We are IntechOpen, the world’s leading publisher of Open Access books
Built by scientists, for scientists

4,200
Open access books available

116,000
International authors and editors

125M
Downloads

154
Countries delivered to

TOP 1%
Our authors are among the most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™
Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com
Chapter 10

Finite Element Models of Elastic Earthquake Deformation

Sui Tung, Timothy Masterlark and Daniel Sai Huen Lo

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.76612

Abstract

The Earth’s surface deforms in response to earthquake fault dislocations at depth. Deformation models are constructed to interpret the corresponding ground movements recorded by geodetic data such GPS and InSAR, and ultimately characterize the seismic ruptures. Conventional analytical and latest numerical solutions serve similar purpose but with different technical constraints. The former cannot simulate the heterogeneous rock properties and structural complexity, while the latter directly tackles these challenges but requires more computational resources. As demonstrated in the 2015 M7.8 Gorkha, Nepal earthquake and the 2016 M6.2 Amatrice, Italy earthquake, we develop state-of-art finite element models (FEMs) to efficiently accommodate both the material and tectonic complexity of a seismic deformational system in a seamless model environment. The FEM predictions are significantly more accurate than the analytical models embedded in a homogeneous half-space at the 95% confidence level. The primary goal of this chapter is describe a systematic approach to design, construct, execute and calibrate FEMs of elastic earthquake deformation. As constrained by coseismic displacements, FEM-based inverse analyses are employed to resolve linear and nonlinear fault-slip parameters. With such numerical techniques and modeling framework, researchers can explicitly investigate the spatial distribution of seismic fault slip and probe other in-depth rheological processes.

Keywords: FEM, earthquake, deformation, inverse model

1. Introduction

With the wealth of geological and geodetic information accumulated around seismogenic zones over the past decades, we are posed to ask: in what way we can unify and take advantage of
these data to study the earthquake hazard of those areas? The rate of interseismic creeping/slow slip [1, 2], coseismic slip [3], and afterslip [4] are usually estimated with fault dislocation models that predict surface deformation from in-depth fault slip motions. Customary analytical (Okada) solutions analyze rectangular slip in an isotropic half-space [5] and serve as a good initial approximation for inferring fault behaviors which are critical for assessing regional strain accumulation related to seismic hazard [6, 7]. However, the more we study, the more we find that the shallow part of the crust (especially the upper crust) is not as simple as, or even far beyond, a uniform half-space (Figure 1) [8]. The major shortcomings of an Okada solution rest on its assumptions of homogeneous crust (HOM) and a rectangular fault dislocation [5] which are inadequate according to in situ geological observations [9]. Failure of simulating the realistic crustal domain could induce fundamental uncertainties in predicting faulting-induced displacements, which could propagate into the interpretations of related earthquake studies [10]. For instance, we found that ignoring the heterogeneous crust (HET) in deformation models could yield to considerable prediction errors when simulating seismic deformation of the 2015 Gorkha, Nepal earthquake (Figure 1) [3]. This can be explained by the lateral and vertical material variations across the epicentral area, which poses a technical challenge for conventional analytical solutions (Figure 1). The importance of HET has also been suggested by many other colleagues. Hearn and Bürgmann [11] develop a finite element model for the 1999 Izmit Turkey earthquake and show that the GPS-recovered seismic moment is up to 40% greater for models incorporating depth-dependent shear modulus than it is for uniform elastic half-space models. The corresponding Coulomb stress change in the lower crust is ~300% larger than a model domain using a homogenous shear modulus. They conclude that models of co-seismic ruptures and postseismic viscoelastic relaxation associated with large strike-slip earthquakes should

Figure 1. (a) Topography-shaped FEM domain of the 2015 M7.8 Gorkha, Nepal earthquake. (b and c) Nearfield meshes are refined within the central Himalayas [7]. Domain spatial distribution of (d) Young’s Modulus and (e) Poisson’s ratio are derived from CRUST2.0 [34].
incorporate depth-dependent elasticity, particularly for the triggering studies beneath the lower crust. Williams and Wallace [12] build finite element models (FEMs) in conjunction with a New Zealand-wide seismic tomography model to assign elastic properties. They find that these heterogeneous models typically require ~20% less slip than homogeneous models where the slip is deep or there is reasonable geodetic coverage above the slipping region. In cases where the slip is shallow (and mostly offshore) and there is little geodetic coverage directly above the slipping region, the heterogeneous models can predict significantly larger amounts of slip (~42%). These changes in the predicted magnitude of slip have important implications for quantifying slip budgets accommodated by slow slip at subduction zones worldwide [12]. The sensitivity of fault-slip solutions to HET is also demonstrated by Trasatti et al. [13] for the 2009 L’Aquila earthquake, showing up to 20% of discrepancy between Okada and FEM-based heterogeneous solutions which reveal new fault-slip features near the epicenter. Tung et al. [3, 10] also find the co-seismic GPS displacements are significantly better recovered by a HET model than a HOM model at the 95% confidence level. This model uncertainty is generally larger than those inherited in the geodetic measurements. The advantages of using FEMs over analytical solutions for simulating fault deformation are also exemplified among other earthquake studies [3, 10, 13–25].

