Missouri University of Science and Technology

Scholars' Mine

Geosciences and Geological and Petroleum Engineering Faculty Research & Creative Works

Geosciences and Geological and Petroleum Engineering

01 Sep 2023

Quantification Of Geodetic Strain Rate Uncertainties And Implications For Seismic Hazard Estimates

Jeremy Maurer
Missouri University of Science and Technology, jlmd9g@mst.edu

Kathryn Materna

Follow this and additional works at: https://scholarsmine.mst.edu/geosci_geo_peteng_facwork

 Part of the Geological Engineering Commons, and the Petroleum Engineering Commons

Recommended Citation

J. Maurer and K. Materna, "Quantification Of Geodetic Strain Rate Uncertainties And Implications For Seismic Hazard Estimates," *Geophysical Journal International*, vol. 234, no. 3, pp. 2128 - 2142, Oxford University Press; Royal Astronomical Society, Sep 2023.
The definitive version is available at https://doi.org/10.1093/gji/ggad191

This Article - Journal is brought to you for free and open access by Scholars' Mine. It has been accepted for inclusion in Geosciences and Geological and Petroleum Engineering Faculty Research & Creative Works by an authorized administrator of Scholars' Mine. This work is protected by U. S. Copyright Law. Unauthorized use including reproduction for redistribution requires the permission of the copyright holder. For more information, please contact scholarsmine@mst.edu.

Geophysical Journal International

Geophys. J. Int. (2023) **234,** 2128–2142
Advance Access publication 2023 May 05
GJI Seismology

https://doi.org/10.1093/gji/ggad191

Quantification of geodetic strain rate uncertainties and implications for seismic hazard estimates

Jeremy Maurer[1] and Kathryn Materna[2]

[1]*Missouri University of Science and Technology, Geological Sciences and Geological and Petroleum Engineering, 1400 N Bishop, Rolla, MO 65401, USA. E-mail: jmaurer@mst.edu*
[2]*U.S. Geological Survey, Earthquake Science Center, Moffett Field, CA* 94035, USA

Accepted 2023 May 2. Received 2023 March 16; in original form 2022 July 9

SUMMARY

Geodetic velocity data provide first-order constraints on crustal surface strain rates, which in turn are linked to seismic hazard. Estimating the 2-D surface strain tensor everywhere requires knowledge of the surface velocity field everywhere, while geodetic data such as Global Navigation Satellite System (GNSS) only have spatially scattered measurements on the surface of the Earth. To use these data to estimate strain rates, some type of interpolation is required. In this study, we review methodologies for strain rate estimation and compare a suite of methods, including a new implementation based on the geostatistical method of kriging, to compare variation between methods with uncertainty based on one method. We estimate the velocity field and calculate strain rates in southern California using a GNSS velocity field and five different interpolation methods to understand the sources of variability in inferred strain rates. Uncertainty related to data noise and station spacing (aleatoric uncertainty) is minimal where station spacing is dense and maximum far from observations. Differences between methods, related to epistemic uncertainty, are usually highest in areas of high strain rate due to differences in how gradients in the velocity field are handled by different interpolation methods. Parameter choices, unsurprisingly, have a strong influence on strain rate field, and we propose the traditional *L*-curve approach as one method for quantifying the inherent trade-off between fit to the data and models that are reflective of tectonic strain rates. Doing so, we find total variability between five representative strain rate models to be roughly 40 per cent, a much lower value than roughly 100 per cent that was found in previous studies (Hearn *et al.*). Using multiple methods to tune parameters and calculate strain rates provides a better understanding of the range of acceptable models for a given velocity field. Finally, we present an open-source Python package (Materna *et al.*) for calculating strain rates, Strain_2D, which allows for the same data and model grid to be used in multiple strain rate methods, can be extended with other methods from the community, and provides an interface for comparing strain rate models, calculating statistics and estimating strain rate uncertainty for a given GNSS data set.

Key words: Plate motions; Satellite geodesy; Seismic cycle; Statistical methods; Earthquake hazards; Kinematics of crustal and mantle deformation.

1 INTRODUCTION

Space geodesy has transformed the study of tectonic processes by providing direct measurement of surface strain rates (e.g. Ward 1994; Shen *et al.* 1996; Shen-Tu *et al.* 1999; Kreemer *et al.* 2000; Beavan & Haines 2001; Shen *et al.* 2007; Wu *et al.* 2011; Haines & Wallace 2020). Strain rate has been linked to seismicity rate and moment accumulation rate (e.g. Kostrov 1974; Savage & Simpson 1997; Stevens & Avouac 2016; Wu *et al.* 2021); thus, robust estimation of the surface strain rate tensor provides valuable constraints on seismic hazard.

Despite the growing availability of surface velocity data from geodetic data such as Global Navigation Satellite Systems (GNSS), uncertainty still exists when computing surface strain rates that are related to tectonic deformation. This is because station locations are not distributed uniformly in space, and in many places are sparse relative to deformation wavelengths. The result is that some form of interpolation is required to calculate surface strain rate at every point in space. Numerous methods and algorithms have been proposed for this purpose, but few published studies to date have systematically explored the differences between methods and the consequent implications for quantifying crustal strain rate and

© The Author(s) 2023. Published by Oxford University Press on behalf of The Royal Astronomical Society. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from https://academic.oup.com/gji/article/234/3/2128/7153326 by Missouri University of Science & Technology user on 16 August 2023

seismic hazard. Approaches can be broken into several broad classes of methods:

(i) Direct calculation of strain from baselines, typically involving some type of triangulation like Delaunay (e.g. Savage *et al.* 2001; Cai & Grafarend 2007). This method was the first method developed for strain rate calculation (e.g. Savage & Burford 1970).

(ii) Fitting the velocity field using various basis functions such as polynomials, splines, wavelets, etc. (e.g. Feigl 1993; Haines & Holt 1993; Beavan & Haines 2001; Hackl 2009; Tape *et al.* 2009; Wang & Wright 2012; Weiss *et al.* 2020; Okazaki *et al.* 2021).

(iii) Fitting based on elasticity theory (e.g. Noda & Matsu'ura 2010; Haines *et al.* 2015; Sandwell & Wessel 2016). This class of methods could be considered a subset of (2), using the elasticity functions ($\ln(r)$, $1/r^2$) as basis functions. The log functions do not have compact support like the other methods in (3), which can lead to growing solutions outside the domain of the data, similar to polynomial interpolators.

(iv) Neighbourhood methods that use either a locally weighted mean or local regression around an estimation point (e.g. Shen *et al.* 1996, 2015; Tarayoun *et al.* 2018; Handwerger *et al.* 2019; Huang *et al.* 2022). This includes methods that weight based on spatial covariance (e.g. least-squares collocation, LSC, El-Fiky *et al.* 1998; Kato *et al.* 1998; Goudarzi *et al.* 2015; Qu *et al.* 2019).

Any method for spatial interpolation requires a decision about at least three factors (Fig. 1): (1) how to handle measurement noise or error in the observations; that is, the 'point-wise' uncertainty, (2) the maximum spatial length scale over which observations are correlated, which in general could vary in direction and space and (3) the behaviour of the interpolated velocity field at short distances; that is, degree of smoothness. Note that (2) and (3) represent the endmember behaviours; that is, (2) is the length scale beyond which observations are not correlated, while (3) relates to near-field behaviour (Fig. 1). Many published methods require an *ad hoc* decision about these parameters, and/or do not report how they decided on a particular set of parameters. The non-uniqueness of the interpolated velocity field implies a trade-off between higher smoothness, which is often more consistent with tectonic strain rates, and fit to the data. Methods for determining model parameters that produce a reasonable range of these metrics include the so-called '*L*-curve' (Hansen 1992), cross-validation and Akaike's Information Criterion (Sakamoto *et al.* 1986), among others, but are not often employed in studies of strain rate.

The different assumptions used by different strain rate methods present a challenge for side-by-side comparison. Few studies present side-by-side comparison of methods, and as different studies use different data sets, it is difficult to construct a consistent picture of method performance. Wu *et al.* (2011) compared LSC with Delaunay triangulation and spline fitting and found that LSC performed best in synthetic tests. Delaunay suffers from noise artefacts if the input data contain outliers, since strain is calculated directly from baselines between stations (Savage *et al.* 2001). Hearn *et al.* (2010) and Sandwell *et al.* (2016) compared a suite of 17 strain rate methods in southern California and found that variability between methods was up to 100 per cent of the signal in some places; however, in that case, the data sets used in each method were different. A follow-up analysis of the same 17 models by Xu *et al.* (2021) showed that on average all models were highly correlated at long wavelengths (>100 km) and approached a correlation of zero at lengths scales of roughly 30 km. These studies compared both model-based and data-based methods for calculating strain rates; however, model-based strain rates are usually limited to explaining strain accumulating on faults, and the presence of significant off-fault strain (e.g. Johnson 2013) could result in a bias between model- and data-based methods.

Not only do different methods make different assumptions about how data are correlated, but they also estimate the interpolated field differently. As discussed further in the next section, some of the methods listed above solve a weighted least-squares problem for a locally linear velocity field, others use a weighted average of the existing data and solve for optimal weights based on the spatial structure function (Fig. 1), and others solve an even-determined problem for body force magnitude at each GNSS station. These lead to further differences in how data noise, sparse data areas and strong velocity gradients (i.e. high strain rates) are handled in the interpolation scheme. A completely different approach was proposed by Pagani *et al.* (2021) to constrain strain rate uncertainties using a hierarchical Bayesian approach. They use the Delaunay method and sample over many possible data grid node configurations, and by averaging over many thousands of models they obtain smooth strain rate fields. However, this approach is extremely computationally demanding (Pagani *et al.* 2021), and may lose short wavelength information due to averaging many thousands of models.

In this study, we present a side-by-side comparison of several representative strain rate methods, including representatives for each of the broad classes listed above, for the well-studied region of southern California. We use a consistent set of velocity data for southern California (see Data and Resources) to highlight the similarities and differences between methods. In this study, we restrict our analysis to data-centred methods that do not rely on specifying a fault model. We discuss issues related to differentiating between tectonic and non-tectonic strain rate, comparing methods using different parametrizations and how methodological differences impact the estimated strain rate field and propose some best practices. We present a new open-source software tool, Strain_2D (Materna *et al.* 2021), and show how it can be used for consistent strain rate calculation and comparison.

