GEOPHYSICAL TRANSACTIONS 1993
Vol. 38. No. 4. pp. 239-249

COMPARISON OF THE KARHUNEN-LOÈVE STACK WITH THE CONVENTIONAL STACK

Leif BRULAND*

Several applications of the Karhunen-Loève (KL) transform to seismic data are known, among which is the use of the first principal component as an alternative stack — the KL stack. On analysing and comparing the KL stack with the conventional stack, it was found that the KL stack is more influenced by noise, especially coherent noise, than the conventional one. With approximately the same signal amplitudes from trace to trace, the conventional stack is therefore the better choice. On the other hand, if the signal amplitudes vary and the noise is uncorrelated with approximately constant energy on all traces, the KL stack should be preferred.

It has been claimed that the KL stack is relatively insensitive to small time shifts of the signals, and that correction for residual statics may be unnecessary when the KL stack is used. It is confirmed here that the KL stack generally gives the better signal-to-noise ratio in such cases. However, the time shifts may seriously distort the output signal, and the distortion is found to be very sensitive to changes in the time shifts, in view of which it is important to correct for residual statics even if the KL stack is used.

Keywords: seismic, stacking, Karhunen-Loève Transformation

1. Introduction

The Karhunen-Loève Transform (KLT) is used to represent a set of, say, M input vectors or traces by a particular set of M orthogonal vectors called principal components. The principal components are linear combinations of the input vectors constructed in such a way that most of the coherent energy is contained in the first component, or in the first few components. The KLT can therefore be used to express information in a compact way. The principal components have long been used in multivariate statistical analysis both for data reduction and in interpretation.

* Institute of Solid Earth Physics, University of Bergen, Allegaten 41, N-5007 Bergen, Norway
Manuscript received: 18 June, 1993.

Since the first principal component, which can be looked upon as a weighted stack, usually contains most of the coherent energy from the input data, it may be used as an alternative stack. This was demonstrated by HEMON and MACE [1978], who initially suggested the application of the KLT to seismic data. Several other applications of the KLT to seismic data were later presented by ULRYCH et al. [1983], LEVY et al. [1983], JONES, LEVY [1987], YEDLIN et al. [1987] and FREIRE, ULRYCH [1988].

In this paper we are mainly concerned with the use of the first principal component as an alternative stack, hereafter called a KL stack. After a short introduction to the theory of the KLT, the properties of the KL stack are explored and compared with those of the conventional stack.

2. The Karhunen-Loève Transform

Let the data be given as

$$\bar{x}_i = (x_{i1}, x_{i2}, \ldots, x_{iN})^T, \quad i = 1, 2, \ldots, M \tag{1}$$

where M is the number of traces, and N the number of samples per trace, $M<N$. All traces are assumed to have zero mean values.

We now search for a vector $\bar{y}$ as a linear combination of the $\bar{x}$'s

$$\bar{y} = \sum_{i=1}^{M} a_i \bar{x}_i = X\bar{a} \tag{2}$$

where $X = \{\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_M,\}$, and $\bar{a} = (a_1, a_2, \ldots, a_M)^T$.

The energy (or variance) of $\bar{y}$ is then

$$V(\bar{y}) = \bar{y}^T \bar{y} = \bar{a}^T X^T X\bar{a} = \bar{a}^T C\bar{a}. \tag{3}$$

where $C=X^T X$ is the covariance matrix of the data.

The first principal component is defined as the vector $\bar{y}$ that maximizes $V(\bar{y})$ under the restriction

$$\bar{a}^T \bar{a} = \sum_{i=1}^{M} a_i^2 = 1. \tag{4}$$

Maximizing (3) subject to (4) is equivalent to maximizing

$$f(\bar{a}, \lambda) = \bar{a}^T C\bar{a} + \lambda (1-\bar{a}^T \bar{a}), \tag{5}$$

where λ is a Lagrange multiplier. Differentiation of $f(\bar{a}, \lambda)$ and equating the result to 0 leads to:

$$\frac{\partial f(\bar{a}, \lambda)}{\partial \bar{a}} = 2C\bar{a} - 2\lambda\bar{a} = 0$$

or

$$(C - \lambda I)\,\bar{a} = 0. \tag{6}$$

From (6) it follows that λ must be an eigenvalue and $\bar{a}$ the associated eigenvector of C. Therefore we must have

$$\bar{a}^T C\bar{a} = \bar{a}^T(\lambda\bar{a}) = \lambda,$$

and the solution to the maximization problem is the eigenvector corresponding to the largest eigenvalue of C (all eigenvalues of C are ≥ 0).

