Supplementary Figure Captions

Supplementary Figure 1 Caption: Map showing earthquake epicenters (black stars), seismic station locations (red triangles), ScS reflection points at the CMB (blue dots), and 12 lines along which cross-sections are made through the migration image volume in the lower mantle (green lines) in Supplementary Figures 2-4, and box indicating the area spanned by migration images at constant depth in Supplementary Figure 5. Profiles A-A' and B-B' in Figure 1 correspond to profiles D-D' and Z-Z', respectively, here in the supplementary material. Harvard CMT focal mechanisms for the 15 earthquakes are also shown. Earthquake magnitudes vary from 6.0 to 7.2 and depths vary from 218 to 610 km. Source-receiver distances vary from 65 to 85 degrees. This is the same data set used by (refs. 11, 13). All data have been deconvolved by a source ScS wavelet, low-pass filtered below 0.3 Hz and amplitudes have been normalized to the S phase. Neither normalization to ScS nor signal to noise ratio weighting significantly alter migration images. Corrections for lithospheric anisotropy were applied in rotating the SH signals. All of the events have stable SH radiation patterns between S and ScS to California stations, with radiation coefficients varying from 0.6 to 1.0. The ellipses indicate the Fresnel zone dimensions at a depth of 2650 km based on optics for a horizontal reflector (outer ellipse) and reduced for the effective region of contribution of non-specular reflections based on finite frequency wave theory (inner ellipse). Sub-Fresnel zone variations can be manifested in the data for complex structures.

Supplementary Figure 2 Caption: Vertical cross-sections along profile D-D' (Supp. Fig. 1) through the image volumes obtained using data binned according to earthquake location. The grid dimensions are –6 to 20° N, 257 to 283° E, and 2310 to 2910 km, in latitude, longitude, and depth, respectively, sampled every 0.2° along the great circle path, and 10 km in depth. Travel times are calculated using the 1-D PREM velocity model. Left panels use ScS as the reference phase while the right panels use S as the reference phase. In all migration images in the manuscript and supplementary information, the data stack is weighted according to the hitcount, S and its surface multiples (sS, SS, etc.) have been censored by applying symmetric 4 s tapers around the phase in the deconvolved waveforms, the colour scale has been saturated to highlight stable features and there is no vertical exaggeration. While the absolute depths of features change by ~50 km, due to an imperfect 1-D reference model, the abrupt jump in the D" discontinuity remains, which confirms that it is a robust feature. The number of waveforms used in the South, Central, and North earthquake bins are 54, 180, and 36 respectively. The migration image in the North bin shows the most streaking due to the low number of waveforms. The image features in the north also appear weak because the source-receiver distances are on average shorter (pre-critical for the D" reflector) than those in the Central and South bins. The step feature is associated with the source clusters, and has abrupt image terminations, as found in double-array stacking analysis[11]. The spatial variation of the number of seismograms contributing data to the stack at each grid point is shown at the bottom. Gridpoints near the step have over 200 seismograms contributing. The abruptness of the D" discontinuity image terminations in the South and Central images is a result of constraining data to have source-to-gridpoint and gridpoint-to-receiver distances less than the ray turning point at about 45 degrees. The amount of overlap and the jump from the deeper to the shallower reflector are only resolved to be less than ~100 km.

Supplementary Figure 3 Caption: Vertical cross-sections along profiles A-A’ to F-F’ (a) through the image volume (Supp. Fig. 1) using the PREM velocity models and ScS as a reference phase (left two columns) or S as a reference phase (right two columns), with S energy suppressed by 4 s wide tapers. ScS is suppressed by 4 s wide tapers in the second and fourth columns. (b) is similar to (a) except the migration images are along profiles U-U’ to Z-Z’.

Supplementary Figure 4 Caption: Vertical cross-sections along profiles A-A’ to F-F’ through the image volume (Supp. Fig. 1) migrated relative to ScS (a and b) or S (c and d) using the PREM, UT, CIT and UCB velocity models, with S (a, b, c and d) and ScS (b and d) energy suppressed by 4 s wide tapers. The reference phase used for migration as well as whether ScS was included or censored is indicated at the top of each of the figures. Measurements of ScS-S differential times for this data set indicate an overall south to north increase in D" shear velocity[11,15], so we favour the results found using the UT and UCB models, as those models are consistent with the trend in the differential times. This is the first use of 3D shear wave velocity models in deep mantle migration.

The tomography models are relatively long wavelength compared to our image volume, but differential times for some scattering positions relative to reference phases of up to 3 s are predicted between our 1D model (PREM) and the 3D models. While such small differential times seem insignificant, they can give rise to 100-200 km differences in image depths due to the grazing ray geometry of the data.

Supplementary Figure 5 Caption: Horizontal cross-sections at 7 different depths through the image volume (Supp. Fig. 1) migrated relative to ScS using the PREM velocity model, with S (a, b, c and d) and ScS (a and c) energy suppressed by 4 s wide tapers. The left two columns are the stacked image at each grid point at each depth. The right two columns indicate the number of seismograms contributing to each grid point at each depth for the two cases on the left.

Supplementary Figure 6 Caption: Our Kirchhoff migration code (similar to the one used in ref. 17) assumes isotropic single-scattering: when an incident SH wave encounters a point scatterer, a scattered signal propagates to each receiver without further interaction with any other scatterers. This is valid for weak scattering, with heterogeneities of less than a few percent. To quantify the extent of smearing along dominant raypaths due to the clustering of both sources and receivers in the data, we migrate synthetic seismograms that include arrivals from point scatterers. Two point scatterers are placed along profile D-D’ (a) and roughly correspond to the middle of the bright red features in the data stacks (Fig. 2a) offset by 100 km vertically. The data have been binned according to earthquake location to show which data illuminate regions of the image volume. The migration using all of the synthetic seismograms show significant lateral streaking. The migrations of data (Fig. 2a profile A-A’) show much less streaking and less overlap of the discontinuity; destructive interference has enhanced the topography of the discontinuity because destructive interference has reduced the amount of lateral streaking. Vertical cross sections along profile U-U’ (b) encounter only the deeper point scatterer and show that streaking is roughly symmetric. S_1* (Figure 2b) only exists in the centre and to the Southwest of the image and is not present towards the Northeast suggesting that it must be displaced to the Southwest**.**

Supplementary Figure 7 Caption: (a) Deconvolved tangential component displacement waveforms from our data set, which show S, ScS and a negative polarity scattered phase whose moveout is inconsistent with a specular reflection, suggesting its origin is an out of plane scatterer. All seismograms have corrections for lithospheric anisotropy applied in rotating the SH signal. The effects of ray parameter dependence are very small for the receiver term, given the common assumption of horizontal hexagonal symmetry axis. We do not have corrections for the source terms, however we only use deep focus events as much as possible to minimize the importance of source side effects. The net effect is an overall shift of the SH and SV waveforms, with little coupling between them. The effects on differential travel times within the SH waveforms (our data) are negligible. (b) Reflectivity synthetic seismograms for shear velocity models shown in (c) demonstrate the weak negative amplitude arrival expected for velocity decreases within the D" layer. Model Bin 1 (from ref. 11) has a 2.6% velocity increase 264 km above the CMB, which gives rise to the Scd and Sbc arrivals. The LVZ models have a negative velocity discontinuity 90 km above the CMB with varying strength that gives rise to reflection SdS. The reflection from the deeper discontinuity is not large enough to see in individual seismograms unless the decrease in velocity is at least 5-10%, and it is still a factor of 3 to 4 smaller than Scd. The reflection anticipated from a velocity decrease caused by post-perovskite converting back to perovskite (model LVZ 1) is much weaker than the observed negative amplitude arrivals. A scatterer, with efficient focusing could amplify negative amplitude arrivals by factors comparable to those for critical angle reflections, possibly matching the data. As the precise geometry is not constrained, this will still require velocity reductions of 5-10% with efficient focusing of the wavefield.

Supplementary Figure 8 Caption: Synthetic seismograms, courtesy of Mike Thorne, computed for a step discontinuity model using an axisymmetric 2.5D finite-difference method were used for test migrations. The basic model is shown at the top. Representative synthetics sampling the southern and northern regions of the model are given by synthetic seismograms 1 and 2, respectively. These are compared to a synthetic for a 1D model with a 2% shear velocity discontinuity. The synthetics show discrete reflections, indicating small effective Fresnel zones (Supp. Fig. 1), and this is consistent with the observations, which tend not to show two arrivals in a single waveform**.**

Supplementary Figure 9 Caption: Migration of axisymmetric 2.5D finite difference synthetics computed using a velocity model (a) with an abrupt step of 100 km in the D" discontinuity (dVs = +1.5%). (b) and (c) show migration images using S or ScS, respectively, as the reference phase. Three sets of synthetic seismograms were used in the migration, simulating the average depth and distance to the step in the D" discontinuity for each of the earthquake clusters (North, Central and South in Supp. Fig. 1). We use 3 sets of synthetic seismograms with station spacing that approximately mimics that of the data in order to understand the artefacts introduced by migration of data for a laterally varying structure and to assess the general ability of migration to recover a step. Each cluster predominantly images the local portion of the reflector most densely sampled, and while there is some streaking in the noise-free synthetics, the step is clearly imaged by abrupt changes in reflector from group to group. Migrating a (computationally very expensive) set of synthetic seismograms, which fully replicate the source-receiver geometry, is not realistic given the limited constraints on the overall geometry of the model. All of the images of the D" discontinuity in (b) and (c) exhibit similar behaviour to our data migrations (Supp. Fig. 2), although they under predict the step, adding confidence that it is a robust feature. However, this feature could still be an artefact of using a homogeneous reference model to compute the travel times for a wave field sampling an inhomogeneous structure. The limited sample density of our data set precludes the use of a reference discontinuity model and a migration method that can explicitly account for triplicated arrivals, such as Gasussian beamlet migration (Hill, 1989) or a finite difference migration (Claerbout, 1971). In order to flatten the discontinuity feature, an abrupt and strong SE to NW decrease in velocity within D" is required (for the migration relative to ScS) combined with a strong SE to NW increase in velocity above D" (for the migration relative to S).

