Transform faults and lithospheric structure: insights from numerical models and shipboard and geodetic observations

Author
Takeuchi, Christopher S.

Publication Date
2012

Peer reviewed|Thesis/dissertation
Transform Faults and Lithospheric Structure: Insights from Numerical Models and Shipboard and Geodetic Observations

A dissertation submitted in partial satisfaction of the requirements for the degree
Doctor of Philosophy

in
Earth Sciences

by
Christopher S. Takeuchi

Committee in charge:
Professor Yuri Fialko, Chair
Professor John Sclater
Professor Jeff Gee
Researcher Donna Blackman
Professor Xanthippi Markenscoff

2012
The dissertation of Christopher S. Takeuchi is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California, San Diego

2012
DEDICATION

To my family.
TABLE OF CONTENTS

Signature Page ......................................................... iii
Dedication ............................................................... iv
Table of Contents ..................................................... v
List of Figures .......................................................... viii
List of Tables ........................................................... xvi
Acknowledgements .................................................... xvii
Vita and Publications ................................................ xix
Abstract of the Dissertation .......................................... xx

Chapter 1  Introduction ................................................. 1

Chapter 2  Segment-scale and intrasegment lithospheric thickness and melt variations near the Andrew Bain megatransform fault and Marion hot spot: Southwest Indian Ridge, 25.5°E - 35°E ......................... 4
  2.1  Introduction ......................................................... 6
  2.2  Tectonic Setting and Regional Background ..................... 8
  2.3  Data Processing and Reduction .................................. 10
      2.3.1  Multibeam Bathymetry ...................................... 10
      2.3.2  Gravity Data ................................................... 10
      2.3.3  Magnetic Data .................................................. 12
      2.3.4  Earthquake Locations ......................................... 14
  2.4  Survey Area AP: 32°E-35°E ....................................... 14
      2.4.1  Morphology ..................................................... 14
      2.4.2  Mantle Bouguer Anomaly ...................................... 17
      2.4.3  Magnetization and Magnetic Anomalies ...................... 20
      2.4.4  Earthquake Locations ......................................... 20
  2.5  Survey Area DA: 25.5°E-28°E .................................... 20
      2.5.1  Morphology ..................................................... 20
      2.5.2  Mantle Bouguer Anomaly ...................................... 25
      2.5.3  Magnetization and Magnetic Anomalies ...................... 25
      2.5.4  Earthquake Locations ......................................... 25
  2.6  Ridge Segmentation and Plate Boundary Location .............. 26
  2.7  Interpretation of Observations .................................. 27
      2.7.1  Definitions and Cautions ...................................... 27
      2.7.2  Seafloor Texture as an Indicator of Melt Supply and Delivery ......................................................... 28
      2.7.3  Observations of Asymmetries in Melt Delivery and Lithospheric Thickness ................................................. 28
4.3 Results . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 82
  4.3.1 Surface velocity patterns . . . . . . . . . . . . . . . . . . . 82
4.4 Discussion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 85
  4.4.1 The effect of shear zones on postseismic deformation . . 85
  4.4.2 Postseismic deformation in the Mojave Desert . . . . . . 87
4.5 Conclusions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 93
4.6 Appendix A: Model Validation . . . . . . . . . . . . . . . . . . 94

Appendix A Abaqus Tips and Tricks . . . . . . . . . . . . . . . . . . . . . . . . . 97
  A.1 Restarts . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 97
  A.2 Outputting Fields in ASCII . . . . . . . . . . . . . . . . . . . . 100
  A.3 Defining boundary conditions in a user subroutine . . . . . . 102
  A.4 Defining self-updating boundary conditions . . . . . . . . . . 104
  A.5 Terminating an analysis based on a field value . . . . . . . . 108
  A.6 Implementation of linear viscoelastic rheologies with Prony series109
  A.7 Avoiding thermal runaway in coupled thermomechanical calcula-
     tions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 113
  A.8 Locking a fault without displacement/velocity boundary condi-
     tions or constraint equations . . . . . . . . . . . . . . . . . . 114

Bibliography . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 121
Figure 2.1: Shipboard bathymetry merged with the 2003 release of the GEBCO 1 min grid [Fisher and Goodwille, 1997] showing the SWIR between 24°E and 37°E. Du Toit, Andrew Bain, Marion, and Prince Edward transform faults are labeled. Black lines indicate track lines of cruise KN145L16. Red boxes show survey areas DA and AP (as labeled) of the present study. Purple lines indicate locations of magnetic anomaly profiles used in Figure 2.2. The profile at the top shows along-axis bathymetric profile of the SWIR between 24°E and 37°E. Transform domains bounded by dashed lines and labeled as follows: DT, Du Toit; AB, Andrew Bain; M, Marion; PE, Prince Edward. Ridge segments are labeled as defined in section 6. In the inset, thick black line denotes crest of SWIR; grey lines denote SWIR fracture zone traces. Thin grey lines indicate 4000 m bathymetric contour. A, AB, DCR, M, and MP denote the Astrid Fracture Zone, Andrew Bain transform fault, Del Cano Rise, Mozambique Escarpment, and Madagascar Plateau, respectively. Red star denotes location of Marion Island. Blue box shows area covered by Figure 2.1.

Figure 2.2: Identification of magnetic anomalies along profiles A, B, and C. Locations are indicated by purple lines in Figure 2.1. Reduced-to-the-pole magnetic anomalies are plotted in black, and synthetic anomalies are plotted in blue. Synthetic anomaly profiles are calculated from a two-dimensional block model based on the Cande and Kent [1995] geomagnetic time scale, with 13-16.2 km/Myr full spreading rates and 0%-25% spreading asymmetries. The forward model assumes a 1 km thick magnetic source layer with a magnetization intensity of 10 A/m for the Brunhes period and 2 A/m off axis. S and N refer to south and north.
Figure 2.3: (a) The 1 min gridded bathymetry of the SWIR between 32°E and 35°E, survey area AP. Bathymetric contour interval is 500 m. Black dots denote locations of earthquake solutions [Engdahl et al., 1998], with focal mechanisms plotted where available. Green dots indicate locations of dredge hauls of Mahoney et al. [1992], both of which recovered normal incompatible-depleted mid-ocean ridge basalts. (b) The 1 min gridded mantle Bouguer anomaly. Shipboard bathymetry has been extended beyond the survey area with the GEBCO 1 min grid [Fisher and Goodwille, 1997 for MBA calculations. MBA contours are in black, interval is 5 mGal, and heavy black line indicates zero MBA contour. White contours are 1000 m bathymetric contours. (c) Reduced-to-the-pole magnetic anomalies. Positive amplitudes to the west-northwest. Purple symbols indicate the center of positive polarity intervals 1 (0-0.78 Ma, diamond), 2a (2.581-3.58 Ma, triangle), 3a (5.894-6.567 Ma, circle), and 4 (7.432-8.072 Ma, inverted triangle) [Cande and Kent, 1995. (d) The 1 min gridded magnetization intensity. Magnetization contours are in black, interval is 1.25 A/m, and heavy black line indicates zero magnetization contour. White contours are 1000 m bathymetric contours. Purple symbols indicate magnetic anomaly picks as in Figure 2.3c. Red lines in Figures 2.3a, 2.3b, and 2.3d indicate the location of the plate boundary. Transform fault locations are labeled in all plots as follows: AB, Andrew Bain; M, Marion; PE, Prince Edward.

Figure 2.4: Along-axis bathymetry, MBA, and magnetization data for survey area AP. Fields are sampled along the plate boundary, plotted in red in Figures 2.3a, 2.3b, and 2.3d. Magnetization is only plotted at crossing points between the plate boundary and ship tracks due to the lack of along-axis ship track coverage within this survey area. Transform fault locations are labeled in all plots as follows: AB, Andrew Bain; M, Marion; PE, Prince Edward. (a) Bathymetry (black) and MBA (red) between the Andrew Bain and Marion transform faults. (b) Bathymetry (black) and MBA (red) between the Marion and Prince Edward transform faults. (c) Bathymetry (black) and magnetization (blue) between the Andrew Bain and Marion transform faults. (d) Bathymetry (black) and magnetization (blue) between the Marion and Prince Edward transform faults.

Figure 2.5: (a) High-resolution shipboard bathymetry (50 m grid interval) of the SWIR between 32°E and 35°E, survey area AP, showing volcanic terrain and abyssal hill fabric. Illumination is from the west-northwest. Grey areas are between-track data gaps. Transform fault locations are labeled in Figure 2.5a as follows: AB, Andrew Bain; M, Marion; PE, Prince Edward. (b) Southern flank and (c) northern flank of the ridge axis between the Andrew Bain and Marion transform faults. (d) Northern flank and (e) southern flank of the ridge axis between the Marion and Prince Edward transform faults.
Figure 2.6: (a) High-resolution shipboard bathymetry (50 m grid interval) of the SWIR between 32°E and 35°E, survey area AP, showing the positive gradient in volcanic cone density from east to west in and around the rift valleys. Flat-topped mounds interpreted to be volcanic cones are circled in white. Illumination is from the south-southwest. Grey areas are between-track data gaps. Transform fault locations are labeled in Figure 6a as follows: AB, Andrew Bain; M, Marion; PE, Prince Edward. (b) Rift valley of segment MP-1. Note the abundance of volcanic cones and narrow volcanic ridge in the west, immediately adjacent to the Marion transform fault. In the east, shipboard bathymetric coverage does not extend all the way to the Prince Edward transform fault (located just off the eastern end of the plotted area; it connects to the ridge axis at the northeastern limit of the plate boundary plotted in red); however, the eastern limit of bathymetric coverage shows relatively fewer volcanic cones in and around the rift valley, at a greater distance from the Prince Edward transform fault. (c) Rift valley of segment AM-1, again showing a greater abundance of volcanic cones within a given distance of the Andrew Bain transform fault than within the same distance from the Marion transform fault.

Figure 2.7: (a) The 1 min gridded bathymetry of the SWIR between 25.5°E and 28°E, survey area DA. Bathymetric contour interval is 500 m. Black dots denote locations of earthquake solutions [Engdahl et al., 1998], with focal mechanisms plotted where available. (b) The 1 min gridded mantle Bouguer anomaly. Shipboard bathymetry has been extended beyond the survey area with the GEBCO 1 min grid [Fisher and Goodwille, 1997] for MBA calculations. MBA contours are in black, interval is 5 mGal, and heavy black line indicates zero MBA contour. White contours are 1000 m bathymetric contours. (c) Reduced-to-the-pole magnetic anomalies. Positive amplitudes to the west-northwest. Purple symbols indicate the center of positive polarity intervals 1 (0-0.78 Ma, diamond), 2a (2.581-3.58 Ma, triangle), 3a (5.894-6.567 Ma, circle), and 4 (7.432-8.072 Ma, inverted triangle) [Cande and Kent, 1995]. (d) The 1 min gridded magnetization intensity. Magnetization contours are in black, interval is 1.25 A/m, and heavy black line indicates zero magnetization contour. White contours are 1000 m bathymetric contours. Purple symbols indicate magnetic anomaly picks as in Figure 2.7c. Red lines indicate the location of the plate boundary. Dashed red line indicates that the location of the plate boundary in this area is unknown. Transform fault locations are labeled in all plots as follows: DT, Du Toit; AB, Andrew Bain.
Figure 2.8: Along-axis bathymetry, MBA, and magnetization data for survey area DA. Fields are sampled along the plate boundary, plotted in red in Figure 2.7. Transform fault locations are labeled in both plots as follows: DT, Du Toit; AB, Andrew Bain. (a) Bathymetry (black) and MBA (red) between the Du Toit and Andrew Bain transform faults. (b) Bathymetry (black) and magnetization (blue) for the same area. Gaps in data in both plots indicate that the location of the plate boundary in this area is unknown.

Figure 2.9: (a) High-resolution shipboard bathymetry (50 m grid interval) of the SWIR between 25.5°E and 28°E. Illumination is from the west-northwest. Grey areas are between-track data gaps. Transform fault locations are labeled in Figure 2.9a as follows: DT, Du Toit; AB, Andrew Bain. Red line delineates areas interpreted to be smooth nonvolcanic seafloor. Green lines indicate areas showing characteristics of corrugated seafloor. (b) Abyssal hill fabric on the northern flank of the western rift valley. (c) Hummocky volcanic terrain on the Du Toit inside corner high. Southwestern area shows weak spreading-parallel lineations. (d) Smooth, broad ridge with no evidence of volcanic cones or fault scarps south of the eastern rift valley. (e) Hummocky volcanic terrain on the Andrew Bain inside corner. (f) Lobate high on the northern flank of the eastern rift valley showing clear spreading-parallel lineaments on its western flank.

Figure 2.10: The 1 min GEBCO bathymetry of the SWIR between 14°E and 42°E [Fisher and Goodwille, 1997]. Red arrows indicate vectors of ridge migration in the hot spot reference frame, calculated from the absolute plate motion model HS3-Nuvel1a [Gripp and Gordon, 2002]. Arrow lengths are proportional to rate of migration; migration for segment AM-1 is 7 mm/yr.

Figure 2.11: (a) Segment length versus axial relief, (b) axial relief versus MBA amplitude, (c) segment length versus MBA amplitude, and (d) mean axial depth versus MBA amplitude for seven areas along the SWIR, listed from west to east: the 15.5°E-25°E supersegment containing 14 individual accretionary segments [Grindlay et al., 1998] (circles), the western rift valley of segment DA-1 (current study, diamond), segments AM-1 and MP-1 (current study, stars), segment 27 (50.47°E) [Sauter et al., 2001; Mendel et al., 2003] (square), segments 20-22 (54.1°E-56.67°E) [Sauter et al., 2001; Mendel et al., 2003] (triangle, one record for all three segments), and segments 14 and 11 (~61.4°E and 63.9°E, respectively) [Cannat et al., 1999] (inverted triangles). Values plotted for segments 20-22 represent the central values of the ranges listed in Table 1.
Figure 2.12: Cartoon (not to scale) showing our proposed mechanism for the along-axis distribution of melt and resulting observed asymmetries, which are strongly controlled by the transform edge effect. Passively upwelling mantle material (solid red arrows) undergoes partial melting and releases melt (dashed red arrows) to the top of the asthenosphere. This melt is then focused away from the bounding transform faults, up the sloping base of the lithosphere (dashed white arrows) [Magde and Sparks, 1997]. However, melt production is suppressed and melt focusing is enhanced by the strong cooling and thickening of the lithosphere towards the long-offset Andrew Bain transform to the east. The majority of what little melt is produced in the mantle near the Andrew Bain is focused toward the west or trapped in the thick lithosphere; very limited melt reaches shallow depths to create a crust primarily composed of a thin basaltic cap. Melt delivery increases to the west through westward melt focusing and increased melt production (along-axis density of dashed red arrows in the lithosphere indicates the relative amount of melt released from the upper asthenosphere that becomes crust).

Figure 3.1: Model geometry.

Figure 3.2: Temperature anomaly produced by viscous dissipation for thermomechanically coupled power law models versus distance from fault and depth for (a) dry and (b) wet rheologies. Dashed black line indicates the elastic-ductile transition (12 km depth); solid black line indicates the prescribed Moho (30 km depth). (c and d) Magnitude of the temperature anomaly versus distance from fault at a depth of 20 km for several epochs during the thermal fault spin-up.

Figure 3.3: Predicted fault-parallel shear stress versus earthquake cycle number for models (a) M2000, (b) M20, (c) DNK, (d) WNK, (e) DCK, (f) WCK, (g) DCS, and (h) WCS. Stresses are plotted at the end of the interseismic period of each numbered cycle. Curves indicate shear stress at x = 0 km, y = 6 km (solid line) at the midpoint of the upper crust on the fault; x = 0 km, y = 12 km (dotted line) at the prescribed brittle-ductile transition on the fault; x = 0 km, y = 18 km (dashed line) in the lower crust 1 km beneath the fault tip; and x = 142 km, y = 6 km (dash-dotted line) well off-fault in the middle of the upper crust, where x is the fault-perpendicular coordinate and y is depth. z = 0 for all.
Figure 3.4: Predicted fault-parallel shear stress versus distance from fault and depth. Model configurations are the same as in Figure 3. Stress is cycle-invariant for (a, b, and e-h) Maxwell and coupled power law models and plotted after 1000 cycles for (c and d) non-coupled power law models. Stress is plotted at the end of an interseismic period, immediately preceding the next earthquake. For all plots, z = 0. Dashed black line indicates the prescribed elastic-ductile transition (12 km depth); solid black line indicates the prescribed Moho (30 km depth).

Figure 3.5: Predicted fault-parallel shear strain rate versus distance from fault and depth. Model configurations are the same as in Figure 3.3. Strain rate in each plot is plotted at the same time as the corresponding plot in Figure 3.4. For all plots, z = 0. Dashed black line indicates the prescribed elastic-ductile transition (12 km depth); solid black line indicates the prescribed Moho (30 km depth).

Figure 3.6: Predicted fault-parallel surface velocity versus distance from fault for several time steps following an earthquake. Model configurations are the same as in Figure 3.3. To illustrate cycle invariance, we plot velocities separated by 50 cycles (solid and dashed lines); the difference is invisible in most cases.

Figure 3.7: Geodetic observations of shear strain rate following the 1906 San Francisco earthquake [Kenner and Segall, 2003] (symbols) and model predictions (solid lines). Shades of gray denote different time periods. Late interseismic strain rates have been subtracted from model predictions.

Figure 3.8: Predicted surface heat flow (lines) for dynamic coupled power law models versus distance from fault, compared with borehole heat flow observations [Benfield, 1947; Fulton et al., 2004; Lachenbruch and Sass, 1980; Sass et al., 1971; Sass et al., 1982] (symbols) from the San Andreas fault.

Figure 3.9: Predicted fault-parallel shear stress versus depth at x = 0 (on the fault plane, dashed lines) and x = 46 (off fault, solid lines), at the end of the interseismic period for dynamic coupled power law models. Gray lines indicate DCS model, black lines indicate WCS model.

Figure 3.10: Shear stress supported by the lithosphere as a function of loading rate for dynamic coupled power law models. Stress is plotted at the end of the interseismic period.

Figure 3.11: Temperature anomaly and predicted fault-parallel velocity at 20 km depth versus distance from fault for dynamic coupled power law models. Both fields are plotted at the end of the interseismic period.

Figure 4.1: Model geometry.
Figure 4.2: Map view of postseismic surface velocities in mm/yr, 6 months after rupture for models a) AFT (afterslip), b) MS (Maxwell), c) BI (biviscous), d) WFB (wet power law, Freed and Bürgmann [2004] thermal regime), e) DNS (dry power law, no shear zone), f) WNS (wet power law, no shear zone), g) DSZ (dry power law, shear zone), h) WSZ (wet power law, shear zone). Colors indicate vertical velocity, arrows indicate horizontal velocity (scales are indicated in each panel). Heavy black line denotes the fault plane.

Figure 4.3: Postseismic surface velocities in mm/yr, 2 years after rupture. Panels as indicated in Figure 2 caption. Colors indicate vertical velocity, arrows indicate horizontal velocity. Heavy black line denotes the fault plane.

Figure 4.4: Postseismic surface velocities in mm/yr, 10 years after rupture. Panels as indicated in Figure 2 caption. Colors indicate vertical velocity, arrows indicate horizontal velocity. Heavy black line denotes the fault plane.

Figure 4.5: Postseismic surface velocities in mm/yr, 50 years after rupture. Panels as indicated in Figure 2 caption. Colors indicate vertical velocity, arrows indicate horizontal velocity. Heavy black line denotes the fault plane.

Figure 4.6: 13 years of surface displacement in mm following the simulated Hector Mine rupture. Colors indicate vertical displacement, arrows indicate horizontal displacement. SCIGN continuous GPS stations with time-series in Figures 4.7 and 4.8 shown by black circles, with station identifiers as indicated. Reference station (for horizontal displacements) RDMT denoted by green circle. Heavy black line denotes the simulated Hector Mine rupture. Black star denotes the epicenter of the Hector Mine earthquake.

Figure 4.7: Panels a and b: horizontal displacements in mm following the Hector Mine rupture. Displacements referred to continuous GPS station RDMT (green dot) and shown by black (modeled) and blue (observed) arrows, respectively. Black asterisks denote GPS station locations, black star denotes the epicenter of the Hector Mine earthquake, and red line denotes the simulated Hector Mine rupture. Other panels show timeseries of observed (grey stars) and modeled (solid black line = biviscous model, dashed black line = power law model) horizontal displacement, referred to station RDMT, for stations indicated with black circles in panels a and b. Dotted black line denotes displacement due to modeled station secular velocity referred to simulated station RDMT.
Figure 4.8: Panels a and b: vertical displacement in mm following the Hector Mine rupture. Modeled and observed displacements shown by colors and vertical black arrows, respectively. Black asterisks denote GPS station locations, black star denotes the epicenter of the Hector Mine earthquake, and heavy black line denotes the simulated Hector Mine rupture. Other panels show timeseries of observed (grey stars) and modeled (solid black line = biviscous model, dashed black line = power law model) vertical displacement plotted for stations indicated with black circles in panels a and b.

Figure 4.9: Observed (panel a) and simulated (panel b = afterslip + poroelastic; panel c = biviscous; panel d = power law) ERS-2 satellite line-of-sight displacement in mm. Black line in panel a indicates true Hector Mine surface rupture, and black star the Hector Mine epicenter; heavy black lines in panels b-d indicate simulated Hector Mine rupture.

Figure 4.10: 20 years of surface displacement for benchmarking strike-slip fault simulations incorporating linear Maxwell rheology, as predicted by four codes: a) Abaqus; b) Relax; c) VISCO1D; and d) PSGRN. Colors indicate vertical displacement, arrows indicate horizontal displacement. Heavy black line denotes the fault plane.

Figure 4.11: Surface velocities 6 months (a and b) and 2 years (c and d) following coseismic rupture or benchmarking strike-slip fault simulations incorporating biviscous Burgers rheology, as predicted by Abaqus (a and c) and VISCO1D (b and d). Colors indicate vertical velocity, arrows indicate horizontal velocity. Heavy black line denotes the fault plane.
### LIST OF TABLES

Table 2.1: Segment Length, Mean Axial Depth, Along-Axis Relief, Across-Axis Relief, Along-Axis MBA Variation, and Regional Basalt Sodium Content for Seven Areas Along the SWIR .......................... 34

Table 3.1: Rheological Properties of Rocks From Laboratory Measurements ... 47
Table 3.2: Model Configurations ........................................... 52

Table 4.1: Laboratory-derived material properties of rocks. .................... 77
Table 4.2: Model Configurations ........................................... 79
ACKNOWLEDGEMENTS

My first thanks go to the two advisors I’ve had while at SIO, John Sclater and Yuri Fialko. I thank John for introducing me to the grand adventure of shipboard data collection, which was the primary reason I came to San Diego in the first place. Additionally, I thank John for his boundless enthusiasm and guidance as I plowed my way through the ins and outs of marine data processing and the composition of my first ever published paper. And while John got things started for me, I thank Yuri for bringing them home. Yuri’s reliability, thoroughness, and ability to simultaneously see both the forest and the trees were of immeasurable assistance, as was his steadfast serenity while I rode the unending fluctuations of confidence and composure that all graduate students must endure. I am also thankful to Jeff Gee and Donna Blackman for their practical advice on many different subjects, particularly those related to keeping my eyes forward and my feet moving. I am grateful to Steve Constable for providing me with additional opportunities to go to sea. I am obliged to all of the teachers I had while at SIO, who laid the foundation for everything I’ve achieved as a graduate student.

Special thanks must be given to my fellow students in IGPP, who changed from colleagues to friends in a matter of moments. I thank Danny, Robin, and Leah for providing steady evidence that progress is being made even if you’re not aware of it. My classmates, Dave, Joseph, and Sylvain all helped keep me afloat when I felt like I couldn’t paddle anymore, and I may never have made it out of my first year without the constant support of Deb and Lynn. And the rabblers, Andy, Brendan, Brent, and Samer, with whom I’ve had many a discussion about the relative merits of IPA versus tripel- I thank them for providing me with balance and non-snobbery.

I thank my parents and my brother, who provided me with everything a growing boy could want or need to make his way in the world. And lastly, I thank Asia, who showed me that true love is the greatest thing in the world, except for a nice MLT.

Chapter 2, in full, is a reformatted version of material as published in Geochemistry, Geophysics, Geosystems (Takeuchi, C.S., J.G. Sclater, N.R. Grindlay, J.A. Madsen, and C. Rommevaux-Jestin (2010), Segment-scale and intrasegment lithospheric thickness and melt variations near the Andrew Bain megatransform fault and Marion hot spot: Southwest Indian Ridge, 25.5°E-35°E, Geochem. Geophys. Geosys., 11(7), Q07012, doi:10.1029/2010GC003054), reprinted with permission. I was the primary investigator and author of this paper.
Chapter 3, in full, is a reformatted version of material as published in the Journal of Geophysical Research (Takeuchi, C.S. and Y. Fialko (2010), Dynamic models of interseismic deformation and stress transfer from plate motion to continental transform faults, J. Geophys. Res., 117, B05403, doi:10.1029/2011JB009056), reprinted with permission. I was the primary investigator and author of this paper.

Chapter 4, in full, is currently being prepared for submission to Geophysical Journal International. I am the primary investigator and author of this paper.
VITA

2002 B. S. in Astrophysical and Planetary Sciences *with distinction*,
University of Colorado, Boulder

2005-2012 Graduate Research Assistant, University of California, San Diego

2011 Teaching Assistant, University of California, San Diego

2012 Ph.D. in Earth Sciences, University of California, San Diego

PUBLICATIONS

(2010), Segment-scale and intrasegment lithospheric thickness and melt variations near
the Andrew Bain megatransform fault and Marion hot spot: Southwest Indian Ridge,

Takeuchi, C.S. and Y. Fialko (2010), Dynamic models of interseismic deformation and
stress transfer from plate motion to continental transform faults, *J. Geophys. Res.*, 117,
ABSTRACT OF THE DISSERTATION

Transform Faults and Lithospheric Structure: Insights from Numerical Models and Shipboard and Geodetic Observations

by

Christopher S. Takeuchi

Doctor of Philosophy in Earth Sciences

University of California, San Diego, 2012

Professor Yuri Fialko, Chair

In this dissertation, I study the influence of transform faults on the structure and deformation of the lithosphere, using shipboard and geodetic observations as well as numerical experiments. I use marine topography, gravity, and magnetics to examine the effects of the large age-offset Andrew Bain transform fault on accretionary processes within two adjacent segments of the Southwest Indian Ridge. I infer from morphology, high gravity, and low magnetization that the extremely cold and thick lithosphere associated with the Andrew Bain strongly suppresses melt production and crustal emplacement to the west of the transform fault. These effects are counteracted by enhanced temperature and melt production near the Marion Hotspot, east of the transform fault.

I use numerical models to study the development of lithospheric shear zones underneath continental transform faults (e.g. the San Andreas Fault in California), with a particular focus on thermomechanical coupling and shear heating produced by long-term fault slip. I find that these processes may give rise to long-lived localized shear zones, and that such shear zones may in part control the magnitude of stress in the lithosphere. Localized ductile shear participates in both interseismic loading and postseismic relaxation, and predictions of models including shear zones are within observational constraints provided by geodetic and surface heat flow data. I numerically investigate the effects of shear zones on three-dimensional postseismic deformation. I conclude that the presence of a thermally-activated shear zone minimally impacts postseismic deformation, and that thermomechanical coupling alone is unable to generate sufficient localization for postseismic relaxation within a ductile shear zone to kinematically resemble that by aseismic fault creep (afterslip). I find that the current record geodetic observations of postseismic
deformation do not provide robust discriminating power between candidate linear and power-law rheologies for the sub-Mojave Desert mantle, but longer observations may potentially allow such discrimination.
Chapter 1

Introduction

The recognition that motion could be 'transformed' between pairs of mid-ocean ridges and/or subduction zones [Wilson, 1965] was a key observation in the development of the theory of plate tectonics. Transform faults were subsequently recognized as a third type of plate boundary, responsible for accommodating divergence at mid-ocean ridges and convergence at subduction zones without requiring relative motion between these other plate boundary faults. Further surveying of the world’s oceans revealed the detailed structure of many transform faults on the various oceanic plate boundaries. On the fast-spreading East Pacific Rise, transform faults are well-approximated by vertical planar strike-slip faults oriented parallel to relative plate motion, with deformation zones a few kilometers wide [Fox and Gallo, 1984]. However, at slower spreading rates (e.g. the Mid Atlantic Ridge), the structure of transform faults may be much more complex due to the fact that that the lithosphere through which the fault must cut is thick relative to that for a similarly long fault on a faster spreading ridge [Ligi et al., 2002]. At long transform age offsets (i.e. the transform fault length multiplied by the relative slip rate), the structure of a transform plate boundary may diverge significantly from a vertical planar fault approximation, with multi-fault systems, transtensional and/or transpressional features, curved fault patterns, and secondary non-transform offsets within the plate boundary zone. In addition, the cold, thick lithosphere juxtaposed against a mid-ocean ridge segment by a long-age offset transform fault may also influence the formation of faults and accretionary processes within the ridge’s plate boundary zone [Fox and Gallo, 1984].

With broad, multi-fault deformation zones, long age-offset oceanic transform faults are akin to continental transform faults, which must cut through thick continental lithosphere and are composed of broadly distributed, complex systems of interlocking faults. While oceanic transform faults are buried beneath several kilometers of ocean at
great distances from centers of population, continental transform faults may lie in close proximity to or even within major cities, e.g. the San Andreas Fault in California and the North Anatolian Fault in Turkey. Though they present significant societal hazard, the locations of continental transform faults also allow for greater ease and frequency of study. In particular, geodetic observations (e.g. Global Positioning System and/or synthetic aperture radar data) allow for an earthquake to be used as the loading source for an \textit{in situ} rock deformation experiment. In the period between earthquakes (the interseismic period), a fault is loaded by regional tectonic motion and/or by localized frictionally-stable aseismic fault creep, which produces elevated strain rates near the fault. The loading triggers an earthquake (the coseismic period). The stress perturbation produced by the coseismic rupture generates a relaxation response, which leads to spatiotemporally evolving postseismic surface deformation that may be recorded by GPS stations and satellite surveying. Comparison of the observations to the predictions of models of postseismic relaxation may allow for constraints on the properties of the lithosphere in the vicinity of the given rupture. There are three primary postseismic relaxation mechanisms: a) poroelastic relaxation, in which the coseismic rupture induces a pressure gradient in the crust, leading to fluid flow from regions of high pressure to low pressure; b) afterslip, in which aseismic fault creep on frictionally-stable sections of the fault relax the coseismic stress, and c) viscoelastic relaxation, in which stresses are relaxed by bulk flow in the ductile lithosphere. The degree to which each mechanism contributes to postseismic relaxation may vary from location to location and earthquake to earthquake. However, the contrast between the endmember aseismic fault creep/afterslip and tectonic/viscoelastic flow models of fault loading and postseismic relaxation illuminates a critical question of whether continental plate motion is generally accommodated by highly localized shear or broadly distributed flow in the lithosphere.