The next principal component is found from (6) when $\bar{a}$ is the eigenvector associated with the next largest eigenvalue, and so on. We can thus write

$$Y = XA \tag{7}$$

where

$$A = \{\bar{a}_1, \bar{a}_2, \ldots, \bar{a}_M,\}$$

Since C is symmetric, the eigenvectors are orthogonal, and $A^T = A^{-1}$. Multiplication of (7) by A^T gives

$$X = YA^T, \tag{8}$$

which is then the inverse transformation.

The variance of the i^{th} trace is

$$V(\bar{x}_i) = \bar{x}_i^T \bar{x}_i.$$

The variance of $\bar{y}_i$ is

$$V(\bar{y}_i) = \bar{y}_i^T \bar{y}_i = \bar{a}_i^T X^T X\bar{a}_i = \lambda_i,$$

and the eigenvalues are therefore just the energy or variance of the principal components.

The total energy of the input data is

$$\sum_i V(\bar{x}_i) = \text{Trace}\,[X^T X] = \sum_i \lambda_i = \sum_i V(\bar{y}_i). \tag{9}$$

From this it follows that the total energy is invariant under the transformation.

Since $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_M$, most of the energy is contained in the first principal components. We can therefore approximate X by a linear combination of the principal components with largest energy, say the first $P<M$ components:

$$X \approx Y^{N\times P} \cdot A^{P\times M}. \tag{10}$$

The amount of reconstructed energy can be calculated from

$$E(P) = \frac{\sum_{i=1}^{P} \lambda_i}{\sum_{i=1}^{M} \lambda_i}. \tag{11}$$

3. KLT and Singular Value Decomposition (SVD)

An SVD of the data matrix X also leads to the matrix of coefficients, A, and the matrix of principal components, Y. To see this we start with the matrix

$$B = \begin{bmatrix} 0 & X \\ X^T & 0 \end{bmatrix}, \; B^T = B \tag{12}$$

The eigenvalue problem for this matrix can be written

$$\begin{bmatrix} 0 & X \\ X^T & 0 \end{bmatrix} \begin{bmatrix} \bar{u} \\ \bar{v} \end{bmatrix} = l \begin{bmatrix} \bar{u} \\ \bar{v} \end{bmatrix}, \tag{13}$$

where $\bar{u}$ is an $(N\times 1)$ vector, $\bar{v}$ is an $(M\times 1)$ vector, and l is an eigenvalue. Since B is symmetric, l will be real. From (13) we get

$$X\bar{v} = l\bar{u}, \; X^T\bar{u} = l\bar{v} \tag{14}$$

Premultiplication of the two equations by X^T and X, respectively, gives

$$X^T X\bar{v} = l^2\bar{v}, \; XX^T\bar{u} = l^2\bar{u} \tag{15}$$

We thus see that $\bar{v}$ is an eigenvector of $C=X^TX$ and $\lambda=l^2$ the associated eigenvalue, while $\bar{u}$ is an eigenvector of XX^T with the eigenvalue λ.

For convenience we assume the rank of C and XX^T to be M. C is then a positive definite matrix, and therefore all eigenvalues are greater than zero. We order the eigenvalues so that $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_M \geq 0$, and let the corresponding eigenvectors be normalized so that $\bar{v}_i^T \bar{v}_i = 1$, $\bar{u}_i^T \bar{u}_i = 1$, i=1, 2, ..., M. We then define the matrices V and U as

$$V = \{\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_M\}$$

$$U = \{\bar{u}_1, \bar{u}_2, \ldots, \bar{u}_M\}$$

and define Λ as the matrix with the eigenvalues as its diagonal elements with zeroes elsewhere.

