Geophysical Journal International

Geophys. J. Int. (2013) **194,** 1042–1049
Advance Access publication 2013 May 3

doi: 10.1093/gji/ggt144

Finite-frequency sensitivity kernels for two-station surface wave measurements

Denise de Vos, Hanneke Paulssen and Andreas Fichtner

Department of Earth Sciences, Utrecht University, PO Box 80.021, 3508 *TA Utrecht, the Netherlands. E-mail: D.deVos@uu.nl*

Accepted 2013 April 8. Received 2013 April 2; in original form 2012 November 20

SUMMARY

We present a method for the computation of finite-frequency sensitivity kernels for two-station surface wave measurements. It is based on the combination of spectral-element modelling of seismic wave propagation, adjoint techniques and traveltime estimates from two-station cross-correlations of seismograms. The analysis of sensitivity kernels for a 1-D earth model resulted in two major conclusions: (1) The finite-frequency sensitivity is zero along the interstation ray path for group velocity measurements obtained by cross-correlations. It follows that interstation group velocity measurements should not be made by cross-correlation if a ray-based interpretation is used. (2) Although sensitivity along the interstation ray path is dominant for phase velocity measurements, sensitivity far from the ray path can be large, depending on the details of the source–receiver geometry. The complexities of finite-frequency two-station sensitivity kernels should be taken into account to avoid misinterpretations and to improve the quality of tomographic inversions.

Key words: Surface waves and free oscillations; Seismic tomography; Theoretical seismology; Wave scattering and Diffraction; Wave propagation.

1 INTRODUCTION

The two-station method (Sato 1955) is often used for regional surface wave tomography and local-scale interstation measurements of phase velocity (e.g. Yao *et al.* 2006; Zhang *et al.* 2007; Endrun *et al.* 2008). The outstanding advantage of this method lies in the reduction of wave propagation effects from the source to the nearest receiver. Assuming that ray theory is valid and that waves propagate along the exact great circle, effects of unknown 3-D structure away from the interstation ray path cancel completely, thereby enabling detailed and reliable investigations of the Earth's local structure.

In the hypothetical absence of lateral heterogeneity, the two-station method would be exact for pure great circle propagation. The phase velocity between two stations could then be estimated from the cross-correlation of seismograms recorded at both stations and filtered within a narrow frequency band. However, the presence of 3-D heterogeneity in the real Earth and the finite frequency content of seismic waves complicate the application of the two-station method. Several studies have shown evidence for propagation off the great circle (e.g. Baumont *et al.* 2002; Alvizuri & Tanimoto 2011). Also, phase velocity curves estimated from events at one side of the station pair can differ from those calculated from events at the other side of the station pair (Zhang 2009). This suggests that velocity perturbations away from the interstation ray path can affect the surface wave arrivals at the receivers. Furthermore, it is practically impossible to find earthquakes that are located exactly on the great circle of interest. Not taking the entire travel path into account, as is common in the two-station method, might therefore result in wrong conclusions concerning the phase velocity structure of the interstation area. Little research has been done to check whether the interpretation of the two-station method is adequate. Pedersen (2006) studied the effect of non-plane waves on phase velocity curves. For random heterogeneities far from the station pair, she found that only five to ten different events are needed to obtain an average phase velocity curve with less than 1 per cent error. However, these results are too specific for general applications because the non-random nature of real, unknown velocity structure and the radiation patterns of the events were not taken into account.

To further improve tomographic inversions based on the two-station method, the effects of 3-D heterogeneity on the measurement of a particular event should be quantified in the form of properly calculated finite-frequency sensitivity kernels. Sensitivity kernels provide information on how specific measurements react to perturbations of elastic properties (e.g. *P* or *S* velocity) anywhere in the earth model volume. They can be computed efficiently via the interaction of the 'forward' wavefield and an 'adjoint' wavefield (e.g. Tarantola 1988; Tromp *et al.* 2005; Fichtner *et al.* 2006; Liu & Tromp 2006; Fichtner 2011). The adjoint wavefield travels backward in time from the receiver to the source, and is excited by an adjoint source, located at the receiver position. In recent years, sensitivity kernels have been computed for various types of measurements and source–receiver geometries (e.g. Tromp *et al.* 2005; Liu & Tromp 2006; Fichtner *et al.* 2008; Bozdağ *et al.* 2011), as well as for interstation ambient noise cross-correlations (Tromp *et al.* 2010).

Earlier attempts to study the sensitivity of interstation phase velocity measurements were based on the calculation of 2-D sensitivity kernels (Kuo *et al.* 2009). However, they lack a direct relation to Earth structure in the form of 3-D perturbations in velocities or elastic parameters. The 3-D sensitivity kernels for interstation measurements presented by Chevrot & Zhao (2007) are based on phase difference, rather than the cross-correlation as used in the two-station method. Here, we properly define interstation traveltime measurements in terms of frequency-dependent cross-correlations between seismograms recorded at two stations. The resulting sensitivity kernels can be directly used in 3-D tomographic inversions of seismic velocity.

In the following, we will discuss the method used for calculating shear wave sensitivity kernels and present the results of the simulations. The influence of different parameters will be investigated, including the interstation distance, radiation pattern, frequency content, as well as the number and distribution of sources.

2 THEORY

Sensitivity kernels can be calculated via the adjoint method in the form of a product that involves the forward and adjoint wavefields (e.g. Tarantola 1988; Tromp *et al.* 2005; Fichtner *et al.* 2006; Liu & Tromp 2006; Fichtner 2011). The adjoint wavefield is excited by the adjoint source which is fully determined by the measurement. Consequently, the adaptation of the general adjoint method to a specific case reduces to the calculation of the adjoint source that corresponds to the particular measurement of interest. Our derivation of the adjoint source, as described below, is based on the operator formulation of the adjoint method (Fichtner 2011), and it can be seen as an adaptation of the wave equation traveltime inversion of Luo & Schuster (1991) to the two-station scenario.