2 METHODOLOGY

2.1 Strain rate for seismic hazard

The surface strain rate tensor is a symmetric, positive-definite tensor with three independent components represented by

$$\dot{\varepsilon}_{ij} = 0.5\left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i}\right) \tag{1}$$

where i and j represent components of coordinates x and velocities v, respectively. Here, we focus on surface strain rates, so $i,j = 1,2$. The principal strain rates $\dot{\varepsilon}_1$ and $\dot{\varepsilon}_2$ are given by the eigenvalues of the strain rate tensor:

$$\dot{\varepsilon}_1, \dot{\varepsilon}_2 = \frac{(\dot{\varepsilon}_{11} + \dot{\varepsilon}_{22})}{2} \pm \sqrt{\left(\frac{\dot{\varepsilon}_{11} - \dot{\varepsilon}_{22}}{2}\right)^2 + (\dot{\varepsilon}_{12})^2} \tag{2}$$

Because strain rate is a tensor, invariants are often used for plotting. Of particular importance are the trace of the strain rate tensor, which is defined as the areal strain rate or dilatation rate (rate of change in area):

$$\dot{\Delta}A = (\dot{\varepsilon}_{11} + \dot{\varepsilon}_{22}) = \dot{\varepsilon}_1 + \dot{\varepsilon}_2 \tag{3}$$

the maximum shear strain rate:

$$\dot{\tau} = \frac{1}{2}\sqrt{(\dot{\varepsilon}_{11} - \dot{\varepsilon}_{22})^2 + 4\dot{\varepsilon}_{12}^2} = \frac{1}{2}(\dot{\varepsilon}_1 - \dot{\varepsilon}_2) \tag{4}$$

Downloaded from https://academic.oup.com/gji/article/234/3/2128/7153326 by Missouri University of Science & Technology user on 16 August 2023

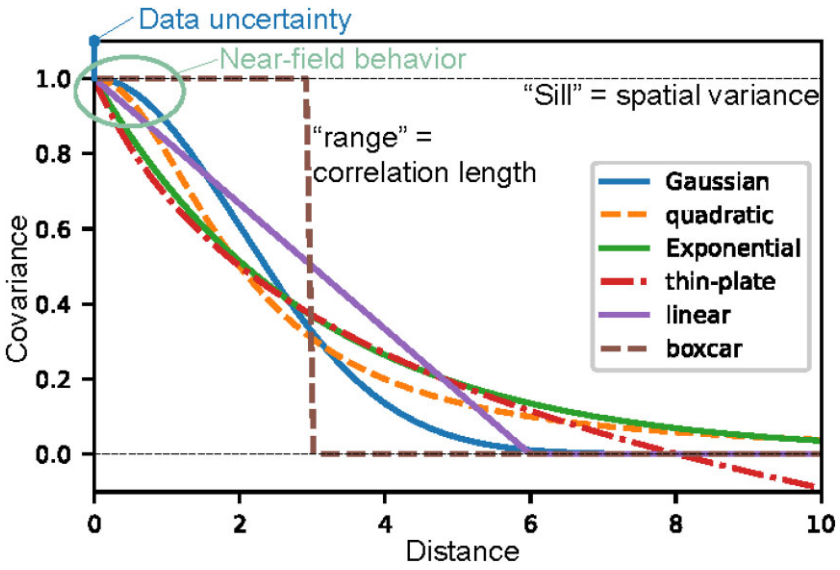

Figure 1. Parameters and/or decisions required for spatial interpolation. The different covariance models represent different weight functions that can be used to estimate a velocity at a new location. Several of the covariance models are used in the strain-rate methods considered in this study (see Table 1 for details). Specifically, Gaussian model: *geostats*/*visr*; quadratic model: *visr*; Exponential model: *geostats*; thin-plate model: *gpsgridder*; boxcar model: *Local average gradient*.

and the second invariant, a measure of the overall magnitude of strain:

$$I_2 = \frac{1}{2}\left[(\dot{\varepsilon}_{11} * \dot{\varepsilon}_{22}) - \dot{\varepsilon}_{12}^2\right] \tag{5}$$

With the goal of relating strain rate to seismic hazard, Savage & Simpson (1997) derived a relation between the strain rate and the minimum seismic moment accumulation rate assuming a fixed seismogenic depth H and shear modulus G based on Kostrov (1974):

$$\dot{M}_d \geq 2GHA_s\,|\dot{\varepsilon}^{\max}| \tag{6}$$

where A_s is the area at the surface of the crustal volume considered and $|\dot{\varepsilon}^{\max}| = \max\{|\dot{\varepsilon}_1|, |\dot{\varepsilon}_2|, |\dot{\varepsilon}_1 + \dot{\varepsilon}_2|\}$ is the so-called 'maximum principal strain rate.' Eq. (6) provides a lower bound on the seismic moment accumulation rate (Savage & Simpson 1997). Note that if detailed spatial information on G and H over the region of interest is available, eq. (6) can be extended to a volume integral to use these data.

2.2 Strain rate estimation methods

2.2.1 Strain_2D package

We implement several strain rate methods in the open-source package Strain_2D (Materna *et al.* 2021). The package is designed to be able to easily interpolate the same velocity field and produce strain rate outputs on the same geographic grid for any of the implemented methods. Outputs of the code include a NETCDF file (Rew & Davis 1990) containing the interpolated velocity field, strain rate components and invariants and corresponding plots. We can also directly calculate the mean, residual, and standard deviation of strain rate fields derived using different methods. Below we further describe each class of methods in general and the specific methods we implement in Strain_2D.

2.2.2 Direct strain rate calculation through baselines

This class of methods uses baselines (changes in length between two stations) to directly calculate strain. The Delaunay triangulation method directly calculates strain rates using baselines between three adjacent GNSS stations. Although simple to compute, this method is highly sensitive to noise in the observations, because of the lack of data redundancy (Kreemer *et al.* 2018). Strain rate is constant within each triangle, so strain rate maps have a discretized appearance that reflects the station geometry. Wu *et al.* (2011) showed that without noise this method can have similar results to others, but with random noise can result in complete loss of the underlying signal. In the Strain_2D library, we implement both flat- and spherical-earth versions of Delaunay.

A variation of the Delaunay method was developed by Kreemer *et al.* (2018), using a kind of hierarchy of triangles around an estimation point, and keeping only those triangles that have good geometric properties (e.g. not overly skinny or skewed) and located such that the estimation point lies within the inscribed circle of the triangle. Strain rate is calculated for each triangle and the median taken of the resulting suite of strain rates to determine the final strain rate at the estimation point. This method benefits from robust strain rate estimation and automatically reflects the network resolution above a threshold based on the degree of noise allowed in each triangle.

2.2.3 Spatial basis functions

This class of methods is typically formulated as solving a least-squares problem for determining optimal coefficients to basis functions to fit the velocity field (e.g. Haines & Holt 1993; Kreemer *et al.* 2000; Beaven & Haines 2001; Tape *et al.* 2009; 2014). Interpolation is applied to a regular grid, and numerical or analytic derivatives are then used to calculate strain rates. Constraints other than GNSS data such as Interferometric Synthetic Aperture Radar (InSAR; Wang & Wright 2012) or seismic moment tensors (Kreemer *et al.* 2000) can be easily incorporated due to the least-squares formulation.

Downloaded from https://academic.oup.com/gji/article/234/3/2128/7153326 by Missouri University of Science & Technology user on 16 August 2023

The equations to solve can be written in the following example form:

$$\begin{bmatrix} d_{\mathrm{GPS}} \\ d_{\mathrm{InSAR}} \\ 0 \end{bmatrix} = \begin{bmatrix} G_{\mathrm{GPS}} & 0 \\ G_{\mathrm{InSAR}} & G_p \\ G_R & 0 \end{bmatrix} \begin{bmatrix} v \\ \alpha \end{bmatrix} \tag{7}$$

where $\mathbf{G}_{\mathrm{GPS}}$ and $\mathbf{G}_{\mathrm{InSAR}}$ represent the operators relating the data to surface velocities, $\mathbf{G}_{\mathrm{p}}$ represents a reference frame correction for InSAR, $\mathbf{G}_{\mathrm{R}}$ is a regularization operator, $\mathbf{d}_{\mathrm{GPS}}$ and $\mathbf{d}_{\mathrm{InSAR}}$ are the observations, and $\boldsymbol{v}$ and $\boldsymbol{\alpha}$ are the velocities and/or coefficients to be estimated. For example, in Wang & Wright (2012) both GNSS and InSAR velocities are used, $\mathbf{G_p}$ with parameters $\boldsymbol{\alpha}$ rotates the InSAR observations into the GNSS reference frame, and $\mathbf{G_R}$ imposes Laplacian smoothing on the estimated velocities $\boldsymbol{v}$. Using the variance–covariance matrix of the errors allows residual tropospheric noise in InSAR to be accounted for (Wang & Wright 2012; Weiss *et al.* 2020).

Another example is the approach originally proposed by Haines & Holt (1993) and more recently developed by Okazaki *et al.* (2021), which solves for coefficients to bicubic splines instead of the velocities directly. In this case, $\boldsymbol{v}$ are the spline coefficients, and the strain rates can be computed analytically by multiplying the coefficients by the spatial derivatives of the basis functions. $\mathbf{G_p}$ and the $\boldsymbol{\alpha}$ coefficients are not needed when only GNSS observations are used. The system is solved using linear least squares. Regularization through $\mathbf{G_R}$ (e.g. Laplacian smoothing) is typically used to suppress high-frequency noise and provide a smooth field by some metric. There is a trade-off between fitting the observations and having a smoother field (e.g. Shen *et al.* 2015), so a decision must be made about how strongly to regularize. For tectonic studies of strain rate, the specific strength of regularization is often not quantitatively justified. We will explore the so-called 'L-curve' below as one option for how to choose the strength of regularization.

From this class of methods, the Strain_2D package implements the Tape *et al.* (2009) *wavelets* method through a Python wrapper of the code *surfacevel2strain*, written in MATLAB and publicly available (Tape *et al.* 2009). This method uses a suite of wavelets of varying degree as basis functions; the parameters to specify include the maximum degree of wavelet to use.

2.2.4 *Elasticity-based methods*

This class of methods uses constraints from elasticity, using a 2-D thin-plate approximation to interpolate 2-D vector data (Haines *et al.* 2015; Sandwell & Wessel 2016; Haines & Wallace 2020). The elasticity constraint provides a physical basis for the spatial correlation length scale and couples the two horizontal velocity components (depending on the Poisson's ratio). Horizontal body forces applied at specified locations deform the surface and are optimized to match the surface velocities. Haines *et al.* (2015) developed an approach called 'Vertical Derivatives of Horizontal Stress' (VDoHS) that uses a finite element model to solve for the body forces using the governing equations, while Sandwell & Wessel (2016) calculated Green's functions for a point source. Both methods co-locate the body forces with the stations, which results in an even-determined problem that can be solved exactly for the body force magnitudes. Strain_2D implements the *gpsgridder* method of Sandwell & Wessel (2016) by wrapping the built-in functionality in GMT (Wessel *et al.* 2019).

2.2.5 *Weighted neighbourhood methods*

This class of methods defines a local neighborhood around an estimation location and uses either a weighted mean or locally weighted linear least squares to estimate the velocity at the query point. The weighted average approach can be written as:

$$v^*(x_i) = \sum_j w_j v(x_j) \tag{8}$$

where $v(x_j)$ are the observed velocities, w_j are weights assigned to each velocity, and $v*(x_i)$ is the estimated velocity at the unobserved location x_i. The goal is to calculate the weights w_j such that the estimated velocities reflect the spatial correlations in the observations. This general class of methods has many common forms; for example, nearest-neighbour sets all weights equal to zero except for the closest point, which takes a weight of one. Another common approach is to use weights proportional to inverse distance to some power (Inverse Distance Weighting): $w_j \propto (1/r_{ij})^{\mathrm{p}}$, where r_{ij} are the distances between each observation at x_j and the unobserved location x_i. The exponent p can be adjusted based on desired properties of the interpolating surface (e.g. Shepard 1968).

