

# A Framework for Online Inversion-Based 3D Site Characterization

Volkan Akçelik<sup>1</sup>, Jacobo Bielak<sup>1</sup>, George Biros<sup>2</sup>, Ioannis Epanomeritakis<sup>1</sup>,  
Omar Ghattas<sup>1</sup>, Loukas F. Kallivokas<sup>3</sup>, and Eui Joong Kim<sup>4</sup>

<sup>1</sup> Carnegie Mellon University, Pittsburgh PA, USA,  
{volkan,jbielak,ike,omar}@cmu.edu

<sup>2</sup> University of Pennsylvania, Philadelphia PA, USA, biros@seas.upenn.edu

<sup>3</sup> University of Texas at Austin, Austin TX, USA, loukas@mail.utexas.edu

<sup>4</sup> Duke University, Durham NC, USA, ekim@duke.edu

**Abstract.** Our goal is to develop the capability for characterizing the three-dimensional geological structure and mechanical properties of individual sites and complete basins in earthquake-prone regions. Toward this end we present a framework that integrates in situ field testing, observations of earthquake ground motion, and inversion-based modeling.

## 1 Geotechnical Site Characterization

An important first step for forecasting strong ground motion during future earthquakes in seismically-active regions is to characterize the three-dimensional mechanical properties and geological structure of sites and basins within those regions. Characterizing a site refers to our ability to reconstruct as faithfully as possible the soil profile in terms of a limited set of material parameters, such as shear and compressional wave velocities, density, attenuation, and the slow velocity for poroelastic media.

Geological and geotechnical materials, soil and rock, impact the performance of the built environment during earthquakes. They play a critical role in the generation of ground motion, and, consequently, on determining the spatial extent and severity of damage during earthquakes. Yet, they are not well-investigated, even though they are the most variable and least controlled of all materials in the built environment. Since soils cannot be accessed easily, their properties can be inferred only indirectly. Currently, geomaterials are characterized with essentially the same general testing methods that were used 25 years ago. These methods rely on testing a small number of specimens in the laboratory and conducting a limited number of small-strain field tests. There is a critical need to advance beyond current methods to reduce the level of uncertainty that currently exists in the estimation of geological and geotechnical material properties.

A variety of techniques for reconstruction of earth properties from noninvasive field tests have been pursued, notably within the gas and oil exploration communities. However, the goals of our work are distinctly different, both in terms of the nature of the problem (e.g. complete material profile reconstruction

vs. estimates of material contrast) and in the models employed (e.g. all elastic waves vs. just acoustic waves).

Reconstruction of the soil model results in a time-dependent or time-harmonic wave propagation inverse problem. Solution of this inverse problem represents an enormous challenge from the theoretical, algorithmic, and computational points of view. An added challenge is the lack of a systematic approach to in situ measurements. Such measurements are often decoupled from the needs of the computational process, due to cost considerations, equipment limitations, or the adoption of ad-hoc and simplified analysis procedures.

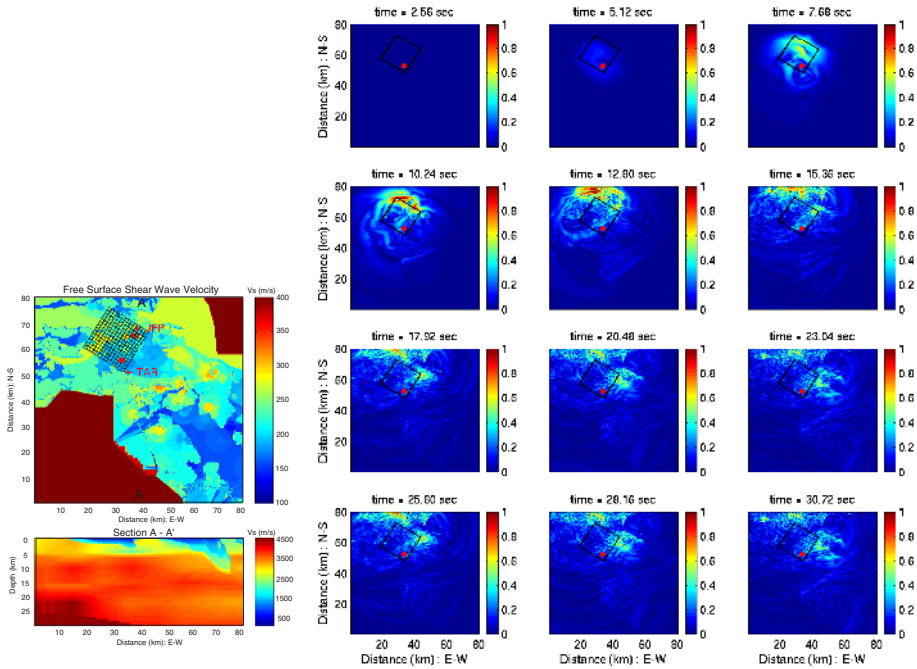
With current test equipment, the volume of soil that can be excited from a single location is somewhat limited because the maximum loads that can be applied are restricted and the response amplitude decays exponentially with distance, frequency, amount of soil damping, and slowness of the propagating waves. The advent of large-scale test equipment makes it possible to apply much larger loads over a wide range of frequencies, and thus excite a larger volume of soil from a single location than has been possible till now.

