

UTILISATION DE LA TRANSFORMATION DE KARHUNEN-LOÈVE COMBINÉE AVEC UN FILTRE DE POLARISATION. APPLICATIONS EN SISMIQUE DE SURFACE ET EN SISMIQUE DE PUIITS.

Mohammed BENCHIKH * et Mabrouk DJEDDI**

RÉSUMÉ

La transformation de Karhunen-Loeve, connue aussi comme analyse en composantes principales, est une technique d'analyse du signal utilisée en traitement de l'image. Elle extrait l'information cohérente d'une manière optimale à partir des données d'entrée au sens des moindres carrés.

Appliquée à la sismique, elle permet de séparer convenablement les différents événements sismiques cohérents quand le bruit de fond n'est pas significatif; c'est le cas en sismique de puits où elle permet d'isoler les ondes montantes et en sismique de surface pour l'amélioration du rapport signal sur bruit avec notamment l'atténuation des réflexions multiples.

Sensible à la présence de bruit à forte énergie, et partant de la qualité des résultats, cette méthode a nécessité un prétraitement basé sur un filtrage de polarisation obtenu à partir des matrices d'inertie.

L'action simultanée des filtres de Karhunen-Loeve et de polarisation a permis d'occulter cette difficulté et d'apporter une amélioration à la qualité des résultats sur des données synthétiques et laissant prévoir une application prometteuse sur des données sismiques réelles.

Mots clefs - Transformation de Karhunen-Loeve - Composantes principales - Evénements cohérents, Bruit de fond - Filtrage de polarisation.

* Institut Algérien du Pétrole (I.A.P- Sonatrach), Boumerdes, 35 000 (Algérie).

** Université de Boumerdes (Algérie).

- *Manuscrit déposé le 05 Janvier 2000, accepté après révision le 07 janvier 2001.*

USE OF THE KARHUNEM-LOEVE TRANSFORM COMBINED WITH A POLARIZATION FILTERING. APPLICATION TO SURFACE SEISMIC AND IN WELL SURVEY.

ABSTRACT

The Karhunen-Loeve transform also known as analysis in main components, is a technique of wavelet processing used in image processing.

It extracts coherent data in an optimum way from input data in the sense of least squares.

Applied to seismics, it allows a convenient separation of the different coherent seismic events when the ground noise is not important. It is the case in well seismic where it may isolate upgoing waves and in surface seismics it may improve the signal to noise ratio particularly with the reduction of secondary reflections.

This method, which is sensitive to the presence of high energy noise, and hence to the quality of the results, has required a pre-processing based upon a polarization filtering obtained from inertial matrices.

The simultaneous effect of K-L transform and polarization filtering has led to the improvement of the quality of results on synthetic data which can be a promising application on seismic data.

Key words - Karhunen-Loeve transform - Main components - Coherent events - Ground noise- Polarization filtering.

I - INTRODUCTION

Ces dernières années, une large variété de méthodes ont été utilisées pour séparer les différents types d'ondes en sismique. Ces méthodes de filtrage ont été classées (Glangeaud et Mari, 1993) en trois familles: les méthodes à gabarit (filtrage f-k et représentation tau-p), les méthodes d'inversion (méthode paramétrique) et les méthodes matricielles (méthode interspectrale, méthode Décomposition en Valeurs Singulières DVS et méthode Karhunen-Loeve K-L). C'est cette dernière méthode, la transformation Karhunen-Loève, qu'on s'est proposé d'élucider dans cet article.

L'objet de la sismique étant la recherche de la réponse impulsionnelle du sous-sol, le principe est donc d'envoyer un signal aussi bref que possible pour caractériser au mieux cette réponse.

Par la suite, le problème se pose en terme de filtrage. On ne s'intéresse alors qu'au signal utile (les réflexions primaires) et on élimine les autres événements considérés comme bruit à l'occasion. Pour cela, différentes techniques sont utilisées pour extraire l'information utile.

Traditionnellement on applique la couverture multiple, les filtres f-k, les filtres de cohérence, etc...

Nous considérons la transformation Karhunen-Loève comme étant une méthode qui vient s'ajouter à la famille des filtres utilisés en sismique conventionnelle.

C'est une technique d'amélioration du signal utilisée en traitement d'image. Elle extrait l'information cohérente d'une façon optimale à

partir de données d'entrée au sens des moindres carrés. Elle a la particularité de concentrer les informations d'une fonction aléatoire sur un nombre réduit de fonctions déterministes.

Dans cet article, on donnera une approche sur les composantes principales. On présentera quelques exemples sur des données de surface synthétiques et réelles ainsi que sur des données de puits.

La méthode devenant inefficace en présence d'un bruit de fond important, on a appliqué un prétraitement basé sur le filtrage de polarisation obtenu à partir des matrices d'inertie .

II - POSITION DU PROBLÈME

On considère une matrice X qui représente un film constitué par un ensemble de points miroirs de p traces à n échantillons par trace sur lesquelles les corrections statiques et dynamiques ont été réalisées.

$$X(p,n) = \begin{vmatrix} x_{11} & \dots & x_{1j} & \dots & x_{1n} \\ \cdot & & & & \\ x_{i1} & \dots & x_{ij} & \dots & x_{in} \\ \cdot & & & & \\ x_{pn} & \dots & x_{pj} & \dots & x_{pn} \end{vmatrix}$$

Soit N le nuage de points formé par les points x_{ij} dans l'espace à p dimensions. On suppose les variables centrées c'est à dire l'origine O et le centre de gravité g du nuage de points confondus (fig 1.).

Le but est de trouver la droite (u) la plus proche de l'ensemble des points au sens de l'erreur quadratique moyenne . En d'autre terme , la droite (u) est celle qui maximise la somme des projections des points du nuage.

Dans le cas où on a un seul vecteur x_j le vecteur \vec{u} optimisant cette projection sera colinéaire à X_j.

