

# Imaging of aseismic fault slip transients recorded by dense geodetic networks

Jeffrey J. McGuire<sup>1</sup> and Paul Segall<sup>2</sup>

<sup>1</sup>Department of Geology and Geophysics, Woods Hole Oceanographic Institution, Woods Hole MA, 02543, USA

<sup>2</sup>Department of Geophysics, Stanford University, Stanford, CA 94305, USA. E-mail: segall@pangea.stanford.edu

Accepted 2003 May 1. Received 2003 April 29; in original form 2002 July 8

## SUMMARY

The large-scale continuous GPS networks that have emerged in the last decade have documented a wide-range of aseismic deformation transients that resulted from physical processes such as aseismic fault slip and magma intrusion. In particular, a new class of ( $M \sim 7$ ) slow earthquakes with durations ranging from days to months have been observed with GPS arrays located above the downdip portion of subduction zone thrust interfaces. Interpretation of the displacement time-series resulting from these events is not straightforward owing to the contaminating effects of multiple contributing signals such as fault-slip, local benchmark motion, seasonal effects, and reference frame errors. We have developed a time-dependent inversion algorithm based on the extended Kalman filter which can separate the various signals and allow the space–time evolution of these slow-slip transients to be studied in detail. We applied the inversion algorithm to the 1999 Cascadia slow earthquake. This event had two primary episodes of moment-release separated by a two week period in which relatively little moment-release occurred. The Cascadia event and other slow earthquakes share numerous similarities with both ordinary earthquakes and afterslip transients suggesting that they may represent fault slip under a conditionally stable regime.

**Key words:** aseismic fault slip, geodetic networks, slow earthquakes.

## 1 INTRODUCTION

Most subduction zones accommodate more than half of their fault-slip aseismically (Pacheco *et al.* 1993). This large component of plate-motion has been generally inferred to occur as steady-state fault creep. However, as dense, continuous geodetic networks have started to record data on the overriding plates of a number of subduction zones it has become clear that much of the aseismic deformation occurs during relatively brief periods. There are now large-scale permanent GPS arrays above subduction zones in Japan (GEONET) (Miyazaki *et al.* 1997; Mazzotti *et al.* 2000), the Pacific Northwest (PANGA) (Miller *et al.* 2001), and elsewhere. The episodes of rapid aseismic deformation observed by these arrays are sometimes associated with earthquake afterslip (Kawasaki *et al.* 1995; Burgmann *et al.* 2001), but increasingly aseismic slip events are being identified that are not associated with any seismic earthquake making them ‘silent earthquakes’ (Beroza & Jordan 1990). Silent earthquakes have been observed by continuous GPS networks along numerous subduction zones including southwest Japan (Hirose *et al.* 1999; Ozawa *et al.* 2001), Cascadia (Dragert *et al.* 2001), the middle America Trench (Lowry *et al.* 2001), and the Aleutians (Frey Mueller *et al.* 2001). Prior to the advent of large scale

GPS arrays, aseismic slip was documented on the San Andreas Fault with strain and creep meters (Linde *et al.* 1996). Aseismic deformation transients resulting from both fault-slip (Cervelli *et al.* 2002) and dyke intrusion (Aoki *et al.* 1999) have also been observed in volcanic areas. Thus, with the advent of continuous geodetic networks a new class of deformation events has become available for detailed study.

There are numerous important reasons for engaging in detailed studies of these events which go beyond the initial documentation stage. Aseismic slip can have a strong influence on the moment budget of faults which needs to be quantified for accurate seismic hazard estimates. Aseismic slip also redistributes stress in the crust affecting the locations of future earthquakes. Transient deformations preceded some volcanic eruptions (e.g. Linde *et al.* 1993; Owen *et al.* 2000), and deep aseismic slip, similar to the recently observed transients, preceded the largest subduction zone thrust earthquake ever, the 1960 Chile Mw 9 earthquake (Kanamori & Cipar 1974; Cifuentes & Silver 1989; Linde & Silver 1989). Finally, detailed imaging of aseismic fault-slip transients may help constrain fault zone constitutive laws.

There are major challenges in developing inverse methods that can determine detailed information about the space–time variations

in the aseismic slip-rate on faults from geodetic data. First, we do not know *a priori* the nature of the time variations that we are trying to detect. Secondly, space geodetic measurements contain spatially and temporally correlated errors owing to atmospheric path delays, and in the case of GPS multipath and random benchmark motions (Wyatt 1989). This is a fundamentally different situation than the earthquake seismology inverse problem where the only significant contribution to the ground-displacement time-series is that resulting from fault-slip. Thirdly, owing to the imperfect resolution of the data, any inversion procedure will involve some spatial and temporal smoothing. It is very desirable that the amount of smoothing be determined by some rigorous procedure, not simply the analysts prejudice about what the slip-distribution should look like.

Segall & Matthews (1997) developed a recursive Kalman filter algorithm for determining the space–time variation in fault slip-rates from a set of geodetic time-series. This method, referred to as the Network Inversion Filter (NIF), incorporates several features, such as a stochastic description of local benchmark motion, that are advantageous when inverting geodetic data. The principal advantages of the NIF over other methods include: a non-parametric description of the slip rate as a function of time, and the fact that the filter operates on raw position time-series rather than derived quantities such as average displacement rate. Subsequent improvements in the technique have included separating the spatial and temporal smoothing of the slip-rate distribution (Segall *et al.* 2000) and estimation and removal of reference frame errors in the GPS data (Miyazaki *et al.* 2003). These studies have utilized linear descriptions of the underlying physical system and measurement process. The linear formulation has led to two primary limitations of the filter in its application to real data sets. First, non-linear constraints on the model, in particular enforcing slip-rate non-negativity, have not been incorporated often leading to the need to over smooth the solutions to avoid unphysical solutions (i.e. left-lateral slip on the San Andreas Fault). Secondly, the hyper-parameters which control the extent of spatial and temporal smoothing have had to be estimated separately using a maximum likelihood procedure. Using a prediction error decomposition, Segall & Matthews (1997) showed that the likelihood could be computed using a recursive filter that inverted the state-covariance matrix at each epoch. While relatively fast, compared to inverting the full data covariance matrix, this approach proved to be computationally burdensome—the full data set needed to be run through a forward filter for every choice of smoothing (hyper) parameters. Moreover, the maximum-likelihood estimates of the smoothing hyper parameters were often disregarded in favour of a significantly oversmoothed solution as a means of enforcing slip-rate positivity. In this paper we seek to correct these two related problems by employing a non-linear system and measurement model. Our new formulation allows the incorporation of non-negativity constraints through the use of pseudo-data, and incorporates the hyper-parameters directly into the state vector so that they are estimated by the filter simultaneously with the slip-rate. The new system and measurement models are solved using an extended Kalman Filter. This non-linear approach provides significantly improved spatial and temporal resolution, lower computational cost, and is equivalent to the maximum-likelihood estimate in high signal-to-noise ratio (SNR) cases which include the larger aseismic transients observed in real data.

## 2 METHOD

We model GPS station positions as a function of time  $t$ , as

$$\begin{aligned} \mathbf{X}(t) = & \sum_k H(t - t_{eq(k)}) \mathbf{X}^{\text{cos}(k)} \\ & + \int_A s_p(\boldsymbol{\xi}, t - t_0) G_{pq}^r(\mathbf{x}, \boldsymbol{\xi}) \mathbf{n}_q(\boldsymbol{\xi}) dA(\boldsymbol{\xi}) \\ & + F\mathbf{f}(t) + L(\mathbf{x}, t - t_0) + \epsilon \end{aligned} \quad (1)$$

where the first term on right hand side represents coseismic offsets  $\mathbf{X}^{\text{cos}(k)}$  at times  $t_{eq(k)}$ , and  $H(t)$  is the Heavyside function. The second term on the right hand side represents deformation due to transient aseismic slip. We relate the time dependent site motions to slip on faults in an elastic medium as a function of position  $\boldsymbol{\xi}$  and time,  $s_p(\boldsymbol{\xi}, t - t_0)$  through Green's functions  $G_{pq}^r(\mathbf{x}, \boldsymbol{\xi})$ . To date, Green's functions relating slip to surface displacement have been computed using analytical expressions for dislocations in uniform half-spaces, although there is no impediment to computing Green's functions in heterogeneous or layered media. In (1)  $p, q, r = 1, 2, 3$ , and summation on repeated indexes is implied,  $\mathbf{n}_q(\boldsymbol{\xi})$  is the unit normal to the fault surface  $A(\boldsymbol{\xi})$ .

In (1) the fault slip includes both steady-state and transient components. In some cases where modelling the steady-state deformation field is not possible (at least to sufficient accuracy), we have found it useful to separate the steady station velocities by adding a term  $\mathbf{V}_{\text{sec}} \cdot (t - t_0)$  to (1). In this case  $s_p$  in (1) represents only the transient component of fault slip. Given sufficient data it is possible to estimate the secular velocity  $\mathbf{V}_{\text{sec}}$  along with the transient fault slip. In other cases we have estimated  $\mathbf{V}_{\text{sec}}$  separately and subtracted it from the position time-series prior to filtering.

