Uncertainties in Finite-Fault Slip Inversions: To What Extent to Believe? (A Critical Review)

by Igor A. Beresnev

Abstract   The matrix inversion of seismic data for slip distribution on finite faults is based on the formulation of the representation theorem as a linear inverse problem. The way the problem is posed and parameterized involves substantial, and often subjective, decision making. This introduces several levels of uncertainty, some of them recognized and some not adequately addressed. First, the inverse problem must be numerically stabilized and geologically constrained to obtain meaningful solutions. It is known that geologically irrelevant solutions also exist that may even better fit the data. It is also known that obtaining a stabilized, constrained solution compatible with the data does not guarantee that it is close to the true slip. Second, once the scheme has been set up, there still remains significant uncertainty in seismological parameterization. Synthetic tests have consistently shown that incorrect assumptions about the parameters fixed in the inversions, such as the rupture speed, fault geometry, or crustal structure, generate geologic artifacts, which are also dependent on array geometry. Third, solving the inverse problem involves numerical approximation of a continuous integral, with the generally grid-dependent result. Fourth, maintaining a linear inverse problem requires that the final slip on each subfault be the only variable to solve for. The slip functions used in the inversions are typically integrals of triangles or boxcars; they all involve a second parameter, slip duration, which has to be fixed. The effect of chosen duration cannot be disregarded, especially when frequencies higher than 0.1–0.5 Hz in the data are modeled. Fifth, the spectra of triangles and boxcars are sinc functions, whose relevance to realistically observed spectra is problematic.

How close then could an inverted slip image be to the true one? There are reasons to believe that the fine structure resolved is often an artifact, dependent on the choice of a particular inversion scheme, variant of seismological parameterization, geometry of the array, or grid spacing. This point is well illustrated by examining the inversions independently obtained for large recent events, for example, the 1999 İzmit, Turkey, earthquake. There is no basis currently available for distinguishing between artificial and real features. One should be cautioned against any dogmatic interpretation of inhomogeneous features on inverted slips, except their very gross characteristics.

Introduction

Inversion of seismic data for the distribution of slip on finite faults has become a popular tool for the reconstruction of faulting processes during large earthquakes. Slip distributions have been published for many major recorded events of the past two decades. In the interpretation of these slip patterns, much confidence is often given to the details of inverted slip, both spatial and temporal, to the extent that they become a standard on which new models of source processes are built (e.g., Heaton, 1990; Mendoza and Hartzell, 1998a; Somerville *et al.*, 1999). Although the derived generalizations may be consistent with known theoretical

scaling relationships (Somerville *et al.*, 1999), the fact is often overlooked that the slip inversions are obtained as a result of the solution of an inverse geophysical problem, which is inherently nonunique. To formulate a resolvable and physically meaningful slip-inversion problem, one has to resort to a number of assumptions and constraints that form the rigid framework of each particular application. Since these sets of constraints cannot be uniquely defined and involve a great deal of arbitrary decision making, any particular implementation chosen virtually controls the resulting solution. In many published applications, little or no

effort has been spent on ascertaining that the resolved image is at all relevant to the true one. Certain assumptions made to formalize elemental source processes may also be so simplistic that they depart from physical reality.

The purpose of this article is to specifically focus on the problem of the uncertainties characterizing the inversion of seismic data for slip distribution on faults. I will dwell on the nonuniqueness, sensitivity, and resolution issues that have been addressed in the past, as well as discuss other uncertainty sources that have not been adequately addressed and the problems that every inversion task will still face. My goal is to explore the reliability of the variations of slip that appear on finite-fault inversions, from which important geological inferences are often made.

Theoretical Background

The representation theorem for the earthquake source as a slip discontinuity in an elastic medium forms the basis for developing a formalized slip inversion (Spudich, 1980; Olson and Apsel, 1982). I will not reproduce the general problem formulation here but will use a simple form that reveals the physical issues that I would like to address.

A compact form of the representation theorem is given by Aki and Richards (1980 [equation 3.17]):

$$u_n(\mathbf{x},t) = \iint_{\Sigma} [u_i]v_j c_{ijpq} * \frac{\partial}{\partial \xi_q} G_{np} d\Sigma, \qquad (1)$$

where $u_n(\mathbf{x},t)$ is the nth component of the displacement at an observation point $\mathbf{x}$, $[u_i]$ is the ith component of the slip (displacement) discontinuity across the fault surface Σ, v_j is the jth component of the unit normal to the fault surface, c_{ijpq} is the tensor of elastic constants, G_{np} is the Green's tensor for the geometry of interest, the ξ variable belongs to the fault surface, the asterisk denotes convolution in time, and the summation is done over repeating indices. Following Aki and Richards (1980) and to obtain a physically transparent solution, I will assume a homogeneous isotropic space and analyze the far-field displacement for simplicity. Further assuming that the fault surface is planar and the direction of displacement discontinuity does not change along the fault, one obtains from equation (1) that the displacement waveforms of P and S waves are described by an integral of the form

$$\Omega(\mathbf{x},t) = \iint_{\Sigma} \Delta\dot{u}(\xi, t - r/c) d\Sigma, \qquad (2)$$

where $\Delta\dot{u}(\xi,t)$ is the time derivative of the slip function at a point on the fault (source time function), r is the distance from this point to the observation point, and c is the wave-propagation velocity (Aki and Richards [1980], equation 14.7).

Simple equation (2) illustrates the basis for the inversion of observed seismic data $\Omega(\mathbf{x},t)$ for the slip function $\Delta u(\xi,t)$. Approximating the continuous integral (2) as a sum, one can rewrite it as

$$\Omega(\mathbf{x},t) \approx \overline{\Delta\dot{u}}_1 \Delta\Sigma_1 + \overline{\Delta\dot{u}}_2 \Delta\Sigma_2 + \dots, \qquad (3)$$

where summation is done over all discrete elements (subfaults), and $\overline{\Delta\dot{u}}_i$ is the slip velocity averaged over the subfault. If we have enough observations $\Omega(\mathbf{x},t)$, the system of linear equations (3) can be solved for $\overline{\Delta\dot{u}}_i$. Note that, to achieve an accurate representation of the integral by the sum, the surface elements $\Delta\Sigma_i$ must be sufficiently small.

So far I have considered the case of unbounded homogeneous medium, for which the Green's function is simply a δ-function. For a realistic inhomogeneous half-space, the slip functions in equation (3) will be multiplied by the corresponding geometry-specific Green's functions (Olson and Apsel, 1982), while the essence of the inverse problem formulation will remain the same. In the most typical implementation, the problem is posed as a matrix equation equivalent to equation (3), in which the left-hand side is a matrix of observed waveforms and the right-hand side is the product of the matrices of Green's functions (synthetic waveforms) and slips. Note that, to formulate a linear inverse problem, the slip on each subfault should be represented by a single number (slip weight). The system of equations is then solved, in a minimum-norm sense, for slip weights on all subfaults. This is an important assumption whose validity will be discussed later.

Equation (2) also reveals fundamental nonuniqueness in the slip distributions that could be obtained using the described inversion approach. Let us take the limiting case of one observation station. Clearly, an infinite number of slip distributions over the fault could be found that would provide the same value of the integral. This uncertainty will decrease with the increasing number of observations available, although it would be hard to quantify how much of it exists for a given set of stations. What is clear is that meaningless results of inversion may be obtained based on a few stations only, and, to quantify the uncertainty, a study of the sensitivity of inversion to dropping or adding subsets of data must be a necessary part of every inversion process. In practice, this kind of analysis is seldom done, leaving the uncertainty nonquantified in many cases. The exceptions to this practice are rare and will be addressed in the text.

Uncertainty in Inversions: Issues Addressed

Uncertainties in Formulating a Resolvable Inverse Problem

Historically, the first attempt to apply the representation theorem to the inversion for slip on finite faults was apparently made by Trifunac (1974), who applied the method to five strong-motion records of the 1971 San Fernando, Cali-

fornia, earthquake. The author used a full-space geometry and a simple trial-and-error approach to fit the limited data. The study made no secret of the extreme nonuniqueness accompanying the inversion process and in fact suggested ways of considerable improvement of the inversion methodology by considering a stratified half-space geometry and larger databases. These improvements were implemented in the early, trial-and-error applications by Heaton (1982) and Hartzell and Helmberger (1982). These studies were still forthright in recognizing the nonunique and controversial character of deriving solutions from the limited data sets, as well as in acknowledging large arbitrariness in the assumptions that allowed one to constrain the method, a motive that seems to have all but disappeared in many later works. Crucial assumptions, for example, included postulating a certain form and duration of an elemental source time function $\Delta u(t)$, such as in equation (2) (to be able to parameterize it), and fixing the velocity of rupture propagation along the fault (to ensure correct timing of subfault triggerings). Not only do their different choices trade off with the resulting slip distribution, but the common choice of the slip function that allows the problem to be parameterized as a linear inverse problem is also physically problematic, as I will discuss later.

