2001

Use of Principal Component Analysis to Determine Down-Hole Tool Orientation and Enhance SH-Waves

Paul Michaels
Boise State University

Use of Principal Component Analysis to Determine Down-hole Tool Orientation and Enhance SH-Waves

P. Michaels
Center for Geophysical Investigation of the Shallow Subsurface, Boise State University, 1910 University Drive, Boise, Idaho 83725
pm@cgiss.boisestate.edu

ABSTRACT

A common problem in down-hole shear wave surveys is the determination of the tool rotation relative to the seismic source polarization direction. This paper reports an algorithm which has been developed to solve for this angle in the horizontal plane. The method employs Principal Component Analysis (PCA) to determine the angle from the large motion of the SH-wave (not first motion). Once the angle is known, the horizontal component data may be rotated so that one component is aligned with the source polarization, and the other orthogonal to it.

Significant findings include the following. First, subtraction of oppositely polarized source efforts improves the PCA formulation by enhancing the S-waves. Second, the Swave enhancement is improved by scaling the oppositely polarized source efforts by the vertical (not horizontal) component data, as this provides a better attenuation of non-Swave energy. Third, observation of the tool orientation as it exits the borehole is helpful in resolving the 180 degree azimuthal ambiguity in the PCA approach. Finally, the polarization of the source radiation may drift relative to the axis of the source during a survey, and should not be assumed invariant.

Introduction

Determination of subsurface seismic tool orientation can be achieved either by hardware design or by the inversion of observations taken from the recorded seismic data.

Hardware solutions have focused on the tool deployment mechanism. One example of a deployment strategy is the use of a rigid bar to lower the geophone down the borehole (Mok *et al.,* 1988). The operator manually orients the geophone by observing markings on the rigid bar. While useful in shallow situations, the technique becomes more difficult with deeper boreholes. Another approach has been to design a tool with electronic fluxgate compass sensors (Crice, 1996). The fluxgate transducer signal gives the tool orientation with respect to magnetic north. A servo-mechanism rotates the elements in the tool to a desired azimuth relative to magnetic North (specified by the operator at the beginning of the survey).

A hybrid technique which mixes hardware with observations of the seismic signal has been demonstrated on seismic cone penetrometer surveys (Brettmann *et al.,* 1996). The geophone signals are observed while repeatedly activating the source and rotating the push rod with a pipe wrench until a maximum signal amplitude is observed on the component to be aligned with the shear (S-wave) source.

While the above techniques have merit, they also have disadvantages. Deployment rods or manipulation of a push rod both slow down data acquisition. In the case of the downhole compass, the polarization direction of the source radiation may drift from a desired magnetic azimuth. Defects in the source construction (or use) and variations in the soil under the source can twist the actual radiation out of the intended polarization direction.

Using the seismic data itself to determine down-hole tool orientation has a number of advantages. First, data acquisition is fast, as no time is spent trying to orient the tool. Second, the orientation of the tool is relative to the actual (not intended) SH-wave polarization. Finally, the process can be automated with software, requiring only a moderate amount of manual intervention at points where the tool might have been released to clear an obstruction in the borehole.

Field Data Collection Method

The specific details of this PCA method are linked to the author's data collection protocols. The author conducts down-hole vertical seismic profile (VSP) surveys for dynamic soil properties using a repeatable hammer source that delivers blows at 135 degrees from the vertical (Michaels, 1998). Figure 1 shows a schematic drawing of the source which is nailed to the soil (no hold-down weight other than the 20 kg mass of the source itself). Additional clamping to the soil is provided by transverse angle iron cleats on

JEEG, December 2001, Volume 6, Issue 4, pp. 175–183

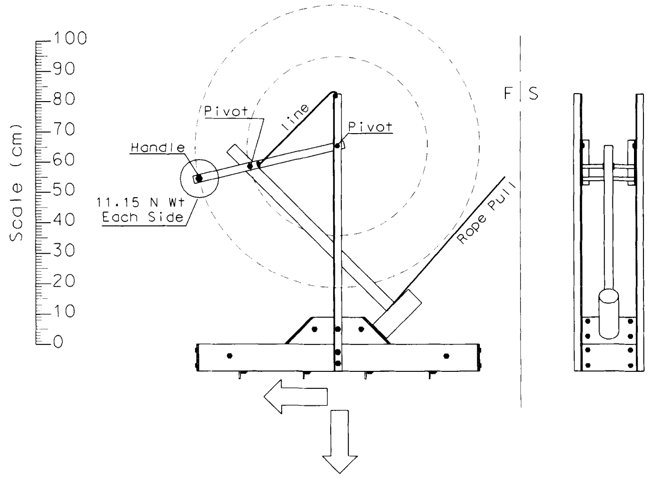

Figure 1. Front and side view of the seismic source. Hammer pivots to deliver blows of opposite horizontal polarity.