An improvement on Inverse Distance Weighting can be had by solving for the optimal weights in eq. (8) given some constraints. Using spatial covariance, this becomes the classic geostatistical problem of kriging. First, variogram analysis or some other method is used to determine the spatial structure function for calculating covariance between pairs of observations (e.g. Matheron *et al.* 1962; and Fig. 1). This structure function is then used to determine the covariance matrix of the observations ($\sum$) and the covariance between the observations and the unsampled estimation locations (σ). Then one can solve directly for the optimal weights w using these covariances. We use the ordinary kriging equations (e.g. Wackernagel 2003; Chilès & Delfiner 2012) to find the optimal weights given an unknown mean:

$$\begin{bmatrix} \Sigma & 1 \\ 1^T & 0 \end{bmatrix} \begin{bmatrix} \lambda \\ \nu \end{bmatrix} = \begin{bmatrix} \sigma_0 \\ 1 \end{bmatrix} \tag{9}$$

where here we use the conventional symbol $\boldsymbol{\lambda}$ for the vector of kriging weights, and ν is a Lagrange multiplier. Eq. (9) is derived by assuming a solution in the form of eq. (8) and requiring unbiasedness and minimum estimation variance. Because $\boldsymbol{\Sigma}$ is a positive semi-definite square matrix, (9) can be directly inverted to solve for $\boldsymbol{\lambda}$ and ν. The weights are substituted into eq. (8) and used to calculate the prediction at the new location x_i. In its simplest form, this method assumes no long-wavelength trends (homogeneous) and no directional variations (isotropic). For GNSS velocities, two separate systems are solved for the two horizontal velocity components.

Shen *et al.* (2015) use a slightly modified version of these weighted-average type methods, but instead of calculating the interpolated velocity directly from weights they solve a locally weighted least-squares problem assuming a locally bilinear field, using the covariance and spatial data coverage as weights. Handwerger *et al.* (2019) and Huang *et al.* (2022) take a similar approach and perform a local linear regression in a moving window with fixed radius, but without any additional weights on the data.

In Strain_2D, we implement a standard ordinary kriging approach, separately interpolating each velocity component, as the *geostats* method. Shen *et al.* (2015) made their code *visr* publicly available as a Fortran executable, so we also provide a Python wrapper to *visr* in Strain_2D. Finally, we implement the Huang *et al.* (2022) moving window algorithm in Python, called 'Local average gradient.'

Downloaded from https://academic.oup.com/gji/article/234/3/2128/7153326 by Missouri University of Science & Technology user on 16 August 2023

2.3 Strain_2D implementation and moment rate calculation

The open-source Strain_2D program (Materna *et al.* 2021; see Data and Resources) provides the 'strain_rate_compute.py' and 'strain_rate_comparison.py' command-line programs for use in calculating strain rates. Explanations of the parameters for each method, any external codes required, and how to get started are available on the code repository page. Strain_2D has been set up using a modular approach to facilitate adding new models. The 'Strain_2D' base class includes a description of how to set up a new strain rate model to use with the code, so users can add their own methods and submit them back to the main repository.

3 APPLICATION TO SOUTHERN CALIFORNIA

3.1 GNSS velocity data set

For this study, we merge two publicly available GNSS velocity data sets together for the purposes of generating a dense velocity field in southern California (Fig. 2). The first velocity field is derived from the processing of continuous and semi-continuous GNSS data at University of Nevada Reno (UNR, Blewitt *et al.* 2018) and using the Median Interannual Difference Adjusted for Skewness (MIDAS) algorithm for estimating velocities (Blewitt *et al.* 2016). We take stations in the geographic domain [−121°, −114°, 32°, 36.5°] that have more than 2 yr of time-series duration, giving us 778 station velocities in the southern California domain.

The second velocity field is an update of the Southern California Earthquake Center (SCEC) Community Geodetic Model (CGM) version 1.0 (Sandwell *et al.* 2016), containing 1585 GNSS velocities from both continuous and campaign GNSS stations. We retain 1496 stations that have formal horizontal uncertainties less than 10 mm yr^{-1} and remove the rest.

We combine the two velocity fields and their uncertainties by determining a seven-parameter Helmert transformation based upon the 549 common stations, and then applying the transformation to convert the MIDAS velocities into the reference frame of the SCEC velocity field (Fig. S1 and Table S1, Supporting Information). The resulting velocity field has 1688 total velocities across southern California in the reference frame of the SCEC velocity field (Fig. 2a). Fig. 2(b) shows station spacing in the region as colour by taking the length scale of a Delaunay-type grid of the station locations.

3.2 Strain rate calculations

We used the merged GNSS data set described above and applied five of the methods (*visr*, *gpsgridder*, *local average gradient*, *wavelets* and *geostats*) that are implemented in the Strain_2D package to calculate velocities and strain rates on a regular grid in southern California and generate max shear, dilatation, and second invariant (I2) maps. GNSS velocities contain some level of noise, which is not accounted for by the direct baseline approach. Therefore, we do not use direct triangulation in the strain rate comparisons below but we provide the figures in the Supporting Information (Fig. S2).

3.2.1 Parameter selection by L-curve method

Parameter values for each method must be determined, as each method can be adjusted by its parameters to produce smoother and rougher models. One traditional approach for parameter tuning in tectonic studies is to use an '*L*-curve,' that is, a curve that shows the trade-off between model fit to the observations and some norm of the model, for example, a smoothness norm (Hansen 1992). The problem is underdetermined; we do not have enough observations to uniquely determine the strain rate at all points in space. In general, we expect strain rates due to tectonic processes (fault locking or fault slip) to be relatively smooth and act over relatively long spatial wavelengths. Thus, rougher solutions tend to fit better (have lower misfit) but likely correspond to noisier strain rate maps, while overly smooth solutions will likely not fit the data well.

We explored the parameter space for each strain rate method and quantified the results in a series of *L*-curves. We used the same input velocity field and grid spacing (0.02°×0.02°) in each model run. Fig. 3 shows *L*-curve plots of misfit (eq. 10) versus total moment (eq. 6) for the strain rate methods considered. Note that the wavelets method does not have a well-defined *L*-curve due to its parametrization so we plot those results as points rather than a line.

$$\sigma = \text{median}\left\{\left(\frac{v_{\text{obs}} - v_{\text{pred}}}{\sigma_{\text{obs}}}\right)^2\right\} \tag{10}$$

For Fig. 3, total moment rate on the *x*-axis is calculated from eq. (6) assuming a shear modulus of 30 GPa, seismogenic depth of 11 km, and only including strain on land. On the *y*-axis, we calculate misfit by using nearest-neighbor interpolation from the gridded velocity field to the data locations. The data contain some obvious outliers, so we use the median reduced chi-square of the misfits in the east and north components. We also excluded the largest outliers (with >10 mm yr^{-1} residuals), a process that removed less than 3 per cent of the data.

3.2.2 Modelling results and comparison

Based on the *L*-curve analysis, we can select a subset of models (highlighted by black squares in Fig. 3) as representative models from each method (Fig. 4) and compare their similarities and differences. Table 1 shows the final parameter values for each method used in Fig. 4. Table 2 shows the total moment calculated using Kostrov summation (eq. 6, Kostrov 1974; Savage & Simpson 1997) and misfit to the observations for each of the highlighted models.

Maximum shear strain rates are high all along the San Andreas fault (SAF) and San Jacinto fault systems. Low levels of shear strain rate are present in the entire Eastern California Shear Zone, the Garlock Fault and faults in the Transverse Ranges. A hotspot of shear strain rate in the vicinity of Landers–Hector mine shows up in all models, so may be related to residual post-seismic strain in the GNSS velocity field (e.g. Pollitz *et al.* 2000). Dilatation rates are consistent with contraction across Transverse Ranges. Extension occurs almost exclusively east of the SAF except for north-central Baja California and near the central SAF. The largest dilatation feature in all the maps is the contraction–dilatation pattern in northern Baja California that is spatially associated with the Cerro Prieto Geothermal Field (CPGF); comparison to vertical GNSS velocities shows widespread subsidence surrounded by uplift. This large-scale signal could be due to geothermal fluid extraction (e.g. Sarychikhina *et al.* 2011) or potentially to cooling of a shallow magma body (e.g. Hamling *et al.* 2022).

The mean and standard deviation of the strain rate invariants across all five methods are shown in Fig. 5, with strain rates masked where they do not exceed the standard deviation across all methods.

Downloaded from https://academic.oup.com/gji/article/234/3/2128/7153326 by Missouri University of Science & Technology user on 16 August 2023

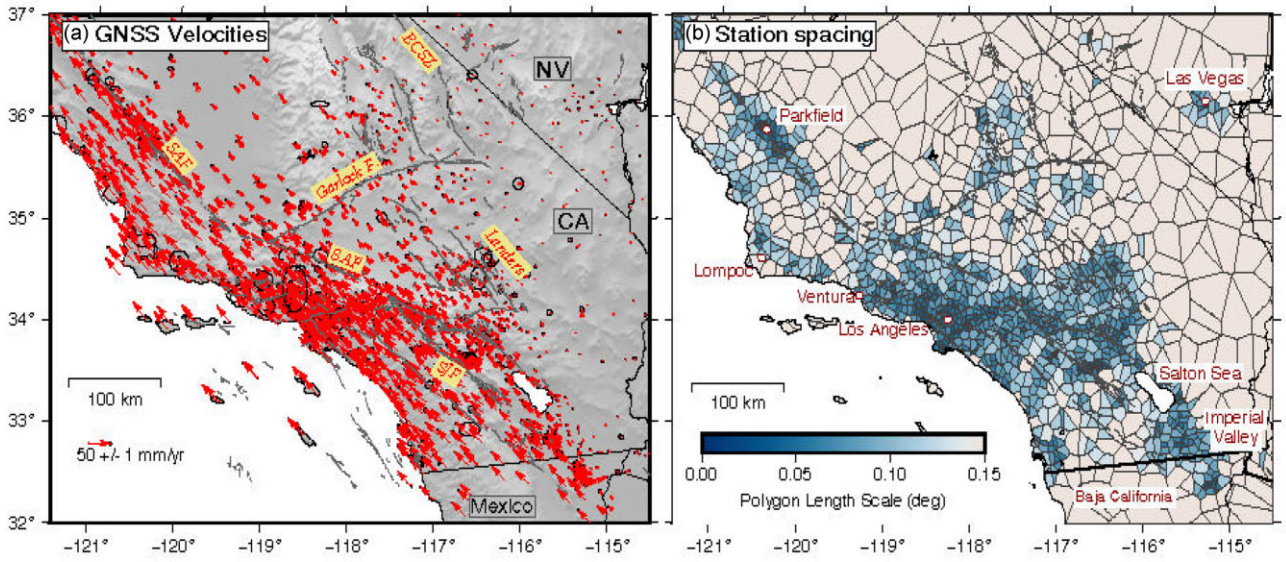

Figure 2. (a) GNSS horizontal velocities used in this study with respect to North America, showing 1688 velocities and their 98 per cent confidence intervals. Holocene fault traces from the USGS Quaternary Faults and Folds Database are shown in dark grey. SAF: San Andreas fault; ECSZ: Eastern California Shear Zone and SJC: San Jacinto fault. (b) Station spacing shown by using Voronoi polygons with stations at cell centres and colouring the polygons by the square root of their area, that is, the interstation length scale.