Admittedly, a milestone work in the development of finite-fault slip inversions was published by Olson and Apsel (1982). The value of this study was in that the authors were not as much interested in obtaining a slip distribution for a particular event as in the rigorous analysis of the stability and uniqueness of the solutions, using the data from the 1979 Imperial Valley earthquake as an example. To the extent of my knowledge, it was the first study in which the problem of slip inversion was considered on a formal basis of the linear inversion theory, as opposed to the trial-and-error waveform fitting that involved subjective judgment and did not appear to even address nonuniqueness in any quantitative way. Most subsequent slip inversions based on the representation theorem were virtually modifications of the method adopted by Olson and Apsel (1982). To support its conclusions, this work also used the largest strong-motion database to date, including records from 26 stations. For these reasons, the main results of this work are worth reformulating in the context of this analysis.

As discussed in the article, the solutions of the generalized least-squares inversion are unstable and nonunique. To ensure stability, sets of equalities are appended to the original matrix equations (equivalent to equation 3) that suppress large variations in the solution caused by small variations in data. By doing this, the original equations are being modified; the goodness of fit is thus sacrificed for the sake of stability. An important point to make is that the stabilization is not required to achieve a nearly perfect fit to the data; however, the obtained results could be geologically meaningless. The stabilization in fact degrades the match between synthetics and observations (see Olson and Apsel [1982] for specific examples). This inference virtually dis-

allows the trial-and-error approach as a meaningful way of solving the inverse problem.

Furthermore, the stabilization itself, while suppressing large variations in the solutions caused by noise in the data, does not still guarantee that the stabilized solution is geologically or physically meaningful. To ensure reasonableness from the seismologist's standpoint, the solution has to be further constrained by some inequalities following from common sense. These constraints may vary. For example, Olson and Apsel (1982) used the slip-velocity positivity condition, which allows slip on every cell to only increase (not reverse direction). Hartzell and Heaton (1983), who used a similar least-squares inversion technique, added a condition of smoothly variable slip to the positivity constraint: the difference between slips on adjacent cells is forced to nearly equal zero. Other typical constraints consist in looking for the most uniform or, conversely, most concentrated slip distributions, or those with the smallest average slip or average slip velocity over the fault, which would still satisfy the data; other variations could also be devised. They limit the range of allowable solutions to the most plausible ones but do not make them unique; the definition of plausibility is always a matter of choice.

A common constraint ensuring reasonableness of the inverted slip distribution is its match with the observed surface displacement for the events that ruptured the surface. Many events, though, do not provide surface ruptures. One more constraint, which, to the extent of my knowledge, has never been consistently implemented, is the requirement that the slip vanish toward the buried edges of the fault. This boundary condition is physically well reasoned, since it avoids creation of an infinite strain at the edges of the rupture and ensures that some physical mechanism actually caused the faulting to die down at depth. Many published inversions contain large slips terminating abruptly at fault edges, which seems unrealistic. One may conjecture that the implementation of this constraint could significantly change the slip distributions obtained without it, since a given boundary condition would inevitably propagate to the rest of the fault.

Olson and Apsel (1982) found four different solutions (least squares, stabilized least squares, constrained least squares, and stabilized constrained least squares), all of which provided satisfactory fit to the data, while bearing little resemblance to each other. Ironically, the best-fitting solution was the least-squares one without any stabilization and inequality constraints, being seismological nonsense. This conclusion was reiterated in many subsequent works. These results showed that the solution of a least-squares inverse problem became virtually a function of the adopted inversion scheme. Also, while a set of imposed conditions may ensure stability and reasonableness of the solution from the geological standpoint, it by no means will guarantee that the solution is even close to the true one.

Very similar conclusions were reached by Das and Kostrov (1990, 1994). For example, Das and Kostrov (1994) investigated the effect of commonly used plausibility con-

straits that ensured a geologically meaningful solution; they found that most of the localized slip concentrations (asperities) were simply artifacts of a constraining scheme. A dramatic demonstration was that even different earthquake mechanisms could be deduced from the same data, based on a particular chosen constraining scheme. For example, either a crack-propagation model (slip occurring in the vicinity of a localized propagating front) or an asperity model (slip continuing at a point after the front has passed) could be inferred from the same data.

Olson and Apsel (1982) and Hartzell and Heaton (1983) used very similar approaches to invert the data from the 1979 Imperial Valley earthquake, with the difference that the latter work used roughly half as many strong-motion stations and added teleseismic data. In spite of using the same technique, the resolved slip distributions were quite different (Olson and Apsel [1982], their figure 7; Hartzell and Heaton [1983], their figure 16). They were only consistent in the sense of overall gross value of slip, constrained by the moment. For example, Hartzell and Heaton (1983) deduced a localized asperity, while Olson and Apsel (1982) explicitly indicated that no asperities were required by the data.

One can conceptualize that there are two distinct levels of uncertainty in a formalized slip inversion. We could call them the method and the parametric uncertainty. The method uncertainty lies at the most fundamental level: it stems from the fact that there is no unique way of constructing an inversion scheme that would satisfy reasonable constraints imposed by both numerical stability and physics. We learn that, depending on which set of constraints is chosen and how it is implemented, the results of inversion may dramatically change. Once the method has been defined, the parametric uncertainty begins to affect the results. It is defined as the sensitivity of the results to a particular choice of the parameters fixed in the inversion (parameterization scheme). For example, to formalize the problem as a linear matrix inversion, one is generally constrained to fix the rupture-propagation velocity (some deviations will be discussed), fault geometry, crustal structure, and the duration of a source time function (a specific discussion of the typical choices of source time functions is also given later). Hartzell and Heaton (1983), Hartzell and Langer (1993), Das and Suhadolc (1996), Das *et al.* (1996), Saraò *et al.* (1998), and Henry *et al.* (2000) investigated this type of parametric uncertainty and found large variability in the inversion results depending on whether the parameters were fixed or allowed to vary in some specified way. Hartzell and Langer (1993) even concluded that false results could be obtained in the case of fixed parameters, casting doubt on prior works that used this assumption. The inferences from the studies by Das and Suhadolc (1996), Das *et al.* (1996), and Saraò *et al.* (1998), which deal with synthetic tests, are summarized in the next section.

The previous comments also apply to the studies originating from the original method of Hartzell and Heaton (1983) (e.g., Mendoza and Hartzell, 1989; Wald and Heaton, 1994; Wald *et al.*, 1996; and other investigations).

Uncertainties Revealed by Synthetic Tests

The soundness of obtained slip inversions is best tested if the inversion results are compared with the actual distribution of slip on the fault, which is impossible for natural earthquakes. The interpretations of many published inversions are then in effect based on the belief that they are close to reality, and one already could see that there are serious grounds for doubting this premise. In the absence of a possibility to compare the inversion to the true solution, the only way of testing the inversion algorithm would be to apply it to synthetic data obtained from the solution of a forward problem based on the representation theorem (a synthetic earthquake). The likeness of the inversion and the known solution would support the credibility of inversions of real earthquake data. There are surprisingly few published investigations of this kind (Olson and Anderson, 1988; Das and Suhadolc, 1996; Das *et al.*, 1996; Saraò *et al.*, 1998; Henry *et al.*, 2000; Graves and Wald, 2001), in which the authors were not as much interested in geologic interpretations of particular inversions as in looking into the accuracy of the method itself, which differentiates these studies from the majority of other applications. The results in fact detract significantly from the credibility of the method.

In one of the first thorough sensitivity studies, Olson and Anderson (1988) used a modification of the matrix inversion, in which the problem was transferred into the frequency domain, allowing greatly simplified bookkeeping and significant savings in computational time. The method was otherwise mathematically equivalent to the standard time-domain inversion. The authors investigated a simple problem of a uniform Haskell rupture propagating in a homogeneous space, which allowed analytical solution. It is not my purpose here to discuss the shortcomings of the model used by the authors; their study achieved its goals by providing a tool for the analysis of the relationship between the inverted slip images and the true solution, on which I will dwell.

The authors generated synthetic time histories from the model rupture for three characteristic near-fault array geometries, composed of 13–16 stations, looking into the effect of network geometry on the strong-motion inversion results. All three data sets were inverted using the standard method and compared with the true solution. The results, perhaps for the first time, revealed stunning uncertainties. A look at the inverted images of just the final slip (Olson and Anderson [1988], their figures 5–7) demonstrates that none of them reproduced the real static slip with any acceptable degree of accuracy. The model static slip was a uniform strike-slip offset with zero dip component. Systematic biases were introduced into the inverted images, which in addition depended on the array geometry. One array generated artificial asperities, nonexistent in the model. All three arrays generated false depth dependence of slip. All three generated a

spurious dip component. Importantly, all this occurred while the observed and simulated seismograms matched exactly. If one tried to draw practical conclusions from Olson and Anderson's (1988) analyses, it would be that most of the fine features on the inverted slip were actually array-dependent artifacts. Another conclusion is that, within the limits of the geometries considered, it would be hard to define an optimum array that would ensure better results than the others. For practical purposes, it is also important to note that there is no way of distinguishing between the artifacts and the real features of faulting. Even gross features, such as the total average slip or peak slip velocity, were recovered with errors. Note that these conclusions were drawn for the idealized problem, which involved uniform slip across the fault and no noise in the data.