In this dissertation, I investigate the influence of transform faults on the structure and deformation of the lithosphere. In Chapter 2, I use shipboard observations of topography, gravity, and magnetics to explore 25.5°-35°E on the ultra-slow spreading Southwest Indian Ridge (SWIR), a previously unmapped section of the global mid-ocean ridge system. In this area, two accretionary ridge segments are separated by the Andrew Bain transform fault, which has the longest age offset of any transform plate boundary in the world. I use observations of morphology, seafloor texture, and gravity and magnetic anomalies to infer that the extreme cooling effects of the lithosphere juxtaposed against
the end of a SWIR segment by the Andrew Bain severely restrict the processes of melt generation in the subaxial lithosphere and crustal production. In Chapters 3 and 4, I turn to continental transform faults. Geological and seismic observations indicate that discrete localized shear zones exist below major strike-slip faults, as postulated by the aseismic fault creep/afterslip model of fault loading and postseismic relaxation. The nature of such shear zones is not well-constrained, nor are the processes that lead to their initiation and sustenance. In Chapter 3, I use two-dimensional numerical experiments incorporating laboratory-derived rheological laws and material parameters to investigate whether steady-state localized shear zones may develop in the ductile lithosphere in response to shear heating driven by fault slip over geologic time. In Chapter 4, I use the results of Chapter 3 to examine whether a viscoelastic relaxation model incorporating a thermally-activated shear zone may predict a three-dimensional postseismic surface deformation pattern similar to that predicted by an afterslip relaxation model. In Appendix A, I present the techniques I used to develop the finite element models used in Chapters 3 and 4, in the hopes that other researchers may avoid reinventing the wheel.
Chapter 2

Segment-scale and intrasegment lithospheric thickness and melt variations near the Andrew Bain megatransform fault and Marion hot spot:
Southwest Indian Ridge, 25.5°E - 35°E

Christopher S. Takeuchi, John G. Sclater, Nancy R. Grindlay, John Madsen, and Céline Rommevaux-Jestin

Abstract. We analyze bathymetric, gravimetric, and magnetic data collected on cruise KN145L16 between 25.5°E and 35°E on the ultraslow spreading Southwest Indian Ridge, where the 750 km long Andrew Bain transform domain separates two accretionary segments to the northeast from a single segment to the southwest. Similar along-axis asymmetries in seafloor texture, rift valley curvature, magnetic anomaly amplitude, magnetization intensity, and mantle Bouguer anomaly (MBA) amplitude within all three segments suggest that a single mechanism may produce variable intrasegment lithospheric thickness and melt delivery. However, closer analysis reveals that a single mechanism is unlikely. In the northeast, MBA lows, shallow axial depths, and large abyssal hills indicate that the Marion hot spot enhances the melt supply to the segments. We argue that along-axis asthenospheric flow from the hot spot, dammed by major transform faults, produces the inferred asymmetries in lithospheric thickness and melt delivery. In the southwest, strong rift valley curvature and nonvolcanic seafloor near the Andrew Bain transform fault indicate very thick subaxial lithosphere at the end of the single segment. We suggest that cold lithosphere adjacent to the eastern end of the ridge axis cools and thickens the subaxial lithosphere, suppresses melt production, and focuses melt to the west. This limits the amount of melt emplaced at shallow levels near
the transform fault. Our analysis suggests that the Andrew Bain divides a high melt supply region to the northeast from an intermediate to low melt supply region to the southwest. Thus, this transform fault represents not only a major topographic feature but also a major melt supply boundary on the Southwest Indian Ridge.
2.1 Introduction

Extensive knowledge of the processes by which oceanic lithosphere and crust is formed has been gained from studies of the fast spreading East Pacific Rise [e.g. Macdonald et al., 1984; Sempéré and Macdonald, 1986; Madsen et al., 1990] and the slow spreading Mid-Atlantic Ridge (MAR) [e.g. Kuo and Forsyth, 1988; Lin et al., 1990; Blackman and Forsyth, 1991; Rommevaux et al., 1994; Detrick et al., 1995]. Recent studies on the Southwest Indian Ridge (SWIR) and Gakkel Ridge have led to the definition of a new spreading regime with unique lithospheric and crustal characteristics: the ultraslow (<20 mm/yr full rate) spreading ridge [Dick et al., 2003; Cannat et al., 2003].

Topography, gravity, and geochemical sampling have all provided evidence for the highly variable nature of melt supply on the SWIR at both regional (>200 km) and segment length (<200 km) scales. The regional melt supply of a spreading ridge is largely controlled by mantle temperature, composition, and upwelling velocity, the last of which is directly tied to the spreading rate [Reid and Jackson, 1981]. At spreading rates below 20 mm/yr (i.e., at ultraslow spreading ridges), conductive cooling strongly inhibits melting and thus crustal production [Bown and White, 1991; White et al., 2001; Dick et al., 2003], though this effect is likely tempered at such rates due to the focusing and acceleration of mantle upwelling [White et al., 2001]. Regional ridge obliquity also influences melt supply; increased obliquity decreases the effective spreading rate (the component of the spreading velocity orthogonal to the ridge) and thus also the mantle upwelling rate and melt supply [Dick et al., 2003; Montési and Behn, 2007].

Regional axial depth and the sodium content of basaltic crust (and other geochemical indicators of partial melting) may be interpreted to reflect variable regional melt supply to a spreading ridge [Klein and Langmuir, 1987; Cannat et al., 2008]. Mean axial depth, if assumed to be the product of isostatically balanced crust, and basaltic sodium content both decrease as melt supply increases [Klein and Langmuir, 1987]. These observations have been invoked to infer a broad-scale decrease in melt supply from the shallow, low sodium content SWIR axis between the Prince Edward and Discovery II transform faults (35.55°E - 41.53°E) to the deep, high sodium content axis near the Rodrigues Triple Junction [Cannat et al., 2008].

While regional melt supply is largely a function of upwelling mantle velocity, temperature, and composition, the delivery of melt to individual segments is heavily influenced by a number of shallow processes. The most prominent of these processes in-
volves discontinuities, both transform and nontransform, that offset the ridge axis. The juxtaposition of cold lithosphere across discontinuities and against the ends of ridge segments cools and thickens the subaxial lithosphere, which can suppress melting processes underneath the ridge [Fox and Gallo, 1984; Magde and Sparks, 1997]. Melt can also be strongly focused away from the discontinuities along the sloping base of the lithosphere, with longer discontinuities focusing melt more intensely [Magde et al., 1997; Magde and Sparks, 1997]. The juxtaposition also produces a weld between the lithospheric sections on either side of the discontinuity. The weld produces a rotation of the horizontal stress field at the RTI due to the interplay between normal stresses at the ridge axis and shear stresses along the adjacent transform fault [Fujita and Sleep, 1978; Phipps Morgan and Parmentier, 1984]. As a result, rift valley normal faults, which initiate orthogonally to the axis of the maximum tensile stress, form at angles increasingly oblique to the spreading direction as the transform fault is approached, which causes the rift valley to curve into the transform fault.

On the SWIR, segment obliquity also affects lithospheric thickness and the amount of melt delivered to the ridge axis. Thick lithosphere under oblique segments can prevent fracturing, which keeps melt from reaching the seafloor [Montési and Behn, 2007]; this melt is either trapped in the lithosphere or migrates along axis to an orthogonal segment [Cannat et al., 2003; Cannat et al., 2008]. This segment-to-segment melt migration represents a fundamental difference in melt delivery variations between the SWIR and MAR; while MAR segments all receive a volume of melt roughly equal to the regional average, SWIR segment melt delivery is highly spatially variable, with some segments receiving more melt than the regional average, some receiving less [Cannat et al., 2008].

In February and March 1996, investigators conducted the first geophysical survey of the SWIR between 15°E and 35°E [Grindlay et al., 1996]. Bathymetric, gravimetric, and magnetic data were collected over roughly 1500 km of the SWIR on cruise KN145L16. Grindlay et al. [1998] studied the bathymetric and gravimetric segmentation of the nearly linear, transform fault-free supersegment between 15°E and 25°E. Dulaney [2002] investigated the axial morphology of the entire survey region. Sclater et al. [2005] examined the Andrew Bain megatransform fault, which lies between 28°E and 32°E. In this study, we investigate the SWIR immediately adjacent to the major topographic boundary of the Andrew Bain fracture zone. We incorporate two surveys on either side of the Andrew Bain, between 25.5°E and 28°E, and between 32°E and 35°E. This
data fills in the last major section of the SWIR not previously reported (see references of Cannat et al. [2008]). We use the morphology, mantle Bouguer anomaly, magnetic anomalies, and magnetization intensity to establish the segmentation patterns of the two areas surveyed. We utilize the observations to infer mechanisms that may produce variations in lithospheric thickness and the volume of melt delivered to and emplaced within segments. We also use our data to establish the Andrew Bain transform fault as a major melt supply boundary.

2.2 Tectonic Setting and Regional Background

The SWIR is the divergent plate boundary between the Antarctic and Nubian/Somalian plates. The SWIR has a total length of 7700 km; it is offset, almost exclusively in a left-stepping sense, by transform faults and higher-order discontinuities. The ridge has an average full spreading rate of $\sim 12–18$ mm/yr [Caress and Chayes, 1999], the slowest rate of any readily accessible mid-ocean ridge axis. Only the ice-covered Gakkel Ridge spreads more slowly, diverging at a full rate of 8-13 mm/yr [Cochran et al., 2003; Dick et al., 2003].

Between 25$^\circ$E and 35$^\circ$E, the SWIR is partitioned four times by the right-lateral, left-stepping Du Toit, Andrew Bain, Marion, and Prince Edward transform faults. These four transform faults offset the SWIR over 1100 km [Fisher and Goodwille, 1997] (Figure 2.1). This is the only section all along the SWIR where the offset is achieved almost exclusively by transform faults. At other locations on the SWIR, the northward offset of the ridge axis is achieved by transform faults, oblique spreading [Mendel et al., 1992; Dick et al., 2003], nontransform discontinuities [Grindlay et al., 1998; Sauter et al., 2001], or any combination thereof [Baines et al., 2007].

With a length of $\sim 750$ km and a maximum width of 120 km, the Andrew Bain transform domain dominates the morphology of this section of the SWIR. The Andrew Bain transform domain is second in offset length to only the Romanche transform fault on the central Mid-Atlantic Ridge [Sclater et al., 2005]. Ligi et al. [2002] classified the Andrew Bain as a megatransform; Caress and Chayes [1999] and Lemaux et al. [2002] have suggested that the plate boundary between the Nubian and Somalian plates intersects the SWIR in the vicinity of the Andrew Bain.
Figure 2.1: Shipboard bathymetry merged with the 2003 release of the GEBCO 1 min grid [Fisher and Goodwille, 1997] showing the SWIR between 24°E and 37°E. Du Toit, Andrew Bain, Marion, and Prince Edward transform faults are labeled. Black lines indicate track lines of cruise KN145L16. Red boxes show survey areas DA and AP (as labeled) of the present study. Purple lines indicate locations of magnetic anomaly profiles used in Figure 2.2. The profile at the top shows along-axis bathymetric profile of the SWIR between 24°E and 37°E. Transform domains bounded by dashed lines and labeled as follows: DT, Du Toit; AB, Andrew Bain; M, Marion; PE, Prince Edward. Ridge segments are labeled as defined in section 6. In the inset, thick black line denotes crest of SWIR; grey lines denote SWIR fracture zone traces. Thin grey lines indicate 4000 m bathymetric contour. A, AB, DCR, M, and MP denote the Astrid Fracture Zone, Andrew Bain transform fault, Del Cano Rise, Mozambique Escarpment, and Madagascar Plateau, respectively. Red star denotes location of Marion Island. Blue box shows area covered by Figure 2.1.
2.3 Data Processing and Reduction

Cruise KN145L16 included two surveys between 25.5°E and 35°E, at either end of the longoffset Andrew Bain transform; the southern covered the SWIR between 25.5°E and 28°E and the northern between 32°E and 35°E. These surveys examined the axis of the SWIR and the flanking seafloor using ~105-120 km long across-axis profiles spaced 8 km apart. In addition, the southern survey included three along-axis tracks. We denote the survey areas according to the transform faults that bound them; we thus refer to the southern survey area as survey area DA, and the northern survey area as AP.

2.3.1 Multibeam Bathymetry

Bathymetric data were acquired with a Sea-Beam 2112 multibeam system, operating with a source frequency of 12 kHz. The system allowed athwartships swath coverage of 90°-120°; swath widths were routinely 11-12 km. Navigation was derived from continuous Global Positioning System fixes. The spacing of ship tracks and swath widths allowed for 90%-100% ensonification of the seafloor. The multibeam data was ping edited to remove spurious returns; noise was significant during periods of bad weather. The outer 3-5 beams on either side of each ping were usually edited out due to their high noise level.

The ping-edited bathymetry data were gridded using the open source swath sonar data processing code MB-System [Caress and Chayes, 1996]. Using MB-System, we calculated the value of each grid cell by taking a Gaussian weighted average of the values at surrounding data points. The bathymetry data were gridded with an interval of 50 m to allow for a thorough analysis of seafloor texture. We utilized a 2-D thin plate spline interpolation to fill in the gaps between individual pings, which were frequent due to the rough weather common to the area. However, we did not fill in gaps between ship tracks in order to be certain that our seafloor texture analysis was controlled by data rather than interpolation artifacts.

2.3.2 Gravity Data

Gravity data were collected using a Bell BGM-3 gravimeter. The data were merged with the GPS navigation and center beam bathymetry to provide synchronized gravity and bathymetry data. The Eötvös correction was then applied, and a reference
gravity value was removed to obtain free-air anomalies. This reference gravity value was calculated using

\[
ref_{\text{gravity}} = 978049.0 \times (1.0 + 0.0052884\sin^2(lat) - 0.0000059\sin^2(2lat))
\]

where \(lat\) was the ship's latitude in degrees. The data were edited to remove spurious points and points recorded during ship turns, smoothed by the application of a 13-point (~2.3 min of time) running mean filter, and linearly interpolated to evenly spaced 1 min (time) values. To obtain estimates of the gravimeter drift, measurements were taken at base stations in Durban and Cape Town, South Africa (the departure and arrival ports, respectively), and then referenced to the values recorded on the ship. A DC correction of -2.1 mGal was also applied to the data to account for the difference between the gravity measured at the base station in Durban and that recorded on the ship prior to departure. Crossover error analysis was completed using the GMT supplement xsystem software [Wessel and Smith, 1998]. The gravity data had crossover errors with a mean of 2.0 mGal and a standard deviation of 2.6 mGal.

We calculated mantle Bouguer anomalies (MBAs) following the method of Prince and Forsyth [1988]. We added additional bathymetric spline interpolation to fill in data gaps between ship tracks. We then extended the shipboard bathymetry data beyond the two ridge axis survey areas using the 2003 release of the GEBCO 1 min grid of Fisher and Goodwille [1997]; the composite bathymetry data was also gridded with a 1 min interval to match the resolution limit of the GEBCO grid. We utilized the Parker [1973] Taylor series method to convert the topography to the gravitational signal due to interfaces between layers of seawater (1 g cm\(^{-3}\)), upper crust (2 km thick, 2.4 g cm\(^{-3}\)), lower crust (4 km thick, 2.7 g cm\(^{-3}\)), and mantle (3.3 g cm\(^{-3}\)). In order to minimize edge effects caused by the periodicity assumed in the Fast Fourier Transform (FFT) calculations of the Parker [1973] method, we mirrored the topography at the edges of our domain. The calculated gravity contributions of the interface topography were then gridded at an interval of 1 min using the GMT software adjustable curvature algorithm surface [Wessel and Smith, 1998] with a tension factor of 0.25. We then sampled the gravity contribution grids at locations along ship tracks at which free air gravity data existed; the sampled values were then removed from the free air gravity, producing values of MBA along ship tracks. The MBA data were gridded at 1 min resolution using surface [Wessel and Smith, 1998] with a tension factor of 0.25.
2.3.3 Magnetic Data

Magnetic data were collected with a Geometrics G-886 Marine Magnetometer. The International Geomagnetic Reference Field (IGRF 6th Generation [Langel, 1992]; the magnetic reduction was completed immediately following the cruise) was removed from the raw magnetic field data to produce magnetic anomalies. The magnetometer logging interval changed during the cruise, so the magnetic anomalies were interpolated using a weighted mean to produce evenly spaced 1 min (time) values. Crossover error analysis was again completed using xsystem [Wessel and Smith, 1998]; crossover errors had a mean of 0.5 and a standard deviation of 25.8 nT.

We utilized two-dimensional forward modeling in order to identify the magnetic anomaly reversal pattern. We reduced the observed magnetic anomalies to the pole to remove the effects of skewness due to latitude. We ran a forward model using a 1 km thick magnetized layer, bounded on top by the bathymetry, with a magnetization distribution based on a square wave function constructed with the Cande and Kent [1995] geomagnetic time scale. We assumed a 10 A/m magnetization for the Brunhes normal polarity period (0-0.78 Ma) and a 2 A/m intensity off axis. The best fit between the synthetic and reduced-to-the-pole anomaly patterns was achieved using a 13-16.2 km/Myr spreading rate with 0%-25% spreading asymmetries (Figure 2.2).

We performed a three-dimensional inversion of the observed magnetic anomaly data using the two-dimensional method of Parker and Huestis [1974], extended to three dimensions by Macdonald et al. [1980]. We used a 1 km-thick magnetic source layer, bounded on top by the same composite bathymetry grid used in the MBA calculations. A cosine taper band-pass filter with a low-cut taper from 150 to 300 km and a high-cut taper from 4 to 8 km assured convergence of the inversion. As with the MBA calculation, we mirrored the topography at the edges to minimize edge effects caused by the assumed periodicity of the FFT calculations in the inversion. We assumed a crustal magnetization parallel to a geocentric axial dipole. The magnetization data were gridded at a resolution of 1 min to match the resolution limit of the composite bathymetry grid.

We did not add any magnetic annihilator to the result of our inversion for a number of reasons. As shown by Tivey and Tucholke [1998], the water depth filter and narrow polarity reversal spacing preclude the ability of sea surface magnetic observations to resolve well the amplitude of magnetization anomalies on seafloor created at slow spreading rates. As such, for slow spreading rates, balancing positive and negative
Figure 2.2: Identification of magnetic anomalies along profiles A, B, and C. Locations are indicated by purple lines in Figure 2.1. Reduced-to-the-pole magnetic anomalies are plotted in black, and synthetic anomalies are plotted in blue. Synthetic anomaly profiles are calculated from a two-dimensional block model based on the Cande and Kent [1995] geomagnetic time scale, with 13-16.2 km/Myr full spreading rates and 0%-25% spreading asymmetries. The forward model assumes a 1 km thick magnetic source layer with a magnetization intensity of 10 A/m for the Brunhes period and 2 A/m off axis. S and N refer to south and north.
amplitudes across reversal boundaries is not an appropriate technique for calculating the amount of annihilator to add to a magnetization solution. However, the technique proposed by Tivey and Tucholke [1998] of adding annihilator so that the magnetization zero crossings on each track match those of the magnetic anomalies was also fruitless, as no constant multiple of the annihilator could successfully accomplish the matching of the zero crossings within each individual ridge segment. As such, we will focus on the relative amplitude variations of the magnetization instead of the absolute values in our analysis.

2.3.4 Earthquake Locations

To further understand the plate boundary geometry, we supplement our shipboard data with earthquakes from the relocated epicenter database of Engdahl et al. [1998] (EHB). The approximately 100,000 global earthquakes of this database are a combined set of International Seismological Centre (ISC) and National Earthquake Information Center (NEIC); they cover the time period from January 1964 through March 2010. The initial ISC and NEIC hypocentral location estimates of this data set were improved using arrival times from phases PKiKP, PKPd, pP, pwP, and sP, in addition to S and P phases, in the location procedure. Where available, we have plotted focal mechanisms for these earthquakes as calculated by the Global Centroid Moment Tensor (GCMT) catalog. Within our survey areas, locations for EHB events before January 1999 have an average epicentral location uncertainty of $13.0 \pm 3.7$ km; uncertainties for later earthquakes have not been published. However, assuming that the post-January 1999 events have a similar order of magnitude of uncertainty as the pre-January 1999 events, the uncertainties preclude our ability to correlate earthquake locations to specific morphological features. We thus only utilize them to determine the general location of active spreading.

2.4 Survey Area AP: $32^\circ$E-$35^\circ$E

2.4.1 Morphology

Rift Valley

Survey area AP shows two very similar morphological sectors bounded by three transform faults (Figure 2.3). The southwesterly of these two sectors, lying between
Figure 2.3: (a) The 1 min gridded bathymetry of the SWIR between 32°E and 35°E, survey area AP. Bathymetric contour interval is 500 m. Black dots denote locations of earthquake solutions [Engdahl et al., 1998], with focal mechanisms plotted where available. Green dots indicate locations of dredge hauls of Mahoney et al. [1992], both of which recovered normal incompatible-depleted mid-ocean ridge basalts. (b) The 1 min gridded mantle Bouguer anomaly. Shipboard bathymetry has been extended beyond the survey area with the GEBCO 1 min grid [Fisher and Goodwille, 1997 for MBA calculations. MBA contours are in black, interval is 5 mGal, and heavy black line indicates zero MBA contour. White contours are 1000 m bathymetric contours. (c) Reduced-to-the-pole magnetic anomalies. Positive amplitudes to the west-northwest. Purple symbols indicate the center of positive polarity intervals 1 (0-0.78 Ma, diamond), 2a (2.581-3.58 Ma, triangle), 3a (5.894-6.567 Ma, circle), and 4 (7.432-8.072 Ma, inverted triangle) [Cande and Kent, 1995. (d) The 1 min gridded magnetization intensity. Magnetization contours are in black, interval is 1.25 A/m, and heavy black line indicates zero magnetization contour. White contours are 1000 m bathymetric contours. Purple symbols indicate magnetic anomaly picks as in Figure 2.3c. Red lines in Figures 2.3a, 2.3b, and 2.3d indicate the location of the plate boundary. Transform fault locations are labeled in all plots as follows: AB, Andrew Bain; M, Marion; PE, Prince Edward.
Figure 2.4: Along-axis bathymetry, MBA, and magnetization data for survey area AP. Fields are sampled along the plate boundary, plotted in red in Figures 2.3a, 2.3b, and 2.3d. Magnetization is only plotted at crossing points between the plate boundary and ship tracks due to the lack of along-axis ship track coverage within this survey area. Transform fault locations are labeled in all plots as follows: AB, Andrew Bain; M, Marion; PE, Prince Edward. (a) Bathymetry (black) and MBA (red) between the Andrew Bain and Marion transform faults. (b) Bathymetry (black) and MBA (red) between the Marion and Prince Edward transform faults. (c) Bathymetry (black) and magnetization (blue) between the Andrew Bain and Marion transform faults. (d) Bathymetry (black) and magnetization (blue) between the Marion and Prince Edward transform faults.

The northeastern sector, lying between the Marion and Prince Edward transform faults, is very similar morphologically to the southwestern sector (Figure 2.3a). The ~85 km long rift valley in the northwestern sector very nearly parallels that of the southwestern sector. It strikes ~102° at its western end and intersects nearly orthogonally with...
the Marion transform fault. The valley begins curving northward ∼25 km away from the Prince Edward transform fault on its eastern end. As with the southwestern sector, the valley is filled by a large central axial high. However, this central high is shallower than that of the southwestern sector, rising to a depth of ∼1700 m (Figure 2.4b). The rift valley also has greater relief than that of the southwestern sector, deepening by ∼2800 m to the west and by ∼3000 m to the east. Again, the bathymetric profile is asymmetric, with (1) the summit of the axial high located slightly (∼8 km) west of the center point between the bounding transform faults and (2) overall greater depths on the eastern end of the valley. The mean depth of the rift valley is ∼3500 m. A small ridge lies at the far western end (Figure 2.4b).

**Seafloor Texture**

Rift valley-parallel abyssal hill fabric dominates the topography flanking both rift valleys of survey area AP (Figure 2.5). These abyssal hills are quite large, rising 500-1000 m above the surrounding seafloor. The seafloor is composed exclusively of volcanic terrain, in the form of pebbly textures composed of small (<500 m diameter), roughly circular constructs akin to hummocky pillow lava flows seen observed with TOBI side scan imagery [Sauter et al., 2002; Sauter et al., 2004b] (hereafter referred to as hummocky terrain), flat-topped mounds, and ridge axis-parallel scarps. Along-axis asymmetries in the density of mounds are observed, with more found near the western bounding transform faults than near the eastern transform faults (Figure 2.6).

### 2.4.2 Mantle Bouguer Anomaly

The gravity data show a broad, roughly circular MBA low centered near the central axial high of each rift valley of survey area AP (Figure 2.3b). The MBA lows are somewhat elongated perpendicular to these valleys. Between the Andrew Bain and Marion transform faults, the minimum of -35 mGal is located near the summit of the central axial high. The MBA increases westward by 22 mGal and eastward by 30 mGal (Figure 2.4a). Between the Marion and Prince Edward transform faults, the gravity data show a broad elliptical -47 mGal MBA low centered near the summit of the central axial high (Figure 2.4b). The MBA increases westward by 30 mGal and eastward by 38 mGal. The MBA of the northeastern sector thus has greater (i.e., more negative) amplitude and greater along-axis variation than that of the southwestern sector. Both MBA lows
Figure 2.5: (a) High-resolution shipboard bathymetry (50 m grid interval) of the SWIR between 32°E and 35°E, survey area AP, showing volcanic terrain and abyssal hill fabric. Illumination is from the west-northwest. Grey areas are between-track data gaps. Transform fault locations are labeled in Figure 2.5a as follows: AB, Andrew Bain; M, Marion; PE, Prince Edward. (b) Southern flank and (c) northern flank of the ridge axis between the Andrew Bain and Marion transform faults. (d) Northern flank and (e) southern flank of the ridge axis between the Marion and Prince Edward transform faults.
Figure 2.6: (a) High-resolution shipboard bathymetry (50 m grid interval) of the SWIR between 32°E and 35°E, survey area AP, showing the positive gradient in volcanic cone density from east to west in and around the rift valleys. Flat-topped mounds interpreted to be volcanic cones are circled in white. Illumination is from the south-southwest. Grey areas are between-track data gaps. Transform fault locations are labeled in Figure 6a as follows: AB, Andrew Bain; M, Marion; PE, Prince Edward. (b) Rift valley of segment MP-1. Note the abundance of volcanic cones and narrow volcanic ridge in the west, immediately adjacent to the Marion transform fault. In the east, shipboard bathymetric coverage does not extend all the way to the Prince Edward transform fault (located just off the eastern end of the plotted area; it connects to the ridge axis at the northeastern limit of the plate boundary plotted in red); however, the eastern limit of bathymetric coverage shows relatively fewer volcanic cones in and around the rift valley, at a greater distance from the Prince Edward transform fault. (c) Rift valley of segment AM-1, again showing a greater abundance of volcanic cones within a given distance of the Andrew Bain transform fault than within the same distance from the Marion transform fault.
have asymmetric profiles inversely correlated to the topography of the two sectors, with lower overall amplitudes in the west relative to the east.

2.4.3 Magnetization and Magnetic Anomalies

Survey area AP has strong central magnetic anomaly highs and central anomaly magnetization highs (CAMHs [Klitgord, 1976]) over each of two rift valleys, with the northeastern sector showing a greater peak and average amplitude (Figures 2.3c and 2.3d). Both the magnetic anomalies and magnetization distribution generally show identifiable magnetic lineations corresponding to alternating polarity intervals. However, asymmetries exist that parallel those observed in the bathymetry and gravity data. The magnetic lineation patterns are clear within 5 km of the western bounding transform faults of both the southwestern and northeastern sectors. However, on the eastern end of each sector, the lineation patterns disappear and amplitudes become flat near where the rift valleys begin to curve into the eastern bounding transform faults (Figures 2.3c and 2.3d). Following this pattern, the intensities of the CAMHs of both sectors decrease by $\sim 9 \text{ A/m} \sim 25 \text{ km}$ from the eastern bounding transform faults (Figures 2.4c and 2.4d).

2.4.4 Earthquake Locations

Of the 27 relocated earthquake epicenters of Engdahl et al. [1998], 22 are located within or near the Andrew Bain, Marion, or Prince Edward transform domains (Figure 2.3a). Focal mechanisms in these locations indicate the expected dextral strike-slip faulting, except for a single normal faulting mechanism in the Marion transform domain. Five events are located within or near the rift valley between the Andrew Bain and Marion transform faults, including one normal faulting event.

2.5 Survey Area DA: 25.5°E-28°E

2.5.1 Morphology

Rift Valley

Survey area DA is characterized by two overlapping rift valleys (Figure 2.7a). The western of the two (hereafter referred to as the western rift valley) is a low-relief (500 m), $\sim 70 \text{ km}$ long valley striking $\sim 116^\circ$. At its western end, it begins to curve south $\sim 20 \text{ km}$ from its intersection with the northern end of the Du Toit transform fault. The
Figure 2.7: (a) The 1 min gridded bathymetry of the SWIR between 25.5°E and 28°E, survey area DA. Bathymetric contour interval is 500 m. Black dots denote locations of earthquake solutions [Engdahl et al., 1998], with focal mechanisms plotted where available. (b) The 1 min gridded mantle Bouguer anomaly. Shipboard bathymetry has been extended beyond the survey area with the GEBCO 1 min grid [Fisher and Goodwille, 1997] for MBA calculations. MBA contours are in black, interval is 5 mGal, and heavy black line indicates zero MBA contour. White contours are 1000 m bathymetric contours. (c) Reduced-to-the-pole magnetic anomalies. Positive amplitudes to the west-northwest. Purple symbols indicate the center of positive polarity intervals 1 (0-0.78 Ma, diamond), 2a (2.581-3.58 Ma, triangle), 3a (5.894-6.567 Ma, circle), and 4 (7.432-8.072 Ma, inverted triangle) [Cande and Kent, 1995]. (d) The 1 min gridded magnetization intensity. Magnetization contours are in black, interval is 1.25 A/m, and heavy black line indicates zero magnetization contour. White contours are 1000 m bathymetric contours. Purple symbols indicate magnetic anomaly picks as in Figure 2.7c. Red lines indicate the location of the plate boundary. Dashed red line indicates that the location of the plate boundary in this area is unknown. Transform fault locations are labeled in all plots as follows: DT, Du Toit; AB, Andrew Bain.
Figure 2.8: Along-axis bathymetry, MBA, and magnetization data for survey area DA. Fields are sampled along the plate boundary, plotted in red in Figure 2.7. Transform fault locations are labeled in both plots as follows: DT, Du Toit; AB, Andrew Bain. (a) Bathymetry (black) and MBA (red) between the Du Toit and Andrew Bain transform faults. (b) Bathymetry (black) and magnetization (blue) for the same area. Gaps in data in both plots indicate that the location of the plate boundary in this area is unknown.

