

**Three-dimensional thermo-mechanical modeling of deformation at plate
boundaries: case study San Andreas Fault System**

Dissertation
zur Erlangung des akademischen Grades
"doctor rerum naturalium"
(Dr. rer. nat.)
in der Wissenschaftsdisziplin "Geophysik"

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät
der Universität Potsdam

vorgelegt von
Anton A. Popov

Potsdam, October 2008

Gutachter:

Prof. Dr. Michael Weber (GeoForschungsZentrum & Universität Potsdam)

Dr. habil. Taras Gerya (ETH Zurich)

Prof. Dr. Romain Bousquet (Universität Potsdam)

Tag der Disputation: 30 April 2009

This work is licensed under a Creative Commons License:
Attribution - Noncommercial - Share Alike 3.0 Germany
To view a copy of this license visit
<http://creativecommons.org/licenses/by-nc-sa/3.0/de/deed.en>

Published online at the
Institutional Repository of the University of Potsdam:
<http://opus.kobv.de/ubp/volltexte/2009/3187/>
[urn:nbn:de:kobv:517-opus-31875](http://nbn-resolving.org/urn:nbn:de:kobv:517-opus-31875)
[<http://nbn-resolving.org/urn:nbn:de:kobv:517-opus-31875>]

Abstract

It has always been enigmatic which processes control the accretion of the North American terranes towards the Pacific plate and the landward migration of the San Andreas plate boundary. One of the theories suggests that the Pacific plate first cools and captures the uprising mantle in the slab window, and then it causes the accretion of the continental crustal blocks. The alternative theory attributes the accretion to the capture of Farallon plate fragments (microplates) stalled in the ceased Farallon-North America subduction zone. Quantitative judgement between these two end-member concepts requires a 3D thermomechanical numerical modeling. However, the software tool required for such modeling is not available at present in the geodynamic modeling community.

The major aim of the presented work is comprised basically of two interconnected tasks. The first task is the development and testing of the research Finite Element code with sufficiently advanced facilities to perform the three-dimensional geological time scale simulations of lithospheric deformation. The second task consists in the application of the developed tool to the Neogene deformations of the crust and the mantle along the San Andreas Fault System in Central and northern California.

The geological time scale modeling of lithospheric deformation poses numerous conceptual and implementation challenges for the software tools. Among them is the necessity to handle the brittle-ductile transition within the single computational domain, adequately represent the rock rheology in a broad range of temperatures and stresses, and resolve the extreme deformations of the free surface and internal boundaries. In the framework of this thesis the new Finite Element code (SLIM3D) has been successfully developed and tested. This code includes a coupled thermo-mechanical treatment of deformation processes and allows for an elasto-viscoplastic rheology with diffusion, dislocation and Peierls creep mechanisms and Mohr-Coulomb plasticity. The code incorporates an Arbitrary Lagrangian Eulerian formulation with free surface and Winkler boundary conditions.

The modeling technique developed is used to study the aspects influencing the Neogene lithospheric deformation in central and northern California. The model setup is focused on the interaction between three major tectonic elements in the region: the North America plate, the Pacific plate and the Gorda plate, which join together near the Mendocino Triple Junction. Among the modeled effects is the influence of asthenosphere upwelling in the opening slab window on the overlying North American plate. The models also incorporate the captured microplate remnants in the fossil Farallon subduction zone, simplified subducting Gorda slab, and prominent crustal heterogeneity such as the Salinian block.

The results show that heating of the mantle roots beneath the older fault zones and the transpression related to fault stepping, altogether, render cooling in the slab window alone incapable to explain eastward migration of the plate boundary. From the viewpoint of the thermomechanical modeling, the results confirm the geological concept, which assumes that a series of microplate capture events has been the primary reason of the inland migration of the San Andreas plate boundary over the recent 20 Ma. The remnants of the Farallon slab, stalled in the fossil subduction zone, create much stronger heterogeneity in the mantle than the cooling of the uprising asthenosphere, providing the more efficient and direct way for transferring the North American terranes to Pacific plate.

The models demonstrate that a high effective friction coefficient on major faults fails to predict the distinct zones of strain localization in the brittle crust. The magnitude of friction coefficient inferred from the modeling is about 0.075, which is far less than typical values 0.6 – 0.8 obtained by variety of borehole stress measurements and laboratory data. Therefore, the model results presented in this thesis provide additional independent constrain which supports the “weak-fault” hypothesis in the long-term ongoing debate over the strength of major faults in the SAFS.

Zusammenfassung

Seit jeher rätselhaft sind die Prozesse, die die Akkretion der Nordamerikanischen Terranen in Richtung der Pazifischen Platte sowie die Wanderung der Plattengrenze der San-Andreas-Verwerfung in Richtung Festland bestimmen. Eine Theorie besagt, dass sich die Pazifische Platte erst abkühlt und den aufsteigenden Mantel im „Slab Window“ fängt und somit die Akkretion der kontinentalen Krustenblöcke bewirkt. Die andere Theorie geht von einer Akkretion durch das Fangen von Teilen der Farallon-Platte (Mikroplatten) aus, die in der inaktiven nordamerikanischen Farallon-Subduktionszone fest stecken. Die quantitative Beurteilung dieser beiden gegensätzlichen Konzepte erfordert eine thermomechanische numerische 3-D-Modellierung. Das dafür benötigte Software Tool steht jedoch der geodynamischen Modellierung derzeit noch nicht zur Verfügung.

Das Hauptziel der vorliegenden Arbeit umfasst im Wesentlichen zwei miteinander verbundene Aufgaben. Die erste besteht in der Entwicklung und Erprobung des Finite-Element-Codes, dessen Eigenschaften den hohen Anforderungen an die Ausführung der dreidimensionalen Simulationen lithosphärischer Deformation auf geologischer Zeitskala gerecht werden müssen. Die zweite Aufgabe ist die Anwendung des entwickelten Tools auf die neogenen Deformationen der Kruste und des Mantels entlang der San-Andreas-Verwerfung in Zentral- und Nordkalifornien.

Die Modellierung auf geologischer Zeitskala lithosphärischer Deformation bringt für die Software Tools in Bezug auf Konzept und Durchführung zahlreiche Herausforderungen mit sich. Unter anderem gilt es, den Brittle-Ductile-Übergang in einem einzigen Modell sowie die Gesteinsrheologie in einer breiten Spanne unterschiedlicher Temperaturen und Spannungen adäquat darzustellen und die extremen Deformationen der freien Oberfläche und internen Grenzen aufzulösen. Im Rahmen der vorliegenden Arbeit erfolgte die erfolgreiche Entwicklung und Erprobung des neuen Finite-Element-Codes (SLIM3D). Dieser Code beinhaltet eine gekoppelte thermomechanische Behandlung von Deformationsprozessen und ermöglicht eine elasto-visko-plastische Rheologie mit Diffusion, Dislokation, Peierls Kriechmechanismen und

Mohr-Coulomb-Plastizität. Der Code verbindet eine Arbitrary Lagrangian-Eulerian kinematische Formulierung mit freien Oberflächen- und Winkler-Randbedingungen.

Das entwickelte Modellierungsverfahren wird für die Untersuchung der Aspekte verwendet, die die neogene lithosphärische Deformation in Zentral- und Nordkalifornien beeinflussen. Die Modellanordnung konzentriert sich auf die Interaktion zwischen drei großen tektonischen Elementen in dieser Region: die Nordamerikanische Platte, die Pazifische Platte sowie die Gorda-Platte, die sich in der Mendocino-Triple-Junction treffen. Unter anderem verdeutlicht die Modellierung den Einfluss des Aufsteigens der Asthenosphäre in das sich öffnende „slab window“ der überlagerten Nordamerikanischen Platte. Die Modelle beziehen auch die angelagerten Überreste der Mikroplatten in der fossilen Farallon-Subduktionszone, die vereinfachte subduzierende Gorda-Platte sowie markante Heterogenitäten der Kruste, wie beispielsweise den „Salinian Block“, mit ein.

Die Ergebnisse zeigen, dass die Erwärmung der Mantellithosphäre unter den älteren Störungszonen sowie die Transpression eine Abkühlung im „Slab Window“ als alleinige Begründung für die Ostwärtsbewegung der Plattengrenze nicht zulassen. Aus Sicht der thermomechanischen Modellierung bestätigen die Ergebnisse das geologische Konzept, welches durch das mehrmalige Fangen von Mikroplatten den Hauptgrund für die Wanderung der Plattengrenze der San-Andreas-Verwerfung in Richtung Festland über die letzten 20 Millionen Jahre sieht. Die Überreste der Farallon-Platte, die in der fossilen Subduktionszone gefangen sind, verursachen im Mantel eine wesentlich stärkere Heterogenität als die Abkühlung der Asthenosphäre und stellen somit den effizienteren und direkteren Weg für die Anlagerung der nordamerikanischen Gebiete an die Pazifische Platte dar.

Die Modelle demonstrieren, dass ein hoher effektiver Reibungskoeffizient an großen Störungen nicht in der Lage ist, die eindeutigen Zonen der Dehnungslokalisierung in der spröden Kruste vorherzusagen. Die Größe des Reibungskoeffizienten, die sich aus der Modellierung ableitet, beträgt etwa 0,075 und ist damit wesentlich kleiner als die durch unterschiedliche Bohrlochmessungen und Labordaten ermittelten Spannungswerte zwischen 0,6 und 0,8. Daher liefern die in dieser Arbeit präsentierten Ergebnisse der Modelle in der seit langem geführten Debatte über die Stärke von großen Störungen in der San-Andreas-Verwerfung eine zusätzliche unabhängige Begründung der „Weak-Fault“-Hypothese.

Contents

1	Introduction	1
2	Dynamics of the San Andreas Fault System	5
2.1	Plate tectonics history	5
2.2	Slab window/gap	10
2.3	Stress field and fault strength	13
2.4	Lithospheric structure	17
2.5	Present day plate motion	21
3	Review of previous modeling	23
3.1	Thin-shell approximation	23
3.2	Two-dimensional models	29
3.3	Three-dimensional models	33
3.4	Extended two-dimensional models	39
4	Numerical modeling techniques	46
4.1	Introduction	46
4.2	Physical models	49
4.3	Numerical algorithms	55
4.4	Benchmarks and examples.	70
4.5	Summary and discussion	82
5	Three-dimensional thermo-mechanical models	84
5.1	Setup of the models	84
5.2	Model results with Gorda slab and slab window	92
5.3	Model results with Monterey microplate	98

6	Summary and discussion	114
6.1	Summary	114
6.2	Discussion	115
6.3	Outlook	117
	Appendix	118
	References	119
	List of Figures	126
	Acknowledgments	128
	Curriculum Vitae	129

Chapter 1

Introduction

Earth does not have a ridged lid unlike the other terrestrial planets. The system of lithospheric plates separated by the special zones, plate boundaries (Fig. 1.1), participates in the global mantle flow (e.g. Bullard et al., 1965; Wilson, 1965; McKenzie and Parker, 1967; Le Pichon, 1968; Morgan, 1968). From the mechanical point of view, the plate boundaries play a major role in controlling the character and style of plate tectonics (e.g. Tackley, 1998).

Complex deformation processes at the plate boundaries are characterized by occurrence of earthquakes, volcanic eruptions, and mountain building. The geological time scale evolution of the subduction zones, continental transform faults and mid-ocean ridges is characterized by inherently three-dimensional style of deformation. This includes formation of normal, reverse, and strike-slip faults, transpression, transtension, and folding. Each of these types of tectonic activity is frequently accompanied by variation of crustal thickness, topography elevation and associated erosion processes. The complicated rheological effects such as nonlinear creep and strain weakening as well as metamorphic reactions and melt formation emerge mostly at the plate boundaries. All these complexities render numerical modeling of the lithospheric plate boundary evolution a challenging task. At present, there are very few numerical tools designed for 3D modeling of the plate boundaries. Recent advancements in this direction include 3D creeping flow model (DOUAR) with mesh adaptation and free surface (Braun et al. 2008), 3D versions (under development) of the 2D method (I2ELVIS) for modeling of visco-elasto-plastic problems with phase transformations (Gerya and Yuen, 2007), and the finite element tool (LAPEX3D) based on the explicit time stepping (Babeyko et al., 2002; Petrunin and Sobolev, 2006, 2008). Modeling is an ongoing effort. Despite noticeable progress, not all the difficulties posed by lithospheric-scale problems are resolved at present. For instance, neither of the above numerical codes can be readily applied for modeling of the long term (tens of millions years) 3D deformation with generation of new faults in the lithospheric domains of 1000 km size.

This work is aimed at developing of a similar advanced numerical technique, which can be applied for studying the complex three-dimensional deformation at the plate boundaries on

Crustal plate boundaries

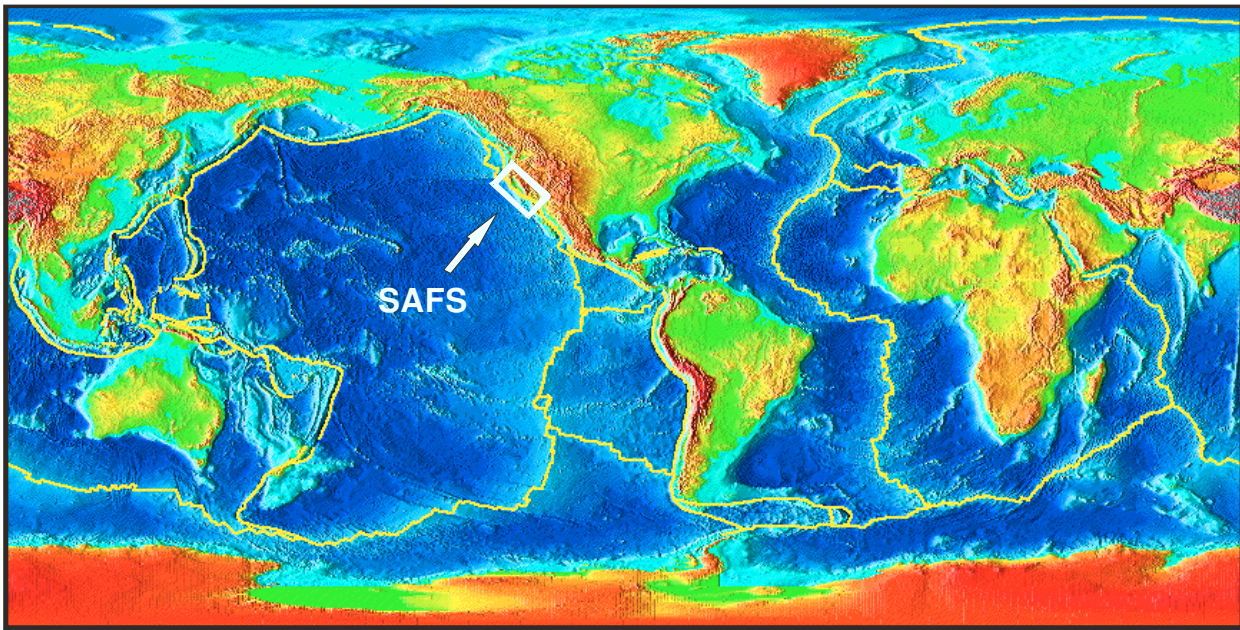


Figure 1.1 Global maps of crustal plate boundaries. Positions of plate boundaries are indicated by yellow lines. Also shown is the location of the San Andreas Fault System. The background map is taken from Earth's Interior & Plate Tectonics website (<http://www.solarviews.com>).

geological time scale and in regional and global spatial scales. The new code should include the thermo-mechanical coupling and free surface deformation. It should also be able to resolve strain localization with nonlinear plastic and stress-dependent creep rheology, and allow for large deformations. The development of such advanced technique can hardly be performed in isolation from a real challenging geodynamic application. In present thesis the Neogene deformation of the San Andreas Fault System in central and northern California is selected as such application. This modeling constitutes a second major task of the thesis.

The San Andreas Fault System (SAFS) is characterized by many complicated tectonic effects. In particular it includes subducting Gorda plate to the north of Mendocino triple Junction, and strike-slip deformation on the numerous sub-parallel faults throughout the entire California western margin (Fig 1.2). For example at the latitude of San Francisco, the entire relative strike-slip velocity between the plates is accommodated on Hayward, Calaveras, San Andreas, and San Gregorio faults (Prescott et al., 2001). Similar situation is observed in southern California, where the San Jacinto and Elsinore faults are active. Another peculiar feature of the SAFS is a strong transpressive bend, near the Frazier Park in Southern California, which is called “big-bend”.

At present the SAFS serves as a boundary between the North America and Pacific plates. The system formed in response to collision of the Farallon ridge with North America plate margin during the Oligocene (e.g. Atwater, 1989). In the context of ridge-trench collision, two

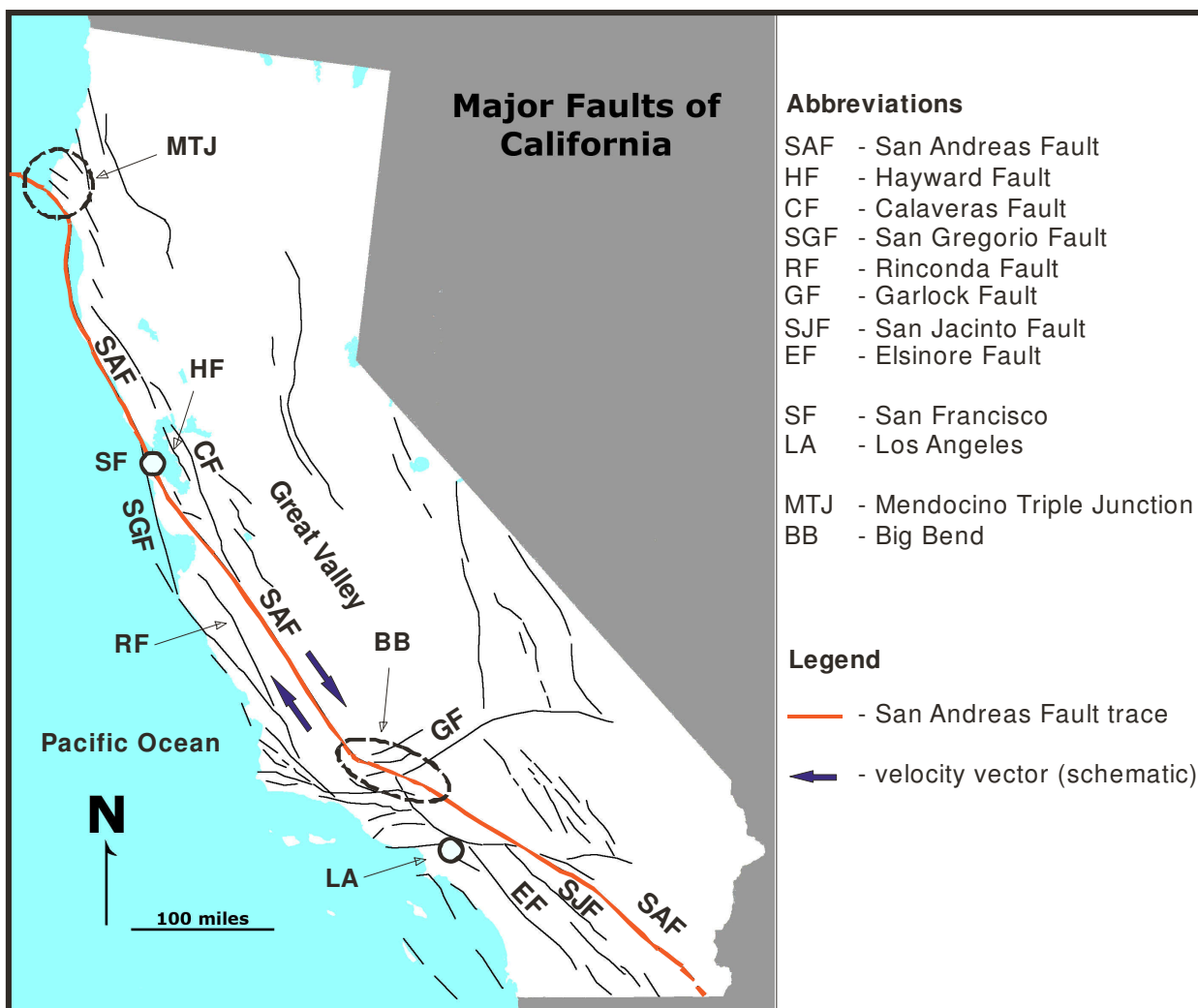


Figure 1.2 Map of major faults of the San Andreas Fault System in California. The background map is taken from USGS Education website (<http://education.usgs.gov/california/resources.html>).

basic concepts can be considered: the so-called “slab window” hypothesis (Dickinson and Snyder, 1979) and the “stalled slab” hypothesis (Bohannon and Parsons, 1995). The difference between the concepts consists in a fact whether or not the remnants of subducting Farallon plate get stuck in the fossil subduction zone. Accordingly, two different scenarios can be proposed to explain the accretion of the North America terranes to Pacific plate. In the slab window model, the accretion is caused by conductive cooling of the upwelling mantle (Furlong et al., 1989), while the microplate model explains terrane accretion by the increasing coupling between the captured microplate and the continent via extinct subduction interface (Nicholson et al. 1994). This thesis is aimed at qualifying both these concepts from the viewpoint of advanced three-dimensional thermo-mechanical modeling. Among the other considered effects are the style of material flow in the slab window and the magnitude of the effective friction on the fault.

The structure of this thesis is as follows.

Chapter 2 presents the aspects of the San Andreas Fault System dynamics, which are important in the context of this study. This includes the tectonic history, crustal and lithospheric

structure, stress field and the plate velocities.

Chapter 3 briefly summarizes the results of the previous similar models.

Chapter 4 extensively describes the implementation details and benchmarking of the three-dimensional numerical modeling technique. This chapter is based on the results published by Popov and Sobolev (2008).

Chapter 5 demonstrates, interprets and discusses the modeling results of the Neogene deformation of the San Andreas Fault System in central and northern California. In particular, the analysis of the “slab window” and the “stalled slab” concepts is carried out. This chapter also shows the style of material flow in the slab window inferred from the modeling, and provides independent constraint on the effective friction on major faults in the SAFS.

Chapter 6 presents general conclusion, and outlook of the thermo-mechanical modeling.

Chapter 2

Dynamics of the San Andreas Fault System

2.1 Plate tectonics history

The present-day essentially transform regime of the San Andreas plate boundary preceded by the subduction of the Farallon plate beneath the edge of the North American continent in the Mesozoic and Cenozoic eras (e.g., Atwater, 1989). The major geological manifestations of this process in central California include granitic batholith belt, the foothills metamorphic belt, the Great Valley belt of forearc sediments, and the Franciscan subduction accretionary wedge complex (e.g. Irwin, 1990). At about 27 Ma the first segment of the East Pacific Rise intersected the trench and the transition from Farallon-North America subduction to the Pacific-North America transtensional margin began (Atwater, 1989; Atwater and Stock, 1998). Since then, the entire plate boundary has evolved from a short simple contact at the subduction interface to the present complex zone (over 2300 km long) that partially resides within the continent and partially includes a small new ocean in the Gulf of California (Atwater and Molnar, 1973; Lonsdale, 1991).

Microplate capture

During collision of the East Pacific Rise with western margin of North America, the Farallon plate began to fragment into various microplates, including the Monterey (27 – 19 Ma), Arguello (20 - 18 Ma), Guadalupe (20 – 14 Ma), Magdalena (14 – 12 Ma) and Riviera plates (5 Ma - present) (Lonsdale, 1991; Stock and Lee, 1994). The evolution of the SAFS plate boundary can be represented as a series of plate-capture events (Nicholson et al., 1994). The simplified tectonic model is shown in (Fig. 2.1). Each time a ridge segment died offshore, the corresponding partially subducted fragment of the Farallon plate assumed the rapid west motion of the Pacific plate. The mechanism of such capture, guessed by Nicholson et al. (1994), was subsequently considered from the quantitative point of view by Bohannon and Parsons (1995). The positive

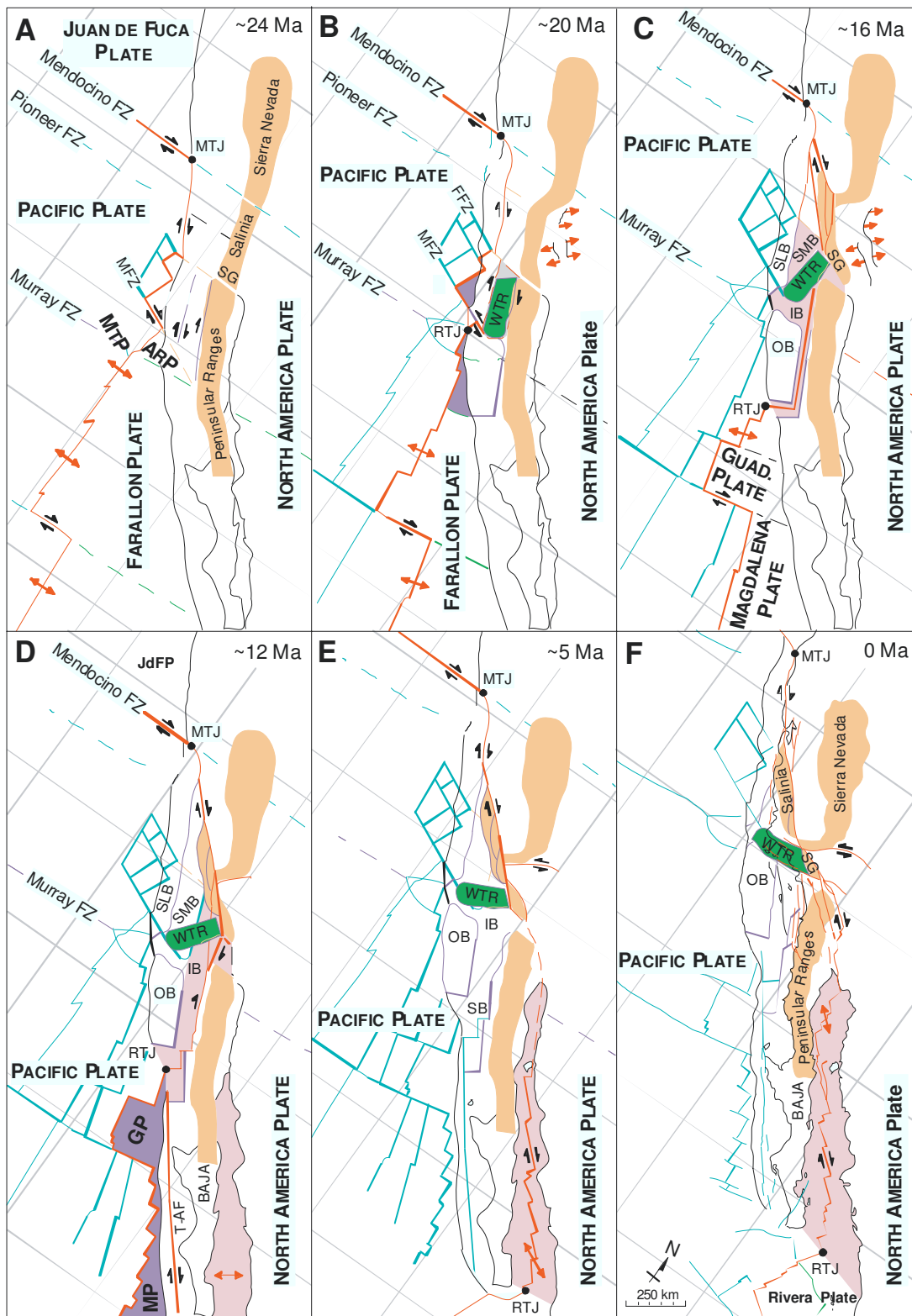


Figure 2.1 Simplified tectonic model of North America Pacific interactions since 24 Ma; modified after Nicholson et al. (1994). Model assumes constant rate and direction of Pacific Plate motion and constant rate of Western Transverse Ranges (WTR) rotation. When partially subducted Monterey (20 Ma), Arguello (17.5 Ma), Guadalupe and Magdalena (12 Ma) are captured, part of North America upper plate is transferred to Pacific plate. Fine gray lines provide reference grid tied to fixed North America. The abbreviations are: JdFP – Juan de Fuca plate; MtP – Monterey plate; ArP – Arguello plate; GP – Guadalupe plate; MP – Magdalena plate; SG – San Gabriel block; SLB – Santa Lucia Bank; SMB – Santa Maria basin; IB, OB, SB – inner, outer and southern borderland, respectively; T-AF – Tosco Abrejos Fault.

buoyancy of young partially subducted microplate and its increased coupling with overriding continent leads to detachment from older sinking slab. The detachment causes cessation of subduction as well as seafloor spreading. Thus, the microplate becomes effectively “welded” to Pacific plate in just a few Myr by conductive cooling over ceased ridge segment (Bohannon and Parsons, 1995).

The microplate capture model provides a direct mechanism for the transfer of the North America continental crust to the Pacific plate. The numerical models of Sobolev and Babeyko (2005) infer that effective strength of typical subduction interface is relatively low compared to laboratory measurements of the rock strength. This might be related to the intensive deformation and/or presence of the fluids in subduction channel. When microplate subduction ceases or substantially decelerates, both these factors diminish, which leads to strengthening of the extinct subduction interface. This mechanism causes pieces of the North America that overlay captured microplate to join the motion of Pacific plate. Thus a boundary between Pacific and North America propagate further landward.

The first event in a series was capture of Monterey microplate, which was completed at about 19 Ma. Nicholson et al. (1994) suggested that this event was responsible for Neogene rotation of Western Transverse Ranges as large coherent crustal block (Luyendyk et al, 1980, Luyendyk, 1991) as well as capture of the Santa Lucia Bank and Outer Borderland blocks by Pacific. At about 19-18 Ma, with the stalling of the offshore Monterey microplate, Pacific-North America plate boundary was almost 700 km in length, stretching from Cape Mendocino to Santa Lucia Bank (Nicholson et al., 1994; Atwater and Stock, 1998). Capture of Arguello microplate occurred about 17-18 Ma (Lonsdale, 1991), shortly after Monterey microplate, which resulted in a rifting and translation of the Outer Borderland from northern Baja California. The complete capturing process of the Outer Borderland took around 2 Myr. During this time it became separated from Santa Lucia Bank by around 100 km (Nicholson et al., 1994).

About 12 Ma, the southern end of the plate boundary was transferred farther to the south when spreading and subduction stalled offshore of southern Baja California. This was related to capture of the Guadalupe and Magdalena microplates that were partially subducted beneath Baja California. However the plate boundary remained on the Pacific side of Baja another 7 Myr. During this period the transtensional motion between Pacific plate and North America plate was essentially partitioned into strike-slip motion along Tosco-Abrejos fault and extension within incipient Gulf of California (Nicholson et al., 1994).

About 5 Ma capturing of Baja by Pacific was completed and plate boundary jumped eastward in the Gulf of California (e.g. Lonsdale, 1991). Simultaneously, the tectonic regime of southern California began to change from transtensional to transpressional when Baja, moving

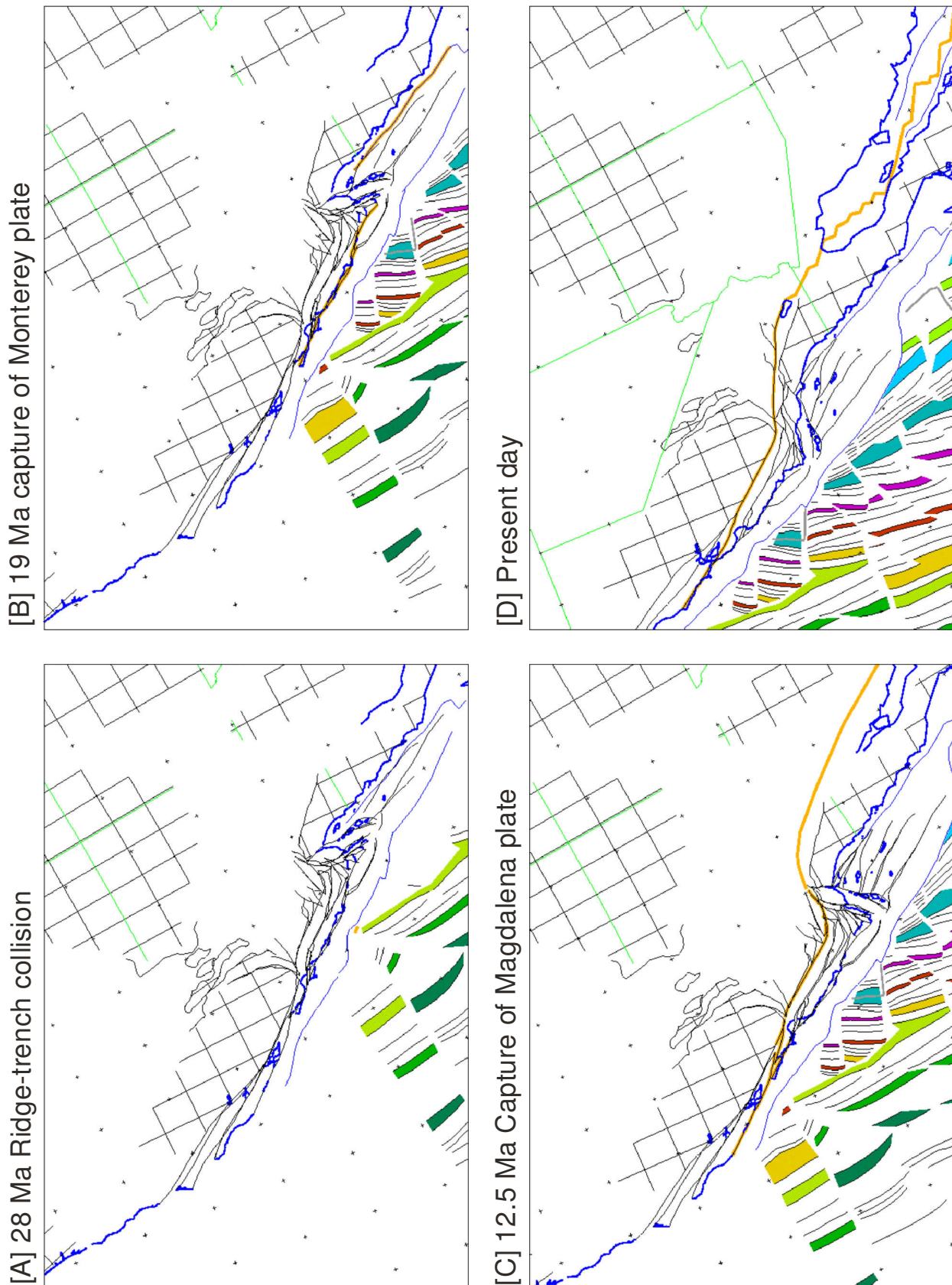


Figure 2.2 Snapshots from the animated reconstructions by Wilson et al. (2005). Stable North America is held fixed. Bold line highlights the fastest slipping fault in continental crust active or becoming active at the reconstruction time. Colors indicate position of the isochrones (similar to Fig. 2.3). Major reorganizations are interpreted to be simultaneous with capture of the Monterey plate at 19 Ma [B] and of the Magdalena plate at 12.5 Ma [C]. In this model, the Big Bend in the San Andreas Fault System formed during the 12.5 Ma reorganization.

synchronously with Pacific, began to ram into its southern edge. Present southern segments of the San Andreas Fault System, including two transpressional bends (the Big Bend and the San Bernardino Mountains bend) and left-lateral Garlock fault, were formed. Since 12 Ma the Baja California rotated about 10° as it was rifted and transported away from the North America. The Plate capture model predicts that relative motion with respect to North America constitutes about 400 km, with about 150 km occurring prior to capture of Baja (5 Ma) (e.g. Nicholson et. al., 1994). At present the full Pacific-North America plate motion accommodates by active ridge system in the Gulf of California.

Plate reconstructions

A number of reconstructions of the tectonic evolution of the San Andreas plate boundary are available in literature (Stock and Molnar, 1988; Atwater and Severinghaus, 1989; Bohannon and Parsons, 1995; Atwater and Stock, 1998, Wilson et al., 2005). The methodology of such reconstructions is based on combination of offshore marine magnetic anomalies with global plate circuit solutions. Additional constraints are provided from the analysis of the onshore volcanic record (Dickinson, 1997; Wilson et al., 2005). Results of the reconstructions show that Pacific-North American relative plate motion was never coast-parallel but rather always included significant offshore component. At about 8 Ma it changed from $N60^\circ W$ to $N37^\circ W$, such that it became closer to coast-parallel but still offshore (Atwater and Stock, 1998)

There are two animated reconstructions of the San Andreas Fault System, proposed by Atwater and Stock (1998) and Wilson et al. (2005), respectively. Both reconstructions have no principal differences and correlate with each other relatively well. The most significant difference contains in cumulative translation of the Baja California with respect to stable interior of the North America. Reconstruction of past positions of costal volcanic units from 27 to 12 Ma of Wilson et al. (2005) strongly supports at least 500 km of displacement. A direct consequence of this interpretation is that the approximate geometry of the current Pacific-North American plate boundary in California and western Mexico developed at about 10–12 Ma.

Fig 2.2 Show an example of snapshots from the reconstruction by Wilson et al. (2005). This reconstruction is based on more rigorous restoration of continental deformation. Their model for North America includes a few major blocks generally recognized as mobile but nearly rigid, namely, the Colorado Plateau, Sierra Nevada-Great Valley, Baja California, and the Sierra Madre Occidental. Many individual blocks bounded by major to moderate faults are tracked in the areas where the volcanic rocks are observed. For the Basin and Range extension, which is a characteristic inland deformation zone, Wilson et al. (2005) consider only the motion between the major blocks and the North America craton. They used an updated global plate circuit based

on recent documentation of non-rigid behavior of the African and Antarctic plates (e.g. Cande et al., 2000).

The reconstruction by Atwater and Stock (1998) placed various constraints on the time space budget of deformation within western North America. The most important aspect is related to the fact that the overlap between the oceanic and continental lithosphere is prohibited. If the North America is kept in its present shape, the reconstruction results in significant overlaps. The magnitude of this overlap can be used as a constraint for the intra-plate deformation.

The displacements of many continental blocks and the offsets on many faults in the western North America are poorly known. Constructing a complete model of continental deformation still requires numerous subjective judgments. With presently available amount of objective constraints derived from various observations, these uncertainties are well resolved in the works of Atwater and Stock (1998) and Wilson et al. (2005).

2.2 Slab window/gap

The evolution of the plate boundary can be alternatively viewed in the context of migrating-triple-junction paradigm (Ingersoll, 1982; Severinghaus and Atwater, 1990; Atwater, 1970; Atwater and Molnar, 1973). The contact between the Pacific and North America plates lengthens as the Mendocino triple junction (MTJ) (northern end) and Rivera triple junction (RTJ) (southern end) migrate in the opposite directions along the coastline. It was proposed that a slab-free zone or slab window (slab gap) developed beneath North America following its contact with the Pacific plate (Dickinson and Snyder, 1979; Zandt and Furlong, 1982; Furlong et al., 1989; Severinghaus and Atwater, 1990). Existence of the void in the mantle which is filled with the hot uprising asthenospheric material follows directly from the plate kinematics and the rigid plate assumption (e.g. Furlong and Schwartz, 2004). Main stages of the slab window formation can be explained as shown in the Fig. 2.3. The process begins with the subduction of the segments of the mid-ocean ridge. At certain depth the spreading terminates and the subducting slab south of the triple junction detaches and sinks in the mantle. Another possibility for the slab is to get stuck beneath the overriding plate and begin to stretch (e.g. ten Brink et al., 1999). However it may be only important in the early stages of the slab window formation. As the slab north of the triple junction continues to subduct and move northward with respect to the overriding plate, the slab window (gap) progressively grows in size and get filled with the mantle material flowing from below.

The emplacement of cold subducting slab with the hot mantle material should have a direct influence on the thermal state, magmatic activity and the deformation of western North America

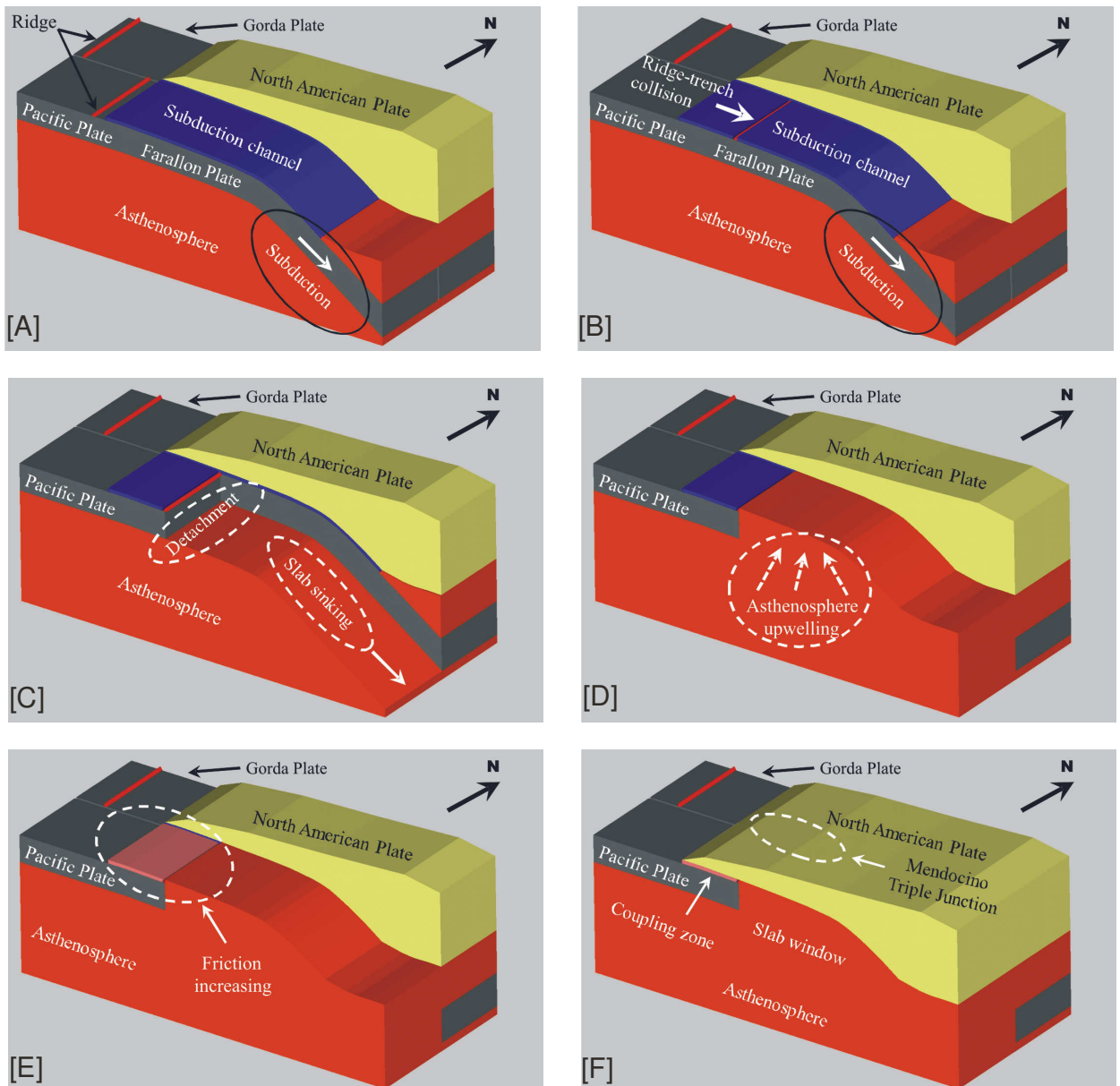


Figure 2.3 Three-dimensional cartoon showing major stages of the slab window formation. [A] Plate tectonic setting just before the ridge-trench collision. Portion of the North American lithosphere removed and underlying subduction channel material is shown. [B] As the subduction continues, the ridge-trench collision takes place. [C] At certain depth the spreading terminates, the portion of the slab detaches along the zone of minimal strength (probably the ridge itself) and sinks in the mantle driven by the slab pull leaving the gap behind. [D] Hot asthenospheric material upraises and fills the gap coming in contact with thin lithosphere. [E] Termination of the subduction leads to increase of the coupling in the extinct subduction interface due to fluid escape and healing processes. This in turn causes the capture of the westernmost North American crust by the Pacific plate. [F] Formation of the slab window finishes. As the MTJ migrates northward, the larger pieces of North American crust become accreted to the Pacific. During the evolution of the SAF boundary, the portions of the Pacific lithosphere just south of the MTJ advance in age and become the controlling feature of the system due to its increasing strength.

lithosphere. For example, the arc magmatism related to the Cascadia subduction zone should terminate at the triple junction. Elevation of the Coast Ranges should increase south of the MTJ due to increased buoyancy of the hot asthenospheric material; elevation should decrease further south as this material cools. The asthenospheric upwelling should also result in high heat flow

and crustal magmatism, and the maximum in heat flow and the surface expression of volcanism should follow the highest topography. Velocity anomalies in the crust and upper mantle should be associated with heating, melting, and magmatic underplating. These and others consequences of the MTJ migration are relatively well supported by the existing geophysical observations (see e.g. the review paper by Furlong and Schwartz, 2004).

Slab window and lithospheric deformation

Another important aspect is the link between the slab window evolution and the deformation of the northern California lithosphere. It can be suggested that the transform deformation along the plate boundary developed simultaneously with thermal re-equilibration of the lithosphere. Throughout central and northern California, the SAFS is coincident with the Coast Ranges, an area of elevated heat flow and uplift (Fig. 2.4). The principal locus of deformation, the San Andreas Fault, is located along the western edge of the Coast Ranges near Cape Mendocino where the fault is born and migrates eastward across the Coast Ranges as the age of the SAFS increases to the south. It has been argued that many aspects of the present day SAFS have resulted from migration of transform deformation eastward within the NA plate (Dickinson, 1997). According to this hypothesis, the SAF jumped to the east (one or more times) until reaching its present position in central and southern California. From analyses of the evolution of lithospheric strength based on the thermal model of the SAFS (Furlong, 1984), it was suggested (Furlong et al., 1989; Furlong, 1993) that this migration of the transform deformation and jumping of major fault was due to the coupling of the thermal processes related to the opening of the slab window below NA plate with the transform displacement at the P-NA plate boundary. On the basis of their 3-D thermo-mechanical modeling of northern California, Furlong et al. (1989) concluded that the lithospheric boundary between the Pacific and North American plates, as well as the crustal structure in the area, must evolve as the Mendocino triple junction travels northward. They inferred that the plate boundary will tend to migrate eastward to the Maacama and Bartlett Springs fault zones (Fig. 2.4).

Furlong and Govers (1999) associated observed crustal thickening in the vicinity of MTJ with the conductive cooling of material that upwells in the slab window. They suggest that this cooling creates basal coupling between the North American plate to the Gorda slab. As a result the North American crust is transported northward causing thickening (contraction) in advance of the triple junction and thinning (extension) in a region extending 300 km to the south of the triple junction. As the slab progresses north, the zone of thickened crust also migrates with it. The underlying process was called the Mendocino Crustal Conveyor (MCC). The MCC model has important consequences for the observed regional heat flow. North of the MTJ the thickening

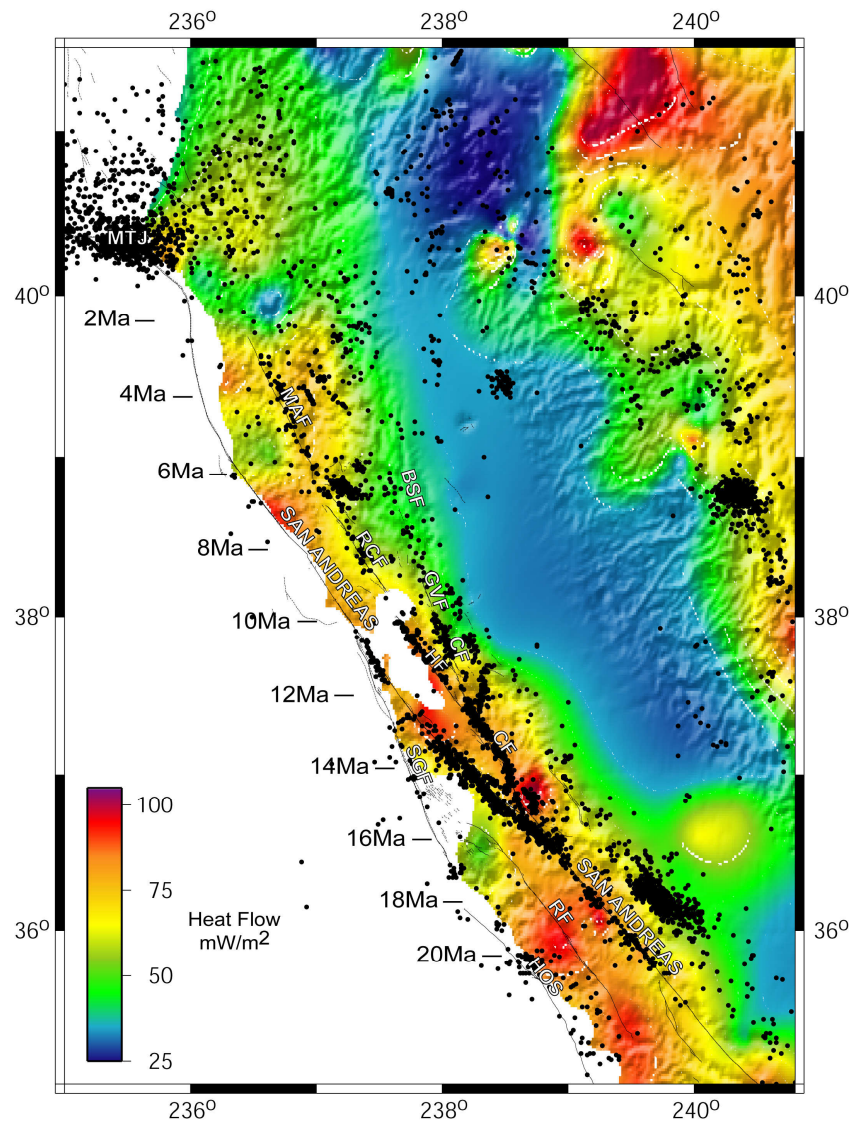


Figure 2.4. Map illustrating major features of the SAFS in central and northern California (from Sobolev et al., 2003). Background colors represent surface heat flow observations from the USGS data base (<http://quake.wr.usgs.gov/heatflow/index.html>). Also shown are the distribution of the seismicity over the past decade (Northern California Data Center), topography and major faults. The numbers to the left side of the San Andreas Fault (shown in red) indicate the positions of the Mendocino Triple Junction (MTJ) relative the Great Valley block (Atwater and Stock, 1998). Other faults shown are the Maacama (MAF), Bartlett Springs (BSF), Rogers Creek (RCF), Green Valley (GVF), Calaveras (CF), Hayward (HF), San Gregorio (SGF), Hosgri (HOS) and Rinconada (RF).

of the North American crust causes regional crustal cooling that correlates to the relatively low heat flow values measured near Cape Mendocino (Lachenbruch and Sass, 1980). South of the MTJ the crustal thinning coupled with the heating from the emplaced asthenospheric material can account for the sharp rise in the heat flow values seen to the south of the triple junction (Fig. 2.4).

2.3 Stress field and fault strength

Borehole stress measurements and earthquake focal plane mechanisms are the main sources of information about relative stress magnitude and orientation. Stress measurements made in deep (>1 km) boreholes in a variety of tectonic settings have shown that stresses in the crust require the friction coefficients to be in the range 0.6 - 0.7 with nearly hydrostatic pore-pressure gradients (e.g. Zoback and Healy, 1992). This range of friction coefficients agrees with laboratory measurements for a wide variety of rock types (Byerlee, 1978). The information about the magnitude and orientation of the stress can be directly used to infer the strength of the faults. The weak-fault hypothesis makes two predictions: (1) the magnitude of a shear stresses measured near a fault will be low; and (2) because the fault must be close to a principal plane, the direction of the maximum horizontal stress (S_{Hmax}) will rotate to a nearly fault-normal orientation as it approaches the fault. In the strong-fault hypothesis the predictions are exactly the opposite: (1) the magnitude of the shear stresses will be high near the fault; and (2) S_{Hmax} will rotate as it approaches the fault to a more acute orientation with it.

In general, a high angle (60-80°) between the direction of the maximum horizontal stress and the local fault orientation is observed along the entire length of the San Andreas Fault (e.g., Zoback et al., 1987). This provides independent evidence of the weakness of the fault, forcing the average shear stress on the fault to be smaller than 10 MPa between depths of 0 and 15 km (Mount and Suppe, 1987). This also should imply low effective friction coefficient of the faults.