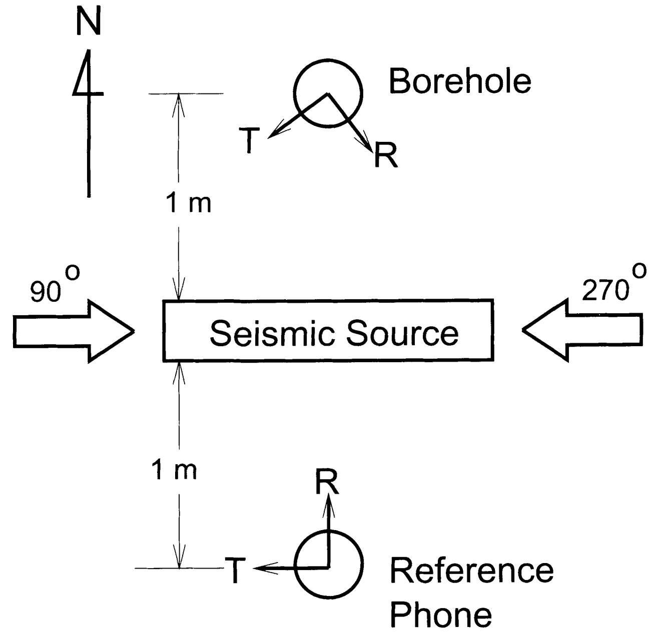

Figure 2. Plan view of a typical survey. Source and transverse (T-component) of reference geophone align East-West. PCA determines borehole geophone orientation.

the base, and the fact that a dynamic hold-down force is generated from the vertical component of the blow. The hammer pivots to deliver blows from opposite sides, permitting separate recordings of opposite initial motion for each geophone level occupied in the borehole. The radiated SH-wave polarization is generally aligned with the long axis of the source, but may become skewed depending on the exact point of impact. Further, small scale (relative to the source base) variations in soil stiffness may also result in a shear couple that alters the polarization direction.

Figure 2 shows the surface layout in plan view. The long axis of the source is orthogonal-to a line drawn from the borehole to the source position. For each down-hole station, two opposite blows are recorded (arrows labeled 270° and 90° indicate the azimuth of the impact). A fixed reference, 3-component geophone remains planted at the surface to monitor variations in the source waveform, and its orientation is as shown, with the transverse component being defined parallel to the long axis of the source. This reference geophone is usually covered with soil to reduce airwave interference. Also shown are the two horizontal components (transverse, T-component and radial, R-component) of the down-hole tool in one of many possible orientations. Determination of the specific orientation downhole is the object of the PCA analysis.

Data collection begins by lowering the tool (Crice, 1996, BHG-2 without a compass) to the bottom of the hole, and then clamping the tool with the electric motor driven bowspring clamp. The tool is then dragged up the hole in a clamped condition to occupy recording stations at 0.25 m intervals. As the tool is dragged up the hole, it may rotate slowly, depending on cable twist and variations in the borehole shape, but being clamped, does not spin freely. Thus, we have the expectation that if the tool changes orientation, it does so only gradually, unless the clamp is released to clear an obstruction.

Upon exiting the hole, the bow-spring orientation is observed and recorded. This provides a control point needed to resolve the 180 degree ambiguity inherent in the PCA analysis.

Field data shown in this paper are taken from downhole surveys acquired at the Boise Hydrologic Research Site (BHRS). The BHRS is located on a gravel bar adjacent to the Boise River. The soil profile is generally sand and gravels from the surface to a depth of about 20 m. The water table is typically at about 2 meters below the surface. To permit the introduction of hydrologic instrumentation, boreholes were completed with 0.1016 m (4 inch) PVC pipe (20 slot screen below a depth of 1.5 m). The surrounding soil was allowed to collapse against the casing as an alternative to backfill. Further details on the site may be found in (Barrash *et al.,* 1999).

Data Processing Method

The mathematical details of the PCA method are given in Appendix A. In short, at each down-hole station one computes a covariance matrix from the large amplitude motion recorded on the horizontal geophone elements in the borehole. The eigenvector associated with the largest eigenvalue gives the direction of the major axis of the SH-

