Trajectory models and reference frames for crustal motion geodesy

Michael Bevis · Abel Brown

Abstract We sketch the evolution of station trajectory models used in crustal motion geodesy over the last several decades, and describe some recent generalizations of these models that allow geodesists and geophysicists to parameterize accelerating patterns of displacement in general, and postseismic transient deformation in particular. Modern trajectory models are composed of three sub-models that represent secular trends, annual oscillations, and instantaneous jumps in coordinate time series. Traditionally the trend model invoked constant station velocity. This can be generalized by assuming that position is a polynomial function of time. The trajectory model can also be augmented as needed, by including one or more logarithmic transients in order to account for typical multi-year patterns of postseismic transient motion. Many geodetic and geophysical research groups are using general classes of trajectory model to characterize their crustal displacement time series, but few if any of them are using these trajectory models to define and realize the terrestrial reference frames (RFs) in which their time series are expressed. We describe a global GPS reanalysis program in which we use two general classes of trajectory model, tuned on a station by station basis. We define the network trajectory model as the set of station trajectory models encompassing every station in the network. We use the network trajectory model from the each global analysis to assign prior position estimates for the next round of GPS data processing. We allow our daily orbital solutions to relax so as to maintain their consistency with the network polyhedron. After several iterations we produce GPS time series expressed in a RF similar to, but not identical with ITRF2008.

We find that each iteration produces an improvement in the daily repeatability of our global time series and in the predictive power of our trajectory models.

Keywords GPS · Station trajectory models · Polynomial · Logarithmic transients · Reference frames · Time series analysis

1 Introduction

The problems of realizing a terrestrial reference frame (RF), and expressing the position and displacement history of geodetic stations in that frame are very strongly coupled (e.g. Bevis et al. 2012a). In this work, we focus on one particular coupling: the mathematical models used by geodesists and geophysicists to characterize the position of geodetic stations as a function of time. We refer to purely kinematical models of this kind as *station trajectory models,* or trajectory models for short. For the geodesist, these models are useful because they provide a simple and compact means to predict the position of a given station on a given day. This task constitutes the very essence of RF realization: constraining the coordinates of some stations at a given epoch in order to determine, via measurements of relative position, the coordinates of other stations at that same epoch. The more precisely and more consistently (over time) one can predict the geometry of a reference network, the more consistent and stable the associated RF becomes. This is a major concern for most geodesists. Geophysicists are usually more concerned about station trajectories because of the insights they give into the dynamical behavior of the solid earth. But everyone benefits from improved RF stability, since daily 'jitter' of the frame (or, more properly its realization) increases the scatter or noise in station coordinate time series, thus diminishing

M. Bevis (✉) · A. Brown
Division of Geodetic Sciences, School of Earth Sciences,
Ohio State University, Columbus, OH, USA
e-mail: mbevis@osu.edu

our ability to resolve and characterize the phenomenology of station displacement.

Geophysicists will go to great lengths to examine, analyze and model crustal motion time series, following Playfair's (1802) dictum that it is wiser to inquire into nature's secrets than to guess at them. A geophysical research group might spend a year analyzing in exhaustive detail the precise nature of the crustal displacements observed in a given region in the aftermath of a single great earthquake, and what can be inferred from those observations. It is not surprising then, that geophysicists have driven much of the recent innovation in trajectory models. Not all of these innovations will be widely adopted by geodesists, especially those engaged in operational analysis, because for geodesists a station trajectory model is really useful only if it is easy to deploy, computationally inexpensive, and fairly widely applicable.

Although many geodetic and geophysical research groups are using 'modern' trajectory models to characterize their crustal displacement time series, very few if any of them are using these trajectory models to define (or refine) and realize the frames in which their time series are expressed. We have been incorporating modern, rather general classes of trajectory model into the 'upstream' as well as the 'downstream' segments of our geodetic workflow for more than 3 years. Upstream we use them to generate the prior station coordinates required by our GPS data processing engine GAMIT/GLOBK (Herring et al. 2010), thereby influencing our daily solutions for the GPS satellite orbits and for the geometry of the global GPS network. After each global analysis or reanalysis is completed, we update the station trajectory models for each station in the network. After several iterations the prior coordinates being injected to the GPS processing engine are no longer consistent with the predictions of ITRF2008 (Altamimi et al. 2011). The differences are modest but they are systematic, and therefore our network time series are actually expressed in a slightly different RF that we call OSU08.

There are two main themes in this paper. First, we wish to review the topic of station trajectory models in a manner that is easily accessible to non-specialists and to students entering our field. This review—perhaps tutorial would be a better term—is not exhaustively comprehensive, and it does emphasize the classes of trajectory model we use at Ohio State University (OSU). Towards the end of this survey, we discuss our experience incorporating the logarithmic transient formula, now widely used to characterize postseismic transients, into our trajectory models. Unlike most discussions of this topic, our approach is pragmatic and focused primarily on the *geodetic* utility of this class of trajectory model, not its geophysical significance.

We then develop our second theme—the potential impact of modern trajectory models on the way in which we define and realize reference frames (RFs). At high school we learned

that a set of axes allows us to give coordinates to a set of points. But for geodesists, it is the giving of coordinates to a set of reference stations that, in effect, defines the axes. Therefore, when a RF is to operate not just at a single epoch, but continuously over an extended period of time, we have to invoke this axis system by specifying the coordinates of the reference stations as functions of time. That is, we must specify the trajectories of these reference stations. If we define a *network trajectory model* as a set of station trajectory models for every station in a network, then, in its operational context, a RF is a network trajectory model for a set of reference stations. It follows that if modern station trajectory models are improving our ability to characterize displacement time series expressed in standard RFs, then these improved trajectory models should also allow us devise and realize more consistent (and 'stable') RFs.

Again, we hope to develop this theme in a fashion easily understood by non-specialists. We will invoke the concept of *inner geometry*, meaning RF-independent geometry. This fundamental idea pervades modern space geodesy, but its ubiquity is sometimes obscured by the variety of language in which it is expressed. To understand the potential impact of modern trajectory models on the RF and the time series expressed in these frames, it is crucial to distinguish between the geometrical alignment or stacking of a network time series, and the subsequent transformation of the aligned time series into a target RF. This two-stage approach is widely used in the space geodetic community, but its nature sometimes proves difficult for non-specialists to grasp. Therefore, we seek to explain the posterior approach to RF realization, and its nexus with trajectory modeling, using consistent and easily understood language.

2 The evolution of station trajectory models

About 50 years ago, before Alfred Wegener's theory of continental drift (or plate tectonics) was widely accepted, and well before modern space geodetic techniques had been developed, geodesists realized their RFs and described their geodetic networks by assigning three spatial coordinates to each geodetic station. In principle, these coordinates were constant over time. Static coordinates are best suited to a static earth, but also serve on a dynamic earth so long as extant positioning methods lack the precision to detect tectonic motions over extended periods of time (say, one decade). But with the advent of very long baseline interferometry (VLBI), satellite laser ranging (SLR), and global positioning system (GPS) geodesy, plate motions were easily resolved over a year or two, and eventually over even shorter periods of time. This level of positioning accuracy rendered static global coordinate systems permanently obsolete. The simplest conceivable reference system for a dynamic earth invokes a *constant*

velocity model (CVM) in which the geocentric Cartesian coordinates (X, Y, Z) for each geodetic station are expressed as a linear function of time, thus

$$
\begin{aligned}
X(t) &= X_R + V_x(t - t_R) \\
Y(t) &= Y_R + V_y(t - t_R) \\
Z(t) &= Z_R + V_z(t - t_R)
\end{aligned}
\tag{1}
$$

or in vector form

$$
\mathbf{x}(t) = \mathbf{x}_R + \mathbf{v}(t - t_R)
\tag{2}
$$

where t_R is a reference time (adopted by convention), $\mathbf{x}_R = \mathbf{x}(t_R) = [X_R\, Y_R\, Z_R]'$ is the reference position, and $\mathbf{v} = [V_x\, V_y\, V_z]'$ is the station velocity. In this model, the trajectory of each station, i.e. its position as a function of time, is completely described by its reference position, $\mathbf{x}_R$, and its velocity, $\mathbf{v}$. That is, six parameters are used to describe each station, instead of three as before. Some groups and services, e.g. the International Earth Rotation and Reference Systems Service (IERS), adopt a single reference time for all stations used to define a given reference frame, while others (including us) prefer station-specific reference times set by convention to the mean epoch (measurement time) of the time series used to estimate the six parameters of the CVM. This convention minimizes the statistical correlation between the reference position and velocity estimates.

As networks of continuous GPS (CGPS) stations proliferated around the world in the late 1980s, it very soon became clear that the CVM was of limited utility in that it could not accommodate the sudden offsets or 'jumps' observed in many geodetic time series. In some cases these jumps were coseismic displacements produced by local or large regional earthquakes. But most coordinate jumps were caused by changes in the GPS hardware, especially the antenna or its radome, and thus were artificial discontinuities in the time series since the ground itself had not really moved. Artificial or not, these sudden jumps in coordinates must be accounted for. One way to do this is to rename the station so as to establish a new reference position. But the more elegant and more manageable approach is simply to model each jump using a Heaviside function, sometimes referred to as the 'unit step' function, defined thus

$$
\begin{aligned}
H(t) &= 0 && \text{for } t < 0 \\
H(t) &= 1/2 && \text{for } t = 0 \\
H(t) &= 1 && \text{for } t > 0
\end{aligned}
\tag{3}
$$

The CVM was modified to accommodate one or more jumps at specified times $\{t_j\}$, $j = 1 : n_J$, thus:

$$
\mathbf{x}(t) = \mathbf{x}_R + \mathbf{v}(t - t_R) + \sum_{j=1}^{n_J} \mathbf{b}_j H(t - t_j)
\tag{4}
$$

where $\mathbf{b}_j$ characterizes the jump which occurs at time t_j as an instantaneous displacement vector specified in geocentric Cartesian coordinates. That is, the three components of $\mathbf{b}_j$ are the jumps that occur in the X, Y, and Z coordinates of the station at the instant of the jth jump. The velocity $\mathbf{v}$ remains constant. This model (Eq. 4) is also referred to as the CVM, but it can be called the 'CVM with jumps', if the distinction from Eq. (2) is important.

This modified trajectory model (Eq. 4) was widely adopted by the international geodetic community by the late 1980s, and it provided an adequate description of the great majority of geodetic CGPS time series until the late 1990s, when geophysicists and geodesists began to resolve annual oscillations in their geodetic time series. These oscillations were usually much larger in the vertical direction than in the horizontal. With the notable exception of Dr. M. Murakami in Japan, who argued they were real, these seasonal cycles were largely ignored or dismissed as artifacts (e.g. due to mismodeling of atmospheric delays) until Heki (2001) vividly demonstrated that they were actual oscillations of the ground manifesting earth's elastic response to seasonal changes in the loads imposed upon the lithosphere by the overlying environment—especially the loads associated with surface water, snow and ice (see also Mangiarotti et al. 2001; Van Dam et al. 2001; Blewitt et al. 2001). Dong et al. (2002) demonstrated that for most GPS stations these annual oscillations could be well approximated by a four-term Fourier series consisting of two annual and two semi-annual terms. They also argued that it was desirable to include this truncated Fourier series in the trajectory model so as to prevent non-integer numbers of displacement cycles affecting, or "leaking into" estimates of a station's reference position and velocity. They implied that this could, in principle, improve the stability of the terrestrial RF. Today, many, if not all, crustal motion geodesy groups routinely include annual displacement cycles in their station trajectory models, which usually take the form:

$$
\mathbf{x}(t) = \mathbf{x}_R + \mathbf{v}(t - t_R) + \sum_{j=1}^{n_J} \mathbf{b}_j H(t - t_j)
$$
$$
+ \sum_{k=1}^{n_F} [\mathbf{s}_k \sin(\omega_k t) + \mathbf{c}_k \cos(\omega_k t)]
\tag{5}
$$

where n_F is the number of frequencies used to model the annual displacement cycle, and

$$
\omega_k = \frac{2\pi}{\tau_k}, \quad \text{and } \tau_1 = 1\,\text{year},
$$
$$
\tau_2 = 1/2\,\text{year}, \quad \tau_3 = 1/3\,\text{year, etc.}
$$

This trajectory model, which superimposes both Heaviside jumps and an annual cycle on a constant velocity trend, is linear with respect to its parameters or coefficients, which number $3 * (2 n_F + n_J + 2)$ in all. That is, $(2 n_F + n_J + 2)$ trajectory
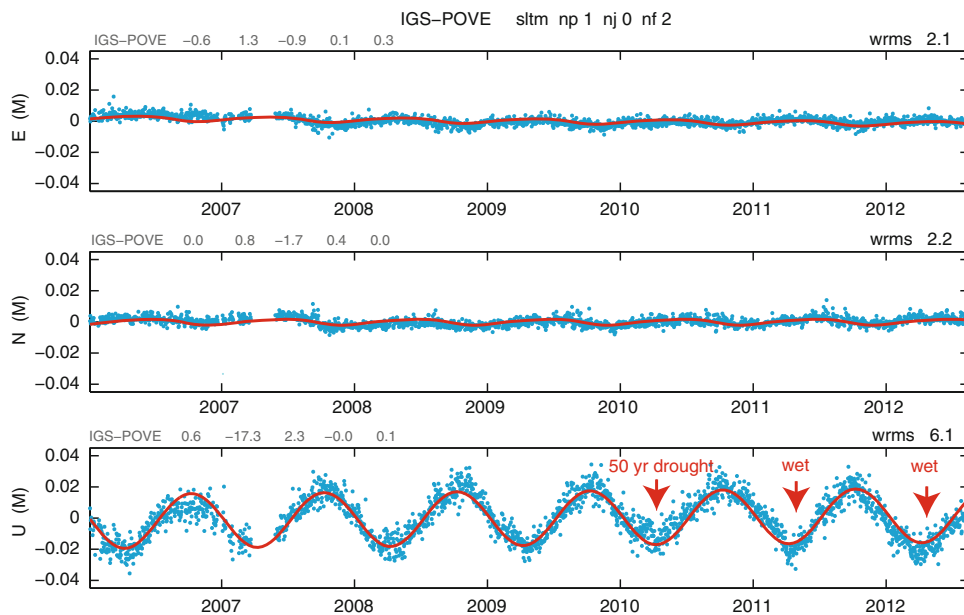
Fig. 1 Daily displacement time series (*blue dots*) at CGPS station POVE in Puerto Vehlo, Brazil, and its best-fit trajectory model (*red curves*). The three *plots* depict the east (*E*), north (*N*) and up (*U*) components of displacement in meters. Annual displacement cycles at POVE are driven by seasonal variations in the mass of water residing in the Central Amazon Basin (e.g. Bevis et al. 2005). The vertical cycle has a peak to peak amplitude close to 40 mm, about an order of magnitude larger than the horizontal cycles. Note that interannual variations in the vertical cycle can be correlated with droughts or unusually wet rainy seasons. The five numbers above the *left side* of each plot are the velocity in mm/year and coefficients (s1, c1, s2, c2) of the truncated Fourier series in mm. The WRMS misfit of data and model is given in mm at *top right*. The super-label at the *top* of the page specifies the structure of the SLTM. All displacements are expressed in a RF attached to the stable core of the South American plate

parameters are needed to describe the temporal evolution of each station coordinate X, Y, or Z. In our experience, it is only rarely necessary to set $n_F > 2$, but setting $n_F = 2$ usually produces a significantly better fit to the observed time series than does setting $n_F = 1$. Figures 1 and 2 depict rather striking examples of seasonal oscillations at GPS stations POVE, in Brazil, and NETT, in New Zealand.

While this trajectory model (Eq. 5) is well known and fairly widely employed, it has never, to our knowledge, been used to *define* a RF. Before focusing on this issue, we describe two further generalizations of the station trajectory model that we frequently use in our work at OSU. The first generalization allows us to model stations whose displacement trends involve significant and sustained accelerations. The second generalization allows a trajectory model to follow the displacements of stations affected by post-seismic transients.

3 Incorporating non-steady displacement trends

The trajectory model presented in Eq. (5) is composed of three sub-models:

$$\mathbf{x}(t) = \mathbf{x}_{\text{trend}} + \mathbf{x}_{\text{jumps}} + \mathbf{x}_{\text{cycle}} \tag{6}$$

with the first sub-model, which accounts for the multi-year trend in position, being the CVM (Eq. 2). How-

ever, sustained accelerations are being observed at increasing numbers of CGPS stations, especially near or within active ice sheets (Khan et al. 2010; Jiang et al. 2010; Bevis et al. 2012b), but also in other settings such as active volcanoes. We can accommodate such behavior by replacing the CVM (Eq. 2) with the polynomial trend model

$$\mathbf{x}_{\text{trend}} = \sum_{i=1}^{n_P+1} \mathbf{p}_i (t - t_{\text{R}})^{i-1} \tag{7}$$

where n_P is the order or maximum power of the polynomial. If $n_P = 1$, then this model reduces to the CVM, with $\mathbf{p}_1 = \mathbf{x}_{\text{R}}$ and $\mathbf{p}_2 = v$. If $n_P = 2$, it becomes the quadratic trend or 'constant acceleration' model in which the acceleration vector $\mathbf{a} = 2\mathbf{p}_3$. No matter what value is assigned to n_P, $\mathbf{p}_1$ always corresponds to the reference position $\mathbf{x}_{\text{R}}$. In practice n_P should not be set larger than 5 or 6, unless the trend model is reformulated to use orthogonal polynomial functions. In our experience it is only rarely necessary to set $n_P > 3$. For a very large majority of CGPS stations, it is perfectly adequate to set $n_P = 1$.