Stress-heat flow paradox

As the rheology of lithosphere is strongly linked to its thermal state, the heat flow of the SAFS provides a key constraint on the assessment of fault strength. Heat flow along the SAFS has two primary characteristics: (1) A broad heat flow anomaly with a mean value of 80 mW/m² is associated with the SAFS. Heat flow decreases significantly to the east when approaching the Great Valley, where typical values are 40 mW/m²; (2) there is no narrow heat flow anomaly associated with the trace of the SAF (Fig. 2.5). This has been interpreted as a low shear heating in the frictional part of the fault zone between depths of 0 and 15 km (Brune et al., 1969; Lachenbruch and Sass, 1980). If a coefficient of friction of 0.6 – 0.8 and hydrostatic pore pressure are assumed, which are typical for crystalline crust (Townend and Zoback, 2000), then an average shear stress of around 100 MPa for the seismogenic zone would be expected. Such a high value of shear stress would lead to a narrow heat flow anomaly of 60 mW/m², which is not observed. The lack of the narrow heat flow anomaly, associated with the anomalous stress orientation at the SAFS, constitutes the *stress-heat flow paradox* of the SAF (Lachenbruch and Sass, 1992). Steady-state conductive model of frictional heating limits the mean value of shear stress on the seismogenic fault to be no larger than ca. 20 MPa. This model, however, did not

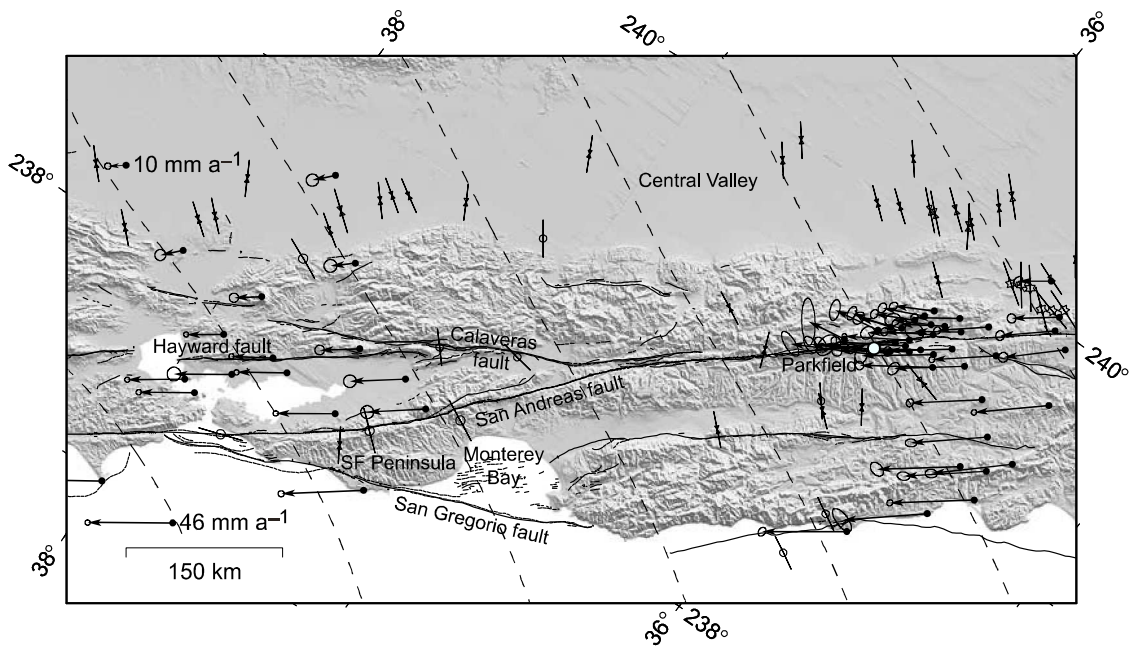


Figure 2.5 Maximum horizontal compression (S_{Hmax}) and crustal velocity data from central California; modified after Townend and Zoback (2004). S_{Hmax} directions determined from borehole breakouts are shown by inward-pointing arrows, from hydraulic fracturing experiments by stars, and from earthquake focal mechanism inversions by split circles (Fuchs and Muller, 2001; Townend and Zoback, 2001). The dashed trajectories show regional S_{Hmax} directions calculated using a model of lithospheric buoyancy and plate interaction (Flesch et al., 2000), and the vectors illustrate crustal velocity data relative to North America (Murray and Segall, 2001). The SAFOD drillhole location is marked by a white circle. SF – San Francisco.

consider shear heating in the ductile portion of the crust and mantle lithosphere. With this additional heating mechanism taken into account, the stress should be even smaller.

Stress orientation

Townend and Zoback (2004) compared different indicators of tectonic stress directions in central and southern California obtained from focal mechanism inversions (Townend and Zoback, 2001) and a variety of sources represented in the World Stress Map Database (Fuchs and Muller, 2001). Throughout central and southern California, a uniform direction of observed S_{Hmax} is remarkably consistent with the superposition of stresses arising from lateral variations in lithospheric buoyancy in the western United States, and far-field Pacific-North America plate interaction (Flesch et al., 2000). In central California (Fig. 2.5), the axis of S_{Hmax} lies at a high angle to the San Andreas Fault. Despite relatively few observations near (± 10 km) the fault, observations in the greater San Francisco Bay area indicate an angle of as much as 85° implying extremely low fault strength. In southern California, observations of stress orientations near the SAF are rotated slightly counter clockwise with respect to the regional field ($68 \pm 7^\circ$ along 400 km of the fault). This indicates that the SAF has moderately low frictional strength in southern California.

Hickman and Zoback (2004) reported the results of the stress measurements from the 2.2-km-deep vertical pilot borehole drilled in summer 2002 at the same site as SAFOD (San Andreas Fault Observatory at Depth). This site locates 1.8 km southwest of the San Andreas Fault near Parkfield, CA (Fig. 2.5), on a segment of the fault that moves through a combination of aseismic creep and repeating microearthquakes. Ultrasonic borehole televiewer logs acquired in the SAFOD borehole reveal extensive stress-induced borehole breakouts at depths from 0.8 to 2.2 km. Borehole breakouts and drilling-induced tensile fractures in the SAFOD borehole, indicate significant local variations in the direction of the maximum horizontal compressive stress, but show a generalized increase in the angle between S_{Hmax} and the San Andreas Fault with depth. This angle ranges from a minimum of $25 \pm 10^\circ$ at 1000 – 1150 m to a maximum of $69 \pm 14^\circ$ at 2050 – 2200 m. The lowermost stress orientation is in fairly good agreement with S_{Hmax} directions seen at much greater distances (>10 – 20 km) from the San Andreas Fault in central California.

Based on the stress orientations derived from the earthquake focal plane mechanisms, Provost and Houston (2003) suggested that the SAFS is mechanically stronger with a greater effective frictional strength in the northern portion than in the creeping section of the SAF, with the Bay Area in an intermediate state. They interpreted this situation to result from the evolution of the plate boundary toward lower effective frictional strength with increasing slip. The underlying slip-weakening mechanism might be responsible for this observation.

Fault weakness speculations

Numerous theories have been proposed over the past decade specifically related to the weakness of the San Andreas Fault. Four general classes of explanations have been suggested: (i) elevated fluid pressures, (ii) intrinsically low coefficients of friction, (iii) solution-transport reactions and (iv) dynamic weakening mechanisms. The high fluid pressure alone is able to satisfy the heat-flow constraint, but fails to explain the stress rotation near the fault (directional constraint) (e.g. Zoback et al., 1987; Scholz, 1989). Intrinsically low coefficients of friction can be due to presence of clays or other weak minerals in the fault gauges (e.g., Janecke and Evans, 1988; Wintsch et al., 1995). Solution-transport mechanisms such as pressure solution, fluid-assisted mineral reactions and crack healing may be important in determining the rheology of fault zones (e.g., Sibson, 1983; Blanpied et al., 1995; Chester, 1995). The dynamic weakening mechanisms include shear heating during slip, leading to transient high fluid pressures (Lachenbruch, 1980) or melting (Spray, 1987); reductions in normal stress accompanying the propagation of dilatational waves along the fault (Brune et al., 1993); and the fluidization of fault-zone materials due to the channelling of co-seismic acoustic energy (Melosh, 1996).

2.4 Lithospheric structure

Seismic tomography and the seismic survey methods are the primary sources of information about composition and characteristics of the crust and mantle. They provide key information that required for constraining geological time-scale models of lithospheric deformation. Broad range of papers investigating lithosphere and crustal structure in western North America has been published in the literature.

Upper mantle

Humphreys and Dueker (1994) performed the three-dimensional inversion of upper mantle P wave velocity structure from travel time residuals recorded within the western United States. They obtained images of the upper mantle beneath this region. This study revealed that the P wave velocities in the upper mantle are slow on average with respect to North America craton, and the structure is very heterogeneous. Certain structures such as Juan de Fuca and Gorda slabs penetrate to depths greater than 200 km and are imaged with good lateral resolution. However the most upper mantle heterogeneity locates at depths less than 200 km. Within first 250 km of the Pacific Coast they imaged long wavelength mantle structures that trend parallel to the surface topography and young tectonic structures.

Relatively high-velocity upper mantle is imaged beneath the westernmost California and the offshore region, extending south from the Cape Mendocino to the San Francisco Bay region (Fig. 2.6). This feature is in principle coincident with the present inferred location of the westernmost part of the slab window. This high-velocity feature does not suit the slab window concept since lower velocity is expected from the hot asthenosphere. Humphreys and Dueker (1994) suggested that it might be due to either the remnants of Farallon plate or the Pacific plate that more recently underthrust beneath North America associated with shortening near the California margin. Alternative interpretation might be the conductive cooling of the asthenosphere in the slab window from the relatively cold Pacific plate. Beneath significant part of the Coast Ranges (from westernmost transverse ranges to the Olympic Mountains on the north) and the adjacent forearc sedimentary basins including western Great Valley, Humphreys and Dueker (1994) imaged a large volume of low-velocity upper mantle (Fig. 2.6). These low velocities probably represent the slab window associated with northward migration of the MTJ. However, significant variations of the velocities occur that are not easily interpreted by simple slab removal.

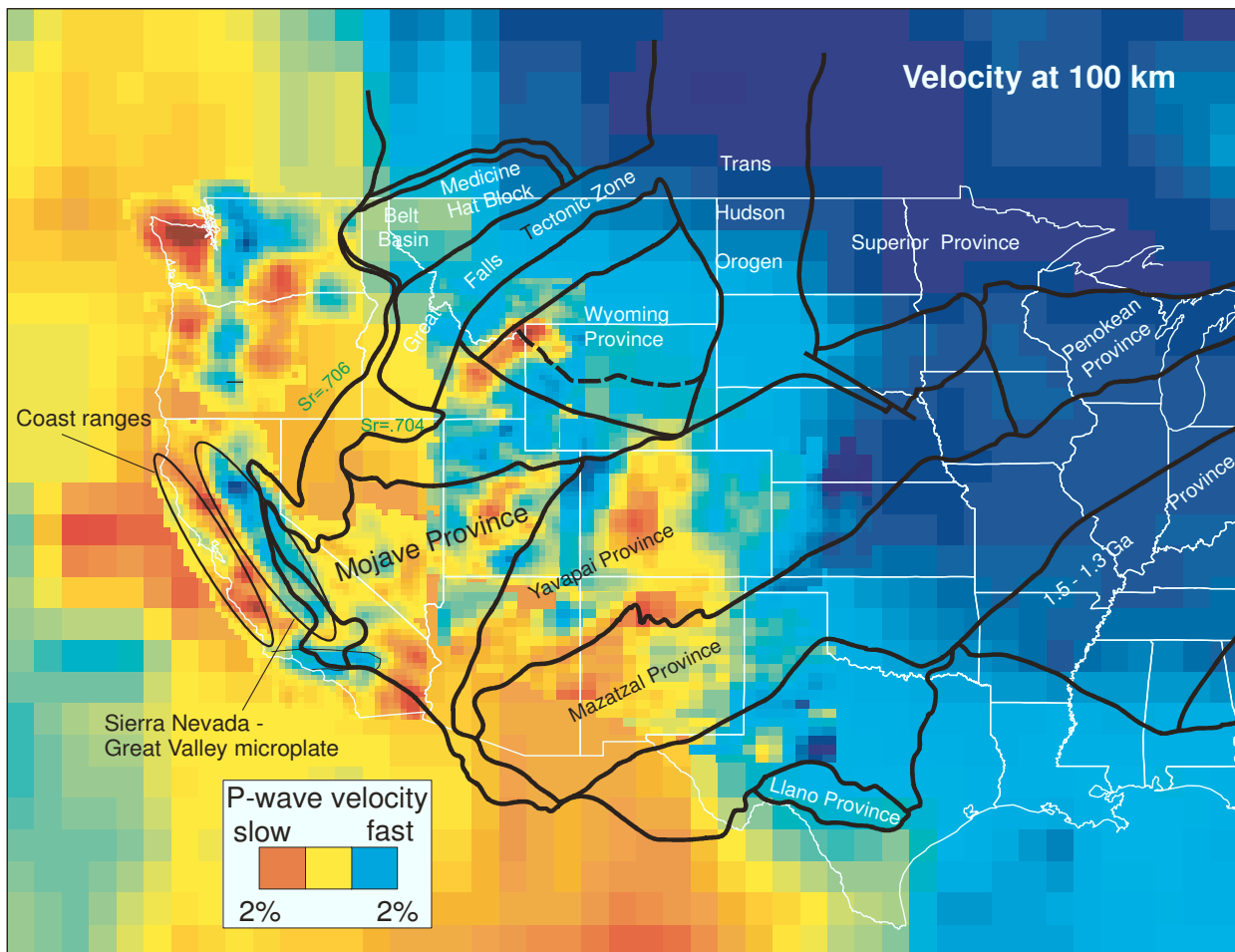


Figure 2.6 P-wave seismic tomography of western United States upper mantle at 100 km depth (from Humphreys and Dueker 1994). Colors indicate P-wave velocity variation, white lines show contours of the states, black lines separate tectonic provinces. Ellipses show approximate locations of the Coast Ranges and Sierra Nevada - Great Valley microplate.

The steeply dipping Juan de Fuca and Gorda slabs are well resolved beneath central Washington, but resolution is poor beneath northwestern California. The most prominent feature imaged in the inversion of Humphreys and Dueker (1994) is the high velocity structure that extends south as far as southern Great Valley and partially resides beneath the Sierra Nevada Great Valley block (Fig. 2.6). They attributed this anomaly to a subducted slab beneath the Pacific Northwest and a downwelling of North America mantle lithosphere. Another anomaly that resolved within 100 – 240 km depths is the east-trending high-velocity feature beneath the Transverse Ranges in southern California. This feature is interpreted as a downwelling lithosphere consistent with the crustal convergence in the Transverse Ranges.

Benz et al. (1992) determined the lithosphere P wave velocity structure beneath northern California. In the upper mantle (30-70 km) Benz et al. (1992) image higher velocities beneath the coastal region from Cape Mendocino to the San Francisco Bay, bounded by lower velocities in a region beneath the Klamath Mountains and the northern Coast Ranges. The cross-section near Clear Lake showed a prominent low-velocity body beneath the northern Coast Ranges that is

juxtaposed with the high velocity anomaly beneath Great Valley. Benz et al. (1992) infer that the southward tapering low-velocity body (30 - 110 km depth) beneath the Coast Ranges is probably due to warm asthenosphere that has moved into the slab window created by the northward migration of the MTJ. The western edge of the anomaly may define the boundary at depth between the Pacific plate and the asthenosphere of the slab window. In a simple rigid plate model, this boundary would coincide with the surface plate boundary, the SAF. However, with deformable plates, the subsurface plate boundary can be offset relative to the surface expression of the San Andreas. Further support for this interpretation and its implications for plate dynamics are presented by Furlong et al. (1989).

Available constraints for the thickness of the lithosphere are also important for the formulation of the adequate thermo-mechanical models. Li et al. (2007) revealed the existence of a mantle discontinuity with velocity reduction downward based on the *S* receiver functions from 67 broad-band seismic stations in the western United States. They interpreted this observation as the lithosphere–asthenosphere boundary (LAB). The boundary is relatively sharp with the average depth about 70 km. The boundary is more prominent south of the Mendocino Triple Junction, where the Farallon Plate has completely subducted. This may indicate partial melts at the base of the lithosphere caused by the upwelling of the asthenospheric flow through the slab window. The LAB map of Li et al. (2007) generally agrees with previous P- and S-wave tomography works (e.g. Humphreys and Dueker, 1994). The depth of the LAB is relatively homogeneous in the central part of the area. It is less than 70 km beneath much of California and Nevada and increases to the north and to the east to more than 80 km. Observations of Li et al. (2007) indicate a sharper LAB with strong velocity contrast beneath SAF zone south of the MTJ, which they attribute to the processes in the slab window. The emplacement of hot asthenospheric mantle at shallow levels would be expected to generate melts at the base of the lithosphere, consequently increasing the LAB contrast.

Crust

A 30-km thick crust with an intra-crustal interface or lowermost crustal layer (LCL) at a depth of 10 km is an average characteristic of the crustal structures in the western US. However in many regions in California this average figure is significantly biased by the intensive deformation of the crust and underlying mantle in the SAFS. Henstock et al. (1997) imaged the offsets in the lower crust and the Moho discontinuity near the SAF and Maacama faults (MF) in northern California. The results were obtained from the 1993 – 1994 Mendocino Triple Junction seismic experiments (e.g. Beaudoin et al., 1996) along an east-west line extending from the Pacific Ocean basin to the eastern side of the Coast Ranges. They found that the SAF and MF may cut

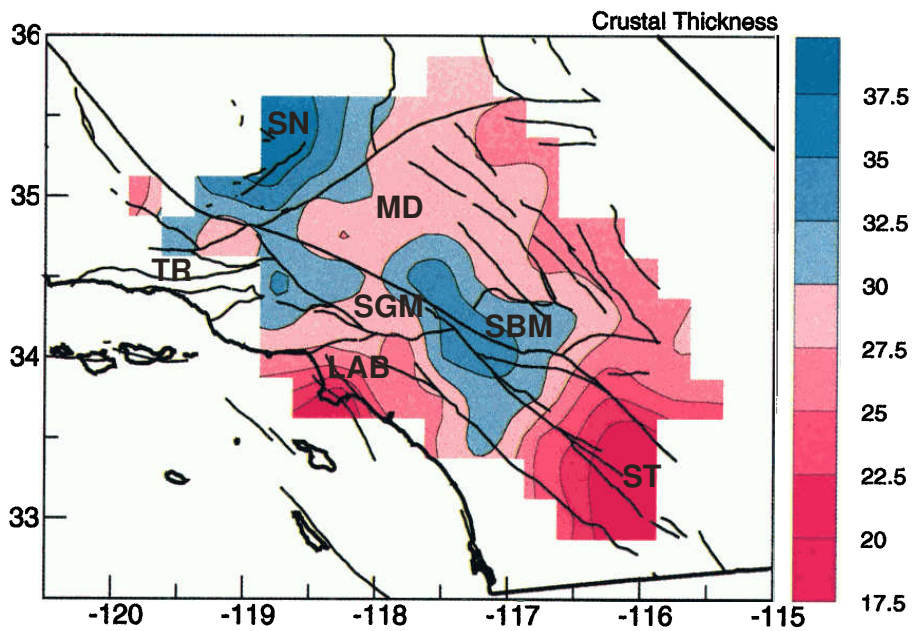


Figure 2.7 Contour plots of crustal thickness estimates in southern California from the stacking *PmP* arrivals (from Richards-Dinger and Shearer, 1997). Crustal thicknesses are in kilometers. The abbreviations are: SN – Sierra Nevada; MD – Mojave Desert; SBM – San Bernardino Mountains; ST – Salton Trough; LAB – Los Angeles Basin; SGM – San Gabriel Mountains; TR – Transverse Ranges.

through the crust to the upper mantle in a relatively narrow zone (less than 10 kilometers wide). They interpreted the vertical offsets in velocity at the top and bottom of the LCL beneath the SAF and MF as extensions of these strike-slip faults into the mantle. Possible interpretation for the observed higher-velocity LCL may be the following: (i) Pacific plate oceanic crust overridden by the North American plate following the migration of the MTJ; (ii) Farallon plate crust captured by the Pacific plate; (iii) unsubducted Gorda plate fragments; (iv) ophiolitic rocks linked with the Great Valley basement rocks; or (v) igneous rocks solidified in situ from mantle melts. Henstock et al. (1997) suggest that the most likely origin of the LCL is underplated Gorda plate crust, which is subsequently intruded by small volumes of melt.

Beaudoin et al. 1996 presented three seismic refraction-reflection profiles in the framework of the MTJ seismic experiment and compared crust and upper mantle of the North American margin before and after it is modified by passage of the MTJ. They interpreted their data to indicate tectonic thickening of the Franciscan complex in response to passage of the MTJ and an associated thinning of these rocks south of the triple junction due to assimilation into melt triggered by upwelling asthenosphere. Their data reveal that southern edge of the Gorda slab and/or the overlying North American crust may have fragmented because of the change in stress presented by the edge.

In southern California crustal thickness varies from about 20 to 35 km. Due to presence of long-term transpressional bends such as Big bend (San Emigdio bend), and the San Bernardino bend, the surface topography exhibits uplifts. Correspondingly, the Moho discontinuity exhibits

depressions due to crustal thickening. From the stacking *PmP* arrivals Richards-Dinger and Shearer (1997) estimated crustal thickness in southern California. They obtained an average crustal thickness of 28 km. Estimated Moho depth ranges between 18 km in the Salton Trough to 33 km beneath the eastern Transverse Ranges and 36 km beneath the southernmost Sierra Nevada (Fig 2.7). They also revealed nearly fiat Moho at about 30 km in the Mojave Desert, thinning slightly to the east, very thin (less than 20 km) crust in the Salton Trough and thin crust offshore of the Los Angeles Basin.

2.5 Present day plate motion

The present-day strike-slip component of motion between the stable North American plate (i.e., the Colorado plateau) and the Pacific plate is estimated to be 49 mm/yr (DeMets et al., 1990). Geodetic studies indicate that as much as one fourth of this relative motion is absorbed within the Basin and Range province to the east (Dixon et al., 1995; Thatcher et al., 1999). As the Sierra Nevada and Great Valley crustal blocks seem to behave rigidly, the SAFS itself accommodates a strike-slip differential motion of 39 ± 2 mm/yr (Argus and Gordon, 2001). Neotectonic studies of the SAFS show a different pattern of strain distribution in central and northern California (Brown, 1990). In most of the study area considered here, relative motion between the P and NA plates is accommodated by slip along several sub-parallel strike-slip faults spaced ~30-60 km apart. For example, at the latitude of San Francisco bay, both geologic (Brown, 1990) and geodetic (Prescott et al., 2001) studies demonstrate that nearly all of the relative plate motion is accommodated by slip on the San Gregorio, San Andreas, Hayward and Calaveras faults. In central California, the long-term strike-slip motion is almost completely restricted to the San Andreas Fault (SAF) itself. The rate of contraction across the SAFS is by more than one order of magnitude smaller than the fault parallel motion. Estimates of the shortening rate across the SAFS based on geological reconstructions (Thompson, 1999) and direct geodetic measurements (Argus and Gordon 2001) correlate relatively well and range between 0 and 3 mm/yr.

Argus and Gordon (2001) used geodetic results from very long baseline interferometry (VLBI), satellite laser ranging (SLR), and the Global Positioning System (GPS) to estimate angular velocities between the Sierra Nevada – Great Valley microplate, Pacific plate, and North American plate. The angular velocities show that the SAFS and central California Coast Ranges accommodate motion of 39 ± 2 mm/yr, mainly by strike-slip faulting (Fig. 2.8). Fault-normal motion is small, is mainly convergent (at rates up to 3.3 ± 1.0 mm/yr), and varies along the coast, but is divergent (at 2.6 ± 1.2 mm/yr) across San Pablo Bay. The mountain ranges tend to be

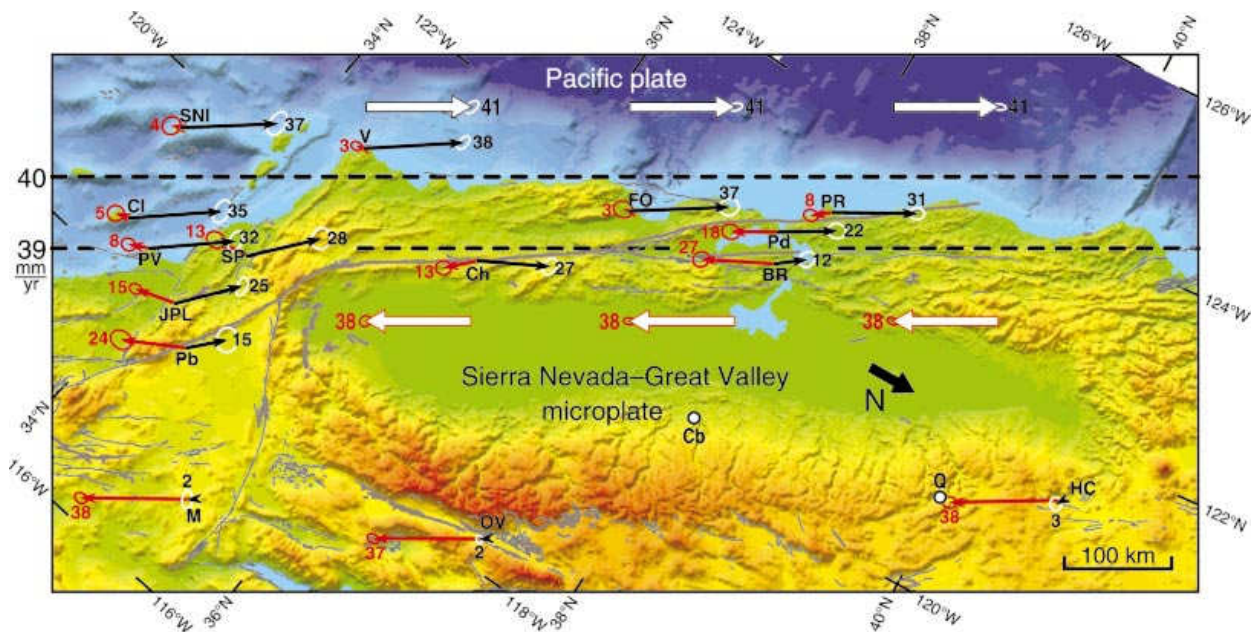


Figure 2.8 Present Plate velocities across SAFS in central California (from Argus and Gordon, 2001). Map is shown in an oblique Mercator projection about the Sierra – Pacific pole of rotation. The velocity of the Pacific plate relative to the Sierra Nevada – Great Valley microplate (SGM) is horizontal everywhere in the projection. Two dashed horizontal lines serve as reference lines for relative speeds of 39 and 40 mm/yr. The wide, black-outlined, white-filled arrows depict the velocity of the Pacific plate relative to the SGM; the narrow black arrows give the velocity of the individual sites relative to the SGM. The wide, red-outlined, white-filled arrows depict the velocity of the SGM relative to the Pacific plate; the narrow red arrows give the velocity of the individual sites relative to the Pacific plate. Error ellipses are 95% confidence limits.

larger where the fault-normal convergence rates are larger.

The uncertainty in the direction of Pacific-Sierra motion varies between $\pm 1.4^\circ$ and $\pm 3.1^\circ$. Argus and Gordon (2001) suggest that significant inelastic deformation, and/or significant slip occur on the offshore faults such as the Hosgri–San Gregorio fault. The best fit straight line approximation of the SAF (from the middle of the Carrizo Plain to point Arena) strikes $\sim 6^\circ$ counterclockwise of the predicted direction of Pacific plate – Sierran microplate motion. Therefore, the space geodetic data provide direct evidence that motion along the plate boundary cannot be purely strike slip and that some convergence is required across it.

Chapter 3

Review of previous modeling

3.1 Thin-shell approximation

In this type of models (both numerical and analytical) the lithosphere is approximated as a shell of finite thickness. To assure adequacy of this approximation the horizontal span of the modeled lithosphere must be significantly larger than its thickness. The geometry of the shell is either assumed to be planar (e.g. Bird and Piper, 1980) or spherical (e.g. Humphreys and Coblenz, 2007). All models in this category make an assumption of a two-dimensional flow (plane stress), such that the horizontal components of the velocity field are independent of depth (e.g. Bird and Piper, 1980). The variation of rheological parameters and stresses with depth are parameterized and explicitly included in the formulation of the thin-shell models. The assumed structure of lithospheric rheology spans from a single-layer elastic model with laterally uniform strength (Humphreys and Coblenz, 2007) to a layered model with horizontally variable strength that resolve frictional sliding on the faults and nonlinear dislocation (power law) creep in the lower crust (e.g. Bird and Kong, 1994). Fig. 3.1 shows an example of the deviatoric stress distribution with depth in the layered lithosphere model (Bird and Piper, 1980) subjected to a typical horizontal shortening. The other common assumptions of the thin-shell modes are: lithostatic vertical stress and isostasy, incompressibility, vertically integrated stresses, and continuum approximation (e.g. Bird and Kong, 1994). Certain models (e.g. Humphreys and Coblenz, 2007; Liu and Bird, 2002) deal with the non-zero mantle drag forces on the base of the lithosphere.

Effective fault strength

Many important issues regarding the dynamics of the western United States with the emphasis put on the SAFS has been addressed in the literature. For example, Bird and Kong (1994) contributed to the long term debate over the strength or weakness of faults. Their finite element model included two types of the elements: triangular continuum elements and the curvilinear

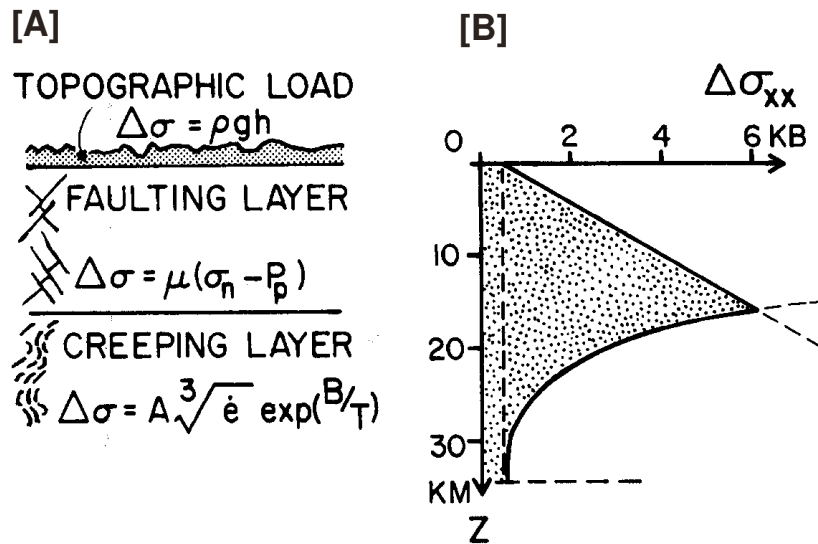


Figure 3.1 Rheological structure of the lithosphere assumed in a typical thin-shell model (from Bird and Piper, 1980). [A] Subdivision of the lithosphere into three layers in which different mechanisms control the deviatoric stress level. [B] Deviatoric stress as a function of depth calculated for a typical horizontal shortening with the strain rate of 10^{-14} s^{-1} .

fault elements. Each fault was represented by its trace on the reference plane and assigned a dip angle. To constrain the temperature-dependent rheology of the lower crust, the model required heat flow distribution on the surface as input data. The major difference with respect to previous study (Bird, 1989) was that horizontal velocities at the top of the mantle were partially assigned and not computed.

Based on the systematic parameter search of the time-averaged friction coefficient on the faults and activation energy of the dislocation creep in the lower crust, Bird and Kong (1994) concluded that relative strength of the faults must be lower than it is assumed by laboratory data. Their best-fit model of California region that closely reproduced the observed slip rates on major faults, constrained the friction coefficient by the range 0.17 – 0.25. This figure constitutes only 20% – 30% of the value 0.85 that is generally assumed for the crustal rocks. This reduction of the fault strength might be attributed to the slip on the fault. Bird and Kong (1994) conducted a set of model experiments to investigate whether fault weakening is uniform or proportional to the amount of the net slip of the fault. The stress data was matched better if only SAF was made weak, as it has the largest slip among the other faults in the system. Thus, these results supported the hypothesis that all weakening is slip-dependent. The inclusion of mantle shear tractions in the model with uniform friction on the faults did not affect the general conclusion that the fault strength must be low. The study of Bird and Kong (1994) extends the finding of very low friction from subduction zone to a dominantly strike-slip system in which underthrusting of wet sediments is not wide-spread.

Lithospheric average viscosity

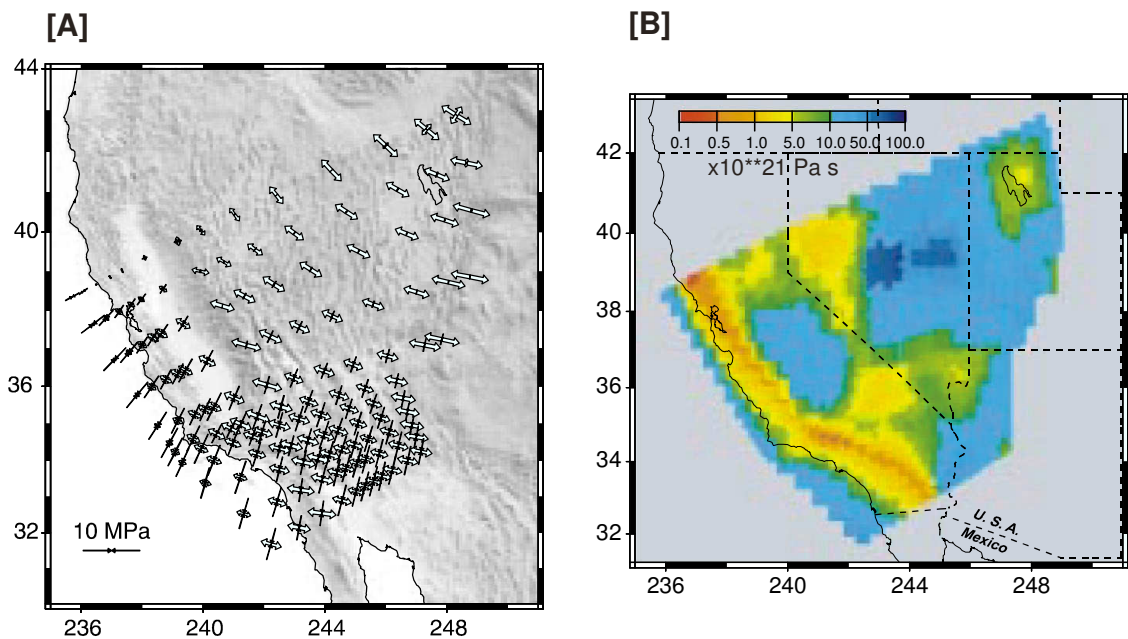


Figure 3.2 Basic results of the thin sheet model of the western United States lithosphere (after Flesch et al., 2000). [A] The magnitude and orientation of the total deviatoric stress field due to gravitational potential energy variations estimated with the geoid data (Smith and Milbert, 1999) and the corresponding best fit stress field boundary conditions associated with Pacific – North America Plate interaction. [B] The vertically averaged effective viscosity (over $L = 100$ km) for the western United States determined by dividing the magnitude of the total deviatoric stress by the magnitude of the strain rate determined from the self-consistent kinematic model.

Flesch et al. (2000) estimated the vertically averaged deviatoric stress field within the western United States using topographic and geoid data, recent GPS observations, and strain rate magnitudes and styles from Quaternary faults. A solution for the equation of steady state motion was obtained using the variational principle. The lithosphere was approximated as a viscous thin shell of 100-km thickness. The deviatoric stress was assumed to be averaged over the thickness of the lithosphere and the basal tractions were ignored. They quantified the individual stress field contributions from gravitational potential energy (GPE) variations within the lithosphere (buoyancy forces) and Pacific – North America plate interaction (boundary forces). The GPE variations were estimated using topographic and geoid data, while the stress boundary conditions were found by comparing the tensor styles of the total deviatoric stresses to the tensor styles of the observed strain rate. Additionally, they provided important constraint for the effective lithospheric strength. They calculated vertically averaged effective viscosities for the western United States using the magnitude of the strain rate tensor (Shen-Tu et al., 1999) and the magnitude of the total deviatoric stress tensor.

The vertically averaged total deviatoric stress field estimated by Flesch et al. (2000) (Fig. 3.2A) predicts the direction and style of all major deformation within the western United States in accord with other stress estimates (e.g. Zoback et al., 1987). In their model compressional stresses on the SAF were rotated 60° to the fault strike along the entire length of the fault. They

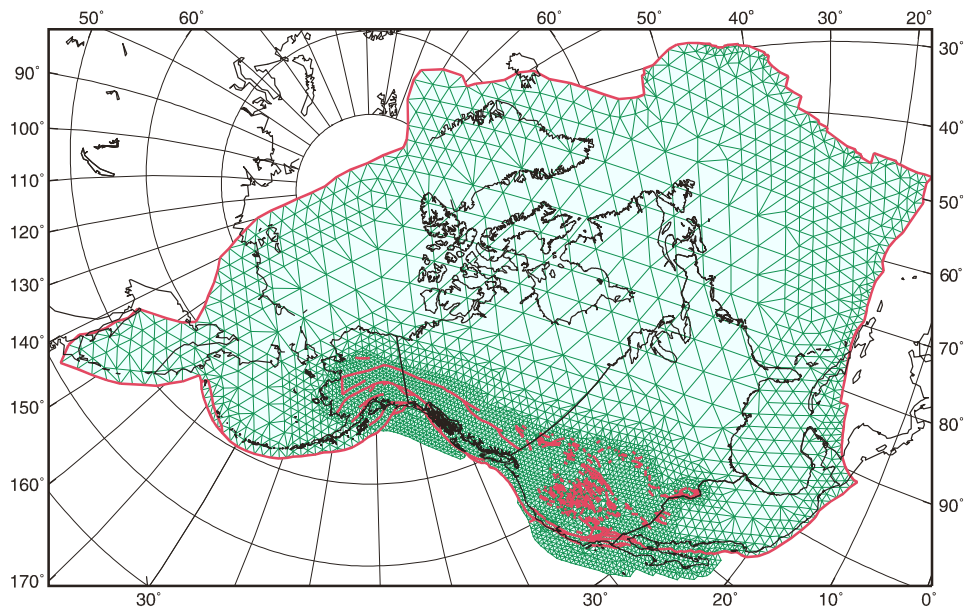


Figure 3.3 Oblique Mercator projection of the thin-shell finite element grid used to study the influence of the mantle drag on the North America plate (from Liu and Bird, 2002). The model domain is the entire North America plate plus adjacent deforming regions of the Pacific plate. Heavy lines are fault elements; light lines are the boundaries of continuum elements.

showed that GPE variations were responsible for orienting compressional deviatoric stresses at a high angle to the coastline and for producing some of the convergent motion between the Sierra Nevada block and the Pacific Plate. The effective viscosity of the lithosphere is found to vary by three orders of magnitude over the western United States (Fig. 3.2B). The SAFS is generally characterized with the lowest viscosities of $\sim 5 \times 10^{20}$ Pa·s. The variation of the viscosity within the Great Valley–Sierra Nevada block is consistent with tomography results of Humphreys and Dueker (1994). The depleted upper mantle in the central Great Basin may explain relatively high viscosities (3 to 7.5×10^{22} Pa·s), while relatively low viscosity in the western Great Basin (2 to 8×10^{21} Pa·s) may be due to partial melt in the upper mantle. The eastern California shear zone is also characterized with low vertically averaged viscosity (1 to 5×10^{21} Pa·s).

Mantle drag

The influence of the mantle drag on the deformation of the North America plate was investigated by Liu and Bird (2002). The laterally-varying thin-shell model with the fault elements at the plate boundaries was computed. This model incorporated ridge-push and slab-pull forces via three-dimensional density distribution and had prescribed velocities along the boundaries with adjacent plates. The model also had the effective fault friction and subduction zone shear traction as the free rheological parameters. They used the triangular finite element grid (Fig. 3.3) with the element sizes ranging from 500 km in the craton to smaller elements (60–120 km) at the plate margins. The fault dips were also assigned differently depending on the type of the fault

(transform, normal or subduction interface). The mantle drag force was either omitted in these models or estimated based on the assumption that the lower mantle is static with respect to Africa or moving 10% faster than North America. The models were constrained by the sea-floor spreading rates (De Mets et al., 1990), geodetic velocities (Bennett et al., 1999), stress directions from the World Stress Map, and NUVEL-1A plate motion (De Mets et al., 1994).

Systematically varying the friction of the faults, strength of the subduction interface and the style of mantle drag, Liu and Bird (2002) concluded that only those models that incorporate active mantle drag were able to successfully fit the constraints. Thus, their study also supported the general conclusion of Bird (1998) that slow-moving continents are typically linked to, driven by, and indicators of the average lower mantle flow beneath them. However, the exact azimuth and pattern of basal drag on the North America plate remained undetermined.

Tectonic forces

Recently Humphreys and Coblenz (2007) quantified the forces acting on the margins and the base of the North American plate, as well as the internal loads arising from the gravitational potential energy, from the two-dimensional (thin-shell) finite element model. Flesch et al. (2001) noted that the stress field remains relatively insensitive to variation in mechanical parameters. Humphreys and Coblenz (2007) employed this observation to justify the rheological simplifications. The rheologically complex lithosphere was represented with a simplest possible linear elastic rheology. Both viscous and plastic-like deformations were omitted, as well as the variations in North American plate stiffness. This approximation appears to be quite adequate for the rate independent model of Humphreys and Coblenz (2007) since they made no conclusions about the strain rates, velocities, or the fault slips (slip rates). However, as it is discussed in e.g. the present work (see Chapter 5), the lithosphere rheology and lateral strength variation are of primary importance for the strain partitioning on the geological time-scale. The advantage of exploiting the linear elastic model is its computational efficiency and appropriateness for the parametric search among many load combinations to determine the best-fit model.

Humphreys and Coblenz (2007) used the finite element mesh of Liu and Bird (2002) (Fig. 3.3) to approximate the geometry of the North American plate. The relatively complete set of applied loads (Fig. 3.4b) consisted of 30 boundary, basal, and internal loads including the global flow tractions of Becker and O'Connell (2001) and gravitational potential energy estimates. The completely dynamic system was considered (i.e. no kinematical loads). The modeling was constrained with the stress data from the World Stress Map (<http://www.world-stressmap.org>). For each load the basis stress field was calculated; the best-fit model was found in the form of linear combination of the basis fields using a simulated annealing algorithm. As a measure of the

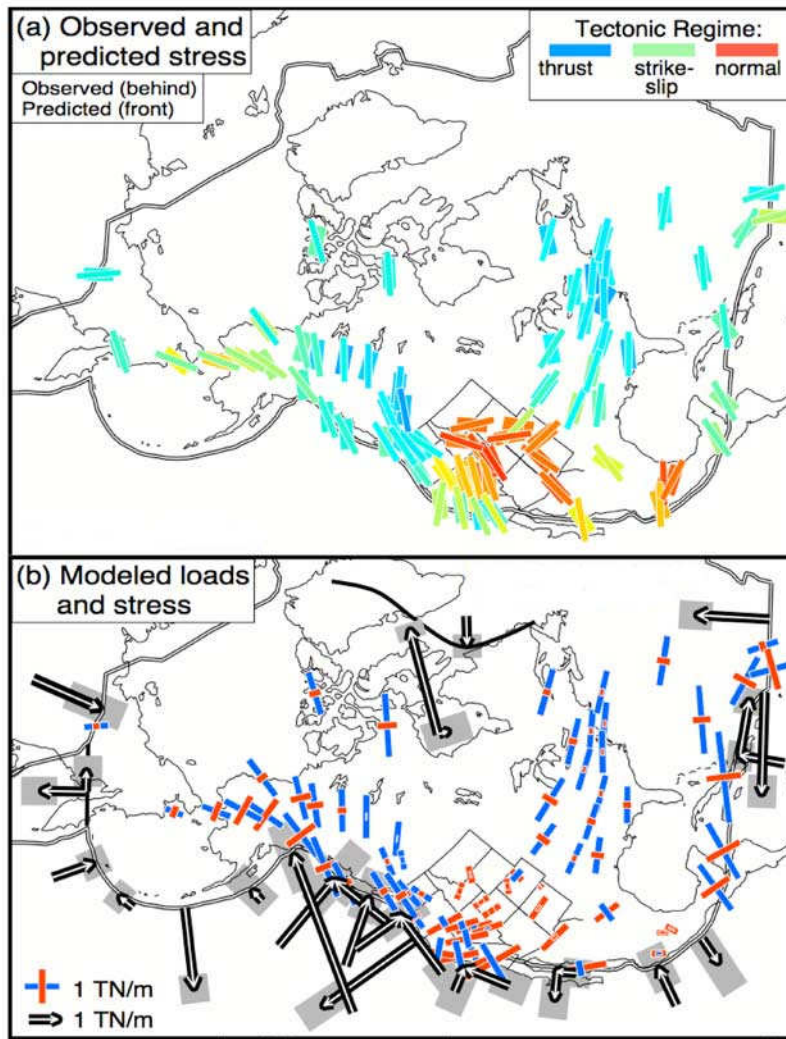


Figure 3.4 Results of thin-shell finite element model of the North American plate (from Humphreys and Coblenz, 2007). (a) Comparison of observed and modeled stress. (b) Modeled loads and stress. Red and blue bars show modeled horizontal principal deviatoric stresses, with red indicating tension and blue indicating compression relative to pressure. Boundary loads are given in TN/m. Uncertainties in load magnitudes are shown with bars at vector head or tail.

misfit Humphreys and Coblenz (2007) used a difference between the observed and predicted stress orientation angle and tectonic regime.

Fig 3.4a shows the results of the best-fit model which successfully predicts the observed stress field over North America plate. The model derived the conclusion that the boundary loads are the most important, followed by internal and basal loads. Estimated average basal traction is very low (0.4 MPa), suggesting that North America is separated from a relatively static deep Earth mantle by a weak asthenosphere. According to models of Humphreys and Coblenz (2007) the integrated San Andreas strength is about 1.5 TN/m; the estimated effective friction coefficient is 0.10 ± 0.05 . These values are similar to those obtained by Bird and Kong (1994) and Townend and Zoback (2004) for the San Andreas fault.

3.2 Two-dimensional models

This type of models considers thermal and/or mechanical processes in the vertical lithospheric cross-sections assuming validity of the plain-strain approximation. Practically it means that the velocity and the heat flux component normal to the cross-section of the model are assumed to be zero. A few interesting models relevant for the scope of this work were published previously.

Fault “jumps” and mantle plate boundary

Furlong (1984, 1993) studied the influence of the thermal processes in the mantle related to the migration of the Mendocino Triple Junction on the lithospheric deformation in the San Andreas Fault System. He noticed that conductive cooling of the upwelled asthenosphere in the wake of MTJ must cause the formation of relatively narrow plate boundary within the mantle lithosphere which is initially offset from the crustal fault zone. This mismatch between crustal and mantle parts of the plate boundary leads to the development of additional faults within the North American crust. The manifestation of this process can be viewed as a discrete eastward “jump” in the location of the crustal plate boundary. Particularly familiar fault configuration is observed in the San Francisco bay area where the activity of the older San Andreas Fault is being taken over by the younger Hayward and Calaveras faults (e.g. Furlong 1993). The important task is to explain which parameters control the spacing between these sub-parallel faults and timing of their simultaneous existence and activity. Furlong (1993) only discussed the influence of temperature and grain size evolution on the creeping flow in the mantle. These thesis further addresses thermo-mechanical effects of the San Andreas plate boundary evolution, including the fault “jumps”, both from the conceptual and modeling viewpoints (see Chapter 5).

The conclusions of Furlong (1993) were mostly based on the 2D numerical model of Furlong (1984). The model considered the time-varying temperature distribution in an idealized slab-window using the so-called method of lines, which employs the finite difference to approximate spatial derivatives. Fig. 3.5a-b show the setup and the boundary conditions of the model. The cross-section shown in Fig. 3.5b corresponds to initial specified temperature distribution in the wake of MTJ, while Fig. 3.5c-d demonstrate the calculated state of temperature after 5 Myr and 10 Myr, respectively. These cross-sections are attributed to the points, where MTJ located 5 Ma and 10 Ma, which is shown on map (Fig. 3.5a).

The time-varying temperature distribution in the slab window causes eastward migration of the locus of maximum mantle temperature (Fig. 3.5c-d). Since the rheology of the rocks is temperature-dependent, the locus of the minimal lithospheric strength also migrates eastward, to

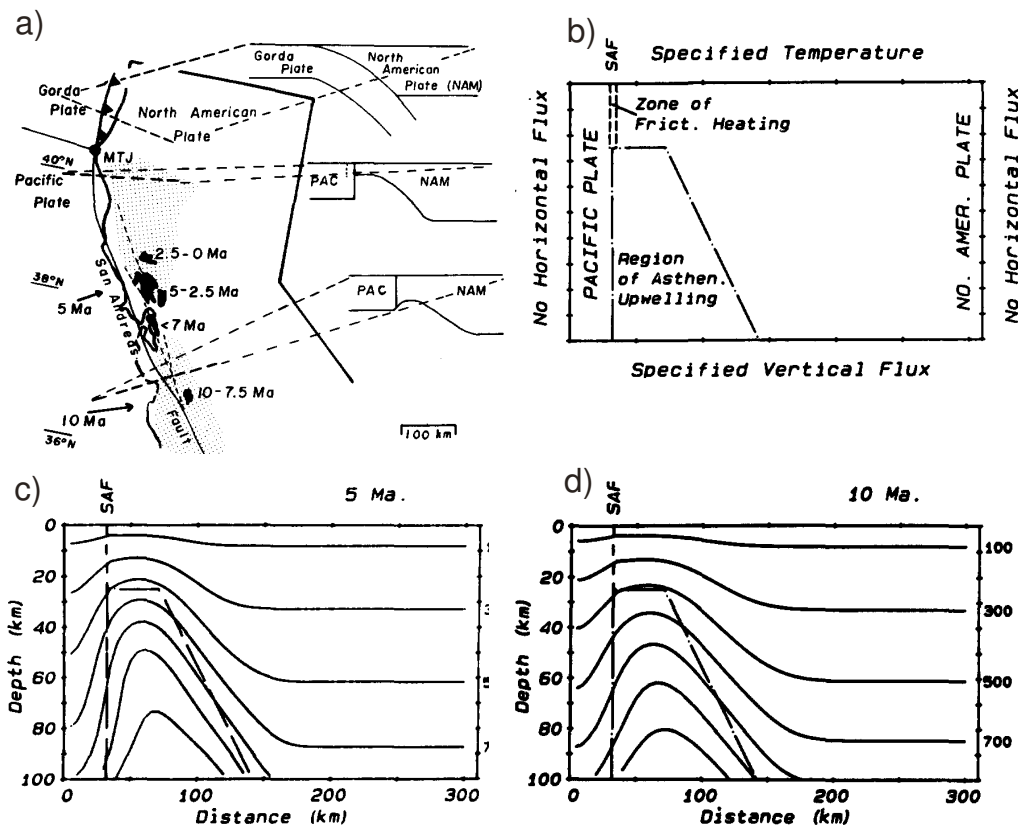


Figure 3.5 Results of the thermal modeling of the Mendocino Triple Junction migration (after Furlong, 1984). (a) Region of California, affected by MTJ passage. Black regions show the volcanic fields related to triple junction. Previous positions of MTJ indicated by the arrows. The abbreviations are: PAC – Pacific Plate; NAM – North America Plate. Schematic lithospheric cross-sections show the structure of the plate boundary. Shaded region contains elevations above 300 m in California Coast ranges. (b) Boundary and initial conditions of the model. The temperature in the upwelling asthenosphere is assumed to be approximately 1300°C. Dashed line represents initial position of the SAF. (c) - (d) Evolution of the temperature field following the MTJ passage at 5Ma and 10Ma, respectively. Temperature contours are labeled in °C every 200 °C beginning with 100°C.

which and Furlong (1993) attribute the fault jumps. However, the actual displacements of the transform plate boundary both in the mantle and in the crust are influenced by many other parameters. These aspects are discussed and modeled in detail in the Chapter 5 of this thesis. The model of Furlong (1984) assumes zero horizontal heat fluxes between the slab window and the Pacific and the North American plates. This simplification strongly slows the rate of cooling in the slab window. Therefore, the models presented here use the more sophisticated thermal boundary conditions, which include conductive heat exchange between the slab window and surrounding plates.

Mendocino Crustal Conveyor (MCC)

Furlong and Govers (1999) related the observed variations of the crustal thickness in northern California to migration of the Mendocino triple junction. The crust in the vicinity of MTJ is first significantly thickened and then equivalently thinned over a distance of a few

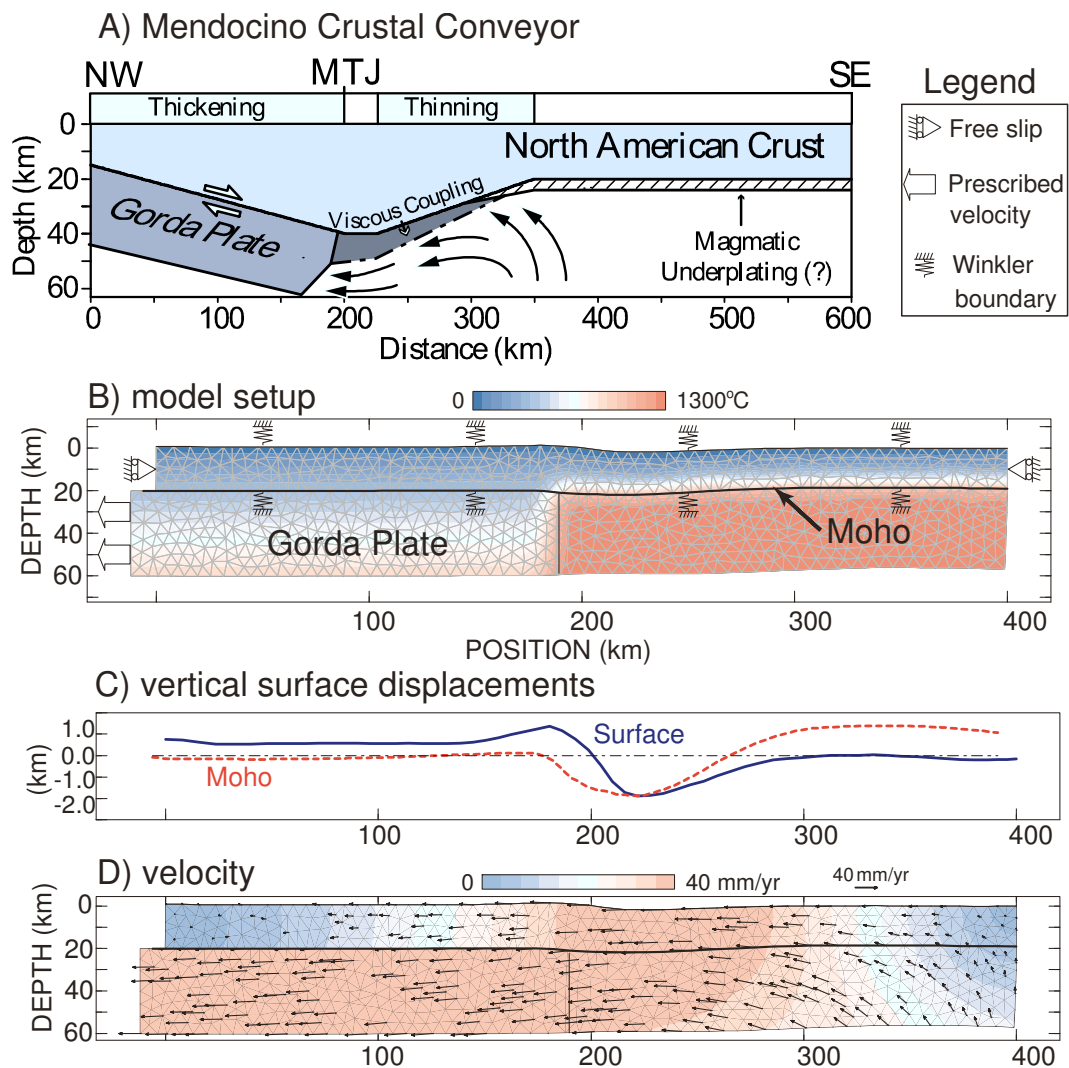


Figure 3.6 Thermomechanical model of the slab removal from beneath the North American plate developed by Furlong and Govers (1999). Shown are the model configuration and results after 300 k.y. of evolution. (A) Thermal structure, boundary conditions, and finite element discretization. Applied velocity is 40 mm/yr. (B) vertical displacement of surface and Moho, relative to initial conditions. (C) Resulting velocity distribution. (D) Schematic diagram of the Mendocino Crustal Conveyor model showing the extent of crustal thickening and thinning as driven by viscous coupling between the migrating Gorda Plate and overlying North American crust (after Guzowski and Furlong, 2002). Cooling of the upwelling asthenosphere causes magmatic underplating at the base of the crust south of the triple junction.

hundred kilometers. According to the “Mendocino Crustal Conveyor” (MCC) conceptual model (Fig 3.6A) of Furlong and Govers (1999), the thickness variations are due to viscous coupling between the northward-migrating Gorda slab and the base of North America south of the triple junction. They suggested that mantle material accreted to the southern edge of Gorda slab due to conductive cooling is in the mechanical continuity with mantle accreted to the base of North America south of the triple junction. In this way, the Gorda slab causes the thickening of overlying North American crust in advance of the triple junction and then thinning of that thickened crust after triple junction passage.