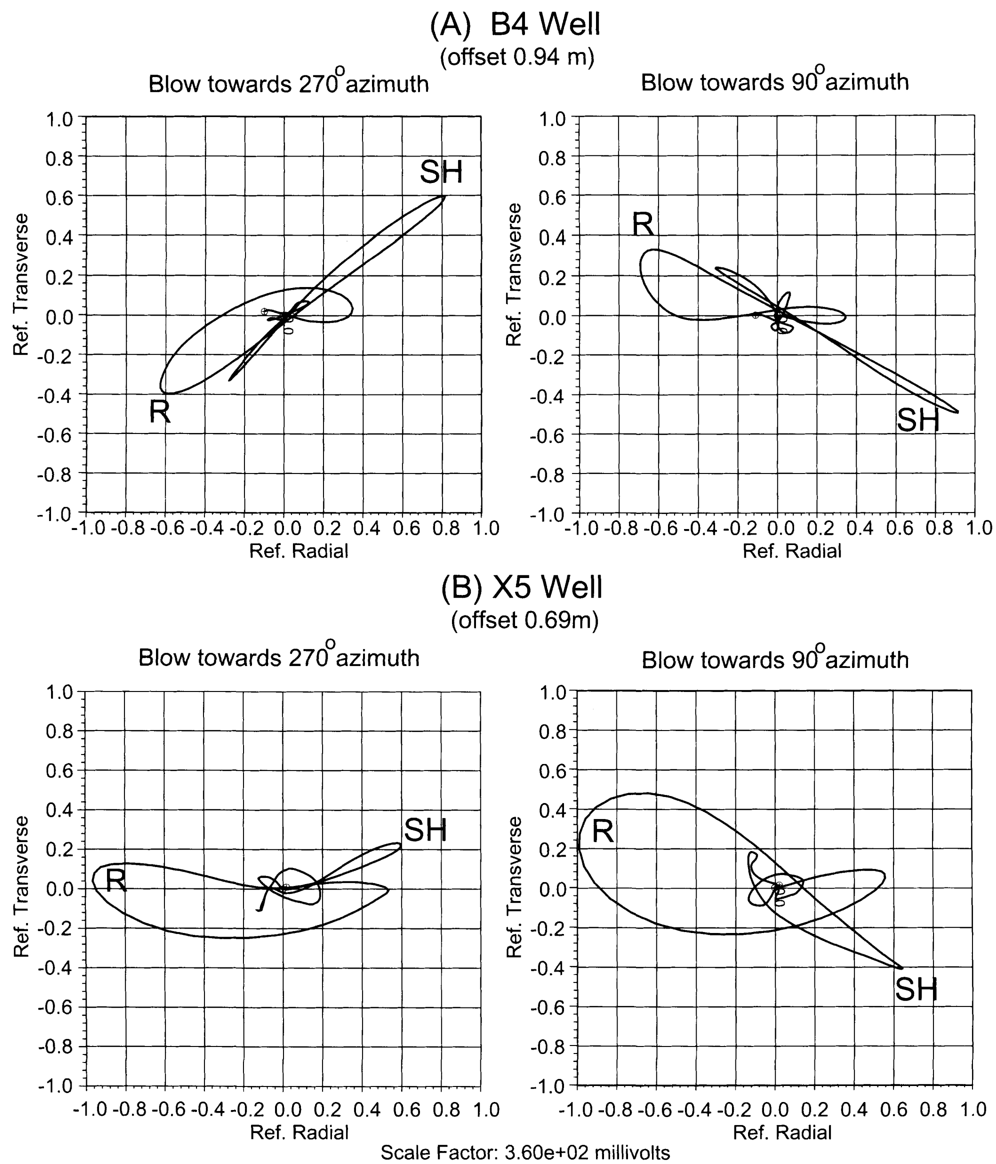

Figure 3. Hodograms from two different boreholes. The B4 hole (A) shows a strong SH-wave. At the X5 hole (B), the Rayleigh wave dominates.

wave polarization ellipse with respect to the coordinate system defined by the geophone elements.

The PCA problem is best posed when only SH-waves are recorded. However, the source generates a number of wavefields in addition to SH-waves (Rayleigh and P-waves for example). To enhance the SH-waves at each down-hole station, the two recordings of opposing first motion polarity are subtracted in the usual way. Since the SH-waves have opposite initial motion, subtraction enhances S-waves relative to P- and Rayleigh waves (which maintain a consistent initial motion).

Enhancement of the SH-wave improves with the degree to which the other waves are nulled out. Since the two recordings at each level may vary in amplitude, the stationary reference geophone is used to scale the data to a consistent peak amplitude before subtraction. Initially, the T-component of the reference geophone was used for scaling the data. While this was satisfactory in many cases, an improved strategy became necessary.

One problem is that the reference geophone is located in the near field of the source. Depending on minor changes in offset (and perhaps variation in soil conditions as well), P-SV motion can overwhelm the SH-wave.

Figure 3 illustrates the problem with hodograms in the horizontal plane taken from two different holes. A hodogram is a display of particle motion as viewed in a two dimensional plane. In case (A), the SH-wave dominates the motion in the horizontal plane. On the other hand, at a different borehole, as shown in (B), the Rayleigh wave dominated the horizontal motion. It became clear that focusing one's attention on the desired signal (SH-waves) was not the best approach.

Attention was then turned to the vertical component of the reference geophone as an alternative basis for scaling the two recordings. By scaling on the noise (Rayleigh waves), the strategy shifted from direct enhancement of signal to direct nulling of noise. That is, a more perfect nulling of noise results in an improved enhancement of the desired signal. Another advantage of using the vertical component of the reference geophone to scale the subtraction is that the vertical component is not sensitive to any horizontal rotation of the source radiation over the course of the experiment.

In summary, the scaling procedure is to scale all data channels (both down-hole and reference) by a single factor for the *i-th* shot record. The factor is

$$f_i = \frac{\max|S_n|}{\max|s_i(t)|} \tag{1}$$

where $\max|S_n|$ is the peak absolute value observed on the vertical reference geophone for the last (shallowest) shot record, and $\max|s_i(t)|$ is the peak absolute value on the vertical reference geophone for the i-th shot record to be scaled. All traces for the i-th record (both down-hole and reference geophones) are scaled by multiplication with the single normalizing factor, f_i.

Figure 4 shows the result of subtracting the opposing polarity records of Fig. 3 to null out the Rayleigh and P-wave energy. The enhancement of SH-waves occurs in both cases (a) and (b), despite the different balance between the Rayleigh and S-wave amplitudes for the two wells.

Demonstration of Method on Field Data

The BHRS site has 18 boreholes, all of which have been surveyed for S-wave properties. The following examples will serve to illustrate the PCA method and demonstrate the robust nature of the process.

The PCA algorithm displays a hodogram using the signals from the two horizontal geophones. The ideal particle motion for an SH-wave (assuming viscous damping is present) will be an ellipse with the major axis aligned with the source's long axis.

Figure 5 illustrates ideal hodograms for two extreme cases of possible subsurface geophone orientation. In (A), the transverse (T-component) is aligned with the source, and produces a particle motion history indicated by the ellipse. At the other extreme, (B), the radial (R-component) is aligned with the source, and the major axis of the motion ellipse is recorded on the "R".