Further sensitivity analyses were conducted by Das and Suhadolc (1996), Das *et al.* (1996), and Saraò *et al.* (1998) using a similar Haskell-type rupture with uniform slip, and Henry *et al.* (2000) using simple inhomogeneous slip distributions, as their synthetic models. The approach was to fix certain model parameters at their true values and investigate the effect of incomplete knowledge of others on the ability of the inversion to retrieve the true fault image. In particular, the authors investigated (1) the effect of incorrect assumption of the rupture speed, fault geometry, and crustal structure (Das and Suhadolc, 1996; Saraò *et al.*, 1998), (2) the effect of particular near-fault station distribution (Saraò *et al.*, 1998), and (3) the effect of adding noise to synthetic data (Saraò *et al.*, 1998; Henry *et al.*, 2000). One finds rather pessimistic conclusions as the outcome of these studies. For example, it was found that the near-field station geometry virtually predetermined the resulting solution and that adding extra stations could sometimes even worsen it (Saraò *et al.*, 1998). This inference echoed the conclusion of Olson and Anderson (1988) that it was hard to define the geometry of an optimum array. It should be noted, though, that this conclusion may not be true for the teleseismic inversion, for which Hartzell *et al.* (1991) did not find as much variability in the solutions based on station configuration. However, one also should keep in mind that the inversions of teleseismic data do not provide as much resolution power as strong-motion inversions; at teleseismic distances, most faults could safely be considered point sources.

Incomplete knowledge of crustal structure could ruin the inversion so that no realistic part of the real fault was recovered (Das and Suhadolc, 1996; Saraò *et al.*, 1998). Incorrect assumptions about the parameters that were varied or the addition of noise to the synthetics produced geologic artifacts, such as nonexisting asperities, spurious fault inhomogeneity, or ghost (secondary) rupture fronts (Das and Suhadolc, 1996; Henry *et al.*, 2000).

Sekiguchi *et al.* (2000) and Graves and Wald (2001) also specifically addressed the effect of incomplete knowledge of crustal structure (Green's function in equation 1) on their ability to reconstruct the true synthetic images. Sekiguchi *et al.* (2000) concluded that a mere 3.5% misestima-

tion of the average seismic velocity in the crustal profile generated large errors in the resulting solution. Note that, in reality, crustal velocities will be much more uncertain. Graves and Wald's (2001) conclusions are not as pessimistic, although they pointed out that only gross features of slip can be recovered in case of incomplete knowledge of velocity structure. Cohee and Beroza (1994) similarly indicated that, if the Green's function fails to explain the wave propagation in the inversion passband, the source-inversion results may be expected to have significant errors.

It is important to keep in mind that any studies of this kind, which attempt to investigate the effects of model parameters one at a time, inevitably leave out the question of their complex interaction, when, as in the real problem, all of them are variable and none are exactly known. With a large number of governing parameters, the full gamut of their potential trade-offs could not even be approached.

The synthetic tests, such as described, provide important guidance in the evaluation of the accuracy of slip inversions. Das *et al.* (1996) and Henry *et al.* (2000) nevertheless emphasized their limited practical value. The synthetic earthquake models used in the tests involve smooth and simple ruptures; most of them have dealt with uniform constant slips. The conclusions obtained from such tests may be irrelevant to the inversion of real data. For example, the synthetic tests do not provide a tool to address the problem of inadequate discretization of the continuous problem, discussed in detail in the next section. As I noted earlier, the approximate value of the representation integral (3) will generally depend on the cell size for any realistic slip distribution. Clearly, for a uniform or smooth synthetic slip, the choice of $\Delta\Sigma$ will not realistically matter as the cell size may be rather crude, which precludes a representative sensitivity analysis. Das *et al.* (1996, p. 176) wrote, "This paper demonstrates the difficulties we encounter even in the simple case of a Haskell-type faulting model. Clearly more realistic models . . . would present even greater difficulties and the current approach of solving the inverse problem used here may not even be usable." Echoing similar conclusions of Olson and Anderson (1988), Das and Kostrov (1994) and Das and Suhadolc (1996) cautioned that only gross features may be reliable in realistic inversions and that these persisting gross features could in principle be obtained by varying the parameters and observing which properties of the solution remained unchanged. We find similar cautioning statements in Hartzell and Liu (1996). However, this process, although the only reasonable way to establish which features of the inversion could be real and which could be artifacts, will still not guarantee that such gross features can in fact be obtained. As more parameters are varied and none of them exactly known, there may be no repetition of gross features at all.

Henry *et al.* (2000) advocated the smoothest solutions, provided by the use of the positive and smoothly variable slip constraints, as the most reliable ones and argued against any complexities in the inverted models. The same approach

was taken, for example, by Mendoza and Hartzell (1988b) and Hartzell *et al.* (1991). Henry *et al.* (2000, pp. 16,111, 16,113) suggested, "The fact of this freedom [a very large number of unknown parameters] might lead to the expectation of a highly complex rupture model, with overfitting of noise. However, . . . our preferred solution has a very simple rupture process." This inference virtually excludes, as in the work by Olson and Anderson (1988), the derivation of asperities as complexities in the solutions, which, as the body of the reviewed tests has consistently shown, are in most cases the artifacts.

Uncertainty in Inversions: Issues Not Addressed

Discretization of Continuous Problem

As equations (2) and (3) demonstrate, the formulation of the linear inverse problem based on the representation theorem approximates a continuous integral over the fault plane by a sum, which, to be accurate, requires sufficiently small subfault areas. How small the grid size in equation (3) must be depends on how irregular the actual slip distribution is, which is an unknown function. If the cell area $\Delta\Sigma$ is not chosen small enough, the approximation (equation 3) may have little to do with the exact value of the integral. The result of the slip inversion will therefore generally depend on the subfault size. It follows that the rigorously conducted inversion should involve a sensitivity analysis that would reduce the subfault size until the convergence of the sum is achieved. This process does not guarantee that the convergence will be obtained at the sizes that are still computationally tractable; however, if the convergence is not obtained and the computational limitations are exceeded, the results of the inversion may be erroneous.

This point seems to have been overlooked in most finite-fault slip inversions. In the early implementation of the method, one does find the analyses of the effect of gridwork spacing on the calculated sum. For example, Hartzell and Helmberger (1982) reduced the cell size until there is no further change in the sum. The authors deduced the maximum spacing of 0.5 km, which was not allowed to increase. This approach guaranteed that there were no errors in the solution caused by the inaccurate representation of the integral. Olson and Anderson (1988) also specifically mentioned the need to achieve sufficient accuracy of the discretization and used an even smaller size of 0.2×0.2 km. This cell size is the smallest that I have been able to find in the published inversions for large faults, and in the case of this particular sensitivity study, it was well justified, because the true slip distribution was known and was uniform over the fault. The gridwork size sufficient to accurately approximate the integral (2) could then be roughly estimated. Such estimation would be impossible for real earthquakes. The technical implementation of a finer grid in the work by Olson and Anderson (1988) was made possible because of the

treatment of the problem in the frequency domain, which considerably reduced computer-storage requirements.

Hartzell and Helmberger (1982) and Hartzell and Heaton (1983) both inverted the data for the Imperial Valley earthquake. However, Hartzell and Heaton (1983) already utilized a subfault size of 3×2.5 km, much larger than 0.5×0.5 km reported by them earlier, which seemed to have violated their former work's conclusion that spacing greater than 0.5 km did not ensure convergence. There is no mentioning of the convergence analyses in the latter study, either. Note that the study by Hartzell and Heaton (1983) served as the basis for a series of similar inversions, in which subfault sizes ranged from 1.29×1.71 km (Wald *et al.*, 1996) to 15×13.9 km (Mendoza and Hartzell, 1989). In the absence of any indication of the convergence analyses, these values seem to have been arbitrarily chosen, perhaps on the grounds of the need to keep the problem computationally tractable. The effect of such choices on the accuracy of the derived inversions remains unclear.

Das *et al.* (1996) and Saraò *et al.* (1998) conducted the sensitivity studies in their synthetic fault models, in which they used finer discretization in the forward problem to construct synthetic data and coarser discretization to invert these data. This was done to infer the effect of the inversion algorithm not knowing the cell size used to generate synthetics. Note that these studies provide a first insight into the problem but do not exactly address the effect of inadequate discretization on the inversion of real data that I discuss here. The difference is that, as I noted at the end of the previous section, because the synthetic tests used very simple, uniform slip distributions, they did not realistically suffer from the effects of insufficient discretization. The realistic slip distributions are different in a sense that their integral sums (equation 3) are dependent on the cell size in some unknown way, which every application should attempt to investigate.

