fay, N.P., Bennett, R.A. & Hreinsdóttir, S., 2008. Contemporary vertical velocity of the central Basin and Range and uplift of the southern Sierra Nevada, *Geophys. Res. Lett.,* **35,** L20309, doi:10.1029/2008GL034949.

Feigl, K.L., King, R.W. & Jordan, T.H., 1990. Geodetic measurement of tectonic deformation in the Santa Maria fold and thrust belt, California, *J. geophys. Res.,* **95**(B3), 2679–2699.

Flesch, L.M., Holt, W.E., Silver, P.G., Stephenson, M., Wang, C.-Y. & Chan, W.W., 2005. Constraining the extent of crust–mantle coupling in central Asia using GPS, geologic, and shear wave splitting data, *Earth planet. Sci. Lett.,* **238,** 248–268.

Freeden, W. & Windheuser, U., 1996. Spherical wavelet transform and its discretization, *Adv. Comput. Math.,* **5**(1), 51–94.

Freeden, W. & Windheuser, U., 1997. Combined spherical harmonic and wavelet expansion—a future concept in earth's gravitational potential determination, *Appl. Comput. Harmonic Anal.,* **4,** 1–37.

Golub, G.H. & Van Loan, C.F., 1989. *Matrix Computations,* 2nd edn, Johns Hopkins Univ. Press, Baltimore, MD, USA.

Golub, G.H., Heath, M. & Wahba, G., 1979. Generalized cross-validation as a method for choosing a good ridge parameter, *Technometrics,* **21,** 215–223.

Guilloux, F., Faÿ, G. & Cardoso, J.-F., 2009. Practical wavelet design on the sphere, *Appl. Comput. Harmonic Anal.,* **26**(2), 143–160.

Haines, A.J. & Holt, W.E., 1993. A procedure for obtaining the complete horizontal motions within zones of distributed deformation from the inversion of strain rate data, *J. geophys. Res.,* **98**(B7), 12 057–12 082.

Hastie, T. & Loader, C., 1993. Local regression: automatic kernel carpentry, *Stat. Sci.,* **8**(2), 120–143.

Heki, K., 1997. Silent fault slip following an interplate thrust earthquake at the Japan Trench, *Nature,* **386,** 595–598.

Holschneider, M., 1996. Continuous wavelet transforms on the sphere, *J. Math. Phys.,* **37**(8), 4156–4165.

Holschneider, M., Chambodut, A. & Mandea, M., 2003. From global to regional analysis of the magnetic field on the sphere using wavelet frames, *Phys. Earth planet. Inter.,* **135,** 107–124.

Holt, W.E., Chamot-Rooke, N., Pichon, X.L., Haines, A.J., Shen-Yu, B. & Ren, J., 2000. Velocity field in Asia inferred from Quaternary fault slip rates and Global Positioning System observations, *J. geophys. Res.,* **105**(B8), 19 185–19 209.

Hsu, Y.-J., Yu, S.-B., Simons, M., Kuo, L.-C. & Chen, H.-Y., 2009. Interseismic crustal deformation in the Taiwan plate boundary zone revealed by GPS observations, seismicity, and earthquake focal mechanisms, *Tectonophysics,* in press.

Hudnut, K.W. *et al.*, 2002. Continuous GPS observations of postseismic deformation following the 16 October 1999 Hector Mine, California, earthquake (M_W 7.1), *Bull. seism. Soc. Am.,* **92**(4), 1403–1422.

Jennings, C.W., 1994. Fault activity map of California and adjacent areas, with locations and ages of recent volcanic eruptions, Calif. Div. Mines and Geology, Geologic Data Map No. 6, map scale 1:750,000.

Kovacevic, J. & Chebira, A., 2007a. Life beyond bases: the advent of frames (Part I), *IEEE Signal Process. Mag.,* **24**(4), 86–104.

Kovacevic, J. & Chebira, A., 2007b. Life beyond bases: the advent of frames (Part II), *IEEE Signal Process. Mag.,* **24**(5), 115–125.

Kreemer, C., Holt, W.E. & Haines, A.J., 2003. An integrated global model of present-day plate motions and plate boundary deformation, *Geophys. J. Int.,* **154,** 8–34.