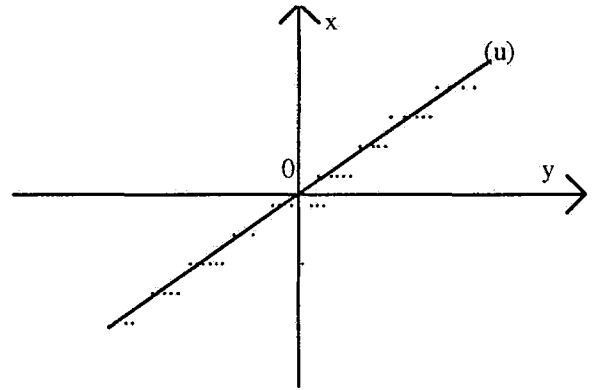


Fig. 1 - Cas de la dimension 2
Case of the 2 D

Si toutes les composantes x_j de X_j sont identiques, \vec{u} définit la bissectrice.

Si on est en présence de bruit , la bissectrice n'a plus de position privilégiée.

La droite (u) qui optimise la projection au sens des moindres carrés est la direction propre de cette forme quadratique associée à la plus grande valeur propre.

Il est donc nécessaire de calculer toutes les valeurs propres λ_i de cette forme quadratique ainsi que les vecteurs propres qui leur sont associées.

Le vecteur \vec{u} associé à λ_1 qui est la plus grande valeur propre est appelé premier facteur principal. Les valeurs propres sont rangées par ordre décroissant :

$$\lambda_1 \geq \lambda_2 \geq \lambda_3 \dots \geq \lambda_p$$

La valeur $Y_{kj} = u_k^t X_k^j$ est appelée k ième composante principale.

Si on désigne par :

T : la matrice des vecteurs propres et T la transposition, on déduit que:

$$Y_j = T^T X_j \text{ avec } 1 \leq j \leq n$$

En faisant varier j , on obtient le tableau Y des composantes principales qui est la transformée du tableau X initial. La transformation ainsi réalisée est celle de Karhunen-Loève.

Il y'a une autre façon de définir la transformation K-L. L'approche de Kramer (1956) donnée par Ulyrch et Jones (1987) est révélatrice. Le problème est posé de la façon suivante: on se donne un ensemble de n signaux réels $x_i(t)$ ($i=1, \dots, n$), on définit l'ensemble transformé $y_j(t)$ et la matrice de transformation A telle que (Jones et Levy, 1987):

$$\psi_j(t) = \sum_{i=1}^n a_{ij} x_i(t); \quad j=1, \dots, n$$

où a_{ij} sont les éléments de la matrice A . Les signaux $\psi_j(t)$ sont choisis de telle sorte qu'ils forment une base orthogonale, de sorte que chaque signal $x_i(t)$ peut être exprimé ainsi:

$$x_i(t) = \sum_{j=1}^n b_{ij} \psi_j(t), \quad i=1, \dots, n. \quad (1)$$

ou approximativement ainsi:

$$x'_i(t) = \sum_{j=1}^m b_{ij} \psi_j(t), \quad i=1, \dots, m; \quad m < n. \quad (2)$$

où $x'_i(t)$ est le i ème signal conjugué reconstruit, b_{ij} sont les éléments de B , la matrice de transformation inverse, et m est le nombre de fonctions de base utilisées dans le développement tronqué.

L'objectif est de reconstruire $x_i(t)$ en utilisant le plus petit nombre possible de signaux de base.

Pour un m donné les matrices de transformation A et B doivent minimiser l'erreur quadratique moyenne:

$$\phi(m) = \sum_{i=1}^n \int_0^T (x_i(t) - x'_i(t))^2 dt \quad (3)$$

Kramer et Matthews (1956) ont montré que les colonnes de la matrice A sont composées des vecteurs propres normalisés de la matrice de covariance Γ définie comme suit:

$$\gamma_{ij} = \int_0^T x_i(t) x_j(t) dt$$

où γ_{ij} sont les éléments de Γ .

La matrice de covariance Γ est symétrique, semi-définie positive et donc décomposable :

$$\Gamma = R \Lambda R^T$$

avec $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ avec $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$.

Les colonnes de R contiennent les vecteurs propres normalisés, où $\Gamma_{\sigma} = \lambda_{\sigma} r_{\sigma}$.

En utilisant les éléments de base définis dans (2), l'erreur de troncature à partir de (3) est:

$$\phi(m) = \sum_{j=m+1}^n \lambda_j$$

Comme les valeurs propres de Γ sont arrangées dans un ordre décroissant, il résulte que la première fonction de base peut être utilisée pour reconstruire davantage d'énergie du signal total qu'aucune autre fonction de base. Elle est définie comme la 1ère composante principale. De même, la 2^{ème} fonction de base est comme la 2^{ème} composante principale, etc.

La transformation Karhunen-Loeve est une transformation qui garde l'énergie totale de l'enregistrement constante. La somme des valeurs propres de la matrice d'autocorrélation représente l'énergie totale des données.

Dans les transformations KL, les éléments qui composent la matrice de transformation et qui sont les vecteurs propres, sont obtenus à partir des données d'entrée, contrairement aux autres

transformations où les vecteurs de base sont fixes et les données indépendantes.

La valeur de la i -ème valeur propre est la mesure de la quantité d'énergie présente dans la i -ème composante principale. Par conséquent, reconstruire les signaux originaux en n'utilisant que les composantes principales associées aux plus grandes valeurs propres équivaut à reconstruire l'énergie cohérente présente dans les données originales.

Inversement, reconstruire les données originales en utilisant les composantes principales associées aux plus petites valeurs propres équivaut à reconstruire la part d'énergie la moins cohérente des données originales.

