Semiautomatic fracture zone tracking

Paul Wessel1, Kara J. Matthews2, R. Dietmar Müller3, Aline Mazzoni3, Joanne M. Whittaker4, Robert Myhill5, and Michael T. Chandler1

1Department of Geology and Geophysics, School of Ocean and Earth Science and Technology, University of Hawaii, Múná, Honolulu, Hawai, USA, 2EarthByte Research Group, University of Sydney, Sydney, New South Wales, Australia, 3Department of Geophysics, University of São Paulo, São Paulo, Brazil, 4Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, Tasmania, Australia, 5Bayerisches Geoinstitut, University of Bayreuth, Bayreuth, Germany

Abstract Oceanic fracture zone traces are widely used in studies of seafloor morphology and plate kinematics. Satellite altimetry missions have resulted in high-resolution gravity maps in which all major fracture zones and other tectonic fabric can be identified, and numerous scientists have digitized such lineaments. We have initiated a community effort to maintain low-cost infrastructure that allows seafloor fabric lineaments to be stored, accessed, and updated. A key improvement over past efforts is our processing software (released as a GMT5 supplement) that allows for semiautomatic corrections to previously digitized fracture zone traces given improved gridded data sets. Here we report on our seafloor fabric processing tools, which complement our database of seafloor fabric lineations, magnetic anomaly identifications, and plate kinematic models.

1. Introduction

A fracture zone (FZ) is a linear tectonic scar caused by strike-slip motion between two oceanic plates along a transform fault [Wilson, 1965]. Its morphology reflects changes in plate motions that promote extension or compression along the fault [e.g., Caness et al., 1988; Kostens, 1987; Menard and Atwater, 1969; Pockalny et al., 1996] as well as vertical tectonics due to differential subsidence and thermal stresses [e.g., Hall and Gunis, 2005; Parmentier and Haxby, 1986; Sandwell, 1984; Wessel and Haxby, 1990]. Generations of scientists interested in plate kinematics have laboriously digitized trends of FZs. For example, the PLATES project at the University of Texas involved numerous scientists and resulted in a well-regarded database of seafloor tectonic fabric and isochrons [Gahagan et al., 1988]. Similarly, Cande et al. [1989] compiled seafloor tectonic fabric (fracture zones, trenches, ridges, etc.) and isochrons in map form; these were later digitized to yield a global database. These data have numerous uses in marine research such as constraining relative plate motion models [e.g., Cande et al., 1995], aiding in the characterization of seamounts [e.g., Kim and Wessel, 2011], and enabling analysis of non-FZ (i.e., abyssal hill) seafloor roughness [e.g., Whittaker et al., 2008], to name just a few. Old FZs are possible candidates for the formation of subduction zones [e.g., Hall et al., 2003], and both topography and geoid offsets across them yield information on the thermal evolution of the lithosphere [e.g., DeLaughter et al., 1999]. Since FZs reflect the history of transform fault offsets, the recorded variation in FZ spacing may constrain models of fracture zone formation and intraplate strain [e.g., Sandwell and Fialko, 2004]. Finally, plate tectonic reconstructions form the primary framework for understanding recent geologic history. FZ traces form an important component of such studies but here magnetic anomaly picks take precedence: conjugate isochron picks allow us to solve for a total reconstruction rotation and its uncertainties [e.g., Chang et al., 1990]. Global reconstruction models combined with FZ and isochron lineations enable the construction of global crustal age grids and derived products [e.g., Müller et al., 1997].

However, the usefulness of specific FZ traces decay with time as new data become available. For instance, the efforts of Gahagan et al. [1988] and Cande et al. [1989] relied on early processing of widely spaced, stacked 1-D SEASAT and GEOSAT profiles. Today, we have extensive 2-D coverage of higher accuracy that allows ever subtler fracture zone trends to be found [Sandwell et al., 2014]. Furthermore, digitized FZ traces are usually not updated, or perhaps only some FZs are redigitized, leading to a heterogeneous, partly global data set reflecting different eras of data qualities and personal styles and procedures for digitizing (i.e., did...
the author follow the FZ trough minimum, the zero crossing, or the steepest gradient normal to the trend?). Such inconsistencies are detrimental to systematic global studies as they blur any evidence of subtle variations, and furthermore, it is simply inefficient for all interested scientists to redigitize the same features (in slightly different ways). Finally, the sheer extent of such a global endeavor presents a formidable barrier to entry into exciting exploratory research.