Malvern, L.E., 1969. *Introduction to the Mechanics of a Continuous Medium,* Prentice-Hall, Upper Saddle River, NJ, USA.

Matheron, G., 1963. Principles of geostatistics, *Econ. Geol.,* **58,** 146–1266.

McCaffrey, R. *et al.*, 2007. Fault locking, block rotation and crustal deformation in the Pacific Northwest, *Geophys. J. Int.,* **169,** 1315–1340.

McGuire, J.J. & Segall, P., 2003. Imaging of aseismic fault slip transients recorded by dense geodetic networks, *Geophys. J. Int.,* **155,** 778–788.

McLaren, M.K., Hardebeck, J.L., van der Elst, N., Unruh, J., Bawden, G.W. & Blai, J.L., 2008. Complex faulting associated with the 22 December 2003 M_W 6.5 San Simeon, California, earthquake, aftershocks, and postseismic surface deformation, *Bull. seism. Soc. Am.,* **98**(4), 1659–1680.

Meade, B.J. & Hager, B.H., 2005. Block models of crustal motion in southern California constrained by GPS measurements, *J. geophys. Res.,* **110,** B03403, doi:10.1029/2004JB003209.

Meade, B.J., Hager, B.H., McClusky, S.C., Reilinger, R.E., Ergintav, S., Lenk, O., Barka, A. & Ozener, H., 2002. Estimates of seismic potential in the Marmara Sea region from block models of secular deformation constrained by Global Positioning System measurements, *Bull. seism. Soc. Am.,* **92,** 208–215.

Melbourne, T.I., Webb, F.H., Stock, J.M. & Reigber, C., 2002. Rapid postseismic transients in subduction zones from continuous GPS, *J. geophys. Res.,* **107**(B10), 2241, doi:10.1029/2001JB000555.

Menke, W., 1989. *Geophysical Data Analysis: Discrete Inverse Theory,* Academic Press, San Diego, CA, USA.

Miyazaki, S., McGuire, J.J. & Segall, P., 2003. A transient subduction zone slip episode in southwest Japan observed by the nationwide GPS array, *J. geophys. Res.,* **108**(B2), 2087, doi:10.1029/2001JB000456.

Miyazaki, S., Segall, P., Fukuda, J. & Kato, T., 2004. Space time distribution of afterslip following the 2003 Tokachi-oki earthquake: Implications for variations in fault zone frictional properties, *Geophys. Res. Lett.,* **31**, L06623, doi:10.1029/2003GL019410.

Mogi, K., 1958. Relations between the eruptions of various volcanoes and the deformations of the ground surfaces around them, *Bull. Earthq. Res. Inst.,* **36,** 99–134.

Nadin, E.S. & Saleeby, J.B., 2009. Quaternary reactivation of the Kern Canyon fault system, southern Sierra Nevada, California, USA, *Geol. Soc. Am. Bull.,* in press.

Oh, H.-S. & Li, T.-H., 2004. Estimation of global temperature fields from scattered observations by a spherical-wavelet-based spatially adaptive method, *J. R. Statist. Soc. B,* **66,** 221–238.

Ozawa, S., Murakami, M., Kaidzu, M., Tada, T., Sagiya, T., Hatanaka, Y., Yarai, H. & Nishimura, T., 2002. Detection and monitoring of the ongoing aseismic slip in the Tokai region, Central Japan, *Science,* **298,** 1009–1012.

Pollitz, F.F., Wicks, C. & Thatcher, W., 2001. Mantle flow beneath a continental strike-slip fault: Postseismic deformation after the 1999 Hector Mine earthquake, *Science,* **293,** 1814–1818.

Pritchard, M.E. & Simons, M., 2006. An aseismic slip pulse in northern Chile and along-strike variations in seismogenic behavior, *J. geophys. Res.,* **111**, B08405, doi:10.1029/2006JB0042580.

Reguzzoni, M., Sansó, F. & Venuti, G., 2005. The theory of general kriging, with applications to the determination of a local geoid, *Geophys. J. Int.,* **162,** 303–314.