The remaining terms are related to measurement and reference frame errors. The term  $F\mathbf{f}(t)$  represent reference frame errors, where  $F$  is a linearized Helmert transformation and  $\mathbf{f}(t)$  is a vector of rigid body translations, rotations, and a scale factor (Miyazaki *et al.* 2003). The fourth term on the right hand side of (1),  $L(\mathbf{x}, t - t_0)$ , represents random benchmark wobble, which we model as a Brownian random walk with scale parameter  $\tau$  (units length time<sup>-1/2</sup>) (Wyatt 1989). While some studies have advocated a flicker noise model for benchmark motion (Mao *et al.* 1999; Zhang *et al.* 1997), Langbein & Johnson (1997) using over 10 yr of data show that a random walk fits the data quite well. Without a long time-series it is very difficult to distinguish between flicker noise (with a  $1/f$  spectral decay) and random walk (with a  $1/f^2$  spectral decay). A key feature of the Network Inversion Filter is that it can distinguish spatially correlated transient signal from site specific coloured noise. The reason for this is that elastic deformation causes a spatially coherent signal, whereas the local benchmark motions are, by definition spatially incoherent. The fact that the local benchmark motions are spatially incoherent may be more important than the precise form of the spectral decay. The final term,  $\epsilon$ , represents observation error, which we take to be normally distributed with zero mean and covariance  $\sigma^2 \Sigma^{\text{GPS}}$ , where  $\Sigma^{\text{GPS}}$  is the covariance matrix of the GPS positions, and  $\sigma^2$  is a scale factor to account for unmodelled errors such as multipath, or azimuthally varying path delays.

To estimate the time-dependant quantities  $s_p(\boldsymbol{\xi}, t - t_0)$ ,  $\mathbf{f}(t)$ , and  $L(\mathbf{x}, t - t_0)$  we incorporate (1) into a state–space model of the underlying system and measurement processes. The state–space system is most generally specified by a non-linear measurement equation:

$$\mathbf{d}_k = \mathbf{h}(\mathbf{x}_k) + \epsilon_k, \quad \epsilon_k \sim \mathbf{N}(\mathbf{0}, \mathbf{R}_k) \quad (2)$$

and a non-linear state-transition equation:

$$\mathbf{x}_{k+1} = \mathbf{t}(\mathbf{x}_k) + \delta_{k+1}, \quad \delta_{k+1} \sim \mathbf{N}(\mathbf{0}, \mathbf{Q}_{k+1}) \quad (3)$$

where the  $k$  subscripts indicate observation epoch,  $\mathbf{d}_k$  is the data vector at epoch  $k$ ,  $\mathbf{x}_k$  is the state vector at epoch  $k$  which contains all the quantities necessary to describe the system. The state transition equation predicts the state of the system at time  $k + 1$  given an estimate of the state at time  $k$ . The measurement and process noise are described by the covariance matrices  $\mathbf{R}$  and  $\mathbf{Q}$  respectively (see Appendix A).

Following Segall & Matthews (1997) we model the local benchmark motion at the  $j$ th station as a random walk with scale parameter  $\tau$

$$\mathbf{L}_j(t) = \tau \int_0^t dw(t') \quad (4)$$

where  $dw$  is formal white noise with unit variance. In this paper,  $\tau$  is assumed to have the same value for all sites, usually around 1–2 mm yr<sup>-1/2</sup> (Langbein & Johnson 1997). Similarly the slip-rate of the  $i$ th subfault is modelled as a random walk with scale parameter  $\alpha$

$$\dot{\mathbf{s}}^i(t) = \alpha \int_0^t dw(t'), \quad \mathbf{s}^i(t) = \int_0^t \dot{\mathbf{s}}^i(t') dt'. \quad (5)$$

Thus, the cumulative slip on any subfault is an integrated random walk and the slip accelerations between observation epochs are modelled as a white noise process. The hyper-parameter  $\alpha$  controls the temporal roughness of the slip-rate function (Segall & Matthews 1997).

## 2.1 Pseudo-data

There are two types of constraints we would like to enforce on the solution. The first is that the slip-rate distribution be one-sided (i.e. no left-lateral slip on the San Andreas fault). The second is that the slip-rate distribution be spatially smoothed. Following Segall *et al.* (2000) we incorporate these constraints into the state estimation process through the use of pseudo-data

$$\nabla^2 \dot{\mathbf{s}}_k = 0 + \epsilon_k, \quad \epsilon_k \sim N(0, \gamma^2 I) \quad (6)$$

$$\dot{\mathbf{s}}_k^i \geq 0 + \kappa_k, \quad \kappa_k \sim N(0, \rho^2 I). \quad (7)$$

Eq. (6) has been used successfully in previous implementations of the NIF to enforce spatial smoothing (Segall *et al.* 2000; Burgmann *et al.* 2002; Miyazaki *et al.* 2003). Eq. (7) can be reformulated as an equality constraint by introducing the dummy variable  $\lambda_k^i$  which is estimated simultaneously with  $\dot{\mathbf{s}}_k^i$  (see Hel-Or & Werman 1996)

$$\dot{\mathbf{s}}_k^i - \lambda_k^i = 0 + \eta_k, \quad \eta_k \sim N(0, \rho^2 I). \quad (8)$$

The hyper-parameters  $\gamma$  and  $\rho$  control the extent to which the two constraints are enforced in this stochastic formulation. Since one-sidedness is a physical constraint we typically fix  $\rho$  at some very small value ( $\sim 1/1000$  of the expected slip-rates) to strictly enforce it. However, in cases where the secular velocity has been subtracted from the data time-series we are actually solving for deviations in slip-rate relative to ‘steady state’. In these cases, values as low as  $-1$  times the plate motion rate are physically allowable as they would be interpreted as a region that is normally fully uncoupled becoming completely coupled. The hyper-parameter  $\gamma$  which controls the extent of spatial smoothing applied at any epoch has been estimated in

previous NIF studies by maximum-likelihood using a prediction error decomposition (Segall & Matthews 1997). This method is ideal in a low signal-noise-ratio case, but it is computationally expensive. In this paper,  $\gamma$  is estimated directly by the filter algorithm as discussed below.

## 2.2 Direct estimation of hyper-parameters

In previous linear formulations of the NIF the hyper-parameters which control spatial and temporal smoothing of the slip-rate distribution,  $\gamma$  and  $\alpha$ , have appeared in the covariance matrices of the pseudo-data noise and the process noise respectively. Since we will need to use a non-linear state-space model to incorporate the positivity constraint in (8) there is also the possibility of directly estimating  $\gamma$  and  $\alpha$  in the filter. Hyper-parameters which control the process noise, such as  $\alpha$  and  $\tau$  can be incorporated into the measurement equation by defining the associated random-walk process as:

$$\dot{\mathbf{W}}^i(t) = \int_0^t dw(t'), \quad \mathbf{W}^i(\mathbf{t}) = \int_0^t \dot{\mathbf{W}}^i(t') dt'. \quad (9)$$

with the slip and slip-rate amplitude given by  $s = \alpha W$  and  $\dot{s} = \alpha \dot{W}$ . The contribution of slip on the  $i$ th subfault to displacement at the  $j$ th station is  $G_{ij} s_i$  (see appendix A).  $\alpha$ ,  $\dot{W}^i$ , and  $W^i$  are all included in the state vector and estimated by the filter.

Hyper-parameters which specify the covariance matrices of the data and pseudo-data, such as  $\gamma$ ,  $\rho$ , and  $\sigma$ , can be directly incorporated in the observation equation by simply dividing equations like (6) by the appropriate standard deviation

$$\frac{1}{\gamma} \nabla^2 \dot{\mathbf{s}}_k = 0 + \epsilon_k, \quad \epsilon_k \sim N(0, \mathbf{I}) \quad (10)$$

Using these direct relationships between the hyper-parameters and the data/pseudo-data the hyper-parameters can then be considered part of the state-vector (see Appendix A). Augmenting the state-vector with the unknown parameters of the system model and solving them using a linearized (extended) filter is a common process in the data assimilation literature referred to as adaptive estimation (see Wunsch 1996, section 6.6.2).

## 2.3 Extended Kalman filter algorithm

The non-linear state-space system described by eqs (2) and (3) (and detailed in appendix A) can be solved for estimates of the state vector at each of the observation epochs using an algorithm known as the extended Kalman filter (see Gelb 1974). The filter is initialized with an *a priori* estimate of the state vector,  $\mathbf{x}_{0|0}$  and its covariance  $\Sigma_{0|0}$  (where the subscripts  $x_{k|j}$  indicate the estimate of the state vector at epoch  $k$  given data through epoch  $j$ ). Typical values of these priors will be discussed in the synthetic and data examples in subsequent sections. The *a priori* estimates are then predicted forward one time step (there is no requirement that subsequent time-steps be of the same size). The predicted state is then compared with the data from the first epoch and any residual is used to update the estimate of the state at epoch 1. The prediction and update process are then iterated through the entire data set.