As I mentioned earlier, one problem with a consistent implementation of such sensitivity analyses is that large grids may not be computationally feasible. Another pitfall is that, as discussed by several authors, the stability of the inverse problem decreases as the size of the grid increases (e.g., Das *et al.*, 1996; Saraò *et al.*, 1998). The way the inverse problem is numerically formulated may even preclude rigorous investigations of the grid-size effects because of the growing instability. Note that this problem was encountered by Saraò *et al.* (1998) for a uniform-slip synthetic model; the difficulties may be further exacerbated in more realistic inversions.

A class of studies seems to have supplemented the investigation of the effect of refining the grid by distributing a large number of point sources (Green's functions in equation 1) over the area of each subfault. To avoid massive computations, Green's functions were interpolated between a precalculated master set. This, again, does not address the effect of inadequate fault discretization I discuss in this section. The refinement of Green's functions improves spatial discretization for a better approximation of the propagation-

path effect (and thus should be considered in the context of improving inadequate representation of crustal structure), but not the fault discretization grid to achieve an accurate integral sum (3). This is illustrated by equation (2), in which, for simplicity, a full-space geometry has been assumed with the same Green's function for every point on the fault (no need for interpolation). Green's function is absent altogether from this equation, but the problem of inadequate fault-plane discretization to achieve an accurate sum remains.

The need for fine spatial discretization will arise for more complicated (such as stratified) crustal structures and will add another source of uncertainty to the inversion. How dense should the spatial sampling of Green's functions be to avoid this type of uncertainty? The authors of relevant studies seem to have addressed this issue in an empirical manner, in which the number of point sources distributed over a subfault could range from 25 (Wald *et al.*, 1996) to 225 (Mendoza and Hartzell, 1989), depending on the sub-fault size used. There seems to be no quantitative basis established to warrant a specific choice, and sampling density will obviously depend on the complexity of the crustal model. The need for quantifying spatial sampling of Green's functions certainly exists; however, it is not related to the need for adequate fault discretization emphasized in this section.

Finally, a view exists in the literature that, if low-pass filtered ground motions are inverted, a coarse fault-discretization scheme can be used. This view is misleading as it assumes that the high-frequency ground motions result from the small-scale irregularities of slip on the fault. This would, for example, imply that ruptures with spatially uniform slip would not radiate high-frequency ground motions, which is obviously not the case.

The Fourier transform of equation (2) is

$$\Omega(\mathbf{x},\omega) = \iint\limits_{\Sigma} \Delta\dot{u}(\xi,\omega) \exp(-i\omega r/c)d\Sigma, \quad (4)$$

where ω is the angular frequency and $\Delta\dot{u}(\xi,\omega)$ is the Fourier spectrum of the time derivative of the source time function. It follows that the spectral properties of the source time function, not smoothness of slip distribution, control the spectra of ground motions. This can be directly seen by performing the integration of equation (4) for simple fault geometries (such as a rectangular fault) in the assumption of unidirectional (Haskell-type) rupture propagation (Aki and Richards [1980], equation 14.18). The integration provides the ground-motion spectrum that is governed by $\Delta\dot{u}(\xi,\omega)$, while the exponential factor in the integrand (4) gives the directivity pattern.

In the low-frequency limit, $\Delta\dot{u}(\xi,\omega)$ yields $U(\xi)$, where U is the final slip value (see discussion of equations 5–7 below; also, Aki and Richards [1980, p. 806]), and the integral (4) is reduced to $\iint\limits_{\Sigma} U(\xi)d\Sigma$. This shows that, even to accurately reproduce the low-frequency spectra, the fault-

discretization scheme should be sufficiently fine if $U(\xi)$ is highly heterogeneous [of course many equivalent slip distributions $U(\xi)$ may produce the same value of the seismic moment; this only emphasizes nonuniqueness in inverting the low-frequency spectra].

Choice of Source Time Function

Inadequate Underlying Spectrum. To maintain a liner inverse problem in equations (2) and (3) (and their analogs for more complicated geometries), the dislocation time history $\Delta u(t)$ has to be chosen in such a way that it is controlled by only one parameter, the total dislocation at the cell. This allows formulation of the matrix equation that is then solved for dislocation weights at each cell. This approach assumes that elementary slip history at a cell is uniquely defined by its final value (Olson and Apsel, 1982; Hartzell and Heaton, 1983).

In most implementations of the finite-fault inversions, the source time function is chosen as an integral of a triangle with a certain duration t_0 (Hartzell and Heaton, 1983), typically an isosceles triangle (Wald and Heaton, 1994; Wald *et al.*, 1996). Let us consider the latter case for simplicity. It is easy to verify that the dislocation time history at each cell will have the form

$$\Delta u(t) = \begin{cases} 2U(t/t_0)^2, & 0 \le t \le t_0/2 \\ U[1 - 2(1 - t/t_0)^2], & t_0/2 < t \le t_0 \end{cases}, \quad (5)$$

where U is the total (static) dislocation value, reached at $t = t_0$. The exact form of the triangle, as a function of U and t_0, is given by the differentiation of equation (5):

$$\Delta\dot{u}(t) = \begin{cases} (4U/t_0^2)t, & 0 \le t \le t_0/2 \\ (4U/t_0^2)(t_0 - t), & t_0/2 < t \le t_0 \end{cases}. \quad (6)$$

The finite-fault radiation is calculated by substituting equation (6) into equation (2).

The radiated signal in a form of triangle is clearly not strictly physical, which leads to its certain distinct spectral features. The modulus of the Fourier transform of equation (6) is the squared sinc function,

$$\Delta\dot{u}(\omega) = U\left[\frac{\sin(\omega t_0/4)}{\omega t_0/4}\right]^2. \quad (7)$$

Its first null occurs at the frequency $f = 2/t_0$. It is agreed upon in seismology, based on the success of the Aki–Brune source model in explaining earthquake spectra, at least at small to moderate magnitude levels, that the radiated spectra more closely follow the ω^{-2} function, $U/[1 + (\omega\tau)^2]$, where $1/\tau$ is the corner frequency (Aki, 1967; Brune, 1970; Boore, 1983). Assuming that the ω^{-2} function is a more realistic model for the observed spectra, one makes an error in approximating them by function (7). To estimate the amount

of this error, one can directly compare the two spectral shapes for the typical values of parameters.

Let us take $t_0 = 1$ sec (Wald and Heaton, 1994); we then have to find τ of the equivalent ω^{-2} source, having the same duration, to compare the two spectra. The time function of the dislocation that radiates the ω^{-2} spectrum has the form $\Delta u(t) = U[1 - (1 + t/\tau) \exp(-t/\tau)]$ (e.g., Beresnev and Atkinson [1997], equation 6). This function formally has unlimited duration. However, we can define the duration (T) as the time it takes for the dislocation to reach 90% of the total value. We then have a simple equation $(1 + T/\tau) \exp(-T/\tau) = 1 - 0.9$, from which $T/\tau \approx 4$, or $\tau \approx T/4$, which relates τ to source duration. The ω^{-2} dislocation, equivalent to the triangular function with duration $t_0 = 1$ sec, then has $\tau \approx t_0/4 = 1/4$ sec.

Let us compare the spectrum (7) ($t_0 = 1$ sec) with the ω^{-2} spectrum ($\tau = 1/4$ sec) by dividing the former by the latter. Figure 1 plots the resulting ratio. The ratio oscillates and is not significantly different from unity at very low frequencies only. Wald and Heaton (1994) used the triangular function to model radiated pulses up to the frequency of 0.5 Hz. If the radiation follows the ω^{-2} spectral law, the maximum error made is about 30% at 0.5 Hz. However, serious problems arise if one attempts to model slightly higher frequencies with the same triangular function, since its spectrum vanishes toward 2 Hz. In this case, the forward model with near-zero spectral energy around the node will try to reproduce finite energy in the realistically observed ground motions; the consequences of this are hard to predict. This may be the case in the study by Hartzell and Helmberger (1982), who used a 1-sec triangular function to model frequencies up to 2 Hz. The upper frequencies of their synthetic wave field thus have zero spectral energy. Another potential difficulty is found in the study by Mendoza and Hartzell (1988b). The authors reported problems in inverting frequencies higher than 0.5–1 Hz in solving for slip distribution for the 1985 Michoacan, Mexico, earthquake, which forced them to exclude short-period data from the inversion. The authors used a 2-sec triangular function, which has its first spectral node at 1 Hz. One could hypothesize that this could be the reason for the unsuccessful short-period inversion, since there was no energy around 1 Hz in their model-generated waveforms. Finally, Hartzell et al. (1991) reported a significantly degraded match of the data by the inversion that used a 3-sec duration triangle compared with the one with a 1-sec triangle. The upper frequency in their data was 1 Hz. One finds that, in the latter case, the first spectral node in the forward model was at 2 Hz, or beyond the frequency range of the data, while it fell at approximately 0.7 Hz in the former case, which was in the frequency range of the data. This again might explain the degradation in the quality of inversion using the 3-sec duration.