Si les données originales $x_i(t)$ sont identiques, à un facteur près, alors toutes les valeurs propres de Γ sont nulles à l'exception de la première (λ_1).

Si les données originales ne présentent aucune cohérence d'une trace à une autre, la matrice d'autocorrélation (Γ) est diagonale et la décroissance des valeurs propres est régulière. Dans ce cas, la méthode devient peu efficace dans la mesure où l'information n'est pas concentrée sur les premières composantes mais répartie sur l'ensemble des composantes.

III - FILTRE DE POLARISATION.

En présence d'un taux de bruit important, la méthode semble inefficace. Ceci est dû au fait qu'on n'est plus en présence d'un ellipsoïde mais plutôt d'une sphère dans l'espace (Benhama, 1984). La direction de la droite la plus proche représentant les données d'entrées devient difficile à chercher. Néanmoins et afin de pallier cet inconvénient, on a appliqué un filtre utilisant la définition de l'inertie et qui est basé sur le taux de polarisation donné par la relation suivante:

$$\tau = \frac{n(\text{trace}(c))^2 - (\text{trace}(c))^2}{(n-1)(\text{trace}(c))^2} \quad (\text{Samson-1973})$$

avec:

(c) = matrice de covariance .

n = ordre de couverture ou dimension du système.

trace (c) = somme des éléments de la 1^{ère} diagonale.

Le but principal de ce filtre est d'atténuer les événements non cohérents et d'amplifier relativement ceux qui présentent une corrélation.

Le principe est de calculer la trace de chaque matrice représentant une partie des données d'entrées et le coefficient de pondération t . Par la suite chaque échantillon considéré au milieu de la fenêtre est multiplié par le coefficient de pondération t . Pour éviter les effets de bord les échantillons avant le milieu de la première plage et après le milieu de la dernière plage, ont été remis à zéro.

IV - APPLICATIONS ET RÉSULTATS.

1 - Application sur des données synthétiques.

Afin de voir l'efficacité de la méthode, on a effectué une étude sur des données synthétiques pour pouvoir comparer entre la trace couverture multiple et la première composante principale.

Pour cela, on s'est donné un film de 24 traces regroupées en point miroir commun (CDP gather) à 256 échantillons chacune, sur lesquelles les corrections statiques et dynamiques ont été réalisées. Le pas d'échantillonnage est de 4 ms.

On a considéré 4 réflecteurs situés respectivement à 220 ms, 420 ms, 620 ms et 820 ms.

On a utilisé une ondelette de Ricker de fréquence centrale 25 hz pour générer ces réflecteurs. Dans un premier temps, on a fait varier le taux

de bruit. Par la suite, on a simulé de mauvaises corrections statiques en restant dans les normes tolérées dans l'industrie, à savoir une fourchette de 20 ms. Ceci étant, on a pu voir comment la transformation KL peut s'affranchir des corrections statiques résiduelles. Les deux méthodes se basent sur le principe de la cohérence.

Pour chaque cas, la démarche suivante a été adoptée:

- 1- Génération des données d'entrées par la simulation d'un modèle synthétique simplifié (corrections dynamiques supposées réalisées).
- 2- Calcul de la matrice d'autocorrélation des traces du modèle.
- 3- Calcul des valeurs et vecteurs propres de la matrice d'autocorrélation.

4- Calcul des composantes principales.

5- Comparaison de la première composante principale avec la trace sommée en couverture multiple.

La figure 2 montre que la transformation KL représentée par les composantes principales (1^{ère} cp et 2^{ème} cp) donne de meilleurs résultats par rapport à la sommation en couverture multiple (CM) quand le bruit est faible (10%). Dans cet exemple (fig.2), les traces sont bien corrigées statiquement ce qui correspond à un décalage nul. Avec un bruit relativement important (90%, fig.3), la sommation en couverture multiple (CM) est préférée à la transformation KL (1^{ère} cp). Ici, on a décalé les traces les unes par rapport aux autres dans une fourchette de 20 ms.