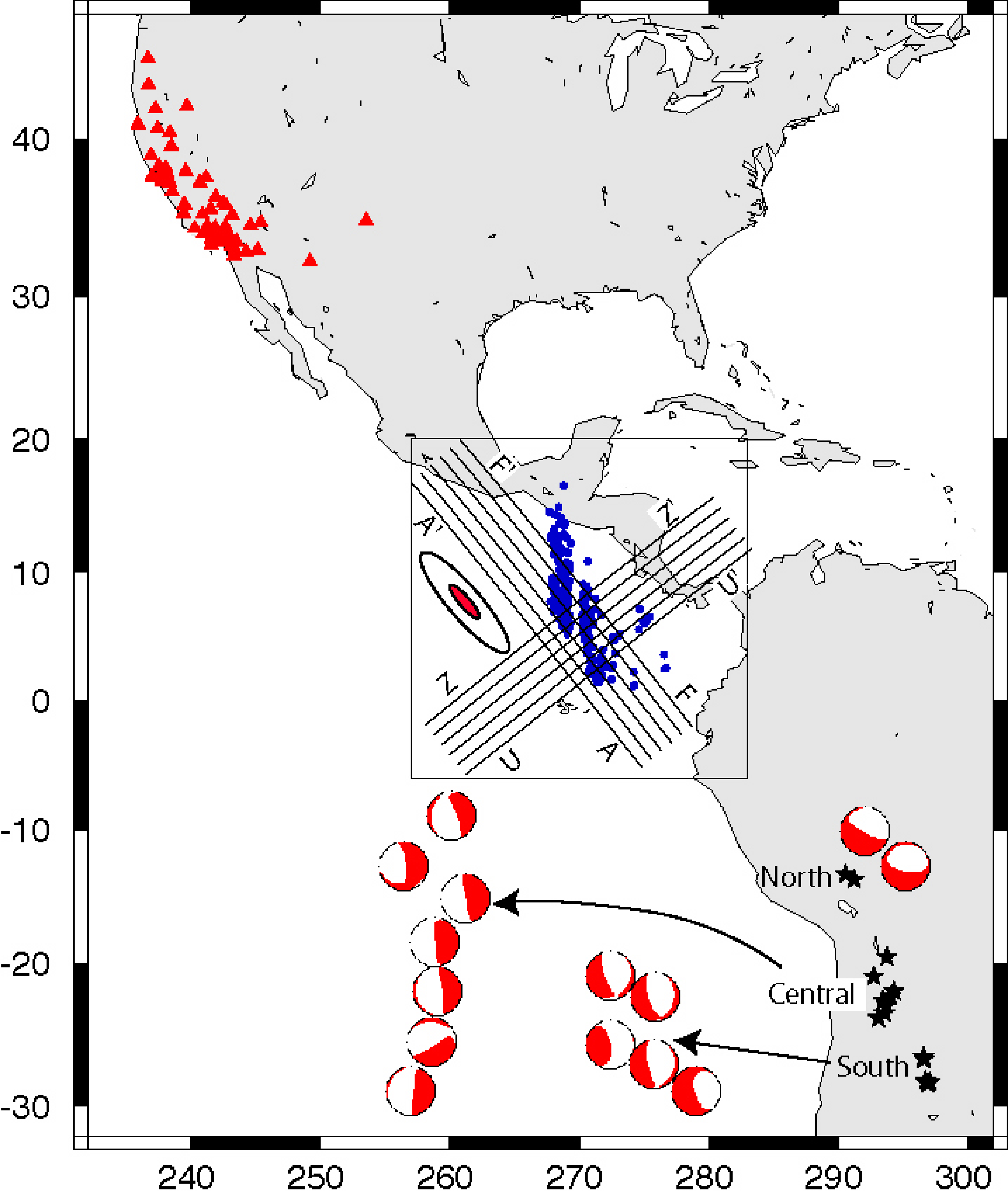
Claerbout, J.F., 1971, Toward a unified theory of reflector mapping: Geophysics, 36, 467-481.

Hill, N.R., 1989, Gaussian beam migration: Geophysics, 55, 1416-1428.

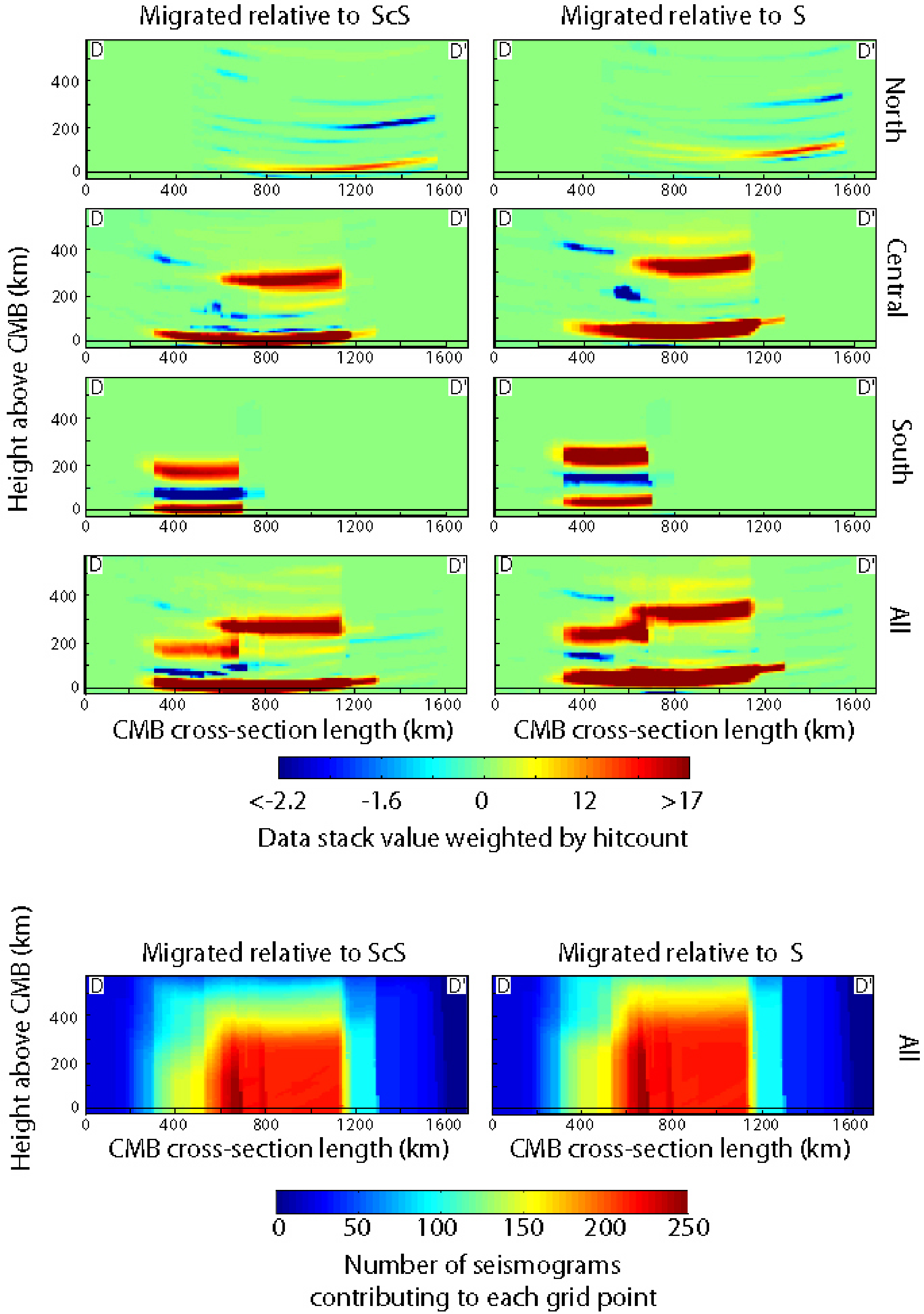
Supplementary Figure 10 Caption: Migration results when systematic shifts in earthquake locations are applied to the southern cluster of 5 earthquakes that illuminate the deeper, southeastern portion of the discontinuity. Data stack values have been weighted according to the hitcount. Shifts of 40 km, a very large mislocation error for these well recorded events, are applied in either depth or epicentre. In all cases, the step remains present. Random location errors or the use of different earthquake location catalogs (e.g. Harvard CMT or EHB) affect the migration images even less. Images using S as the reference phase are more affected than those using ScS, and even when the step in the image formed relative to S reduces in amplitude, there is not a corresponding reduction in the step in the image formed relative to ScS. Our confidence in the step image is partly based on this consistency (supp. Fig. 3). Changing the depth shifts the travel time curve (with its triplication), forward or backward in distance, and this gives the same sign of change of the reflected arrival time relative to S or ScS (rather than opposite signs, as needed to simultaneously flatten out the separate migration images). This is also the challenge for accounting for the step by volumetric heterogeneity. One must change the velocities above and below the discontinuity with opposite sign in order to coherently move the image location formed for reference phases above and below the discontinuity. There is no support for this whatsoever in the tomographic models.