Ideally, a digitized FZ trace should be used as a starting point when new data become available, with the retracing being done automatically. Such an approach would allow the process to be repeated quickly. The revised digital traces would later serve as new starting points for future improvements, and so on. Our approach builds on prior digitization efforts to address the long-term maintenance of plate-tectonic data via a community-driven effort accessible from a simple website. Our digitized seafloor fabric lineations have been added to our database for community-maintained FZ traces and magnetic anomaly picks via a dedicated website (www.soest.hawaii.edu/PT/GSFML). All lineaments are stored using version control allowing for complete documentation of the evolution of the database, and standard data formats are used for download. The website gives all users access to the latest and best estimates of seafloor fabric traces, a GMT5 [Wessel et al., 2013] supplement with our software for further data processing (Appendix A), and the complete documentation, including instructions for scientists who wish to contribute to this community resource. Here we focus on the methodology for digitizing and processing seafloor fabric lineations, in particular, FZ traces.

2. Identification of FZ Traces

Our approach to FZ trace digitization differs from previous attempts in that we recognize that the data sets used to infer FZ locations will improve with time. Thus, it is not the final trace of the FZ that is the fundamental information stored in the database but rather the terminal beginning and end points of each FZ plus intermediate points to define a smooth trace and any relevant metadata. From this information, moderate adjustment of FZ traces and estimates of their widths can be determined by guided tracking through the latest vertical gravity gradient (VGG) grids. Trusted users may supply new lineations, which may be revised as work is focused in certain geographic areas. Typically, these traces are digitized in Google Earth based on gridded KML overlays of free-air anomalies (FAA) and VGG and previously digitized FZ traces. We have completed a preliminary assessment of global FZs using Google Earth. Six types of lineaments have been identified, such as FZ, FZ of lower quality (FZLC), propagating rifts (PR), discordant zones (DZ), V-anomalies (VANOM), and unclassified V-anomalies (UNCV). For further description of these classifications, see Matthews et al. [2011].

3. Semiautomatic FZ Tracking

FZ expressions are varied and can change along the trace. Because the seafloor contains many lineated features as well as more isolated constructs (e.g., seamounts), a fully automated tracker is unlikely to succeed. We, therefore, implemented a semiautomatic FZ tracker, which analyzes each FZ candidate and determines the best modified trace given the VGG grid and the supplied metadata. We prefer vertical gravity gradients to free-air anomalies since they amplify short-wavelength signal associated with tectonic boundaries. We use GMT to sample the VGG grid along cross profiles that are equidistantly spaced along a digitized trace. Such VGG cross profiles show a range of behaviors between two distinctly different anomaly shapes. One end-member is the symmetric trough anomaly (Figure 1a) which is believed to reflect isostatically uncompensated crust of similar age [Kastens, 1987]. Such signals are common in the Atlantic Ocean where numerous FZs exist, many with relatively small age contrasts. The other end-member is the asymmetric edge-effect signal (Figure 1b), which is believed to reflect more-or-less isostatically compensated lateral changes in crustal age [Sandwell, 1984]; such anomalies are very common along large-offset FZs in the Pacific Ocean. Finally, plate motion changes are known to place segments of FZs under extension or compression. This may lead to an amplification of the peripheral bulges (Figure 1c). Theoretical models predict this range of behavior [e.g., Parmentier and Haxby, 1986; Sandwell, 1984].

For FZ tracking, we are less concerned with the actual mechanisms and more concerned with the typical shapes of anomalies. We want to match a wide variety of anomaly shapes reflecting different amounts of compression and spatial/temporal offsets. Based on observations, we have chosen three functions that are
combined to match a given cross profile using nonlinear least squares fitting. These are the Gaussian function \(G_0\), its first derivative \(G_1\), and the high-order component of the second derivative, i.e.,
\[
G_0 = e^{-\xi^2}, \quad G_1 = \xi e^{-\xi^2}, \quad G_2 = \xi^2 e^{-\xi^2},
\]
and visualized in Figures 1d–1f. Here \(\xi\) normalizes the distance \(x\) along a profile orthogonal to the FZ, i.e.,
\[
\xi = \frac{x - \Delta}{w},
\]
where \(\Delta\) is any offset of the FZ from its digitized location \((x = 0)\) and \(w\) is the trough’s half-width. Because of long-wavelength, low-amplitude regional signals in the VGG, we chose to model each cross profile as a linear trend perturbed by a sum of the components in equation (1), i.e.,
\[
z(x) = ax + b + A(mG_1 + (1-m)(uG_2 - G_0)),
\]
Specifically, we solve for the parameters \(\Delta\), (FZ offset) \(w\) (half width), \(A\) (amplitude), \(m\) (normalized asymmetry in 0–1 range), \(u\) (normalized compression in 0–1 range), and the aforementioned linear regional trend. The dominant components \(G_{0-2}\) depend on \(m\) and \(u\) and these may be used to quantify how and where a FZ signature changes. For example, finding significant \(G_2\) amplitudes (large \(u\)) may indicate a FZ segment under compression, whereas a strong \(G_1\) signal (large \(m\)) may indicate a relative large age contrast. This information may be useful in tectonic studies.