Assuming the S-enhancement subtraction process has been performed, SH-wave amplitudes will be significantly

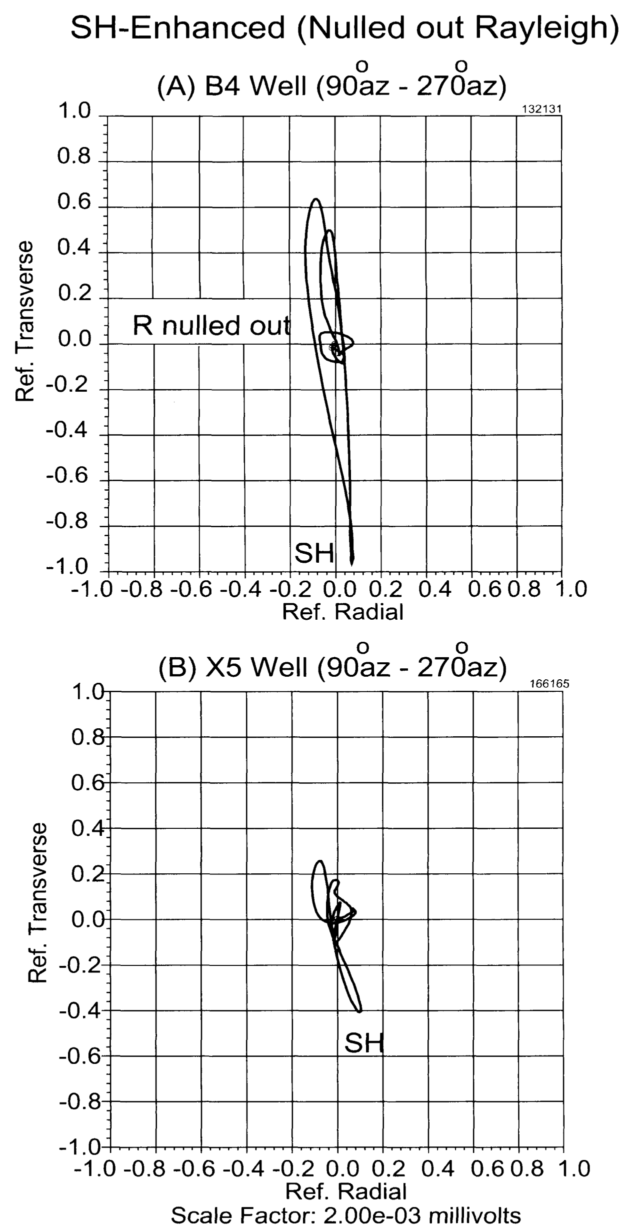

Figure 4. Result of subtracting the opposite polarity hodograms of Fig. 3. The resulting enhanced SH-waves are subjected to PCA.

larger than P- and Rayleigh wave amplitudes. The PCA algorithm permits the analysis of a large amplitude subset, and it is only these amplitudes which are shown on the following plots. A circle with a dot in the center plots the position of each large amplitude time sample.

Figure 6 presents the results for the X5 well. The PCA-determined azimuth is plotted for the T-component in (A). The horizontal geophone hodogram, (B), shows the motion for amplitudes larger than 50% of the peak motion recorded at the 12.5 m deep station. Superimposed on the hodogram is the solution for the T- and R-component azi-

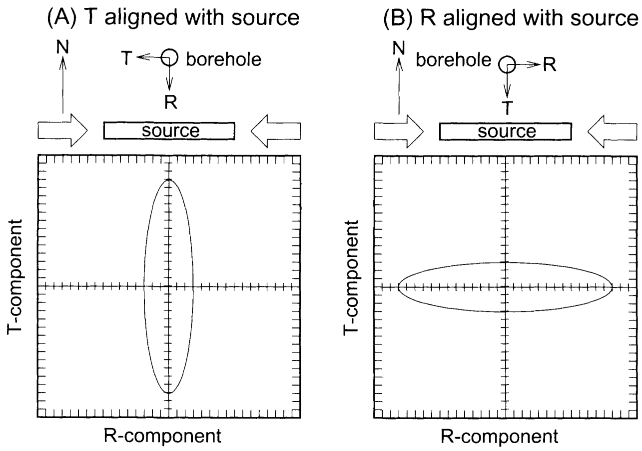

Figure 5. PCA determines the major axis of the elliptical particle motion. Shown are limiting cases of the source aligned (A) with the T-component, and (B) with the R-component.

muths. The 180 degree ambiguity inherent in eigenvector determination was resolved by observing the orientation of the tool when exiting the hole (see Appendix A).

Drift of Source Polarization Direction

Because both the reference geophone and the source are clamped to the soil, one would expect that application of PCA to the reference geophone would result in a constant orientation angle. While significant effort has been expended to guarantee a well-centered and constant strike point for the hammer, some variation will occur. Perhaps more significantly, the small-scale variations in soil stiffness below the source can result in a nonuniform resistance to the blow in shear. The resulting couple twists the radiation ellipse out of alignment with the long axis of the source. This condition appears to decrease as the survey proceeds, perhaps due to the dynamic compaction of the soil (which produces a more uniform, compacted soil below the source).