Furlong and Govers (1999) tested the MCC hypothesis using relatively short-term (500 Kyr – 1 Myr) plain-strain thermomechanical finite element model. The setup (Fig 3.6B) was

developed using the finite element code TECTON (Govers and Wortel, 1995) and included elasto-visco-plastic material rheology. The Gorda slab was driven kinematically with the applied velocity 40 mm/yr. The North American crust could move vertically, but was not allowed to change in total length. The initial temperature distribution was assigned consistently with the three-dimensional modeling results of Goes et al. (1997). The model included elasto-visco-plastic material behaviour with the mantle and crustal viscosities calculated according to temperature-dependent power-law creep.

The results in Fig 3.6C-D show the instantaneous picture of a steady-state crustal deformation process near the MTJ (surface displacements and velocity). The authors suggest that these results can be integrated in time to simulate the long-term pattern of crustal structure produced by migrating triple junction. The model succeeded in matching the various observations. In particular the predicted pattern of crustal deformation is in agreement with the reflection and wide-angle refraction results (line 9 from Beaudoin et al., 1996). The average crustal thickening (thinning) rates range in the model about 4-5 mm/yr, which creates the potential for 20 km variations of the North American crustal thickness as the triple junction traverses the region. The model-predicted earthquake activity is attributed to crustal shortening north of the MTJ. Simultaneously, the low-seismicity zone just south of the MTJ (Castillo and Ellsworth, 1993) is reproduced in the model as a zone of low strain rates. Further south there is a region of crustal thinning, which is also consistent with the observed seismicity pattern (Castillo and Ellsworth, 1993).

Guzofski and Furlong (2004) complemented the MCC tectonic model with a one-dimensional finite-difference thermal model to study the long-term (15 Myr) implications of Gorda slab migration on the Coast Range surface heat flow. They used background heat flow of 75 mW m^{-2} and adjusted the crustal thickening/thinning rates of Furlong and Govers (1999) to obtain a best fit model. The effects of the mantle upwelling and the crustal thickness variation were estimated separately. Guzofski and Furlong (2004) showed that the thickening (from 15 to 40 km) of the North American crust initially produces a decrease in the surface heat flow compatible with the observations in northern California. Subsequently (after 5 Myr) the emplacement of the asthenosphere in the slab window beneath the thinning crust leads to the sharp increase of the surface heat flow south of MTJ. The high heat flow values near the San Francisco area are attributed to the further conductive heating of thin (20km) North American crust from the emplaced asthenospheric material. They also showed that lower temperatures of the asthenospheric upwelling (900 C) are able to more precisely fit the observed heat flow values. This result is attributed to the cumulative effect of crustal thinning, which was ignored by previous models (e.g. Goes et al. 1997).

3.3 Three-dimensional models

To date there is a lack of a coupled long-term ($> 1\text{Myr}$) thermomechanical model for evolution of the North America lithosphere and the San Andreas Fault System. The 3D models published in the literature consider only the relatively short-term deformation ($< 1\text{Myr}$), or just study the thermal and mechanical processes separately, without a feedback. The 3D models have substantial advantage over the both 2D and thin-shell approximations. The velocity and/or temperature fields (if modeled) contain no assumptions about their spatial variation. The general lack of a long-term model of a strike-slip fault, such as the San Andreas Fault, is explained by difficulties related to handling of the material advection in the highly deformed lithospheric domain. The adequate model must resolve lithospheric rheology, strain localization, surface erosion and sedimentation, and motion of the free surface (e.g. Gerya and Yuen, 2007; Popov and Sobolev, 2008; Braun et al., 2008). This thesis is an attempt to make a progress in creating such a model. However, the published works must be analyzed to be used as a basis of the 3D formulations. Some of the presented models are discussed below.

3D thermal models

Van Wijk et al. (2001) studied the stalled-slab and slab-window scenarios (e.g. Bohannon and Parsons, 1995) for cessation of the Farallon plate subduction and the post-subduction history of Monterey and Arguello microplates. They calculated two three-dimensional thermal models using the finite difference code of Goes et al. (1997). In the stalled-slab scenario the subduction ceases, and the microplates become attached to Pacific Plate, while in the slab window scenario the subducting slabs break off and sink into the mantle. Both concepts are schematically shown in the Fig 3.7b (see also Chapter 2). In both model simulations the plate velocity fields were prescribed. Other important aspects of the thermo-mechanical coupling, such as the shear heating, were also ignored. The major shortcoming of this approach is that it cannot adequately account for the mantle flow in the slab window, which is difficult to assume in three-dimensions.

The models of Van Wijk et al. (2001) included the time interval between 19 Ma and present. The initial location of the model domain is shown together with schematic plate tectonic evolution of the studied region in Fig. 3.7a. The present-day thermal structures in central and southern California resulting from both the slab-window model (left) and the stalled-slab model (right) are shown in Fig 3.7b. The differences between the models are relatively moderate. The maximum temperature difference reaches the magnitude 150 C at depths about 100 km. The cold

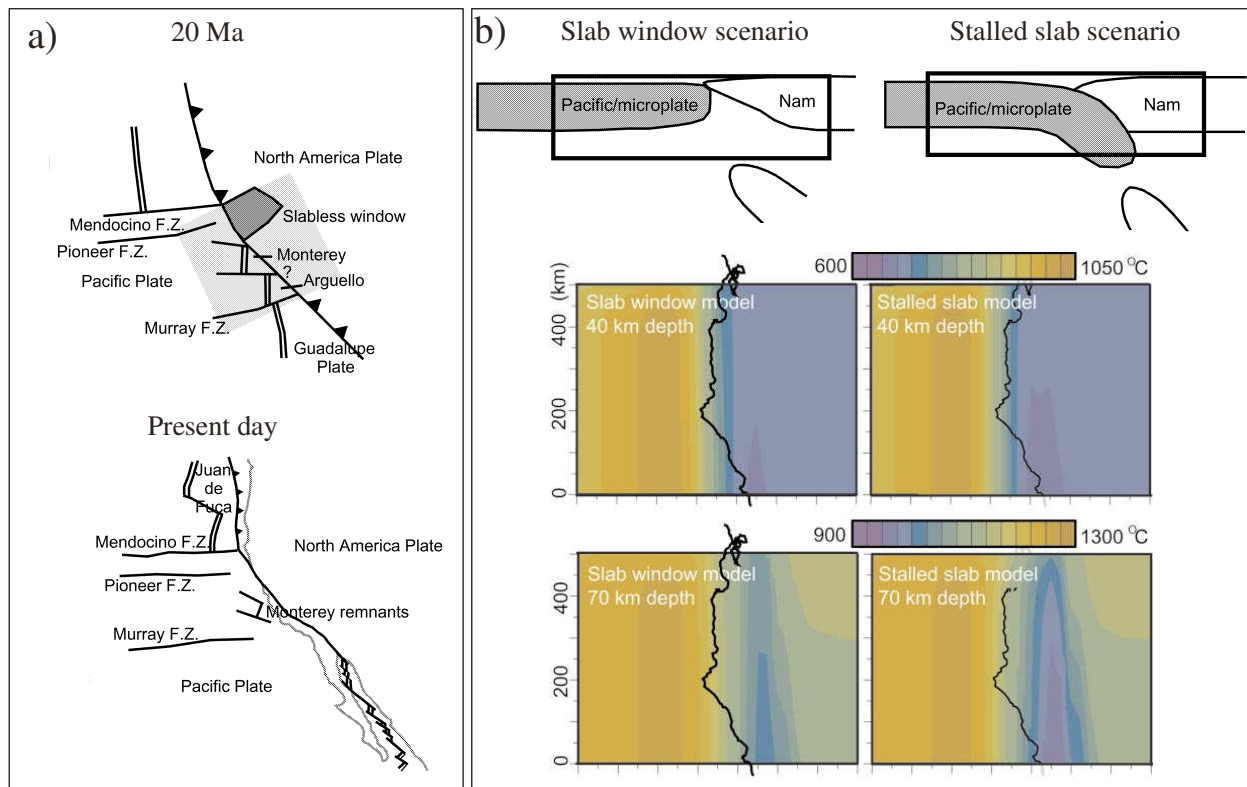


Figure 3.7 Three-dimensional thermal model of the California upper mantle (after Van Wijk et al., 2001). (a) Schematic plate tectonic history of offshore California. At about 20 Ma (top figure) Monterey and Arguello microplates probably cease to subduct. A slab window (dark gray) has opened after subduction of a ridge segment. The light gray rectangle indicates the domain of the numerical model. Present-day structure of the Pacific – North America Plate boundary (bottom figure) includes remnants of the Monterey and Arguello microplates offshore southern California. (b) Sketch and the present-day temperature distribution at depths 40 km and 70 km for the slab window model (left panel) and the stalled slab model (right panel). In the slab window scenario, the slab is absent. The space left behind is filled with asthenospheric material. In the stalled slab scenario, the slab remains attached. The kinematically required gap in the slab is at sub-lithospheric levels.

thermal regions located in the former slab window (or slab) are explained by the influx of the relatively warmer North American material from the north.

The important result of the models developed by Van Wijk et al. (2001) is that the slab-window and the stalled-slab scenarios cannot be discriminated based solely on the thermal constraints. Both models satisfy the heat flow data and tomography observations. This implies that the present day large-scale upper mantle temperature field is relatively insensitive to the assumed microplate scenarios. The models considered in this thesis (Chapter 5) agree with this conclusion in general. However, the presence of the microplate thermal heterogeneities in the upper mantle on the initial stages of the plate capture can significantly affect the deformation of the overlying continent. The local temperature field may be significantly biased by the long-term (> 1 Myr) persisting fault zone.

The similar models considering the thermal aspect of the Pacific, North American, and Juan de Fuca plates interaction in the MTJ was presented previously by Goes et al. (1997). The model

incorporated the rather complicated kinematical effects of an asthenospheric slab window, changes in relative plate motions and the migration of the Juan de Fuca-Pacific spreading ridge. Two end-member scenarios of the slab window kinematics were placed in focus. The hot asthenosphere upwelling was prevented either by underthrusting of the Pacific plate beneath North America prior to modeled period (15 Ma to present) or by accretion of the asthenosphere to the Pacific plate eastern margin during the model evolution. The goal was to explain the wedge-shaped low velocity anomaly in the uppermost mantle layer (30-70 km) of the P-wave tomography model published by Benz et al. (1992), (see also Chapter 2).

The modeling results, well-constrained by the heat flow and seismicity observations, fail to discriminate between the underthrusting and accretionary scenarios. The results of synthetic tomography, which take into account only temperature variations, are inconsistent with the results of Benz et al. (1992). The models presented in Chapter 5 confirm the importance of the asthenosphere conductive cooling from the Pacific plate. However the rate of this cooling is significantly influenced by the thermal boundary conditions. In particular, we suggest that the assuming zero heat flow through the sides of the box is inappropriate (especially for the Pacific side of the box), unless the size of the box is significantly larger than the slab window. The absence of variations in thermal structure outside of the model box does not necessarily imply the absence of heat exchange between the box and the outside material.

3D mechanical models

Braun and Beaumont (1995) pioneered this branch of models by presenting a fully three-dimensional numerical experiment of a strain partitioning at the oblique plate boundaries. Despite a relatively simple idealization (a uniformly thick elasto-plastic crustal layer, driven from below by the obliquely convergent lithospheric mantles) (Fig. 3.8a), the authors were able to apply some of their results for the interpretation of the geological observations from southern Coast Ranges and central South Island in New Zealand. Using a large-strain finite element model, and iteratively solving the force balance equation, Braun and Beaumont (1995) studied the occurrence of the strain partitioning without pre-existing faults or regions of weakness, and in the absence of strain softening. They considered two basic setups: either purely compressive experiment (Fig 3.8b) or purely transcurrent (strike-slip) experiment (Fig 3.8c).

In both cases the modeling results demonstrated excellent agreement with the observed geological structures: thrust faults in the compressive case, and a set of quasi-vertical planes in the strike-slip experiment (the so called flower structure). Both models appear to be relevant for the present day slightly transpressive strike-slip deformation at the SAFS (Argus and Gordon, 2001). Unfortunately, the assumed rigidity of the mantle lithosphere renders the results of Braun

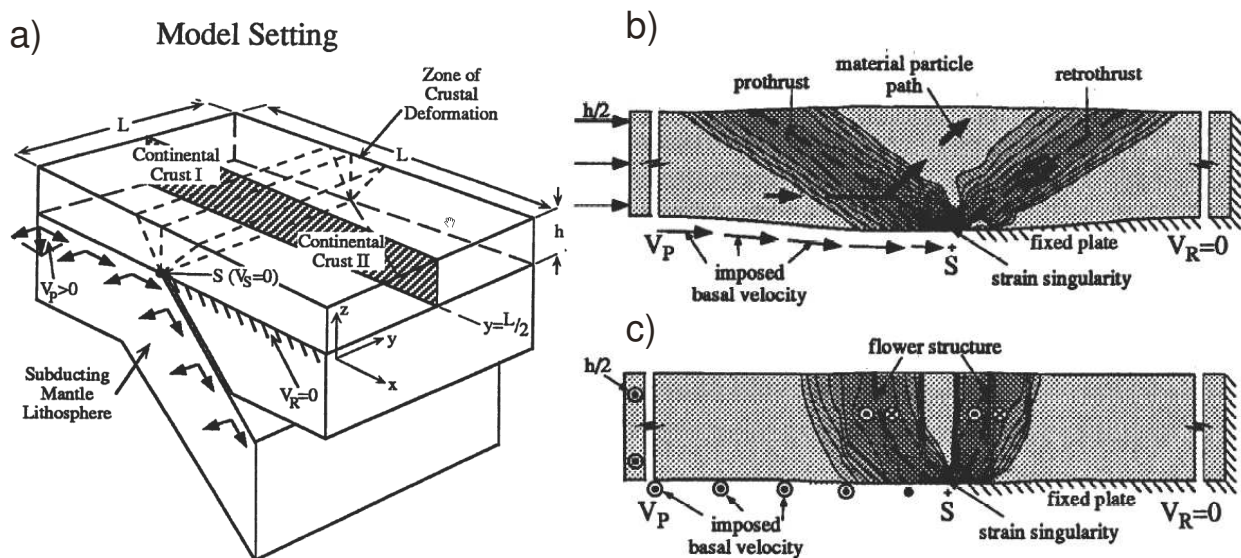


Figure 3.8 Three-dimensional finite element model of crustal deformation over the two near-rigid obliquely-converging lithospheric mantles (from Braun and Beaumont, 1995). Model was used to study strain partitioning at oblique plate boundaries (particularly in the southern Coast Ranges). (a) Geometry and boundary conditions. The model includes only the laterally uniform crustal layer comprising the pro-region (Continental Crust I) and retro-region (Continental Crust II), the basal boundary conditions, and vertical reaction forces to achieve a local isostatic balance. The crust undergoes shortening and shearing focused in the region above S (dashed lines) in a manner that depends on the obliquity of the subducting mantle velocity (V_p). (b) and (c) are the results of pure contraction and pure transcurrent models, respectively. The grayscale shading shows the distribution of the horizontal compressive incremental strain. In the contraction case (b) the deformation is concentrated along two planar thrusts. In the transcurrent case (c) the deformation is accommodated by quasi-vertical shear planes arranged in a so-called “flower” structure.

and Beaumont (1995) inapplicable for the long-term ($> 1\text{Myr}$) evolution of the SAFS, which is strongly affected by the mantle processes. However, this fact does not diminish implications of the results to the crustal deformation.

The multi-scale finite element models of Liu et al. (2007) demonstrate a current state of the short-term ($< 1\text{Ma}$) three-dimensional mechanical modeling. These models consider both the present day deformation of the entire western United States including the major tectonic provinces, and the strain partitioning at the SAFS on the time scales of up to 1 kyr. The large-scale model (Fig. 3.9a) was used basically to study the influence of the various driving forces and the 3D lithospheric strength distribution on the deformation in the western US. The power law viscous rheology was assumed for the 200 km thick lithosphere. The fault zones were represented by the regions of the relatively low viscosities and the plates were driven kinematically. The models were calculated using the fully parallel finite element modeling technique, which extensively utilizes the domain decomposition method and the Message Passing Interface.

Fig. 3.9b-d shows the major results of the large-scale models developed by Liu et al. (2007). They separately considered the effects of gravitational potential energy, plate boundary forces,

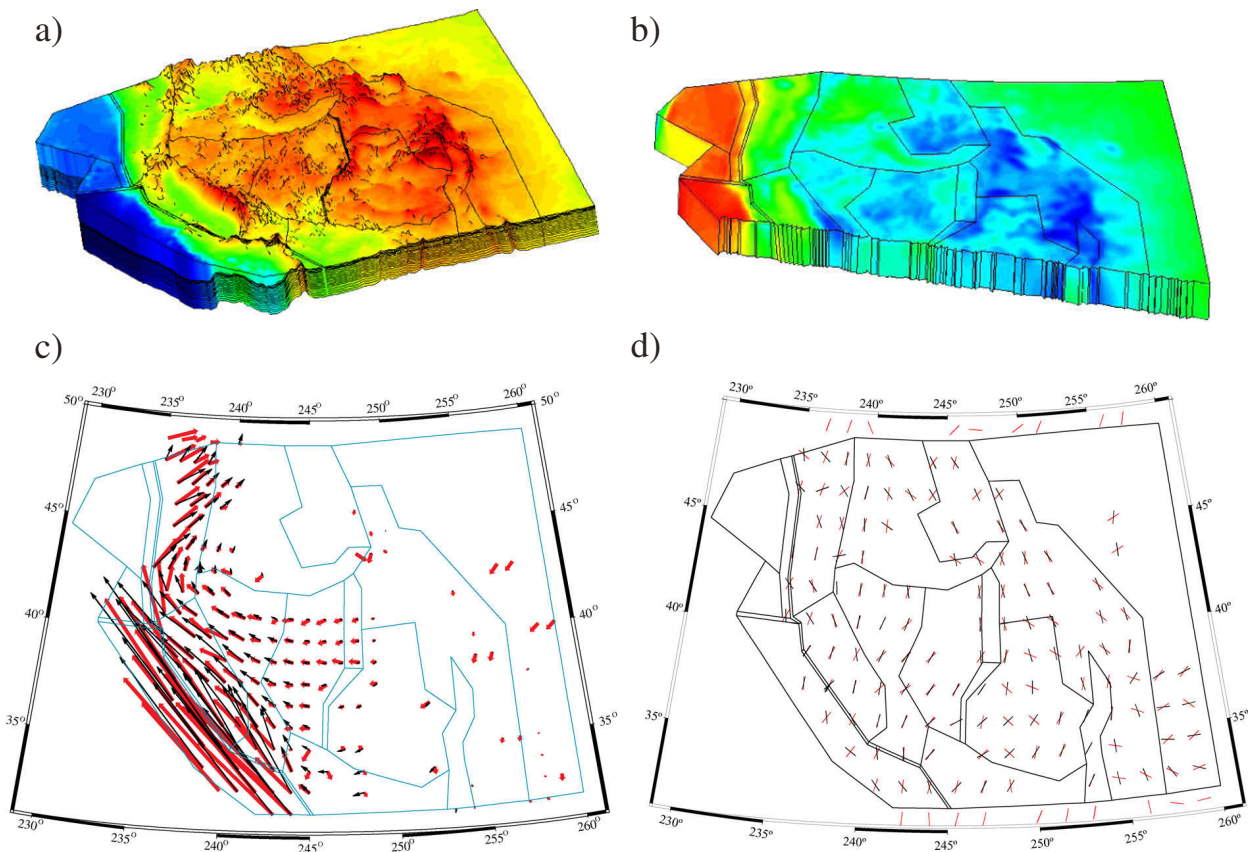


Figure 3.9 Three-dimensional finite element models for active tectonics in the western US (from Liu et al. 2007). (a) Finite element discretization of the model. The mesh consists of 693522 triangle prism elements and 371260 nodes. The model spans to 200 (km) depth, contains 18 mesh layers, and includes boundaries of major tectonic units. The topography is shown in the figure with vertical exaggeration. (b) Contour plot of the predicted vertical stress, which is largely dependent on topographic loading. (c) Comparison of the predicted surface velocities (black arrows) with the averaged GPS velocities (red arrows). (d) Comparison of the predicted directions of the maximum horizontal compressive stresses (black bars) with those from the World Stress Map (red bars).

and the basal shear. In this preliminary model the gravitational potential energy resulted mainly from the topography variations. The calculated vertical stress shown in Fig. 3.9b represents the fact that the topographic loading is the major contributor to the vertical stresses. Fig. 3.9c-d show the calculated surface velocity and the S_{Hmax} stress orientations, respectively, compared with the GPS data and the World Stress Map (<http://www-wsm.physik.unikarlsruhe.de>). This typical model considered the effect of both the boundary forces and the gravitational potential energy and included moderate viscosities on the plate boundaries ($10^{18} - 10^{21}$). The average stress orientation misfit constituted up to 31° , and the average surface velocity misfit was about 5–8 mm/year.

Despite the significant misfits in the modeling results, the work of Liu et al. (2007) represents a relatively new trend in the 3D numerical modeling. In future this approach may completely replace the thin-shell approximation and bring the advantage of the depth-varying velocity and lithospheric strength. At present it is not fully obvious whether authors prefer using

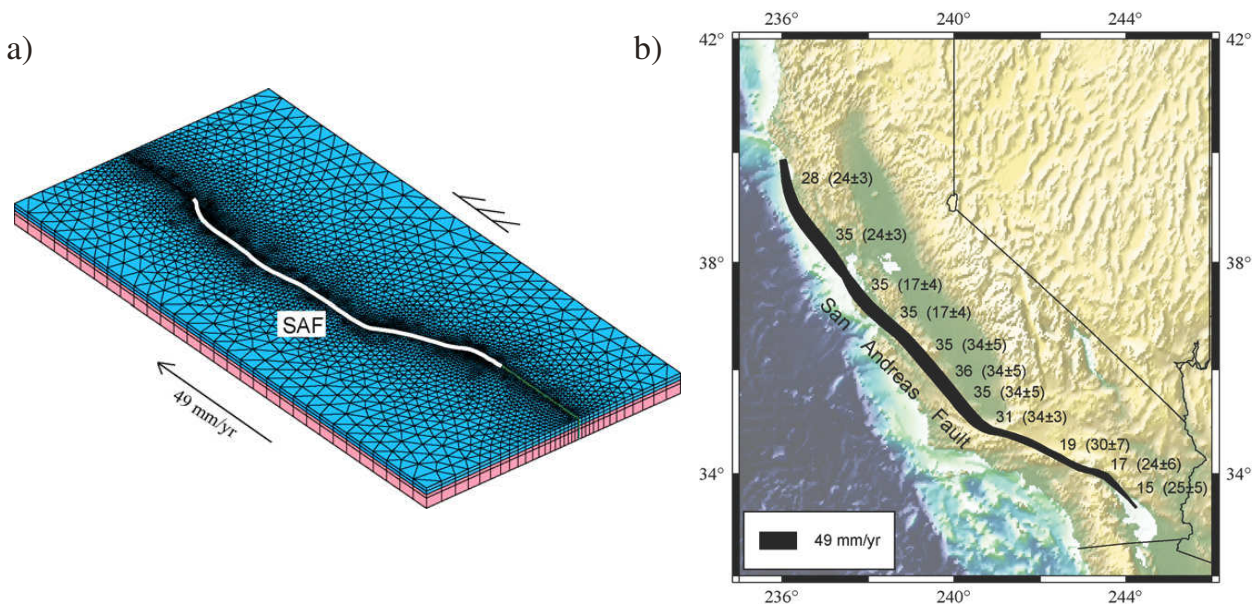


Figure 3.10 Three-dimensional visco-elasto-plastic model of the San Andreas Fault System (from Liu et al. 2007). (a) Numerical mesh and boundary conditions of the finite element model. The entire San Andreas Fault (black line) is explicitly included in the model. (b) Comparison between the predicted long-term slip rates along the San Andreas Fault and the geological slip rates (numbers in parenthesis) from the California Geological Survey <http://www.consrv.ca.gov/CGS/rghm/psha/index.htm>.

the vertically averaged or depth-varying viscosities. The 3D nonlinear formulation of their model is clearly advantageous over the similar thin-shell approximation (e.g. Humphreys and Coblenz, 2007). The latter, however, enjoys the linearity of the assumed rheology, which allows considering the superposition of loads.

The regional-scale model of Liu et al. (2007) was used to study the strain localization and the slip rates along the San Andreas Fault. The geometrical setup and the finite element mesh of this two-layered model are shown in Fig. 3.10a. The model included a 20-km thick upper crust with an elasto-plastic rheology, and a 40-km thick viscoelastic layer representing both the lower crust and the uppermost mantle. Details of the modeling technique are given in (Li and Liu, 2006). The model included a predefined San Andreas Fault with much lower friction and cohesion than surrounding crust.

The comparison between the observed and calculated steady-state surface slip rates are shown in Fig. 3.10b. Shown are the time-integrated continuous creeping rates that effectively characterize the stick-slip process occurring over shorter time-scales. The correlation between the observed and calculated values is to a great extent the function of the lithospheric rheology and thermal and compositional heterogeneity. Another important aspect concerns the existence of the multiple sub-parallel faults (Hayward, Calaveras, Elsinore, San Jacinto), which accumulate significant slip rates together with the SAF. Both their existence and initiation is controlled by distribution of the lithospheric strength which is influenced by many parameters. The results of the relatively simple model of Liu et al. (2007) are in fairly good agreement with

the observations. This thesis considers the aspects influencing the slip distribution between the sub-parallel faults in more details (see Chapter 5).

3.4 Extended two-dimensional models

The extended 2D formulation is somewhat intermediate between the 2D plane-strain and the complete 3D models. The vertical cross-section in the extended model is oriented approximately perpendicular to the strike-slip direction. Unlike the plane strain models, the along-strike velocity components are assumed to be non-zero in the extended formulation. In general, this assumption allows for the material advection through the cross section of the model. However, the along-strike variations of the velocities and the entire properties of the advected material, such as stresses, temperatures, composition, etc., are ignored. This formulation, in principle, is suitable to model a single infinitely long strike-slip fault or a system of sub-parallel faults without junctions or significant along-strike lithospheric strength variations. The junction points create the state of transpression/transension in the system, which should be handled in detail with a full 3D formulation. The advantage of the extended model is essentially a two-dimensional geometry that substantially saves the computational effort and allows for higher resolution. Recently, the aspects of the deformation at the SAFS have been studied using the extended 2D formulation.

Short-term evolution of the crust

Chery et al. (2001) analyzed the relations between the rheological properties of upper lithosphere and the short-term (50 kyr) deformation of the San Andreas Fault using a finite element model. Fig. 3.11a shows a sketch illustrating the setup of the model. The initial finite element discretization is shown in Fig. 3.11c. Despite the fact that a complete 3D formulation was used, the model actually falls into the extended 2D category, because of the ignored along-strike variation of the entire fields. The authors considered an equally thick (25 km) crustal layer obeying rather comprehensive elasto-visco-plastic rheology with temperature-driven brittle-ductile transition. The domain included a predefined weak zone (the San Andres Fault) with the friction angle lower than that in the adjacent material. The buoyant crust was assumed floating over the inviscid mantle with zero shear tractions, which appears to be an adequate approximation at the initial stages of the slab window formation. The deformation in the crust was driven by the applied relative velocities between the Pacific plate (left side) and the Great Valley (right side) (Fig. 3.11a). The initial temperature variation was varied to match either the

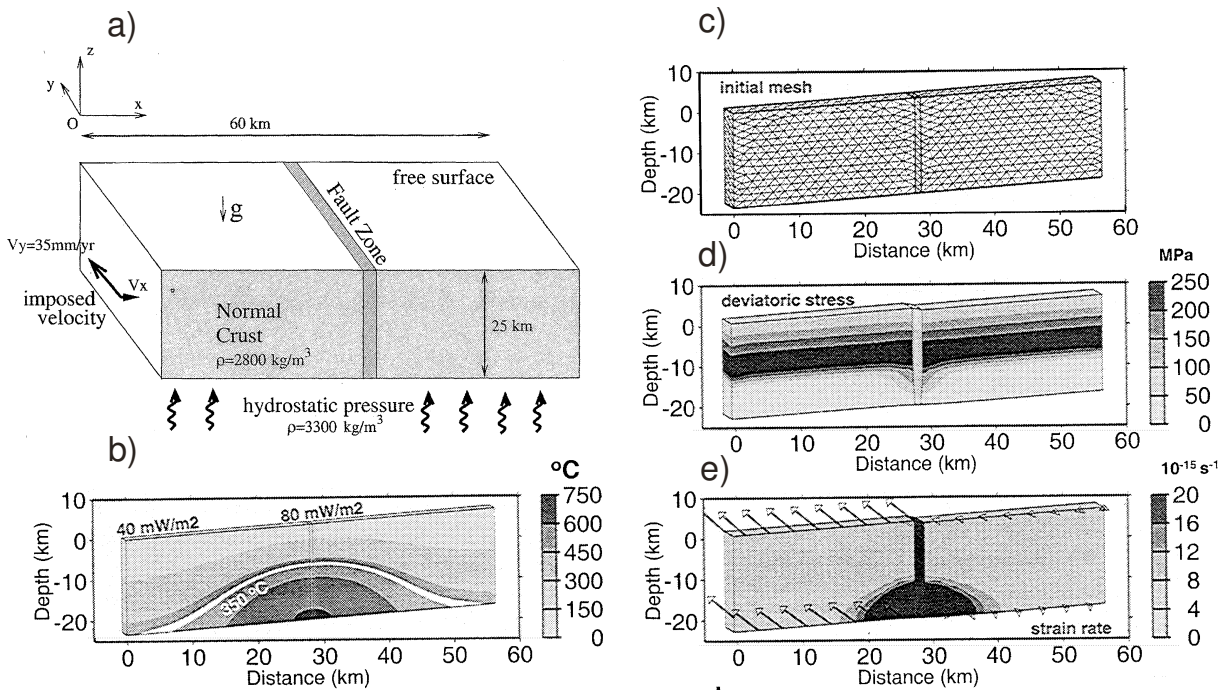


Figure 3.11 Extended two-dimensional mechanical model of the San Andreas Fault (from Chery et al., 2001). (a) Geometry and boundary conditions of the model. The model has no along-strike variation of the geometry and thermal conditions. The left side corresponds to the Pacific Plate and has the relative velocity components (V_x and V_y) with respect to the right side (Great Valley). Only 25 km of crust is included. The mantle contribution to the integrated strength is ignored. The fault zone is predefined by the friction angle lower than surrounding crust. (b) Temperature distribution with strong lateral heat flow variation used in the modeling. The value of 80 mW m^{-2} is typical for the SAFS. The value of 40 mW m^{-2} is typical for the Great Valley. (c) Initial mesh. (d) Deviatoric stress norm. Stress increase in the upper crust reflects the frictional behavior. Stress decrease in the lower crust reflects the temperature-dependent viscous behavior. (e) Total strain rate. Strain rate is localized in the upper crust of the fault zone and diffuses in the lower crust below.

uniform heat flow typical of SAF (80 mW m^{-2}), or the typical values of Great Valley (40 mW m^{-2}) on the flanks of the model. The latter case is shown in Fig. 3.11b. The evolution of the temperature field and the thermo-mechanical coupling were ignored due to relatively short modeled period (50 kyr). Among the other varied parameters are the frictional strength of the SAF, and the slip-related softening in the crust.

Fig. 3.11d-e shows the distribution of the deviatoric stress and the total strain rate, respectively, from the typical model of Chery et al. (2001). The deviatoric stress results (Fig. 3.11d) show presence of the brittle-ductile transition in the model. Below the depth of the relative stress peak, the viscous deformation takes over, while above this depth, the stress profile is controlled by the frictional elasto-plastic rheology. The depth of brittle-ductile transition is controlled by the temperature distribution. The deviatoric stress on the fault is very low, which is consistent with its assumed low frictional strength. In the brittle layer, the strain rate concentrates mostly at the fault (Fig. 3.11d), while the adjacent crust behaves mostly rigidly. In the ductile

part there is a wide shear zone with the elements of sub-horizontal detachment below the brittle layer.

By varying the frictional strength of the fault and changing the initial temperature distribution Chery et al. (2001) obtained the different profiles of the strike-slip velocity across the model. They showed that the temperature increase in the ductile roots beneath the fault can significantly increase the slip rate on the fault. On the other hand, the less is the frictional contrast between the fault and the adjacent crust, the more slip “leaks” from the fault and accommodates in the adjacent crust. This important conclusion results from the fact that the deformation is attracted towards the regions of lower lithospheric strength, which is a sum of the brittle and ductile components. Generally, only the frictional strength is discussed (e.g. Scholz, 2000), whereas the ductile strength remains completely out of focus. These results, however, explicitly emphasize the importance of the ductile component of the fault strength. Chery et al. (2001) also studied a combination of the moderately strong fault and a strain softening. For this case they were able to localize the slip rate in the surrounding crust into a distinct sub-parallel fault zone. This is mainly due to fact that the strain softening acts as an additional localization factor.

The models of Chery et al. (2001) present a substantial step forward in understanding of the strike-slip fault dynamics, such as the SAF. They considered many important factors controlling the partitioning of the slip-rate between the sub-parallel faults and the adjacent crust. However, certain aspects have been left off the consideration. These are, in particular, the influence of the mantle and the slab window, the thermo-mechanical coupling, and the long-term simulation (> 1 Myr). The models described in the next section study these parameters using the extended 2D formulation.

Long-term evolution of the lithosphere

Sobolev et al. (2003) developed a fully coupled long-term (20 Myr) thermo-mechanical model of the SAFS in northern and central California. They used an explicit time marching technique (Babeyko et al., 2002) with the elasto-visco-plastic material behavior. The setup of the model, including the boundary conditions and the formulation (Fig. 3.12a) was essentially similar to that of Chery et al. (2001). However the model included significantly thicker lithospheric domain (80 km), which incorporated two-layered crust of variable thickness and composition, and a significant portion of the upper mantle.

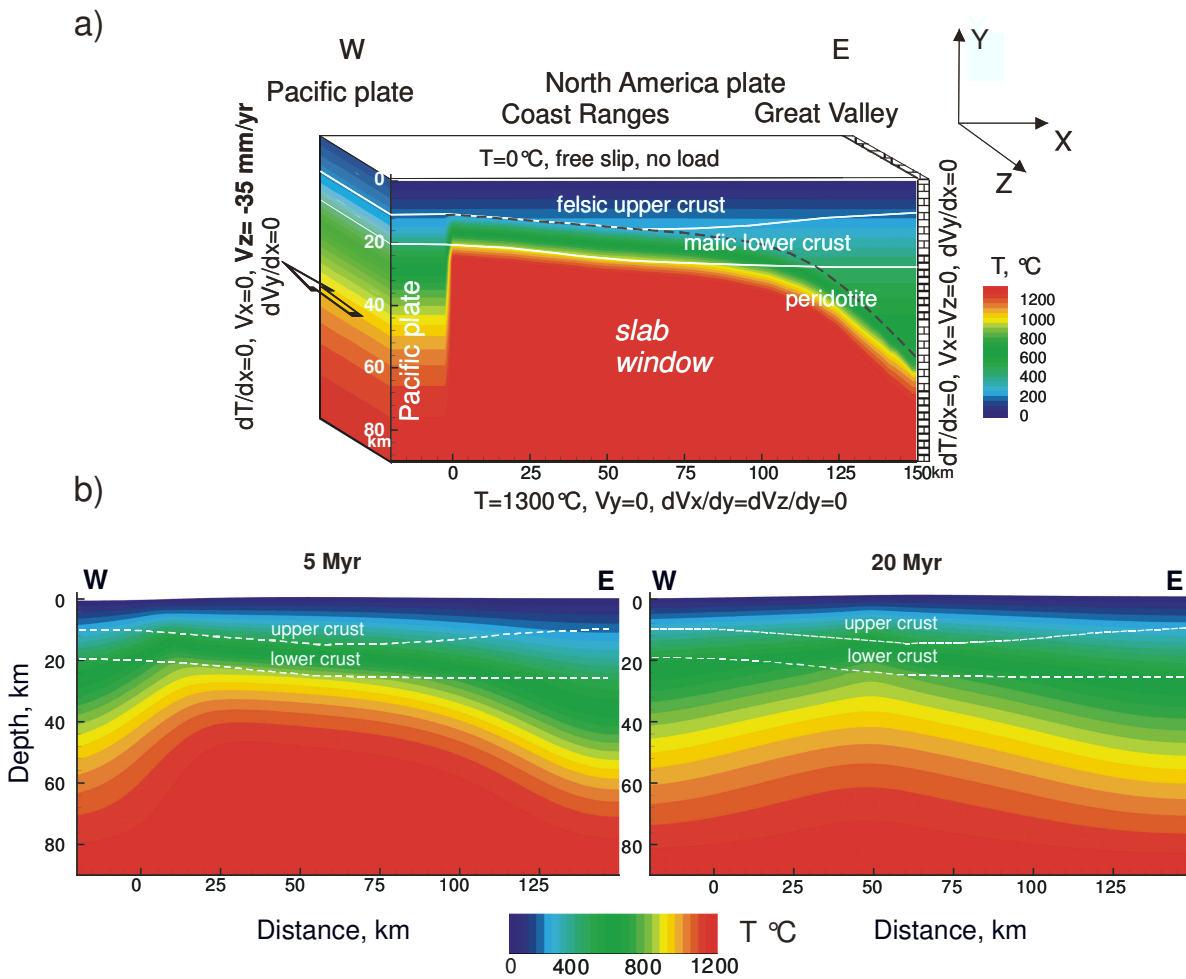


Figure 3.12 Extended two-dimensional geological time scale thermomechanical model of the SAFS evolution in central and northern California (after Sobolev et al. 2003). (a) Setup of the model. Colors show initial temperature distribution (just south of the Mendocino Triple Junction). Model includes two-layered crust. The geometry of the slab window is constrained by the top of former Juan de Fuca slab (dashed line). The approximately northward velocity of the Pacific Plate (right side) is assumed to be 35 mm/year relative to the Great Valley block (left side). Similar to Chery et al. (2001), the along-strike variations are neglected. (b) Thermal evolution of the slab window. Shown are the contour plots of the temperature distribution at 5Myr and 20Myr after MTJ passage.

Unlike the models of Chery et al. (2001), the models of Sobolev et al. (2003) focused mostly on the long-term evolution of the lithosphere and incorporated the important effects of the thermo-mechanical feedback, such as the shear heating. The predefined fault zones were not included in the model, but formed spontaneously following the variation of the lithospheric strength. The models included the strain softening to simulate the reduction of the frictional strength with increased fault slip (slip-weakening). The magnitude of slip weakening was considered as a model parameter to simulate either weak or moderately strong faults.

With this relatively simple extended 2D models, Sobolev et al. (2003) attempted to relate the process of thermal evolution in the slab window following the passage of Mendocino Triple Junction (previously modeled by Furlong, 1984; see Fig. 3.5) with the partitioning of the strike-slip deformation between the Great Valley and Pacific plate. The initial configuration of the slab

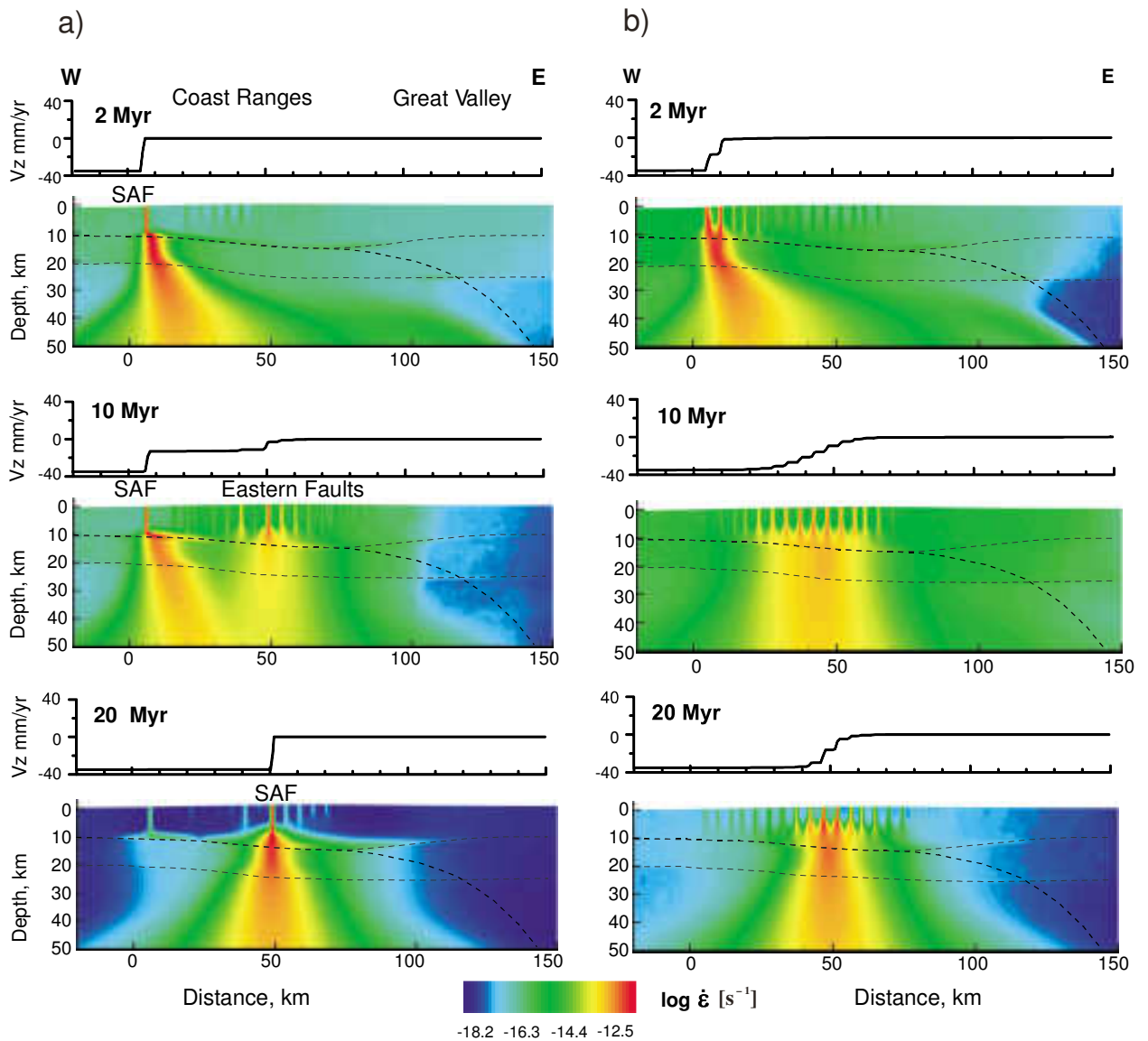


Figure 3.13 Results of the thermomechanical models of the SAFS developed by Sobolev et al. (2003). Shown are the contour plots of the effective logarithmic strain rate and the strike-slip velocity distribution at 2Myr, 10Myr, and 20Myr after MTJ passage. (a) The weak faults model (reduction of the friction angle from 30° to 3° at plastic strain of 2.0). The model generates distinct sub-parallel faults. (b) The moderately strong faults model (reduction of the friction angle from 30° to 20° at plastic strain of 2.0). The model generates numerous closely spaced faults which is not consistent with the structure of the SAFS.

window and the temperature distribution in the wake of MTJ are shown in Fig. 3.12a. After the triple junction passage, the sharp lateral temperature gradients start to equilibrate due to conductive cooling. The distribution of the temperature fields at 5 Myr and 20Myr are shown in Fig. 3.12b. The eastward migration of the lithospheric strength minima, caused by cooling of the mantle in the slab window, should eventually displace the plate boundary in the same direction. The similar mechanism, called eastward “jump” of the SAF, was qualitatively discussed by Furlong et al. (1989) and Furlong (1993) (see also section 3.2). However in the work of Sobolev et al. (2003) this mechanism was tested quantitatively by a numerical model.

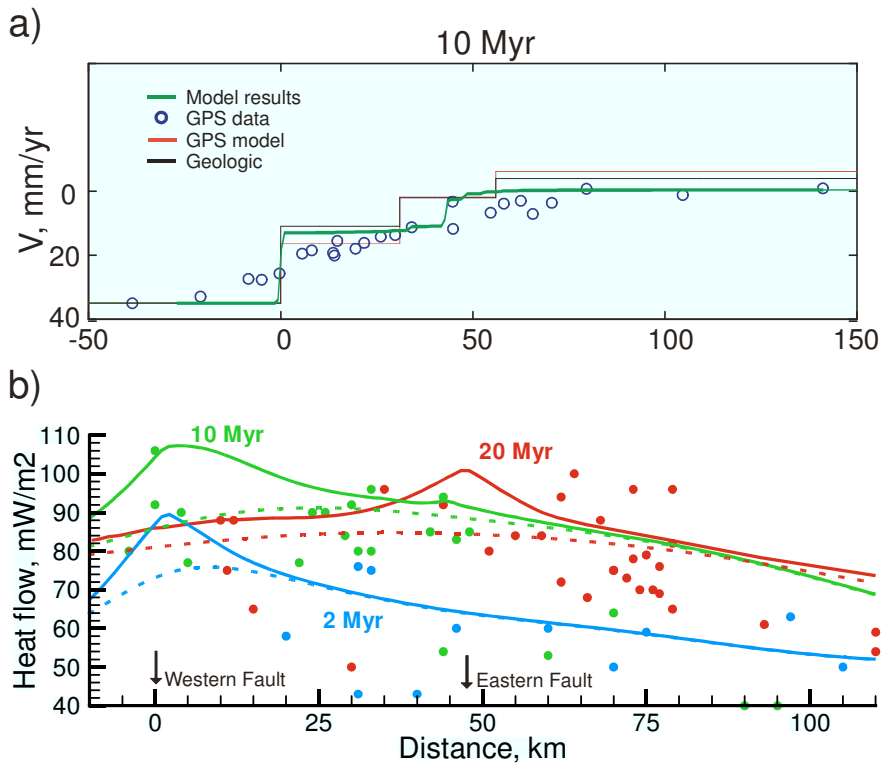


Figure 3.14 Results of the thermomechanical models of the SAFS developed by Sobolev et al. (2003). (a) Comparison of the strike-slip velocities from the model (green lines) at 10 Myr after MTJ passage with geodetic observations and models (Freymueller et al., 1999; Murray et al., 2001), and long-term geologic slip rates. Distance is measured from the westernmost fault in the SAFS and in the model. (b) Distribution of the surface heat flow in the model at 2 Myr (cyan solid curve), 10 Myr (green solid curve) and 20 Myr (red solid curve) versus heat flow data from the regions of comparable average age. Dashed curves show heat flow distributions in the model without incorporation of shear heating. Arrows show locations of the western and eastern faults in the model. Distance is measured from the westernmost fault in the SAFS and in the model.

Fig. 3.13 shows the evolution of the effective strain rate in the lithosphere and the fault slip rate following the migration of the MTJ for both the weak (a) and the moderately strong faults (b). These results emphasized the importance of the slip-related weakening for the generation of the distinct fault zones in the brittle crust. The model with the weak faults successfully predicted the crustal plate boundary migration towards east, and relatively well estimated the distance between the sub-parallel faults. The San Francisco Bay area is coincident with the location of the MTJ at approximately 10 Ma. The eastern sub-parallel faults in Fig. 3.13a (middle) can be interpreted as the Hayward and Calaveras faults. The modeled distance between the faults (50 km) correlates with the observed values. On the other hand, the model with the relatively strong faults is completely incapable to form distinct faults in the brittle crust (Fig. 3.13b). The rather diffuse shear zone freely follows the region of the smallest lithospheric strength. Therefore, the models of Sobolev et al. (2003) independently confirmed the assumption of the weak San Andreas Fault (e.g. Zoback, 2000) from the viewpoint of the thermomechanical modeling.

The comparison between the modeled and the observed distributions of the fault slip-rate (a) and the surface heat flow (b) are shown in Fig 3.14. The modeling results, in general, demonstrate a relatively good correlation with the observations. The predicted present day slip-rate distribution at the San Francisco Bay Area, for instance, closely matches both the geodetic and the long-term geological estimates (Fig 3.14a). The heat flow anomaly in the Coast Ranges, resulting from the upwelling of the hot asthenosphere, is relatively well reproduced by the calculated heat flow values (Fig 3.14b). The heat flow after 10Myr is slightly higher than that after 20 Myr, which reflects the subsequent cooling of the asthenosphere at shallow depths. The model also approximates the eastward decrease of the heat flow towards the Great Valley. For comparison, Fig 3.14b shows the results of the models without shear heating (dashed lines). These results demonstrate that only a small fraction of the modeled heat flow is related to the shear heating on the faults, because of their low frictional strength. The maximum energy release associated with the shear heating occurs in the ductile lower crust and in the upper mantle. The model with the moderately strong faults (not shown here) completely fails to satisfy the heat flow constraint if the shear heating is incorporated to full extent.

The models of Sobolev et al. (2003) represent the first attempt to address the long-term evolution of a sub-parallel fault system. Despite many positive properties described above, the models possess serious drawbacks that prevent their application to the entire SAFS. The major limitation is the absence of the along-strike variation, which prevents occurrence of the fault step induced transpression. This factor, for instance, shortens the period of simultaneous activity of the sub-parallel faults in the Bay Area (0.5-1.0 Myr in the model vs. 5-10 Myr in the observations). The model ignores the influence of the subducting Gorda plate on the north, and the stalled microplates on the south (Nicholson et al. 1994), which substantially influences the deformation of the North American crust. The essentially 3D mantle flow in the slab window is also ignored. The assumed zero heat flux between the model domain and the Pacific plate significantly reduces the rate of asthenosphere cooling in the slab window. All these drawbacks are either removed or significantly reduced in the three-dimensional models presented in the Chapter 5.

Chapter 4

Numerical modeling techniques^{*}

This chapter describes a methodology for the three-dimensional lithospheric-scale modeling of the solid state deformation including strain localization processes. The new code, that we developed in the framework of present work (SLIM3D) includes a coupled thermo-mechanical treatment of deformation processes and allows for an elasto-visco-plastic rheology with diffusion, dislocation and Peierls creep mechanisms and Mohr-Coulomb plasticity. The code incorporates an Arbitrary Lagrangian Eulerian formulation with free surface and Winkler boundary conditions. SLIM3D is developed and implemented using the C++ object-oriented programming language. Here we describe aspects of physical models as well as details of the numerical implementation, including the Newton-Raphson solver, the stress update procedure, and the tangent operator. The applicability of the code to lithospheric-scale modeling is demonstrated by a number of benchmark problems that include: (i) the bending of an elastic plate, (ii) the sinking of a rigid cylinder into a viscous fluid, (iii) the initiation of shear bands in the brittle crust, (iv) triaxial compression test, and (v) lithospheric transpressional deformation. The directions of further development are outlined in the end of the chapter.

4.1 Introduction

On geological time scales, lithospheric rocks deform by three fundamentally different phenomenological mechanisms. These are: (i) elastic mechanism (reversible), which on geological time-scales occur mostly during flexural deformation of the lithospheric plates or bending of slabs in subduction zones; (ii) viscous mechanism, typical of the convective mantle, or for lower crustal flow at plate boundaries; and (iii) plastic mechanism, which simulate brittle

^{*} This chapter is based on the publication: Popov, A.A., Sobolev, S.V., SLIM3D: A tool for three-dimensional thermomechanical modeling of lithospheric deformation with elasto-visco-plastic rheology, *Phys. Earth Planet. Interiors* (2008), doi:10.1016/j.pepi.2008.03.007

failure and manifests itself in the formation of narrow shear zones or faults.

The quantitative proportion between each of these mechanisms depends on many parameters, the most important of which are temperature and stress. Viscosity, for instance, varies by several orders of magnitude with temperature change of a few hundred degrees. At low temperatures, viscous deformation is prohibited which leads to a buildup of elastic stresses during deformation, until the rock fails. This type of behavior characterizes brittle or elasto-plastic deformation. At high temperatures, on the other hand, the rock may accommodate high deformation rates through the viscous mechanism without excessive stress accumulation. This is a typical manifestation of the ductile flow regime.

The lithosphere is not only a region of large compositional heterogeneity, contrasting in this respect with asthenosphere, but it is also a locus of large temperature variations. In this context, it becomes necessary not only to resolve separate brittle and ductile flow regimes, but also to adequately model the brittle-ductile transition. Apart from many other complexities, this single requirement makes lithospheric-scale modeling a challenging task.

This is further complicated by several factors, including: the presence of the Earth's free surface, where erosion and sedimentation processes take place; the occurrence of metamorphic reactions in rocks with changing temperature and pressure; the presence of spontaneously evolving large-scale fault zones; by the inherent three-dimensional nature of most lithospheric-scale problems, and many others. All these complexities pose severe conceptual and implementation problems for lithospheric-scale modeling tools.

In recent decades, the geodynamic modeling community has accumulated significant experience in the development and application of numerical tools designed for modeling various geodynamic processes. These studies include Christensen and Harder (1991), Weinberg and Schmeling (1992), Bercovici (1993), Poliakov et al. (1993), Braun and Sambridge (1994), Bunge and Baumgardner (1995), Fullsack (1995), Trompert and Hansen (1996), Zhong and Gurnis (1996), Schmalholz et al. (2001), Babeyko et al. (2002), Tackley and Xie (2003), Moresi et al. (2003), Sobolev et al. (2005), Muhlhaus and Regenauer-Lieb (2005), Petrunin and Sobolev (2006), O'Neil et al. (2006), Gerya and Yuen (2007), Braun et al., (2008), which have all contributed to the development of geodynamic modeling techniques.

