Geophys. J. Int. (1996) **127**, 363–378

Ensemble inference in terms of empirical orthogonal functions

H. Douma, R. Snieder and A. Lomax

Department of Theoretical Geophysics, University of Utrecht, PO Box 80.021, 3508 *TA Utrecht, The Netherlands*

Accepted 1996 June 28. Received 1996 June 3; in original form 1996 January 29

SUMMARY

Many geophysical problems involve inverting data in order to obtain meaningful descriptions of the Earth's interior. One of the basic characteristics of these inverse problems is their non-uniqueness. Since computation power has increased enormously in the last few years, it has become possible to deal with this non-uniqueness by generating and selecting a number of models that all fit the data up to a certain tolerance. In this way a solution space with acceptable models is created. The remaining task is then to infer the common robust properties of all the models in the ensemble. In this paper these properties are determined using empirical orthogonal function (EOF) analysis. This analysis provides a method to search for subspaces in the solution space (ensemble) that correspond to the patterns of minimum variability. In order to show the effectiveness of this method, two synthetic tests are presented. To verify the applicability of the analysis in geophysical inverse problems, the method is applied to an ensemble generated by a Monte Carlo search technique which inverts group-velocity dispersion data produced by using vertical-component, long-period synthetic seismograms of the fundamental Rayleigh mode. The result shows that EOF analysis successfully determines the well-constrained parts of the models and in effect reduces the variability present in the original ensemble while still recovering the earth model used to generate the synthetic seismograms. Finally, an application of the method to examine the contrast in upper-mantle *S*-wave velocity across the Tornquist–Tesseyre Zone is presented, indicating a significant change in *S*-wave velocity in the upper mantle beneath this zone bordering the East European Platform and Tectonic Europe, and a significantly thicker crust beneath the East European Platform.

Key words: numerical techniques, Tornquist–Tesseyre Zone.

INTRODUCTION

Many geophysical problems involve inverting data in order to obtain meaningful descriptions of the Earth's interior. It is well known that these inverse problems are characterized by their non-uniqueness. In order to deal with this non-uniqueness, forward modelling has become popular in the last few years; one can often create an ensemble of models that all 'fit the data' to a certain tolerance by solving the forward problem repeatedly. The presentation of one model for the earth's interior has been replaced by a search for the robust features shared by all the models in an ensemble of acceptable models.

The conventional approach used to determine one model involves a minimization of a misfit function in a least-squares sense. This method requires a good starting model and therefore needs *a priori* information. Since this information is often not available, there is a risk of converging to a local minimum instead of the global minimum of the objective function.

To perform the forward modelling and to create an ensemble of acceptable models, a number of techniques can be used. Among these methods are genetic algorithms (Goldberg 1989; Nolte & Frazer 1994; Sambridge & Drijkoningen 1992 and Sen & Stoffa 1992), simulated annealing (e.g. Rothman 1985 and Basu & Frazer 1990) and Monte Carlo search techniques (e.g. Mosegaard & Tarantola 1995). In this study, we are interested in the reliable information encompassed by the resulting ensemble of models obtained from such a forward search. In fact we are searching for the joint features shared by the ensemble of models, since they are expected to be the most reliable features. Empirical orthogonal function (EOF) analysis provides a method for investigating the patterns of variability within a set of models and thus can find the basis function that corresponds to the pattern of *smallest* variability within the ensemble. Normally, EOF analysis is used to determine the patterns of *largest* variability. For example, meteorologists are interested in the spatial and temporal variability of physical fields such as the sea surface temperature or an atmospheric pressure field, in order to be able to diagnose

or predict these fields in the future (e.g. Trenberth & Shin 1984; Rinne, Karhila & Järvenoja 1981). Often this technique is also used to decrease the number of parameters needed to describe an ensemble (e.g. Rao 1964). By only using the patterns with the largest variability this number can be effectively reduced, while leaving the total explained variance in the ensemble relatively high.

In this study, empirical orthogonal functions are used to determine the common features of a population of models that fit the data with a prescribed tolerance. By defining a scatter probe which represents the total variability in the ensemble in a certain direction in model space, and searching the ensemble for the directions where this probe is minimized, we find a subspace of model space that corresponds to the pattern of smallest variability that will represent the similarities present in the ensemble—a subspace that encompasses the shared properties of the models. An outline of this method was hinted at by Kennett & Nolet (1978). Although their description differs in some sense from ours, especially with respect to the estimated resolution, the main characteristics of the methods are the same. Vasco, Johnson & Majer (1993) used the method to estimate the dimensionality of the subspace of acceptable models.

Lomax & Snieder (1995a) (hereafter referred to by LS) used a genetic algorithm to study the contrast in upper-mantle shear-wave velocity between the stable Precambrian continental crust in northeast Europe and the tectonically active continental and oceanic regions of central and western Europe and the western Mediterranean, by inversion of fundamental-mode Rayleigh-wave group dispersion curves. The acceptable models obtained in their inversion show regions of large scatter indicating parts of the model space that are not well constrained. Some of these poorly constrained regions can be related to trade-off relationships while other regions are poorly constrained due to lack of data. This means that there are patterns of large variability present in the model space. This study was set up to identify and remove the trade-off relationships and reduce the variability in the regions that are poorly constrained.

THEORY

This section follows the main EOF theory outlined in Preisendorfer (1988, pp. 25–44). Let us consider an ensemble of real-valued functions (or models) $m'(\omega, z)(\omega = 1, \ldots, n, z \in \mathbf{R})$ on the continuous domain $\mathbf{R}$, where ω denotes the model number within the ensemble and z is a spatial coordinate (e.g. depth). These functions can be centred around the mean model $\bar{m}(z) = (1/n)\sum_{\omega=1}^{n} m'(\omega, z)$ (see Fig. 1). The result is a centred ensemble $m(\omega, z) = m'(\omega, z) - \bar{m}(z)$. The *scatter* (or *variability*) Ψ along a vector probe $\mathbf{e}$ in an infinite-dimensional Euclidian space $\mathbf{E}_\infty$ is defined as the sum of the squared projections of all the models onto $\mathbf{e}$:

$$\Psi(\mathbf{e}) = \sum_{i=1}^{n} [\mathbf{m}(\omega), \mathbf{e}]^2 = [\mathbf{e}, [\mathbf{S}, \mathbf{e}]], \tag{1}$$

where $\mathbf{S}$ is the (symmetric) scatter matrix [in fact $(n-1)$ times the covariance matrix]

$$S(z, z') = \sum_{\omega=1}^{n} m(\omega, z) m(\omega, z') \tag{2}$$

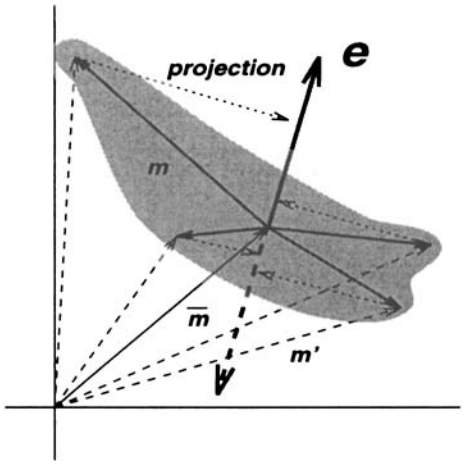

Figure 1. 2-D visualization of EOF analysis. The uncentred ensemble of models m' are centred around the mean model $\bar{m}$, resulting in the centred ensemble of models m. This ensemble is projected onto some vector probe $\mathbf{e}$, which results in the scatter probe $\Psi(\mathbf{e})$.

and the inner product $[f, g]$ is defined as

$$[f, g] \equiv \int_{\mathbf{R}} w(z) f(z) g(z)\, dz, \tag{3}$$

where $w(z)$ is a positive weighting function specific to the problem.

Since we are interested in the patterns in the ensemble corresponding to the smallest variability, we need to determine the local minima of the scatter $\Psi(\mathbf{e})$. Therefore $\mathbf{e}$ needs to be varied in $\mathbf{E}_\infty$ until Ψ reaches these extrema.

Before doing so, we reformulate the problem in the finite-dimensional context of EOF analysis (Preisendorfer 1988, pp. 75–81). Therefore we consider the problem in the subspace spanned by a p-dimensional orthonormal [with respect to the inner product defined in eq. (3)] family of p basis functions $\phi_i(z)$ on $\mathbf{R}$. For a function in this subspace we can expand

$$e_i(z) = \sum_{j=1}^{p} c_j^i \phi_j(z), \tag{4}$$

with the coefficients c_j^i still to be determined. Eq. (1) then becomes

$$\Psi(\mathbf{e}_i) = \sum_{j=1}^{p}\sum_{k=1}^{p} c_j^i \left[\int_{\mathbf{R}}\int_{\mathbf{R}} w(z)\phi_j(z) S(z, z')\phi_k(z') w(z')\, dz\, dz'\right] c_k^i \tag{5}$$

If we now define a $p \times p$ (symmetric) matrix $\mathbf{U}$ as

$$U_{ij} = \int_{\mathbf{R}}\int_{\mathbf{R}} w(z)\phi_i(z) S(z, z')\phi_j(z') w(z')\, dz\, dz', \tag{6}$$

we can write $\Psi(\mathbf{e}_i)$ in the simple form

$$\Psi(\mathbf{e}_i) \equiv \Psi'(\mathbf{c}^i) = \sum_{j=1}^{p}\sum_{k=1}^{p} c_j^i U_{jk} c_k^i. \tag{7}$$

It should be pointed out that, if the $\mathbf{e}_i$ are orthonormal, the vectors $\mathbf{c}^i$, defined by expanding the $\mathbf{e}_i$ with relation (4), are orthonormal with respect to the inner product:

$$[\mathbf{c}^i, \mathbf{c}^j] = \sum_{k=1}^{p} c_k^i c_k^j = \delta_{ij} \tag{8}$$

(see Preisendorfer 1988, p. 76).

Searching for the stationary values of $\Psi'(\mathbf{c}^i)$ under the constraint that the $\mathbf{c}^i$ are orthonormal leads to an eigenvalue problem of the form

$$\sum_{k=1}^{p} U_{jk} c_k^i = \lambda_i c_j^i. \tag{9}$$

Thus the problem of minimizing the scatter probe Ψ and determining the coefficients $\mathbf{c}^i$ is replaced by a simple eigenvalue problem. After eq. (9) has been solved, eq. (4) can be used to determine the functions $\mathbf{e}_i$. These functions are the *empirical orthogonal functions* (in the following they will be called EOFs). It can easily be shown that the (non-negative) eigenvalue λ_i represents the variability of the ensemble $m(\omega, z)$ in the direction $\mathbf{e}_i$ in the subspace of E_∞ spanned by the orthonormal set $\phi_i(z)$. Therefore the EOFs with the small eigenvalues span the subspace of patterns with small variability within the ensemble. This subspace encompasses the joint features shared by the models in the ensemble.

