Modeling of global seismic wave propagation on the Earth Simulator

Seiji Tsuboi1*, Dimitri Komatitsch2, Chen Ji2 and Jeroen Tromp2

1 Institute of Frontier Research on Earth Evolution, Japan Marine Science and Technology Center, 3173–25 Showa-machi, Kanazawa-ku, Yokohama, Kanagawa 236-0001, Japan
2 Seismological Laboratory, 252–21 California Institute of Technology, Pasadena, California 91125, USA

(Received January 26, 2004; Revised manuscript accepted March 17, 2004)

Abstract We use the Earth Simulator to model seismic wave propagation resulting from large earthquakes on a global scale. Simulations are conducted based upon the spectral-element method, a high-degree finite-element technique. We include the full complexity of the Earth, i.e., a three-dimensional seismic wave velocity and density structure, a 3-D crustal model, the elliptical figure of the Earth as well as topography and bathymetry. We use 507 nodes of the Earth Simulator (4056 processors) and model the three-dimensional Earth with 13.7 billion grid points. A total of 7 terabytes of memory is needed. We have reached a vectorization ratio of 99.3%, and attained a total performance of 10 teraflops (32% of the peak performance). The very high resolution of the mesh allows us to calculate theoretical seismic waves for realistic fully three-dimensional Earth model, which are accurate at periods of 3.5 seconds and longer. We have modeled seismic waves generated by recent large earthquakes and show the importance of including three-dimensional structure in the seismic wave simulation. These theoretical seismograms calculated for fully three-dimensional Earth models should improve our understanding of earthquake source mechanisms and the Earth dynamics.

Keywords: Seismic wave propagation, broadband seismograms, Spectral-Element Method, Earth Simulator

1. Introduction

A mission to the center of the Earth might not be science fiction any more [1] but it will still take many years until it becomes reality. Nevertheless, the use of seismic waves will remain a unique way to probe our planet. In this regard, accurate modeling of seismic wave propagation in fully three-dimensional (3-D) Earth models is of considerable interest in seismology in order to determine both the 3-D seismic wave velocity structure of the Earth and the rupture process during a large earthquake. Calculations of theoretical seismograms for spherically symmetric Earth models have been accomplished traditionally by the summation of the normal modes of the Earth [2]. This normal mode summation technique has proved to be a powerful tool for the calculation of theoretical seismograms and has been used in various field of seismology. However, the actual Earth has significant deviations from spherical symmetry, such as the 3-D seismic wave velocity structure inside the solid mantle and laterally heterogeneous crust at the surface of the Earth, which makes an extension of the normal mode approach to this problem a formidable task.

The field of numerical modeling of seismic wave propagation in 3-D structures has been significantly evolved in the last few years due to the introduction of the Spectral-Element Method (SEM), which is a high-degree version of the finite-element method that is very accurate for linear hyperbolic problems such as wave propagation, having very little intrinsic numerical dispersion. In addition, the mass matrix is exactly diagonal by construction, which makes it much easier to implement on parallel machines because no linear system needs to be inverted. The 3-D SEM was first used in seismology for local and regional simulations [3, 4, 5], and more recently adapted to wave propagation at the scale of the full Earth [6, 7, 8, 9]. However, until recently, at the scale of the globe, available computer resources intrinsically limited such large calculations. For instance, on a PC Beowulf Cluster with 150 processors, Komatitsch and Tromp [7] reached minimum seismic periods of 18 seconds (i.e. 1/18 Hz). Such periods are not short enough to capture important effects on wave propagation related to smaller-size 3-D heterogeneity in the Earth.

Here we show that our implementation of the SEM on the Earth Simulator allows us to calculate theoretical seismic waves, which are accurate up to 3.5 seconds and

* Corresponding author: Seiji Tsuboi, Institute of Frontier Research on Earth Evolution, Japan Marine Science and Technology Center, 3173–25 Showa-Machi, Kanazawa, Yokohama, Kanagawa 236-0001, Japan. Email: tsuboi@jamstec.go.jp
longer for fully 3-D Earth models. We include the full complexity of the 3-D Earth in our simulations, i.e., a 3-D seismic wave velocity and density structure, a 3-D crustal model, ellipticity as well as topography and bathymetry [7, 8]. We also include the effect of the oceans on surface wave dispersion [8]. Synthetic waveforms at such high resolution (periods of 3.5 seconds and longer) allow us to perform direct comparisons of arrival times of various body wave phases between observed and synthetic seismograms, which has never been accomplished before. Usual seismological algorithms, such as normal-mode summation techniques that calculate quasi-analytical synthetic seismograms for one-dimensional (1-D) spherically symmetric Earth models [2], are typically accurate down to 8 seconds. In other words, the SEM on the Earth Simulator allows us to simulate global seismic wave propagation in fully 3-D Earth models at periods shorter than current seismological practice for simpler 1-D spherically symmetric models.