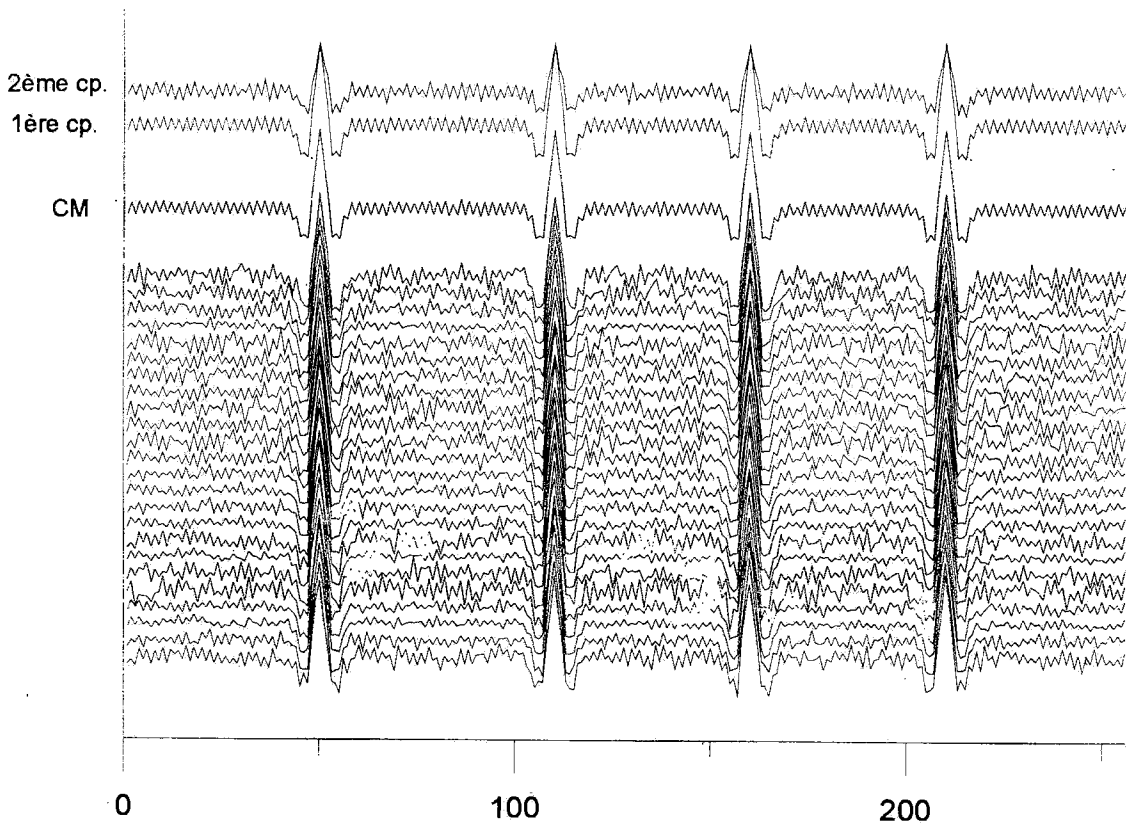


Fig. 2 - Modèle synthétique représentant 24 traces, leur addition et les deux premières composantes principales.
Taux de bruit : 10%. Décalage nul

*Synthetic model representing 24 channels, their stack and their two first mean components.
Noise ratio : 10%*

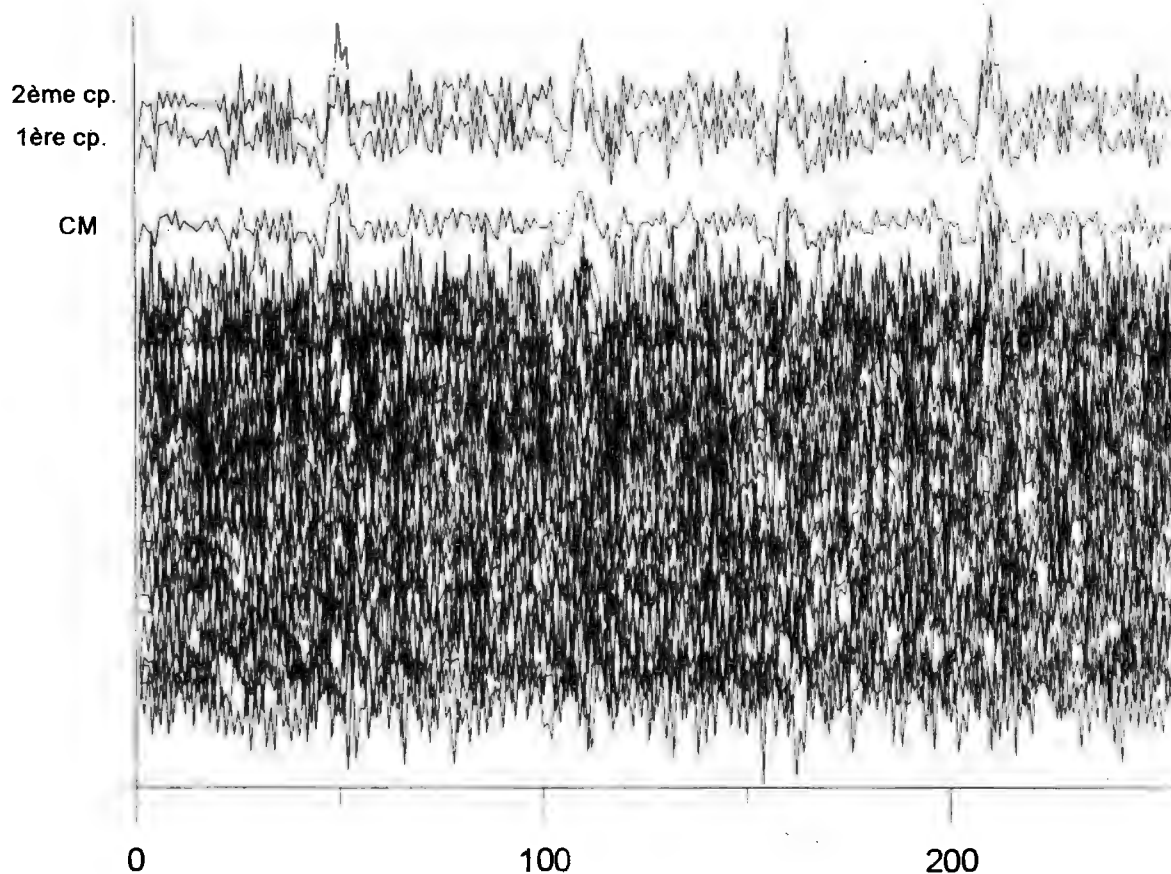


Fig. 3 - Modèle synthétique représentant 24 traces décalées, leur addition et les deux premières composantes principales. Taux de bruit : 90%

Synthetic model representing 24 channels removed, their stack and their two first mean components. Noise ratio : 90%

La figure 4 donne une comparaison des résultats obtenus en utilisant les différents filtres: couverture multiple (CM), polarisation par inertie (PO), transformation KL (TKL) et transformation KL après polarisation (KLPO). on constate une nette amélioration apportée par le filtre par polarisation. On remarque bien l'orthogonalité des deux premières composantes. Pour l'application du filtre par polarisation, on a adopté le principe de la fenêtre glissante. Pour chaque fenêtre d'une longueur de 64 ms (16 échantillons), on a calculé le taux de

polarisation qui sera multiplié par l'amplitude de l'échantillon situé au milieu de la fenêtre. Par la suite, on fait glisser cette fenêtre d'un pas de 4 ms (correspondant à 1 échantillon). τ^n a pour pondération n .