2.1 Definition of measurements and misfits

For the two-station method, we define the measurement as the frequency-dependent traveltime between two stations, estimated from the interstation cross-correlation. Let $s_i(\mathbf{m}; \mathbf{x}_A, t)$ be the *i*- and $s_j(\mathbf{m}; \mathbf{x}_B, t)$ the *j*-component of the synthetic displacement seismograms for an Earth model $\mathbf{m}$ at locations $\mathbf{x}_A$ and $\mathbf{x}_B$, respectively. The cross-correlation between the (potentially filtered) seismograms at A and B is given by

$$C_{ij}(t) = \int_{\tau=-\infty}^{\infty} s_i(\mathbf{x}_A, t+\tau)\, s_j(\mathbf{x}_B, \tau)\, \mathrm{d}\tau\,, \tag{1}$$

where we omitted the dependence on $\mathbf{m}$ in the interest of a condensed notation. The time at which the global maximum of the cross-correlation occurs is defined as the synthetic traveltime, $T_{AB}(\mathbf{m})$, from station A to B. Because the time derivative of the cross-correlation is zero at its maximum, the following relation holds for $t = T_{AB}$:

$$\dot{C}_{ij}(T_{AB}) = \int_{\tau=-\infty}^{\infty} \dot{s}_i(\mathbf{x}_A, T_{AB}+\tau)\, s_j(\mathbf{x}_B, \tau)\, \mathrm{d}\tau = 0\,. \tag{2}$$

Eq. (2) defines T_{AB} implicitly. The measurement of the observed traveltime T_{AB}^{obs} from observed seismograms $s_i^{\mathrm{obs}}(\mathbf{x}_A, t)$ and $s_j^{\mathrm{obs}}(\mathbf{x}_B, t)$ follows the same procedure.

The L_2 misfit between synthetic and observed traveltimes is given by

$$\chi = \frac{1}{2}\left(T_{AB} - T_{AB}^{\mathrm{obs}}\right)^2\,. \tag{3}$$

A small perturbation of the model, $\delta\mathbf{m}$, induces the misfit variation

$$\delta\chi = \left(T_{AB} - T_{AB}^{\mathrm{obs}}\right)\, \delta T_{AB}\,, \tag{4}$$

where δT_{AB} and $\delta\chi$ are related to $\delta\mathbf{m}$ via the measurement sensitivity kernel vector $\mathbf{K}_{T_{AB}}$ and the misfit sensitivity kernel vector $\mathbf{K}_\chi$, respectively:

$$\delta T_{AB} = \int_V \mathbf{K}_{T_{AB}}(\mathbf{x}) \cdot \delta\mathbf{m}\, \mathrm{d}V\,, \tag{5}$$

$$\delta\chi = \int_V \mathbf{K}_\chi(\mathbf{x}) \cdot \Delta\mathbf{m}\, \mathrm{d}V\,. \tag{6}$$

The components of the kernel vectors represent the sensitivity kernels for different variables, such as *P*- or *S*-wave velocity, considered in the inversion. It follows from the combination of eqs (5) and (6) that the misfit and measurement sensitivity kernels are scalar multiples of each other, and the scaling factor is the traveltime difference $(T_{AB} - T_{AB}^{\mathrm{obs}})$:

$$\mathbf{K}_\chi = (T_{AB} - T_{AB}^{\mathrm{obs}})\, \mathbf{K}_{T_{AB}}\,. \tag{7}$$

Eq. (7) implies that the misfit sensitivity kernel $\mathbf{K}_\chi$, used in a tomographic inversion, has the same spatial pattern as the measurement sensitivity kernel $\mathbf{K}_{T_{AB}}$, which is independent of any observed seismograms. This is in contrast to data-dependent misfit sensitivity kernels for various types of phase misfits (e.g. Fichtner *et al.* 2008; Bozdağ *et al.* 2011) or the L_2 norm (e.g. Tarantola 1988), and it allows us to draw general conclusions from the analysis of $\mathbf{K}_{T_{AB}}$ without the need to involve actual observations.

2.2 Derivation of the adjoint source for two-station measurements

We can find the adjoint source $\mathbf{f}^\dagger$ that produces the measurement sensitivity kernel $\mathbf{K}_{T_{AB}}$ by bringing the variation of the measurement, δT_{AB}, into the following canonical form (Fichtner 2011):

$$\delta T_{AB} = \int_V \int_{t=-\infty}^{\infty} \mathbf{f}^\dagger(\mathbf{x}, t) \cdot \delta\mathbf{s}(\mathbf{x}, t)\, \mathrm{d}t\, \mathrm{d}V\,, \tag{8}$$

where $\delta\mathbf{s}$ is the variation of the vectorial displacement field induced by the model perturbation $\delta\mathbf{m}$. Specifically we find (see Appendix)

$$\begin{aligned}\delta T_{AB} = \frac{1}{N}\Bigg[&\int_{t=-\infty}^{\infty} \dot{s}_j(\mathbf{x}_B, t - T_{AB})\, \delta s_i(\mathbf{x}_A, t)\, \mathrm{d}t \\ &- \int_{t=-\infty}^{\infty} \dot{s}_i(\mathbf{x}_A, T_{AB} + t)\, \delta s_j(\mathbf{x}_B, t)\, \mathrm{d}t\Bigg]\,,\end{aligned} \tag{9}$$

with

$$N = \int_{t=-\infty}^{\infty} \ddot{s}_i(\mathbf{x}_A, T_{AB} + t)\, s_j(\mathbf{x}_B, t)\, \mathrm{d}t\,. \tag{10}$$

To isolate the wavefield perturbation $\delta\mathbf{s}$ in eq. (9) we introduce delta functions and unit vectors ($\hat{\boldsymbol{\nu}}_i$ and $\hat{\boldsymbol{\nu}}_j$):

$$\begin{aligned}\delta T_{AB} = \frac{1}{N}\int_V \int_{t=-\infty}^{\infty} \big[&\dot{s}_j(\mathbf{x}_B, t - T_{AB})\, \delta(\mathbf{x} - \mathbf{x}_A)\, \hat{\boldsymbol{\nu}}_i \\ &- \dot{s}_i(\mathbf{x}_A, T_{AB} + t)\, \delta(\mathbf{x} - \mathbf{x}_B)\, \hat{\boldsymbol{\nu}}_j\big] \cdot \delta\mathbf{s}(\mathbf{x}, t)\, \mathrm{d}t\, \mathrm{d}V\,.\end{aligned} \tag{11}$$