The valley is deepest in the nodal basin of the ridge-transform intersection (RTI), reaching a depth of \( \sim 5900 \) m (Figure 8a). It shoals by \( \sim 2500 \) m to the shallowest point of the rift valley, and then deepens by \( \sim 1100 \) m toward an elliptical basin in the east. The valley shoals and narrows east of this elliptical basin before disappearing. The mean depth of the western rift valley between the nodal basin in the west and the elliptical basin in the east is \( \sim 4500 \) m. No single distinct axial volcanic ridge (AVR) can be identified within the rift valley (Figure 9). Rather, a shallow, broad, low-relief swell topped by a series of small ridges and troughs lies in the center of the valley.

The western rift valley overlaps with another rift valley to the south (the eastern rift valley) by \( \sim 22 \) km (Figure 2.7a). A \( \sim 1200 \) m high ridge sits between the overlapping valleys. The eastern rift valley lacks an arch-shaped profile; rather, it deepens from west to east (Figure 2.8a), reaching a maximum depth of \( \sim 6550 \) m at the Andrew Bain RTI. The eastern rift valley is \( \sim 65 \) km long, has a mean depth of 4850 m and no discernible AVR (Figure 2.9). The valley is roughly linear south of the valley overlap, striking \( \sim 92^\circ \) before curving northward \( \sim 40 \) km from the Andrew Bain transform fault. The eastern rift valley deepens from west to east by \( \sim 1800 \) m (Figure 2.8a).
Seafloor Texture

The seafloor texture of survey area DA is widely variable (Figure 2.9). The primary observation is the dearth of clearly identifiable volcanic cones anywhere in the area. Rather, the most abundant texture is that of hummocky volcanic terrain. The seafloor flanking the western rift valley is predominantly this hummocky terrain, both in the organized spreading-perpendicular abyssal hill fabric north of the rift valley (Figure 2.9b) and the Du Toit inside corner high (ICH) south of the valley (Figure 2.9c). The Du Toit ICH shows clear evidence of widespread faulting, with a number of scarps oriented subparallel to the western rift valley. Weak spreading-parallel lineaments can also been seen on the Du Toit inside corner high (the southwestern area of Figure 2.9c).

North of the overlapping rift valleys, the seafloor shows some ridge-parallel abyssal hill fabric and hummocky terrain; however, this area is dominated by a circular, domal bathymetric high centered at 27°03'E, 52°24'S, which stands ~1 km above the surrounding seafloor (Figure 2.9). South of the valley overlap, the seafloor is covered by a series of long, broad, roughly parallel ridge-trough structures; the eastern rift valley is the northernmost trough in this configuration. This ridge-trough system is in stark contrast to the relatively gentle abyssal hill fabric north of the rift valley in the western area; the ridges have relief of up to 1 km relative to the intervening valleys. The ridges stand ~1500 m deeper than the northern Du Toit ICH, and taper and deepen to the east as they begin to gently curve northward. At the western end of the ridge-trough system, the seafloor shows hummocky terrain similar to that of the Du Toit ICH; however, as the ridges deepen to the east, the terrain becomes very smooth, with little evidence of hummocky textures or fault scarps (Figure 2.9d). This smooth texture continues to the east all the way to the Andrew Bain fracture zone.

Opposing the broad, smooth ridges across the eastern rift valley, on the inside corner of the Andrew Bain RTI, is a complex jumble of terrain. The seafloor is composed largely of rough, hummocky terrain similar to that of the Du Toit ICH (Figure 2.9e). However, the Andrew Bain inside corner shows no dominant directional fabric. A single ridge oriented parallel to the eastern rift valley cuts across the area, while other features are oriented at a variety of angles. Abundant scarps with short horizontal extents (~3 km), interpreted to mark faults, cut through the terrain with no preferred orientation. Contrary to the general disorganization of the area, the western flank of the lobate high centered at 27°10'E, 52°43'S shows structured spreading-parallel lineaments (Figure
Figure 2.9: (a) High-resolution shipboard bathymetry (50 m grid interval) of the SWIR between 25.5°E and 28°E. Illumination is from the west-northwest. Grey areas are between-track data gaps. Transform fault locations are labeled in Figure 2.9a as follows: DT, Du Toit; AB, Andrew Bain. Red line delineates areas interpreted to be smooth nonvolcanic seafloor. Green lines indicate areas showing characteristics of corrugated seafloor. (b) Abyssal hill fabric on the northern flank of the western rift valley. (c) Hummocky volcanic terrain on the Du Toit inside corner high. Southwestern area shows weak spreading-parallel lineations. (d) Smooth, broad ridge with no evidence of volcanic cones or fault scarps south of the eastern rift valley. (e) Hummocky volcanic terrain on the Andrew Bain inside corner. (f) Lobate high on the northern flank of the eastern rift valley showing clear spreading-parallel lineaments on its western flank.
2.9f).

2.5.2 Mantle Bouguer Anomaly

The MBA signature of survey area DA is dominated by a broad, irregularly shaped -40 mGal low centered slightly to the north of the western rift valley (Figure 2.7b). The MBA increases by $\sim$53 mGal to the west along the rift valley. It increases almost monotonically eastward from the low, reaching a positive maximum of $\sim$28 mGal near the eastern end of the northernmost broad ridge south of the eastern rift valley (Figure 2.8a). This rift valley maximum, however, is low relative to the off-axis MBA amplitudes of the eastern section, which are up to 20 mGal higher (Figure 2.7b).

2.5.3 Magnetization and Magnetic Anomalies

The magnetic anomalies and crustal magnetization of survey area DA show a clear elongated central magnetic anomaly/CAMH over the southern flank of the western rift valley (Figures 2.7c and 2.7d). The central magnetic anomaly/CAMH extends the full length of the western rift valley, along the ridge between the overlapping valleys, and over the lobate spreading-parallel-lineated high north of the center of the eastern rift valley. Intensity variations within the CAMH are broadly arch-shaped. They have a maximum near the elliptical basin at the eastern end of the western rift valley, and weaken by $\sim$12 A/m to both the west and east (Figure 2.8b). However, the overall along-axis magnetization and magnetic anomaly distributions are asymmetric. While the CAMH and central magnetic anomaly are clearly visible within the western section of the survey area, up to within $\sim$10 km of the Du Toit transform fault, they decay to the east, disappearing entirely $\sim$35 km from the Andrew Bain transform fault (Figures 2.7c and 2.7d). Similarly, while weak magnetic lineations persist off axis near the Du Toit fracture zone, they do not exist in the eastern section.

2.5.4 Earthquake Locations

A total of 32 relocated earthquake epicenters lie within survey area DA [Engdahl et al., 1998] (Figure 2.7a). Eleven of these are found within the transform domains bounding the ridge axis, four in the Du Toit and seven in the Andrew Bain. The remainder are clustered around the two rift valleys, and follow their overlap and possible offset. Focal mechanisms indicate normal faulting within the rift valleys and dextral
strike-slip faulting in the transform domains, as would be expected. Eleven earthquake epicenters lie within or near the eastern rift valley. Eight of these are clustered around a deep basin within the rift valley at 27°17′E, 52°52′S. Three lie well to the west in the central part of the valley. Ten events are located near the western rift valley, including two strike-slip mechanisms near the nodal basin of the Andrew Bain RTI.

2.6 Ridge Segmentation and Plate Boundary Location

Our observations allow us to define three segments between 25.5°E and 35°E on the SWIR. Two are located in survey area AP, one between the Andrew Bain and Marion transform faults, the other between the Marion and Prince Edward transform faults. These segments are defined primarily on a morphological basis; they both have arch-shaped along-axis bathymetric profiles uninterrupted by higher-order offsets and are thus straightforward first-order segments [Macdonald et al., 1988; Grindlay et al., 1991]. They both have rift valleys, partially or completely filled by a volcanic edifice, deep bulls-eye MBA lows, and high magnetization amplitudes. We follow the naming convention of Hosford et al. [2003] and Baines et al. [2007] and denote these segments AM-1 and MP-1. Since the AVRs within the rift valleys of these two segments are obscured, we place the plate boundary at the line of greatest depth (Figure 2.3).

Between the Du Toit and Andrew Bain transforms faults, the segmentation is not as easily defined, as the plate boundary cannot be located at every point. In the western section, to achieve the best agreement among the morphology, the magnetization, and the earthquake locations, we place the plate boundary on the southern wall of the western rift valley, following the central magnetization high (Figure 2.7). In the eastern section, the magnetization provides little control; as such, we locate the plate boundary along the deepest part of the eastern rift valley, in agreement with the earthquake locations. Between these two sections, the location of the plate boundary is unclear (dashed line in Figure 2.7).

If a ridge axis offset within survey area DA does indeed exist, then there would clearly be two individual segments by the basic definition of ridge segmentation. The earthquake locations would support this definition, as they appear to step to the right, following the locations of the two rift valleys (Figure 2.7a). However, we cannot definitively locate a clear offset in the morphology. The magnetization and magnetic anomalies show no evidence of an offset in the central magnetic anomaly/magnetization high (Fig-
ures 2.7c and 2.7d). The gravity data show a single bulls-eye MBA low, evidence of a
single zone of melt supply and/or mantle upwelling (Figure 2.7b). Hence, while rift valley
morphology and earthquake locations support the designation of two distinct segments,
the magnetic and gravity data support the designation of only one. As section 2.7 will
show, our analysis does not depend on a possible offset of the ridge axis. Thus, we will
utilize a single gravimetric segment, DA-1, for the remainder of this study.

2.7 Interpretation of Observations

2.7.1 Definitions and Cautions

For the sake of clarity, we introduce some terminology that distinguishes between the
various factors that influence the amount of melt that ridge segments receive. The terms melt supply and melt production will refer to the volume of melt produced by partial melting of the upwelling mantle, with enhanced upwelling producing more melt. Melt focusing refers to the lateral migration of melt within or between individual segments, away from discontinuities and/or oblique segments and toward orthogonal segment centers. Melt delivery refers to the amount of melt emplaced within a given ridge segment, and is thus a sum of melt supply, melt focusing, and any other processes that influence shallow melt emplacement (e.g., crustal magma plumbing). Variable regional melt supply produces variations in regional axial depth and basalt sodium content [Klein and Langmuir, 1987; Cannat et al., 2008]. Melt focusing causes melt delivery to and within individual segments to vary [Magde et al., 1997; Magde and Sparks, 1997; Cannat et al., 2003; Cannat et al., 2008]. This produces segment-scale and intrasegment variation in crustal and extrusive volcanic layer thickness.

In the following analysis, we use gravimetric and magnetic data in our investigation of segment-scale and intrasegment melt delivery variations. However, caution must be taken with this approach for a variety of reasons. MBA amplitudes include signals from crustal thickness, crustal density, and mantle density; as we lack seismic crustal thickness data, our gravity data are unconstrained. Similarly, our magnetization data cannot distinguish between intensity variations and lateral variations in the thickness of the magnetic source layer.
2.7.2 Seafloor Texture as an Indicator of Melt Supply and Delivery

Fortunately, seafloor texture is a strong indicator of melt delivery [Cannat et al., 2006]. On the SWIR, melt delivery variations produce three different types of seafloor [Cannat et al., 2006]. The first type is MAR-like volcanic seafloor, which occurs in times and/or locations of relatively high melt delivery. Magmatic seafloor is characterized by numerous volcanic cones, hummocky terrain, and fault scarps ∼50-500 m high bounding ridgeparallel abyssal hills.

The second type of seafloor is unique to ultraslow spreading ridges: smooth seafloor, which occurs at times and/or locations of low melt delivery. Smooth seafloor is predominantly found at oblique sections of the SWIR. These sections are typically marked by an axial trough with high MBA amplitudes, low magnetization intensity, and no evidence of volcanic edifices, flanked by broad, smooth ridges [Dick et al., 2003; Sauter et al., 2004b; Cannat et al., 2006]. Dredging of smooth seafloor reveals compositions of primarily mantlederived peridotite, with thin and/or scattered basalts and gabbros [Dick et al., 2003; Seyler et al., 2003].

The third type, corrugated seafloor, occurs very infrequently. Corrugated seafloor is characterized by dome-shaped features with spreading-parallel lineations <1 km wide and 30-100 m high [Cannat et al., 2006]. This morphology has been interpreted to reflect the exposed footwall of a low-angle detachment fault [e.g. Tucholke and Lin, 1994]. While it was originally thought that corrugated seafloor reflects low melt supply [Tucholke and Lin, 1994], recovery of large volumes of gabbro in cores drilled into the domes [e.g. Dick et al., 2000; Blackman et al., 2006] and numerical experiments [Tucholke et al., 2008] indicate that corrugated seafloor forms at times and/or locations with a narrow window of intermediate melt delivery.

2.7.3 Observations of Asymmetries in Melt Delivery and Lithospheric Thickness

Our analysis of the seafloor texture of segments AM-1 and MP-1 has revealed exclusively volcanic terrain, including hummocky terrain and numerous circular flat-topped mounds, which we interpret as volcanic cones (Figures 2.5 and 2.6). Our magnetic analysis shows that magnetization intensity decreases from segment centers to segment ends. This is a typical observation of SWIR segments, and has been attributed to a reduction in the frequency of volcanic eruptions away from segment centers, which
thins the extrusive volcanic layer and thus the magnetic source layer [Sauter et al., 2004a]. The westward skew of magnetization intensity thus means that the magnetic source layer is relatively thick on the western ends of the segments. This conclusion is supported by an along-axis gradient in the observed density of volcanic cones in and around the rift valleys of segments AM-1 and MP-1, with more volcanic cones near the western bounding transform faults than the eastern (Figure 2.6). Similarly, asymmetries in MBA amplitudes show that the western ends of segments AM-1 and MP-1 have a mass deficit relative to their eastern ends.

Segment DA-1 also shows a very prominent asymmetry in seafloor textures. The three seafloor textures described earlier [Cannat et al., 2006] are all present within segment DA-1. The western end is composed almost exclusively of volcanic seafloor, both in the abyssal hill fabric north (Figure 2.9b) of the ridge axis and the Du Toit ICH south of the axis (Figure 2.9c). We interpret the broad ridges south of the axis on the eastern end of the segment as smooth seafloor; they lack any evidence of volcanic cones or other volcanic terrain (Figure 2.9d). This smooth seafloor is opposed by volcanic seafloor north of the axis (Figure 2.9e). In the center of the segment, the ridges south of the axis have an increasingly volcanic texture from east to west. Three areas have characteristics of corrugated seafloor. The western flank of the lobate high centered at $27^\circ10'E, 52^\circ43'S$ shows clear spreading-parallel corrugations (Figure 9f); one small area of possible corrugated seafloor is also observed on the Du Toit ICH (the southwestern area of Figure 2.9c). The large high centered at $27^\circ04'E, 52^\circ26'S$ has the characteristic domal shape of corrugated seafloor; however, the flanks of the dome show no evidence of corrugations and a bathymetric data gap prevents us from observing the texture on the summit. The segment thus shows an east-west gradient in seafloor texture, with nearly equal volumes of smooth and volcanic seafloor on the eastern end of the segment transitioning to almost entirely volcanic seafloor on the western end.

MBA amplitude variations parallel the gradient in seafloor texture within segment DA-1, with high amplitudes in the east where smooth seafloor dominates and low amplitudes in the west where volcanic seafloor prevails (Figures 2.7b and 8a). Caution must be taken in the interpretation of the MBA. The assumption that MBA variations can largely be attributed to crustal thickness variations, frequently utilized for MAR segments, does not apply for the case of segment DA-1, as the prevalence of nonvolcanic seafloor indicates that lateral density variations in both the crust and mantle certainly
contribute to the MBA signal. Similarly, the along-axis variation in magnetic anomaly amplitude and magnetization intensity (Figures 2.7c and 2.8b) undoubtedly contains signal from both variable source layer thickness and intensity, as highly variable seafloor textures, and thus compositions, and pervasive faulting will create disorganized and unpredictable source layer properties.

Another striking asymmetry common to all three segments is that of rift valley curvature. At their western ends, the rift valleys of segments AM-1 and MP-1 intersect the bounding transform faults nearly orthogonally (Figure 2.5). Abyssal hills curve southward only within 5-7 km of the fracture zones. At their eastern ends, the rift valleys begin curving north just east of the shallowest points of the segments, 20-30 km from the bounding transform faults. The western rift valley of segment DA-1 begins to curve south ∼20 km from the Du Toit transform fault, while the eastern rift valley begins to curve north ∼40 km from the Andrew Bain transform fault (Figure 2.9).

Our observations of the west-east asymmetries in these three segments indicate three effects: (1) magmatic activity is higher in the west, producing greater volumes of volcanic seafloor textures, thicker extrusive volcanic layers and/or higher magnetization, and higher magnetic anomaly amplitudes; (2) the lithospheric welds controlling rift valley curvature are weak in the west relative to the east; and (3) relative mass deficits exist in the west, as reflected in lower MBA amplitudes there. The asymmetries of volcanic seafloor and extrusive volcanic layer thickness may be most easily explained by intrasegment variations in melt delivery, i.e., a greater volume of melt emplaced at shallow depths at the western end of the segments than at the eastern ends. The asymmetries in rift valley curvature may be explained by thin subaxial lithosphere and thus weaker lithospheric welds beneath the western ends of the segments. Relative mass deficits at the western ends of the segments may be explained by a combination of the above mechanisms: low MBA amplitudes could result from both thicker crust (i.e., greater melt delivery) and thin subaxial lithosphere (i.e., greater volumes of hot, relatively low density asthenospheric mantle at shallow depths). That all three segments show broadly similar asymmetries suggests that they may have a single common mechanism. However, obvious differences do exist, most prominently the presence of nonvolcanic seafloor within segment DA-1 versus the lack thereof within segments AM-1 and MP-1. In section 2.8, we will investigate possible explanations for the observed asymmetries, each of them involving relative motion between the spreading ridge and the underlying mantle.
2.8 Models of Asymmetric Melt Delivery

2.8.1 Ridge Migration

The first mechanism investigated involves migration of the spreading ridge in the hot spot reference frame. This migration results in an increase in melt supply and melt delivery, reflected in shallower depths, at leading ridge segments [Carbotte et al., 2004]. It is possible, then, that westward migration of the SWIR over the underlying mantle could produce asymmetric mantle upwelling under, and increased melt delivery to, the western ends of segments DA-1, AM-1, and MP-1. To test this hypothesis, we calculate the migration of the SWIR at several points between 14°E and 42°E using the absolute plate motion model HS3-Nuvel1a [Gripp and Gordon, 2002] (Figure 2.10). The direction of motion of the SWIR in this region is predominantly south. As such, ridge migration cannot be responsible for our observed asymmetries.
2.8.2 Subaxial Asthenospheric Flow: Segment DA-1

A broad along-axis asthenospheric flow, produced by gradients in mantle temperature, has been invoked to explain along-axis asymmetries in axial depths and MBA amplitudes observed between 98°E and 112°E on the Southeast Indian Ridge (SEIR) [West and Sempéré, 1998]. This mechanism involves flow entering the upstream end of a ridge segment from beneath the older, thicker adjacent lithosphere [see West and Sempéré [1998], Figure 6]. Similar asymmetries have been observed within the 15.5°E-25°E supersegment of the SWIR [Grindlay et al., 1998]. Relatively shallow depths and low MBA values on the western end of the supersegment have been ascribed to enhanced mantle upwelling, likely due to hotter mantle temperatures underneath this part of the ridge.

Eastward subaxial flow could also explain weakening of the lithospheric weld responsible for rift valley curvature on the upstream segment end. Enhanced upstream mantle upwelling would advect heat into this area, warming and thinning the subaxial lithosphere, thus reducing the strength of the lithospheric weld. This model could thus potentially explain the asymmetries observed in segment DA-1; enhanced mantle upwelling and melting would then be focused on the western, upstream end of the segment, producing axial depth and MBA amplitude patterns mirroring those observed in the 15.5°E-25°E SWIR supersegment and the 98°E-112°E SEIR region, as well as the observed stronger magnetization intensity and weaker lithospheric weld in the west. In principle, segment DA-1 could be on the distal end of the same mantle upwelling asymmetry proposed by Grindlay et al. [1998] to explain their along-axis variations in axial depth and MBA. However, eastward subaxial flow model does not explain the presence of nonvolcanic seafloor and low melt delivery at the eastern end of segment DA-1, a feature that does not appear at the eastern ends of segments within the 15.5°E-25°E SWIR supersegment nor the 98°E-112°E region of the SEIR.

2.8.3 Subaxial Asthenospheric Flow: Segments AM-1 and MP-1

Segments AM-1 and MP-1 lie ~250 and ~330 km away from Marion Island, the current seafloor expression of the Marion hot spot (Figure 2.1, inset). Seismic tomography has revealed the presence of a regional negative shear wave velocity anomaly at 75 km depth in the vicinity of the hot spot, which has been interpreted to reflect a hotter mantle in this area [Debayle et al., 2005; Sauter et al., 2009]. Anomalously shallow
bathymetry [Sauter et al., 2009] and low regional MBA amplitudes [Georgen et al., 2001] between the Andrew Bain and Discovery II transform faults (\(\sim 32^\circ E\) to \(41.75^\circ E\)) have been interpreted to reflect regional-scale thick crust due to the presence of the Marion hot spot [Georgen et al., 2001] and thus increased regional melt supply along this section of the ridge [Cannat et al., 2008].

If segments AM-1 and MP-1 are indeed within the sphere of influence of the Marion hot spot, a subaxial asthenospheric flow could be explained via either channelized flow along the ridge [e.g. Vogt and Johnson, 1975; Georgen and Lin, 2003], or diffuse radial flow outward from the hot spot [e.g. Ribe et al., 1995; Georgen et al., 2001]. We cannot definitively discount the possibility of an eastward mantle flow. However, due to the location and orientation of the segments with respect to the Marion hot spot, either channelized subaxial flow or diffuse radial flow would far more likely be directed to the west, with enhanced melt delivery on the downstream rather than the upstream ends of the segments. As such, it would appear that the asymmetries observed on either side of the Andrew Bain result from different mechanisms.

2.8.4 Influence of the Marion Hot Spot

Before we investigate how the Marion hot spot could produce the observed asymmetries, we must first show that segments AM-1 and MP-1 are influenced by the hot spot. Unfortunately, the limited dredge sampling within these two segments (green dots in Figure 2.3a [Mahoney et al., 1992]) have not conclusively revealed the presence or lack of Marion basalts at the ridge; Marion basalts have a geochemical signature more similar to that of MORB than that of incompatible-rich ocean island basalts (OIB) [Meyzen et al., 2005]. However, we may utilize our bathymetric and gravimetric observations to infer that segments AM-1 and MP-1 receive a high supply of melt, as they should if they are influenced by the hot spot. To demonstrate that these two segments have high melt supply, we compare various melt supply indicators to those of other segments of the SWIR (Table 2.1). Based on regional axial depth and basalt sodium content, Cannat et al. [2008] have placed the segments in Table 2.1 other than AM-1 and MP-1 within various melt supply regions: segment 27 is within a high melt supply region; the 15.5°E-25°E supersegment is a moderate melt supply region, as is the region containing segments 20-22. Segments 11 and 14 are within low melt supply regions. Segments AM-1 and MP-1 are on the western edge of a high melt supply region which is bounded by the
Table 2.1: Segment Length, Mean Axial Depth, Along-Axis Relief, Across-Axis Relief, Along-Axis MBA Variation, and Regional Basalt Sodium Content for Seven Areas Along the SWIR

<table>
<thead>
<tr>
<th>Location</th>
<th>Segment Length (km)</th>
<th>Mean Axial Depth (m)</th>
<th>Along-Axis Relief (m)</th>
<th>Across-Axis Relief (Abyssal Hill Relief (m)</th>
<th>MBA variation (mGal)</th>
<th>Regional Basalt Na&lt;sub&gt;8.0&lt;/sub&gt; (wt %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>15.5°E-25°E supersegment</td>
<td>21.5-77.5</td>
<td>3747</td>
<td>640-1180</td>
<td>200-600</td>
<td>7.5-31</td>
<td>~3.25</td>
</tr>
<tr>
<td>DA-1w</td>
<td>70</td>
<td>4500</td>
<td>2500</td>
<td>-</td>
<td>53</td>
<td>-</td>
</tr>
<tr>
<td>AM-1</td>
<td>85</td>
<td>3650</td>
<td>2100</td>
<td>500-800</td>
<td>35</td>
<td>-(2.71)</td>
</tr>
<tr>
<td>MP-1</td>
<td>85</td>
<td>3200</td>
<td>2900</td>
<td>500-1000</td>
<td>48</td>
<td>-(2.71)</td>
</tr>
<tr>
<td>Segment 27</td>
<td>85</td>
<td>2700</td>
<td>1900</td>
<td>500-900</td>
<td>51</td>
<td>2.61</td>
</tr>
<tr>
<td>Segments 20-22</td>
<td>40-45</td>
<td>4200</td>
<td>900-1300</td>
<td>100-350</td>
<td>42-33</td>
<td>3.03</td>
</tr>
<tr>
<td>Segment 14</td>
<td>65.1</td>
<td>4742</td>
<td>2352</td>
<td>-</td>
<td>40.4</td>
<td>3.89</td>
</tr>
<tr>
<td>Segment 11</td>
<td>60.4</td>
<td>4542</td>
<td>2672</td>
<td>-</td>
<td>53.1</td>
<td>3.72</td>
</tr>
</tbody>
</table>

Listed from west to east: the 15.5°E-25°E orthogonal supersegment [Grindlay et al., 1998], the western rift valley of segment DA-1 (current study, denoted as DA-1w), segments AM-1 and MP-1 (current study), segment 27 (50.47°E) [Sauter et al., 2001; Mendel et al., 2003], segments 20-22 (54.1°E-56.67°E) [Sauter et al., 2001; Mendel et al., 2003], and segments 14 and 11 (~61.4°E and 63.9°E, respectively) [Cannat et al., 1999]. Regional basalt sodium content values from all regions are from Cannat et al. [2008], except for the 15.5°E-25°E supersegment [Standish et al., 2008]. Sodium contents in parentheses for segments AM-1 and MP-1 indicate values published for the adjacent melt supply region [Cannat et al., 2008]; no values for these specific segments have been published. Across-axis relief values for the western rift valley of DA-1 have not been calculated because the Du Toit inside corner high produces very asymmetric topography.
Figure 2.11: (a) Segment length versus axial relief, (b) axial relief versus MBA amplitude, (c) segment length versus MBA amplitude, and (d) mean axial depth versus MBA amplitude for seven areas along the SWIR, listed from west to east: the 15.5°E-25°E supersegment containing 14 individual accretionary segments [Grindlay et al., 1998] (circles), the western rift valley of segment DA-1 (current study, diamond), segments AM-1 and MP-1 (current study, stars), segment 27 (50.47°E) [Sauter et al., 2001; Mendel et al., 2003] (square), segments 20-22 (54.1°E-56.67°E) [Sauter et al., 2001; Mendel et al., 2003] (triangle, one record for all three segments), and segments 14 and 11 (∼61.4°E and 63.9°E, respectively) [Cannat et al., 1999] (inverted triangles). Values plotted for segments 20-22 represent the central values of the ranges listed in Table 1.

We compare segment length versus along-axis relief, axial relief versus along-axis MBA variation, segment length versus along-axis MBA variation, and mean axial depth versus along-axis MBA variation for the segments listed in Table 2.1 (Figure 2.11). Weak correlations between length and axial relief, and length and MBA amplitude, for the segments in the moderate melt supply 15.5°E-25°E supersegment have been attributed to weak shallow melt focusing [Grindlay et al., 1998]. For segments 20-22, higher MBA amplitudes for similar length, axial relief, and regional melt supply to the 15.5°E-25°E segments reflects a higher degree of melt focusing, most reasonably
due to larger bounding offsets between segments and thus thicker lithosphere at the segment ends [Sauter et al., 2001]. The high relief and MBA amplitude of segment 27 has been attributed to high melt supply produced by a melting anomaly carried to the SWIR from the Crozet hot spot [Sauter et al., 2009]. This conclusion is supported by the large abyssal hills observed at this segment; abyssal hill relief has recently been correlated with crustal thickness, with greater relief abyssal hills occurring in segments with greater crustal thickness and melt supply [Mendel et al., 2003]. High relief and high MBA amplitudes at segments 14 and 11, coupled with relatively long segment lengths, have been attributed to strong shallow melt focusing; this focusing explains why these segments have similar relief and MBA amplitudes to segment 27 despite having low regional melt supply [Cannat et al., 1999].

Segments AM-1 and MP-1 have a greater similarity to segments 11, 14, and 27 than they have to either the 15.5°E-25°E segments or segments 20-22 (Figures 2.11a-2.11c), meaning that they have a high degree of melt delivery. However, if we are to suggest a Marion hot spot influence, we must demonstrate that segments AM-1 and MP-1 have high melt supply, regardless of the amount of melt focusing that occurs. The axial depths of segments AM-1 and MP-1 are considerably shallower than those of the low melt supply segments 11 and 14 (Figure 2.11d). Also, the abyssal hills on the flanks of segments AM-1 and MP-1 are similar in size to those of high melt supply segment 27 (Table 1). Given these observations, the inferred relation of regional axial depth and melt supply, and the location of segments AM-1 and MP-1 on the edge of a high melt supply region [Cannat et al., 2008], we conclude that the high relief and MBA amplitudes of segments AM-1 and MP-1 result predominantly from high melt supply rather than strong melt focusing.

2.8.5 Flow from the Marion Hot Spot

It is not known whether Marion hot spot flow to the ridge is channelized or diffuse/radial. However, the observation that the depth and gravity anomalies associated with the Marion hot spot are confined between the Andrew Bain and Discovery II transforms suggest that hot spot material is channeled along the ridge. Numerical modeling suggests that hot spot material delivered to the SWIR near the Eric Simpson Fracture Zone (∼40°E) could successfully negotiate its way along the ridge to segments MP-1 and AM-1 before stagnating near the Andrew Bain transform fault [Georgen and Lin, 2003].
Channelized subaxial asthenospheric flow along a ridge from a hot spot has been previously investigated by Vogt and Johnson [1975]. In their model, partial melts are entrained in a shallow pipelike asthenospheric flow pool against the thick lithosphere abutting the end of the ridge segment across the discontinuity. The pooled melts then propagate upward through the ridge axis, producing thicker crust and constructional volcanic features on the downstream end of the segment (see Figure 1 of Vogt and Johnson [1975]).