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

Savage, J.C. & Langbein, J., 2008. Postearthquake relaxation after the 2004 *M*6 Parkfield, California, earthquake and rate-and-state friction, *J. geophys. Res.,* **113**, B10407, doi:10.1029/2008JB005723.

Savage, J.C. & Prescott, W.H., 1976. Strain accumulation on the San Jacinto fault near Riverside, California, *Bull. seism. Soc. Am.,* **66**(5), 1749–1754.

Shao, J., 1993. Linear model selection by cross-validation, *J. Am. Stat. Assoc.,* **88,** 486–494.

Shao, J., 1997. An asymptotic theory for linear model selection, *Stat. Sin.,* **7,** 221–242.

Shearer, P.M., 1999. *Introduction to Seismology,* Cambridge Univ. Press, Cambridge, UK.

Shen, Z.-K., Jackson, D.D. & Ge, B.X., 1996. Crustal deformation across and beyond the Los Angeles basin from geodetic measurements, *J. geophys. Res.,* **101**(B12), 27 957–27 980.

Shen, Z.-K. *et al.*, 2003. *The SCEC Crustal Motion Map*, version 3.0, available online at http://epicenter.usc.edu/cmm3/.

Shen, Z.-K., King, R.W., Agnew, D.C. & Wang, M., 2006. The California Crustal Motion Map, version 1.0, in *SCEC 2006 Annual Meeting Abstracts,* Palm Springs, Calif. (Sept. 11–13, 2006).

Shen-Tu, B., Holt, W.E. & Haines, A.J., 1999. Deformation kinematics in the western United States determined from Quaternary fault slip rates and recent geodetic data, *J. geophys. Res.,* **104**(B12), 28 927–28 955.

Silverman, B.W., 1986. *Density Estimation for Statistics and Data Analysis*, Vol. 26 of Monographs on Statistics and Applied Probability, Chapman and Hall, New York.

Simons, F.J., Dahlen, F.A. & Wieczorek, M.A., 2006. Spatiospectral concentration on a sphere, *SIAM Rev.,* **48**(3), 504–536.

Simons, M. & Hager, B.H., 1997. Localization of the gravity field and the signature of glacial rebound, *Nature,* **390,** 500–504.

Simons, M., Solomon, S.C. & Hager, B.H., 1997. Localization of gravity and topography: constraints on the tectonics and mantle dynamics of Venus, *Geophys. J. Int.,* **131,** 24–44.

Spakman, W. & Nyst, M.C.J., 2002. Inversion of relative motion data for estimates of the velocity gradient field and fault slip, *Earth planet. Sci. Lett.,* **203,** 577–591.

Tarantola, A., 2005. *Inverse Problem Theory and Methods for Model Parameter Estimation,* SIAM, Philadelphia, PA, USA.

Teza, G., Pesci, A. & Galgaro, A., 2008. Grid_strain and grid_strain3: software packages for strain field computation in 2D and 3D environments, *Comput. Geosci.,* **34**(9), 1142–1153.

Thatcher, W., 1995. Microplate versus continuum description of active tectonic deformation, *J. geophys. Res.,* **100**(B3), 3885–3894.

Wahba, G., 1981. Spline interpolation and smoothing on the sphere, *SIAM J. Sci. Stat. Comp.,* **2**(1), 5–16.

Walcott, R.I., 1973. Structure of the earth from glacio-isostatic rebound, *Annu. Rev. Earth planet. Sci.,* **1,** 15–37.

Wang, Z. & Dahlen, F.A., 1995. Spherical-spline parameterization of three-dimensional Earth, *Geophys. Res. Lett.,* **22,** 3099–3102.

Ward, S.N., 1998a. On the consistency of earthquake moment release and space geodetic strain rates: Europe, *Geophys. J. Int.,* **135,** 1011–1018.

Ward, S.N., 1998b. On the consistency of earthquake moment rates, geological fault data, and space geodetic strain: the United States, *Geophys. J. Int.,* **134,** 172–186.

Weisberg, S., 2005. *Applied Linear Regression,* 5th edn, Wiley, Hoboken, NJ, USA.