Supplementary Figure 11 Caption: Results of applying double array stacking[11,12] using two bins corresponding to the southern and northern regions on either side of the step in the D" discontinuity, as shown in the map. The large amplitude at 0 km is produced by ScS, and the prominent shallower features around 200 or 300 km above the CMB are produced by Scd+Sbc energy from the D" discontinuity. The fact that simple D" reflector features appear at apparent reflector depths offset by about 100 km using *either* S or ScS as a reference phase, provides strong support for the step in the discontinuity imaged by our migrations.

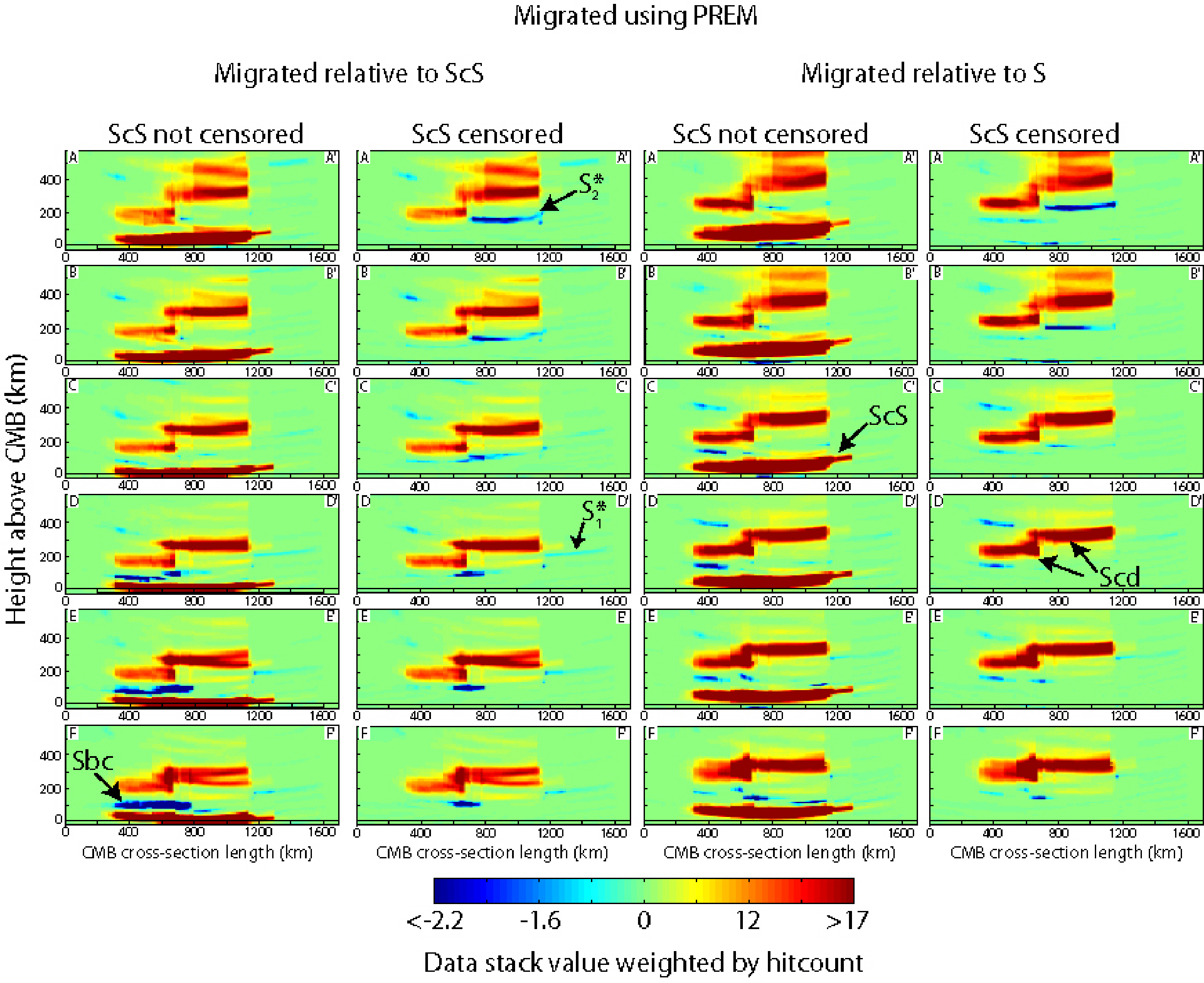
Supplementary Figure 12 Caption: Cross sections through the mantle S-wave velocity tomography model UT[2]. Slab-like features in the mid-mantle extend from the Caribbean to under the central U.S., but in the deep mantle, these can be interpreted to fold and distort strongly, connecting to high velocity material that extends well westward from the site of down-welling. We highlight possible slab folds that may correspond to the feature imaged in this study. Interpreting tomography images is highly subjective; our purpose here is to connect our well-resolved migration image to the broader context provided by mantle tomography.



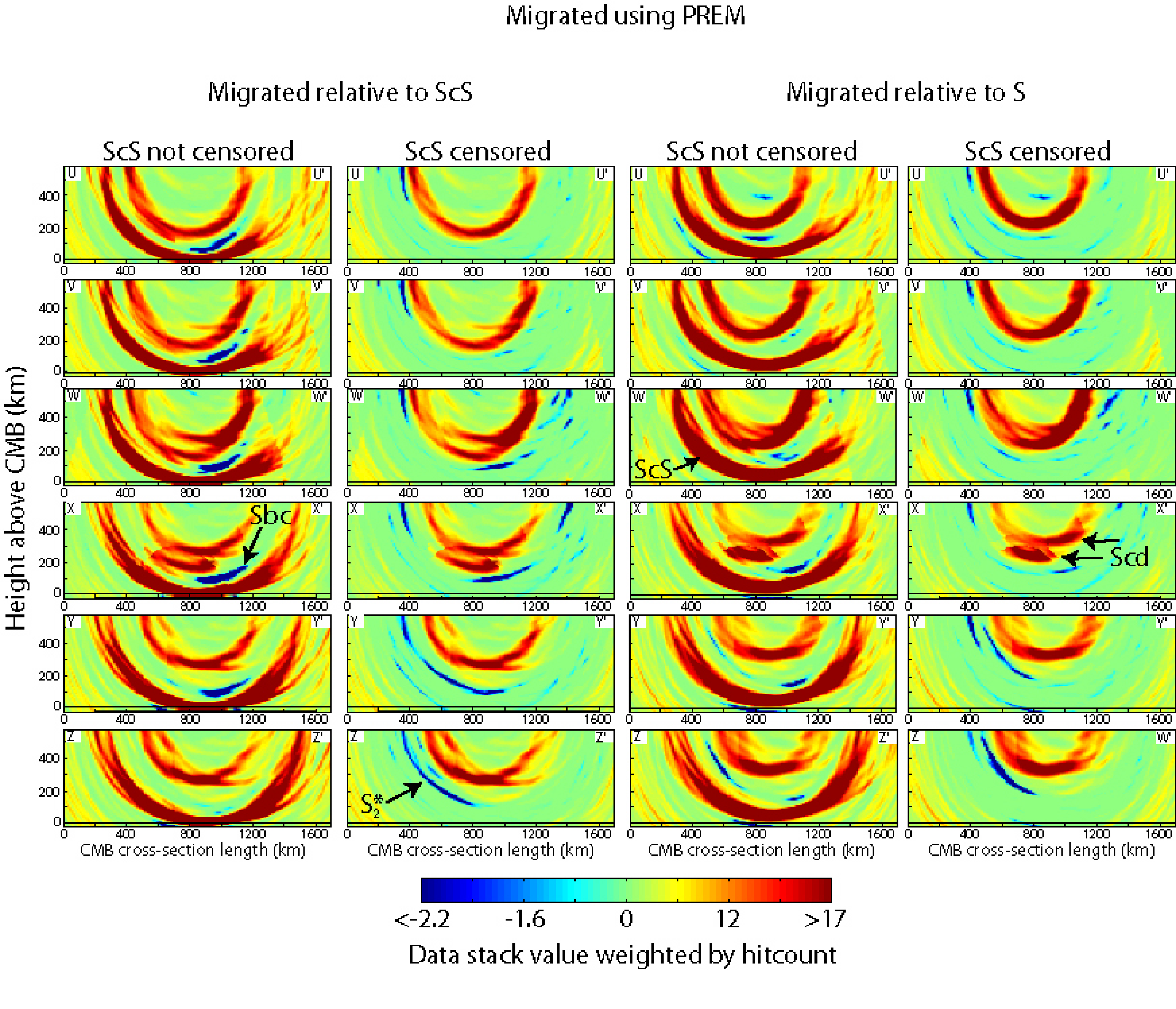
Supplementary Figure 1 [Hutko et al., 2006]