To effectively extract the desired information from the field test data, we need robust, efficient, and scalable forward and inverse three-dimensional wave propagation solvers. We have developed such methods and fine-tuned them for the analysis of earthquake ground motion in large basins. Our forward and inverse modeling methods are overviewed in Sections 2 and 3, respectively. Finally, in Section 4, we present an on-line framework for local site characterization that integrates steerable field experiments with inverse wave propagation.

## 2 Forward Elastic Wave Propagation Modeling

Our forward elastic wave propagation simulations are based on wavelength-adaptive mesh algorithms, which overcome many of the obstacles related to the wide range of length and time scales that characterize wave propagation problems through heterogeneous media [1,2,3,4,5,6,7]. In highly heterogeneous media such as sedimentary basins, seismic wavelengths vary significantly, and wavelength-adaptive meshes result in a tremendous reduction in the number of grid points compared to uniform meshes (e.g. a factor of 2000 in the Los Angeles Basin). Our code includes octree-based trilinear hexahedral elements and local dense element-based data structures, which permit wave propagation simulations to substantially greater resolutions than heretofore possible.

We have validated our code using the Southern California Earthquake Center (SCEC) LA Basin model and an idealized model of the 1994 Northridge earthquake. To illustrate the spatial variation of the 1994 Northridge earthquake ground motion, Fig. 1 presents snapshots at different times of an animation of the wave propagation through the basin. The left part of the figure shows a plan view and cross-section of the basin, as defined by the distribution of shear wave velocity. The projections of the fault and hypocenter are also shown. The directivity of the ground motion along strike from the epicenter and the concentration of motion near the fault corners are response patterns of the actual earthquake



(a) Shear wave velocity distribution

(b) Northridge simulation

**Fig. 1.** Basin profile and basin response snapshots

that are reproduced in the simulation. Other effects, however, such as strong ground motion in the southeast portion of the San Fernando Valley (Sherman Oaks), in the Santa Monica area, and La Cienega, are not. This discrepancy is due, in part, to the simplified nature of the source, but also because the current SCEC model does not yet provide the necessary fidelity.

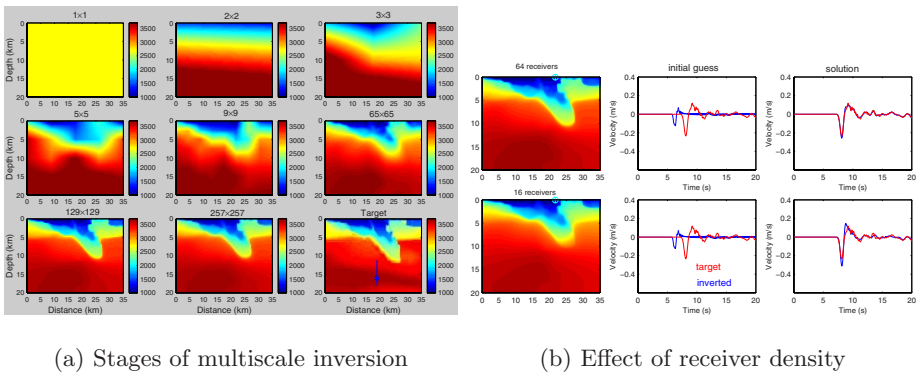
### 3 Inverse Elastic Wave Propagation Modeling

The discrepancy between observations and simulation illustrated in the previous section underscores the need to invert for basin material properties, either from records of past earthquakes or from observations of ground motion generated by seismic testing equipment. The inverse problem is formulated as a nonlinear optimization problem, with an objective function that represents a norm misfit between observations and predictions, and constraints in the form of the seismic wave propagation initial-boundary value problem. The inverse wave propagation problem is significantly more difficult to solve than the corresponding forward problem, because (1) forward solution is just a subproblem of the inverse problem, (2) the inverse problem is often ill-posed despite the well-posedness of the

forward problem, (3) the inverse operator couples the entire time history of the system's response and leads to large dense matrices, and (4) the objective function is often highly oscillatory, with basins of attraction that shrink with the wavelength of the propagating waves, thus entrapping most optimization methods at local minima.