Wiaux, Y., Jacques, L. & Vandergheynst, P., 2005. Correspondence principle between spherical and Euclidean wavelets, *Astrophys. J.,* **632**(1), 15–28.

APPENDIX A: CONSTRUCTION OF A FRAME OF DOG SPHERICAL WAVELETS

The 'Difference of Gaussians' (DOG) spherical wavelet, of scale a, centred at pole $\mathbf{x} \in S^2$, defined by Bogdanova *et al.* (2005) is given by the following expression

$$\psi_{\mathbf{x},a}(\mathbf{x}') \;=\; \lambda_a(\gamma)^{1/2} \exp\left[-\frac{\tan^2(\gamma/2)}{a^2}\right] - \frac{1}{\alpha}\lambda_{\alpha a}(\gamma)^{1/2} \exp\left[-\frac{\tan^2(\gamma/2)}{(\alpha a)^2}\right], \tag{A1}$$

where $\alpha > 1$ is a fixed number, γ is the angle between poles $\mathbf{x}$ and $\mathbf{x}'$ and the parametric function λ is given by

$$\lambda_a(\gamma) \;=\; \frac{4a^2}{[(a^2-1)\cos\gamma + (a^2+1)]^2}. \tag{A2}$$

A DOG spherical wavelet $\psi_{\hat{\mathbf{z}},a}$ with parameter $\alpha = 1.25$, centred at the North Pole $\hat{\mathbf{z}}$, is plotted in Fig. 1(b), for three different scales. Fig. 1(c) shows the corresponding profiles, for a fixed longitude ϕ. Fig. 1(d) displays their spectra (i.e. the spherical harmonics coefficients $l \mapsto |\Psi_{\hat{\mathbf{z}},a}(l,0)|$, where $\Psi_{\hat{\mathbf{z}},a}(l,m) = \int_{S^2} \psi_{\hat{\mathbf{z}},a} Y_{l,m} d\Omega$).

A frame of DOG spherical wavelets can be obtained by appropriate discretization of position $\mathbf{x}$ and scale a. Let $\mathcal{G}_q$ denote a grid of vertices on the sphere, at a scale q. We are interested in defining a set of grids $\{\mathcal{G}_q, q = 0, 1, \ldots\}$ having the following properties.

(i) The geodesic distance between neighbouring vertices is nearly uniform. This provides nearly uniform coverage for a given scale.

(ii) $\mathcal{G}_q \subset \mathcal{G}_{q+1}$: vertices in the grid at scale q are also vertices in grids at finer scales. This is a requirement for multiscale analysis.

A grid generated by successive subdivisions of an initial icosahedron satisfies these properties. Given an initial icosahedron $\mathcal{G}_0$, each triangle is subdivided into four new equilateral triangles, and the new vertices are projected into the sphere. As examples, we show meshes $\mathcal{G}_2$, $\mathcal{G}_3$ and $\mathcal{G}_4$ in Fig. 1(a). The geometric properties of the spherical-triangular grids are summarized in Table 2. For each scale q, the average side arc-length $\bar{L}_q$ measures the sampling density on the sphere; the average sampling frequency is then given by $2\pi/\bar{L}_q$. Assuming that this sampling is uniform (the approximation is reasonable for our meshes), for a spherical function $f \in L(S^2)$ to be well sampled by the grid $\mathcal{G}_q$, it follows from Nyquist criterion that its bandwidth cannot exceed the value $\pi/\bar{L}_q$. The maximum bandwidth $\bar{l}_q = \text{int}(\pi/\bar{L}_q)$ associated to each grid scale level q is shown in the rightmost column of Table 2.