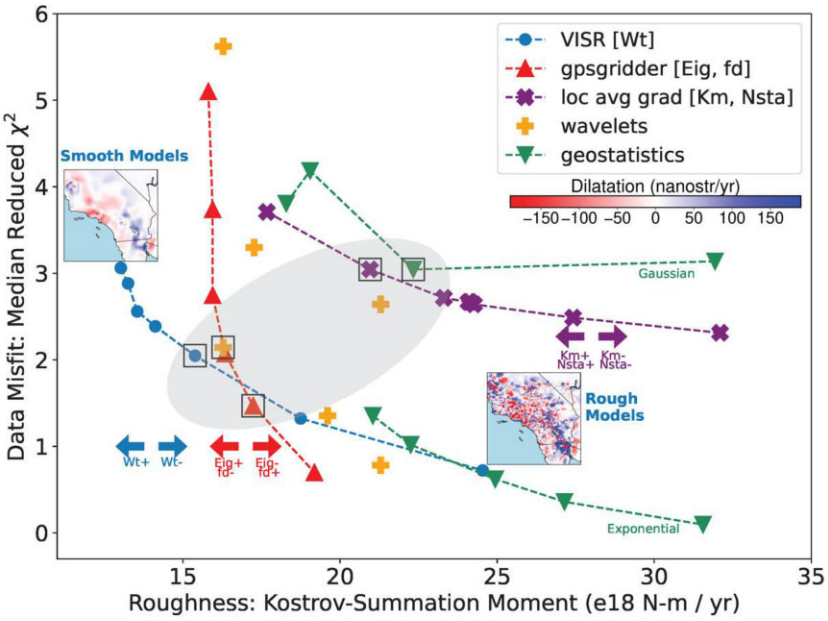

Figure 3. Data misfit metric versus total moment for various model parameter choices in each of the strain rate methods considered in this study (Table 1). Models near the corners of each method's *L*-curve, outlined by black boxes, were selected for subsequent analysis. The grey oval represents the region in roughness/misfit space containing all the representative models.

The residuals from the mean are shown in Fig. 6. Maximum shear strain rates are generally consistent across all models, with differences mainly in the degree of smoothness of the estimated field. The two principal exceptions are the northwest corner of the map near Parkfield and the area south of the Imperial Valley in southern California, where the SAF transitions to the Baja California spreading ridge. These two areas show large variability between methods even though they have high observation density. In comparison to the max shear strain rate, dilatation and I2 rates are much more variable between methods. The standard deviation of the dilatation rate is as large or larger than the signal in many places. Figs S3 and S4 (Supporting Information) show scatter cross-plots and histograms for maximum shear and dilatation across all methods.

Fig. 7 shows how strain rate varies along three profiles across the SAF. Max shear strain rates are highest in the south in the Imperial Valley profile, with lower and broader shear strain rates occurring in the two northern profiles. The profile plots are instructive because (1) they show that the variability between methods is less than the strain signal in many places, (2) they demonstrate the range of behaviours of the different methods and (3) they clearly show that strain rates from some methods are not smooth, at least for the choice of parameters reflected in Figs 3

Downloaded from https://academic.oup.com/gji/article/234/3/2128/7153326 by Missouri University of Science & Technology user on 16 August 2023

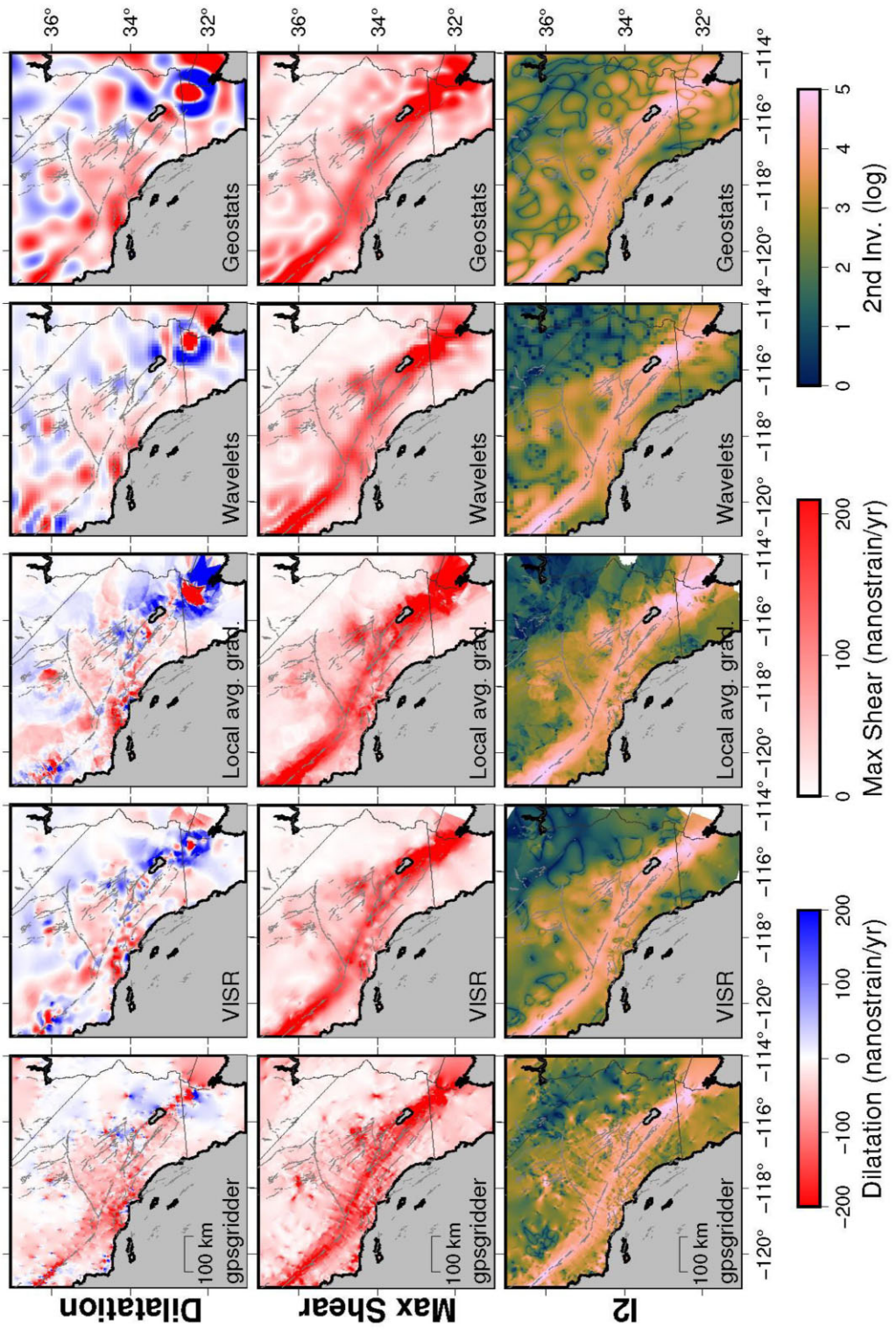

Figure 4: Dilatation, maximum shear strain rate and second invariant from five strain rate techniques using the same interpolation grid ($0.02^\circ \times 0.02^\circ$) and same input GNSS velocities. The maps are not ordered by any metric. Negative dilatation (red colours) corresponds to shortening.

and 4. *geostats*-Gaussian and *visr* are relatively smooth, *geostats*-Exponential and *local average gradient* are relatively rough, and *gpsgridder* and *wavelets* are somewhat intermediate. Strain rates are highest south of the Salton Sea, highlighting the critical seismic hazard of this region. Note that we included the *geostats*-Exponential model to show how the *geostats* method can produce a range of behaviours that encompass the other methods depending on parameter choices.

Downloaded from https://academic.oup.com/gji/article/234/3/2128/7153326 by Missouri University of Science & Technology user on 16 August 2023

Table 1. Parameters selected for each of the strain rate methods shown in the black boxes in Fig. 3. The 'Value' column shows the parameter choices for the five strain rate methods used to calculate the strain rate maps shown in Fig. 4.

Method	Class	Parameter	Value
Āll methods		range_strain	−121/−114/31/37
		range_data	−122/−113/30.5/37.5
		xinc/yinc	0.02°/0.02°
gpsgridder (Sandwell& Wessel 2016)	Elasticity	Poisson	0.5
		Fd	0.01
		eigenvalue	0.0001
VISR (Shen *et al.* 2015)	Neighbourhood	distance_weighting	Gaussian
		spatial_weighting	Voronoi
		min_max_inc_smooth	5/80/5
		weighting_threshold	10
		uncertainty_threshold	0.5
Local avg. grad. (Huang *et al.* 2022)	Neighbourhood	EstimateRadiusKm	120
		nstations	22
wavelets (Tape *et al.* 2009)	Basis functions	qmin	3
		qmax	8
		qsec	8
geostats (This publication)	Neighbourhood	Model_type	Gaussian
		Sill_east	1.1
		Range_east	0.55°
		Nugget_east	0.18
		Sill_north	1.1
		Range_north	0.55°
		Nugget_north	0.18
		trend	0

Table 2. Total moment rate computed from each strain rate model in black boxes in Fig. 3. Median deviation of the modelled velocities with respect to the observed velocities is computed using eq. (10) from a trimmed velocity data set that does not consider misfits > 10 mm yr^{-1}, to remove outliers from stations on islands, etc.

Method	Type of method	Moment accumulation rate, in 1e18 N·m yr^{-1} and equivalent Mw yr^{-1}	Median absolute deviation (mm yr^{-1})
gpsgridder	Elasticity	17.24 ($M_w = 6.79$)	1.5
VISR	Weighted neighbourhood	15.39 ($M_w = 6.76$)	2.1
Local avg grad	Weighted neighbourhood	20.96 ($M_w = 6.85$)	3.1
wavelets	Spatial basis function	16.28 ($M_w = 6.77$)	2.1
geostats—Gaussian covariance	Weighted neighbourhood	22.33 ($M_w = 6.87$)	3.0

4 DISCUSSION

4.1 Characterizing uncertainty in geodetic strain rate

Part of the epistemic uncertainty in strain rate can be represented by the centre column of panels in Fig. 5, which is calculated by taking the standard deviation of the models shown in Fig. 4. An additional source of epistemic uncertainty is the range of reasonable values for the parameters chosen for a given method, which can be constrained based on the *L*-curve. In our case of multiple types of strain rate methods, we have multiple *L*-curves, and by comparing all the *L*-curves together we chose the models highlighted by boxes in Fig. 3.