La figure 5 illustre bien la difficulté d'appliquer KL avec un bruit trop élevé. On remarque que le filtre par polarisation a permis d'identifier légèrement les réflexions (PO et KLPO) alors que les résultats ne sont pas bons avec la transformation KL.

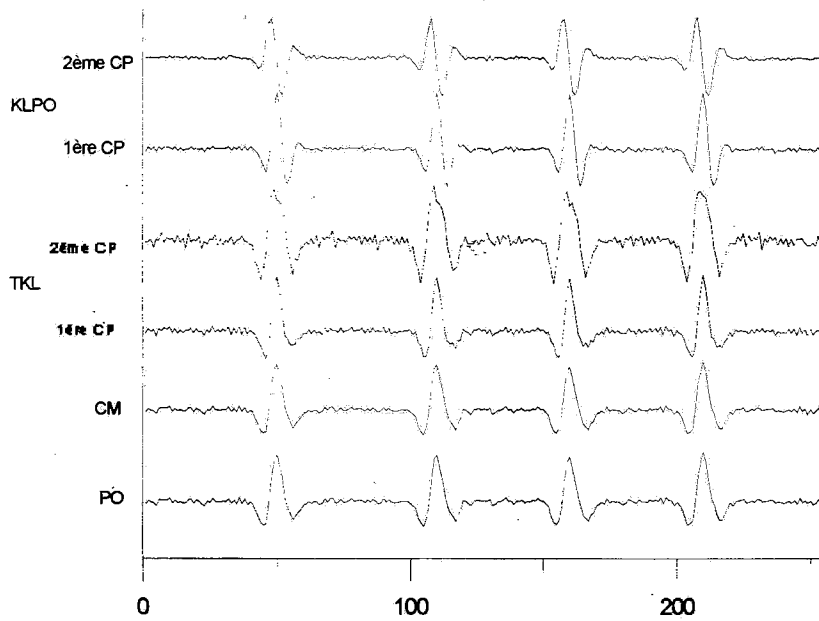


Fig. 4 - Comparaison des résultats avec différents filtres

Comparison of results with different filters

KLPO : transformation de Karhunen-Loève (TKL) après polarisation (PO) (*Karhunen-Loève transform (TKL) after polarization (PO)*); CM : couverture multiple. Fenêtre: 64 ms, bruit:10%, pondération:1 (*Commun Depth point stack. Window: 64ms, noise: 10%, ponderation:1*).

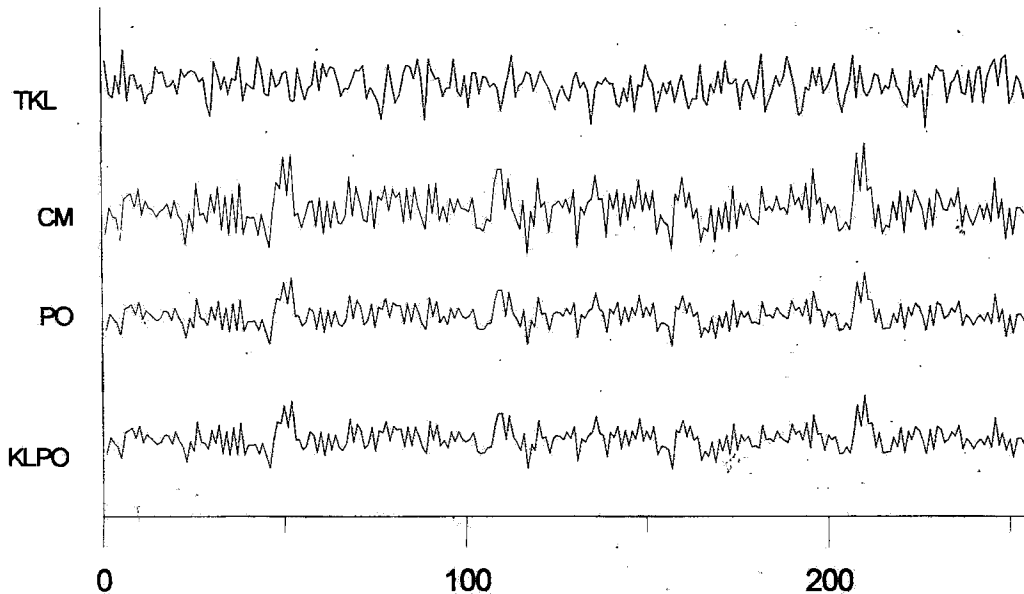


Fig. 5 - Comparaison des résultats avec différents filtres

Comparison of results with different filters

KLPO : transformation de Karhunen-Loève (TKL) après polarisation (PO) (*Karhunen-Loève transform (TKL) after polarization (PO)*); CM : couverture multiple. Fenêtre: 64 ms, bruit:90%, pondération:1 (*Commun Depth point stack. Window: 64ms, noise: 90%, ponderation:1*).

2 - Application sur des données réelles.

Un ensemble de 24 points miroirs (CDP) composé chacun de 24 traces nous a été fourni par le Centre de calcul de la Division Exploration de l'entreprise Sonatrach. Ces traces sont corrigées statiquement et dynamiquement.

On a choisi un CDP et on lui a appliqué les différents filtres (fig. 6). L'amélioration ici est surtout visible sur les résultats donnés par la TKL après polarisation (PLK).

Enfin la fig. 7 montre une différence notable entre CM et TKL après filtrage par polarisation. Ceci est particulièrement visible sur la fin de la trace.