The problem is further exacerbated if the elemental boxcar (rectangular) functions are used instead of triangles to model $\Delta \dot{u}(t)$. The amplitude spectrum of a boxcar of width t_0 is again a sinc function, whose first two nodes occur at
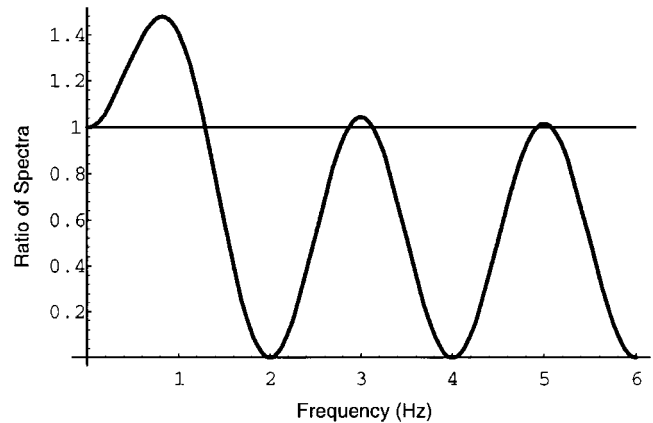

Figure 1. Ratio of the amplitude spectrum of isosceles triangle to the spectrum of ω^{-2} dislocation of equivalent duration. The width of the triangle is 1 sec.

$f = 1/t_0$ and $2/t_0$. Hartzell and Langer (1983) utilized $t_0 = 2$ sec and included frequencies up to 1 Hz in their inversion, which consequently cover both spectral nodes at 0.5 and 1 Hz. Das and Kostrov (1990) utilized boxcars with $t_0 = 5$ sec (first node at 0.2 Hz) and included frequencies up to 0.5 Hz, and Hartzell and Liu (1996) utilized boxcars with $t_0 = 1$ sec (first node at 1 Hz) and included frequencies up to 5 Hz. In the work by Bouchon et al. (2002), boxcar t_0's are permitted to vary between 0.25 and 5 sec, yielding frequencies of the first nodes between 0.2 and 4 Hz. With these synthetic functions, they match unfiltered strong-motion data with frequencies up to 25 Hz. As one can see, several spectral nodes may occur in the modeled frequency range if the boxcar elemental functions are used. It remains to be seen how this deficit in synthetic spectral energy will be handled by a formal inversion algorithm, when it attempts to match real data with null spectra in certain frequency intervals.

This analysis shows that there is interplay between the shape and duration of the assumed subfault time functions and the frequency range of the data that cannot be ignored. The shape and duration and the frequencies modeled cannot be considered independent of each other. This point may have been overlooked in some of the published inversions.

I have used the ω^{-2} elemental spectrum to illustrate the problems of inadequate underlying spectrum. As I stated earlier, this spectrum is admittedly the closest representation of the observed spectra among other possible shapes; however, any other realistic theoretical function (e.g., Ji et al., 2002, their equation 2) would serve the same illustration purpose. The conclusions of this analysis, merely emphasizing the unphysical lack of energy near nodal frequencies for the triangular and boxcar elemental functions, which extends to the modeled frequency range, will remain the same.

Dependence of Underlying Spectrum on Two Parameters. I have just discussed the possible implications for finite-fault inversion of using the underlying spectrum that is probably

not the one realistically observed. Another problem arises from treating the underlying spectrum (equation 7) as a function of only one parameter U instead of both U and t_0. As we saw earlier, this expedient arises from the need to formulate a system of linear equations in which the only variables to solve for are the dislocation weights U at subfaults. The parameter t_0 is assumed constant. Clearly, the function (7) is not independent of t_0; its shape will rapidly change proportionally to t_0^2. The inversion results will thus be generally dependent on its specific choice.

One could evaluate the effect of different choices of the triangle width t_0 on the shape of the spectrum by dividing the spectra (7) calculated for different t_0. Figure 2 shows the ratio of the spectrum calculated at $t_0 = 2$ sec to that calculated at $t_0 = 0.6$ sec. These widths are those used by Mendoza and Hartzell (1989) and Wald *et al.* (1996), respectively.

Figure 2 shows that the underlying spectra assumed in the inversion are approximately independent of t_0 (ratio equal to 1) at very low frequencies only. The parameter t_0 is in total control of higher-frequency spectra. At 0.5 Hz, the uncertainty caused by the two different choices is around 50%, and the uncertainty becomes infinite at 1 Hz where the spectrum at $t_0 = 2$ sec vanishes.

This difficulty again becomes more pronounced if the boxcar time functions are assumed. Their widths in the inversions for large earthquakes can range from 1 sec (Hartzell and Liu, 1996) to 5 sec (Das and Kostrov, 1990), any of which would be equally plausible for real earthquakes. If we divide the spectra of boxcar functions with $t_0 = 5$ sec and 1 sec, we obtain the ratio plotted in Figure 3. The spectra in this case are roughly independent of the choice of t_0 at extremely low frequencies below 0.1 Hz only; at any frequency above, the result of the inversion could be expected to depend heavily on which particular value of the boxcar width had been chosen.

Since there is no single reasonable choice of the width of the dislocation time history, the uncertainty of this kind will enter the result of the inversion. If all inversions had been performed for frequencies below approximately 0.5 Hz for Figure 2 or 0.1 Hz for Figure 3, the uncertainty would probably still have been tolerable. However, the current tendency is to extend the frequencies treated by the inversions into the range where the synthetic pulse becomes a function of both U and t_0. The only remedy in this case is to treat both U and t_0 as free parameters and solve for both, which no longer allows a linear inverse-problem formulation. The inversion may be achieved, for example, by using a different, grid-search-type approach, but not through the traditional linear matrix inversion. The point emphasized by this analysis is that it may sometimes have been overlooked that the extension of formalized inversion into the high-frequency range is not just a matter of using more powerful computers; the parameterization scheme must be revised in this case to make sure it is valid for higher frequencies as well.

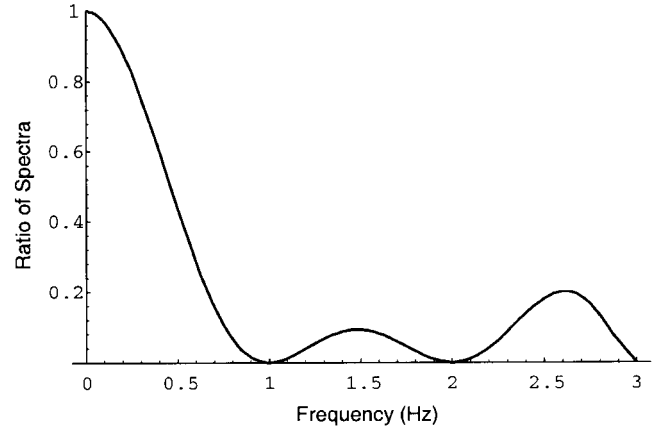Some later works using the linear inversion approach

Figure 2. Ratio of the amplitude spectra of isosceles triangles with different widths. The spectrum calculated for $t_0 = 2$ sec is divided by the spectrum calculated for $t_0 = 0.6$ sec.
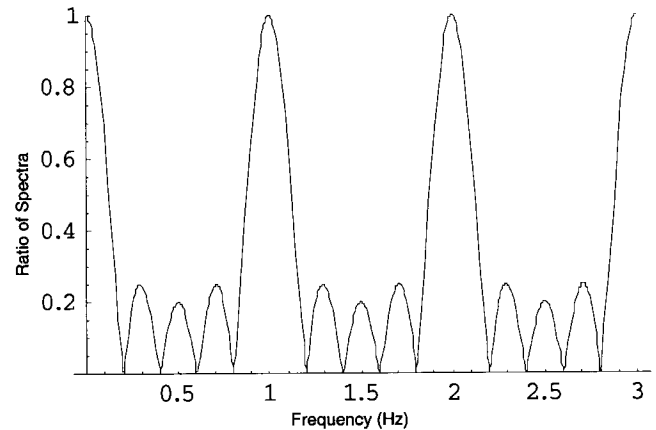

Figure 3. Ratio of the amplitude spectra of boxcar functions with different widths. The spectrum calculated for $t_0 = 5$ sec is divided by the spectrum calculated for $t_0 = 1$ sec.

have modified pulse parameterization by subdividing the subfault rise time into multiple windows, allowing separate elemental slip in each. The total subfault slip is thus represented as a sum of delayed triangles (e.g., Wald and Heaton, 1994; Wald *et al.*, 1996; Cho and Nakanishi, 2000; Chi *et al.*, 2001) or boxcars (Hartzell and Langer, 1993; Sekiguchi *et al.*, 2000; Sekiguchi and Iwata, 2002). The authors' rationale behind using this technique is that in such a way the dislocation time function of apparently arbitrary complexity and variable duration can be constructed, if the number of elementary blocks is great enough (10 windows were allowed by Hartzell and Langer [1993]). Since the first window does not have to have a nonzero slip, this approach also allows delayed subfault triggering mimicking a locally variable rupture-propagation velocity, although the velocity is a constant inversion parameter. Note that this approach still has elemental slip weights in the subfault windows as the

only variables to solve for, not changing anything in the linear matrix-inversion formulation (Wald *et al.*, 1996). The cost of the multiwindow modification is the greatly increased number of variables, though, which becomes the number of subfaults times the number of windows.