Together, the EOFs form a complete basis for the original ensemble of models $m(\omega, z)$ within the subspace spanned by the $\phi_i(z)$. Thus this ensemble can be projected onto the determined EOFs.

$$m(\omega, z) = \sum_{i=1}^{p} a_i(\omega) e_i(z) \quad \omega = 1, \ldots, n, \quad z \in \mathbf{R}, \tag{10}$$

with

$$a_i(\omega) \equiv [\mathbf{m}(\omega), \mathbf{e}_i] = \int_{\mathbf{R}} w(z) m(\omega, z) e_i(z)\, dz. \tag{11}$$

However, we only want to use the EOFs with small eigenvalues since they encompass the common features within the ensemble. We therefore have to define a cut-off point p' in the eigenvalue spectrum such that we can create a new ensemble or population, $v^{p'}(\omega, z)$, which consists of the projection of the original population on the subset of the EOFs that correspond to the patterns of smallest variability within the ensemble:

$$v^{p'}(\omega, z) = \bar{m}(z) + \sum_{i=1}^{p'} a_i(\omega) e_i(z) \quad \omega = 1, \ldots, n, \quad z \in \mathbf{R}. \tag{12}$$

This new population will be called the filtered ensemble.

If we use $p' = p$ EOFs to project, the filtered ensemble will be identical to the original ensemble since the EOFs form a complete set of basis functions (within the finite-dimensional subspace). If, however, we use the first $p' < p$ EOFs in order to reduce the variability and to reveal the common robust properties in the ensemble, the basis functions for the filtered ensemble will no longer be complete. This results in a smearing-out effect of the information and therefore in a loss of resolution. To quantify this, we can project a delta function onto the p' EOFs that are accounted for in the filtered ensemble. The deviation of this filtered delta function from the true delta function is a measure of the resolution that can be obtained. Therefore we define a resolution kernel as

$$R^{p'}(z, z') = \sum_{j=1}^{p'} [\mathbf{D}(z'), \mathbf{e}_j] e_j(z), \tag{13}$$

with

$$D(z, z') = \delta(z - z'). \tag{14}$$

The inner product is defined in eq. (3) and $D(z, z')$ are the coefficients of the vector $\mathbf{D}(z')$.

SYNTHETIC TESTS

In order to get a feeling of how EOF analysis works in ensemble inference and how the results should be interpreted, we first present a synthetic test. The test is set up to analyse an ensemble of models that satisfies only one constraint, viz:

$$\int_0^1 m'(x)\, dx = 1. \tag{15}$$

The ϕ_i from eq. (4) are defined as block functions:

$$\phi_i(x) = \begin{cases} 1 & \text{if } (i-1)\Delta x \le x < i\Delta x, \\ 0 & \text{otherwise}, \end{cases} \tag{16}$$

with $\Delta x = 0.1$ and $i = 1, \ldots, 10$. If we set the weighting function in eq. (3) to $w(x) = 1/\Delta x$, the ϕ_i are orthonormal with respect to the inner product defined in eq. (3). A Monte Carlo search provides an initial set of models with the search range for each $m'(x_i)$ set from -1.0 to 3.0 (symmetrically around 1.0). In order to be able to judge whether the models fit the data, a misfit function is defined as

$$\delta = \left[1 - \int_0^1 f(x)\, dx\right]^2. \tag{17}$$

Only models with a misfit $\delta \le 0.02$ with the data are accepted, resulting in an ensemble $m'(\omega, x)$ with $\omega = 1, \ldots, 430$.

Fig. 2 shows the calculated EOFs, and Table 1 the eigenvalues of the matrix $\mathbf{U}$ and the integrals of the EOFs over $[0, 1]$. It can clearly be seen that $\mathbf{e}_1$, the eigenvector with the smallest eigenvalue, is almost exactly equal to unity over the interval $[0, 1]$, while the other eigenvectors show a strongly oscillating behaviour; their integrals over $[0, 1]$ are almost equal to zero (Table 1). Since the ensemble of models was designed to have only one joint feature, i.e. the integral of the function over the interval $[0, 1]$, we expect the only reasonable information we can extract from this ensemble to be a constant (the average). Because the smallest eigenvector $\mathbf{e}_1$ is almost

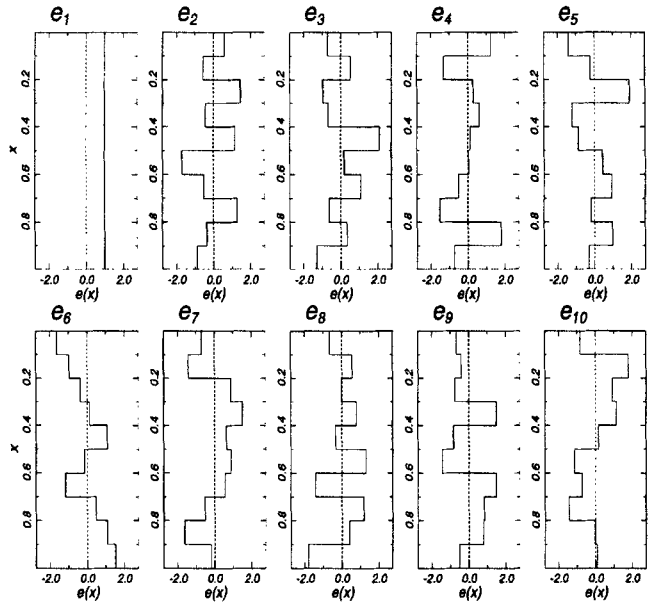

Figure 2. The EOFs for the synthetic example with only one constraint. The eigenvalues increase monotonically from $\mathbf{e}_1$ to $\mathbf{e}_{10}$.

Table 1. Eigenvalues and integrals of the EOFs calculated for an ensemble satisfying only one constraint.

i	*eigenvalue* λ_i	$\int_0^1 e_i(x)dx$
1	0.044	1.000
2	50.082	0.000
3	51.635	-0.003
4	55.270	0.001
5	58.014	0.001
6	60.703	-0.001
7	62.043	0.001
8	69.162	0.001
9	72.150	-0.002
10	75.175	-0.003

constant (1.0) it can be stated that the EOF with the smallest eigenvalue indeed represents the common robust feature present in the ensemble.

As an explanation of why the first EOF corresponds to the single data kernel of the problem, consider a 2-D example of an ensemble that satisfies the constraint that the sum of the parameters almost equals unity. In this 2-D case, the models will sample a narrow subspace around the line $y = 1 - x$ (see Fig. 3). If the ensemble contains very little noise (in other words the accepted misfit level is very low), the direction of minimum variability ($\mathbf{e}_1$) will be close to the direction $y = x$, meaning that the main information we can extract from this ensemble is indeed almost a constant.

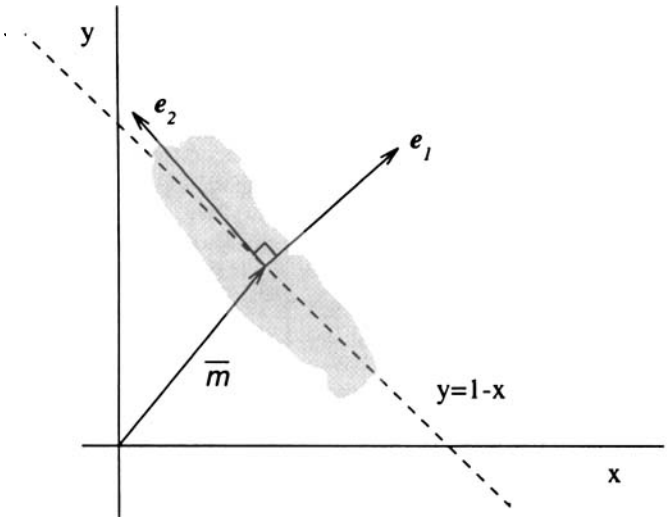

Figure 3. 2-D visualization of EOF analysis for an ensemble satisfying the constraint that the sum of two parameters equals unity.

In order to calculate the subspace spanned by $\mathbf{e}_1$ we can perform a projection of the ensemble onto this EOF, i.e. we can calculate the filtered ensemble $v^{p'}(x)$ from eq. (12) for $p' = 1$. Fig. 4 shows the filtered ensembles for $p' = 1, \ldots, 4$. We can clearly see that, by using only $\mathbf{e}_1$ as a basis, we do indeed recover the average which is the robust information put into the ensemble by the one constraint (eq. 15). Using more than this one EOF to calculate $v^{p'}(x)$ results in an enormous increase of the variability.

In general, an ensemble satisfies more than one constraint. In this case the joint features of the models will form a multi-dimensional subspace in model space. In order to verify whether EOF analysis is able to discover such a multi-dimensional subspace of well-constrained model features successfully, we present a synthetic test with two constraints instead of one. The constraints are set to

$$\int_{-1}^{1} m'(x)\,dx = 0, \tag{18}$$

$$\int_{-1}^{1} xm'(x)\,dx = 1. \tag{19}$$

The ϕ_i from eq. (4) are again defined as block functions:

$$\phi_i(x) = \begin{cases} 1 & \text{if } (i-1)\Delta x - 1 \le x < i\Delta x - 1, \\ 0 & \text{otherwise}, \end{cases} \tag{20}$$

with $\Delta x = 0.2$ and $i = 1, \ldots, 10$. The weighting function in eq. (3) is again $w(x) = 1/\Delta x$, resulting in an orthonormal set ϕ_i. A Monte Carlo search generates the initial model space with the search range for each $m'(x_i)$ set from -4.0 to 4.0, symmetrically around zero. The misfit function is defined as

$$\delta = \frac{1}{2}\left\{\left[1 - \int_{-1}^{1} xm'(x)\,dx\right]^2 + \left[\int_{-1}^{1} m'(x)\,dx\right]^2\right\}. \tag{21}$$

Models with a misfit $\delta \le 0.05$ are accepted, leaving an ensemble $m'(\omega, x)$ with $\omega = 1, \ldots, 170$ (see Fig. 5).

Since eqs (18) and (19) define the projections of the models onto the kernels $k_1(x) = 1$ and $k_2(x) = x$, these constraints restrict the projection of the model space onto the subspace of linear functions $f(x) = a + bx$. Therefore we expect the EOF analysis to produce eigenfunctions $\mathbf{e}_1$ and $\mathbf{e}_2$, corresponding to the smallest eigenvalues, which span this linear subspace.