Supplementary Figure 2 [Hutko et al., 2006]

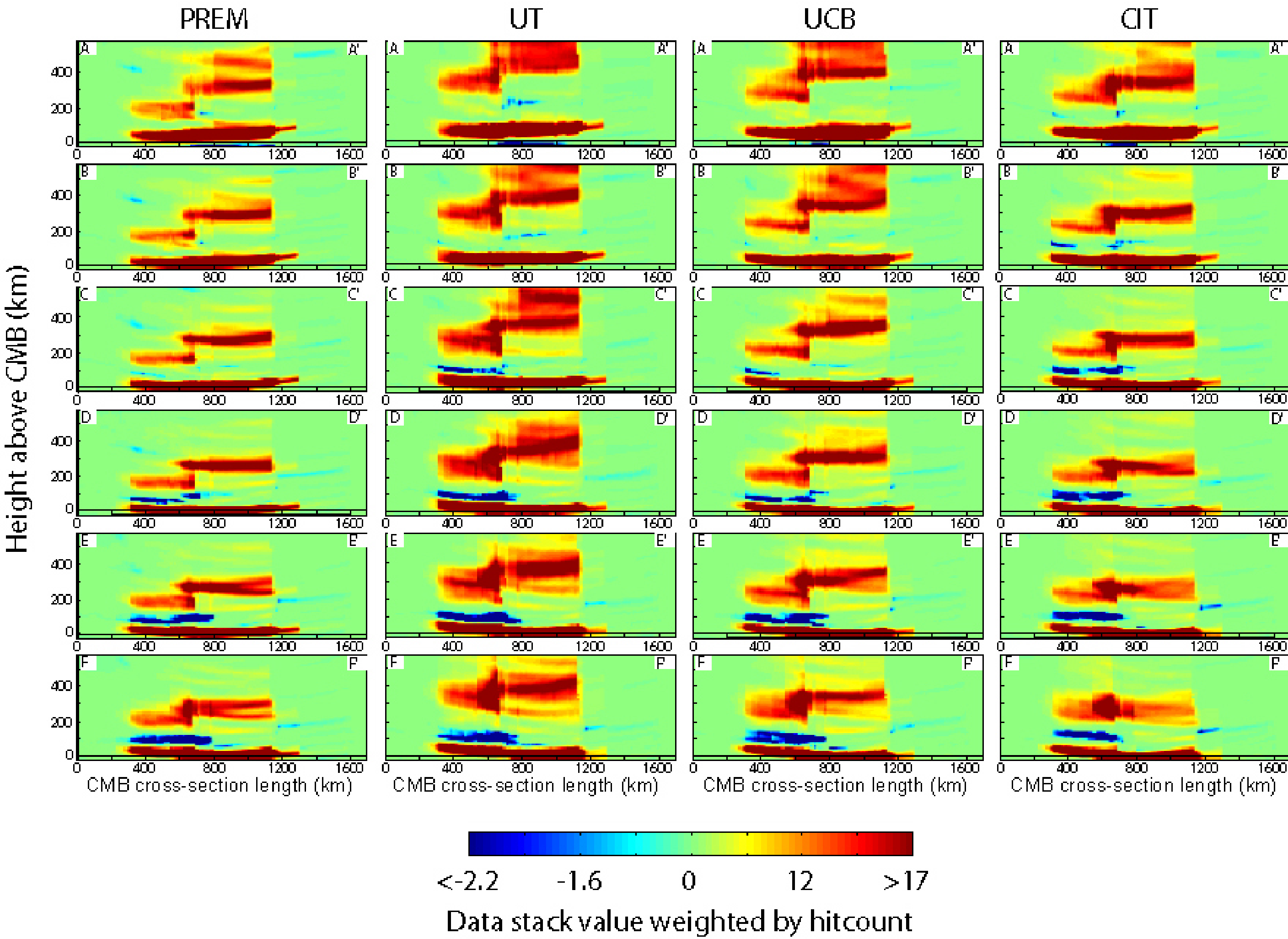


Supplementary Figure 3a [Hutko et al., 2006]

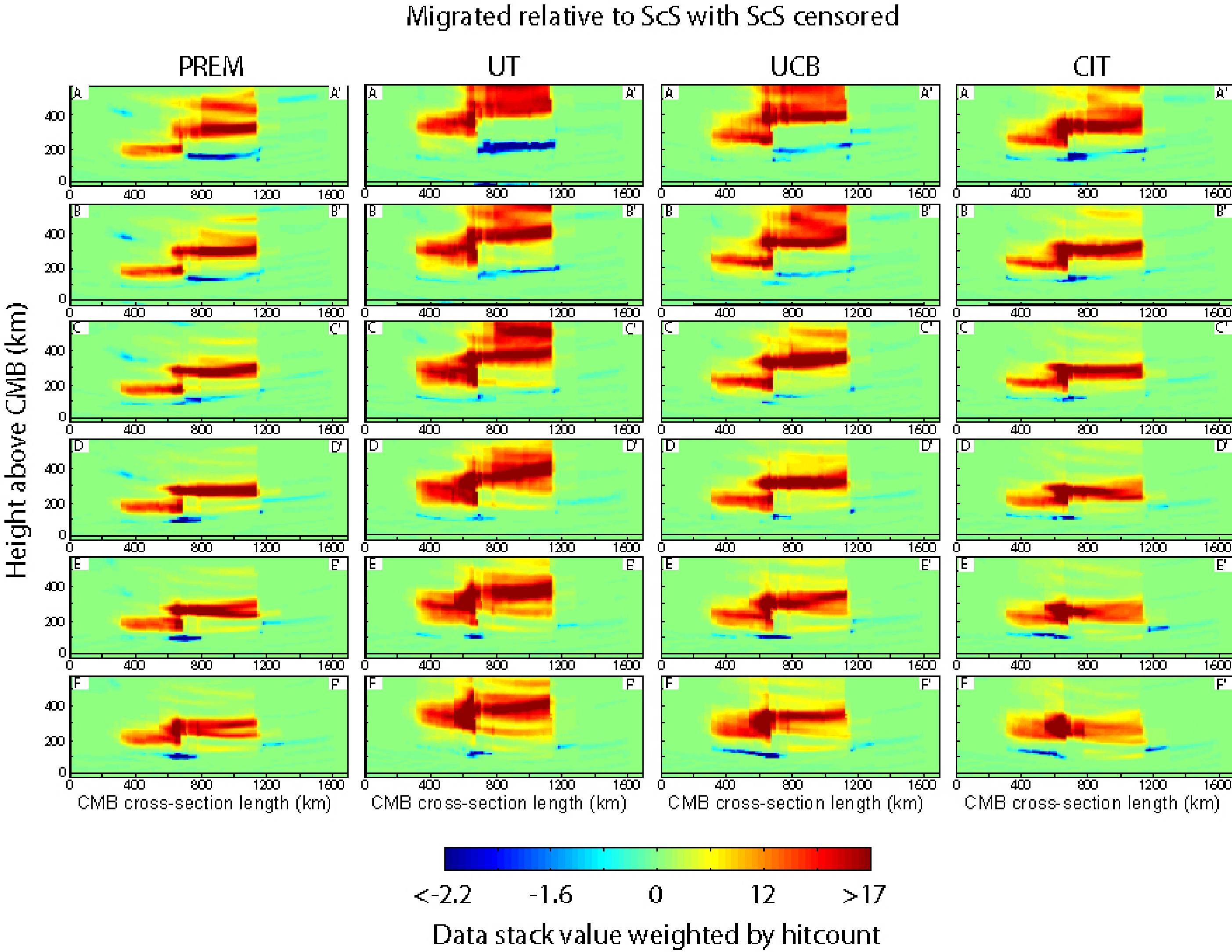


Supplementary Figure 3b [Hutko et al., 2006]