In our research we are mostly interested in highly dynamic deformation processes occurring at plate boundaries involving the combination of transform and compressional/extensional deformation. Such processes are essentially 3D and operate on a temporal scale of a few to several hundred million years and a spatial scale of hundreds to thousands of kilometers. Specific examples include continental collision processes in Tibet and transform deformation at the San Andreas Fault System and the Dead Sea Transform. Numerical techniques to handle such

processes must be 3D, should operate with an elasto-visco-plastic rheology, as this is the most adequate rheology to model lithospheric deformation, and must be suitable to model geological time-scale processes. Unfortunately, none of the tools available at present can be directly applied to these problems. The most suitable 3D codes either still do not include elastic deformation processes (Braun et al., 2008) or, being explicit codes (Sobolev et al., 2005), require very small steps for time integration and therefore are not efficient for modeling long-term deformation. The best candidate among existing tools is the 3D version of visco-elasto-plastic code I2ELVIS, which is currently under development (Gerya and Yuen, 2007). We note, however, that this code is based on a purely Eulerian approach and uses fully staggered finite-difference discretization. Staggered schemes are inherently stable (Shih et al., 1989), but the Eulerian approach substantially complicates treatment of the free surface.

With this in mind, we propose here a new code for lithospheric-scale modeling (SLIM3D). It is an implicit Arbitrary Lagrangian Eulerian Particle-in-Cell Finite Element code with a free surface, designed specifically for the thermo-mechanical modeling of deformation processes involving an elasto-visco-plastic lithospheric rheology on geologic time scales. SLIM3D is being designed to model deformation at plate boundaries (the focus of the recently formed geodynamic modeling group in GFZ-Potsdam). It is intended to complement the capability of the explicit code LAPEX3D already used in the same group for similar problems (Petrunin and Sobolev, 2006). SLIM3D is developed and implemented using the C++ object-oriented programming language.

We emphasize that the potential advantage of SLIM3D over the other similar codes is in its three-dimensional nature combined with the adequate description of lithospheric deformation (rheology, strain localization, free surface etc.). To our knowledge, such a combination is unique at present. At the same time, we note that the code is still limited in certain aspects, which are already implemented in the other codes. These aspects are, for example, the multigrid solver (e.g. Moresi et al., 2003) and the adaptive mesh (e.g. Braun et al., 2008). Despite the present disadvantages, we suggest that SLIM3D has a good potential to be substantially improved in the future versions.

This chapter introduces the numerical method in general with several benchmarks and examples. The structure of the chapter is as follows. In the section 4.2, we describe the basic physical and rheological framework that we utilize for lithospheric scale modeling. Section 4.3 outlines the implementation details of the code. In section 4.4, we test our code with five benchmark problems which demonstrate aspects of elastic, viscous and plastic deformation mechanisms both separately and in combination. Finally (section 4.5), we briefly conclude our work and describe the direction we are planning to further develop our technique.

4.2 Physical models

Conservation equations

Lithospheric-scale deformation can be effectively characterized as quasi-static thermo-mechanically coupled deformation process. Assuming a continuous media approximation, we can describe this process by the conservation equations of momentum:

$$\frac{\partial \sigma_{ij}}{\partial x_j} + \rho g \hat{z}_i = 0, \quad (4.1)$$

and thermal energy:

$$\frac{DU}{Dt} = -\frac{\partial q_i}{\partial x_i} + r. \quad (4.2)$$

Here, x_i ($i=1, 2, 3$) denote Cartesian coordinates, σ_{ij} is the Cauchy stress tensor, ρ is the material density, g is the gravitational acceleration, \hat{z}_i is the unit vector of the vertical axis pointing downward, U is internal energy, D/Dt is the material time derivative, q_i is the heat flux vector, r is volumetric heat sources. In the above equations and in the rest of this chapter, we use indicial notation and apply Einstein summation convention over repeated indices. Since there is no difference between covariant and contravariant components in the Cartesian coordinate system, all tensor indices are written as subscripts. For convenience we additionally explain notation, meaning, and dimension of basic quantities used in this chapter in Table 4.1.

Deviatoric-volumetric decomposition

The thermo-rheological behavior of the rocks is more conveniently formulated in terms of the deviatoric-volumetric strain (stress) decomposition (e.g. Bonet and Wood, 1997). For the Cauchy stress tensor we may write:

$$\tau_{ij} = \sigma_{ij} + p\delta_{ij}, \quad p = -\frac{1}{3}\sigma_{ii}, \quad (4.3)$$

where τ_{ij} is the Cauchy stress deviator and p is hydrostatic pressure (positive in compression).

The deviatoric strain rate tensor and the rate of volume change, respectively, may be written directly as:

$$\dot{\epsilon}_{ij} = \frac{1}{2}(\partial v_i / \partial x_j + \partial v_j / \partial x_i) - \frac{1}{3}(\partial v_k / \partial x_k)\delta_{ij}, \quad \dot{\theta} = \partial v_i / \partial x_i, \quad (4.4)$$

where v_i is the spatial velocity vector. We adopt the second (Euclidean) norm as the effective scalar measure of deviatoric tensorial quantities. For an arbitrary tensor a_{ij} the second norm is

Table 4.1
Nomenclature

Variable	Meaning	Variable	Meaning		
x_i	m	Cartesian coordinates	τ_p	Pa	Peierls stress
t	s	Time	φ	(°)	Friction angle
v_i	$\text{m} \cdot \text{s}^{-1}$	Velocity vector	c	Pa	Cohesion
T	K	Temperature	α	K^{-1}	Thermal expansivity
g	$\text{m} \cdot \text{s}^{-2}$	Gravitational acceleration	C_p	$\text{J} \cdot \text{kg}^{-1} \cdot \text{K}^{-1}$	Specific heat
ρ	$\text{kg} \cdot \text{m}^{-3}$	Density	λ	$\text{W} \cdot \text{m}^{-1} \cdot \text{K}^{-1}$	Thermal conductivity
K	Pa	Bulk modulus	A	$\text{W} \cdot \text{kg}^{-1}$	Radiogenic heat production
G	Pa	Shear modulus	q_i	$\text{W} \cdot \text{m}^{-2}$	Heat flux vector
B_L	$\text{Pa}^{-1} \cdot \text{s}^{-1}$	Diffusion creep constant	U	J	Internal energy
H_L	$\text{J} \cdot \text{mol}^{-1}$	Diffusion creep enthalpy	p	Pa	Pressure
B_N	$\text{Pa}^{-n} \cdot \text{s}^{-1}$	Dislocation creep constant	θ	–	Volumetric strain
H_N	$\text{J} \cdot \text{mol}^{-1}$	Dislocation creep enthalpy	τ_{ij}	Pa	Deviatoric stress tensor
n	–	Dislocation creep exponent	$\dot{\epsilon}_{ij}$	s^{-1}	Deviatoric strain rate tensor
B_p	s^{-1}	Peierls creep constant	η_{eff}	$\text{Pa} \cdot \text{s}$	Effective viscosity

expressed as:

$$a_{II} = (a_{ij}a_{ij})^{1/2}. \quad (4.5)$$

Continuity equation

We include the effects of elastic compressibility and thermo-elasticity. In this case, the continuity equation can be conveniently coupled with the constitutive equation for hydrostatic pressure:

$$\frac{Dp}{Dt} = -K \left(\dot{\theta} - \alpha \frac{DT}{Dt} \right). \quad (4.6)$$

Here, K is the bulk modulus and α is the coefficient of thermal expansion. The presence of volumetric deformations evokes the following corrections of material density:

$$\rho = \rho_0 \left[1 - \alpha(T - T_0) + \frac{p}{K} \right], \quad (4.7)$$

where ρ_0 is the density at reference temperature and zero pressure and T_0 is the reference temperature. Note that Eq. (4.7) can be easily replaced by the equation of state applicable for high pressure and temperature for models involving deep portions of the mantle.

Additive decomposition and elasticity

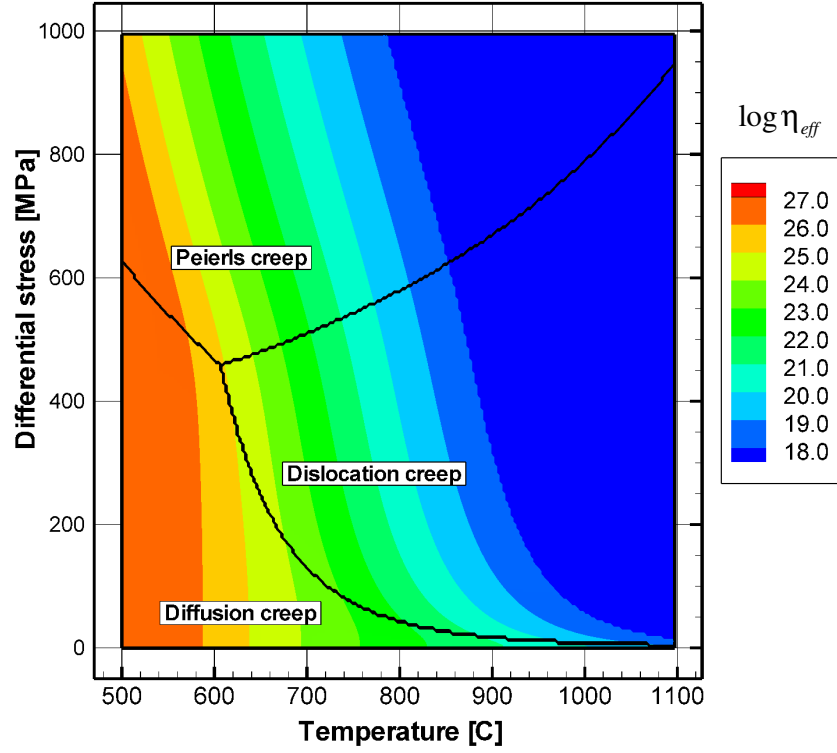


Figure 4.1 Logarithm of effective viscosity for dry olivine calculated using parameters from Kameyama et al. (1999). Temperature-stress domains in which each particular creep mechanism produces the largest strain rate are labeled. Black solid lines separate the domains. Viscosity is truncated to a reasonable range $10^{18} - 10^{27}$ [Pa·s].

Taking benefit from the smallness of *elastic* strains, we adopt conventional additive decomposition (e.g. Simo and Hughes, 2000) of total deviatoric strain rate. The elastic, viscous and plastic components, respectively, can be written as follows:

$$\dot{\epsilon}_{ij} = \dot{\epsilon}_{ij}^{el} + \dot{\epsilon}_{ij}^{vs} + \dot{\epsilon}_{ij}^{pl} = \frac{1}{2G} \hat{\tau}_{ij} + \frac{1}{2\eta_{eff}} \tau_{ij} + \dot{\gamma} \frac{\partial Q}{\partial \tau_{ij}}, \quad (4.8)$$

where G is the elastic shear modulus, $\hat{\tau}_{ij}$ is the objective stress rate (e.g. Bonet and Wood, 1997), η_{eff} is the effective creep viscosity, $\dot{\gamma}$ is the plastic multiplier, and Q is the plastic potential function (e.g. Simo & Hughes, 2000). Note that for numerical integration of elastic stresses over the finite time step, we use the incrementally objective scheme of Hughes and Winget (1980). For more details see the next section of this chapter.

Ductile creep

We adopt a detailed description of the ductile deformation component. The total viscous strain rate is additively decomposed into three temperature- and stress-dependent creep mechanisms, namely diffusion creep, dislocation creep and Peierls creep (see e.g. Kameyama et al., 1999). The corresponding effective creep viscosity is given by:

$$\eta_{eff} = \frac{1}{2} \tau_{II} (\dot{\epsilon}_L + \dot{\epsilon}_N + \dot{\epsilon}_P)^{-1}, \quad (4.9)$$

where τ_{II} is effective differential stress (see Eq. (4.5)), and $\dot{\epsilon}_L$, $\dot{\epsilon}_N$ and $\dot{\epsilon}_P$ are the effective scalar strain rates due to the diffusion, dislocation and Peierls mechanisms, respectively. Specific expressions for each strain rate can be written as follows:

$$\dot{\epsilon}_L = B_L \tau_{II} \exp\left(-\frac{H_L}{RT}\right), \quad (4.10)$$

$$\dot{\epsilon}_N = B_N (\tau_{II})^n \exp\left(-\frac{H_N}{RT}\right), \quad (4.11)$$

$$\dot{\epsilon}_P = B_P \exp\left[-\frac{H_P}{RT}(1-\beta)^2\right] \left(\frac{\tau_{II}}{\beta\tau_p}\right)^s, \quad (4.12)$$

where

$$s = 2\beta(1-\beta) \frac{H_P}{RT}. \quad (4.13)$$

In the above equations, B_L , B_N , B_P and H_L, H_N, H_P denote the creep parameter and activation enthalpy, respectively, of each correspondent mechanism, R is the gas constant, n is the power law exponent, τ_p is the Peierls stress, and $0 < \beta < 1$ is an adjustable approximation parameter. Note that in Eq. (4.12) we have adopted an asymptotic approximation of Peierls mechanism (see Kameyama *et al.* 1999), since the original equation is inappropriate for stresses below $O(10^2)$ MPa.

Here we assume that the creep parameter and activation enthalpy are constant for all creep mechanisms. In the general case, the creep parameter can depend on grain size (Karato *et al.*, 2001), and activation enthalpy may be sensitive to pressure (Regenauer-Lieb and Yuen, 2004). If necessary, these complexities can be easily introduced in the presented numerical formulation.

In Fig. 4.1, we plot the effective logarithmic viscosity versus temperature and differential stress, assuming parameters for dry olivine from Kameyama *et al.* (1999). Each creep mechanism dominates over the others (produces a higher strain rate) in different temperature-stress domains. In Fig. 4.1, these domains are separated by black solid lines and labeled. Note that for olivine, most of effective viscosity reduction from extremely high (10^{27} Pas) to extremely low (10^{18} Pas) values occurs within a relatively narrow temperature range between 500 and 1000 C for all stress levels.

Brittle failure

We describe brittle failure of rocks by the classical Mohr–Coulomb plasticity model (see e.g.

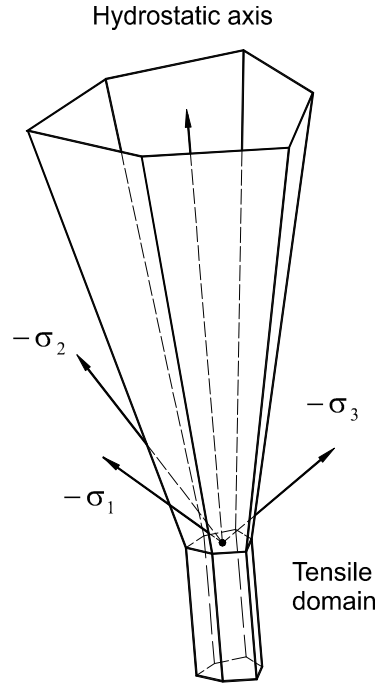


Figure 4.2 Schematic representation of the yield surface in principal stress space. In the tensile domain we use Tresca criterion.

Vermeer, 1990). The expression for the Mohr–Coulomb yield surface can be written as:

$$F = \frac{1}{2}(\sigma_{\max} - \sigma_{\min}) + \frac{1}{2}(\sigma_{\max} + \sigma_{\min}) \sin \varphi - c \cos \varphi \leq 0, \quad (4.14)$$

where σ_{\max} and σ_{\min} are the maximum and minimum principal stresses (negative in compression), φ is the material angle of friction, c is cohesion, $\frac{1}{2}(\sigma_{\max} - \sigma_{\min})$ is the maximum differential stress, and $\frac{1}{2}(\sigma_{\max} + \sigma_{\min})$ is the normal stress.

In the principal stress space, the Mohr-Coulomb yield surface can be represented as a hexagonal pyramid with a singular apex point in the tensile domain. Physically, this singularity means that the Mohr-Coulomb yield surface is inappropriate to model tensile failure of the rocks. A more adequate description in the context of lithospheric deformation would require coupling with the continuity equation to account for large plastic dilatation. Such sophisticated treatment of the tensile failure is not yet available. Therefore in this chapter we adopt a standard ad-hoc approach and approximate the Mohr-Coulomb yield surface in the tensile domain with the Tresca criterion which is given by:

$$F = \frac{1}{2}(\sigma_{\max} - \sigma_{\min}) - c \cos \varphi \leq 0. \quad (4.15)$$

The resulting composite yield surface is shown in Fig. 4.2.

Associative Mohr–Coulomb in 3D is essentially a multi-surface plasticity model which has so-called corner regions (Simo et al., 1988) in both yield surface and flow potential. Integration of such a model constitutes a particularly nontrivial task (Sloan and Booker, 1986; Larsson and Runesson, 1996; Borja et al., 2003). Another issue is related to the associativity of the flow rule,

which significantly overestimates plastic dilatation of the rocks (e.g. Alejano and Alonso, 2005). We resolve both these issues by adopting the purely deviatoric corner-free Prandtl-Reuss flow rule (e.g. Zienkiewicz and Taylor, 2000), which is non-associative with the Mohr–Coulomb yield surface. In this case the plastic potential function takes the following simple form:

$$Q = \tau_{II}. \quad (4.16)$$

The Prandtl-Reuss flow rule assumes complete plastic incompressibility (i.e. dilatation angle is zero), which is a suitable approximation for rocks in the large strain regime. At the same time, the integration of the plasticity model becomes considerably simpler due to a lack of corners in the plastic potential.

We approximate the degradation of the strength in faults and gouge zones by the strain softening model. The friction angle is assumed in the following product form:

$$\varphi = \varphi_0 D(\kappa), \quad (4.17)$$

where φ_0 is initial friction and $D(\kappa)$ is the function that controls degradation of the friction angle with progressive increase of the accumulated plastic strain. We assume $D(\kappa)$ in piecewise linear form. The accumulated plastic strain is given by:

$$\kappa = \int_t \left(\dot{\epsilon}_{ij}^{pl} \dot{\epsilon}_{ij}^{pl} \right)^{1/2} dt. \quad (4.18)$$

We note that the adopted plasticity model with strain softening does not incorporate any length scale for the strain localization (e.g. Muhlhaus and Aifantis, 1991). Thus the theoretical thickness of the shear band is zero. In the numerical model the thickness of the shear band is limited from below by the element size. Therefore in the context of an adaptively refined grid one must incorporate the length-scale or certain regularization (e.g. Belytschko and Tabbara, 1993) to prevent the element size from reducing to very small values. In this chapter we do not explicitly address this mesh-dependence issue, since at present we only consider uniform non-adaptive grids.

Heat flow

At present, we neglect latent heat effects due to phase change and assume internal energy in the form of linear function of temperature, i.e.:

$$U = C_p T, \quad (4.19)$$

where C_p is specific heat. We define the heat flux vector according to Fourier law as follows:

$$q_i = -\lambda \delta_{ij} \frac{\partial T}{\partial x_j}, \quad (4.20)$$

where $\lambda\delta_{ij}$ is the isotropic thermal conductivity tensor. The volumetric heat sources include radiogenic heat and heat produced by visco-plastic deformation, i.e.:

$$r = \rho A + \chi \tau_{ij} (\dot{\epsilon}_{ij}^{vs} + \dot{\epsilon}_{ij}^{pl}), \quad (4.21)$$

where A is radiogenic heat per unit mass and $0 \leq \chi \leq 1$ is the constant regulating degree of thermo-mechanical feedback. This coupling constant is an ad-hoc approach that effectively accounts for the processes which we do not include in our models, such as heat transport by the fluids, etc.

Boundary conditions

The conservation equations (4.1) and (4.2) must be complemented with initial and boundary conditions. This may include specified velocity, temperature, stress and heat flux. Additionally, in the context of lithospheric-scale modeling it becomes necessary to approximate the boundary conditions at the base of the lithosphere. In this chapter we use the classical Winkler foundation (see e.g. Fig. 5 in Regenauer-Lieb, 2006), which assumes zero viscous drag forces and takes into account buoyancy forces. The boundary stress tensor of the Winkler foundation can be parameterized as follows:

$$\bar{\sigma}_{ij} = -[p_0 + \rho_{ext} g (z - z_0)] \delta_{ij}, \quad (4.22)$$

where p_0 is pressure at the reference surface, ρ_{ext} is the density of the external material, and z and z_0 are the vertical coordinates of the bottom boundary and the reference surface, respectively. Typically, we use the initial position of the bottom boundary as the reference surface. The pressure at the reference surface is assigned as total weight divided by the depth of the model.

Other non-standard boundary conditions include more complicated effects such as erosion on the free surface and material in-flux and out-flux. These effects are considered in more detail in the next section.

4.3 Numerical algorithms

Spatial discretization

We build numerical models of lithospheric-scale deformation on the basis of the Finite Element Method (e.g. Hughes, 1987; Zienkiewicz and Taylor, 2000; Belytschko et al., 2000). Following the Galerkin procedure (e.g. Belytschko et al., 2000), we transform the momentum balance

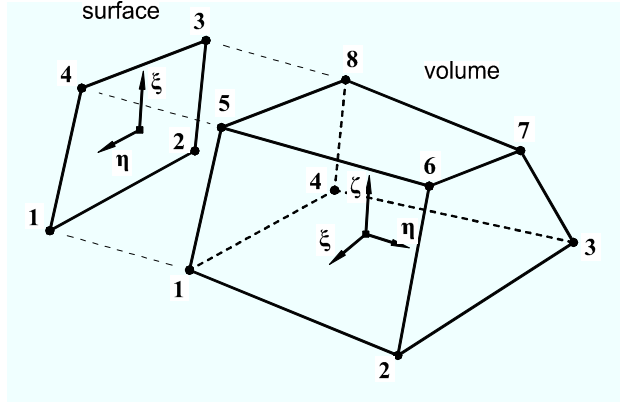


Figure 4.3 Finite element discretization of volume (hexahedrons) and surface (quadrilaterals). Shown are the local coordinate systems and local numbers of the nodes.

equation (4.1) into a corresponding nodal force residual:

$$f_{il} = \int_{\Omega} \frac{\partial N_I}{\partial x_j} \sigma_{ij} d\Omega - \int_{\Omega} N_I \rho g \hat{z}_i d\Omega - \int_{\Gamma} N_I \bar{\sigma}_{ij} n_j d\Gamma = 0. \quad (4.23)$$

Similarly, we approximate the energy balance equation (2) by a nodal power residual:

$$w_l = \int_{\Omega} N_I \rho \frac{DU}{Dt} d\Omega - \int_{\Omega} \frac{\partial N_I}{\partial x_i} q_i d\Omega - \int_{\Omega} N_I r d\Omega + \int_{\Gamma} N_I \bar{q}_i n_i d\Gamma = 0. \quad (4.24)$$

Here N denotes the nodal shape function (Zienkiewicz and Taylor, 2000), I is the nodal index, $\bar{\sigma}_{ij}$ and \bar{q}_i are surface stress and surface heat flux, respectively, n_i is the outward unit normal, and Ω and Γ stand for the volume and surface of the domain.

We employ hexahedral finite elements with linear interpolation functions (e.g. Zienkiewicz and Taylor, 2000) to approximate volume integrals in the above equations, and similar quadrilateral elements to approximate surface integrals (see Fig. 4.3). Element integrals are expressed in terms of parametric coordinates and evaluated over the unit cube or square using numerical integration (e.g. Belytschko et al., 2000). This stage is directly followed by assembly of the global residual equations (4.23) and (4.24) from element contributions (e.g. Hughes, 1987).

In each element, the shape function derivatives with respect to global coordinates are computed using standard coordinate transformation (summation over repeated nodal indices is implied):

$$\frac{\partial N_I}{\partial x_i} = J_{ij}^{-1} \frac{\partial N_I}{\partial \xi_j}, \quad J_{ij} = \frac{\partial N_I}{\partial \xi_i} x_{jI}, \quad (4.25)$$

where ξ_i are the local coordinates and J_{ij} denotes Jacobian matrix. The unit outward normal vector is evaluated according to:

$$n_i = J^{-1} \hat{n}_i, \quad J = (\hat{n}_i \hat{n}_i)^{1/2}, \quad (4.26)$$

where J stands for surface Jacobian and \hat{n}_i is the cross-product between two surface tangent

vectors, which can be expressed as follows:

$$\hat{n}_i = \varepsilon_{ijk} t_j^{(1)} t_k^{(2)}, \quad t_i^{(a)} = \frac{\partial N_I}{\partial \xi_a} x_{iI}, \quad a=1,2. \quad (4.27)$$

Here $t_i^{(a)}$ are the surface tangent vectors and ε_{ijk} is the Levi-Civita permutation symbol. To ensure a positive determinant of the Jacobian matrix and outward orientation of the unit normal vector, we perform appropriate control over the numbering of nodes in the element.

Locking and Hourglass

In nearly incompressible problems the finite element mesh is prone to locking, which manifests itself in severe underestimation of velocities/displacements (e.g. Belytschko et al., 2000). Locking is cured either by suitable under-integration of stress terms (Malkus and Hughes, 1978) or by using higher interpolation order for displacements (velocities) than for the stresses (e.g. Zienkiewicz and Taylor, 2000). Both approaches are essentially equivalent.

We suppress locking by evaluating constitutive equations only in one point per element (Flanagan and Belytschko, 1981). This approach implies that strains (strain rates) and other spatial gradients are evaluated using the following element-average derivatives of the shape functions:

$$b_{iI} = \frac{1}{V_E} \int_{\Omega} \frac{\partial N_I}{\partial x_i} d\Omega, \quad (4.28)$$

where V_E is the element volume. For element integrals, we use standard quadrature.

The hexahedral elements employed here violate the so-called LBB stability condition named after Ladyzhenskaya (1969), Babuska (1973) and Brezzi (1974), which leads to mesh artefacts known as ‘‘hourglass modes’’ (e.g. Flanagan and Belytschko, 1981) (see also Fig. 4a). Despite this, we use them together with careful monitoring of the stress and displacements fields, because the more stable quadratic element is computationally very expensive in 3D problems. With the same element resolution, quadratic interpolation requires roughly an order of magnitude more nodes (degrees of freedom) than the linear interpolation.

We have tested some of the anti-hourglass techniques available for hexahedral (quadrilateral) and tetrahedral (triangular) elements (see e.g. Flanagan and Belytschko, 1981; Liu et al., 1998; Bonet et al., 2001; Reese, 2003; Puso and Solberg, 2006). In certain problems, such as the Rayleigh-Taylor instability with large abrupt viscosity variations (see Fig. 4.4), none of stabilization technique we tested was able to prevent the hourglass modes. We have inferred that the critical factor is the relatively high confining pressure compared to deviatoric stress in the low viscosity domain.

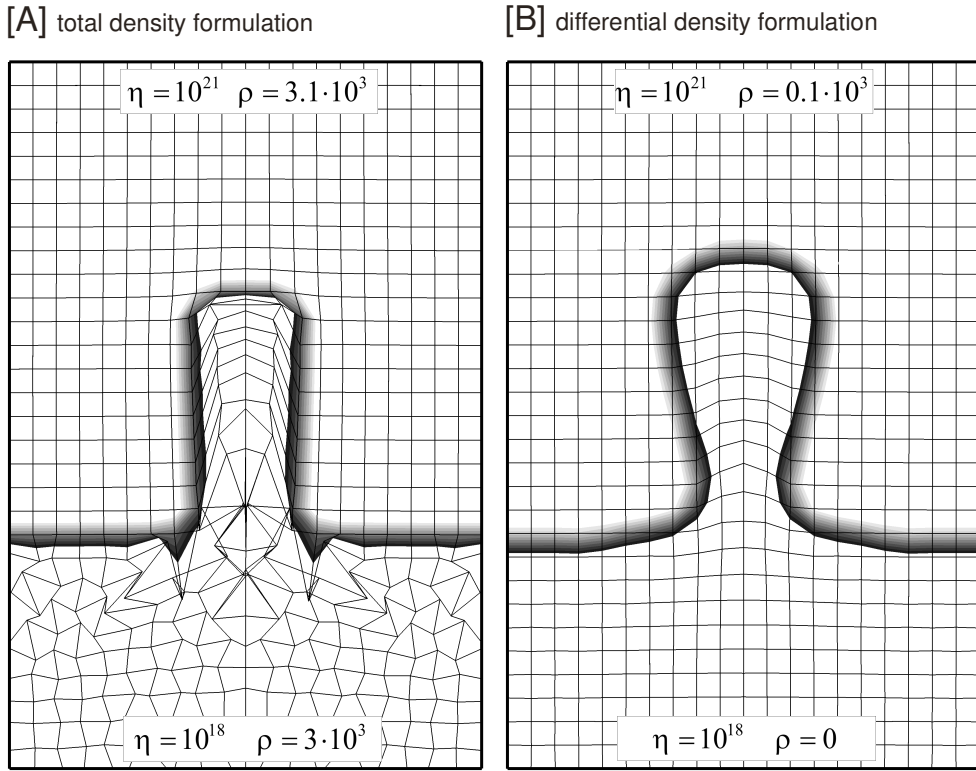


Figure 4.4 Grid distortion for a single time step of the Rayleigh-Taylor instability problem with large (3 orders of magnitude) abrupt viscosity variation. (a) total density formulation, (b) differential density formulation. Gradient curve indicates boundary between the layers. Shown are the viscosity and density of upper and lower layers.

The magnitude of the confining pressure can be reduced by replacing the material density with the differential density $\Delta\rho = \rho - \rho_{ref}$. Here ρ_{ref} denotes the arbitrary chosen constant reference density. With the differential density approach, the numerical solution produces the dynamic pressure Δp . To evaluate the constitutive equations properly, it becomes necessary to augment the dynamic pressure with a lithostatic component $\rho_{ref} g (z - z_{free})$. Here z and z_{free} denote the vertical coordinates of the integration point and the free surface, respectively. The density of the external material must also be set to the differential value $\Delta\rho_{ext} = \rho_{ext} - \rho_{ref}$. Similarly, the Winkler condition must be applied at the Lagrangian free surface with the differential density equal to negative reference density i.e. $\Delta\rho_{air} = -\rho_{ref}$.

According to our experience, this simple rearrangement of the computational scheme is sufficient to remove the hourglass modes even in the Rayleigh-Taylor instability problem mentioned above. Figs. 4a and 4b show the mesh distortion during a typical time step for the total and differential density formulations, respectively. The total density formulation exhibits uncontrollable oscillation while differential density formulation behaves stably.

We note, however, that despite the fact that the differential density technique can extend stability margins in certain problems, there is no guarantee that it will work in other problems.

Therefore, we suggest that this technique should be used only in conjunction with careful monitoring of the stress and displacements fields. Additionally, one can use spatial averaging over the neighboring nodes (e.g. Fulsack, 1995) to enhance the stability.

As an alternative to finite elements with bilinear interpolation for velocities and constant pressure, fully staggered finite difference discretization can be used (Harlow and Welsh, 1965), which appear to be non-oscillatory (for discussion, see Shih et al., 1989). Successful application of this discretization in geodynamic problems is presented in Gerya and Yuen (2007). However this approach, despite stability, is restricted to pure Eulerian formulations. In lithospheric scale modeling, we consider this restriction as unfavourable because it complicates tracking of the free surface.

Time discretization and primary variables

We use the backward Euler method as the primary time discretization algorithm for both the momentum and energy balance equations. This approach is referred to as implicit and first order accurate. The first order accuracy brings advantageous stability, which is missed by the second order accurate trapezoidal rule (Ortiz and Popov, 1985). In the adopted context, integration of all quantities becomes particularly simple, e.g. for velocities we can write:

$$\Delta u_i = v_i \Delta t, \quad (4.29)$$

where Δu_i is incremental displacement vector and $\Delta t = {}^{n+1}t - {}^n t$ is the time step. Here the left superscript indicates the time step index. In the remainder of this chapter we omit the index of the current time step for notational clarity, e.g. we simply write: $\Delta t = t - {}^n t$.

The simplicity of the integration scheme allows us to choose either velocities or *incremental* displacements as primary kinematical variables, irrespective of the employed rheology or kinematical formulation. Indeed, it is clear from the integration formula (4.29) that both approaches are essentially equivalent. For convenience we prefer incremental displacements. Note that all results presented below are readily extendable to the velocity formulation merely by simple scaling.

We do not separate pressure as an independent variable, since we use a compressible formulation. However, in typical problems the near-incompressible behaviour is recovered in the asthenosphere domain. We treat this additional difficulty during the linear solution stage by appropriate damping of the stiffness parameters (see subsection “linearization” for more details). For the heat balance equation we use temperature as the primary variable.

Kinematical formulation and solution scheme

We employ the Arbitrary Lagrangian-Eulerian kinematical formulation (Hirt et al., 1974) to account for material advection. A typical calculation step is subdivided into three major stages:

- (i) In the first stage, we solve the discretized residual equations (4.23) and (4.24) by the Newton Raphson iterative method (e.g. Belytschko et al., 2000). During iterations we conveniently treat advection terms implicitly by means of the Updated Lagrangian formulation (e.g. Bathe et al., 1975). The advantage is that Lagrangian treatment does not limit the time step by the amount of rigid body motion, in contrast with the Eulerian approach. In the Lagrangian context, only the amount of material straining poses an actual kinematic limit for the time step.
- (ii) In the second stage, we perform mesh adaptation (regridding) such that the new mesh fits the free surface, the moving (stretching) calculation window and simple material interfaces without overturns or self-intersections. The flexibility of the approach is that regridding can be done either in each step or after a certain number of steps, which is favourable in the case that advection is not the dominant type of nonlinearity in the system (e.g. the onset of plastic localization). Regridding is also the easiest way to implement boundary material fluxes (erosion, sedimentation, etc.)
- (iii) In the third and final stage, we perform consistent remapping of all solution variables onto the updated mesh. Our algorithm is based on the particle-in-cell approach which was initially developed by Harlow and Welsh (1965) and subsequently has been widely used in geodynamic applications (Moresi et al., 2003; Tackley and Xie, 2003; O'Neil et al., 2006; Gerya and Yuen, 2007).

We explain each calculation stage in more detail in the following sections. The overall computational flowchart is summarized in Table 4.2.

Hughes-Winget scheme

A constitutive model of advection dominated flow that includes memory effects (e.g. Fullsack, 1995) should be objective, i.e. it should ensure conservation of tensor quantities (e.g. Muhlhaus and Regenauer-Lieb, 2005). This can be formally achieved by using time-continuous (infinitesimal) objective stress rates (e.g. Bonet and Wood, 1997). However, since our practical interest is implicit large-step integration of the flow, we use the Hughes-Winget scheme (Hughes-Winget 1980) to ensure objectivity in an incremental, rather than an infinitesimal, sense.

The central point of the scheme is the following second order accurate approximation for an incremental displacement gradient:

Table 4.2

Overall computational flowchart

Newton-Raphson solution

1. Update stresses & heat fluxes in the elements
2. Assemble & solve linear systems $\delta \mathbf{u} = -\mathbf{K}_{(k)}^{-1} \mathbf{f}_{(k)}, \quad \delta \mathbf{T} = -\mathbf{E}_{(k)}^{-1} \mathbf{w}_{(k)}$
3. Update displacement and temperature $\Delta \mathbf{u}_{(k+1)} = \Delta \mathbf{u}_{(k)} + \delta \mathbf{u}, \quad \mathbf{T}_{(k+1)} = \mathbf{T}_{(k)} + \delta \mathbf{T}$
4. Update coordinates $\mathbf{x}_{(k+1)} = \mathbf{x}_{(k)} + \delta \mathbf{u}$
5. Check convergence $\|\delta \mathbf{u}\| < \varepsilon \|\Delta \mathbf{u}_{(k+1)}\|, \quad \|\delta \mathbf{T}\| < \varepsilon \|\mathbf{T}_{(k+1)}\|$
6. go to step (1) if necessary

Regridding and Remapping

1. Save history increment to markers
2. Advect markers by mesh
3. Adapt mesh to fit the free surface
4. Apply erosion boundary condition
5. Map markers onto adapted mesh
6. Count number of markers per cell
7. Insert/delete markers where necessary
8. Interpolate properties and history to mesh

$$h_{ij} = \frac{\partial \Delta u_i}{\partial x_j^{n+1/2}}, \quad {}^{n+1/2}x_j = {}^n x_j + \frac{1}{2} \Delta u_j, \quad (4.30)$$

where ${}^{n+1/2}x_j$ is the mid-point material configuration. In each element, we compute h_{ij} using the mid-point element-average derivatives of shape functions (see Eq. (4.28)) as follows:

$$h_{ij} = \Delta u_{il} \, {}^{n+1/2}b_{jl}. \quad (4.31)$$

We note that the mid-point scheme in Eq. (4.30) effectively eliminates the influence of rigid body motion on the strain components (Hughes and Winget, 1980). This property contrasts with pure Eulerian formulation, which is restricted to differentiation only with respect to the current configuration.

Deviatoric and volumetric strain *increments* to be used in constitutive equations are computed, respectively, as follows:

$$\boldsymbol{\varepsilon}_{ij} = \frac{1}{2}(h_{ij} + h_{ji}) - \frac{1}{3}h_{kk}\delta_{ij}, \quad \theta = h_{ii}. \quad (4.32)$$

Correspondent rate quantities are approximated by:

$$\dot{\boldsymbol{\varepsilon}}_{ij} = \frac{\boldsymbol{\varepsilon}_{ij}}{\Delta t}, \quad \dot{\theta} = \frac{\theta}{\Delta t}. \quad (4.33)$$

According to the Hughes-Winget scheme, rotation of stress from the previous time step is governed by the following orthogonal tensor (for more details see Hughes and Winget, 1980):

$$\mathbf{R}_{ij} = \left(\boldsymbol{\delta}_{ik} - \frac{1}{2} \boldsymbol{\omega}_{ik} \right)^{-1} \left(\boldsymbol{\delta}_{kj} + \frac{1}{2} \boldsymbol{\omega}_{kj} \right), \quad \boldsymbol{\omega}_{ij} = \frac{1}{2} (h_{ij} - h_{ji}). \quad (4.34)$$

The approximations (4.30) – (4.34), despite their simplicity, produce stable and accurate results even for large strain and rotation increments (e.g. Rashid, 1993).

Incremental constitutive equations

Incremental expansion of the continuity equation (4.6) gives the following updated pressure, if the reactional (phase transformation) volume changes are neglected:

$$p = {}^n p - K\theta + K\alpha(T - {}^n T). \quad (4.35)$$

To update the deviatoric stress, we use a two-step predictor-corrector procedure (e.g. Simo and Hughes, 2000). In the first step, the plastic multiplier in equation (4.8) is assumed to be zero and the correspondent trial visco-elastic stress is evaluated. In the second step, the Mohr-Coulomb yield surface is checked (Eq. (4.14)), and if violated, the trial stress is modified to remain on the yield surface. This gives rise to a nonzero plastic multiplier and corresponding plastic strain increment.

Combining the Hughes-Winget scheme with the analytical integration of equation (4.8) over the time step we obtain the following expression for trial deviatoric stress:

$$\tau_{ij}^{tr} = 2\eta_{CR}\dot{\epsilon}_{ij} + \alpha_{CR} R_{ik} {}^n \tau_{kl} R_{jl}, \quad (4.36)$$

where ${}^n \tau_{kl}$ is the Cauchy stress deviator from previous time step, R_{ij} is the incremental rotation tensor (see Eq. (4.34)), and η_{CR} and α_{CR} are the effective viscosity and relaxation ratio of visco-elastic creep, respectively, given by:

$$\eta_{CR} = \eta_{eff} (1 - \alpha_{CR}), \quad \alpha_{CR} = \exp\left(-\frac{\Delta t}{t_M}\right), \quad t_M = \frac{\eta_{eff}}{G}. \quad (4.37)$$

Here t_M stands for Maxwell time.

Equation (4.36) shows that the magnitude of Maxwell time does not pose the upper limit for the time step (see also Bailey, 2006). Indeed, in typical calculations we choose time steps equal to tens of thousands years. In the asthenosphere, where temperature is high enough to sufficiently decrease the viscosity, the following purely viscous limit is recovered:

$$\frac{\Delta t}{t_M} \gg 1, \quad \lim_{\alpha_{CR} \rightarrow 0} \tau_{ij}^{tr} = 2\eta_{eff}\dot{\epsilon}_{ij}. \quad (4.38)$$

In the upper crustal domain, on the other hand, the relatively low temperatures make the viscosity so large that Equation (4.36) degenerates into the corresponding elastic end-member:

$$\frac{\Delta t}{t_M} \ll 1, \quad \lim_{\alpha_{CR} \rightarrow 1} \tau_{ij}^{tr} = 2G\epsilon_{ij} + R_{ik} {}^n \tau_{kl} R_{jl}. \quad (4.39)$$

However, due to the presence of the Mohr-Coulomb stress limiter, most of deformation in the

upper crustal domain manifests itself in the form of plastic localization.

Apart from temperature dependence, the effective viscosity depends also on stress, and this introduces nonlinearity into the equation (4.36). This issue can be resolved, in principle, with good stress estimation from the previous time step. However, we have concluded from our experiments that this explicit approach severely restricts stable calculation steps due to stress oscillations. The alternative is, therefore, implicit treatment of nonlinearity by solving the nonlinear equation (4.36) for a trial deviatoric stress.

The simplest fixed point iteration cannot be used to this end, since in our case it may produce infinite loops without convergence. Hence, we have adopted the simple and reliable root-solving algorithm FZERO (Shampine and Watts, 1970), which is based on a combination of the bisection and secant methods.

Prior to solution we rewrite the tensorial equation (4.36) in scalar residual form by taking the difference between norms of its left- and right hand sides. The solution is achieved when the residual is reduced to a sufficiently small relative tolerance. Typically, very few iterations (around 5-7) are required by FZERO to converge.

For numerical stability reasons, we remove stress dependence in the effective viscosity at differential stress levels higher than $O(10^3)$ MPa, which is far higher than expected in typical problems. Finally, in the numerical simulations we truncate effective viscosity to the reasonable range $10^{18} - 10^{27}$ Pas.

After the visco-elastic stress predictor step, we check the yield surface to determine whether a plastic stress correction is required. In the following, we introduce basic elements of the procedure, whereas more details are outlined in Popov & Sobolev (2008).

Plastic stress correction starts with spectral decomposition of the trial deviatoric stress tensor by the optimized 3D Jacobi eigenvalue algorithm (Press et al., 2002):

$$\tau_{ij}^{tr} = \sum_{A=1}^3 \tau_A^{tr} m_{ij}^{(A)}. \quad (4.40)$$

Here τ_A^{tr} denotes the trial principal deviatoric stresses, $m_{ij}^{(A)}$ are the correspondent spectral directions. In the next stage, we evaluate the trial yield surface (Popov & Sobolev, 2008):

$$F^{tr} = \frac{1}{2}(\tau_{\max}^{tr} - \tau_{\min}^{tr}) + \frac{1}{2}(\tau_{\max}^{tr} + \tau_{\min}^{tr}) \sin \varphi - p \sin \varphi - c \cos \varphi. \quad (4.41)$$

For simplicity, the friction angle in the above equation is computed using the magnitude of accumulated plastic stain from previous time step, i.e. $\varphi = \varphi_0 D(^n \kappa)$. As soon as the trial deviatoric stress departs from the yield surface ($F^{tr} > 0$ condition is met), we undertake the following correction:

Table 4.3
Stress update in the integration point

1. Compute displacement gradient	$h_{ij} = \Delta u_{il}^{n+1/2} b_{jl}$
2. Evaluate strain increments	$\epsilon_{ij} = \frac{1}{2}(h_{ij} + h_{ji}) - \frac{1}{3} h_{kk} \delta_{ij}, \quad \theta = h_{ii}$
3. Update pressure	$p = {}^n p - K\theta + K\alpha(T - {}^n T)$
4. Find trial deviatoric stress by FZERO	$\tau_{ij}^{tr} = 2\eta_{CR} \dot{\epsilon}_{ij} + \alpha_{CR} R_{ik} {}^n \tau_{kl} R_{jl}$
5. Decompose trial deviatoric stress	$\tau_{ij}^{tr} = \sum_{A=1}^3 \tau_A^{tr} m_{ij}^{(A)}$
6. Check trial yield surface	$F^{tr} < 0 \Rightarrow Exit$
7. Compute plastic scaling ratio	$\alpha_{PL} = \frac{2(p \sin \varphi + c \cos \varphi)}{(1 + \sin \varphi) \tau_{max}^{tr} - (1 - \sin \varphi) \tau_{min}^{tr}}$
8. Update stress	$\tau_{ij} = \alpha_{PL} \tau_{ij}^{tr}$
9. Update plastic strain	$\kappa = {}^n \kappa + \frac{\Delta t}{2\eta_{CR}} (1 - \alpha_{PL}) \tau_{II}^{tr}$

$$\tau_{ij} = \alpha_{PL} \tau_{ij}^{tr}, \quad (4.42)$$

where $0 < \alpha_{PL} < 1$ is the plastic scaling ratio, which is given by:

$$\alpha_{PL} = \frac{2(p \sin \varphi + c \cos \varphi)}{(1 + \sin \varphi) \tau_{max}^{tr} - (1 - \sin \varphi) \tau_{min}^{tr}}. \quad (4.43)$$

For simplicity, we assume that both the effective viscosity of visco-elastic creep and the trial stress remain unaltered during plastic stress correction.

The magnitude of plastic strain rate is readily computed as follows:

$$\dot{\gamma} = \frac{1}{2\eta_{CR}} (1 - \alpha_{PL}) \tau_{II}^{tr}. \quad (4.44)$$

Integrating the above expression by the Backward Euler algorithm we obtain the following relation for the plastic strain increment:

$$\Delta \kappa = \frac{\Delta t}{2\eta_{CR}} (1 - \alpha_{PL}) \tau_{II}^{tr}, \quad (4.45)$$

which produces asymptotically exact results for both viscous and elastic cases. Corresponding limit expressions are given by:

$$\lim_{\alpha_{CR} \rightarrow 0} \frac{\Delta t}{2\eta_{CR}} = \frac{\Delta t}{2\eta_{eff}}, \quad \lim_{\alpha_{CR} \rightarrow 1} \frac{\Delta t}{2\eta_{CR}} = \frac{1}{2G}. \quad (4.46)$$

Accumulated plastic strain for the next time step is updated by:

$$\kappa = {}^n \kappa + \Delta \kappa. \quad (4.47)$$

For the Updated Lagrangian kinematical formulation, the material and partial time

derivatives are coincident. We can therefore approximate the rate of internal energy in each integration point by the following first order difference:

$$DU / Dt = C_p (T_I N_I - {}^n T_I N_I) / \Delta t. \quad (4.48)$$

For consistency with the Hughes-Winget scheme, we evaluate the spatial temperature gradient with respect to the mid-point material configuration (see Eq. (4.30)). Thus, the finite element approximation for the heat flux vector can be written as (see Eq. (4.28)):

$$q_i = -\lambda T_I {}^{n+1/2} b_{ii}. \quad (4.49)$$

Heat source due to dissipation of mechanical energy can be re-expressed as:

$$\tau_{ij} (\dot{\epsilon}_{ij}^{vs} + \dot{\epsilon}_{ij}^{pl}) = \tau_{II} \left(\frac{\tau_{II}}{2\eta_{eff}} + \dot{\gamma} \right). \quad (4.50)$$

The basic steps of the stress update procedure are summarized in Table 4.3. We note that this sequence is repeated every iteration of every time step in each integration point.

Nonlinear solver

At each time step, we solve the coupled system of discretized residual equations (4.23) and (4.24) by the full Newton-Raphson iterative method (e.g. Belytschko et al., 2000). We neglect coupling terms in the Jacobian matrix. Therefore, at each iteration we solve the following two systems of linear equations:

$$\delta \mathbf{u} = -\mathbf{K}_{(k)}^{-1} \mathbf{f}_{(k)}, \quad \Delta \mathbf{u}_{(k+1)} = \Delta \mathbf{u}_{(k)} + \delta \mathbf{u}. \quad (4.51)$$

$$\delta \mathbf{T} = -\mathbf{E}_{(k)}^{-1} \mathbf{w}_{(k)}, \quad \mathbf{T}_{(k+1)} = \mathbf{T}_{(k)} + \delta \mathbf{T}. \quad (4.52)$$

Here, the right subscript (k) denotes iteration index, \mathbf{f} is the out-of-balance nodal force vector (Eq. (4.23)), \mathbf{w} is the out-of-balance nodal power vector (Eq. (4.24)), $\Delta \mathbf{u}$ is the incremental displacement (total velocity) vector, \mathbf{T} is the nodal temperature vector, $\mathbf{K} = \partial \mathbf{f} / \Delta \mathbf{u}$ is the mechanical tangent matrix, $\mathbf{E} = \partial \mathbf{w} / \partial \mathbf{T}$ is the thermal tangent matrix, and $\delta \mathbf{u}$ and $\delta \mathbf{T}$ are iterative correction vectors for incremental displacement and temperature, respectively. At each iteration, we update nodal coordinates similarly to incremental displacements: $\mathbf{x}_{(k+1)} = \mathbf{x}_{(k)} + \delta \mathbf{u}$, while all the element integrals are evaluated over the latest updated coordinates (Updated Lagrangian formulation).

We achieve a coupled thermo-mechanical solution by:

- (i) Calculating effective viscosity and density using the latest updated temperature,
- (ii) Calculating mechanical dissipation using the latest available stress and strain rate,
- (iii) Simultaneously solving the force balance and energy balance equations.

In the context of the Newton-Raphson method, it is necessary to carefully select the initial

guess, since it can significantly reduce computation time. At the first time step we take zero displacements and an initial temperature distribution. For each subsequent time step we use the previous converged solution.

Constrained degrees of freedom are removed from the tangent matrices. To facilitate convergence *only* at the first iteration of the first time step, we add corrections related to removed degrees of freedom to the right hand side vectors.

We employ the so-called line search procedure (see e.g. Press et al., 2002; Crisfield, 1983) to stabilize iterative solution of the mechanical equation. The displacement update formula with line search is modified to the following form:

$$\Delta \mathbf{u}_{(k+1)} = \Delta \mathbf{u}_{(k)} + \alpha \delta \mathbf{u}. \quad (4.53)$$

Here $0 < \alpha < 1$ is the damping parameter, which is chosen to satisfy the following condition:

$$\|\mathbf{f}_{(k+1)}\| < \|\mathbf{f}_{(k)}\|. \quad (4.54)$$

We bisect the time step and restart the overall equilibrium iteration in case the above criterion cannot be fulfilled. This technique forms the core of our adaptive time-stepping algorithm. Initially we specify the uniform time step; the algorithm then performs bisections whenever convergence problems are experienced. When iteration becomes stable again, the algorithm tries to increase the time step. However it keeps in memory the last converged solution, to restore stable iteration in the case of unsuccessful time step increases. To further enhance convergence in the strain localization problems, more sophisticated time-stepping methods may be used such as the method of subplane control functions (Geers, 1999).

The Newton-Raphson iteration is terminated as soon as the following criteria are satisfied:

$$\|\delta \mathbf{u}\| < \varepsilon \|\Delta \mathbf{u}_{(k+1)}\| \text{ and } \|\delta \mathbf{T}\| < \varepsilon \|\mathbf{T}_{(k+1)}\|, \quad (4.55)$$

where ε is a small tolerance, typically $\varepsilon = 10^{-3}$.

Linearization and linear solver

We assume the following simplified approximations for the tangent matrices (Popov & Sobolev, 2008):

$$K_{ijkl} = \int_{\Omega} \frac{\partial N_I}{\partial x_j} C_{ijkl} \frac{\partial N_J}{\partial x_l} d\Omega + \int_{\Gamma} N_I (\rho_{ext} g n_z \hat{z}_i \hat{z}_k) N_J d\Gamma, \quad (4.56)$$

$$E_{IJ} = \int_{\Omega} N_I \frac{\rho C_p}{\Delta t} N_J d\Omega + \int_{\Omega} \frac{\partial N_I}{\partial x_i} \lambda \frac{\partial N_J}{\partial x_i} d\Omega. \quad (4.57)$$

Here, C_{ijkl} is the material tangent operator (Simo and Taylor, 1985), n_z is the vertical component of outward normal vector, and \hat{z}_i are the components of the unit vector of the vertical

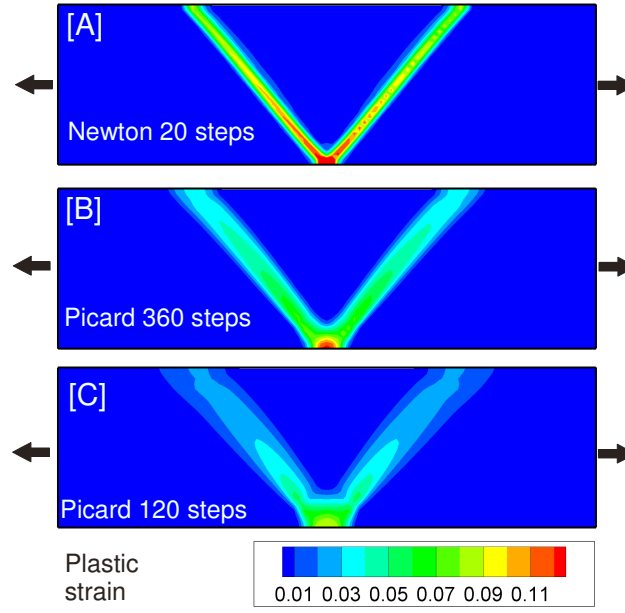


Figure 4.5 Contour plots of accumulated plastic strain for the shear band initiation problem (see section 4.4); (a) calculated by Newton-Raphson solver in 20 increments, (b) calculated by Picard solver in 360 increments, (c) calculated by Picard solver in 120 increments. The same amount of extension is applied in all models. In case (a) the shear band has minimum possible width constrained by the cell size. In cases (b) and (c) the shear band is significantly wider and more diffuse slowly decreasing its width with increasing number of time increments.

axis (downward positive). The first term in the Eq. (4.56) corresponds to linearization of the internal stress term, while the second term is caused by the Winkler boundary condition.

Depending on type of the flow (visco-elastic or elasto-visco-plastic), the material tangent operator takes the following simple form:

$$C_{ijkl} = K \delta_{ij} \delta_{kl} + \begin{cases} 2G_{CR} I_{ijkl}^D & \text{if } F^{tr} \leq 0 \\ 2G_{PL} I_{ijkl}^D + \tau_{ij}^{tr} \hat{g}_{kl} & \text{if } F^{tr} > 0 \end{cases} \quad (4.58)$$

Here $I_{ijkl}^D = \frac{1}{2}(\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}) - \frac{1}{3} \delta_{ij} \delta_{kl}$ is the fourth order unit deviatoric tensor, $G_{CR} = \eta_{CR} / \Delta t$ is the effective visco-elastic shear modulus, and $G_{PL} = \alpha_{PL} G_{CR}$ is the effective elasto-visco-plastic shear modulus. The non-dimensional tensor \hat{g}_{kl} is given in Popov & Sobolev (2008).