Concernant les données de puits, La méthode a été appliquée pour séparer les ondes montantes des ondes descendantes. Sur les 80 traces mises à notre disposition, seules 40 traces ont été considérées. Le but consiste à horizontaliser les ondes montantes puis à les

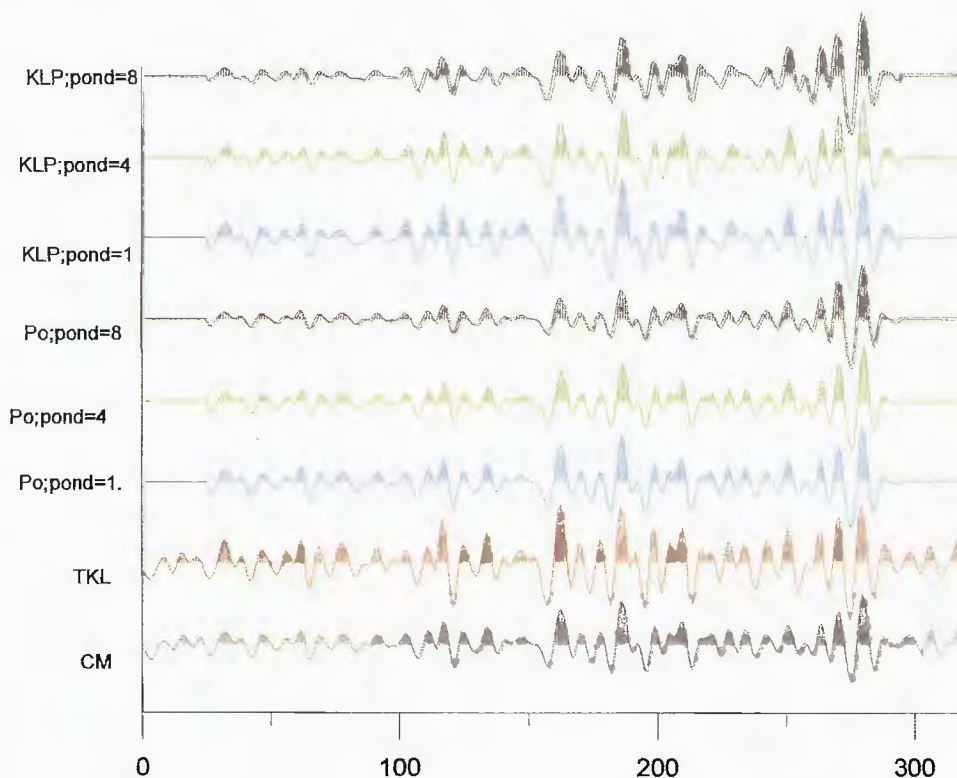


Fig. 6 - Comparaison entre les différents filtres. Fen = 200ms

Comparison between different filters

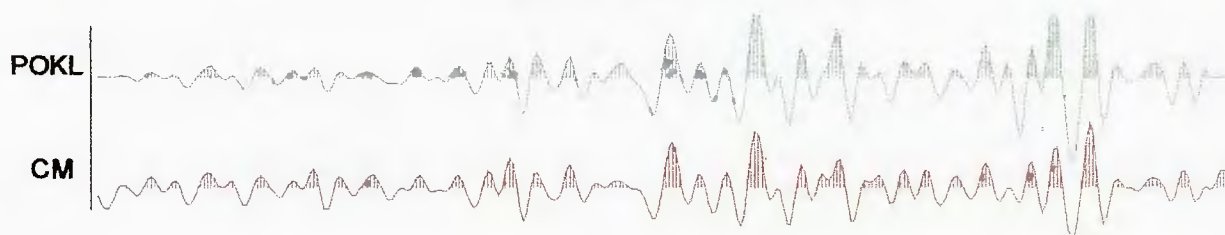


Fig. 7 - Comparaison entre le filtre CM et KL après Poline; Fen = 50ms; pond = 1

Comparison between stack filter and KL method; win : 50ms; pond = 1

additionner. Les figures 8, 9 et 10 représentent des données de puits réelles où on constate une nette amélioration apportée par le filtrage par polarisation (fig 10). Enfin sur la figure 11, on constate le déplacement du spectre vers l'axe des ampli-

tudes ce qui démontre la concentration de l'énergie sur les premières valeurs propres dans le cas où on applique le filtre de polarisation. Par contre, la décroissance des valeurs propres est régulière quand les données ne sont pas filtrées. [G1]