With the advancement of computation power, FEM and large data acquisition techniques such as space geodesy, remote sensing, and imaging, we are now able to study the seismic activities on large-scale tectonic plates across continents with unprecedented detail and precision. For finite elastic deformation, elastoplastic analysis over a large domain, based on the Hellinger-Reissner and the Hu-Washizu functionals, 3D solid enhanced assumed strain formulations are among the most efficient and stable finite elements [26–35]. For high accuracy FE solutions over complicated domains of curved boundary, however, we could also use quadratic solid elements such as 10-node and 20-node tetrahedral elements, 20-node and 27-node hexahedral elements, etc. [36, 37]. Such methods are demonstrably useful for simulating a variety of complex science and engineering systems. Nonlinear contact problem of a hip joint is analyzed using T4, T10, H8 and H20 elements [38]. Fluid-saturated, inelastic, pressure-sensitive porous solid medium subjected to dynamic large deformation is analyzed by the mixed theory formulation using solid quadratic H27 elements [39].

Following the advanced numerical simulations, much work has been done recently on the deformation, stress distribution, faults, ruptures, dynamics, and wave propagation of tectonic plates by FEMs. An elastic plane stress FEM incorporating realistic rock parameters was used to calculate the stress field, displacement field, and deformation of the plate interactions in the eastern Mediterranean [40]. A 3D FE model of ~3000 hexahedral elements and nodes is set up by Lu et al. [41] for the surface topology, major active fault zones and the stress field of the Chinese continent to study the mechanism of the long-distance jumping migration over active seismogenic areas. Shear zones are identified over regional-scale tectonic plates by 2D FEMs of faults and boundaries of tectonic plates [42]. By means of cascaded FE simulations, glacial isostatic adjustment is extended to investigate the relationship between glacial loading/unloading and fault movement due to the spatial–temporal evolution of stresses [43]. Lithospheric pressure and density fields are determined by novel FEM-based gravity inversion which is implemented within the open-source escript modeling environment [44]. Sophisticated 3D
FEMs highlight that surface processes acting on normal-fault-bounded mountain ranges may sustain fault slip for millions of years even after regional extension has stopped [45]. By combining several datasets, a FE module was developed to estimate the gravitational potential energy of the lithosphere and calculate stresses acting on the real (non-planar) geometry of African plate [46]. A 3D FEM is also employed to simulate the original co-seismic Coulomb stress patterns through space and time as modified by post-seismic viscoelastic flow [47]. FEM-derived solutions are also integrated into the regularized linear inversion of InSAR data over the volcano surface to image the 3D deformation field and pressure distribution [48, 49].

Finite element generation is an important step for advancing 3D large-scale numerical modeling, as almost three quarters of the overall analysis time is devoted to mesh generation and the related geometrical analysis. A comprehensive account of various mesh generation techniques is described and discussed in the textbook “Finite Element Mesh Generation” by Lo [50]; and in general, unstructured meshes are generated by the Delaunay triangulation, the advancing-front approach and the quadtree/Octree techniques etc., whereas structured meshes of hexahedral elements can be synthesized by some mapping and sweeping processes. Transition quadrilateral and hexahedral elements [51] and universal connection hexahedral elements [52] have also been developed for adaptive refinement analysis. However, in conjunction with the popular mesh generation methods mentioned here, other techniques could also be employed for specific applications to broad-scale earthquake problems. A full waveform inversion method that incorporates seismic data on a wide range of space-temporal scales on both crustal and upper-mantle structure is developed with the multi-grid FE scheme [53]. Furthermore, a non-conforming octree-based scheme on a fictitious domain for the numerical modeling of earthquake induced ground motion of realistic surface topology of the Earth’s crust was presented by Restrepo and Bielak [54]. Other interdisciplinary examples are the adaptive multi-material grids generated from image data for biomedical fluid–structure simulations [55], and the conformal finite element/volume meshes derived from 3D measurements of the propagation of small fatigue cracks [56].

2. Data

2.1. Seismic tomography

The propagation of earthquake waves is a function of rock material properties within the crustal layer that hosts the waves [57]. These material properties alter the traveling velocities of the P and S wave subjected to the local elastic rock properties. A tomography model refers to a velocity model that describes a 3D distribution of P-wave velocity \(V_p\), S-wave velocity \(V_s\), which is interchangeable with the spatial distribution of elastic moduli, namely, Young’s Modulus \(E\) and Poisson’s ratio \(v\), as formulated by [3, 10]:

\[
E = \frac{\rho V_s^2 \left(3V_p^2 - 4V_s^2\right)}{V_p^2 - V_s^2} \quad \text{and} \quad v = \frac{V_p^2 - 2V_s^2}{2 \left(V_p^2 - V_s^2\right)}
\]
In general, $E$ increases as a function of depth, while $v$ decreases in the deeper crust and converges to 0.25 for mantle rocks [58] (Figure 1). It is noteworthy that the homogeneous elastic half-space assumptions, which are commonly used in the models of earthquake deformation [5], assume that these moduli are uniform in space and generally conflict with the tomographic observation. Tung et al. [3, 10] quantifies the implications of ignoring seismic tomography in elastic deformation models for the 2015 M7.8 Gorkha, Nepal earthquake (Figure 1) and the 2016 M6.2 Amatrice, Italy earthquake. In these examples, over-simplifying the relatively weak materials near the surface translates to substantial prediction errors of InSAR and GPS signals. This underpins the necessity of modeling seismic deformation within the HET domain of FEMs which are one of the few existing methods capable of simulating 3D crustal rock heterogeneity.

