The velocity structure and heterogeneity of the upper mantle *

B.L.N. Kennett and J.R. Bowman

Research School of Earth Sciences, Australian National University, G.P.O. Box 4, Canberra, A.C.T. 2601 (Australia)

(Received March 20, 1988; revision accepted February 13, 1989)

In a particular region a smoothed and averaged representation of the velocity structure in the upper mantle can be derived from long-period body wave and higher-mode surface wave observations. The vertical resolving power of such techniques is limited by the relatively long wavelengths used. In contrast, short-period studies show considerable complexities in postulated one-dimensional models which are unresolvable in the long-period band. However, these may be associated with the mapping of lateral heterogeneity into a vertical profile.

To assess the extent to which lateral heterogeneity in the mantle can be resolved, we have attempted to model the features seen on recordings of earthquakes in the Indonesia/New Guinea region at large-aperture (400–1000 km) arrays of portable seismic recorders. With horizontal heterogeneity scales of the order of 300–500 km and velocity perturbations up to 2%, the amplitude deviations from those predicted for a stratified model can be considerable and fit in well with the pattern of observations. For long-period waves such laterally heterogeneous structures have only a modest effect, with weak inter-mode coupling resulting from the heterogeneity. As a result, a one-dimensional model may still be a useful approximation for long-period studies, but will not be adequate for broad-band or short-period work.

1. Introduction

Over the last 30 yr substantial effort has been expended by many workers on determining the structure of the upper mantle. The dominant approaches have been to use long- and short-period body wave observations and the dispersion of surface waves. In the body wave studies, attention is directed to the energy return to the surface from the structure in the upper mantle. In the surface wave work, on the other hand, the waves are generally observed after a longer propagation path during which multiple interactions with the structure have occurred. Thus body wave methods generally give an idea of local structure, and surface wave dispersion is dependent on a horizontally averaged structure. Recently, the distinctions between these approaches have begun to blur as advances in the calculation of theoretical seismograms have led to the exploitation of more of the seismic waveform, particularly for S waves. Body wave studies have included surface multiple phases (Grand and Helmerger, 1984), and multi-mode surface wave analyses have included many body wave phases (Nolet, 1977; Lerner-Lam and Jordan, 1983; Nolet et al., 1986). However, the best-developed models are dependent on the assumption of a stratified upper mantle, even though it has long been recognized that there has to be noticeable regional variation in the seismic velocities in the upper mantle (Herrin, 1969).

By treating the lateral variability in the Earth's structure as a perturbation to a radially stratified...
model, Woodhouse and Dziewonski (1984) have been able to map the large-scale structure in the upper mantle. They have matched global observations of shear wavetrains with theoretical seismograms to derive a spherical harmonic expansion of the heterogeneity in the upper mantle. The horizontal resolution available is of the order of 2000 km, and the major features such as continental shields are clearly visible in the shallow parts of the model. The horizontal scale currently attainable by this approach is thus of the same order as the distance over which energy is returned from the top 800 km of the mantle.

In this paper we will be concerned with the character of the heterogeneity spectrum on a regional scale that lies beneath the resolving power of the global studies. Nearly all previous attempts to look at regional heterogeneity in the upper mantle have been based on tomographic methods (e.g., Romanowicz, 1980; Spakman, 1986). Such approaches are limited in their horizontal resolution by the distribution of seismic stations and work well when there are major velocity contrasts present which may, for example, arise from the tectonic evolution of a region. These studies rely on the travel time differences in transmission from the source, with the assumption that any heterogeneity is concentrated in the upper mantle. Grand (1987) has shown how such work can be combined with studies of the travel times of multiply reflected body waves on long-period records, to give a more uniform data distribution and thus enhance resolution.

The use of permanent seismic stations means that, in general, it is not possible to track the detailed evolution of the seismic wavefield with position, which would be diagnostic of the character of the mantle heterogeneity field. However, as we shall see in the next section, the use of arrays of portable seismic recorders can fill in the detail and so provide improved constraints on the nature of the spatial variation of the seismic velocity field.

We can obtain some idea of the degree of heterogeneity in the upper mantle by comparing different stratified velocity models. In Fig. 1 we illustrate a representative velocity model for a shield region (S8—Burdick, 1981) and have tried to indicate by the shaded region the zone of velocity variation spanned by different velocity models. The zone above 200 km shows the largest variability associated with the different tectonic environments sampled in different regions, but significant variation persists to considerable depth in models derived from shorter-period data. Those models generated by the modelling of long-period body wave arrivals show much less spread beneath 300 km and their velocity gradients are generally concordant with spherically averaged models such as PREM (Dziewonski and Anderson, 1981). The major velocity transitions near 400 and 650 km stand out even in the envelope of a suite of models, but individual models have significant variations in the velocity contrast and the character of the transition. The vertical resolving power of long-period data is limited and different velocity models can give rise to very similar waveforms (Ingate et al., 1986).

