Wave Inversion Technology, Report No. 2, pages 297-307

Coherency Analysis and Correlation Procedures

Robert Mauch[1]

keywords: *coherence measures, crosscorrelation, random processes*

ABSTRACT

Coherency measures are widely used in the analysis of seismic data. The foundations of many coherency measures are the concepts of linear correlation and stochastic processes. These concepts are briefly reviewed and, based on them, a standard coherency measure, the statistically normalized crosscorrelation measure, is derived. The possible uses of coherency measures in seismic data processing are various. Each time, however, the input data, a theoretical moveout model and a coherency measure are combined to analyze the data. With a synthetic example it is shown how coherency analysis can be used to detect multiple reflections in CSP gathers.

INTRODUCTION

With the use of multichannel data coherency analysis was introduced to the standard processing sequence of seismic data. With it, statistical concepts were applied to multiple channels (Schneider and Backus, 1968; Taner and Koehler, 1969). The formulations of early coherence functions constitute generalizations to existing statistical correlation coefficients (Bendat and Piersol, 1986). With a brief review, the underlying statistical concepts are explained. Based on these concepts the statistical normalized crosscorrelation measure is derived. As a second important utility the formulation of a theoretical traveltime model goes into the examination. The combination of input data, moveout model and coherency measure constitutes the framework of coherency analysis. One possible combination is applied to detect multiple reflections in CSP gathers. Various other possible applications are mentioned.

BASIC CONCEPTS

In the next three sections I give a brief review of the concepts of *linear correlation*, *stochastic processes* and *time-series analysis* which are fundamental in the development

[1]**email:** robert.mauch@phys.uni-karlsruhe.de

of coherency measures, such as *normalized crosscorrelation* or *semblance*.

Linear Correlation

In probability the concept of *random variables* is introduced. A random variable is a function that assigns a real number to each outcome of a *random experiment*. A random experiment is characterized by the fact that its possible output is not predictable. The *mean* or *expected value* of a discrete random variable X is

$$\mu_X = E(X) = \sum_x x\, p_X(x)$$

where $\sum_x$ is the sum over the set of all possible values of the random variable and $p_X(x)$ is the *probability mass function* from this set.
The *variance* σ^2 of a random variable X is defined as

$$\sigma_X^2 = E(X-\mu_X)^2 = \sum_x (x-\mu_X)^2\, p_X(x)$$

If we consider two random variables a measure of the possible relationship between them is the *covariance*:

$$\sigma_{XY} = E[(X-\mu_X)(Y-\mu_Y)]$$

whereas

$$E[(X-\mu_X)(Y-\mu_Y)] = \sum_R (x-\mu_X)(y-\mu_Y)\, p_{XY}(x,y)$$

is the weighted average of the product of the deviations of the two random variables from their mean values. The weight is here the *joint probability mass function* $p_{XY}(x,y)$.
Another measure of the relationship between two random variables is the *correlation coefficient* ρ_{XY}. It is just a scaled version of the covariance.

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} \qquad -1 \le \rho_{XY} \le +1$$

Two random variables with nonzero correlation are said to be *correlated*. Note that both covariance and correlation are measures of the *linear relationship* between random variables.

In practical situations where a finite sample of measurements is observed the accuracy of a linear relationship can be estimated by the *sample correlation coefficient*

$$R = \frac{S_{XY}}{\sqrt{S_{XX} S_{YY}}} \qquad -1 \le R \le +1 \tag{1}$$

with $S_{XY} = \sum_{i=1}^n (x_i - \overline{x}) y_i$, $S_{XX} = \sum_{i=1}^n (x_i - \overline{x})^2$ and $S_{YY} = \sum_{i=1}^n (y_i - \overline{y})^2$, where n is the sample size.

In the next section the concept of *stochastic processes* is introduced which constitutes the statistical basis for *time-series analysis* and makes it possible to apply the concept of linear correlation to time series.

Stochastic Processes

The concept of stochastic or random processes is used to describe data representing random physical phenomena. The hypothetical structure of the data of a stochastic process $\{X(t)\}$ consists of a infinitely large ensemble of time-history records or time functions. The continuous time functions reach from minus infinity to plus infinity in time. At any instant of time the values over the ensemble represent all possible outcomes of a random experiment. Thus, at any instant of time (t_i) a *random variable* $X(t_i)$ can be defined. The statistical properties of each random variable are computed by constituting the average over the ensemble of time functions.