2.2. Geodetic data

2.2.1. GPS data

The time series of Earth positioning are collected by thousands of GPS receiver stations using radio-wave signals from the constellations of Global Positioning System (GPS) satellites (Figure 2). Generally, these data provide a 3D displacement field of a station location with uncertainties close to 1 mm depending the atmospheric noise and other data processing errors [59]. Some GPS stations sample the ground positions continuously, while others are re-visited periodically through multiple surveying campaigns [11, 60]. Furthermore, some of the former become able to provide real-time or near-real-time data feed with automatic data processing procedures and web-based data sharing platforms, such as EarthScope-PBO-UNAVCO, USGS-NEIC and NSF-Cascadia Initiative [19]. Continuous GPS sites record systematic positioning data and generally require more considerations such as sustainable power supply, data logging protocols and
secure station design, while campaign-style measurements rely on more labor-intensive surveying strategies and are of lower temporal resolution over a longer time period. The technical details of GPS survey implementation are far beyond the scope of this work. For our purposes, we mainly focus on GPS data to constrain the three-component displacement field before and after an earthquake at given GPS stations (Figure 2).

2.2.2. InSAR data

In 1993, Massonnet et al. [61] presented the first Interferometric Synthetic Aperture Radar (InSAR) image to map the displacement field of the 1992 Landers earthquake. This image is derived from the changes in the phase distribution of radar scenes acquired respectively from two separate satellite passes over the epicentral area. An InSAR image unfolds the LOS displacement field that represents the difference between the positions of surface location at the time of the second and first satellite passes. By comparing the differential radar phase arrivals recorded before and after the earthquake, the spatial distribution of phase interference estimates the displacements parallel to the look direction of the satellite in unwrapped InSAR images (Figure 2). No information is available about the displacement between the two image acquisition times. The technical details of InSAR processing are far beyond the scope of this chapter, but aspects relevant to signal modeling are described here. The process of unwrapping InSAR data is to integrate the spatial phase data to map the line-of-sight (LOS) displacements. For our purposes, the obtained displacement field refers to that induced by earthquake dislocations [61]. FEMs are designed to predict these unwrapped phase data and hence characterize seismic sources [3, 10]. Moreover, InSAR observations are susceptible to artifacts caused by atmospheric noises and mismodeled orbital effects [62]. The former can be avoided via reducing the temporal baseline separation between two satellite passes, while the latter can be accounted for with linear inverse methods, as discussed in Section 4. Due to these artifacts, each pixel of an InSAR image is not completely independent so that a data covariance matrix is involved to empirically weight each pixel [63]. Alternatively, geospatial reduction techniques such as quadtree decomposition may be applied to filter unwanted signals, account for covariance and improve computational efficiency of matrix inversion.

2.3. Topography and bathymetry

Unlike the conventional HOM assumption [5], our FEMs and corresponding meshing regimes are capable of calculating the fault deformation over surface topography [3]. This surface is configured as a stress-free surface because we assume that there are only minimal normal stress variations and shear resistance. It is well known that the shape of such free surface affects deformation predictions, especially for tsunami modeling studies [64]. We can visualize this aspect by considering how the calculated deformation field would be affected by the limiting case of a vertical cliff near the rim of continental shelf. In this case, the ground surface is orthogonal to the assumed flat surface of an HOM domain. Matsuyama et al. [65] underlines the importance of including non-uniform topography and bathymetry in fault deformation model to assess the tsunami hazard and coastal impact upon tsunamigenic events. Subjected to the ongoing tectonic movements and irregular structural settings, seismogenic/tsunamigenic...
zones usually attain a variable topography or bathymetry, which can be well accommodated by our FEMs for better accuracy of source characterization and tsunami wave predictions (Figure 1) [3, 66].