Figure 7 illustrates how the source radiation drifts toward the desired orientation during the course of the survey for the X5 well (see Fig. 6 for the down-hole analysis). Shown in Fig. 7(A) is the PCA analysis for the fixed reference geophone (T-component). Note that the first source effort at the bottom of the figure is rotated away from the desired alignment of 270°. As the survey continues (and the soil compacts), the polarization direction drifts toward the desired orientation. While the algorithm was designed to display the geophone orientation, note that the result is the relative angle between the geophone and the source radiation. Since the geophone orientation was fixed and observable at the surface, the display actually represents a drift in the SH-wave radiation ellipse. Figure 7(B) is a plan view that brings together the results for the source, its radiation, and both down-hole and reference geophones. Figure 7(C) shows the hodogram for the reference geophone recording when the down-hole geophone was at 12.5 m (see Fig. 6). Thus, the 37° rotation determined in Fig. 6 for the

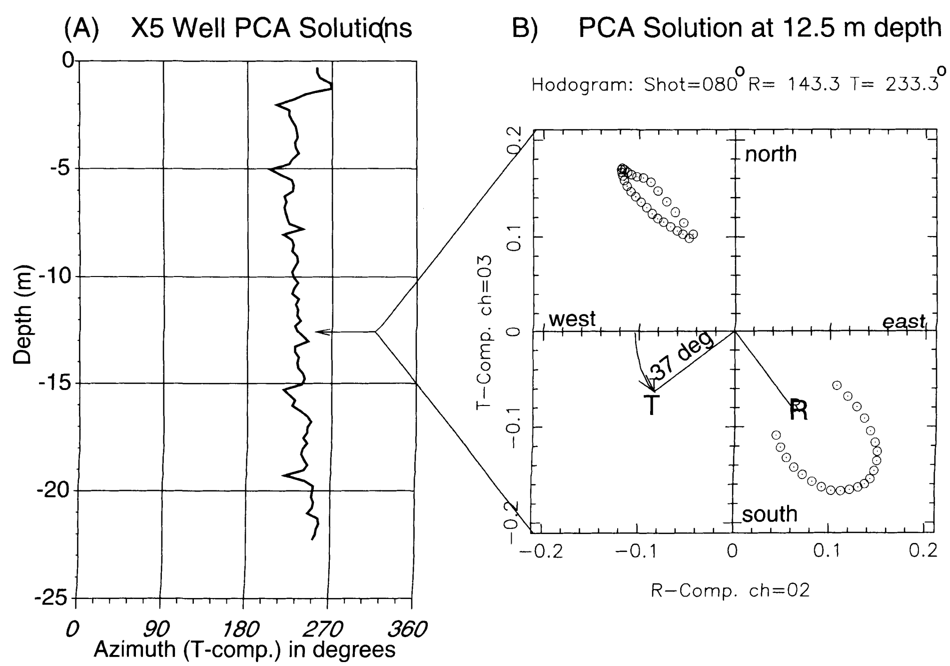

Figure 6. Results of PCA on down-hole geophone for the X5 hole. Azimuth of T-component shown in (A). Sample PCA analysis is shown in (B) for station at 12.5 m depth.

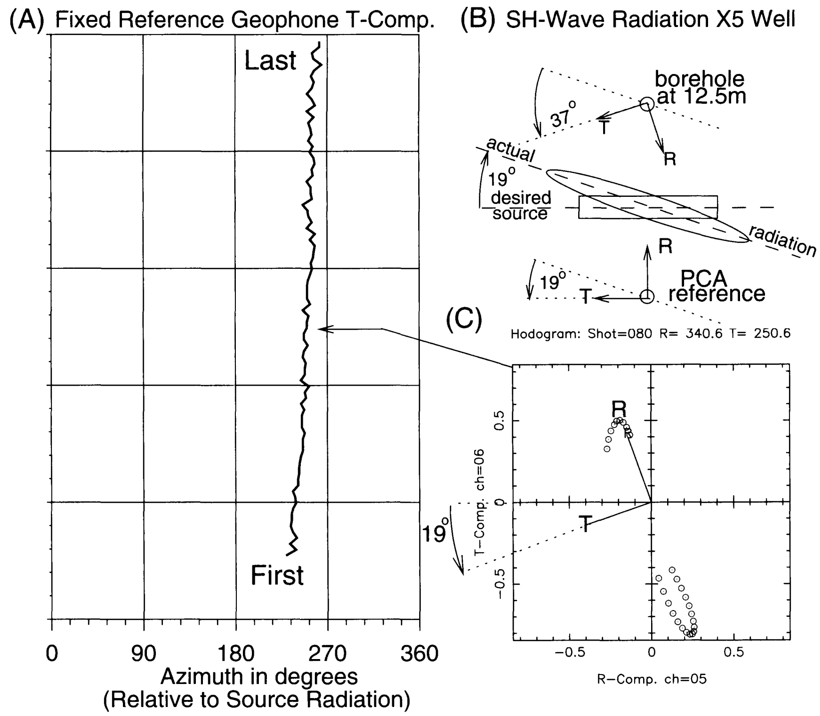

Figure 7. Results of PCA on fixed reference geophone for X5 hole. Azimuth of T-component shown in (A). Plan view (B) showing source radiation angle drifted 19 degrees out of alignment with source axis as determined in hodogram (C).

down-hole geophone is relative to the radiation pattern which was found to be rotated 19° from the long axis of the source.