We infer from our numerical experiments that anisotropic stress terms in the tangent operator (Eq. (4.58)) have little influence on the convergence rate of the Newton-Raphson iteration. Therefore, we suggest that they can be omitted whenever necessary. With this assumption, the effective elasto-visco-plastic shear modulus (G_{PL}) becomes essentially equivalent to the effective viscosity (see e.g. Eq. 37 of Moresi et al. 2003, or Fig. 4 of Fullsack, 1995).

Both these papers employ the fixed point method or direct (Picard) iteration (e.g. Zienkiewicz and Taylor, 2000) as a nonlinear solver. Even omitting differences in the tangent operator, there are two substantial differences between the Newton-Raphson and Picard

methods:

- (i) The Picard method has only external forces and stresses from the previous time step in the right-hand-side, while the Newton-Raphson method has complete out-of-balance forces.
- (ii) The Picard method operates on total velocity (incremental displacement), while the Newton-Raphson method operates on iterative corrections.

We have tried both Picard and Newton-Raphson methods. Fig. 4.5 shows a comparison between the two solvers for the shear band initiation benchmark problem, discussed in the next section. We calculated an identical setup by both methods with a gradually increasing number of increments in which we applied a constant amount of extension. Fig. 4.5 shows the severe sensitivity of the Picard method to increment size. At the same time, the Newton-Raphson method produces accurate results even for comparatively large increments. Other authors (e.g. Muhlhaus and Regenauer-Lieb, 2005) have also pointed out superiority of the Newton-Raphson method over the Picard method.

At the current development stage, we use the Preconditioned Conjugate Gradient (PCG) iterative method (Hestenes, Stiefel, 1952) for both the mechanical and thermal linear systems (Eq. (4.51) & Eq. (4.52)). For the mechanical system we adopt a symmetric version of the Incomplete LU factorization with Threshold (ILUT) preconditioner (Saad, 1994), equipped with near-optimal nested dissection ordering (George, 1973). The thermal system is preconditioned by the inverse of diagonal of thermal tangent matrix.

The above algorithm, in general, is not numerically scalable (Farhat et al., 2000), i.e. the number of iterations grows nonlinearly with increasing size of the linear system. At present the algorithm can only run in sequential mode. However, it works well with the roughly 10^5 grid nodes in typical 3D regional-scale models on ordinary PCs with 2 GB RAM.

Note that the linear solver may fail in the low-viscosity domains (e.g. asthenosphere), where effective the visco-elastic modulus becomes many orders of magnitude lower than the bulk modulus. We treat this issue by supplementing the mechanical linear solver with the fictitious compressibility technique (Zienkiewicz and Taylor, 2000). The fictitious bulk modulus is made at maximum two orders of magnitude larger than the shear modulus in the integration point. The algorithm then iteratively removes spurious volumetric strains until the original continuity equation is satisfied.

Regridding and Remapping

After the Newton Raphson step, we resolve the advection terms of the flow. This step is implemented by means of regridding and remapping. The computational sequence of adopted

procedure is summarized in the second part of the Table 4.2.

We employ a specific type of the Arbitrary Lagrangian Eulerian kinematical formulation (Hirt et al., 1997). It is characterized by using a structured Cartesian grid whose nodes are only allowed to move in the vertical direction. With this flexibility, we track the free surface or certain material interfaces inside the model domain. Using the same mechanism we implement erosion and sedimentation processes on the free surface. The bottom surface of the grid is equipped with the Winkler boundary condition.

Lateral boundaries of the grid may be either fixed in space or move parallel to the coordinate axis with constant or variable velocity. This is a particularly useful feature, since it creates a flexible computation window which can track some portion of space, or even stretch or contract laterally. When the length of the mesh in a certain direction changes by the size of the original element, we add (or remove) an entire vertical slice of the elements. Our solvers are designed to treat a variable number of elements during computation.

To track material properties and solution history we use the Particle-in-Cell method (Harlow and Welsh, 1965). Each marker has an individual history, including coordinates, material number, pressure, deviatoric stress tensor, accumulated plastic strain, total displacements, and temperature. During the Newton-Raphson step, the history accumulates in cell centers. When markers are advected by mesh, all markers that belong to the cell obtain the same *increment* of history. When markers are mapped back to the mesh, both material properties and history are averaged among all markers in the updated cell. This incremental approach minimizes spurious numerical diffusion.

We advect markers in the displacement field by consistent interpolation from the mesh. Since the mesh remains rectangular in horizontal plain, the mapping of markers to updated cell positions becomes particularly simple. Certain difficulties exist in regions where a population of markers becomes too dense or too sparse (Moresi et al., 2003). We treat this issue by inserting or deleting markers in cells where the total number of markers falls beyond the limits. We treat material fluxes on the boundaries with the same methodology. When a marker leaves the domain, we simply delete it. When a marker is generated in the boundary cell it acquires material properties and solution history from the closest neighbours. We suggest, therefore, that material extends beyond all boundaries of the domain except the free surface.

We note that the remapping procedure inevitably produces perturbations in the global force balance. We treat these perturbations as initial residuals to be removed on the next time step. This explicit approach to a certain extent reduces the accuracy of the numerical solution (stresses, displacements). Nevertheless, it is common in geodynamics to use similar techniques with the argument that the time step can be reduced if higher accuracy in the time integration is

required (e.g. Moresi et al. 2003). In section 4.4 we demonstrate that the remapping procedure does not significantly affect the accuracy of the velocity field, which is estimated analytically, even in the benchmark with hundreds of remapping cycles.

4.4 Benchmarks and examples

In this section, we present five simplified problems to verify computer implementation and applicability of the adopted physical models and numerical techniques. The following problems are included:

- (i) Bending of an elastic plate
- (ii) Sinking of a rigid cylinder in a viscous fluid
- (iii) Initiation of shear bands in the brittle crust
- (iv) Triaxial compression test
- (v) Lithospheric transpressional deformation

The first two problems are designed to test the code's ability to separately handle elastic and viscous rheological mechanisms in 2D. In the first problem, the 2D formulation (plane strain) is chosen to facilitate comparison with the analytical solution, while in the second problem, the choice is based on the resolution issue. With the sequential solver and without mesh adaptation, the solution of the free-falling Stokes sphere problem with the same resolution would require too much computation time. The next two problems demonstrate quasi-brittle elasto-plastic strain localization in 2D and 3D, respectively. The final problem deals with the complete elasto-viscoplastic rheology in a typical 3D lithospheric-scale setup. Since the plain strain formulation is not currently available in SLIM3D, we discretize entire 2D setups by a single layer of hexahedral elements with zero velocity in the out-of-plane direction. Wherever necessary we introduce and track markers in the mid-surfaces of the elements.

Bending of an elastic plate

We consider bending of an elastic plate which has infinite length in one of the lateral dimensions. The setup of the problem is depicted in Fig. 4.6, and the geometry, loading and elastic constants of the plate are summarized in Table 4.4. Similar to a cantilever beam, the plate has one end fixed and the other end free. It has slightly negative buoyancy relative to surrounding material, which allows it to sink and bend. Since we are interested in modeling elastic deformation of the plate, we exclude the surrounding viscous material. Calculations are done in the Lagrangian mode without remeshing.

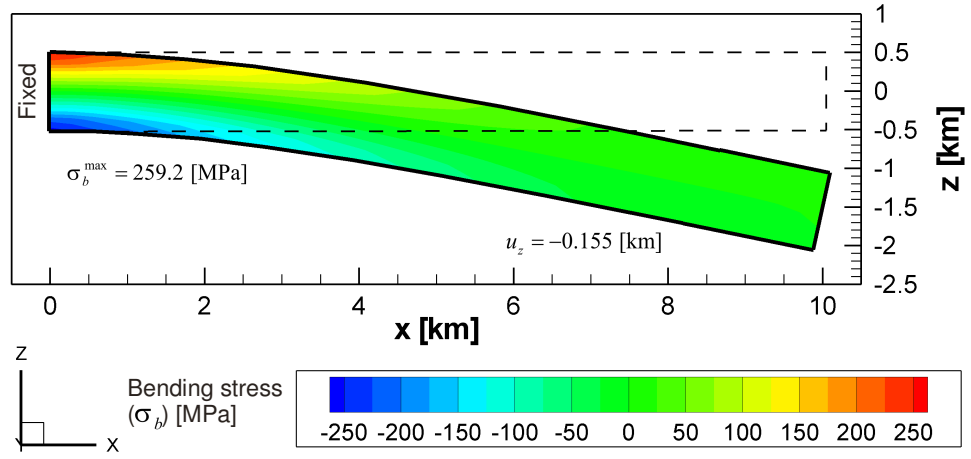


Figure 4.6 Initial (dashed line) and deformed (solid line) shape of the elastic plate. Vertical displacements are exaggerated by a factor of 10. Contour plot shows distribution of normal bending stress. Shown are calculated values for maximum bending stress and vertical displacement.

In this benchmark, we compare the maximum vertical deflection of the plate and the maximum bending stress with the analytical solution of the Euler-Bernoulli beam equation. Application of this equation requires displacements and thickness to be small compared with length of the plate. We fulfil this requirement by appropriately selecting parameters of the setup. Assuming validity of the Euler-Bernoulli approximation, the maximum deflection of plate can be expressed in terms of parameters in Table 4.4 as (Popov & Sobolev, 2008):

$$w = \frac{3\Delta\rho gl^4}{2Eh^2}. \quad (4.59)$$

For comparison with the numerical solution, we calculate analytical bending stresses using the following relation (Popov & Sobolev, 2008):

$$\sigma_b = \pm \frac{3\Delta\rho gl^2 (h - \Delta h)}{h^2}, \quad (4.60)$$

Table 4.4
Plate benchmark

Parameter	Meaning	Value	Dimension
$\Delta\rho$	Differential density	100	$\text{kg} \cdot \text{m}^{-3}$
E	Young's modulus	990	kbar
$l \times h$	Size of the domain	10×1	km
$n_l \times n_h$	Discretization	35×7	elements
w_{theor}	Theoretical deflection	0.151	km
w_{calc}	Calculated deflection	0.155	km
err_w	Rel. deflection error	2.6	%
σ_{theor}	Theoretical stress	257.2	MPa
σ_{calc}	Calculated stress	259.2	MPa
err_σ	Rel. stress error	0.8	%
σ_{cycl}	Cyclic stress residual	0.1	%

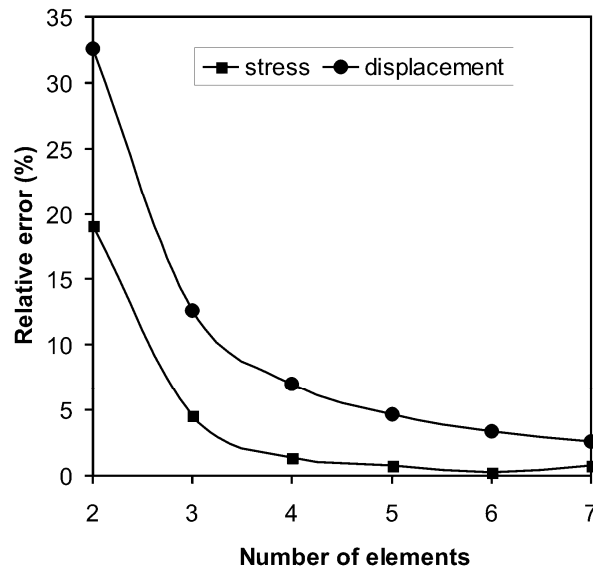


Figure 4.7 Relative error in approximation of stresses (rectangles) and vertical displacements (circles) versus number of elements across the elastic plate thickness.

where Δh is the element size along the height of the beam.

We apply and remove the gravity load multiple times in this model. Each time, when the load is removed, we allow the plate to restore its original shape and simultaneously release stored elastic stresses. During the solution, we monitor spurious residual stresses, which tend to accumulate during multiple load cycles. This deficiency is related to the rate formulation of elasticity adopted in SLIM3D (see Eq. (4.8)). Unlike exact large strain elasticity models (e.g. Bonet and Wood, 1997), this formulation produces nonzero dissipation of elastic energy along the closed deformation path (e.g. Belytschko et al., 2000). The applicability of the model to rocks is based on the fact that elastic strains remain small, while elastic stresses constantly dissipate either by plastic or viscous rheological mechanisms. In this benchmark, we numerically assess spurious dissipation of elastic energy.

Fig. 4.6 shows the shape of the deformed plate (exaggerated by factor of 10) along with a contour plot of normal bending stresses. Key results of the benchmark are presented in Table 4.4. We investigate influence of discretization on the accuracy of the solution using different numbers of elements along the plate thickness starting with just two elements (since we have a single integration point per element). In all cases, the 1:2 (thickness vs. length) element aspect ratio is maintained within the vertical plane. Fig. 4.7 shows the relative error in the approximation of stresses and displacements versus number of elements along the plate thickness. Note that discretization converges relatively fast, providing acceptable accuracy (around 5.0% error) with just 4-5 elements across the plate. The results in Table 4.4 correspond to discretization by 7 elements across the plate. Further increasing the number of elements did not yield a substantial increase in the accuracy in neither the vertical deflection (2.6% error) nor

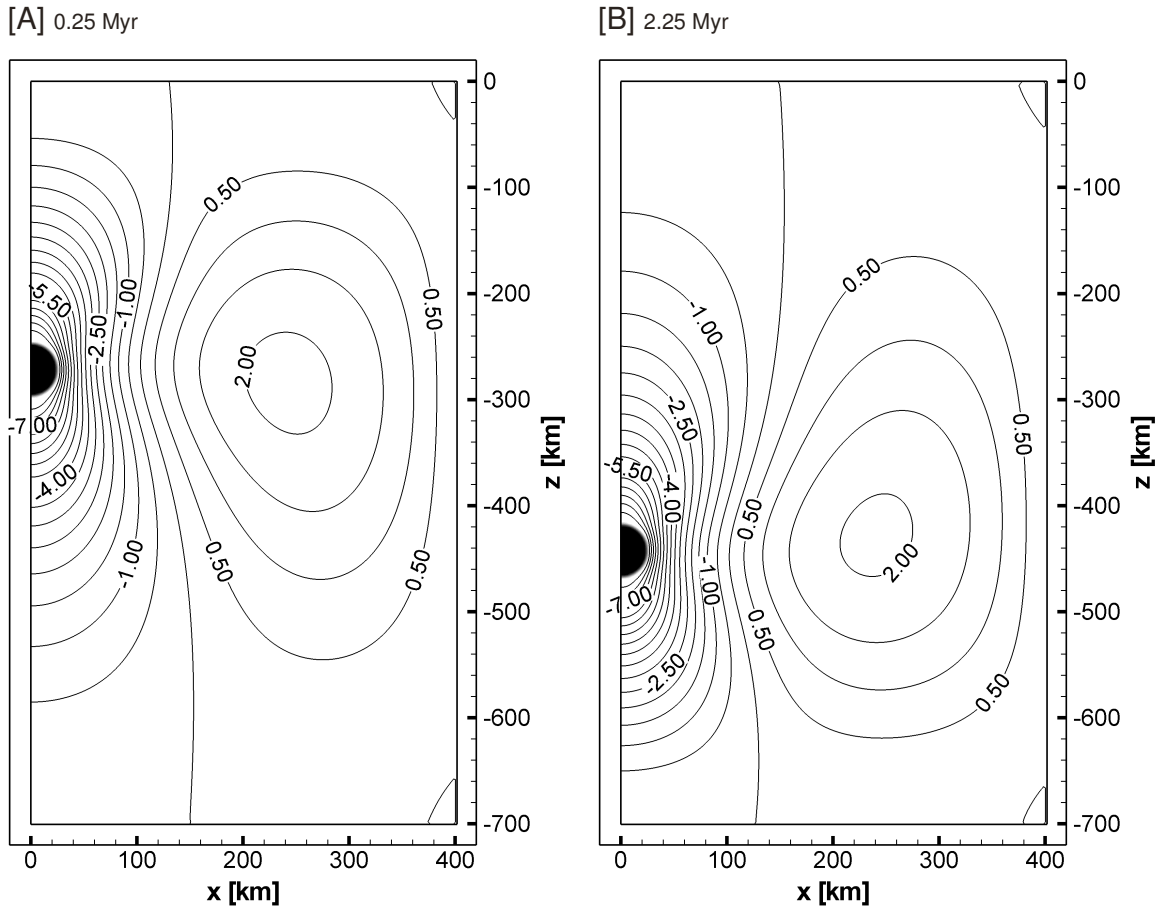


Figure 4.8 Isolines of vertical velocity for the cylinder problem. (a) after 0.25 Myr, (b) after 2.25 Myr. Black circles indicate positions of the particles which represent the cylinder.

in the bending stresses (0.8% error). We attribute this fact partially to inaccuracies of the approximate analytical solution and partially to the presence of shear locking (e.g. Belytschko et al., 2000) of the employed elements. After 10 cycles of cyclic loading, the residual stress remains about 0.1% of peak bending stress (0.3 MPa vs. 300.0 MPa). Thus, we conclude that the stress integration algorithm has sufficient accuracy for practical applications.

Sinking of a rigid cylinder in a viscous fluid

The purpose of this benchmark is twofold: (i) we test the ability to simulate purely viscous flow with large abrupt changes of the viscosity; (ii) we simultaneously assess the quality of the remeshing algorithm (treatment of markers and numerical diffusion issue). Setup of the problem includes one half of a rectangular domain which is 800 km wide and 700 km deep (see Fig. 4.8) occupied by viscous fluid. We impose sticking between the domain walls and the fluid. On the left side of the domain we impose a symmetry condition. An infinite cylinder with 25 km radius is initially placed at 250 km depth. Due to negative buoyancy the cylinder sinks into the fluid. We maintain relative rigidity of the cylinder by a 4 orders of magnitude viscosity contrast with the surrounding fluid. Parameters of the setup are summarized in Table 4.5. Numerical solution

Table 4.5
Cylinder benchmark

Parameter	Meaning	Value	Dimension
$\Delta\rho$	Differential density	100	$\text{kg} \cdot \text{m}^{-3}$
r	Radius of cylinder	25	km
$l \times h$	Size of the domain	400×700	km
$n_l \times n_h$	Discretization	200×350	elements
η_f	Viscosity of fluid	10^{20}	$\text{Pa} \cdot \text{s}$
η_c	Viscosity of cylinder	10^{24}	$\text{Pa} \cdot \text{s}$
v_{theor}	Theoretical velocity	8.47	$\text{cm} \cdot \text{yr}^{-1}$
v_{calc}	Calculated velocity	8.61	$\text{cm} \cdot \text{yr}^{-1}$

is obtained with a time step 10^4 years.

In this benchmark, we compare analytical and numerical estimation of the cylinder velocity. The analytical solution for resisting force due to motion of the cylinder in a reservoir of finite size can be found in Slezkin (1955). By equating resistance force with Archimedes force one can obtain the following expression for the velocity of the cylinder (Popov & Sobolev, 2008):

$$v_c = \frac{1}{4} \left(\ln k - \frac{k^2 - 1}{k^2 + 1} \right) \frac{r^2 \Delta\rho g}{\eta_f}, \quad k = \frac{b}{r}, \quad (4.61)$$

where η_f is fluid viscosity, r is the radius of cylinder and b is the characteristic distance between the cylinder and rigid wall of the domain. In our setup this distance is equal to 375 km.

We plot the distribution of vertical velocity together with positions of cylinder markers at two different times in Fig.4.8. Note that through the multiple remeshing steps, cylinder shape shows completely no sign of numerical diffusion. The cylinder remains perfectly circular and undisturbed through the course of computation. We therefore conclude an acceptable quality of adopted remeshing algorithm. The average velocity of the cylinder constitutes around 8.5 cm per year. The maximum relative difference between the analytical and numerical estimations is less than 1.7% (see Table 4.5). Error in the numerical solution for the velocity field remains relatively low, despite hundreds of remapping cycles and a large time step (10^4 years). Increase of viscosity contrast up to 7 orders of magnitude did not result in any substantial changes in the numerical solution. We conclude that the adopted numerical formulation allows handling of purely viscous flow with large abrupt variations of the viscosity.

Initiation of shear bands in the brittle crust

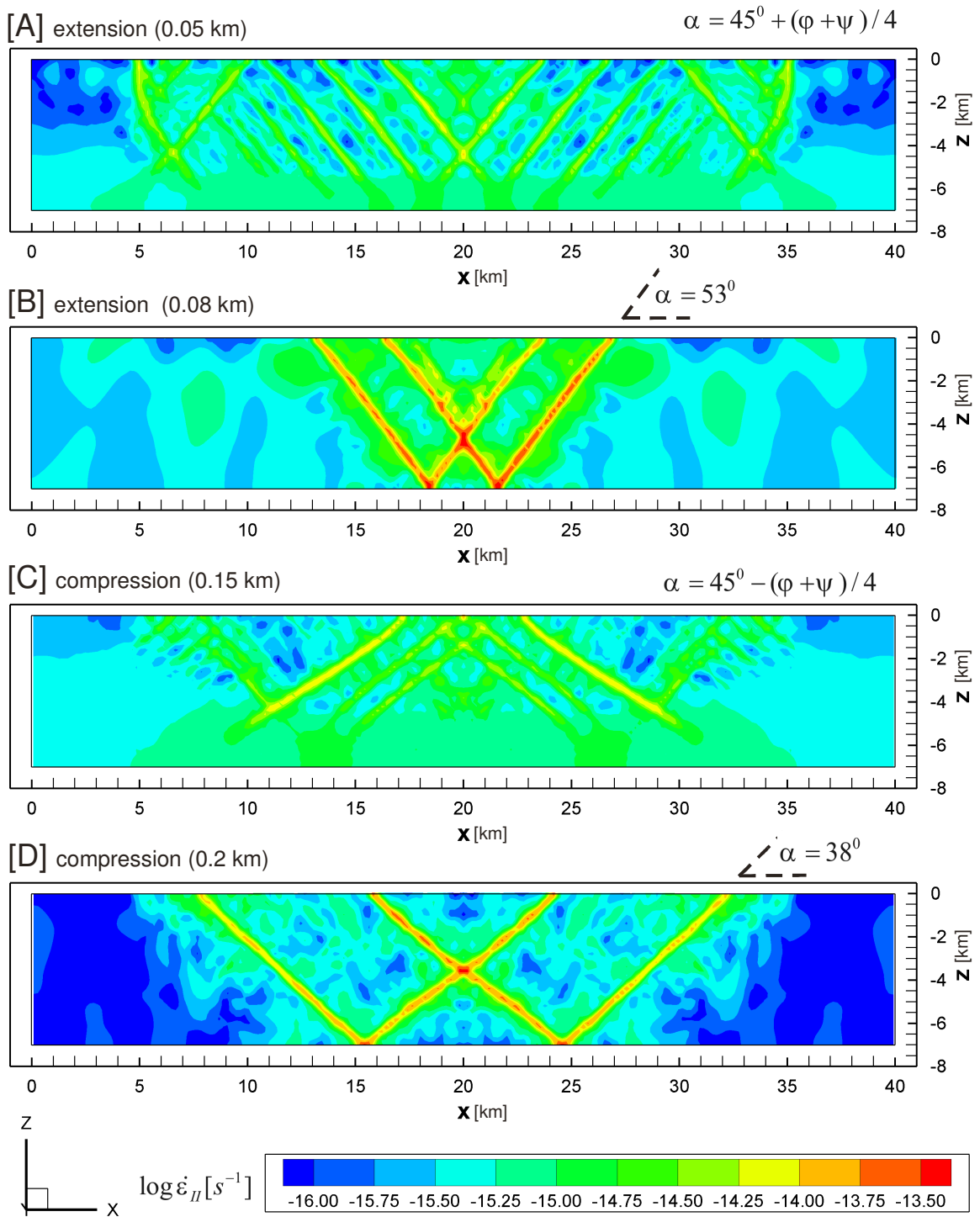


Figure 4.9 Contour plot of effective logarithmic strain rate for the shear band initiation problem. (a) 0.05 km of extension, (b) 0.08 km of extension, (c) 0.15 km of compression (d) 0.2 km of compression. Dashed lines indicate expected orientation of shear bands according to Arthur et al. (1977) (expression shown in the figure). Also shown are the values of Arthur angles between the shear bands and horizontal axis.

Formation of zones of localized shear (shear bands) during planar deformation has been the subject of extensive research (see e.g. Arthur et al., 1977; Muhlhaus and Vardoulakis, 1987; Vermeer, 1990; Borja and Aydin, 2004; Buitter et al., 2006; Kaus and Podladchikov, 2006). One of the most important questions concerns preferred orientation of the shear bands. Three basic

Table 4.6
Shear bands benchmark

Parameter	Meaning	Value	Dimension
ρ	Density	3000	$\text{kg} \cdot \text{m}^{-3}$
K	Bulk modulus	630	kbar
G	Shear modulus	400	kbar
φ	Friction angle	30	($^{\circ}$)
ψ	Dilatation angle	0	($^{\circ}$)
c	Cohesion	20	MPa
$l \times h$	Size of the domain	40×7	km
$n_l \times n_h$	Discretization	200×35	elements
α_{ext}	Inclination angle (extension)	53	($^{\circ}$)
α_{comp}	Inclination angle (compression)	38	($^{\circ}$)

models are generally discussed (Coulomb, 1773; Roscoe, 1970; and Arthur et al., 1977). For a pressure-sensitive material with friction angle φ and dilatation angle ψ , these models predict the following angles between the shear band and direction of minor principal stress:

$$\begin{aligned}
 \alpha &= 45^{\circ} + \varphi/2 && \text{(Coulomb),} \\
 \alpha &= 45^{\circ} + \psi/2 && \text{(Roscoe),} \\
 \alpha &= 45^{\circ} + (\varphi + \psi)/4 && \text{(Arthur).}
 \end{aligned}
 \tag{4.62}$$

The theoretical analysis for the onset of shear banding presented by Vermeer (1990) also yields the Roscoe and Coulomb angles as limits for possible shear-band orientation. However, due to elastic unloading the shear band can, in principle, assume any orientation within this range.

In this benchmark, we tested the capability to model initiation of shear bands and compare inclination angles of modeled shear bands with analytical estimations and numerical solutions obtained by other codes. Since only the onset of shear bands is considered, calculations are done in the small strain regime. The large strain plastic localization in 3D is additionally demonstrated in the next subsection.

The model setup includes a crustal domain which is 40 km wide and 7 km thick (see Fig. 4.9). On the top boundary we impose a free surface, while the bottom boundary has zero normal velocity and tangential free-slip. We consider two loading cases. On the lateral boundaries, we prescribe 1 cm per year normal velocity such that it causes either shortening or extension of the domain. We assume an elasto-plastic rheology typical of mafic crust (see Table 4.6). Since we employ the purely deviatoric Prandtl-Reuss flow rule, the dilatation angle is zero in our formulation.

We measure inclination angles of the shear bands from the horizontal axis. In the case of compression the orientation of the minor principal stress is vertical. Therefore the angle with respect to horizontal axis should be computed with a minus sign; for example, the Coulomb

angle should be $\alpha = 45^\circ - \varphi/2$. In the case of extension, the minor principal stress is horizontal; therefore sign should be plus i.e. $\alpha = 45^\circ + \varphi/2$.

Fig. 4.9 shows the contour plot of effective strain rate for both extension (a, b) and compression (c, d) cases at different stages of the process. The shear bands start to propagate from the surface where strength is minimal. We note that it takes significantly less extension than shortening for shear bands to cut the entire domain. The initial pattern of shear bands is highly chaotic. Some of them have inclination angles close to Coulomb angles ($30^\circ, 60^\circ$) at this stage. At later stages, only a few shear bands survive. They begin to accommodate entire deformation caused by shortening / extension such that other shear bands become completely inactive. The mature post-failure shear bands precisely match the Arthur angles ($38^\circ, 53^\circ$). In both cases (compression or tension), shear bands at the Roscoe angles (45°) did not occur in the model.

Our results are in good agreement with results of the other codes. For example, Poliakov and Herrmann (1994) solved a similar problem using a non-associative Mohr-Coulomb rheology and explicit FLAC technique. In their solution, the shear bands formed spontaneously and assumed various inclination angles ranging between the Roscoe and Coulomb limits. The results obtained with I2ELVIS (Gerya and Yuen, 2007) for the numerical sandbox experiment (Buiter et al., 2006) revealed that the majority of shear bands in the numerical solution tend to assume the Arthur angle. We conclude that our code is able to effectively simulate origination and propagation of localized shear zones.

Triaxial compression test

Elasto-plastic strain localization in the large strain regime is further demonstrated by numerical simulation of the triaxial compression test. We consider a laterally unconfined parallelepiped specimen of size $1 \times 1 \times 3$ km discretized into $20 \times 20 \times 60$ hexahedral elements, subjected to kinematical axial compressional loading (Fig. 4.10a and Fig. 4.11a). Material parameters are assumed from the previous benchmark (see Table 4.6). We specify two different types of heterogeneity: either small random seeds uniformly distributed over the specimen (Fig. 4.10a) or an oblique gash located near the edge as shown in Fig. 4.11a. In both cases, localization is induced by relatively low cohesion of the seeds with respect to surrounding material (down to 20% of nominal).

The type of heterogeneity significantly impacts the pattern of deformation observed in the model. Thus, the random seeding induces formation of multiple shear bands as shown in Fig. 4.10b – Fig.4.10d. Despite that the specimen exhibits fragmentation into small pieces, one can

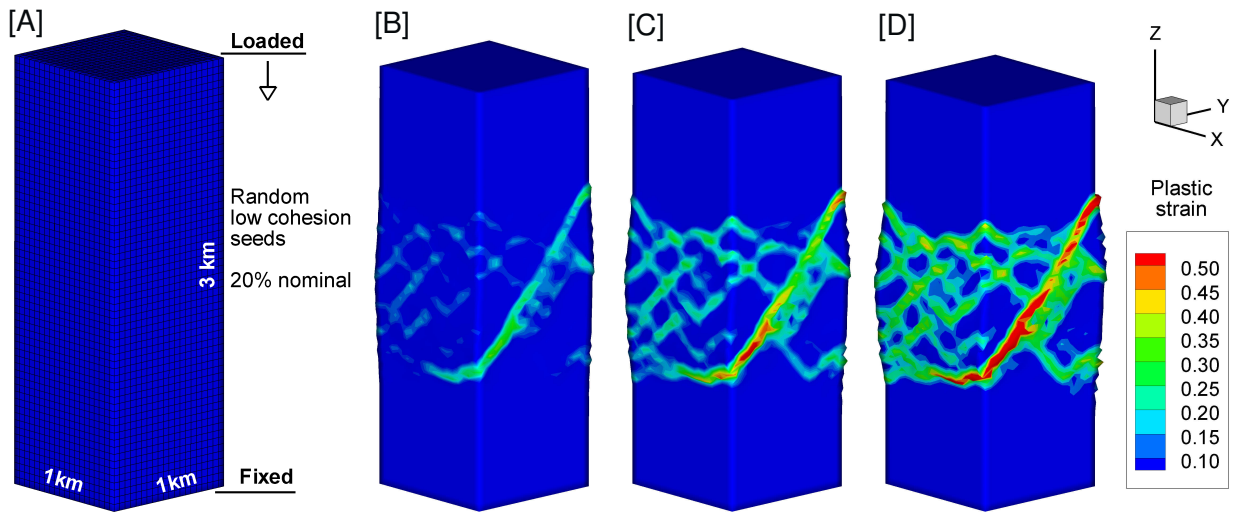


Figure 4.10 Setup (a) and results (b – d) of the triaxial compression numerical test with uniform random seeding (see text). Cases b – d correspond to 2.5%, 4.2%, 5.8% of axial deformation, respectively. Magnitude of accumulated plastic strain is shown in colors.

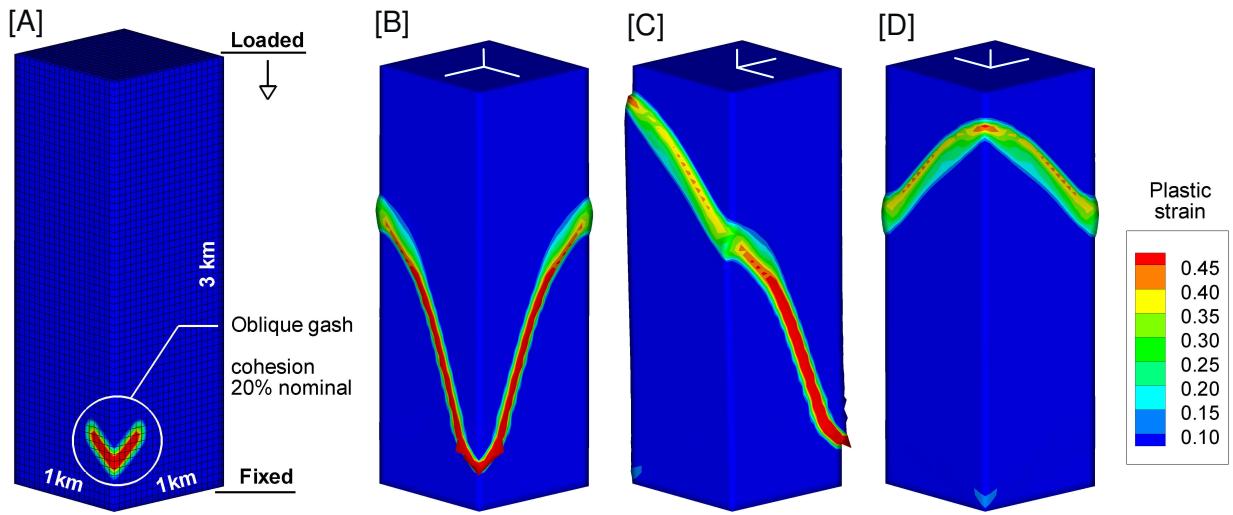


Figure 4.11 Setup (a) and results (b – d) of the triaxial compression numerical test with oblique gash seeding (see text). Cases b – d differ by amount of specimen rotation around vertical axis (0°, 90° and 180°, respectively) and all correspond to 2.7% of axial deformation. White lines on the top face show coordinate axis for rotational reference. Magnitude of accumulated plastic strain is shown in colors.

easily distinguish major failure plane, which gradually evolves though the course of loading. In contrast, the gash-type seeding immediately produces a complex curvilinear failure plane, which separates the entire specimen into two parts (Fig. 4.11b – Fig.4.11d). The majority of the shear band inclination angles also coincide with the Arthur angles. Results of the triaxial compression test support the conclusions of previous benchmark for the 3D case and the large strain regime.

Lithospheric transpressional deformation

In this example, we present a typical 3D lithospheric-scale setup which includes effects of elastovisco-plastic rheology with a temperature controlled brittle-ductile transition. The model

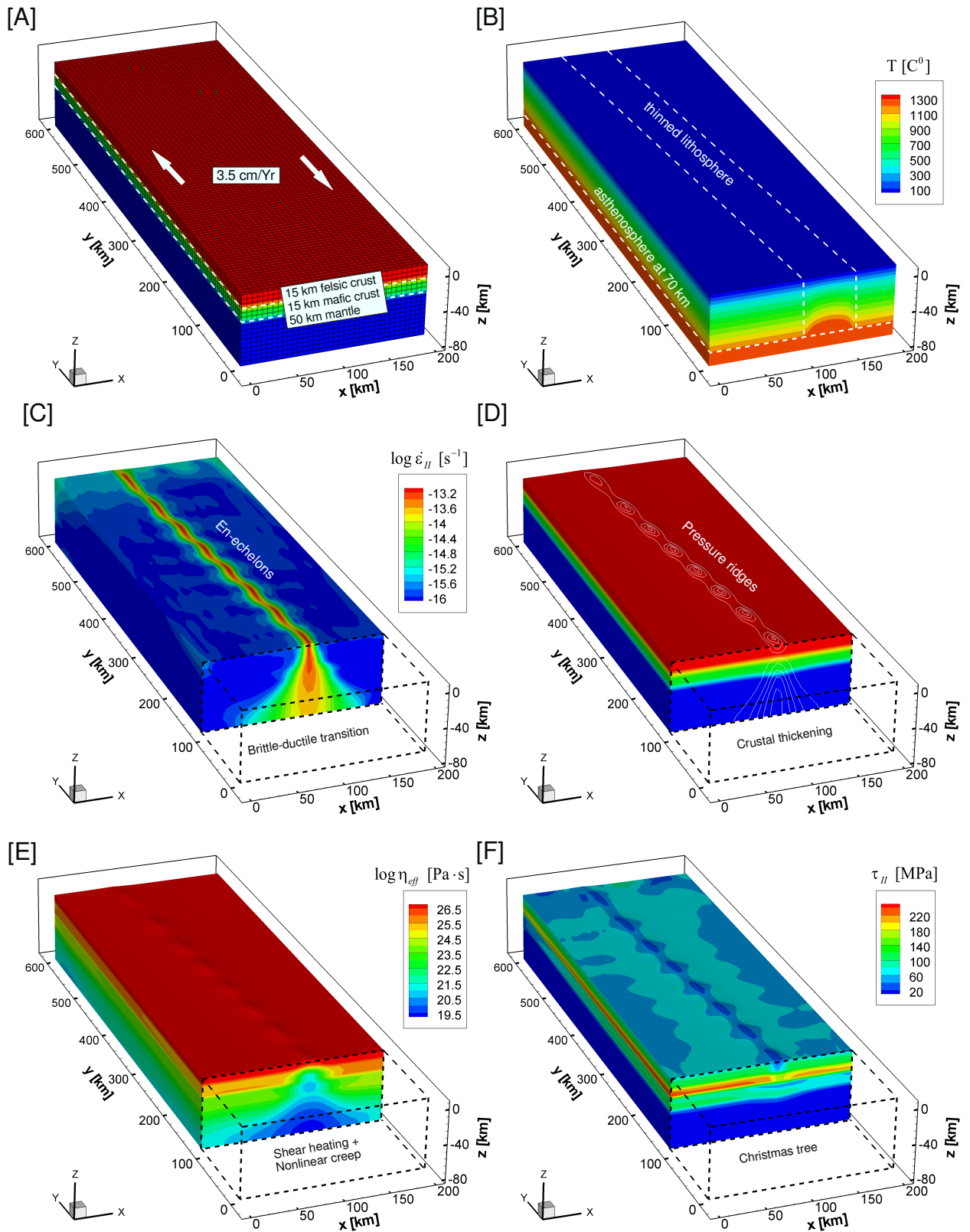


Figure 4.12 Lithospheric transpressional deformation problem. (a) setup and discretization, (b) initial thermal model, (c) distribution of effective strain rate on the surface and in the cross-section after 3Myr , (d) material phases (colors) and vertical displacements (white lines), (e) distribution of effective logarithmic viscosity (f) effective stress.

consists of a lithospheric domain which has lateral dimensions 200 by 600 km and depth 80 km, as shown in Fig. 4.12a. The domain is discretized into $36 \times 108 \times 16$ hexahedral elements (68561

Table 4.7
Rheological and thermal parameters used in transpression problem

Parameter	Dimension	Felsic Upper Crust	Mafic Lower Crust	Mantle peridotite
ρ	$\text{g} \cdot \text{cm}^{-3}$	2.7	3.0	3.3
K	kbar	550	630	1220
G	kbar	360	400	740
$\log B_N$	$\text{Pa}^{-n} \cdot \text{s}^{-1}$	-28.0	-21.05	-16.3
H_N	$\text{kJ} \cdot \text{mol}^{-1}$	223	445	535
n	—	4.0	4.2	3.5
φ	(°)	30	30	30
c	MPa	20	20	20
α	K^{-1}	$2.7 \cdot 10^{-5}$	$2.7 \cdot 10^{-5}$	$3.0 \cdot 10^{-5}$
C_p	$\text{kJ} \cdot \text{kg}^{-1} \cdot \text{K}^{-1}$	1.2	1.2	1.2
λ	$\text{W} \cdot \text{m}^{-1} \cdot \text{K}^{-1}$	2.5	2.5	3.3
A	$\text{nW} \cdot \text{kg}^{-1}$	1.0	0.1	0.0

grid points). Compositional heterogeneity includes equally thick (15 km) felsic upper crust and mafic lower crust and a 50 km thick peridotite mantle. Rheological and thermal parameters used in the modeling are presented in Table 4.7. At elongated vertical boundaries we impose zero normal velocities as well as 3.5 cm/year of right-lateral strike-slip velocity. The upper and lower boundaries are the free surface and the Winkler support boundary, respectively. At the short vertical boundaries we impose lithostatic pressure along with a free-slip condition.

The initial temperature distribution is shown in Fig.4.12b. We use different temperature gradients in the upper crust, and the lower crust with lithospheric mantle to account for radiogenic heat. In the asthenosphere, we assume a temperature of 1300°C. The lithosphere-asthenosphere boundary is initially placed at 70 km depth. To initiate the state of transpression in the model, we use a zone of slightly thinned lithosphere (10 km of thinning), which is oriented obliquely to the strike-slip direction.

Fig.4.12c shows a contour plot of the effective strain rate on the free surface and in a typical cross-section after 3 Myr of evolution and 105 km of strike-slip. The far-field strike-slip deformation tends to localize in a narrow zone located above the initially thinned lithosphere where strength is minimal. This mechanism leads to the self-consistent formation of the boundary between two lithospheric plates. The structure of the plate boundary changes with depth from a very narrow shear zone in the brittle upper crust to progressively more diffuse and wide shear zone in the ductile lower crust and upper mantle. The brittle upper crust tends to form regular secondary structures (en-echelon faults). Fig. 4.12d shows a contour plot of material phases and isolines of vertical displacements. The transpressional character of strike-slip deformation leads to surface uplift on pressure ridges related to en-echelon structures as well as

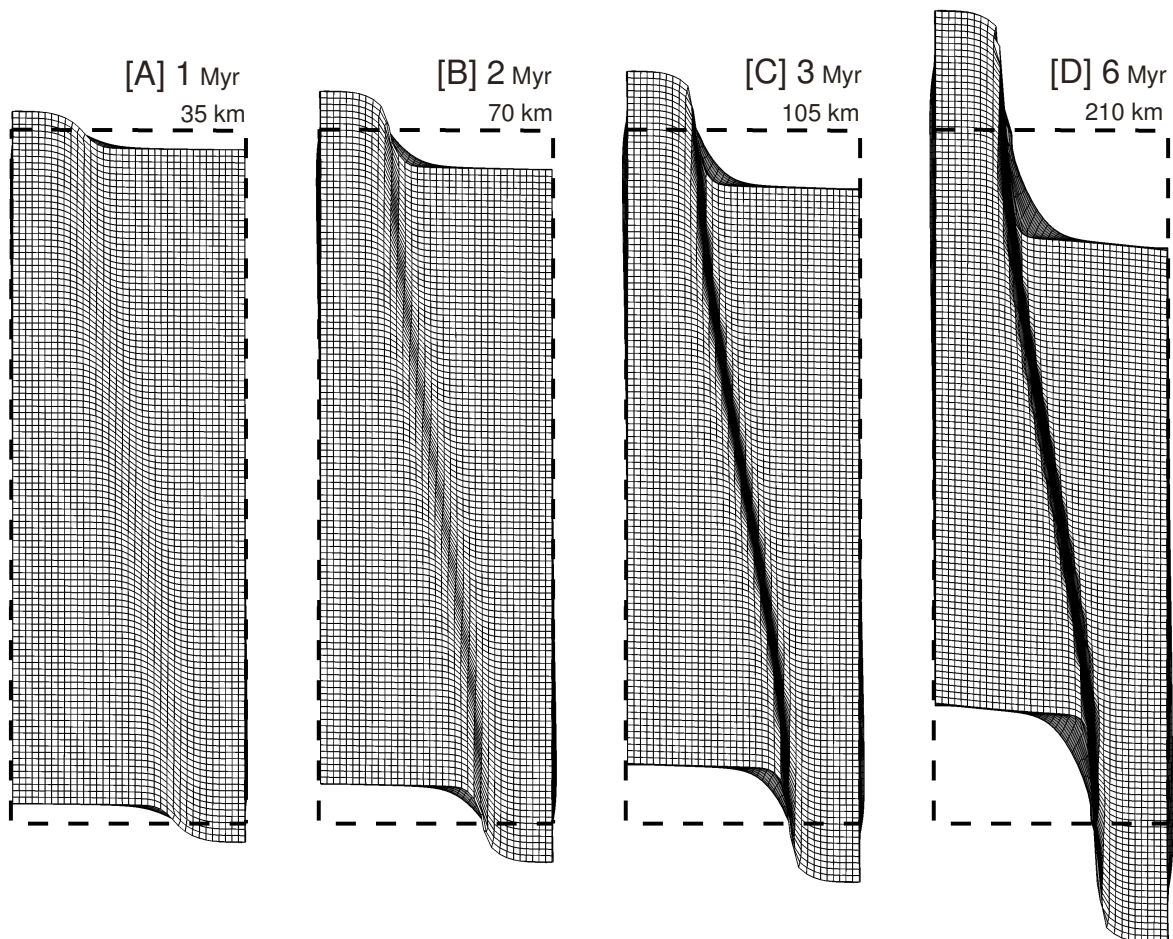


Figure 4.13 Evolution of the Lagrangian mesh embedded in the material for the lithospheric transpressional deformation problem. View from the top of the model. Thick dashed lines show the lateral boundaries of calculation domain shown in Fig. 4.12. Note that material fluxes on the boundaries of the domain are nonzero.

corresponding crustal thickening along the entire plate boundary.

Fig. 4.12e shows the influence of temperature and stress on the effective viscosity distribution. The overall viscosity variation in this model is about 7 orders of magnitude, with most of this variation caused by gradual temperature increase with depth and compositional heterogeneities. Among the local effects, we note the effects of shear heating and nonlinear creep which come into play across the entire plate boundary. Relative stress peaks associated with both the upper-lower crust boundary and the Moho, commonly referred to as “christmas-trees” (e.g. Brace and Kohlstedt 1980), can be recognized in Fig. 4.12f. Note that due to strain softening, the strength of major faults becomes low, which is also visible from the stress distribution on the surface (Fig. 4.12f).

Fig. 4.13 show a 6 Myr evolution of the Lagrangian mesh embedded in the material. We note that material fluxes on the boundary of computational box are nonzero. Incoming material acquires the properties and solution history from the closest material inside the box. The magnitude of the strike-slip displacement is large (210 km) and comparable to the characteristic size of the model. The transpressional plate boundary accumulates significant shear strains

during the evolution. Summarizing these results we suggest that our code, in principle, can be applied for 3D modeling of geological scale evolution of lithospheric plate boundaries.

4.5 Summary and discussion

Outlook

Currently, the code SLIM3D already includes many techniques necessary for modeling deformation processes on lithospheric and geologic scales. Namely, we have implemented coupled thermo-mechanical solutions, realistic elasto-visco-plastic rheology, and a kinematic framework that is able to handle large deformations. However, in this section we discuss the techniques which are still missing in the code. We also point out the directions of the future development.

The major disadvantage is that code is still severely restricted in size of the model. This fact is explained mostly by the following: (i) the code is still sequential (ii) the code utilizes a linear solver which is not numerically scalable (Farhat et al., 2000). To remove these disadvantages, we plan to extend our code in the following mainstream directions of linear solver technology:

- (i) Direct parallel multifrontal solvers (Gupta et al., 1997; Amestoy et al., 2000)
- (ii) Multigrid methods with adaptation for plasticity (Adams, 2000; Ekevid et al., 2004)
- (iii) Dual-primal domain decomposition methods (Farhat et al., 2000)

At the same time, we are planning to parallelize the entire finite element routines using the MPI package.

Deformation processes in the lithosphere occur on highly variable spatial scales. To sufficiently resolve the material length scale in lithospheric models, the element size should be of the order of hundreds of meters (Regenauer-Lieb and Yuen, 2003). Therefore, a successful code for lithospheric-scale modeling should allow variable spatial resolution. In SLIM3D we are presently restricted to a structured non-adaptive grid. A possible way to introduce adaptivity is to switch to unstructured tetrahedral mesh generators (e.g. Rassineux, 1998) or octree-based methods (e.g. Shephard et al., 1991; Braun et al., 2008). These will be investigated in the future.

Apart from advancements in numerical methods, we are planning to further develop physical models of rocks introducing advanced damage rheological models and incorporating metamorphic reactions and melting as well as coupled porous flow. Furthermore, we plan to introduce coupling between deformation processes at regional and global spatial scales as well as at seismic-cycle and geological time scales.

Summary remarks

This chapter presents a new tool (SLIM3D) for three-dimensional lithospheric-scale modeling. The code includes a coupled thermo-mechanical treatment of deformation process and allows complicated elasto-visco-plastic rheology with diffusion, dislocation and Peierls creep and Mohr-Coulomb plasticity. The code incorporates an Arbitrary Lagrangian Eulerian formulation with free surface and Winkler boundary conditions.

We have developed practical implementation of elasto-visco-plastic rheology and shown that Maxwell time does not pose an upper limit for the integration time step in the numerical scheme.

We have applied the full Newton-Raphson method for solution of discrete balance equations and demonstrated its superiority over the Picard method, which is widely used in geodynamics.

For plasticity model we have obtained simplified expressions for the tangent operator. Anisotropic stress terms in tangent operator have little influence on the convergence rate of the Newton-Raphson iteration. Therefore they can be omitted whenever necessary. With this assumption, the difference between effective viscosity and tangent operator concepts vanishes.

With linear sequential solvers, which are not numerically scalable, we are able to treat problems with up to hundred thousand of grid points on ordinary PCs. It is expected that capabilities of the method can be drastically expanded by implementing numerically scalable parallel solvers (e.g. multigrid) as well as an adaptive mesh.

Chapter 5

Three-dimensional thermo-mechanical models

As it was mentioned in Chapter 3, the extended two-dimensional models failed to reproduce the basic features of the San Andreas Fault System. The two-dimensional idealization and the lack of along-strike variations of the velocities, temperature and stresses is a serious drawback that substantially limits the realism of the simulation. This chapter briefly describes the formulation and results of the more sophisticated three-dimensional models designed to diminish previous limitations, and to reproduce complicated tectonic effects substantially better. In particular, the model (i) resolves transpression on the curved strike-slip fault, (ii) contains idealization of the subducting Gorda slab and material flow in the slab window, and (iii) allows variation of the entire fields in three dimensions without restrictions. Additionally, the model can be used to estimate the influence of the stalled microplates in the fossil subduction zone on the long-term deformation of the overlying continent. Technical details of the Finite Element code SLIM3D used for the modeling are thoroughly described in Chapter 4.

5.1 Setup of the models

The framework of the long-term models presented in this chapter is based on the reconstructions of the Neogene tectonic history of south-western United States of Atwater and Stock (1998). The position and orientation of the model calculation domain with respect to the reconstructed plate configuration is shown in Fig. 5.1. The model time starts at about 18 Ma, shortly after the completion of the Monterey microplate capture (Nicholson et al., 1994). Selection of the spatial location of the calculation domain allows considering interaction between major tectonic elements in the region. North-western part includes the portion of subducting Juan de Fuca (presently Gorda) plate. Western part is occupied by the Pacific plate. On the eastern boundary there is a Great Valley block. In certain models we also consider the influence of the Salinian crustal heterogeneity. Finally, the southern part contains the captured Monterey microplate.

Model setup

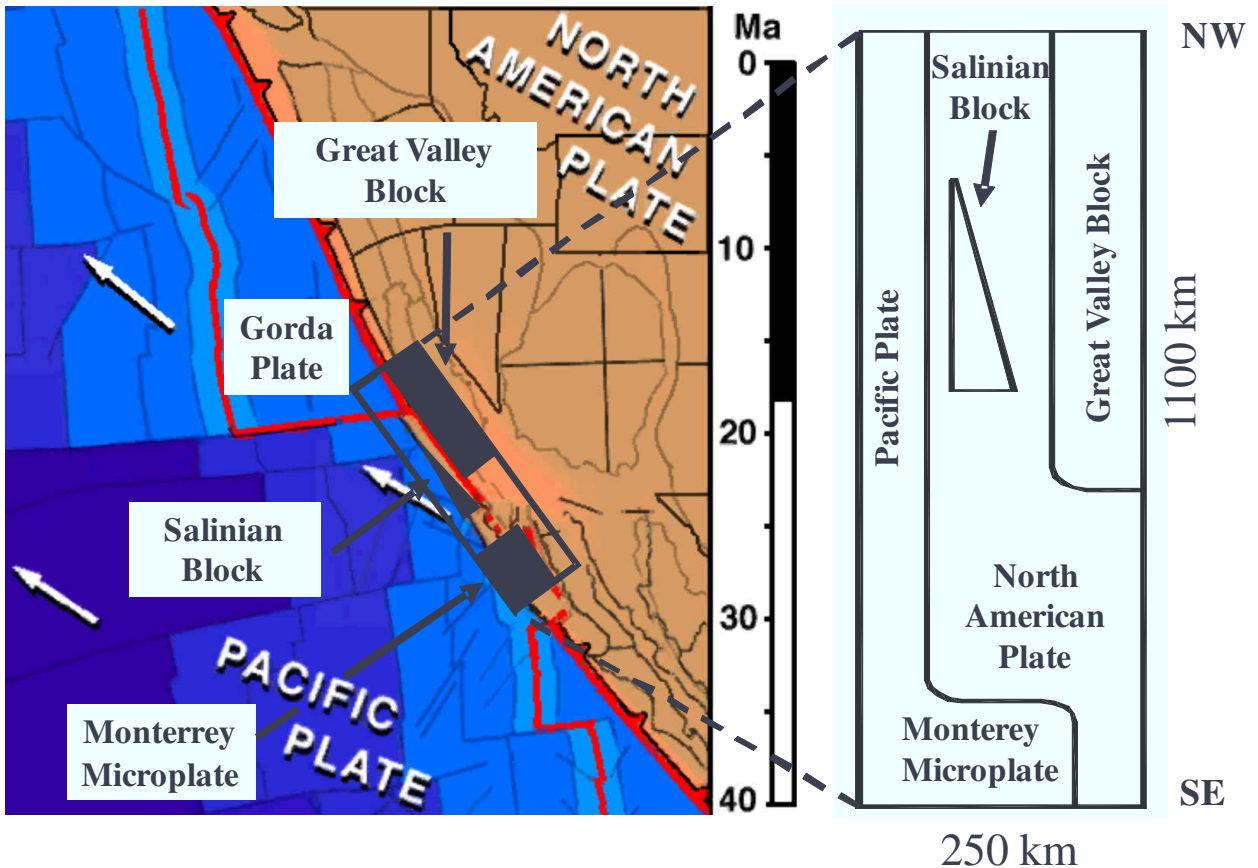


Figure 5.1 Positioning of the model domain with respect to reconstructed plate configuration (Atwater and Stock, 1998). Background snapshot is taken from animation by T. Atwater. Vertical bar shows time in Ma. Also shown are the major mantle (Gorda slab, Monterey microplate, and Great valley) and crustal (Salinian block) heterogeneities included in the model.