### 2.3.1 Prediction step

$$\mathbf{x}_{k+1|k} = \mathbf{t}(\mathbf{x}_{k|k}) + \delta_{k+1}, \quad \delta_{k+1} \sim \mathbf{N}(\mathbf{0}, \mathbf{Q}_{k+1}) \quad (11)$$

$$\Sigma_{k+1|k} = \mathbf{T}_k * \Sigma_{k|k} * \mathbf{T}_k^T + \mathbf{Q}_{k+1} \quad (12)$$

### 2.3.2 Update step

$$\mathbf{x}_{k+1|k+1} = \mathbf{x}_{k+1|k} + \mathbf{K}_{k+1} * [\mathbf{d}_{k+1} - \mathbf{h}(\mathbf{x}_{k+1|k})] \quad (13)$$

$$\Sigma_{k+1|k+1} = [\mathbf{I} - \mathbf{K}_{k+1} \mathbf{H}_{k+1}] \Sigma_{k+1|k} \quad (14)$$

$$\mathbf{K}_{k+1} = \Sigma_{k+1|k} \mathbf{H}_{k+1}^T [\mathbf{H}_{k+1} \Sigma_{k+1|k} \mathbf{H}_{k+1}^T + \mathbf{R}_{k+1}]^{-1} \quad (15)$$

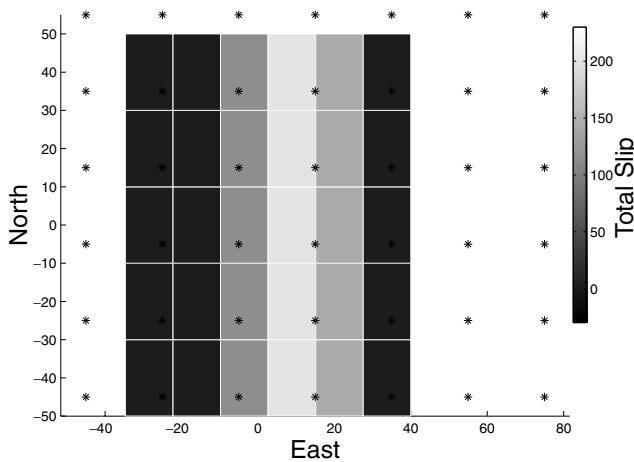
where

$$\mathbf{T}_k = \left. \frac{\partial \mathbf{t}(\mathbf{x})}{\partial \mathbf{x}} \right|_{\mathbf{x}_k} \quad \text{and} \quad \mathbf{H}_k = \left. \frac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}} \right|_{\mathbf{x}_k} \quad (16)$$

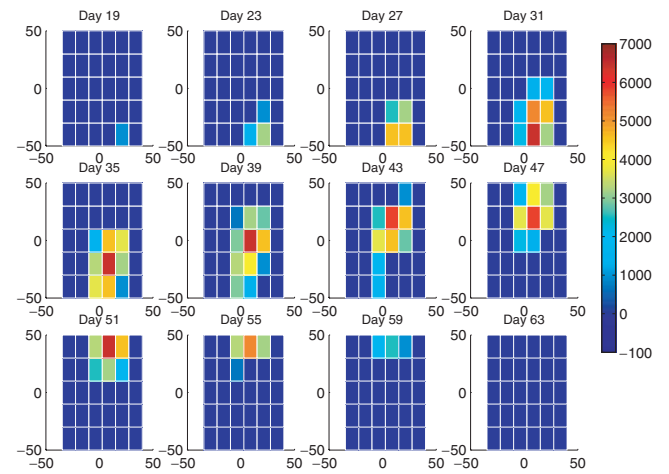
After reaching the last observation epoch we have estimates of the state vector at each epoch given data up to that epoch,  $\mathbf{x}_{k|k}$ . To obtain estimates of the state vector at each epoch given all the data,  $\mathbf{x}_{k|N}$  (where  $N$  is the number of observation epochs), it is necessary to run a similar recursive procedure backward in time. This algorithm is known as smoothing (see Rauch *et al.* 1965; Segall & Matthews 1997).

## 3 SIMULATION

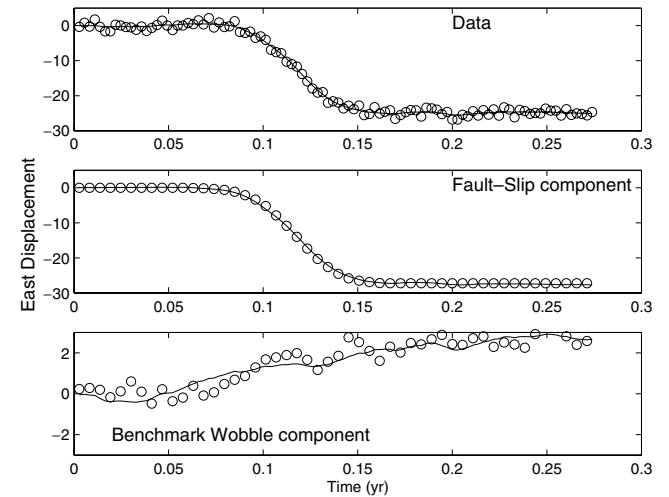
In this section we present the results of inverting a data set for a simulated aseismic slip transient on a dipping thrust fault recorded by an array of GPS receivers on the overriding plate (see Fig. 1). The transient begins near the southern downdip corner of the fault and propagates both updip and along strike with a rupture velocity of  $3 \text{ km d}^{-1}$  and a rise-time at any one point of 17 d. The total duration of the event is about 45 days. Data time-series are computed for the East and North components for the 42 sites shown in Fig. 1. Each data time-series is the sum of the signal resulting from the fault-slip, the contribution from local benchmark motion, and the contribution from measurement error (assumed to be a white noise process). While the absolute scale of the slip and resultant displacements (mm, cm, etc.) is arbitrary, we have chosen the relative size of the fault-slip, benchmark motion, and measurement error to be



**Figure 1.** Map of the station distribution (asterisk) and fault grid (coloured rectangles) for our synthetic tests. The shading indicates the final slip distribution. The fault strikes North-South and dips  $15^\circ$  to the East. The upper (Western) edge is at 10 km depth and the lower (Eastern) edge is at 30 km depth.



**Figure 2.** Maps of the true slip-rate distribution every 4 d from the start of the transient to its end. The event nucleates near the southern downdip edge of the fault and propagates updip and to the North.

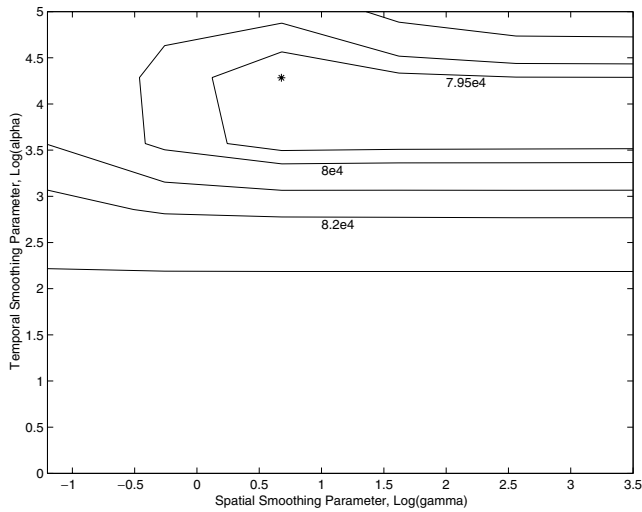


**Figure 3.** The top panel shows the East component data (blue asterisks) for the station located at  $(-5 \text{ km}, 15 \text{ km})$  in Fig. 1 which is the sum of the fault-slip contribution (middle panel), the benchmark wobble contribution (bottom panel), and white noise measurement error. The black lines in each panel show the fit from running the extended filter on the entire data set.

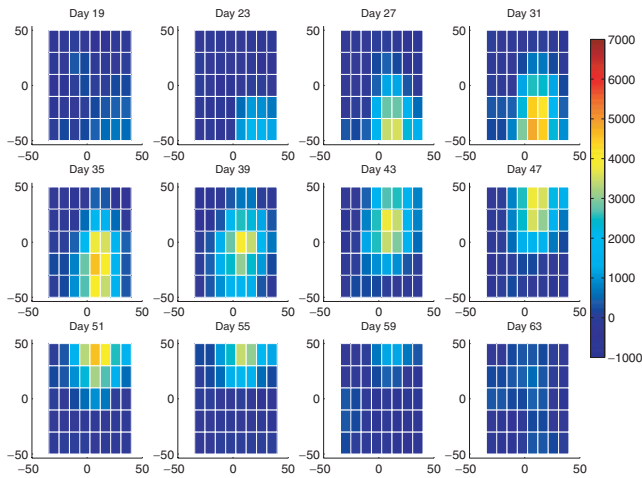
representative of the best recorded transients (e.g. Miyazaki *et al.* 2003). The measurement errors are normally distributed with a scale of 1, the benchmark motion is a random walk with a scale parameter of  $5 \text{ yr}^{-1/2}$ , and the maximum surface displacement from fault slip is 29. An example of a data time-series and its contributions from fault-slip and benchmark motion is shown in Fig. 3. The following sections discuss the inversion of this data set with positivity constraints using both the original maximum likelihood algorithm of Segall & Matthews (1997) and the new algorithm which estimates the hyper-parameters directly by including them in the state vector.