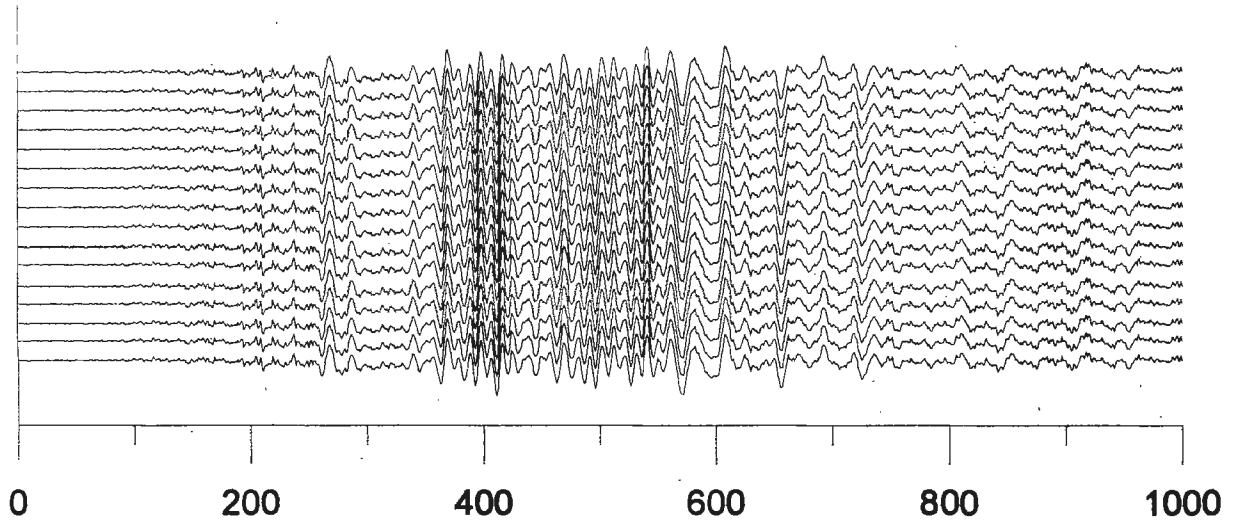


Fig. 8 - Ondes montantes additionnées
Upping waves stacked

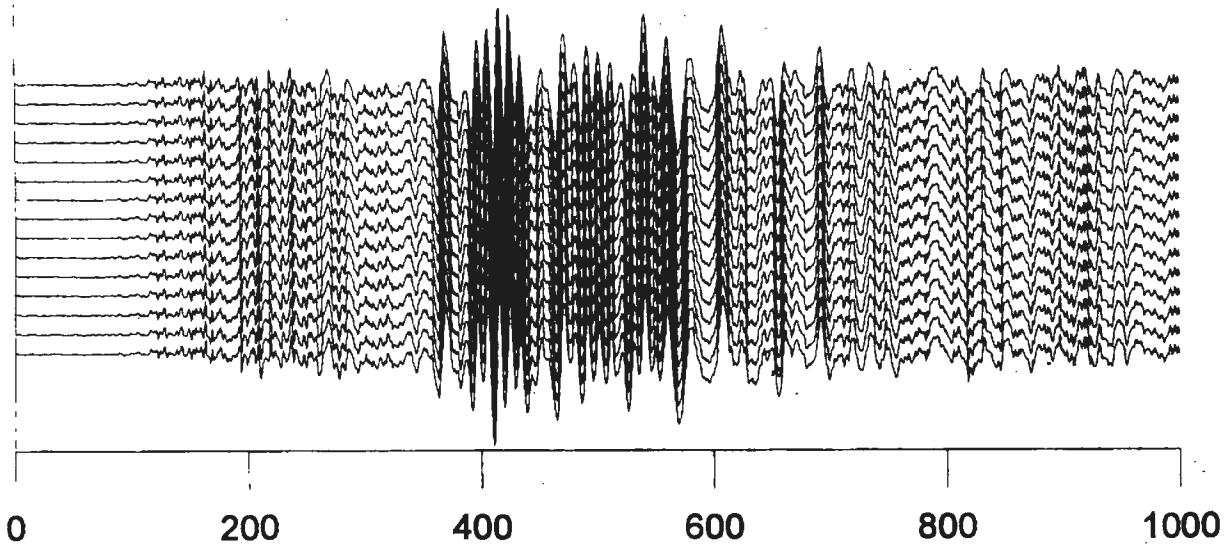


Fig. 9 - Ondes montantes traitées par la méthode KL
Upping waves processed by KL method

UTILISATION DE LA TRANSFORMATION DE KARHUNEN-LOÈVE COMBINÉE AVEC UN FILTRE DE POLARISATION

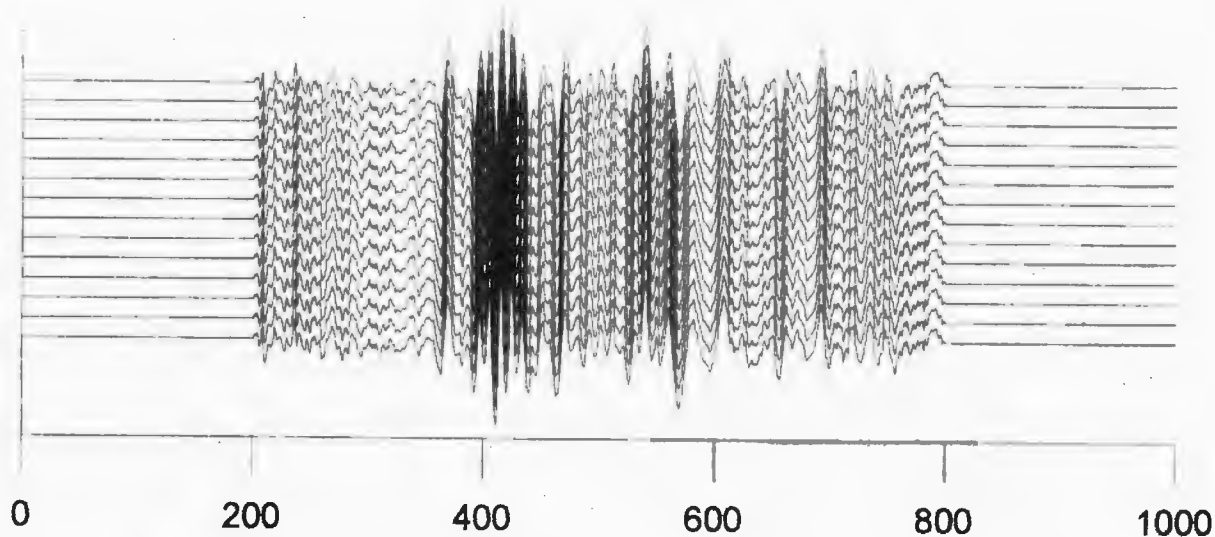


Fig. 10 - Ondes montantes traitées par la méthode KL après le filtrage par polarisation
Upping waves processed by KL method filter after polarisation filtering