One can notice that the spectrum of each elemental subevent in a window is still one of a triangle or a boxcar; the modified scheme thus does not correct the problem of an inadequate underlying spectrum. The inverted dislocation time histories are composed of a number of start–stop phases, which generate the same artifacts in the spectra. The inversion dependence on the choice of t_0 for each elemental window is not removed either, which is the greater the higher the frequency.

A Recent Example: The 1999 İzmit, Turkey, Earthquake

In the context of evaluating the uncertainties in finite-fault slip inversions, it would be instructive to analyze a case of a recent large earthquake, for which inversions have been independently obtained by a number of different researchers. I use the case of the catastrophic *M* 7.4 1999 İzmit, Turkey, earthquake, which was extensively studied and for which five inversions of seismic data were recently published (Bouchon *et al.*, 2002; Delouis *et al.*, 2002; Gülen *et al.*, 2002; Li *et al.*, 2002; Sekiguchi and Iwata, 2002). Specific inversions used teleseismic data (Gülen *et al.*, 2002; Li *et al.*, 2002), strong-motion data (Bouchon *et al.*, 2002; Sekiguchi and Iwata, 2002), and jointly geodetic, teleseismic, and strong-motion data (Delouis *et al.*, 2002).

Figure 4 combines the five inversions on the same horizontal scale, aligned to have a common hypocenter. The vertical reference lines are drawn through −40, 0, and 40 km distances along the strike. The figures were taken from the original articles without editing; the individual slip scales were retained.

Perusal of the five slip distributions reveals that they have in fact little in common, as far as the specific patterns of slip distribution or the size and location of individual asperities are concerned. For example, large slips extend to maximum depths in Figure 4a,c,d, while slip is mostly surficial in Figure 4b. Figure 4e places most of the moment release in the hypocenter, while the hypocentral area is entirely free of slip in Figure 4b. Figure 4c reveals three principal asperities (in addition terminating abruptly at the edges of the fault), while Figure 4a contains only one, located in the area where there are no asperities in Figure 4c. The average size of the asperity would also be different if calculated separately from the distributions a–e.

Each slip distribution interpreted separately is liable to lead to much different geological inferences. An example of mutually exclusive interpretations is given by the fact that Bouchon *et al.* (2002) and Sekiguchi and Iwata (2002) invoked supershear rupture propagation (rupture velocity exceeding shear-wave velocity) as a result of their inversion, while Delouis *et al.* (2002) stated that no supershear is needed to explain the data. The only consistent feature among the slip distributions is the maximum value of slip on the fault; however, this value is largely constrained by the seismic moment.

It would be hard to conclude which one of the five inversions in Figure 4 is more reliable and which one is less. Individual studies invert different data sets using different parameterization schemes and sets of constraints; each combination may have brought about its own set of uncertainties and possible artifacts. Among the five studies, only one contains an extensive study of artifacts and the reliability of inversion using synthetic tests before the inversion of real data is shown (Delouis *et al.*, 2002). It is not my purpose to advocate their particular solution (for example, the authors do not use smoothing constraints strongly recommended by many others), but rather to emphasize the implications for the inversion of the artifacts revealed by this study. Delouis *et al.* (2002) used a synthetic slip distribution to test resolvability of its features using their inversion algorithm, inverting the three components of their database (geodetic, teleseismic, and strong-motion data) separately and jointly. Their conclusions are that (1) the teleseismic inversion gives an entirely wrong picture (the shape and even the presence of all asperities cannot be retrieved), (2) the strong-motion inversion does marginally better in resolving the shape but does not resolve the deeper part of the fault, and (3) the joint inversion works significantly better but still does not resolve the deeper part of the fault. The match between the synthetic and observed waveform is perfect in all cases and is even better in separate inversions. Note that the result of the joint inversion by Delouis *et al.* (2002) is very different from all other inversions in Figure 4. The resolution findings prompted the authors to recommend resolution and sensitivity tests for any kinematic inversion of real data, as the only way to avoid interpretation of artifacts, reemphasizing the point raised by Olson and Anderson (1988) more than a decade earlier.

Later inversion studies tend to use combined data sets (e.g., seismic and geodetic as in Delouis *et al.*) in order to provide more constraints on the resulting solutions. Of special interest in the context of this analysis are the related sensitivity studies utilizing synthetic ruptures, which present inversion results of each data set independently and compare them with the combined inversion to assess what may be reliably interpreted (e.g., Wald *et al.*, 1996; Wald and Graves, 2001; Delouis *et al.*, 2002). Yet other studies add observed surface faulting (if present) as an observational constraint. Delouis *et al.* (2002, their figures 11 and 12) provided a comparison of the results of both independent and combined inversions with and without surface-offset constraint. A common thread of these studies is that, all other conditions being equal, the addition of geodetic constraints to seismic data helps improve resolution of static slip (Cohee and Beroza, 1994; Wald and Graves; 2001; Delouis *et al.*, 2002). However, adding geodetic data does not play a piv-
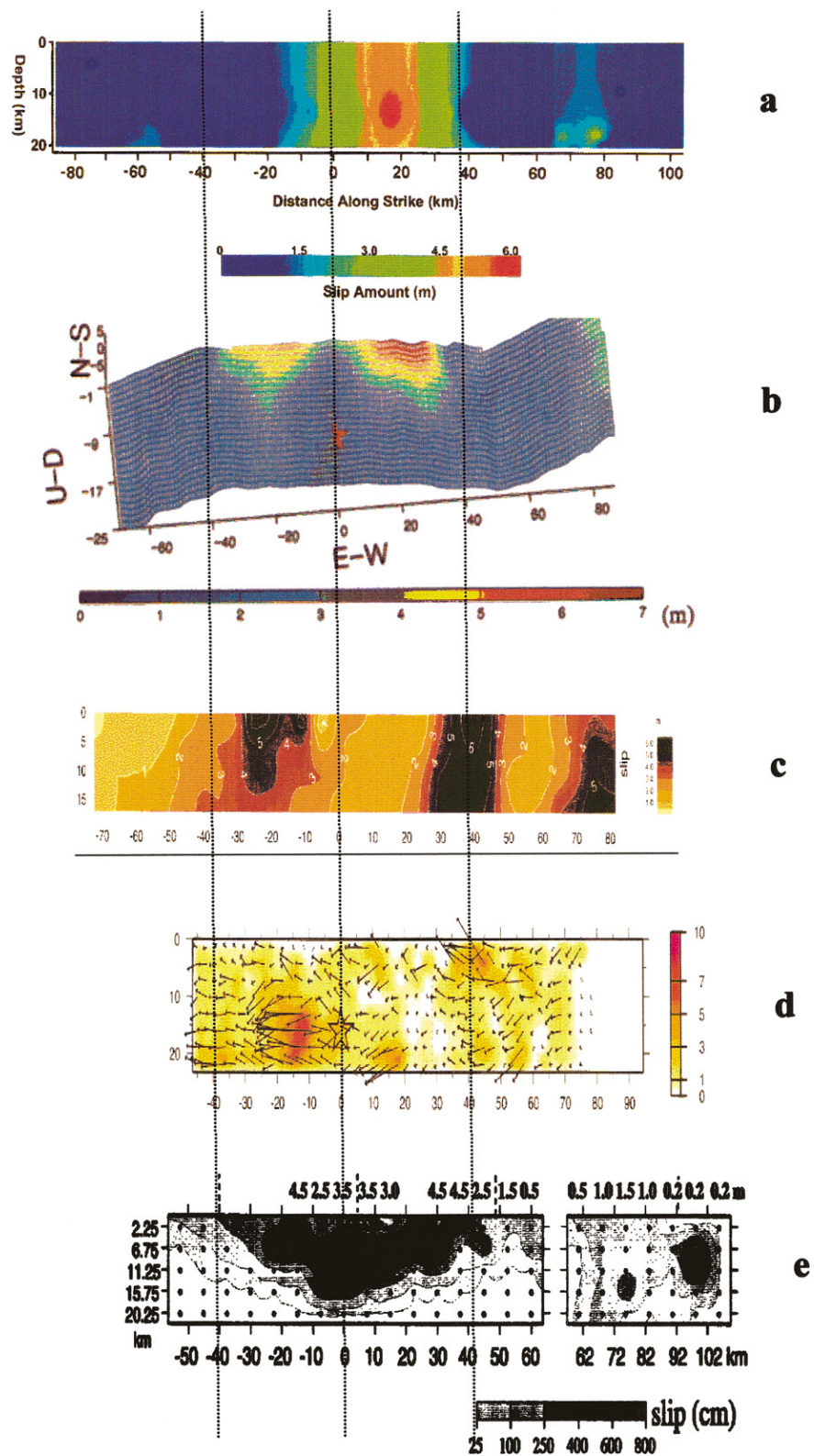
Figure 4. Finite-fault inversions for the distribution of slip on the rupture of the *M* 7.4 1999 İzmit, Turkey, earthquake. (a) Gülen *et al.* (2002, their figure 13); (b) Li *et al.* (2002, their figure 10); (c) Bouchon *et al.* (2002, their figure 3); (d) Sekiguchi and Iwata (2002, their figure 6), and (e) Delouis *et al.* (2002, their figure 12).

otal role in reducing the overall uncertainty in slip inversions discussed in this article (e.g., caused by inadequate knowledge of crustal structure as in Wald and Graves [2001]), which still dominates the inversion results.