### 3.1 Maximum likelihood estimation of hyper-parameters

In previous implementations of the Network Inversion Filter, the hyper parameters which control the spatial and temporal smoothing of the slip-rate distribution,  $\gamma$  and  $\alpha$ , have been estimated before



**Figure 4.** Likelihood surface ( $-2L$ ) as a function of  $\gamma$  and  $\sigma$ . Asterisk marks the minimum, corresponding to the maximum likelihood estimate.



**Figure 5.** Maps of the estimated slip-rate using maximum-likelihood estimates of the hyper-parameters, similar to Fig. 2.

running the filter using maximum likelihood (Segall & Matthews 1997). The best estimates of the hyper parameters are found by maximizing the log likelihood  $L$  over the set of hyperparameters  $\Theta = [\sigma, \alpha, \gamma]$ . Maximizing  $L$  for the simulated data set results in an estimate of  $\alpha = 1.9e4$  and  $\gamma = 4.8$  (see Fig. 4).

Running the synthetic data through a version of the filter which used hyper-parameters fixed to the maximum likelihood values and incorporated positivity constraints, produced the slip-rate estimates shown in Fig. 5. In general the estimated slip and slip-rate distributions are similar to, but somewhat more smoothed than, the true distribution. We measured the accuracy of the inversion result in terms of the summed squared residual between the true slip and slip-rate distributions and the recovered distributions:

$$r_1 = \frac{\sum_i \sum_j (s_{i,j} - \hat{s}_{i,j})^2}{\sum_i \sum_j s_{i,j}^2} \quad (17)$$

$$r_2 = \frac{\sum_i \sum_j (s_{i,j} - \hat{s}_{i,j})^2}{\sum_i \sum_j s_{i,j}^2} \quad (18)$$

where  $s$  and  $\hat{s}$  are the true distributions and  $\hat{s}$  and  $\hat{\hat{s}}$  are the estimated ones. Thus the quantities  $r_1$  and  $r_2$  are analogous to a vari-

ance reduction. For the maximum likelihood inversion  $r_1 = .24$  and  $r_2 = .04$  (a value of 0 indicates perfect recovery of the true distribution).

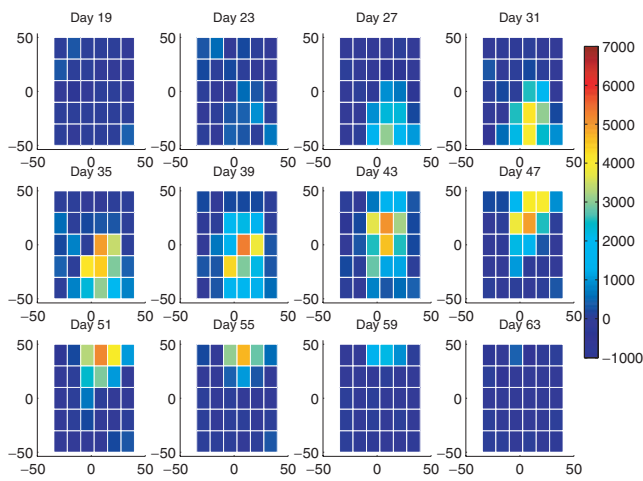
### 3.2 Direct estimation of hyper-parameters

As described in Section 2.2 and Appendix A, an alternate approach to [maximum-likelihood] estimation of  $\gamma$  and  $\alpha$  is to directly incorporate them into the state vector so that they are determined as part of the estimation process. This is useful because the filter need only be run once rather than once for each combination of hyperparameters,  $\Theta_i$ , to be tested. The key additional piece of information that is necessary to include the hyper-parameters in the state vector is an *a priori* estimate,  $x_{0|0}$  and its covariance,  $\Sigma_{0|0}$ . We assign *a priori* values of  $\alpha$  and  $\gamma$  that would result in a very smooth slip-rate distribution, but also assign very large variances so that the *a priori* information will contribute little to the final estimate. We also assume that the hyper parameters do not vary as a function of time during the observation period, i.e. there is no process noise for these state vector elements and that there is no *a priori* covariance between them. Thus the best estimate of the hyperparameters comes from the state vector at the final epoch,  $\mathbf{x}_{N|N}$ . A similar approach is taken with the prior information for the slip-rate which is set to an *a priori* value of zero with a large variance. The hyper-parameters which control the benchmark motion,  $\tau$ , and the scale parameter for the GPS covariance matrix,  $\sigma$ , are assumed known (say based on an analysis of a quiet portion of the time-series), and fixed to their correct values in these simulations as was also done in the maximum-likelihood simulation above. In principle there may be sufficient information in a long time-series to estimate the values of these parameters as well, but we have not experimented with this.

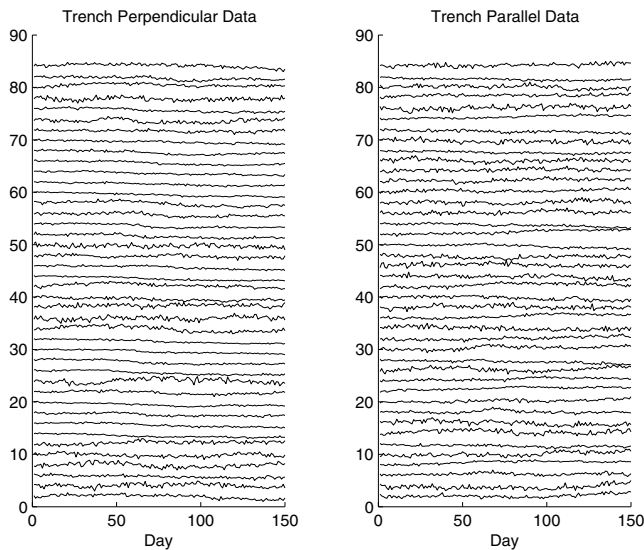
For the synthetic data set we first ran the filter with an *a priori* values of  $\alpha = 1 \pm 1000$  and  $\gamma = 1 \pm 1000$ , this produced estimates of  $\alpha = 2.2e3$  and  $\gamma = 116$  indicating that our prior variance on  $\alpha$  was too small. So the filter was rerun with an *a priori* values for  $\alpha$  and  $\gamma$  of  $1e3$  and  $1e2$ , and *a priori* variances of  $1e6$  and  $1e6$  respectively. This resulted in estimates of  $\alpha = 1.3e4 \pm 402$  and  $\gamma = 780 \pm 2e3$ . (runs with higher *a priori* values of  $\alpha$  and  $\gamma$  did not result in significantly different final estimates). These results are similar to the maximum likelihood estimate in that  $\alpha$  estimates are similar and that  $\gamma$  is poorly constrained (see Fig. 4). The estimates of the slip-rate distribution from the extended filter are shown in Fig. 6. The slip-rate estimates from this filter provide a better recovery of the true slip-rate values ( $r_1 = .14$ ,  $r_2 = .07$ ) than the maximum-likelihood filter as evidenced by the higher peak slip-rates in Fig. 6 than in Fig. 5. The datafits for one station are shown in Fig. 3. Although the input signal was purely thrust slip, we inverted for both the thrust and strike-slip components of motion on each subfault. The estimated erroneous amount of strike-slip motion estimated by the filter corresponds to only about 3 per cent of the total slip indicating that by using both horizontal components of GPS displacement, we are able to effectively resolve the azimuth of slip. In general, the filter can separate the spatial coherent signal due to fault-slip from the spatially incoherent signal resulting from benchmark-motion. Thus, the extended filter provides nearly optimal estimates of the slip-rate distribution in the high signal-to-noise ratio case without the computational cost of the maximum likelihood filter.

### 3.3 Low signal-to-noise ratio example

The previous simulation incorporated a transient with a high signal to noise ratio which is representative of the largest aseismic

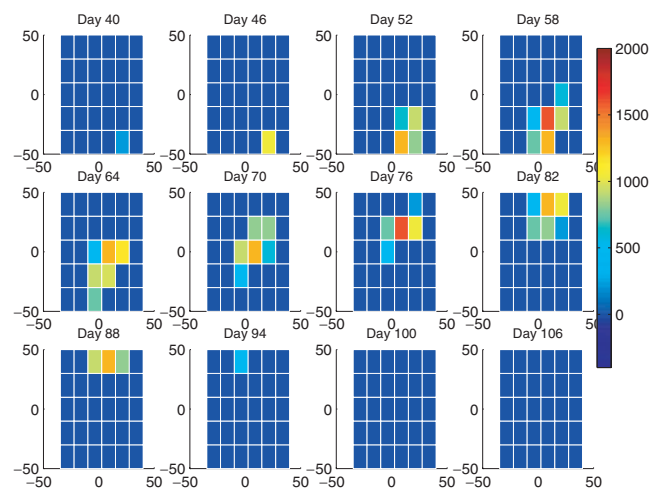


**Figure 6.** Maps of the estimated slip-rate using estimates of the hyper-parameters determined by the filter, similar to Fig. 2.