Fig. 6 shows the resulting EOFs. As expected, the two eigenfunctions $\mathbf{e}_1$ and $\mathbf{e}_2$ corresponding to the smallest eigenvalues are indeed almost perfect linear functions, which is not the case for the rest of the EOFs. If we consider the eigenvalues in Table 2 we see that there is a clear cut-off in the eigenvalue spectrum between $\mathbf{e}_1$ and $\mathbf{e}_2$ on one side, and the other EOFs on the other, meaning that the variability for the linear functions is significantly smaller than for the other functions in the ensemble. Note that $\mathbf{e}_3$ to $\mathbf{e}_{10}$ have a strongly oscillating character. For this example, EOFs $\mathbf{e}_1$ and $\mathbf{e}_2$ do indeed represent the shared properties in the ensemble—the projection of the models on the subspace of linear functions $f(x) = a + bx$.

The data kernels of the two constraints in eqs (18) and (19) are $k_1(x) = 1$ and $k_2(x) = x$. Of course the subspace determined by the two constraints is spanned by these two data kernels. Since there is no *a priori* reason to expect EOF analysis to

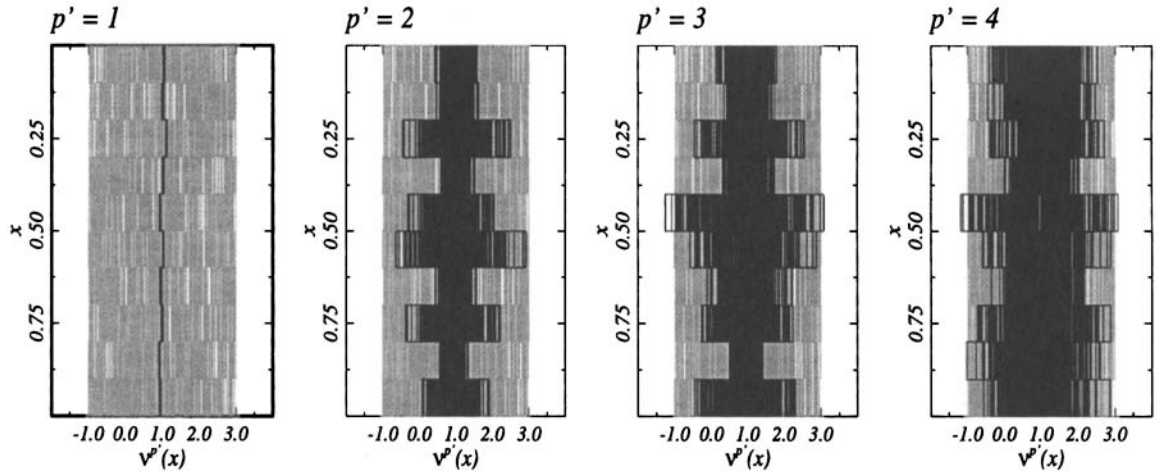

Figure 4. The filtered ensembles (black) with $p' = 1$–4 compared with the original ensemble (grey) for the synthetic example with only one constraint.

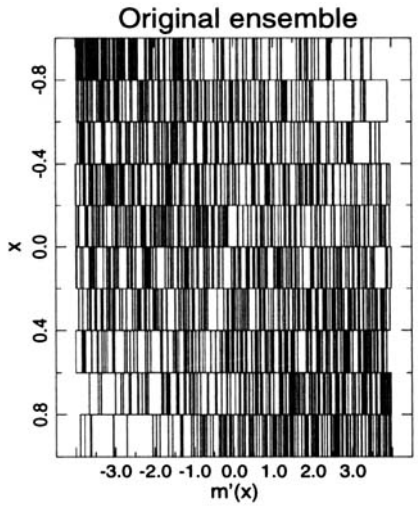

Figure 5. Ensemble resulting from the Monte Carlo search with the two constraints (18) and (19) and a misfit level up to 0.05.

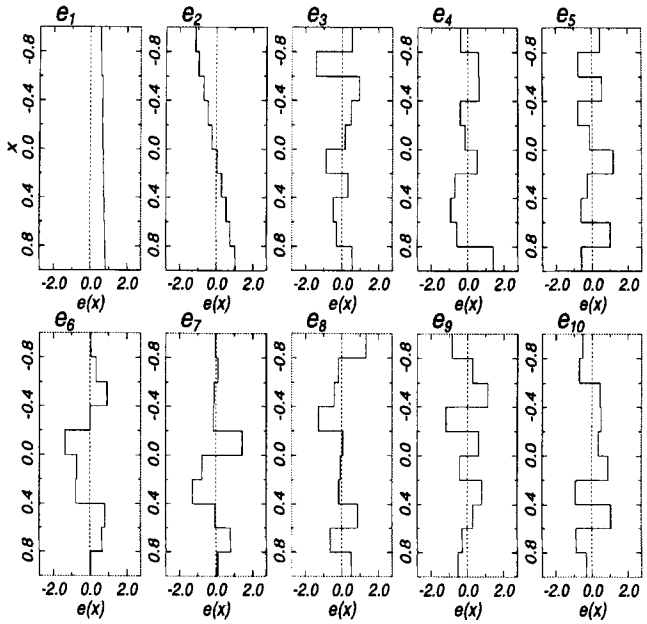

Figure 6. The EOFs for the synthetic run with two constraints. The eigenvalues increase monotonically from $\mathbf{e}_1$ to $\mathbf{e}_{10}$.

produce two eigenfunctions corresponding to the smallest eigenvalues equal to the data kernels $k_1(x)$ and $k_2(x)$, we expect the analysis to find these eigenfunctions to be a linear combination of the data kernels.

Given both constraints we can calculate the solution of the form $f(x) = a + bx$ that satisfies the constraints (18) and (19): $f(x) = 3x/2$. In order to see if this function is embedded in the subspace spanned by $\mathbf{e}_1$ and $\mathbf{e}_2$ we can calculate the filtered ensemble $v^{p'}(x)$ from eq. (12). Fig. 7 shows the filtered ensembles $v^{p'}(x)$ with $\omega = 1, \ldots, 170$ and the cut-off point p' increasing up to 5. We can see clearly that if we project only on the first two eigenfunctions ($p' = 2$), the function $f(x) = 3x/2$ is well embedded in the filtered ensemble, which means that the information that was put into the ensemble by the two constraints is recovered by the analysis if we only use the two EOFs with the smallest eigenvalues. If we use more than these two eigenfunctions, the variability increases enormously. The eigenvalue spectrum shown in Table 2 is a valuable tool for deciding how many EOFs to use for the filtered ensemble.

INVERSION OF GROUP-VELOCITY DISPERSION FOR THE *S* VELOCITY

LS performed a group-velocity inversion with noisy synthetic data using a genetic algorithm. The data consisted of vertical-component, long-period seismograms of the fundamental Rayleigh mode, generated by a shallow double-couple source recorded at a distance of 30° in the IASP91 model (Kennett & Engdahl 1991). Multiple-filter analysis (Dziewonski, Bloch & Landisman 1969) provided the group-velocity dispersion estimates (realistic noise was added to these estimates) that could be inverted by the genetic algorithm. A definition of the

Table 2. Eigenvalues of the EOFs calculated for an ensemble satisfying two constraints.

i	*eigenvalue* λ_i
1	1.383
2	4.992
3	133.435
4	151.516
5	164.625
6	181.53
7	193.639
8	211.243
9	253.834
10	271.901

misfit function and the ensuing definition of acceptable models (LS) led to an ensemble of *S*-velocity models, in which depths where the velocity was poorly constrained showed up as depths where the variability in the ensemble was large. In this section, we apply the same method to calculate the synthetic data, but the ensemble is created using a Monte Carlo search technique. The EOFs are then used to remove the poorly constrained features and reveal the well-constrained robust features present in the ensemble.

Model parametrization

In order to create earth models for the shear-wave velocity, we need to specify a parametrization of depth. LS used a nodal parametrization with four 'crustal' and 14 'mantle' nodes. The bottom 'crustal' and the top 'mantle' nodes were located at the same depth, allowing a step discontinuity between the crust and the mantle in order to model the Moho discontinuity. The crustal thickness was variable between 15 and 70 km. Therefore in the study of LS the resulting models prescribed the shear-wave velocity β as a function of the *node number,* while other parameters specified the node depth.

EOF analysis uses a centring around the average to move the origin into the ensemble. A problem arises when we consider the models generated with the nodal parametrization as earth models showing the shear-wave velocity β as a function of *depth.* Calculating the mean earth model $\bar{\beta}(z)$ and centring the model space around this mean would then involve comparing velocities at different depths, since every model contains

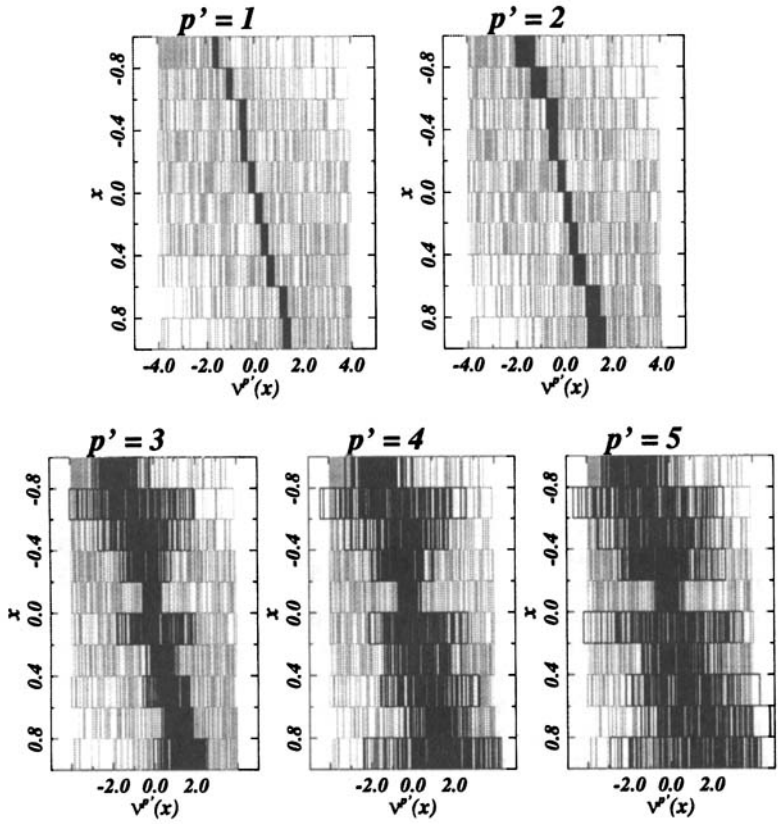

Figure 7. Filtered ensembles with $p' = 1$–5 for the synthetic test with two constraints.

a different depth parametrization. In fact, the nodal parametrization with a variable crustal thickness prevents the calculation of the scatter matrix **S**. Therefore we created a new ensemble with a different depth parametrization. To create the ensemble we applied the above method that LS used to generate their ensemble, except that we used a Monte Carlo search technique and a fixed-depth parametrization. Table 3 shows the fixed-depth model. The thickness of the layers increases with depth since the resolution of the fundamental-mode Rayleigh waves decreases with the penetration depth.