The frame of spherical wavelets will be defined as the set $\mathcal{F} = \{\psi_{\mathbf{x}_{(q,j)},a_q}(\mathbf{x}),\ \mathbf{x}_{(q,j)} \in \mathcal{G}_q,\ \ q \in \mathbb{N}\}$. The set is built by considering, for each grid $\mathcal{G}_q$, all the wavelets with scale a_q centred at its gridpoints. The choice of the scale a_q associated to each grid resolution q is not arbitrary. Let $\mathcal{B}_Q$ denote the set of functions of bandwidth $B \leq B_Q$. We want to define the set $\mathcal{F}_Q = \{\psi_{\mathbf{x}_{(q,j)},a_q}(\mathbf{x}),\ \mathbf{x}_{(q,j)} \in \mathcal{G}_q,\ q = 0, \ldots, Q\}$ in such a way that $\text{Span}(\mathcal{F}_Q) = \mathcal{B}_Q$. This requirement can only be satisfied if the union of the spectral support of all wavelets with scale a_q, $q \leq Q$, lies in the interval $[0, B_Q]$. Since we want this condition to be true for all $Q \geq 1$, it follows that the spectral support of the wavelets with scale a_q has to be the interval $[b_Q, B_Q]$, for some $b_Q \leq B_{Q-1}$. In other words, scale should be discretized in such a way that the family of wavelets in the set $\mathcal{F}$ covers the spectrum, and the cardinal of $\mathcal{F}$ is not too large. Alternatively, if we view the wavelets as interpolating functions, the wavelet scale a_q associated to the grid $\mathcal{G}_q$ cannot oscillate more than what is permitted by the grid resolution sampling frequency $\bar{l}_q$.

In the case of DOG spherical wavelets, the spectral support is not compact, but most of the energy (at least 99.0 per cent) is concentrated in a relatively small number of spherical harmonic coefficients, corresponding to the lowest spherical harmonics degrees (see Fig. 1d). Table 3 shows, for DOG wavelets with scales $a_q = 2^{-q}$, $q = 0, \ldots, 10$, the spherical harmonic degrees at which the energy reaches 99.0 per cent and 99.9 per cent of the total energy. By comparing these values with the value $\bar{l}_q$ in the rightmost column in Table 2, one can see that the choice $a_q = 2^{-q}$ is well adapted to the grid resolution. Fig. 1(d) also illustrates the good spectral coverage obtained with this scale discretization.

APPENDIX B: SELECTION OF THE REGULARIZATION PARAMETER

The least-squares solution via Tikhonov regularization to the overdetermined problem is given by

$$\mathbf{m}' = \left(\mathbf{G}'^T\mathbf{G}' + \lambda^2\mathbf{I}\right)^{-1}\mathbf{G}'^T\mathbf{d}'. \tag{B1}$$

We review that (B1) is a general form of (13). Using the substitutions

$$\mathbf{m}' = \mathbf{S}^{1/2}\,\mathbf{m}, \tag{B2}$$

$$\mathbf{G}' = \mathbf{Z}^{1/2}\,\mathbf{G}\,\mathbf{S}^{-1/2}, \tag{B3}$$

$$\mathbf{d}' = \mathbf{Z}^{1/2}\,\mathbf{d}, \tag{B4}$$

where $\mathbf{S}$ and $\mathbf{Z}$ are symmetric, square-root invertible matrices, we transform (B1) into (13)

$$\mathbf{S}^{1/2}\,\mathbf{m} = [(\mathbf{Z}^{1/2}\,\mathbf{G}\,\mathbf{S}^{-1/2})^T(\mathbf{Z}^{1/2}\,\mathbf{G}\,\mathbf{S}^{-1/2}) + \lambda^2\mathbf{I}]^{-1}(\mathbf{Z}^{1/2}\,\mathbf{G}\,\mathbf{S}^{-1/2})^T\mathbf{Z}^{1/2}\,\mathbf{d} \tag{B5}$$

$$\mathbf{m} = \mathbf{S}^{-1/2}[\mathbf{S}^{-1/2}\,\mathbf{G}^T\,\mathbf{Z}\,\mathbf{G}^T\,\mathbf{S}^{-1/2} + \lambda^2\mathbf{I}]^{-1}\mathbf{S}^{-1/2}\,\mathbf{G}^T\,\mathbf{Z}\,\mathbf{d} \tag{B6}$$

$$= [\mathbf{S}^{1/2}(\mathbf{S}^{-1/2}\,\mathbf{G}^T\,\mathbf{Z}\,\mathbf{G}^T\,\mathbf{S}^{-1/2} + \lambda^2\mathbf{I})\mathbf{S}^{1/2}]^{-1}\mathbf{G}^T\,\mathbf{Z}\,\mathbf{d} \tag{B7}$$

$$= (\mathbf{G}^T\,\mathbf{Z}\,\mathbf{G}^T + \lambda^2\,\mathbf{S})^{-1}\mathbf{G}^T\,\mathbf{Z}\,\mathbf{d}. \tag{B8}$$