Conclusions

The representation theorem allows formulation of a linear inverse problem that solves for total-dislocation vectors on the discretized fault plane. Maintaining a stable linear problem, satisfying reasonable geologic constraints on the resulting slip distribution, involves a substantial amount of arbitrary decision making. It has been shown that numerous solutions can be found that equally well satisfy the data, including the solutions of nonstabilized, nonconstrained schemes. If anything can be said with confidence, it is that the fact of a particular solution matching the data well does not guarantee that this solution is close to the true one.

The uncertainty in the inversions starts at the level of imposing a set of constraints that ensure that the result of a mathematically defined scheme is physically meaningful. These constraints limit the range of possible solutions to a particular subset, which still contains significant nonuniqueness, especially considering that various forms of constraints could equally well be implemented. For a given set of constraints, ambiguities remain at the level of problem parameterization. The assumptions about the true values of the parameters, needed to formulate a resolvable problem, virtually control the solution obtained.

Historically, there have been a surprisingly greater number of published studies that sought geologic interpretation of particular earthquake solutions than of those that investigated the reliability of the solutions based on synthetic tests, although it logically should have been the other way around. It seems that every reliability study, for example, the work conducted with the best resolution by Olson and Anderson (1988) and similar subsequent works, has consistently led to discouraging results. The important inferences made are that incorrect assumptions about the rupture speed, fault geometry, or crustal structure lead to incomplete or erroneous results, and these results are also dependent on array geometries. The addition of noise to synthetic data has the same consequences. Note that all of these conditions apply to any realistic inversion. Artifacts and biases are typically introduced, such as spurious asperities, nonexisting spatial variations of slip, ghost ruptures, or false components of slip vector. It is true that these distinctive features are mostly prone to being geologically interpreted. These conclusions were drawn despite the fact that the authors used simple, uniform slip models and idealized data sets.

The real problems will be more challenging. First, there is a limited value in the studies that fix certain parameters, assuming their true values are known, and investigate the effect of incomplete knowledge of others, as these conditions never materialize in practice. The issues of parameter interaction and unknown trade-offs are not addressed by this approach. A synthetic test that would truly mimic the inversion of a real earthquake's data would be to carry out an inversion in which all assumed parameters are perturbed from their true values and see whether this would still allow reasonable slip recovery. Such a test has yet to be performed. In light of the results of the existing fixed-parameter sensitivity studies, the ability of real inversions to recover earthquake slip is questionable.

Second, conclusions drawn from simple synthetic models of slip may not be applicable to realistic faults with slip distributions of unknown complexity. There is no framework currently available for determining which features of the solutions are real and which are artificial. A sensitivity study, involving dropping or adding some stations to see how the solution changes, accompanying each particular inversion could help recognize the real features but will in no way guarantee uniqueness. Such studies are lacking; one example is the work by Delouis et al. (2002), in which the authors examine the features of the İzmit earthquake inversion controlled by a single isolated station or identify parts of the fault that are in particular responsible for fitting the data (their figure 16).

It has been pointed out that gross features of the inversions, persisting from one parameter variation to the other, could be considered the real features of the images. This approach is reasonable; however, with more parameters varied and more variants of numerical and physical constraints used, there could be no repetition of gross features at all. This point is well illustrated by the recent example of five different inversions for the same event shown in Figure 4, where it would be hard for a user to judge what were indeed the real features of faulting.

There also are uncertainty issues not fully addressed in the published inversions. First, the formulation of the inverse problem based on the representation theorem uses the approximation of a continuous integral of unknown slip by the sum. This approximation will generally be cell-size dependent. The studies of convergence of the sum to some fixed value, which could be performed by repeating the inversion for progressively reduced sizes until the results did not change, are not presented in most inversion applications. Although these studies are desirable, the reduction in cell size is limited by both computing resources and the growing numerical instability as the number of unknowns increases; one should then be ready to accept that some of the obtained solutions will not be accurate representations of the continuous problem. Again, even a proven convergence will not guarantee uniqueness, which is not simply caused by inaccurate numerical approximation.

Second, maintaining a linear inversion problem requires parameterization of the dislocation time function through a single parameter, the final dislocation value. The functions used are integrals of simple shapes such as triangles or boxcars. The amplitude Fourier spectrum of these waveforms is a sinc function that correctly describes realistic spectra at very low frequencies only and is totally irrelevant to the real

spectra at its nodal points. The underlying spectrum used in the inversions is thus problematic. Also, any source time function, including the simple forms used, is a function of two parameters: both the static dislocation U and the duration t_0. Allowing the former to vary while fixing the latter makes the inversion t_0 dependent; this dependence cannot be neglected if one attempts to model frequencies even as low as 0.1–0.5 Hz or higher. Ideally, the inversion should solve for both parameters; however, this precludes the problem formulation as a linear matrix inversion.

Lately, there has been growing use of the inversion algorithms alternative to the traditional linear matrix inversion, based on a grid search in the parameter space. In one of the applications (simulated annealing), the search is carried out using a Monte-Carlo-type random walk to find a global minimum of the objective function (the difference between the observed and model-predicted waveforms); an optimal search algorithm (annealing) starting from a random initial model is prescribed (e.g., Liu *et al.*, 1995; Hartzell and Liu, 1996). An obvious advantage of this approach is that it can theoretically accommodate as many free parameters (in addition to subfault slip weights) as necessary, limited only by practicality issues, and thus avoids the need to assume the values of poorly known parameters to maintain a linear matrix problem. Although some authors still keep subsource durations t_0 fixed (e.g., Hartzell and Liu, 1996), other recent works used smoother pulse shapes (as opposed to triangles and boxcars) and allowed their durations t_0 to be free parameters (e.g., Ji *et al.*, 2002). The latter variant avoids difficulties with underlying spectrum outlined in this article.

The simulated-annealing algorithms have not been adequately explored yet and may suffer from their own deficiencies. For example, being in essence algorithms of structured random search, they rely on a subjective choice of initial model and the prescribed search algorithm, which imparts them a somewhat heuristic character. In the case of many free parameters allowed, the topography of the objective function may become so complicated as to make finding the global minimum problematic; the algorithm may unpredictably fall into one of the local minima, providing a completely incorrect solution. As I pointed out in this article, the increased number of variables also aggravates the problem instability. Finally, the approach of bluntly increasing the amount of unknown variables as a way of avoiding the need to assume some of them does not seem to be productive in general. Quoting from Graves and Wald (2001, p. 8764), "Allowing more complexity in the source . . . would . . . greatly improve the fit. However, this is simply mapping inadequacy in the GFs [Green's functions (any other parameter could fill their place)] back into the source." The linear matrix inversions, having rigorous mathematical basis, still dominate the published case histories.

Because there are several levels of uncertainty in the linear matrix inversions, to what extent could the reality of the specific features on the inverted images for major earth-quakes be trusted? As one can conclude, this question has no answer for many published inversions. It cannot be answered unless each application is accompanied by thorough sensitivity and resolution tests using synthetic data, with all the caveats summarized in this article. What fault information could possibly always be trusted? This probably is crude slip in some average sense, as exemplified in Figure 4, which could also be obtained from the seismic moment. All other details are likely to be artifacts dependent on the choice of a particular inversion scheme, variant of seismological parameterization, geometry of observational array, and grid spacing. There is a sufficient amount of evidence to support this view.

Realistic images of slip on rupturing faults have large implications for seismic-hazard analysis and earthquake physics. However, one should be cautioned against any dogmatic interpretation of slip distributions that are obtained without having these considerations in mind.

Acknowledgments

References

Aki, K. (1967). Scaling law of seismic spectrum, *J. Geophys. Res.* **72,** 1217–1231.