Having made this substitution we obtain the trajectory model known at OSU as the standard linear trajectory model (SLTM):
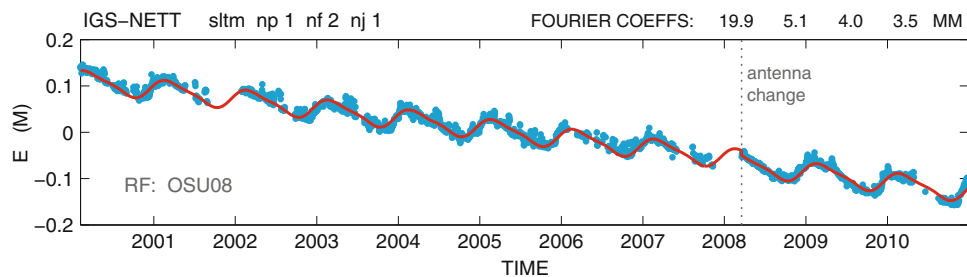
Fig. 2 The CGPS station NETT in the Southern Alps of New Zealand exhibits one of the largest horizontal displacement cycles observed anywhere on earth. This cycle is much stronger in the *E* component of displacement (shown) than in the other components (not shown). Note that the SLTM incorporates a single Heaviside jump which accommodates an artificial displacement associated with an antenna change. This time series is expressed in ITRF2008

$$\mathbf{x}(t) = \sum_{i=1}^{n_P+1} \mathbf{p}_i (t - t_R)^{i-1} + \sum_{j=1}^{n_J} \mathbf{b}_j H(t - t_j)$$
$$+ \sum_{k=1}^{n_F} \mathbf{s}_k \sin(\omega_k t) + \mathbf{c}_k \cos(\omega_k t) \qquad (8)$$

This vector equation can be thought of as a system of three scalar equations describing the temporal evolution of the *X*, *Y*, and *Z* coordinates, respectively. Indeed, operationally it is most practical to work with the three scalar equations one at a time, because, in analyzing the displacement history at a given station, we can solve for the three sets of trajectory parameters independently, given that the *X*, *Y*, and *Z* axes are orthogonal. Note that the trajectory model for each Cartesian coordinate involves a total of $(2 n_F + n_J + n_P + 1)$ parameters or coefficients. In estimating the parameters of this model by inverting a station position time series $\{t_i, \mathbf{x}_i\}, i = 1 : m$, it is useful to keep the architecture of the model (Eq. 6) in mind. For instance, we build the design matrix (which we partition) from the individual design matrices associated with each submodel, and, after the inversion, we frequently decompose the trajectory model into its component parts.

By way of an example, we show a trajectory model fitted to our coordinate times series for the West Greenland CGPS station KELY at Kellyville. This time series, and all others shown in this paper, were obtained using GAMIT/GLOBK software (Herring et al. 2010) at OSU. We fit this time series with a trajectory model composed of a quadratic trend (by setting $n_P = 2$), one Heavyside jump ($n_J = 1$), and the standard four-term Fourier series ($n_F = 2$). As usual, it is instructive to transform the data and the trajectory model into the local (topocentric) cartesian axis system [*E, N, U*] in which *E* is east, *N* is north and *U* is ellipsoid-normal up. We then see that the jump, the annual oscillations and the accelerations are all larger in the vertical direction (Fig. 3). The jump, which occurred at 17:18 Z on 2001/09/14, was associated with a change of receiver, antenna and radome as well as a change in the antenna reference point. This produced an

apparent downward jump of 61.1 mm, and horizontal jumps of 3.5 mm W and 3.4 mm S. Since this jump is artificial in the sense that the earth's crust did not really jump at this time, it is useful to remove the Heaviside jump from the data and the trajectory curves so as to isolate actual crustal motion (Fig. 4). Clearly, KELY has reversed the direction of its vertical movement during the course of its lifetime (Jiang et al. 2010).

We quantify these rate changes in Fig. 5, which shows the 'velocity trend', that is the station velocity history once the oscillatory velocity changes associated with the seasonal displacement cycle have been excluded. This is achieved by taking the first derivative of the polynomial model for displacement trend. We see that the *U* component of velocity changed from about −2.8 mm/year in 1996.0 to about +4.1 mm/year in 2010.4. The acceleration estimate is 0.49±0.02 mm/year². The accelerations are much smaller in the *E* component (−0.10 ± 0.01 mm/year²) and negligible in the *N* component (−0.04 ± 0.02 mm/year²). As KELY has accelerated upwards in response in accelerating ice loss to its east (Khan et al. 2010; Jiang et al. 2010), it also accelerated horizontally, nearly westwards. This 'up and away' acceleration is the classic 'Boussinesq response' for a point located outside of a region of accelerating mass loss (Becker and Bevis 2004; Bevis et al. 2012b).

It is possible to characterize accelerating patterns of displacement without using polynomial trend models. Some recent works have invoked a *linear spline* or *polyline trend model* instead (e.g. Khan et al. 2010). This divides the total time span encompassing a station time series into two or more adjoining intervals, and invokes a CVM in each time interval, but requires these line segments to connect at the boundaries between time intervals. The resulting polyline is continuous, but its gradient can be and usually is discontinuous at the boundaries of the time intervals.

Note that the linear spline model is distinct from the *piecewise constant velocity model* (PCVM) sometimes invoked by the ITRF, since the line segments in this case are not required

Fig. 3 The crustal displacement time series at station KELY in West Greenland, and a standard linear trajectory model invoking a quadratic trend, a single Heaviside jump, and a four-term Fourier series (*red curve*). The WRMS misfit of data and model is shown (in mm) on the *top right* side of each *plot*. This time series is expressed in a frame that minimizes horizontal motion at long-lived CGPS stations within Greenland
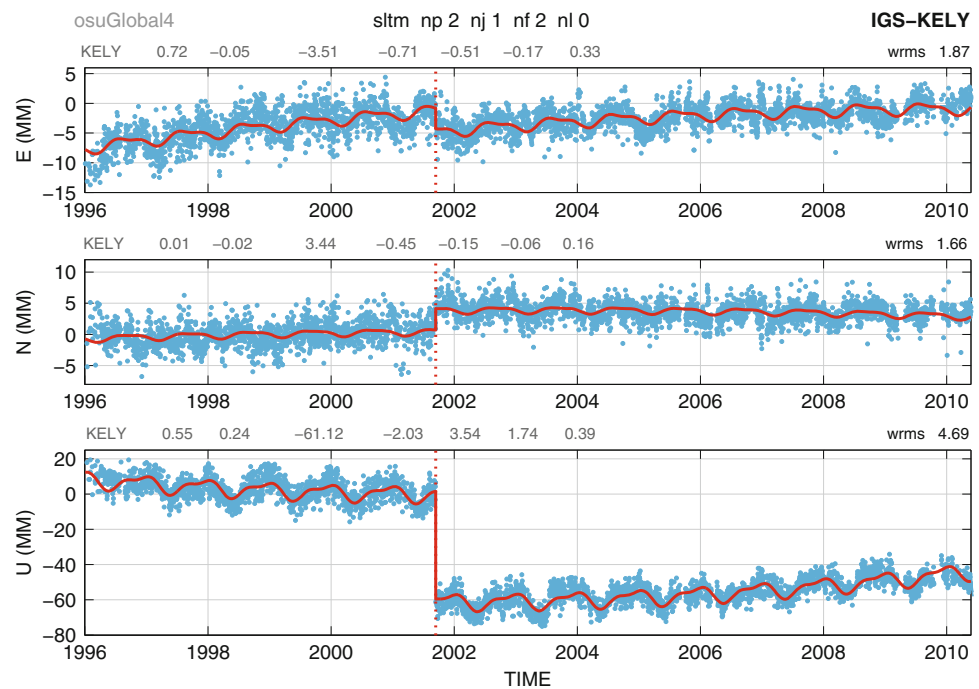

Fig. 4 The displacement time series and trajectory model shown in Fig. 3 after the Heaviside jump has been removed from the data and the model. The location of the jump, which was an artifact of an antenna and radome change, is shown by the *dashed red line*. These plots now depict the actual motion of the ground. Clearly the direction of vertical motion has reversed during the course of KELY's lifetime
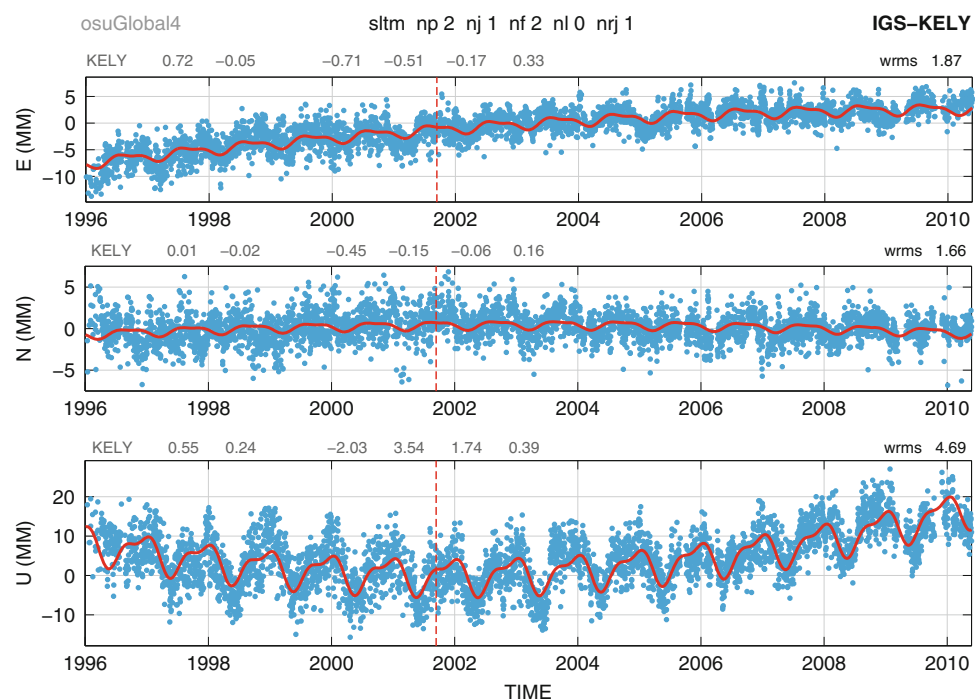
to join at the boundaries between time segments, and so the PCVM often produces a discontinuous representation of a secular trend. Indeed, ITRF sometimes invokes the PCVM using time segments that overlap, or which are separated by gaps in which the station position is undefined. See Sect. 6.2 for further discussion of the PCVM.

Yang et al. (2013) model both secular trends and annual cycles recorded in Greenland using a more complicated approach than any discussed so far. Their approach uses a

spline model to address how annual 'cycles' might vary from year to year. This is perhaps the most specialized trajectory model devised to date for use in tectonically stable areas: its structure reflects the authors' understanding of climate cycles and climate change in a polar ice sheet. It is probably too specialized for wholesale adoption by geodesists and geophysicists engaged in global GPS analysis. We shall make no further reference to this interesting class of trajectory model.
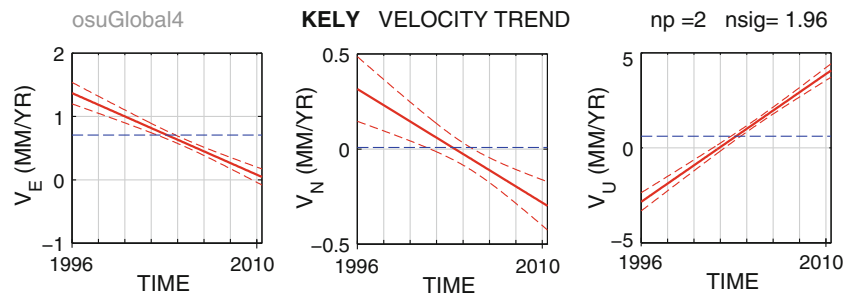
Fig. 5 The *E*, *N* and *U* components of the velocity trend at KELY (*solid red line*) and their nominal 95 % confidence intervals (*dashed red lines*). Since this station was modeled using a quadratic trend, this is a 'constant acceleration trajectory.' The *dashed blue line* represents the lifetime average station velocity. Strongly accelerating patterns of displacement in Greenland cannot be explained in terms of postglacial rebound (PGR). They manifest instantaneous elastic rebound driven by accelerating rates of ice loss (Bevis et al. 2012b)

4 Accommodating postseismic deformation

A significant fraction of CGPS stations worldwide are located in seismogenic areas, and therefore, over the last 15–20 years, many CGPS stations have recorded coseismic jumps followed by pronounced postseismic transients. This produces a significant challenge for groups engaged in designing or realizing global RFs valid from the early 1990s to the present day, because there are numerous regions of the world, e.g. western South America, Indonesia, Turkey, California, Japan, Taiwan, Alaska and the Southwest Pacific, in which many or even all long-lived CGPS stations have recorded one or more episodes of vigorous postseismic deformation. Unless the trajectories of these stations can be approximated using a suitably modified trajectory model, then these stations cannot easily be used to define or realize a RF, and the remaining global reference stations will have a significantly poorer spatial distribution. Polynomial trend models (Eq. 7) are not well suited to this task, because postseismic transients often begin 'in the middle' of a time series, with the fastest accelerations of all just after the earthquake, and no transient acceleration at all immediately before the earthquake.

Postseismic deformation is widely thought to be driven by some combination of

(1) *poroelastic rebound*, which is deformation caused by pore fluids flowing in response to the stress perturbations produced by the earthquake (Peltzer et al. 1998; Wang 2000; Masterlark and Wang 2002; Jónsson et al. 2003; Fialko 2004; Wang et al. 2004),

(2) *afterslip* on the fault or plate boundary that generated the earthquake (Smith and Wyss 1968; Bucknam et al. 1978; Marone et al. 1991; Heki et al. 1997; Marone 1998; Perfettini and Avouac 2007; Perfettini et al. 2010, Lin et al. 2013), and

(3) bulk *viscoelastic relaxation* of material surrounding (in map view) and beneath the fault (e.g. Thatcher and Rundle 1979; Wahr and Wyss 1980; Deng et al. 1998; Pollitz et al. 2000; Freed 2007; Freed et al. 2007; Wang et al. 2012).

While poroelastic rebound may not always be a strong contributor to postseismic deformation, when it is clearly present, it is distinct from the other two mechanisms in that it is relatively short-lived (often being weak or even unresolvable after a few weeks), and it is best expressed in the near-field of the main event. Afterslip and viscoelastic relaxation persist for many years, and they may affect a much wider region. Many geophysical theorists believe that viscoelastic relaxation persists much longer than does afterslip (e.g. Wang et al. 2012), and that present day postseismic deformation near the rupture areas of the giant 1960 Chile and 1964 Alaska megathrust earthquakes can be explained purely in terms of viscoelastic relaxation (Wang et al. 2007).

Of the three mechanisms identified above, only the second mechanism is easy to model using a simple mathematical formula. Afterslip is known to generate transient displacements that are well approximated over several years by the logarithmic transient formula $A \log(1 + \Delta t/T)$ suggested by rate-and-state friction models (Marone et al. 1991; Perfettini and Avouac 2007; Perfettini et al. 2010). In contrast, poroelastic effects and bulk viscoelastic relaxation are typically modeled using complicated numerical methods (e.g. finite element models), and there are no closed-form, analytical expressions for the transients generated at each geodetic station.

Given that multiple, quite distinct mechanisms contribute to postseismic deformation, it does not seem likely, at first glance, that we could develop a simple means of augmenting the SLTM so as to account for postseismic transient motion. But we will show that it is possible to achieve this goal, provided that (1) we focus directly on the trajectories of the geodetic stations affected by postseismic deformation, not on the causative phenomena taking place underground, and (2) we accept that this simple model often fails to account for the especially rapid deformation that takes place in the first few weeks or months following the main event. We achieve this apparently unlikely outcome by assuming that the logarithmic transient formula can model postseismic transient displacement at any GPS station, even though this for-

mula is theoretically associated with the afterslip mechanism alone. The frequent failure of this formula to account for the first few weeks or months of postseismic deformation might arise because actual short-term deformation is produced in large part by a mechanism other than afterslip, most likely poroelastic deformation. (We should keep in mind that the fluctuations of well water levels sometimes observed immediately following nearby earthquakes, provides direct evidence of postseismic pore-pressure transients, e.g. Jónsson et al. 2003; Wang et al. 2004.) However, the inability of the logarithmic transient formula to fit the *entire* time series could also arise because GPS station displacement caused by afterslip is sensitive to spatial variation of the frictional properties of the fault or plate boundary engaged in afterslip. While the logarithmic transient formula might be a valid expression for afterslip in every small patch of the fault, the value of parameter T might vary from patch to another, and since all such patches contribute to displacement at a given GPS station, it is not really valid to assume that station displacement follows the logarithmic form with a single value of T. That is, the logarithmic formula may be overly simplistic when applied to a single GPS station as opposed to a small fault patch undergoing afterslip.

However, as we shall see, if we exclude (or downweight) the several week to several month period immediately following the earthquake, the subsequent success of the logarithmic transient formula in accounting for GPS station displacement is so striking that it suggests to us that afterslip nearly always dominates viscoelastic relaxation even a decade or more after the primary seismic event, and not just for a several years, as is already widely acknowledged—see Lin et al. (2013) for a brief review of the relevant literature.

Indeed, the claim we shall make below is rather more remarkable than what we have suggested so far. The logarithmic transient formula has two parameters, A and T, the second of which appears inside the logarithm, and thus implies a nonlinear inversion will be required to estimate the parameters or coefficients for each station. But we will demonstrate that a useful first approximation is obtained simply by setting $T = 1$ year, and estimating A alone, in which case the inversion is purely linear. We do recommend tuning the T parameter by station, usually one station at a time, but only as a non-urgent refinement undertaken as time permits, and not in the context of global time series analysis and RF realization. The insensitivity to T is such that, having refined the estimate of T at some later date, using a station-specific nonlinear estimation procedure, it is not necessary to repeat this refinement process very often, at least once the postseismic transient has been measured for ∼2 years or longer.