From (14) we get

$$XV = U\Lambda \tag{16}$$

Comparing (16) with (7) we find that

$$A = V$$

$$Y = U\Lambda \tag{17}$$

Postmultiplication of (16) with V^T gives the decomposition of X

$$X = U\Lambda V^T = YA^T$$

which according to (8) is equivalent to the inverse transformation.

4. Comparison of the KL stack with the conventional stack

Some properties of the KL stack are more easily revealed by observing that the principal components can be derived in a different way.

From (17) it follows that the principal components are scaled versions of the first M eigenvectors of XX^T. These eigenvectors can also be found from a maximization problem, viz.

$$\max (\bar{u}^T XX^T \bar{u}) \tag{18}$$

under the restriction

$$\bar{u}^T \bar{u} = 1$$

Since this leads to exactly the same sort of problem as was defined by (5), only with $X^T X$ replaced by XX^T, $\bar{u}$ will be the eigenvector of XX^T that is associated with the largest eigenvalue λ. The first principal component is just $\bar{y} = \lambda\bar{u}$. But expression (18) can be written

$$\max (\bar{u}^T XX^T \bar{u}) = \max \sum_{i=1}^{M} (\bar{x}_i^T \bar{u})^2 \tag{19}$$

Thus, the normalized first principal component maximizes the sum of the square of the inner products between this component and the traces. It can be easily shown that the normalized conventional stack maximizes the sum of the inner products. In summary,

The normalized first principal component, $\bar{u}$, maximizes $\sum_i (\bar{x}_i^T \bar{u})^2$

The normalized conventional stack, $\bar{s}$, maximizes $\sum_i (\bar{x}_i^T \bar{s})$

From these properties of $\bar{u}$ and $\bar{s}$ we can draw some conclusions:

If one or more traces are reversed in polarity, this will have no influence on the KL stack. This is shown by *Fig. 1.* where exactly half of the traces have been reversed in polarity so that the conventional stack becomes a zero trace.

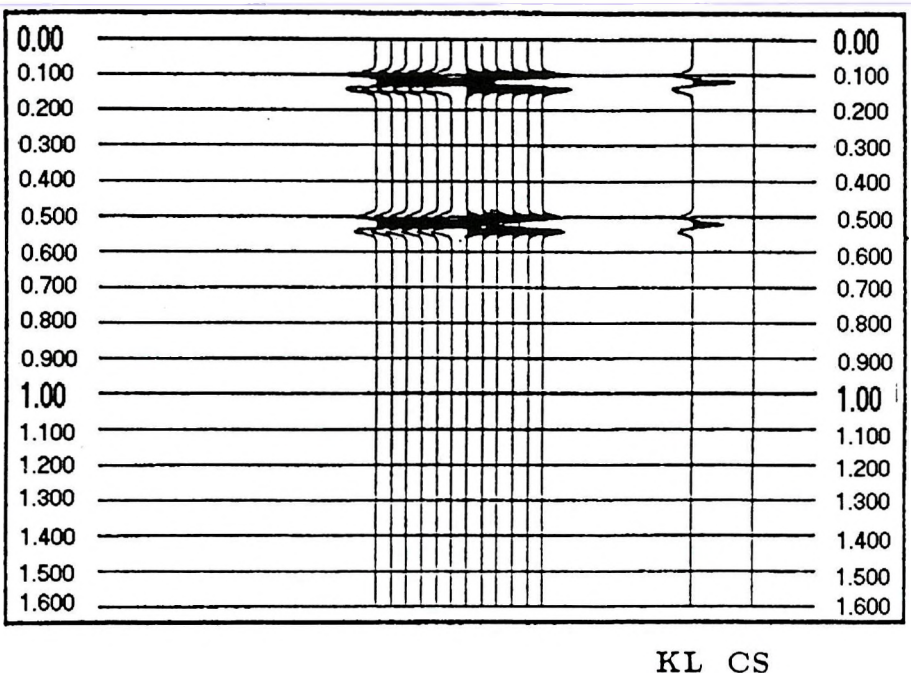

Fig. 1. Traces may be reversed in polarity without affecting the KL stack. In this example 6 out of 12 identical traces have been reversed in polarity. The comventional stack (CS) gives a zero trace, the KL stack (KL) reproduces the input trace