Other Field Examples

Figure 8 displays the PCA solutions for 3 different wells that span the range of tool rotation typically observed

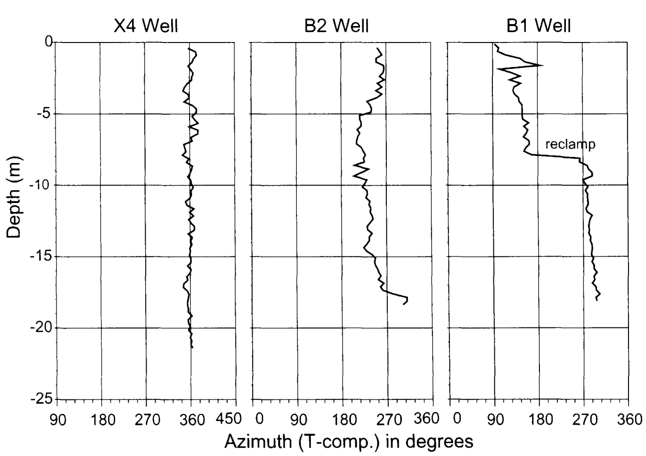

Figure 8. Sample PCA solutions for down-hole T-component from three different boreholes shows the range of typical results. The clamp was released to clear an obstruction in the B1 hole.

at the BHRS site. The simplest result is for the X4 well. The combination of straight round casing and minimal cable twist resulted in little rotation of the tool as it was drawn up the hole.

In the case of the B2 well, more variation in tool orientation is seen as it is drawn up the hole. However, since the tool remained clamped, the variation is continuous and limited to within about 90 degrees.

The B1 well illustrates the discontinuous behavior that results when an obstruction is encountered and the tool released. In this case, the tool was released at 8 meters depth to clear an obstruction (possibly a casing edge at a couple point or some other casing defect). Once released, the tool is free to spin, producing the discontinuity in orientation. The PCA solution starts at the surface, working down-hole until a re-clamp point is reached (Fig. 8). Then a new guide vector is specified (see Appendix A) for the next segment of depths to be analyzed.

B1 Hole in Detail

Figure 9 shows selected hodograms from the PCA analysis algorithm. The top three hodograms are for levels above the release point. The bottom three hodograms correspond to depths below the release point. While most hodograms display elongated elliptical motion, there are instances, like (D), where the motion becomes more open

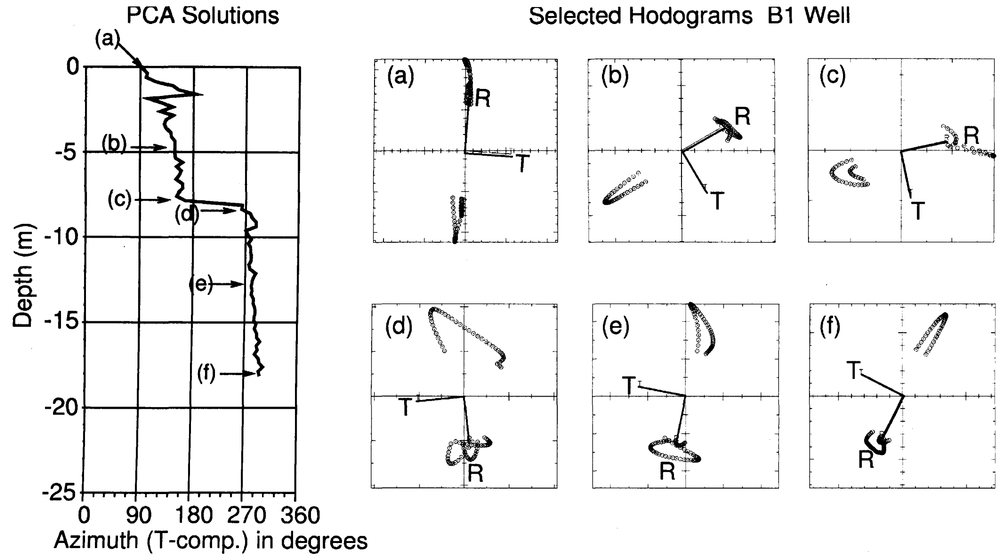

Figure 9. Detailed plots of PCA for B1 well. Data were analyzed in two sets (first above re-clamp point, then below).

(perhaps when scattered waves mix with the direct arrival). In any case, the PCA analysis appears to produce robust results, even for motion which is not purely elliptical.

Figure 10 displays the waveforms as originally recorded on the T-component. The effect of the re-clamping is quite evident in the waveforms. The PCA solution was then used to rotate the data into a standard alignment.

Shown in Fig. 11 are the data after coordinate rotation. The resulting waveforms appear as what would have been recorded if the T-component had remained parallel to the source radiation. The S-wave enhancement subtracted the 270° source azimuth recording from the 90° azimuth recording. Thus, the first motion S-wave is expected to be a peak (black) corresponding to a positive voltage (SEG polarity convention).

Conclusions

The determination of tool orientation by principal component analysis of SH-wave down-hole recordings has been demonstrated. Enhancement of S-waves by the usual subtraction of alternate source polarity recordings improves the formulation by nulling out P- and Rayleigh wave interference. The enhancement process benefited by scaling the two alternate source efforts by the vertical component peak amplitudes. Determination of tool orientation by the large amplitude motion avoids difficulties that might occur with first motion studies (such as residual interference from earlier arriving P-waves). While other hardware-based methods exist, PCA analysis has two advantages. First, it does not slow down data acquisition. Second, it determines the tool rotation relative to the actual source radiation (and not the intended polarization direction). PCA analysis of the fixed reference geophone suggests that the polarization direction of S-wave radiation may drift, even when the source is tightly clamped to the soil. This drift may be due to variations in soil stiffness under the source which diminish over time as the soil becomes compacted by the repeated blows.