2. Implementation of the Spectral-Element Method on the Earth Simulator

The SEM is based upon a weak formulation of the equations of motion, obtained after dotting the classical differential linear wave equation with a so-called test vector (in finite-element parlance) and integrating by parts over the volume of the model. The SEM leads to an exactly diagonal mass matrix, therefore no linear system needs to be inverted and the method lends itself well to calculations on large parallel machines with distributed memory, such as the Earth Simulator.

The SEM combines the flexibility of the finite-element method with the accuracy of the pseudospectral method. It uses a mesh of hexahedral finite elements on which the wave field is interpolated by high-degree Lagrange polynomials on Gauss-Lobatto-Legendre (GLL) integration points [10, 11, 7]. The mesh honors all first and second-order discontinuities in the reference Earth model [12]. In the mantle, crust and solid inner core of the Earth, we solve the anelastic seismic wave equation in terms of the displacement vector, whereas in the liquid outer core we solve the acoustic wave equation in terms of a fluid potential [7]. The SEM has been extensively benchmarked against semi-analytical reference solutions for 1-D reference Earth models in previous work [7, 8]. To represent the 3-D Earth, we use mantle model S20RTS [13], which expands the 3-D shear wave velocity structure in spherical harmonics with a maximum degree of 20, crustal model CRUST2.0 [14], and global topography and bathymetry model ETOP05 (from the U.S. National Oceanic and Atmospheric Administration). In this formulation, what are not included in our approximations based

![Fig. 1](image_url)
on the 3-D Earth model are effects of the ocean layer and the perturbations in the gravitational potential, an approximation known as the Cowling approximation.

Figure 1 shows a global view of the spectral-element mesh at the surface of the Earth. The sphere is meshed using hexahedra only, based upon an analytical mapping from the six sides of a unit cube to a six-block decomposition of the surface of the sphere, which is called the ‘cubed sphere’ [15, 16, 6, 7]. The figure illustrates that each of the six sides of the ‘cubed sphere’ is divided into 26 × 26 slices, shown with different colors, for a total of 4056 slices. We use 4056 processors, i.e., 507 nodes out of 640 of the Earth Simulator and allocate one slice per processor. Each slice is subdivided with a mesh of 48 × 48 spectral-elements at the surface of each slice. Within each surface element we use $5 \times 5 = 25$ GLL grid points, which translates into an average grid spacing of 2.0 km (i.e., 0.018 degrees) at the surface. The total number of spectral elements in this mesh is 206 million, which corresponds to a total of 13.8 billion global grid points, since each spectral element contains $5 \times 5 \times 5 = 125$ grid points, but with points on its faces shared by neighboring elements. This in turn corresponds to 36.6 billion degrees of freedom (the total number of degrees of freedom is slightly less than 3 times the number of grid points because we solve for the three components of displacement everywhere in the mesh, except in the liquid outer core of the Earth where we solve for a scalar potential). Using this mesh, we can calculate synthetic seismograms that are accurate down to seismic periods of 3.5 seconds. Such simulations use a total of approximately 7 terabytes of memory. As a comparison, we can simulate global wave propagation only down to seismic wave periods of 18 seconds on a PC cluster computer with 150 processors (733 MHz Pentium-III) and 75 gigabytes of memory. The mesh files, created once and for all by our in-house parallel mesh generator, are stored on the Earth Simulator’s large capacity Mass Data Processing System, to avoid having to recreate the mesh every time we start a new simulation.

Our SEM solver is based upon a pure MPI implementation, combined with loop vectorization. The SEM algorithm is not ‘embarrassingly parallel’, because of the need to assemble internal force contributions between neighboring slices, and because of the pattern of communications needed to assemble such slices on the edges and corners of the six blocks of the cubed-sphere mesh, as can be seen in Figure 1 (for instance the valence of most surface points is 4, but it is three at the corners of the six blocks). However, because the mass matrix is exactly diagonal, processors spend most of their time performing actual computations, and the amount of communications is comparatively small. Because of the regular mesh pattern in each of the six blocks of the cubed sphere, load balancing is very good by construction in our SEM algorithm. Historically, this SEM code was initially developed in 1996 on a Thinking Machine CM-5. Then it was ported to MPI on a Linux PC cluster. Therefore, it took no effort to parallelize the code on the Earth Simulator, because we already had a portable MPI implementation. The current package is written in Fortran95 + MPI.