1. ábra. A csatornák polraitása megfordulhat anélkül, hogy a KL stacking eredményét megváltoztatná. A példán 12 azonos csatornából 6 ellentétes polraitású. A hagyományos stacking (CS) zéró csatornát eredményez, a KL stacking (KL) a bemeneti csatornát adja vissza

If the noise is uncorrelated from trace to trace, and all traces have identical signals and the same signal-to-noise ratio, the conventional stack is the optimum (weighted) stack. In this case the weights in (2) will also be equal, and therefore the KL stack is also optimum. Now, if the noise energy varies from trace to trace, the KL stack will be most influenced by the traces with highest noise energy. This is true whether the noise is correlated or not, but the effect will be more pronounced if the noise is coherent over two or more traces. This result is illustrated in *Fig. 2.*

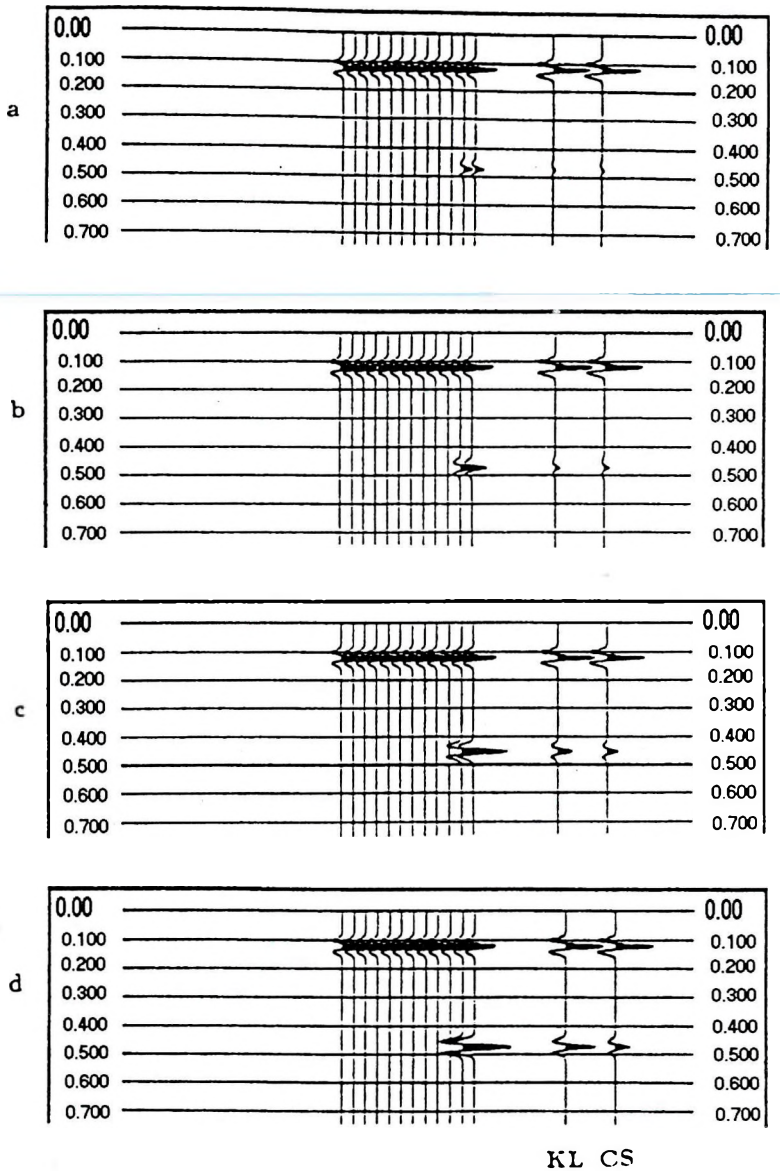

Fig. 2. Illustration of the effect of a 'noise signal' present on 2 traces (in this case). When the noise energy is small compared with the signal energy, there are no discernible differences between the KL stack (KL) and the conventional stack (CS) (a and b). With an increasing relative amount of noise energy, the differences become quite pronounced (c and d). The stacks have been scaled to equal signal amplitudes

2. ábra. Két csatornán jelen lévő „zavar jel" hatása. Ha a zaj energiája a jel energiájához képest kicsi, a KL stacking (KL) és a hagyományos stacking (CS) között nincsenek észrevehető különbségek (a és b). A zaj energiája relativ hányadának növekedésével a különbségek meglehetősen hangsúlyozottak lesznek (c és d). A stackingeket az egyenlő jel amplitudókhoz igazítottuk

The synthetic input data to the left of Fig. 2a, contain one coherent signal (identical on all traces) and a 'noise' signal present on only two out of twelve traces. The KL stack and the conventional stack are shown to the right. The stacks have been normalized to the signal amplitude. There is no visually discernable difference between these stacks but the traces with noise were given slightly higher weights than the other traces in the KL stack. If we increase the noise energy on the two input traces, these traces will be given successively higher weights in the KL stack (Fig. 2b-2d).