Discretization and boundary conditions

The typical three-dimensional model domain spans laterally 250 by 1100 km and 80 km in depth. The finite element discretization of the domain containing about 70000 elements is shown in Fig. 5.2. The entire plate velocities used in the model are calculated with respect to the fixed Great Valley reference frame. We assume the Pacific plate velocity to be about 3.5 cm/yr relative to the Great Valley block (Argus and Gordon, 2001). The y-axis of the model coordinate system is coincident with the direction of the Pacific plate velocity. The convergent velocity component between the Pacific and Great Valley is small at present (< 3 mm/yr) (Argus and Gordon, 2001); therefore it is assumed to be zero. The significant change of the Pacific plate direction about 8 Ma (Atwater and Stock, 1998) is also ignored in these models for the sake of simplicity. The kinematical condition from both the Pacific plate and the Great Valley are applied on the western and eastern boundaries of the domain above the 40 km depth. Below this level and on the other vertical boundaries we impose lithostatic pressure. On the top boundary of the domain we model the deformation of the free surface. On the bottom boundary the Winkler boundary condition is enforced, which implies zero mantle drag forces and isostatic equilibrium with asthenosphere.

Model setup and discretization

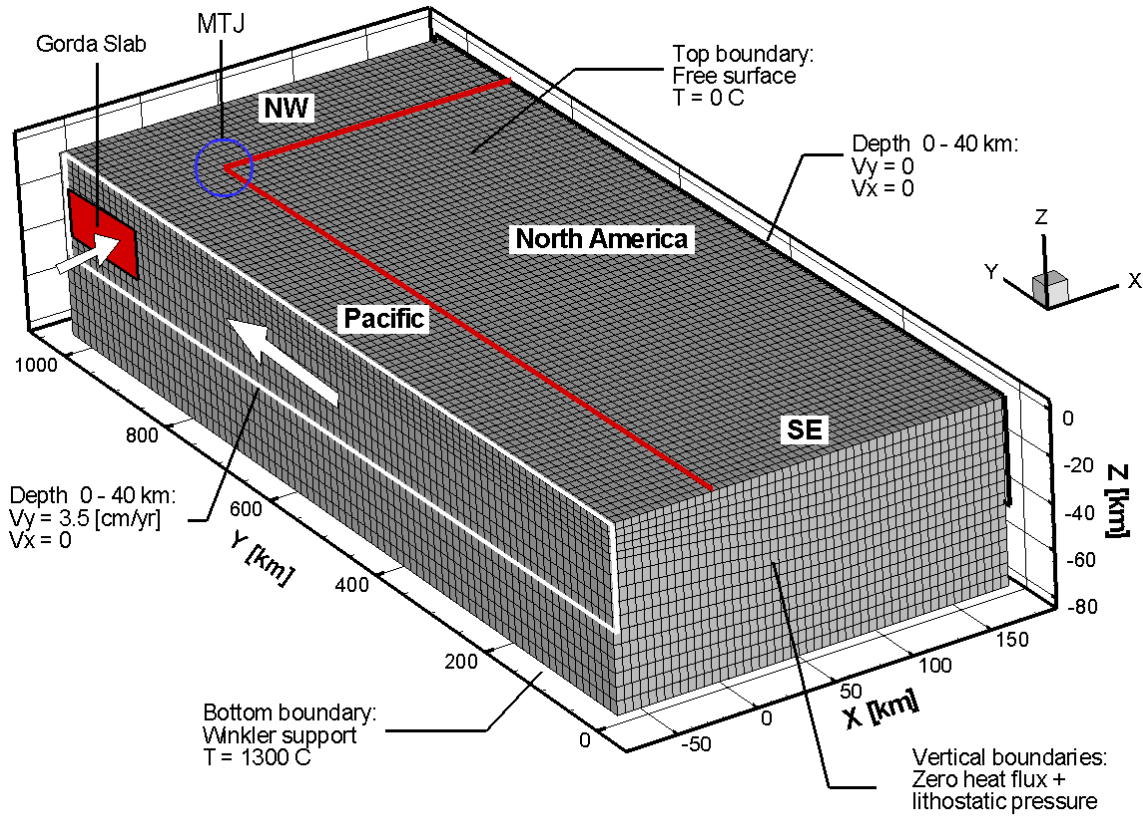


Figure 5.2 Finite element discretization and boundary conditions of the three-dimensional model. Thick red lines show surface traces of the boundaries between Pacific, North America, and Gorda plates. White arrows indicate applied plate velocities. Velocities are given in the fixed North America reference frame. MTJ is the Mendocino Triple Junction.

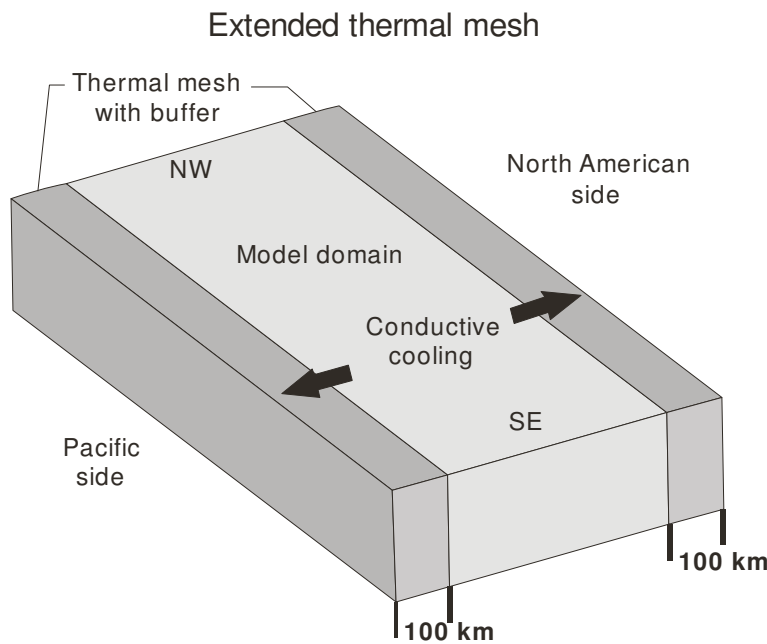


Figure 5.3 Block diagram showing extended domain for solution of the energy balance equation. Thermal mesh covers the area 100 km larger than the model domain to the East and West. Rigid body motion is assumed outside of the model domain. The heat flow on the eastern and western boundaries is nonzero, allowing conductive heat exchange between the model and Pacific and North American plates.

Model scenarios

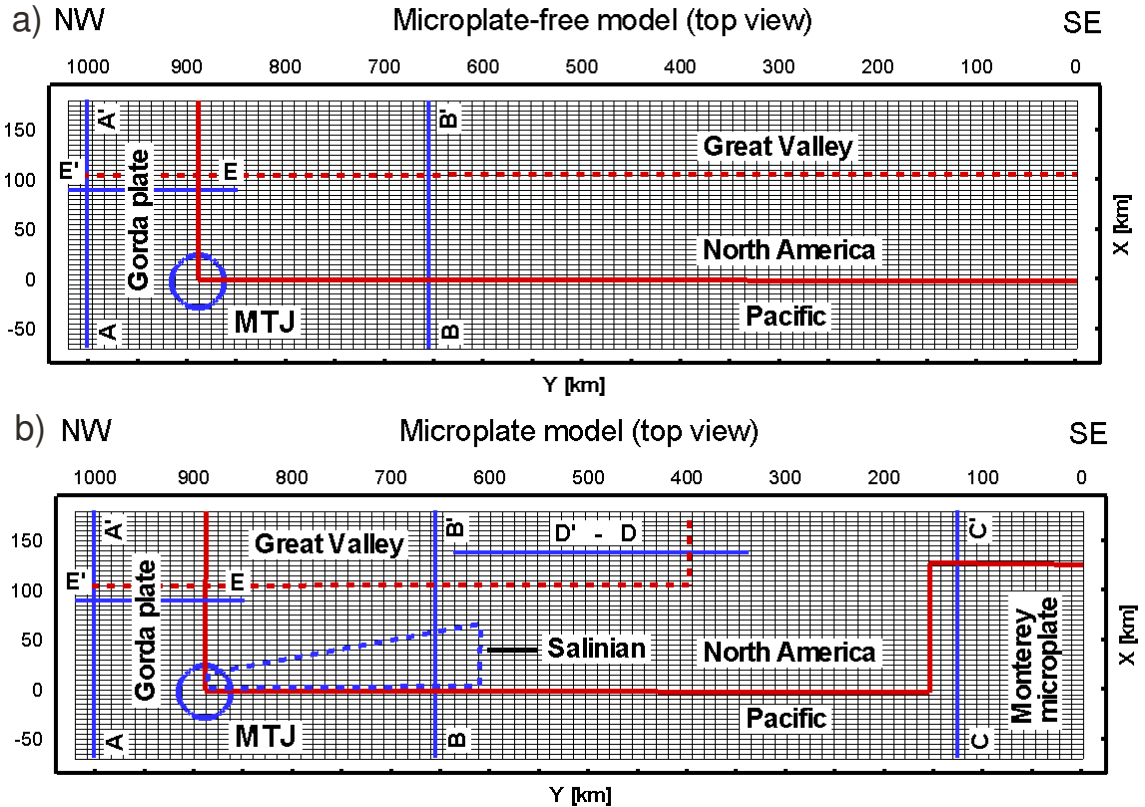


Figure 5.4 Top views on the discretized finite element models (model scenarios). a) Model with the slab window and subducting Gorda slab. b) Model with the thermal heterogeneity in the slab window on the South (the Monterey microplate). Dashed red lines indicate boundary of the Great Valley Block. Dashed blue lines indicate boundaries of the Salinian crustal heterogeneity. Solid blue lines indicate the profiles used in the Figures (5.5) – (5.7).

The formulation of the thermal boundary condition includes prescribed temperatures on the top and bottom boundaries: 0 and 1300 °C, respectively and zero heat flow on the north-western and south-eastern boundaries. To include the influence of heat exchange with Pacific plate and with Great Valley block, we employ the composite model shown in Fig. 5.3. The thermal equations are solved on a mesh which is 100 km larger than model domain to the west and east. The thermal properties of the material and kinematics outside the mechanical model are assumed from corresponding boundaries of the domain. Such rearrangement of the computational scheme substantially better reproduces the character of interaction between the modeled region and surrounding plates.

Reference frame

The plate velocities in the model are measured with respect to fixed Great Valley block. The model box is fixed in the reference frame of the Mendocino Triple Junction. Material can enter and leave the model box through the north-western and south-eastern boundaries and from below. We assume material properties and history variables for the incoming material from the

closest material inside the box (see Chapter 4). The output for the figures is provided in the fixed Pacific plate reference frame.

Model scenarios

Two substantially different families of models are considered in this chapter. (i) The first family of model aims to study the influence and character of the terrane transfer from the North American to Pacific plate due to thermal equilibration in the slab window (Furlong et al., 1989). This setup is shown in Fig. 5.4a. It does not include the microplate and along-strike variation of the Great valley structure. To reproduce the character of the material flow in the asthenosphere upwelling zone more realistically, the model contains the idealization of the Gorda plate subduction. (ii) The second family of models is substantially more complicated (Fig. 5.4b). Besides all features of the previous type, these models contain the Monterey microplate on the southern end of the domain. In certain models of this type we also include the structural heterogeneity which we interpret as Salinian crustal block. The major aim of this type of models is to estimate the influence of the mantle and crustal heterogeneities (microplates, crustal blocks) on the style of terrane accretion along a transform plate boundary.

Thermal and crustal models

Initial temperature structure in the models is substantially three-dimensional (Fig 5.5). All the tectonic elements such as the Great Valley or Monterey microplate are basically specified as the thermal heterogeneities in the mantle. The major element of the thermal model is the uniform initial temperature gradient assumed about 15 °C/km to mimic the typical low surface heat flow in the forearc basin. To the south of MTJ (Fig 5.5b) we initially assume direct upwelling of hot (1300 °C) asthenosphere material to the depth 10 km below the top of Gorda slab. The geometry of Juan de Fuca plate (Gorda) slab is simplified from the compilation of geologic and geophysical data in the Coast Ranges (Page et al., 1998). In the 10 km transition zone we assume linear temperature change between the asthenospheric temperature and the temperature on the top of Gorda slab. The same method is used to specify temperature distribution in the Pacific plate, Gorda Slab and the Monterey microplate. The initial thicknesses of these tectonic elements are assumed to be 40 km for Pacific and Monterey (Fig 5.5c), and 25 km for Gorda plate (Fig 5.5a). The 10 km transition zone in the slab window is introduced for stability reasons. The initial configuration of the slab window has no along-strike variation in the microplate-free models (Fig 5.5b). In the models with microplate, the top of asthenosphere becomes shallower at the southern margin of Great Valley, gradually inclining from 50 km to 25 km (Fig 5.5d).

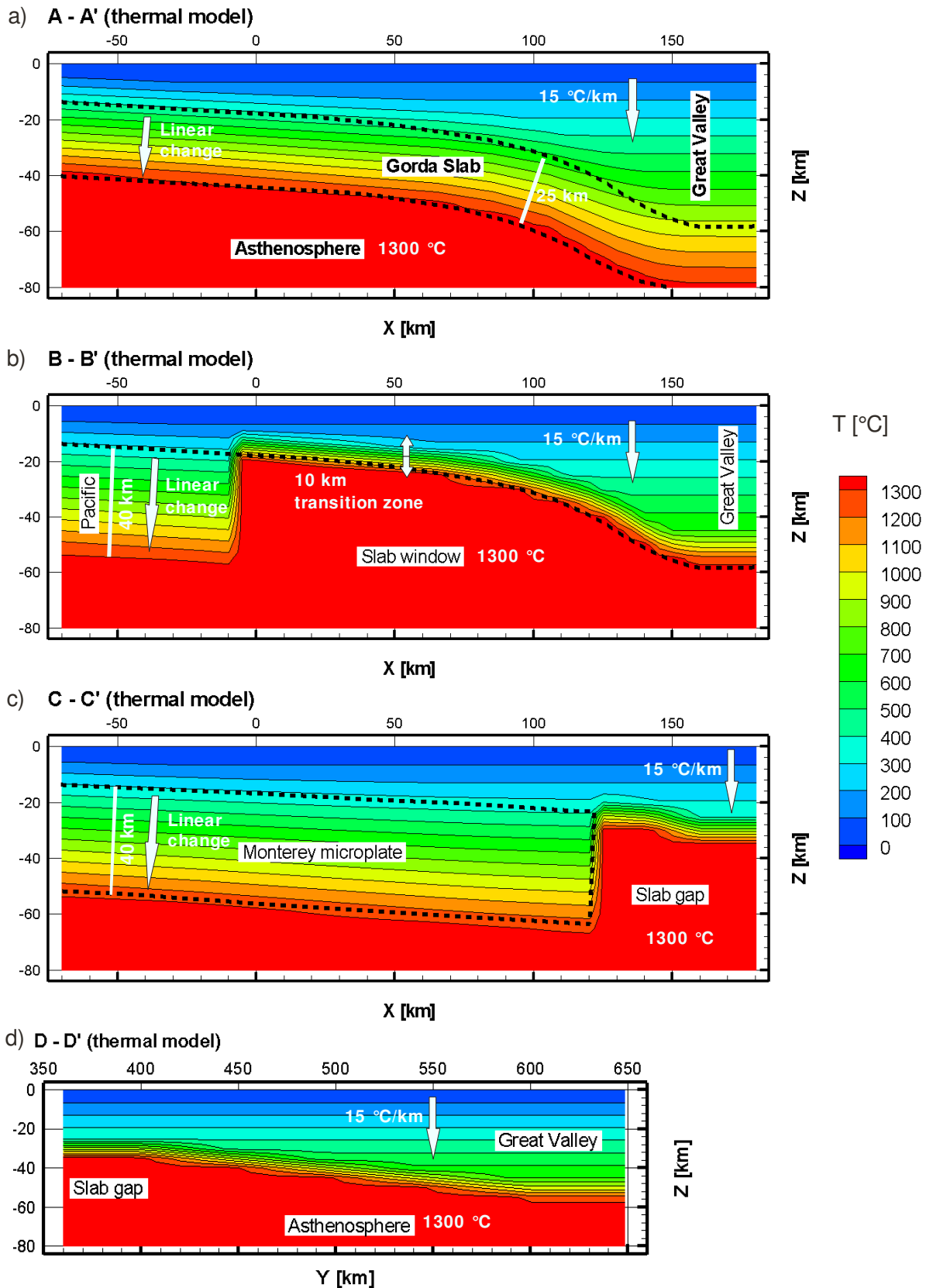


Figure 5.5 Initial thermal structure of the model. Colours indicate temperature magnitudes in $^{\circ}\text{C}$. The uniform gradient $15\text{ }^{\circ}\text{C}/\text{km}$ is assumed for the crustal layer. Asthenosphere temperature in the slab window is $1300\text{ }^{\circ}\text{C}$. Temperature distribution in the oceanic slabs is linear along the slab thickness. 10 km transition zone is assumed between the lithosphere and the slab window. a) Temperature profile in the Gorda slab (profile A-A'). b) Initial temperature distribution in the slab window (profile B-B'). c) Temperature profile in the Monterey microplate (profile C-C'). d) Temperature profile at the southern margin of the Great Valley (profile D-D'). The profiles are marked in Figure (5.4).

The configuration of the Moho discontinuity is simplified after compilation of geologic and geophysical data in the Coast Ranges (Page et al., 1998). Fig. 5.6 shows the initial crustal model in the strike-perpendicular profile. We neglect strike-parallel variations of the crustal thickness in the initial crustal model. However the thickness of crustal layer is allowed to vary during the solution. These variations are tracked by the finite element discretization as shown in Fig. 5.6. For simplicity, the upper and lower crustal layers are not differentiated by the mechanical properties. The material characteristics used for both the mantle and the crust are summarized in the Table A.1 (Appendix). The Salinian crustal block is 15 km thick and differs from the surrounding crust by a lack of strain softening. The friction coefficient within Salinian block remains always equal to ambient value of 0.6. This approach precludes generation of faults within the block and allows it to move as relatively rigid.

Gorda slab subduction idealization

One of the distinctive features of presented models is the incorporation of the simplified Gorda slab subduction. The model idealization is shown in Figure 5.7. In the real tectonic situation near the MTJ, the Gorda slab remains in dynamic equilibrium with the mantle and the surrounding plates. Moreover, the direction and the velocity of the plate motion have been changing since at least 6.5 Ma (Riddihough, 1984). The plate itself behaves non-rigidly and accommodates significant deformations (Chaytor et al., 2004). In our models we assume that Gorda plate moves north-westward in accord with Pacific plate with the velocity 3.5 cm/yr (Figure 5.7b) and converges towards the North America with the velocity 1.5 cm/yr (perpendicular direction) (Figure 5.7a). Velocities and temperatures in the entire grid points that belong to Gorda slab are prescribed. The deformation within the slab is limited to bending in the subduction zone. To adequately model the interaction of the Gorda slab with the overriding North America plate, we include a weak subduction channel located above top of Gorda slab. The thickness of the subduction channel is assumed to be 5 km with effective friction coefficient 0.02 (e.g. Sobolev and Babeyko, 2005). Behind of Gorda slab we include a 10 km transition zone with linear increase of the friction coefficient to standard value (Figure 5.7b). This low effective friction can be attributed to excessive fluid pressure in the subduction channel.

Effective friction on the faults

The models presented here do not include coupling between the deformation and the porous flow processes. It also does not directly describe the generation of clay minerals in the fault gouge. Therefore, the weak rheology on the major fault relative to surrounding crust (e.g. Zoback, 2000)

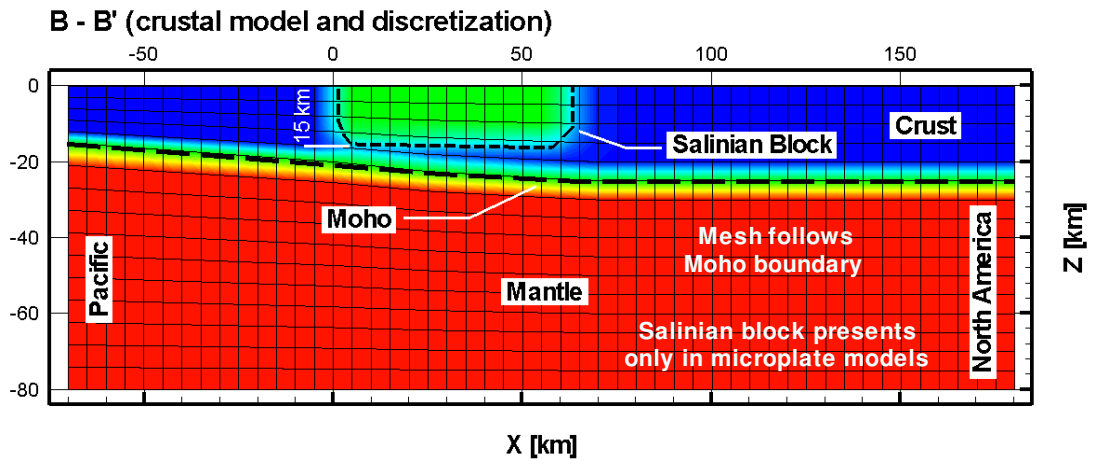


Figure 5.6 Crustal model and discretization in the vertical direction shown in the profile B-B' (Figure 5.4). Colours show different material phases: red (mantle), blue (crust), green (Salinian block). The Salinian block remains relatively strong relative to surrounding crust due to deactivated strain softening. The effective friction within Salinian block remains equal to 0.6 (ambient value). Salinian block is only included in certain models with microplate. The two-dimensional Moho structure is simplified after Page et al. (1998). The finite element mesh follows the configuration of the Moho discontinuity during the solution.

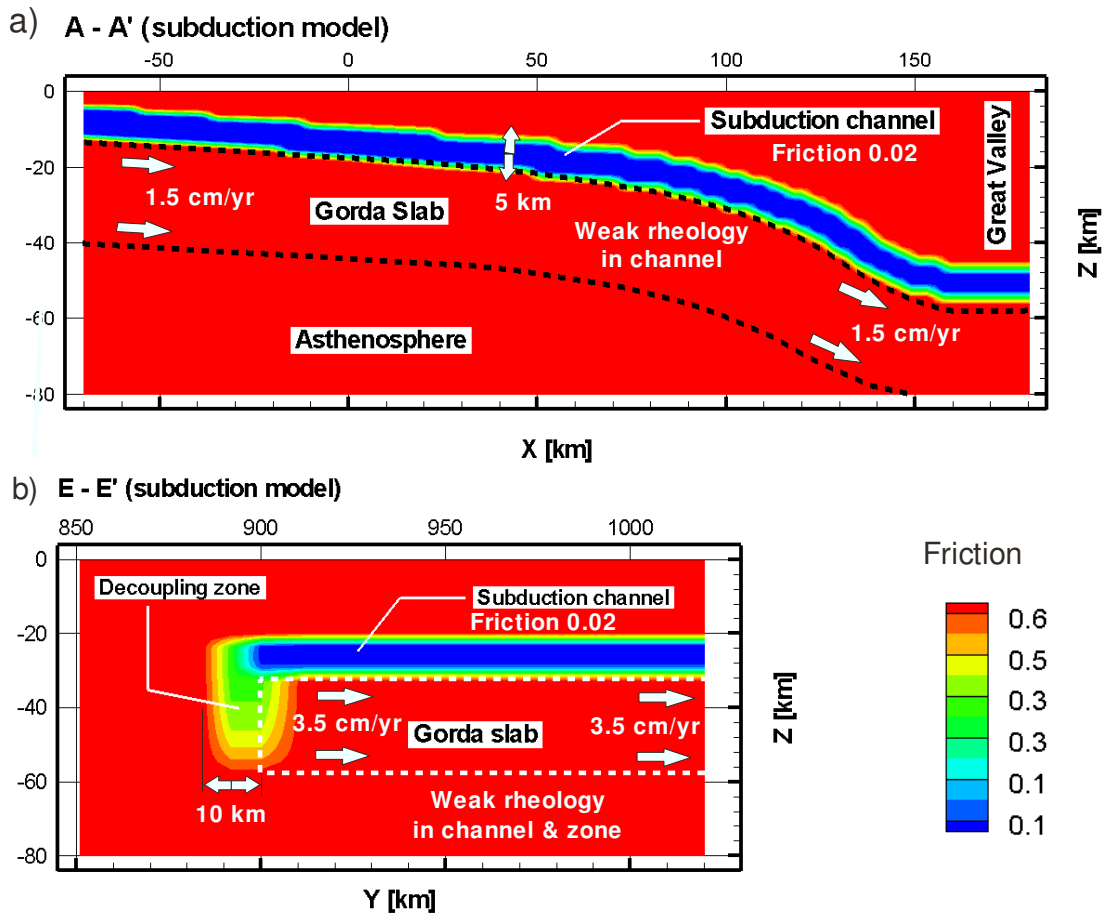


Figure 5.7 Idealization of the Gorda plate subduction. Colours indicate magnitudes of the effective friction coefficient. a) strike-perpendicular profile (A-A'). b) strike-parallel profile (E-E'). The profiles are marked in Figure 5.4. The friction coefficient in the subduction channel is assumed to be 0.02. In the decoupling zone behind the slab, the variation is linear, up to ambient value 0.6. The motion of the Gorda slab is prescribed kinematically with 3.5 cm/yr component in the direction of the Pacific plate and 1.5 cm/yr in perpendicular direction. The profiles are marked in Figure (5.4). Slab configuration is simplified after Page et al. (1998).

cannot be properly simulated in our models by elevated fluid pressure and intrinsically low coefficients of friction. We parameterize both these effects with a simple strain-softening model. The physical friction is replaced with an effective friction which is made dependent on the accumulated plastic strain. The friction coefficient reduces with an increase of the fault strain, making the fault relatively weaker than surrounding crust. As the model parameter we specify the minimum friction coefficient which is allowed to be reached on the mature strike-slip fault. This parameter is varied between the models to provide a better insight on the strength of the strike-slip fault.

It should be noted that the elasto-visco-plastic rheology used in the Finite Element code SLIM3D is capable to produce strain localization into narrow zones (faults). Therefore, the faults in our models are not prescribed but rather self-generated in the vicinity of minimal lithospheric strength zones.

In the formulation of the models it is more convenient to operate with the friction coefficient than with friction angle. The magnitudes of the friction coefficient are uniquely converted to the Mohr-Coulomb friction angles for internal use in the code. The technique can be found in e.g. Zienkiewicz and Taylor (2000).

5.2 Model results with Gorda slab and slab window

The models of this type are aimed to verify numerically the influence of the thermal conduction processes in the slab window and the asthenosphere upwelling on the lithospheric deformation in the San Andreas Fault System (Furlong, 1993; Furlong et al., 1989). According to hypothesis of Furlong et al. (1989), the conductive cooling in the slab window should cause the migration of the plate boundary to the east first in the mantle and then in the crust (“fault jump”). Using a simple 2D thermal model, Furlong (1993) considered only the influence of temperature and grain size evolution on the creeping flow in the mantle. In this chapter we extend consideration towards the complete thermo-mechanical three-dimensional setup shown in Fig 5.4a. The advantage of these models with respect to previous studies (Sobolev et al. 2003) is in presence of along-strike variations and completely three-dimensional material flow in the slab window.

Thermal equilibration in the slab window

In the vicinity of MTJ the temperature distribution is highly unstable due to large horizontal and vertical temperature gradients. As the triple junction migrates north-westward the temperature gradients begin to decrease due to conductive cooling. This process is shown in Fig. 5.8.

Thermal equilibration in the slab window

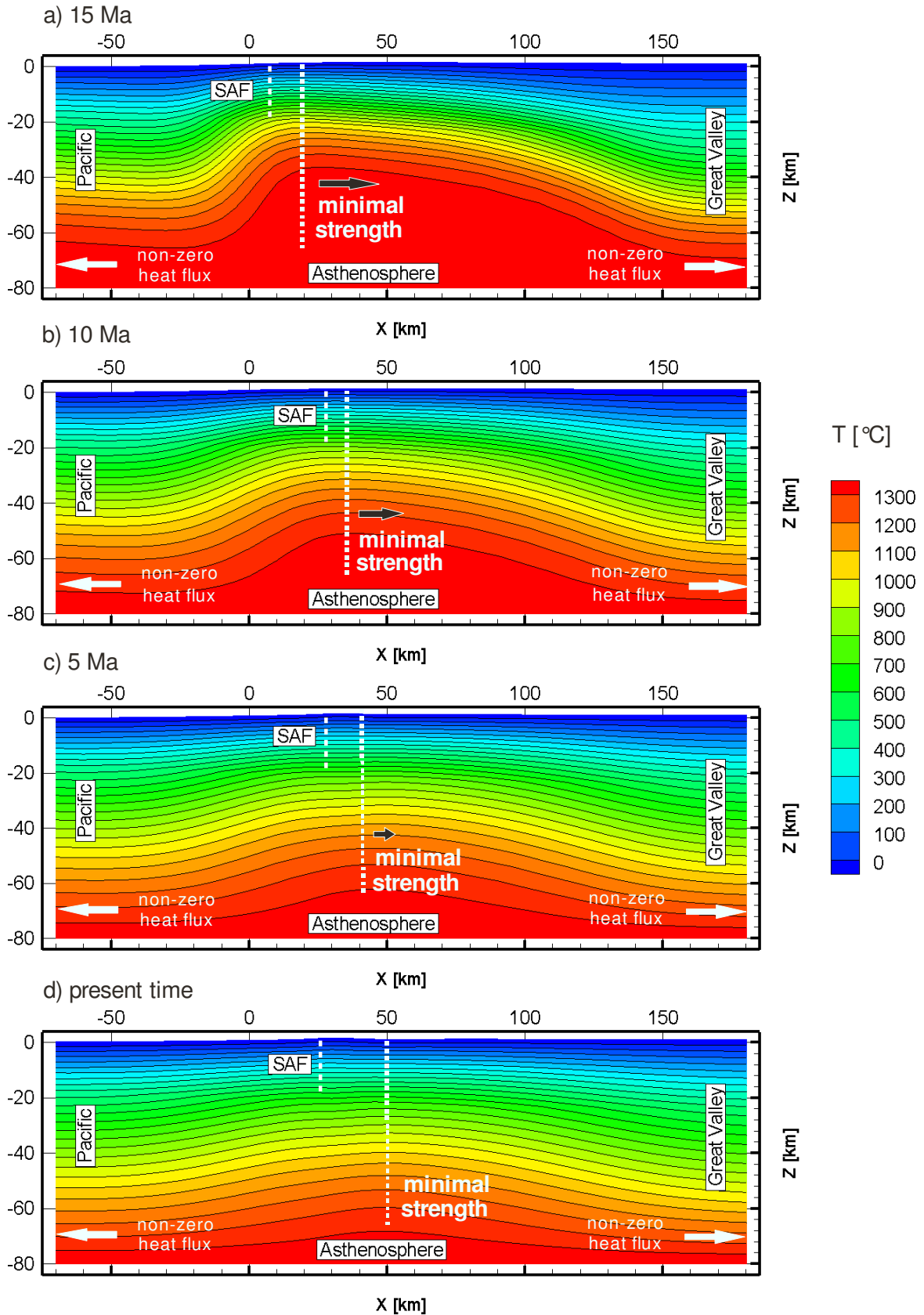


Figure 5.8 Distribution of the temperature at times a) 15Ma, b) 10 Ma, c) 5Ma, and d) present time in the model with 10% hear heating efficiency. Colors indicate temperature magnitudes in °C. The profile is located 700 km to the South of the Gorda plate. The locus of the minimum integrated lithospheric strength is indicated with dashed vertical line, which migrates eastward. Black arrow indicates relative rate of migration. White arrows indicate the nonzero heat flux on the vertical boundaries of the model. Also shown are the current locations of the major fault in the San Andreas System.

The equilibration causes the locus of maximum mantle temperature and minimum lithospheric strength to move in the eastward direction. Hence, the plate boundary both in the mantle and in the crust also tends to move in the eastward direction. The initial location of the plate boundary is coincident with the eastward margin of the Pacific plate. As it is shown in Fig. 5.8, the eastward migration of the minimal strength profile is fastest in the initial moment due to the highest lateral gradients. Then it gradually slows down with decrease of the gradients. After 18 Myr from the model initiation the migration virtually stops at the point located about 50 km to the east of the initial location. The lithosphere far from the MTJ has almost uniform thickness at that moment. Therefore even with realistic slab configuration, the estimated maximum potential for displacing the plate boundary due to thermal equilibration is no more than 50 km. Note that the models presented here are free from the assumption of zero heat flux between the slab window and surrounding plates (e.g. Furlong, 1984). We suggest that it can significantly accelerate the cooling rate of the asthenospheric material.

The actual migration of the plate boundary and surface faults is a function of multiple parameters. As it is show in the next sections, the most important of those are shear heating, and jump-induced transpression. Also important is the fault friction. However, to minimize the parametric search we set the minimal friction on the faults to be 0.12 according to the best-fit estimate for the San Andreas slip-dependent friction (Bird and Kong, 1994). The influence of the fault friction is considered in the context of more sophisticated models with Monterey microplate and Salinian block.

Effect of shear heating

The heating processes related to active shear deformation can produce significant local perturbations in the temperature field. Hartz and Podladchikov (2008) showed that this mechanism can cause up to tenfold reduction of the effective strength of typical lithosphere. It is important to note that major part of this reduction (heat release) is due to viscous deformation in the lower crust and in mantle, rather than frictional deformation in the crust if friction coefficient at major fault is low (like 0.12 assumed in this model). In the cold brittle crust the viscosities are too high to be influenced by the shear heating, while the brittle rheology itself is temperature-independent. Therefore, the shear heating process can significantly increase stability of the long-term persisting fault zone due to the heating of mantle roots. The slip-weakening process can redouble this effect, since the older the fault the lower the frictional strength. Both these mechanisms act like inhibitors of the natural migration of the plate boundary in the direction of the thermal impulse of slab window. As a result, the persistence of the older faults is substantially better motivated energetically, than initiation of younger faults.

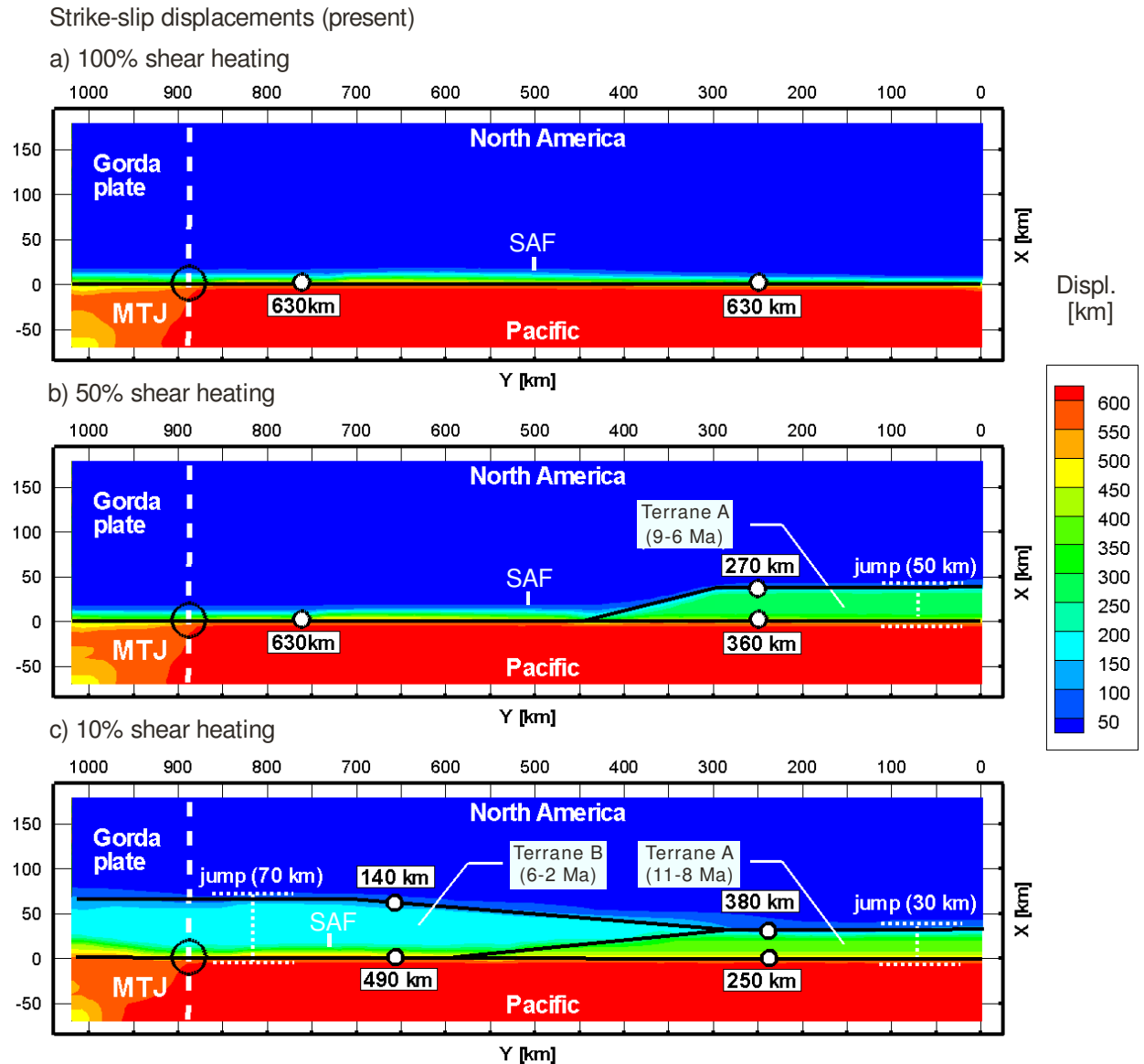


Figure 5.9 Comparison of the total strike-slip displacements in the direction of Pacific plate (shown in colours) between the models with a) 100%, b) 50%, and c) 10% shear heating efficiency. Output is in the fixed Pacific reference frame. White circles and labels indicate the offsets on major faults. Also shown are the times of terrain accretion events from North America to Pacific and correspondent magnitudes of the eastward fault jumps.

The results of the three-dimensional models generally confirm these considerations. We calculated a series of similar setups with variable magnitude of shear heating efficiency parameter, which is the part of dissipated mechanical energy directly converted into heat. The heat is produced not only due to slip on the fault (Lachenbruch and Sass, 1980), but also due to viscous deformation in the ductile part of the lithosphere. Fig. 5.9 demonstrates the influence of the shear heating efficiency parameter on the fault migration and terrane accretion. If the shear heating is activated to the full extent, the model fails to displace the crustal plate boundary during the entire 18 Myr of model time (Fig. 5.9a). Gradual reduction of the shear heating efficiency causes intensification of the eastward fault migration. At 50% efficiency (Fig. 5.9b)

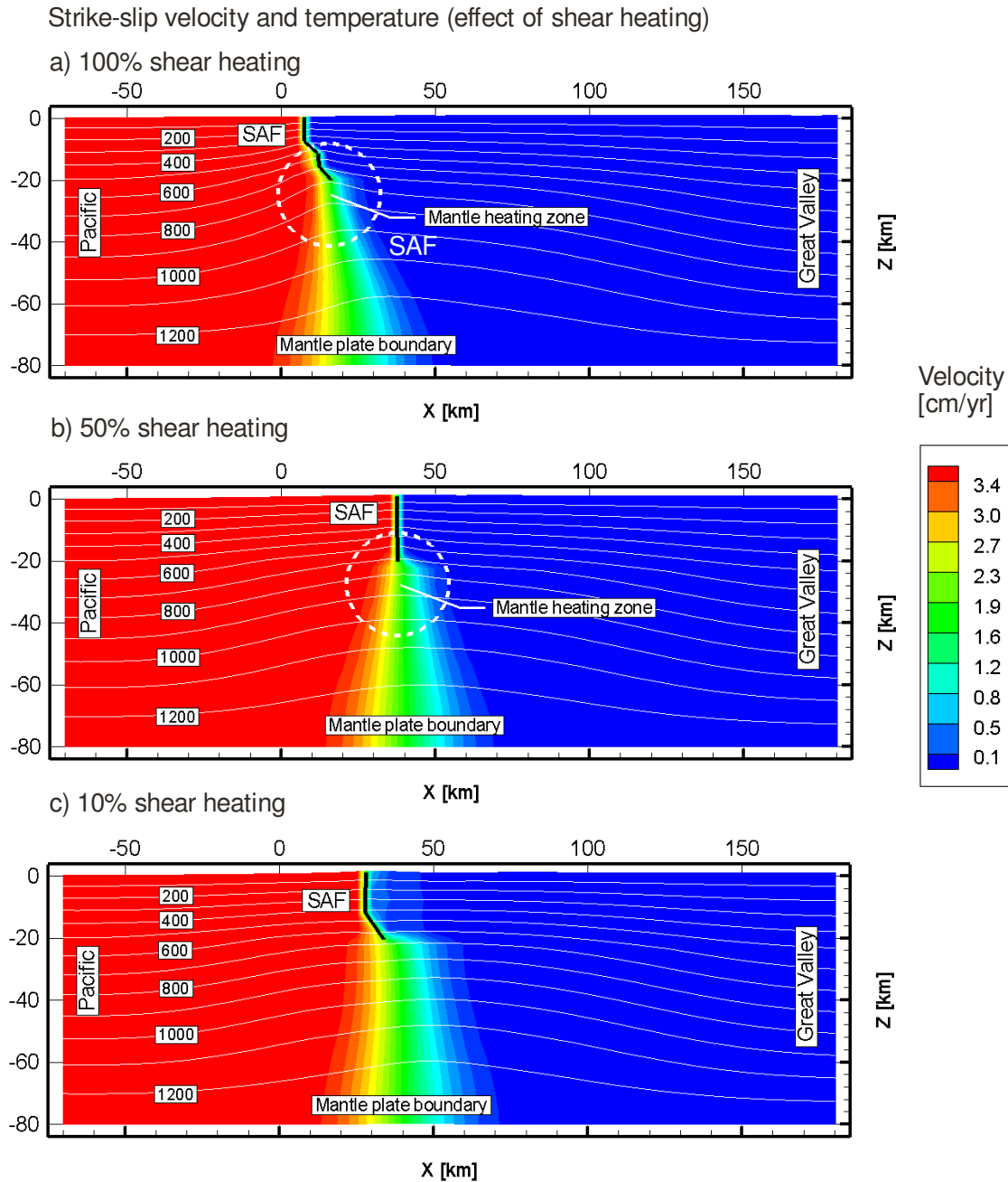


Figure 5.10 Comparison of the fault configuration and temperature distribution between the models with a) 100%, b) 50%, and c) 10% shear heating efficiency. The profile is located 700 km to the South of the Gorda plate. Colours indicate strike-slip velocities in the direction of the Pacific plate. White labeled lines show the temperature distribution. The dashed circle indicates the zone of significant local temperature perturbation caused by both brittle and ductile deformation processes in the crust and mantle.

the model predicts a 50 km eastward jump and accretion of one terrane (Terrane A) within the 9-6 Ma period. At 10% efficiency (Fig. 5.9c) the model predicts accretion of two terranes: Terrane A with 30 km eastward jump during 11-8 Ma and Terrane B with 70 km eastward jump during 6-2 Ma.

The configuration of the plate boundary at present time in the southern part of the domain is shown in transects in Fig 5.10. The models can in principle resolve the discrepancy in the

location of the plate boundary in the crust and mantle. There is a direct correlation of the thermal anomaly in the ductile part of the plate boundary and the shear heating efficiency. The largest anomaly is observed in the model shown in Fig 5.10a (100% shear heating). Local temperature perturbation in the mantle beneath the fault causes the viscosity to decrease, thus reducing the lithospheric strength and making the older fault more optimal locus for the plate boundary.

Among the models shown in Fig 5.9, only 10% shear heating efficiency case (Fig 5.9c) can be compared with the fault configuration near the San Francisco Bay Area. However, it is not clear how that scenario could be realized in practice. The part of the dissipated energy could be taken away by migrated fluids. It is, however, unlikely that almost the entire dissipated energy can be withdrawn in such a way, especially from the mantle, where the effect of shear heating is the most important. Altogether it renders the conductive cooling of the asthenosphere in the slab window incapable to explain the eastward migration of the fault. We suggest, therefore, that another mechanism is required to explain the inland jump of the plate boundary.

Effect of transpression

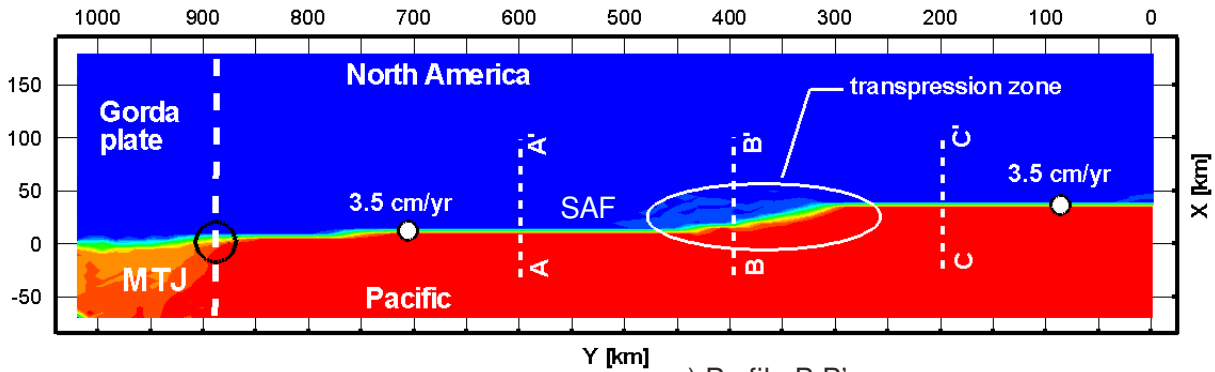
In the context of migrating plate boundary we should mention another highly important effect: which is transpression or transtension induced by the fault stepping. If a strike-slip fault tends to form a restraining fault step, the fault stepping is inhibited by the additional tectonic load related to transpression. On the other hand, if the fault forms a releasing step, the resulting transtension may significantly accelerate the fault stepping. The magnitude of additional tectonic forces depends on the magnitude of the offset.

The models discussed in this chapter, are characterized by restraining fault steps. Therefore, the transpression poses another mechanism that works against the natural migration of the plate boundary in the eastward direction. The transpression is a highly three-dimensional effect. All previous models (e.g. Furlong et al., 1989; Sobolev et al., 2003) could not resolve it due to inherent restrictions (lack of the mechanical equation; 2D formulation). However in the presented models this effect is adequately represented. Fig. 5.11 shows the distribution of the strike-slip velocity on the surface (a), on the north (b), across (c) and on the south (d) of the transpression zone. The shown situation corresponds to the long-term persisting transpression zone formed in attempt to capture the Terrane B in the model with 50% shear heating efficiency. In southern profile, the plate boundaries in the crust and mantle are coincident. In the northern profile there is large discrepancy between the boundaries with an intensive sub-horizontal detachment zone beneath the crust. In the transpression zone both the surface uplift and corresponding crustal thickening are observed. Note that the strike-slip fault is tilted at depth.

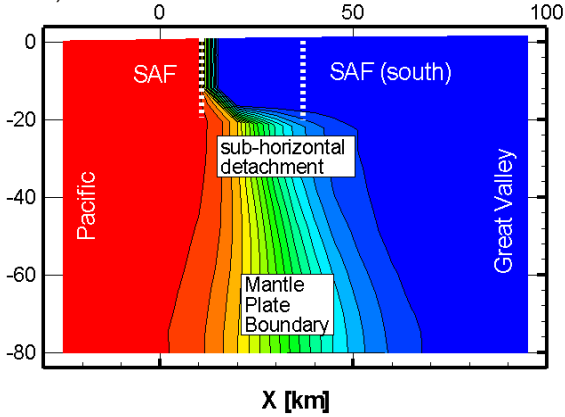
It can be concluded that the sufficiently advanced three-dimensional thermo-mechanical

Strike-slip velocity (effect of transpression) (50% shear heating, present)

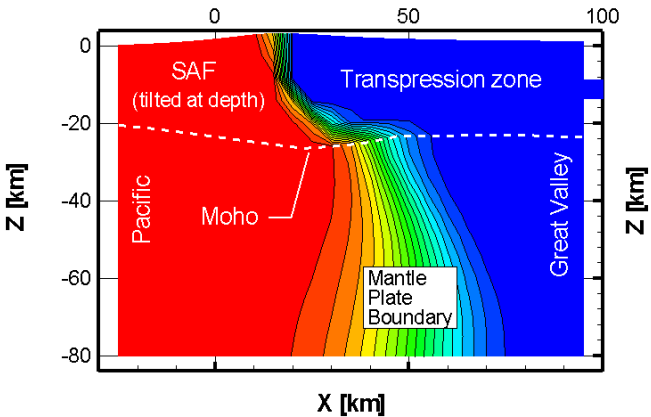
a) top view



b) Profile A-A'



c) Profile B-B'



d) Profile C-C'

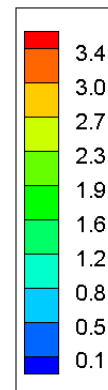
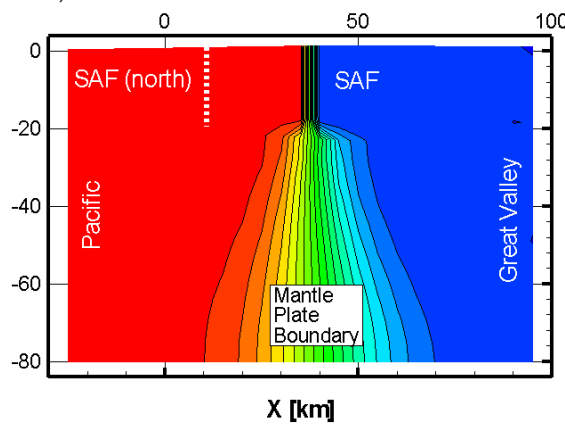


Figure 5.11 Distribution of the strike-slip velocities (colours) in the model with 50% shear heating efficiency at present time. a) top view. b), c), and d) show the profiles A-A', B-B', and C-C', respectively. White labeled circles indicate the present long-term slip rates on the major fault. Also shown are the locations of major faults with respect to profiles and the Moho configuration in the active transpression zone.

models show that both the shear heating of the mantle roots beneath the mature faults, together with the transpression, make the thermal equilibration process in the slab window alone insufficient to explain the inland migration of the San Andreas plate boundary.

5.3 Model results with Monterey microplate

In this section we investigate a geological concept of Nicholson et al. (1994) from the viewpoint of the thermomechanical modeling. This concept states that a series of microplate capture events has been the primary reason of the inland migration of the SAFS plate boundary over the recent 20 Ma. The model setup used here (Fig 5.4b) is a gradual extension of the model considered in the previous section. The entire features including Gorda plate subduction and the slab window are inherited. In addition we specify a low-temperature perturbation in the southern part of the domain which we interpret as the Monterey microplate lithosphere (Fig 5.5c). The Monterey plate subduction and spreading from the Pacific are assumed to have been ceased before the model initiation. To ensure proper coupling between the North America and Monterey microplate we specify regular friction (0.6) in the fossil subduction channel. Rapid increase of the effective friction on the subduction interface also adequately characterizes the ceased subduction process. Since the low shear heating efficiency scenario was ruled out in the previous section, we assume that the entire dissipated mechanical energy is converted into heat in the models with Monterey microplate.

Influence of microplate on terrane accretion

The remnants of the Farallon slab, stalled in the fossil subduction zone, create much stronger heterogeneity in the mantle than the cooling of the uprising asthenosphere, providing the more efficient and direct way for transferring the North American terranes to Pacific plate. After the completion of the microplate capture and cessation of subduction, the integrated lithospheric strength in the zone covered by microplate significantly increases. This factor dramatically decreases or even completely prohibits the generation of faults in the continent above the microplate. Hence the plate boundary can instantly jump inland up to the microplate margins. This behaviour is confirmed by the three-dimensional models presented here. Since the beginning of model time the San Andreas plate boundary form a large restraining bend which tends to bypass the microplate obstacle. The results of the models with different friction on the faults are shown in Figures 5.12 and 5.13. In the northern and central parts of the domain the plate boundary locates near the edge of the slab window. Further south it curves and strikes around the eastern margin of the microplate (Fig. 5.12a, Fig. 5.13a). The transpression resulting from this fault bend transmits the tectonic forces from the Pacific plate directly to the North America crustal blocks to the north of microplate. Thus the terrane accretion is greatly intensified by the captured microplate.