This model can explain the observed asymmetries of segments AM-1 and MP-1. Melt pooling against the Andrew Bain and Marion transform faults, on the downstream end of a possible westward channelized flow in the asthenosphere under segments AM-1 and MP-1, respectively, could account for thicker crust and a thicker extrusive volcanic layer on the western ends of the segments. The excess melt could also possibly explain the presence of the narrow ridge just east of the Marion transform fault in segment MP-1 (Figure 2.6b); this ridge resembles the constructional volcanic ridges discussed by Vogt and Johnson [1975]. The interaction of hot channelized flow from the Marion hot spot with the downward sloping lithosphere at the downstream end of the segments could also serve to heat and thin this lithosphere, reducing the strength of the lithospheric weld.

Georgen and Lin [2003] have suggested that Marion hot spot material may enter the SWIR somewhere near the Eric Simpson fracture zone. If this is the case, channelized subaxial flow would proceed in both directions along the ridge; as such, we would expect to see similar enhanced melt delivery on the downstream side of both the Prince Edward and Discovery II transforms. However, no high-resolution shipboard data for this region have been published. As such, we lack the ability to assess seafloor texture, rift valley curvature, and magnetic anomaly/magnetization variations within the region as we have for our survey area. MBAs, potentially calculated from the GEBCO bathymetry data [Fisher and Goodwille, 1997] and gravity data from satellite altimetry [Sandwell and Smith, 2009] cannot constrain lithospheric thickness and melt delivery variations by themselves.

2.8.6 Segment DA-1: Influence of the Long-Offset Andrew Bain Transform Fault

Smooth nonvolcanic seafloor is ordinarily found within oblique segments of the SWIR. While we cannot discount the effects of innate ridge obliquity on melt delivery
within segment DA-1, there is a simple mechanism that could account for thick lithosphere, low melt delivery, and even ridge obliquity at the eastern end of the segment. The Andrew Bain transform domain is composed of a 450 km long southern transform valley connected to a 100 km long northern transform valley by a series of extensional relay basins [Sclater et al., 2005]. Using the southern transform valley length, a half spreading rate of 9 mm/yr (to get a minimum estimate of the age offset), and the relation for lithospheric thickness as a function of age of Parker and Oldenburg [1973],

\[ Z(t) = 9.4t^{1/2} \text{ km} \]

the Andrew Bain juxtaposes \(~66\) km thick lithosphere against the eastern end of DA-1. This extremely thick, cold lithosphere strongly cools and thickens the subaxial lithosphere of segment DA-1. At the northern end of the Andrew Bain, significantly thinner lithosphere (\(~31\) km) abuts segment AM-1; the cooling effects of this lithosphere are not only weaker, but also mitigated by the alongaxis asthenospheric flow from the Marion hot spot.

Thus, we envision that the low melt delivery and nonvolcanic seafloor at the eastern end of segment DA-1 are effects of extremely thick subaxial lithosphere due to the Andrew Bain transform fault (Figure 2.12). In this model, the cooling of the subaxial lithosphere at the eastern end of segment DA-1 by the cold, thick lithosphere abutting it strongly suppresses melt production. The majority of what little melt is produced under the eastern end of segment DA-1 flows westward along the sloping base of the lithosphere [Magde and Sparks, 1997] to the western end of the segment, or is trapped in the deep lithosphere. The remainder, a small fraction of the total melt produced, forms thin crust and/or and irregular extrusive volcanic layer. As the lithosphere thins toward the western end of the segment more melt is able to reach shallow depths. This increases the abundance of hummocky volcanic terrain, and thickens the crust and magnetic source layer. The large distance (\(~40\) km or more) at which the rift valley of segment DA-1 begins curving into the Andrew Bain reflects the strength of the lithospheric weld across the transform fault; it is entirely possible that the apparent obliquity of the rift valley at the eastern end of segment DA-1 is actually rift valley curvature produced by the weld. If this model is correct, the eastern end of segment DA-1 represents an unusual, if not unique, accretionary regime: nonvolcanic accretion produced by a combination of ultraslow spreading and the cooling effect of a long-offset transform fault.
Figure 2.12: Cartoon (not to scale) showing our proposed mechanism for the along-axis distribution of melt and resulting observed asymmetries, which are strongly controlled by the transform edge effect. Passively upwelling mantle material (solid red arrows) undergoes partial melting and releases melt (dashed red arrows) to the top of the asthenosphere. This melt is then focused away from the bounding transform faults, up the sloping base of the lithosphere (dashed white arrows) [Magde and Sparks, 1997]. However, melt production is suppressed and melt focusing is enhanced by the strong cooling and thickening of the lithosphere towards the long-offset Andrew Bain transform to the east. The majority of what little melt is produced in the mantle near the Andrew Bain is focused toward the west or trapped in the thick lithosphere; very limited melt reaches shallow depths to create a crust primarily composed of a thin basaltic cap. Melt delivery increases to the west through westward melt focusing and increased melt production (along-axis density of dashed red arrows in the lithosphere indicates the relative amount of melt released from the upper asthenosphere that becomes crust).

2.9 Andrew Bain Transform Fault: A Major Melt Supply Boundary

The axial relief, MBA variation, and length of the western rift valley of segment DA-1 (i.e., the fully magmatic portion of the segment) are very similar to those of segments 11 and 14 near the Rodrigues Triple Junction (Table 1 and Figure 2.11). This observation, combined with large axial depth, suggests that segment DA-1 receives a relatively low volume of melt. The large relief and MBA amplitude of the segment are thus the result of strong melt focusing, which agrees with our previous analysis of seafloor textures and lithospheric thickness. As stated in section 8.2., it is reasonable to presume that segment DA-1 may be on the distal end of the gradient in mantle upwelling responsible for depth and MBA asymmetries within the 15.5°E- 25°E supersegment.
We demonstrated above that segments AM-1 and MP-1 receive a high supply of melt, likely due to the influence of the Marion hot spot. Our observations thus allow us to fill the gap in the regional melt supply analysis of Cannat et al. [2008]. Segment DA-1 may be added to the 14.18°E-25.14°E intermediate-to-low melt supply region, and segments AM-1 and MP-1 may be added to the 35.55°E-41.53°E high melt supply region [see Cannat et al. [2008], Table 1]. The Andrew Bain transform fault thus represents not only a large topographic boundary (profile in Figure 2.1), but a large melt supply boundary as well.

2.10 Conclusions

1. Within the two segments northeast of, and the one segment southeast of, the Andrew Bain transform fault, we observe similar asymmetries in seafloor texture, MBA amplitude, magnetization, and rift valley curvature. We interpret these asymmetries to reflect variations in (1) the amount of melt delivered to and emplaced along the ridge axis and (2) the subaxial lithospheric thickness, which controls the strength of the lithospheric welds responsible for rift valley curvature. Despite the observed similarities between the three segments, we argue that the asymmetries observed at either end of the Andrew Bain transform fault are not created by the same mechanism.

2. The two segments northeast of the Andrew Bain transform fault have high melt supply due to the proximity of the Marion hot spot. The asymmetries within these two segments result from channelized subaxial asthenospheric flow from the hot spot. In this model, partial melts entrained in the asthenospheric flow pool against the western bounding transform faults, producing thicker crust and a thicker extrusive layer in these areas relative to areas near the eastern bounding transform faults. The interaction of the hot subaxial flow with the thickening lithosphere at the western end of these segments heats and thins this lithosphere. This reduces the strength of the lithospheric weld and hence rift valley curvature.

3. An along-axis gradient in seafloor texture, from equal amounts of volcanic and nonvolcanic seafloor in the east to almost entirely volcanic seafloor in the west, reflects an east-west gradient in melt delivery in the segment southwest of the Andrew Bain transform fault. Rather than being a result of ridge obliquity, low melt delivery on the eastern end of this segment is likely due to the strong influence of the long-offset
Andrew Bain transform fault. Extremely thick and cold lithosphere abutting the end of the segment thickens the subaxial lithosphere, strengthens the lithospheric weld and rift valley curvature, reduces melt production, and focuses melt to the west. This limits the amount of melt that reaches shallow levels to produce crust and extrusive volcanics.

4. The Andrew Bain transform fault represents a major melt supply boundary on the SWIR, separating a high melt supply region to the northeast from an intermediate-to-low melt supply region to the southwest.

Acknowledgments

We thank Donna Blackman, Jeff Gee, Daniel Sauter, Graham Baines, and two anonymous reviewers for their helpful comments on the manuscript. This work was supported by National Science Foundation grant OCE-0329565.

This chapter, in full, is a reformatted version of material as published in Geochemistry, Geophysics, Geosystems (Takeuchi, C.S., J.G. Sclater, N.R. Grindlay, J.A. Madsen, and C. Rommevaux-Jestin (2010), Segment-scale and intrasegment lithospheric thickness and melt variations near the Andrew Bain megatransform fault and Marion hot spot: Southwest Indian Ridge, 25.5°E-35°E, Geochem. Geophys. Geosys., 11(7), Q07012, doi:10.1029/2010GC003054), reprinted with permission. Christopher Takeuchi was the primary investigator and author of this paper.
Chapter 3

Dynamic models of interseismic deformation and stress transfer from plate motion to continental transform faults

Christopher S. Takeuchi and Yuri Fialko

Abstract. We present numerical models of earthquake cycles on a strike-slip fault that incorporate laboratory-derived power law rheologies with Arrhenius temperature dependence, viscous dissipation, conductive heat transfer, and far-field loading due to relative plate motion. We use these models to explore the evolution of stress, strain, and thermal regime on geologic timescales ($\sim 10^6$-$10^7$ years), as well as on timescales of the order of the earthquake recurrence ($\sim 10^2$ years). Strain localization in the viscoelastic medium results from thermomechanical coupling and power law dependence of strain rate on stress. For conditions corresponding to the San Andreas fault (SAF), the predicted width of the shear zone in the lower crust is $\sim 3$-5 km; this shear zone accommodates more than 50% of the far-field plate motion. Coupled thermomechanical models predict a single-layer lithosphere in case of dry composition of the lower crust and upper mantle, and a jelly sandwich lithosphere in case of wet composition. Deviatoric stress in the lithosphere in our models is relatively insensitive to the water content, the far-field loading rate, and the fault strength and is of the order of $10^2$ MPa. Thermomechanical coupling gives rise to an inverse correlation between the fault slip rate and the ductile strength of the lithosphere. We show that our models are broadly consistent with geodetic and heat flow constrains from the SAF in Northern California. Models suggest that the regionally elevated heat flow around the SAF may be at least in part due to viscous dissipation in the ductile part of the lithosphere.
3.1 Introduction

Geodetic observations indicate that major plate boundary faults are associated with elevated strain rates throughout the earthquake cycle [e.g. Thatcher, 1975; McClusky et al., 2000; Beavan and Haines, 2001; Fialko, 2006]. Two classes of models were proposed to explain the observed interseismic deformation: a buried dislocation in an elastic half-space representing localized aseismic shear at depth [e.g. Savage and Burford, 1970; Thatcher, 1975; Savage, 1990], and a fault in an elastic plate overlying a viscoelastic substrate [e.g. Elsasser, 1969; Nur and Mavko, 1974; Savage and Prescott, 1978]. These models have proven to be equally capable of reproducing geodetic observations of interseismic deformation, as well as some aspects of time-dependent postseismic deformation, at least in the case of a two-dimensional (2-D) fault geometry and kinematic boundary conditions [Savage, 1990].

While the elastic half-space and layered viscoelastic earthquake cycle models can produce identical surface deformation, they represent fundamentally different mechanisms of stress transfer from plate motion to a seismogenic fault. The elastic half-space models postulate that faults in the upper brittle crust are loaded by localized shear at depth. Such a shear is usually prescribed as a boundary condition without consideration of the mechanisms of localization and the behavior of the ambient ductile rocks. The layered viscoelastic models stipulate postseismic stress transfer from a relaxing viscoelastic substrate back into the brittle crust. Interseismic localization of surface strain in such models is thus a memory of past earthquakes, and the effective rheology of the ductile substrate is usually chosen to match available geodetic data. Fault slip in such models is often imposed (rather than solved for), which as we argue below leads to unrealistic stresses in the seismogenic layer. The layered viscoelastic models typically predict fairly broad and diffuse viscous flow below the brittle-ductile transition. Geological and seismic observations, including exposed mylonite zones [Poirier, 1980; White et al., 1980; Rutter, 1999; Norris and Cooper, 2003], offsets of the Moho [e.g. Lemiszki and Brown, 1988; Stern and McBride, 1998; Zhu, 2000; Brocher et al., 2004; Weber et al., 2004], seismic velocity contrasts across faults at depth [Eberhart-Phillips et al., 2006; Thurber et al., 2006; Tape et al., 2009], and deep tremors on the downward extension of major faults [Nadeau and Dolenc, 2005; Shelly, 2010] indicate that localized shear zones do exist below the brittle-ductile transition, although the depth extent, the degree of strain localization (as a function of lithology, temperature regime, and fault slip rate), and the
rheology of such fault roots are poorly understood [e.g. Bürgmann and Dresen, 2008; Wilson et al., 2004].

In this paper we consider self-consistent models of the earthquake cycle that use laboratory-derived rheologies of rocks in the lower crust and upper mantle, typical geothermal gradients, and far-field loading (i.e., representing relative plate motion) to investigate the evolution of stress and strain as a function of fault age, plate velocity, composition of the ductile substrate, and thermal regime. We focus on the case of mature continental strike-slip faults such as the San Andreas fault (SAF) in California. We consider two mechanisms of strain localization, thermomechanical coupling and (implicitly) grain size reduction, and demonstrate that viscoelastic models that employ realistic rheologies become kinematically similar to elastic half-space models. Coupled thermomechanical models can be used to infer the magnitude of absolute stress (the effective strength) of the ductile part of the lithosphere, the temperature anomaly at depth and heat flow at the Earth’s surface associated with long-term fault slip. There is a debate in the literature regarding the magnitude of stress in the lithosphere and shear heating below the brittle-ductile transition. For example, theoretical estimates of a thermal anomaly due to a strike-slip fault range from several kelvins [e.g. Lyzenga et al., 1991; Savage and Lachenbruch, 2003] to several hundred kelvins [e.g. Thatcher and England, 1998; Leloup et al., 1999]. We demonstrate that nonsingular models of transform faults constrained by experimental data require ductile stresses of the order of $10^2$ MPa and temperature perturbations of the order of $10^2$ K.

3.2 Model Description

All numerical calculations presented in this study were performed using the finite element software Abaqus/Simulia (http://www.simulia.com/products/abaqus_fea.html). We simulate the earthquake cycle by applying a far-field velocity boundary condition representing tectonic loading, and allowing the fault to instantaneously slip in the upper crust to make up the displacement deficit accrued during the previous interseismic periods.

3.2.1 Model Geometry

We consider an infinitely long strike-slip fault, which simplifies the problem to a two-dimensional antiplane-strain formulation. The model domain is composed of three
rheological layers: a 12 km thick elastic upper crust underlain by an 18 km thick viscoelastic lower crust (30 km total crustal thickness) and a 45 km thick viscoelastic mantle (Figure 3.1). A fault cuts entirely through the upper crust and terminates within the lower crust at a depth of 17 km. We make use of the symmetric nature of deformation with respect to the fault plane to reduce the computational burden.

The finite element mesh consists of two 50 km thick along-strike ($z$ direction) element layers, each composed of 75 elements in depth ($y$ direction) and 17 elements in the fault-perpendicular ($x$) direction, for a total of 2550 elements. The node spacing varies in the fault-perpendicular direction from 0.5 km on the fault to 93.58 km in the far field. Antiplane strain conditions are enforced by ensuring that each along-strike nodal layer deforms identically. The solution is insensitive to the chosen element sizes, as confirmed by simulations using meshes with finer nodal spacing.

3.2.2 Rheology

We use four rheological models of the lower crust and upper mantle. Two of these models assume classical linear Maxwell viscoelastic rheology, and the other two assume temperature-dependent power law viscoelastic rheology. Both crustal and mantle materials incorporate elastic behavior defined by the linear isotropic Hookes Law, with
a Youngs modulus and Poissons ratio of 80 GPa and 0.25, respectively. The upper crust in all models is composed of a purely elastic material with the same parameters. For simplicity, we assume that the elastic moduli do not vary with depth.

We define the dynamic viscosities of the Maxwell viscoelastic materials by using ratios of the characteristic relaxation time to the earthquake recurrence interval. The characteristic relaxation time is given by \( \tau_r = \mu / G \) where \( G \) is the elastic shear modulus, and \( \mu \) is the dynamic viscosity. Model M2000 represents a relatively strong Maxwell material with a relaxation time of 2000 years (\( \mu = 2.02 \times 10^{21} \text{ Pa s} \)), and model M20 represents a weak Maxwell material with a relaxation time of 20 years (\( \mu = 2.02 \times 10^{19} \text{ Pa s} \)). In the case of linear Maxwell models, no distinction is made between the lower crust and upper mantle.

For the power law viscoelastic materials, the steady state constitutive relation between deviatoric strain rate \( \dot{\varepsilon}_d \) and deviatoric stress \( \sigma_d \) is

\[
\dot{\varepsilon}_d = A\sigma_d^n \exp\left( -\frac{Q}{RT} \right)
\]

where \( Q \) is the activation energy, \( R \) is the universal gas constant, and \( A \) and \( n \) are rheological parameters [e.g. Kirby and Kronenburg, 1987]. One can define the effective viscosity \( \eta_{\text{eff}} \), such that

\[
\sigma_d = \eta_{\text{eff}} \dot{\varepsilon}_d; \quad \eta_{\text{eff}} = \frac{1}{A\sigma_d^{n-1}} \exp\left( \frac{Q}{RT} \right).
\]

We assume mafic composition of the lower crust [Rudnick and Fountain, 1995] and ultramafic composition of the upper mantle [Anderson and Bass, 1984; Karato and Wu, 1993]. To allow for variations in composition and water content, we consider two end-member models of wet and dry lower crust and upper mantle. Laboratory-derived parameters of these rheological models are summarized in Table 3.1.

3.2.3 Thermal Regime

Because we allow for temperature dependence in our power law viscoelastic materials, the material properties also include thermal parameters (Table 3.1). Over the duration of each power law simulation, we maintain 10°C at the Earth’s surface and 1510°C at 75 km depth (the bottom of the model domain). Zero heat flux boundary conditions are applied on all remaining faces of the domain. We divide our power law
Table 3.1: Rheological Properties of Rocks From Laboratory Measurements

<table>
<thead>
<tr>
<th>Rock Type</th>
<th>A</th>
<th>n</th>
<th>Q (kJ mol(^{-1}))</th>
<th>(\rho) (kg m(^{-3}))</th>
<th>k (W m(^{-1})K(^{-1}))</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dry (Maryland) diabase</td>
<td>8.0</td>
<td>4.7</td>
<td>485</td>
<td>2850</td>
<td>2.1</td>
<td>Mackwell et al. [1998]</td>
</tr>
<tr>
<td>Wet diabase</td>
<td>2.2 \times 10(^{-4})</td>
<td>3.4</td>
<td>260</td>
<td>2850</td>
<td>2.1</td>
<td>Shelton and Tullis [1981]</td>
</tr>
<tr>
<td>Dry olivine</td>
<td>1.1 \times 10(^{4})</td>
<td>3.5</td>
<td>535</td>
<td>3320</td>
<td>3.0</td>
<td>Hirth and Kohlstedt [2004]</td>
</tr>
<tr>
<td>Wet olivine</td>
<td>3.6 \times 10(^{5})</td>
<td>3.5</td>
<td>480</td>
<td>3320</td>
<td>3.0</td>
<td>Hirth and Kohlstedt [2004]</td>
</tr>
</tbody>
</table>

Elastic moduli for all materials are Youngs modulus = 80 GPa and Poissons ratio = 0.25. All temperature-dependent calculations assume a specific heat \(c_p\) of 1000 J K\(^{-1}\) kg\(^{-1}\); thermal diffusivities are 7.37 \times 10\(^{-7}\) and 9.04 \times 10\(^{-7}\) m\(^2\) s\(^{-1}\) for diabase and olivine, respectively. The elastic upper crust has a conductivity of 2.5 W m\(^{-1}\) K\(^{-1}\) in all coupled power law models.

models into two classes that aim to explore the efficiency of thermomechanical coupling [e.g. Yuen et al., 1978; Brun and Cobbold, 1980] as a mechanism for long-term strain localization. Therefore in each case we performed two sets of simulations, one excluding the feedback between viscous dissipation and temperature, and another allowing for full coupling.

For the first (noncoupled) set of power law models, we prescribe a temperature profile that varies linearly between 10\(^\circ\)C at the top surface (\(y = 0\)) and 1510\(^\circ\)C at the base of the model (\(y = 75\) km). This amounts to a conductive geothermal gradient of 20 C/km, typical of the upper continental crust [e.g. Turcotte and Schubert, 2002, p. 133]. The assumed geotherm may be appropriate for a tectonically active crust, but likely overestimates temperature in the lower crust and upper mantle in the stable continental lithosphere.

For the second (fully coupled) class of power law models, viscous dissipation and heat conduction modify the thermal structure and thus ductile properties of the lower crust and upper mantle. For each finite element in the ductile regions, conservation of
energy states that

\[ \rho H + k \nabla^2 T = \rho c_p \frac{\partial T}{\partial t}, \]  

(3.3)

where \( \rho \) is density, \( k \) is the thermal conductivity, \( c_p \) is the specific heat, and \( H \) is the internal heat production rate per unit mass. For the noncoupled models, we assume a linear temperature distribution with depth and do not solve equation 3.3. For the coupled models, viscous dissipation contributes internal energy equal to the scalar product of stress and strain rate tensors,

\[ \rho H = \sigma_{ij} \dot{\varepsilon}_{ij} \]  

(3.4)

where the repeating indices imply summation. During each model time increment \( \Delta t = t_f - t_i \), heat conduction produces a temperature increment in each element

\[ \Delta T_c = \int_{t_i}^{t_f} \kappa \nabla^2 T \, d\tau \]  

(3.5)

where \( \kappa = k/\rho c_p \) is the thermal diffusivity of the material; viscous dissipation contributes a temperature increment

\[ \Delta T_v = \int_{t_i}^{t_f} \frac{1}{\rho c} \sigma_{ij} \dot{\varepsilon}_{ij} \, d\tau. \]  

(3.6)

The model evaluates the total temperature increment \( \Delta T = T_c + T_v \) in each element for each time increment \( \Delta t \), and adds \( \Delta T \) to the element temperature at the end of the previous time increment. The updated temperature field then modifies the effective viscosity through equation 3.2. In the upper crust, \( H = 0 \) and temperature evolution is governed by heat conduction alone. Thermomechanical coupling requires that the energy equation 3.3 and stress equilibrium equations

\[ \sigma_{ij,j} = 0 \]  

(3.7)

(in the absence of body forces) be solved simultaneously. In equation 3.7, the comma operator denotes differentiation.

Shear heating has been investigated as a potential mechanism of strain localization in the ductile regime [Yuen et al., 1978; Brun and Cobbold, 1980; Fleitout and Froidevaux, 1980; Chery et al., 1991; Leloup et al., 1999; Montesi and Zuber, 2002; Sobolev et al., 2005; Kaus and Podladchikov, 2006]. A preexisting weakness is required to initiate a positive feedback between thermal softening and localized shear. Here we
are focused on well-developed equilibrium shear zones and are not concerned with the onset of localization. Note that stress concentration at the bottom tip of seismic ruptures provides a natural seed for strain localization in the ductile substrate. Other processes, such as dynamic recrystallization [Rutter, 1999, Montési and Hirth, 2003], may also contribute to strain localization. For an equilibrium grain size (reflecting a balance between dynamic recrystallization and static grain growth), and comparable contributions of dislocation and diffusion creep [De Bresser et al., 1998], the constitutive flow law governing viscous deformation with grain size reduction has the same stress exponent as dislocation creep [e.g Montési and Hirth, 2003]. Therefore we use the thermally activated power law rheology (equation 3.1) as a proxy for all strain-weakening mechanisms, assuming that the considered range of rheologic parameters such as the stress exponent $n$ and the premultiplying factor $A$ will account for potential contributions of other mechanisms.

Coupled thermomechanical models need to be evolved to generate a temperature anomaly that reflects a balance between conductive heat loss and viscous dissipation in the lower crust and upper mantle in response to farfield plate motion and fault slip over geologic time. In our models this is achieved by applying the plate velocity both in the far field ($x = 300$ km) and on the fault in the elastic layer (0-12 km). We cosine taper the slip rate on the fault from the far-field rate at 12 km depth to zero at 17 km depth. The temperature structure is initially linear and one-dimensional (1-D) ($20^\circ$C/km), and the model is kinematically driven over a given period of time, or until the temperature approaches a two-dimensional steady state. Note that because our models include variations in thermal conductivity with depth (Table 3.1), the corresponding steady state temperature gradient is no longer constant, and varies between the layers. These variations are calculated as part of the thermal evolution. We assume perturbations in temperature and strain rate due to viscous dissipation can develop and grow spontaneously, and do not consider conditions that lead to their initial development. Unless otherwise noted, all coupled simulations discussed below were evolved using a total plate velocity of 40 mm/yr for a slip duration of 20 Myr, comparable to the SAF slip history [e.g. Lisowski et al., 1991]. We simulate the long-term slip history and thermal evolution of the fault using adaptive time stepping without resolving individual earthquake cycles, which would otherwise be computationally prohibitive. The calculated thermal structure for each rheological end-member is then used as an initial condition for the respective earthquake cycle simulations.
The approach described above worked well for the fully coupled model with weak (wet) end-member rheology. Coupled simulations using strong (dry) power law rheology generated extremely high stresses (>10^{10} \text{ Pa}) in the lower crust, resulting in eventual thermal runaway, an instability involving a rapid temperature increase and a complete stress drop [Gruntfest, 1963; Anderson and Perkins, 1974; John et al., 2009], after which the model evolves to a new steady state. To avoid initial instabilities, we applied a perturbation to the 1-D temperature field. We sought the smallest initial temperature perturbation that ensured a quasi-steady solution in the ductile domain. Numerical tests showed that an initial perturbation of 250°C applied within 1.5 km of the fault plane in the depth interval from 10 to 17 km, and linearly decreasing to zero toward both the top (0 km) and bottom (75 km) of the model domain was sufficient to prevent unstable behavior during the kinematically driven thermal evolution. The initial temperature perturbation does not affect the structure of the steady state solution. In particular, in the absence of thermomechanical coupling the initial perturbation diffuses away to negligible values over a 20 Myr period. We subtracted the conductive contribution of the initial perturbations from the model predictions of temperature and heat flow discussed below.

Figures 3.2a and 3.2b show the temperature anomalies generated by viscous dissipation for each end-member rheology. The maximum temperature increase varies from \sim 160°C in the case of wet composition, up to \sim 375°C in the case of dry composition of the lower crust and upper mantle. As one can see in Figures 3.2c and 3.2d, near the fault the temperature field approaches a steady state after \sim 10 Myr.

### 3.2.4 Simulations of Earthquake Cycles

Simulations of earthquake cycles using coupled power law models were performed for both kinematic (displacement-controlled) and dynamic (stress-controlled) boundary conditions on the rupture surface. In kinematic models (typical of most previous studies of interseismic deformation), we apply an instantaneous coseismic slip on the fault surface such that the slip is constant (8 m) in the elastic layer (0-12 km depth), and cosine tapered to zero from 12 km to 17 km depth. We then lock the entire fault (0-17 km) for a period of 200 years (the earthquake recurrence interval). A constant velocity of 20 mm/yr, corresponding to the long-term half-slip rate, is applied at the fault-perpendicular far edge of the model ($x = 300$ km). The near edge ($x = 0$ km) has
zero-displacement boundary conditions below the rupture tip (depths greater than 17 km).

In dynamic models, the far-field loading is applied until the shear stress on the fault at a depth of 6 km (halfway through the elastic layer) exceeds a critical threshold, or the average fault strength $\sigma_s$. We use $\sigma_s$ of 30 MPa [Brune et al., 1969; Zheng and Rice, 1998; Fialko et al., 2005; Fay and Humphreys, 2006] in most calculations described below. Once a critical threshold is reached, the fault is allowed to slip and is locked again once the shear stress at 6 km depth reaches 25 MPa, corresponding to a static stress drop of 5 MPa. In these simulations, the earthquake recurrence interval and coseismic slip are calculated as part of the solution rather than imposed a priori.
<table>
<thead>
<tr>
<th>Model Configuration</th>
<th>Lower Crustal Rheology</th>
<th>Upper Mantle Rheology</th>
<th>Thermomechanical Coupling Active</th>
<th>Coseismic Rupture</th>
</tr>
</thead>
<tbody>
<tr>
<td>M20 (Maxwell)</td>
<td>$\mu = 2.02e19$ Pa s</td>
<td>$\mu = 2.02e19$ Pa s</td>
<td>No</td>
<td>Kinematic</td>
</tr>
<tr>
<td>M2000 (Maxwell)</td>
<td>$\mu = 2.02e21$ Pa s</td>
<td>$\mu = 2.02e21$ Pa s</td>
<td>No</td>
<td>Kinematic</td>
</tr>
<tr>
<td>DNK (Power law)</td>
<td>Dry diabase</td>
<td>Dry olivine</td>
<td>No</td>
<td>Kinematic</td>
</tr>
<tr>
<td>WNK (Power law)</td>
<td>Wet diabase</td>
<td>Wet olivine</td>
<td>No</td>
<td>Kinematic</td>
</tr>
<tr>
<td>DCK (Power law)</td>
<td>Dry diabase</td>
<td>Dry olivine</td>
<td>Yes</td>
<td>Kinematic</td>
</tr>
<tr>
<td>WCK (Power law)</td>
<td>Wet diabase</td>
<td>Wet olivine</td>
<td>Yes</td>
<td>Kinematic</td>
</tr>
<tr>
<td>DCS (Power law)</td>
<td>Dry diabase</td>
<td>Dry olivine</td>
<td>Yes</td>
<td>Stress controlled</td>
</tr>
<tr>
<td>WCS (Power law)</td>
<td>Wet diabase</td>
<td>Wet olivine</td>
<td>Yes</td>
<td>Stress controlled</td>
</tr>
</tbody>
</table>

Eight model configurations utilized in this study. Each pair of rows down the table represents an additional level of complexity in turn: temperature-dependent power law rheology, thermomechanical coupling, and stress-controlled rupture. For power law rheology, nomenclature is as follows: the first letter denotes water content (D = dry; W = wet); the second letter denotes whether or not thermomechanical coupling is active (C = coupled; N = noncoupled); the third letter denotes the mechanism of coseismic rupture (K = kinematic; S = stress controlled).