3. Model configurations

3.1. Domain and mesh configurations

Since FEMs are designed to simulate the crustal body of the seismogenic zone in a scale of few tens to thousands of kilometers, one of the initial decisions is to define a particular model coordinate system and units. A FEM is an assembly of numerous finite-volume elements stitched together to form a broader modeling domain (Figure 1). Those elements may attain different degrees of freedom (DOF) and geometry. For instance, a linear (p = 1) 4-node tetrahedral T4 element having 4 vertices attains a DOF of 4, while a linear (p = 1) 8-node hexahedral H8 element comprises 8 vertices and inherits 8 DOF. The latter could be further improved by the enhanced assume strain to be very competitive in regular simple geometry and structural shell problems. Furthermore, DOF applied to the solution variables may, for example, have 3 displacement components (DOF = 3) plus an additional pore pressure DOF. The meshing schemes of these elements are generally divided into two main categories, namely, structured meshing and free meshing. The former requires the meshes to be created according to a certain degree of uniformity. The element orientation, volume and nomenclature are defined in a structured manner, which is favorable for low-level modeling and solving procedures. On the contrary, the latter loosens all these criteria to let mesh “fill up” the model domain with the least number of elements. The choice of element type and meshing scheme heavily depends on the nature of the problems researchers are going to resolve. When earthquake slip is along a complex fault curvature, tetrahedral elements are preferred with regards to their smaller interior angles and thus ability to effectively tessellate a sharply-turning geometry such as listric faults and abruptly-changing topography (Figure 1). The prediction differences between the tetrahedral and rectangular elements become negligible when the fault is planar and the surface is flat. Given the same element-edge length and constant model domain, the tetrahedral mesh aggregation usually contains more elements than the rectangular aggregation as the volume of an individual tetrahedral element is smaller than that of a rectangular element. Hence, the computational time is longer for the former. Similarly, the free meshing algorithm allows more efficient and flexible tessellation of complex geometry than the structured approach, requiring more computational power. The modeling accuracy could be further boosted by incorporating quadratic elements (p = 2) instead of linear elements (e.g., T4 and H8). Quadratic Tetrahedral T10 element, which is one of the most versatile elements for both flexibility and accuracy, can fill up most complicated domains using an automatic mesh generation scheme, while the corresponding hexahedral H20 element provides another accurate formulation for simple geometry. As expect, using quadratic elements not only substantially improves simulation accuracy but also increases the number of domain nodes and hence computing time.
This leads the researcher to cautiously consider a fundamental trade-off problem between the FEM approximations and the limitations of the available computing resources. There might be cases in which differences between 1D model and a 2D model are negligible for a smaller study area. However, the computational time of model configuration and execution of a 3D domain is at least several tens to thousands times longer than that of a 2D domain, depending on the adopted meshing scheme and element/seed size. A 3D domain subjected to tetrahedral random meshing with the largest number of elements is compensated by a maximum flexibility of simulating the tectonic and lithospheric environment. For a given size of domain space, a large number of smaller elements translates to larger solution matrix of algebraic operations that may become numerical unfeasible when the computing time is too long or the calculation process is non-accomplishable. Alternatively, a small number of larger elements satisfies a smaller matrix problem that only requires nominal computing facilities, but at a cost of losing precision to resolve the equations of elasticity. Thus, apart from a general adaptive refinement analysis [3, 10], a common approach is to tessellate the near field region with a relatively small element size which gradually increases near the far-field boundaries [3, 10, 62, 67]. This radially-decaying meshing strategy satisfies the need for a refined resolution of nearfield areas expected with a relatively higher strain gradient (Figure 1), while the far-field boundary conditions are still connected numerically through large elements between the deformation source (i.e., the earthquake fault(s)) and the outer lateral surfaces exhibiting relatively low strain gradients. When installing the heterogeneous distribution of rock material into the FEM domain, elements of similar elastic properties (similar values of $E$ and $v$ ideally with respect to their integration points) are grouped into discrete element sets such that the entire FEM is a representation of multiple element sets (Figure 1) [3, 10]. As such, the resolution of rock heterogeneity is controlled by the element size as well as the discretization of elastic parametric values. From our modeling experiences, we default the number of element set to be about 100 for describing both regional and local crustal material variations (Figure 1).

3.2. Governing equations of elasticity

The governing equations regulate the physical behavior of a system. The governing equations for the elastic materials in a heterogeneous domain are [5, 58]:

$$\frac{\partial}{\partial x} \left[ G(x) \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \right] + \frac{\partial}{\partial x} \left[ \lambda(x) \frac{\partial u_i}{\partial x_k} \right] \delta_{ij} = 0 \quad \& \quad \delta_{ij} = \begin{cases} 0, i \neq j \\ 1, i = j \end{cases}$$

(2)

where $x$ is a spatial component of coordinate axes $x$; $u$ refers to the corresponding displacement; $G$ and $\lambda$ are respectively the shear modulus and Lamé’s parameter; $\delta$ is the Kronecker delta; component indices $i$ and $j$ span over orthogonal axes 1, 2, and 3 for a 3D domain such that $x_1$, $x_2$, and $x_3$ are equivalent to Cartesian coordinates $x$, $y$, and $z$. The subscript $k$ represents summation over all these three components. These equations describe elastic behavior in a domain comprising a spatial distribution of isotropic elastic material properties $G$ and $\lambda$ which can be derived from $E$ and $v$ [68]. Noting that when the elastic properties are taken outside of the spatial derivatives, along with appropriate initial and boundary conditions, Eq. (2) is reduced to the Navier formulation [69] and becomes a description of a HOM space that is
commonly assumed in deformation models [5]. However, some seismogenic zones such as subduction margins require localized complexity, or a distribution of material properties, \( G(x), \lambda(x) \) or \( E(x), \nu(x) \). As we incorporate seismic topographical data (e.g., CRUST2.0 [8]) into the calculation of rock material (Figure 1), Eq. (2) describes elastic behavior of a 3D elastic domain inheriting a spatial distribution of isotropic elastic properties. If researchers want to include elastic anisotropy, they can replace scalar elastic moduli by tensors. FEMs are, so far, the best mathematical tool that satisfy these elastic equations over arbitrary crustal domains.