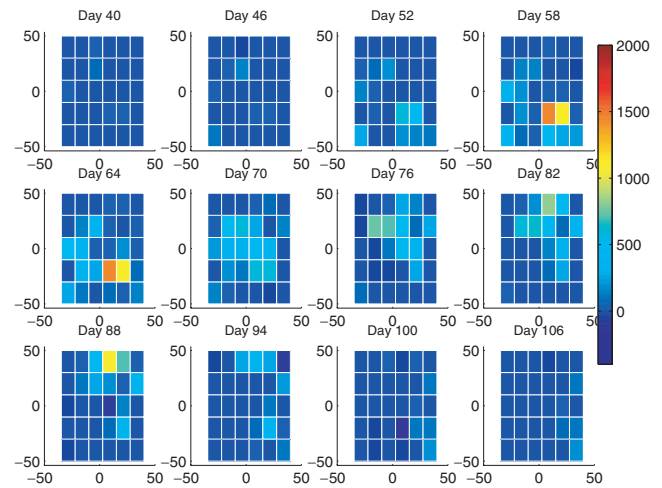


**Figure 7.** Data time-series for the low signal-to-noise ratio simulation.

transients. In this section we address the problem of detecting a transient in a noisy data set where the fault-slip signal is not unambiguously present in the GPS time-series by eye. We applied the filter to a case where the other time-dependent contributions to the GPS time-series are larger than the fault-slip contribution. The amplitudes of the fault-slip and benchmark wobble terms were chosen to have approximately equal maximum amplitude contributions to the GPS time-series. The duration of the slip-transient was also increased in this episode to increase its similarity to the benchmark wobble signals. The resulting data is shown in Fig. 7. The transient is contained between days 50 and 100 and is visible in the trench perpendicular component of some of the most sensitive stations, but numerous stations exist where the net displacement over the 150 d time period is of the opposite sign than what would be expected from fault slip owing to the contribution from the benchmark wobble term. In this low signal to noise ratio case, the filter is more unstable than the previous case in that using very large *a priori* uncertainties on  $\alpha$  and  $\gamma$  can result in (unphysical) negative values of these hyper-parameters. Thus it was necessary to start with a smooth prior for  $\alpha$  and  $\gamma$  with low uncertainty. This resulted in an estimate



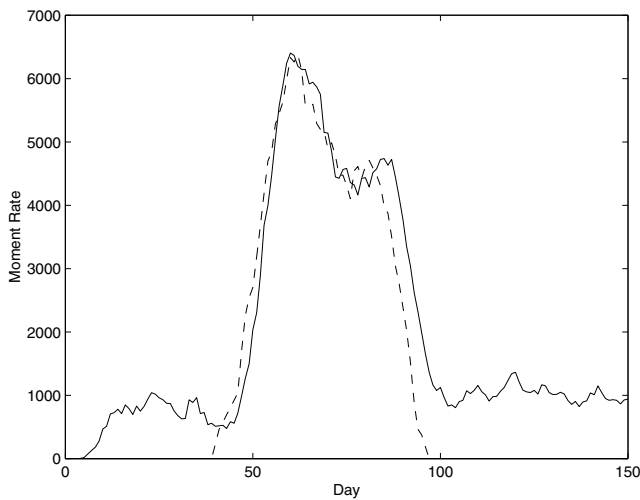
**Figure 8.** Input slip-rate distribution for the low SNR simulation.



**Figure 9.** Recovered slip-rate distribution for the low SNR simulation.

of  $\alpha$  that was larger than  $2\sigma$  above the prior, so the filter was rerun a few times gradually increasing the *a priori* estimate of  $\alpha$  and its uncertainty. The final run, when the estimate of  $\alpha$  was within  $1\sigma$  of the *a priori* values, was initiated with prior values for  $\alpha$  and  $\gamma$  of 4000 and 100 with associated uncertainties of  $1e3$ . The final estimates of the hyper parameters were  $\alpha = 4466 \pm 320$  and  $\gamma = 226 \pm 549$ .

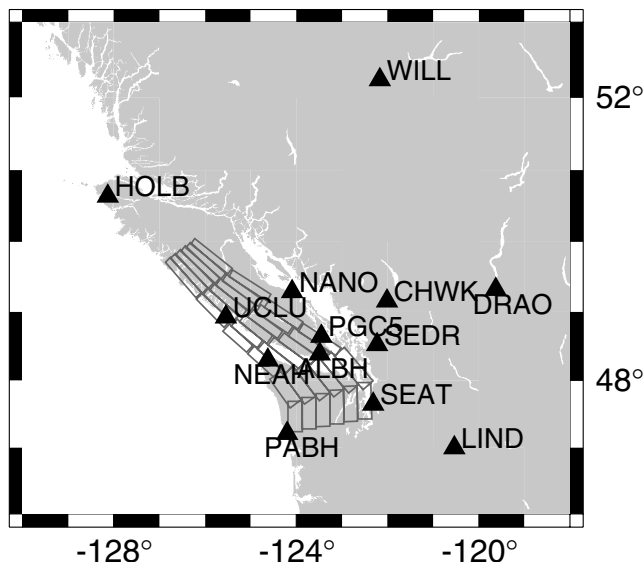
The true and estimated slip-rate distributions resulting from the low SNR test are shown in Figs 8 and 9. The estimated slip-rate distribution is qualitatively quite similar to the true distribution although the peak slip-rates are not quite recovered ( $r_1 = .71$ ,  $r_2 = .69$ ). The estimated moment-rate distribution captures the duration and even some of the internal structure of the transient (Fig. 10), but a small amount of the benchmark wobble and measurement noise is mapped into the fault-slip components of the state vector when no transient is present. Again, the input signal was purely thrust slip erroneous amount of strike-slip motion estimated by the filter remained small (about 10 per cent of the total slip). Thus even in the case with a SNR less than 1 for the overall data set, the filter is capable of pulling out an accurate and useful description of the transient's spatial extent, duration, sense of slip, and propagation.



**Figure 10.** Input (dashed) and recovered (solid) source-time functions for the low SNR simulation.

#### 4 APPLICATION TO THE 1999 CASCADIA SLOW EARTHQUAKE

In 1999 a cluster of continuous GPS stations run by the Geological Survey of Canada and the Pacific Northwest Geodetic Array (PANGA, (Miller *et al.* 2001)) network recorded a 40 day period of anomalous displacements, on the order of a few mm, resulting from slow-slip on the deep portion of the Cascadia thrust interface (Dragert *et al.* 2001). The stations which showed the anomalous displacements are contained within a  $\sim 200$  km by  $\sim 200$  km region extending from Seattle to the central portion of Vancouver Island (see Figs 11 and 12). Dragert *et al.* (2001) matched the anomalous displacements with a forward model consisting of a uniform zone of 2.1 cm of slip between about the 30 and 40 km depth contours of the thrust interface with slip tapering to zero by about  $\sim 25$  km depth. This model did not attempt to fit the time dependence of the signal beyond breaking the slip-event into three time-periods



**Figure 11.** Map showing the station distribution (black triangles and 4 letter labels) and fault-geometry (rectangles) for the 1999 Cascadia slow earthquake.

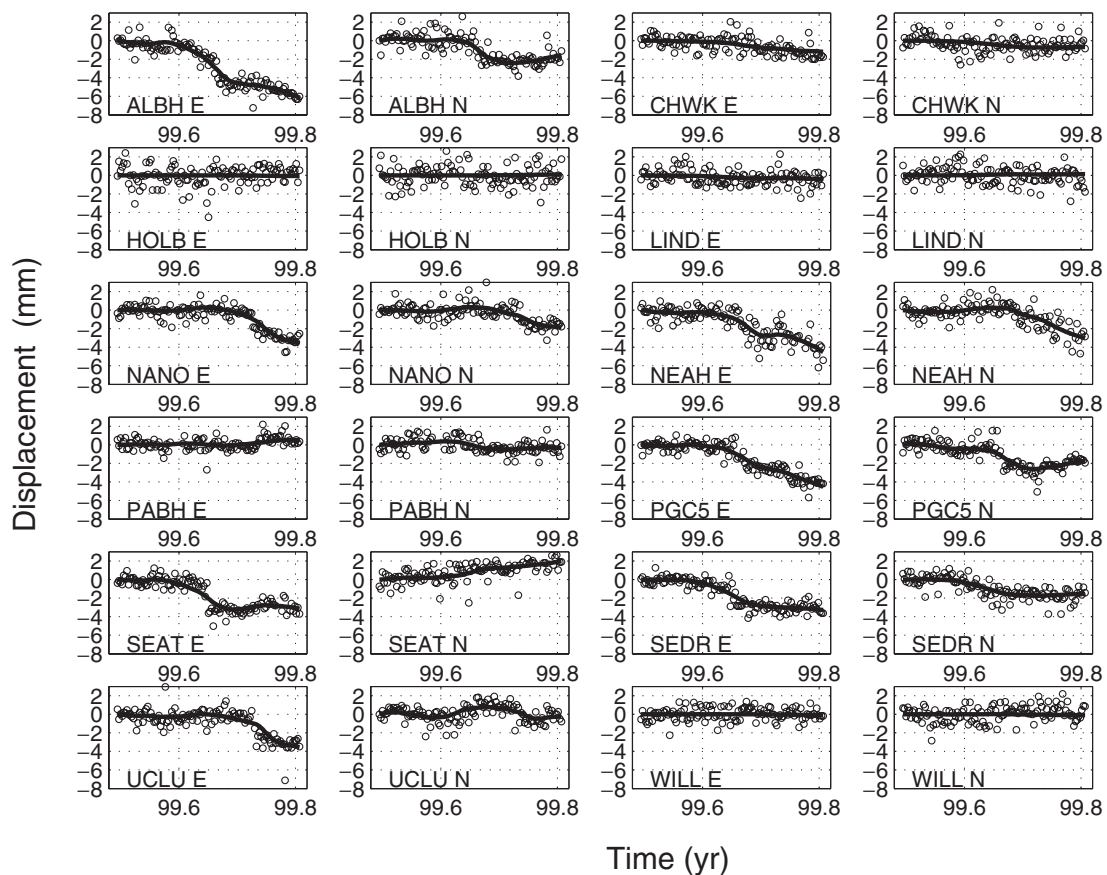
which showed general northward propagation consistent with the observed delay in the arrival of the event at the northern stations (Dragert *et al.* 2001). Here we present the results of applying the version of the NWIF that includes both positivity constraints and direct estimation of the hyper-parameters to the Cascadia data set.