Eq. (11) is now in the canonical form (8), with the adjoint force given by

$$\begin{aligned}\mathbf{f}^\dagger(\mathbf{x}, t) = \frac{1}{N}\big[&\dot{s}_j(\mathbf{x}_B, t - T_{AB})\, \delta(\mathbf{x} - \mathbf{x}_A)\, \hat{\boldsymbol{\nu}}_i \\ &- \dot{s}_i(\mathbf{x}_A, T_{AB} + t)\, \delta(\mathbf{x} - \mathbf{x}_B)\, \hat{\boldsymbol{\nu}}_j\big]\,.\end{aligned} \tag{12}$$

The adjoint source consists of two separate parts, each being located at one of the receivers. The time evolution of the adjoint source at location A is based on the velocity seismogram at location B, but shifted forward in time by T_{AB}. At location B, the velocity seismogram of location A is used, now shifted backward in time by T_{AB}. Note that this structure of the adjoint source is similar to the one encountered for noise cross-correlation (Tromp *et al.* 2010).

3 FORWARD AND ADJOINT MODELLING

For the simulation of forward and adjoint wavefields we used the spectral-element solver SES3D by Fichtner & Igel (2008). SES3D numerically solves the 3-D seismic wave equation for heterogeneous earth models. Other methods to calculate wavefields could have been used as well. A normal mode approach (Zhao & Chevrot 2011a,b), for instance, would have been less time consuming. However, it can not be generalized to 3-D heterogeneous media, which might be useful for iterative inversions. Also, the near-field contributions of the adjoint sources located at the receivers, which have a large effect on the sensitivity kernels computed in this paper, are omitted in a normal mode approach. In the spectral-element approach used here these terms are computed correctly. In this study, we limited ourselves to the isotropic variant of the reference earth model PREM (Dziewonski & Anderson 1981). Only vertical components of seismograms were used. Furthermore, we restricted ourselves to the analysis of S velocity kernels, because the sensitivities with respect to P velocity and density are negligibly small when Rayleigh waves are considered.

To calculate sensitivity kernels, two wavefield simulations had to be performed: First, the forward wavefield was calculated for an earthquake source, given by a moment tensor and for a limited frequency band. Ideally, tests should be done for a single frequency, because then the traveltime difference between the stations would give the exact frequency dependent phase velocity [$c(f) = \Delta x_{AB}/\Delta T_{AB}(f)$]. However, this is practically impossible, because the seismograms are calculated using a time-domain approach. Therefore, the seismograms are calculated for a certain frequency band through a bandlimited source time function. Narrowing the frequency band results in a better approximation of single-frequency (i.e. phase velocity) measurements, but also requires a longer source time function. In this study, sensitivity kernels were calculated for both a wide and narrow frequency band. The source signals for the wide frequency band were obtained by bandpass filtering (Butterworth filter, 4 poles) between cut-off periods of 5 s below and above the centre period. For the phase velocity measurements, on the other hand, we filtered a monochromatic source signal with a Gaussian filter (with a width of 0.0028 Hz) resulting in a much narrower amplitude spectrum. In the following, the terms 'group' and 'phase' velocity relate to interstation cross-correlation measurements with wide and narrow frequency bands, respectively. The group velocity kernels may therefore differ from the group velocity kernels obtained by frequency differentiation of the phase delay (e.g. Dahlen & Zhou 2006).

For the second simulation, we excited an adjoint field by the adjoint force given in eq. (12). The traveltime difference of the surface waves, T_{AB}, was obtained from the maximum of the cross-correlation between the vertical components of synthetic seismograms at both receivers, calculated in the forward simulation. Only events with strong Rayleigh wave excitation were used. This resulted in an adjoint wavefield as illustrated in Fig. 1; according to the mathematical description of the adjoint source, the adjoint wavefield is excited at both receiver locations, with a time difference of T_{AB}.

4 SENSITIVITY KERNEL GALLERY

As stated before, we calculated sensitivity kernels for both a wide and narrow frequency band, which is referred to as group and phase velocity measurements, respectively. Fig. 2 shows horizontal sections at 100 km depth through the shear velocity sensitivity kernel, for a source at an epicentral distance of 28.5° and 34.5° from the first and second receiver, respectively. The centre period was 30 s. Both group and phase velocity kernels have a cigar-shaped structure of alternating bands of positive and negative sensitivity (Fresnel zones); this structure is present at all depths. The group velocity kernel (Fig. 2 a) has a large sensitivity between the source and the first receiver, whereas it shows a gap of zero sensitivity in the interstation area. This remarkable result is exactly opposite to that expected from ray theory. The phase velocity kernel (Fig. 2b), on the other hand, shows the expected negative sensitivity between the stations and reduced sensitivity in surrounding areas. The 'ringy' structure of the kernels is due to the limited frequency bands of the wavefields: the narrower the frequency band, the more oscillatory the kernel becomes. Although the interstation sensitivity is not fully restricted to the vicinity of the great circle path, the importance of using a narrow frequency band in the two-station method is clear.

4.1 Interstation distance

We investigated the variations in sensitivity for several interstation distances. To get useful results, a minimum value of two wavelengths is required for the interstation distance, which in turn depends on the frequency content of the measurements. The effect of changing the interstation distance on the sensitivity kernels with a centre period of 30 s is shown for both group and phase velocity measurements in Figs 3 and 4, respectively. Again, the kernels based on group velocity measurements (Fig. 3) show a gap of zero sensitivity in the interstation area, which becomes larger with increasing interstation distance. This is similar to the reduced sensitivity of surface wave phase shift measurements along the source–receiver ray path (Spetzler *et al.* 2002), and the character of banana-doughnut body wave traveltime kernels measured from cross-correlations (Hung *et al.* 2000). The phase velocity kernels, on the other hand, have a fairly homogeneous sensitivity within the interstation area. The assumption that, when using the two-station method, the phase velocity estimation gives an average velocity for the entire interstation area is therefore more reliable for measurements with a narrow frequency band. However, due to the strong sensitivity outside the interstation ray path, the ray approximation can still be poor. In the following, only sensitivity kernels based on phase velocity measurements will be shown, because those are most relevant for the purpose of this study.