The style of published models differ considerably. Some models are constructed to fit the major features of the travel time curves and so have rather simple character, as, for example, in the work of King and Calcagnile (1976). However, when an attempt is made to match a suite of observations in detail, rather complex velocity models result (Hales et al., 1980).
For a given region, one should therefore view the shaded zone in Fig. 1 as providing outside limits on any particular vertical velocity profile drawn from the ensemble which collectively defines a laterally heterogeneous model. The allowable deviations in seismic velocity away from a smooth, stratified, reference model would generally be less than ±3%, but we need additional information to constrain the horizontal scale of variation. We shall see subsequently that horizontal heterogeneity scales of the order of 200–400 km at depths below 200 km can explain much of the observed seismic behaviour.

2. Observations on the Australian shield

To place strong constraints on upper mantle structure we need to have a relatively close network of short-period observations. We recall that features which may appear sharp from long-period observations will not necessarily be sharp when subjected to the greater vertical resolution of shorter-period waves. However, the shorter wavelengths are also more sensitive to the presence of lateral heterogeneity, and such effects will tend to degrade the attainable vertical resolution. With sufficient data control we may hope to obtain a good representation of the mean velocity structure in a region as well as information on the character of the deviations from that structure.

To this end, the Research School of Earth Sciences at the Australian National University has deployed a number of arrays of portable seismic recorders over the last 12 yr. These arrays were designed to record the abundant earthquakes from the Indonesian and New Guinea seismic belts, with propagation paths which lie almost entirely within the Indo-Australian plate. New Guinea is separated from Australia by the shallow Arafura

![Temporary Array Studies](image-url)

Fig. 2. Location map for the portable array stations and the earthquakes displayed in the record sections. The NWB and TCT portable arrays are indicated by the open symbols and the position of the permanent Warramunga array (WRA) is also shown. The solid pentagon marks the event used with the TCT array in Fig. 3 and the solid triangles the events used in the composite section of Fig. 4.
Sea, which is underlain by continental material, and a broad continental shelf reaches to almost the Indonesian arc structures (Fig. 2).

The arrays have been deployed at a variety of spacings to give a total aperture ranging from ~400 km (e.g., NWB) to ~1000 km (e.g., TCT). In each case, a larger aperture can be synthesized by using multiple sources. The size of these receiver arrays is comparable to the depth of the upper mantle and so provides an excellent opportunity to investigate the heterogeneity scales within a single region of the upper mantle. The duration of recording for each array is ~3 months, which provides a reasonable number of medium-sized events, sufficiently large to be effectively located but not so big as to have very complex source radiation ($4 \leq m_b \leq 5.5$). The early, large-aperture, arrays consisted entirely of direct recording analogue units, but the more recent medium-aperture arrays have been supplemented by an array of portable digital units with an aperture of ~100 km. In addition, the observations for a number of the portable arrays can be linked to the permanent Warramunga array (WRA), which has an L-shaped configuration and a 25 km aperture (Cleary et al., 1968). WRA lies near the junction of the north–south and east–west arms of the NWB array in Fig. 2 and close to the western end of the TCT array. The combination of the various arrays gives a detailed data set which covers scales from 25 to 1000 km for propagation paths from the same general location.

If we look at the records for a single event, we find that a number of different seismic phases can often be traced over considerable distances (Fig. 3). However, the relative amplitude of different parts of the seismogram can differ markedly across the record section, with a very significant effect on the visibility of later phases. For example, in Fig. 3, for an Indonesian event of 66 km depth recorded at the TCT array, there is a prominent later phase which is especially clear at ranges from 1600 to 1800 km and again from 2200 to 2500 km. Over the intervening region this arrival is much less clear and the time–distance curve is more subjective. The timing would fit with an arrival refracted below the 400 km transition. Part of the amplitude variation may be associated with interference with the reflection from the transition. However, the change in waveform over only a 50 km interval is very difficult to explain with a stratified model.

Such a variation in the clarity of a phase is very
common and the behaviour bears no simple relation to the particular recording site. We cannot therefore ascribe such effects to local geological structure, but rather must look for a deeper cause for variability on a 50–300 km scale.