3.3. Loading conditions and kinematic constraints

The loading conditions can be viewed as the impulse that triggers the fault model to deform. For our purposes of simulating fault-slip deformation and consistency with analytical solutions, the loading conditions are assigned with a set of kinematic constraints developed by Masterlark et al. [70]. The fault discontinuity in FEMs is meshed with multiple node pairs which consist of two overlapping nodes sharing the same initial geographic location. A quasi-static fault slip is applied to these node pairs by locally offsetting these two node members, node n1 and n2 of each pair along the rake, \( \theta_{\text{rake}} \). The loading condition specifies the subfault dislocation, \( \Delta m \), of each node pair through three equations of motion along the orthogonal axes:

\[
\begin{align*}
m_{\text{strike}}^{n1} - m_{\text{strike}}^{n2} &= \Delta m_{\text{strike}} = \Delta m \times \cos (\theta_{\text{rake}}) \\
m_{\text{dip}}^{n1} - m_{\text{dip}}^{n2} &= \Delta m_{\text{dip}} = \Delta m \times \sin (\theta_{\text{rake}}) \\
m_{\text{normal}}^{n1} - m_{\text{normal}}^{n2} &= 0
\end{align*}
\] (3)

where \( m_{\text{strike}}^{n1} \) and \( m_{\text{strike}}^{n2} \) refer the along-strike motion of node n1 and n2 respectively; \( m_{\text{dip}}^{n1} \) and \( m_{\text{dip}}^{n2} \) are the along-dip motion of node n1 and n2 respectively; \( m_{\text{normal}}^{n1} \) and \( m_{\text{normal}}^{n2} \) denote the fault-normal motion of node n1 and n2 respectively. The null fault-normal displacements of Eq. (5) ensure the footblock and hanging block are welded together along the fault-normal axis throughout the entire calculation. Non-slipping node pairs are constrained with \( \Delta m = 0 \) equivalent to a welded condition. For the FEM of the 2016 M6.2 Amatrice, Italy earthquake [10], 665 node pairs are assigned to assemble the fault so that the loading conditions and kinematic constraints contain \( \sim 2000 \) equations of motion. The predicted earthquake deformation with these equations are proved consistent with half-space analytical solutions.

4. Model calibration

The primary purpose of seismic source characterization is to resolve the spatial and temporal distribution of fault dislocations during earthquakes. Fault deformation models reveal fundamental elastic behavior of fault slip to interpret the observed quasi-static earthquake displacements. Geodetic data that map the surface deformation of an earthquake, are used to quantify the slip directionality, \( \theta_{\text{rake}} \) and magnitude, \( \Delta m \) of each subfault node (Figure 3) [63, 70]. These
data are regarded as the calibration targets, while the slip parameters are deemed as calibration parameters. To implement, the calibration procedure involves both forward modeling and inverse modeling [71]. The former yields predictions with respect to a particular input slip pattern, where the latter calibrates the pattern retroactively by fitting the geodetic observations [71]. The forward model is an essential component for inverse modeling as it signifies the deformation field induced by each individual subfault of an earthquake rupture (Figure 3) [3, 10]. Characterizing an earthquake source from the coseismic deformation field first involves finding the fault location, orientation and eventually the detailed slip distributions (Figure 3) [10]. The choice of calibration parameters strongly depends on the availability of data constraints and background information. Well-studied earthquakes having more data warrant more calibration parameters than those having relatively sparse information. There are some calibration parameters (e.g., \( \Delta m \)) which change linearly with the calibration data (e.g., \( u \)), while others (e.g., fault strike, dip and depth) vary nonlinearly with the geodetic data. For instance, doubling the slip magnitude doubles the ground movement observed in the data, whereas doubling the depth or dip of an earthquake fault of a constant slip magnitude does not double the resulting surface deformation [5]. The former could be easily calibrated via matrix inversion methods [71], while the latter necessitate exclusive sampling of a multidimensional nonlinear parameter space [10]. In particular, the latter require exponentially more samples when the loading entity has to be adequately characterized by a large number or DOF of nonlinear parameters. The number of random samples required to explore a suite of nonlinear parameters, thus, depends on the degree of nonlinearity and the algorithmic efficiency of stochastic optimization. A rule of thumb could start from bisecting the parametric space. For instance, a model of \( 7^2 \) freedom may requires \( 2^7 = 128 \) samples for each iteration of stochastic sampling [10]. While experience plays an important role in selecting a specific suite of nonlinear parameters, the calibration process and its solution convergence quantitatively control the precision of these calibration parameters [10]. The details are described in the following paragraphs.
4.1. Forward model

The predicted three-component displacement, \( d_i \), for location \( i \) due to the slip, \( m_j \), of subfault, \( j \), is:

\[
d_i = G_{ij} \ m_j \tag{6}
\]

where \( d_i \) is a three-component displacement vector \([d_{x,i}, d_{y,i}, d_{z,i}]^T\); \( m_j \) is a two-component slip vector \([m_{\text{dip,slip,}j}, m_{\text{strike,slip,}j}]^T\); the superscript T denotes the matrix transpose operator; \( G_{ij} \) is the unit-slip-displacement Green’s function between location \( i \) and subfault \( j \) based on the FEM predictions of unit fault slip embedded in a domain having a specified distribution of material properties (Figure 1). Note that the deformation is a linear function of slip and nonlinear function of the fault location, dip and strike as well as the distribution of domain materials.