We model the daily station positions (see Fig. 12) determined by (Dragert *et al.* 2001) as the sum of the contributions from transient fault slip, reference frame error, secular velocity, benchmark wobble, and measurement error. Secular velocities at these stations vary on the time-scales of years (Dragert *et al.* 2001) so to avoid mapping any of this variability into our slip-rate estimates we inverted data from a relatively short time period centered on the transient, 1999.5 to 1999.85, in which the secular-rate could be considered constant. The secular velocities were determined ahead of time using a Kalman filter that estimated the velocity, reference frame error, and benchmark wobble components of the time-series in the period between 1999.2 and 1999.6. The contribution from secular velocity only was then extrapolated and subtracted from the data time-series.

We used a 3-D fault geometry derived from the model of the Cascadia thrust interface published by Flueck *et al.* (1997) (see Fig. 11). Rectangular fault segments were fit to a set of points specifying the thrust interface at approximately 5 km depth intervals between 20 and 50 km depth. The location and dip of the interface are specified in a manner that best approximates an elastic half-space by accounting for topography and preserves the dip-angle of the thrust interface (Flueck *et al.* 1997). There were 72 subfaults with average lengths of about 25 km and widths of about 15 km. Green's functions were then calculated for the subfaults assuming a homogeneous isotropic half-space (Okada 1985). To account for station DRAO being held fixed in the GPS processing (Dragert *et al.* 2001), the Green's function for this station was subtracted from those of other stations. Despite station DRAO being held fixed in the GPS processing, short term random translations of the network are still present in the time-series (Dragert *et al.* 2001). To account for these errors, we estimated the translations of the GPS reference frame, but owing to the relatively small scale of the network, we did not solve for rotations or scale. Additional process noise was introduced in the benchmark wobble terms at a number of epochs to account for known antenna/hardware changes (Dragert, personal communication). The temporal smoothing parameter,  $\alpha$  was initialized at a value of  $50 \pm 50$  and had a final estimate of  $63 \pm 30$ . Larger values of the *a priori* variance of  $\alpha$  led to unstable hyper parameters, so the *a priori* estimate and its uncertainty were gradually increased until the final estimate was within  $1\sigma$  of the *a priori* value, similar to the low SNR simulation. The spatial smoothing parameter was initialized with a value of  $.1 \pm 1000$  and had a final estimate of  $0.8 \pm 72$ . The other hyper-parameters,  $\tau$ ,  $\sigma$ ,  $\rho$ , were fixed at  $2 \text{ mm yr}^{-1/2}$ , 10, and  $1e-4 \text{ (m yr)}^{-2}$  respectively.

The results of the filter are shown in Figs (12 and 13). The slip transients begins at depths of 30–35 km just south of station ALBH. It then expanded to the south before propagating both updip and along-strike to the north. The peak slip was about 8 cm in the region just south and updip of ALBH, while the peak slip-rates were about  $1 \text{ m yr}^{-1}$  (though this value trades off with grid-size). We estimated both the thrust and strike-slip component of slip on each subfault, the magnitude and direction of the final slip is shown in Fig. 14. On most faults the direction of slip is resolved to be roughly parallel to plate motion direction, but for the deepest row of subfaults, resolution of the azimuth (and amplitude) of slip is poorer. The moment-release distribution was not smooth in time. There are two peak periods of high slip-rates centered around day 240 and 270 separated by a roughly a 15–20 d period between about Julian day 250 and 265





**Figure 12.** East-component time-series (asterisks) for the time-period used in the filter estimation. The contributions from secular velocity, reference frame error, and benchmark wobble have been removed. The solid-lines show the fit to the data from the estimated slip distribution.

where slip-rates temporarily decreased below the  $0.5 \text{ m yr}^{-1}$  level on all subfaults (roughly 1999.71 in Fig. 13). The inferred late moment release results from the late changes at stations UCLU and NANO (Fig. 12). The total moment-release is equivalent to about an  $M_w$  6.9 event and the average stress-drop was approximately .01 MPa. The duration of the period of high slip-rate on any one subfault ranged from about 5–30 d.

## 5 DISCUSSION

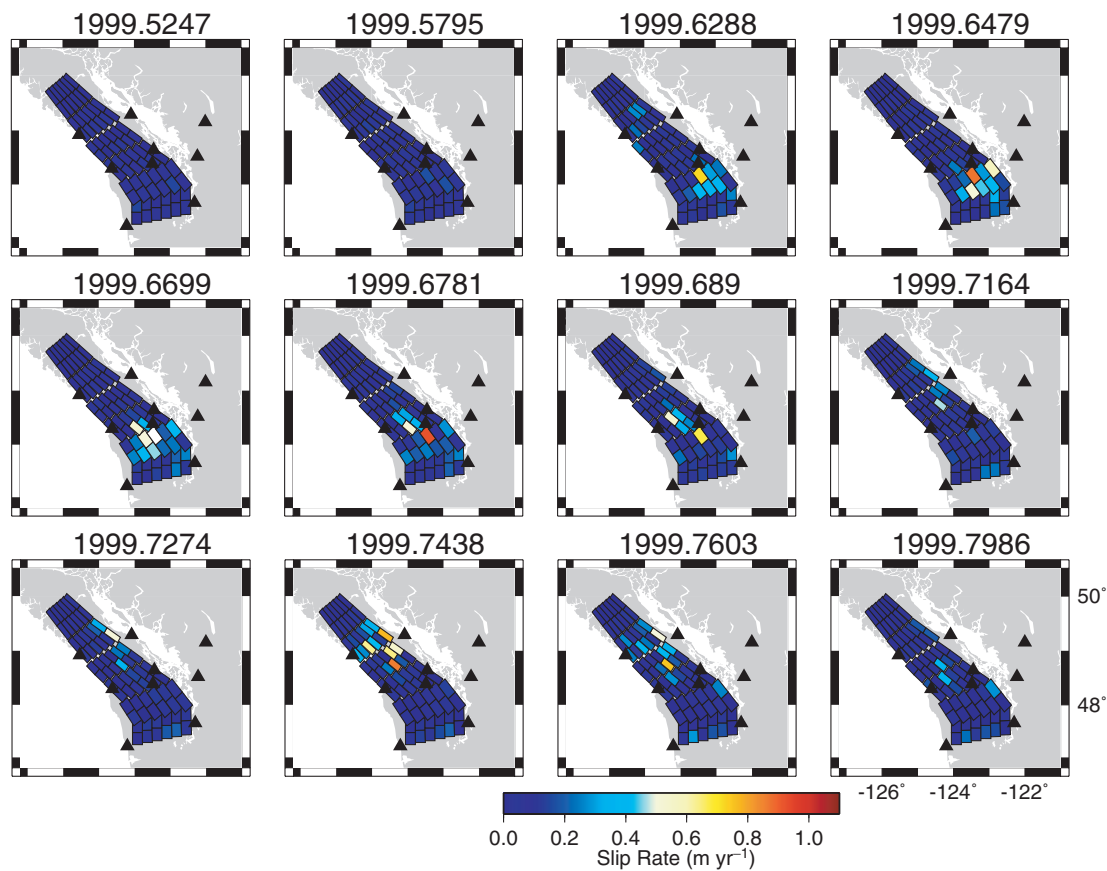
The algorithm described in the previous sections has several advantages for interpreting continuous geodetic data and may be suited for real-time fault monitoring. In particular, the inclusion of a positivity constraint on slip-rate increases the spatial resolution of slip-rate anomalies over previous versions of the NIF. This feature can be accommodated with either direct or maximum-likelihood estimation of the filter hyper-parameters. The incorporation of the hyper-parameters into the state-vector may become quite useful in real-time applications with high sample-rate data when the maximum-likelihood algorithm is too expensive to implement. Including the hyper-parameters may also suggest an event detection algorithm, as they are only forced away from the smooth *a priori* values during periods of anomalous slip.