Making the substitution $\mathbf{Z} = \mathbf{C}_\text{D}^{-1}$, we obtain (13). Thus, the inclusion of the weighting matrix for observations and the regularization matrix for the model parameters is equivalent to the solution in (B1) with the transformations in (B2)–(B4). In other words, the selection of λ may be discussed in the context of (B1) rather than (13).

Our choice of the regularization parameter is based on cross-validation (Golub *et al.* 1979; Shao 1993, 1997), which assume Gaussian uncertainties of the observations, and is based upon the fact that a good regularization parameter should provide good estimates for withheld data. We used 'ordinary cross-validation', also known as 'leave-one-out cross-validation' or 'delete-1 cross-validation.' This involves removing one observation, computing a model vector, and then computing the cross-validated residual, which is the residual between the value at the removed point and the estimated value at the removed point. It turns out that n separate regressions are not needed to compute the cross-validation prediction error (e.g. Weisberg 2005, p. 291-292). The ith cross-validation residual can be written as

$$\hat{e}_{(i)} \ \equiv\ d_i - \hat{d}_{(i)} \ =\ \frac{d_i - \hat{d}_i}{1 - N_{ii}(\lambda)}, \tag{B9}$$

where

$$\mathbf{N}(\lambda) = \mathbf{G}(\mathbf{G}^T\mathbf{G} + \lambda^2\mathbf{I})^{-1}\mathbf{G}^T \tag{B10}$$

is the regularized data resolution matrix for the overdetermined least-squares problem (e.g. Menke 1989), λ is the regularization parameter, $\hat{d}_i$ is the estimated value from a model obtained by using all n observations, and $\hat{d}_{(i)}$ is the estimated value from a model obtained by using $n-1$ observations.

The ordinary cross-validation function is defined in terms of eq. (B9) as

$$H(\lambda) = \frac{1}{n}\sum_{i=1}^{n} \hat{e}_{(i)}^2 = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{d_i - \hat{d}_i}{1 - N_{ii}(\lambda)}\right]^2. \tag{B11}$$

For our examples, the curve $H(\lambda)$ provided a single minimum from which we selected the regularization parameter. Fig. 2 shows $\log_{10} H(\lambda)$ for three components of the velocity field.

APPENDIX C: DEFORMATION GRADIENT FOR ELASTIC RHEOLOGY

Here we show the steps used to derive eqs (20) and (21). For an isotropic linear elastic solid, the stress tensor $\boldsymbol{\sigma} = \mathbf{c} : \boldsymbol{\varepsilon}$ reduces to (e.g. Shearer 1999, eq. 2.24)

$$\boldsymbol{\sigma} = \begin{bmatrix} \lambda \,\mathrm{Tr}(\boldsymbol{\varepsilon}) + 2\mu\varepsilon_{11} & 2\mu\varepsilon_{12} & 2\mu\varepsilon_{13} \\ \sigma_{12} & \lambda \,\mathrm{Tr}(\boldsymbol{\varepsilon}) + 2\mu\varepsilon_{22} & 2\mu\varepsilon_{23} \\ \sigma_{13} & \sigma_{23} & \lambda \,\mathrm{Tr}(\boldsymbol{\varepsilon}) + 2\mu\varepsilon_{33} \end{bmatrix}. \tag{C1}$$

Combining the constraints in eq. (19) with eq. (C1), and solving for the strain components, we obtain

$$\varepsilon_{11} = F\,(\varepsilon_{22} + \varepsilon_{33}), \qquad \varepsilon_{12} = 0, \qquad \varepsilon_{13} = 0. \tag{C2}$$

where F is a constant (eq. 22). Hence, the strain tensor for an isotropic solid is given by

$$\boldsymbol{\varepsilon}_0 = \begin{bmatrix} F\,(\varepsilon_{22} + \varepsilon_{33}) & 0 & 0 \\ 0 & \varepsilon_{22} & \varepsilon_{23} \\ 0 & \varepsilon_{23} & \varepsilon_{33} \end{bmatrix}, \tag{C3}$$

where the 0-subscript denotes that it is evaluated at the surface.