We are not suggesting that geophysicists engaged in the study of earthquake physics, fault mechanics, and crustal and mantle rheology, could consider such a simplistic approach sufficient or adequate. Rather, we are suggesting that a very

simple augmentation of the SLTM, based on the logarithmic transient formula, provides an adequate means for geodesists to predict the positions of almost *any* geodetic station subject to postseismic transient deformation, for all but a few months at worst, with centimeter or sub-centimeter accuracy. This is much preferable to eliminating all such stations from the set of reference stations used to realize a RF (at least for all epochs after the first earthquake that produces a postseismic transient at that station). Modeling such transients allows many more long-lived CGPS stations to be used in RF realization (specifically in the time series alignment step discussed in Sect. 6.3), including stations in regions that would otherwise be almost devoid of representation. This promotes a 'stiffer' RF and less noisy estimates of station motion (including postseismic transients) within that frame (see Sect. 6.4). Indeed, such an approach ultimately benefits the geophysicist who will analyze station displacements in a much more sophisticated way, because it will largely suppress the RF jitter and RF drift that is often precipitated by very large earthquakes. Even so, our immediate interest here is in predicting the station trajectory, not in understanding its physical causes.

5 Adding logarithmic transients to the standard linear trajectory model

In this section, we will discuss how we augment the SLTM (Eq. 8) by adding one or more logarithmic transients, as needed, in order to accommodate postseismic deformation, at least in an approximate sense.

5.1 Some preliminary considerations

A logarithmic transient displacement has the form

$$d = A \log(1 + \Delta t / T) \tag{9}$$

where Δt is time since the earthquake occurred. Since this formula applies only after the earthquake has occurred, we are restricted to the domain $\Delta t \geq 0$. The scalar d might refer to any one of the geocentric Cartesian coordinates (X, Y, or Z) or topocentric Cartesian coordinates (E, N, and U) used to describe a geodetic time series.

In Appendix 1, we demonstrate that the SLTM augmented with the logarithmic transient formula (Eq. 9) is surprisingly insensitive to the value of the nonlinear parameter T in the sense that if one assigns a moderately erroneous value to T, and allows A to adjust accordingly, one can retain a very good fit to the great majority of the data. There is a subtlety here in that this insensitivity is not inherent to the logarithmic transient formula itself, but arises when it is used in conjunction with the SLTM (Eq. 8). That is, the coefficients of the SLTM,

as well as the linear coefficient *A*, *all* adjust so as to permit a good fit to the data even when *T* is assigned an inaccurate value (though *A* usually adjusts the most). If one's goal is fitting or characterizing an observed station trajectory, rather than isolating the individual physical contributions to station displacement, this trade-off between the various coefficients is highly beneficial. It also provides the geodesist with a useful option, further discussed below, i.e. to simply assign a 'reasonable' or default value for *T*, and fit an augmented trajectory model using a completely *linear* LS approach. The value of *T* can be refined later, on a station by station basis, using a nonlinear estimation approach. The practical advantage of this option is that during the global process of time series analysis and reference frame realization, one is not engaged in estimating many dozens or even hundreds of non-linear parameters simultaneously. This task can be deferred, and later undertaken in a non-global context. It is often preferable to estimate say 100 non-linear parameters one or two at a time, in separate inversions, than to estimate all of these parameters simultaneously (along with thousands of linear parameters) in one large inversion.

5.2 The extended trajectory model: nomenclature and patterns of usage

We define the *extended trajectory model* (ETM) as the combination of a SLTM and one or more logarithmic transients. Explicitly, the ETM is

$$\mathbf{x}(t) = \sum_{i=1}^{n_P+1} \mathbf{p}_i (t - t_R)^{i-1} + \sum_{j=1}^{n_J} \mathbf{b}_j H(t - t_j)$$
$$+ \sum_{k=1}^{n_F} \mathbf{s}_k \sin(\omega_k t) + \mathbf{c}_k \cos(\omega_k t)$$
$$+ \sum_{i=1}^{n_T} \mathbf{a}_i \log(1 + \Delta t_i / T_i) \tag{10}$$

where n_T is the number of logarithmic transients. For each transient caused by an earthquake at time t_{EQ}, we define $\Delta t = 0$ for $t < t_{EQ}$, and otherwise $\Delta t = t - t_{EQ}$. This is one way to ensure that the transient is restricted to the time period after the primary or causative seismic event.

When we engage in station trajectory analysis and reference frame realization using a global time series involving >1,200 GPS stations and a combined or total timespan of ∼18 years (e.g. Bevis et al. 2012a), we are typically modeling ∼100 stations affected by postseismic transients. We routinely perform a provisional analysis for any newly observed transient in which T_i is treated as a known constant (metadata for a particular station and earthquake), so that the amplitude coefficients, $\mathbf{a}_i$, can be estimated using a linear LS approach. In this case we call the model (Eq. 10) the *extended linear*

trajectory model (*ELTM*). We rarely fail to obtain a useful first approximation by assigning *T* a default value of 1 year.

We have modified our global 'jump table', which was previously used to indicate the timing of Heaviside jumps at each station, and nothing else, so that following the station code and the jump time there is now a third, numerical entry. If this entry is set to zero, it indicates that no logarithmic transient is invoked. (This is always the case for artificial jumps associated with equipment changes.) If this entry is set to any positive number, it indicates that a logarithmic transient should be 'attached to' the Heaviside jump, and that the value of this entry should be assigned to the parameter *T*. By convention, $T = 1$ is the (default) value entered when a transient has been recognized for the first time, and its optimal value is unknown. If any value other than 1 is found in the jump table, it indicates that this value that has already been 'tuned' by a non-linear analysis of this station's time series, usually in the calm aftermath of a large global analysis or 'rerun'. As noted above, in our experience, this tuning does not need to be revisited very often, unless one is truly focused on *T* as a geophysical quantity, rather than just as a means to predict station position. (The exception is for those stations for which little time has passed since an earthquake initiated a transient. Until the transient is characterized by about 2 years of postseismic observations, the best fit value of *T* may change significantly with each new increment of postseismic data.)

Non-linear estimation of *T* implemented on a single station basis is particularly simple if (as usual) the *ETM* invokes only a single transient. In this case *T* can be estimated using an iterative one-dimensional grid search. It is also possible to estimate an optimal *shared* value of *T* for a group of CGPS stations that all record the postseismic transient deformation produced by a single earthquake. This level of detail is of more interest to a geophysicist studying earthquake source physics than to a geodesist whose immediate interest is to be able to predict the coordinates of a given station on a given day.

5.3 Case studies

We now present some examples of postseismic transients and the extent to which we can model them using the ETM or the ELTM. We shall include the latter approach, because we wish to demonstrate that remarkably good results can be obtained even without estimating the parameter *T* using a nonlinear inversion.

We first consider the case of the M_w 7.6 'Centennial' earthquake that occurred near the South Orkney islands on 4 August 2003 (Plasencia 2007; Smalley et al. 2007). The epicenter of this shallow, transtensional event was located about 76 km from the GPS station BORC at Base Orcadas. So far, the postseismic transient at BORC has been observed

Fig. 6 The displacement time series at station BORC, in the South Orkney islands, fit with an ELTM with T at its default value of 1 year. Note that the first few months after the earthquake, the E time series is systematically down-weighted by our robust inversion algorithm. The down-weighted points are shown in *orange*. The down-weighting factor, s, is largest in the first few days of the transient, but progressively weakens as the transient is better modeled by the ELTM. Eventually, down-weighting ceases. In contrast, most of the orange 'outliers' visible in the N and U time series have been only gently down-weighted. They represent noisy measurements, not model problems
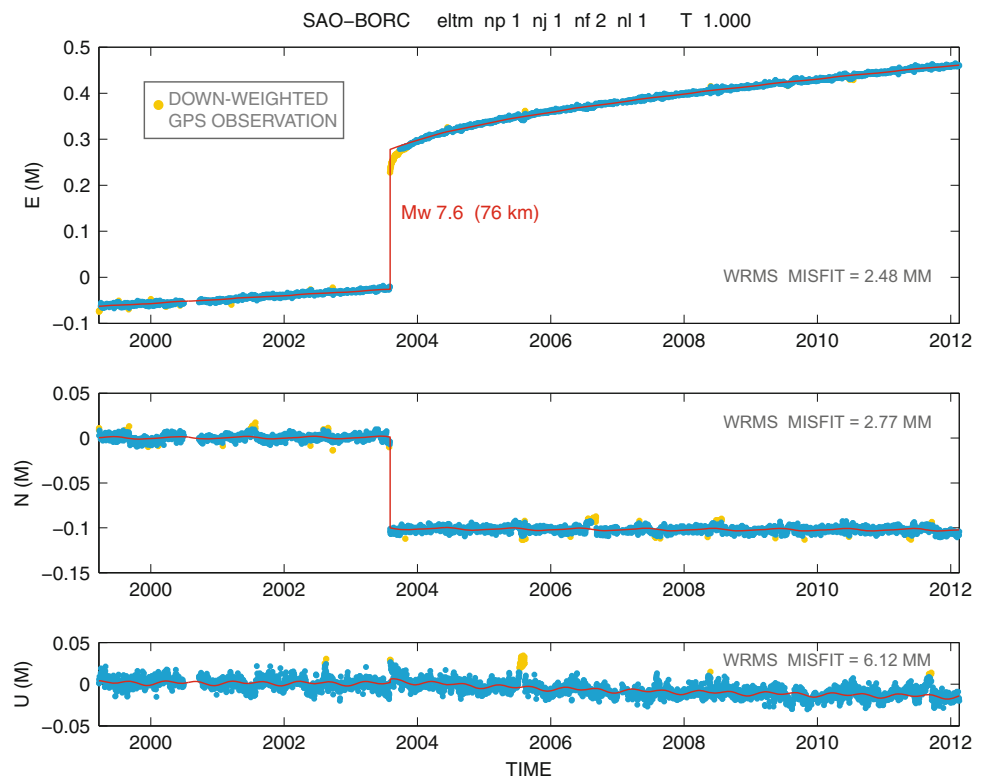
for a total of 8.5 years (Fig. 6). It is most strongly developed in the E component of displacement. The best fit ELTM, with T defaulting to 1 year, is shown in Figure 6. A second, nonlinear analysis finds $T = 0.283$ years, leading to a modestly improved fit (Fig. 7). In both cases our iterative reweighting scheme, which is designed to reduce the impact of unusually noisy measurements (i.e. outliers), systematically downweights the observations during the early portion of the transient (Fig. 8). Note that model problems rather than data problems are responsible for this misfit. However, even the ELTM with T defaulting to 1 year, produces a very good fit to the observations for the last ∼8 years of the time series.

Our second example is the postseismic transient observed by GPS station AREQ in Arequipa, Peru (Fig. 9). This transient was precipitated by an M_w 8.4 megathrust earthquake which occurred on June 23, 2001 about 230 km from the GPS station. We present the best fit ELTM obtained when T was set to its default value of 1 year. We accounted for a M_w 7.6 aftershock which occurred on July 7, 2001, by adding a second Heaviside jump, but we did not invoke a second logarithmic transient. The observed time series are well fit by the resulting trajectory curve, except in the first few weeks following the main event. In this particular case, subsequent tuning of the value of T produced only a very minor improvement to the fit.

Our last case study, which focuses on the displacements recorded by GPS station SAMP in Sumatra, is unusual in that this station has recorded three distinct postseismic transients to date. The first transient followed the M_w 9.1 Sumatra–

Andaman earthquake of 2004/12/26, which produced the highly destructive Indian Ocean Tsunami that killed over 230,000 people in 14 countries. The epicenter of this megathrust event was located about 305 km from the GPS station. This earthquake may have triggered the M_w 8.6 megathrust event which occurred on 2005/03/28 off the Sumatran island of Nias, about 246 km from SAMP. Finally, this GPS station recorded the M_w 7.8 event off the Banyak islands near Sumatra on 2010/04/06. Our trajectory model for SAMP invokes three Heaviside jumps and three logarithmic transients. The fit obtained in Fig. 10 was obtained with refined estimates for T for the first two events (0.45 and 1.45 years, respectively), but because we did not have much data available after the third event, we allowed T to retain its default value of 1 year. Note that the RMS misfit of the data and model is higher but not greatly higher than what we attain at many tropical stations—and most of those are not undergoing postseismic deformation. The largest misfit at SAMP occurs in the Y component, which almost corresponds to the local up (U) direction.

6 Trajectory models and reference frame design and realization

In this section we discuss the ways in which utilizing general trajectory models impacts the design, definition or realization of terrestrial RFs. The *designing and defining* of RFs is a specialized activity performed by a very small segment of

Fig. 7 The same displacement time series for station BORC as shown in Fig. 6, but now fit by an ETM with an optimized value for *T* (0.283 years). Far fewer of the *E* measurements immediately following the earthquake are down-weighted in this analysis, as seen in Fig. 9
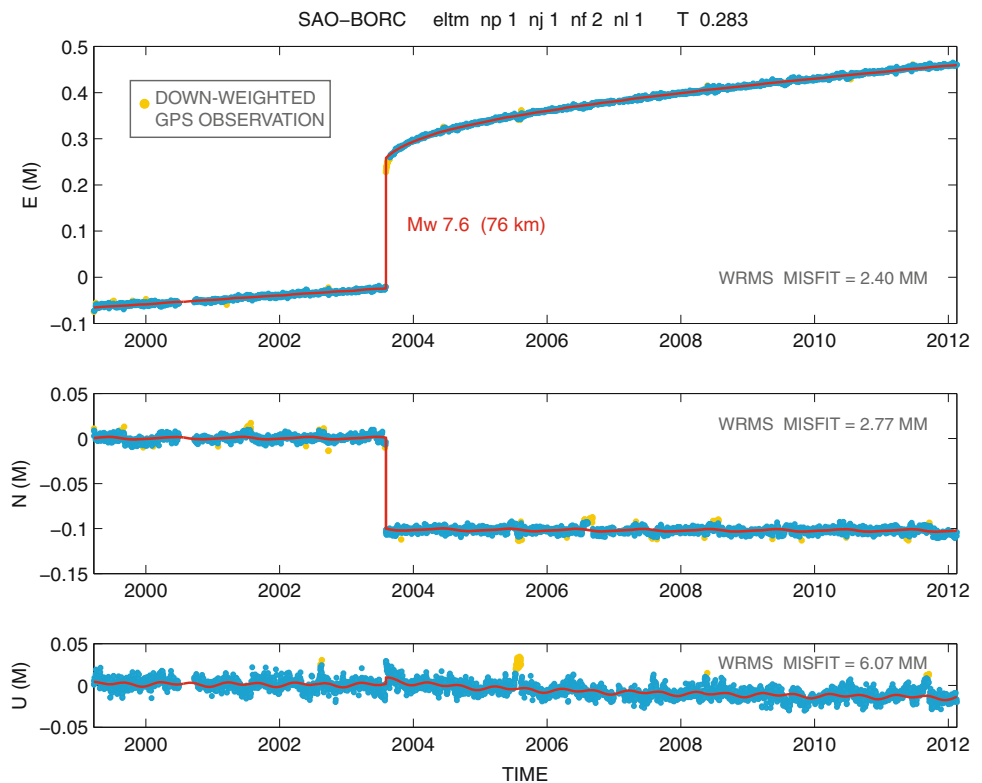

Fig. 8 A detailed comparison of the fits already seen in Figs 7 and 8, which zooms in on the time period immediately following the earthquake. **a** When *T* takes it default value of 1 year, the ELTM significantly and systematically misfits the observed transient for a little over 90 days. **b** When *T* is optimized by non-linear estimation, the period of poor model fit is reduced to ∼3 weeks
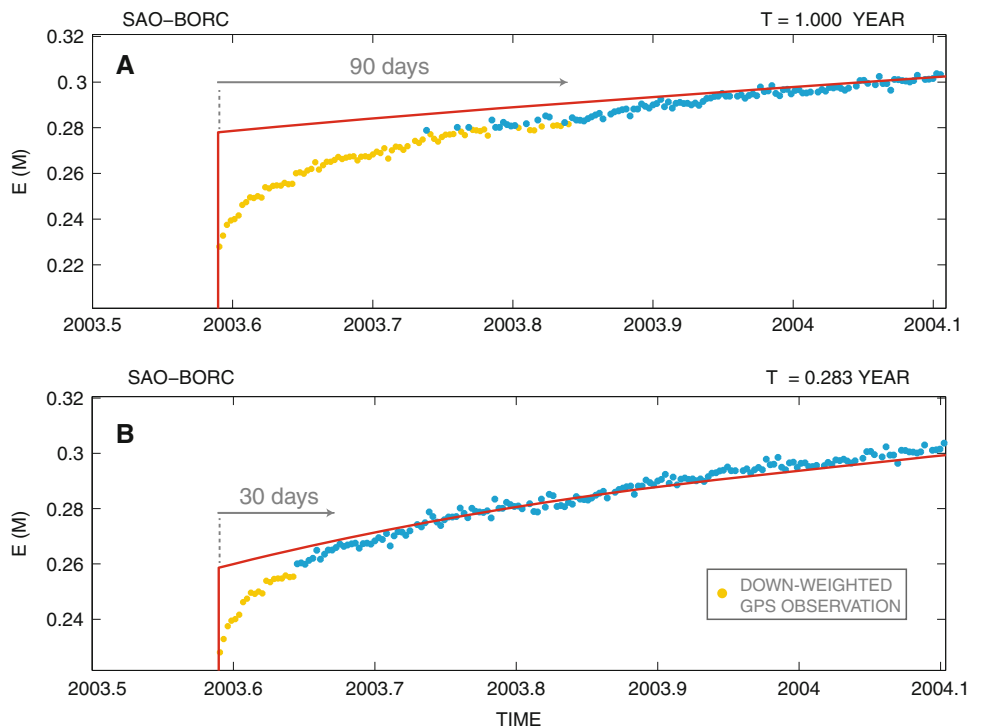
the geodetic community. But RF *realization* is a routine step for anyone engaged in the analysis of GPS networks, since station positions and velocities are almost always desired and stated with respect to a given or target RF. We discuss RFs in the context of 'pure GPS' networks because (i) we, like the majority of geophysicists, geologists, glaciologists, meteorologists, etc., engaged in applied geodesy, and even many 'pure' geodesists, analyze geodetic networks consisting only of GPS stations, and (ii) this allows us to simplify our discussion, because we have already demonstrated that