Migrated relative to ScS with ScS not censored

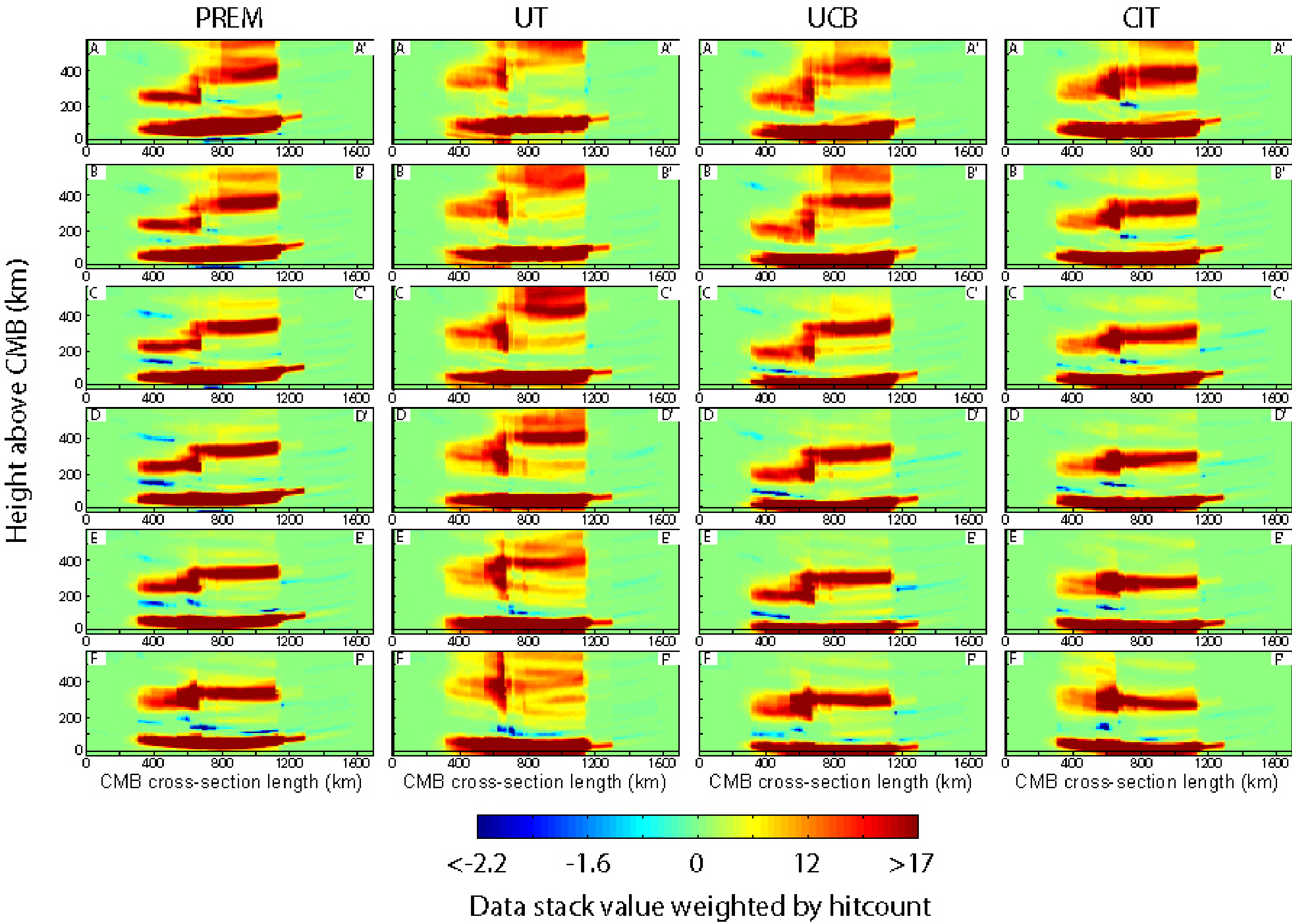


Supplementary Figure 4a [Hutko et al., 2006]

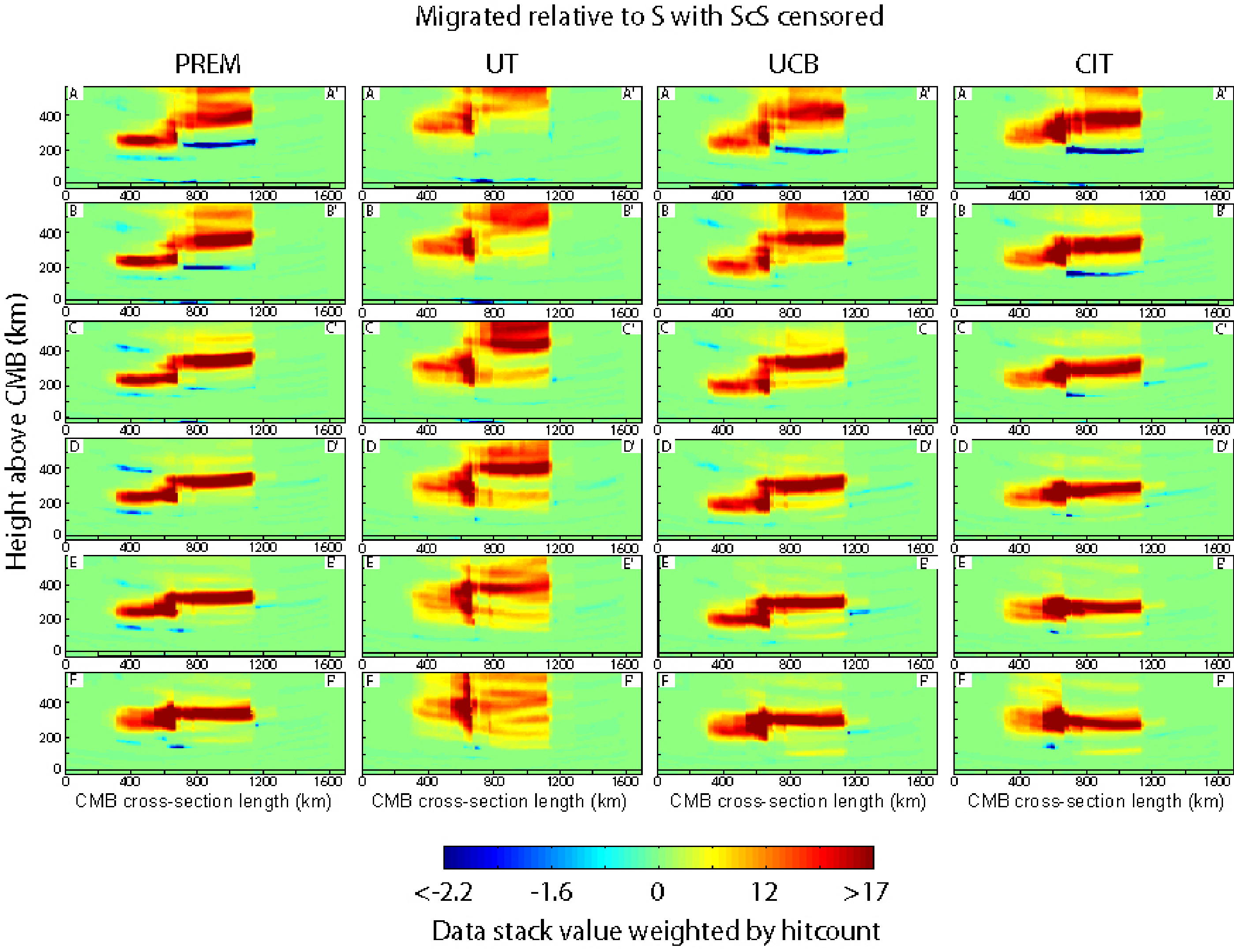


Supplementary Figure 4b [Hutko et al., 2006]

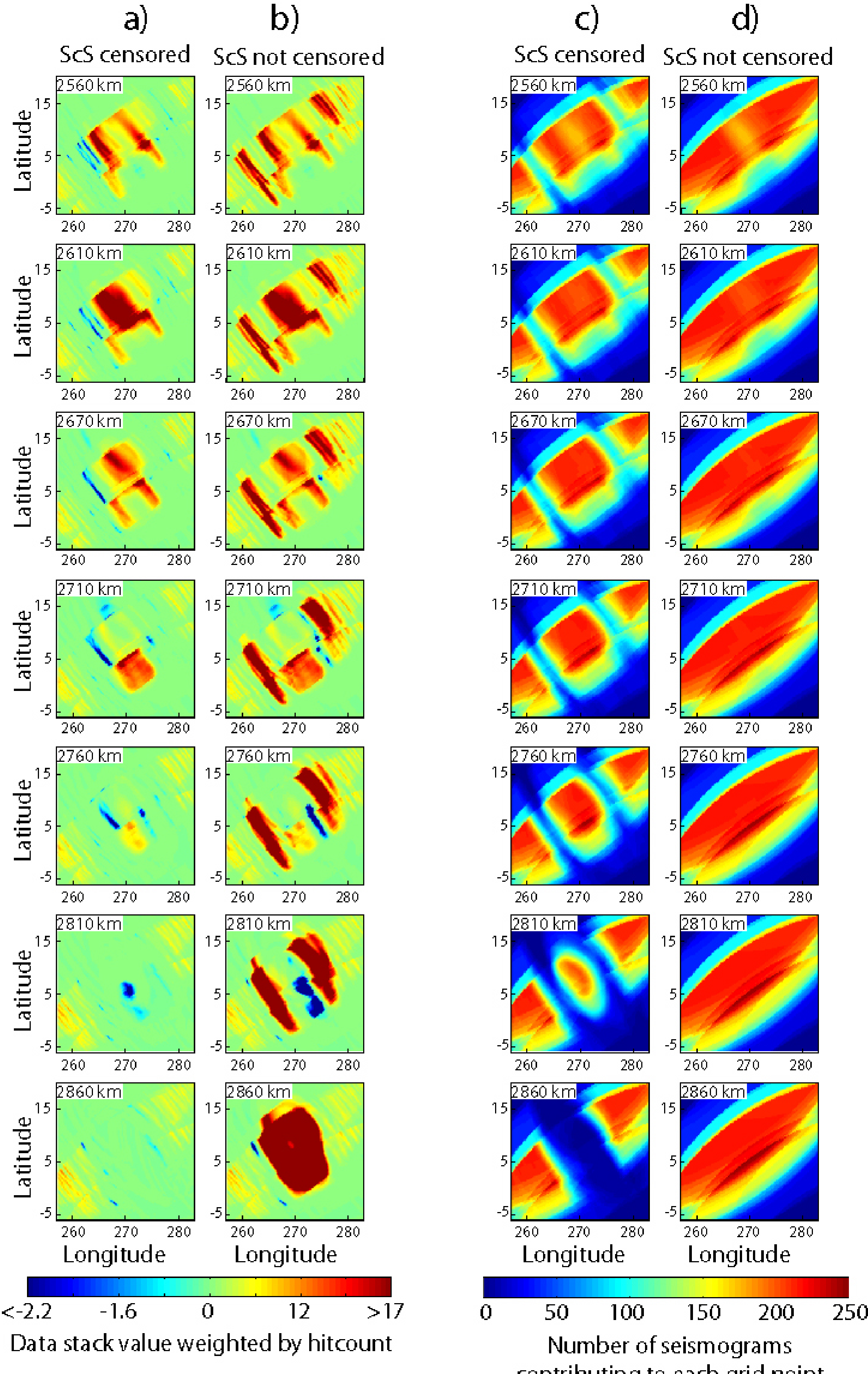
Migrated relative to S with ScS not censored