For example, suppose the ensemble values of the random variable $X(t_1)$ at time t_1 are $x_1(t_1), x_2(t_1), \ldots, x_N(t_1)$. Then the *expected value* or *mean value* is obtained computing the *ensemble average*

$$\mu_X(t_1) = E(X(t_1)) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} x_k(t_1)$$

assuming the ensemble values to be equally likely.

In general, the ensemble values may come from a continuous sample space with infinite extension and may also be weighted by a *probability density function* $p(x)$. In this case we define the *expected value* of a random variable $X(t)$ at arbitrary fixed values of t in the form

$$\mu_X(t) = E(X(t)) = \int_{-\infty}^{\infty} x(t)\, p_X(x(t))\, dx(t) \tag{2}$$

A *stochastic process* is completely defined by constituting all available statistical quantities over the ensemble. In general, the values of these quantities are different at different times. In the case where the statistical quantities do not vary with time the stochastic processes are called *stationary*.

In the case of stationarity there exists a subclass of stochastic processes for which the statistical quantities computed as *time averages* are equal to the corresponding ensemble averages. These *ergodic processes* constitute the statistical foundation to analyze time series in seismic applications.

The *time average* of a stochastic process is performed by computing the average over the time range of a single realization from the ensemble of time functions. That is, only one time record is necessary to describe the statistical structure of the whole process. For example, in Equation 2 the mean value of the stationary stochastic process $\{X(t)\}$, $-\infty \leq t \leq \infty$, was defined as ensemble average. For ergodic processes this definition is equivalent to the *time average definition* of the *mean*

$$\mu_x = \lim_{T\to\infty} \frac{1}{2T+1} \int_{-T}^{T} x(t)\, dt \tag{3}$$

Note that the statistical quantities are now subscribed by lower-case letters to denote that they are resulted from averaging one possible value of all random variables $X(t)$, $-\infty \leq t \leq \infty$, instead of averaging all possible values of one random variable $X(t)$, t fixed.

Time-Series Analysis

In time-series analysis we refer to a *stationary time series* to be a realization of a discrete time-history record coming from a ergodic stationary random process. That is, one single stationary time series implies all underlying probabilistic properties of the stochastic process it comes from.

For example, the *mean value* of a random process is computed as in Equation 3 defined, but in discrete form:

$$\mu_x = \lim_{T\to\infty} \frac{1}{2T+1} \sum_{t=-T}^{T} x_t$$

In practical situations only values over a finite time interval are obtainable. Such a limited section of a stationary time series is called a *sample of a stationary time series*. We use a *time gate* or *time window,* $N+1$ time samples wide, to define a sample of a stationary time series. The sequence of signal values within a time gate may be given by

$$x_{k-(N/2)}, x_{k-(N/2)+1}, \ldots, x_k, \ldots, x_{k+(N/2)-1}, x_{k+(N/2)}$$

where k denotes the center of the time gate. The *mean value* over this time gate is computed by

$$\hat{\mu}_x = \frac{1}{N+1} \sum_{t=k-\frac{N}{2}}^{k+\frac{N}{2}} x_t$$

The mean value $\hat{\mu}$ is now regarded as an *estimate* of the mean value of the random process since only a finite sample of observations is used to determine this quantity. That is, using samples of stationary time series yields to statistical errors in all statistical quantities describing the structure of the stochastic process.

Now, equivalent to the *sample correlation coefficient* (Equation 1) in correlation analysis, a statistical estimator may be introduced to judge the linear relationship between the values of two traces. To do so, we first define the *sample crosscorrelation* *function* for a pair of samples of stationary time series:

$$\hat{R}_{xy}(\tau) = \begin{cases} \frac{1}{N+1-\tau} \sum_{t=k-\frac{N}{2}}^{k+\frac{N}{2}-\tau} x_t y_{t+\tau} & \text{for } \tau = 0, 1, \ldots, k + \frac{N}{2} \\ \hat{R}_{xy}(-\tau) & \text{for } \tau = -1, -2, \ldots, -(k + \frac{N}{2}) \\ 0 & \text{for } \tau < -(k + \frac{N}{2}) \text{ and } \tau > k + \frac{N}{2} \end{cases} \quad (4)$$

Note that τ denotes the *lag number* in this definition, i.e. the number of time sample points record y is shifted with respect to record x.

This estimate provides us with an estimate for the *normalized crosscorrelation function* which is defined by

$$\hat{\rho}_{xy}(\tau) = \frac{\hat{R}_{xy}(\tau)}{\sqrt{\hat{R}_{xx}(0)\hat{R}_{yy}(0)}} \qquad -1 \leq \hat{\rho}_{xy}(\tau) \leq +1 \quad (5)$$

The *zero-lag value* of this function is equivalent to the *sample correlation coefficient*. This means, a value of zero indicates no correlation between the values of two traces and a value of one indicates a perfect linear correlation.