The LOS displacement, \( d_{\text{LOS},i} \), for the \( i \)th InSAR pixel is a linear combination of contributions from the fault slip and a plane shift:

\[
d_{\text{LOS},i} = d_i V_i^T + p_1 x_i + p_2 y_i + p_3 \tag{7}
\]

where \( V_i \) is the line-of-sight (LOS) unit vector of the \( i \)th InSAR pixel; \( p_i \) are coefficients of a plane to account for the plane-shift displacements attributed to errors in modeling orbital effects [62].

The generalized matrix formulations for Eqs. (6) and (7) are respectively:

\[
d = G \ m \& d_{\text{LOS}} = [G \ V^T, x, y, 1] \ [m, p_1; p_2; p_3]^T \tag{8}
\]

where \( d \) is the column vector of displacement data; \( G \) is the integrated Green’s function matrix; \( m \) is the slip vector; \( d_i \) is the InSAR data column vector; \( V \) is the LOS unit column vector; \( x \) and \( y \) are the pixel location column vectors; \( 1 \) is a unity vector; The complete data vector that includes both GPS and InSAR data could be constructed by appending the matrices given in Eq. (8).

4.2. Inverse model

The common goal of inverse model is to estimate the calibration fault parameters based on the observed seismic data. While recognizing that a forward model is the linkage between the calibration data and the calibration parameters, inverse models step forward to optimize the calibration parameters and minimize the prediction errors against the calibration data. As mentioned above, those linear and non-linear calibration parameters are analyzed differently based their relations with the earthquake deformation. Our FEMs primarily contribute to the calculation of the Green’s function matrix, \( G \) (Eq. (8)), depending on the characteristic model configurations that may include nonlinear calibration parameters. Variations of the nonlinear calibration parameters typically exert asymmetric and non-systematic influence on \( G \), whereas those of linear calibration parameters multiply matrix entries with a constant factor.

4.2.1. Linear inverse analysis

With the consideration of both strike-slip (ss) and down-dip (dd) component, a complete Green’s function matrix becomes \( G = [G_{ss}, G_{dd}]^T \) and has dimensions of \( 2 \times M \times N \), given there
are \( M \) dislocating nodes-pairs (within the fault-slip region) and \( N \) displacement data points. Similarly, the dislocation vector has dimensions of \( 2M \) so that \( m = [m_{dd}, m_{ss}]^T \). Given \( m \) is a vector representing \( M \) dislocating nodes-pairs and \( d \) is a vector of \( N \) displacements, the under-determined problem of linear inversion (when \( 2M > N \)) always poses non-unique solutions of fault-slip models (Figure 3). Elastic dislocation problems for multiple slip patches generally have both over- and under-determined aspects - a given patch influences all data, and each datum constrains all patches. A physical solution can be hence resolved by simultaneously [71]: 1) estimating the slip distribution that minimizes misfit to geodetic data, 2) damping spurious solution oscillations, and 3) accounting for the uncertainties of geodetic data. To do so, first, we pre-multiply Eq. (6) by a weight matrix, \( W \) to account for the geodetic data uncertainties [71]:

\[
WG_m = Wd = G_wm = d_w \quad \text{and} \quad W_{ij} = \begin{cases} 
1/\sigma_i & \text{if } i = j \\
0 & \text{if } i \neq j 
\end{cases} \tag{9}
\]

where \( W \) is an \( N \times N \) weight matrix formulated from the reported 1-sigma uncertainties of geodetic measurements, \( \sigma_i \). \( \sigma_i \) is the uncertainty of the \( j \)th element of \( d \). Alternatively, \( W \) can be derived from data covariance matrix, \( C_d = (W^TW)^{-1} \) through Cholesky decomposition. Neglecting the data uncertainties implies that \( C_d \) and \( W \) are an identity matrix, for which the uncertainties and weights are unity for all data. Second, we can reconfigure Eq. (9) using second-order Tikhonov Regularization to damp the null space of the data kernel (smooth the fault-slip distribution, \( m \)) by:

\[
(G_w^TG_w + \beta^2L^TL)m = G_w^Td_w \quad \text{where} \quad L = \begin{bmatrix} L_{dd} & 0 \\
0 & L_{ss} \end{bmatrix} \tag{10}
\]

where \( L \) is a \( 2M \times 2M \) matrix of damping. For curved fault configurations, \( L \) is literally substituted by the global conductance matrix, \( L_{GCM} \) referring to finite element approximation of Laplacian operator, \( \nabla^2 m = 0 [10, 17] \). This global conductance matrix, \( L_{GCM} \) is the only mathematical tool to impose Laplacian regularization over slip locations of irregular fault geometries, for example, associated with the 2015 M6.2 Amatrice, Italy earthquake [10]. Using such Laplacian operator for smoothing allows us to conveniently impose and test Dirichlet (null; \( x = 0 \)) and/or Neumann (\( \partial m / \partial x = 0 \)) specifications along the boundaries of the rupture surface [10]. With desired geodetic points, \( d \), the Green’s function matrix, \( G \), and the global conductance matrix, \( L_{GCM} \) and weight matrix, \( W \) all essential components of the slip inversion are ready in Eq. (10) to solve the least-squares solution of \( m \) (Figure 3):