We can also approximate aleatoric strain rate uncertainties by propagating the uncertainties in the velocity field obtained from the *geostats* method through the strain rate calculation, recognizing that the aleatoric uncertainty estimate depends upon selecting reasonable model parameters (see below). We then use a linear Taylor expansion to obtain uncertainties in the shear strain rate from uncertainties in velocity (see Supporting Information). Fig. 8 shows this aleatoric uncertainty as well as the estimate of epistemic uncertainty described above. As expected, aleatoric uncertainty is high where there is no data and low in areas of dense station spacing. In contrast, epistemic uncertainty is high where strain rates are high and station spacing is either higher or lower than average across the scene. Epistemic uncertainty is slightly higher for dilatation rate, while aleatoric uncertainty is slightly higher for max shear strain rate. Note that the colour pattern artefacts in the lower-right panel in Fig. 8 are due to the linearization of the max shear strain rate to estimate uncertainty.

There is a trade-off between epistemic and aleatoric uncertainty that is related to the parameter choice made for each model using the *L*-curve (Fig. 3). With reference to Figs 1 and 3, models that assume a longer correlation length scale and/or a lower total variability would tend to have lower aleatoric uncertainty because the observations impose stronger constraints on the unknown estimates (i.e. results in a smoother model, which by definition has a lower aleatoric uncertainty). A shorter length scale implies that the data provide weaker constraints on the models. In other words, 'dense' station spacing versus 'sparse' station spacing is a function of the distance over which observations are correlated; a 5-km station spacing is sparse if observations are only correlated over 1 km. For tectonic problems and assuming elasticity, the relevant length scale is the fault locking depth, which varies in southern California but is in the range 5–25 km (e.g. Smith-Konter *et al.* 2011; Johnson 2013; Maurer *et al.* 2018). Thus, in southern California, station spacing <10 km is relatively 'dense' and aleatoric uncertainty is low, while

Downloaded from https://academic.oup.com/gji/article/234/3/2128/7153326 by Missouri University of Science & Technology user on 16 August 2023

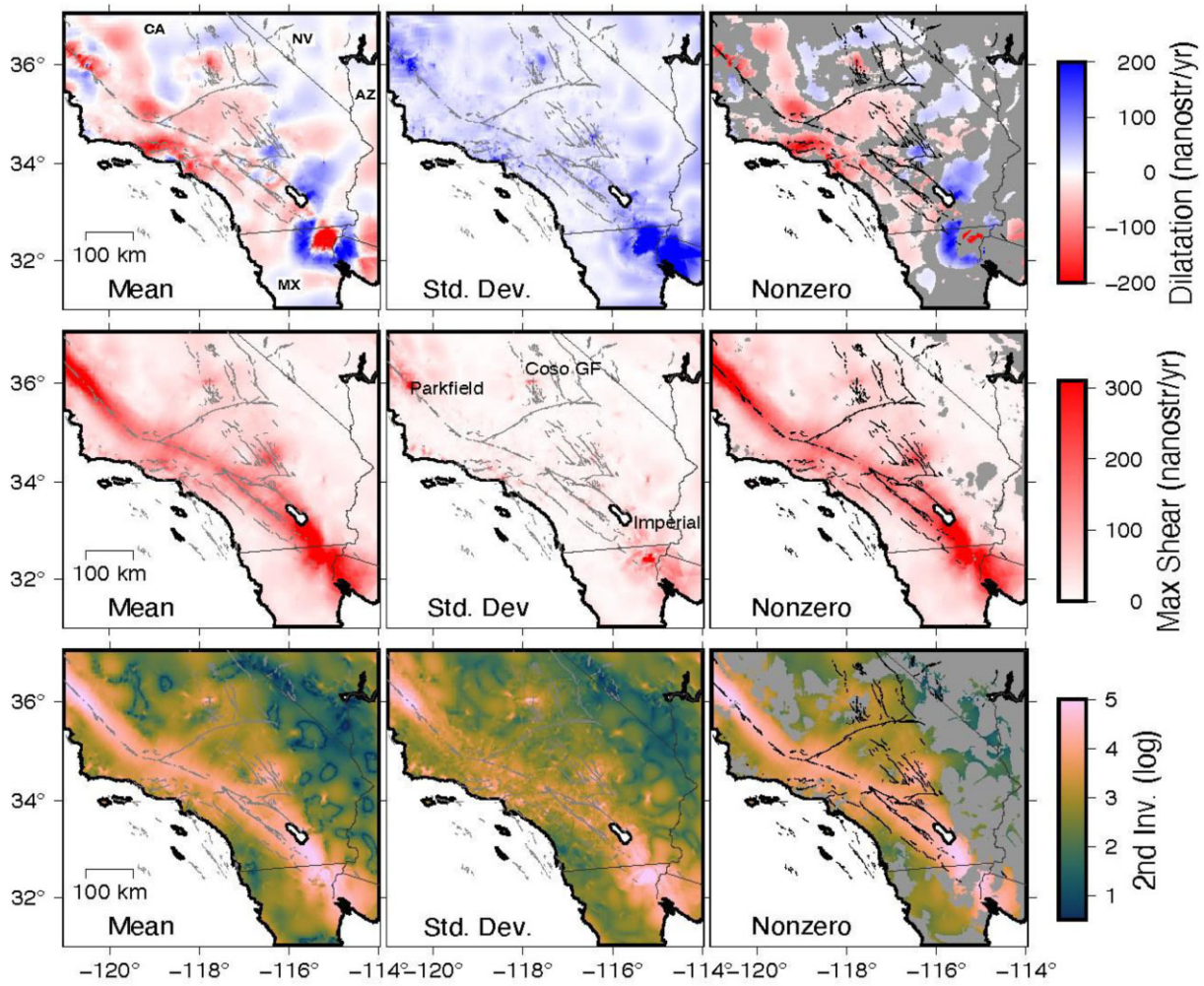

Figure 5: Mean and standard deviation of derived quantities from the five strain rate models shown in Fig. 4. Features such as the Transverse Ranges, Parkfield, CA, the Coso GF and the CPGF are labelled. The right-hand column (Nonzero) shows the mean strain rates (left-hand column: mean) that have been masked where the standard deviations from the central column are larger than the mean values themselves, making them statistically indistinguishable from zero.

>25 km spacing is 'sparse' and aleatoric uncertainty is high (cf. Figs 2b, 6 and 8).

Fig. 3 and Table 2 show that, in comparison to the *L*-curve for each individual method, when comparing between different methods moment and misfit do not systematically trade-off. For example, the *visr* method is the smoothest (in the sense of having lowest total Kostrov moment) but has a higher misfit than the *gpsgridder* method. Both the local average gradient and the *Geostats*-Gaussian methods have high moment and high misfit. Moment varies ∼40 per cent between all methods (15.4–22.3 × 10^{18} N·m yr^{-1}) and the misfit varies by ∼50 per cent. With reference to Fig. 3, these characteristics are due to systematic shifts in the location of the *L*-curve in misfit-moment space, which are related to the strain rate and/or velocity field parametrizations used by each method. This is illustrated by comparing the *geostats*-Gaussian model with the *geostats*-exponential model (Fig. 3); the two methods are the same but in one case parametrizes the velocity field using an exponential spatial structure function (spatial covariance) versus the Gaussian structure function. Fig. 7 shows how the different assumptions of each method get reflected in profiles of strain rate and suggests that epistemic uncertainty for strain rate will depend in part on the degree of roughness considered to be plausible by the practitioner.

4.2 Contributions to uncertainty in geodetic strain rate related to data and interpolation errors

What is the source of uncertainty shown in Fig. 8 and discussed above? Comparing Fig. 2(b) with Figs 6 and 8 highlights that much of the variability between methods occurs at length scales roughly at or below the station spacing. As noted in the Introduction, every method for interpolation must implicitly or explicitly make three assumptions about the underlying field: how to handle measurement noise in the observations, the interpolated field's behavior at short wavelengths and the length scale over which the observations are correlated. Every strain rate method makes different choices for these three characteristics. Not surprisingly, in our experiments, some of the large epistemic strain rate discrepancies occur in areas with sparse station spacing and highly variable or noisy data (e.g. Cerro Prieto in northern Mexico). However, several of the large discrepancies in modelled strain rate are co-located with regions with dense station spacing, such as Parkfield and the SAF south of the Salton Sea (cf. Fig. 2b with Figs 4–6), some of which represent velocities from GNSS campaigns. Higher station density occurs along the SAF north of the Salton Sea, the Imperial Valley in CA, the San Jacinto fault, the Parkfield section of the SAF, the region between Los Angeles and Ventura, the Coso GF and the Highway

Downloaded from https://academic.oup.com/gji/article/234/3/2128/7153326 by Missouri University of Science & Technology user on 16 August 2023

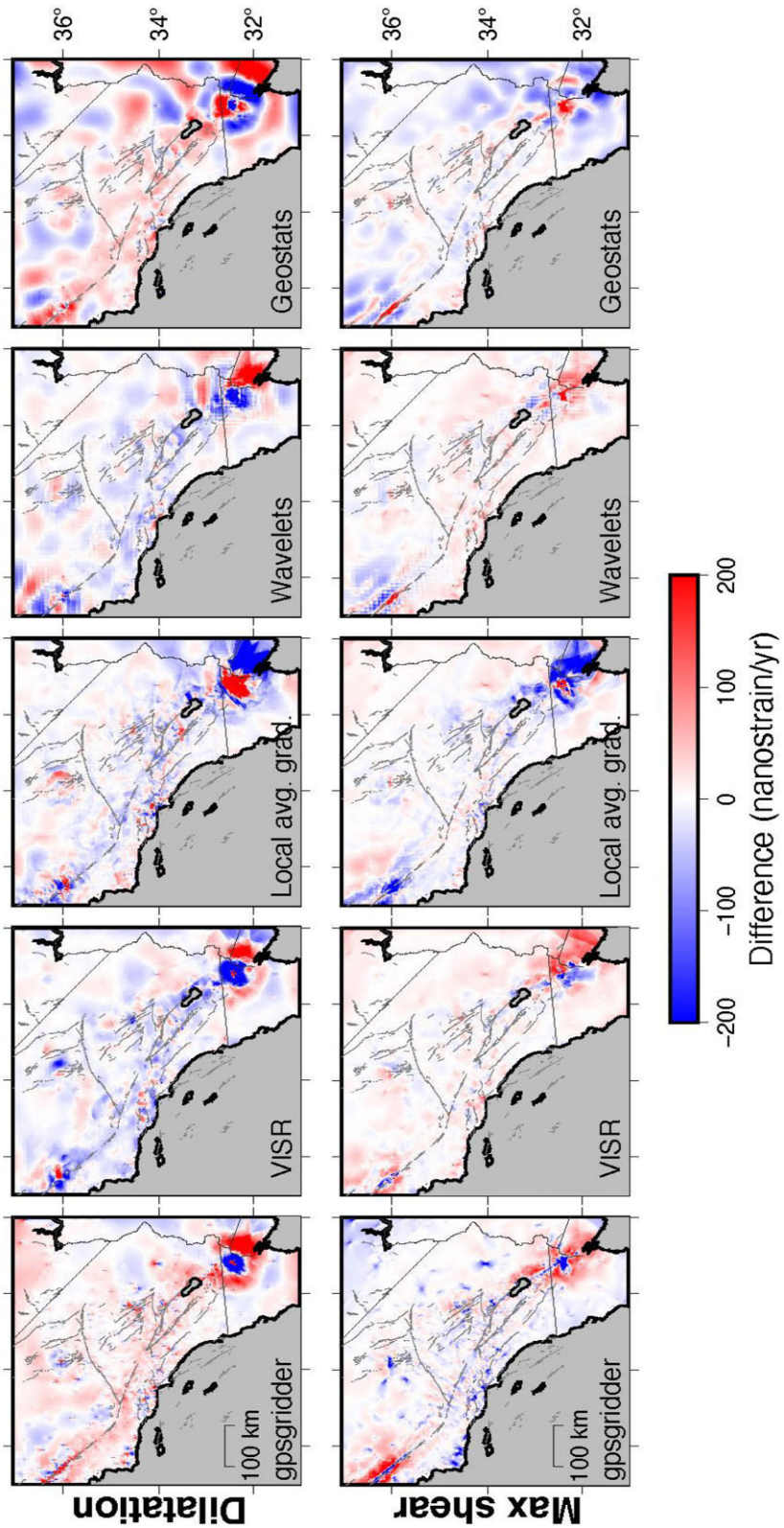

Figure 6: Residual strain rates (i.e. mean in Fig. 5 minus individual solution in Fig. 4) for dilatation (top row) and max shear (bottom row).