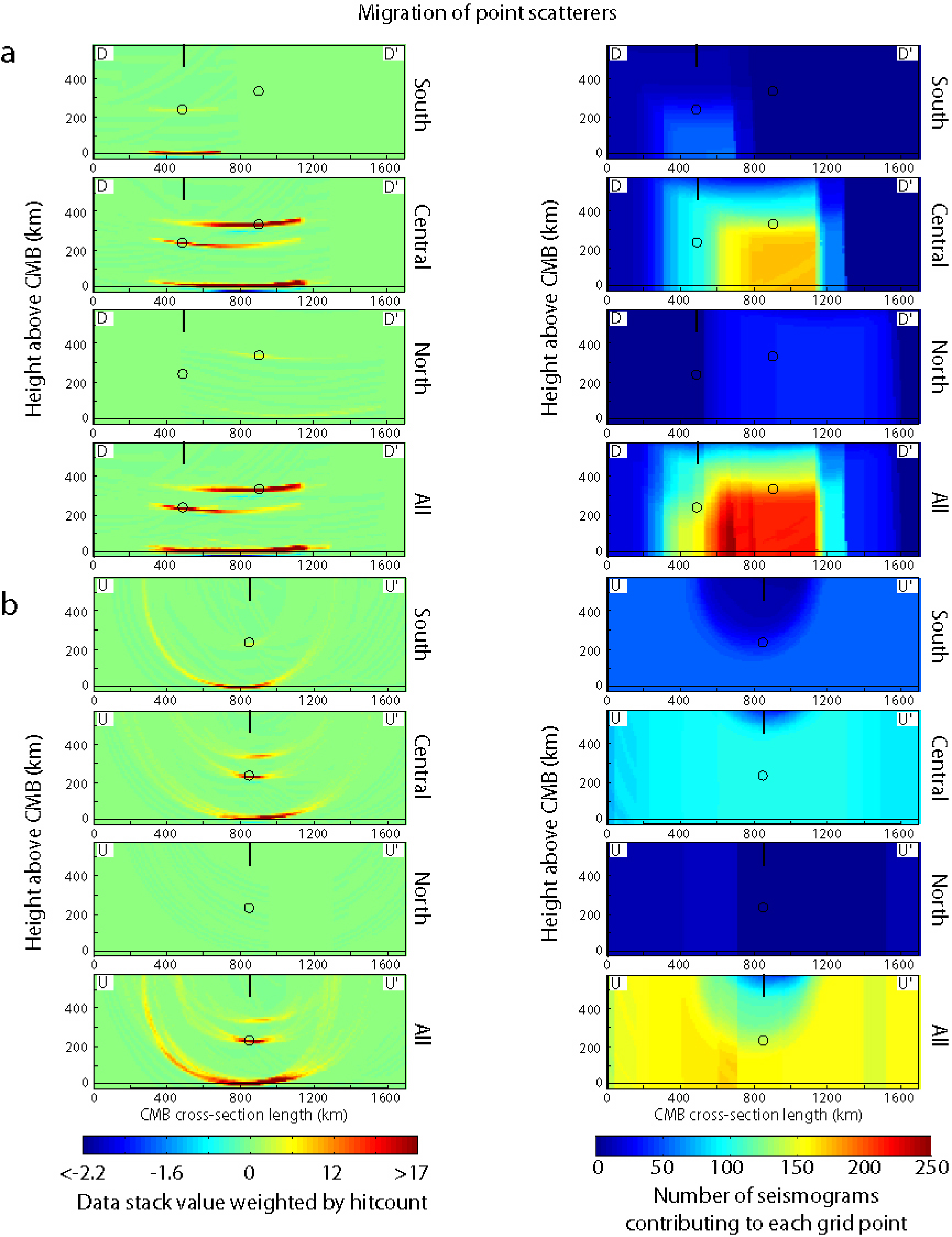
Supplementary Figure 4c [Hutko et al., 2006]



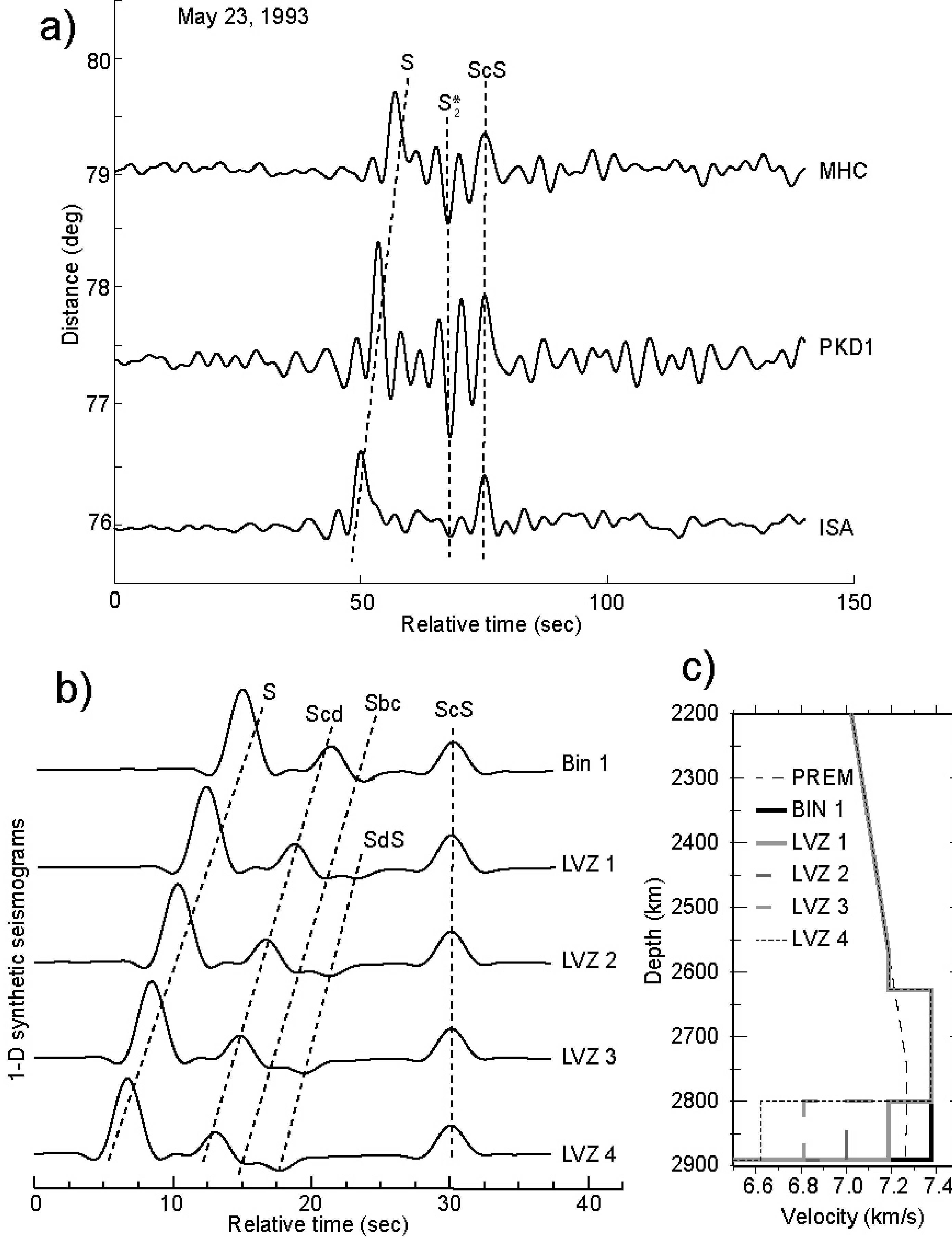
Supplementary Figure 4d [Hutko et al., 2006]