4.2 Frequency dependence

An important factor in seismic velocity estimation is the frequency content of the seismograms. Lower frequencies sample the Earth deeper than higher frequencies, and the use of high-frequency data can result in a higher spatial resolution. Using a centre period of 70 s, rather than 30 s, results in the sensitivity kernel shown in Fig. 5. As expected, using a lower frequency causes wider Fresnel zones in the sensitivity kernel.

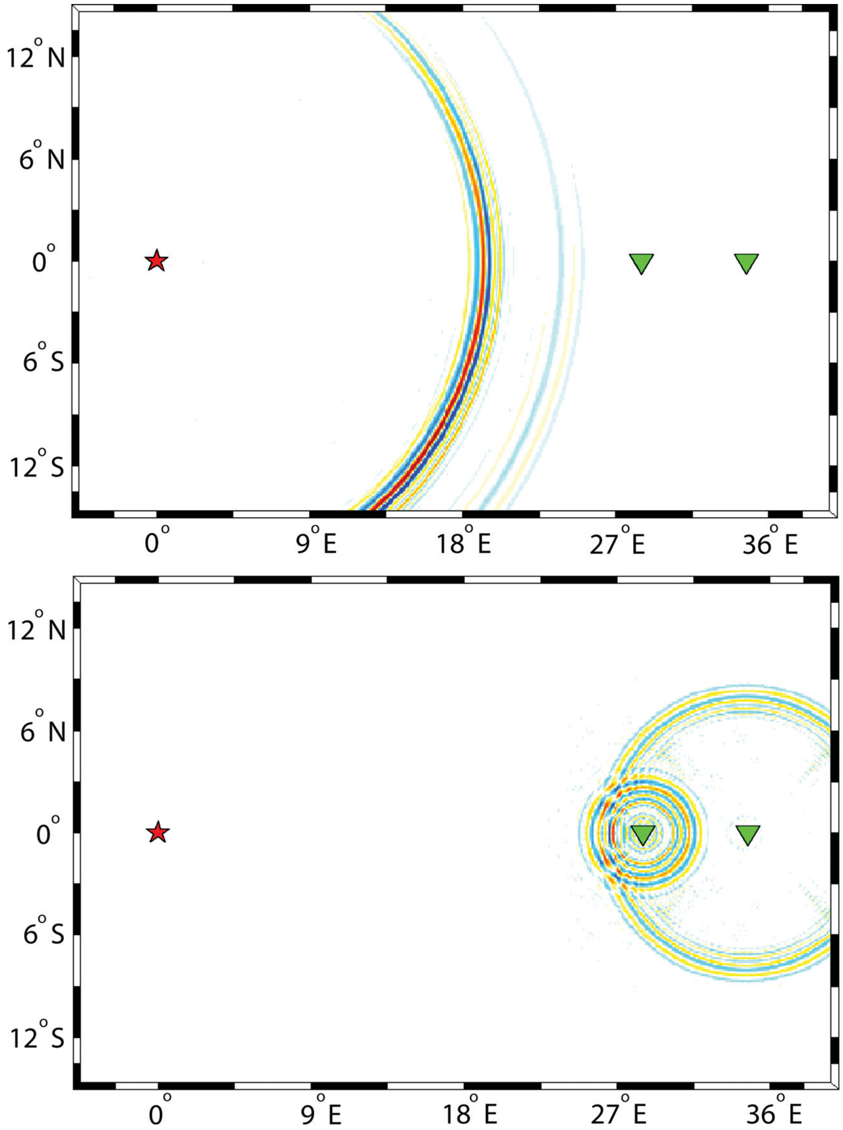

Figure 1. Illustrative example of the vertical components of the forward (top) and adjoint (bottom) velocity wavefields at 600 s after excitation of both wavefields. The large-amplitude Rayleigh waves are mostly visible. The stars represent the source, triangles represent receivers. The source is located at 28.5° and 31.5° from the first and second receiver, respectively. A period range of 25–35 s was used.

4.3 Source effects: seismic moment, focal mechanism and source location

The source mechanism and location are expected to have a significant impact on the sensitivity kernels. The seismic moment, however, does not affect the results. This is because an amplitude increase in the forward wavefield results in a decrease of equal size in the amplitude of the adjoint wavefield (see eq. 12), and the net result after interaction of both wavefields is therefore zero.

We performed tests, using several focal mechanisms, epicentral distances and focal depths. An example of a sensitivity kernel, using a centre period of 30 s, for a focal mechanism and location different from those used in the previous figures, is shown in Fig. 6. Clearly, a shorter epicentral distance results in a sensitivity kernel with a smaller spatial extent. Also, the use of a different focal mechanism results in a different radiation pattern of Rayleigh waves. This causes small-scale variations in the pattern of the sensitivity kernel. However, the overall shape and sign of the sensitivity kernels (cigar-shaped; negative sensitivity between the stations) remained the same for all configurations. The focal depth was found to have a negligible effect for relatively shallow events with strong Rayleigh wave excitation.

4.4 Number and distribution of sources

A practical difficulty in using the two-station method is finding events that are located exactly on the great circle of interest. For a plane wave propagating at an angle that differs less than 5° from the great circle, the effect on the traveltime difference between both stations is less than 0.4 per cent. Events that satisfy this condition are therefore often used in the two-station method. Furthermore, multiple events are often included to improve the signal-to-noise ratio. To test whether this increases the dominance of the interstation sensitivity we can simply add kernels for specific source locations.