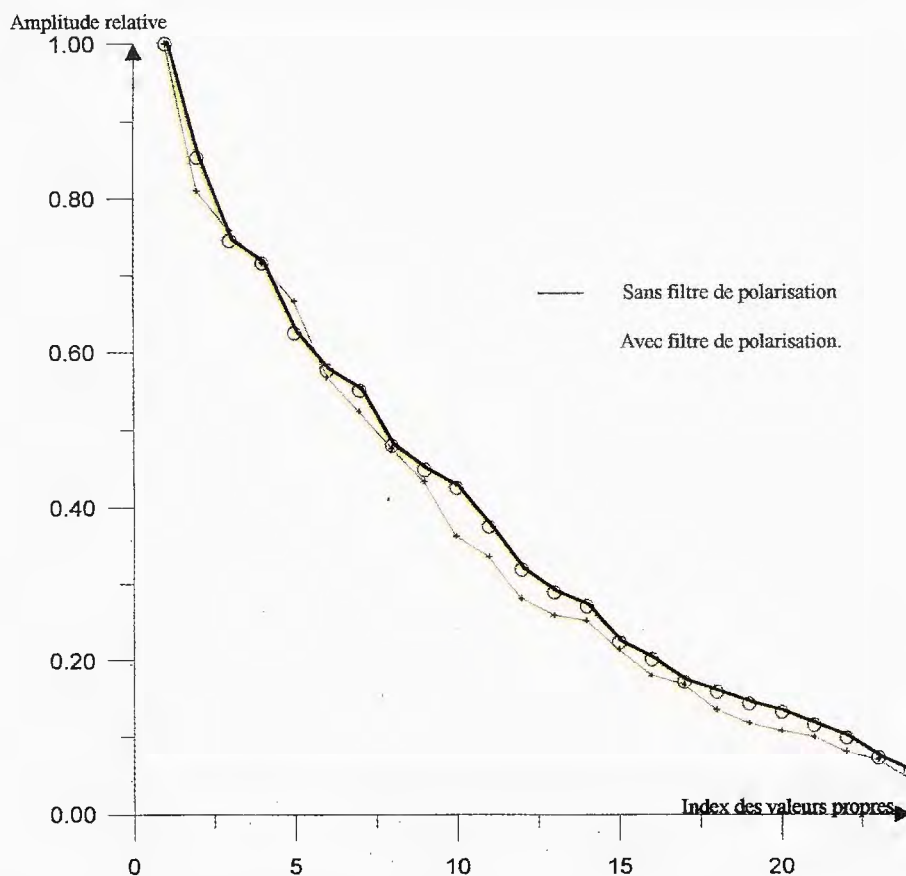


Fig. 11 - Spectre des valeurs propres d'un point miroir commun traité par le filtre de polarisation; fenêtre: 100 ms
Eigen values spectrum of CDP processed by the polarization filter; Amplitude relative window: 100 ms

V - CONCLUSION

La transformation de Karhunen-Loève, connue aussi comme analyse en composantes principales, est un outil puissant pour séparer les ondes montantes des ondes descendantes dans le cas d'un profil sismique vertical ou pour augmenter le rapport signal sur bruit en sismique de surface, surtout quand le niveau de bruit n'est pas trop élevé. Malgré les résultats satisfaisants qu'elle donne, elle bute toujours sur deux obstacles majeurs:

- Le temps de calcul relativement élevé, ce qui a amené les analystes à lui préférer d'autres transformations rapides telle que la transformation de Fourier.

- Son impuissance à extraire les événements incohérents, autrement dit à supprimer le bruit aléatoire quand ce dernier devient trop important.

Si le premier problème est supposé résolu puisque l'on dispose actuellement d'algorithmes performants et de machines assez puissantes, il n'en demeure que le bruit de fond reste toujours un handicap pour pouvoir utiliser cette technique comme étape préliminaire en traitement sismique.

Nous avons voulu pallier cet inconvénient en appliquant le filtre de polarisation, construit à partir des matrices d'inertie, avant même de filtrer par la transformation de Karhunen-Loève. Les résultats sur des données synthétiques en sismique conventionnelle sont bons et ceux relatifs aux données réelles en sismique de puits (P.S.V) semblent prometteurs.

Il est bien évident que, sur le plan économique et pour des résultats identiques, on préfère la couverture multiple à la transformation de Karhunen-Loève. Néanmoins, la méthode peut être appliquée à des cas particuliers, là où les autres filtres donnent des résultats peu probants. Elle peut être utilisée sur une portion de section de mauvaise qualité, un PSV, etc...

Bibliographie

- BENCHIKH, M., 1993.** Applications de la transformée de Karhunen - Loève au traitement sismique, *Thèse de Magister en géophysique appliquée* - Université de Boumerdès.
- BENCHIKH, M., 1998.** De l'utilité d'un filtre de polarisation en sismique réflexion. *Thèse de Docteur de l'Université Pierre et Marie Curie* (Paris 6),
- BENHAMA, A., 1984.** Le mouvement réel des particules en sismique: polarisation détection automatique des arrivées, *Thèse de Docteur-Ingénieur de l'Univ. Pierre et Marie Curie* (Paris 6).
- FREIRE, S.L. AND ULRICH, T.J., 1988.** Application of singular value decomposition to vertical seismic profiling, *Geophysics* 53, 778-785.
- GLANGEAUD, F. ET MARI, J.L., 1993.** Wave separation, *Edition Technip*- Institut Français du Pétrole.
- HEMON, CH. ET MACE, D, 1978.** Essai d'une application de la transformation de Karhunen-Loève au traitement sismique, *Geophysical Prospecting* n° 26 page 600-620,
- JONES, I.F. AND LEVY, S., 1987.** Signal to noise ratio enhancement in multichannel seismic data via the Karhunen-Loève transform, *Geophysical Prospecting* n° 35, 12-32.
- MARI, J.L. ET COPPENS, F., 1989.** La sismique de puits, *Editions technip*-Institut Français du Pétrole.
- OPPENHEIM, A.V, 1978.** Applications of Digital Signal Processing, Prentice-Hall.
- SAMSON, J.C. 1973.** Descriptions of the polarization states of vector processes: applications to ULF magnetic fields. *Geophys.J.R.Astr.Soc.*, n° 34, 403-419.
- ULTRYCH ET JONES, 1987.** Use of the Karhunen-Loève transformation in seismic data processing, *Geophysical Prospecting*, 26, p 600-625.