The slip-rate anomalies shown in Fig. 13 show several characteristics that are analogous to slip in ordinary earthquakes, but with longer time-scales. The slip-rate on a portion of the fault takes a finite amount of time, from a few to  $\sim 20$  d, to reach its peak. This property is much more similar to unstable earthquake slip than the

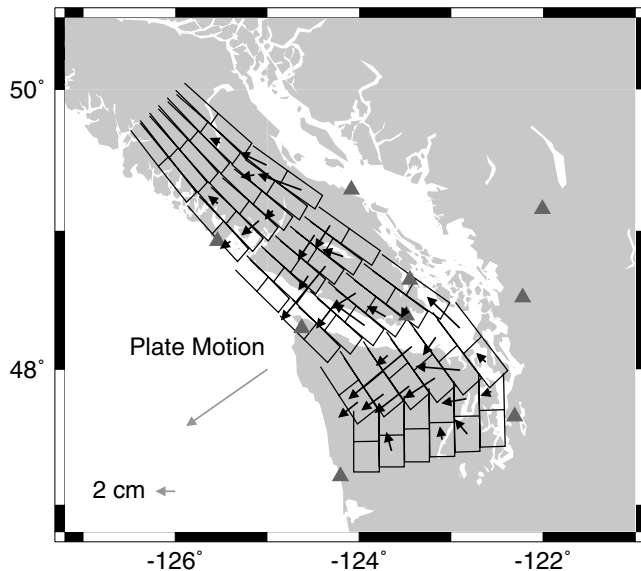
exponential decay that is usually observed in stable earthquake afterslip. The 1999 Cascadia slow-event also shows something like slip-pulse propagation where the southern portion of the fault had finished slipping before the northern portion began. The key differences between the Cascadia event and a normal earthquake are: (1) the longer time-scales, both for the total event and local rise-time and associated the lower propagation and slip velocities. (2) The much lower average stress drop resulting from the relatively small amount of slip over a large area. However, these two characteristics are very similar to typical afterslip distributions. Thus if earthquakes represent fault-slip governed by unstable velocity-weakening friction and afterslip is an example of fault-slip governed by stable velocity-strengthening friction, then silent earthquakes appear to have behaviour somewhat in-between, sharing important characteristics with both.

The aseismic slip transients that have been observed recently at numerous subduction zones Miyazaki *et al.* (2003), Dragert *et al.* (2001), Lowry *et al.* (2001), Freymueller *et al.* (2001) have several features in common. Primarily that they all rupture the portion of a subduction zone thrust interface immediately below the ‘locked’ zone, and that while the fault slip-rate clearly accelerates over a period of time, there is a limit to the acceleration and the slip does not becoming unstable. This portion of the thrust interface is predicted to be ‘conditionally stable’ in the rate-state friction formulation owing to the temperature approaching the onset of plasticity in feldspar at about  $450^\circ\text{C}$  (Scholz 1998). The conditionally stable regime of rate-state friction has been well characterized analytically and in single degree of freedom numerical simulations (Rice & Gu 1983).





**Figure 13.** Maps of the estimated slip-rate as a function of time and station distribution (black triangles).



**Figure 14.** Final slip distribution. Arrows denote magnitude and direction of slip on each fault segment for which the slip was larger than 1.0 cm. Faults with less slip than this are considered unresolved. In general the slip vectors are closely aligned with the plate-motion vector except for the deepest row of subfaults, which are poorly resolved. Grey triangles show station locations.

Under steady-state loading, faults in the conditionally-stable regime slip aseismically unless there is a sudden perturbation to the system. If the perturbation is large, the fault can be pushed over a stability boundary causing it to become unstable and slip seismically. However, if the perturbation is such that the fault approaches but does not cross the stability boundary, a period of stable sliding with increased velocity will follow as the fault evolves back to its steady-state behaviour. This type of behaviour represents one potential mechanism for explaining slow earthquakes. However the types of triggering events that could lead to a perturbation in the state of the fault, primarily the shear and normal stress, are hard to constrain for slow events. There were no nearby large earthquakes to trigger either the Cascadia slow event, and there is no clear connection between the 1996,  $M_w$  6.7, Hyuganada earthquakes and the 1997–1998 Bungo-Channel slow earthquake (Miyazaki *et al.* 2003). However, stress-state changes that initiate the slow event could result from a number of other mechanisms that are difficult to observe from the surface, such as changes in pore fluid pressure. There is also a portion of the conditionally stable rate-state parameter space where oscillatory behaviour between stable sliding and locked behaviour is observed in 1-D simulations. This has been proposed as a potential explanation for the creeping section of the San Andreas fault (Scholz 1998) where repeated creep events have been observed. If the deep slow-slip transients on subduction zones are roughly periodic events as has been suggested for Cascadia (Miller *et al.* 2002), and the 1-D simulations can be extrapolated to real faults this may provide a mechanism that doesn't require a triggering event. Additionally, recent simulations of rate-state friction on a 2-D subduction

interface have generated slow earthquakes confined to the conditionally stable region below the seismogenic zone (Shibazaki & Iio 2002).

## ACKNOWLEDGMENTS

This work was supported by NSF-EAR0074181 and NASA SENH NAG5-6167. We thank Herb Draggert for giving us access to his processed station coordinates which were derived from GPS data collected by the Geological Survey of Canada and the PANGA network. We thank Stephane Mazzotti and Takeshi Sagiya for helpful reviews.

## REFERENCES

- Aoki, Y., Segall, P., Kato, T., Cervelli, P. & Shimada, S., 1999. Imaging magma transport from inversion of deformation data: The 1997 seismic swarm off the Izu Peninsula, Japan, *Science*, **286**, 927–930.
- Beroza, G.C. & Jordan, T.H., 1990. Searching for slow and silent earthquakes using free oscillations, *J. geophys. Res.*, **95**, 2485–2510.
- Burgmann, R., Kogan, M.G., Levin, V.E., Scholz, C.H., King, R.W. & Steblov, G.M., 2001. Rapid aseismic moment release following the 5 december, 1997 kronotsky kamchatka, earthquake, *Geophys. Res. Lett.*, **28**, 1331–1334.
- Burgmann, R., Ergintav, S., Segall, P., Hearn, L., McClusky, S., Reilinger, R., Woith, H. & Zschau, J., 2002. Time-space variable afterslip on and deep below the izmit earthquake rupture, *Bull. seismol. Soc. Am.*, **92**, 126–137.
- Cervelli, P., Segall, P., Johnson, K., Lisowski, M. & Miklius, A., 2002. Sudden aseismic fault slip on the south flank of kilauea volcano, *Nature*, **415**, 1014–1018.
- Cifuentes, I.L. & Silver, P., 1989. Low-frequency source characteristics of the great 1960 Chilean earthquake, *J. geophys. Res.*, **94**, 643–663.
- Dragert, H., Wang, K. & James, T.S., 2001. A silent slip event on the deeper cascadia subduction interface, *Science*, **292**, 1526–1528.
- Flueck, P., Hyndman, R.D. & Wang, K., 1997. Three-dimensional dislocation model for great earthquakes of the cascadia subduction zone, *J. geophys. Res.*, **102**, 20 539–20 550.
- Frey Mueller, J., Zweck, C., Fletcher, H., Hreinsdottir, S., Cohen, S.C. & Wyss, M., 2001. The great alaska ‘earthquake’ of 1998–2001, *EOS, Trans. Am. Geophys. Un.*, **82**.
- Gelb, A., 1974. *Applied Optimal Estimation*, MIT Press, Cambridge, MA.
- Hel-Or, Y. & Werman, M., 1996. Constraint fusion of recognition and localization of articulated objects, *International Journal of Computer Vision*, **19**, 5–28.
- Hirose, H., Hirahara, K., Kimata, F., Fujii, N. & Miyazaki, S., 1999. A slow thrust slip event following the two 1996 Hyuganada earthquakes beneath the Bungo Channel, southwest Japan., *Geophys. Res. Lett.*, **26**, 3237–3240.
- Kanamori, H. & Cipar, J., 1974. Focal process of the great Chilean earthquake, *Phys. Earth planet. Inter.*, **9**, 128–136.
- Kawasaki, I., Asai, Y., Tamura, Y., Sagiya, T., Mikami, N., Okada, Y., Sakata, M. & Kasahara, M., 1995. The 1992 Sanriku-Oki, Japan, ultra-slow earthquake, *J. Phys. Earth*, **43**, 105–116.
- Langbein, J. & Johnson, H., 1997. Correlated errors in geodetic time-series: Implications for time-dependent deformation, *J. geophys. Res.*, **102**, 591–603.
- Linde, A.T. & Silver, P., 1989. Elevation changes and the great 1960 Chilean earthquake; support for aseismic slip, *Geophys. Res. Lett.*, **16**, 1305–1308.
- Linde, A.T., Agustsson, K., Sacks, S.I. & Stefansson, R., 1993. Mechanism of the 1991 eruption of Hekla from continuous borehole strain monitoring, *Nature*, **365**, 737–740.
- Linde, A.T., Gladwin, M.T., Johnston, M.J.S., Gwyther, R.L. & Bilham, R.G., 1996. A slow earthquake sequence on the San Andreas fault, *Nature*, **383**, 65–68.
- Lowry, A.R., Larson, K.M., Kostoglodov, V. & Bilham, R., 2001. Transient fault slip in Guerrero, southern Mexico, *Geophys. Res. Lett.*, **28**, 3753–3756.