Fig. 7 shows the result after summing the sensitivity kernels of three different sources. Compared to Figs 2(b) and 6, the relative interstation sensitivity is increased significantly, whereas the values close to the sources are reduced. This indicates that using multiple events partly compensates for effects from

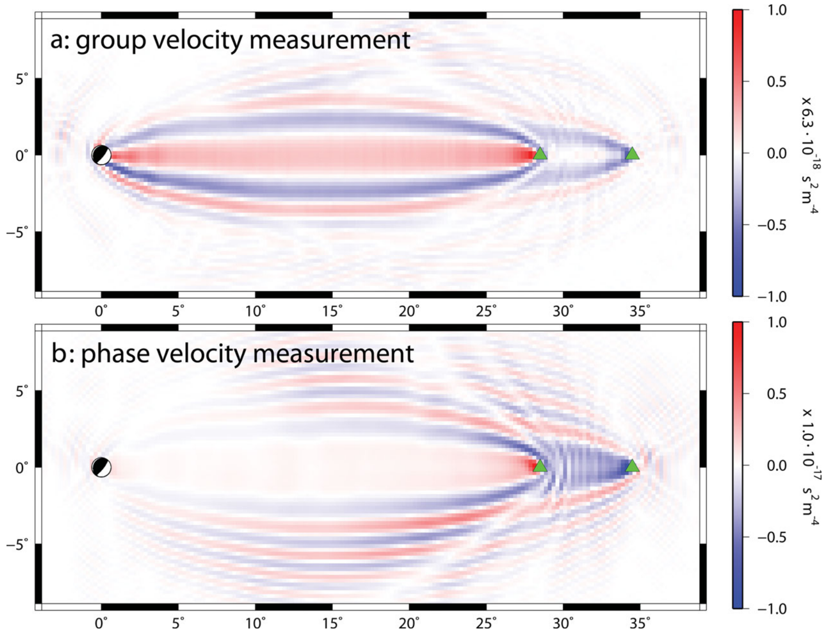

Figure 2. Horizontal sections through shear velocity sensitivity kernels at 100 km depth, for a source at 28.5° and 34.5° from the first and second receiver, respectively. (a) Group velocity kernel with a period range of 25–35 s. (b) Phase velocity kernel with a centre period of 30 s. Moment tensor components: $M_{\theta\theta} = -0.300 \cdot 10^{18}$ N m; $M_{\phi\phi} = -0.800 \times 10^{18}$ N m; $M_{rr} = 1.100 \times 10^{18}$ N m; $M_{\theta\phi} = -0.560 \times 10^{18}$ N m; $M_{\theta r} = 1.050 \times 10^{18}$ N m; $M_{\phi r} = 1.250 \times 10^{18}$ N m.

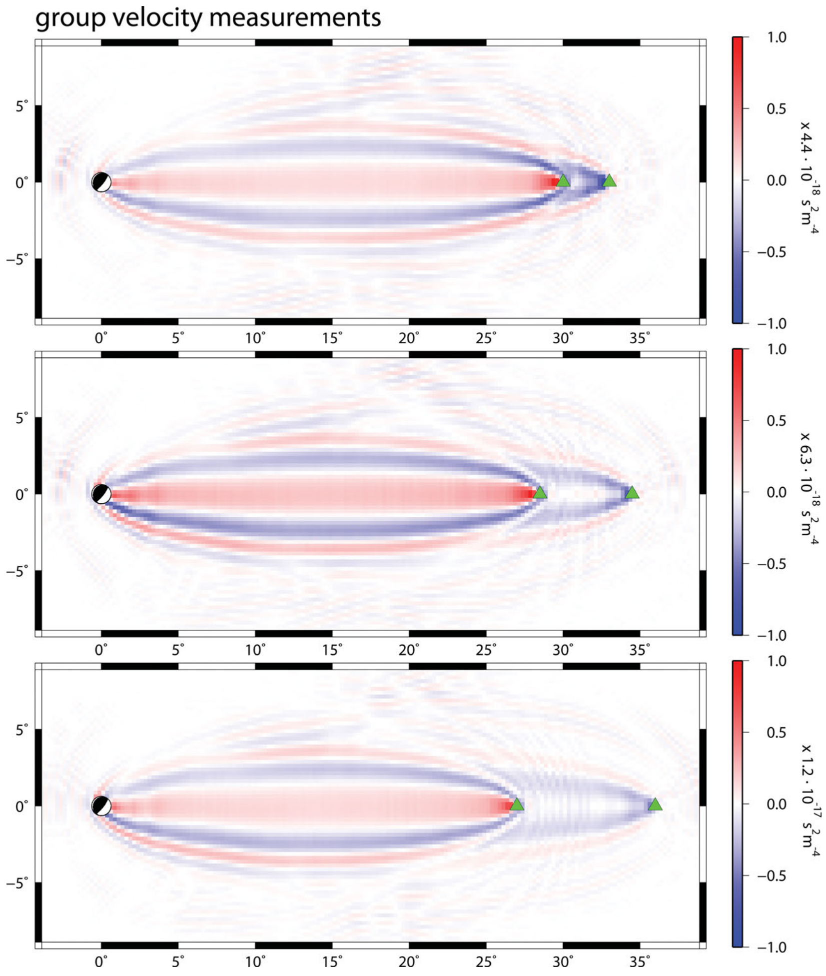

Figure 3. Horizontal sections through shear velocity sensitivity kernels at 100 km depth. The sources are located at 30° and 33° (top); 28.5° and 34.5° (middle); 27° and 36° (bottom) from the first and second receiver. The kernels are based on group velocity measurements with a period range of 25–35 s. Other parameters were similar to Fig. 2(a).