The vectorization was a more difficult issue to address. Unfortunately, our existing Fortran95 MPI source code could not be efficiently vectorized automatically by the compiler because of the fact that the main routines consist of matrix-vector products in each spectral element, but these elements, and therefore the related matrices and vectors, are very small ($5 \times 5 \times 5$ grid points as mentioned above). We therefore manually restructured and inlined most of these loops (in particular critical inner loops). The key issue was to reorder loop index variables to make sure that vector register access became continuous. Fortunately, only two small routines had to be rewritten (the code spends 80 to 90 percent of the CPU time in these critical routines, which are small in terms of source code size). After manual loop restructuring, all critical loops in the SEM code were fully vectorized, and we reached a vectorization ratio of 99.3%, measured using the MPI Program Runtime Performance Information system tool.

Using 4056 processors, a simulation of 5 minutes of actual seismic wave propagation accurate up to a period of 3.5 seconds and longer requires about 2 hours of wall-clock time (using 6000 time steps of 5 ms each for the explicit time integration scheme of the SEM algorithm). Total performance of the code, again measured using the MPI Program Runtime Performance Information was 10 teraflops, which is about one third of expected peak performance for this number of nodes (507 nodes × 64 gigaflops = 32 teraflops). This performance level is not as high as that reached by other applications on the Earth Simulator, mostly because of the short vector lengths involved in the matrix-vector products in the SEM algorithm. As mentioned above, each spectral-element contains $5 \times 5 \times 5 = 125$ points, which is about half the size of the vector registers on the Earth Simulator (256). As a result, loop operations performed at the level of the elements do not fully take advantage of the vector pipes.

Before we graduated to 507 nodes of the Earth Simulator, we used 243 nodes (1944 processors) to calculate theoretical seismic waves accurate up to 5 seconds or longer. For the 243 nodes case, the total performance we achieved was about 5 teraflops, which also is about one third of the peak performance. The fact that when we
double the number of nodes from 243 to 507 the total per-
formance also doubles from 5 teraflops to 10 teraflops
shows that this SEM code exhibits an excellent scaling
relation with respect to performance. We may expect that
we can calculate seismic waves accurate up to 1 second
and longer by increasing the number of processors by a
factor of ten without significant loss in the total perf-
formance.

3. Seismic wave propagation resulting from
large earthquakes in the full 3-D Earth

We first simulate seismic waves generated by a deep
earthquake of magnitude 6.7 that occurred in Colombia
on September 2, 1997, at a depth of 213 km. Because
there are more than 200 seismic recording stations
equipped with broadband digital seismometers now per-
manently installed worldwide, giving us access to three-
component broadband seismograms, we can perform a
direct comparison between our synthetic seismograms
and real data from recorded earthquakes. Such records are
composed of compressional body waves (P-waves), shear
body waves (S-waves) and surface waves and their
reflected and converted phases caused by the discontinu-
ities of the Earth model.

3-D models of the seismic wave velocity structure of
the Earth are traditionally built based upon a combination
of travel-time anomalies of short-period body waves and
long-period surface waves. However, independent valida-
tion of such existing 3-D Earth models has never been
attempted before, by lack of an independent numerical
way of computing the seismic response in such models.
Figure 2 [17] shows the first attempt to verify such agree-
ment; we compare the vertical component of displace-
ment from synthetic seismograms calculated using 243
nodes of the Earth Simulator, which are accurate up to 5
second and longer, and observed records for several seis-
mic recording stations with an epicentral distance of
about 75 to 85 degrees. The first pulses in these traces are
the direct P waves. Because the Colombia earthquake was
a deep event (located at a depth of 213 km), two so-called
depth phases are present in the records after the P
arrivals. The first one is called the pP phase, which means

---

**Fig. 2** Left: Broadband data and synthetic displacement seismograms for the 1997 Colombia earthquake, bandpass-filtered with a
two-pass four-pole Butterworth filter between periods of 5 and 150 seconds. Vertical component data (black) and synthetic
(red) displacement seismograms aligned on the arrival time of the P wave are shown. For each set of seismograms the azimuth
is printed above the records to the left, and the station name and epicentral distance are printed to the right. Right: Similar
seismograms at other seismic stations for a total duration of 2000 seconds instead of 200 seconds.
that the P wave from the hypocenter first travels upward and then reflects off the surface of the Earth as a compressional wave. The second phase is called sP, which means that the S wave from the hypocenter first travels upward and reflects at the surface of the Earth and then propagates to the seismic station as a converted P wave.