Fig. 9 The displacement time series at station AREQ in Arequipa, Peru, fit with an ELTM with T at its default value of 1 year. This accounts for the logarithmic transient produced by the M_w 8.4 earthquake of 2001/6/23. The M_w 7.6 aftershock of 2001/7/7 (which was closer to the GPS station) is accommodated with a pure Heaviside jump, because it too closely followed the main event to warrant a separate transient. The *inset figures* show the period before and after the second earthquake in greater detail
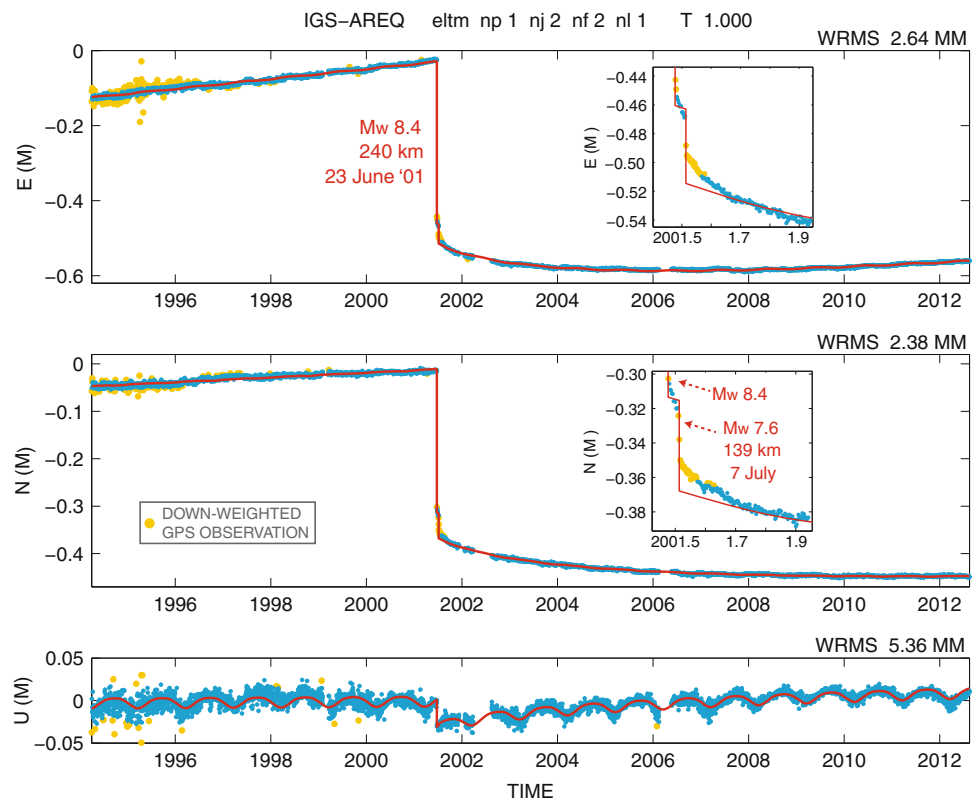

Fig. 10 The station SAMP in Sumatra has recorded jumps and seismic transients associated with two great and one major megathrust earthquakes. The time series is modeled using three logarithmic transients, the last of which assumes the default value (1 year) for T, since there is not enough data following this event to warrant nonlinear estimation of this parameter
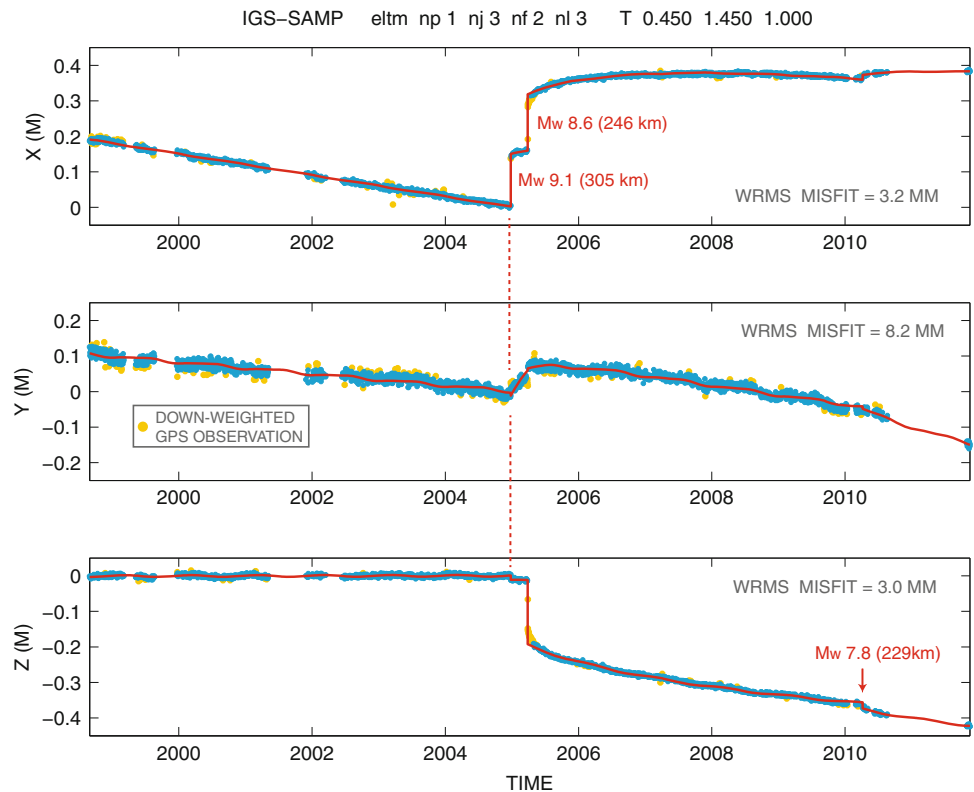
there is no evidence for *significant* temporal changes in the scale associated with GPS measurements, provided that the entire time series is analyzed using modern satellite antenna models, and geodetic software and processing protocols are not changed during the course of the analysis (Bevis et al. 2012a). This allows us to invoke a single and time-invariant metric or scale. As a result, the coordinate transformations associated with a change of RF can be completely described in terms of the three translations and the three rotations that relate one RF (or axis system) to another one. These are the parameters of the 6-parameter Helmert transformation H6 (see Appendix 2 for additional discussion).

6.1 Basic concepts

We often refer to a GPS network as a polyhedron, meaning that the GPS stations of the network constitute the vertices of this polyhedron. Typically we quantify the geometry of a polyhedron by assigning coordinates (X, Y, Z) to each of its vertices. Those coordinates refer to a specific RF that we normally think of as a particular, right-handed Cartesian axis system [X, Y, Z]. Most scientists and engineers were trained, when young, to think of a Cartesian axis system as an entity that allows us to assign coordinates to points. But operationally, the reverse is true. In space geodesy, it is the assigning of coordinates to points that, in effect, defines the axes. A terrestrial RF such as ITRF2008 (Altamimi et al. 2011) is a *conventional model* (Appendix 2) that can be used to predict the coordinates of a certain set F of GPS stations at any given epoch or time t, and thereby invoke a specific axis system. We view the term 'reference frame' as having two distinct meanings: (i) a particular and named axis system, and (ii) a model that predicts the coordinates for a global set of GPS stations, allowing us to use these stations to invoke or realize this axis system. When we wish to refer to a RF in this second, operational sense, we can call it a *RF model*. A RF model is a *network trajectory model* consisting of a set of station trajectory models (i.e. including the values of the parameters of these models)—one for each reference station incorporated into the definition of the RF. When we use a RF model to predict the coordinates of a network (or subnetwork) at a single epoch t, we call this prediction the *model polyhedron* at time t. As we shall explain below, model polyhedra are often used as *alignment targets* in geodetic time series analysis.

Suppose we wish to position a general set or network N of GPS stations, some subset C of which are also members of the frame-defining station set F. (F is the set of *reference stations* used to define a given RF). The stations in C are common to N and F. Because GPS geodesy is a *differential* positioning technique, it determines the *inner* geometry of a polyhedron (i.e. its size and shape) much more accurately that it can directly determine its location or orientation. But we

can use the RF model to predict the coordinates of the common stations (in subset C), and arrange for the GPS solution (polyhedron) to be re-positioned and re-oriented so as to conform with the predicted coordinates for the common stations, or very nearly so. Since the shape and size of the polyhedron is very strongly constrained by the GPS measurements, the entire polyhedron is now positioned and oriented in the target frame, and the coordinates of all stations (including those in N but not in F) are determined. In other words, we transform the coordinates of polyhedron N so as to express it in the target RF by aligning the observed (GPS) sub-polyhedron C with the model sub-polyhedron C.

We invoked above the concept of 'inner' geometry—i.e. the geometry that can be defined purely in terms of the lengths between points, such as the vertices of a polyhedron. The inner geometry of a polyhedron makes no reference to axis systems, and is entirely independent of them. Inner geometry depends only on the scale or metric associated with measurements in physical space. A physicist might say that inner geometry is *invariant* under rotation and/or translation of the RF. Those readers not familiar with the concept of inner geometry (or our terminology) are referred to Appendix 2 for an extended discussion.

We use the 6-parameter Helmert transformation, H6, to shift and reorient a polyhedron in space so that some subset of its vertices take on the coordinates predicted by a given RF (or very nearly do). We can certainly think of this transformation as a general, rigid-body motion of the polyhedron relative to the axes. But in geodesy we usually adopt the opposite but equivalent point of view that it is the axes that are moving rather than the polyhedron. Imagine that a preliminary solution for the polyhedron has the correct inner geometry but is expressed in an arbitrary RF rather than the target RF. All we must do is estimate the 6-parameters of an H6 transformation that changes the coordinates for the stations in C so that they correspond as closely as possible to the coordinates predicted by the target RF. That is we transform the GPS polyhedron by aligning some subset of its stations with a model polyhedron generated using the network trajectory model that defines the RF in the operational sense. We then use those H6 parameter values to transform the coordinates of the entire GPS polyhedron, which is then expressed in or referred to the target frame.

A RF model is most useful if the inner geometry of its predictions are consistent (or very nearly consistent) with the actual inner geometry of the corresponding GPS stations over an extended period of time. This is a key means of characterizing the 'geometrical consistency' of a RF. This is because the inner geometry of a polyhedron is in some deep sense its most fundamental and irreducible geometry. Two very different sets of station coordinates might refer to exactly the same polyhedron because they differ only in the choice of the RFs used to express the coordinates. But two

sets of station coordinates that imply quite different inner geometries do not and cannot refer to the same polyhedron (Appendix 2).

In practice, when an H6 transformation is used to align or re-align a GPS polyhedron as closely as possible with a model polyhedron, this alignment will almost never be perfect. A measure of the residual misfit that characterizes the lengths of the deviation or residual vectors between the two sets of vertices is said to characterize their *inner coordinate misfit* or *inner coordinate scatter* (Appendix 2). This is a RF-independent or 'frameless' measure of the geometrical inconsistency of the polyhedra. If a GPS network time series composed of one estimated polyhedron per day is aligned with the model polyhedra predicted by two distinct RFs and the inner coordinate scatter (for the entire time series) associated with RF#1 is smaller than that associated with RF#2, then RF#1 is said to be more consistent with the GPS measurements, and it would normally be preferred over RF#2.

Standard RFs such as ITRF2008 normally use equation (4)—i.e. the CVM augmented, as necessary, by jumps—to predict the position coordinates of the great majority of GPS stations incorporated into their definition. For the rest of this paper we use the abbreviation CVM whether the model invokes no, one or more jumps. This has long been the favored station trajectory model for ITRF. For GPS stations at which postseismic transient deformation produces large changes in velocity over time, ITRF2008 invokes the PCVM instead. The PCVM is further discussed and illustrated in Sect. 6.2. The key point we wish to make here is that to be operationally useful, a RF must be associated with a network trajectory model which can predict the model polyhedra used as alignment targets, and it is highly desirable that these predictions are geometrically consistent with the GPS measurements (i.e. have very similar inner geometries), but there is no reason why these station trajectory models should be limited to particular classes of model such as the CVM or the PCVM. We prefer to employ more general classes of trajectory models—i.e. the SLTM and the ETM or ELTM—because they produce more consistent predictions of station geometry, as illustrated in Sect. 6.2, and this leads to a more consistently realized RF, as discussed in Sect. 6.3.

In our view, the RFs we produce in this way are modest refinements of an existing parent frame. In recent years our parent frame is ITRF2008 (from which we inherit our GPS scale), and we designate our variant as OSU08. Loosely speaking, OSU08 has been aligned as closely as possible with ITRF2008, both in position space and in rate or velocity space. Indeed, for many practical purposes, such as mapping and engineering surveys, the differences between the two frames are insignificant. We prefer OSU08 over ITRF2008 because the time series that we express in OSU08 have lower levels of coordinate scatter.
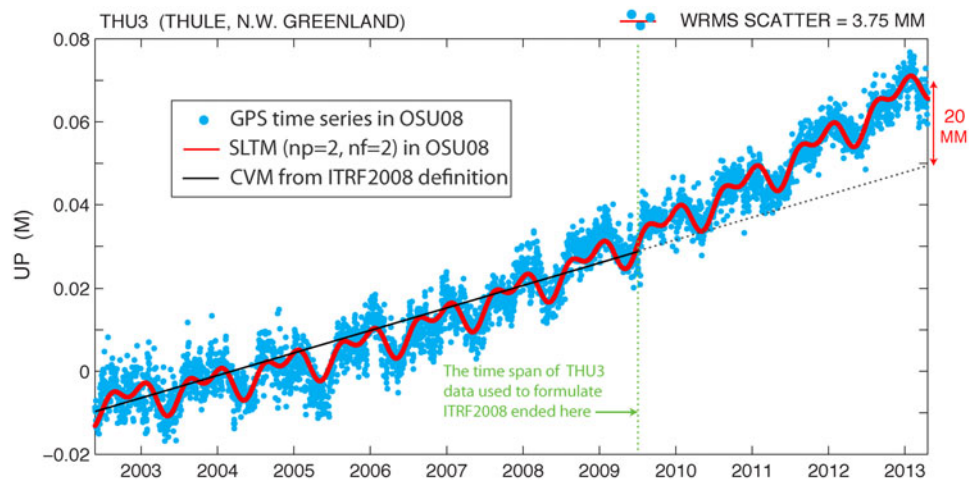
6.2 Contrasting general station trajectory models with those employed by ITRF

We begin by noting that even if we position a network N consisting only of stations which are well characterized by the CVM, and even if we invoked only the CVM in our own geodetic analysis, should we use a longer time series than that available to Altamimi et al. (2011) to formulate and define ITRF2008, then our estimates of station velocity and reference position would almost certainly differ from the canonical predictions of ITRF2008 RF model, and therefore our realization of ITRF2008 would not be completely consistent with ITRF2008 as it was originally defined. This is sometimes referred to as 'reference frame realization error' (Dietrich et al. 2001)—a very useful term, though perhaps the word 'error' is rather harsh if the dataset being used to realize ITRF2008 is several years longer than that used to design and define it, and the extended GPS times series is now less consistent with the predictions of ITRF2008.

A similar situation arises if the secular trend of many stations in N was truly constant-velocity, but we inferred these velocities using Eq. (5) rather than the Eq. (4). We would do this because, for stations with fairly short time series in particular, not estimating annual oscillations (as part of the trajectory model) can cause them to alias to 'leak into' geodetic estimates of station velocity (Dong et al. 2002; Collilieux et al. 2012). Adding cycles to all our station trajectory models improves the geometrical consistency of our velocity estimates, but may slightly degrade our post-alignment fits with the predictions of ITRF.

A more dramatic version of this problem arises when we wish to process a station manifesting large accelerations (i.e. very significant changes in velocity over time). By mid-2013, we could clearly resolve accelerations at about half of the GPS stations in Greenland. In Fig. 11 we show a recent solution for the upwards (U) component of vertical displacement at station THU3 in N.W. Greenland, and the best fitting SLTM (with $n_P = n_F = 2$), both expressed in OSU08. We also show the CVM for THU3 predicted by ITRF2008. The CVM fits our solutions reasonably well within the time-span of the THU3 dataset available to Altamimi et al. (2011) when they designed and defined ITRF2008, although the CVM clearly does not fit the data as well as the SLTM (or even just the quadratic trend component of this SLTM). Beyond the time span of the data available to Altamimi et al. (2011), i.e. after 2009.51, the CVM trajectory (indicated by the dotted line) systematically diverges from the actual trajectory of this station. The key point is that the actual trend of $U(t)$ is curved, whereas the CVM invoked by ITRF2008 is not. In other words, the ITRF2008 prediction for THU3 becomes increasingly inconsistent with reality as time passes. We have found it necessary to invoke polynomial trends to adequately model stations subject to volcanic deformation, and even ground

water withdrawal as well as elastic displacements driven by accelerating changes of ice mass in nearby glaciers or ice sheets.

For some stations, particularly those affected by postseismic transients, Altamimi et al. (2011) accommodate pronounced changes in station velocity by invoking a PCVM rather than their usual CVM. For example, their SINEX file entries for station AREQ in Arequipa, which we discussed in Sect. 5.3, provides predictions for 6 time segments by invoking a CVM model for each segment (Table 1). In Fig. 12a, we focus on what this RF model implies for the east (E) component of displacement at AREQ, and compare it with our GPS measurements referred to OSU08. Since our model is the ELTM (with a default value for T), the non-cyclical part of the E trajectory model is determined by just 3 coefficients. In contrast, the PCVM requires $2 \times 6 = 12$ coefficients. A modest difference between the ITRF velocity prior to the M_w 8.4 earthquake, and the velocity implied by our GPS solutions and ELTM in this time period (Fig 12b), probably manifests the non-identity of ITRF2008 and OSU08. At the other end of the time series, beyond the end date of the AREQ data originally available to Altamimi et al. (2011), the ITRF2008 prediction progressively diverges from our GPS solutions and our SLTM model, and these discrepancies are far larger than those that occur prior to the M_w 8.4 event. This is because even after 2010.0, when the postseismic transient has been decaying for more than 8 years, velocity is still changing too rapidly to allow the prediction produced using the sixth CVM to be projected for much more than about 1 year. The ITRF2008 PCVM does out-perform our ELTM in the several month period immediately following the two earthquakes—indeed, it concentrates about half of its degrees of freedom in this time period for that very purpose. But apart from that limited period, the ELTM model does rather better, despite having far fewer degrees of freedom, and after about 2011.5 it does much better than the ITRF2008 prediction.