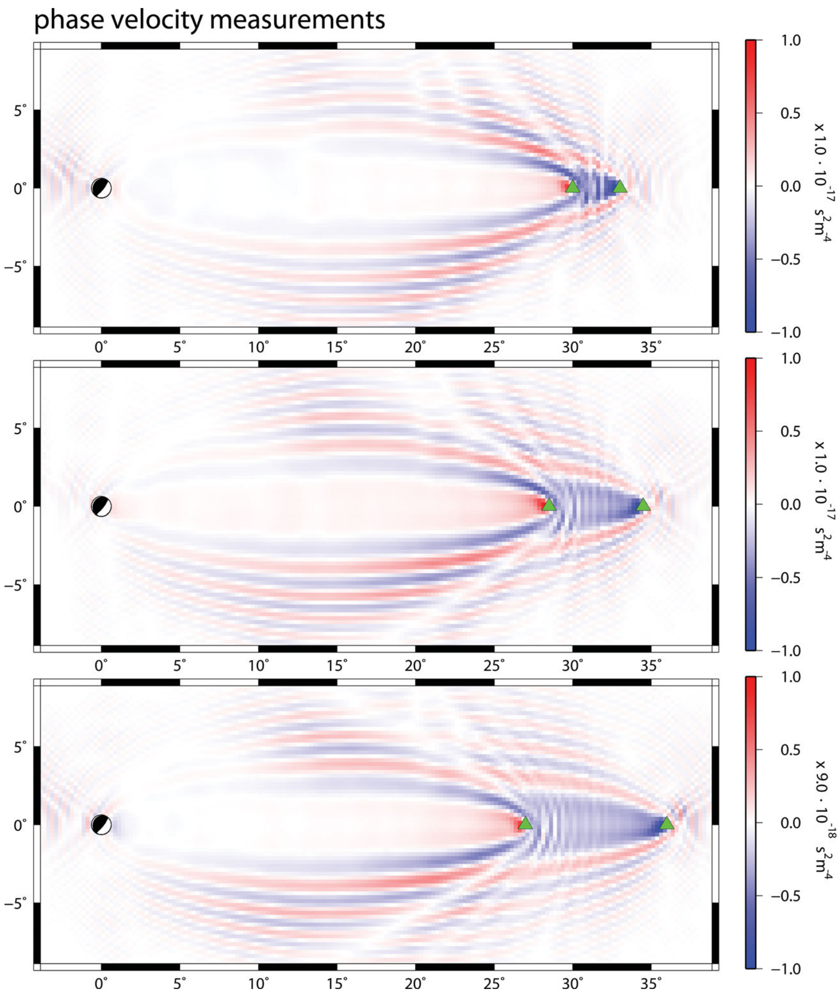

Figure 4. Horizontal sections through shear velocity sensitivity kernels at 100 km depth. The sources are located at 30° and 33° (top); 28.5° and 34.5° (middle); 27° and 36° (bottom) from the first and second receiver. The kernels are based on phase velocity measurements with a centre period of 30 s. Other parameters were similar to Fig. 2(b).

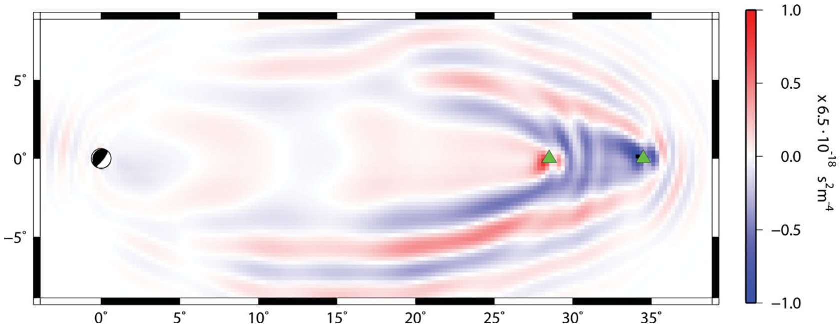

Figure 5. Horizontal section through shear velocity sensitivity kernel at 100 km depth, for a source at 28.5° and 34.5° from the first and second receiver, respectively. The kernel is based on phase velocity measurements with a centre period of 70 s. Other parameters were similar to Fig. 2(b).

non-interstation areas and that the cumulative traveltime difference for all sources is mainly affected by the interstation area. However, no matter how many sources are included, the sensitivity between the first station and the nearest source cannot be completely removed, and the value of T_{AB} therefore also contains information from outside the interstation area. A further improvement can be obtained by using events from both sides of the station pair. As shown in Fig. 8, this increases the interstation sensitivity relative to the surrounding areas. However, additional sensitivity is introduced at the other side of the station pair.

5 DISCUSSION AND CONCLUSIONS

We have calculated sensitivity kernels for surface wave two-station traveltime measurements, using the cross-correlation between seismograms at two stations. Sensitivity kernels for wide band (group velocity) measurements show a strong sensitivity between the source and the first receiver. Also, a gap of zero sensitivity along the great circle is present in the interstation area. This indicates that ‘group velocity’, measured by cross-correlation for a relatively wide frequency band around the centre frequency, does not adequately

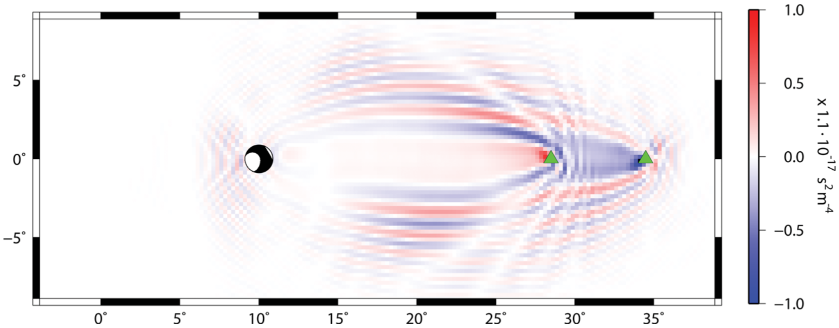

Figure 6. Horizontal section through shear velocity sensitivity kernel at 100 km depth, for a source at 18.5° and 24.5° from the first and second receiver, respectively. The kernel is based on phase velocity measurements with a centre period of 30 s. Moment tensor components: $M_{\theta\theta} = 0.710 \times 10^{19}$ N m; $M_{\phi\phi} = -0.356 \times 10^{19}$ N m; $M_{rr} = -0.355 \times 10^{19}$ N m; $M_{\theta\phi} = 0.800 \times 10^{19}$ N m; $M_{\theta r} = 0.315 \times 10^{19}$ N m; $M_{\phi r} = -1.150 \times 10^{19}$ N m.