Inversion using the Monte Carlo search technique

LS used smoothing in their inversion to suppress node-to-node oscillation in the solution. Smoothing is a widely known method used to prevent the solution from having too many artificial artefacts. This study focuses on identifying well-constrained information and reducing poorly constrained information. The introduction of smoothing in the inversion results in a loss of unconstrained information. Thus to some extent it hampers EOF analysis from retrieving the physically well-constrained features from the ensemble, because in the presence of smoothing the pattern of the smallest variability is caused by a combination of the physics of the problem and the smoothness constraint. Therefore smoothing is not implemented in the construction of the ensemble of models that fit the group-velocity dispersion data.

In the section on 'theory' we defined the inner product and introduced a weighting function $w(z)$ specific to the problem. The fixed-depth parametrization we use consists of layers that increase in thickness with depth. Introducing the inner product as a simple integral over depth (e.g. $w(z) = 1$) would then result in an unequal weighting of the parameters; velocities in deep layers would affect the projection more strongly than velocities in shallow layers. This would contradict the fact that, for large depths, the velocity does not affect the surface wave very much. The original layers were chosen in such a way that each layer is of more or less the same importance in determining the group velocity of the surface waves. To maintain this in the EOF analysis we define

$$w(z) \equiv w(z_i) = \frac{1}{\Delta z_i}, \tag{22}$$

so that

$$[f, g] \equiv \int w(z) f(z) g(z) = \sum_{i=1}^{p} w(z_i) f(z_i) g(z_i) \Delta z_i = \sum_{i=1}^{p} f(z_i) g(z_i), \tag{23}$$

with Δz_i the thickness of layer i.

In order to pose the problem in a finite-dimensional context, we introduced a finite set of orthonormal functions ϕ_i in the theory section. Since the continuous depth domain is parametrized into layers, the models in the ensemble are block-shaped functions. Therefore we can again define ϕ_i as normalized block functions with respect to the inner product defined in eq. (23):

$$\phi_i(z) = \begin{cases} 1 & \text{if } z_{i-1} \le z < z_i, \\ 0 & \text{otherwise.} \end{cases} \tag{24}$$

In order to clarify our figures, we represent the models using a linear interpolation between the velocities at the centres of the layers rather than using the block functions. The EOFs and resolution kernels are also plotted in this way for consistency.

Fig. 8 shows the ensemble generated by a Monte Carlo inversion of synthetic fundamental-mode Rayleigh-wave dispersion data. The ensemble contains 143 models that show strong node-to-node oscillations. The IASP91 model used to generate the synthetic seismograms of the fundamental Rayleigh mode is contained within the scatter of the ensemble. We now apply EOF analysis to this ensemble to see if we can in effect remove the unconstrained information. Fig. 9 shows the resulting EOFs.

The first five eigenfunctions clearly show that the most robust information is concentrated in the crust and upper mantle. The rest of the eigenfunctions mainly contain information about the deeper part of the earth. This is exactly what is to be expected, considering the data employed. The synthetic

Table 3. Depth parametrization for inversion of group-velocity dispersion for the S velocity.

i	*depth* z_i *(km)*	*thickness* Δz_i *(km)*	i	*depth* z_i *(km)*	*thickness* Δz_i *(km)*
1	5.0	5.0	10	301.8	79.9
2	10.3	5.3	11	400.5	98.7
3	17.8	7.5	12	519.8	119.3
4	29.3	11.5	13	662.0	142.2
5	47.1	17.8	14	829.0	167.0
6	73.2	26.1	15	1022.9	193.9
7	109.6	36.4	16	1245.9	223.0
8	158.5	48.9	17	1500.0	254.1
9	221.9	63.4			

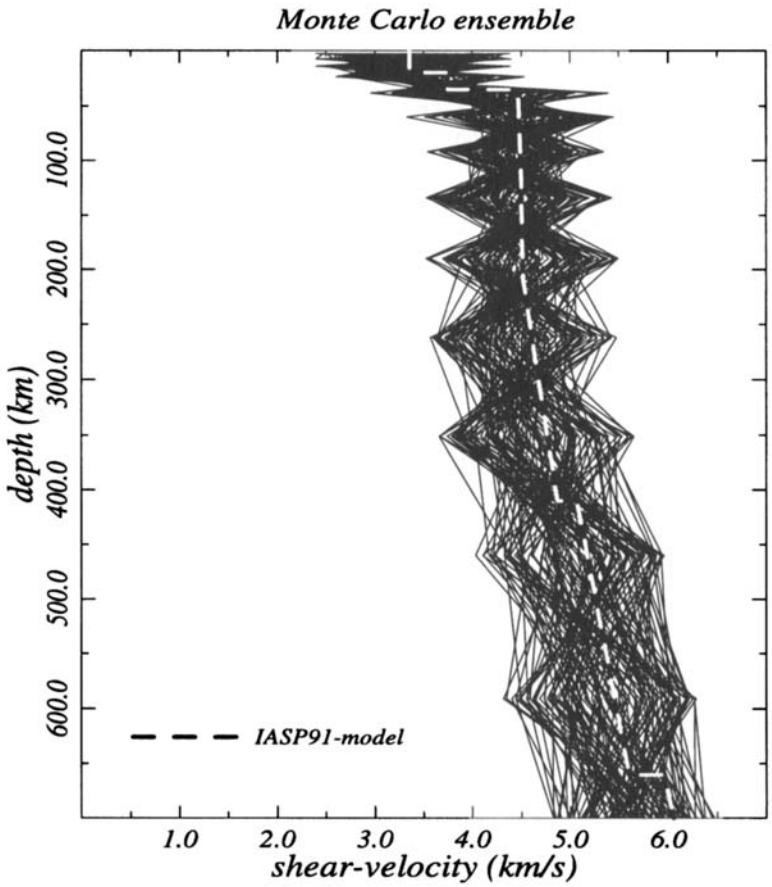

Figure 8. Ensemble resulting from the Monte Carlo inversion of fundamental-mode Rayleigh-wave group-velocity dispersion curves.

seismograms needed to generate the group-velocity estimates were generated using fundamental Rayleigh modes with periods larger than 10 s and less than 300 s. Given the penetration depth of the fundamental Rayleigh wave, we expect the models resulting from the Monte Carlo search to be poorly constrained at depths greater than about 400 km. Since the data do not constrain the model at these depths, the scatter in the ensemble is a result of the search algorithm, the search bounds and the parametrization. This deep scatter is almost entirely represented by the EOFs $\mathbf{e}_6$–$\mathbf{e}_{17}$.

The eigenvalue spectrum (Fig. 10) shows a clear cut-off between EOFs $\mathbf{e}_5$ and $\mathbf{e}_6$: $\lambda_5 = 8.9$ and $\lambda_6 = 23.8$, while $\lambda_{\min} = \lambda_1 = 1.0$ and $\lambda_{\max} = \lambda_{17} = 112.2$. As already mentioned in the section on theory, the eigenvalue λ_i reflects the variability (or scatter) of the models in the direction $\mathbf{e}_i$. Since there is a large increase in the variability from $\mathbf{e}_5$ to $\mathbf{e}_6$, we can conclude that the first five EOFs do indeed reflect the well-constrained information (or shared properties) of the models in the ensemble. The eigenvalue spectrum therefore clearly indicates the transition from well-constrained to poorly constrained information.

To verify whether the EOFs with small eigenvalues are stable with respect to the number of models in the ensemble, we calculate the EOFs resulting from different ensembles with a comparable misfit level but an increasing number of models. Fig. 11 shows the results for both an EOF with a small eigenvalue ($\mathbf{e}_2$) and one with a large eigenvalue ($\mathbf{e}_{15}$). Since the shape of $\mathbf{e}_2$ tends to converge rapidly with the size of the ensemble (N), $\mathbf{e}_2$ seems to be stable with respect to the variation in the number of models in the ensemble (N), whereas the overall change in shape of $\mathbf{e}_{15}$ does not indicate stability at all. This is to be expected, since $\mathbf{e}_2$ is supposed to encompass part of the robust information present in the ensemble, and therefore should not depend heavily on the number of models in the ensemble.

Based on the shapes of the EOFs and their eigenvalues we can decide to use only the first five EOFs to calculate the filtered ensemble $v^{p'}(\omega, z)$ with $\omega = 1, \ldots, 143$ and $p' = 5$. The next issue to be considered, however, is the depth resolution that can be obtained with these five eigenfunctions only.

The resolution kernels as calculated here depict the sensitivity at a certain depth of the filtered ensemble to a perturbation in the original ensemble of models. In other words they estimate the smearing-out in the filtered ensemble due to a perturbation in the original ensemble. Since the analysis is performed in a finite-dimensional subspace spanned by the functions $\phi_i(z)$, we cannot obtain perfect resolution as defined by the delta function in eq. (14). Therefore we redefine the resolution kernels in eq. (13) to denote the degree to which the basis functions ϕ_i can be reconstructed by using the first p' EOFs only:

$$R_i^{p'}(z) = \sum_{j=1}^{p'} [\mathbf{D}_i, \mathbf{e}_j] e_j(z), \tag{25}$$

with

$$D_i(z) = \phi_i(z). \tag{26}$$

In this context the best resolution that can be obtained is a triangle function with its maximum (equal to unity) at the

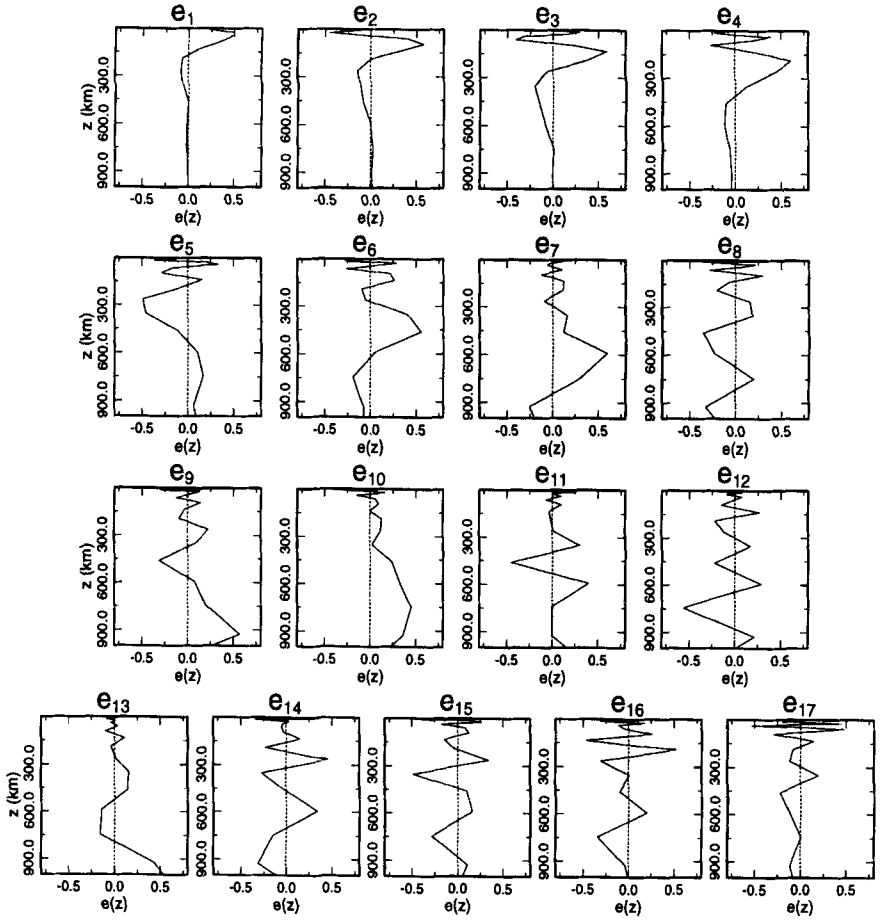

Figure 9. EOFs for the ensemble resulting from the Monte Carlo inversion of fundamental-mode Rayleigh-wave group-velocity dispersion curves.