It may be illustrative to calculate the weights for traces with and without noise in a case like the last one.

Let the traces be given as

$$\bar{x}_i = \bar{s}, \qquad i = 1, 2, ..., m$$
$$\bar{x}_i = \bar{s} + \bar{n}, \quad i = m+1, m+2, ..., M$$

We assume $\bar{s}^T \bar{n} = 0$ (i.e., no overlap between coherent noise and signal), and denote $\bar{s}^T \bar{s} = a$, $\bar{n}^T \bar{n} = b$ and $(\bar{s}+\bar{n})^T (\bar{s}+\bar{n}) = a+b = c$. We then have

$$X^T X = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}, \tag{20}$$

where A is an $m \times m$ matrix, B is $m \times (M-m)$ and C is $(M-m)\times(M-m)$. The elements in A and B are all equal to a, and those of C are all equal to c. The eigenvalue-eigenvector problem is then

$$\begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \begin{bmatrix} \bar{u} \\ \bar{v} \end{bmatrix} = \lambda \begin{bmatrix} \bar{u} \\ \bar{v} \end{bmatrix}, \tag{21}$$

where $\bar{u}$ is an $m \times 1$ and $\bar{v}$ an $(M-m) \times 1$ vector. The eigenvector associated with the largest λ has only two different elements, since all elements in $\bar{u}$ must be equal, and so must all elements in $\bar{v}$. These values, which we denote u and v, respectively, are the weights given to traces without noise and with noise in the calculation of the KL stack.

The system is now reduced to

$$\begin{bmatrix} ma & (M-m)\, a \\ ma & (M-m)\, c \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \lambda \begin{bmatrix} u \\ v \end{bmatrix}, \tag{22}$$

Solving for λ gives

$$\lambda = \frac{1}{2}\left[(M-m)\, c+ma\right] + \frac{1}{2}\sqrt{\left[(M-m)\, c+ma\right]^2 - 4ma\,(M-m)\,(c-a)}, \tag{23}$$

and the ratio v/u becomes

$$\frac{v}{u} = \frac{\lambda}{\lambda - b\,(M-m)}$$

Since the total energy of the traces is $E = (M-m)\,c + ma$ and the total noise energy is $N = b\,(M-m)$, we get the following inequality for the ratio v/u

$$\frac{E}{E-2N} > \frac{v}{u} > \frac{E}{E-N} \tag{24}$$

We thus see that the noise traces will always get larger weights in the calculation of the KL stack, even if the noise is present on only one trace. It may be concluded that as long as the signal is completely coherent with constant amplitudes from trace to trace, the conventional stack should be preferred to the KL stack irrespective of the noise structure.