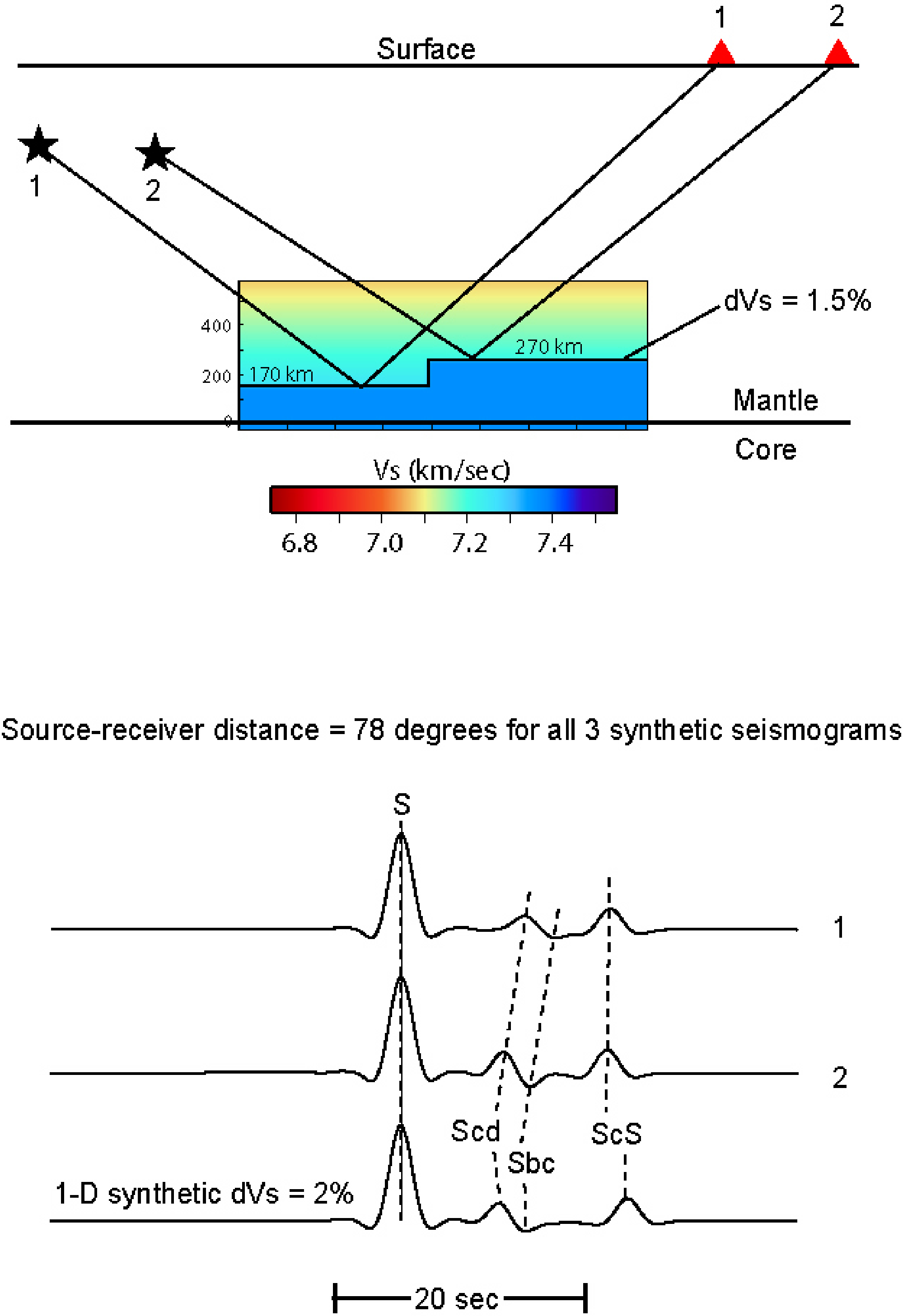
Supplementary Figure 5 [Hutko et al., 2006]



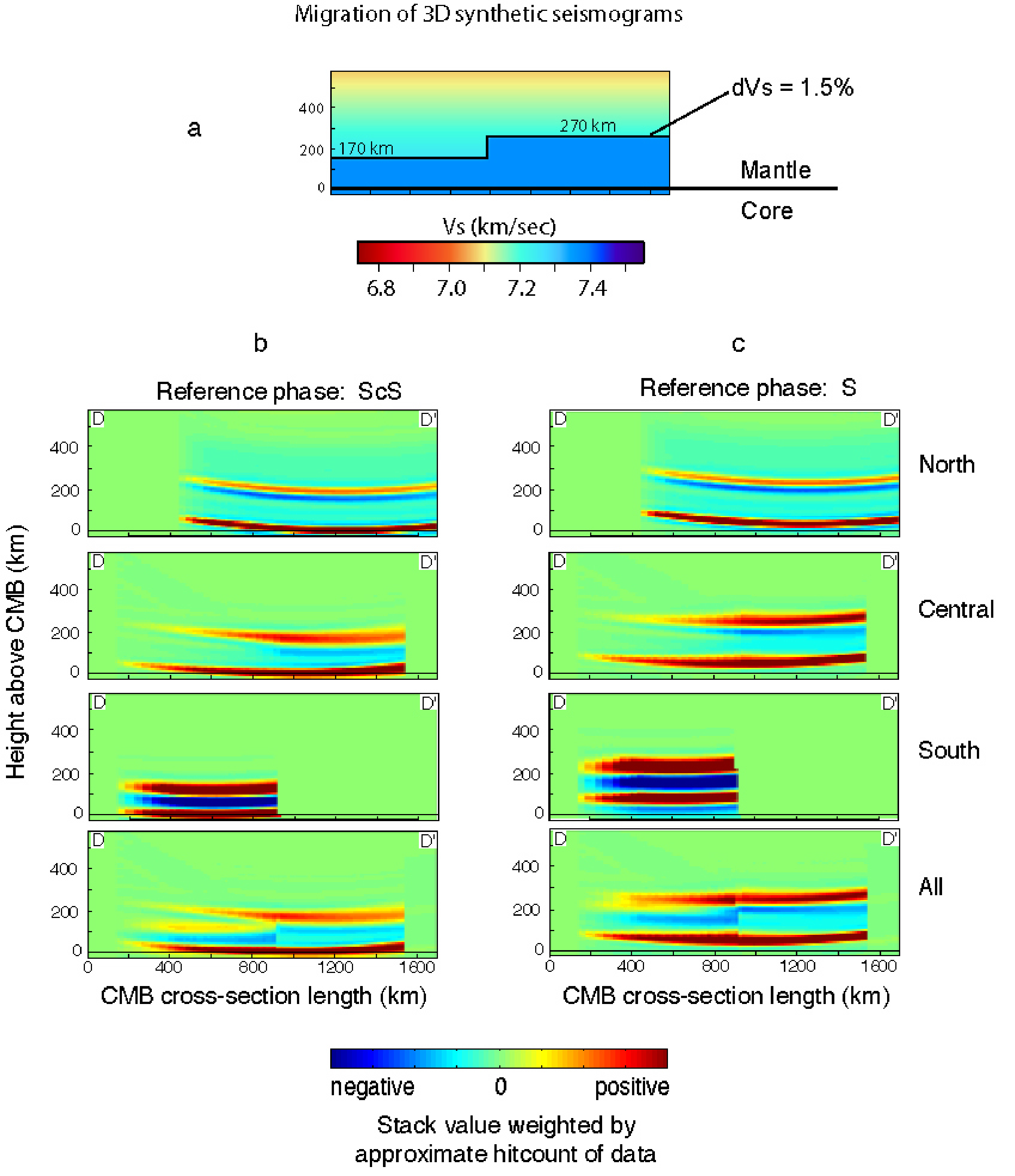
Supplementary Figure 6 [Hutko et al., 2006]



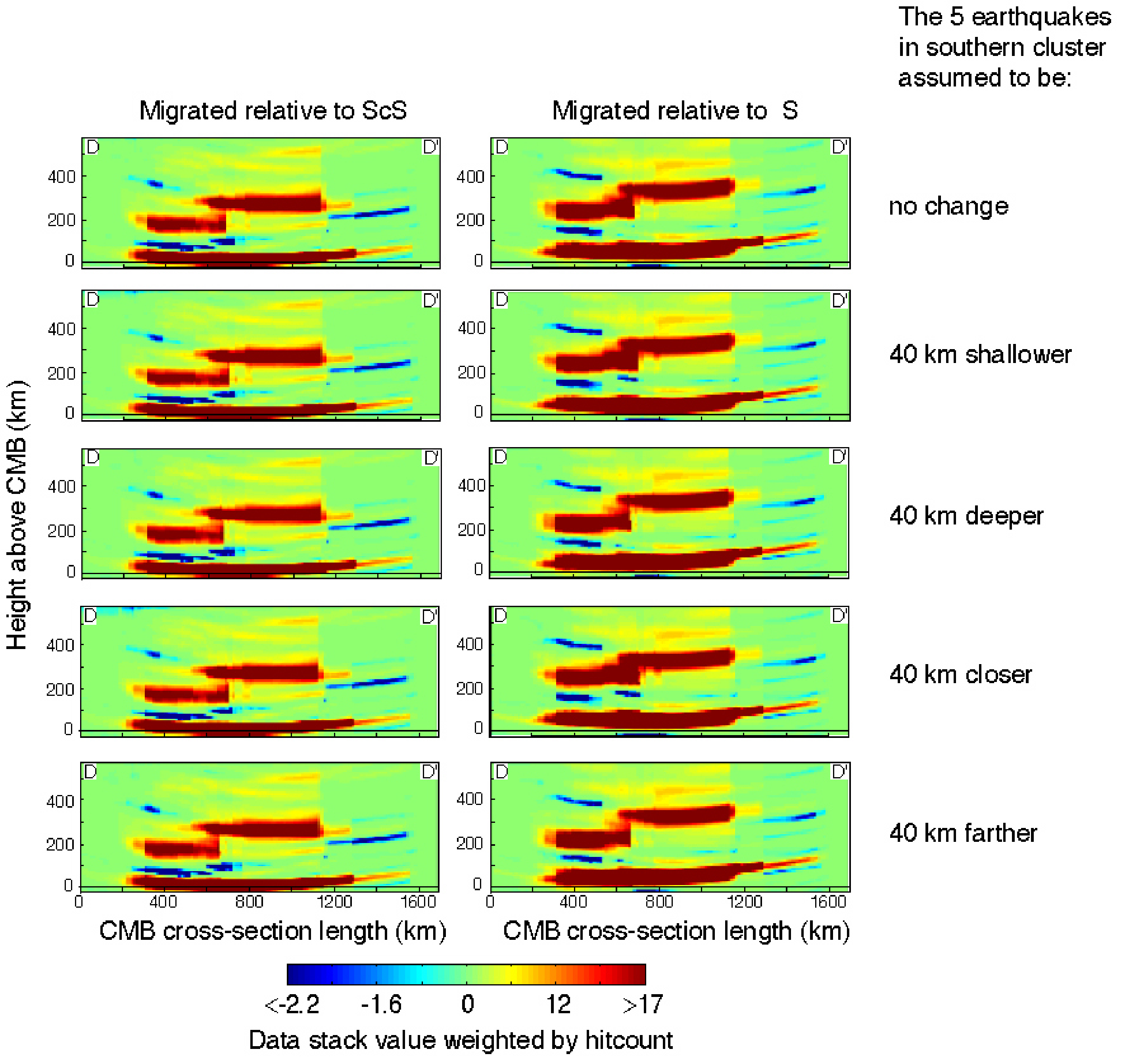
Supplementary Figure 7 [Hutko et al., 2006]