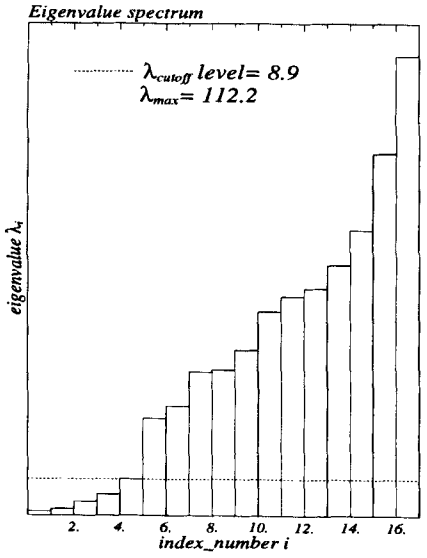

Figure 10. Eigenvalue spectrum for the ensemble resulting from the Monte Carlo inversion of fundamental-mode Rayleigh-wave group-velocity dispersion curves.

centre of a layer. First of all it should be noted that, if all EOFs are used to generate the filtered ensemble $v^{p'}(\omega, z)$, the original ensemble created by the Monte Carlo search is found again. The bottom panel of Fig. 12 shows the resolution kernel at a depth of about 135 km and the resulting filtered ensemble $v^{p'}(\omega, z)$ using all eigenfunctions. If all the EOFs are used to calculate the filtered ensemble, the original ensemble is indeed recovered (*cf.* Fig. 8) and the input function (a triangle centred at $z \approx 135$ km) can be perfectly described using all the EOFs.

If we only use five eigenfunctions, we are interested in the resolution that can be obtained at depths where the shear velocity is well constrained. Since we know that, based on the data used, the upper-mantle part ($z < 400$ km) is well constrained, we display the resolution kernel for a depth of about 135 km. Fig. 12 shows the resolution kernels and the resulting filtered ensembles for $p' = 4$, 5, 6 and 17. It can clearly be seen that if we increase p' the resolution improves but the variability at depths where the velocity is poorly constrained increases. This reflects the trade-off relation between resolution and variability. Therefore by choosing $p' = 5$ we effectively reduce the variability at depths where the velocity is poorly constrained (>400 km) while still retaining good resolution at depths where the velocity is well constrained.

The filtered ensemble using the first five EOFs ($p' = 5$) is shown in Fig. 13, together with the original ensemble. The filtered ensemble shows an effective reduction of the overall scatter, especially for the deeper (>400 km) parts of the models. The IASP91 model used to create the synthetic seismograms is clearly embedded within the filtered ensemble up to a depth of about 400 km. One should remember that the parametrization we used was different from the one Kennett & Engdahl (1991)

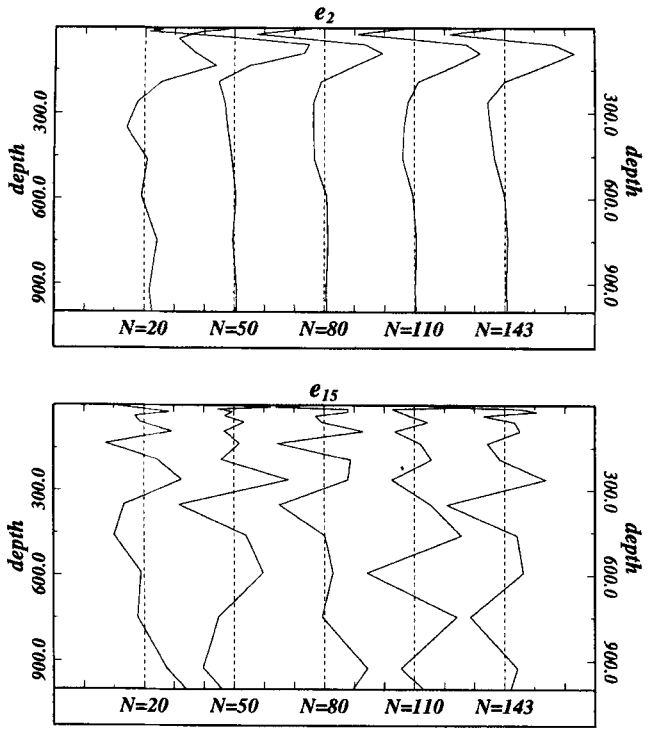

Figure 11. Estimated stability with respect to the number of models in the ensemble (N) for both an EOF with a small eigenvalue ($\mathbf{e}_2$) and one with a large eigenvalue ($\mathbf{e}_{15}$). Both EOFs are calculated for an increasing number of models N.

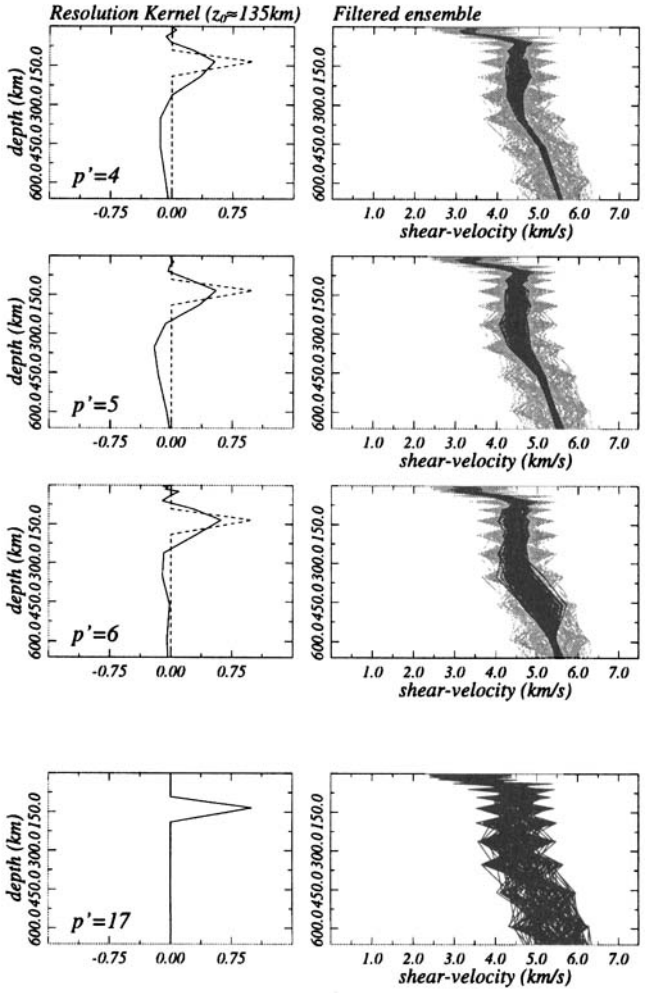

Figure 12. (Left) Resolution kernels calculated at 135 km depth using eqs (25) and (26). The dashed line represents the resolution kernel calculated using all EOFs and the solid line the kernel calculated using only the first p' EOFs. (Right) The filtered ensembles for $p' = 4$, 5, 6 and 17. The grey ensemble represents the initial ensemble.

used to determine the IASP91 model. Therefore it is possible that the IASP91 model at some depths falls slightly out of the scatter of the filtered ensemble.

Fig. 14 shows the filtered ensemble and two resolution kernels calculated for depths of 190 and 590 km respectively with $p' = 5$. The resolution kernel for 190 km indicates that the velocity at this depth is indeed well constrained. The filtered ensemble shows a broad velocity range at this depth. This range can be interpreted as an estimation of the variance of the velocity averaged over a depth range defined by the resolution kernel. The resolution kernel for 590 km is almost equal to zero. This is due to the fact that the first five EOFs all practically vanish at this depth. This reflects the fact that the filtered ensemble does not constrain the model at depths greater than the penetration depth of the surface waves. The fact that the variance in the filtered ensemble is small at a depth of 590 km should not be interpreted as a sign that the velocity at this depth is well constrained; it is due to the fact that the filtered ensemble does not constrain the velocity at this depth at all. At well-constrained depths ($z \lesssim 400$ km), the filtered ensemble reflects an estimation of the variance of the velocity at these depths.

APPLICATION TO THE CONTRAST IN UPPER-MANTLE *S* VELOCITY ACROSS THE TORNQUIST–TESSEYRE ZONE

The Tornquist–Tesseyre zone (TTZ) is a NW–SE-trending suture zone which separates the stable Precambrian East European Platform (EEP) in the east from the younger, tectonically active areas of central and western Europe, referred to as Tectonic Europe (TE), in the west (Fig. 15). Some geophysical studies have indicated a change in crustal thickness of 25–35 km beneath TE and 40–55 km under EEP (Guterch *et al.* 1986; Meissner 1986; Blundell, Freeman & Mueller 1992). Other studies using *S* body- and surface-wave data (Snieder 1988; Zielhuis & Nolet 1994; Lomax & Snieder 1995a) indicate a significant contrast in *S*-velocity structure in the uppermost mantle in the region of the TTZ, with higher velocities beneath EEP. In order to show the power of the method of EOF analysis, we present an application of the method to examine this contrast in the upper-mantle *S*-wave velocity across the TTZ. The observations we use consist of digital seismograms and are the same as the ones used by LS. For a detailed description of the data employed we therefore refer the reader to LS.

