Earth and Planetary Science Letters 308 (2011) 151–160

Contents lists available at ScienceDirect

Earth and Planetary Science Letters

journal homepage: www.elsevier.com/locate/epsl

Tectonic regionalization without a priori information: A cluster analysis of upper mantle tomography

Vedran Lekic [a,b,*], Barbara Romanowicz [a]

[a] Berkeley Seismological Laboratory, Berkeley, CA 94720, USA
[b] Geological Sciences, Brown University, Providence, RI 02912, USA

ARTICLE INFO

Article history:
Received 13 December 2010
Received in revised form 25 May 2011
Accepted 27 May 2011
Available online 21 June 2011

Editor: P. Shearer

Keywords:
global tomography
continental lithosphere
hotspots
cluster analysis
waveform modeling

ABSTRACT

Global mantle tomography can be improved through better use of data and application of more accurate wave propagation methods. However, few techniques have been developed for objective validation and exploration of the resulting tomographic models. We show that cluster analysis can be used to validate and explore the salient features across such models. We present a cluster analysis of a global upper mantle radially anisotropic model SEMum developed using full waveform tomography and the Spectral Element Method. Applied to SEMum down to 350 km depth, the cluster analysis reveals that absolute shear wave velocity (Vs) depth profiles naturally group into families that correspond with known surface tectonics. This allows us to construct a global tectonic regionalization based solely on tomography, without the help of any a priori information. We find that the profiles of stable platforms and shields consistently exhibit a mid-lithospheric low velocity zone (LVZ) between 80 and 130 km depth, while the asthenosphere is found at depths greater than 250 km in both regions. This global intra-continental-lithosphere low velocity zone agrees with recent receiver function studies and regional tomographic studies. Furthermore, we identify an anomalous oceanic region characterized by slow shear wave speeds at depths below 150 km. Hotspots are found preferentially in the vicinity of this anomalous region. In the Pacific Ocean, where plate velocities are largest, these regions have elongated shapes that align with absolute plate motion, suggesting a relationship between the location of hotspots and small-scale convection in the oceanic upper mantle.

1. Introduction

Until now, global mantle tomography has relied on approximate seismic wave computational tools that provide robust images of long wavelength mantle structure. Resolving smaller structure, especially in low velocity regions, remains a challenge for two reasons. First, the uneven sampling of the mantle by commonly analyzed phases — those well separated on the seismogram — must be overcome. This can be done by full-waveform modeling, which can extract the complete information contained in seismic records. Second, more accurate 3D wave propagation tools need to be employed. This is because ray approximations break down as the wavelength of the sought-after structure approaches that of the input waveforms (Spetzler et al., 2002). Furthermore, unmodeled effects of crustal structure can obscure the mantle signal (Bozdağ and Trampert, 2008; Lekic et al., 2010). Fortunately, the advent of new, fully numerical codes like the Spectral Element Method (SEM) enables accurate calculation of wave propagation through highly heterogeneous structures, including the crust (Komatitsch and Vilotte, 1998).

We developed SEMum (Lekic and Romanowicz, 2011), a high resolution model of upper mantle structure, using a fully numerical wave propagation code C-SEM (Capdeville et al., 2003) that is capable of accurately representing both the scattering and (de)focusing of seismic waves by elastic heterogeneity, and, with some approximation, the effects of the oceans, topography/bathymetry, ellipticity, gravity, rotation and anelasticity (Komatitsch and Tromp, 2002). C-SEM allows for efficient computations by restricting the SEM numerical computation to a region of the globe (here the mantle), through coupling with a fast 1D mode calculation (here in the core). We optimized data utilization through the use of full-waveform modeling of long period waveforms, with a cut-off period of 60 s to keep computational costs realistic. We minimized crustal contamination by including constraints from both long period waveforms and higher-frequency group velocity dispersion maps. We also keep computational costs reasonable by computing finite-frequency Frechet kernels — relating structure perturbations to waveform perturbations — using approximate, non-linear 2D finite-frequency kernels based on normal mode perturbation theory (Li and Romanowicz, 1995), which brings out the ray character of overtones. While the approximate partial derivatives may slow down convergence, our use of C-SEM ensures that the cost function — and therefore the tomographic model itself — is calculated more accurately than has previously been possible. Data used, parameterization, forward

* Corresponding author at: Berkeley Seismological Laboratory, Berkeley, CA 94720, USA.
E-mail address: lekic@seismo.berkeley.edu (V. Lekic).

0012-821X/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.epsl.2011.05.050

modeling and inversion scheme, and treatment of crustal structure are described in detail in Lekic and Romanowicz (2011), and the model is available at http://www.seismo.berkeley.edu/~ekic/SEMum.html. Here, we focus on the application of a cluster analysis to the upper 350 km of SEMum.

Despite the proliferation of global tomographic velocity models (Romanowicz, 2003), few tools exist for quantitative exploration, comparison, and validation of these models. Cluster analysis allows classification of a dataset into several groups (clusters), whose members tend to be similar in some fashion (see, e.g. Romesburg, 1984). The classification is objective in the sense that the groups emerge spontaneously, and are not chosen by an operator; indeed, the only way to influence the results of the clustering is by defining the metric that quantifies similarity between individual and groups of data points. Cluster analysis has been applied across physical and social sciences. In geophysics, it has been used with success to classify structures based on a variety of data (e.g. Dumay and Fournier, 1988; Tronicke et al., 2004); in global seismology, its use has been confined to time series analysis (Houser et al., 2008). Here, we discuss the results of cluster analysis applied to the SEMum tomographic model itself, represented by isotropic shear wave speed V_S profiles in the uppermost 350 km. The goals of the analysis are twofold: 1. to identify geographical regions that share common shear velocity structure; 2. to objectively define and investigate representative velocity profiles characteristic of each of these geographic regions.