We may also synthesize a large-aperture array by superimposing the records from a number of events at a medium-aperture array. With multiple earthquakes, the variation in the character of the records from event to event can be quite striking, even when the propagation paths are similar. A composite record section from the NWB array is shown in Fig. 4 for a number of shallow Indonesian earthquakes, whose position are indicated by the solid triangles in Fig. 2. Each trace in the section is normalized to the same peak amplitude. This normalization has the effect of suppressing the very-low-amplitude first arrivals and so all the prominent features in the section arise from later phases. The prominent arrival at ~20 s reduced time near 1700 km can be tracked to both shorter and longer ranges, but many of the traces with a distinct arrival come from the same event.

With the high density of data coverage achieved in Fig. 4 one can begin to correlate phases across many seismograms, even where arrivals are only weakly represented on individual traces. However, the interpretation is likely to be strongly influenced by the dominant traces. The events which comprise Fig. 4 have all been recorded at the same array so that structure beneath the array should have a comparable effect on all events. The propagation paths from these Indonesian earthquakes to the NWB array cover a relatively narrow range of azimuths, and the turning point for the upper mantle phases should lie in a zone of ~200–300 km width beneath the northwestern coast of Australia. Some part of the change in character between events may arise from the source radiation pattern. Strike-slip events along the arc give particularly efficient P radiation in the direction of the NWB array, but typical normal and thrust

---

Fig. 4. Composite record section of events from the Indonesian arc recorded at the NWB array. To reduce clutter the traces have only been displayed at 2 km intervals, with normalization to the same peak amplitude. The prominent later arrival from the 400 km transition shows a variable aspect across the section.
mechanisms are not as favourable. However, it is difficult to arrange the source mechanism so as to selectively suppress an arrival from the 400 km transition and allow shallower propagating energy to arrive. We therefore need to look for some structural explanation for the variability of the relative amplitudes of different upper mantle phases in these short-period studies.

On smaller scales, Kennett (1987) and Korn (1988) have shown that the complexity of the recordings of Indonesian events at the 25 km aperture Warramunga array is reduced as the depth of the source increases. Scattering effects are particularly prominent for paths lying above ~ 80 km depth. The simplest records are obtained for source depths > 200 km, for which the propagation path includes only a relatively short leg in any shallow zone of concentrated heterogeneity.

### 3. Models of upper mantle heterogeneity

In the previous section we have seen that detailed short-period studies of the upper mantle display a number of features which cannot be described by an individual stratified model. In particular, we have identified amplitude anomalies in records of a single event on scales of 150–300 km, and significant variation in the character of seismic records between different events.

Nevertheless, we must recall that for many long-period studies a radially stratified model appears to give a good description of the seismic structure of the upper mantle. For multi-mode surface wavetrains with periods > 20–30 s, a stratified model will allow a detailed fit to both dispersion and waveforms. Any model which we propose must therefore be such that it is transparent to such horizontally travelling long-period waves. In addition, we would like to explain the success of modelling long-period body wave phases by matching theoretical seismograms calculated for stratified Earth models.

#### 3.1. Laterally heterogeneous models

As we are not seeking to explain one particular set of observations but rather the general character of many observations, we will not attempt to construct a specific model but will endeavour to propose a class of laterally heterogeneous models which have the required properties.

We require that any suitable model should appear to be stratified for low-frequency propagation and so it is effective to parameterize the laterally heterogeneous structure as perturbations away from a stratified reference model. Such perturbations should not be greater than ~ 2% positive or negative excursions from the reference, to avoid venturing too far outside the bounds derived from a wide range of proposed mantle models (Fig. 1).

After some experimentation, it proved possible to find a class of models which have the required properties: these models are characterized by horizontal variation on a scale of 300–500 km with a vertical scale which increases with depth from ~ 100 km scale at 200 km deep to ~ 200 km scale at 900 km depth. Three two-dimensional simulations of such a class of models are shown in Fig. 5. Each of these cross-sections of the perturbation to the reference model was generated by randomly perturbing the velocities specified at 100 km horizontal intervals with a uniform random distribution, allowing up to ±2% change. The models were then smoothed with a 500 km moving window to remove rapid oscillations in velocity values. This imposes a preferred scale length of ~ 400 km with a perturbation typically < 1%. The velocity gradients, both horizontally and vertically, are strongest in relatively narrow belts and have a strong effect on short-period wave propagation. The model in three dimensions would have a similar modulation of the reference velocity structure with horizontal scales again of the order of 400 km.