We have developed a scalable, parallel algorithm for the inverse wave propagation problem that addresses these difficulties [8]. To treat multiple local minima, we employ multiscale continuation, in which a sequence of initially convex, but increasingly oscillatory, approximations to the objective function are minimized over a sequence of increasingly finer discretizations of state and parameter spaces, which keeps the sequence of minimizers within the basin of attraction of the global minimum. Total variation (TV) regularization is used to address ill-posedness and preserve sharp material interfaces such as between geological layers. To overcome the difficulties of large, dense, expensive to construct, indefinite Hessians, we combine Gauss-Newton linearization with matrix-free inner conjugate gradient iterations. These require good preconditioners in the case of the TV regularization operator, which presents considerable difficulties due to its strong heterogeneity and anisotropy. Our preconditioner is based on the limited memory BFGS algorithm to approximate curvature information using changes in directional derivatives, initialized by second order stationary iterations applied to the TV operator.

Here we provide evidence of the feasibility of this approach for finely-parameterized models. We illustrate first with a two-dimensional sedimentary basin undergoing antiplane motion. The seismic model under consideration comprises a portion of the vertical cross-section of the Los Angeles basin in Fig. 1, as shown at the bottom of the right frame in Fig. 2a, where the trace of a strike-slip fault and the hypocenter of an idealized source are identifiable. In our numerical experiments, we use waveforms synthesized from the target model on the free



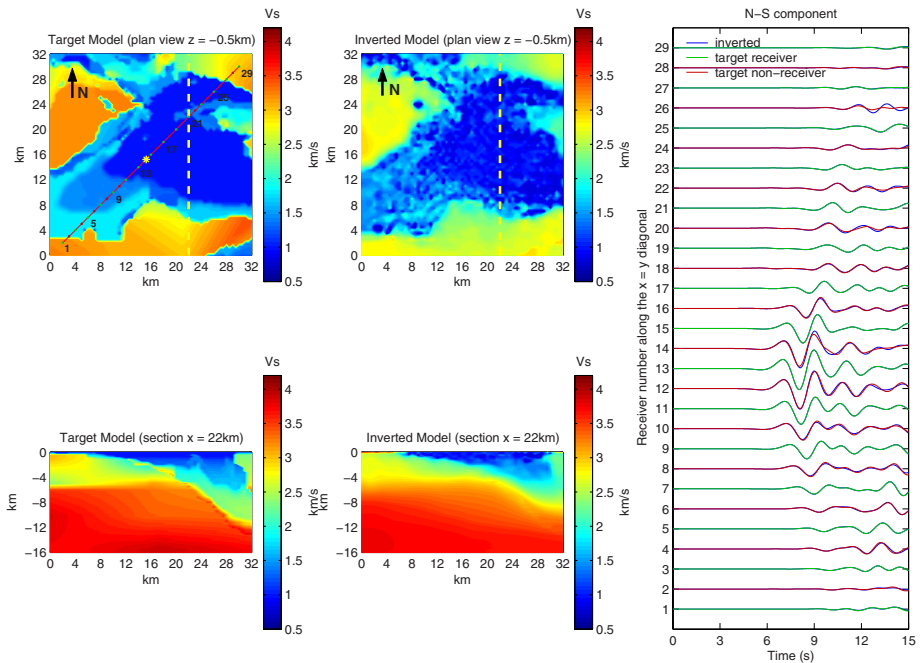
(a) Stages of multiscale inversion

(b) Effect of receiver density

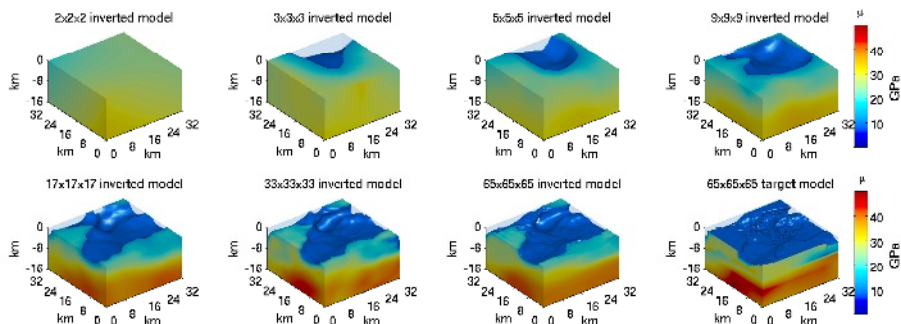
**Fig. 2.** Stages of the multiscale inversion, starting from a homogeneous guess (*left*); comparison of wave velocity profiles and time histories (*red*: target; *blue*: inverted) obtained using different numbers of receivers (*right*)