Theoretical modeling of large-offset Pacific FZs have shown that the actual location of the FZ matches the steepest slope in the geopotential fields [e.g., Sandwell, 1984], whereas for symmetric trough profiles, the FZ location coincides with the trough itself [e.g., Müller et al., 1991]. Given noise and other features present in

Figure 1. Examples of VGG anomalies across fracture zones and model components. (a) Symmetrical trough anomaly observed across the Pioneer FZ, (b) asymmetrical edge anomaly across the large-offset Udintsev fracture zone. (c) Symmetric trough anomaly with amplified bulges (Pioneer FZ). (d) Symmetric trough component modeled by a Gaussian function, and (e) asymmetry is modeled by its first derivative. (f) Amplified flank bulges are modeled with the non-Gaussian component of its second derivative.
the data, the only feature that can be traced with any degree of certainty is the continuity of the trough itself. Digitizing the steepest slope in noisy data is not feasible, even if that location may be more appropriate for some FZs. We decided to always digitize the trough location and let software determine if the FZ trace should be shifted toward the steepest slope. Identifying troughs was also the principle used by Nankivel [1997].

3.1. Synthetic Modeling
For testing, we created a synthetic FZ trace along the Equator but allowed its character to oscillate from symmetric to asymmetric along trace (i.e., we systematically varied the parameter $m$ from 0–1). We built two test VGG grids: one with constant amplitude and one where the amplitude decayed exponentially toward the east (to simulate signal degradation due to sedimentation and upward continuation). Finally, we considered cases with and without simulated, correlated noise. We imported the synthetic grid with constant amplitude into Google Earth and digitized it rather carelessly. In this case, we know the correct FZ location is at zero degrees latitude but because the FZ character meanders so does our digitized latitude (since we digitized the trough). This trace was then used to analyze the grid further. Figure 2 shows the four test VGG grids used; the yellow line indicates the actual FZ trace while the red line is our digitized trace.

Using GMT5’s new options in grdtrack [Wessel et al., 2013], we automatically generated numerous cross profiles normal to the digitized FZ trace at an equidistant spacing; such profiles are shown in Figure 3a for the case introduced in Figure 2c (correlated noise and constant FZ amplitude). The fzanalyzer module determines, on a profile-by-profile basis, the best model parameters using equation (3). Key model parameters are presented in plots of individual cross sections; here (Figure 3b) we just show two representative profiles ($#48$ and $#84$) to illustrate the process. Two model curves are plotted on top of the data (red circles): the best symmetric trough model (green, with $m$ forced to be 0) and the best overall model (blue, with optimal $m$), which may be asymmetrical. For each model, we find the optimal trough location (triangles) as well as steepest slope (blue circle) for both model. FZ widths $w$ are also resolved and shown on Figure 3b. The optimal FZ locations (data trough, red; symmetric model trough, green; and asymmetrical trough, blue) determined for each cross profile are plotted in map view (Figure 3a, smallest circles).