COHERENCY MEASURES

So far, we only have analyzed the correlation between *two* seismic traces. To analyze the coherent signal content within particular time gates of *multiple channels* there is nothing more to do than to calculate the two-trace correlation for each combination of the channels and to average the sum of the correlation values by the number of combinations. The position of each time window is defined by the *moveout model*, that is, we parameterize the center positions of the time windows by the time values coming from the moveout formula of interest.

For example, we consider the case of a multichannel CMP gather. We want to analyze this gather for coherent signals following the theoretical lag trajectory

$$t_k(i) = \sqrt{t_k^2 + \frac{x_i^2}{v^2}}$$

where $t_k(i)$ are the calculated two-way times for discrete offsets x_i at zero-offset time t_k, where k denotes the time sample of the center of the time gate, and velocity v.

For a time gate $N + 1$ time samples wide we get a set of $N + 1$ symmetrically disposed trajectories

$$t_j(i) = t_k(i) + (j - k)\Delta t \qquad k - N/2 \leq j \leq k + N/2$$

where t_k is the trajectory laying along the centers of the time gates and Δt is the sampling rate.

So defined, the zero-lag value of the crosscorrelation function (Equation 4) for two channels i and i' is calculated by

$$R_{ii'} = \frac{1}{N+1} \sum_{j=k-\frac{N}{2}}^{k+\frac{N}{2}} u_{x_i,t_j(i)} u_{x_{i'},t_j(i')} \tag{6}$$

when we denote the signal values on the traces with u. The multichannel equivalent to the normalized crosscorrelation function (Equation 5) is

$$NCC = \frac{2}{(M-1)M} \sum_{i=1}^{M} \sum_{i'>i} \frac{R_{ii'}}{\sqrt{R_{ii} R_{i'i'}}} \qquad -1 \leq NCC \leq +1 \tag{7}$$

where M is the number of channels. This coherency measure is the *(statistically) normalized crosscorrelation measure*. Its value reaches unity if the phase and shapes of the signal within the time gates are identical. In other words, if the corresponding signal values of different channels take a fixed linear combination over the entire time window the linear correlation is a maximum (comp. with the two-trace correlation of Equation 5).

Further coherency measures base on this statistical concepts, e.g., *unnormalized crosscorrelation* (Schneider and Backus, 1968) or *semblance* (Taner and Koehler, 1969; Neidell and Taner, 1971). Other coherency measures were designed in different domains or with different normalization schemes and differ in their ability to record variations in the alignment of signals or their sensitivity to amplitude or sign changes (Garotta and Michon, 1967; Taner and Koehler, 1969; Neidell and Taner, 1971; Morozov and Smithson, 1996; Mauch, 1999).

Whichever measure one chose, they must always be considered in connection with the type of data set and the traveltime model they shall applied to.

MOVEOUT MODELS

In seismic moveout models are traveltime functions. A whole of theoretical traveltime models exists for every type of gathered data. The models differ in their theoretical concepts and in the set of parameters used to parameterize the model.

For example, for Common Shot Point gathers (CSP) a moveout model based on the Homeomorphic Image theory (HI) (Gelchinsky, 1989) was proposed by Keydar et al. (1996). In their description the moveout is parameterized by the *radius of curvature* of the wave front r_0 and the *angle of emergence* of the central ray β_0

$$t(x) = t_0 + \frac{(r_0^2 + 2r_0 x \sin\beta_0 + x^2)^{1/2} - r_0}{v_0} \tag{8}$$

where the near surface velocity v_0 is mostly assumed to be known and therefore no additional parameter. Once, a set of parameters is chosen and a zero-offset time (t_0) is fixed, the traveltime $t(x)$ can be calculated in dependence of the receiver distance x. The result of a coherency analysis may be a set of coherency values computed for a range of radii, angles and zero-offset times. In the next section such a data cube is shown for this moveout model (see Figure 3).

In general, coherency analysis aims to extract that parameter combinations for which the chosen model fits best to the data. The parameters may then be used to stack the data or to analyze the structure of a coherency spectrum to detect primary and multiple reflections. In addition, if there were multiple gathers to analyze, the parameter values itself may be plotted in a simulated zero-offset section to reveal the structure of the subsurface.

To illustrate the application of the coherency technique two simulated shot gathers are analyzed in the next section.