surface as pseudo-observed data. We assume that the source, defined by its rupture velocity and slip function, is known, and invert for the shear wave velocity distribution of the geological structure. We have assumed that the material is lossless and that the density distribution is known. Five percent random noise has been added artificially to the synthetic seismogram. Fig. 2a shows a sequence of inverted material models, corresponding to increasingly finer inversion grids. Inversion was performed with 64 observers distributed uniformly on the free surface. The high fidelity of the results is noteworthy. Fig. 2b compares results for 64 and 16 receivers. As expected, the resolution of the inverted model for 16 receivers is not as great as for 64 receivers, yet it still closely approximates the target model. The synthetics in Fig. 2b show the velocity at one non-receiver location for the two models, both for the initial guess and for the inverted solution.

Corresponding results for a three-dimensional simulation of the San Fernando Valley are shown in Figs. 3 and 4. The material in this model is linearly elastic and lossless, with density and Lamé parameter  $\lambda$  fixed. We invert for the shear modulus using a grid of  $15 \times 15$  receivers placed on the free surface and uniformly distributed throughout the valley. The source is a point-source double-couple located at a 16 km depth. Fig. 3 displays sections of the distribution of target and inverted shear wave velocity. The right panel shows the N-S component of



**Fig. 3.** LA Basin 3D wave-based inversion: target and computed material profiles; *top row*: plan view at 500m depth; *bottom row*: section view



**Fig. 4.** Convergence of computational to target model via inversion grid refinement

the particle velocity at the points located on the diagonal line shown in the top left panel. Here again the high fidelity of the results is notable; major features of the target model are captured by the inverted model, and the agreement of the target and inverted waveforms is excellent, even at the non-receiver locations. Figure 4 illustrates the different levels of the multiscale algorithm. In this figure we show 12 GPa isosurfaces of the shear modulus. The similarity between the finest-grid inverted model and the target model is evident by comparing the last two images. Current efforts are aimed at extending the method to incorporate inversion for attenuation properties and nonlinear constitutive parameters of the soil, both of which are important for assessing response of geotechnical systems under strong ground motion.

#### 4 A Framework for Online Inversion-Based 3D Site Characterization

The integration of field tests with seismic inversion presents some unique challenges and opportunities, both from the computational as well as experimental point of view. Our goal is to close the loop between experiment and simulation in the context of characterizing individual sites and basins. Toward this end we are pursuing an online framework for site characterization that integrates inversion-based modeling and steerable field experiments. The idea is to repeatedly: (1) excite the soil and collect surface and downhole ground motion data at a particular field location, (2) transmit the data back to high-end supercomputing facilities, (3) perform inversion to predict the distribution of material properties, (4) identify regions of greatest remaining uncertainty in the material properties, and (5) provide guidance on where the next set of field tests should be conducted.

The first key issue will be to turn around the inversion calculations in a time frame appropriate for the field experiments. Getting data to and from the compute nodes will not be a problem; two-way satellite connection will be able to keep up with the sensor ground motion data that is sent, and the material model

that is received, on a daily basis. On the other hand, the run times needed to compute the 3D inversions in the previous section (on the order of a day on 512 processors) would be prohibitive in the online context. However, there are several improvements and simplifications we can make. First, we will invert for significantly smaller regions than the San Fernando Valley example considered above, typically just a square kilometer surrounding the test site. Second, whereas the inversions presented above began with homogeneous initial guesses, here we will incorporate priors based on an *offline phase* in which records of past earthquakes are inverted for an initial model of the basin. With the incorporation of a prior, we can expect faster convergence to the optimum. Third, ongoing improvements to the Hessian preconditioner are expected to yield a significant speedup in iterations. With these improvements, we can expect to comfortably turn around an inversion overnight on a reasonable number of processors.