The reference model chosen for the construction of specific laterally heterogeneous sections was the isotropic version of the PREM model (Dziewonski and Anderson, 1981) with a superimposed continental crust, shown as the continuous line in Fig. 7. The basic perturbation was constructed for P waves, and the S wave model was derived assuming that the percentage change from the reference model was the same as for the P waves, which may represent an underestimate of
the effect of the heterogeneity. Within the heterogeneous region, the displacement and traction fields are represented on each depth section as a sum of the appropriate eigenfunctions for the reference structure. In a stratified region, the coefficients in the modal expansion at a particular horizontal location would comprise a fixed part determined by the modal content at the entry location, multiplied by a phase increment determined by the phase velocity for each mode. In the presence of heterogeneity it is worthwhile to continue to extract the dominant phase roll term for each mode, but now the remaining part of the modal coefficient is position dependent. These modal coefficients are determined by a set of coupled first-order differential equations whose entries depend on the effect of the deviations from the reference model integrated over the whole depth section. To avoid numerical problems with the specification of boundary conditions, it is convenient to work directly with reflection and transmission matrices for modal interaction. These matrices satisfy a nonlinear Ricatti differential equation with a simple boundary condition, which can readily be integrated numerically starting from the far side of the heterogeneous region. In effect, a larger region of the heterogeneity model is exposed with each step of the integration. Initially, there is no reflection and full transmission, but the matrices are modified by the properties of the heterogeneous region.

For our purpose we are interested in the transmitted modal field after passage through the models shown in Fig. 5. To ensure that all S body wave interactions down to 1000 km as well as surface wave effects are included, we have taken the modal expansion to cover all modes with phase velocities \(< 7 \text{ km s}^{-1}\) at each frequency. The number of relevant modes rises nearly linearly with increasing frequency: at 0.047 Hz a total of nine modes are needed; by 0.070 Hz this has risen to 14; and 28 modes are required at 0.141 Hz. As the computation time rises as the square of the number of modes, substantial computational effort is required at the higher frequencies.

For frequencies of \(< 0.05 \text{ Hz}\) the modal transmission is almost complete, with any transfer be-
between modes involving <1% of the incident energy. Thus for periods longer than 20 s, the seismograms after passage through the heterogeneous region would be essentially indistinguishable from those for an equivalent stratified zone.

As the frequency rises, the transfer of energy between modes becomes more important, with consequent distortions of the propagation pattern. At 0.141 Hz there can be a substantial reduction in the proportion of an incident mode transmitted through the heterogeneous model without change of mode type, to only 70% or so of the incident energy. The remaining part is redistributed into other modes, predominantly into the immediate neighbours of the original mode. This transfer of energy is selective depending on the particular heterogeneity model, but can give significant coupling between modes 3 and 6 associated with shallow structure, and also between groups of higher modes which would interfere to give the S wave phases returned from the upper mantle. The distortion of the incident modal distribution is considerable but the behaviour is dominated by straight transmission. This frequency (0.141 Hz) is at the upper limit of those included in modelling body wave phases on long-period records, so we can see why a stratified medium can constitute a useful description of longer-period phases.

At even higher frequencies (and thus shorter wavelengths) the number of modes becomes very large and there is complex interaction between modes as the heterogeneity effects become dominant. In this regime the modal description is rather cumbersome and it is preferable to look directly at the body wave propagation.

3.3. Body wave analysis

In Fig. 6 we consider a comparison between the travel-time behaviour for the reference model C (marked by open symbols) and the heterogeneous model G (with the solid symbols) for a source at 10 km inside the left-hand edge of the model at a depth of 60 km. The take-off angles at the source were designed to give a relatively uniform sampling of the travel time branches in the case of the reference medium, but the pattern of symbols is noticeably distorted in the presence of heterogeneity. The clumping and thinning of the ray pattern is associated with the development of local caustics, generated by the velocity gradients at depth, of the type discussed by Nye (1985). These caustics will be associated with significant changes in amplitude, with maxima in those parts of the seismic section where solid symbols are concentrated. Once we allow for diffraction effects, the amplitude anomalies have a scale of the order of 150–300 km, in agreement with the observations from the portable arrays.

In addition to arrivals representing distorted but recognizable travel time branches from the reference model, the combination of horizontal
and vertical gradients in the heterogeneous model can combine to produce new features. For example, in the case of model G, an apparent reflection from 500 km depth arises from the gradient structure near 1200 km range, which leads to a complex interference of travel time curves near 2200 km in the time section. Because of our assumption of comparable heterogeneity scales in the transverse direction we can anticipate that there could be abrupt changes in the character of an arrival across a spatially distributed array arising from the pattern of caustics at the surface. Such effects are seen for some events at the digital subarray of the NWB experiments.