\[
m = (G_w^TG_w + \beta^2L_{GCM}^TL_{GCM})^{-1}G_w^Td_w \tag{11}
\]

where \( \beta \) is the regularization parameter controlling the tradeoff between minimizing misfit, \( \epsilon_w^T \epsilon_w \) (given \( \epsilon_w = d_w - G_wm \)) and satisfying \( \nabla^2 m = 0 [71] \). Some other workers recast the linear inversion through a Bayesian method to arrive at a physical solution, \( m [2, 72] \). A Markov-chain Monte Carlo (MCMC) method is used to sample numerous combinations of subfault slips, \( m_i \) and smoothing coefficient, \( \beta_i \) (and possibly relative data weights, \( \alpha_i \)) to create posterior probability density functions which in terms provide estimates of the above parameters
The uncertainty of the calibration parameters is characterized by the parameter covariance matrix, $C_m$ [71]:

$$C_m = (G_w^T G_w)^{-1} G_w^T d_w G_w (G_w^T G_w)^{-1} \quad (12)$$

Eqs. (11) and (12) provide a mechanism for providing estimates of central tendency and uncertainties for linear calibration parameters, in a way that accounts for the data uncertainties. From the 2015 M7.8 Gorkha, Nepal earthquake, the HOM domain without considering heterogeneous rock properties in calculating $G$ significantly degrades the fidelity of predicted GPS displacements beyond the data uncertainties (Figure 3), suggesting that a HET FEM domain is necessary for improving model predictions [3].

4.2.2. Nonlinear inverse analysis

This procedure is specially designed for nonlinear deformational parameters such as fault location, width, length, dip and strike to quantify the geometry and location of earthquake rupturing faults. As such, nonlinear inverse analyses are always conducted before the inverting for linear fault-slip parameters [10]. The solutions of those nonlinear parameters then later influence the accuracy of the linear slip solutions. For instance, uncertainties in fault dip propagate into the magnitude of subfault slip components such that a larger dip mistakenly resolved by the nonlinear analysis gives rise to larger slip magnitude predicted by the linear solutions. The nonlinear inverse method constitutes perturbing a nonlinear parameter and examining its impact on $G$ with numerous forward model predictions. The ultimate goal of such analysis is to resolve a set of nonlinear parameters that minimizes $e_w^T e_w$ [10]. There are many different sampling approaches, including classical grid search and probabilistic random search [73]. The former conducts the parameter search over a predefined grid to find an optimal solution [63]. However, this strategy is usually biased by the researcher’s expectations and achieves a poor solution resolution. On the contrary, the probabilistic type of Monte Carlo sampling randomly perturbs the solutions of a nonlinear parameter with more sophisticated and dynamic sampling strategies [10]. These regimes require re-computation of $G$ upon each suite of sampled parameters, and hence is more applicable to the earthquake models with few (e.g., less than 10) calibration parameters. Directed stochastic sampling methods such as Monte Carlo Markov Chain (MCMC) and Monte Carlo Simulated Annealing (MCSA), combine the effectiveness of gradient methods and adaptive random sampling to calibrate nonlinear fault parameters of earthquake sources [10].

In the 2016 M6.2 Amatrice, Italy earthquake (AE), Tung et al. [10] used the MCSA method to calibrate a few thousands of nonlinear parameters in FEM-based models of seismic deformation (Figure 4). In particular, both a planar and listric dislocation are examined through a series of nonlinear analysis to invert the InSAR data obtained by ESA Sentinel-1 A/B and JAXA ALOS-2 satellite, assuming a uniform slip distribution. On one hand, seven nonlinear parameters, namely, fault dip, $\delta$, strike, $\phi$, length, $L$, width, $W$, and fault-center location, $[x_c, y_c, z_c]$ are used to designate the geometry and location of a planar fault (Figure 4) [10]. On the other hand, the listric fault geometry is constrained by a set of 6 parameters, namely, listric parameter, $[a, b]$, locking depth, $D_m$, fault horizontal width, $H_x$, fault length, $L$ and fault location, $x_c$ [10].
Once these geometric parameters are fixed, the planar and listric models of distributed coseismic slip can then be derived linearly (Figure 5). The nonlinear analysis searches through a few thousands of uniformly-slipping fault models to minimize the weighted error misfit $e_w^T e_w$. This MCSA method combines simulated annealing [74] and nested Monte Carlo method [73] to search for a set of fault parameters that minimize $e_w^T e_w$. The cooling schedule of the MCSA algorithm is described by

$$T_i = T_o \left(\frac{N-i}{N}\right)^k + T_o \omega_{min}$$

$$v_{pi} = \left[ \left( \frac{T_i}{T_o} \right)^{k_p} + v_{pmin} \right] v_{bounded\_range}$$

where $T_o$ and $T_i$ refer to the temperature of initial step and $i^{th}$ step (0,1,2,...N-1) respectively; $\omega_{min}$ denotes the scaling factor of minimum temperature, $T_{min} = \omega_{min} T_o$, $N$ is number of iteration (step);