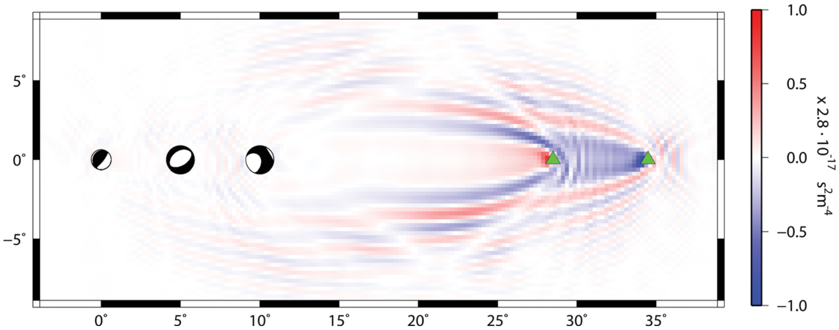

Figure 7. Horizontal section through the sum of shear velocity sensitivity kernels at 100 km depth, for three different sources. The kernel is based on phase velocity measurements with a centre period of 30 s.

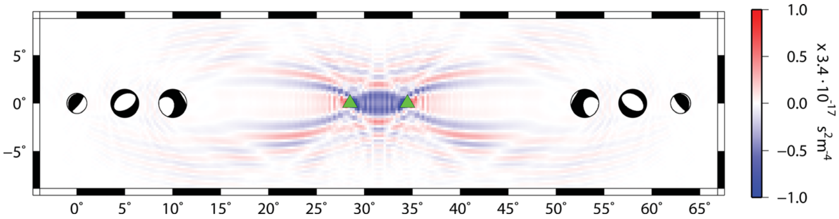

Figure 8. Horizontal section at 100 km depth through the sum of shear velocity sensitivity kernels, for six different sources, distributed on both sides of the station pair. The kernel is based on phase velocity measurements with a centre period of 30 s.

represent the average interstation group velocity along the ray path. However, when phase velocity measurements are approached by using a narrow frequency band, the interstation sensitivity is dominant and the sensitivity closer to the source is reduced. The use of phase velocity measurements is therefore more useful in the context of the two-station cross-correlation method.

Although the use of multiple sources on both sides of the station pair clearly improves the concentration of sensitivity along the interstation ray path, pronounced streaks of sensitivity far from the interstation area and off the great circle remain. Perturbations in these areas can have a large effect on the traveltime difference between the stations. For large-scale perturbations (compared to the frequency) the effects might cancel out, but this will not be the case for small-scale perturbations. This strong sensitivity to small-scale structure far from the interstation ray path can result in wrong interpretations of the measurement. A study that encountered this problem was performed by Zhang (2009), who used the two-station method for surface wave tomography in the Gulf of California. For a station pair located on the Baja-California Peninsula, with a great circle along the strike of the coast, Zhang (2009) found a discrepancy between the phase velocity curves for the two propagation directions (i.e. from station *A* to *B* and from station *B* to *A*) for frequencies above 30 mHz. For events from the northeast, the area of higher sensitivity outside the interstation area coincides with the ocean-continent transition. Because this transition is not included in the reference model, this causes large phase velocity perturbations. As a result, the phase velocity curves are affected by an area outside the interstation area.

In conclusion, using two-station phase velocity measurements in a tomographic inversion based on ray theory can be appropriate, as long as the measurements and inversion parameters are chosen with care. The importance of using as many sources as possible is clear. When choosing a station pair and events on the corresponding great circle, it is important to identify the regions of high sensitivity outside the interstation area. If, in those areas, there are indications for large anomalies compared to the reference model, it is better not to include these events.

However, to avoid misinterpretations and to improve the quality of tomographic models, correct 3-D sensitivity kernels should be used. This means that for each combination of stations and events, a sensitivity kernel has to be calculated and used subsequently in a tomographic inversion. Due to the complexity of the sensitivity kernels, a fine tomographic grid should be used (Chevrot & Zhao 2007). Although being computationally expensive, this improvement is needed to further advance tomographic inversions based on two-station measurements.

ACKNOWLEDGEMENTS

We would like to thank Jeannot Trampert, Jeroen Tromp, Sébastien Chevrot and an anonymous reviewer for useful comments and suggestions for this research. Also many thanks to Theo van Zessen, for his help using the computer cluster and to Suzanne Atkins, for proof reading this paper. Most of the figures were produced using Generic Mapping Tools (GMT). This study was financed by The Netherlands Research Centre for Integrated Solid Earth Sciences (ISES), project number 2010-57.

REFERENCES

Alvizuri, C. & Tanimoto, T., 2011. Azimuthal anisotropy from array analysis of Rayleigh waves in Southern California, *Geophys. J. Int.,* **186,** 1135–1151.

Baumont, D., Paul, A., Zandt, G., Beck, S.L. & Pedersen, H., 2002. Lithospheric structure of the central Andes based on surface wave dispersion, *J. geophys. Res.,* **107**(B12), doi:10.1029/2001JB000345.

Bozdağ, E., Trampert, J. & Tromp, J., 2011. Misfit functions for full waveform inversion based on instantaneous phase and envelope measurements, *Geophys. J. Int.,* **185,** 845–870.

Chevrot, S. & Zhao, L., 2007. Multiscale finite-frequency Rayleigh wave tomography of the Kaapvaal craton, *Geophys. J. Int.,* **169,** 201–215.

Dahlen, F. & Zhou, Y., 2006. Surface-wave group-delay and attenuation kernels, *Geophys. J. Int.,* **165,** 545–554.

Dziewonski, A. & Anderson, D., 1981. Preliminary reference earth model, *Phys. Earth planet. Inter.,* **25,** 297–356.

Endrun, B., Meier, T., Lebedev, S., Bohnhoff, M., Stavrakakis, G. & Harjes, H.-P., 2008. S velocity structure and radial anisotropy in the Aegean region from surface wave dispersion, *Geophys. J. Int.,* **174,** 593–616.

Fichtner, A., 2011. *Full Seismic Waveform Modelling and Inversion, Advances in Geophysical and Environmental Mechanics and Mathematics,* Springer, Berlin, Heidelberg.

Fichtner, A., Bunge, H.-P. & Igel, H., 2006. The adjoint method in seismology, *Phys. Earth planet. Inter.,* **157**(1–2), 86–104.