In Fig. 6, for clarity, we have shown only the time section for one of the heterogeneous models from our favoured class. If, however, the time sections for the three models shown in Fig. 5 are superimposed, the scatter of ~1 s around each travel time branch is similar to that observed from upper mantle phases (Simpson, 1973), and the amplitude distributions show a markedly different character.

To investigate the effects of neglecting the presence of lateral heterogeneity, we have undertaken inversion of the travel time curves for models G, M and N to produce stratified Earth models. The inversion was undertaken in the tau—p domain, using the reference model as the starting point (Kennett, 1976). It proved difficult to construct stratified models which would give a satisfactory fit to data from the heterogeneity suite, even with an allowance for generous 'picking' errors. Representative models derived from the inversion are illustrated in Fig. 7, together with the reference model C shown as a continuous line. There was little constraint on the behaviour below the 670 km discontinuity and this is reflected in the structure produced for model M.

The most noticeable feature of Fig. 7 is that each inversion for a heterogeneous case tends to wobble about the reference model with alternate zones of positive and negative gradients and perturbations of the order of 1% from the reference. Such an alternation is often found in previous models which attempted to explain the fine structure of a record section in terms of a radially stratified model. Although Fig. 7 is presented in the form of a depth section, it should be recalled that different parts of a heterogeneous model are sampled in constructing the travel time curves. The models in Fig. 7 should therefore be viewed as actually sampling along a trajectory corresponding to the turning points of rays in the reference model. When we trace out this path we find that there is a reasonable correlation between the inferred deviations from the reference model in the stratified inversions and the actual perturbation at the corresponding depth in the neighbourhood of the turning point. For example, the small low-velocity zone above the 670 km discontinuity in the inversion for model M correlates with the alternation of high and low velocities in this region around 1200 km in the original heterogeneous model. Thus with a number of different profiles in a region it is possible to make some progress towards inferring the character of the heterogeneity field from one-dimensional inversions.

4. Discussion

In this paper, we have shown that within a single region of the upper mantle such as the zone
beneath the northern Australian shield, there is significant heterogeneity in seismic properties on scales of hundreds of kilometres. This scale of variation is too small to be detectable by global studies and fits within the scale of a single unit in global heterogeneity models. The size of the features we propose to occur in the upper mantle is such that they have the greatest effect on short-period waves, whereas long-period waves can pass through the structure with comparatively little interference.

The nature of the observations which lead to the development of our laterally heterogeneous models of the upper mantle is that they are most sensitive to features at the 100–400 km scale level. There is evidence from the smaller, permanent, Warramunga array of complications in the seismic wavefield on somewhat smaller scales (Kennett, 1987). However, it is interesting to note that an entirely different analysis of the complexities of surface wave trains in Europe by Snieder (1987) also infers horizontal scales of the order of 300 km.

A full model of the lateral heterogeneity in the upper mantle must take into account a hierarchy of scales from tens to thousands of kilometres. The shortest-scale features are most evident near the surface and it appears likely that the dominant horizontal and vertical scales of heterogeneity increase with depth. In the 200–800 km depth range a model with horizontal variations at a scale of 200–400 km appears to give a good representation of the major behaviour of the seismic wavefield. Smaller-scale features may be superimposed, which give rise to scattering and hence the complex observed waveforms, but these are concentrated above 200 km (Korn, 1988).

Analysis of the detail of a set of short-period observations from such a heterogeneous region in terms of a stratified model is likely to be misleading. However, it is reasonable to infer a reference structure which fits the main behaviour and then seek to map out the distribution of heterogeneity. Such a laterally heterogeneous model will be needed to provide a full description of the wavefield for both short-period and broad-band studies.

Within a single major unit of the upper mantle, a stratified model can provide a very effective summary description of the dominant character of the seismic velocity distribution (see, e.g., Grand and Helmberger, 1984). However, the highest-frequency waves used in the study of long-period body wave phases are beginning to be sensitive to the detailed nature of the heterogeneity field and this could give some mild bias to the resulting models. Strong horizontal velocity gradients within the mantle have been recognized from the modelling of the body waves on long-period records (Rial et al., 1984) corresponding to regions with different tectonic histories. The even longer-period studies of multi-mode surface waves could be reasonably well described by a stratified velocity model within one region. However, the long propagation paths commonly analysed may have sampled a number of such regions and so be influenced by larger-scale heterogeneity.

Acknowledgements

We are grateful to D.R. Christie and the staff of the Warramunga array for their support of the portable array program, and to J. Hulse and G. McAllister for their efforts in the processing of the array data.

References