We are well aware that it is in some sense unfair to compare the predictions of ITRF2008 and those we achieved using our

favored trajectory models, because our data set was larger than that available to Altamimi et al. (2011), and our fit used all of our data, whereas the predictions of ITRF had access to less data per station. But almost everyone analyzing a data set using a RF such as ITRF2008 has longer if not much longer time series for many of the GPS stations that they have in common. We must express our more up-to-date time series in some frame, and it is reasonable to use or produce a variant of ITRF2008 rather than ITRF2008 itself if this leads to a more consistently-realized RF and less 'noisy' time series.

In the next section we point out that using more general trajectory models than those used by the designers of ITRF is not just about improving the way in which we characterize the stations our network has in common with those used to define ITRF, but rather improving the way in which we characterize the displacements occurring at *all* GPS stations in our networks.

6.3 The advantage of imposing the RF after fitting the trajectory models

Geodesists and geophysicists can realize a target RF in a variety of ways. It is useful to distinguish between two general approaches that we refer to as *prior* versus *posterior* imposition of the RF.

6.3.1 Prior imposition of the RF

Suppose, as before, that we are analyzing a set N of GPS stations, some subset C of which were also used to define a target RF such as ITRF2008. Perhaps the simplest of all approaches to RF realization, which we call 'prior imposition' of the frame, is pursued during the daily geodetic analysis of the network N. GPS data processing software requires the user to specify a prior position estimate (X, Y, Z) for each station in the polyhedron (for each and every epoch or day), and associated prior position constraints, i.e. the standard errors $(\sigma_X, \sigma_Y, \sigma_Z)$ indicating the extent to which the

Table 1 Model parameters for CGPS station AREQ (in Arequipa, Peru) extracted from the SINEX file describing ITRF2008 (Altamimi et al. 2011)

ITRF-2008 piecewise constant velocity model for AREQIPA station AREQ

SEG#	Starts	Ends	Mean epoch		
AREQ 1	96:363:00000	01:176:00000	99:087:00000		
AREQ 2	01:188:00000	01:225:00000	01:206:43200		
AREQ 3	01:223:00000	01:260:00000	01:241:43200		
AREQ 4	01:258:00000	02:231:00000	02:062:00000		
AREQ 5	02:264:00000	07:337:00000	05:117:43200		
AREQ 6	07:342:00000	09:187:00000	08:265:00000		
Th	**TREF**	**XREF (M)**	**Sigma (M)**	**XVEL (M/Y)**	**Sigma (M/Y)**
AREQ 1	05:001:00000	1942826.8243	0.72953E−03	0.012736	0.80417E−04
AREQ 2	05:001:00000	1942825.4645	0.84181E−01	−0.238831	0.24440E−01
AREQ 3	05:001:00000	1942825.8945	0.53804E−01	−0.113263	0.16017E−01
AREQ 4	05:001:00000	1942826.1127	0.36669E−02	−0.046701	0.11804E−02
AREQ 5	05:001:00000	1942826.2123	0.63327E−03	−0.003089	0.64376E−04
AREQ 6	05:001:00000	1942826.1976	0.98817E−03	0.002727	0.20859E−03
Th	**TREF**	**YREF (M)**	**Sigma (M)**	**YVEL (M/Y)**	**Sigma (M/Y)**
AREQ 1	05:001:00000	−5804070.2300	0.11404E−02	0.001613	0.16782E−03
AREQ 2	05:001:00000	−5804070.3802	0.18773E+00	−0.031151	0.54510E−01
AREQ 3	05:001:00000	−5804070.4999	0.11958E+00	−0.065838	0.35602E−01
AREQ 4	05:001:00000	−5804070.3063	0.79765E−02	−0.007141	0.25927E−02
AREQ 5	05:001:00000	−5804070.3120	0.81051E−03	−0.006449	0.13747E−03
AREQ 6	05:001:00000	−5804070.3145	0.18335E−02	−0.005281	0.44987E−03
Th	**TREF**	**ZREF (M)**	**Sigma (M)**	**ZVEL (M/Y)**	**Sigma (M/Y)**
AREQ 1	05:001:00000	−1796893.8460	0.75175E−03	0.013794	0.74854E−04
AREQ 2	05:001:00000	−1796894.8069	0.74558E−01	−0.171979	0.21648E−01
AREQ 3	05:001:00000	−1796894.5339	0.47409E−01	−0.091980	0.14114E−01
AREQ 4	05:001:00000	−1796894.3207	0.31852E−02	−0.026622	0.10178E−02
AREQ 5	05:001:00000	−1796894.2552	0.67426E−03	0.002723	0.62008E−04
AREQ 6	05:001:00000	−1796894.2632	0.93833E−03	0.007518	0.18162E−03

The first block describes the six time intervals invoked by the PCVM for this station. The next three blocks provide the parameters of the CVM invoked for each time interval (1–6), for the *X*, *Y* and *Z* coordinates, respectively. This PCVM is also depicted in Fig. 12

An *M* 8.4 EQ occurred on 2001/06/23 at 20:33:14 UTC. Jump time is 01:174:73994 in YY:DOY:SOD format

An *M* 7.6 EQ occurred on 2001/07/07 at 09:38:43 UTC. Jump time is 01:188:34094 in YY:DOY:SOD format

Time format YY:DOY:SOD = year:day of year:second of day

posterior position estimate is expected, desired or 'allowed' to deviate from the prior estimate. If we very tightly constrain the daily prior positions for all stations in C to the positions predicted by the target RF model by setting the associated constraints to something very small, such as 2–3 mm, while allowing much looser constraints (perhaps 20–30 cm) for the remaining stations (i.e. those in N but not in C), then each daily polyhedron solution is automatically 'aligned' to the target frame, and the velocities obtained by analyzing the time series of coordinates for each station are in consequence tied to that same frame. This is the general approach favored most surveyors.

6.3.2 Posterior imposition of the RF

Most GPS geodesists, including us, use an alternative approach to geodetic time series analysis involving posterior imposition of the RF. All stations appearing in the daily GPS analysis (including the stations in set C) are only *loosely constrained* to the prior coordinates predicted by the target RF. We then manipulate this time series of loosely constrained or 'loose' polyhedron solutions in two stages, which we call *time series alignment* or just *alignment* for short, followed by *post-alignment imposition of the RF*. The alignment process unfolds in an arbitrary and drifting axis system. The target

Fig. 12 Contrasting the PCVM invoked by ITRF2008, and our GPS solutions and the ELTM (expressed in OSU08), for the east component of displacement at AREQ. **a** Shows the entire time series, allowing the 6 time segments (see Table 1) to be resolved. **b** The interseismic period prior to the M_w 8.4 event. **c** The end of the time series, emphasizing the inability of the CVM adopted for the 6th time segment to accurately predict the displacement trend after about mid 2010
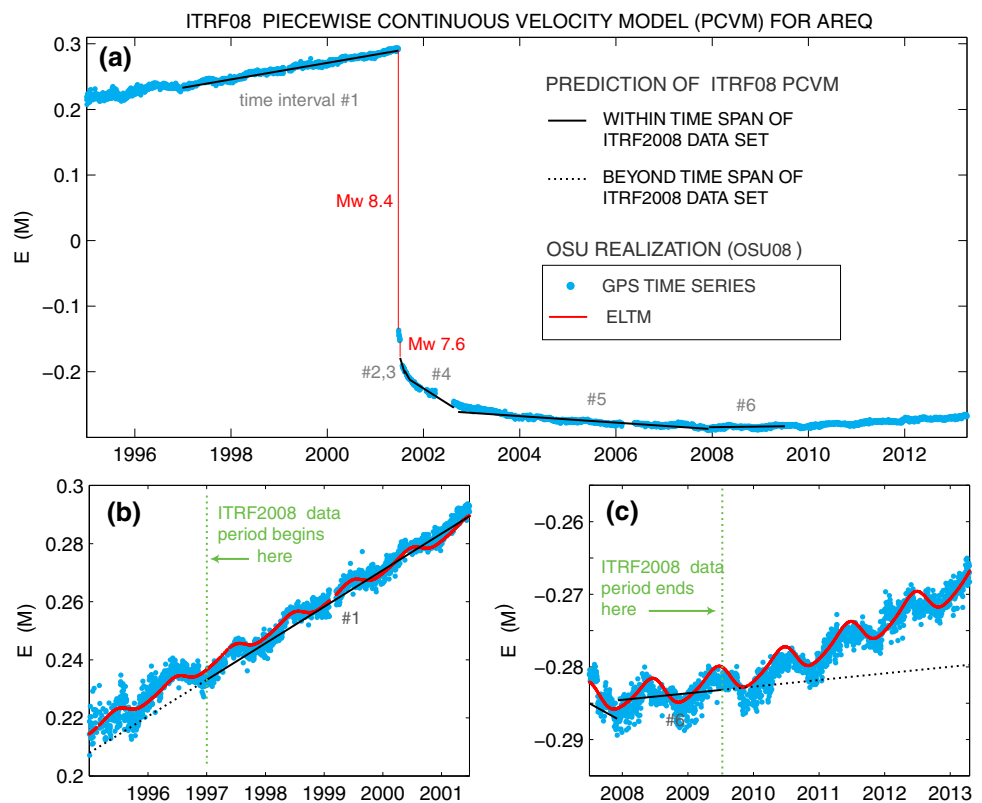
frame is invoked only after alignment is complete. Both of these stages involve the use of Helmert transformations that rigorously preserve the inner geometry of each daily solution for the GPS polyhedron, while manipulating their outer coordinates. The first and most time-consuming stage, alignment, involves a great many H6 transformations and repeated estimation of the parameters of all station trajectory models. The second stage simply transforms these results into the target RF using a single 12-parameter Helmert transformation, H12. This can be thought of as two H6 transformations operating in concert—one in position space and the other in velocity space (Appendix 2).

In implementing this two-stage approach, we could use the CVM favored by the authors of ITRF, or the more general trajectory models that we favor. No matter what trajectory models are used, time series alignment is accomplished in an inversion that (i) estimates the 6-parameters of each daily H6 transformation which align each loose polyhedron solution with the daily predictions of the network trajectory model, and (ii) estimates the coefficients of the trajectory models so as to produce the closest possible agreement of these models with the aligned time series. We do this by minimizing the inner coordinate scatter (i.e. collective misfit *length*) of the time series of transformed polyhedra about the corresponding series of model polyhedra. We emphasize that both the outer coordinates of the aligned time series of GPS polyhedra, and the parameters of the various station trajectory models change with each iteration of the inverse algorithm.

This process stops (typically after 7–12 iterations) only when the inner coordinate scatter ceases to decline by a significant amount. Because we take this approach, the station reference positions and station velocities estimated by this algorithm are ultimately expressed in no particular reference frame. (Though, because the daily transformations and infinitesimal rotations associated with the daily Helmert transformations (H6) produce only small changes in daily station coordinates, the final reference position and velocities are stated in an arbitrary frame that still lies 'fairly close to', or is weakly aligned with the target RF model used to generate the prior position estimates during GPS data processing.)

Once the alignment process discussed above is complete, we move on to the next step, and transform the RF-sensitive coefficients of the station trajectory models and the associated time series of (aligned) polyhedra, by estimating the 12-parameters of a generalized Helmert transformation (H12) consisting of three translations, three translation rates, three rotations and three rotations rates that specify the relative motion of the arbitrary frame and the target frame (Appendix 2). Those interested in why we utilize the H6 and H12 transformations, rather than the H7 and H14 transformations that also include scale change and scale change rate parameters, are referred to Appendix 1 of Bevis et al. (2012a).

At first look, this two-step approach may seem to be unnecessarily complicated relative to the simpler approach, described earlier, when the RF was imposed ('upstream') on a day-by-day basis during daily GPS data processing. Indeed,

if every station in the network N is also included in the definition of the target RF (so that set C = set N), then one might expect a similar result. But when, as is typical, many stations in N are not included in C, then the two approaches are not, in fact, equivalent. *This is because during the alignment process, the daily Helmert parameters used to align polyhedra onto the network trajectory model, are estimated using* <u>*all*</u> *stations in* N, *and not just those in* C. That is, we are exploiting the fact that we know the general form of the station trajectory at *every* station, not just those stations incorporated in the definition of the target RF. This leads to better and sometimes much better constrained estimates for the Helmert parameters, and this is reflected in the quality of the final solution.

6.4 Trajectory prediction error and RF instability

The 'up front' approach to RF realization discussed in Sect. 6.3.1 provides a simple context in which to explain the concept of *RF jitter*. Suppose we have produced GPS solutions for a sparse regional network or polyhedron for 7 days in a row. Suppose our network of 20 stations contains four of the reference stations associated with our favorite RF. Suppose we tightly constrained the positions of these four stations at the 2 mm level in order to invoke our target RF. Since very little plate motion occurs in 7 days, the prior coordinates adopted for these reference stations changed very little from day to day. Suppose the RF model we use to generate the prior position estimates for our reference stations was formulated using the CVM (Eq. 4). Suppose that 3 of these RF stations are very well characterized by the CVM, but the 4th is not, because it is located next to the Amazon river and it experiences a very large annual vertical displacement cycle (Bevis et al. 2005) that is not incorporated into its trajectory model. Let us suppose that this cycle happens to be near its maximum during our study period, and the station location is actually 25 mm above the long term secular displacement trend invoked by the RF's CVM. The entire GPS network solution will be very strongly influenced by this Amazonian station since it will tend generate a far larger residual vector than the other three reference stations, and any least squares process seeks to balance the square of the magnitude of the normalized residual vectors. If we had used the other 3 reference stations alone we might have obtained nearly perfect 'alignment' of prior and posterior position estimates. But when the Amazonian reference station is included, the alignment parameters estimated internally by the GPS processing package produce a much-degraded fit at these three stations so as to reduce the enormous residual at the Amazonian station. The entire alignment process will have been biased by this one station. But suppose the Amazonian station had a technical problems on days 4 and 5, its data were not available, and do not appear in our time series. Then this bias

in the alignment process would not have occurred on these days. That is the alignment process on days 1, 2, 3, 6 and 7 would be distinct from that on days 4 and 5. This effect would generate artificial scatter in the coordinates of each and every station in the time series. We can think of this as the polyhedron jerking about, but actually it was the realization of the RF that was jerking or jittering.

RF jitter is a stochastic process (think of Brownian motion) in which day-to-day variations in the (inner geometrical) misfit between the GPS polyhedron and the model polyhedron (for the reference stations) produce statistical fluctuations in the alignment process associated with RF realization. These variations are often modulated by the presence or absence of poorly modeled stations in the time series. Temporal variability in trajectory model error can also arise in response to the model not capturing rapidly changing displacements such as postseismic displacement in the days or weeks following a great earthquake, or unrecognized jumps in station time series.

Returning to our thought experiment, it is fairly obvious that the RF realization error (jitter) precipitated by the uneven appearance of the Amazonian station could be reduced by (i) using more realistic trajectory models (so as to reduce the number of large station-day residuals), and/or (ii) using many rather than fewer reference stations to control the alignment process, so that one or two highly problematic stations do not have so great an influence (this is the *defendit numerus* strategy advocated by Bevis et al. 2012a).

Of course, the prediction errors generated by oversimplified trajectory models can also develop slowly over months or years, slowly but relentlessly distorting the alignment process, producing *RF drift* rather than high frequency RF jitter. If 'RF realization error' (Dietrich et al. 2001) is not a constant bias, but a time varying error or discrepancy, then RF jitter relates or refers to its high frequency component and RF drift refers to its low frequency component. Failure to incorporate significant and sustained accelerations in a trajectory model, as discussed in Sect. 6.2 and illustrated in Fig. 11, can lead to RF drift. For example, suppose two GPS groups (A and B) analyze another sparse regional network, mostly located in North America, over more than one decade, using the same target RF. Suppose both groups use five reference stations to realize this frame, but they have only four of these stations are in common. Let us suppose that these four reference stations are well characterized by the CVM invoked in the definition of the target frame, as is the fifth reference station invoked by geodesy group A. But let's suppose geodesy group B chooses a Greenland station for its fifth reference station, and this station has been subject to a large and nearly constant rate of acceleration throughout its lifetime, but the target RF definition invokes a CVM for this station. As discussed earlier, the inconsistency of this CVM and the actual trajectory means that the inner geometry

of the model polyhedra used as alignment targets by group B will be inconsistent with the actual inner geometry of this sub-network, and this inconsistency will vary systematically over time. As a result, the two GPS groups will find that their 'aligned' time series have different velocities at nearly every station, and will probably conclude that despite their best intentions, their results are expressed in somewhat different RFs. We call this RF drift, in contrast to RF jitter, because it derives from a discrepancy whose impact systemically grows over time.

6.5 The broader context: repeated reanalysis

For the last ∼3 years we have processed and reprocessed >1,200 unique CGPS stations (up to 700 per day) for a 15–18 year time period (Bevis et al. 2012a). We perform the daily GPS data analysis using GAMIT/GLOBK software (Herring et al. 2010). In the first processing pass of our most recent series of solutions, all prior coordinates were similar to those used by IGS analysis centers, based on the ITRF2008 standard, and we used IGS orbital solutions expressed in ITRF2008. We then fit each station time series using an SLTM or ETM/ELTM. Once the alignment step was completed, we transformed our trajectory models (and the associated GPS time series) into a frame that was as closely aligned as possible with ITRF2008, while preserving the inner geometry of the aligned time series and the network trajectory model. We then used our trajectory model to generate the prior coordinate estimates for the next iteration (or 'rerun') of GPS data processing. Because the prior positions are no longer completely consistent with ITRF2008, we allowed the orbits to relax so as to maintain consistency with the global polyhedron. We stacked or aligned the new GPS time series, transformed the aligned time series and the associated network trajectory model so as to nearly match the predictions of ITRF2008, and then used the updated network trajectory model to generate the prior coordinates for the next processing iteration, etc. During the early iterations we used fairly loose constraints (∼30 cm) even for the 'best' of the stations used to define ITRF2008, and in subsequent iterations we gradually tightened these constraints, but always ensured that the station constraints (i.e. the (σ_X, σ_Y, σ_Z) discussed in Sect. 6.3) remained 5–10 times larger than the WRMS scatter levels produced by the previous analysis for the corresponding stations. What we noticed during this iterative process was that the inner coordinate scatter of the entire GPS time series around the network trajectory model declined with each new iteration. This includes the last iteration, so our reanalysis effort has yet to converge.