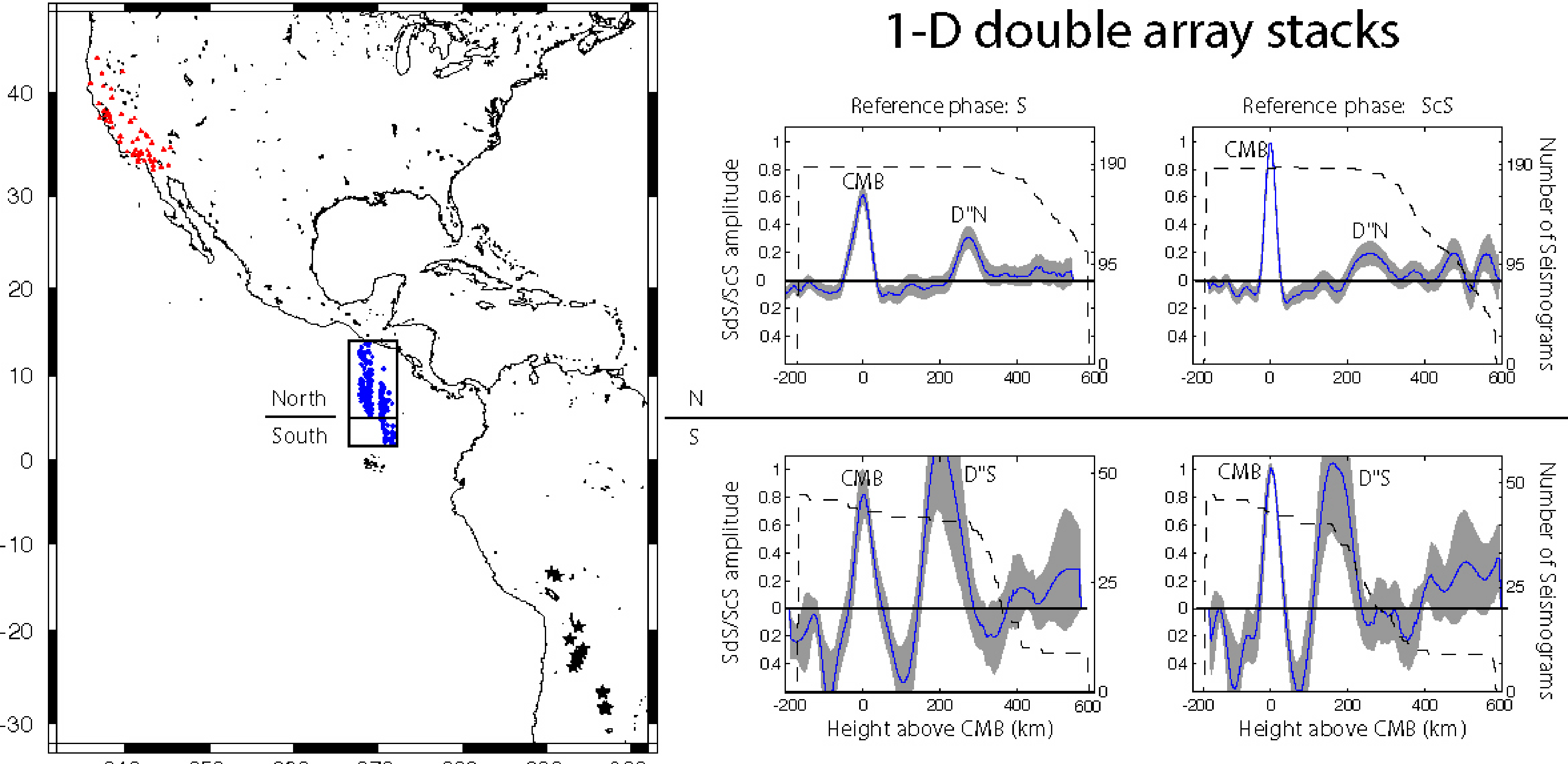
Supplementary Figure 8 [Hutko et al., 2006]



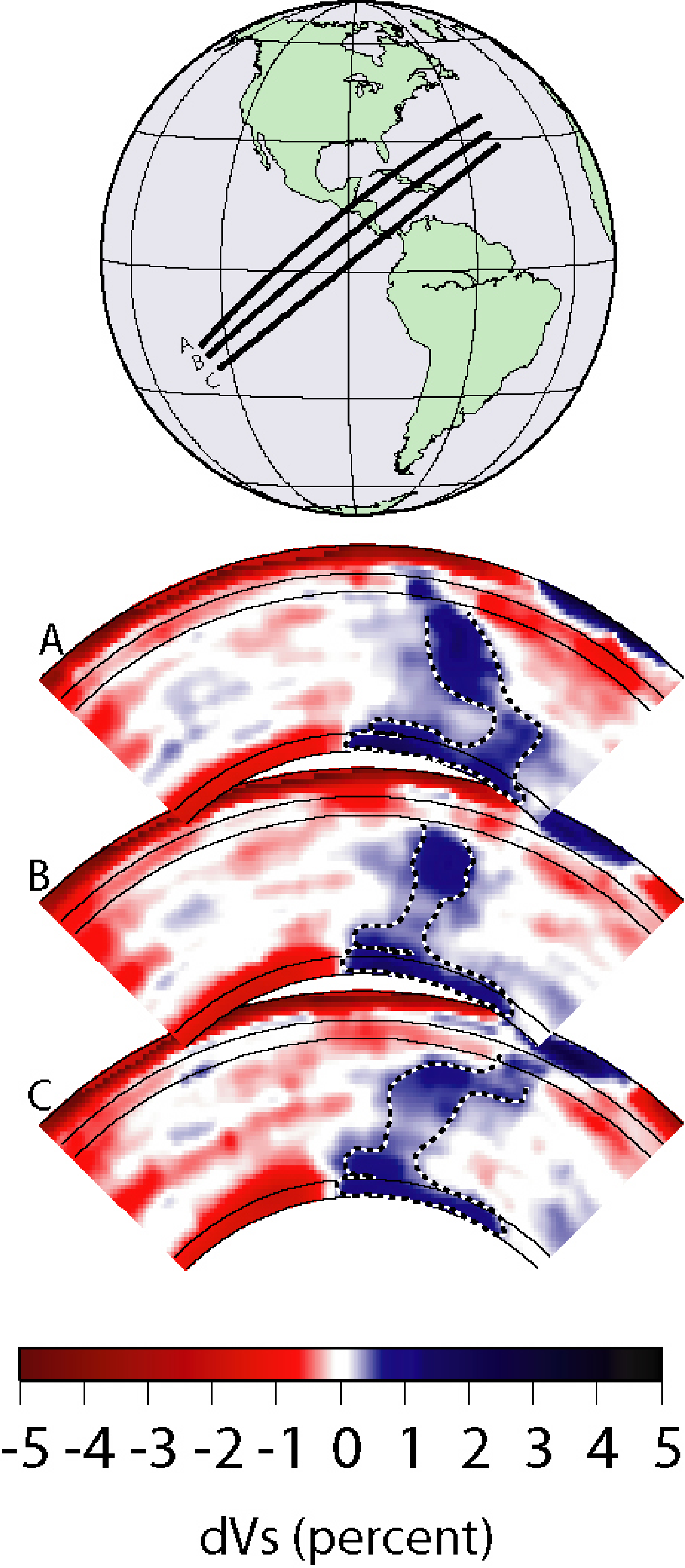
Supplementary Figure 9 [Hutko et al., 2006]



Supplementary Figure 10 [Hutko et al., 2006]



Supplementary Figure 11 [Hutko et al., 2006]



Supplementary Figure 12 [Hutko et al., 2006]