However, this profile-by-profile processing is merely an intermediate step that completely ignores the fact that a key characteristic of FZ traces is their along-track continuity. A second module (fzblender) was developed to filter the results along the FZ trace (Appendix B). We implemented a two-step filtering procedure where a robust filter (e.g., median) is run first (to exclude outliers overly influenced by seamounts or other
Figure 3. (a) Cross profiles every 10 km normal to digitized FZ trace (red). Larger green/yellow circles on Equator are best fit FZ locations, dashed orange lines reflect inferred FZ width. The analysis recovered the synthetic trace exactly. (b) Cross profiles show data (red circles) and best fitting symmetrical (green) and general (blue) models. Triangles indicate best FZ location for the models (red is observed trough); error bar represents inferred FZ width. Blue circle represent the FZ location based on steepest slope. Asymmetry ($m$) and compression ($u$) for both models are listed. (c) Correction $D$ to digitized fracture zone locations (in km) measured along the FZ. Here $D = 0$ represent original digitized trace. The three solutions (raw data, trough model, and blend model) generally agree. Large departures toward end of the trace (~900 km) reflect poor digitization of the original trace and detected by all three approaches.
non-FZ fabric), optionally followed by a Gaussian filter that results in smoothly varying trace parameters. We also assign a quality index to points making up the final FZ trace; however, these are preliminary and will require further calibration and testing (described later). The final trace is shown as green and yellow circles along the Equator in Figure 3a and connected by a cyan line. For this synthetic case we completely recovered the actual trace (Equator), even in the presence of realistic noise, without any along-track filtering. Here filtering (by stacking) would have strengthened the along-track correlations.

Figure 3c shows the result for $\Delta$, the across-track distance from the digitized FZ to the optimal FZ location, as a function of distance along trace. This plot shows we systematically digitized the trough too far south in addition to being way off track in the easternmost section. In map view (Figure 3a), the digitized, actual, and best traces are shown, demonstrating that we are able to improve on the raw digitized trace provided the data signal exceeds the noise.

Figure 4 repeats the exercise of Figure 3 using correlated noise superimposed on an exponentially decaying FZ signal (Figure 2d). Clearly, as data quality is degraded (i.e., amplitude drops below noise threshold) the modeling becomes unreliable (but so is the digitizing). Here we applied along-track filtering to the model results. The filtered trace (cyan) can be followed until $\sim 1^\circ$E–$2^\circ$E where the FZ signal loses all coherency. Thus, it is possible to recognize subtle FZ traces as long as they have continuity and amplitude comparable to the background noise level. This noise floor, reflecting seafloor roughness, varies geographically [e.g., Smith, 1998].

3.2. Modeling of Observed Traces

The purpose of this paper is to introduce the community resource and demonstrate how the tracker analysis works; it is not to present a global analysis of FZs. We have thus only tested the tracker on a few FZs to determine the strengths and weaknesses of the chosen approach.

Matthews et al. [2011] performed a preliminary digitization of the global set of seafloor lineated fabric. These data form the bulk of the lineations available from our web site but has been recently updated by new data from the Ellice Basin [Chandler et al., 2012] and revised further in light of improved altimetry [Sandwell et al., 2014]. The revised altimetry allowed deeply buried FZs to be extended further in various regions (Figure 5). While we refer you to Matthews et al. [2011] for details, we highlight their results for two different tectonic regions. The cross-profile technique was employed on two prominent fracture zones: an equatorial Atlantic FZ (Four North; east strand) and a South Pacific FZ (Udintsev FZ; west strand). The VGG data, profile spacing, and other particulars are outlined in Figure 6. For each area, representative profiles and best fitting models are shown in Figure 7. Of particular interest is the difference between the digitized trace obtained manually in Google Earth (using VGG image overlays) and the refinements offered by the software. Figures 7c and 7d
show the offsets \( \Delta \) determined for the two fracture zones, as well as the mean absolute deviation (MAD) of these offsets. For the entire Four North-east FZ, the MAD between the hand-digitized FZ trough and the raw VGG minima (red) is only 0.75 km, with a maximum offset of 2 km.

Similarly, for the Udintsev-west FZ, the MAD of their hand-digitized trough positions from raw VGG minima (red) is 2.9 km, and from the location of the maximum slope derived from the best fitting blend model (blue) is 2.28 km. The maximum offset from the VGG minima is 6 km. The maximum offset from the blend model-derived FZ trace exceeds 10 km for profiles 42–45, yet away from this part of the FZ the maximum offset remains less than 6 km (Figure 7d). A secondary seafloor structure may be overprinting the VGG signal for this segment of the Udintsev-west FZ, particularly as fZanalyzer placed the maximum slope position so far to the north onto older seafloor. As we hand-digitized the VGG trough it is more likely that the maximum slope position was south of our trace, toward the younger seafloor; this is the dominant trend for the rest of the FZ.