1 corridor near Lompoc, CA (Fig. 2b). Each of these regions is characterized by higher discrepancy between the strain rate models in maximum shear strain rate and/or dilatation rate. Many, but not all, of these regions have high strain rates.

The locations of these discrepancies and their correspondence to non-random station placement suggests that locally densifying stations to improve resolution of strain rate in one small part of a larger area will not improve the calculation of strain rate without

Downloaded from https://academic.oup.com/gji/article/234/3/2128/7153326 by Missouri University of Science & Technology user on 16 August 2023

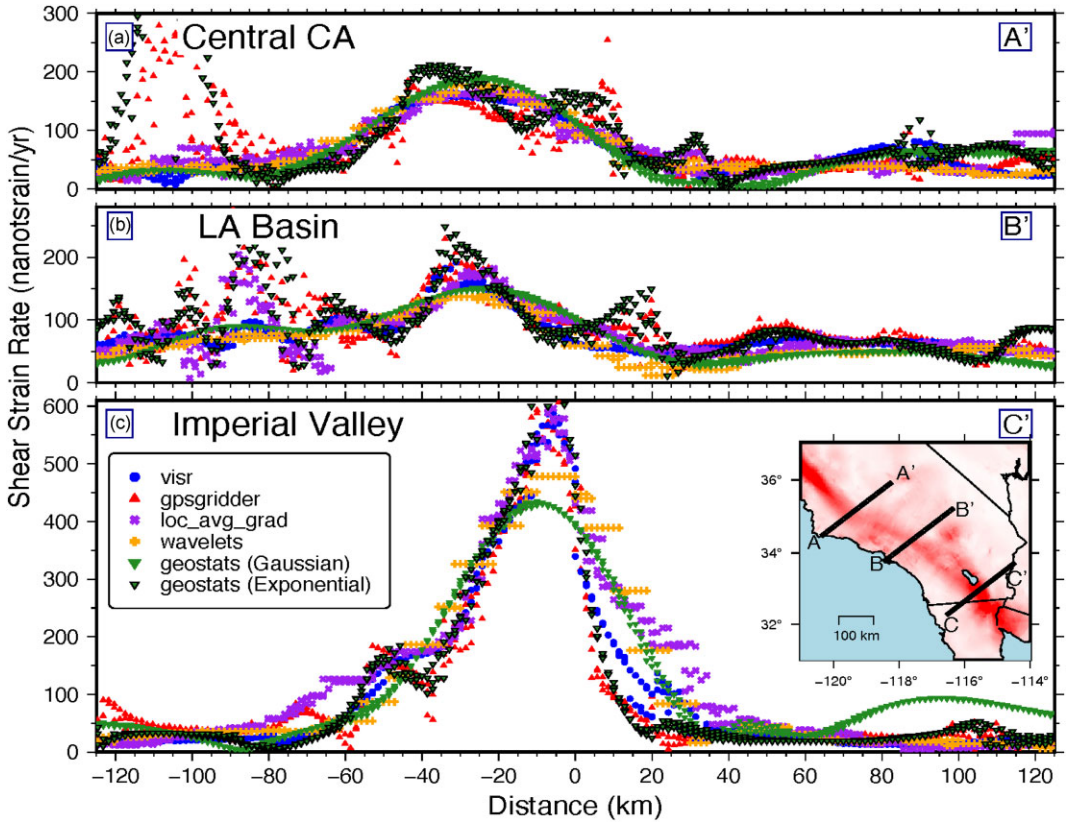

Figure 7. Three max shear strain rate profiles across the SAF across the five representative models from Fig. 4 and also including the *geostats*-Exponential model. Top: cross-section along the central SAF; middle: cross-section at about 34° N across the Los Angeles (LA) basin and bottom: cross-section across the Imperial Valley region. Inset map shows the maximum shear strain rate from the average of the five representative models (same as in Fig. 5).

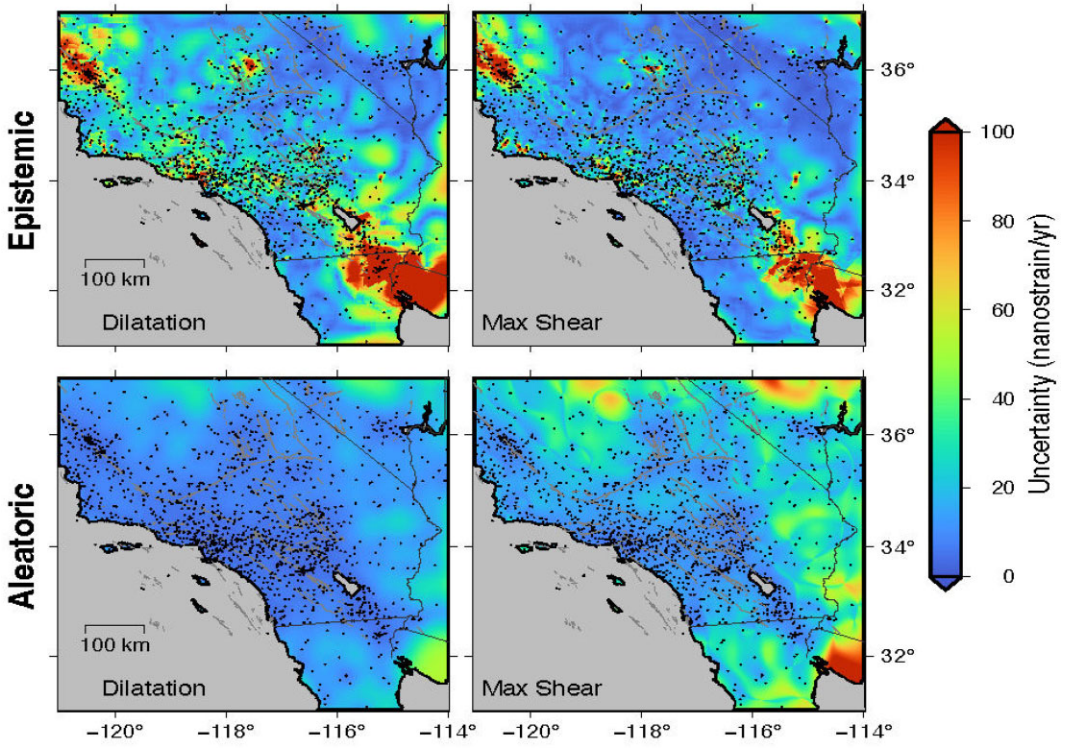

Figure 8. Top row: epistemic uncertainty (1σ level) derived from five methods of strain calculation from Fig. 4. Bottom row: uncertainty in dilatation rate and maximum shear strain rate obtained by propagating velocity interpolation uncertainties derived from the geostatistical method presented in this paper. Black dots show GNSS station locations.

Downloaded from https://academic.oup.com/gji/article/234/3/2128/7153326 by Missouri University of Science & Technology user on 16 August 2023

consideration of the different spatial wavelengths able to be targeted with the method used. In other words, the addition of data without the ability to adjust the spatial parameters corresponding to the local station density is a limitation in many of the strain methods presented here.

Many of the methods implemented in Strain_2D use a single set of parameters to describe the spatial variability in the observations; regions with denser station spacing need different parameter choices based on the local variability than is appropriate for the whole model domain. One way to do this would be to use targeted strain rate calculations on smaller domains, or to use a multiscale method like *wavelets* (Tape *et al.* 2009) or *visr* (Shen *et al.* 2015). However, the multiscale parametrization in these existing methods is not fully automatic; it requires some hand-tuning. The median filter-based method of Kreemer *et al.* (2018) does automatically adjust resolution for the region, but with a limited minimum resolution much larger than the relevant spatial length scale in southern California.

In summary, these findings suggest that (1) computational methods generally agree at long wavelengths, consistent with Xu *et al.* (2021), and also in areas with low overall strain rates, and (2) strong gradients in strain rate and/or noisy data are handled differently by the different methods, resulting in larger epistemic uncertainties in those areas. Importantly, more data do not automatically improve the resolution of strain rate, because the strain rate methods do not automatically account for a changing spatial correlation length scale.

4.3 Implications for using strain rates to constrain seismic hazard

Numerous studies have applied strain rates to estimate seismic hazard (e.g. Ward 1994; Shen *et al.* 2007; Bird *et al.* 2010; Bird & Kreemer 2014; D'Agostino 2014, 2015; Zeng *et al.* 2018; Zheng *et al.* 2018; Stevens & Avouac 2021), often using eq. (6) or equivalent to convert strain rates to moment rates (Kostrov 1974; Savage & Simpson 1997). Variability in total moment rate across the study area varies by roughly 40 per cent across the five models selected for inclusion in Fig. 5. Over 150 yr, roughly the time since the 1857 Fort Tejon earthquake, this difference accumulates to an M_w 8 earthquake. In comparison, Xu *et al.* (2021) reported strain rate uncertainties from different models approaching 100 per cent of the strain signal in high-strain areas such as along the SAF. This larger uncertainty could be due in part to different underlying GNSS velocity fields, the inclusion of strain rate models based on elastic fault models (Hearn *et al.* 2010), or to independent parameter tuning by each individual researcher who produced the models. Our results in Fig. 3 suggest that comparing multiple methods together provides additional insight into parameter selection that would be missed using only a single method.

Clearly, quantifying and, if possible, reducing strain rate uncertainty are critical for seismic and geodetic moment rate estimates (e.g. Mazzotti *et al.* 2011; D'Agostino 2014; Stevens & Avouac 2016) and could have a major impact on seismic hazard forecasts based on strain rates. A single quantitative metric for determining the 'best' strain rate model does not account for the fact that some real strain may be unrelated to tectonic deformation processes; a practitioner may therefore prefer a smoother model with less small-scale variability that fits the data worse than a better-fitting model, and the *L*-curve approach provides one method for selecting reasonable model parameters given such preferences. Strain_2D provides the information needed for practitioners to evaluate the inherent trade-off between fit to the observations and model norm for different strain rate methods and to decide the range of parameters that makes the most sense given their particular research questions.