Model configurations used in this study are summarized in Table 3.2. Each pair of rows from top to bottom adds an additional layer of complexity: temperature-dependent power law rheology, thermomechanical coupling, and stress-controlled ruptures. In figures, we refer to each model configuration by a shorthand term. Models involving linear Maxwell rheology are referred to by their associated Maxwell relaxation time, M20 ($\tau_r = 20$ years) and M2000 ($\tau_r = 2000$ years). Power law models are referred to using a three letter acronym, in which the first letter indicates the effective water content (D = dry, W = wet), the second letter indicates whether or not thermomechanical coupling is included (N = noncoupled, C = coupled), and the third letter indicates the type of boundary condition on the fault (K = kinematic, S = stress controlled or dynamic).
3.2.5 Cycle invariance

Viscoelastic models of earthquake cycles often need to be spun up (i.e., run over multiple cycles) to ensure that predicted surface velocities are cycle invariant; i.e., do not depend on the number of cycles since the model initiation [e.g. Hetland and Hager, 2006]. The number of cycles required to accomplish this invariance scales with the ratio of the effective relaxation time to the recurrence interval. As we demonstrate below, power law models require significantly longer spin-ups compared to linear viscoelastic models. Furthermore, we show that models that achieve strain rate invariance (such that the history of surface velocities does not change from cycle to cycle) may not achieve stress invariance, with important implications for the mechanics of loading of seismogenic faults. Models that account for thermomechanical coupling may require spin-up times that are longer still because of the large timescales required to achieve thermal equilibrium. After a sufficient number of cycles, the incremental heat generation and bulk rheological change during a single cycle are negligible and stress and strain rate become effectively cycle invariant.

3.3 Results of Numerical Simulations

In this section we present model predictions for fault-parallel shear stress and fault-parallel shear strain rate at the end of an interseismic period, immediately preceding the next slip event. We also show the predicted time-dependent surface velocities between two slip events (for two interseismic periods separated by 50 cycles, to illustrate cycle invariance).

3.3.1 Fault-Parallel Shear Stress

The number of cycles required to achieve stress cycle invariance varies widely depending on the rheology of the ductile substrate. Figure 3.3 illustrates the evolution of stress at the end of repeated seismic cycles (i.e., immediately preceding the next earthquake) in different locations within the computational domain for the eight tested configurations. Maxwell models (Figures 3.3a and 3.3b) achieve cycle invariance in fewer than 100 cycles, with M20 reaching invariance almost immediately. In contrast, power law models that do not account for viscous heating fail to produce converging stresses even after many thousands of cycles (here we show stresses for the first 1000 cycles; we
Figure 3.3: Predicted fault-parallel shear stress versus earthquake cycle number for models (a) M2000, (b) M20, (c) DNK, (d) WNK, (e) DCK, (f) WCK, (g) DCS, and (h) WCS. Stresses are plotted at the end of the interseismic period of each numbered cycle. Curves indicate shear stress at x = 0 km, y = 6 km (solid line) at the midpoint of the upper crust on the fault; x = 0 km, y = 12 km (dotted line) at the prescribed brittle-ductile transition on the fault; x = 0 km, y = 18 km (dashed line) in the lower crust 1 km beneath the fault tip; and x = 142 km, y = 6 km (dash-dotted line) well off-fault in the middle of the upper crust, where x is the fault-perpendicular coordinate and y is depth. z = 0 for all.

note that surface strain rate invariance is indeed reached in these calculations). Furthermore, predicted stresses are unrealistically high (Figures 3.3c and 3.3d). Results shown in Figures 3.3e-3.3h were obtained by applying a temperature field from a 20 Myr simulation (described in section 3.2.3) to a new (undeformed) mesh. The inclusion of thermomechanical coupling mitigates the high stresses predicted by the noncoupled models, though a significant number of earthquake cycles are still required to achieve cycle invariance (Figures 3.3e-3.3h). In the case of dynamic ruptures (Figures 3.3g and 3.3h) the stresses approach cycle invariance within a few hundred cycles.

Figure 4 shows the distribution of shear stress as a function of depth and distance from the fault after reaching stress cycle invariance for Maxwell and coupled power law models, and during a 1000 cycle spin-up for noncoupled power law models. Stresses are plotted at the end of the interseismic period. Model M20 shows essentially negligible stress in the entire domain, as all of the coseismic stress change is relaxed by the end of the interseismic period (Figure 3.4a). All kinematically driven models are associated
Figure 3.4: Predicted fault-parallel shear stress versus distance from fault and depth. Model configurations are the same as in Figure 3. Stress is cycle-invariant for (a, b, and e-h) Maxwell and coupled power law models and plotted after 1000 cycles for (c and d) non-coupled power law models. Stress is plotted at the end of an interseismic period, immediately preceding the next earthquake. For all plots, $z = 0$. Dashed black line indicates the prescribed elastic-ductile transition (12 km depth); solid black line indicates the prescribed Moho (30 km depth).

with large negative (i.e., opposite to the sense of far-field loading) stress around the seismogenic fault (Figures 3.4a-3.4f). This is clearly unphysical, as the fault is forced to slip in a sense opposite to that of the resolved shear stress. Increasing stress on a fault to a positive value in such models requires an additional shear of the entire domain. As a result, shear stress in the elastic layer off of the fault is always higher than the stress acting on the fault, and may be in fact higher than the yield strength of the intact upper crust, depending on the magnitude of the developed negative stress anomaly on the fault (Figures 3.4a-3.4f). The stress-controlled models (Figures 3.4g and 3.4h) are self consistent in that they produce sustained earthquake cycles driven by far-field plate motion while incorporating laboratory-derived rheologies and realistic geothermal gradients. As expected, stresses are lower for wet (weak) compositions compared to dry (strong) compositions, although the difference is moderate. Stresses are also lower for coupled models compared to noncoupled models. Coupled kinematic and dynamic models produce similar stresses below the brittle-ductile transition.
### 3.3.2 Fault-Parallel Shear Strain Rate

The linear Maxwell rheologies give rise to a broadly distributed viscous flow, as evidenced in the respective shear strain rate fields at the end of the interseismic period (Figures 3.5a and 3.5b). High ratios of the relaxation time to the recurrence interval (M2000) are required to maintain strain rate anomalies throughout the interseismic period (Figure 3.5a). As expected, a weak substrate model (M20) fails to produce a localized strain rate anomaly near the fault late in the interseismic phase, as most of the coseismic stress change is completely relaxed (Figure 3.5b). Noncoupled power law models show the effects of stress-dependent weakening, with a noticeable localization of strain rates in the crust and mantle within $\sim 25$ km of the fault plane. Localization is more robust for the weak rheology (wet composition), with higher nearfault strain rates than in the case of strong rheology (dry composition). Also notable is the lobe of negative strain rate for the wet rheology, with highest magnitude $\sim 10$ km away from the fault at a depth of $\sim 17$ km, decaying away into the upper mantle (Figure 3.5d). This is also observed in the coupled power law models, and is especially prominent in the case of dry rheology (Figures 3.5e and 3.5g). These features are surprising, given that the inferred sense of shear is opposite to the sense of shear stress (Figure 3.4). They likely represent nonlinear viscoelastic effects. In particular, no backward flow is observed in models that impose a constant slip rate or a stress-free boundary condition on the fault. Coupled power law configurations also illustrate the enhanced strain localization produced by thermomechanical coupling (note the logarithmic color scale in Figure 3.5). The stress-controlled models produce strain rate fields that are nearly indistinguishable from those predicted by kinematic power law simulations.

### 3.3.3 Surface Velocities

Spatiotemporal evolution of surface velocities is of interest, as it can be used to constrain rheological properties of the Earth’s crust and upper mantle [Thatcher, 1975; Li and Rice, 1987; Pollitz et al., 2000; Kenner and Segall, 2003; Freed and Bürgmann, 2004]. Results shown in Figures 3.6a and 3.6b illustrate that simple linear Maxwell models fail to reproduce key geodetic observations, namely postseismic velocity transients and permanently elevated interseismic strain rates near the fault. In particular, the highviscosity model (M2000) is able to generate an arctangent-like velocity profile throughout the interseismic period, but does not produce postseismic transients (Figure
Figure 3.5: Predicted fault-parallel shear strain rate versus distance from fault and depth. Model configurations are the same as in Figure 3.3. Strain rate in each plot is plotted at the same time as the corresponding plot in Figure 3.4. For all plots, $z = 0$. Dashed black line indicates the prescribed elastic-ductile transition (12 km depth); solid black line indicates the prescribed Moho (30 km depth).

3.6a). The low-viscosity model (M20) generates robust postseismic transients, but no strain rate anomaly around the fault late in the interseismic period (Figure 3.6b). Figures 3.6c and 3.6d illustrate that power law models are able to achieve surface velocity invariance while stress invariance is not reached (Figures 3.3c and 3.3d). All power law models produce both postseismic transients and arctangent-like profiles at the end of the interseismic period (Figures 3.6c-3.6h). Wet compositions result in more robust early postseismic transients, as one might expect. We note that the wavelength of the transient velocity peak is nearly the same in all power law models, and is considerably larger than that due to the low-viscosity Maxwell model (Figure 3.6b). We interpret this result as indicating that the high-stress lid (Figures 3.4c-3.4h) does not contribute much to postseismic deformation, and the latter is controlled primarily by viscous relaxation in the weak substrate. Simulations using the same rheologic parameters as in model M20, but assuming the thickness of the elastic layer of 30 km (instead of 12 km) produced a wavelength of the surface velocity profiles comparable to that seen in Figures 3.6c-3.6h.

We also investigated to what extent the predicted surface velocities depend on the size of the computational domain. In particular, we performed simulations in which
the domain size in the fault perpendicular direction was increased by a factor of 3 (from 300 to 900 km, see Figure 3.1). A common feature of all models driven by a velocity boundary condition applied on the sides is a small but non-vanishing strain rate in the far field. The magnitude of this far-field strain rate does depend on the domain size (large domains giving rise to smaller strain rates at the end of an interseismic period). However, the near-field strain rate (within several locking depths from the fault trace) is relatively insensitive to the assumed size of the computational domain.

3.4 Comparison With Observations

In this section we compare predictions of our models to available observations. We do not tailor the models to specific earthquake scenarios, as we are interested in overall qualitative features of the models.

3.4.1 Geodetic Observations

Unfortunately, few observations exist of surface deformation spanning the entire cycle of great earthquakes on a mature strike-slip fault. Here we use a data set col-
Figure 3.7: Geodetic observations of shear strain rate following the 1906 San Francisco earthquake [Kenner and Segall, 2003] (symbols) and model predictions (solid lines). Shades of gray denote different time periods. Late interseismic strain rates have been subtracted from model predictions.

lected over a period of 87 years following the 1906 San Francisco earthquake [Kenner and Segall, 2003]. This data set contains GPS, trilateration, and triangulation measurements of surface strain rate. While the available data are too sparse and imprecise to discriminate between candidate rheologies, they may be sufficient to test whether the models produce reasonable surface deformation patterns. We compare modeled surface strain rates with the observed rates (Figure 3.7) at several epochs after the earthquake. The data shown in Figure 3.7 were corrected for interseismic deformation [Kenner and Segall, 2003]; correspondingly, we subtracted the late interseismic strain rates from each model prediction.

All models except M2000 and DNK produce postseismic transients that are reasonably consistent with observations. Of the eight models, only model M20 produces an 11.5 year transient that is comparable in magnitude to the data point at the respective time. However, this model lacks any significant strain rate signature later in the interseismic period, as discussed in section 3.3.3. We note that the early strain rate anomaly inferred from triangulation/trilateration data might be affected by shallow afterslip, resulting in a spuriously high near-fault strain rate amplitude. Therefore we do not consider the high apparent strain rates in the early phase of postseismic relax-
ation as a strong model discriminant. Both noncoupled power law models (Figures 3.7c and 3.7d) and models including thermomechanical coupling (Figures 3.7e-3.7h) can be deemed to be within the measurement errors. Kenner and Segall [2003] and Johnson and Segall [2004] argued that models incorporating viscoelastic shear zones on the downdip extensions of faults provide the best fit to the data. We note that shear zones in the models of Kenner and Segall [2003] and Johnson and Segall [2004] were introduced ad hoc, while in our models they are generated as part of the solution. While we cannot discriminate between candidate rheologies due to a considerable scatter in the data, one may conclude that stress-controlled coupled power law models (Figures 3.7g and 3.7h) are within the available geodetic constraints.

3.4.2 Surface Heat Flow Observations

Here we compare the surface heat flow predicted by our coupled power law models to borehole heat flow observations from several areas around the SAF (Figure 3.8)-Parkfield, the Elk Hills, and the San Joaquin Valley [Benfield, 1947; Lachenbruch and Sass, 1980; Sass et al., 1971; Sass et al., 1982; Fulton et al., 2004]. These borehole sites have been selected such that there are no other faults between the site and the SAF, so as to avoid potential thermal contributions from other faults. The selected borehole data also encompass a wide range of distances from the SAF and provide information about both the magnitude and the wavelength of the observed SAF heat flow anomaly.

It is well known that the SAF lacks the near-fault heat flow anomaly that would be expected from frictional heating above the brittle-ductile transition, assuming a coefficient of friction of 0.6-0.8 [e.g. Lachenbruch and Sass, 1980]. However, there is a broader heat flow anomaly in the California Coast Ranges approximately centered on the SAF [Lachenbruch and Sass, 1973]. The proposed explanations for the regionally elevated heat flow include the slab window [Dickinson and Snyder, 1979], advective transport of the frictionally generated heat on the SAF by fluid flow through permeable upper crust [e.g. Scholz et al., 1979], and viscous dissipation in the underlying plastosphere [Lachenbruch and Sass, 1980; Molnar, 1991; Thatcher and England, 1998]. Our results lend support to the suggestion that the observed heat flow anomaly may be at least partially due to shear heating in the ductile substrate. Both end-member rheologies in our coupled power law models satisfy constraints provided by heat flow measurements (Figure 3.8).
Differences in the predicted heat flow maxima (∼20 mW m$^{-2}$) may be large enough for the heat flow data to provide some discrimination between candidate rheologies of the ductile substrate. Such a discrimination would hinge on the contribution from frictional heating in the brittle crust, which is ignored in our models. If the contribution of frictional heating is significant, the observations are more consistent with a weak wet composition, as a strong dry composition and high friction would result in a heat flow anomaly greater than that observed (Figure 3.8). If the heat flow from frictional heating is insignificant [Brune et al., 1969; Lachenbruch and Sass, 1980; Fulton et al., 2004], then the strong dry rheology may be favored. Mature strike-slip faults likely have a transition zone from highly localized frictional slip to ductile shear that includes a transition from velocity-weakening to velocity-strengthening friction [Marone et al., 1991; Dieterich, 1992; Scholz, 1998]. The depth extent of the velocity-strengthening slip, and the associated effective coefficient of friction, are not well understood, and are not explicitly included in our models. We partially account for a possible occurrence of localized creep by extending the slip interface by 5 km into the viscoelastic medium (12-17 km depth). If stable sliding occurs over a greater depth range and under a low effective
normal stress (e.g., due to elevated pore pressure), the predicted heat flow anomaly due to viscous dissipation might be lower than that shown in Figure 3.8.

3.5 Discussion

3.5.1 Model Comparisons

Prior investigations of lower crustal and upper mantle rheology using layered viscoelastic cycle models have shown that surface deformation patterns at a single post-earthquake epoch can be used to infer the effective viscosity of the ductile substrate at that particular time; however, the effective viscosity appears to change throughout the cycle [e.g. Kenner and Segall, 2003; Pollitz, 2003; Freed and Bürgmann, 2004; Hearn et al., 2009]. Therefore, a univiscous Maxwell rheology was deemed to be inadequate for the lower crust and/or upper mantle. Our results for univiscous Maxwell models agree with these findings, in that such models are unable to produce both arctangent-like interseismic velocity profiles and transient postseismic deformation (Figures 3.7a and 3.7b). Proposed alternatives include biviscous or multiviscous [e.g. Pollitz et al., 2001; Pollitz, 2003; Pollitz, 2005; Kenner and Segall, 2003; Hetland and Hager, 2005; Hearn et al., 2009] or nonlinear (e.g., power law) [e.g. Reches et al., 1994; Freed and Bürgmann, 2004] rheologies. The latter are motivated by laboratory experiments indicating that under high stress and temperature, ductile rocks deform by power law creep [Kirby and Kronenburg, 1987; Karato and Wu, 1993]. Previous studies incorporating power law creep reported that the effective viscosities of the lower crust and upper mantle inferred from fitting the geodetic data must be far lower than those suggested by laboratory experiments [Lyzenga et al., 1991; Reches et al., 1994], or that temperatures below the brittle-ductile transition must be higher than those suggested by surface heat flow data [Freed and Bürgmann, 2004]. Our results show that models assuming laboratory-derived power law parameters and normal geotherms, but neglecting thermomechanical coupling, give rise to unrealistically large stresses in the lithosphere, as viscous dissipation is unable to keep up with the build up of elastic stress (Figures 3.3c and 3.3d). Models that account for a feedback between viscous dissipation and the effective viscosity remove this problem and predict reasonable (i) stresses in the lithosphere (Figures 3.3e-3.3h and 3.4e-3.4h), (ii) surface velocities throughout the earthquake cycle (Figure 3.7e-3.7h), and (iii) surface heat flow anomalies (Figure 3.8), at least for mature faults. Note that our models do not require anomalous temperatures or unusual rheologies below the brittle-ductile
transition.

3.5.2 Thickness and Strength of the Mechanical Lithosphere

Models presented in section 3.3 satisfy basic conservation laws (in particular, conservation of energy and momentum), and may allow for predictions of the thickness of the mechanical lithosphere, defined as the portion of the model that supports high deviatoric stress. A high-stress lid extends well below the elastic-ductile transition in dynamic coupled power law simulations (Figures 3.4g and 3.4h). The lithosphere in these models thus develops self-consistently for the assumed rheology and loading conditions. The stress in the lithosphere away from the fault is nearly constant down to some characteristic depth $H_l$ that can be associated with the effective mechanical thickness of the lithosphere. The inferred magnitude of shear stress is of the order of 50-125 MPa, consistent with petrological estimates of shear stress below the brittle-ductile transition in tectonically active continental crust [e.g. Hirth et al., 2001; Behr and Platt, 2011]. We did not consider brittle failure in the bulk of the upper crust, so that stress is overestimated in the uppermost $\sim 10$ km. The magnitude of stress in the lithosphere only weakly, if at all, depends on the assumed static strength of the fault $\sigma_s$, as well as the assumed composition of the ductile substrate. Our models predict nearly identical magnitudes of stress in the lithosphere away from the active fault for $\sigma_s$ between 30 and 90 MPa, and for the wet and dry endmember rheologies of the lower crust and upper mantle (Figure 3.9). The main effect of water content is variable thickness of the lithosphere. For the SAF-like loading rates and the assumed initial geothermal gradients of 20 C/km, the effective thickness of the lithosphere decreases from $\sim 30$ km for the dry composition to $\sim 20$ km for the wet composition (Figure 3.9). Also, the wet composition gives rise to a jelly sandwich structure (strong middle crust and upper mantle separated by a weak lower crust, Figures 3.4h and 3.9).

Figure 3.10 shows the shear stress supported by the lithosphere as a function of the far-field loading rate. For the models illustrated in this figure, we evolved the previously described solution for each end-member rheology by changing the loading rate by a factor of 3 (to 13.3 mm/yr and 120 mm/yr) and applying the respective velocity on both the fault plane in the elastic layer (0-12 km depth) and in the far field. The slip rate on the fault plane was again cosine tapered to zero from 12 to 17 km depth. The model was kinematically driven in this manner until a new quasi-steady thermal
state was established, which required $\sim 5$ Myr in all cases. We then simulated dynamic earthquake cycles as before, using the new thermal states as initial conditions, until full cycle invariance was achieved.

As Figure 3.10 shows, the stress supported by the lithosphere decreases with increasing loading rate for both endmember rheologies. This relation is somewhat non-intuitive, as one might expect that absolute stresses scale with rates of relative plate motion. The inferred inverse proportionality between stress and loading rate is a consequence of thermomechanical coupling, which allows the localized shear zone to thermally soften and accommodate higher strain rate at lower shear stresses [e.g. Fialko and Khazan, 2005]. At lower loading rates, the dissipative thermal anomaly is relatively small and only weakly promotes localization, resulting in more distributed shear and higher stresses.

3.5.3 The Magnitude of Temperature Increases Due to Viscous Heating

Self-heating in the ductile lithosphere due to a longterm motion on a strike-slip fault has been investigated in several studies that reached very different conclusions. For
example, Thatcher and England [1998] and Leloup et al. [1999] suggested that the dissipative temperature perturbation should be of the order of hundreds of degrees Celsius, while Lyzenga et al. [1991] and Savage and Lachenbruch [2003] argued for much smaller temperature increases (under similar loading conditions) of the order of 1-10°C. Given that the magnitude of the dissipative temperature anomaly has important implications for the effective strength of the lithosphere and the surface heat flow, tighter constraints on the effects of viscous heating are certainly warranted.

Savage and Lachenbruch [2003] proposed that the high-end (order of 10^2°C) temperature anomaly deduced by previous studies [e.g. Thatcher and England, 1998] stems from an unphysical stress singularity at the bottom of the elastic layer. Indeed, models of Thatcher and England [1998] and Leloup et al. [1999] assume a constant slip rate in the elastic layer, and an abrupt termination of slip at the brittle-ductile transition. Models of Lyzenga et al. [1991] and Savage and Lachenbruch [2003] avoid the stress singularity by tapering the fault slip below the brittle-ductile transition, or introducing a yield threshold near the fault tip, respectively. Our models assume a tapered slip distribution, similar to that of Lyzenga et al. [1991], so that stresses are bounded every-
where, regardless of the grid size. The predicted temperature anomaly is of the order of a few hundreds of degrees (Figure 3.2), similar to the values obtained by Thatcher and England [1998] and Leloup et al. [1999], and much larger than the values obtained by Lyzenga et al. [1991] and Savage and Lachenbruch [2003]. We note that results presented in Figure 2 are based on the assumption of a mafic composition of the ductile substrate, while the low-end values of Savage and Lachenbruch [2003] were inferred for the case of granitic composition. To test to what extent the predicted temperature anomaly depends on composition, we modified our model to include a granitic middle crust in the depth range of 12-20 km, with creep law parameters \( n = 3.3, Q = 186.5 \text{ kJ mol}^{-1}, \) and \( A = 2.11 \times 10^{-5} \text{ MPa}^{-n} \text{s}^{-1} \) [Carter and Tsenn, 1987], and a lower crust in the depth range of 20-30 km with the creep law parameters \( n = 3.1, Q = 243 \text{ kJ mol}^{-1}, \) and \( A = 8.0 \times 10^{-3} \text{ MPa}^{-n} \text{s}^{-1} \) corresponding to those of felsic granulite [Wilks and Carter, 1990]. In these simulations we assumed a wet olivine rheology of the upper mantle (see Table 1). As one might expect, the respective temperature anomaly is lower than those predicted for the mafic composition (Figure 3.2), but still in excess of 75°C. The predicted temperature anomaly also strongly depends on the assumed geothermal gradient. For a typical geotherm in the continental crust [e.g. Turcotte and Schubert, 2002, pp. 143-144], the felsic end-member model described above predicts a dissipative temperature increase of 270°C, comparable to predictions of the mafic end-member models assuming higher temperatures below the brittle-ductile transition (Figure 3.2 ). We speculate that the relatively low magnitude of thermal perturbations deduced by Lyzenga et al. [1991] and Savage and Lachenbruch [2003] was due to their neglect of thermomechanical coupling, and underestimation of the background stress.

### 3.5.4 Strain Localization

Our coupled power law simulations demonstrate that strain localization at depth is produced on the downdip extension of the fault due to a positive feedback between shear heating and temperature-dependent rheology (Figures 3.5e-3.5h and 3.11). For the parameters used in this study, the predicted width of the shear zone in the lower crust is several km (Figure 3.11), in good agreement with some geological observations of exposed lower crustal shear zones [Leloup and Kienast, 1993; Dumond et al., 2008]. Thus much of the relative plate motion is accommodated by a deep fault root that extends into the lower crust, and possibly into the upper mantle (Figures 3.5e-3.5h and 3.11).
This might provide some physical justification for the use of elastic half-space models of interseismic deformation [Savage and Burford, 1970; Lapusta et al., 2000], although one needs to systematically compare predictions of the elastic half-space models incorporating rate-state friction to predictions of nonlinear viscoelastic models (e.g., Figure 3.6) to understand similarities and differences between the two classes of models. This will be addressed in future work. The degree of strain localization depends on the host rock composition, water content, ambient temperature, and fault slip rate. Stiffer rheologies, lower ambient temperatures, and higher slip rates all give rise to narrower shear zones. Such dependence may be in part responsible for the ongoing debate on the localized versus distributed nature of deformation in the ductile part of the continental lithosphere. For example, Wilson et al. [2004] argued for broadly distributed deformation below the Wairau and Awatere faults in New Zealand based on the absence of Moho offsets and regional seismic anisotropy. However, these faults have relatively low slip rates (several mm/yr), young age (several Myr), and are part of a complex system of subparallel, closely spaced faults [Bourne et al., 1998], so that the region of elevated strain rate at the base of the crust may indeed be broad and possibly overlapping between neighboring fault zones. Higher silica content and elevated geotherms may also contribute to broad deformation zones in the lower crust. On the other hand, fast-moving mature faults such as the SAF may be associated with fairly deep and localized shear zones, consistent with available data [Poirier, 1980; White et al., 1980; Rutter, 1999; Stern and McBride, 1998; Zhu, 2000; Thurber et al., 2006, Tape et al., 2009; Nadeau and Dolenc, 2005; Shelly, 2010].

3.5.5 Implications for Field Observations of Ductile Shear Zones

To date, few cases of thermally controlled localized ductile shear have been reported for strike-slip fault systems [e.g. Leloup and Kienast, 1993; Camacho et al., 2001; Dumond et al., 2008]. It is usually assumed that the relict thermal indicators of a shear zone generated through shear heating, such as metamorphic grade, overprinted mineral assemblages, isotopic closure temperature, deformation microstructures, etc., should vary over a distance comparable to the width of the shear zone. However, our results show that for mature faults the width of the shear zone (defined by the region of high strain rate) may be considerably (by as much as an order of magnitude) narrower than the associated thermal anomaly (Figure 3.11). The predicted temperature anomaly
Figure 3.11: Temperature anomaly and predicted fault-parallel velocity at 20 km depth versus distance from fault for dynamic coupled power law models. Both fields are plotted at the end of the interseismic period.

shows variations of only 15-25°C within a few km of the fault, which may be too small a difference to produce an obvious signature in thermal indicators within the shear zone compared to ambient rocks. The corollary is that field evidence for thermomechanical coupling may be subtle, as expressed e.g., in a regional (tens of km wide) paleotemperature anomaly centered on a shear zone (assuming that the exposure allows one to identify and track the same paleodepth), or anomalously high temperature of the shear zone with respect to the normal geotherm (assuming that the paleodepth of the exposure can be determined independently).

3.6 Conclusions

We have considered models of earthquake cycles on a mature strike-slip fault. Models that incorporate laboratory-derived temperature-dependent power law rheology, viscous heat generation, and conductive heat transfer predict the development of shear zones in the middle and lower crust, gradually widening in the upper mantle. These shear zones localize strain in the interseismic period, resulting in stress transfer from the
relative plate motion to seismogenic faults in the upper brittle crust. Shear zones also participate in postseismic transients by relaxing coseismic stress changes. For the SAF-like loading rates, the predicted temperature anomaly below the brittle-ductile transition is of the order of 200-400°C, and the width of the shear zone is of the order of several kilometers, consistent with geological observations in exposed deep shear zones worldwide [Leloup and Kienast, 1993; Dumond et al., 2008]. Our numerical simulations suggest that the water and silica content in the lower crust and upper mantle do not appreciably affect the shear stress in the lithosphere, but do control the thickness of a high-stress lid. The stress in the lithosphere is found to be of the order of 50-125 MPa, in agreement with petrological evidence [Hirth et al., 2001; Behr and Platt, 2011]. The lithospheric stress decreases with increasing rate of relative plate motion due to enhanced thermal weakening in the shear zone. Thermomechanical coupling is thus a viable mechanism by which stress perturbations in the viscoelastic lower crust and upper mantle may spontaneously generate localized ductile shear zones that ultimately control the effective strength of the continental lithosphere. Mature (∼10^7 yrs) shear zones generated in the temperature-dependent power law lower crust and upper mantle may be an order of magnitude narrower than the associated thermal anomalies, implying that field evidence for thermally induced localized shear may be subtle unless paleotemperature indicators are mapped over considerable (kilometers to tens of kilometers) distances away from the zone of enhanced ductile shear. Our modeling results suggest that the broad heat flow anomaly around the San Andreas fault in Northern California may in part reflect viscous heating in the deep fault root extending into the lower crust and possibly the upper mantle.

Acknowledgments

We thank two anonymous reviewers and editor Tom Parsons for their helpful comments and suggestions. This work was supported by the National Science Foundation (grant EAR-0944336) and the Southern California Earthquake Center (SCEC). Finite element meshes were created using code APMODEL, courtesy of S. Kenner.

permission. Christopher Takeuchi was the primary investigator and author of this paper.
Chapter 4

Numerical investigation of the role of lithospheric shear zones in postseismic deformation

Christopher S. Takeuchi and Yuri Fialko

Abstract. We present three-dimensional (3-D) numerical models of postseismic deformation following earthquakes on a vertical strike-slip fault. Our models incorporate linear Maxwell, Burghers, and temperature-dependent power-law rheology for the lower crust and upper mantle. We investigate effects of viscous shear zones that result from thermomechanical coupling [Takeuchi and Fialko, 2012] and potential kinematic similarities between viscoelastic models incorporating shear zones, and elastic models incorporating rate-strengthening friction on a deep aseismic fault root. We find that the thermally-activated shear zones have little effect on postseismic relaxation. In particular, the presence of shear zones does not change the polarity of vertical displacements in cases of rheologies that are able to generate robust postseismic transients. Stronger rheologies do give rise to an opposite polarity of vertical displacements, but the amplitude of the predicted transient deformation is generally negligible. We conclude that additional (to thermomechanical coupling) mechanisms of strain localization are required for a viscoelastic model to produce a surface deformation pattern similar to that due to afterslip on a deep extension of a fault. We also investigate the discriminating power of models incorporating Burghers and power law rheology. These rheologies were proposed to explain postseismic transients following large ($M7$) earthquakes in the Mojave desert, Eastern California [e.g., Pollitz, 2003; Freed and Burgmann, 2004].
Numerical simulations indicate that it may be difficult to distinguish between these rheologies even with high-quality geodetic observations for observation periods less than a decade. Longer observations, however, may potentially allow discrimination between the competing models, as illustrated by the model comparisons with available GPS and InSAR data.
4.1 Introduction

Large earthquakes are followed by spatially- and temporally-varying deformation as the lithosphere responds to the stress perturbation produced by coseismic slip. Imaging of this transient deformation has dramatically improved in recent years as the density of GPS and InSAR observations have increased in both time and space around major faults. The primary mechanisms invoked to explain postseismic deformation include viscoelastic relaxation [Elsasser, 1969; Nur and Mavko, 1974, Savage and Prescott, 1978], aseismic fault creep [Ruina, 1983; Tse and Rice, 1986], and poroelastic rebound [Booker, 1974; Jonsson et al., 2003]. Poroelastic relaxation and shallow afterslip are upper crustal processes that have predominantly near-field effects, while deeper afterslip and viscoelastic relaxation occur mainly in the lower crust and upper mantle and thus produce broad-ranging surface displacements.