To invert the fundamental-mode Rayleigh-wave group-velocity estimates obtained from these digital seismograms we use a Monte Carlo search technique (MC) rather than the genetic algorithm that LS used. Since this latter method tends to cluster very quickly around some optimal solution, the model space is poorly sampled, whereas MC samples the model space very well (Lomax & Snieder 1995b). By using

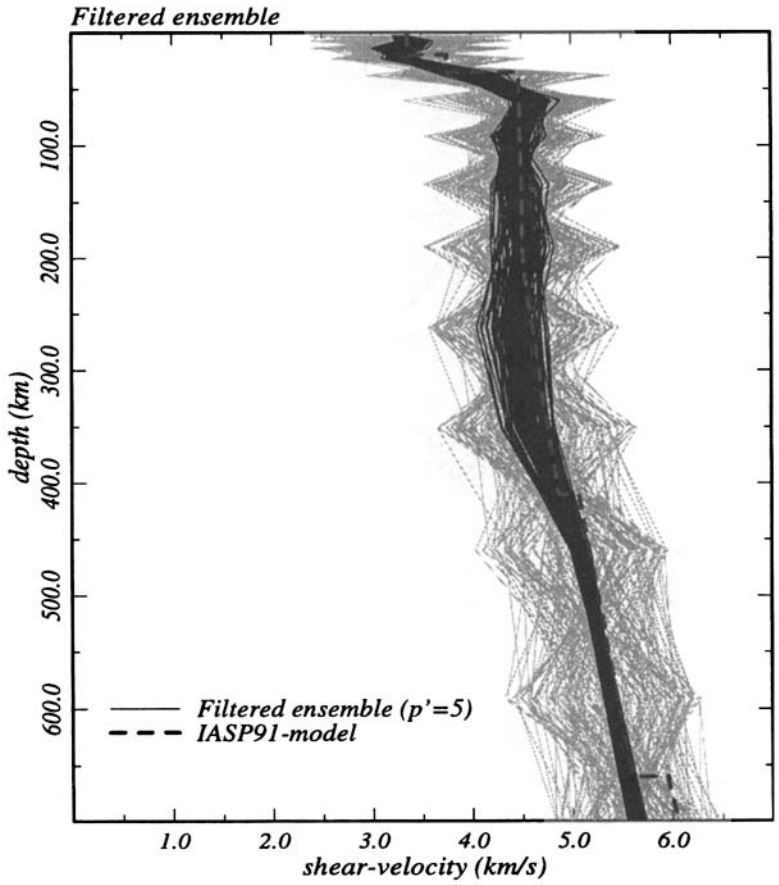

Figure 13. The filtered ensemble obtained using only the first five EOFs. The grey ensemble represents the initial ensemble of *S*-velocity models.

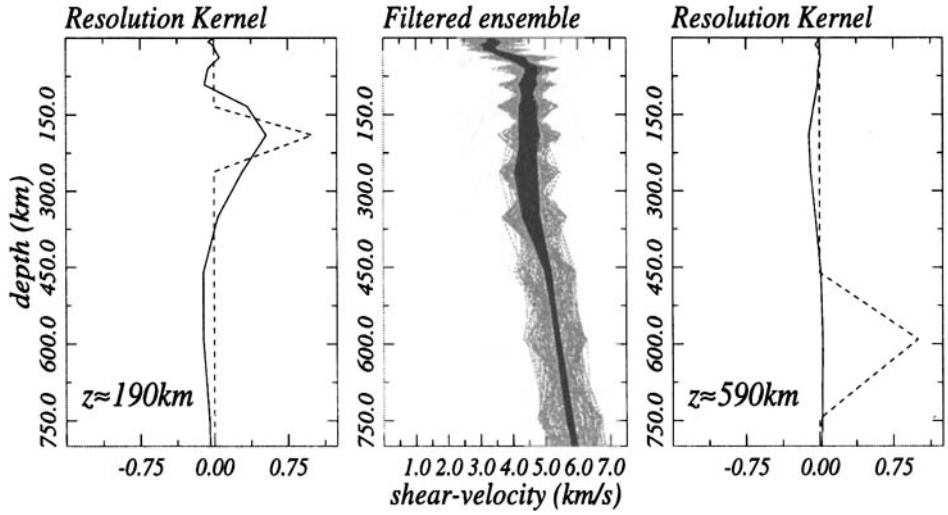

Figure 14. Resolution kernels for $z = 190$ km and $z = 590$ km (calculated using eqs 25 and 26) related to the filtered ensemble for $p' = 5$. The dotted kernels are the kernels calculated using all EOFs. At depths where the velocity is well constrained (e.g. 190 km) we have good resolution and the filtered ensemble gives an estimation of the variance of the velocity averaged over a depth range defined by the resolution kernel. For depths where the velocity is poorly constrained (e.g. 590 km) the resolution kernels are almost equal to zero. From the filtered ensemble we therefore cannot make an estimation of the variance of the velocity at these depths.

MC we create 100 000 models, and, by defining a misfit acceptance level comparable with the acceptance level resulting from the definition of acceptable models that LS used, two ensembles of 300 and 132 models are created for the EEP and TE respectively. These ensembles will be referred to as the original ensembles in the context of EOF analysis. The depth parametrization we use is identical to the one in the synthetic example (Table 3).

Fig. 16 shows the original ensembles for both EEP (in grey) and TE (in black). Since no smoothing is used, both ensembles

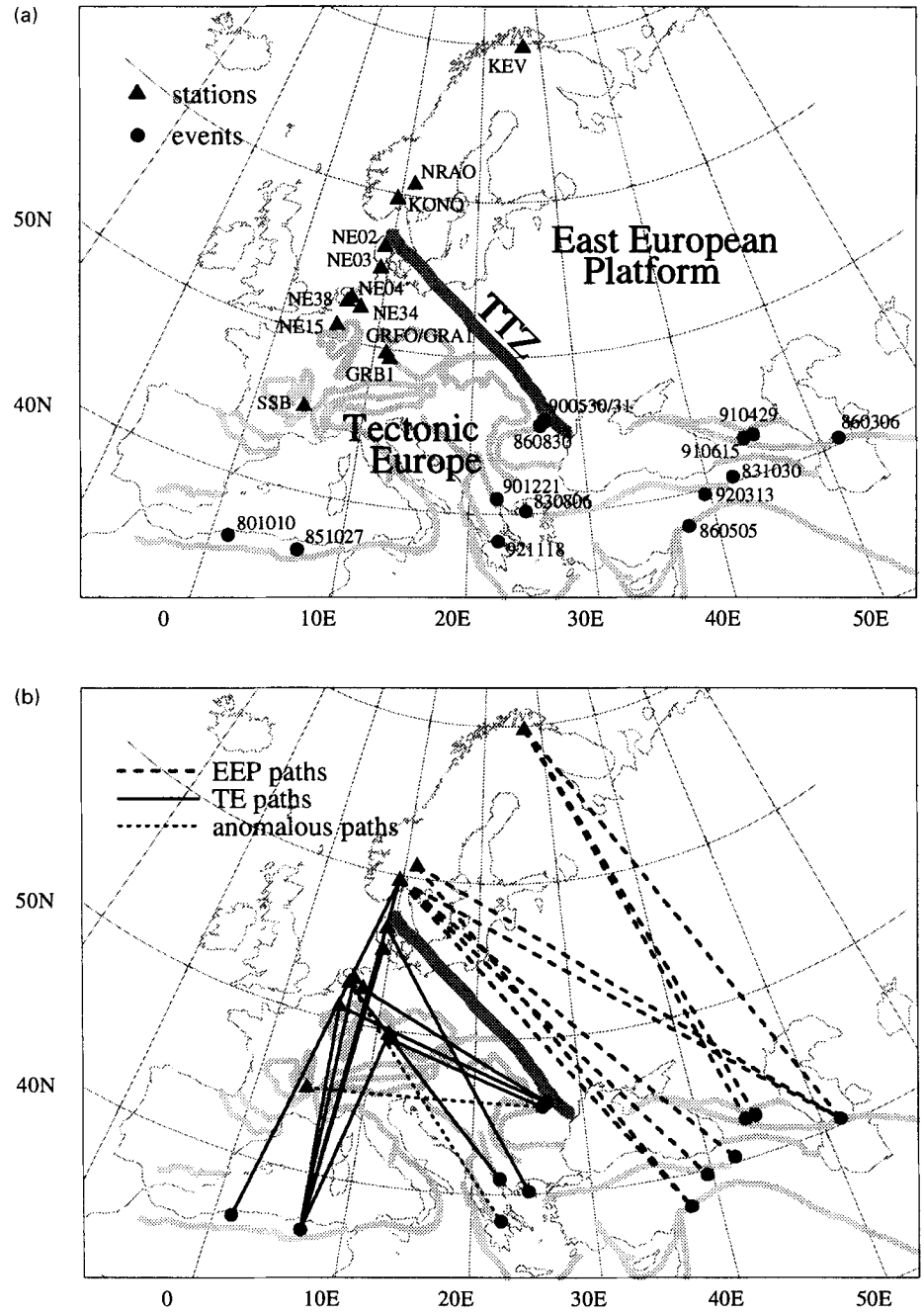

Figure 15. Figure taken from Lomax & Snieder (1995a). (a) Map showing major tectonic boundaries (medium-width grey lines), Tornquist–Tesseyre Zone (TTZ, thick grey line), stations (▲) and events (●). (b) Great-circle paths between sources and receivers for observed seismograms for the group-velocity estimates in this study. Long-dashed lines indicate paths used for East European Platform (EEP) inversion; solid lines indicate paths used for Tectonic Europe (TE) inversion. The short-dashed lines in TE indicate the paths that produce anomalous dispersion estimates.

show strong node-to-node oscillations as in the synthetic example, and they overlap to a very large degree. From these original ensembles it is virtually impossible to extract any direct information about the contrast in upper-mantle S velocity. There is a slight indication of a somewhat lower S velocity beneath TE, but we need to perform EOF analysis to examine the robust S-velocity patterns present in both ensembles. The eigenvalue spectra (Fig. 17) indicate a cut-off in eigenvalues between $\mathbf{e}_3$ and $\mathbf{e}_4$ for both EEP and TE: for EEP $\lambda_{\text{cut-off}} = 8.1$ and for TE $\lambda_{\text{cut-off}} = 4.8$. Therefore, for both EEP and TE, EOFs $\mathbf{e}_1$ to $\mathbf{e}_3$ are expected to encompass the robust shear-wave velocity patterns in both regions. The difference in $\lambda_{\max}$ (see Fig. 17) between EEP and TE is caused by the difference in the number of models in the original ensembles.