Our results show a way to quantify strain rate uncertainty arising from data and interpolation errors; however, we have not addressed a potential third source of epistemic uncertainty in strain rates, that related to signals that are not of interest to the practitioner. Strain rates in southern California include contributions from tectonics, hydrology, and magmatic systems and small-scale signals such as landslides. Practitioners using strain rates to estimate seismic hazard could try to remove non-tectonic signals prior to hazard analysis; however, it is unclear how best to do so. For example, magmatic systems result in strain rate that are not related to tectonic plate motion, but could result in seismicity (e.g. Hamling *et al.* 2022). The same might be true of hydrologic signals in some cases (e.g. Johnson *et al.* 2017; Hsu *et al.* 2021). In this case, what signals can or should be removed will be a case-by-case basis and, for seismic hazard, will need to reflect the relevant timescales involved; that is, over what time period is seismic hazard of interest? Studies focused solely on the long-term loading of faults will of course want to remove as many other types of signals as possible, while studies of short-term seismic hazard may want to retain some non-tectonic strain sources. Currently there are no systematic methods for dealing with such a decision process; this is an area ripe for additional research.

4.4 Will InSAR solve all these problems?

One may reasonably wonder whether the rising quantity and quality of InSAR data for velocity determination will resolve all the problems related to interpolation of sparse geodetic measurements. In particular, the high density of surface observations provided by InSAR could potentially allow for calculation of strain rates at high resolution without interpolation. However, there are several challenges with using InSAR for strain analysis: (1) only line-of-sight velocities are obtained, (2) current satellite geometries limit sensing of the north–south component of deformation and (3) the presence of residual tropospheric noise results in unreal strain signals that are difficult to remove. There have been several successful attempts to calculate strain rates using InSAR by combining with GNSS measurements (e.g. Wang & Wright 2012; Tymofyeyeva & Fialko 2018; Weiss *et al.* 2020; Xu *et al.* 2021). These approaches all use relatively dense GNSS observations and filtering to correct the InSAR velocities for tropospheric noise, and required coherent, high accuracy InSAR velocities. Achieving these conditions may be difficult in many areas around the world, in particular where GNSS velocities are sparse, and also in tropical regions where InSAR decorrelation is a significant problem. The use of InSAR for strain rate estimation is expected to improve strain rate estimation when sufficient GNSS data is available, and without GNSS data it is expected to be limited. GNSS velocity interpolation for strain rate calculation will thus likely remain an important problem for at least the foreseeable future. The use of InSAR to improve strain rate calculations in locations with spatially sparse GNSS data is a critical future research direction that fully takes advantage of the strengths and weaknesses of each data set.

Downloaded from https://academic.oup.com/gji/article/234/3/2128/7153326 by Missouri University of Science & Technology user on 16 August 2023

4.5 Recommendations and future directions

Considering the above observations, we can make a few practical suggestions for practitioners using strain rates:

- Strain rate uncertainty using the methods presented here will likely only be as good as is possible with the average station spacing. Regions of higher station density will not be better resolved simply because there is more data there. For areas with more data, a separate analysis of that smaller region using appropriate parameters may be warranted.
- When determining the parameters for a given method, assessing the data fit and smoothness of the model can aid in parameter selection. An *L*-curve provides one means of quantifying the trade-off between misfit and model roughness.
- Comparing several different methods and/or parameter combinations to quantify the full space of models that represent a reasonable trade-off between misfit and roughness and then averaging the results together may produce a more robust strain estimate. This is most easily done with a consistent way to handle the data and gridding such as the open-source Strain_2D package. Differences between methods will alert the user to areas that need further attention, and the user can decide which suite of models best represents the needs of the study.
- Using the geostatistical method or a similar approach allows for estimates of aleatoric uncertainty.
- Investigations of how to best combine GNSS and InSAR data, in particular for scenarios without dense GNSS velocity fields, should be pursued.

Future development of strain rate methods should address automatic local determination and application of spatial structure parameters. Such algorithms exist but need to be further developed (Tape *et al.* 2009) or have not been applied to strain rate estimation (e.g. Gribov & Krivoruchko 2020). Such algorithms also should account for spatial anisotropy typical of GNSS velocities (such as is currently done in the *visr* method) and ideally would also easily incorporate constraints from elasticity, seismicity or other types of strain measurements (e.g. VLBI and InSAR, e.g. Haines & Holt 1993; Sandwell & Wessel 2016; Weiss *et al.* 2020). Such a method would allow for reliable multiscale estimation of strain rates globally while also retaining local details. In addition, better integration of InSAR velocities when available that accounts for the spatially correlated uncertainties of both data sets has great potential for reducing strain rate uncertainties, both aleatoric and the component of epistemic uncertainty related to interpolation method. The Strain_2D code provides a potential platform for these future developments through its modular design and use of a consistent input format. Integrating additional strain rate calculation methods from the scientific community could improve estimates of epistemic uncertainties in the Strain_2D workflow.

5 CONCLUSIONS

We have presented an application of the Strain_2D software package (Materna *et al.* 2021) for estimating strain rate and its epistemic and aleatoric uncertainties in southern California using a new combined GNSS velocity field. We find that maximum shear strain rates are highest along the SAF system, as expected, with the overall highest strain rates occurring around the Imperial Valley. Dilatation rates are generally consistent with strike-slip tectonics except in northern Baja California, where dilatation patterns suggest deformation potentially related to either geothermal development or a shallow magmatic system. We construct our mean strain rate model using five separate strain rate methods, and for each method, we systematically choose appropriate model parameters through the construction of an *L*-curve. By comparing the five representative strain rate models, we find epistemic variability to be roughly 40 per cent of the strain rate signal, a much lower value than roughly 100 per cent that was found in previous studies (e.g. Hearn *et al.* 2010). Comparison between the models also suggests that denser geodetic data do not automatically improve the resolution of strain rate because most strain rate methods do not automatically account for spatially varying spatial structure. The Strain_2D package provides a convenient means to calculate and average results from multiple strain rate methods and can be used to determine both aleatoric and epistemic uncertainties to improve understanding of strain rates in new areas.

DATA AND RESOURCES

The Strain_2D library and its tutorial and examples is publicly available (Materna *et al.* 2021). GNSS velocities from the UNR through the MIDAS algorithm were accessed at http://geodesy.unr.edu/velocities/midas.NA.txt on 2020 November 7. The combined post-Helmert-transformation GNSS velocity table and input parameter file used to reproduce all figures in this study are provided as part of the Electronic Supplement.

ACKNOWLEDGMENTS

We thank the contributors of the different strain methods, including Mong-Han Huang and Carl Tape, for providing their methods and/or adjusting them so they could be included in this comparison, and Zheng-Kang Shen, Paul Wessel and David Sandwell for making their codes publicly available. We are grateful to Mike Floyd for providing the updated SCEC velocity field. We are grateful to Bill Hammond, Tim Clements and one anonymous reviewer for constructive reviews that greatly improved the manuscript. Some figures were made with Generic Mapping Tools version 6 and PyGMT (Wessel *et al.* 2019; Uieda *et al.* 2021). Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. government.

CONFLICT OF INTEREST

We have no known conflicts of Interest.

SUPPORTING INFORMATION

Supplementary data are available at *GJI* online.

MaurerMaterna_SI_revised.pdf

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the paper.

REFERENCES

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.: Solid Earth,* **106**(B1), 741–770.

Downloaded from https://academic.oup.com/gji/article/234/3/2128/7153326 by Missouri University of Science & Technology user on 16 August 2023

Bird, P., Kreemer, C. & Holt, W., 2010. A long-term forecast of shallow seismicity based on the global strain rate map, *Seismol. Res. Lett.,* **81**(2), 184–194.

Bird, P. & Kreemer, C. 2014. Revised tectonic forecast of global shallow seismicity based on version 2.1 of the global strain rate map, *Bull. seism. Soc. Am.,* **105**(1), 152–166.

Bird, P., Jackson, D., Kagan, Y., Kreemer, C. & Stein, R., 2015. GEAR1: a global earthquake activity rate model constructed from geodetic strain rates and smoothed seismicity, *Bull. seism. Soc. Am.,* **105**(5), 2538–2554.

Blewitt, G., Kreemer, C., Hammond, W. & Gazeaux, J., 2016. MIDAS robust trend estimator for accurate GPS station velocities without step detection, *J. geophys. Res.-Solid Earth,* **121,** 2054–2068.

Blewitt, G., Hammond, W. & Kreemer, C., 2018. Harnessing the GPS data explosion for interdisciplinary science, *Eos,* **99**.

Cai, J. & Grafarend, E., 2007. Statistical analysis of geodetic deformation (strain rate) derived from the space geodetic measurements of BIFROST Project in Fennoscandia, *J. Geodynam.,* **43**(2), 214–238.

Chilès, J. & Delfiner, P., 2012. *Geostatistics: Modeling Spatial Uncertainty: Second Edition*. John Wiley & Sons.

D'Agostino, N., 2014. Complete seismic release of tectonic strain and earthquake recurrence in the Apennines (Italy), *Geophys. Res. Lett.,* **41**(4), 1155–1162.

El-Fiky, G. & Kato, T., 1998. Continuous distribution of the horizontal strain in the Tohoku district, Japan, predicted by least-squares collocation, *J. Geodynam.,* **27**(2), 213–236.

Feigl, K. *et al.*, 1993. Space geodetic measurement of crustal deformation in central and southern California, 1984–1992, *J. geophys. Res.: Solid Earth,* **98**(B12), 21677–21712.

Goudarzi, M., Cocard, M. & Santerre, R., 2015. GeoStrain: an open source software for calculating crustal strain rates, *Comput. Geosci.,* **82,** 1–12.

Gribov, A. & Krivoruchko, K., 2020. Empirical Bayesian kriging implementation and usage, in *Science of The Total Environment,* Vol. **722,** p. 137290.

Hackl, M., Malservisi, R. & Wdowinski, S., 2009. Strain rate patterns from dense GPS networks, *Natural Hazards Earth Syst. Sci.,* **9**(4), 1177–1187.

Haines, A. & Holt, W., 1993. A procedure for obtaining the complete horizontal motions within zones of distributed deformation from the inversion of strain rate data, *J. geophys. Res.: Solid Earth,* **98**(B7), 12057–12082.

Haines, A., Dimitrova, L., Wallace, L. & Williams, C., 2015. *Enhanced Surface Imaging of Crustal Deformation: Obtaining Tectonic Force Fields using GPS Data*. Springer.

Haines, A. & Wallace, L., 2020. New Zealand-wide geodetic strain rates using a physics-based approach, *Geophys. Res. Lett.,* **47**(1), e2019GL084606.

Handwerger, A., Huang, M.-H., Fielding, E., Booth, A. & Bürgmann, R., 2019. A shift from drought to extreme rainfall drives a stable landslide to catastrophic failure, *Sci. Rep.,* **9**(1), 1569.

Hansen, C., 1992. Analysis of discrete ill-posed problems by means of the L-curve, *SIAM Rev.,* **34**(4), 561–580.