Figure 4. Solution convergence of resolving a planar source geometry in a HOM and HET domain for the 2016 M6.2 Amatrice, Italy earthquake [22]. The configurations of a planar dislocation specified by (a) width, (b) length, (c) dip, (d) strike and (e, f and g) the location of fault center $[x_c, y_c, z_c]$ are optimized by minimizing (h) the model misfit, $\chi^2$ between the predicted and observed LOS displacements. The samples over HOM and HET are denoted as gray and black dots respectively, while the solution convergence is denoted by blue and red lines respectively. The brackets denote the 1-sigma parameter range. The HET solution yielding smaller misfit is more compatible with the fault model (yellow dashed line) suggested by [5].
V_p refers to the p\textsuperscript{th} parameter’s search range in the i\textsuperscript{th} step; V_pmin and V\textsubscript{bounded range} are the last-step scaling factor and the bounded parameter range of the p\textsuperscript{th} parameter respectively; k and k_p respectively indicate the cooling/decaying coefficient of temperature and the search range of the p\textsuperscript{th} parameter. The cooling schedule T\textsubscript{i} is an array of decreasing number/temperature guiding the search of parameter at each step as well as the solution precision (Figure 4). Sharply changing (larger k and k_p) schedules narrows down the search range earlier, and allow the solutions to converge more rapidly. The unique advantage of MCSA over other Bayesian approaches like MCMC is its effectiveness of searching for the global optimal solution over the local minima of the model misfit. Its resistance to local minima is reinforced by acceptance criteria:

\[
\chi^2_i < \chi^2_{i-1} \tag{15}
\]

\[
\text{e}^{-\left(\chi^2_i - \chi^2_{i-1}\right)/\tau} > r \tag{16}
\]

where \(\chi^2_i\) and \(\chi^2_{i-1}\) refer to the mean error of the i\textsuperscript{th} step and i-1\textsuperscript{th} step respectively; r denotes a random parameter between 0 and 1; \(\tau\) is a normalization factor. The criterion described in Eq. (15) always accepts a better solution of the subsequent step, while the criterion in described Eq. (16) accepts a worse solution only when the misfit discrepancy, \(\chi^2_i - \chi^2_{i-1}\) between the subsequent steps is small. The cooling schedule T\textsubscript{i} is initially designed with a MCMC analysis that does not include the acceptance criteria described in Eqs. (15) and (16). We then retrospectively fine-tune the cooling schedule and the acceptance criteria in a manner until a consistent solution is obtained. A general principle of MCSA is fix the values of T\textsubscript{i} and \(\tau\) so that the initial few steps are able to escape the local minima and arrive at a desired solution precision after all calculation steps (Figure 4).

5. Conclusions

The innovative modeling protocols of FEMs are developed to satisfy the need of simulating realistic elastic earthquake systems. By taking advantage of the increasingly data availability of seismic and tomographic studies, complex fault geometry and distributed rock materials are revealed especially within the upper crust. The customary half-space models of fault deformation, which assume a homogeneous domain and rectangular dislocations, cannot fully
account for such shallow-crust complexity and hence induce prediction uncertainties when imaging earthquake sources with geodetic observations. New generations of fault model are fashioned in the framework of finite elements such that arbitrary lithological and structural heterogeneity can be accommodated when modeling seismic ruptures, which is particularly essential for earthquake locations of drastically changing lithology such as subduction margins. The modeling results of FEMs are found significantly more accurate than those of the conventional analytical solutions in nonlinear fault-geometry analyses and linear inversion for detailed slip distributions. This chapter, for the first time, describes the basic principles of constructing a sophisticated FEM for modeling elastic dislocation and elaborate how other auxiliary geophysical and geodetic data can be fed into the numerical domain and associated inverse analyses respectively. The resolution of governing equations and the corresponding validations are also discussed to ensure the reliability of the proposed FEM method. The modeling capacities of FEMs can further be extended beyond to simulate earthquake-induced poroelastic [75, 76] and viscoelastic [77] coupling processes which render physical mechanisms of triggering aftershocks and post-seismic surface deformation, summarizing the exceptional advantages of using FEMs for a wide range of earthquake research.

Acknowledgements

This work is supported jointly by a NASA grant NNX17AD96G, NSF grant 1316082, NSF grant OCE-1636653-subaward-4(GG013106-1) and NASA JPL subcontract 1468758. We would also like to acknowledge support from JAXA research program (0414001PI#3357).

Author details

Sui Tung1*, Timothy Masterlark1 and Daniel Sai Huen Lo2

*Address all correspondence to: sui.tung@sdsmt.edu

1 Department of Geology and Geological Engineering, South Dakota School of Mines and Technology, Rapid City, SD, USA
2 Department of Civil Engineering, The University of Hong Kong, Hong Kong SAR

References