The trajectory models produced by our fifth and most recent global rerun in May 2013 can model the GPS time series (expressed in OSU08) at almost every station with a daily WRMS misfit of 1–3.5 mm in the horizontal, and

2–7 mm in the vertical, once something like 0.5–2% of the observations were down-weighted by our robust inversion algorithms. The WRMS misfit levels for all stations and all epochs combined were 2.3, 2.2 and 4.7 mm in E, N, and U respectively. These scatter levels represent a >50 % reduction from those achieved in the first iteration. The posterior re-alignment of our daily polyhedra and trajectory models with the predictions of ITRF2008 was achieved using 84 common stations rather than all common stations. These stations were selected on the basis of (a) high solution quality, (b) a very long time series uninterrupted by any jumps, and (c) no suggestion that velocity is changing significantly as a function of time. The WRMS difference between the velocities assigned to these stations in ITRF2008 and in our solution (i.e. in OSU08) was 0.55, 0.64 and 0.55 mm/year in the X, Y and Z directions respectively. The WRMS differences in the predicted positions of these stations at the ITRF2008 reference epoch was 5.9, 4.3 and 8.2 mm, respectively.

It is said that 'the proof of the pudding is in the eating', and we are persuaded of the utility of our overall approach by the steady reduction of the WRMS scatter levels associated each iteration of our global analysis. Only recently, for example, have we observed WRMS misfit levels in the horizontal components of displacement fall to the ∼1 mm level, or better, at many higher latitude stations—a recent solution for KAGA in Greenland (Fig. 13) provides an example. Since (data-model) misfits reflect both measurement error and mis-specification (usually oversimplification) of the trajectory model, then the random component of our horizontal GPS positioning error is probably rather better than this. (Water vapor dynamics still prevents us reaching this level of repeatability at mid and low latitudes).

7 Discussion

Geodesists engaged in daily GPS data processing traditionally resist or avoid the application of tight prior constraints on station position because this might inject mis-information into the analysis of satellites orbits and the global polyhedron. (One might tightly constrain stations to incorrect or inconsistent sets of prior coordinates. Tight constraints would be justified if we knew the 'correct' solution in advance, but in this case we would not need to engage in geodetic measurement!). Because we avoid tight prior constraints, the GPS-estimated coordinates (outer geometry) of the polyhedron scatters much more, from day to day, than does the implied inner geometry. That is, the daily solutions for the polyhedron have nearly the correct shape and size, but are nearly randomly displaced and reoriented by noise which differs from one day to the next. The alignment process suppresses this outer coordinate noise, by re-aligning the time series of polyhedra. We could approach this task by aligning each GPS

Fig. 13 The displacement history at GNET station KAGA, fit with an ELTM that invokes a quadratic trend. Plots **a–c** show the *U*, *E* and *N* components of displacement, respectively. Plot **d** depicts the implied temporal variation in uplift rate. **e** shows the location of station KAGA
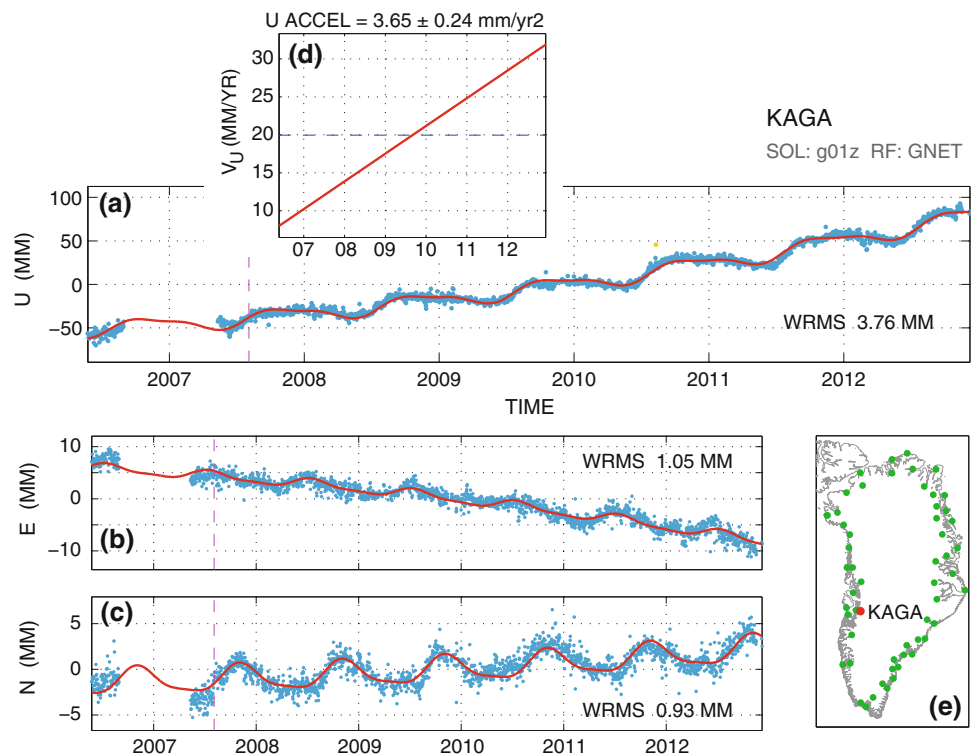
polyhedron onto the polyhedron for the previous day. We do not do so because the station trajectory model is typically a better predictor of the polyhedron's inner geometry at a given epoch than is the inner geometry estimated (using just one day's data) the day before. The beauty of the iterative alignment process is that well-tuned station trajectory models lead to better alignment of the GPS time series, and better aligned time series lead to improved station trajectory models. This alignment or stacking-in-time process is the central task of geodetic time series analysis, and posterior imposition of a target RF is a relatively minor post-process. The greatest benefit of more realistic trajectory models is improvement of time series alignment.

Our course, improved trajectory models also generate improved geophysical insight. Five years ago and more, geophysicists routinely plotted GPS-derived vertical velocity vectors on maps of Antarctica, Greenland and elsewhere, and then attempted to explain them in terms of glacial isostatic adjustment. Recently it has been recognized that uplift rates in Greenland and in parts of Antarctica are so variable in time that the lifetime average uplift rate is a limited representation of what is happening, and it is much more fruitful to think in terms of accelerations as well as average rates of uplift (Khan et al. 2010; Jiang et al. 2010; Bevis et al. 2012b). The vertical time series at KAGA provides a particularly striking example. Here the uplift rate has increased from ∼7.5 mm/year in mid-2006 to ∼32 mm/year at the end of 2012. It should be obvious looking at Fig. 13 (and Fig. 4) that as technical improvement in the alignment process reduces the level of

scatter of the GPS measurements about the trajectory model, this improves our ability to resolve curvature in the trend component of this model.

Although many geodetic and geophysical research groups are using the more general classes of trajectory models to characterize their crustal displacement time series, very few of them are using these trajectory models to estimate orbital solutions or realize (and modify) the frames in which their network time series are expressed. Because a RF is, for operational purposes, a network trajectory model, and our RF model now differs from that of ITRF2008 for the stations we have in common, our final RF—in which we estimate orbital solutions as well as the network time series—is no longer identical to ITRF2008. Our final RF inherited its GPS scale from ITRF2008, and is not intended to serve as a shared, general purpose RF. Indeed, as OSU08 varies from one iteration of our global analysis to the next one, it can hardly be thought of as 'standard' RF in the traditional sense.

We believe that the steady improvements we are attaining as we repeat and refine our global analyses mainly result from (i) the increasing geometrical consistency between our network time series, our trajectory models, the prior position estimates we inject into GPS data processing, and our orbital solutions, and (ii) ongoing refinement of our trajectory models for each station we process. Lowering the scatter levels in each new solution often helps us identify small and previously unidentified jumps in the time series, usually traceable to an overlooked change in the antenna or radome, or pre-

viously unresolved phenomenology such curvature (acceleration) in the secular trends. This causes us to change the number of parameters available to a trajectory model, or even invoke a new class of trajectory model. For example, newly resolved jumps are accommodated by adding another Heaviside jump to the trajectory model, and newly resolved curvature might be accommodated by invoking a quadratic rather than a linear secular trend.

Although OSU08 plays a special role in our reanalysis effort because the prior coordinates injected to daily GPS data processing are always stated in that frame (so as to 'stay close' to ITRF2008), in the scientific studies that utilize the products of our latest reanalysis, we nearly always transform into other frames, typically on a project by project basis, as previously discussed by Bevis et al. (2012a). This is because as geophysicists we place no particular importance on the no-net-rotation (NNR) aspect of ITRF, nor any other purely conventional aspects (Appendix 2) of this RF. Some readers may be worried about our apparent disinterest in monitoring the position and motion of Earth's center of mass (CM) in our frame. But, this is because we choose to estimate relative frame motion later on, in the context of specific comparisons between GPS uplift rate, model GIA (or PGR) rates, or observed rates of sea level rise (Bevis et al. 2012a,b).

There is still a lot of interesting research to be done with regard to trajectory models. In addition to devising and testing new classes of trajectory model, there are open questions as to how best to deploy or utilize extant models. When we invoke a SLTM with a polynomial trend, for example, how should we select the value we assign to the maximum power n_P and thus the number of degrees of freedom available to our trend sub-model and the SLTM? Obviously the more degrees of freedom one assigns to the polynomial trend, the better it can fit a given station position time series, but at some point one begins to model 'noise' rather than 'signal'. The stability and thus the predictive power of a model will fall given too many degrees of freedom simply because it has started to model the noise. Given the role of a trajectory model in providing prior coordinates in GPS data processing, etc., we are very concerned with the model's predictive power—each global re-analysis tends to include a significant amount of 'new' GPS data. A graph of the WRMS misfit of data and trajectory model versus the number of degrees of freedom assigned to the model can help one assess this trade-off between resolution and reliability (Jackson 1972; Lawson and Hanson 1974). One approach is to increase n_P until the WRMS misfit matches ones expectation for the standard deviation of the measurement noise at that station. But in practice, one has no exact prior knowledge of this statistic. As a result, there is a lot of work in the numerical analysis community associated with identifying natural 'break points' in this or related trade-off curves such as the 'L-curve' (Lawson and Hanson 1974; Hansen 1992).

If the analyst is prepared to assume that the noise is drawn from a Gaussian distribution, it is also possible to use a formal statistical hypothesis test to determine if each increment in the value of n_P produces a statistically significant reduction in the scatter of the data about the model. In order to implement this hypothesis test one must adopt a given confidence level. Changing the confidence level at which the hypothesis accept/reject decision is made can change the value finally selected for n_P. Thus the decision on the best value of n_P remains somewhat arbitrary. We do not use this approach because we do not believe that the noise process in GPS positioning is Gaussian (Bevis et al. 2012a) or corresponds to any other named statistical distribution. Indeed, we do not believe that the GPS positioning noise process is stationary in space or time.

Our approach to choosing n_P for each station trajectory model tends to be heuristic and conservative. We are very reluctant to increase n_P from 1 to 2, and even more reluctant to increase it from 2 to 3. When we fit an SLTM to a time series, we usually plot the E, N and U residuals as a function of time. Only when these residuals suggest strong and systematic structure, do we consider increasing the value of n_P. The great majority of the station models in our global analysis have $n_P = 1$, and the great majority of those with $n_P > 1$ have $n_P = 2$. We prefer to 'under-model' the data rather than 'over-model' it. This is a subjective decision, if a conservative one. But the 'regularization' of inverse problems and empirical trend analysis of almost any kind tends to be something of an art (Hanke and Hansen 1993).

We admit to a rather pragmatic attitude towards our trajectory models. As geodesists we tend to deploy them opportunistically, focusing on improving the consistency of our frame, our orbital solutions and the daily repeatability of our time series, with a blithe disregard to physical meaning. But subsequently, as geophysicists, we feel no compunction in abandoning a trajectory model (as a mere first approximation) as we seek to probe the physical causes of crustal displacement and deformation.

Geophysicists and geodesists typically play complementary roles in crustal motion geodesy. Geophysicists use crustal motion geodesy—often in combination with additional observations—to study a wide range of geodynamic phenomena. Their progress in discerning and explaining the phenomenology of crustal motion, will sometimes help geodesists better formulate their station trajectory models. Better trajectory models will lead to better predictions of station position, improved orbital solutions, more stable reference frames and less noisy times series of crustal displacement. These 'cleaner' displacement time series would then lead geophysicists to resolve new dynamic phenomena (e.g. the 2010 'uplift anomaly' in Greenland, Bevis et al. 2012b) or improve their characterization of already recognized phenomena, allowing this synergetic cycle to continue. Up to

now, however, new classes of trajectory models have had very little impact on the way in which standard or named RFs are designed and defined. The trajectory models used to define the ITRF, in particular, have not changed for more than a decade. We expect this to change in the next few years.

8 Appendix 1: The insensitivity of the ELTM to the transient timescale parameter T

Here we demonstrate our contention that the SLTM augmented with the logarithmic transient formula is surprisingly insensitive to the value of the nonlinear parameter T in the sense that if one assigns a moderately erroneous value to T, and allows A to adjust accordingly, one can retain a very good fit to the great majority of the data. As we shall show here, this insensitivity is not inherent to the logarithmic transient formula itself, but arises when it is used in conjunction with the SLTM (Eq. 8).

We begin by generating a simulated data set that we can then analyze in several different ways. Suppose a transient displacement is given by Eq. (9), where Δt is time since the earthquake occurred, $T = 1$ year and $A = 100$ mm. Note that this formula applies only after the earthquake has occurred, so we are restricted to the domain $\Delta t \geq 0$. The scalar d might refer to any one of the geocentric Cartesian coordinates (X, Y, or Z) or topocentric Cartesian coordinates (E, N, and U) used to describe a geodetic time series. We computed the value of d, using this formula, once per day for the first 5 years of the transient, and then added zero-mean Gaussian noise with sigma = 3 mm so as to simulate observations incorporating positioning noise (the blue dots in Fig. 14).

First, we invert this data set using the transient formula alone, or in isolation. If we invert for A under the (correct) assumption that $T = 1$ year, our least squares (LS) estimate for $A = 100.03$ mm, which produces the dotted black curve passing through the middle of the data (Fig. 14a). The RMS deviation of the data about this curve is 2.99 mm, closely matching the designed level of noise. Next we 'erroneously' set $T = 3$ years, and obtained a LS estimate $A = 196.4$ mm, producing the dashed red curve (Fig. 14a). The RMS devia-

tion of the data about this curve is 9.2 mm. Finally, we erroneously set $T = 1/3$ years, and estimated $A = 60.0$ mm, producing the solid magenta curve (Fig. 14 a) and an RMS misfit of 9.7 mm. We twice assigned a value for T that was off by a factor of 3 (either too big, or too small), and resulting approximation to the transient behavior was rather poor, and the RMS misfit was more than three times larger than the measurement noise.

However, we will rarely estimate the parameters of the transient model (Eq. 9) in isolation, since any station recording a postseismic transient is almost certain to have recorded a coseismic jump as well, and this jump would be modeled using a Heaviside function. This means our transient model, operating in conjunction with the SLTM (Eq. 8), can 'misappropriate' part of this jump, so as to reduce the misfit of the combined model to the transient displacement data. Therefore in our second set of inversions we invoke a two-parameter model in which we estimate both A and a constant offset, or d-axis intercept, d_0, that mimics the impact of the Heaviside function. This offset parameter has very little effect when T is assigned its correct value of 1 year, but when T is assigned values of 3 and 1/3 year, the estimates of A are modified to 172.4 and 70.8 mm, respectively, and the RMS misfit is reduced to 5.8 and 5.6 mm, respectively. The improved fit to the data (Fig. 14b) was achieved using curves that now predict d values of 16.8 and -23.7 mm when $\Delta t = 0$. These offsets correspond to that portion of the Heaviside jump being used to improve any imperfectly formulated transient model rather than represent a coseismic offset.

When we use logarithmic transients operationally, we are prepared to down-weight the observations occurring in the first few weeks or even the first few months after the earthquake (either a priori, or via an iterative re-weighting approach, Holland and Welsch 1977) because we believe that short-lived poroelastic rebound (should it occur) is unlikely to produce a logarithmic transient. This down-weighting can also impact our fit to the transient when T has been assigned an incorrect value, since much of the misfit associated with an incorrect T value is concentrated early in the time series (Fig. 14b). We illustrate this effect by repeating our LS analysis, without using the simulated displacement data when $\Delta t < 1$ month. Down-weighting or eliminating these early observations has very little impact to our fit when T is assigned its correct value, but if T is assigned a value of 3 or 1/3 year, the resulting estimates for A change to 169.9 and 72.1 mm, respectively, and the RMS deviations (for $\Delta t > 1$ month) are further reduced to 5.2 and 5.1 mm, respectively. The modified curves now predict d values of 18.4 and -26.6 mm, respectively, when $\Delta t = 0$.