The individual contributions of these mechanisms to postseismic relaxation following a given event may be difficult to identify, largely due to the non-uniqueness of inverse models. For instance, for an infinitely long rupture undergoing uniform coseismic displacement, viscoelastic relaxation predicts surface deformation indistinguishable from that due to an appropriately configured elastic dislocation model [Savage, 1990]. It was proposed that in the case of three-dimensional (3-D) deformation due to finite ruptures, the afterslip and viscoelastic relaxation mechanisms may in principle be distinguished as these mechanisms predict vertical postseismic velocity patterns that are opposite in polarity [e.g. Pollitz et al., 2001]. In addition, viscoelastic relaxation models typically predict relatively large fault-normal velocities up- and down-strike from the finite rupture, while such velocities are small or absent in afterslip models [Hearn, 2003]. However, if multiple mechanisms contribute to postseismic relaxation, even high-quality observations may not allow for robust discrimination between the candidate mechanisms. For example, postseismic deformation in the Eastern California Shear Zone following the 1992 $M_{w}7.3$ Landers and 1999 $M_{w}7.1$ Hector Mine earthquakes has been ascribed to viscoelastic relaxation [Pollitz, 2003, Freed and Bürgmann, 2004], a combination of poroelastic relaxation and afterslip [Fialko, 2004a; Peltzer et al., 1998], poroelastic and viscoelastic relaxation [Masterlark and Wang, 2002], as well as other mechanisms [Massonnet et al., 1996; Jacobs et al., 2002].

Afterslip on the deep extension of the rupture plane may be considered kinematically analogous to an infinitely narrow ductile shear zone, with shear deformation
governed by the constitutive equation for stress-driven frictional afterslip on a planar interface rather than ductile flow within a finite volume [e.g. Barbot et al., 2009]. Several previous investigations of viscoelastic relaxation have allowed for strain localization within a viscoelastic framework by incorporating a tabular region of reduced effective viscosity around the fault plane [e.g. Kenner and Segall, 2003; Freed et al., 2007; Hearn et al., 2009]. However, the prescription of such shear zones is somewhat \textit{ad hoc} and neglects details of how such features formed in the first place. In a previous study, we demonstrated that shear heating and thermomechanical coupling in the ductile substrate give rise to long-lived localized shear zones beneath mature strike-slip faults [Takeuchi and Fialko, 2012]. These shear zones participate both in loading faults interseismically and relaxing coseismic stress perturbations. Here, we extend our 2-D results to 3-D finite-rupture scenarios. We use these models to test the hypothesis that, under assumptions of laboratory-derived rheologies and far-field loading, a model incorporating highly localized ductile shear zone produced by shear heating over geologic time may produce a postseismic deformation field similar to that predicted by a frictional afterslip model. In this case, variable patterns of postseismic deformation might be expected depending on the effective fault age and slip rate, with afterslip-like localized viscoelastic shear dominating postseismic relaxation in the case of a mature fault and diffuse viscoelastic relaxation dominating in the case of an immature fault. We find that, contrary to the hypothesis, thermally-induced shear zones have little effect on postseismic relaxation. It follows that the degree of shear localization necessary for a viscoelastic relaxation model to mimic postseismic surface deformation due to frictional afterslip model requires additional localization mechanisms.

4.2 Model Description

4.2.1 Geometry

Most numerical simulations presented in this study were carried out using the finite element software Abaqus/Simulia (www.simulia.com/products/abaqus_fea.html). The model domain is a 600 km (fault-normal, \( x \) coordinate) \( \times \) 600 km (along-strike, \( y \) coordinate) \( \times \) 75 km (vertical, \( z \) coordinate) rectangular block (Figure 4.1). The domain is composed of three horizontal rheological layers: a 12 km-thick elastic upper crust overlying an 18 km-thick viscoelastic lower crust and a 45 km-thick viscoelastic upper mantle. A 600 km-long (\( y = -300 \) to 300 km) vertical planar fault is introduced within
the domain at $x = 0$. The fault penetrates through the entire upper crust and roots in the lower crust at a depth of 17 km.

The model domain is discretized into 612,000 elements, with 34 element layers in the fault-normal direction, 240 layers in the strike direction, and 75 vertical layers. Fault-normal node spacing decreases towards the fault, from 93.58 km in the far-field to 0.5 km on the fault. Along-strike node spacing varies from 19.57 km in the far-field to 0.5 km for $|y| > 35$ km; nodes are spaced by 0.5 km within $|y| \leq 35$ km. Nodes are spaced vertically by 1 km.

4.2.2 Rheology

We explored four candidate rheologies of the viscoelastic lower crust and upper mantle. Two models incorporate linear rheologies for these layers. The first model incorporates a Maxwell rheology for the entire ductile substrate. The mantle has a viscosity $\eta_m$ of $1.6 \times 10^{17}$ Pa s, corresponding to the transient rheology of Pollitz [2003], and a shear modulus $\mu_m$ of 70 GPa. The lower crust in this model has a viscosity $\eta_c$ of $3.2 \times 10^{19}$ Pa s and a shear modulus $\mu_c$ of 38 GPa. Poisson’s ratio is 0.25 in all model layers. We refer to this model as MS.

The second model incorporating linear rheology assumes a biviscous Burghers body rheology, in which Maxwell and Kelvin viscoelastic elements are in series. The shear relaxation modulus for a Burghers rheology with steady-state and transient viscosities $\eta_1$
and $\eta_2$, respectively, and steady-state and transient shear moduli $\mu_1$ and $\mu_2$, respectively [Findlay et al., 1989], is

$$G_R(t) = \frac{1}{A}[(q_1 - q_2 r_1) \exp(-r_1 t) - (q_1 - q_2 r_2) \exp(-r_2 t)] \quad (4.1)$$

where

$$q_1 = 2\eta_1, \ q_2 = 2\eta_1\eta_2/\mu_2, \ r_1 = p_1 - A/2p_2, \ r_2 = p_1 + A/2p_2. \quad (4.2)$$

with

$$p_1 = \eta_1/\mu_1 + \eta_1/\mu_2 + \eta_2/\mu_2, \ p_2 = \eta_1\eta_2/\mu_1\mu_2, \ A = \sqrt{p_1^2 - 4p_2} \quad (4.3)$$

The shear relaxation modulus may be represented by a dimensionless Prony series expansion

$$g_R(t) = \frac{G_R(t)}{G_0} = 1 - \sum_{i=1}^{N} \bar{g}_i^P (1 - \exp(-\frac{t}{\tau_i^G G})). \quad (4.4)$$

by recognizing that if $C = \frac{q_1 - q_2 r_1}{A}$ and $D = \frac{q_2 r_2 - q_1}{A}$, equation 4.1 may be expressed as

$$G_R(t) = (C + D) - [C(1 - \exp(-r_1 t)) + D(1 - \exp(-r_2 t))]. \quad (4.5)$$

Normalizing by $G_0 = C + D$ yields the two-term dimensionless Prony series with $\bar{g}_1^P = C/(C + D), \ \bar{\tau}_1^G = 1/r_1, \ \bar{g}_2^P = D/(C + D), \ \text{and} \ \bar{\tau}_2^G = 1/r_2$.

We select Burghers body parameters $\eta_1 = 4.6 \times 10^{18}$ Pa s, $\eta_2 = 1.6 \times 10^{17}$ Pa s, $\mu_1 = 70$ GPa, and $\mu_2 = 70$ GPa, which were the best-fitting values of Pollitz [2003] for the rheology of the mantle beneath the Mojave Desert, California. We also use the best-fitting model of Pollitz [2003] for the lower crust, which incorporates a Maxwell rheology with a viscosity $\eta_c = 3.2 \times 10^{19}$ Pa s and a shear modulus $\mu_c = 38$ GPa (this value represents an average of the depth-varying shear modulus of Pollitz [2003] over our lower crustal depth range of 12-30 km). The lower crust for the biviscous-mantle model and the Maxwell-mantle model are thus identical. Poisson’s ratio is 0.25 for all materials in the biviscous-mantle model. We refer to this model as BI.

The remaining two models incorporate non-linear, temperature-dependent rheology. In these models, the deviatoric strain rate $\dot{\varepsilon}_d$ and deviatoric stress $\sigma_d$ in each finite element are related by the constitutive equation

$$\dot{\varepsilon}_d = A\sigma_d^n \exp(-\frac{Q}{RT}) \quad (4.6)$$
Table 4.1: Laboratory-derived material properties of rocks.

<table>
<thead>
<tr>
<th>Rock Type</th>
<th>A (MPa$^{-n}$ s$^{-1}$)</th>
<th>n</th>
<th>Q (kJ mol$^{-1}$)</th>
<th>$\rho$ (kg m$^{-3}$)</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dry diabase</td>
<td>8.0</td>
<td>4.7</td>
<td>485</td>
<td>2850</td>
<td>1</td>
</tr>
<tr>
<td>Wet diabase</td>
<td>$2.2 \times 10^{-4}$</td>
<td>3.4</td>
<td>260</td>
<td>2850</td>
<td>2</td>
</tr>
<tr>
<td>Dry olivine</td>
<td>$1.1 \times 10^{4}$</td>
<td>3.5</td>
<td>535</td>
<td>3320</td>
<td>3</td>
</tr>
<tr>
<td>Wet olivine</td>
<td>$3.6 \times 10^{5}$</td>
<td>3.5</td>
<td>480</td>
<td>3320</td>
<td>3</td>
</tr>
</tbody>
</table>

$1$ Mackwell et al. [1998]; $2$ Shelton and Tullis [1981]; $3$ Hirth and Kohlstedt [2004]

where the power law pre-multiplier $A$, stress exponent $n$, and activation energy $Q$ are empirically-determined constants ($R$ is the universal gas constant) [Kirby and Kronenburg, 1987]. The constitutive relation yields a stress- and temperature-dependent effective viscosity $\eta_{eff}$ for each finite element

$$\eta_{eff} = \frac{\sigma_d}{\dot{\varepsilon}_d} = \frac{1}{A\sigma_d^{n-1}}\exp\left(\frac{Q}{RT}\right). \quad (4.7)$$

We assume mafic (diabase) and ultramafic (olivine) composition for the lower crust and upper mantle, respectively [Rudnick and Fountain, 1995; Karato and Wu, 1993]. To account for variability in ductile strength, we consider end-member models of hydrated (weak) and dry (strong) mineral compositions. Laboratory-determined material parameters for these models are provided in Table 4.1. In all power-law models, the elastic behavior of both the upper crust and the ductile substrate is governed by the linear isotropic Hooke’s Law, with a shear modulus and Poisson’s ratio of 32 GPa and 0.25, respectively.

4.2.3 Thermal Regime

To pinpoint the effects of thermally-activated strain localization, we consider three different temperature regimes in our power-law models. These temperature regimes modify the effective viscosity through the exponential temperature dependence of the flow law (Equations 4.6 and 4.7). Strain will localize if the temperature below the fault is higher than that in the surrounding material.

The first thermal model assumes a one-dimensional (1-D, i.e. only varying with depth) linear geothermal gradient of $20^\circ$C/km. The top and bottom surfaces have tem-
temperatures of 10°C and 1510°C, respectively. This geotherm is used for both models incorporating power-law rheology; we refer to these models as WNS (wet/weak, no shear zone) and DNS (dry/strong, no shear zone). The second thermal model uses a 1-D piecewise-linear gradient corresponding to the maximum (in terms of the temperature at a given depth) geotherm of Freed and Bürgmann [2004], with top and bottom surfaces temperatures of 10°C and 1331°C, respectively. This model corresponds to the best-fitting model of Freed and Bürgmann [2004] for the ductile substrate underlying the Mojave Desert, with a wet olivine mantle, a wet diabase lower crust, and a temperature of 1300°C at 50 km depth. The geotherm used in this model predicts higher temperatures and thus weaker ductile material at all depths above ~66 km relative to the linear 20°C/km geotherm. We refer to this model as WFB.

The third thermal model includes localized shear zones around the downdip extension of the fault. The shear zones for each power law composition are generated using a model of long-term fault slip, during which heat conduction and viscous dissipation modify the initially 1-D 20°C/km geotherm. During each modeled time increment in the third thermal model, viscous dissipation generates internal energy within each finite element,

$$\rho H = \sigma_{ij} \dot{\varepsilon}_{ij}$$

(4.8)

where $\rho$ is density, $H$ is the internal heat production per unit mass, and $\sigma_{ij}$ and $\dot{\varepsilon}_{ij}$ are the stress and viscous strain rate (total strain rate less elastic strain rate) tensors, respectively (repeating indices imply summation). This dissipative energy term then contributes to changes in temperature within each element as governed by conservation of energy,

$$\sigma_{ij} \dot{\varepsilon}_{ij} + k \nabla^2 T = \rho c_p \frac{\partial T}{\partial t}$$

(4.9)

where $k$ is the thermal conductivity and $c_p$ is the specific heat capacity of the material in the element. This energy balance produces a temperature increment $\Delta T = \Delta T_v + \Delta T_c$, with a contribution $\Delta T_v$ from viscous dissipation

$$\Delta T_v = \int_{t_i}^{t_f} \frac{1}{\rho c} \sigma_{ij} \dot{\varepsilon}_{ij} d\tau$$

(4.10)

and a contribution $\Delta T_c$ from heat conduction

$$\Delta T_c = \int_{t_i}^{t_f} \kappa \nabla^2 T d\tau$$

(4.11)
### Table 4.2: Model Configurations

<table>
<thead>
<tr>
<th>Model Configuration</th>
<th>Lower Crustal Rheology</th>
<th>Upper Mantle Rheology</th>
<th>Thermal Regime</th>
</tr>
</thead>
<tbody>
<tr>
<td>MS (Maxwell)</td>
<td>$\eta = 3.2 \times 10^{19}$ Pa s</td>
<td>$\eta = 1.6 \times 10^{17}$ Pa s</td>
<td>N/A</td>
</tr>
<tr>
<td>BI (Burghers)</td>
<td>$\eta = 3.2 \times 10^{19}$ Pa s</td>
<td>$\eta_1 = 4.6 \times 10^{18}$ Pa s $\eta_2 = 1.6 \times 10^{17}$ Pa s</td>
<td>N/A</td>
</tr>
<tr>
<td>WFB (Power law)</td>
<td>Wet diabase</td>
<td>Wet olivine</td>
<td>Piecewise linear$^1$</td>
</tr>
<tr>
<td>WNS (Power law)</td>
<td>Wet diabase</td>
<td>Wet olivine</td>
<td>20°C/km</td>
</tr>
<tr>
<td>WSZ (Power law)</td>
<td>Wet diabase</td>
<td>Wet olivine</td>
<td>Shear zone$^2$</td>
</tr>
<tr>
<td>DNS (Power law)</td>
<td>Dry diabase</td>
<td>Dry olivine</td>
<td>20°C/km</td>
</tr>
<tr>
<td>DSZ (Power law)</td>
<td>Dry diabase</td>
<td>Dry olivine</td>
<td>Shear zone$^2$</td>
</tr>
</tbody>
</table>

Seven viscoelastic model configurations utilized in this study. Rheological parameters for diabase and olivine are summarized in Table 4.1. $^1$1-D geotherm of Freed and Bürgmann [2004]; $^2$shear zone temperature structure of Takeuchi and Fialko [2012].

where $\kappa = k/\rho c$ is the thermal diffusivity and $t_i$ and $t_f$ are the initial and final times of the time increment, respectively. The temperature increment adds to the total temperature in the finite element and updates the effective viscosity of the material in the element through equation 4.7. The thermal evolution has a duration of 20 Myr, and thus simulates the thermal conditions expected for a mature (i.e. plate boundary) fault. Further details of the setup simulating the long-term thermal evolution may be found in Takeuchi and Fialko [2012]. Takeuchi and Fialko [2012] made use of symmetries of the 2-D problem. For the current study, the thermal structures generated by Takeuchi and Fialko [2012] for each power law model are extruded along-strike, so that the $x-z$ planes at each along-strike ($y$) coordinate in the model domain used in this study are identical. As viscous dissipation and heat conduction are symmetric with respect to the fault plane, the one-sided temperature structures are mirrored across the fault to provide temperature conditions for each node in the mesh for each power law model. These symmetric temperature distributions are then used as thermal boundary conditions for the respective power law earthquake cycle models. We refer to these models as WSZ (wet/weak, with shear zone) and DSZ (dry/strong, with shear zone). Table 4.2 summarizes the rheological and thermal models used in this study.
4.2.4 Earthquake Simulations

In this study we employ kinematically-driven earthquake cycles, i.e. coseismic fault slip is prescribed as a boundary condition. Such models generate unphysical stresses in the lithosphere; in particular, repeated kinematically-prescribed earthquakes produce a large negative (i.e. having sense opposite to that of fault slip) stress concentration in the seismogenic layer that increases without bound in the case of power law rheology [Takeuchi and Fialko, 2012]. However, surface velocities are relatively insensitive to the type of boundary condition on a fault surface [Takeuchi and Fialko, 2012]. We include gravitational stresses in our models because they influence surface deformation in viscoelastic earthquake models, primarily by suppressing long-wavelength vertical displacements late in the interseismic period [Pollitz, 1997]. Gravity is implemented as a body force in the stress equilibrium equations,

$$\sigma_{ij,j} + \rho g_i = 0 \quad (4.12)$$

where $\rho$ is density, $g_i=(0,0,g)$ is the gravitational acceleration, and the comma in the first term represents differentiation. We also apply an initial lithostatic stress field

$$\sigma_l = \int_0^h \rho(z)gdz \quad (4.13)$$

where $h$ is the depth at the center of a given finite element; the initially piecewise-constant lithostatic stresses are equilibrated in an initial model step. We simulate tectonic loading by imposing constant slip rate of 40 mm/yr within the upper 12 km (corresponding to the thickness of the elastic layer) of the 600-km fault for 100 kyr. Velocities of 20 mm/yr are prescribed at the far-field boundaries of the model domain ($x = \pm 300$ km). Fault slip is cosine-tapered from the full slip rate at 12 km to zero at 17 km depth in order to prevent a stress singularity due to the abrupt termination of fault slip at the base of the elastic layer.

This initial model is followed by ten earthquake cycles on the entire 600 km-long fault. In these earthquake cycles, we prescribe an instantaneous coseismic slip of 8 m on the upper 12 km of the fault. Fault slip is again cosine-tapered to zero at 17 km depth. We then lock the fault for an interseismic period of 200 years. The total relative plate velocity (40 mm/yr) is maintained on the far-field boundaries of the model domain for the duration of the earthquake cycle sequence. Because deformation is anti-plane strain,
only fault-parallel motion is allowed.

Following the final interseismic period of the last system-wide earthquake, we simulate a finite rupture on a 70 km-long segment in the middle of the domain ($|y| \leq 35$ km). Slip amplitude is the same as in previous system-wide events. The respective moment is $1.08 \times 10^{20}$ N m, similar to the moment estimates for the 1992 $M_w$7.3 Landers rupture [Kanamori et al., 1992; Sieh et al., 1993; Fialko, 2004b], and slightly higher than those for the 1999 $M_w$7.1 Hector Mine rupture [Dreger and Kaverina, 2000; Simons et al., 2002]. During the coseismic step, the fault is not allowed to move in the fault-perpendicular or vertical directions. Following the finite rupture, the entire fault is then locked for a duration of 200 years during which coseismic stress changes are allowed to relax. During the finite rupture and following relaxation period, no vertical motion is allowed on the far-field model boundaries; normal displacements are also not allowed on the fault-parallel sides of the domain, while relative far-field velocities are maintained on these boundaries ($x = \pm 300$ km). Antiplane-strain conditions are also maintained on the model boundaries orthogonal to the slip direction ($y = \pm 300$ km). Deformation in all three dimensions is otherwise freely permitted.

4.2.5 Afterslip model

For comparison, we also perform simulations of stress-driven afterslip of the deep extension of the fault using the fictitious body force code RELAX [Barbot and Fialko, 2010]. In this model, a fault of identical dimensions to those in the viscoelastic models is embedded within a 512 km x 512 km x 512 km elastic domain. An afterslip plane lies beneath and oriented with the fault; the plane is centered at $x=0, y=0$, is 150 km in length and 50 km downdip ($|y| \leq 75$ km, $17 \leq z \leq 67$ km). The slip velocity $V$ on the afterslip plane is governed by rate-strengthening friction,

$$V = 2\dot{\gamma}_0 \sinh \frac{\Delta\tau}{(a - b)\sigma}$$

where $(a - b)\sigma$ and $\dot{\gamma}_0$ are constitutive parameters, the former controlling the "slipperiness" of the afterslip plane and the latter the timescale of postseismic slip, and $\Delta\tau$ is the coseismic stress change. $a$ and $b$ correspond to the parameters of rate- and state-dependent friction [e.g. Ruina, 1983; Dieterich, 1978]. Because in situ values of the model parameters are not well-constrained, the timescale of relaxation is rather arbitrary. The afterslip model will thus only be used for qualitative comparison. We adopt
\[(a - b)\sigma = 0.9 \text{ MPa and } \gamma_0 = 1 \text{ km/yr.}\] The coefficient of friction \(\mu_0\) at the reference slip rate \(\dot{\gamma}_0\) is 0.6. We refer to this model as AFT.

### 4.3 Results

#### 4.3.1 Surface velocity patterns

As geodetic observations provide the dominant means for identifying and distinguishing various relaxation mechanisms, in the following section we focus on theoretical predictions of surface deformation and the sensitivity of surface deformation to the assumed rheologies. We generate maps of postseismic velocity, from which contributions of secular deformation have been removed. We study the first 50 years of deformation, as postseismic velocities after this time are likely to fall below the detection limit, and there are few examples of geodetically documented long-lasting (>50 years) postseismic deformation transients. Figures 4.2-4.5 illustrate modeled postseismic surface velocities within 100 km of the fault for several epochs following coseismic rupture. Six months after the earthquake (Figure 4.2), each model shows a clear four-quadrant pattern of uplift and subsidence, though the polarity and amplitude of deformation varies. Model MS predicts the largest amplitudes of vertical velocity at this time epoch, as expected due to the short relaxation timescale of the substrate. Wet power law models all have similar vertical velocity patterns, though WFB has larger amplitudes and smaller velocity lobes compared to models WNS and WSZ. Model DNS shows a similar pattern to the wet power law models, though amplitudes are an order of magnitude smaller, and relatively large vertical velocities are seen in the far-field. Model DSZ shows similar amplitudes and far-field polarity as DNS, but the near-field polarity is reversed relative to all other viscoelastic models. Model AFT is the only other model showing near-field subsidence northeast and southwest of the finite rupture, and uplift northwest and southeast of the rupture.

Horizontal velocities six months following the earthquake show broadly similar qualitative patterns for all models, with quadrant patterns of combined right-lateral fault-parallel velocities and left-lateral fault-perpendicular velocities. As with vertical velocities, model MS predicts the largest amplitudes of horizontal velocity. Wet power law models are nearly indistinguishable except for the relatively large amplitudes for model WFB. Dry power law models predict horizontal velocities an order of magnitude smaller than wet power law models. While horizontal velocity patterns are generally
similar in all models, minor deviations from the overall patterns do exist for certain models. For instance, model AFT predicts horizontal velocities that have smaller relative amplitudes in the far-field compared to viscoelastic models, as well as vanishing fault-perpendicular velocities near the down- and up-strike extensions of the fault. Model MS predicts negligible fault-parallel velocities down- and up-strike from the fault. Model DSZ predicts large far-field velocities relative to the near-field, as well as a relatively small fault-perpendicular component of velocity near the fault plane.

All models predict a reduction in the amplitude of both vertical and horizontal surface velocities from six months to two years (Figure 4.3) as the transient deformation begins to decay. The percentage reduction is largest for model MS, which incorporates the weakest substrates. Velocity patterns and vertical polarity are maintained for all models other than MS; near-field vertical polarities remain reversed for models AFT and DSZ relative to all other models. MS predicts that maximum vertical velocities within the four quadrants move away from the fault, illustrating a "strain wave" due to stress diffusion.
Figure 4.3: Postseismic surface velocities in mm/yr, 2 years after rupture. Panels as indicated in Figure 2 caption. Colors indicate vertical velocity, arrows indicate horizontal velocity. Heavy black line denotes the fault plane.

Predicted amplitudes of both horizontal and vertical velocity are further decreased for all models ten years after the rupture (Figure 4.4). Velocity patterns and polarity are largely unchanged for model BI and all wet power law models. Models AFT and DNS illustrate expanded near-fault vertical velocity lobes. Models MS and DSZ show the most notable changes in vertical and horizontal velocity. MS predicts continued outward movement of the initial vertical velocity quadrants and the initiation of a reversed near-fault vertical polarity pattern. Model DSZ predicts enhanced near-field vertical velocities compared to far-field velocities; near-field fault-parallel velocities are also enhanced relative to the far-field.

Fifty years after the rupture, all models except DNS, WFB, and WNS demonstrate significant evolution of the respective surface velocity patterns (Figure 4.5). Vertical velocity lobes in model AFT have expanded significantly towards the far-field, reflecting propagation of afterslip towards deeper parts of the fault. Model BI predicts a low-amplitude reversed-polarity near-field vertical velocity pattern. Model DSZ shows significantly enhanced near-field vertical velocities relative to far-field velocities, as well as reduced far-field fault-parallel velocity and negligible fault-parallel velocity up- and down-strike from the rupture. Model MS predicts expansion of the 10-year reversed polarity quadrant pattern towards the far-field. Model WSZ shows the initiation of a
reversed near-field vertical pattern. Vertical velocities at this epoch are very small (<1 mm/yr), likely below the geodetic detection limits. However, models BI and MS maintain appreciable horizontal velocities, reflecting ongoing substrate response to coseismic stress changes even 50 years after the event.

4.4 Discussion

4.4.1 The effect of shear zones on postseismic deformation

The postseismic fields presented in figures 4.2-4.5 allow for a straightforward comparison of models that do and do not incorporate long-lived shear zones. Early in the postseismic period, both the polarity and magnitude of vertical velocity are nearly indistinguishable for models WNS and WSZ. Horizontal velocities are also very similar for these two models within the first ten years after the earthquake. Only later in the postseismic period, fifty years and more after the earthquake, do the velocity fields of these two models diverge, with a near-fault vertical polarity reversal and enhanced near-field fault-parallel velocities developing for model WSZ. However, after such a lengthy time period following the earthquake, horizontal and vertical postseismic transient velocities are \(\sim 1 \text{ mm/yr}\) or less and would thus be difficult to observe with GPS or InSAR.
Figure 4.5: Postseismic surface velocities in mm/yr, 50 years after rupture. Panels as indicated in Figure 2 caption. Colors indicate vertical velocity, arrows indicate horizontal velocity. Heavy black line denotes the fault plane.

The maximum magnitude of vertical velocity predicted for the reversed-polarity lobes in model WSZ never exceeds 0.3 mm/yr, and the respective deformation is thus essentially negligible.

Models DNS and DSZ deviate to a far greater extent than do models WNS and WSZ. The polarity of vertical velocity is reversed at all times for model DSZ relative to DNS. Model DSZ also shows relatively high near-field fault-parallel velocity at all times. However, vertical velocities for these models never exceed 1 mm/yr. As such, the qualitative differences between models DNS and DSZ are likely unidentifiable from an observational standpoint, and these models fail to produce postseismic transients of sufficient amplitude.

Model DSZ is the only viscoelastic model to predict the same vertical polarity pattern as the frictional afterslip model at all times in the postseismic period, illustrating that a thermally-induced shear zone may qualitatively resemble the behavior of a frictional afterslip model. However, the magnitudes of postseismic velocity transients, particularly in the vertical direction, are negligible. Wet power law models do not qualitatively resemble the frictional afterslip model at any time during the postseismic period, save for a polarity reversal late in the postseismic period for model WSZ that would also be difficult to detect with available geodetic techniques. It follows that even
though thermally localized shear zones are able to produce afterslip-like postseismic de-
formation, such deformation can only occur for very strong ductile materials that yield
negligible postseismic deformation. Weaker ductile materials are able to produce robust
transients, but the strain localization in such models is insufficient to alter surface ve-
locity patterns (in particular, the polarity of vertical velocities). This indicates that, for
the range of rheological properties considered in this study, an additional strain localiza-
tion mechanism must be invoked for viscoelastic model to produce transient deformation
similar to that predicted by a frictional afterslip model.

Potential candidates for such localization mechanisms are e.g. dynamic recrystal-
lization and fabric anisotropy. Recrystallization and resulting grain size reduction have
been observed in ductile shear zones in nature [e.g. Fitz Gerald and Stünitz, 1993; Jin
et al., 1998]; and laboratory experiments [e.g. Tullis and Yund, 1985; Rutter, 1995]. In
our simulations of shear zone development, we have assumed that all viscous deformation
is governed by dislocation creep (equation 4.6), and that thermal weakening serves as a
proxy for all localization mechanisms. Explicit consideration of dynamic recrystallization
requires that a diffusion creep term be added to the constitutive relation, such that

$$
\dot{\varepsilon}_d = \dot{\varepsilon}^{\text{dis}}_d + \dot{\varepsilon}^{\text{diff}}_d = A_{\text{dis}} \sigma_d^{n_{\text{dis}}} \exp\left(-\frac{Q}{RT}\right) + A_{\text{diff}} d^{-m} \sigma_d^{n_{\text{diff}}} \tag{4.15}
$$

where $\dot{\varepsilon}^{\text{dis}}_d$ is the dislocation creep strain rate, $\dot{\varepsilon}^{\text{diff}}_d$ is the diffusion creep strain rate,
$A_{\text{dis}}$, $n_{\text{dis}}$, and $Q$ are the dislocation creep parameters (as in section 4.2.2), $d$ is the
grain size, $m$ is the grain size exponent, $A_{\text{diff}}$ is the diffusion creep premultiplier, and
$n_{\text{diff}} = 1$ [Poirier, 1985]. In this formulation, diffusion creep is grain size dependent,
and dislocation creep is temperature dependent. Grain size evolution and shear heating
are stress-dependent processes [Poirier, 1985; equation 4.8], and the dominant creep and
localization mechanism in the combined creep law (equation 4.15) depends largely on
stress as well. The inclusion of grain-size dependent diffusion creep in the flow law would
thus likely alter the evolution of shear in the substrate. Therefore additional numerical
experiments are needed to evaluate the effect of grain size reduction on strain localization
within ductile shear zones.