EXAMPLE OF A COHERENCY ANALYSIS

In this section I use the above mentioned traveltime model (Equation 8) to analyze two shot gathers. The shot gathers were generated with an elastic FD algorithm — one with a free surface producing surface multiples and one with a surface having absorbing properties (see Figure 2). The shots were generated for the SEG/EAGE Salt Model (Figure 1) at a distance of 5,960 meters. For coherency analysis the above derived normalized crosscorrelation measure (Equation 7) was used. A coherency value was calculated for every zero-offset time t_0 over a range of radii from 0 to 10,000 meters and a range of angles from -80 to 80 degree. The calculation was performed in a

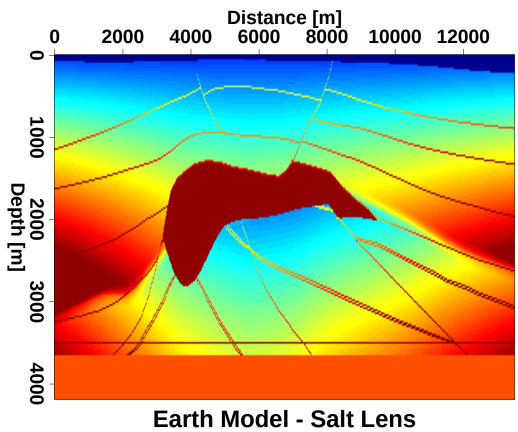

Figure 1: SEG/EAGE Salt Model.

parallel environment using a HPF coded program. The results can be visualized by data cubes which are best examined in a virtual environment.

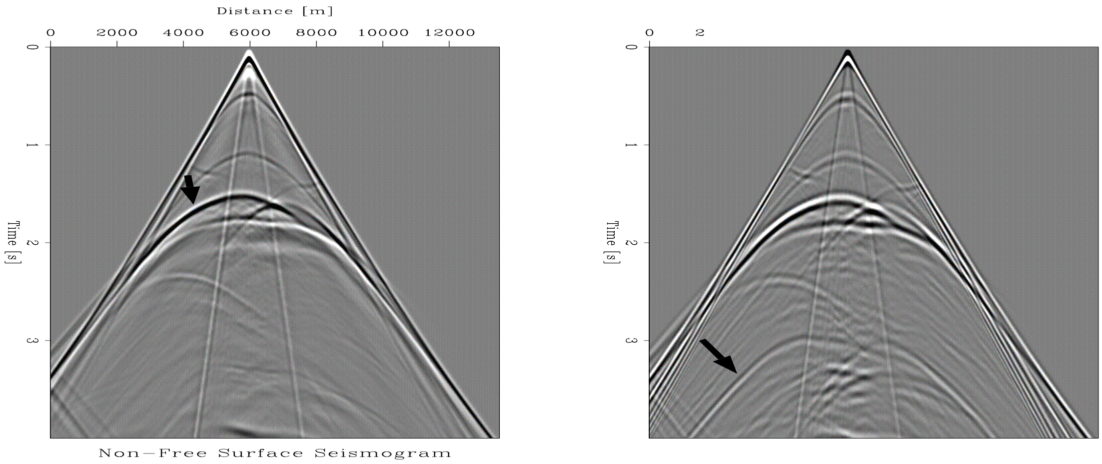

Figure 2: Synthetic CSP-gathers: Absorbing Surface (left) and Free Surface (right).

Figure 3 shows three slices of this cube at zero-offset time $t_0 = 1.67$ s, angle $\beta_0 = 1.6$ deg and radius $r_0 = 5,300$ m. This is the parameter combination which yielded the highest coherency value. Comparing this result with the figures of the salt model and the seismogram it can be observed that it corresponds with the strong primary reflection from the top of the salt lens (indicated by an arrow in Figure 2, left).

Figure 3: Result of a coherency analysis: This section through the coherency cube shows the focused high coherency value of the primary reflection from the top of the salt lens.

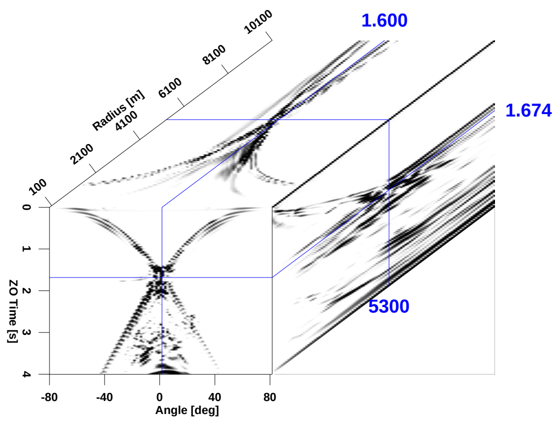

Figure 4 shows a comparison of the two data cubes at zero-offset time $t_0 = 2.87$ s, angle $\beta_0 = 3.2$ deg and radius $r_0 = 5,900$ m. Since the values for angle and radius are only slightly different from the values of the primary reflection and the zero-offset time is approximately doubled, this parameter combination corresponds with the surface multiple of the reflection from the salt top (see arrow in Figure 2, right). As expected, this multiple can be detected at a high coherency value in the data of the Free Surface Seismogram (Figure 4, right) while it is not visible in the data of the Non-Free Surface Seismogram (Figure 4, left).