If we want to examine the contrast in S velocity between EEP and TE we should compare only the robust subspaces of the model spaces for both areas. Therefore the filtered ensembles calculated using only EOFs $\mathbf{e}_1$ to $\mathbf{e}_3$ should be compared. It should be noted that these EOFs vanish at larger depths ($\gtrsim$400 km). Considering the data employed, this is to

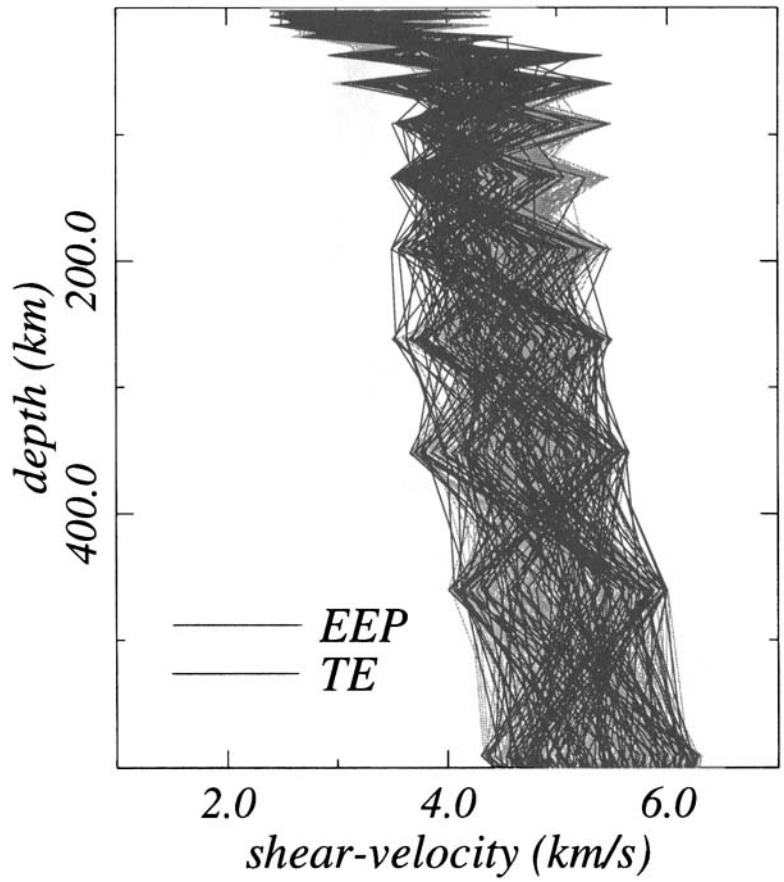

Figure 16. Original ensembles resulting from the group-velocity inversion for both EEP (grey) and TE (black). Both ensembles tend to overlap, although there is a slight indication of a somewhat lower S-velocity beneath TE.

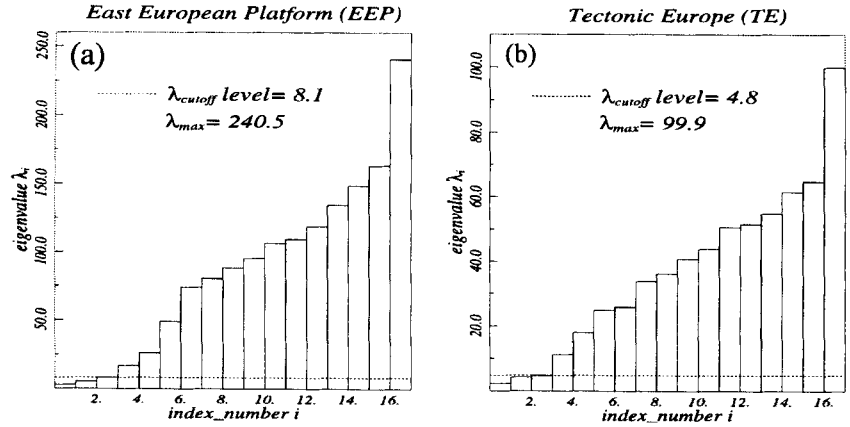

Figure 17. Eigenvalue spectra for EEP (a) and TE (b). A cut-off in the eigenvalues can be identified for both regions at $\lambda = 8.1$ and $\lambda = 4.8$ respectively. These levels correspond to the first $p' = 3$ EOFs that encompass the robust patterns in the ensembles for both regions.

be expected since the group-velocity estimates were created using only fundamental-mode Rayleigh waves with periods of between 10 and 300 s for EEP and between 7 and about 150 s for TE. Therefore the S velocity beneath both regions is poorly constrained at depths larger than about 400 km.

Fig. 18 shows the filtered ensembles calculated using only the first $p' = 3$ EOFs for both EEP and TE, and the resolution kernels (calculated from eqs 25 and 26) for both regions at depths of about 135 and 460 km. It is clearly seen that there is a significant contrast in S velocity in the upper mantle between about 50 and 225 km, with higher velocities beneath EEP and a reasonable corresponding resolution for both areas (see left-hand figure). This confirms earlier results from studies by Snieder (1988), Zielhuis & Nolet (1994) and Lomax & Snieder (1995a). At larger depths the velocities also differ slightly. The right-hand figure, however, indicates that the resolution vanishes at these depths, which means that the velocity is not constrained and that the small scatter in models at these depths should not be interpreted as an indication of a well-constrained difference in S velocities in either region.

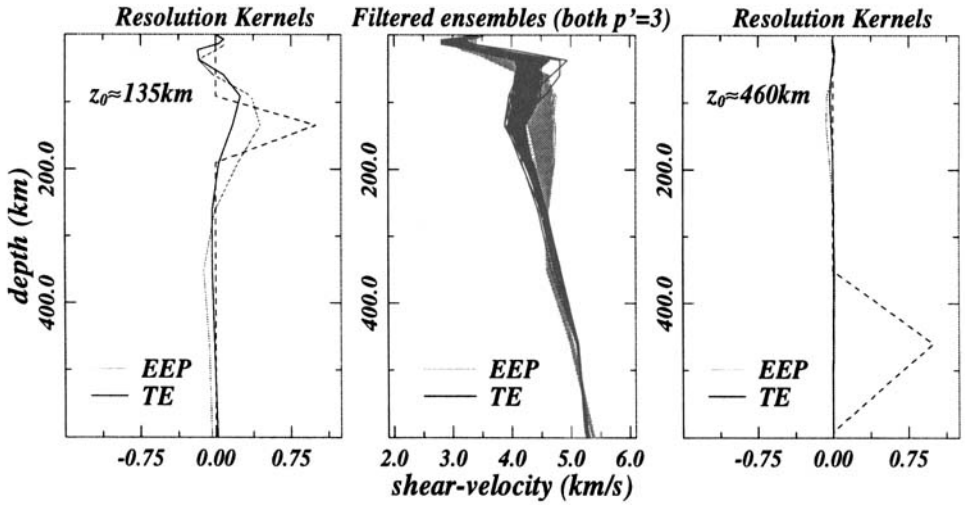

Figure 18. (Left) Resolution kernels for EEP and TE calculated at $z_0 \approx 135$ km. (Middle) Filtered ensembles using $p' = 3$ EOFs for both EEP and TE. (Right) Resolution kernels for both regions calculated at $z_0 \approx 460$ km.

The left-hand figure shows a slightly better resolution at shallower depths ($z \lesssim 400$ km) for EEP than for TE. This is probably caused by the difference in the number of accepted models present in the original ensembles. The middle panel of Fig. 18 also indicates a shallower Moho discontinuity beneath TE than under EEP: for TE, $z_{\text{MOHO}} \sim 30$ km, and for EEP, $z_{\text{MOHO}} \sim 50$ km. These depths agree with results from earlier studies by Guterch *et al.* (1986), Meissner (1986) and Blundell *et al.* (1992).

CONCLUSIONS AND DISCUSSION

The results of the synthetic tests and the application of the analysis to the geophysical problem of group-velocity inversion show the effectiveness of EOF analysis in determining the well-constrained information hidden in an ensemble of models. In the synthetic tests it is shown that the constraints that were imposed upon the model space are recovered and that EOF analysis can effectively separate this constrained information from the unconstrained information in the ensemble. The application to the group-velocity inversion shows the same thing; the deep ($\gtrsim 400$ km) scatter reflecting an artificial scatter produced by the search algorithm can in effect be removed (see Fig. 13). However, the EOF decomposition technique cannot always be expected to work well.

The applicability of EOF analysis depends strongly on the pattern of clustering in the population. If we have a problem where the ensemble is formed by one cluster, EOF analysis can effectively determine the directions of minimum (or maximum) variability (see Fig. 19). The centring around the mean is needed in order to analyse the variability within the cluster. If centring were omitted, the direction of maximum variability would be almost equal to the direction of the vector pointing to this mean.

The geophysical problem of group-velocity inversion using the fundamental Rayleigh mode is a fairly linear problem. Therefore the resulting ensemble of S-velocity models generated by MC will probably form a hyperellipsoid in p-dimensional Euclidian space. A weak non-linearity will deform the hyper-ellipsoid to a different form, such as that shown in Fig. 19. In this case, EOF analysis remains applicable. Strong non-linearities can, however, have two complicating effects.

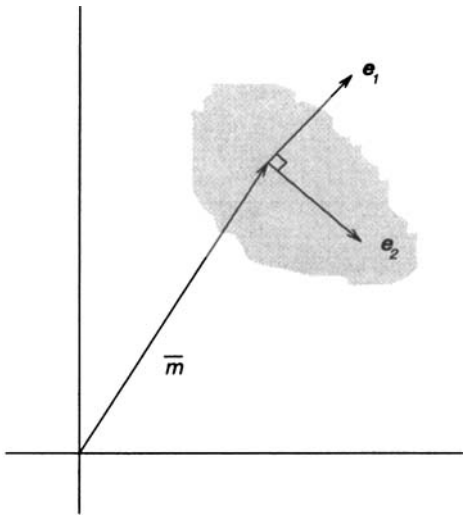

Figure 19. EOF analysis of an ensemble consisting of only one cluster.

First, non-linearity can cause the shape of the solution space to become very different from an ellipsoid, such as the hyper 'banana' in Fig. 20. If EOF analysis were applied to such an ensemble the directions of minimum and maximum variability would not be representative of the variability within the ensemble. The basic orthonormality property of the analysis prevents it from effectively determining the representative directions of minimum and maximum variability. The direction of minimum variability ($\mathbf{e}'_1$) is, however, reasonably well determined compared with the direction of maximum variability ($\mathbf{e}'_2$)—see Fig. 20. If we want to perform the EOF analysis as presented in this paper, we need to reparametrize the model space, in order to transform the curved cluster to a more ellipsoidal cluster. The linearizing transformation introduced by Vasco (1995) can be used for this purpose.