Acknowledgments

Financial support was provided by grant DAAH04-96-1-0318 from the U.S. Army Research Office. Views and conclusions contained herein are those of the author and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Army Research Office or the U.S. Government. The author wishes to thank Dr. W. Barrash for his helpful comments and descriptions of the BHRS site geology, and those students in the BSU geophysics program who helped with the data acquisition. Finally, the author would like to thank the developers of Scilab (INRIA, 2000), Seismic Unix (Cohen and Stockwell, 2000), and Latex whose software was used in the preparation of this work. CGISS contribution number 0117.

REFERENCES

Barrash, W., Clemo, T., and Knoll, M., Mar 1999, Boise Hydrogeophysical Research Site (BHRS): Objectives, design, initial geostatistical results, Proceedings of the Symposium on the Application of Geophysics to Engineering and Environmental Problems 1999, 389–398.

Brettmann, T., Gauer, R., and Yilmaz, R., 1996, Shear wave velocities of gulf coast soils determined from cross hole and seismic cone peneterateion tests, Proceedings of the Symposium on the Application of Geophysics to Engineering and Environmental Problems 1996, 617–624.

Cohen, J., and Stockwell, J., 2000, CWP/SU: Seismic Unix re-

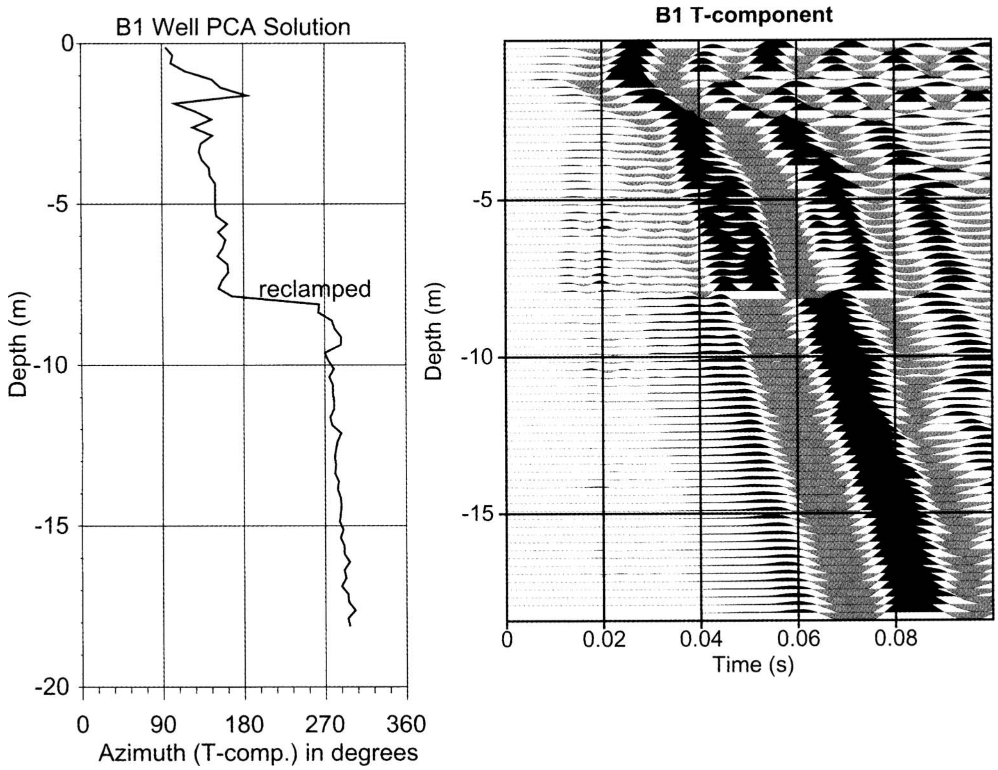

Figure 10. T-component waveforms as recorded for B1 hole.

lease 34: a free package for seismic research and processing: Center for Wave Phenomena, Colorado School of Mines.

Crice, D., 1996, BHG-2, BHG-3 borehole geophone operation manual: GeoStuff, 19623 via Escuela Dr., Saratoga, Ca 95070.

INRIA, 2000, Introduction to Scilab: INRIA Unité de recherche de Rocquencourt Projet Meta2, Domaine de Voluceau Rocquencourt B.P. 105 78153 Le Chesnay Cedex.

Jolliffe, I., 1986, Principal Component Analysis: Springer-Verlag.

Liu, X., 1999, Ground roll supression using the Karhunen-Loève transform: Geophysics, **64**, no. 2, 564–566.

Menke, W., 1989, Geophysical data analysis, discrete inverse theory: Academic Press.

Michaels, P., 1998, In situ determination of soil stiffness and damping: Journal of Geotechnical and Geoenvironmental Engineering, **24**, no. 8, 709–719.

Mok, Y.J., Sanchez-Salinero, I., II, K.S., and Roesset, J.M., 1988, In situ damping measurements by crosshole seismic method, *in* Recent Advances in Ground Motion Evaluation ASCE, Geotechnical Special Publication 20, 305–320.

Sanger, T., 1989, Optimal unsupervised learning in a single-layer linear feed forward neural network: Neural Networks, pages 459–473.