In practice, the trend component of a SLTM at most stations invokes constant velocity, so if we fit an SLTM augmented with a logarithmic transient, it is also likely that the presence of the transient will modify the velocity estimate

Fig. 14 Modeling a logarithmic transient with correct and erroneous values of *T*, using **a** the logarithmic transient formula (LTF) alone, **b** the LTF plus a constant offset parameter, **c** LTF, offset and velocity parameters
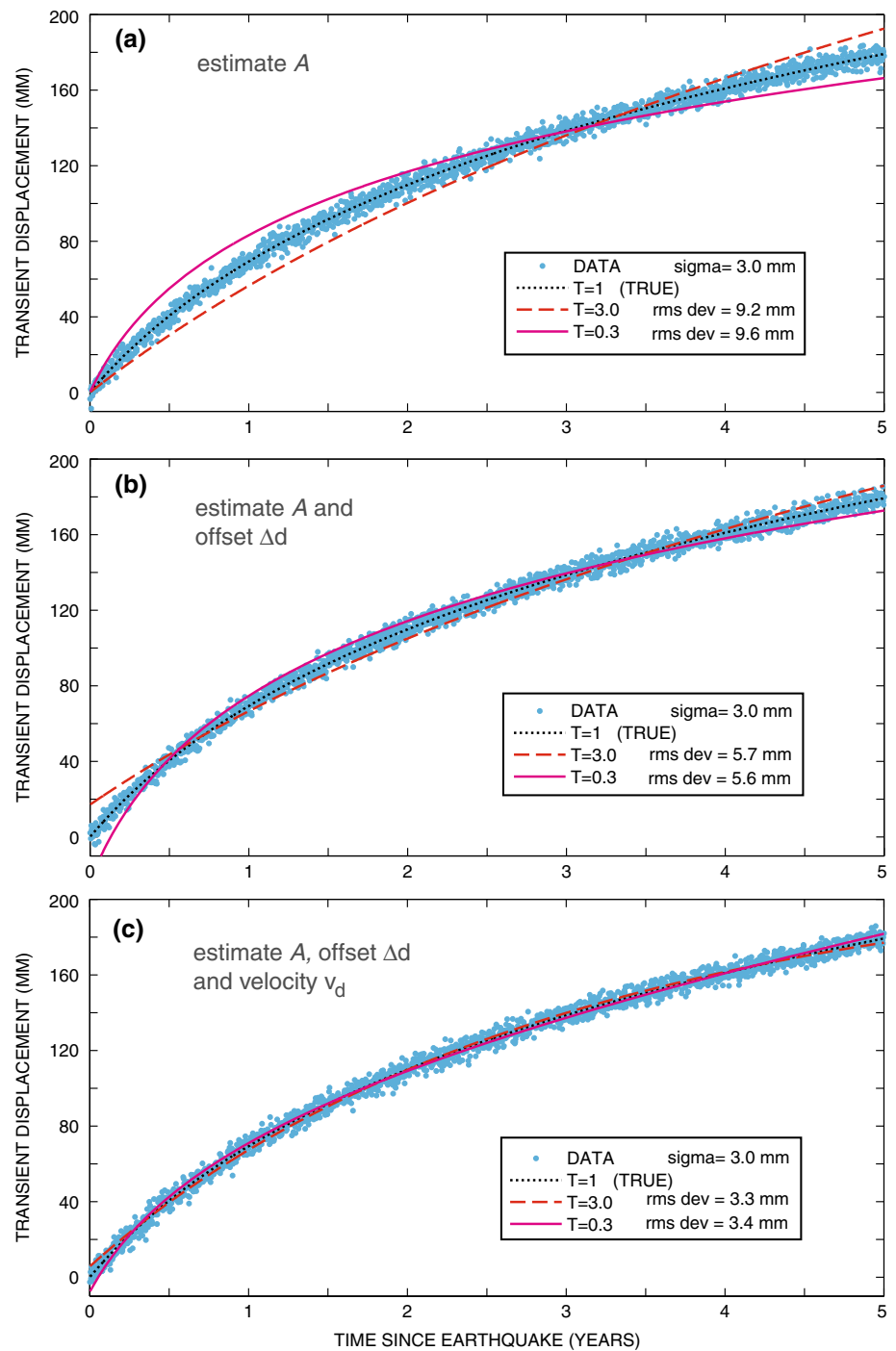
to some degree so as to mitigate misfits of the transient that arise when we assign an incorrect value to parameter T. If that portion of the time series that precedes the earthquake is several years long, the constant velocity estimate is probably quite tightly constrained by the pre-earthquake observations, so the velocity perturbation just discussed is likely to be small. In the event that the earthquake occurred early in the time series, however, a larger perturbation might occur.

We now simulate the extreme case in the constant velocity estimate is entirely determined using data acquired after the onset of the transient. We do this by assigning the value of T, as before, and fitting a three parameter model that estimates an offset, a constant velocity and the parameter A of the logarithmic transient. First we do this using all data, and then we repeat the inversion using only the observations with $\Delta t > 1$ month. When we invert all data while assuming the correct

Table 2 Results of our model-fitting experiments with the simulated transient data (Fig. 14)

Model parameters	Value assigned to T (years)	Time period of analysis	WRMS misfit (mm)
A	1 (correct)	All	3.0
A	3	All	9.2
A	1/3	All	9.7
$A, \Delta d$	1 (correct)	All	3.0
$A, \Delta d$	3	All	5.8
$A, \Delta d$	3	$\Delta t > 1$ month	5.2
$A, \Delta d$	1/3	All	5.6
$A, \Delta d$	1/3	$\Delta t > 1$ month	5.1
$A, \Delta d, v$	1 (correct)	All	3.0
$A, \Delta d, v$	3	All	3.3
$A, \Delta d, v$	3	$\Delta t > 1$ month	3.1
$A, \Delta d, v$	1/3	All	3.4
$A, \Delta d, v$	1/3	$\Delta t > 1$ month	3.2

value for T (i.e. $T = 1$ year), the addition of a velocity parameter to our model makes only a tiny change to the goodness of fit. However, it produces a dramatic improvement to the RMS misfit when T is incorrectly assigned. Setting $T = 3$ years produces an RMS misfit of 3.3 mm, whereas setting $T = 1/3$ year results in an RMS misfit of 3.4 mm. If we model the transient for $\Delta t > 1$ month, the misfits are further reduced to 3.1 and 3.2 mm, which is almost a perfect fit, given that the design or constructed level of RMS measurement noise is 3.0 mm. Whether or not we use the first month of observations, the fits are remarkably good, only 3–13 % higher than the designed level of noise in the data. The results of our experiments with this synthetic data set are summarized in Table 2.

If one's goal is fitting the actual station trajectory, rather than isolating the individual components of this trajectory, the trade-off between the coefficients of the SLTM and the coefficient A of the transient model is highly beneficial if one does not know the correct value of the parameter T. This provides the geodesist with a useful option, discussed in Sect. 5, i.e. to simply assign a 'reasonable' or default value for T, and fit an augmented trajectory model using a completely *linear* LS approach. The value of T can be refined later, on a station by station basis, using a nonlinear estimation approach.

9 Appendix 2: Inner geometry, outer geometry and Helmert transformations

The theme of outer versus inner geometry appears quite frequently in the literatures of classical mechanics and space geodesy, but individual discussions are often framed using different terminologies, e.g. 'fiducial' versus 'fiducial-free' analysis (Heflin et al. 1992), and this tends to obscure the

ubiquity and utility of these concepts. We use this appendix to further explain *our* terminology, both to clarify our meaning in the main text of this paper, and to help students and non-specialists grasp these sometimes subtle but fundamental and very useful concepts.

Einstein once said that 'all geometry is length'. He was emphasizing the central role of the metric tensor in defining and measuring both lengths and angles. But even in classical or Galilean mechanics, this statement remains valid and profound.

Let us assume that we work in 3D physical space equipped with a well-defined scale which is independent of position or orientation (i.e. there is a single scalar metric for the whole space), that this scale does not change over time, and that we are concerned with the geometry of a set of points—representing GPS stations—that constitute the vertices of a polyhedron embedded in this space. The *inner geometry* of the polyhedron addresses its shape and size, but not its location or orientation. The *outer geometry* of the polyhedron encodes or implies its inner geometry, but also addresses the position and the orientation of the polyhedron relative to an external reference frame, which we will normally think of as a Cartesian axis system [X, Y, Z]. The outer geometry of a polyhedron is expressed or defined by measuring or assigning the $3n$ coordinates $\{X_i, Y_i, Z_i\}$, $i = 1 : n$ of the n stations or vertices composing the polyhedron. The inner geometry of the polyhedron is the geometry that can be measured or specified with reference to the polyhedron, but without reference to (or access to) to an axis system or RF. If one knows the lengths of all baselines (inter-vertex distances) in the polyhedron, then the inner geometry is completely determined. Inner geometry is all about length, so, in principle, all one really needs to discuss it or measure it is a length scale.

Since the inner geometry of a polyhedron is totally determined by measuring or specifying the lengths between the

vertices of the polyhedron, this description is invariant with respect to a change of axis system. But the coordinates (X, Y, Z) of these vertices, i.e. the outer geometry of the polyhedron, obviously depend on the particular choice of axis system. It is because the inner geometry is invariant with respect to the choice of axes (as long as the scale or metric is fixed) that the inner geometry is often thought of as being the deeper or more fundamental description of the polyhedron's geometry. (If we redefined the Z-axis so that it passed through New York, the distance between New York and Los Angeles would remain the same).

There is a subtlety of language that can cause confusion. We can refer to the inner geometry of a polyhedron specified by listing its $3n$ coordinates. But we are referring to the shape and size of the polyhedron implied by those coordinates, and not the coordinates themselves. Indeed, one often discusses the inner geometry of a polyhedron with reference to its outer coordinates, because this is much simpler in practice than describing an independent set of baseline lengths. One could specify the inner geometry with a set of lengths and related metadata, and those lengths might be referred to as *inner coordinates*. This is very rarely done in practice because formulating the metadata required to interpret these lengths correctly is quite complicated. Nevertheless, the concept of inner coordinates is a useful one. Let us suppose we have aligned two very similar polyhedra (perhaps a GPS polyhedron and a model polyhedron) as closely as possible, and determined the remaining differences in the coordinates of corresponding vertices $\{\Delta X_i, \Delta Y_i, \Delta Z_i\}, i = 1 : n$. We might characterize total misfit of these polyhedra using the statistic $M^2 = \Sigma_i (\Delta X_i^2 + \Delta Y_i^2 + \Delta Z_i^2)$ which, at first sight, might appear to be an outer coordinate description. But, since $\Delta X_i^2 + \Delta Y_i^2 + \Delta Z_i^2 = L_i^2$, where L_i is the *length* of the residual vector between the i th vertex of the two polyhedra, then the global misfit measure $M^2 = \Sigma_i (\Delta L_i^2)$ is actually an inner geometrical measure since it is depends only on length. The measure M is independent of the RF. For this reason M can be said to quantify the *inner coordinate misfit* or *inner coordinate scatter* of the two polyhedra. If formal estimates of the uncertainties attending each ΔX_i, ΔY_i, and ΔZ_i. are available, they can be used to weight these quantities when evaluating M.

One must specify the lengths of all six sides of a tetrahedron (which has 4 vertices) in order to render it 'rigid' or to completely specify its inner geometry. If we add one new vertex to an already rigid polyhedron, it is necessary to specify the distance between the new vertex and three (non-collinear) existing vertices in order that the enlarged polyhedron remains rigid, and its inner geometry remains fully determined. It follows that we must measure a total of $(3n-6)$ lengths to determine or define the inner geometry of a polyhedron composed of n GPS stations. (Note that if $n > 4$, this is less than the total number of baselines or lengths associated

with a polygon, which number $n (n-1)/2$ in all. This situation arises because not all baselines are linearly independent, and the $(3n-6)$ lengths referred to above must be independent.) We can specify the outer geometry of this polyhedron with respect to a given set of [X, Y, Z] axes by measuring or specifying the $3n$ coordinates $\{X_i, Y_i, Z_i\}, i = 1 : n$ of the n stations in that coordinate system. The difference between needing $(3n - 6)$ parameters to determine the inner geometry of a polyhedron and needing $3n$ parameters to specify its outer geometry amounts to 6 degrees of freedom. These degrees of freedom correspond to the 3 translations (t_x, t_y, t_z) and 3 rotations (r_x, r_y, r_z) that produce a length-preserving (or *isometric*) but otherwise arbitrary displacement of the axis system (or, equivalently, any possible rigid-body displacement of the polyhedron relative to a fixed set of axes). In space geodesy, coordinate transformation associated with the RF change [X, Y, Z] => [X', Y', Z'] is achieved by the 6-parameter Helmert transformation H6 constructed using these translational and rotational parameters (see Bevis et al. 2012a, for a related discussion).

The H6 transformation can be represented in different ways. For example, if some point P has the position or coordinate vector $\mathbf{x}_A$ in frame A, and $\mathbf{x}_B$ in frame B, then

$$\mathbf{x}_B = \mathbf{t} + \mathbf{R} \, \mathbf{x}_A \tag{11}$$

where $\mathbf{t} = [t_x t_y t_z]'$ is the frame translation vector, and $\mathbf{R}$ is the frame rotation matrix. The elements of $\mathbf{R}$ depend only the frame rotation angles r_x, r_y, and r_z, which correspond to rotations about the X, Y, and Z axes, respectively. That is, the rotation matrix $\mathbf{R}$ depends only on the elements of the rotation vector $\mathbf{r} = [r_x r_y r_z]'$. When these angles are very small, the rotations are said to be infinitesimal rotations, in which case matrix $\mathbf{R}$ takes on the especially simple form, and this equation can be restated in matrix form as

$$\begin{bmatrix} X_B \\ Y_B \\ Z_B \end{bmatrix} = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} + \begin{bmatrix} 1 & -r_z & r_y \\ r_z & 1 & -r_x \\ -r_y & r_x & 1 \end{bmatrix} \begin{bmatrix} X_A \\ Y_A \\ Z_A \end{bmatrix} \tag{12}$$

We can also represent the H6 transformation as a nonlinear functional relationship

$$\mathbf{x}_B = H6(\mathbf{x}_A|\mathbf{t}, \mathbf{r}) = H6(\mathbf{x}_A|\mathbf{h}_6) \tag{13}$$

where the symbol '|' occurring between the parentheses means 'given' the parameters that follow this symbol. This formalism emphasizes that the H6 transformation is controlled by the three translations and three rotations of the frame, which can be thought of as the elements of a parameter vector $\mathbf{h}_6$.

To apply this formalism to a polyhedron it is useful to define the *network position vector*, $\mathbf{x}^N$, for a polyhedron com-

posed of n vertices or GPS stations as the column vector of length $3n$ built by stacking the n station position vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots \mathbf{x}_n$. Then

$$
\mathbf{x}_B^N = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ . \\ \mathbf{x}_n \end{bmatrix}_B = \begin{bmatrix} \mathbf{t} \\ \mathbf{t} \\ . \\ \mathbf{t} \end{bmatrix} + \begin{bmatrix} \mathbf{R} & & 0 \\ & \mathbf{R} & \\ & & . \\ 0 & & \mathbf{R} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ . \\ \mathbf{x}_n \end{bmatrix}_A \quad (14)
$$

or

$$
\mathbf{x}_B^N = H6(\mathbf{x}_A^N | \mathbf{h}_6) \quad (15)
$$

These equivalent Eqs. (14 and 15) can be thought of as expressing a forward problem: given knowledge of the frame translation and frame rotation parameters (i.e. the vector $\mathbf{h}_6$), transform the coordinates of a polyhedron (i.e. network position vector) $\mathbf{x}^N$ in frame A so to determine its coordinates in frame B.

The problem of 'aligning' two polyhedra is actually the inverse problem. Suppose we wish to align polyhedron A with a 'target' polyhedron T. Then we seek a particular vector $\mathbf{h}_6$—i.e., a particular change of RF—that minimizes some measure of the (post-transformation) deviation or residual vector, $\mathbf{d}$, defined as

$$
\mathbf{d} = \mathbf{x}_T^N - H6(\mathbf{x}_A^N | \mathbf{h}_6) \quad (16)
$$

The obvious measure or penalty function is that discussed earlier, i.e.

$$
M^2 = \mathbf{d}'^* \mathbf{d} = \sum_{i=1}^{n} \Delta X_i^2 + \Delta Y_i^2 + \Delta Z_i^2 \quad (17)
$$

although, in practice, one typically uses a suitably weighted version of this penalty function, such as $\mathbf{d}'^* \mathbf{C}^{-1} * \mathbf{d}$ where $\mathbf{C}$ is the covariance matrix for $\mathbf{d}$. One can usually get a similar result at less computational cost by minimizing $\mathbf{d}'^* \mathbf{W}^* \mathbf{d}$ where $\mathbf{W}$ is the diagonal matrix whose ith. diagonal element is the reciprocal of the estimated variance for the ith. element of $\mathbf{d}$. This second approach better lends itself to robust least squares inversion via iterative re-weighting (Holland and Welsch 1977).

If polyhedron A and the target polyhedron T had the same inner geometry, then they could be perfectly aligned and M (as given by Eq. 17) would be zero. If M is minimized, but has a non-zero value, this statistic characterizes the differences in the inner geometry of the two polyhedra, i.e. a post-alignment misfit which manifests a fundamental difference in size and shape. Note that this measure of inner coordinate scatter will normally be smaller, sometimes much smaller, than the corresponding outer coordinate scatter statistic

$$
M^2_{\text{pre-alignment}} = (\mathbf{x}_T^N - \mathbf{x}_A^N)'^* (\mathbf{x}_T^N - \mathbf{x}_A^N)
$$

$$
= \sum_{i-1}^{n} \Delta X_i^2 + \Delta Y_i^2 + \Delta Z_i^2 \quad (18)
$$

in which the ΔX, ΔY, and ΔZ are computed at each vertex *before* alignment takes place.

The outer geometry of a polyhedron uniquely implies or determines its inner geometry, but its inner geometry in no way implies or determines its outer geometry. There are an infinite number of outer geometries consistent with a given inner geometry, and all these outer geometries can be explored by searching over the space spanned by the six parameters of the H6 transformation. A key point is that if two observers ascribe different inner geometries to a polyhedron, then they are describing fundamentally different polyhedra, or proposing *inconsistent* descriptions of a single polyhedron. But it is possible that two observers ascribe differing outer geometries (sets of coordinates) for the same polyhedron and yet both have valid descriptions. If their measurements were perfect, this situation would arise if the two observers had adopted (or realized) different RFs, either deliberately or by accident. Two differing sets of coordinates for a GPS network are completely consistent only if they imply the same inner geometry, i.e. the same lengths for all baselines in that polyhedron. An equivalent statement that is often more useful operationally is that two sets of coordinates for a polyhedron are completely consistent if and only if it is possible to find specific values for the six parameters of the H6 transformation that transform one set of vertex coordinates so that it is in complete agreement with the other set.