By identifying geographic regions that share similar V_S profiles in an objective and self-consistent fashion, cluster analysis makes it possible to develop a seismic regionalization without the use of any a priori information. A number of regionalization schemes have been developed previously (Gudmundsson and Sambridge, 1998; Jordan, 1981; Nataf and Ricard, 1996), which divide the Earth's surface into provinces based on geological observations. Because seismic structure correlates with tectonic setting (Romanowicz, 1991), these regionalizations could be used to predict seismic structure. A motivation for doing this was to compensate for the small amplitudes of velocity anomalies in older tomographic models. However, such regionalizations involved assumptions about extrapolations to regions with poor data coverage. Also, as they were dominated by surface observations, the regionalization-based models poorly fit observed long period surface waves, which sample deeper structures (Ekstrom et al., 1997).

We show that, now, global upper mantle V_S structure has been mapped with sufficient accuracy and uniformity to define a tectonic regionalization based solely on tomography. Indeed, a cluster-analysis based regionalization of SEMum shows compelling agreement with regionalizations based on our surface-based inferences on tectonics. Comparison of regionalizations obtained via cluster analysis of different tomographic models offers a new means of exploring tomographic models. Furthermore, inconsistencies and incongruities between these seismic regionalizations and geologic/tectonic inferences can be used as a novel means of validating seismic models and shedding light on regions where the geological structure may not be well indicative of upper mantle structure. We will demonstrate how such arguments can be brought to bear on SEMum and two other recent tomographic models and argue that SEMum more successfully recovers the well known main tectonic provinces. Finally, because the centroid of each cluster specifies a characteristic Vs profile for its corresponding geographic region, cluster analysis provides us with V_S profiles that bring out the salient characteristics of each region. Here, we focus on characteristic V_S profiles to investigate the structure of the continental lithosphere and regions affected by hotspot volcanism.

2. Cluster analysis of global tomography

We apply a k-means clustering scheme to the profiles of absolute shear wave speed (V_S) and radial anisotropy parameter ($\xi = \frac{V_{SH}^2}{V_{SV}^2}$) in SEMum beneath a regular Gaussian grid of points (2° spacing) on the Earth in the 30–350 km depth range (sampled every 10 km). This grid is finer than the nominal model resolution, which is found from resolution tests to be 1500 km laterally and ~50 km in depth (Lekic and Romanowicz, 2011), in order to avoid spatial aliasing. *k-means* is a process well-suited to very large datasets, in which a set of M-dimensional observations (e.g. vectors containing absolute Vs at a discrete number of depths) is partitioned into k sets ("clusters") so that the within-set variance is small. Thus, k-means cluster analysis requires choosing a pre-determined number of clusters (N) and will produce N reference M-dimensional points that define the clusters. MacQueen, 1967 states the procedure clearly and succinctly: *"the k-means procedure consists of simply starting with k groups each of which consists of a single random point, and thereafter adding each new point to the group whose mean the new point is nearest. After a point is added to a group, the mean of that group is adjusted in order to take account of the new point. Thus at each stage the k-means are, in fact, the means of the groups they represent (hence the term k-means)."*

A distance measure is needed to give meaning to concepts *near* and *far*. We explore two simple distance measures: 1. squared Euclidean distance, where profiles of Vs or ξ specified at *m* discrete depths are treated as vectors in *m*-dimensional space; and, 2. correlation distance, where 1 − correlation between two Vs profile vectors defines the distance between them. While correlation is the distance metric that is most-often adopted in cluster analyses of time series, it discards information on the amplitudes of velocity variations. Squared Euclidean distance, on the other hand, depends strongly on the amplitudes of Vs variations.

The starting set of k vectors is itself the result of a clustering of a decimated set of Vs profiles, which is initialized with k randomly selected profiles. Because the k-means procedure is not guaranteed to converge to the set of clusters that minimize the intra-cluster variance, we replicate the entire procedure 5 times, and take the regionalization with smallest intra-cluster variance. Our k-means clustering results are very compatible with those found using agglomerative hierarchical clustering with complete linkage, though the clusters emerge in different order. We use the MATLAB implementation of the k-means algorithm. We also carry out hierarchical agglomerative cluster analysis, and find that complete linkage — where distance between two groups of vectors is taken to be the largest distance between their constituent members — yields very similar results to those obtained from k-means clustering. In contrast, simple or average linkage forms clusters with very different numbers of members, and appears to be strongly affected by outlier profiles whose similarity to one another results in merging otherwise dissimilar clusters.

3. Patterns of upper mantle heterogeneity

3.1. Vs structure

We start with profiles of isotropic shear wave speed and by allowing two clusters to form. The geographic extents of the clusters obtained with a squared Euclidean (left) and correlation-based (right) distance measure are shown in Fig. 1. For both distance measures, the first two clusters (Fig. 1a,i) trace out the continent/ocean dichotomy, confirming that this dichotomy is the dominant pattern of upper mantle structure (Dziewonski, 1970; Kanamori, 1970; Toksöz and Anderson, 1966). One cluster covers ~60% of the earth's surface including most of the oceans as well as several Phanerozoic orogenic and magmatic zones. The other cluster covers areas undisturbed since the Phanerozoic. For the squared Euclidean distance measure, the very oldest ocean in the northwestern Pacific is grouped within the largely continental region. This is due to the fast velocities of the oldest oceanic lithosphere, to which the squared Euclidean distance measure is inherently more sensitive, and is consistent with findings of Okal (1977). Introducing a third cluster (Fig. 1b, j) separates the oceanic region into two according to age: one with a mean age of 40 Ma and