Appendix A: PCA Analysis

Consider only the seismic signals recorded from the two horizontal elements of a 3-component geophone. Each recorded time sample may be viewed as a vector sample, formed from the respective scalar samples taken from the two horizontal geophones. Thus, the j-th vector sample can be written as

$$\xi_j = \begin{pmatrix} x_j \\ y_j \end{pmatrix}$$

where x_j and y_j are the samples taken from the two orthogonally oriented horizontal geophones. For a given S-enhanced seismic recording, there will be a total of N vector samples, ξ. Thus, the collection of vector time samples, ξ, form an N-dimensional vector composed of 2-dimensional samples. Rather than viewing this as a matrix, think of it as N realizations of a vector variable. It is this set of N 2-dimensional samples which is subjected to PCA at any given subsurface station. Only samples whose modulus exceeds a minimum threshold (typically 50% of the maximum) are included to form the set of N vector samples.

There will be no loss of generality if the x_j and y_j values are voltages representing particle displacement, velocity, or acceleration. The only requirement is that they both be of the same type of quantity, with equal scaling of units. The orientations of the x and y geophones establish a basis (coordinate system) for the vector samples, ξ. The basis vectors are directed along the physical axis of each horizontal geophone element, pointed in the direction of a consistent voltage sign convention.

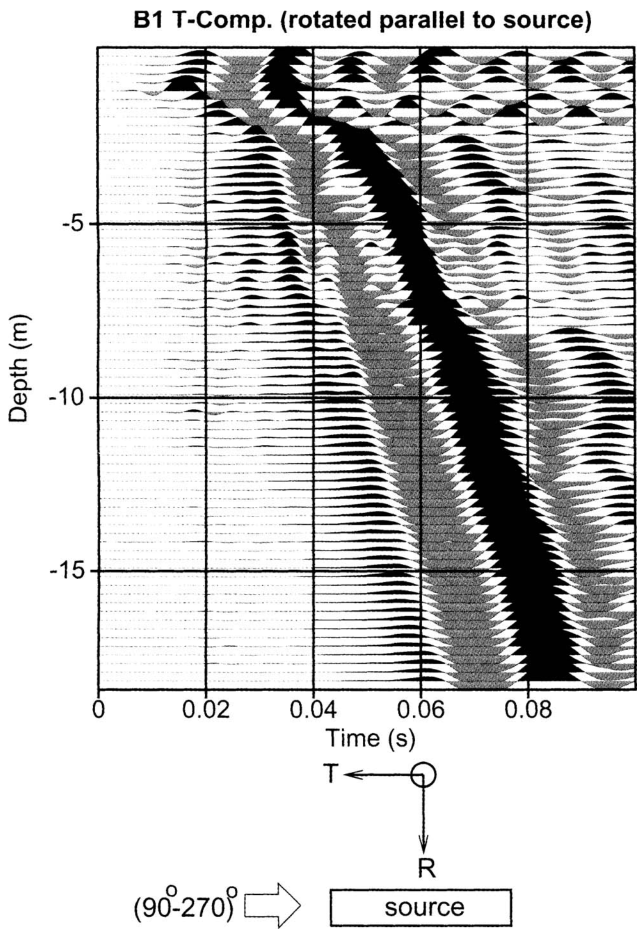

Figure 11. T-component waveforms after rotation into alignment with the source as determined from the Fig. 10 PCA solutions.

The recording of an SH-wave arrival will produce N realizations of vector samples, ξ, which would plot as an elongated distribution of points falling within a generally elliptical pattern (Fig. 5). The direction of the long axis of this ellipse is the polarization direction of the SH-wave relative to the basis established by the geophone elements.

Our objective in this application of PCA analysis is to determine the relative SH-wave polarization direction in the geophone basis. This will be the direction with the greatest variance on a horizontal hodogram, since we define the polarization direction as the major axis of the "elliptical" motion. Details on PCA can be found in Jolliffe (1986). Related topics are found in communication theory (Karhunen-Loève transforms) (Liu, 1999), factor analysis (Menke, 1989), and neural network feature extraction (Sanger, 1989). The following is a brief summary of the mechanics of the computation for this application.

For N realizations of the vector samples, ξ, we will define an average value in the usual way,

$$\bar{\xi} = \frac{1}{\mathrm{N}} \sum \xi_j = \begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix}.$$

Recall that the N realizations of seismic samples also form vectors in an N-dimensional vector space. We will define a scalar inner product in this N-dimensional space as,

$$\langle x, y \rangle = \sum x_j y_j = x^{\mathrm{T}} y.$$

We define the 2 × 2 covariance matrix by the outer product,

$$\underline{\underline{C}} = \sum (\xi_j - \bar{\xi})(\xi_j - \bar{\xi})^{\mathrm{T}}$$
$$= \begin{bmatrix} \langle (x - \bar{x}), (x - \bar{x}) \rangle & \langle (x - \bar{x}), (y - \bar{y}) \rangle \\ \langle (y - \bar{y}), (x - \bar{x}) \rangle & \langle (y - \bar{y}), (y - \bar{y}) \rangle \end{bmatrix}.$$

The principal component directions are the eigenvectors of $\underline{\underline{C}}$. The polarization direction of the SH-wave is the eigenvector associated with the largest eigenvalue. Since there will be an azimuthal ambiguity of 180° in the eigenvector determination, practical algorithms require additional input (guide vector) to resolve that issue. This can be done by examining the data for consistent polarity or by direct observation of the source and geophone at the ground surface.