### 4.4.2 Postseismic deformation in the Mojave Desert

One of the vivid examples of non-unique interpretations of geodetic data is de-
formation following the 1992 $M_w 7.3$ Landers and 1999 $M_w 7.1$ Hector Mine earthquakes
in the Mojave Desert (eastern California). Pollitz et al. [2000] and Pollitz et al. [2001] argued that the pattern of postseismic vertical velocities around the fault is consistent with relaxation in the upper mantle, and opposite to that expected for afterslip. Pollitz et al. [2000] and Pollitz et al. [2001] used models assuming linear Maxwell rheology, while Freed and Bürgmann [2004] argued that the non-exponential character of timeseries of postseismic displacements can be indicative of power-law rheology. Pollitz [2003] proposed that a biviscous rheology can also explain observations. The best fitting rheologies in these studies were chosen based on model fits to <3 years of GPS data collected in the Eastern California Shear Zone. Alternatively, it was proposed that InSAR and GPS observations of the two events can be interpreted in terms of a combination of afterslip and poroelastic rebound, such that afterslip dominates horizontal deformation, and poroelastic rebound dominates vertical deformation [Peltzer et al., 1998; Fialko, 2004a]. Here we investigate whether longer observations can provide discrimination between the previously proposed rheologies. For this purpose, we consider extended records of GPS and InSAR measurements of postseismic deformation following the Hector Mine earthquake.

We use a simplified model of the Hector Mine rupture that consists of a 54 km-long vertical planar fault. We apply a uniform slip in the depth interval 0-9 km and cosine-taper the slip to zero at 13 km depth. The average slip amplitude is chosen to match the seismic moment of the Hector Mine earthquake, \(4 - 6 \times 10^{19}\) N m [Dreger and Kaverina, 2000, Simons et al., 2002]. These simplifications are justified as (i) coseismic stress changes at depth are sensitive to the earthquake moment and not to the details of slip distribution, and (ii) we are interested in the overall qualitative features of viscoelastic deformation. Our models incorporate the viscoelastic rheologies (and thermal regime in the case of the power-law model) used in models BI and WFB, as well as the layered elastic structure used in Pollitz [2003]. We generate a tectonic background stress by applying a constant strain rate of \(0.1 \times 10^{-6}\) yr\(^{-1}\) [Savage et al., 2003] until the shear stress in the ductile substrate is effectively constant. In the case of the power law model, the thickness of the 'elastic' upper crust is in fact controlled by the duration of the tectonic loading. For longer loading periods, layers with larger relaxation times begin to flow, and the effective thickness of the elastic-brittle lid decreases. We apply the tectonic loading in the power law model until the lid is 18 km thick (equal to the prescribed thickness of the elastic crust in the biviscous model) and the stress in the ductile substrate is constant. The model predictions are essentially insensitive to a longer duration of tectonic loading.
Figure 4.6: 13 years of surface displacement in mm following the simulated Hector Mine rupture. Colors indicate vertical displacement, arrows indicate horizontal displacement. SCIGN continuous GPS stations with timeseries in Figures 4.7 and 4.8 shown by black circles, with station identifiers as indicated. Reference station (for horizontal displacements) RDMT denoted by green circle. Heavy black line denotes the simulated Hector Mine rupture. Black star denotes the epicenter of the Hector Mine earthquake.

The tectonic loading is applied parallel to the fault plane, and thus differs in azimuth from the estimated regional direction of shear of N40°W [Savage et al., 2003]. This discrepancy is second-order and has little effect on the qualitative features of the model predictions. Following the tectonic loading, we simulate coseismic rupture and 50 years of postseismic relaxation. The modeled fault is aligned with the surface rupture of the Hector Mine earthquake, with an average strike of N30°W (Figure 4.6). We extract 13-year timeseries of displacement at locations of the continuous GPS stations and reference them to station RDMT (see Figure 4.6).

We compare our model displacement timeseries to post-Hector Mine GPS timeseries obtained from the Scripps Orbit Permanent Array Center (SOPAC, http://sopac.ucsd.edu) (Figures 4.7 and 4.8). The GPS timseries are corrected for station location bias (the removal of which sets the beginning of the timeseries to zero), coseismic offsets, and annual and semi-annual effects. Data referencing to a given site removes the contribution from the regional motion of the ECSZ with respect to a stable reference frame. The referenced GPS data therefore represent contributions due to postseismic and secular deformation only.
Figures 4.7 and 4.8 show the observed horizontal and vertical GPS displacements along with predictions of model configurations BI and WFB (Table 4.2), adjusted for simulation of the Hector Mine earthquake, as discussed above. The comparison of model predictions to GPS data is somewhat difficult due to uncertainties in estimating the contribution of secular deformation to the GPS timeseries. Very few stations in the ECSZ were installed before the Hector Mine earthquake which prevents the robust estimation of secular velocities. In addition, the data may be affected by deformation due to the 1992 Landers earthquake, which is ignored in our model. Fitting a linear trend to post-Hector Mine timeseries is not sufficient to establish whether horizontal GPS velocities have returned to pre-seismic values 13 years following the event. The biviscous model appears to predict constant (but higher than interseismic) horizontal velocities after robust transient deformation within the first $\sim 3$ months, but a comparison to the assumed secular velocity (relative to station RDMT as denoted by the dotted line in Figure 4.7) shows that the coseismic stress changes remain essentially unrelaxed after 13 years of postseismic deformation. In contrast, the power law model seems to indicate that horizontal station velocities nearly returned to secular rates by the end of the 13-year period. While neither model provides a satisfactory fit to the data (possibly due to simplifications discussed above), it is worth noting that the biviscous model generally overpredicts the observed horizontal velocities, while the power-law model generally underpredicts the data. Because the coseismic source is identical between the models, these differences are due to the assumed rheological properties alone.

Vertical displacements are less vulnerable to problems of separating the transient signal from secular deformation and provide a powerful additional discriminant. As one can see from Figure 4.8, a power law model provides a better fit to the vertical GPS timeseries, both qualitatively and quantitatively, than does the biviscous model. The latter overpredicts the observed velocities at all epochs, and the misfit increases with time. Unfortunately, the vertical GPS data are inherently noisy, and may be subject to local site effects.

InSAR observations provide an additional powerful constraint, as they are highly sensitive to vertical motion, and provide a synoptic view of anomalous deformation. Pollitz et al. [2001] used several short-term (<1 yr) postseismic interferograms to argue that the polarity of the observed line-of-sight (LOS) displacements is consistent with a model postulating viscoelastic relaxation in the upper mantle. Unfortunately, InSAR
Figure 4.7: Panels a and b: horizontal displacements in mm following the Hector Mine rupture. Displacements referred to continuous GPS station RDMT (green dot) and shown by black (modeled) and blue (observed) arrows, respectively. Black asterisks denote GPS station locations, black star denotes the epicenter of the Hector Mine earthquake, and red line denotes the simulated Hector Mine rupture. Other panels show timeseries of observed (grey stars) and modeled (solid black line = biviscous model, dashed black line = power law model) horizontal displacement, referred to station RDMT, for stations indicated with black circles in panels a and b. Dotted black line denotes displacement due to modeled station secular velocity referred to simulated station RDMT.

Figure 4.8: Panels a and b: vertical displacement in mm following the Hector Mine rupture. Modeled and observed displacements shown by colors and vertical black arrows, respectively. Black asterisks denote GPS station locations, black star denotes the epicenter of the Hector Mine earthquake, and heavy black line denotes the simulated Hector Mine rupture. Other panels show timeseries of observed (grey stars) and modeled (solid black line = biviscous model, dashed black line = power law model) vertical displacement plotted for stations indicated with black circles in panels a and b.
imaging of postseismic deformation due to the Hector Mine earthquake was hampered by termination of the ERS-1 mission, and failure of gyroscopes on the ERS-2 satellite. Nevertheless, one acquisition made by the ERS-2 satellite in December 2005 from the descending track 127 could be interfered with the immediate post-earthquake acquisition of October 20, 1999. The respective interferogram is shown in Figure 4.9a. Importantly, it does not appear to be affected by atmospheric noise, which allows measurement of the LOS postseismic displacements at a centimeter level. The data show pronounced lobes of LOS displacements to the west of the rupture and little or no deformation to the east of the rupture, identical to the pattern observed following the nearby 1992 Landers earthquake [Fialko, 2004a]. This qualitative similarity implies a common deformation mechanism. The amplitude of LOS displacements in anomalous lobes to the west of the Hector Mine rupture is of the order of $\sim 3$ cm (Figure 4.9a).

We compare the observed LOS displacements accumulated over a period of 6 years following the Hector Mine earthquake to predictions of our forward models assuming various candidate rheologies (Figure 4.9). For comparison, we also calculate LOS displacements for a model combining afterslip and poroelastic relaxation. We estimate relaxation due to poroelastic effects by subtracting surface displacements calculated using drained and undrained Poisson’s ratios (0.25 and 0.31, respectively).

The predictions of the three models are broadly similar, with quadrant patterns showing enhanced areas of range increase in the northwest and southeast quadrants of the rupture and range decrease in the southwest and northeast quadrants. LOS
displacement amplitudes are notably enhanced on the west side of the rupture for all models. However, the size and shape of the LOS displacement quadrants vary. The biviscous model predicts fairly large circular lobes and high LOS displacement amplitudes compared to the observations (Figure 4.9c). The power law model predicts smaller circular lobes and lower amplitudes that slightly underpredict observed displacements (Figure 4.9d). The combined afterslip-poroelastic model appears to fit the observations well both qualitatively and quantitatively, consistent with the findings of previous studies of postseismic deformation due to the Landers earthquake [Peltzer et al., 1998; Fialko, 2004a] (Figure 4.9b).

Given that InSAR LOS displacements are highly sensitive to vertical deformation, such observations may provide the key to discriminating between candidate rheologies of the lower crust and upper mantle underlying the ECSZ. For example, a biviscous model seems to predict continued uplift and subsidence after 10-20 years following the earthquake at rates that are ruled out by the observations (Figures 4.8 and 4.9). Calculations using a power law rheology proposed by Freed and Bürgmann [2004] are in better agreement with the vertical GPS and InSAR data, but significantly underpredict the magnitude of horizontal velocities (Figure 4.7). Far-field transients detected by GPS observations indicate that viscoelastic flow in the mantle may indeed be required to explain postseismic deformation following the Hector Mine earthquake [Freed et al., 2007]. The same observations have been subsequently used to suggest that steady-state power law mantle rheology does not adequately explain the data, and that a transient rheology may be required [Freed et al., 2010], similar to that assumed in the biviscous rheology. As a successful model (or family of models) must be able to explain both near-field and far-field observations, results presented in this study suggest that further work is required to understand the nature of postseismic deformation due to the Mojave Desert earthquakes.

4.5 Conclusions

We have considered models of three-dimensional deformation resulting from finite-length strike-slip coseismic rupture in an elastic crust underlain by ductile lower crust and upper mantle of various assumed rheologies. In some models, we incorporate ductile shear zones generated by thermomechanical coupling and shear heating driven by long-term slip. Models including shear zones thus incorporate the thermal and rheological
conditions expected for a mature (i.e. plate boundary) strike-slip fault, while models lacking shear zones may represent immature faults. We find that thermally-induced shear zones contribute minimally to postseismic deformation following a finite coseismic rupture. A rheologically weak (wet) substrate produces only moderate localization of strain. A strong (dry) substrate generates a high degree of localization, but postseismic velocities at all epochs are negligible for models assuming such a rheology. Comparisons of model predictions to observations of postseismic deformation in the Eastern California Shear Zone suggest that current GPS and InSAR data provide for only weak discrimination between viscoelastic relaxation models incorporating Burghers and power law rheologies. However, differences in the decay timescale of the transient deformation predicted by the two models indicate that discrimination may be possible with continued GPS and InSAR monitoring on decadal timescales.

4.6 Appendix A: Model Validation

We benchmarked our finite element models against several available open-source semi-analytical models. We use three codes, developed to investigate the earth response to earthquakes, to compare to predictions of the finite element model: RELAX [Barbot and Fialko, 2010], VISCO1D [F. Pollitz, http://earthquake.usgs.gov/research/software/], and PSGRN [Wang et al., 2006]. The former two codes use elastic Green’s functions approaches to calculate the deformation response of a layered elastic/viscoelastic half-space due to dislocations; the latter code calculates postseismic deformation of a spherically-stratified Earth model using a spherical harmonic expansion of global deformation modes and a sum of these modes for a given dislocation source. For simplicity, we replicate a published benchmark of strike-slip faulting in an elastic crust underlain by a Maxwell viscoelastic mantle [Barbot and Fialko, 2010]. Each model considered incorporates an identical 30 km-thick elastic crust, with a shear modulus of 30 GPa and Poisson’s ratio of 0.25. The Abaqus finite element model assumes a 270-km thick upper mantle. The Relax and PSGRN/PSCMP models assume half-space mantles. VISCO1D assumes a 6431 km-radius spherical mantle (total earth radius=6371 km). The mantle in all models is composed of a Maxwell viscoelastic material with a relaxation time $\tau_r$ of 10 years.

Figure 4.10 shows the surface displacement predicted by each of the four models at $t = 2\tau_r = 20$ years after the coseismic rupture. Allowing for minor variations due to the different approaches used by each code, good agreement (within 18%) is achieved
Figure 4.10: 20 years of surface displacement for benchmarking strike-slip fault simulations incorporating linear Maxwell rheology, as predicted by four codes: a) Abaqus; b) Relax; c) VISCO1D; and d) PSGRN. Colors indicate vertical displacement, arrows indicate horizontal displacement. Heavy black line denotes the fault plane.

Among the four models in both the horizontal and vertical components of displacement. However, it should be noted that the displacements predicted by PSGRN/PSCMP may range over nearly an order of magnitude through variation in the input parameters controlling wavenumber integration. The predictions of each of the other three models also vary, though over smaller ranges, by adjusting geometrical parameters and grid resolution. As such, care should be taken when using any of these approaches and results should be checked for consistency against other codes.

Figure 4.11: Surface velocities 6 months (a and b) and 2 years (c and d) following coseismic rupture or benchmarking strike-slip fault simulations incorporating biviscous Burghers rheology, as predicted by Abaqus (a and c) and VISCO1D (b and d). Colors indicate vertical velocity, arrows indicate horizontal velocity. Heavy black line denotes the fault plane.

In order to test our implementation of the Burghers body rheology, we also compare the results of an Abaqus model with those from a VISCO1D model, which is the only semi-analytical code of the three tested that is able to implement a Burghers rheology. The models tested duplicate the rheological layering and properties of model
BI, discussed in section 4.2.2. However, in the benchmark models, the fault penetrates to 10 km depth, rooting in the elastic upper crust. The shallower fault tip is necessary because VISCO1D is not designed to model fault slip in a viscoelastic material. Figure 4.11 shows surface velocities plotted at $t = 6$ months, and $t = \tau_1 = 2$ years, the latter being the steady-state relaxation time of the Burghers material. While some differences may be observed between the two models, namely in the far-field vertical velocities, the Abaqus model captures maximum amplitudes of the vertical and horizontal velocities within 15%.

Acknowledgments

We thank Sylvain Barbot and Fred Pollitz for their assistance with their respective semi-analytical codes. Finite element meshes were created using code APMODEL, courtesy of S. Kenner.

This chapter, in full, is being prepared for submission to Geophysical Journal International (Takeuchi, C.S. and Y. Fialko, Numerical investigation of the role of lithospheric shear zones in postseismic deformation). Christopher Takeuchi was the primary investigator and author of this paper.
Appendix A

Abaqus Tips and Tricks

This appendix is intended to describe a few Abaqus tricks I learned during my trials with the software. I will assume that the reader already has a fair bit of familiarity with the structure of the Abaqus input file (*.inp). The first six sections describe how to use a few handy Abaqus functions. The latter two sections illustrate how to apply Abaqus functions to certain problems, in the hope that the descriptions may assist the reader should he or she encounter a similar problem.

A.1 Re却ts

Restarting an Abaqus simulation is handy if the user has multiple different steps within a single Abaqus simulation, or has a very long (computationally speaking) single step. Utilizing the Abaqus restart capability allows the user to break up the model run into pieces, which is very useful if, for instance, the computer on which the software is being run crashes when it’s 85% finished with a week-long job. In such a case, if the user includes a couple of extra lines in the input file, the run may be continued more or less where it left off and not require starting over from the beginning.

To activate the Abaqus restart capability, simply add the line:

*RESTART, WRITE

to a step definition; for simplicity’s sake, locating it directly above *END STEP (near the other output definitions) is useful. Given this keyword, Abaqus will output the necessary data to the .res file. One may alter the frequency of output using the FREQUENCY option, or specify an increment number interval using the NUMBER INTERVAL option. However, *RESTART does not allow time interval or time point specification as *OUTPUT does. Adding the OVERLAY option will tell Abaqus to overwrite the .res file
with each successive restart output. This is useful if you are using many steps and call
*RESTART at the end of each one, as the .res file may get very large, particularly if a
great number of nodes/elements are used in the model mesh. However, use of OVERLAY
only allows you to restart from the last *RESTART output.

To configure an input file for a restart run, add the line:

*RESTART,READ
to the beginning of the input file. In addition to the *RESTART call, step definitions are
all that is required in the restart input file; all other definitions (mesh, node and element
sets, equations, materials, etc.) are carried over from the .res file and thus cannot be
redefined. One may specify a step after which the restart run will begin using the STEP
option, assuming that step was written to the .res file during the previous run. The
INC option may also be used to specify a time increment after which the restart run
will begin. Inclusion of the END STEP option will terminate the step from which the
restart run will initiate, allowing the user to define a new step and thus new boundary
conditions, loads, outputs, etc. If the END STEP option is not included, the previously
configured step will continue.

To execute a restart, use the command:
	nice abaqus job=name-of-restart-job oldjob=name-of-previous-job

The 'oldjob' parameter defines the name of the job from which the new job will
restart. The restart requires that the .mdl, .odb, .prt, .res, and .stt files of the previous
job are located in the directory where the restart job will be executed.

The restart capability may be useful should the user require a set of alternating steps to be run in a repeated sequence- for example, an earthquake cycle sce-
nario. In such a case, one may configure two input files, each containing a single step:
1) the coseismic step, in which fault slip is specified with a displacement or velocity
boundary condition; and 2) the interseismic step, during which the fault is locked. If
*RESTART,WRITE,FREQUENCY=1,OVERLAY is added to the step definitions in
each input file, one may write a shell script to execute the two steps in sequence as many
times as required. For instance, in bash:

ijo=$job

nice abaqus job=$ijo ask delete=off &
sleep 10

pid=`ps -eo pid,args | grep standard.exe | grep $csjob | awk '{print $1}'`
while [-e /proc/$pid ]; do sleep 10; done

while [ -e /proc/$pid ]; do sleep 10; done

ncyc=1000

while [ $a -le $ncyc ]

    nice abaqus job=$isjob oldjob=$csjob ask_delete=off &

    sleep 10

    pid=`ps -eo pid,args | grep standard.exe | grep $isjob | awk '{print $1}'`
    while [-e /proc/$pid ]; do sleep 10; done

    nice abaqus job=$csjob oldjob=$isjob ask_delete=off &

    sleep 10

    pid=`ps -eo pid,args | grep standard.exe | grep $csjob | awk '{print $1}'`
    while [-e /proc/$pid ]; do sleep 10; done

    a=$(( $a + 1 ))

done

This process allows the user to utilize two input files to run 2001 total steps while preventing .odb files from becoming overly large. The downside is that the user will lose the outputs of all steps but the last two. However, the shell script could be modified so that after certain steps the .odb will be copied and saved. In this script, Ijob is the input file for the first step, with mesh, node and element set, equation, etc. information, plus a coseismic step with the appropriate boundary conditions, loads, and output definitions (including *RESTART, WRITE). ISjob contains the interseismic step definition preceded by *RESTART,READ,END STEP. Again, no geometry, equation, etc. information is required in a restart input file, as it will be carried over from the previous job. CSjob contains the coseismic step definition, also preceded by *RESTART,READ,END STEP. Ijob runs first, and every successive job is then restarted from the previous job, alternating between ISjob and CSjob.
The `ask_delete` option is included in the Abaqus execution commands to tell Abaqus to automatically overwrite existing job files; otherwise, Abaqus will ask if the user wants to overwrite these files, which will cause the script to fail. The `pid` lines are necessary to prevent Abaqus from beginning a restart run before the previous job has finished, which will yield an error. The script will search for the process ID of the current Abaqus run (standard.exe), and while that process ID still exists (as checked by the following `while` loop), it will not start the following restart run. The `sleep` lines preceding the `pid` lines are included because Abaqus requires a preprocessing period (during which pre.exe runs), and the script will fail to find the standard.exe process ID during that period. The times used in the `sleep` commands may need to be modified depending on the runtimes of the steps.

A.2 Outputting Fields in ASCII

It is frequently convenient to output certain fields to ASCII for plotting purposes. Otherwise, the user is required to output fields from CAE, which is time-consuming and yields a lower degree of data precision. Abaqus user subroutine URDFIL allows for ASCII output capability through the use of simple Fortran write statements. To activate URDFIL (and any other user subroutine file), include it in a user subroutine file which is then called using the 'user' option in the Abaqus execution command.

URDFIL accesses the results file (.fil) at the end of each increment. As such, it is necessary for the user to specify output of the desired field(s) to the results file within the step definitions in the input file. For example, to output displacement information to the results file:

```fortran
*NODE FILE,FREQUENCY=1
U
*NODE FILE is used because displacement is a nodal variable in Abaqus. *
FILE should be used for the output of element variables (e.g. stress, strain, strain rate); both *
FILE and *
FILE write to the same results file. *
FILE outputs at element integration points by default. In the case that the user desires element variable information at nodes, *
FILE,POSITION=AVERAGED AT NODES, FREQUENCY=1 may be used.

Once output to the results file is specified, URDFIL may access that data in order to write an ASCII file as follows (underlined words are required parts of URDFIL,
as specified in the user’s manual; the following code assumes that all parameter and variable definitions required by the subroutine have already been defined):

```fortran
open(130,file=trim(kfpath)//'/U1time.dat')
open(140,file=trim(kfpath)//'/U2time.dat')
open(150,file=trim(kfpath)//'/U3time.dat')
call posfil(kstep,kinc,jrray,jrcd)
do k1=1,nmax
    call dbfile(0,jrray,jrcd)
    if (jrcd /= 0) go to 110
    key=jrray(1,2)
    if (key == 101) then
        u1=array(4)
        u2=array(5)
        u3=array(6)
        write(130,904) array(3),kstep,kinc,time(1),u1
        write(140,904) array(3),kstep,kinc,time(1),u2
        write(150,904) array(3),kstep,kinc,time(1),u3
    end if
end do
110 continue
```

If the user does not specify that the output be written to the current directory (via `kfpath`), the output files will be placed in the directory containing the Abaqus standard.exe executable. Abaqus utility subroutine POSFIL specifies the location in the results file using `kstep` (the current step, brought in by URDFIL) and `kinc` (the current increment, also brought in by URDFIL). URDFIL then uses utility subroutine DBFILE to read the results file at the position determined by POSFIL. Included in the data read in by DBFILE are the record keys for whatever output fields the user requested with *NODE FILE in the input file (in this case, displacement).

The *if* loop then searches the data read in by DBFILE for records containing record key 101, which is the key for displacement. All record keys are defined in the 'Results File Output Format’ section of the Abaqus Analysis User’s Manual, as are the array structures for each record (i.e. what is contained in each position of 'array’). Every
array of each record in the results file contains the record length in the first position (array(1)) and the record key in the second position (array(2)). For displacement, array(3) is the node number, while array(4), array(5), and array(6) contain the three components of total (cumulative) displacement for that node. These values are extracted from the arrays and are written to file in ASCII format, along with the step number, increment number, and step time (time(1), brought in by URDFIL).

A.3 Defining boundary conditions in a user subroutine

Simple boundary conditions may be defined within the Abaqus input file, but it may be of interest to configure more complex boundary conditions using the Abaqus user subroutine DISP. For instance, boundary condition involving a location-dependent velocity or temperature field may be defined using DISP. Configuring such a boundary condition in the Abaqus input file would require that the user identify all node locations, calculate (outside of Abaqus) the appropriate boundary condition value for each node, and then specify the boundary condition for every individual node within the given analysis step definition in the input file.

To illustrate the use of DISP, consider the case in which the user desires as a thermal boundary condition a temperature field given by the equations for cooling of the oceanic lithosphere:

\[
\frac{T(x, y) - T_0}{T_1 - T_0} = \text{erf} \left( \frac{y}{2\sqrt{\kappa x/u}} \right)
\] (A.1)

where \(x\) is the fault-normal distance, \(y\) is depth, \(T_0\) is the temperature at \(y = 0\), \(T_1\) is the temperature at \(y = \infty\), \(\kappa\) is the thermal diffusivity, and \(u\) is the plate spreading velocity. The derivation of this equation may be found in chapter 4 of Turcotte and Schubert [2002].

To call DISP for a certain node set \(NSET\) and degree of freedom \(DOF\), include the following lines in the analysis step definition in the Abaqus input file:

*BOUNDARY, USER
NSET, DOF, DOF, 1

The second line specifies that DISP will be called for only degree of freedom \(DOF\), and the value 1 specifies the actual magnitude of the variable; the latter may be considered a scaling value by which the magnitude defined by DISP will be multiplied.
To specify that DISP be used to configure the temperature boundary condition for every
node in the model domain (node set=\textit{NALL}), use:

\begin{verbatim}
*BOUNDARY, USER
  NALL,11,11,1
\end{verbatim}

The second line specifies that DISP will be called for the temperature degree of
freedom 11 for each node in node set \textit{NALL}. DISP will be called for each individual node
in \textit{NALL}; if \textit{NALL} has \textit{n} nodes, DISP is called \textit{n} times in each time increment of the
analysis step in which the boundary condition is active.

The basic setup of DISP is quite simple, as shown in the Abaqus User Subroutines
Reference Manual:

\begin{verbatim}
SUBROUTINE DISP(U,KSTEP,KINC,TIME,NODE,NOEL,JDOF,COORDS)
INCLUDE 'ABA_PARAM.INC'
DIMENSION U(3),TIME(2),COORDS(3)
definition of U
RETURN
END
\end{verbatim}

The array \(U\) is the input to DISP and defines the value of the boundary con-
dition for a particular degree of freedom \(JDOF\). \(U(1)\) is the prescribed value of the
boundary condition, while \(U(2)\) and \(U(3)\) are the first and second time derivatives of
\(U(1)\) (these may be ignored for a temperature boundary condition). \(COORDS\) is passed
in by DISP and contains the three coordinates of the node; for this example, assume that
\(COORDS(1)\) is the fault-normal coordinate \(x\), \(COORDS(2)\) is the fault-parallel coordi-
nate \(y\), and \(COORDS(3)\) is vertical coordinate \(z\), positive upwards. Note that \(COORDS\)
contains only the original nodal coordinates; actual nodal locations (i.e. original coor-
dinates plus displacements accrued during previous analysis steps) are not passed into
DISP unless nonlinear geometry is allowed. To define the boundary condition for each
node, assuming values of \(T_0\), \(T_1\), \(\kappa\), and \(u\), and that the fault lies at \(x=0\):

\begin{verbatim}
T0=0
T1=1300
kappa = 31.6 ! in m^2/yr
u=0.04 ! in m/yr
if (COORDS(1) .gt. 0) then
  U(1)=T0+(T1-T0)*erf(-COORDS(3)/(2*sqrt(kappa*COORDS(1)/u)))
\end{verbatim}
else if (COORDS(1) .lt. 0) then
    \[ U(1) = T0 + (T1-T0) \cdot \text{erf}(\frac{-COORDS(3)}{2 \cdot \sqrt{kappa \cdot COORDS(1)/(-u)}}) \]
else if (COORDS(1) .eq. 0) then
    \[ U(1) = T1 \]
end if

The if loop is included because nodes with \( x < 0 \) require a negative velocity \(-u\), and because the error function \( \text{erf} \) may produce a singularity at \( x = 0 \).

Boundary conditions on multiple degrees of freedom may be defined by adding more lines to the *BOUNDARY,USER statement in the input file, for example NALL,1,1,1 to define boundary conditions for displacement degree of freedom 1. In the user subroutine, an if loop may be used to define \( U(1) \) for each required degree of freedom, for instance:

\[
\text{if (JDOF .eq. 1) then}
    \text{definition of } U(1) \text{ for JDOF=1}
\]
else if (JDOF .eq. 11) then
    \text{definition of } U(1) \text{ for JDOF=11}
endif

It should be noted that, if the model is created using Abaqus/CAE rather than APMODEL, it may be possible to configure a similar boundary condition using the CAE Analytical Field toolset. However, DISP is generally more simple because the GUI interface of the Analytical Field toolset may make it difficult to represent complicated equations.

### A.4 Defining self-updating boundary conditions

As with direct specification in the input file, the boundary conditions defined in DISP are static, i.e. they do not change during the course of an analysis step. As such, if boundary conditions are implemented in such manners, a new analysis step must be added to modify a boundary condition on a given node set. For example, if the user desired to configure velocity boundary conditions for each node defined by corner flow, which is position-dependent:

\[
\begin{align*}
    u(x, y) &= \frac{2}{\pi} \left[ \tan^{-1} \left( \frac{x}{y} \right) - \frac{xy}{x^2 + y^2} \right], \\
    v(x, y) &= \frac{-2}{\pi} \left[ \frac{y^2}{x^2 + y^2} \right]
\end{align*}
\]  
(A.2)
where \( x \) is the fault-normal distance (the fault lies at \( x=0 \)), \( y \) is depth, \( u \) is the fault-normal velocity, and \( v \) is the vertical velocity. The approach discussed in section A.3 would allow for straightforward definition of the boundary conditions for the appropriate degrees of freedom using *BOUNDARY, USER, TYPE=VELOCITY in the input file, and in the user subroutine DISP:

\[
\begin{align*}
pi &= 3.141529 \\
kx &= \text{COORDS}(1) \\
ky &= \text{COORDS}(2) \\
\text{if (JDOF .eq. 1) then} \\
\text{if (kx .ge. 0) then} \\
U(2) &= (2/\pi) \left[ \text{atan}(kx/ky) - (kx*ky)/(kx^2+ky^2) \right] \\
\text{else if (kx .lt. 0) then} \\
U(2) &= -(2/\pi) \left[ \text{atan}(kx/ky) - (kx*ky)/(kx^2+ky^2) \right] \\
\text{end if} \\
\text{else if (JDOF .eq. 2) then} \\
U(2) &= -\left( 2/\pi \right) \left[ \text{atan}(kx/ky) - (kx*ky)/(kx^2+ky^2) \right]
\end{align*}
\]

However, the velocities defined by these equations for each node in DISP will remain constant for the duration of the analysis step instead of adjusting to reflect each node’s new position at the end of each time increment. This is because the array COORDS, which is passed into DISP, only contains the initial nodal coordinates unless geometric nonlinearity is accounted for during the step. Since COORDS does not update with each increment, neither do the applied velocities.

Using a combination of URDFIL (discussed in section A.2) and DISP (section A.3), it is possible to configure a self-updating boundary condition, i.e. a boundary condition that will adjust based on the position of each node at the end of each increment. Information cannot be passed from URDFIL to DISP internally since each user subroutine operates independently. URDFIL is thus used to output nodal displacements to a file at the end of each increment, and at the beginning of the next increment, DISP will read the displacement file and use the displacements to update nodal positions and boundary conditions. Note that the displacements output by URDFIL are total displacements, not displacement increments. URDFIL is set up almost identically as discussed in section A.2; however, displacements are output to a single file rather than three separate
files (underlined parts are again required by URDFIL):