The actual influence of the captured microplates on the tectonic evolution of the North America western margin is far more complicated compared to presented models (Nicholson et al., 1994). However, with this simple idealization we can still estimate basic effects relatively

Strike-slip velocity (terrane accretion) (microplate model, friction 0.12)

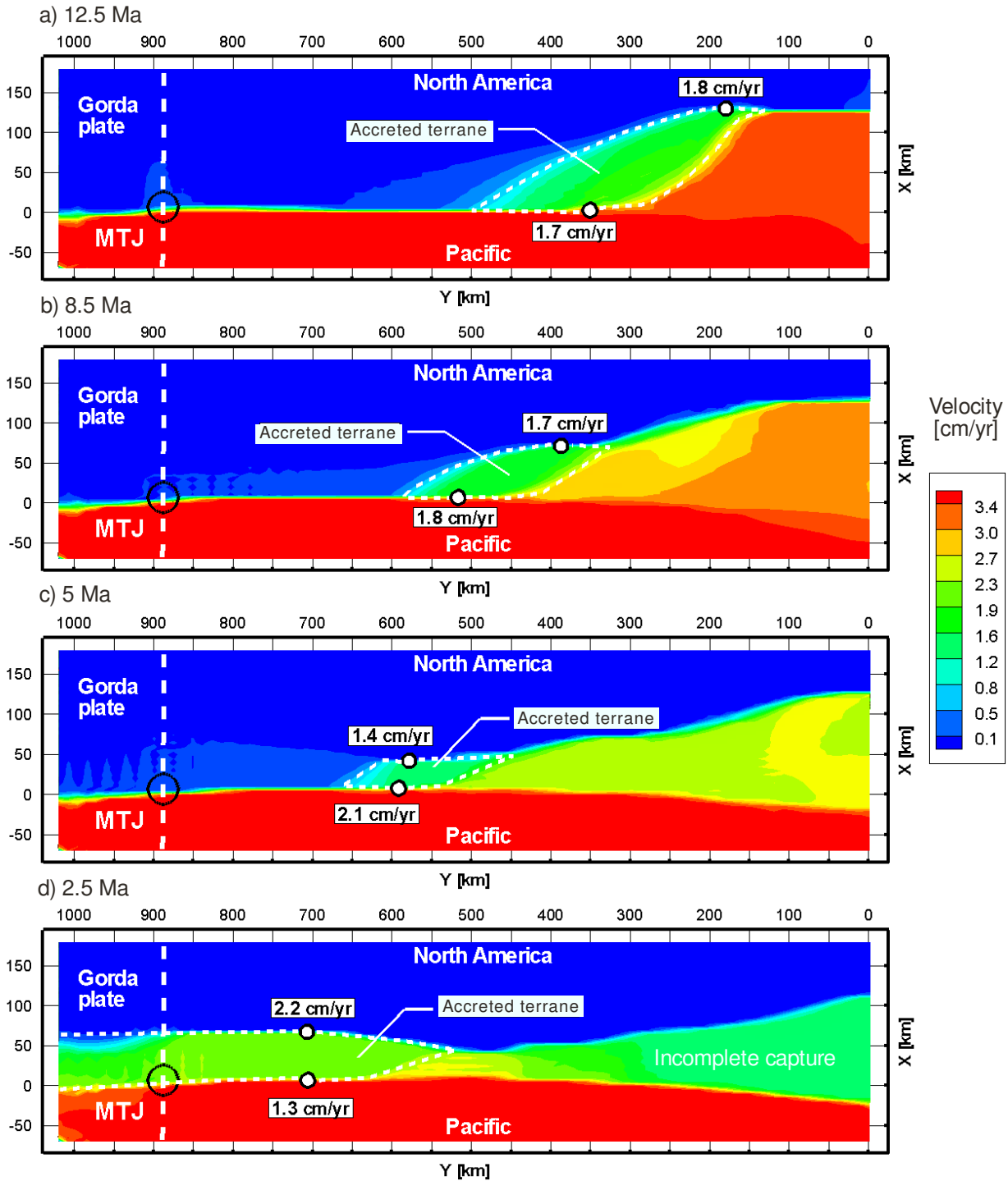


Figure 5.12 Results of the model with Monterey microplate, and effective friction on the faults 0.12. Colours indicate strike-slip velocities in the direction of the Pacific plate. Output is in the fixed Pacific reference frame. Figures a) – d) show the events of terrain accretion from North America to Pacific occurring about 12.5 Ma, 8.5 Ma, 5 Ma, and 2.5 Ma, respectively. White labeled circles indicate the present long-term slip rates on the major fault. Note the incomplete character of the terrain accretion; the maximum velocity of the Pacific plate (3.5 cm/yr) is never reached within the terrains.

well. In particular we can study the effect of the fault friction on the style and rate of terrane accretion. Fig. 5.12 and Fig. 5.13 show the accretion events predicted by the models with effective friction 0.12 and 0.09. Shown are the time moments in which the slip rates on the western and eastern bounding faults of newly captured terrane are comparable. Accretion of at

Strike-slip velocity (terrane accretion) (microplate model, friction 0.09)

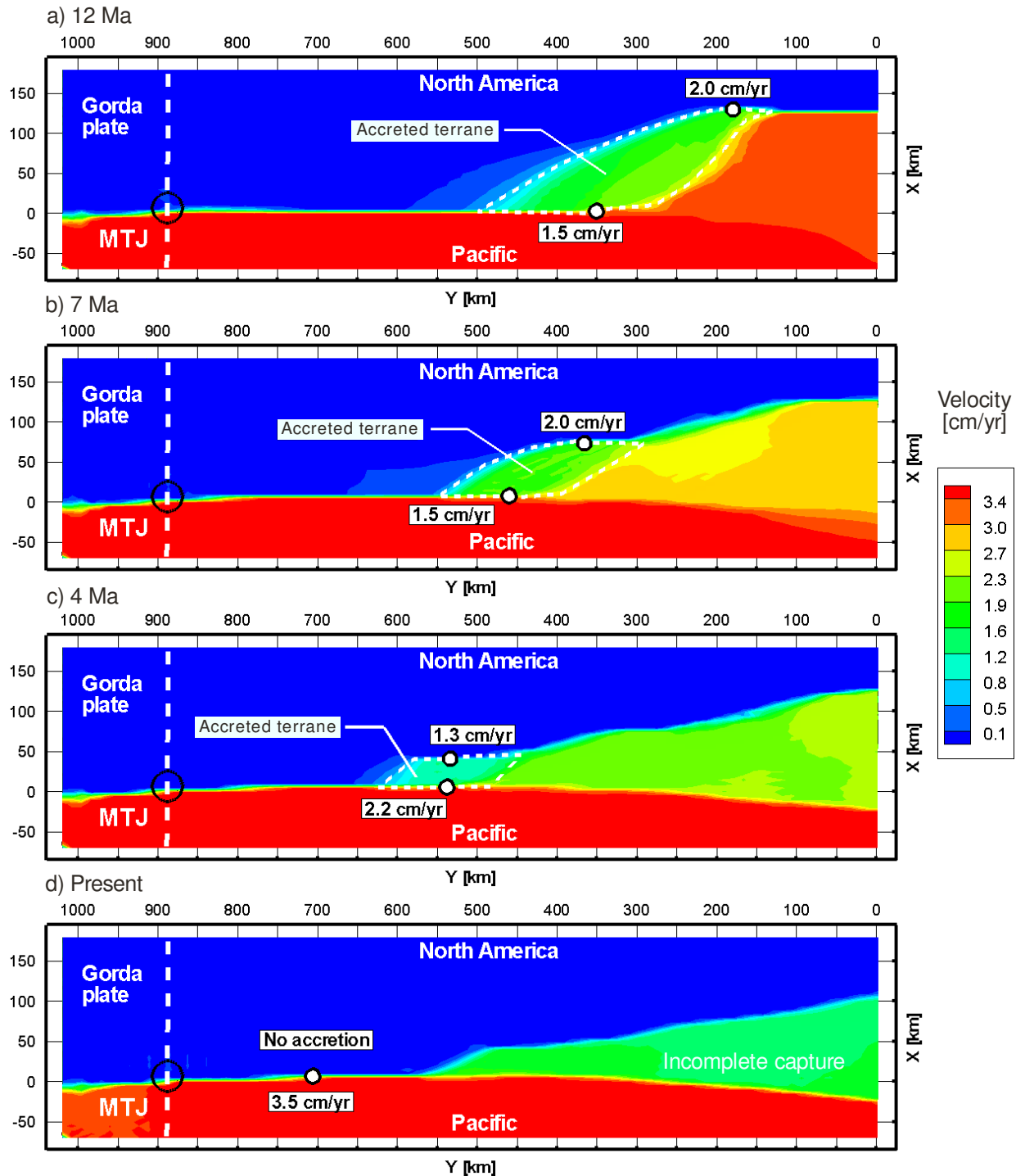


Figure 5.13 Results of the model similar to that shown in Figure (5.12) with effective friction on the faults 0.09. Note that coupling between the terrains and Pacific plate is less effective than with friction 0.12. The terrain accretions take place about a) 12 Ma, b) 7 Ma, and c) 4 Ma. The fourth terrain remains untied from Pacific until present.

maximum four independent terranes can be identified in the models. By comparing the Fig. 5.12 and Fig. 5.13 we can conclude that the higher friction on the faults the greater the coupling between terranes and Pacific. Note that in the model with friction 0.09 the last terrane accretion does not occur until present. The capture of other three terranes is systematically delayed in time.

In both models the terranes are only partially captured by the Pacific. The following

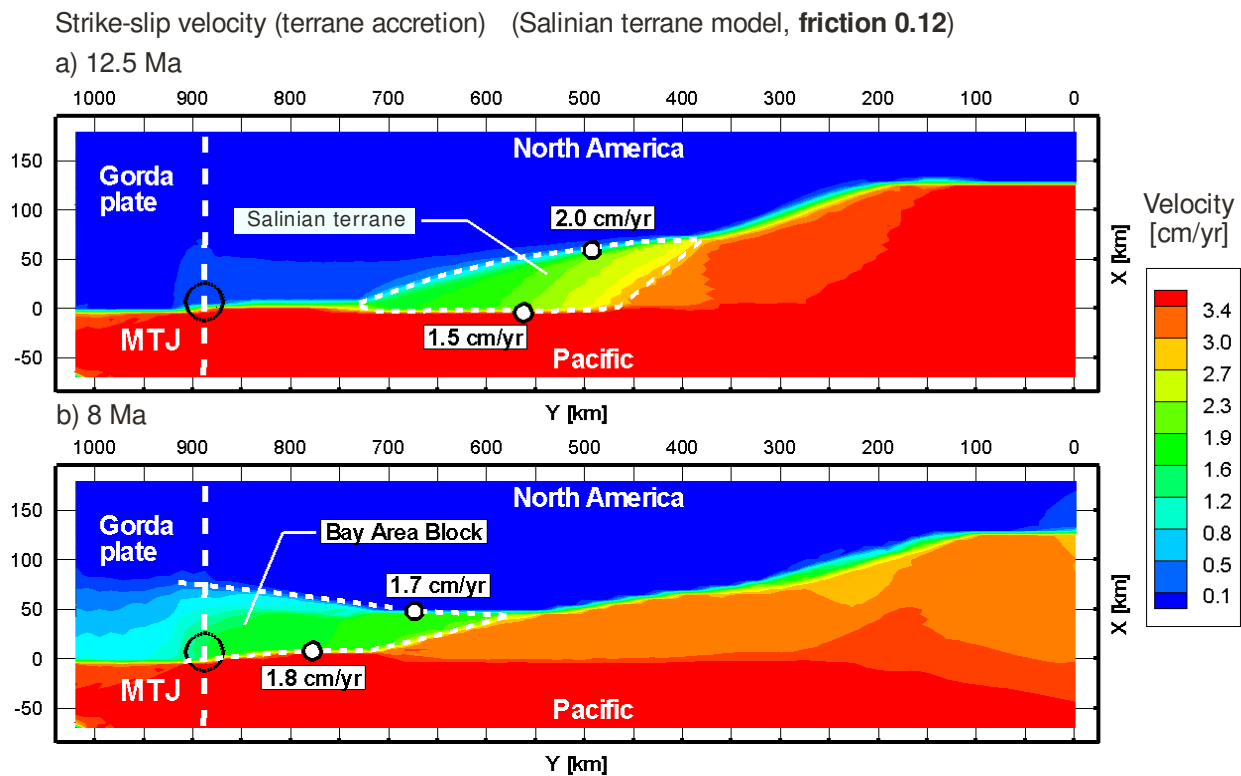


Figure 5.14 Distribution of the strike-slip velocity (colours) in the model with Monterey microplate, Salinian block, and effective fault friction 0.12. Output is in the fixed Pacific reference frame. a) Capture of the Salinian Block occurs about 12.5 Ma. b) Capture of the Bay Area block occurs about 8 Ma. White labeled circles indicate the long-term slip rates on the major fault.

considerations are helpful to explain this result. There is interplay between the size of the microplate in the model and the efficiency of the terrane capture, since the microplate is the most important mechanism which transfers the tractions from Pacific. In the present tectonic situation in southern California and northern Mexico the configuration and number of the microplates is different to what is assumed in presented models. In particular we cannot exclude that the captured Arguello, Guadalupe and Magdalena microplates (e.g. Stock and Lee, 1994) transfer much higher traction from the Pacific than the Monterey microplate alone. Dynamic modeling of the microplate capture in the southern California is beyond the scope of this work. The presented model describes only the hypothetical situation with single microplate. However the model is still useful for estimating the general consequences of microplate capture for the overlying continent.

Model with Salinian block: reference model

The style of crustal terrane accretion can also depend on the heterogeneities in the crust itself. Both the number of captured terrines and degree of coupling can be altered. To investigate this effect we calculated a simplified model which contains a crustal heterogeneity, which we interpret as a Salinian block (Irwin, 1990). The rocks within the Salinian block are assumed to be relatively stronger than surrounding crust. This assumption is in agreement with the geological observations. The geological formations that bounds Salinian block: Franciscan assemblage and

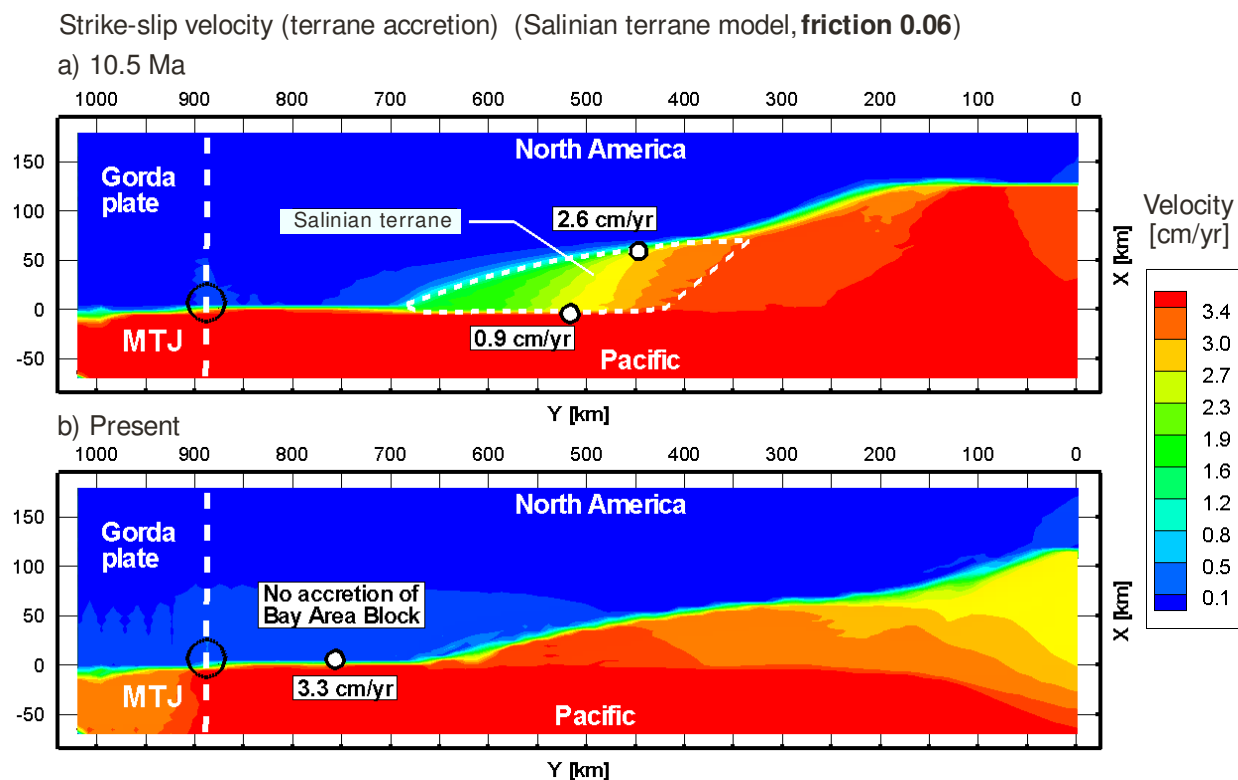


Figure 5.15 Results of the model similar to that shown in Figure (5.14) with effective friction on the faults 0.06. Note that coupling between the terrains and Pacific plate is less effective than with friction 0.12. a) Capture of the Salinian Block occurs about 10.5 Ma. b) The Bay Area block remains untied from Pacific until present.

Great Valley sequence are composed of weaker strata than the plutonic rocks of Salinian block. The presented model is not aimed at explaining the origin or evolution of the Salinian block, but rather at studying the influence of crustal heterogeneities on the continental deformation. In the model the higher strength of the block is achieved by deactivating the strain softening, such that the friction within the block remains equal to background value 0.6. Setup of the model is shown in Fig. 5.4b.

Presence of the Salinian block significantly changes the style of terrane transfer from the North America to Pacific. Instead of the four small terranes in the previous models, the new models identify basically only two major terranes: the Salinian block itself, and the block located to the north, which we interpret as San Francisco Bay Area block. Fig. 5.14 and 5.15 show the influence of the friction on the accretion of North America terranes. Generally the previous influence of friction on the degree of coupling is confirmed. In the model with higher friction (Fig. 5.14) the terranes are captured earlier and more firmly than with lower friction (Fig. 5.15). Note that accretion of Bay Area block does not occur until present with friction 0.06 (Fig. 5.15b). Another conclusion is that presence of the Salinian block significantly intensifies the rate of terrane transfer and degree of coupling. Note that terranes are almost completely captured by Pacific (Fig. 5.14, Fig. 5.15). These results better agree with the observations than the models without Salinian block (Fig. 5.12, Fig. 5.13).

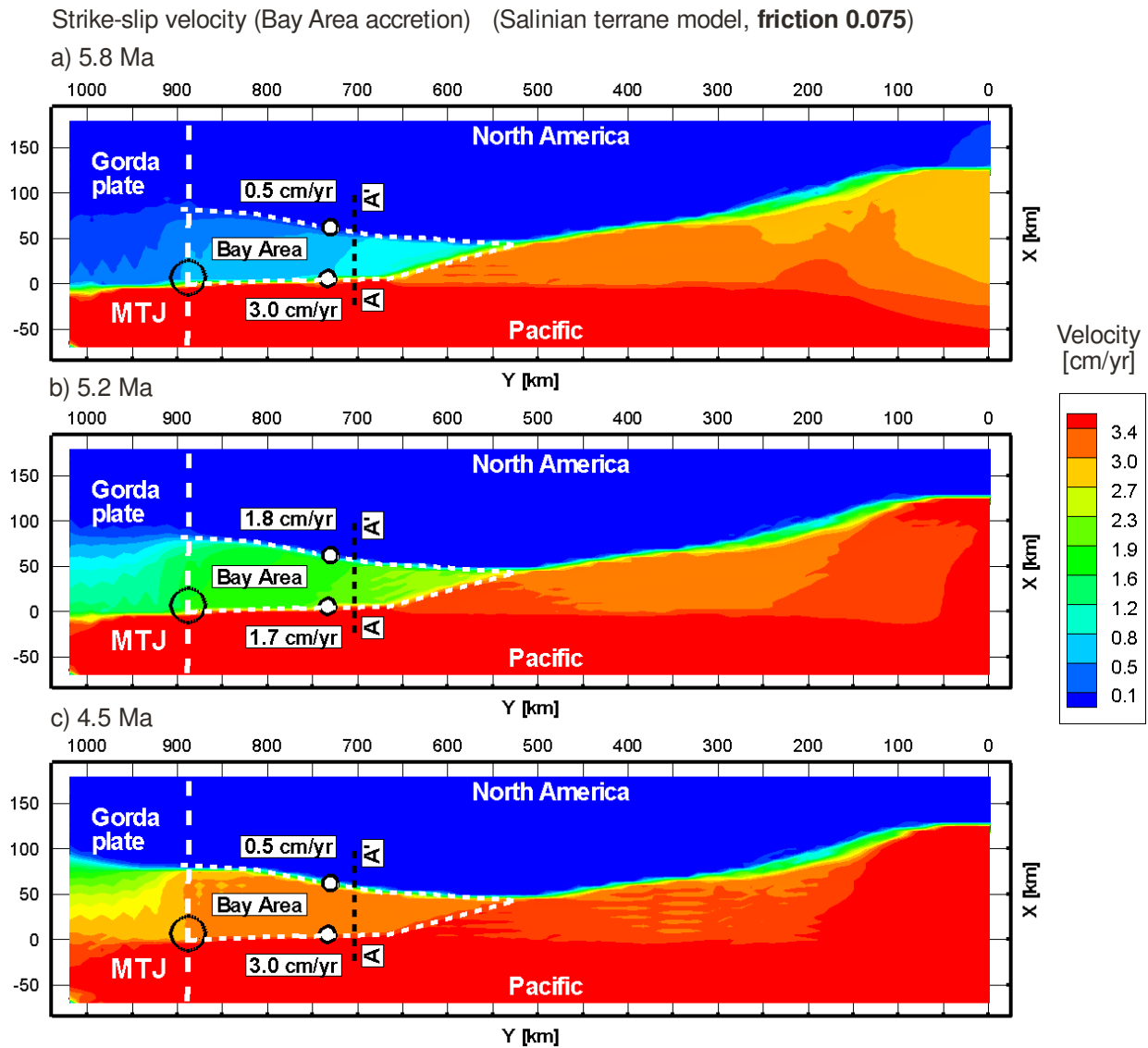


Figure 5.16 Results of the reference model with Monterey microplate, Salinian block, and friction on the faults 0.075 (reference model). Shown are the details of the Bay area accretion process. Colors indicate strike-slip velocities, labeled circles show slip rates. White dashed line shows boundary of the Bay Area block. Black dashed lines mark the profiles shown in the Figure (5.17). a) Initiation of the Calaveras-Heyward fault system (5.8 Ma). b) Equal slip-rates on the Calaveras-Heyward and San Andreas faults (5.2 Ma). c) Reduction of the slip rate on the San Andreas Fault, capture of the Bay Area block (4.5 Ma).

The magnitude of the effective friction on the fault has a two-fold influence on the terrane accretion process in our models. Changing the effective friction can simultaneously alter both the initiation of terrane capture and the rate of terrane velocity increase. In the framework of presented models we have conducted a parametric search with respect to the friction coefficient. We conclude that the time of initiation of the terrane capture can vary in wide range. For example with the friction 0.12, which is equal to the best-fit estimate obtained by Bird and Kong (1994), the accretion of Bay Area block occurs about 8 Ma (Fig. 5.14b). By reducing the friction to 0.06 the capture of the Bay Area can be delayed up to present time.

The rate of terrane velocity increase, on the other hand, is much less sensitive to the friction. Once the capture of the Bay Area block is initiated, it takes about 3-4 Myr to complete the

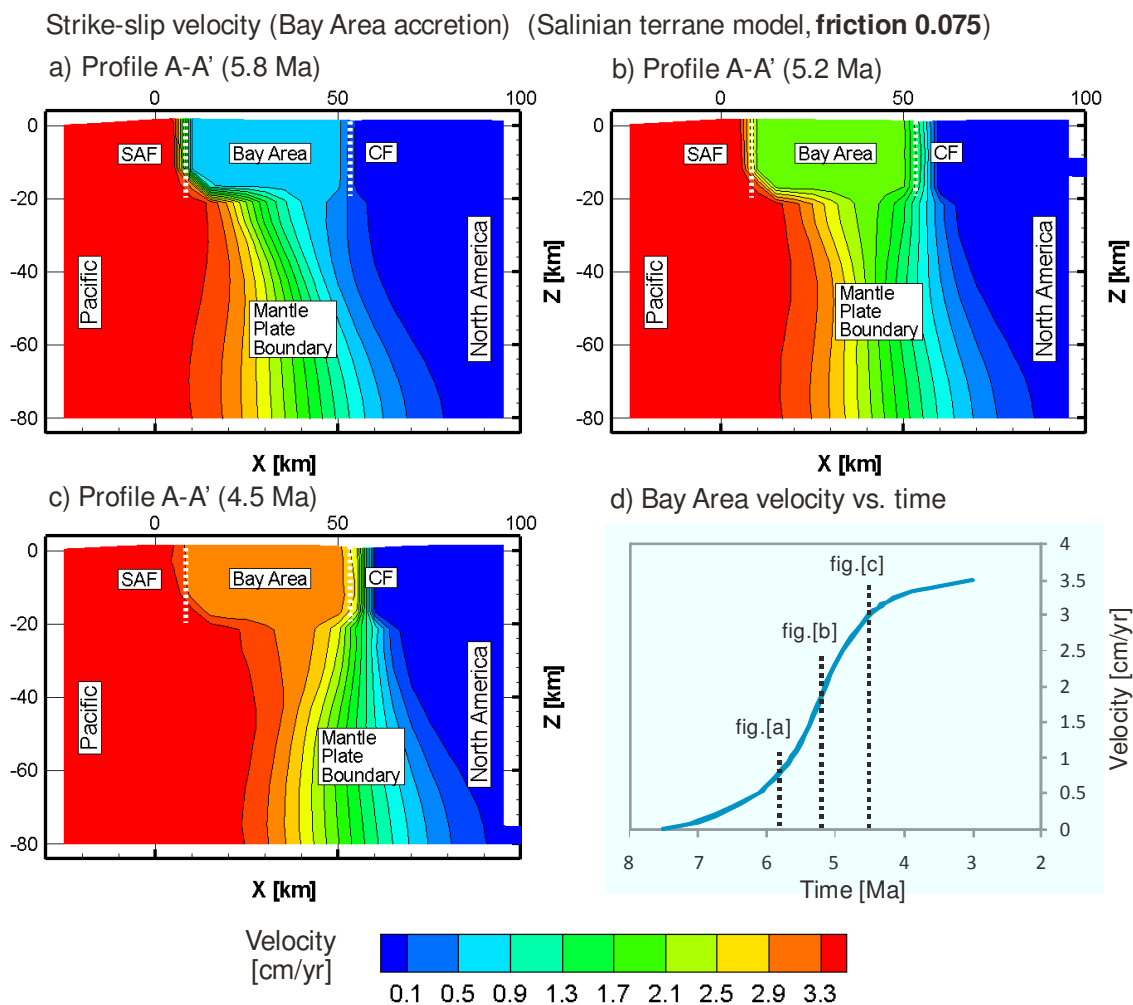
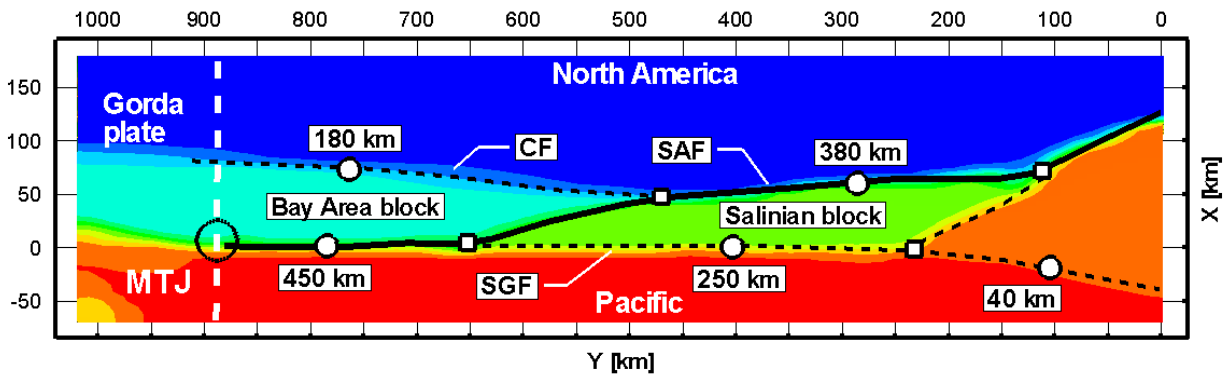


Figure 5.17 Distribution of the strike-slip velocity at depth in the reference model shown in Figure (5.17). a) Mantle plate boundary is displacing eastward relative to SAF below the Bay Area (5.8 Ma). b) Process continues; the eastern and western bounding faults are simultaneously active (5.2 Ma). c) Slip rate on the SAF is reducing; plate boundary moves below the Calaveras-Heyward fault system (4.5 Ma). d) Graph showing time evolution of the Bay Area block velocity in the direction of the Pacific plate relative to fixed North America.

accretion in all the models. The maximal increase of the terrane velocity occurs even in fewer periods about 1.5 Myr. Fig 5.16 and Fig 5.17 show the process of the Bay area accretion in the typical model with friction 0.075, from the top, and in the vertical transects, respectively. Although these results fit much better to the observations than the extended 2D model of Sobolev et al. (2003) (where both western and eastern faults were simultaneously active only during less than 1 Myr), the 3D model is still highly contrasting with the reconstruction of the fault offsets, which predict simultaneous activity of both the eastern and western Bay Area faults since at least 8 Ma (e.g. McLaughlin et al., 1996). Unfortunately, the character of this activity is hardly known. Some authors (e.g. McLaughlin et al., 1996) suggested migration of the plate boundary both eastward and westward with time, due to changing plate velocities and geometry of the plate margin. In the presented models these potential changes are neglected, therefore the plate boundary migrates only in the eastward direction.

Fault offset comparison (Salinian terrane model, friction 0.075) (reference model)

a) Strike-slip displacements



b) Fault offsets near San Francisco Bay Area (Dickinson, 1997)

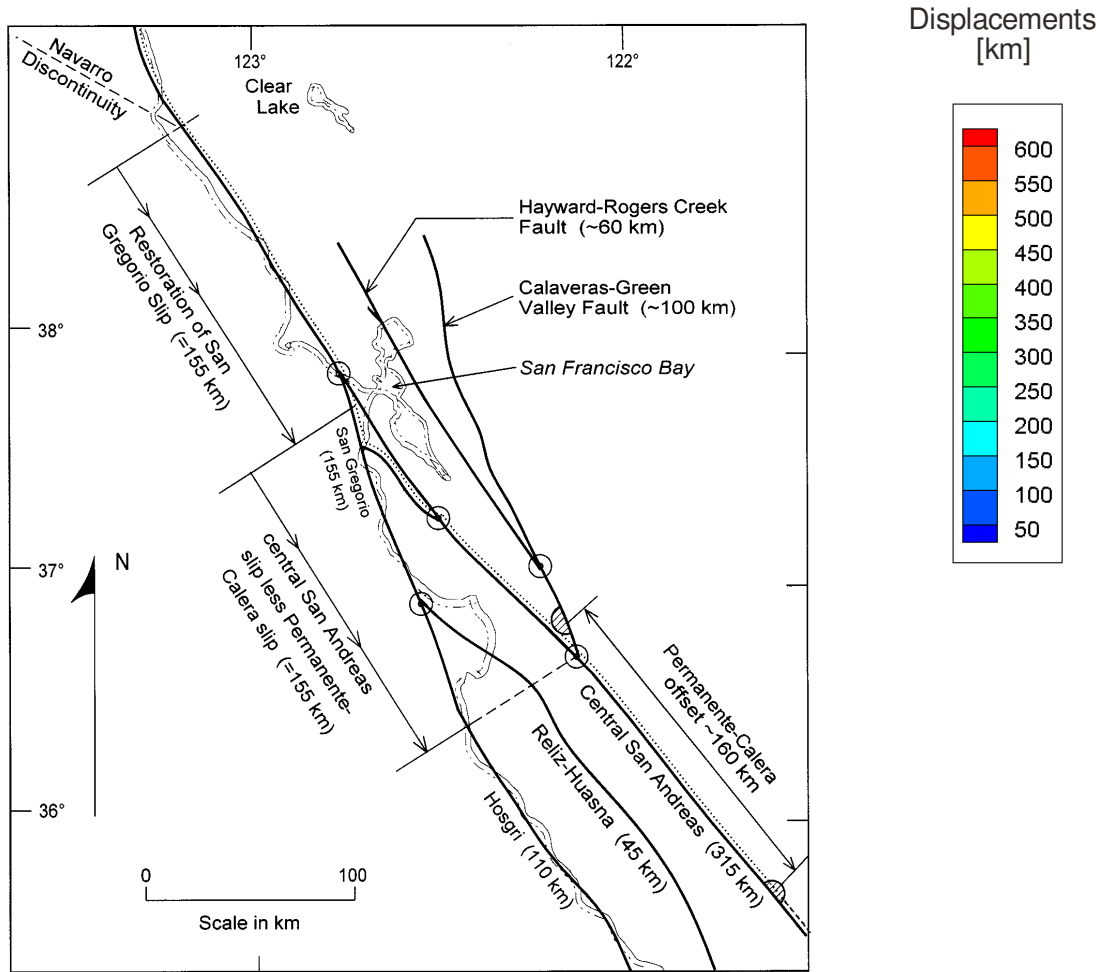


Figure 5.18 Comparison of the major fault offsets between the reference model and the restoration (Dickinson, 1997). a) Results of the reference model. Colours indicate strike-slip displacements. Fault offsets are displayed in the labels near the white circles. White squares show fault junctions. Calaveras and Heyward faults are resolved in the model with one fault marked as CF (Calaveras Fault). The western faults including San Gregorio-Hosgri and Reliz-Huasna are represented by fault marked as SGF (San Gregorio Fault). b) Fault map near the San Francisco Bay area showing the restored magnitudes of the major fault offsets (modified after Dickinson, 1997).

The rate of eastward migration in the results may be significantly biased by the too soft conditions on the northern boundary of the domain. The influence of the terranes located far north from the MTJ may be more pronounced than currently assumed lithostatic pressure.

Another explanation may be related to the increase of the fault effective friction to the east with increasing distance from the fossil Farallon trench, and hence with decreasing amount of fluids. Such eastward-decreasing friction can inhibit generation of eastern faults and extend the period of Bay Area block capture. In any case, these effects are beyond the scope of this work. We suggest that further study is required.

In the framework of presented models we constrain the effective friction on the faults by the amount of slip accommodated on the eastern Bay Area faults without attempting to properly reproduce the time span of Bay Area capture. In our reference model the friction is 0.075, which is even less than the best-fit estimate 0.12 of Bird and Kong (1994). Another most recent similar study (Carena and Moder, 2008) also show comparable values (0.05) for the friction on major faults in the SAFS.

Offsets on major faults

The reference model demonstrates good correlation with the reconstructed offsets on major faults near the San Francisco Bay Area. The pattern of the faults and the present terrane configuration are also reproduced relatively well. However, the model does not resolve certain faults. For example the Hayward-Calaveras fault system is represented with single eastern fault in the model. The western margin of Salinian block is also bounded with single fault, which we interpret here as San Gregorio fault. Fig. 5.18a shows the accumulated terrane displacements in the direction of Pacific Plate in the reference model. The Bay area and Salinian terranes accommodated 180 km and 380 km displacements, respectively, with respect to the fixed Great Valley since 18 Ma. These figures correlate relatively well with the total 160 km for the Hayward-Calavers system and 315 km for the Central San Andreas as shown in Fig. 5.18b (Dickinson, 1997). The 250 km slip on the western bounding faults of the Salinian block predicted by the reference model is larger than the reconstructed 155 km slip on the San Gregorio fault. This discrepancy may be explained by existence of the offshore faults which accommodated slip before or simultaneously with San Gregorio fault.

Asthenosphere upwelling

The three-dimensional flow of the material in the slab window is a distinctive feature of the models presented in this work. Previous studies (e.g. Goes et al., 1997; Van Wijk et al. 2001) contained assumptions regarding the kinematics of the entire plate boundary, due to a lack of mechanical model. In the presented models the dynamic balance between the void formation in the wake of north-westward moving Gorda slab and the material inflow from the asthenosphere

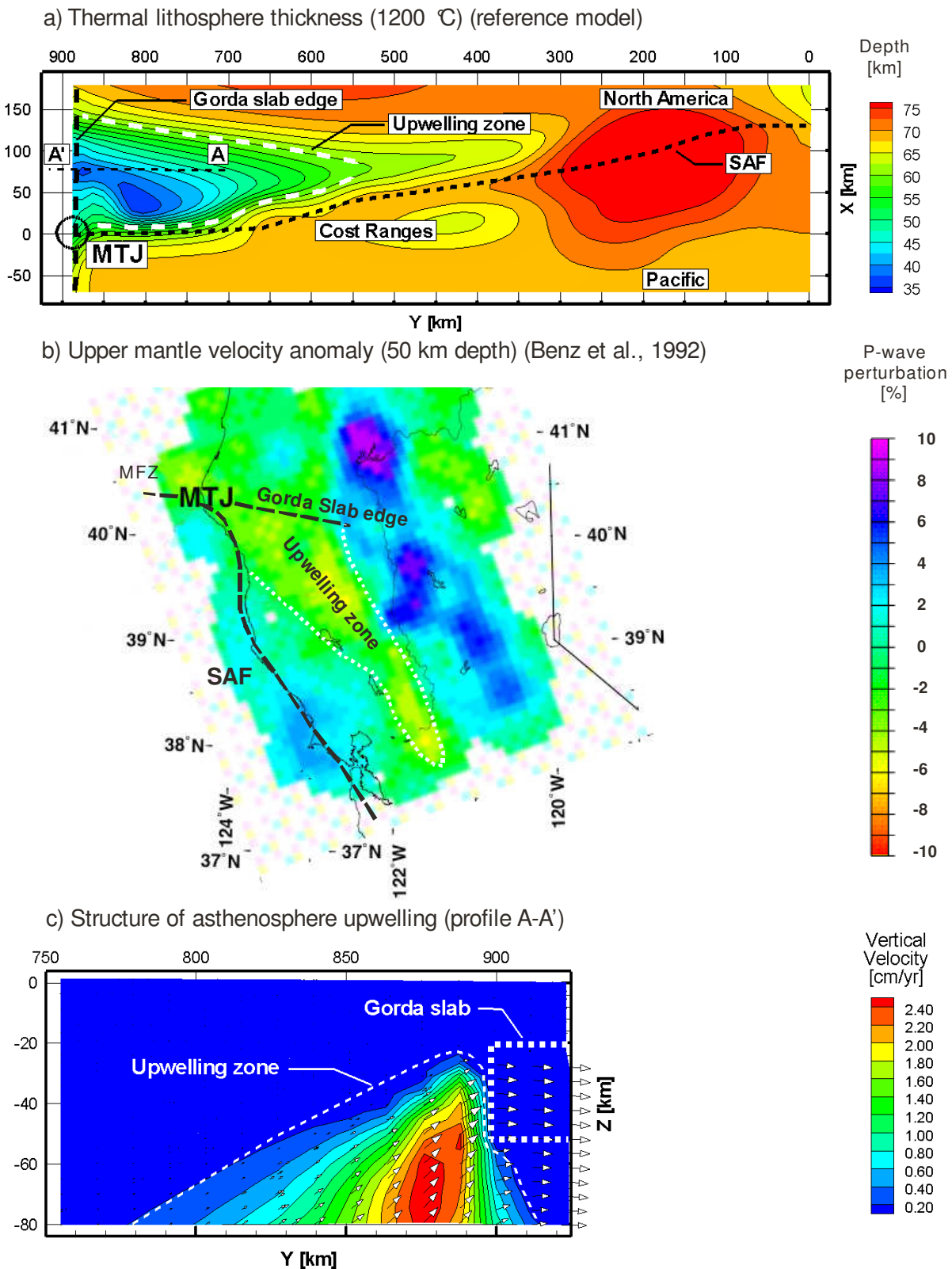


Figure 5.19 Character of the asthenosphere upwelling predicted by the reference model and interpreted from the P-wave seismic tomography (Benz et al., 1992). a) Depth of the thermal lithosphere in the reference model (depth map of the 1200 °C isosurface). Colours indicate depths in km. White dashed lines show the extents of the active upwelling zone. b) P-wave seismic tomography of Northern California upper mantle at 50 km depth (from Benz et al., 1992). Colours indicate P-wave velocity perturbations. Positive velocity anomaly is interpreted as the asthenospheric upwelling in the opening slab window. c) Distribution of vertical velocities in the upwelling zone from the reference model (profile A-A'). Colours show vertical velocities. Vectors indicate the material flow in the upwelling zone behind the Gorda slab (dashed lines).

is naturally maintained due to coupled solution of mass and momentum balance equations. Moreover, the accretion of the asthenosphere material to the Pacific and North America plates is also modeled here without additional assumptions, due to temperature dependent rheology and coupled thermomechanical solution.

The map of the thermal lithosphere thickness in the reference model is shown in Fig. 5.19a. As the bottom of the thermal lithosphere we assume the isosurface of the temperature 1200 °C. The model predicts a relatively confined zone of active asthenosphere upwelling, with the center locating about 100 km to the south of MTJ. This distance is in relatively good agreement with current position of the upwelling center inferred from the fluid sources within the upper mantle (Levander et al., 1998). The minimum thermal lithospheric thickness observed in the model is about 35 km near the upwelling center. The maximum thickness (75 km) is associated with a long-term transpression of the restraining bend in southern part of the model.

The style of material flow in the upwelling zone is shown in Fig. 5.19c. The upwelling zone inferred from the vertical velocities is even more confined than the zone of shallower thermal lithosphere. Almost entire asthenosphere material enters the slab window within the first 50 km from the southern edge of Gorda slab. The discrepancy between the locations of the thermal and mechanical lithosphere can be intuitively explained by upwelling, cooling and gradual accretion of the asthenosphere to the overlying continent.

Fig. 5.19b shows P-wave velocity structure beneath the northern California at 50 km depth (Benz et al., 1992). The low-velocity body beneath the Coast Ranges is interpreted as the zone of high temperature and partial melting within the mantle caused by asthenosphere upwelling. In general, the size and shape of the shallow lithosphere zone in the reference model is comparable with the configuration of the low-velocity body beneath the Coast Ranges. The western edge of the upwelling zone coincides with the active mantle plate boundary in our models.

Topography and Moho depth

The major factors controlling both the crustal thickness and the surface elevation on the time span of presented models are the tectonic shortening and extension. The influence of the long-term transpression near the restraining bend in the southern part of the model, which we interpret as a prototype of the Big Bend in the SAFS, is much more intensive than the uplift due to upwelling of the hot asthenosphere. Fig. 5.20a shows the map of the surface elevation in the reference model in the middle of the Bay Area accretion period (Fig. 5.17d). The elevation of the Coast Ranges varies within the range 0.5 - 1.5 km, which closely corresponds to the observed 1 km of average elevation in the Coast Ranges. The ranges of high mountains with 2.5-4.5 km elevation, coincident with surface trace of San Andreas Fault are related to the transpressional

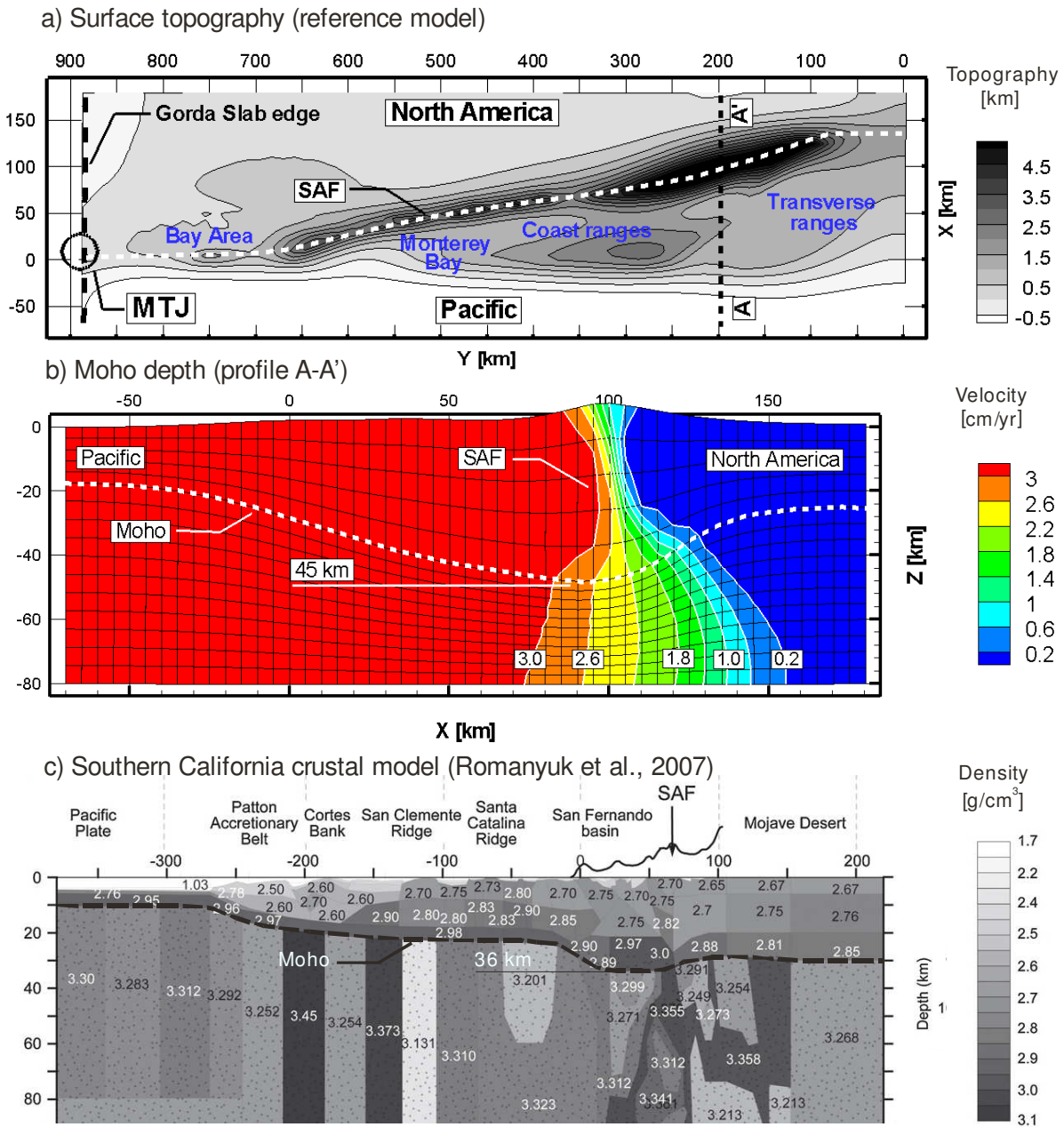


Figure 5.20 Influence of the active deformation processes on the crustal thickness and the surface topography. a) Surface topography from the reference model. Different shades of gray indicate the surface elevation. San Andreas Fault is shown with white dashed line. b) Configuration of the Moho discontinuity (white dashed lines) in the profile A-A'. Maximum Moho depth in the reference model is about 45 km. Colours and labeled lines show the distribution of the strike-slip velocity and configuration of the San Andreas Fault in the vicinity of large restraining bend. Note that the fault is tilted at depth. c) Detailed density transect for the crust and uppermost mantle near the Big Bend (southern California) obtained using a linear gravity inversion (Romanyuk et al., 2007). Maximum Moho depth in the transect is about 35-36 km.

character of the plate boundary. The model succeeds in predicting the low elevation near the Monterey bay, but fails in the Bay Area block. This is probably caused by insufficient resolution and inability to resolve distinct faults in the Hayward-Calaveras system.

The configuration of the Moho boundary and the surface elevation are shown in the vertical profile in Fig. 5.20b. The modeled thickness of crustal layer is compared against the observations

obtained using combined gravity-seismic inversion (Romanyuk et al., 2007) near the Big Bend in southern California (Fig. 5.20c). Both observations and the model results show dramatic dependence of the crustal thickness on the active tectonic shortening. However, the maximum Moho depth in the model is about 45 km compared to the 35-36 km thickness in the observations. The significant overestimation in both the surface elevation and, correspondingly, the Moho depth is likely related to insufficient model size in the strike-perpendicular direction. Due to zero assumed normal velocities on the Pacific and Great Valley sides of the model domain, the tectonic shortening related to transpression produces an exaggerated effect. The real tectonic system in southern California has an ability to compensate an excessive surface elevation by appropriate horizontal movements.

Heat flow

The presented models show that the surface heat flow is a relatively weak constraint to estimate the thermal state of the crust and mantle. The heat flow signals from the mantle can be relatively easily overprinted by the crustal thickening processes. Moreover, even different perturbations in the mantle can give fairly similar effect on the surface. For example, Goes et al. (1997) and Van Wijk et al. (2001) were not able to distinguish different model scenarios based solely on the heat flow observations. However certain effects, such as the upwelling of the asthenosphere in the slab window (Lachenbruch and Sass, 1980), are clearly visible in the surface heat flow. Fig. 5.21 shows the heat flow results for the reference model in the middle of the Bay Area accretion period (Fig. 5.17d). On the surface heat flow map (Fig. 5.21a) there are only minor thermal anomalies ($< 5 \text{ mWm}^{-2}$) related to the shear heating on the San Andreas plate boundary. Despite 100% assumed shear heating efficiency, the associated heat flow increase is very small, due to small effective friction. Therefore, the presented model overcomes the “heat flow paradox” (Lachenbruch and Sass, 1992) and explains absence of broad heat flow anomaly without excessive hydrothermal activity.

In, general the observed heat flow correlates with observations (Lachenbruch and Sass, 1980) relatively well, in both strike-parallel (Fig. 5.21b) and typical strike-perpendicular profiles (Fig. 5.21c). The averaged heat flow over the Coast Ranges initially increases up to 83 mWm^{-2} at 200 km distance from the Mendocino Triple Junction. Further south it decreases to the average value of 75 mWm^{-2} at 300 km distance. Fig. 5.21d shows the dramatic influence of the long-term transpression on the surface heat flow. In the southern part of the model near the restraining bend, the average heat flow value becomes locally overprinted by rapid crustal thickening. The model heat flow in this region is lower than observed due to the overestimation in the model of the transpressional crustal thickening (see above).

Surface heat flow (reference model)

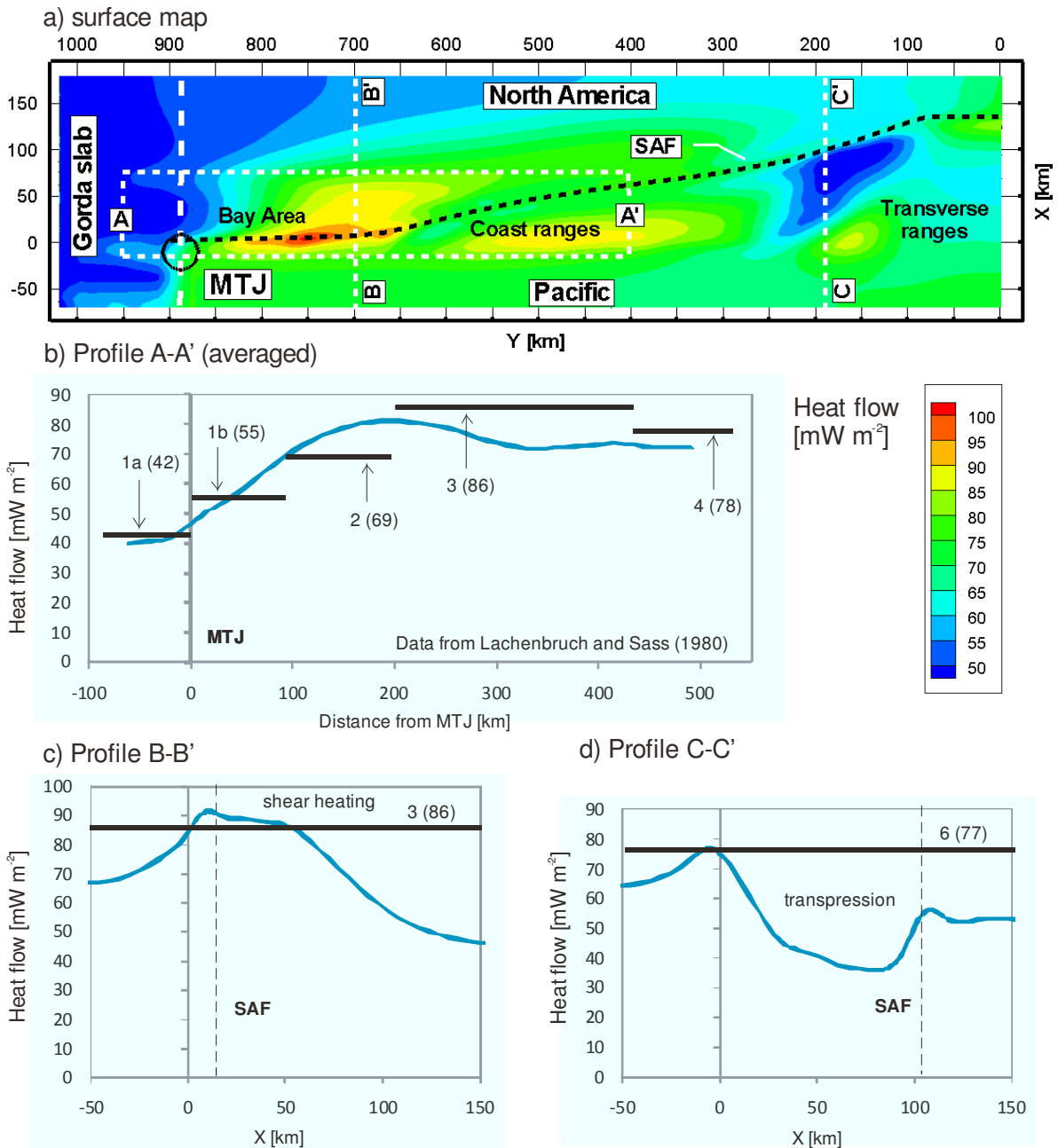


Figure 5.21 Comparison of the calculated and observed surface heat flow. a) Surface heat flow map in the reference model. Colours indicate the heat flow magnitudes. Black dashed line show the trace of the San Andreas Fault. b) Averaged heat flow along the strike-parallel profile A-A'. Averaging is performed over the width of rectangular area shown in panel a). Panels c) and d) show the heat flow distribution along the strike-perpendicular profiles B-B', and C-C', respectively. Black horizontal bars show the averaged heat flow magnitudes for different regions along the San Andreas Fault (Lachenbruch and Sass, 1980).