Considering a time increment, Δt, we have $\boldsymbol{\varepsilon} \to \mathbf{D}$. With these assumptions, for an observation point at the surface ($r = R$), eq. (16) reduces to eq. (20), where we used the constraints

$$\frac{\partial v_r}{\partial r} = F\left[\frac{1}{R}\left(v_r + \frac{\partial v_\theta}{\partial \theta}\right) + \frac{1}{R}\left(v_r + v_\theta \cot\theta + \frac{1}{\sin\theta}\frac{\partial v_\phi}{\partial \phi}\right)\right],$$

$$\frac{1}{2R}\left(-v_\theta + \frac{\partial v_r}{\partial \theta} + R\frac{\partial v_\theta}{\partial r}\right) = 0,$$

$$\frac{1}{2R}\left(-v_\phi + \frac{1}{\sin\theta}\frac{\partial v_r}{\partial \phi} + R\frac{\partial v_\phi}{\partial r}\right) = 0,$$

which can be solved for the radial gradients (at $r = R$) to obtain

$$\frac{\partial v_r}{\partial r} = \frac{F}{R}\left(2v_r + v_\theta \cot\theta + \frac{\partial v_\theta}{\partial \theta} + \frac{1}{\sin\theta}\frac{\partial v_\phi}{\partial \phi}\right), \tag{C4}$$

$$\frac{\partial v_\theta}{\partial r} = \frac{1}{R}\left(v_\theta - \frac{\partial v_r}{\partial \theta}\right), \tag{C5}$$

$$\frac{\partial v_\phi}{\partial r} = \frac{1}{R}\left(v_\phi - \frac{1}{\sin\theta}\frac{\partial v_r}{\partial \phi}\right). \tag{C6}$$

Substituting these equations into eq. (17), the rotation-rate tensor (at $r = R$) reduces to eq. (21).

APPENDIX D: MATRIX NORM

We denote the $N \times N$ tensor $\mathbf{T}$ in its eigenbasis as

$$\mathbf{T} = \mathbf{V}\boldsymbol{\Delta}\mathbf{V}^{-1}, \tag{D1}$$

where $\mathbf{V}$ contains the eigenvectors of $\mathbf{T}$, and $\boldsymbol{\Delta}$ is a diagonal matrix of corresponding eigenvalues λ_k. Using the property $\mathrm{Tr}(\mathbf{AB}) = \mathrm{Tr}\,(\mathbf{BA})$ with eq. (D1), one can obtain

$$\mathrm{Tr}(\mathbf{T}^p) = \sum_{k=1}^{N} \lambda_k^p. \tag{D2}$$

Because the eigenvalues of tensor $\mathbf{T}$ are invariant to coordinate changes, the operation Tr ($\mathbf{T}^p$) is also invariant for any p.

A common matrix norm is the Frobenius norm, defined as (Golub & Van Loan 1989, p. 56)

$$\|\mathbf{T}\|_{\mathrm{F}} = \left(\sum_{i=1}^{N} \sum_{j=1}^{N} |T_{ij}|^2 \right)^{1/2} . \tag{D3}$$

If $\mathbf{T}$ is real-valued, then

$$\|\mathbf{T}\|_{\mathrm{F}}^2 = \mathbf{T} : \mathbf{T}, \tag{D4}$$

where the ':' denotes the scalar product of two tensors (Malvern 1969, p. 35)

$$\mathbf{T} : \mathbf{T} = \sum_{i=1}^{N} \sum_{j=1}^{N} T_{ij}\, T_{ij}. \tag{D5}$$

If $\mathbf{T}$ is symmetric, then

$$\|\mathbf{T}\|_{\mathrm{F}}^2 = \mathbf{T} : \mathbf{T} = tr(\mathbf{TT}) = \sum_{k=1}^{N} \lambda_k^2. \tag{D6}$$