Second, the ensemble can be split up into several clusters because of the presence of secondary minima of the misfit

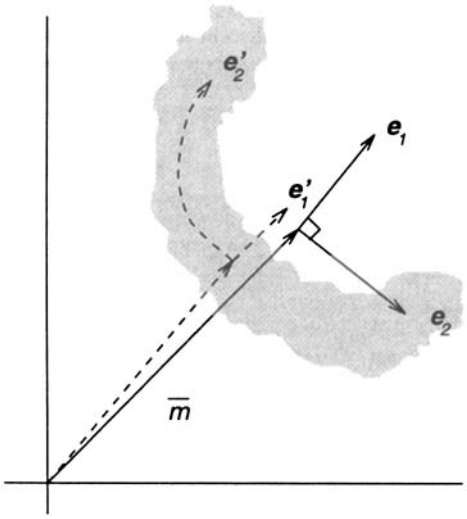

Figure 20. EOF analysis of a 'banana'-shaped ensemble, which can be the result of a strongly non-linear problem.

function. Because of the centring with respect to the mean, the effectiveness of EOF analysis might in this case break down (see Fig. 21). The direction of maximum variability will then lie along the line of centres of the two clusters, with the minimum-variability direction orthogonal to it (in 2-D). Of course these directions in general do not represent the minimum- and maximum-variability directions within the two separate clusters. In order to find the representative directions for both clusters individually, the clusters first need to be identified and separated. One possible way to do this is to use minimal spanning trees (Preisendorfer 1988, pp. 285–287; Preisendorfer & Mobley 1982).

The method of minimal spanning trees (MST) provides a way to detect natural families or clusters in an ensemble. The vectors in the ensemble are normalized to lie on a unit sphere. By starting at any particular seed vector in the ensemble, searching for the nearest neighbour, linking them up, searching for their joint nearest neighbour, linking them up, and so on, recursively different MSTs can be formed for different seed vectors. Each MST can produce a pair of clusters $(A, B)_i$ by

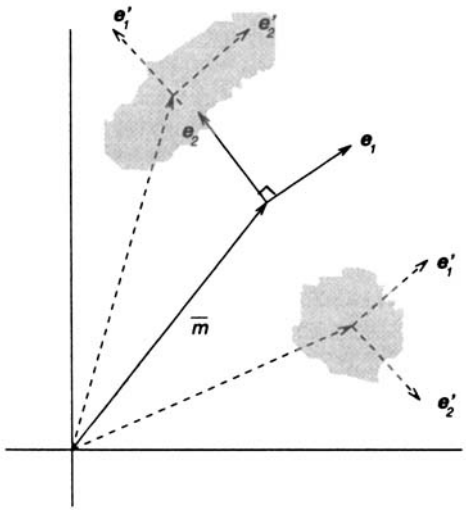

Figure 21. EOF analysis of an ensemble consisting of two clusters. The clustering can be the result of a strongly non-linear problem.

removing the link of maximal length l_i. By defining the ratio $R_i \equiv l_i/d_i$, with d_i the average diameter of the cluster pair $(A, b)_i$ (the diameter can for example be defined as the maximum distance between the vectors in the ensemble), and searching for the cluster pair with the largest ratio R_i, two clusters can be identified.

After the clusters in the ensemble have been recognized, EOF analysis can be applied to the individual clusters in order to determine the shared properties in each cluster. In general, EOF analysis will not be able to detect clusters in the ensemble by itself. Since the EOFs form an orthonormal set of functions, it is unlikely (if not impossible) that the analysis will come up with eigenvectors that point to the different clusters.

EOF analysis appears to be very similar to singular value decomposition (SVD), which is used to analyse linear inverse problems of the form

$$\mathbf{d} = \mathbf{G}\mathbf{m}, \tag{27}$$

where **d** are the data, **G** is a matrix holding the sensitivity kernels for the initial model, and **m** is a vector holding the model deviations from the initial model. The matrix **G** is then decomposed using SVD. The eigenfunctions with the largest eigenvalues are then considered to be the most important ones since they influence the data most strongly. SVD thus calculates the eigenfunctions from the sensitivity kernels **G**, the known physics of the problem. EOF analysis calculates the eigenfunctions from the ensemble of acceptable models represented in the scatter matrix. Therefore it extracts the relevant physics of the problem out of the ensemble, instead of relying directly on physical theory as SVD does.

Nolet & Snieder (1990) proposed a projection technique using Lanczos' method (Lanczos 1950) to solve large linear inverse problems. This method tries to remove the redundancy of a subspace in $\mathcal{M}$, the Hilbert space of earth models spanned by the $G_i(x)$ $(i = 1, \ldots, N, N$ is large) that form the sensitivity kernels of the projection defined as

$$\int_0^a G_i^*(x) m(x)\, dx = d_i. \tag{28}$$

In this way a *sufficient* subspace $\mathcal{K}$ of $\mathcal{M}$, spanned by a different basis $\mu_1(x), \ldots, \mu_K(x)$ with $K \ll N$, is defined such that for arbitrary **d** a model $m(x) \in \mathcal{K}$ can be found that satisfies (28) within the precision of **d**. The Lanczos iteration method provides the basis of the proposed technique which describes a sequence of projections and backprojections. In EOF analysis, the unconstrained part of the model space is effectively filtered, as shown by the synthetic examples, by determining the directions of small variability. This means that the redundancy of the model space can be reduced without explicit knowledge of the sensitivity kernels $G_i(x)$, which is required in the projection method of Nolet & Snieder (1990).

Since the EOFs with the smaller eigenvalues (selected by defining the cut-off level in the eigenvalue spectrum) encompass the robust patterns present in an ensemble of models, they are candidates for reparametrization of the inverse problem. Reparametrization using these functions might increase the efficiency of the inversion for a specific problem in a certain area, since the number of parameters has been reduced. But, more importantly, future inversions in this area will be more focused on the well-constrained properties using the reparametrization, and therefore will be more effective.

Summarizing, we can conclude that EOF analysis provides a method to extract the robust patterns out of an ensemble, that it is able to estimate the variance of the patterns, and that the calculated EOFs can be used for an intelligent reparametrization of the models.

ACKNOWLEDGMENTS

The critical comments of an anonymous reviewer are very much appreciated. This research was supported by the Netherlands Organization for Scientific Research (NWO) through the Pionier project PGS 76-144. This is Geodynamics Research Institute (Utrecht University) publication 96.035.

REFERENCES

Basu, A. & Frazer, L.N., 1990. Rapid determination of the critical temperature in simulated annealing inversion, *Science,* **249,** 1409–1412.

Blundell, D., Freeman, R. & Mueller, S., 1992. *A Continent Revealed, The European Geotraverse,* Cambridge University Press, Cambridge.

Dziewonski, A., Bloch, S. & Landisman, M., 1969. A technique for the analysis of transient seismic signals, *Bull. seism. Soc. Am.,* **59,** 427–444.

Goldberg, D.E., 1989. *Genetic Algorithms in Search, Optimization and Machine Learning,* Addison-Wesley, Reading, MA, pp. 1–88.

Guterch, A., Grad, M., Materzok, R. & Perchuć, E., 1986. Deep structure of the Earth's crust in the contact zone of the Paleozoic and Precambrian platforms in Poland (Tornquist-Teisseyre Zone), *Tectonophysics,* **128,** 251–279.

Kennett, B.L.N. & Engdahl, E.R., 1991. Traveltimes for global earthquake location and phase identification, *Geophys. J. Int.,* **105,** 429–466.

Kennett, B. & Nolet, G., 1978. Resolution analysis for discrete systems, *Geophys. J. Int.,* **53,** 413–425.

Lanczos, C., 1950. An iteration method for the solution of the eigenvalue problem of linear differential and integral operators, *J. Res. Nat. Bur. Stand.,* **45,** 255–281.

Lomax, A. & Snieder, R.K., 1995a. The contrast in upper mantle shear-wave velocity between the East European Platform and tectonic Europe obtained with genetic algorithm inversion of Rayleigh-wave group dispersion, *Geophys. J. Int.,* **123,** 169–182 (LS).

Lomax, A. & Snieder, R.K., 1995b. Identifying sets of acceptable solutions to non-linear, geophysical inverse problems which have complicated misfit functions, *Nonlinear Processes in Geophysics,* **2,** 222–227.

Meissner, R., 1986. *The Continental Crust,* Academic Press, Orlando, FL.

Mosegaard, K. & Tarantola, A., 1995. Monte Carlo sampling of solutions to inverse problems, *J. geophys. Res.,* **100,** 12 431–12 447.

Nolet, G. & Snieder, R.K., 1990. Solving large linear inverse problems by projection, *Geophys. J. Int.,* **103,** 565–568.

Nolte, B. & Frazer, L.N., 1994. Vertical seismic profile inversion with genetic algorithms, *Geophys. J. Int.,* **117,** 162–178.

Preisendorfer, R.W., 1988, *Principal Component Analysis in Meteorology and Oceanography,* Elsevier Science Publishers, Amsterdam, pp. 1–44, 75–81, 285–287.

Preisendorfer, R.W. & Mobley, C.D., 1982. Data intercomparison theory. I. Minimal spanning tree tests for location and scale differences, *NOAA Tech. Memo. ERL PMEL-38,* Pacific Mar. Env. Lab., Seattle, Wash. 98115 (NTIS PB83-182311).

Rao, C.R., 1964. The use and interpretation of principal component analysis in applied research, *Sankhyã Ser.* A, **26,** 329–358.

Rinne, J., Karhila, V. & Järvenoja, S., 1981. The EOF's of the 500 mb height in the extratropics of the northern hemisphere, *Rept. No. 17,* Dept. of Meteorology, Univ. of Helsinki, Finland.

Rothman, D.H., 1985. Nonlinear inversion, statistical mechanics, and residual statistics estimation, *Geophysics,* **50,** 2797–2807.

Sambridge, M. & Drijkoningen, G., 1992. Genetic algorithms in seismic waveform inversion, *Geophys. J. Int.,* **109,** 323–342.

Sen, M.K. & Stoffa, P.L., 1992. Rapid sampling of model space using genetic algorithms: examples from seismic waveform inversion, *Geophys. J. Int.,* **108,** 281–292.

Snieder, R.K., 1988. Large-scale waveform inversions of surface waves for lateral heterogeneity 2. Application to surface waves in Europe and the Mediterranean, *J. geophys. Res.,* **93,** 12 067–12 080.

Trenberth, K.E. & Shin, W.K., 1984. Quasi-biennal fluctuations in sea level pressures over the northern hemisphere, *Mon. Weather Rev.,* **112,** 761–777.

Vasco, D.W., 1995. A transformational approach to geophysical inverse problems, *Geophys. J. Int.,* **123,** 183–212.

Vasco, D.W., Johnson, L.R. & Majer, E.L., 1993. Ensemble inference in geophysical inverse problems, *Geophys. J. Int.,* **115,** 711–728.

Zielhuis, A. & Nolet, G., 1994. Shear-wave velocity variations in the upper mantle beneath central Europe, *Geophys. J. Int.,* **117,** 695–715.