The agreement between the synthetic seismograms and observed records for the P and pP phases in terms of arrival time is excellent for these stations, which means that the 3-D seismic P-wave velocity model [13] used in this simulation is accurate. Therefore, our simulation demonstrates that this 3-D model represents the general picture of the Earth’s interior fairly well. However, the arrival time for the sP phase is not as good for these stations, which turn shows that the S-wave velocity structure of this model is not as accurate around the earthquake hypocentral region. This information could be directly used to further improve the current 3-D seismic wave velocity model. Figure 2 also compares longer records of about 30 minutes of vertical component synthetic and observed seismograms. These records show that for most stations within this epicentral distance range both arrival times and amplitudes of body waves and surface waves are generally well reproduced by our current 3-D seismic wave velocity model. In figure 3, we compare synthetics calculated by using traditional normal mode summation technique and those calculated by SEM. Because the synthetics calculated by normal mode summation are accurate up to 8 seconds and longer, we used 48 nodes of the Earth Simulator to calculate synthetics by SEM. The results shown in figure 3 clearly demonstrate that the agreement of both synthetics and observation is significantly improved by including 3D Earth structure in SEM synthetics.

We next simulate another deep earthquake of magnitude 6.3 that occurred South of Japan on November 12, 2003, at a depth of 382 km. In Figure 4, we compare the vertical component of displacement from synthetic seismograms calculated using 507 nodes of the Earth Simulator and observed records for several broadband...
seismic stations of the F-net array operated by the National Institute of Earth Science and Disaster Prevention in Japan. Because of the proximity of these stations to the epicenter, there are no depth phases in these seismograms. It is amazing that the global 3-D seismic velocity model used in this simulation still produces fairly good agreement with the observations even at periods of 3.5 seconds, because it is supposed that the crustal and mantle structure beneath Japanese Islands are highly heterogeneous and may be difficult to capture by using global 3D Earth model. However, Figure 4 also shows that the theoretical seismograms calculated with 507 nodes of the Earth Simulator do not reproduce some of the fine features in the observation and suggests the limitation of this global 3-D seismic velocity model. For those stations located to the north-east of the epicenter (the azimuth is about 20 degrees), the observed waves show large high-frequency oscillations because the waves travel along the subducting Pacific plate, but this feature is not modeled in the theoretical seismograms. This shows that we need to improve our 3-D seismic wave velocity model to calculate theoretical seismic waves that are accurate at 3.5 seconds and longer.

Next, we model effects caused by the rupture propagating along a finite-size fault on the amplitudes of seismic waves for large earthquakes. We simulate a very large recent earthquake that occurred in Alaska on November 3, 2002 (magnitude 7.9, at a depth of 15 km). This event is the largest strike-slip earthquake in North America since the very destructive April 18, 1906, San Francisco earthquake. It ruptured 220 km of the Denali fault in Central Alaska [18], in five distinct fault segments. To define the finite-size fault model in our numerical simulations, we approximate the fault rupture by a set of 475 sub-events of size 4 km × 5 km.
Fig. 5 Maps on which stations in the Global Seismographic Network, which recorded displacement seismograms for the 2002 Alaska earthquake, are labeled and denoted by black triangles. To the right of each station the data are shown in black and the 3-D SEM synthetic seismograms in red. Both data and synthetic seismograms are bandpass-filtered with a two-pass six-pole Butterworth filter between periods of 5 and 150 seconds. The black star denotes the epicenter, and the black bar at the bottom denotes the time scale. Both data and synthetic seismograms are multiplied by the inverse of the body-wave geometrical spreading factor in an attempt to remove effects associated with epicentral distance. Top: P wave displacement on the vertical component. Bottom: S wave displacement on the transverse component. To compare the P- and S-wave amplitudes in the two figures the S waves need to be multiplied by a factor of four.