It has been claimed that the KL stack is relatively insensitive to small trace-to-trace time shifts of the signal, and therefore residual static correction can often be avoided when the KL stack is used [HEMON, MACE 1978, ULRICH et al. 1983]. However, this is only partly true, as can be seen from the following argument.

One choice of the weights, a_i, in equation (2) which satisfies (4) is $a_k = 1$, $a_i = 0$ when $a \neq k$. Thus the energy of the first principal component is always greater than or equal to the energy of the trace with the highest energy. This means that even if the signal is somewhat out of phase from trace to trace, the signal will not be cancelled by a KL stack as it might be by a conventional stack. With uncorrelated noise, the S/N ratio will therefore be higher in the KL stack than in the conventional stack. However, there is no guarantee that the KL stack will reproduce the signal; in fact it may be highly distorted, and the form of the signal in the KL stack is very sensitive to small changes in the statics. This is illustrated in *Fig. 3*, where quite different signals appear in the KL stack although only one trace has been changed from step to step. If signal distortion is to be avoided, it is therefore necessary to perform residual static correction even if the KL stack is to be applied.

Next we consider the case with varying signal amplitudes across the traces. If the noise is approximately uncorrelated with nearly the same energy on all traces, we can use the arguments of the last example to see that in this case the KL stack is preferable to the conventional stack. This follows from the fact that the energy in the KL stack cannot be less than the energy in the trace of maximum energy, and since the difference in trace energy is due to the difference in signal energy, the S/N ratio will always be higher in the KL stack than in the single traces. This will not always be the case for the conventional stack. If the noise varies from trace to trace, the situation becomes more obscure since the relative amount of signal energy to noise energy will affect the weights in the KL stack. With an increasing amount of coherent noise, the KL stack should again be avoided.

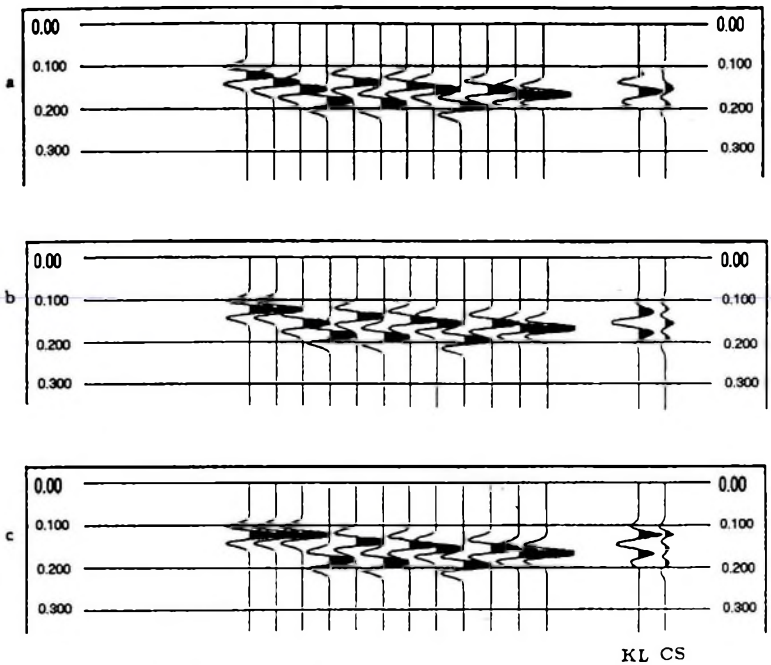

Fig. 3. KL stacks (KL) and conventional stacks (CS) of sets of traces with time shifted signals. Only one of the traces (second from the left) in Fig. 3b is different from those in Fig. 3a. In Fig. 3b and in 3c, only the third trace is different

3. ábra. KL stackingek (KL) és hagyományos stackingek (CS) csatornasorozata időben eltolt jelekkel. A 3b. ábrán csak egyetlen csatorna (balról a második) tér el a 3a. ábrán lévő csatornáktól. A 3b és a 3c ábrák között csak a harmadik csatornában van különbség

5. Conclusions

The properties of the KL stack have been analysed and compared with those of the conventional stack, and the results can be summarized as follows:

Both stacks are optimal in the case of identical signals contaminated by completely uncorrelated noise. With identical signals on all traces the conventional stack is superior to the KL stack in all other cases.