The second key issue is the ability to invert the material model for observations associated with *multiple* excitations generated by the field equipment for a particular location. These multiple sources provide a richer dataset for inversion, and thus result in better resolution of the reconstructed soil model. However, this entails a large increase in necessary computational resources, since at each iteration of the inversion algorithm, a forward model must be solved for *each* independent excitation, and several dozen such excitations may be available. Fortunately, there is considerable coarse granularity available in the algorithm: all of the forward wave propagation simulations can proceed independently in parallel, at each inversion iteration. Therefore, by exploiting multiple, possibly distributed, groups of processors, we can effect inversion for multiple sources within the same clocktime as a single source—perhaps faster, since the inverse problem is better conditioned with the more copious available data, and the number of iterations should be fewer. The processor groups can be tightly coupled, as for example in the PSC AlphaCluster, or else loosely coupled across multiple sites of the NSF NEESGrid/TeraGrid. The inversion algorithm is highly latency tolerant—the coordination among the multiple forward simulations is in the form of broadcasts of material model corrections—done only after the forward (and adjoint) simulations are completed. Thus we can harvest cycles on processor groups wherever they are available.

The third key issue is the ability to steer the experimental equipment towards regions of high uncertainty in basin properties. The opportunity for steering occurs when the inversion is completed, the material model (of the entire basin) updated to reflect the recently acquired information, and the model transmitted back to the field to help direct the next round of field tests. Uncertainty in the estimates of the material parameters, as indicated by their variances, can be used to steer the equipment towards regions where conducting new tests is most profitable. Evaluating the variances exactly is intractable for such large-scale problems, but through the use of improved preconditioners we may be able to generate sufficiently good estimates of the important components of the Hessian, since the misfit term is typically of very low rank. The main task is to balance quality of the estimates against speed in obtaining them, while keeping sight of the goal of on-line experimental steering.



**Acknowledgements.** This work was supported by the National Science Foundation's Knowledge and Distributed Intelligence (KDI) and Information Technology Research (ITR) programs (through grants CMS-9980063, ACI-0121667, and ITR-0122464) and the Department of Energy's Scientific Discovery through Advanced Computation (SciDAC) program through the Terascale Optimal PDE Simulations (TOPS) Center. Computing resources on the HP Alpha-Cluster system at the Pittsburgh Supercomputing Center were provided under NSF/AAB/PSC award BCS020001P.

## References

- [1] Bao H., Bielak J., Ghattas O., Kallivokas L.F., O'Hallaron D.R., Shewchuk J.R., and Xu J. Large-scale Simulation of Elastic Wave Propagation in Heterogeneous Media on Parallel Computers. *Computer Methods in Applied Mechanics and Engineering*, 152(1–2):85–102, January 1998.
- [2] Bielak J., Xu J., and Ghattas O. Earthquake ground motion and structural response in alluvial valleys. *Journal of Geotechnical and Geoenvironmental Engineering*, 125:413–423, 1999.
- [3] Bielak J., Loukakis K., Hisada Y., and Yoshimura C. Domain reduction method for three-dimensional earthquake modeling in localized regions, part I: Theory. *Bulletin of the Seismological Society of America*, 93:817–824, 2003.
- [4] Yoshimura C., Bielak J., Hisada Y., and Fernandez A. Domain reduction method for three-dimensional earthquake modeling in localized regions, part II: Verification and applications. *Bulletin of the Seismological Society of America*, 93:825–840, 2003.
- [5] Kim E., Bielak J., Ghattas O., and Wang J. Octree-based finite element method for large-scale earthquake ground motion modeling in heterogeneous basins. In *Eos Trans. AGU*, 83(47), Fall Meet. Suppl., Abstract S12B-1221, 2002.
- [6] Volkan Akcelik, Jacobo Bielak, George Biros, Ioannis Epanomeritakis, Antonio Fernandez, Omar Ghattas, Eui Joong Kim, Julio Lopez, David O'Hallaron, Tiankai Tu, and John Urbanic. High resolution forward and inverse earthquake modeling on terascale computers. In *Proceedings of SC2003*, Phoenix, 2003. ACM/IEEE.
- [7] E. Kim, J. Bielak, and O. Ghattas. Large-scale Northridge Earthquake simulation using octree-based multiresolution mesh method. In *Proceedings of the 16th ASCE Engineering Mechanics Conference*, Seattle, Washington, July 2003.
- [8] Akcelik V., Biros G., and Ghattas O. Parallel multiscale Gauss-Newton-Krylov methods for inverse wave propagation. In *Proceedings of IEEE/ACM SC2002 Conference*, Baltimore, MD, November 2002. SC2002 Best Technical Paper Award.