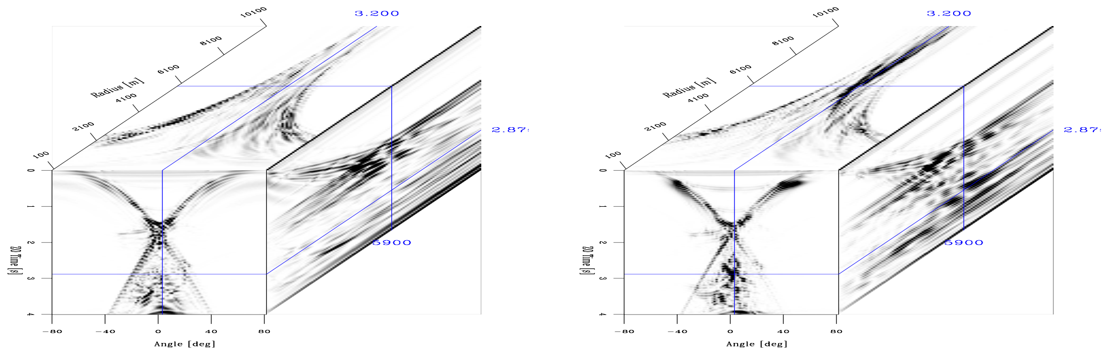

Figure 4: Comparison at the parameter combination of the multiple: Absorbing Surface Data (left) and Free Surface Data (right).

This qualitative example points out only one application of coherency analysis from a multiplicity of possible applications in seismic data processing. For further examples see, e.g., Gelchinsky et al. (1985), Marfurt et al. (1998) or Mauch (1999). Also, it should be noted here that in the macro-model-independent imaging process coherency analysis plays a crucial role concerning the evaluation time and the accuracy of the final stack (for CRS Stack see the other publications in this report).

CONCLUSION

Many coherency measures are based on the statistical concept of linear correlation. The statistically normalized crosscorrelation measure was derived in this paper. In seismic data processing there exists a wealth of possible applications of coherency analysis. The performance of coherency analysis mainly depends on four factors: The type of data to be to analyzed, the proper moveout model, the ability of the coherency measure to reveal coherent signals and the performance of the chosen computer algorithm in connection with the appropriate hardware. Since the use of coherency measures will increase it is worth to further develop these techniques.

REFERENCES

Bendat, J. S., and Piersol, A. G., 1986, Random data: Analysis and measurement procedures: Wiley-Interscience.

Garotta, R., and Michon, D., 1967, Continous analysis of the velocity function and of the move out corrections: Geophys. Prosp., **15**, 584–597.

Gelchinsky, B., Landa, E., and Shtivelman, V., 1985, Algorithms of phase and group correlation: Geophysics, **50**, 596–608.

Gelchinsky, B., 1989, Homeomorphic imaging in processing and interpretation of seismic data — fundamentals and schemes: 59th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 983–988.

Keydar, S., Gelchinsky, B., and Berkovitch, A., 1996, Common shot point stacking and imaging method: J. Seism. Expl., **5**, 261–274.

Marfurt, K. J., Kirlin, R. L., Farmer, S. L., and Bahorich, M. S., 1998, 3-D seismic attributes using a semblance-based coherency algorithm: Geophysics, **63**, 1150–1165.

Mauch, R., 1999, Coherency analysis of seismic data: Master's thesis, Univ. of Karlsruhe.

Morozov, I. B., and Smithson, S. B., 1996, High-resolution velocity determination: Statistical phase correlation and image processing: Geophysics, **61**, 1115–1127.

Neidell, N. S., and Taner, M. T., 1971, Semblance and other coherency measures for multichannel data: Geophysics, **36**, 482–497.

Schneider, W. A., and Backus, M. M., 1968, Dynamic correlation analysis: Geophysics, **33**, 105–126.

Taner, M. T., and Koehler, F., 1969, Velocity spectra — Digital computer derivation and applications of velocity functions: Geophysics, **34**, 859–881.

PUBLICATIONS

Detailed results were published by Mauch (1999).