```fortran
open(130, file=trim(kfpath)//'/displacements.dat')
call posfil(kstep, kinc, jrray, jrcd)
do k1=1, nmax
   call dbfile(0, jrray, jrcd)
   if (jrcd /= 0) go to 110
   key=jrray(1,2)
   if (key == 101) then
      u1=array(4)
      u2=array(5)
      u3=array(6)
      write(130,904) kstep, kinc, array(3), u1, u2, u3
   end if
end do
110 continue
close(130)
```

URDFIL is called at the end of each time increment; closing the displacements file will allow for it to be overwritten in successive increments, while leaving it open will allow for output to be appended. In DISP (to be included in the same Fortran file containing URDFIL, since all user subroutines must be contained in a single file; KINC, NODE, COORDS, and JDOF are variables passed in by DISP):

```fortran
if (KINC .eq. 1) then
   xpos=COORDS(1)
   ypos=COORDS(2)
   zpos=COORDS(3)
else
   open(131, file=trim(kfpath)//'/displacements.dat')
k1=0
109 k1=k1+1
   read(131,910, end=951) kfstep1(k1), kfincl(k1),
kfnodem(k1), kfxdisp1(k1), kfydisp1(k1), kfxzdisp1(k1)
910 format(3(i8,1x),3(f40.20,1x))
```
goto 109
continue
close(131)
kincm1=KINC-1
do 231 kb=1,k1-1
    if (kfnode1(kb).eq.NODE. and.kfinc1(kb).eq.kincm1) then
        xpos=COORDS(1)+kfxdisp1(kb)
        ypos=COORDS(2)+kfydisp1(kb)
        zpos=COORDS(3)+kfzdisp1(kb)
    end if
end do
if (JDOF .eq. 1) then
    if (xpos .ge. 0) then
        U(2)=(2/pi)*[atan(xpos/ypos)-((xpos*ypos)/(xpos^2+ypos^2))]
    else if (kx .lt. 0) then
        U(2)=-(2/pi)*[atan(xpos/ypos)-((xpos*ypos)/(xpos^2+ypos^2))]
    end if
else if (JDOF .eq. 2) then
    U(2)=-(2/pi)*[(ypos^2)/(xpos^2+ypos^2)]
endif

The primary if loop in this code excerpt determines the location (xpos, ypos, zpos) of the NODE for which DISP has been called. The loop separates the first time increment from all successive increments. This is necessary because DISP is called at the beginning of each increment; for the first time increment, URDFIL has not yet been called, the displacements file has not been written, and the nodal positions are simply the initial node coordinates. For each successive time increment, the displacements file written by URDFIL at the end of the previous increment is read in. The current NODE and KINC are checked against those in the displacements file (recall that the increment during which the displacements file is written is KINC-1), and the appropriate displacements are added to the initial nodal coordinates. The resulting nodal locations are then used to calculate the velocity boundary condition for the time increment (note that zpos is not required for the boundary conditions since the corner flow velocity field
is two-dimensional). Since the displacements file is updated when URDFIL is called at the end of each increment, the velocity boundary condition set by DISP at the beginning of each increment is updated as well.

The implementation of such a scheme may produce a significant increase in computation time due to the required output/input to and from the displacements file, particularly if the model mesh contains a high number of nodes. The computational burden may be mitigated by using binary I/O rather than ASCII.

A.5 Terminating an analysis based on a field value

It may also be of interest to force a step to terminate once a specific field output reaches a critical value. For example, the user may want to load a fault until the shear stress at a given location exceeds a critical threshold, after which analysis ends. URDFIL may be used for such a purpose.

As discussed in the previous section, the Abaqus input file must be configured to output fields to the results file in order for URDFIL to be used. In this case, the lines

*EL FILE,POSITION=AVERAGED AT NODES,FREQUENCY=1

should be used to output stresses to the results file. Then, in URDFIL:

\[
\begin{align*}
\text{kcnode} &= 1375 \\
\text{kslim} &= 30e6 \\
\text{call posfil(kstep,kinc,jrray,jrcd)} \\
\text{do k1=1,nmax} \\
\text{call dbfile(0,jrray,jrcd)} \\
\text{if (jrcd /= 0) go to 110} \\
\text{key} &= \text{jrray(1,2)} \\
\text{if (key == 1) then} \\
\text{jnode} &= \text{jrray(1,3)} \\
\text{end if} \\
\text{if (key == 11) .AND. (jnode == kcnode) then} \\
\text{ks13} &= \text{array(7)} \\
\text{if (ks13 > kslim) then} \\
\text{lstop} &= 1 \\
\text{end if}
\end{align*}
\]
In this code excerpt, user-defined variables \textit{kcnode} and \textit{kslim} specify the critical node at which the stress will be extracted and critical stress threshold, respectively. Utility subroutines POSFIL and DBFILE operate as discussed in the previous section. Stresses are element fields, and as such do not carry node numbers in their data records in the results file. As such, the node numbers must be obtained from data record 1, the element header record. The third position of data arrays for record 1 contain node numbers if the user has requested nodal-averaged element values in the \textit{*EL FILE} keyword. Once the node numbers are extracted, the code searches the data for record key 11, the key for stress. Arrays with record key 11 contain the 11, 22, 33, 12, 13, and 23 components of stress in positions 3-8, respectively. In the given example, the code extracts the 1-3 shear stress component \( \text{array}(7) \). This stress is then compared to the user-defined critical stress \textit{kslim}. If the stress exceeds the critical threshold, URDFIL variable \( lstop \) is set to 1, which is a flag used to indicate whether or not an analysis to continue. \( lstop=1 \) tells URDFIL to terminate the analysis.

This approach may be used in conjunction with the Abaqus restart capability to configure an earthquake cycle composed of interseismic periods that load a fault until a critical upper stress threshold is reached, followed by coseismic periods during which slip occurs until the stress is reduced below a lower stress threshold. In this manner a static earthquake stress drop may be defined.

A.6 Implementation of linear viscoelastic rheologies with Prony series

Newtonian and power-law viscous properties of a viscoelastic solid are most simply implemented in Abaqus using the \textit{*CREEP} keyword. However, more complicated linear viscoelastic rheologies such as Kelvin-Voigt, Standard Linear Solid, Burghers, Generalized Maxwell and Generalized Kelvin cannot be implemented with \textit{*CREEP} because the constitutive equations for such rheologies are not simple expressions of strain rate as a function of stress. These rheologies may be incorporated through the use of the \textit{*VISCOELASTIC} keyword with the option \textit{TIME=PRONY} included to designate that Prony series parameters will be entered for time-domain viscoelasticity.

A viscoelastic material may be represented by a Prony series expansion of the...
material’s relaxation modulus. For shear deformation (which will be considered here),
the stress response \( \tau(t) \) to a small shear strain \( \varepsilon(t) \) is

\[
\tau(t) = \int_0^t G_R(t - s) \dot{\gamma}(s) ds \tag{A.3}
\]

where

\[
G_R(t) = G_\infty + \sum_{i=1}^{N} G_i^p \exp\left(-\frac{t}{\tau_i^G}\right) \tag{A.4}
\]

is the shear relaxation modulus, where \( G_\infty \) is the long-term (fully-relaxed) modulus and
\( G_i^p \) and \( \tau_i^G \) are material properties. By noting that \( G_R(0) = G_0 = G_\infty + \sum_{i=1}^{N} G_i \),
equation A.4 becomes

\[
G_R(t) = G_0 - \sum_{i=1}^{N} G_i^p \left[1 - \exp\left(-\frac{t}{\tau_i^G}\right)\right] \tag{A.5}
\]

Equation A.5 may be normalized by \( G_0 \) to yield

\[
g_R(t) = \frac{G_R(t)}{G_0} = 1 - \sum_{i=1}^{N} g_i^p \left[1 - \exp\left(-\frac{t}{\tau_i^G}\right)\right] \tag{A.6}
\]

A similar expression may be formed for the bulk modulus in the case of volumetric strain.
Equation A.6 represents the Abaqus implementation of viscoelasticity by Prony series,
with the user specifying the coefficients \( g_i^p \) and relaxation times \( \tau_i^G \). These parameters
may be obtained from the constitutive equation for the desired material. For illustration
purposes, consider a Burghers body under shear deformation:

\[
\frac{2 \eta_1 \eta_2}{\mu_2} \ddot{\varepsilon} + 2 \eta_1 \dot{\varepsilon} = \frac{\eta_1 \eta_2}{\mu_1 \mu_2} \ddot{\sigma} + \left[\frac{\eta_1}{\mu_1} + \frac{\eta_1}{\mu_2} + \frac{\eta_2}{\mu_2}\right] \dot{\sigma} + \sigma \tag{A.7}
\]

where \( \eta_1 \) and \( \eta_2 \) are the steady-state and transient viscosities, respectively, and \( \mu_1 \) and
\( \mu_2 \) are the steady-state and transient shear moduli, respectively. If a strain step \( \varepsilon_0 \) is
applied at \( t = 0, \varepsilon = \varepsilon_0 H(t) \) where \( H(t) \) is the Heaviside step function, \( \dot{\varepsilon} = \varepsilon_0 \delta(t) \) where
\( \delta(t) \) is the Dirac delta function, and \( \ddot{\varepsilon} = \varepsilon_0 \frac{d\delta(t)}{dt} \). Equation A.7 becomes

\[
q_2 \varepsilon_0 \frac{d\delta(t)}{dt} + q_1 \varepsilon_0 \delta(t) = p_2 \ddot{\sigma} + p_1 \dot{\sigma} + \sigma \tag{A.8}
\]
where
\[ p_1 = \frac{\eta_1}{\mu_1} + \frac{\eta_1}{\mu_2} + \frac{\eta_2}{\mu_2}, \quad p_2 = \frac{\eta_1 \eta_2}{\mu_1 \mu_2}, \quad q_1 = 2\eta_1, \quad q_2 = 2\frac{\eta_1 \eta_2}{\mu_2}, \] (A.9)

Laplace-transforming equation A.8 yields
\[ q_2 \varepsilon_0 s + q_1 \varepsilon_0 = p_2 s^2 \dot{\sigma} + p_1 s \dot{\sigma} + \dot{\sigma}, \] (A.10)

and solving for \( \dot{\sigma} \),
\[ \dot{\sigma} = \frac{\varepsilon_0 (q_1 + q_2 s)}{1 + p_1 s + p_2 s^2} = \varepsilon_0 \hat{G}_R(s) \] (A.11)

where \( \hat{G}_R(s) \) is the Laplace-transformed shear relaxation modulus. Using the quadratic equation, the denominator \( 1 + p_1 s + p_2 s^2 \) becomes
\[ p_2 (s - \frac{-p_1 + \sqrt{p_1^2 - 4p_2}}{2p_2}) (s - \frac{-p_1 - \sqrt{p_1^2 - 4p_2}}{2p_2}) = p_2 (s + r_1)(s + r_2) \] (A.12)

where \( r_1 = \frac{p_1 - A}{2p_2}, \quad r_2 = \frac{p_1 + A}{2p_2}, \) and \( A = \sqrt{p_1^2 - 4p_2} \). \( \hat{G}_R(s) \) may then be expanded by partial fractions:
\[ \hat{G}_R(s) = \frac{q_1 + q_2 s}{1 + p_1 s + p_2 s^2} = \frac{\frac{1}{p_2}(q_1 + q_2 s)}{(s + r_1)(s + r_2)} = X + Y. \] (A.13)

Equating the numerators and combining like terms:
\[ \frac{q_1}{p_2} + \frac{q_2}{p_2} = X(s + r_2) + Y(s + r_1) \Rightarrow \frac{q_1}{p_2} = Xr_2 + Yr_1, \quad \frac{q_2}{p_2} = X + Y. \] (A.14)

Substituting for \( X = \frac{q_2}{p_2} - Y \) and solving for \( Y \) (noting that \( r_1 - r_2 = -\frac{A}{p_2} \)),
\[ \frac{q_1}{p_2} = \left( \frac{q_2}{p_2} - Y \right)r_2 + Yr_1 \Rightarrow Y = \frac{q_1 - q_2 r_2}{p_2(r_1 - r_2)} = -\frac{q_1 - q_2 r_2}{A}. \] (A.15)

Substituting for \( Y \) (noting that \( p_2 = -\frac{A}{r_1 - r_2} \)):
\[ X = \frac{q_2}{p_2} - Y = \frac{q_2}{p_2} + \frac{q_1 - q_2 r_2}{A} \Rightarrow X = \frac{q_1 - q_2 r_1}{A}. \] (A.16)

Inserting the expressions for \( X \) and \( Y \) into equation A.13,
\[ \hat{G}_R(s) = \frac{\frac{q_1 - q_2 r_1}{A}}{s + r_1} + \frac{\frac{q_1 - q_2 r_2}{A}}{s + r_2}. \] (A.17)
Inverse Laplace-transforming equation A.17 yields the shear relaxation modulus:

\[ G_R(t) = \frac{q_1 - q_2 r_1}{A} \exp(-r_1 t) - \frac{q_1 - q_2 r_2}{A} \exp(-r_2 t) = C \exp(-r_1 t) + D \exp(-r_2 t) \]  

(A.18)

where \( C = \frac{q_1 - q_2 r_1}{A} \) and \( D = \frac{q_2 r_2 - q_1}{A} \). Manipulating equation A.18,

\[ G_R(t) = (C + D) - [C(1 - \exp(-r_1 t)) + D(1 - \exp(-r_2 t))]. \]  

(A.19)

Equation A.19 is a two-term Prony series of the form of equation A.5. Normalizing by \( C + D \) yields the dimensionless shear relaxation modulus in the form utilized by Abaqus, equation A.6:

\[ g_R(t) = \frac{G_R(t)}{G_0} = 1 - \sum_{i=1}^{2} g_p^i [1 - \exp(-\frac{t}{\tau_G^i})] \]  

(A.20)

with \( G_0 = C + D \), \( g_p^1 = \frac{C}{C+D} \), \( g_p^2 = \frac{D}{C+D} \), \( \tau_G^1 = \frac{1}{r_1} \), and \( \tau_G^2 = \frac{1}{r_2} \). These parameters may be included in an Abaqus input file under the *VISCOELASTIC, TIME=PRONY keyword to configure a Burghers viscoelastic rheology. Parameters for other linear viscoelastic rheologies may be derived in a similar fashion; Findlay et al. [1989] Table 5.1 contains expressions of the relaxation moduli for some of these rheologies.

Two additional details must be considered in the implementation of Prony series viscoelasticity in Abaqus. First, elastic moduli as defined with the *ELASTIC keyword are by default long-term moduli. The option MODULI=INSTANTANEOUS must be included in order for the specification of \( G_0 \) to accurately reflect the instantaneous relaxation modulus. Secondly, if \( \sum_{i=1}^{N} g_p^i = 1 \), as \( t \to 0 \), \( \sum_{i=1}^{N} g_p^i [1 - \exp(-\frac{t}{\tau_G^i})] \to 1 \), and thus \( g_R(t) \to 0 \), meaning that the material has no long-term strength and behaves as a fluid on long timescales. While zero long-term strength is an intrinsic feature of some viscoelastic materials (e.g. Maxwell and Burghers), an Abaqus/Standard analysis is not built to handle fluid behavior and will crash during preprocessing under such conditions. The error may be bypassed by the simple expedient of slightly decreasing one of the coefficients \( g_p^i \), such that \( \sum_{i=1}^{N} g_p^i \) will be slightly less than unity as \( t \to 0 \) without adversely affecting the finite-time results of the analysis.
A.7 Avoiding thermal runaway in coupled thermomechanical calculations

In faulting simulations incorporating temperature-dependent rheology, simulated long-term fault slip may generate extremely high stresses near the fault tip if the material in that region is mechanically strong (e.g. dry diabase). These stresses may in turn produce extremely large dissipative temperature perturbations. The exponentially-temperature-dependent effective viscosity is then very highly reduced, which leads to extremely high strain rates. As a result, Abaqus may attempt to localize deformation within a very small area that may not be resolvable by a computationally-reasonable finite element mesh. This ‘thermal runaway’ may not actually lead to a model crash via excessive time increment cutbacks, because the dissipative temperature perturbation also generates very large thermal gradients over very short distances. Thermal conduction thus operates very quickly and may effectively counterbalance or even dominate dissipative heating if the time increment is sufficiently small. Such a situation can lead to a repeating mode in which a) Abaqus will calculate in a given time increment a large dissipative temperature perturbation near the fault tip and thus a very large strain increment, and will fail to converge; b) the time increment will be reduced as necessary and thus so too will be the magnitude of temperature perturbation and strain increment, and the increment will converge; c) the dissipative temperature perturbation drives rapid conductive cooling, which reduces the effective viscosity and limits the strain increment. This produces rapid convergence for the shorter time increment and a subsequent time increment increase, leading back to (a). The simulation may cycle within this mode until the user-specified step time is met, ultimately leading to a successfully-completed analysis that produces nonsensical results. An excerpt from the .sta file of such an analysis may look like:

16 4784 1U 0 6 6 4.73e+14 7.39e+09 3.112e+07
16 4784 2 0 1 1 4.73e+14 7.40e+09 7.780e+06
16 4785 1 0 1 1 4.73e+14 7.41e+09 1.556e+07
16 4786 1U 0 6 6 4.73e+14 7.41e+09 3.112e+07
16 4786 2 0 1 1 4.73e+14 7.42e+09 7.780e+06
16 4787 1 0 1 1 4.73e+14 7.44e+09 1.556e+07

The first and fourth lines represent step (a) of the cycle, the second and fifth lines step (b), and the third and sixth lines step (c). Such a situation may be prevented through the prescription of an initial temperature step that will reduce the effective viscosity of the material near the fault tip. For the purpose of stabilizing a simulation of long-
term fault slip with a model incorporating a dry diabase lower crust and a dry olivine upper mantle [Takeuchi and Fialko, 2010] under the assumption of zero heat flux through all model boundaries but the top surface, we applied an initial perturbation of 250°C within a 3 km-wide slab surrounding the fault from 10 to 17 km depth, and linearly decreasing towards the top and bottom of the model domain. This perturbation was added to the initially linear one-dimensional (1D) temperature gradient of 20 °C/km. The particular structure of the initial perturbation is of relatively minor importance, as long as two criteria are met: a) the magnitude of the perturbation is large enough to mitigate thermal runaway and stabilize convergence, and b) the perturbation diffuses to negligible levels over the timescale of the long-term simulation. Criterion (b) may be tested by applying the initial temperature structure to a fully static model (i.e. no motion boundary conditions) with identical heat flux boundary conditions, and allowing heat conduction to adjust the thermal structure over the intended long-term timescale in the absence of fault slip and viscous dissipation.

A.8 Locking a fault without displacement/velocity boundary conditions or constraint equations

Previous earthquake cycle models have generally been antiplane-strain configurations that incorporate infinitely-long faults. Finite-length rupture models used to study three-dimensional postseismic deformation generally utilize a single, embedded coseismic rupture followed by one interseismic period. However, it may be of interest to combine the two approaches- i.e. build a tectonic stress field via a cycle of far-field loading and infinitely-long ruptures with an antiplane model, followed by a single finite-length rupture. Such an approach requires that the entire fault be able to slip during the infinite-length ruptures, but also that the up- and down-strike extensions of the fault be locked during the finite-length rupture.

An Abaqus antiplane simulation may be run with a one-sided fault (i.e. a quarter-space with the fault located on one of the sides) because no fault-perpendicular or vertical deformation is allowed. Fault locking in such models is achieved using a simple zero-displacement or -velocity boundary condition (Abaqus keyword *BOUNDARY) on the fault-parallel displacement degree of freedom. Since the locking is applied with a boundary condition, it may be removed or added as needed from analysis step to analysis step. However, with a finite-length rupture, it is desirable to allow the fault to move in the
three spatial directions even while locked. In particular, fault-normal deformation on the
up- and down-strike extensions of a finite-rupture may be used to distinguish viscoelastic
postseismic relaxation from relaxation by afterslip [Hearn, 2003]. A zero-displacement
or -velocity boundary condition is thus too restrictive to be used to lock the fault in
finite-rupture models.

APMODEL, the Matlab software written by Shelley Kenner for use in designing
Abaqus earthquake models, typically uses linear constraint equations (Abaqus keyword
\*EQUATION) to lock two-sided faults. For example, to lock two adjacent faces of the
down-dip extension of the fault, the linear constraint

\*EQUATION

\[FLTEXT1W,3,1.00,FLTEXT1E,3,-1.00\]

may be used to constrain the z-displacement degree of freedom (DoF 3) for node sets
\(FLTEXT1W\) and \(FLTEXT1E\). This constraint specifies that the z-displacements for each
pair of corresponding nodes in \(FLTEXT1W\) and \(FLTEXT1E\) vary equally during the
simulation. Adding similar constraint equations for the x- and y-displacement degrees of
freedom (DoFs 1 and 2) fully locks the two node sets together. Such constraints on the
down-dip extensions of the fault are acceptable for a finite-length rupture model because
that part of the fault plane never undergoes any relative displacement. Similarly, three-
dimensional constraint equations are acceptable on the up- and down-strike extensions
of the finite rupture, if a single finite-length rupture is simulated. However, if during
any step in an earthquake cycle simulation the up- and down-strike fault extensions are
allowed to slip, but must be locked in another step, constraint equations cannot be used
for these planes. This is because linear constraint equations are model data, and thus
cannot be modified from their initial definitions at any future point in the simulation.

The up- and down-strike fault extensions may instead be locked and unlocked as
needed through the use of contact pairs (Abaqus keyword \*CONTACT PAIR). Contact
pairs may be defined on Abaqus surfaces, whether they be node-based or element based.
In APMODEL, contact pairs are automatically configured for corresponding fault planes
using element-based definitions. For example, to define the element set to be used for
setting up the contact pair:

\*ELSET,ELSET=FLTSRF1E,GENERATE

\[128776,128791,1\]
*ELSET defines the element set, named FLTSRF1E, and each successive line defines the specific elements to be included in the set. Following the FLTSRF1E element set definition,

*SURFACE,TYPE=ELEMENT,NAME=FLTPLN1E

FLTSRF1E

specifies that FLTSRF1E be used to define a free surface FLTPLN1E on the fault-side face of the element set. Once the corresponding fault surface (FLTPLN1W) has also been defined, the contact pair is specified:

*CONTACT PAIR,INTERACTION=FLT_1,SML SLIDING, EXTENSION ZONE=.1

FTPLN1W,FTPLN1E

In this statement, FLTPLN1W is defined as the slave surface, and FTPLN1E as the master surface. The options SMALL SLIDING and EXTENSION ZONE are APMODEL defaults and may be modified or removed with caution. The required INTERACTION parameter defines how the surfaces in the contact pair interact via:

*SURFACE INTERACTION,NAME=FLT_1

*SURFACE INTERACTION allows various properties, among them *FRICTION and *SURFACE BEHAVIOR, which may be used to lock the fault:

*FRICTION,ROUGH

*SURFACE BEHAVIOR, NO SEPARATION

*FRICTION,ROUGH specifies that no slip be allowed between the surfaces in the contact pair, which locks the fault in the vertical and fault-parallel directions. All definitions up to this point (*ELSET, *SURFACE, *CONTACT PAIR, *SURFACE INTERACTION, *FRICTION) are automatically configured by APMODEL. Adding *SURFACE BEHAVIOR, NO SEPARATION to *SURFACE INTERACTION prevents any opening between the surfaces of the contact pair once contact has been established. This locks the fault in the fault-normal direction. Thus the combination of *FRICTION,ROUGH and SURFACE BEHAVIOR, NO SEPARATION in the surface interaction fully locks the fault.

The contact pair governing fault locking may be added or removed within a given
analysis step using *MODEL CHANGE with the option ADD or REMOVE included. For instance:

*MODEL CHANGE,TYPE=CONTACT PAIR,ADD
FLTPLN1W,FLTPLN1E

To illustrate the procedure by example, let FLTPLN1W and FLTPLN1E be the surfaces of the finite rupture, F2PLN1W and F2PLN1E be the surfaces of the up-strike rupture extension, and F3PLN1W and F3PLN1E be the surfaces of the down-strike rupture extension. The full fault is thus the union of the three individual fault patches. During the finite rupture, F2PLN1W/F2PLN1E and F3PLN1W/F3PLN1E must be locked together, and the entire fault length must be locked for the duration of the subsequent relaxation period. However, preceding the finite rupture, a simulation of long-term slip on the full-fault-length is to be run, which requires that the entire fault be unlocked.

The contact pairs for each pair of fault planes should be set up as model data in the input file as described above:

*ELSET,ELSET=FLTSRF1E,GENERATE
...
*ELSET,ELSET=FLTSRF1W,GENERATE
...
*SURFACE,TYPE=ELEMENT,NAME=FLTPLN1E
FLTSRF1E,
*SURFACE,TYPE=ELEMENT,NAME=FLTPLN1W
FLTSRF1W,
*CONTACT PAIR,INTERACTION=FLT_1,
SMALL SLIDING,EXTENSION ZONE=0.1
FLTPLN1W,FLTPLN1E
*SURFACE INTERACTION,NAME=FLT_1
*FRICITION,ROUGH
*SURFACE BEHAVIOR, NO SEPARATION
with similar statements for F2PLN1W, F2PLN1E, F3PLN1W, and F3PLN1E. There will be a total of 7 analysis steps in the simulation:

1) Long-term slip on the full-length fault
2) Establishment of contact
3) Coseismic rupture on the finite-length fault
4) Establishment of contact
5) Postseismic relaxation

Steps 2 and 4 are required for the contact pairs to be correctly engaged. In step 1, the contact pairs established in the model data section must be removed in order for slip to be applied. This is done by adding to the step definition:

```plaintext
*MODEL CHANGE, TYPE=CONTACT PAIR, REMOVE
FLTPLN1W, F2PLN1W, F3PLN1W
FLTPLN1E, F2PLN1E, F3PLN1E
```

In step 2, the contact pairs must be reestablished using:

```plaintext
*MODEL CHANGE, TYPE=CONTACT PAIR, ADD
FLTPLN1W, F2PLN1W, F3PLN1W
FLTPLN1E, F2PLN1E, F3PLN1E
```

In this step, the fault surfaces will be adjusted so they are precisely in contact with each other, such that any overclosures or openings accrued during the previous step are eliminated. This is done using:

```plaintext
*CONTACT INTERFERENCE, OP=NEW, TYPE=CONTACT PAIR
FLTPLN1W, F2PLN1W, F3PLN1W, -1.000E-04
FLTPLN1E, F2PLN1E, F3PLN1E, -1.000E-04
```

The value -1.000e-4 is the allowable interference and may be modified as needed. See the ’Modeling contact interference fits in Abaqus/Standard’ section in the Abaqus Analysis User’s Manual for more details. Contact may be established over multiple increments within the step by specifying that two or more time increments be used in the step definition, for example:

```plaintext
*STEP, AMPLITUDE=STEP, INC=2
*VISCO, CETOL=1.00E-05
5.00E-11, 1.00E-10, 0.000E+00
```

In step 3, the finite-rupture fault must be unlocked using:

```plaintext
*MODEL CHANGE, TYPE=CONTACT PAIR, REMOVE
FLTPLN1W, F2PLN1E
```
Step 4 reestablishes contact on the finite-length fault as in step 2:

*MODEL CHANGE, TYPE=CONTACT PAIR, ADD
FLTPLN1W, FLTPLN1E

*CONTACT INTERFERENCE, OP=NEW, TYPE=CONTACT PAIR
FLTPLN1W, FLTPLN1E, -1.000E-04

In step 5, all sections of the fault are already locked, so no further *MODEL CHANGE or *CONTACT INTERFERENCE statements are required.

Three potential issues with this approach exist that will cause warnings, errors, and/or spurious results if neglected, all involving overconstraints. First, APMODEL configures the fault surfaces so that they intersect (i.e. share nodes) at the ends of the finite rupture. In the example above, FLTPLN1E and F2PLN1E intersect at the up-strike end of FLTPLN1E, and FLTPLN1E and F3PLN1E intersect at the down-strike end of FLTPLN1E. The same follows for FLTPLN1W, F2PLN1W, and F3PLN1W. The contact pairs between corresponding fault plane pairs will then produce overconstraints at these surface intersections. The overconstraints will generate warning messages in the .msg file (including zero pivots and/or Lagrange multipliers) for each node where the overconstraints apply, and the contact pairs will fail at these nodes. To remove the overconstraints, the fault surfaces may be trimmed where they intersect. Trimming the element sets comprising the finite rupture (FLTPLN1E and FLTPLN1W) by one element is logical, because these surfaces will then be concurrent with the node sets upon which coseismic slip is applied (FLTSMC1E and FLTSMC1W).

Secondly, APMODEL automatically configures two node sets FSIDE and BSIDE corresponding to the fault-normal model boundaries. It is frequently desirable to constrain these node sets to antiplane conditions using a linear constraint equation on their fault-parallel degrees of freedom. However, where these node sets intersect with the up- and down-strike fault surfaces (F2PLN and F3PLN in the example above), overconstraints will occur. The overconstraints will yield similar warning messages in the .msg file as the fault plane intersections discussed above, and will also yield failed contact pairs. The overconstraints may be removed by trimming from FSIDE and BSIDE any nodes that lie in the master surfaces of the fault plane contact pairs (F2PLN1W and F3PLN1W in the example above). It should also be noted that antiplane constraints on FSIDE and BSIDE will also conflict with APMODEL-configured linear constraint equations on the downdip extension of the fault plane (F2EXT and F3EXT in the example
above), and as such will produce an error. The offending nodes should be trimmed from \textit{FSIDE} and \textit{BSIDE} as well.

Finally, zero vertical or fault-normal displacement or velocity boundary conditions may also be desired for the fault-normal boundaries (to pin these boundaries, for instance). Where the node sets intersect with the contact pairs on the fault plane, an overconstraint will occur, and warning messages involving prescribed displacements tangent to the master surface will appear in the .msg file. The frictional component of the contact pair (\textit{*FRICITION,ROUGH}) will be changed to the default penalty method, which may lead to the fault being improperly locked. Such overconstraints may be avoided by ensuring that the boundary condition is applied to the same trimmed \textit{FSIDE} and \textit{BSIDE} node sets discussed above.

To summarize the solutions to these potential issues: ensure that any nodes included in the surfaces of a contact pair, with certain degrees of freedom restricted by that contact pair, do not have those degrees of freedom included in another contact pair, a linear constraint equation, or a boundary condition.
Bibliography


121