Fichtner, A. & Igel, H., 2008. Efficient numerical surface wave propagation through the optimization of discrete crustal models - a technique based on non-linear dispersion curve matching (DCM), *Geophys. J. Int.,* **173,** 519–533.

Fichtner, A., Kennett, B.L.N., Igel, H. & Bunge, H.-P., 2008. Theoretical background for continental- and global-scale full-waveform inversion in the time-frequency domain, *Geophys. J. Int.,* **175,** 665–685.

Hung, S.-H., Dahlen, F. & Nolet, G., 2000. Fréchet kernels for finite-frequency traveltimes—II. Examples, *Geophys. J. Int.,* **144,** 175–203.

Kuo, B.-Y., Chi, W.-C., Lin, C.-R., Chang, E.T.-Y., Collins, J. & Liu, C.-S., 2009. Two-station measurement of Rayleigh-wave phase velocities for the Huatung basin, the westernmost Philippine Sea, with OBS: implications for regional tectonics, *Geophys. J. Int.,* **179,** 1859–1869.

Liu, Q. & Tromp, J., 2006. Finite-frequency kernels based on adjoint methods, *Bull. seism. Soc. Am.,* **96**(6), 2383–2397.

Luo, Y. & Schuster, G., 1991. Wave equation traveltime inversion, *Geophysics,* **56**(5), 645–653.

Pedersen, H., 2006. Impacts of non-plane waves on two-station measurements of phase velocities, *Geophys. J. Int.,* **165,** 279–287.

Sato, Y., 1955. Analysis of dispersed surface waves by means of Fourier Transform I, *Bull. Earthquake Res. Tokyo Univ.,* **33,** 33–47.

Spetzler, J., Trampert, J. & Snieder, R., 2002. The effect of scattering in surface wave tomography, *Geophys. J. Int.,* **149,** 755–767.

Tarantola, A., 1988. Theoretical background for the inversion of seismic waveforms, including elasticity and attenuation, *Pure appl. Geophys.,* **128,** 365–399.

Tromp, J., Luo, Y., Hanasoge, S. & Peter, D., 2010. Noise cross-correlation sensitivity kernels, *Geophys. J. Int.,* **183,** 791–819.

Tromp, J., Tape, C. & Liu, Q., 2005. Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels, *Geophys. J. Int.,* **160,** 195–216.

Yao, H., van der Hilst, R. & de Hoop, M., 2006. Surface-wave array tomography in SE Tibet from ambient seismic noise and two-station analysis—I. Phase velocity maps, *Geophys. J. Int.,* **166,** 732–744.

Zhang, X., 2009. The upper mantle beneath the Gulf of California from surface wave dispersion, *PhD thesis,* Utrecht University, Utrecht, The Netherlands.

Zhang, X., Paulssen, H., Lebedev, S. & Meier, T., 2007. Surface wave tomography of the Gulf of California, *Geophys. Res. Lett.,* **34,** doi:10.1029/2007GL030631.

Zhao, L. & Chevrot, S., 2011a. An efficient and flexible approach to the calculation of three-dimensional full-wave Fréchet kernels for seismic tomography—I. Theory, *Geophys. J. Int.,* **185,** 922–938.

Zhao, L. & Chevrot, S., 2011b. An efficient and flexible approach to the calculation of three-dimensional full-wave Fréchet kernels for seismic tomography—II. Numerical results, *Geophys. J. Int.,* **185,** 939–954.

APPENDIX A

Here, we derive the variation in the misfit due to a change in the seismogram, as a consequence of a perturbation of the model. For a single model parameter m (of $\mathbf{m}$), eq. (2) can be rewritten as

$$\dot{C}_{ij}(T_{AB}) = f(T_{AB}(m), m) = 0\,. \tag{A1}$$

Therefore,

$$\frac{\partial f[T_{AB}(m), m]}{\partial m} = 0\,, \tag{A2}$$

or

$$\frac{\partial f}{\partial T_{AB}}\frac{\partial T_{AB}}{\partial m} + \frac{\partial f}{\partial m} = 0\,, \tag{A3}$$

which results in

$$\frac{\partial T_{AB}}{\partial m} = -\frac{\frac{\partial f}{\partial m}}{\frac{\partial f}{\partial T_{AB}}}\,. \tag{A4}$$

Or, with all terms written out explicitly and replacing derivatives by perturbations (δ), this can be written as

$$\delta T_{AB} = -\frac{1}{N}\left[\int_{\tau=-\infty}^{\infty} \delta\dot{s}_i(\mathbf{x}_A, T_{AB}+\tau)s_j(\mathbf{x}_B, \tau)\,\mathrm{d}\tau + \int_{\tau=-\infty}^{\infty} \dot{s}_i(\mathbf{x}_A, T_{AB}+\tau)\delta s_j(\mathbf{x}_B, \tau)\,\mathrm{d}\tau\right], \tag{A5}$$

with

$$N = \int_{\tau=-\infty}^{\infty} \ddot{s}_i(\mathbf{x}_A, T_{AB}+\tau)s_j(\mathbf{x}_B, \tau)\,\mathrm{d}\tau\,. \tag{A6}$$

Partial integration of the first term of eq. (A5), causality and substitution of $t = T_{AB} + \tau$ in the first, and $t = \tau$ in the second term and N lead to

$$\delta T_{AB} = \frac{1}{N}\left[\int_{t=-\infty}^{\infty} \dot{s}_j(\mathbf{x}_B, t - T_{AB})\delta s_i(\mathbf{x}_A, t)\,\mathrm{d}t - \int_{t=-\infty}^{\infty} \dot{s}_i(\mathbf{x}_A, T_{AB}+t)\delta s_j(\mathbf{x}_B, t)\,\mathrm{d}t\right], \tag{A7}$$

with

$$N = \int_{t=-\infty}^{\infty} \ddot{s}_i(\mathbf{x}_A, T_{AB}+t)s_j(\mathbf{x}_B, t)\,\mathrm{d}t\,. \tag{A8}$$

This equals eq. (9), which we needed to prove.