The concept of inner versus outer geometry also applies in 4D (i.e. in classical space–time) to the displacement histories of GPS networks (polyhedra) that are moving and deforming over extended periods of time. The inner description of an evolving polyhedron is a description that can be framed purely in terms of baseline lengths describing the polyhedron at each and every epoch. Alternatively, this description could be framed in terms of the baseline lengths in effect at a reference epoch, and the temporal changes in these lengths that occur throughout the time span of interest. The inner geometry of a geodetic time series describing a deforming network is independent of any RF, and therefore invariant under RF transformation.

We can generalize the H6 transformation which accounts for rotat<u>ion</u> and translat<u>ion</u> of an axis system, by assuming that, over an extended period of time, we are concerned with the relationship between two axis systems wherein one is continuously translat<u>ing</u> and rotat<u>ing</u> relative to the other one. We require that if any station has a constant velocity in RF#1 it will have a (generally different) constant velocity in RF#2. This is achieved by assuming that the rotation and translation vectors appearing in Eqs. (11) and (12) change linearly in

time, such that

$$\mathbf{r} = \mathbf{r}_0 + (t - t_0)\frac{d\mathbf{r}}{dt} = \mathbf{r}_0 + (t - t_0)\dot{\mathbf{r}} \qquad (19)$$

$$\mathbf{t} = \mathbf{t}_0 + (t - t_0)\frac{d\mathbf{t}}{dt} = \mathbf{t}_0 + (t - t_0)\dot{\mathbf{t}} \qquad (20)$$

where t_0 is some reference time, adopted by convention. Note that these equations are invocations of the CVM (Eq. 2), but applied to frame translation and frame rotation. The reader might object that a rotation rate is an angular velocity, and ongoing rotation necessarily produces accelerations, which violates our requirement that every station has a constant velocity in both frames. But, in our context, vector $\mathbf{r}$ expresses infinitesimal rotations (typically $|\mathbf{r}| \sim 1$ nanoradian $= 10^{-9}$), and the same is true of the derivative $d\mathbf{r}/dt$. A typical value for this derivative or rotation rate might be $\sim 10^{-9}$ year^{-1}. Since both RFs are very nearly geocentric, and the displacements implied near the earth's surface are circular arcs, then the amplitude of the associated centripetal or transverse acceleration is given by the well known formula $a = -\omega^2 R_E$, where R_E is the radius of the earth, and ω is angular speed. Because the acceleration depends on the square of $\omega = |d\mathbf{r}/dt|$, and because ω is a tiny number, the implied accelerations are completely negligible.

Thus, in the context of crustal motion geodesy, GPS coordinate times series referred to two distinct axis systems (with a shared and time-invariant scale) can be related by the three rotations, three rotation rates, three translations and three translation rates that describe the motion of one frame relative to the other one. These are the 12 parameters of the generalized Helmert transformation, H12 (see Bevis et al. 2012a, for a related discussion). The H12 transformation can be written down in a variety of ways, but perhaps it is simplest to think of it as Eq. (11) or Eq. (12) in which the frame translations and rotations behave according to Eqs. (19) and (20). We can combine these equations so as to arrive at a single complicated equation, but this is not particularly edifying.

In the event that the trajectory of each station in the network is characterized in RF#1 using a CVM, that is, by stating a reference position (x_R, y_R, z_R) and a velocity (v_x, v_y, v_z) for each station in the network, then the inner geometry equivalent or 'isometric' description in RF#2 is easily found using the appropriate H12 transformation. Two distinct sets of reference positions and station velocities purporting to represent the same polyhedron and the same period of time are *consistent*, i.e. imply the same inner geometry, if and only if we can find specific values for the 12 parameters of the H12 transformation that maps one description $\{x_R, y_R, z_R, v_x, v_y, v_z\}^i$, $i = 1 : n$ onto the other one. If we cannot do this exactly, but we find instead the parameters that bring the two descriptions as closely together as possible, then the remaining differences manifest irreducible length differences in the inner geometry of the two descrip-

tions, and their evolution over time. That is, they manifest the *inconsistency* of the two kinematic descriptions.

Suppose we parameterize a trajectory model using a reference position vector $\mathbf{x}_R$ and a velocity vector $\mathbf{v}$, we can view this description as constituting a position–velocity vector $\mathbf{w}$ defined as

$$\mathbf{w} = \begin{bmatrix} \mathbf{x}_R \\ \mathbf{v} \end{bmatrix} \qquad (21)$$

We can then define the network position–velocity vector for a polyhedron as

$$\mathbf{w}^N = \begin{bmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \\ . \\ \mathbf{w}_n \end{bmatrix} \qquad (22)$$

which is analogous to definition of the network position vector which appeared in Eq. (14), but describes the temporal evolution of the polyhedron or network geometry, not just its instantaneous geometry. And by analogy with Eq. (14), we can represent the 12-parameter generalized Helmert transformation H12 thus:

$$\mathbf{w}_B^N = \text{H12}\left(\mathbf{w}_A^N \,|\, t_0, \mathbf{r}_0, \dot{\mathbf{t}}, \dot{\mathbf{r}}\right) = \text{H12}\left(\mathbf{w}_A^N \,|\, \mathbf{h}_{12}\right) \qquad (23)$$

where the 12 elements of the parameter vector $\mathbf{h}_{12}$ constitute the translation and rotations that relate frames A and B at some frame reference epoch, and the (constant) translation rates and rotation rates, that describe how frame rotation and translation evolve as a function of time. Eq. [23] constitutes the forward problem of transforming the network position–velocity vector (to change from frame A to frame B) given the parameter vector $\mathbf{h}_{12}$.

The problem of posterior imposition of a target RF, described in Sect. 6.3, is essentially an inverse problem in which we estimate the particular vector $\mathbf{h}_{12}$ that transforms the network position–velocity vector (for some sub-network of 'common' stations) expressed in an arbitrary frame so that the transformed vector matches as closely as possible the predictions of the target RF model for that sub-network. Having solved this inverse problem, then the forward equation (23) can be applied to the entire network time series and model.

The reader might wonder how this approach applies if we are using the SLTM or the ETM/ELTM instead of a CVM. In fact this generalization poses no difficulties since translating the RF in no way effects the amplitudes of oscillations or jumps or the amplitude coefficient A in the logarithmic transient formula (Eq. 9), and the same is true for infinitesimal rotations and infinitesimal rotation rates. The rotation vector $\mathbf{r}$, and the rotation rate vector $d\mathbf{r}/dt$, produce discernable changes only when multiplied by a very large number

such as the radius of the earth given in meters. So infinitesimal rotations affect geocentric position vectors, but they have essentially no effect on topocentric vectors (i.e. those 'rooted' at the geodetic station) such as oscillation amplitudes or jump amplitudes with typical magnitudes of 1 mm to several meters. A similar consideration applies in rate or velocity space.

The role of conventions in standard RFs. All geometrical RFs contain purely arbitrary aspects that are resolved by convention. (From the linear algebraic point of view, these conventions introduce constraints that resolve rank deficiencies or degeneracies that would otherwise occur in the matrix description of the RF and/or positioning of stations in that RF). These conventions in no way affect the inner geometry of any entities described in the frame. For example, if we frame a static geographical coordinate system (latitude, longitude) for the earth, the location of the prime meridian (i.e. the locus of zero longitude) is completely arbitrary. At one time this particular ambiguity was resolved by adopting the convention that the prime meridian passes through the Greenwich Observatory in London. Any formula for the distance between two points on an axi-symmetric earth refers only to the *difference* in the longitudes of these points, which is entirely unaffected by the conventional designation of the prime meridian. Since any Cartesian axis system [X, Y, Z] can be related to any similar axis system [X′, Y′, Z′] using six parameters or degrees of freedom, this is the number of degrees of freedom that must be resolved by convention when defining a 'standard' coordinate system or RF. The same is true for terrestrial RFs, such as ITRF2008, that operate continuously in time except that 12 degrees of freedom must be resolved by convention (6 associated with position space and 6 more for velocity or rate space). The fact that ITRF2008 is nominally a NNR frame, for example, is essentially a conventional means to resolve an otherwise arbitrary decision about the rotation rates of this particular frame. The inner geometry of a GPS coordinate time series is invariant with respect to the choice of reference frame, so it is entirely independent of the NNR assumption or any other conventional aspect of ITRF2008. In this sense we might paraphrase Einstein and state that in crustal motion geodesy, all geometry is either inner geometry (length) or convention.

In this paper we analyze GPS networks without direct reference to the other space geodetic techniques used to devise and define ITRF2008, and previous versions of ITRF. So, in this paper we are directly concerned only with the scale associated with GPS measurements. We acknowledge that our GPS metric or scale is inherited from ITRF2008, via the prior constraints applied during daily GPS data processing, and that this scale actually derives from a combination of geodetic techniques. Because of this, our RFs are not independent of ITRF2008, and we think of them as modifications of, or being derived from ITRF2008. We choose to depart from the ITRF2008 standard because we place more importance on the consistency of our geometrical description than in sharing this description with third parties (which requires us to adopt a common and 'named' RF), and because we place more importance on the consistency of our geometrical descriptions than on the convention-related properties of ITRF2008 such as its nominal NNR property. There are applications and contexts in which adopting a *standard* and *shared* reference frame, such as ITRF2008, is more important that any consideration related to modest improvements in the geometrical consistency of the adopted frame. But physics usually places its main emphasis on inner geometry rather than convention.

References

Altamimi Z, Collilieux X, Métivier L (2011) ITRF2008: an improved solution of the International Terrestrial Reference Frame. J Geodesy 85:457–473. doi:10.1007/s00190-011-0444-4

Becker J, Bevis M (2004) Love's problem. Geophys J Int 156:171–178

Bevis M, Brown A, Kendrick E (2012a) Devising stable geometrical reference frames for use in geodetic studies of vertical crustal motion. J Geodesy 87:311–321. doi:10.1007/s00190-012-0600-5

Bevis M et al (2012b) Bedrock displacements on Greenland driven by ice mass variations, climate cycles and climate change. Proc Natl Acad Sci. doi:10.1073/pnas.1204664109

Bevis M, Alsdorf D, Kendrick E, Fortes L, Forsberg B, Smalley R Jr, Becker J (2005) Seasonal fluctuations in the weight of the Amazon River system and Earth's elastic response. Geophys Res Lett 32:L16308. doi:10.1029/2005GL023491

Blewitt G, Lavallée D, Clark P, Nurutdinov K, (2001) A new global mode of earth deformation: seasonal cycle detected. Science 294:2342–2345

Bucknam R, Plafker CG, Sharp R (1978) Fault movement (afterslip) following the Guatemala earthquake of February 4, 1976. Geology 6:170–173

Collilieux X, van Dam T, Ray J, Coulot D, Métivier L, Altamimi Z (2012) Strategies to mitigate aliasing of loading signals while estimating GPS frame parameters. J Geodesy 86:1–14. doi:10.1007/s00190-011-0487-6

Deng J, Gurnis M, Kanamori H, Hauksson E (1998) Viscoelastic flow in the lower crust after the 1992 Landers. Science 282:1689–1692

Dietrich R, Dach R, Engelhardt G, Ihde J, Korth W, Kutterer H-J, Lindner K, Mayer M, Menge F, Miller H, Müller C, Niemeier W, Perlt J, Pohl M, Salbach H, Schenke H-W, Schöne T, Seeber G, Veit A, Völksen C (2001) ITRF coordinates and plate velocities from repeated GPS campaigns in Antarctica—an analysis based on different individual solutions. J Geodesy 74:756–766

Dong D, Fang P, Bock Y, Cheng MK, Miyazaki S (2002) Anatomy of apparent seasonal variations from GPS-derived site position time series. J Geophys Res 107(B4). doi:10.1029/2001JB000573

Fialko Y (2004) Evidence of fluid-filled upper crust from observations of postseismic deformation due to the 1992 Mw7.3 Landers earthquake. J Geophys Res 109:B08401. doi:10.1029/2004JB002985

Freed AM (2007) Afterslip (and only afterslip) following the 2004 Parkfield, California earthquake. Geophys Res Lett 34:L06312. doi:10.1029/2006GL029155

Freed AM, Bürgmann R, Herring T (2007) Far-reaching transient motions after Mojave earthquakes require broad mantle flow

beneath a strong crust. Geophys Res Lett 34:L19302. doi:10.1029/2007GL030959

Hanke M, Hansen PC (1993) Regularization methods for large-scale problems. Surv Math Ind 3:253–315

Hansen PC (1992) Analysis of discrete ill-posed problems by means of the L-curve. SIAM Rev 34:561–580

Heflin MB, Bertiger WI, Blewitt G, Freedman A, Hurst K, Lichten SM, Lindqwister U, Vigue Y, Webb F, Yunck T, Zumberge J (1992) Global geodesy using GPS without fiducial sites. Geophys Res Lett 19:131–134

Heki K (2001) Seasonal modulation of interseismic strain buildup in Northeastern Japan driven by snow loads. Science 293:89–92

Heki K, Miyazaki S, Tsuji H (1997) Silent fault slip following an inter-plate thrust earthquake at the Japan trench. Nature 386:595–598

Herring T, King R, McClusky S (2010) GAMIT and GLOBK reference manuals, release 10.4. Mass. Inst. of Technol., Cambridge

Holland PW, Welsch RE (1977) Robust regression using iteratively reweighted least-squares. Commun Stat Theory Methods 6(9):813–827

Jackson DD (1972) Interpretation of inaccurate, insufficient and inconsistent data. Geophys J R Astr Soc 28:97–109

Jiang Y, Dixon TH, Wdowinski S (2010) Accelerating uplift in the North Atlantic region and an indicator of ice loss. Nature Geosci 3:404–407

Jónsson S, Segall P, Pedersen R, Björnsson G (2003) Post-earthquake ground movements correlated to pore-pressure transients. Nature 424:179–183

Khan SA, Wahr J, Bevis M, Velicogna I, Kendrick E (2010) Spread of ice mass loss into northwest Greenland observed by GRACE and GPS. Geophys Res Lett 37:L06501. doi:10.1029/2010GL042460

Lawson C, Hanson R (1974) Solving least squares problems. Prentice-Hall, Englewood Cliffs

Lin YN, Sladen A, Ortega-Culaciati F, Simons M, Avouac JP, Fielding E, Brooks B, Bevis M, Genrich J, Reitbock A, Vigny C, Smalley R, Soquet A (2013) Coseismic and postseismic slip associated with the 2010 Maule earthquake, Chile: characterizing the Arauco Peninsula barrier effect. J Geophys Res 118:3142–3159. doi:10.1002/jgrb.50207

Mangiarotti S, Cazenave A, Souderaub L, Cretaux J (2001) Annual vertical crustal motions predicted from surface mass redistributions and observed by space geodesy. J Geophys Res 106:4277

Marone C (1998) Laboratory-derived friction laws and their application to seismic faulting. Annu Rev Earth Planet Sci 26:643–696

Marone C, Scholtz C, Bilham R (1991) On the mechanics of earthquake afterslip. J Geophys Res 96:8441–8452

Masterlark T, Wang HF (2002) Transient stress-coupling between the 1992 Landers and 1999 Hector Mine, California, earthquakes. Bull Seismol Soc Am 92:1470–1486

Peltzer G, Rosen P, Rogez F, Hudnut K (1998) Poroelastic rebound along the Landers 1992 earthquake surface rupture. J Geophys Res 103:30,131–30,145

Perfettini H, Avouac J-P (2007) Modeling afterslip and aftershocks following the 1992 Landers earthquake. J Geophys Res 112:B07409. doi:10.1029/2006JB004399

Perfettini H et al (2010) Seismic and aseismic slip on the Central Peru megathrust. Nature 465:78–81

Plasencia M (2007) Lithospheric characteristics and seismic sources in the Scotia Arc through waveform inversion. Ph.D. dissertation, University of Trieste. http://www.openstarts.units.it/dspace/handle/10077/2687

Playfair J (1802) Illustrations of the Huttonian theory of the earth. Cadell and Davies, London

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

Smalley R, Dalziel IWD, Bevis M, Kendrick E, Stamps DS, King E, Taylor FW, Lauría E, Zakrajsek A (2007) Kinematics of the Scotia Arc from GPS geodesy. Geophys Res Lett 34:L21308. doi:10.1029/2007GL031699

Smith SW, Wyss M (1968) Displacement on the san andreas fault subsequent to the 1966 Parkfield earthquake. Bull Seism Soc Am 58:1955–1973

Thatcher W, Rundle JB (1979) A model for the earthquake cycle in underthrust zones. J Geophys Res 84(B10):5540–5556. doi:10.1029/JB084iB10p05540

Van Dam T, Wahr J, Milly P, Shmalkin A, Blewitt G, Lavallée D, Larson K (2001) Crustal displacements due to continental water loading. Geophys Res Lett 28:651–654

Yang Q, Dixon T, Wdowinski S (2013) Annual variation of coastal uplift in Greenland as an indicator or variable and accelerating ice loss. Geochem Geophys Geosyst 14:1569–1589

Wahr J, Wyss M (1980) Interpretation of postseismic deformation with a viscoelastic relaxation model. J Geophys Res 85:6471–6477

Wang HF (2000) Theory of linear poroelasticity with applications to geomechanics and hydrogeology. Princeton Univ, Press, New Jersey

Wang CY, Wang CH, Manga M (2004) Coseismic release of water from mountains: evidence from the 1999 (Mw=7.5) Chi-Chi, Taiwan, earthquake. Geology 32:769–772. doi:10.1130/G20753.1

Wang K, Hu Y, Bevis M, Kendrick E, Smalley R, Vargas RB, Lauria E (2007) Crustal motion in the zone of the 1960 Chile earthquake: detangling earthquake-cycle deformation and forearc-sliver translation. Geochem Geophys Geosyst 8:Q10010. doi:10.1029/2007GC001721

Wang K, Hu Y, He J (2012) Deformation cycles of subduction earthquakes in a viscoelastic earth. Nature 484:327–332. doi:10.1038/nature11032