Hamling, I., Kilgour, G., Hreinsdottir, S., Bertrand, E. & Bannister, S., 2022. Estimating the distribution of melt beneath the Okataina Caldera, New Zealand: an integrated approach using geodesy, seismology and magnetotellurics, *J. Volc. Geotherm. Res.,* **426,** 107549. https://doi.org/10.1029/2021GL096465

Hearn, E., Johnson, K. & Thatcher, W., 2010. Space geodetic data improve seismic hazard assessment in California: workshop on Incorporating Geodetic Surface Deformation Data Into UCERF3; Pomona, California, 1–2 April 2010, *Eos, Trans. Am. geophys. Un.,* **91**(38), 336.

Hsu, Y. *et al.*, 2021. Synchronized and asynchronous modulation of seismicity by hydrological loading: a case study in Taiwan, *Sci. Adv.,* **7**(16), eabf7282. https://doi.org/10.1126/sciadv.abf7282

Huang, M., Lopez, K. & Olsen, K., 2022. Icequake-magnitude scaling relationship along a rift within the Ross Ice Shelf, Antarctica, *Geophys. Res. Lett.,* e2022GL097961. https://doi.org/10.1029/2022GL097961

Johnson, K., 2013. Slip rates and off-fault deformation in Southern California inferred from GPS data and models, *J. geophys. Res.: Solid Earth,* **118**(10), 5643–5664.

Johnson, C., Fu, Y. & Bürgmann, R., 2017. Seasonal water storage, stress modulation, and California seismicity, *Science,* **356**(6343), 1161–1164.

Kato, T., El-Fiky, G., Oware, E. & Miyazaki, S., 1998. Crustal strains in the Japanese islands as deduced from dense GPS array, *Geophys. Res. Lett.,* **25**(18), 3445–3448.

Kostrov, V., 1974. Seismic moment and energy of earthquakes, and seismic flow of rock, *Izv. Acad. Sci. USSR Phys. Solid Earth,* **1,** 23–44.

Kreemer, C., Haines, J., Holt, W., Blewitt, G. & Lavallee, D., 2000. On the determination of a global strain rate model, *Earth Planets Space,* **52**(10), 765–770..

Kreemer, C., Blewitt, G. & Klein, E., 2014. A geodetic plate motion and global strain rate model, *Geochem. Geophys. Geosyst.,* **15**(10), 3849–3889.

Kreemer, C., Hammond, W. C. & Blewitt, G., 2018. A robust estimation of the 3-D intraplate deformation of the North American plate from GPS, *J. geophys. Res.: Solid Earth,* **123**(5), 4388–4412.

Materna, K., Maurer, J. & Sandoe, L., 2021. Strain-2D (version 1.1.1). [Computer software].

Matheron, G. & Blondel, F., Bureau de recherches ge´ologiques et minie`res (France), 1962. Traite´ de ge´ostatistique applique´e, *tome i. Technip*. **OCLC #**: 491866302.

Maurer, J., Johnson, K. & Segall, P., 2018. Bounding the moment deficit rate on crustal faults using geodetic data: application to Southern California, *J. geophys. Res.: Solid Earth,* **123**(12), 11,11-48,61.

Mazzotti, S., Leonard, L., Cassidy, J., Rogers, G. & Halchuk, S., 2011. Seismic hazard in western Canada from GPS strain rates versus earthquake catalog, *J. geophys. Res.: Solid Earth,* **116**(12), 1–17.

Noda, A. & Matsu'ura, M., 2010. Physics-based GPS data inversion to estimate three-dimensional elastic and inelastic strain fields, *Geophys. J. Int.,* **182**(2), 513–530.

Okazaki, T., Fukahata, Y. & Nishimura, T., 2021. Consistent estimation of strain-rate fields from GNSS velocity data using basis function expansion with ABIC, *Earth Planets Space,* **73**(1), 1–22.

Pagani, C., Bodin, T., Métois, M. & Lasserre, C., 2021. Bayesian estimation of surface strain rates from global navigation satellite system measurements: application to the Southwestern United States, *J. geophys. Res.: Solid Earth,* **126**(6), e2021JB021905.

Pollitz, F., Peltzer, G. & Bürgmann, R., 2000. Mobility of continental mantle: evidence from postseismic geodetic observations following the 1992 Landers earthquake, *J. geophys. Res.,* **105**(B4), 8035–8054.

Qu, W., Chen, H., Liang, S., Zhang, Q., Zhao, L., Gao, Y. & Zhu, W., 2019. Adaptive least-squares collocation algorithm considering distance scale factor for GPS crustal velocity field fitting and estimation, *Remote Sens.*, **11,** 2692. https://doi.org/10.3390/rs11222692

Rew, R. & Davis, G., 1990. NetCDF: an interface for scientific data access, *IEEE Comput. Graph. Appl.,* **10**(4), 76–82.

Sakamoto, Y., Ishiguro, M. & Kitagawa, G., 1986. *Akaike Information Criterion Statistics*. D. Reidel, Dordrecht, The Netherlands, 81.10.5555: 26853.

Sandwell, D. & Wessel, P., 2016. Interpolation of 2-D vector data using constraints from elasticity, *Geophys. Res. Lett.,* **43**(20), 10,703–710,709.

Sandwell, D.T., Zeng, Y., Shen, Z.-K., Crowell, B., Murray, J., McCaffrey, R. & Xu, X., 2016. The SCEC community geodetic model V1: horizontal velocity grid, In *SCEC Annual Meeting*.

Sarychikhina, O., Glowacka, E., Mellors, R. & Vidal, F., 2011. Land subsidence in the Cerro Prieto Geothermal Field, Baja California, Mexico, from 1994 to 2005: an integrated analysis of DInSAR, leveling and geological data, *J. Volc. Geotherm. Res.,* **204**(1), 76–90.

Savage, J. & Burford, R., 1970. Accumulation of tectonic strain in California, *Bull. seism. Soc. Am.,* **60**(6), 1877–1896.

Savage, J. & Simpson, R., 1997. Surface strain accumulation and the seismic moment tensor, *Bull. seism. Soc. Am.,* **87**(5), 1345–1353.

Savage, J., Gan, W. & Svarc, J., 2001. Strain accumulation and rotation in the Eastern California Shear Zone, *J. geophys. Res.: Solid Earth,* **106**(B10), 21995–22007.

Downloaded from https://academic.oup.com/gji/article/234/3/2128/7153326 by Missouri University of Science & Technology user on 16 August 2023

Shen, Z.-K., Jackson, D. & Ge, B., 1996. Crustal deformation across and beyond the Los Angeles basin from geodetic measurements, *J. geophys. Res.: Solid Earth,* **101**(B12), 27957–27980.

Shen, Z.-K., Jackson, D. & Kagan, Y., 2007. Implications of geodetic strain rate for future earthquakes, with a five-year forecast of M5 earthquakes in Southern California, *Seismol. Res. Lett.,* **78**(1), 116–120.

Shen, Z., Wang, M., Zeng, Y. & Wang, F., 2015. Optimal interpolation of spatially discretized geodetic data, *Bull. seism. Soc. Am.,* **105**(4), 2117–2127.

Shen-Tu, B., Holt, W. & Haines, A., 1999. Deformation kinematics in the western United States determined from Quaternary fault slip rates and recent geodetic data, *J. geophys. Res.: Solid Earth,* **104**(B12), 28927–28955.

Shepard, D., 1968 A two-dimensional interpolation function for irregularly-spaced data, in *Proceedings of the 1968 ACM National Conference, New York, 27-29 August 1968,* pp. 517–524. https://doi.org/10.1145/800186.810616

Smith-Konter, B., Sandwell, D. & Shearer, P., 2011, Locking depths estimated from geodesy and seismology along the San Andreas Fault System: implications for seismic moment release, *J. geophys. Res.,* **116,** B06401.

Stevens, V. & Avouac, J.-P., 2016. Millenary $M_W > 9.0$ earthquakes required by geodetic strain in the Himalaya, *Geophys. Res. Lett.,* **43**(3), 1118–1123.

Stevens, V. L. & Avouac, J.-P., 2021. On the relationship between strain rate and seismicity in the India–Asia collision zone: implications for probabilistic seismic hazard, *Geophys. J. Int.,* **226**(1), 220–245.

Tarayoun, A., Mazzotti, S., Craymer, M. & Henton, J., 2018. Structural Inheritance Control on Intraplate Present-Day Deformation: GPS Strain Rate Variations in the Saint Lawrence Valley, Eastern Canada, *J. geophys. Res.: Solid Earth,* **123**(8), 7004–7020.

Tape, C., Musé, P., Simons, M., Dong, D. & Webb, F., 2009. Multiscale estimation of GPS velocity fields, *Geophys. J. Int.,* **179**(2), 945–971.

Tymofyeyeva, E. & Fialko, Y., 2018. Geodetic evidence for a blind fault segment at the Southern end of the San Jacinto fault zone, *J. geophys. Res.: Solid Earth*.**123,** 878–891,

Uieda, L. *et al.* , 2021. PyGMT: a Python interface for the Generic Mapping Tools. https://doi.org/10.5281/ZENODO.5607255

Wackernagel, H., 2003. Multivariate geostatistics: an introduction with applications. Germany: Springer.

Wang, H. & Wright, T., 2012. Satellite geodetic imaging reveals internal deformation of western Tibet, *Geophys. Res. Lett.,* n/a–n/a, , **39**(7),.

Ward, S., 1994. A multidisciplinary approach to seismic hazard in Southern California, *Bull. seism. Soc. Am.,* **32**(5), 1293–1309.

Weiss, J. *et al.*, 2020. High-resolution surface velocities and strain for Anatolia from Sentinel-1 InSAR and GNSS data, *Geophys. Res. Lett.,* **47**(17), e2020GL087376.

Wessel, P., Luis, J., Uieda, L., Scharroo, R., Wobbe, F., Smith, W. & Tian, D., 2019. The Generic Mapping Tools Version 6. *Geochem. Geophys. Geosyst.,* **20**(11), 5556–5564.

Wu, Y., Jiang, Z., Yang, G., Wei, W. & Liu, X., 2011. Comparison of GPS strain rate computing methods and their reliability, *Geophys. J. Int.,* **185**(2), 703–717.

Wu, Y., Jiang, Z., Pang, Y. & Chen, C., 2021. Statistical correlation of seismicity and geodetic strain rate in the Chinese Mainland, *Seismol. Res. Lett.,* **93**(1), 268–276.

Xu, X., Sandwell, D., Klein, E. & Bock, Y., 2021. Integrated Sentinel-1 InSAR and GNSS time-series along the San Andreas fault system, *J. geophys. Res.: Solid Earth,* **126**(11), e2021JB022579.

Zeng, Y., Petersen, M. & Shen, Z.-K., 2018. Earthquake potential in California-Nevada implied by correlation of strain rate and seismicity, *Geophys. Res. Lett.,* **45**(4), 1778–1785.

Zheng, G., Lou, Y., Wang, H., Geng, J. & Shi, C., 2018. Shallow seismicity forecast for the india-eurasia collision zone based on geodetic strain rates, *Geophys. Res. Lett.,* **45**(17), 8905–8912..

Downloaded from https://academic.oup.com/gji/article/234/3/2128/7153326 by Missouri University of Science & Technology user on 16 August 2023

© The Author(s) 2023. Published by Oxford University Press on behalf of The Royal Astronomical Society. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.