Correlated noise will always have a greater influence on the KL stack than on the conventional stack. The differences between stacking methods are small as long as the amount of noise energy is small compared with the total signal energy, but they increase rapidly with an increasing relative amount of correlated noise energy. This is true whether the noise is coherent over all traces (except for the case of identical noise 'signals' on all traces) or only a few.

If traces with residual statics are KL stacked, the S/N ratio will normally increase (and never decrease), but the signal may be highly distorted. It is

therefore important to perform residual static correction even if KL stacking is to be applied.

If the signal amplitudes vary across the traces while the noise is uncorrelated and has approximately the same energy on each trace, the KL stack is a better choice than the conventional stack. This may be so even when the noise varies and/or is correlated to some extent, but it would be very difficult to prescribe which method to use in such cases.

REFERENCES

FREIRE S. L. M., ULRYCH T. J. 1988: Application of singular value decomposition to vertical seismic profiling. Geophysics **53**, 6, pp. 778-785

HEMON CH, MACE D. 1978: Essai d'une application de la transformation de Karhunen-Loève au traitement sismique. Geophysical Prospecting **26**, 3, pp. 600-626

JONES I. F., LEVY S. 1987: Signal-to-noise ratio enhancement in multichannel seismic data via the Karhunen-Loève transform. Geophysical Prospecting **35**, 1, pp. 12-32

LEVY S., ULRYCH T. J., JONES I. F., OLDENBURG D. W. 1983: Applications of complex common signal analysis in exploration seismology. Proceedings of the 53rd Annual SEG meeting, Las Vegas, S6.6

ULRYCH T. J., LEVY S., OLDENBURG D. W., JONES I. F. 1983: Applications of the Karhunen-Loève transformation in reflection seismology. Proceedings of the 53rd Annual SEG meeting, Las Vegas, S6.5

YEDLIN M. J., JONES I. F., NAROD B. B. 1987: Application of the Karhunen-Loève transform to diffraction separation. IEEE Transactions on acoustics, speech, and signal processing. ASSP-35, 1

A KARHUNEN-LOÈVE ÉS A HAGYOMÁNYOS STACKING ELJÁRÁS ÖSSZEHASONLÍTÁSA

Leif BRULAND

A szeizmikában a Karhunen-Loève (KL) transzformáció számos alkalmazása ismert, ezek közül az első főkomponensnek alternatív összegzésként való alkalmazása a KL stacking. A KL stackinget és a hagyományos stackinget elemezve és összehasonlítva, megállapítottuk, hogy a KL stackinget a zaj, különösen pedig a koherens zaj jobban befolyásolja, mint a hagyományos stackinget.

Csatornáról csatornára haladva közel azonos jelamplitúdók mellett ezért a hagyományos stacking a jobb választás. Másrészt azonban, ha a jel amplitúdója változik, és a zaj minden csatornán közel azonos energiájú és korreláltalan, a KL stackinget kellene előnyben részesíteni.

Azt állították, hogy a KL stacking viszonylag érzéketlen a jelek kismértékű időbeli eltolódásaira, és a maradék statikus korrekció KL stacking esetében felesleges. Megerősítjük, hogy ilyen esetekben valóban a KL stacking adja a jobb jel/zaj viszonyt. Azonban, az időbeli eltolódások a kimenő jelet lényegesen torzíthatják, a torzulás nagyon érzékeny az időbeli eltolódásokra, és mindezekre való tekintettel, fontos, hogy a KL stacking esetében is végrehajtsuk a maradék statikus korrekciót.