Aki, K., and P. Richards (1980). *Quantitative Seismology: Theory and Methods*, W. H. Freeman, New York, 932 pp.

Beresnev, I. A., and G. M. Atkinson (1997). Modeling finite-fault radiation from the ω^n spectrum, *Bull. Seism. Soc. Am.* **87,** 67–84.

Boore, D. M. (1983). Stochastic simulation of high-frequency ground motions based on seismological models of the radiated spectra, *Bull. Seism. Soc. Am.* **73,** 1865–1894.

Bouchon, M., M. N. Toksöz, H. Karabulut, M.-P. Bouin, M. Dietrich, M. Aktar, and M. Edie (2002). Space and time evolution of rupture and faulting during the 1999 İzmit (Turkey) earthquake, *Bull. Seism. Soc. Am.* **92,** 256–266.

Brune, J. N. (1970). Tectonic stress and the spectra of seismic shear waves from earthquakes, *J. Geophys. Res.* **75,** 4997–5009.

Chi, W.-C., D. Dreger, and A. Kaverina (2001). Finite-source modeling of the 1999 Taiwan (Chi-Chi) earthquake derived from a dense strong-motion network, *Bull. Seism. Soc. Am.* **91,** 1144–1157.

Cho, I., and I. Nakanishi (2000). Investigation of the three-dimensional fault geometry ruptured by the 1995 Hyogo-ken Nanbu earthquake using strong-motion and geodetic data, *Bull. Seism. Soc. Am.* **90,** 450–467.

Cohee, B. P., and G. C. Beroza (1994). A comparison of two methods for earthquake source inversion using strong motion seismograms, *Ann. Geofis.* **37,** 1515–1538.

Das, S., and B. V. Kostrov (1990). Inversion for seismic slip rate history and distribution with stabilizing constraints: application to the 1986 Andreanof Islands earthquake, *J. Geophys. Res.* **95,** 6899–6913.

Das, S., and B. V. Kostrov (1994). Diversity of solutions of the problem of earthquake faulting inversion: application to *SH* waves for the great 1989 Macquarie Ridge earthquake, *Phys. Earth Planet. Interiors* **85,** 293–318.

Das, S., and P. Suhadolc (1996). On the inverse problem for earthquake

rupture: the Haskell-type source model, *J. Geophys. Res.* **101,** 5725–5738.

Das, S., P. Suhadolc, and B. V. Kostrov (1996). Realistic inversions to obtain gross properties of the earthquake faulting process, *Tectonophysics* **261,** 165–177.

Delouis, B., D. Giardini, P. Lundgren, and J. Salichon (2002). Joint inversion of InSAR, GPS, teleseismic, and strong-motion data for the spatial and temporal distribution of earthquake slip: application to the 1999 İzmit mainshock, *Bull. Seism. Soc. Am.* **92,** 278–299.

Graves, R. W., and D. J. Wald (2001). Resolution analysis of finite fault source inversion using one- and three-dimensional Green's functions I. Strong motions, *J. Geophys. Res.* **106,** 8745–8766.

Gülen, L., A. Pinar, D. Kalafat, N. Özel, G. Horasan, M. Yilmazer, and A. M. Işikara (2002). Surface fault breaks, aftershock distribution, and rupture process of the 17 August 1999 İzmit, Turkey, earthquake, *Bull. Seism. Soc. Am.* **92,** 230–244.

Hartzell, S. H., and T. H. Heaton (1983). Inversion of strong ground motion and teleseismic waveform data for the fault rupture history of the 1979 Imperial Valley, California, earthquake, *Bull. Seism. Soc. Am.* **73,** 1553–1583.

Hartzell, S., and D. V. Helmberger (1982). Strong-motion modeling of the Imperial Valley earthquake of 1979, *Bull. Seism. Soc. Am.* **72,** 571–596.

Hartzell, S., and C. Langer (1993). Importance of model parameterization in finite fault inversions: application to the 1974 M_W 8.0 Peru earthquake, *J. Geophys. Res.* **98,** 22,123–22,134.

Hartzell, S., and P. Liu (1996). Calculation of earthquake rupture histories using a hybrid global search algorithm: application to the 1992 Landers, California, earthquake, *Phys. Earth Planet. Interiors* **95,** 79–99.

Hartzell, S. H., G. S. Stewart, and C. Mendoza (1991). Comparison of L_1 and L_2 norms in a teleseismic waveform inversion for the slip history of the Loma Prieta, California, earthquake, *Bull. Seism. Soc. Am.* **81,** 1518–1539.

Heaton, T. H. (1982). The 1971 San Fernando earthquake: a double event? *Bull. Seism. Soc. Am.* **72,** 2037–2062.

Heaton, T. H. (1990). Evidence for and implications of self-healing pulses of slip in earthquake rupture, *Phys. Earth Planet. Interiors* **64,** 1–20.

Henry, C., S. Das, and J. H. Woodhouse (2000). The great March 25, 1998, Antarctic Plate earthquake: moment tensor and rupture history, *J. Geophys. Res.* **105,** 16,097–16,118.

Ji, C., D. J. Wald, and D. V. Helmberger (2002). Source description of the 1999 Hector Mine, California, earthquake, part I: Wavelet domain inversion theory and resolution analysis, *Bull. Seism. Soc. Am.* **92,** 1192–1207.

Li, X., V. F. Cormier, and M. N. Toksöz (2002). Complex source process of the 17 August 1999 İzmit, Turkey, earthquake, *Bull. Seism. Soc. Am.* **92,** 267–277.

Liu, P., S. Hartzell, and W. Stephenson (1995). Non-linear multiparameter inversion using a hybrid global search algorithm: applications in reflection seismology, *Geophys. J. Int.* **122,** 991–1000.

Mendoza, C., and S. H. Hartzell (1988a). Aftershock patterns and main shock faulting, *Bull. Seism. Soc. Am.* **78,** 1438–1449.

Mendoza, C., and S. H. Hartzell (1988b). Inversion for slip distribution using teleseismic P waveforms: North Palm Springs, Borah Peak, and Michoacan earthquakes, *Bull. Seism. Soc. Am.* **78,** 1092–1111.

Mendoza, C., and S. H. Hartzell (1989). Slip distribution of the 19 September 1985 Michoacán, Mexico, earthquake: near-source and teleseismic constraints, *Bull. Seism. Soc. Am.* **79,** 655–669.

Olson, A. H., and J. G. Anderson (1988). Implications of frequency-domain inversion of earthquake ground motions for resolving the space-time dependence of slip on an extended fault, *Geophys. J.* **94,** 443–455.

Olson, A. H., and R. J. Apsel (1982). Finite faults and inverse theory with applications to the 1979 Imperial Valley earthquake, *Bull. Seism. Soc. Am.* **72,** 1969–2001.

Saraò, A., S. Das, and P. Suhadolc (1998). Effect of non-uniform station coverage on the inversion for earthquake rupture history for a Haskell-type source model, *J. Seism.* **2,** 1–25.

Sekiguchi, H., and T. Iwata (2002). Rupture process of the 1999 Kocaeli, Turkey, earthquake estimated from strong-motion waveforms, *Bull. Seism. Soc. Am.* **92,** 300–311.

Sekiguchi, H., K. Irikura, and T. Iwata (2000). Fault geometry at the rupture termination of the 1995 Hyogo-ken Nanbu earthquake, *Bull. Seism. Soc. Am.* **90,** 117–133.

Somerville, P., K. Irikura, R. Graves, S. Sawada, D. Wald, N. Abrahamson, Y. Iwasaki, T. Kagawa, N. Smith, and A. Kowada (1999). Characterizing crustal earthquake slip models for the prediction of strong ground motion, *Seism. Res. Lett.* **70,** 59–80.

Spudich, P. K. P. (1980). The DeHoop–Knopoff representation theorem as a linear inversion problem, *Geophys. Res. Lett.* **7,** 717–720.

Trifunac, M. D. (1974). A three-dimensional dislocation model for the San Fernando, California, earthquake of February 9, 1971, *Bull. Seism. Soc. Am.* **64,** 149–172.

Wald, D. J., and R. W. Graves (2001). Resolution analysis of finite fault source inversion using one- and three-dimensional Green's functions II. Combining seismic and geodetic data, *J. Geophys. Res.* **106,** 8767–8788.

Wald, D. J., and T. H. Heaton (1994). Spatial and temporal distribution of slip for the 1992 Landers, California, earthquake, *Bull. Seism. Soc. Am.* **84,** 668–691.

Wald, D. J., T. H. Heaton, and K. W. Hudnut (1996). The slip history of the 1994 Northridge, California, earthquake determined from strong-motion, teleseismic, GPS, and leveling data, *Bull. Seism. Soc. Am.* **86,** S49–S70.

Department of Geological and Atmospheric Sciences
Iowa State University, 253 Science I
Ames, Iowa 50011-3212
beresnev@iastate.edu