Results of our simulations using 243 nodes of the Earth Simulator are shown in Figure 5 [17, 19]. Because of the fact that the main rupture occurred in a southeasterly direction along large finite-size fault segments [20, 21], recorded seismograms show significantly enhanced ground motions toward the conterminous United States for both body waves and surface waves, a phenomenon referred to as directivity. Figure 5 shows that this directivity is well modeled in our theoretical seismograms [17, 19]. Figure 6 illustrates snapshot of seismic wave propagation and shows this effect of directivity. This shows that it is essential to include finite source effect to model seismic waves generated by this size
Modeling of global seismic wave propagation

Fig. 6 Snapshot of the propagation of seismic waves in the Earth during the November 3, 2002 Denali fault earthquake simulated in Tsuboi et al (2003). Note that large amplification in the western coast of the United States due to the source directivity is modeled well in the simulation.

of earthquake. In Figure 7 [17, 19] we show 12 minutes of vertical component displacement seismograms aligned on the arrival time of the P wave, and 17 minutes of transverse component displacement seismograms aligned on the arrival time of the S wave. This figure illustrates that the theoretical seismograms calculated by 243 nodes of the Earth Simulator for a realistic 3-D Earth model reproduce the entire observed seismic waves, including body and surface wave arrivals. For reference, the same transverse component data are compared to SEM synthetic seismograms for the spherically-symmetric Earth model PREM [12] in Figure 8. The figure demonstrates that agreement between theoretical seismograms and observations is significantly improved by incorporation of the 3-D model. Thus it is essential to incorporate 3-D Earth structure to model the rupture history of large earthquakes. This work [17] has been awarded the 2003 Gordon Bell Prize at SC2003 in Phoenix, Arizona (http://www.sc-conference.org/sc2003/nr_finalaward.html).

Fig. 7 Broadband data and synthetic displacement seismograms for the 2002 Alaska earthquake, bandpass-filtered with a two-pass four-pole Butterworth filter between periods of 5 and 150 seconds. Left: vertical component data (black) and synthetic (red) displacement seismograms aligned on the arrival time of the P wave. Right: transverse component data (black) and synthetic (red) displacement seismograms aligned on the arrival time of the S wave. For each set of seismograms the azimuth is plotted above the records to the left, and the station name and epicentral distance are plotted to the right. The transverse component seismograms need to be multiplied by a factor of ten to compare them with the vertical component seismograms in terms of amplitude.
4. Discussion

We have shown that the use of both the Earth Simulator and the SEM has allowed us to reach unprecedented resolution for the simulation of global seismic wave propagation resulting from large earthquakes. The results demonstrate that given a detailed earthquake source model, precise models of the Earth’s mantle and crust, a precise numerical technique, and a large computer, seismic waves generated by large earthquakes propagating through heterogeneous Earth models, which span an amplitude range that covers several orders of magnitude and a few decades in frequency, can be accurately modeled. In particular, we have successfully attempted for the first time an independent validation of an existing 3-D Earth model. Such 3-D calculations on the Earth Simulator reach shorter periods than quasi-analytical 1-D spherically-symmetric solutions that are current practice in seismology.

It is obvious that the synthetic seismic waveform-modeling tool presented in this work will allow us to further investigate and improve existing Earth models. Tono et al. [22] have calculated synthetic seismograms for a recently deployed dense short-period seismographic array in Japan using the SEM on the Earth Simulator. They modeled seismic waves excited by large deep earthquakes, which occurred under the Japan Sea on June 28, 2002, and compared the results of this simulation with the observation. By using the SEM synthetics calculated for a realistic 3-D Earth model, it is possible to determine differences in the arrival times between theoretical seismograms and observations for those waves that reflected at seismic discontinuities in the mantle. These differences in arrival time can be interpreted as depth variations of the discontinuities and the authors argue that the spatial distribution of the depth variation correlates with seismic velocity anomalies, which has important implications for mantle dynamics. This kind of study would not have been possible without the combination of a precise seismic wave modeling technique, such as the SEM, on a powerful computer, such as the Earth Simulator, and a dense seismic observation network.
Acknowledgments

All the simulations were performed at the Earth Simulator Center of JAMSTEC by S.T. Broadband seismograms used in this study were recorded at Global Seismic Network stations and were obtained from the IRIS Data Management Center (www.iris.washington.edu) and F-net stations operated by the National Institute of Earth Science and Disaster Prevention. Three of the authors (D. K., C. J. and J. T.) were funded in part by the U.S. National Science Foundation. S.T. was partly funded by a Grant-in-Aid for Scientific Research (KAKENHI 13640426) from the Japan Society for the Promotion of Science. We thank Brian Savage from Caltech for providing software that plots seismograms on maps.

(This article is reviewed by Dr. Takeshi Yukutake.)

References