4. Discussion

Thus, preliminary digitization of the global seafloor fabric provides a reasonable starting point for further refinement. We expect that the next generation global gravity grids will benefit from the inclusion of CryoSat and other altimeter data. Over the next 3–5 years, three currently operating satellite altimeters (CryoSat, Jason-1, and Envisat) will provide a wealth of new marine gravity data [Sandwell et al., 2014]. Combined with earlier data from ERS-1 and Geosat, it is expected that accuracy of global marine gravity grids will improve by a factor or 2–4 depending on location [Sandwell et al., 2013]. This should enable significant improvement in FZ digitization and validation and lead to improved tectonic models.

While these preliminary results are encouraging, the approach we have chosen may not work in more complicated regions, such as very densely spaced fracture zones (e.g., Southwest Indian Ridge) and for features whose traces are not very linear (e.g., propagating rifts, some V-anomalies). However, our synthetic tests...
showed that the tracker loses continuity only when the noise level becomes quite high (Figure 2d). This occurs because stacking (via filtering) enhances the signal relative to noise. Because noise is correlated over short distances there is a fundamental problem when the actual trend of a feature is unclear to the digitizer. Obviously, no method will work if there is not even a faint trace to digitize. Our software is not searching unaided: it is actively guided by the initial trace the user digitized. This choice determines the alignment of cross profiles for stacking and means we improve the correlation along a trace once an expert user has done the initial identification of the trace. As more traces are subjected to the analysis and the thresholds required to determine a significant improvement are better established, we expect that the software will perform well enough to be a valuable assistant for experts in digitizing oceanic fracture zones, possibly even in more complicated regions. As Matthews et al. [2011] found, even careful digitization can be biased.

Figure 6. VGG location maps of FZs analyzed with our semiautomatic FZ tracking software. (a) Four North-east FZ in the Central Atlantic (African Plate), and (b) Udintsev-west FZ in the South Pacific (Pacific Plate). FZ traces are pink on yellow, yellow squares are hand-digitized points, cross profiles are light blue, with profiles used in Figure 7 in bold. Inset shows cross profile specifics, with the search range (blue) as the area in which the program will search for the best FZ location [from Matthews et al., 2011].
relative to the optimal trace. The methodology we have implemented offers objectivity, consistency, and repeatability of an approach that traditionally has been particularly subjective.

5. Conclusions
We maintain a community-driven web portal for scientists interested in seafloor fabric, magnetic lineations, and plate tectonic modeling. We have found that Google Earth is well suited for digitization of seafloor.
fabric and for inspecting these global data sets. GMT can produce KML overlays based on any global grid and these allow the digitizer to produce fairly robust initial traces [Matthews et al., 2011]. Such KML tiles for FAA, VGG, topography, potential tilt [Miller and Sing, 1994], and crustal ages [Müller et al., 2008] are accessible from our website. We anticipate that the portal will continue to evolve to better serve the user community as we gain more experience with the process and receive user feedback.

Appendix A: The GSFML Supplement to GMT

GMT5 allows anyone to develop new modules based on the GMT5 API [Wessel et al., 2013]. The tools discussed here constitute a new supplement (GSFML) available from the GSFML website. With GMT5 installed, users can build the supplement for their operating system and place the resulting shared library in the GMT plugin directory. The new modules can thus be accessed via the GMT executable like any other module. In addition, several bash scripts are available to simplify the plotting of the traces and cross profiles.

Appendix B: Blending the FZ Model Candidates

The fznanalyzer module returns four candidate FZ tracks; these are the optimal blend model (b) defined by (3), the best location based on identifying the trough and its width empirically (d), the optimal trough model (t) defined by (3) with \( m = 0 \), and the user’s original digitized trace (u). We use the fbrenderer module to produce a single, blended average of two or more of these candidates. Each candidate is weighted according to its calculated quality weight \( q \); in addition, we may supply custom weights \( w \) to each candidate (Default is \( 1 \)). There are several parameters for which we need a blended result (e.g., longitude, latitude, FZ width, and others); we refer to these here using the single parameter \( P \). The blended result is computed as

\[
P = \frac{\sum_i P_i w_i}{\sum_i w_i}, \quad i = b, d, t, u, \tag{B1}
\]

where the sum only extends over the filtered candidates the user has chosen to include. If the user’s digitized trace (u) is included, then, we must first estimate the quality weight to be assigned to this trace based on the average quality weights used by the chosen candidates, e.g.,

\[
q_u = 1 - \frac{\sum_i w_i}{n}, \tag{B2}
\]

where \( n \) is the number of traces chosen among \( \{b, d, t\} \). In other words, if the candidate quality indices are very low, the blend will be dominated by the user’s digitized trace.

References