- Mao, A., Harrison, G.A. & Dixon, T.H., 1999. *J. geophys. Res.*, **104**, 2, 2797–2816.
- Mazzotti, S., Pichon, X.L., Henry, P. & Miyazaki, S., 2000. Full interseismic locking of the Nankai and Japan-west kurile subduction zones; an analysis of uniform elastic strain accumulation in Japan constrained by permanent GPS, *J. geophys. Res.*, **105**, 13 159–13 177.
- Miller, M., Johnson, D., Rubin, C., Dragert, H., Wang, K., Qamar, A. & Goldfinger, C., 2001. Gps-determination of along-strike variation in cascadia margin kinematics: Implications for relative plate motion, subduction zone coupling, and permanent deformation, *Tectonics*, **20**, 161–176.
- Miller, M., Melbourne, T., Johnson, D.J. & Sumner, W.Q., 2002. Periodic slow earthquakes from the cascadia subduction zone, *Science*, **29**, 295.
- Miyazaki, S., Saito, T., Sasaki, M., Hatanaka, Y. & Iimura, Y., 1997. Expansion of GSI’s nationwide GPS array, *Bull. Geogr. Surv. Inst.*, **43**, 11–22.
- Miyazaki, S., McGuire, J. & Segall, P., 2003. A transient subduction zone slip episode in Southwest Japan observed by the nationwide GPS array, *J. geophys. Res.*, **108**, 10.1029/2001JB000, 456.
- Okada, Y., 1985. Surface deformation due to shear and tensile faults in a half-space, *Bull. seism. Soc. Am.*, **75**, 1135–1154.
- Owen, S., Segall, P., Lisowski, M., Murray, M., Bevis, M. & Foster, J., 2000. The january 30, 1997 eruptive event on kilauea volcano, hawaii, as monitored by continuous gps, *Geophys. Res. Lett.*, **27**, 2757–2760.
- Ozawa, S., Murakami, M. & Tada, T., 2001. Time-dependent inversion study of the slow thrust event in the nankai trough subduction zone, southwest Japan, *J. geophys. Res.*, **106**, 787–803.
- Pacheco, J.F., Sykes, L.R. & Scholz, C.H., 1993. Nature of seismic coupling along simple plate boundaries of the subduction type, *J. geophys. Res.*, **98**, 14 133–14 159.
- Rauch, H.E., Tung, F. & Striebel, C.T., 1965. Maximum likelihood estimates of linear dynamic systems, *AIJA J.*, **3**, 1445–1450.
- Rice, J. R. & Gu, J.C., 1983. Earthquake aftershocks and triggered seismic phenomena, *PAGEOPH*, **121**, 187–219.
- Scholz, C. H., 1998. Earthquakes and friction laws, *Nature*, **391**, 39–42.
- Segall, P. & Matthews, M., 1997. Time dependent inversion of geodetic data, *J. geophys. Res.*, **102**, 22 391–22 400.
- Segall, P., Burgmann, R. & Matthews, M., 2000. Time-dependent triggered afterslip following the 1989 Loma Prieta earthquake, *J. geophys. Res.*, **105**, 5615–5634.
- Shibazaki, B. & Iio, Y., 2002. Modeling silent slip events along the brittle-ductile transition for large thrust earthquakes, *Proceedings of the 3rd international ACES workshop*, p. 101.
- Wunsch, K., 1996. *The Ocean Circulation Inverse Problem*, Cambridge Univ. Press, Cambridge.
- Wyatt, F.K., 1989. Displacement of surface monuments; vertical motion, *J. geophys. Res.*, **28**, 2117–2120.
- Zhang, J., Bock, Y., Johnson, H., Fang, P., Williams, S., Genrich, J., Wdowinski, S. & Behr, J., 1997. Southern California Permanent GPS Geodetic Array: Error Analysis of Daily Position Estimates and Site Velocities, *J. geophys. Res.*, **102**, B8, 18 035–18 055.

## APPENDIX A: STATE VECTOR SYSTEM

This appendix details the non-linear state–space system used in the extended Kalman filter inversion algorithm for the case where the secular velocity term has been subtracted from the data prior to running the filter. The system is described by the observation equation,

$$\mathbf{d}_k = \mathbf{h}(\mathbf{x}_k) + \epsilon_k, \quad \epsilon_k \sim \mathbf{N}(\mathbf{0}, \mathbf{R}_k) \quad (\text{A1})$$

and the state-transition equation:

$$\mathbf{x}_{k+1} = \mathbf{t}(\mathbf{x}_k) + \delta_{k+1}, \quad \delta_{k+1} \sim \mathbf{N}(\mathbf{0}, \mathbf{Q}_{k+1}). \quad (\text{A2})$$

The state vector,  $\mathbf{x}_k$  contains all the parameters necessary to describe the state of the system, i.e. the terms on the right-hand side of eq. (1). The state-vector has elements corresponding to four types of model-parameters, fault-slip ( $W_k^j, \dot{W}_k^j, \lambda_k^j$ ), benchmark motion ( $\beta_k^j$ ), reference frame ( $f$ ), and hyper-parameters ( $\tau, \sigma, \alpha, \gamma, \rho$ ),

$$x_k = [(W_k^1, \dot{W}_k^1, \lambda_k^1), \dots, (W_k^M, \dot{W}_k^M, \lambda_k^M) \\ \beta_k^1, \dots, \beta_k^N, f_k^1, \dots, f_k^P, \tau, \sigma, \alpha, \gamma, \rho]. \quad (A3)$$

There are three types of ‘data’ in the vector  $d_k = [d_k^{\text{GPS}}, d_k^{\nabla}, d_k^+]$ , the GPS observations at epoch  $k$ , the spatial smoothing pseudo-data, and the positivity pseudo-data. The observation equation for each type of data can be written as:

$$\frac{1}{\sigma} \mathbf{d}_k^{\text{GPS}} = \frac{1}{\sigma} \mathbf{G}_{ij} \alpha \mathbf{W}_k^j + \frac{1}{\sigma} \tau \beta_k^i + \frac{1}{\sigma} \mathbf{F} \mathbf{f}(\mathbf{t}) + \epsilon_k^i, \quad \epsilon_k \sim \mathbf{N}(\mathbf{0}, \Sigma_k^{\text{GPS}}) \quad (A4)$$

$$\mathbf{d}_k^{\nabla} = \mathbf{0} = \frac{1}{\gamma} \nabla^2 \alpha \dot{\mathbf{W}}_k + \epsilon_k, \quad \epsilon_k \sim N(0, \mathbf{I}) \quad (A5)$$

$$\mathbf{d}_k^+ = \mathbf{0} = \frac{1}{\rho} \alpha \dot{\mathbf{W}}_k^i - \frac{1}{\rho} \lambda_k^{i^2} + \epsilon_k, \quad \epsilon_k \sim N(0, \mathbf{I}), \quad (A6)$$

where  $i$  and  $j$  indicate the station and subfault index respectively as in Section 2. Thus, the covariance matrix of the concatenated data vector is,

$$\mathbf{R}_k = \begin{pmatrix} \Sigma_k^{\text{GPS}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{pmatrix}. \quad (A7)$$

The extended Kalman filter algorithm uses the full non-linear  $\mathbf{h}(\mathbf{x}_k)$ , i.e. eqs (A4)–(A6), to calculate the residual at each epoch. The update step uses the partial derivatives of eqs (A4)–(A6) with respect to the elements of the state-vector to form the matrix  $\mathbf{H}$ .

The state transition equation is specified by the stochastic models assumed for each process, which are discussed in the main text. In our case, it is actually linear,  $\mathbf{x}_{k+1} = \mathbf{T} \mathbf{x}_k + \delta_{k+1}$ , where

$$\mathbf{T} = \begin{pmatrix} \begin{pmatrix} 1 & \Delta t & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{I} \Delta t & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} \quad (A8)$$

$$\mathbf{Q} = \begin{pmatrix} \begin{pmatrix} \frac{1}{3} \Delta t^3 & \frac{1}{2} \Delta t^2 & 0 \\ \frac{1}{2} \Delta t^2 & \Delta t & 0 \\ 0 & 0 & \frac{4}{5} \sqrt{\Delta t} \end{pmatrix} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{I} \Delta t & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} & \Sigma_f \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} \quad (A9)$$

The upper-left submatrices result from the random-walk and integrated random-walk models assumed for  $\dot{\mathbf{S}}$  and  $\mathbf{S}$  respectively. The values for the elements of  $\mathbf{Q}$  corresponding to the positivity dummy variable  $\lambda$  were determined from numerical simulations of random-walk time-series. The hyper-parameters are assumed to be time-invariant, and  $\Sigma_f$  is the time-independent covariance assumed for the reference frame errors (Miyazaki *et al.* 2003).