Weak vs. strong rheology on the fault

To prevent the broad heat flow anomaly on the SAFS, the effective friction on the faults in our models must be low. Moreover we have constrained the friction coefficient to specific low value since it also influences the generation and accretion of the terranes. In this subsection we provide

“Weak” vs. “Strong” faults comparison

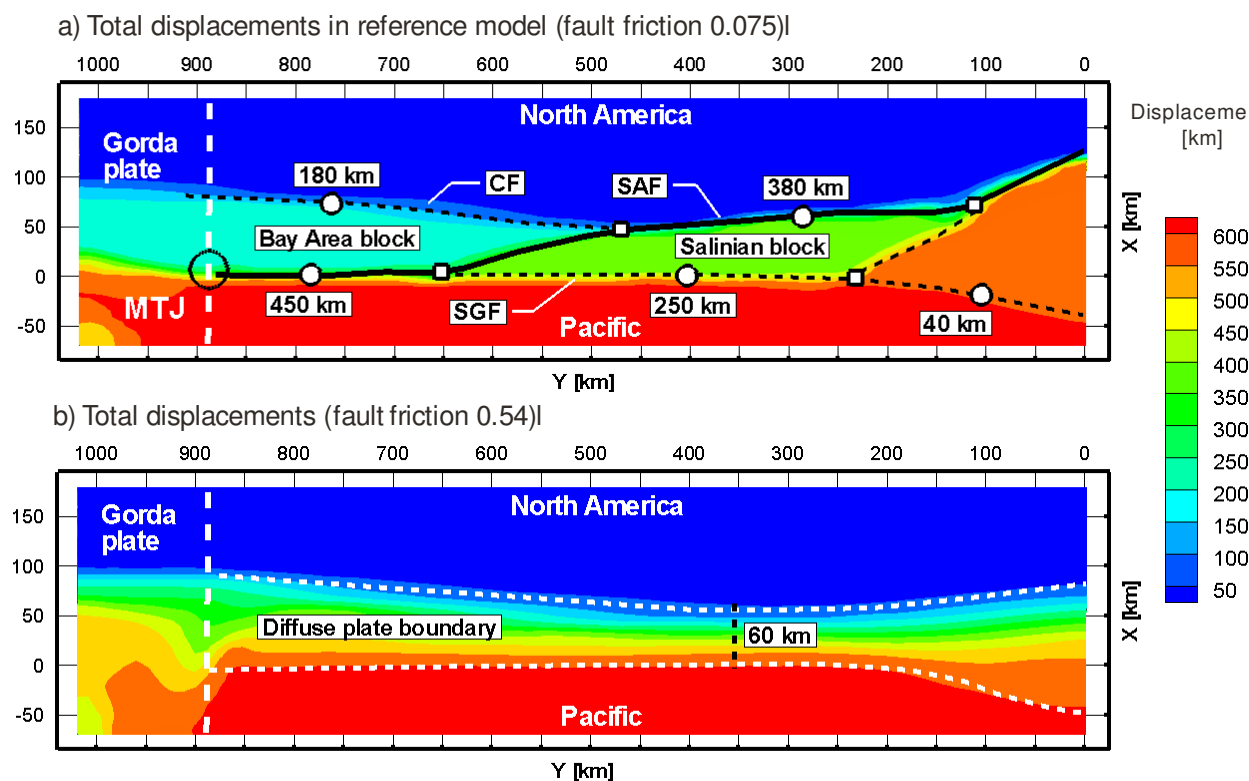


Figure 5.22 Comparison between the models with the weak and moderately strong fault rheologies. a) Fault map of the reference model (Figure 5.18a). Distinct terrains bounded by the sharp faults are marked with different colors. b) Strike-slip displacements in the model equivalent to the reference model with the effective friction on the faults 0.54. As a result of a long term evolution (>10Ma) this model predicts formation of a diffuse 60 km wide shear zone coincident with a zone of minimal lithospheric strength. The model with high effective friction fails to predict the distinct terrains and faults.

additional independent constrain by running the model equivalent to the reference with the friction on the fault 0.54, which is equal to just 90% of standard value. Fig. 5.22 shows the comparison of the strike-slip displacements between the models with weak and strong fault rheology. The model with high friction (Fig. 5.22b) demonstrates that a lack of sufficient slip weakening fails to predict the distinct zones of strain localization in the brittle crust. Instead of this, the model predicts broad zone of distributed shear (60 km wide), which follows the zone of minimal integrated lithospheric strength. Our reference model for the northern California and Bay Area region requires the friction coefficient on major faults to be about 0.075, which is far less than typical values 0.6 – 0.8 obtained by variety of borehole stress measurements (e.g. Zoback and Healy, 1992) and laboratory data (e.g. Byerlee, 1978). Our forward modeling results, therefore, contribute to the long-term debate over the weakness of major faults in the SAFS (e.g. Zoback, 2000; Scholz, 2000) among the other similar contributions (e.g. Bird and Kong, 1994; Carena and Moder, 2008). We side with a “weak fault” concept, and favour importance of the slip-related weakening.

Chapter 6

Summary and discussion

6.1 Summary

The major aim of the presented work is comprised basically of two interconnected tasks. The first task is the development and testing of the research Finite Element code with sufficiently advanced facilities to perform the three-dimensional geological time-scale simulations of lithospheric deformation. The second task consists in application of the tool developed to the Neogene deformations of the crust and the mantle along the San Andreas Fault System in Central and northern California.

The result of the first task was a new numerical modeling tool (SLIM3D), which was described in detail and published in a peer-reviewed paper (Popov and Sobolev, 2008). In the framework of this thesis, the description of the methodology is presented in Chapter 4. The general conclusions and outlook related to the method development are summarised in section 4.5. In the context of the model results presented here, we can additionally conclude that the developed code can be successfully applied to complicated tectonic problem including large 3D compositional, thermal and mechanical heterogeneities in the crust and mantle and complicated non-linear temperature-, stress- and strain-dependent elasto-visco-plastic rheology.

The results of the second task are presented and discussed in the Chapter 5. The conclusions derived from the advanced three-dimensional models of Neogene tectonic evolution of the San Andreas Fault System can be briefly summarized as follows:

- 1) The models show the ability to localize the deformation into narrow shear zones (faults).
- 2) The strain localization occurs in the zones of minimal integrated lithospheric strength.
- 3) The coupled thermo-mechanical deformation on the SAFS produces temperature perturbations in both crust and mantle. The effective viscosity decreases locally in response to temperature increase, making the fault zone generally weaker.
- 4) The relative importance of shear heating is higher in the mantle than in the crust, due to low temperatures in the brittle part and temperature-independence of brittle rheology. With low

effective friction on the fault, the shear heating in brittle part is negligible.

- 5) The eastward migration of the SAFS (e.g. Furlong et al., 1989) is significantly suppressed by the transpression resulting from restraining fault steps.
- 6) Together, the shear heating of the mantle roots beneath the mature faults, and the step-induced transpression, make the cooling and accretion of asthenosphere material in the slab window (Furlong et al., 1989) incapable to explain the eastward migration of the SAFS.
- 7) The stalled microplates in the fossil Farallon subduction zone (e.g. Monterrey microplate) (Nicholson et al., 1994) can cause instant inland migration of the SAFS plate boundary due to higher integrated lithospheric strength in the zone occupied by the microplate.
- 8) The accretion of North America terranes towards the Pacific plate is greatly intensified by presence of the microplate.
- 9) The pattern of faults in the SAFS can be better reproduced in the model that includes strong Salinian crustal block (e.g. Irwin, 1990).
- 10) The effective friction on the mature strike-slip faults in the SAFS must be low. The reference model predicts the effective friction to be about 0.075. Models with high friction on the faults are unable to form discrete fault zones and crustal terranes.
- 11) The location of the asthenosphere upwelling center inferred from the fluid sources within the upper mantle (Levander et al., 1998) is confirmed by the reference model. The model also shows the discrepancy between the thermal and mechanical lithosphere. The actual upwelling zone is much smaller than the zone of elevated temperature. This is attributed to time delays due to conductive cooling.
- 12) Heat flow is a relatively weak constraint to estimate the thermal state of the crust and mantle. Models demonstrate that the regional heat flow can be locally overprinted by rapid crustal thickening. The averaged heat flow values in the Coast Ranges (Lachenbruch and Sass, 1980) are in good agreement with the model results.

6.2 Discussion

Advantages of the models

The models presented in this thesis have numerous substantial improvements with respect to all previous similar studies (Furlong, 1984; 1993; Furlong et al., 1989; Goes et al., 1997, van Wijk et al., 2001; Chery et al, 2001; Sobolev et al, 2003). This work presents the first attempt to consider the long-term tectonic evolution of the SAFS in three-dimensional setup with full thermo-mechanical coupling. In general, the models include relatively sophisticated effects, such

as nonlinear rheology, non-standard boundary conditions, etc. However the actual advantage of the models is that they are based on a relatively simple idealization. Despite this fact, the models are capable of producing meaningful results, which help to study complex deformation processes at the SAFS.

The major advantages of the presented models with respect to previous studies can be combined in the following list:

- 1) Full 3D formulation with thermo-mechanical coupling
- 2) Ability to resolve transpression on a fault bend
- 3) Presence of along-strike variations of the entire fields
- 4) Gorda slab subduction idealization
- 5) 3D material flow in the slab window (asthenosphere upwelling)
- 6) Presence of simplified microplate (Monterey microplate)
- 7) Ability to model deformation including the crustal terranes (Salinian block)

The results of the models are in good agreement with the observations. In particular, the heat flow in the reference model matches the thermal pulse of the slab window (Fig. 5.21b). Both the structure of asthenosphere upwelling (Fig. 5.19), and the terrane configuration in central and northern California (Fig. 5.18) correlate with the observations reasonably well. One of the most important results of the reference model is the independent constrain for the magnitude of the effective friction on the faults (Fig. 5.22).

Weaknesses of the models

Modeling is an ongoing effort. Despite relatively high complexity of the presented three-dimensional models, they can only be used for qualitative estimation of the long-term tectonic evolution. For example, the influence of the effective friction on major faults on the terrane accretion rate can be analysed. However, it is hard to expect the exact one-to-one correlation of the modeling results with the present-day tectonic configuration, because the models still contain numerous simplifications and assumptions both in the conceptual part and in the initial conditions.

The present simplifications cause inability to reproduce certain effects in the model. For example the time span of the Bay Area block accretion is about 4 Myr in the model, vs. at least 8 Myr according to reconstructions (McLaughlin et al., 1996). Two explanations have been proposed in section 5.2. Either the absence of terranes to the north of the model domain may stimulate the terrane capture in the model, or the friction may increase to the east with increasing distance from the fossil Farallon trench, and decreasing fluid concentration.

Another problem is related to systematic overestimation of the surface elevation and the

Moho depth near the fault bend in the southern part of the domain, which is interpreted as Big Bend -like structure. The discrepancy range is between 5.0 km elevation in the model vs. 3.0-3.5 km in San Bernardino and San Gabriel Mountains, and 45 km Moho depth in the model vs. 35-36 km in observations (e.g. Romanyuk et al., 2007). These discrepancies are likely caused by insufficient size of the model in the east-west direction and inability to compensate of the surface elevation by horizontal movements of eastern and western boundaries (see section 5.2 for more details).

Among other disadvantages of the models the following can be mentioned:

- 1) Single-layer crustal model
- 2) Insufficient size of the model domain
- 3) Prescribed velocities on the boundaries
- 4) Kinematically driven Gorda slab subduction (dynamic model is the alternative)
- 5) Lack of coupling with the mantle convection (no basal tractions)

6.3 Outlook

Apart from the planned advancements in the numerical method, which are described in the section 4.5, it would be interesting to substantially increase the resolution, the size and the time span of the presented models. The valuable alternative would be switching to the adaptive discretization to resolve the fine structures, which are unavailable in the present formulation. The formulation of the crustal and thermal models requires systematic improvement, as well as Gorda slab idealization. Potentially it would be also interesting to embed the dynamic subduction and spreading ridge in the current modeling framework. In general this improvement will allow direct simulation of the microplate capture process, which is beyond the scope of presented work. In the global perspective, the entire sequence of microplate capture events (Monterey, Arguello, Guadalupe, and Magdalena) (Lonsdale, 1991) can be modeled. Another important subject for the future detailed 3D modeling could be the influence of Mendocino Triple Junction on the deformation of the North American crust. Fulfillment of this task would require enlargement of the model domain to the north-west, as well as significant improvement of mesh resolution in the crustal layer.

Appendix

Table A.1
Rheological and thermal parameters used in the 3D modeling

Parameter	Meaning	Dimension	Crust	Mantle
ρ	Density	g cm^{-3}	2.85	3.3
K	Bulk modulus	kbar	630	1220
G	Shear modulus	kbar	400	740
$\log B_N$	Dislocation creep constant	$\text{Pa}^{-n} \text{s}^{-1}$	-21.05	-16.3
H_N	Dislocation creep enthalpy	kJ mol^{-1}	445	535
n	Dislocation creep exponent	–	4.2	3.5
μ	Friction coefficient	–	0.6	0.6
c	Cohesion	MPa	20	20
α	Thermal expansivity	K^{-1}	$2.7 \cdot 10^{-5}$	$3.0 \cdot 10^{-5}$
C_p	Specific heat	$\text{kJ kg}^{-1} \text{K}^{-1}$	1.2	1.2
λ	Thermal conductivity	$\text{W m}^{-1} \text{K}^{-1}$	2.5	3.3
A	Radiogenic heat production	nW kg^{-1}	0.5	0.0

References

- Adams, M.F., 2000. Parallel multigrid solvers for 3D unstructured finite element problems in large deformation elasticity and plasticity. *Int. J. Numer. Meth. Engng.*, 48: 1241-1262.
- Alejano, L.R., Alonso, E., 2005. Considerations of the dilatancy angle in rocks and rock masses. *Int. J. Rock Mech. Min. Sci.*, 42: 481-507.
- Amestoy, P.R., Duff, I.S., L'Excellent, J.-Y., 2000. Multifrontal parallel distributed symmetric and unsymmetric solvers. *Comput. Methods Appl. Mech. Engrg.*, 184: 501-520.
- Argus, D.F., Gordon, R.G., 2001. Present tectonic motion across the Coast Ranges and San Andreas fault system in central California. *Geol. Soc. Am. Bull.*, 113: 1580–1592.
- Arthur, J.R.F., Dunstan, T., Al-Ani, Q.A.J., Assadi, A., 1977. Plastic deformation and failure of granular media. *Geotechnique*, 27: 53-74.
- Atwater, T., 1989. Plate tectonic history of the northeast Pacific and western North America. In E.L. Winterer, D.M. Hussong, and R.W. Decker (editors). *The Geology of North America*, vol. N, The Eastern Pacific Ocean and Hawaii. *Geol. Soc. Am., Boulder, Colo.*, 21–72.
- Atwater, T., Severinghaus, J., 1989. Tectonic map of the northeast Pacific Ocean. In E.L. Winterer, D.M. Hussong, and R.W. Decker (editors). *The Geology of North America*, vol. N, The Eastern Pacific Ocean and Hawaii. *Geol. Soc. Am., Boulder, Colo.*, 15–20.
- Atwater, T., Stock, J., 1998. Pacific–North America plate tectonics of the Neogene southwestern United States: An update. *Int. Geol. Rev.*, 40: 375–402.
- Atwater, T.M., 1970. Implications of plate tectonics for the Cenozoic tectonic evolution of western North America. *Geol. Soc. Am. Bull.*, 81: 3513–3536.
- Atwater, T.M., Molnar, P., 1973. Relative motions of the Pacific and North American plates deduced from seafloor spreading in the Atlantic, Indian, and South Pacific Oceans. In R.L. Kovach and A. Nur (editors). *Proceedings of the Conference on Tectonic Problems of the San Andreas Fault System*. Stanford Univ. Publ. *Geol. Sci.*, 13: 136–148.
- Babeyko, A.Yu., Sobolev, S.V., Trumbull, R.B., Oncken, O., L.L. Lavier, 2002. Numerical models of crustal scale convection and partial melting beneath the Altiplano-Puna Plateau. *Earth Planet. Sci. Lett.*, 199: 373-388.
- Babeyko, A.Yu., Sobolev, S.V., Trumbull, R.B., Oncken, O., Lavier, L.L., 2002. Numerical models of crustal scale convection and partial melting beneath the Altiplano-Puna plateau. *Earth Planet. Sci. Lett.*, 199: 373-388.
- Babuska, I., 1973. The finite element method with Lagrangian multipliers. *Num. Math.*, 20: 179-192.
- Bailey, R.C., 2006. Large time step numerical modelling of the flow of Maxwell materials. *Geophys. J. Int.*, 164: 460-466.
- Bathe, K.J., Ramm, E., Wilson, E.L., 1975. Finite element formulations for large deformation dynamic analysis. *Int. J. Numer. Meth. Engng.*, 9: 353-386.
- Beaudoin, B.C., Godfrey, N., Klemperer, S.L., Lendl, C., Trehu, A.M., Henstock, T.J., Levander, A., Holl, J.E., Meltzer, A.S., Luetgert, J.H., Mooney, W.D., 1996. The transition from slab to slabless: Results from the 1993 Mendocino Triple Junction seismic experiment. *Geology*, 24: 195–199.
- Becker, T., O'Connell, R., 2001. Predicting plate velocities with mantle circulation models. *Geochem. Geophys. Geosyst.*, 2(12)., doi: 10.1029/2001GC000171.
- Belytschko, T., Liu, W.K., Moran, B., 2000. *Nonlinear Finite Elements for Continua and Structures*. John Wiley & Sons, Chichester.
- Belytschko, T., Tabbara, M., 1993. H-adaptive finite element methods for dynamic problems, with emphasis on localization. *Int. J. Numer. Meth. Engng.*, 36: 4245-4265.
- Bennett, R.A., Davis, J.L., Wernicke, B.P., 1999. Present-day pattern of Cordilleran deformation in the western United States. *Geology*, 27: 371-374.
- Benz, H.M., Zandt, G., Oppenheimer, D.H., 1992. Lithospheric structure of northern California from teleseismic images of the upper mantle. *J. Geophys. Res.*, 97: 4791–4807.
- Bercovici, D., 1993. A simple model of plate generation from mantle flow. *Geophys. J. Int.*, 114: 635-650.
- Bird, P., 1989. New finite element techniques for modeling deformation histories of continents with stratified temperature-dependent rheologies. *J. Geophys. Res.*, 94: 3967-3990.

- Bird, P., Kong, X., 1994. Computer simulations of California tectonics confirm very low strength of major faults. *Geol. Soc. Am. Bull.*, 106: 159-174.
- Bird, P., Piper, K., 1980. Plane-stress finite-element models of tectonic flow in southern California. *Phys. Earth Planet. Inter.*, 21: 158-175.
- Blake, A., 1985. *Handbook of Mechanics, Materials, and Structures*. Wiley, New York.
- Blanpied, M.L., Lockner, D.A., Byerlee, J.D., 1991. Fault stability inferred from granite sliding experiments at hydrothermal conditions. *Geophys. Res. Lett.*, 18: 609-612.
- Bohannon, R.G., Parsons, T., 1995. Tectonic implications of post-30 Ma Pacific and North American relative plate motions. *Geol. Soc. Am. Bull.*, 107: 937-959.
- Bonet, J., Marriott, H., Hassan, O., 2001. An averaged nodal deformation gradient linear tetrahedral element for large strain explicit dynamic applications. *Commun. Numer. Meth. Engng.*, 17: 551-561.
- Bonet, J., Wood, R.D., 1997. *Nonlinear Continuum Mechanics for Finite Element Analysis*. Cambridge University Press, Cambridge.
- Borja, R.I., Aydin, A., 2004. Computational modeling of deformation bands in granular media, I: Geological and mathematical framework. *Comput. Methods Appl. Mech. Engng.*, 193: 2667-2698.
- Borja, R.I., Sama, K.M., Sanz, P.F., 2003. On the numerical integration of three-invariant elastoplastic constitutive models. *Comput. Methods Appl. Mech. Engng.*, 192: 1227-1258.
- Brace, F.W., Kohlstedt, D.L., 1980. Limits on lithospheric stress imposed by laboratory experiments. *J. Geophys. Res.*, 50: 6248-6252.
- Braun, J., Beaumont, C., 1995. Three-dimensional numerical experiments of strain partitioning at oblique plate boundaries: implications for contrasting tectonic styles in the southern Coast Ranges, California and central South Island, New Zealand. *J. Geophys. Res.*, 100: 18059-18074.
- Braun, J., Sambridge, M., 1994. Dynamical Lagrangian Remeshing (DLR): A new algorithm for solving large strain deformation problems and its application to fault propagation folding. *Earth Planet. Sci. Lett.*, 124: 211-220.
- Braun, J., Thieulot, C., Fullsack, P., DeKool, M., Beaumont, C., Huisman, R., 2008. DOUAR: A new three-dimensional creeping flow numerical model for the solution of geological problems. *Phys. Earth Planet. Inter.*, doi:10.1016/j.pepi.2008.05.003.
- Brezzi, F., 1974. On the existence, uniqueness and approximation of saddle-point problems arising from Lagrange multipliers. *RAIRO*, 8: 129-151.
- Brown, R.D. J., 1990. Quaternary deformation. In R.E. Wallace (editor). *The San Andreas Fault System, California*, U.S. Geol. Surv. Prof. Paper., 1515: 83-114.
- Brune, J.N., Brown, S., Johnson, P.A., 1993. Rupture mechanism and interface separation in foam rubber models of earthquakes: A possible solution to the heat flow paradox and the paradox of large overthrusts. *Tectonophysics*, 218: 59-67.
- Brune, J.N., Henyey, T.L., Roy, R.F., 1969. Heat flow, stress, and rate of slip along the San Andreas fault, California. *J. Geophys. Res.*, 74: 3821-3827
- Buiter, S.J.H., Babeyko, A. Yu., Ellis, S., Gerya, T.V., Kaus, B.J.P., Kellner, A., Schreurs, G., Yamada, Y., 2006. The Numerical Sandbox: Comparison of model results for a shortening and an extension experiment, analogue and numerical modelling of crustal-scale processes. Geological Society, London, Special Publication, 253: 29-64.
- Bullard, E. C., Everett, J. E. Smith, A. G., 1965. The fit of the continents around the Atlantic. *Phil. Trans. R. Soc., Math. Phys. Sci.*, A258: 41-51.
- Bunge, H.-P., Baumgardner, J.R., 1995. Mantle convection modeling on parallel virtual machines. *Comput. Phys.*, 9: 207-215.
- Byerlee, J.D., 1978. Friction of rocks. *Pure Appl. Geophys.*, 116: 615-629.
- Cande, S.C., Stock, J.M., Mueller, R.D., Ishihara, T., 2000. Cenozoic motion between East and West Antarctica. *Nature*, 404: 145-150.
- Carena, S., Moder, C., 2008. Fault strength in California and western Nevada. *Geophys. Res. Abs.*, 10: EGU2008-A-01703.
- Castillo, D.A., Ellsworth, W.L., 1993. Seismotectonics of the San Andreas fault system between Point Arena and Cape Mendocino in northern California: Implication for the development and evolution of a young transform. *J. Geophys. Res.*, 98: 6543-6560.
- Chaytor, J.D., Goldfinger, C., Dziak, R.P., Fox, C.G., 2004. Active deformation of the Gorda plate: Constraining deformation models with new geophysical data. *Geology*; 32: 353-356.
- Chery, J., Zoback, M.D., Hassani, R. 2001. An integrated mechanical model of the San Andreas fault in central and northern California. *J. Geophys. Res.*, 106: 22051-22066.
- Chester, F.M., 1995. A rheologic model for wet crust applied to strike-slip faults. *J. Geophys. Res.*, 100: 13033-13044.
- Christensen, U., Harder, H., 1991. 3-D Convection with variable viscosity. *Geophys. J. Int.* 104: 213-220.
- Coulomb, C.A., 1773. Test on the applications of the rules of maxima and minima to some problems of statics related to architecture (in French). *Mem. de Math. et de Phys.*, 7: 343-382.
- Crisfield, M.A., 1983. An arc-length method including line searches and accelerations. *Int. J. Numer. Meth. Engng.*, 19: 1269-1289.

- DeMets, C., Gordon, R.G., Argus, D.F., Stein, S. 1990. Current plate motions. *Geophys. J. Int.*, 101: 425–478.
- DeMets, C., Gordon, R.G., Argus, D.F., Stein, S., 1994. Effect of recent revisions to the geomagnetic reversal time scale on estimate of current plate motions. *Geophys. Res. Lett.*, 21: 2191-2194.
- Dickinson, W.R., 1997. Tectonic implications of Cenozoic volcanism in coastal California. *Geol. Soc. Am. Bull.*, 109: 936–954.
- Dickinson, W.R., Snyder, W.S., 1979. Geometry of subducted slabs related to San Andreas transform. *J. Geology*, 87: 609–627.
- Dixon, T.H., Robaudo, S., Lee, J., Reheis, M.C., 1995. Constraints on present-day Basin and Range deformation from space geodesy. *Tectonics*, 14: 755–772.
- Ekevid, T., Kettil, P., Wiberg, N.-E., 2004. Adaptive multigrid for finite element computations in plasticity. *Comput. Struct.*, 82: 2413–2424.
- Farhat, C., Lesoinne, M., Pierson, K., 2000. A scalable dual-primal domain decomposition method. *Numer. Linear Algebra Appl.*, 7: 687-714.
- Fish, J., Shek, K., 1999. Computational aspects of incrementally objective algorithms for large deformation plasticity. *Int. J. Numer. Meth. Engng.*, 44: 839-851.
- Flanagan, D.P., Belytschko, T., 1981. A uniform strain hexahedron and quadrilateral with orthogonal hourglass control. *Int. J. Numer. Meth. Engng.*, 17: 679-706.
- Flesch, L.M., Holt, W.E., Haines, A.J., Shen-Tu, B., 2000. Dynamics of the Pacific-North American plate boundary in the western United States. *Science*, 287: 834–836.
- Frey Mueller, J., Murray, M., Segall, P., Castillo, D., 1999. Kinematics of the Pacific-North America plate boundary zone, northern CA. *J. Geophys. Res.*, 104: 7419-7441.
- Fuchs, K., Muller, B., 2001. World stress map of the Earth: A key to tectonic processes and technological applications. *Naturwissenschaften*, 88: 357–371.
- Fullsack, P., 1995. An arbitrary Lagrangian-Eulerian formulation for creeping flows and its application in tectonic models. *Geophys. J. Int.*, 120: 1-23.
- Furlong, K.P., 1984. Lithospheric behavior with triple junction migration: An example based on the Mendocino triple junction. *Phys. Earth Planet. Inter.*, 36: 213–223.
- Furlong, K.P., 1993. Thermal-rheologic evolution of the upper mantle and the development of the San Andreas fault system. *Tectonophysics*, 223: 149–164.
- Furlong, K.P., Govers, R., 1999. Ephemeral crustal thickening at a triple junction: the Mendocino crustal conveyor. *Geology*, 27: 127–130.
- Furlong, K.P., Hugo, W.D., Zandt, G., 1989. Geometry and evolution of the San Andreas Fault zone in Northern California. *J. Geophys. Res.*, 94: 3100–3110.
- Furlong, K.P., Schwartz, S.Y., 2004. Influence of the Mendocino triple junction on the tectonics of coastal California. *Annu. Rev. Earth Planet. Sci.*, 32: 403–433.
- Geers, M.G.D., 1999. Enhanced solution control for physically and geometrically non-linear problems. Part I-the subplane control approach. *Int. J. Numer. Meth. Engng.*, 46: 177-204.
- George, A., 1973. Nested dissection of a regular finite element mesh. *SIAM J. Numer. Analysis*, 10: 345-363.
- Gerya, T.V., Yuen, D.A., 2007. Robust characteristics method for modelling multiphase visco-elasto-plastic thermo-mechanical problems. *Phys. Earth Planet. Inter.*, 163: 83-105.
- Goes, S., Govers, R., Schwartz, S., Furlong, K., 1997. Three-dimensional thermal modeling for the Mendocino triple junction area. *Earth Planet. Sci. Lett.*, 148: 45-57.
- Govers, R., Wortel, M.J.R., 1995. Extension of stable continental lithosphere and the initiation of lithospheric scale faults. *Tectonics*, 14: 1041-1055.
- Gupta, A., Karypis, G., Kumar, V., 1997. Highly scalable parallel algorithms for sparse matrix factorization. *IEEE Trans. Parallel Distrib. Syst.* 8: 502-520.
- Guzofski, C.A., Furlong, K.P., 2002. Migration of the Mendocino triple junction and ephemeral crustal deformation: Implications for California Coast Range heat flow. *Geophys. Res. Lett.*, 29: 12-1 - 12-4.
- Harlow, F., Welsh, J., 1965. Numerical calculation of time-dependant viscous incompressible flow of fluid with free surface. *Phys. Fluids*, 8: 2182-2189.
- Hartz, E.H., Podladchikov, Y.Y., 2008. Toasting the Jelly Sandwich: The effect of shear heating on lithospheric geotherms and strength. *Geology*, 36: 331-334.
- Henstock, T.J., Levander, A., Hole, J.A., 1997. Deformation in the lower crust of the San Andreas Fault System in northern California. *Science*, 278: 650–653.
- Hestenes, M.R., Stiefel, E., 1952. Methods of conjugate gradients for solving linear systems. *J. Res. Nat. Bur. Stand.*, 49: 409-436.
- Hickman, S., Zoback, M. D., 2004. Stress orientations and magnitudes in the SAFOD Pilot Hole. *Geophys. Res. Lett.*, 31, L15S12, doi:10.1029/2004GL020043.
- Hirt, C.W., Amsden, A.A., Cook, J.L., 1974. An arbitrary Lagrangian-Eulerian computing method for all flow speeds. *J. Comput. Phys.*, 14: 227-253.
- Hughes, T.J.R., 1987. *The Finite Element Method*, Prentice-Hall, Englewood-Cliffs, New Jersey.
- Hughes, T.J.R., Winget, J., 1980. Finite rotation effects in numerical integration of rate constitutive equations arising in large-deformation analysis. *Int. J. Numer. Meth. Engng.*, 15: 1862-1867.

- Humphreys, E. D., Coblenz, D.D., 2007. North American dynamics and western U.S. tectonics. *Rev. Geophys.*, 45: RG3001, doi:10.1029/2005RG000181.
- Humphreys, E.D., Dueker, K.G., 1994. Western U.S. upper mantle structure. *J. Geophys. Res.*, 99: 9615–9634.
- Ingersoll, R.V., 1982. Triple-junction instability as cause for late Cenozoic extension and fragmentation of the western United States. *Geology*, 10: 621–624.
- Irwin, W.P., 1990. Quaternary deformation. In R.E. Wallace (editor). *The San Andreas Fault System, California*, U.S. Geol. Surv. Prof. Paper., 1515: 61–80.
- Janecke, S.U., Evans, J.P., 1988. Feldspar-influenced rock rheologies. *Geology*, 16: 1064–1067.
- Kameyama, M., Yuen, D.A., Karato, S.-I., 1999. Thermal-mechanical effects of low-temperature plasticity (the Peierls mechanism) on the deformation of a viscoelastic shear zone. *Earth Planet. Sci. Lett.* 168: 159-172.
- Karato, S.-I., Riedel, M.R., Yuen, D.A., 2001. Rheological structure and deformation of subducted slabs in the mantle transition zone: implications for mantle circulation and deep earthquakes. *Phys. Earth Planet. Inter.*, 127: 83-108.
- Kaus, B.J.P., Podladchikov, Y.Y., 2006. Initiation of localized shear zones in viscoelastoplastic rocks. *J. Geophys. Res.*, 111, B04412, doi: 10.1029/2005JB003652.
- Lachenbruch, A.H., 1980. Frictional heating, fluid pressure, and the resistance to fault motion. *J. Geophys. Res.*, 85: 6097–6112.
- Lachenbruch, A.H., Sass, J.H., 1980. Heat flow and energetics of the San Andreas fault zone. *J. Geophys. Res.*, 85: 6185–6223.
- Lachenbruch, A.H., Sass, J.H., 1992. Heat flow from Cajon Pass, fault strength and tectonic implications. *J. Geophys. Res.*, 97: 4995–5015.
- Ladyzhenskaya, O.A., 1969. *The Mathematical Theory of Viscous Incompressible Flow*, 2nd edn. Gordon & Breach, New York.
- Larsson, R., Runesson, K., 1996. Implicit integration and consistent linearization for yield criteria of the Mohr-Coulomb type. *Mech. Cohes.-Frict. Mat.*, 1: 367-383.
- Le Pichon, X., 1968. Sea-floor spreading and continental drift, *J. Geophys. Res.*, 73: 3661-3697.
- Levander, A., Henstock, T.J., Meltzer, A. S., Beaudoin, B.C., Trehu, A.M., Klemperer, S. L., 1998. Fluids in the lower crust following Mendocino triple junction migration: Active basaltic intrusion?. *Geology*, 26: 171-174.
- Li, Q., Liu, M., 2006. Geometrical impact of the San Andreas Fault on stress and seismicity in California. *Geophys. Res. Lett.*, 33: L08302, doi:10.1029/2005GL025661.
- Li, X., Yuan, X., Kind, R., 2007. The lithosphere–asthenosphere boundary beneath the western United States. *Geophys. J. Int.*, 170: 700–710.
- Liu, M., Yang, Y., Li, Q., Zhang, H., 2007. Parallel computing of multi-scale continental deformation in the Western United States: Preliminary results. *Phys. Earth Planet. Inter.*, 163: 35-51.
- Liu, W.K., Guo, Y., Tang, S., Belytschko, T., 1998. A multiple-quadrature eight-node hexahedral finite element for large deformation elastoplastic analysis. *Comput. Methods Appl. Mech. Engrg.*, 154: 69-132.
- Liu, Z., Bird, P., 2002. North America Plate is driven westward by lower mantle flow. *Geophys. Res. Lett.*, 29: 2164, doi:10.1029/2002GL016002.
- Lonsdale, P., 1991. Structural patterns of the Pacific floor offshore of Peninsular California. In J.P. Dauphin, and B.T. Simoneit (editors). *Gulf and Peninsula Province of the Californias*. Am. Assoc. Petrol. Geol. Mem., 47: 87–125.
- Luyendyk, B.P., 1991. A model for Neogene crustal rotations, transtension, and transpression in southern California. *Geol. Soc. Am. Bull.*, 103, 1528–1536.
- Luyendyk, B.P., Kamerling, M.J., Terres, R.R., 1980. Geometric model for Neogene crustal rotations in southern California. *Geol. Soc. Am. Bull.*, 91: 211–217.
- Malkus, D.S., Hughes, T.J.R., 1978. Mixed finite element methods - reduced and selective integration techniques: a unification of concept. *Comput. Methods Appl. Mech. Engrg.*, 15: 63-81.
- McKenzie, D.P., Parker, R.L., 1967. The North Pacific: an example of tectonics on a sphere. *Nature* 216: 1276 – 1280.
- McLaughlin, R.J., Sliter, W.V., Sorg, D.H., Russell, P.C., Sarna-Wojcicki, A.M., 1996. Largescale right-slip displacement on the East San Francisco Bay region fault system: Implications for location of late Miocene to Pliocene Pacific plate boundary. *Tectonics*, 15: 1–18.
- Melosh, H.J., 1996. Dynamic weakening of faults by acoustic fluidization. *Nature*, 379: 601–606.
- Moresi, L.N., Dufour, F., Muhlhaus, H., 2003. A Lagrangian integration point finite element method for large deformation modeling of viscoelastic geomaterials. *J. Comput. Phys.*, 184: 476-497.
- Morgan, W.J., 1968. Rises, trenches, great faults, and crustal blocks. *J. Geophys. Res.*, 73: 1959-1982.
- Mount, V.S., Suppe, J., 1987. State of stress near the San Andreas fault: Implications for wrench tectonics. *Geology*, 15: 1143–1146.
- Muhlhaus, H.-B., Aifantis, E.C., 1991. A variational principle for gradient plasticity. *Int. J. Solids Struct.* 28: 845–857.
- Muhlhaus, H.-B., Regenauer-Lieb, K., 2005. Towards a self-consistent plate mantle model that includes elasticity: simple benchmarks and application to basic modes of convection. *Geophys. J. Int.*, 163: 788-800.
- Muhlhaus, H.-B., Vardoulakis, I., 1987. The thickness of shear bands in granular materials. *Geotechnique*, 37: 271-

- 283.
- Murray, J., Segall, P., Cervelli, P., Prescott, W., Svarc, J., 2001. Inversion of GPS data for spatially variable slip-rate on the San Andreas Fault near Parkfield, CA. *Geophys. Res. Lett.*, 28: 359-362.
- Murray, M.H., Segall, P., 2001. Modeling broadscale deformation in northern California and Nevada from plate motions and elastic strain accumulation. *Geophys. Res. Lett.*, 28: 4315-4318.
- Nicholson, C., Sorlien, C.C., Atwater, T., Crowell, J.C., Luyendyk, B.P., 1994. Microplate capture, rotation of the western Transverse Ranges, and initiation of the San Andreas transform as a low angle fault system. *Geology*, 22: 239-242.
- O'Neill, C., Moresi, L., Muller, D., Albert, R., Dufour, F., 2006. Ellipsis 3D: A particle-in-cell finite-element hybrid code for modelling mantle convection and lithospheric deformation. *Comput. Geosci.*, 32: 1769-1779.
- Ortiz, M., Popov, E.P., 1985. Accuracy and stability of integration algorithms for elastoplastic constitutive relations. *Int. J. Numer. Meth. Engng.*, 21: 1561-1576.
- Page, B., Thompson, G., Coleman, R., 1998. Late Cenozoic tectonics of the central and southern Coast Ranges. *Geol. Soc. Am. Bull.*, 110: 846-876.
- Perez-Foguet, A., Rodriguez-Ferran, A., Huerta, A., 2000. Numerical differentiation for local and global tangent operators in computational plasticity. *Comput. Methods Appl. Mech. Engng.*, 189: 277-296.
- Petrinin, A., Sobolev, S.V., 2006. What controls thickness of sediments and lithospheric deformation at a pull-apart basin? *Geology*, 34: 389-392.
- Petrinin, A.G., Sobolev, S.V., 2008. Three-dimensional numerical models of the evolution of pull-apart basins. *Phys. Earth Planet. Inter.*, doi:10.1016/j.pepi.2008.08.017
- Poliakov, A.N., Cundall, P.A., Podladchikov, Y.Y., Lyakhovsky, V.A., 1993. An explicit inertial method for the simulation of the viscoelastic flow: an evaluation of elastic effects on diapiric flow in two- and three-layers models. In: D.B. Stone and S.K. Runcorn (Editors). *Flow and creep in the Solar System: observations, modelling and theory*, Kluwer Academic Publishers, 175-195.
- Poliakov, A.N., Herrmann, H.J., 1994. Self-organized criticality of plastic shear bands in rocks. *Geophys. Res. Lett.*, 21: 2143-2146.
- Popov, A.A., Sobolev, S.V., 2008. SLIM3D: A tool for three-dimensional thermomechanical modeling of lithospheric deformation with elasto-visco-plastic rheology. *Phys. Earth Planet. Inter.*, doi:10.1016/j.pepi.2008.03.007
- Prescott, W.H., Savage, J., Svarc, J., Manaker, D., 2001. Deformation across the Pacific-North America plate boundary near San Francisco, CA. *J. Geophys. Res.*, 106: 6673-6682.
- Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 2002. *Numerical recipes in C++: the art of scientific computing*, 2nd edn. Cambridge University Press, Cambridge.
- Provost, A.-S., Houston, H., 2003. Stress orientations in northern and central California: Evidence for the evolution of frictional strength along the San Andreas plate boundary system. *J. Geophys. Res.*, 108(B3), 2175, doi:10.1029/2001JB001123.
- Puso, M.A., Solberg, J., 2006. A stabilized nodally integrated tetrahedral. *Int. J. Numer. Meth. Engng.*, 67: 841-867.
- Rashid, M.M., 1993. Incremental kinematics for finite element applications. *Int. J. Numer. Meth. Engng.*, 36: 3937-3956.
- Rassineux, A., 1998. Generation and optimization of tetrahedral meshes by advancing front technique. *Int. J. Numer. Meth. Engng.*, 41: 651-674.
- Reese, S., 2003. On a consistent hourglass stabilization technique to treat large inelastic deformations and thermo-mechanical coupling in plane strain problems. *Int. J. Numer. Meth. Engng.*, 57: 1095-1127.
- Regenauer-Lieb, K., 2006. Water and Geodynamics. *Rev. Miner. Geochem.*, 62: 451-473.
- Regenauer-Lieb, K., Yuen, D.A. 2003. Modeling shear zones in geological and planetary sciences: solid- and fluid-thermal-mechanical approaches. *Earth Sci. Rev.*, 63: 295-349.
- Regenauer-Lieb, K., Yuen, D.A. 2004. Positive feedback of interacting ductile faults from coupling of equation of state, rheology and thermal-mechanics. *Phys. Earth Planet. Inter.*, 142: 113-135.
- Richards-Dinger, K.B., Shearer, P.M., 1997. Estimating crustal thickness in southern California by stacking PmP arrivals. *J. Geophys. Res.*, 102: 15211-15224.
- Riddihough, R., 1984. Recent movements of the Juan de Fuca plate system. *J. Geophys. Res.*, 89: 6980-6994.
- Romanyuk, T., Mooney, W.D., Detweiler, S., 2007. Two lithospheric profiles across southern California derived from gravity and seismic data. *J. Geodynam.*, 43: 274-307.
- Roscoe, K.H., 1970. The influence of strains in soil mechanics, 10th Rankine Lecture. *Geotechnique*, 20: 129-170.
- Saad, Y., 1994. ILUT: a dual threshold incomplete ILU factorization. *Numer. Linear Algebra Appl.*, 1: 387-402.
- Schmalholz, S.M., Podladchikov, Y.Y., Schmid, D.W., 2001. A spectral/finite difference method for simulating large deformations of heterogeneous, viscoelastic materials. *Geophys. J. Int.*, 145, 199-208.
- Scholz, C.H., 1989. Mechanics of faulting. *Ann. Rev. Earth Planet. Sci.*, 17: 309-334.
- Scholz, C.H., 2000. A fault in the 'weak San Andreas' theory. *Nature*, 405: 234.
- Severinghaus, J., Atwater, T., 1990. Cenozoic geometry and thermal state of the subducting lithosphere beneath North America. In B.P. Wernicke (editor). *Basin and Range extensional tectonics near the latitude of Las Vegas, Nevada*. *Geol. Soc. Am. Mem.*, 176: 1-22.
- Shampine, L.F., Watts, H.A., 1970. FZERO, a root-solving code. Report SC-TM-70-631, Sandia Laboratories.

REFERENCES

- Shen-Tu, B., Holt, W.E., Haines, A.J., 1999. Deformation kinematics of the western United States determined from Quaternary fault slip rates and geodetic data. *J. Geophys. Res.*, 104, 28927-28955.
- Shephard, M.S., Georges, M.K., 1991. Automatic three-dimensional mesh generation by the finite octree technique. *Int. J. Numer. Meth. Engng.*, 32: 709-749.
- Shih, T.M., Tan, C.H., Hwang, B.C., 1989. Effects of grid staggering on numerical schemes. *Int. J. Numer. Meth. Fluids*, 9: 193-212.
- Sibson, R.H., 1983. Continental fault structure and the shallow earthquake source. *J. Geol. Soc. London*, 140: 741–767.
- Simo, J.C., Hughes, T.J.R., 2000. *Computational Inelasticity*, 2nd edn. Springer-Verlag, New York.
- Simo, J.C., Kennedy, J.G., Govindjee, S., 1988. Non-smooth multisurface plasticity and viscoplasticity. Loading/unloading conditions and numerical algorithms. *Int. J. Numer. Meth. Engng.*, 26: 2161-2185.
- Simo, J.C., Taylor, R.L., 1985. Consistent tangent operators for rate-independent elastoplasticity. *Comput. Methods Appl. Mech. Engrg.*, 48: 101-118.
- Slezkin, A., 1955. *Dynamics of viscous incompressible fluid* (in Russian). Gostekhizdat, Moscow.
- Sloan, S.W., Booker, J.R., 1986. Removal of singularities in Tresca and Mohr-Coulomb yield functions. *Commun. Appl. Numer. Meth.*, 2: 173-179.
- Smith, D.A. Milbert, D.G., 1999: The GEOID96 high resolution geoid height model for the United States. *J. Geodesy*, 73: 219-236.
- Sobolev, S.V., Babeyko, A.Yu., 2005. What drives orogeny in the Andes? *Geology*, 33: 617–620 (Highlighted by Nature, vol. 436, 2005, p. 756).
- Sobolev, S.V., Petrunin, A., Garfunkel, Z., Babeyko, A.Y., DESERT Group., 2005. Thermo-mechanical model of the Dead Sea transformation. *Earth Planet. Sci. Lett.*, 238: 78-95.
- Sobolev, S.V., Zoback, M.D, Babeyko, A.Yu., 2003. A thermo-mechanical model of the San Andreas Fault system in central and northern California. *Eos Trans. AGU*, 84(46), Fall Meet. Suppl., Abstract T51A-01 INVITED.
- Spray, J.G., 1987. Artificial generation of pseudotachylite using friction welding apparatus: simulation of melting on a fault plane. *J. Struct. Geol.*, 9: 49–60.
- Stock, J.M., Lee, J., 1994. Do microplates in subduction zones leave a geological record? *Tectonics*, 13: 1472–1487.
- Stock, J.M., Molnar, P. 1988. Uncertainties and implications of the Late Cretaceous and Tertiary position of North America relative to the Farallon, Kula, and Pacific plates. *Tectonics*, 7: 1339–1384.
- Tackley, P.J., 1998. Self-consistent generation of tectonic plates in three-dimensional mantle convection. *Earth Planet. Sci Lett.*, 157: 9-22
- Tackley, P.J., Xie, S., 2003. STAG3D: A code for modeling thermo-chemical multiphase convection in Earth's mantle. In: K.J. Bathe (Editor). *Proceedings of the Second MIT Conference on Computational Fluid and Solid Mechanics*. Elsevier B.V., Amsterdam, 1524-1527.
- ten Brink, U.S., Shimizu, N., Molzer, P.C., 1999. Plate deformation at depth under northern California: slab gap of stretched slab? *Tectonics*, 18: 1084–1098.
- Thatcher, W., Foulger, G., Julian, B., Svarc, J., Quilty E., Bawden, G., 1999. Present-day deformation across the Basin and Range province, western United States. *Science*, 283: 1714–1718.
- Thompson, G.A., 1999. California coast ranges tectonics in the light of geophysical constraints: A tribute to Ben Page. *Int. Geol. Rev.*, 41: 383–390.
- Townend, J., Zoback, M. D., 2000. How faulting keeps the crust strong. *Geology*, 28: 399–402.
- Townend, J., Zoback, M. D., 2001. Implications of earthquake focal mechanisms for the frictional strength of the San Andreas fault system. In R.E. Holdsworth et al. (editors). *The Nature and Significance of Fault Zone Weakening*. *Geol. Soc. London Spec. Publ.*, 186: 13–21.
- Townend, J., Zoback, M. D., 2004. Regional tectonic stress near the San Andreas fault in central and southern California. *Geophys. Res. Lett.*, 31, L15S11, doi:10.1029/2003GL018918.
- Trompert, R.A., Hansen, U., 1996. The application of a finite-volume multigrid method to 3-dimensional flow problems in a highly viscous fluid with a variable viscosity. *Geophys. Astrophys. Fluid Dyn.*, 83: 261-291.
- van Wijk, J. W., Govers, R., Furlong, K.P., 2001. Three-dimensional thermal modeling of the California upper mantle: a slab window vs. stalled slab. *Earth Planet. Sci. Lett.*, 186: 175-186.
- Vermeer, P.A., 1990. The orientation of shear bands in biaxial tests. *Geotechnique*, 40: 223-226.
- Weinberg, R.B., Schmeling, H., 1992. Polydiapirs: multiwavelength gravity structures. *J. Struct. Geol.*, 14: 425-436.
- Wilkins, M.L., 1964. Calculation of elastic-plastic flow. In: B. Alder (Editor). *Methods in Computational Physics*, Academic press, New York, 211-263.
- Wilson, D.S., McCrory, P A., Stanley, R.G., 2005. Implications of volcanism in coastal California for the Neogene deformation history of western North America. *Tectonics*: 24, TC3008, doi: 10.1029/2003TC001621.
- Wilson, J.T., 1965. A new class of faults and their bearing on continental drift. *Nature*, 207: 343-347.
- Zandt, G., Furlong, K.P., 1982. Evolution and thickness of the lithosphere beneath coastal California. *Geology*, 10: 376–381.
- Zhong, S., Gurnis, M., 1996. Incorporation of fault-bound plates in three-dimensional models of mantle flow. *Nature*, 383: 245-247.
- Zienkiewicz, O.C., Taylor, R.L., 2000. *The finite element method*, 5th edn. Butterworth-Heinemann, Oxford.
- Zoback, M.D., 2000. Strength of the San Andreas. *Nature*, 405:31-32.

- Zoback, M.D., Healy, J.H., 1992. In situ stress measurements to 3.5 km depth in the Cajon Pass Scientific Research Borehole: Implications for the mechanics of crustal faulting. *J. Geophys. Res.*, 97: 5039-5057.
- Zoback, M.D., Zoback, M.L., Mount, V.S., Suppe, J., Eaton, J.P., Healy, J.H., Oppenheimer, D., Reasenber, P., Jones, L., Raleigh, C.B., Wong, I.G., Scotti, O., Wentworth, C., 1987. New evidence on the state of stress of the San Andreas fault system. *Science*, 238: 1105-1111.

List of figures

1.1	Global map of crustal plate boundaries	2
1.2	Map of major faults in the SAFS	3
2.1	Simplified tectonic model of North America Pacific interactions	6
2.2	Snapshots from animated reconstructions of SAFS evolution	8
2.3	Slab window formation cartoon	11
2.4	Map of major features of the SAFS in central and northern California	13
2.5	Maximum horizontal compression and crustal velocity in central California	15
2.6	P-wave seismic tomography of western United States upper mantle	18
2.7	Crustal thickness estimates in southern California	20
2.8	Present plate velocities across SAFS in central California	22
3.1	Rheological structure of lithosphere in a thin-shell model	24
3.2	Thin-shell model of western United States	25
3.3	Thin-shell finite element grid for the North America plate	26
3.4	Results of thin-shell finite element model of North American plate	28
3.5	Results of thermal model of Mendocino Triple Junction migration	30
3.6	Thermomechanical model of Gorda slab removal	31
3.7	Three-dimensional thermal model of the California upper mantle	34
3.8	Three-dimensional finite element model of crustal deformation	36
3.9	Three-dimensional finite element models for active tectonics in western US	37
3.10	Three-dimensional visco-elasto-plastic model of the SAFS	38
3.11	Integrated mechanical model of the SAF	40
3.12	Extended two-dimensional thermomechanical model of the SAF	42
3.13	Results of the extended 2D model (strain rate and strike-slip velocity)	43
3.14	Results of the extended 2D model (heat flow).	44
4.1	Logarithm of effective viscosity for dry olivine	51
4.2	Schematic representation of the yield surface	53
4.3	Finite element discretization	56
4.4	Grid distortion for a single time step of the Rayleigh-Taylor instability	58

4.5	Contour plots of accumulated plastic strain	67
4.6	Initial and deformed shape of the elastic plate	71
4.7	Relative error in approximation of stresses and vertical displacements	72
4.8	Isolines of vertical velocity for the cylinder problem	73
4.9	Effective logarithmic strain rate for the shear band initiation problem	75
4.10	Setup and results of the triaxial compression test with uniform random seeding	78
4.11	Setup and results of the triaxial compression test with gash seeding	78
4.12	Lithospheric transpressional deformation problem	79
4.13	Evolution of the Lagrangian mesh embedded in the material	81
5.1	Positioning of the model with respect to reconstructed plate configuration	85
5.2	Finite element discretization and boundary conditions	86
5.3	Block diagram of extended domain for solution of the energy balance equation	86
5.4	Top views on the discretized finite element models (model scenarios)	87
5.5	Initial thermal structure of the model	89
5.6	Crustal model and discretization in the vertical direction	91
5.7	Idealization of the Gorda plate subduction	91
5.8	Temperature evolution in the slab window	93
5.9	Effect of shear heating on the terrane accretion	95
5.10	Fault configuration and temperature distribution in transects	96
5.11	Effect of transpression on strike-slip velocities and crustal thickness	98
5.12	Results of the model with Monterey microplate, and friction 0.12	100
5.13	Results of the model with Monterey microplate, and friction 0.09	101
5.14	Results of the model with Salinian block, and friction 0.12	102
5.15	Results of the model with Salinian block, and friction 0.06	103
5.16	Capture of Bay Area block (top view)	104
5.17	Capture of Bay Area block (transects)	105
5.18	Comparison of major fault offsets in central and northern California	106
5.19	Character of the asthenosphere upwelling	108
5.20	Topography and Moho depth results and comparison	110
5.21	Comparison of the calculated and observed surface heat flow	112
5.22	Comparison between the models with the weak and moderately strong friction	113

Acknowledgments

I would like to express my deepest appreciation to Dr. Stephan Sobolev, head of Geodynamic Modeling Section in GFZ-Potsdam, who helped, supported, and guided me through the course of my PhD study. The door in his office was always opened for me to come and ask any questions. He always had enough patience to explain me geophysical stuff, and treated my pretty much engineering mind with care. I would also like to thank the other colleagues of mine from Geodynamic Modeling Section (Alexey, Andrey, Sascha, Andreas, Ernesto, Javier, and Bernhard) for numerous fruitful discussions and collaboration. Special thanks go to our secretary Marie for helping me with bureaucratic stuff and German language.

I thank my supervisor at the University of Potsdam and head of Geophysical Deep Sounding Section Prof. Dr. Michael Weber, who hosted me in GFZ for almost three and half years. Prof. Weber helped me a lot by providing favourable working atmosphere and perfect scientific management. I also extracted huge benefit from attending lectures by Prof. Patrick O'Brien, Prof. Frank Scherbaum, and Prof. Manfred Strecker. Special thanks go to Professor Romain Bousquet, who kindly accepted my participation in his University course and helped me to obtain some teaching experience.

Rather occasional discussions with Taras Gerya, Stefan Schmalholz, Boris Kaus, Klaus Regenauer-Lieb, Jean Braun, Evgeni Burov, Yuri Podladchikov, Alan Levander, and Mark Zoback helped to improve my understanding of certain things. I was also greatly inspired by the scientific work of these people. Back to Russia, I would like to thank my former colleagues and supervisors Alex Vlasov and Mike Mnushkin. It can be said with no exaggeration that they have introduced me into the world of scientific computing. My mother Galina, aunt Alla, and uncle Peter are thanked for their constant support. Thank you Katja, for all the best moments we had together.

This work was partially financed by the Deutsche Forschungsgemeinschaft. Additional funding was provided by German Research Centre for Geosciences (GFZ Potsdam). This support is gratefully acknowledged.

Curriculum Vitae

First Name: Anton
Middle Name(s): Alexander
Last Name: Popov
Birth Date: 14 June 1980
Place of Birth: Tashkent, Uzbekistan (former USSR)
Nationality: Russian
Marital Status: single

Education

1986 - 1989 Primary school, Zarafshan, Uzbekistan (former USSR)
1989 - 1997 Secondary school, Pyatigorsk suburbs, Russia
1997-2002 Moscow State University of Civil Engineering, Russia
Master degree in Civil Engineering
Specialization in Soil and Rock Mechanics and Foundation Design
2006-2008 University of Potsdam, Germany
PhD study in Geophysics (Geodynamic modeling)

Career

2004-present: Researcher at Department of Physics of the Earth
GeoForschungsZentrum Potsdam (Potsdam, Germany)
2002-2004: Researcher at Institute of Applied Mechanics
Russian Academy of Sciences (Moscow, Russia)
2002: Trainee engineer at Department of Scientific Consulting
Design Institute "HydroProject" (Moscow, Russia)
2000-2001: Assistant at Laboratory of Soil Mechanics
Moscow State University of Civil Engineering (Moscow, Russia)

Tutoring Experience

Assistance in the lecture course "Heat Flow in the Earth"
Department of Geosciences
University of Potsdam (Potsdam, Germany)

