Statistical Averages for Velocity Analysis and Stack: Median vs. Mean

Marta Jo Woodward

Introduction

Multichannel data is combined into approximate zero-offset sections through the operations of velocity analysis, normal moveout and stack. As traditionally implemented, both velocity analysis and stack use the mean as a statistical average—the former in calculating semblance or constant velocity panels and the latter in converting a moveout corrected gather into a zero-offset trace. This paper compares and contrasts mean and median formulations of velocity analysis and stack through examination of two data sets with contrasting noise statistics. The results illustrate the well known ***robust***, spike-rejecting quality of the median while raising questions concerning the statistic's sensitivity to residual normal moveout.

Median vs. Mean Stacks

Without trace balancing

Representative CDP gathers from two data sets are shown in Figure 1: the one on the left is pulled from a split spread, Vibroseis survey on land; the one on the right comes from an off end, marine air-gun survey. Both data sets have been corrected for spherical divergence; neither has been trace balanced. For this study the most important difference between the two gathers is the contamination of the Vibroseis survey with high amplitude noise spikes on the inner offsets. Mean and median stacks of the two data sets are shown in Figures 2 and 3: displays for each data set are plotted with the same absolute clip; the Vibroseis sections were trace balanced after stack for display purposes.

Examination of the figures reveals three differences resulting from the application of the contrasting statistical averages. First, the high frequency content of the median stacks is larger than that of the mean stacks in both examples. Summing across offset, the mean is a linear operator and preserves the anti-alias filtered frequency spectrum of the input data; sampling across offset, the median is a nonlinear operator and introduces frequencies up to the Nyquist. A comparison

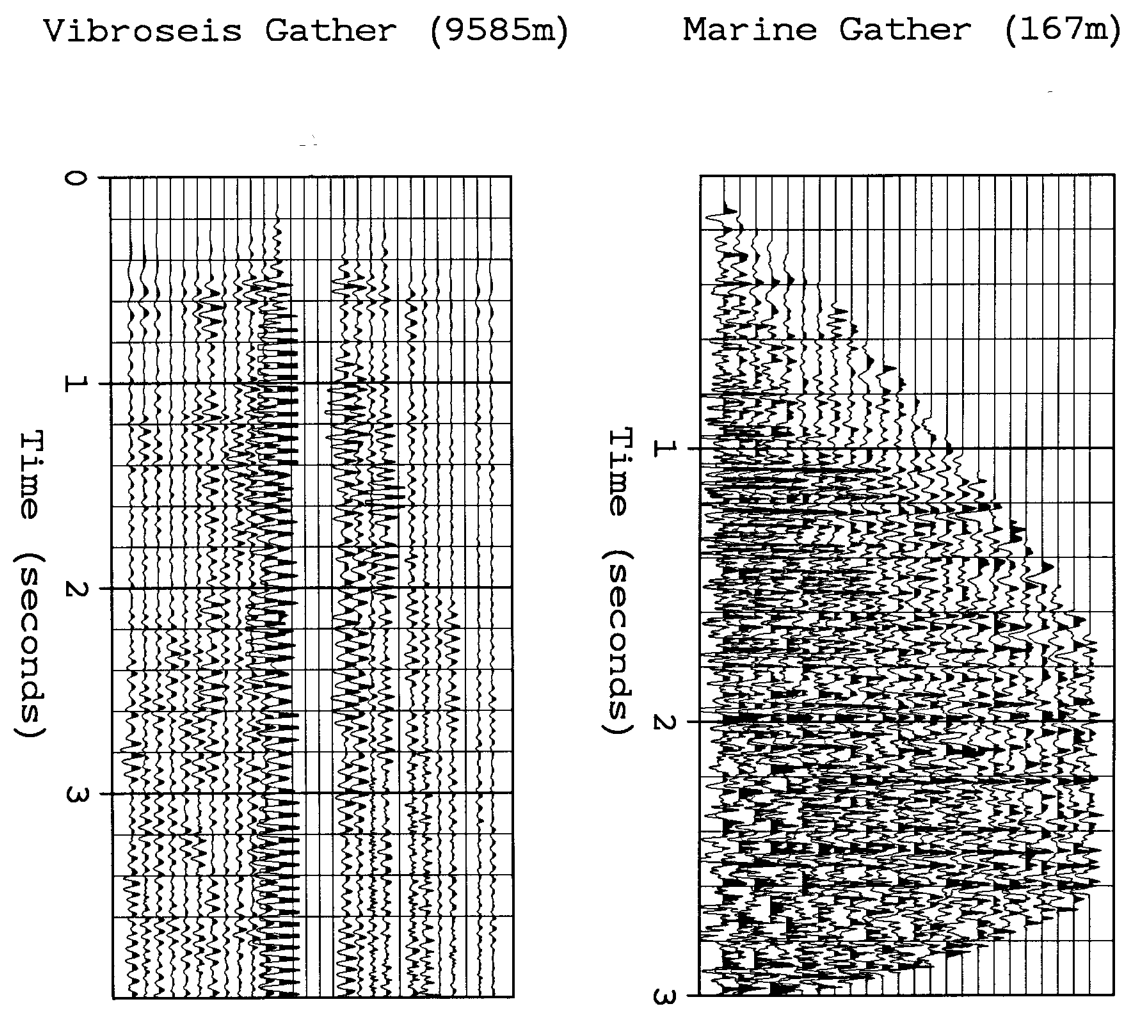

FIG. 1: Moveout corrected gathers from a Vibroseis land survey and a marine air-gun survey.

of the mean median and mean stack amplitude spectra is shown in Figure 4.

Second, the reflector continuity of the Vibroseis median stack is increased relative to the mean stack across traces contaminated by large noise spikes. Note in particular the noisy region around midpoint 50. This result illustrates the fact that the L1 (median) and L2 (mean) norms are the maximum likelihood estimators for the high tail weight exponential and lower tail weight Gaussian probability distributions (Claerbout and Muir, 1973; Gray, 1979), respectively; alternatively phrased, it demonstrates that the median functions as a robust, automatic editor by rejecting erratic data values.

Third, for both data sets the median stack shows less total power than the mean stack; median stack power is 86% of mean stack power for the marine example and 89% for the land example (calculated before trace balancing). Since the mean and the median should estimate the same average for symmetric distributions, this consistent difference suggests both data sets

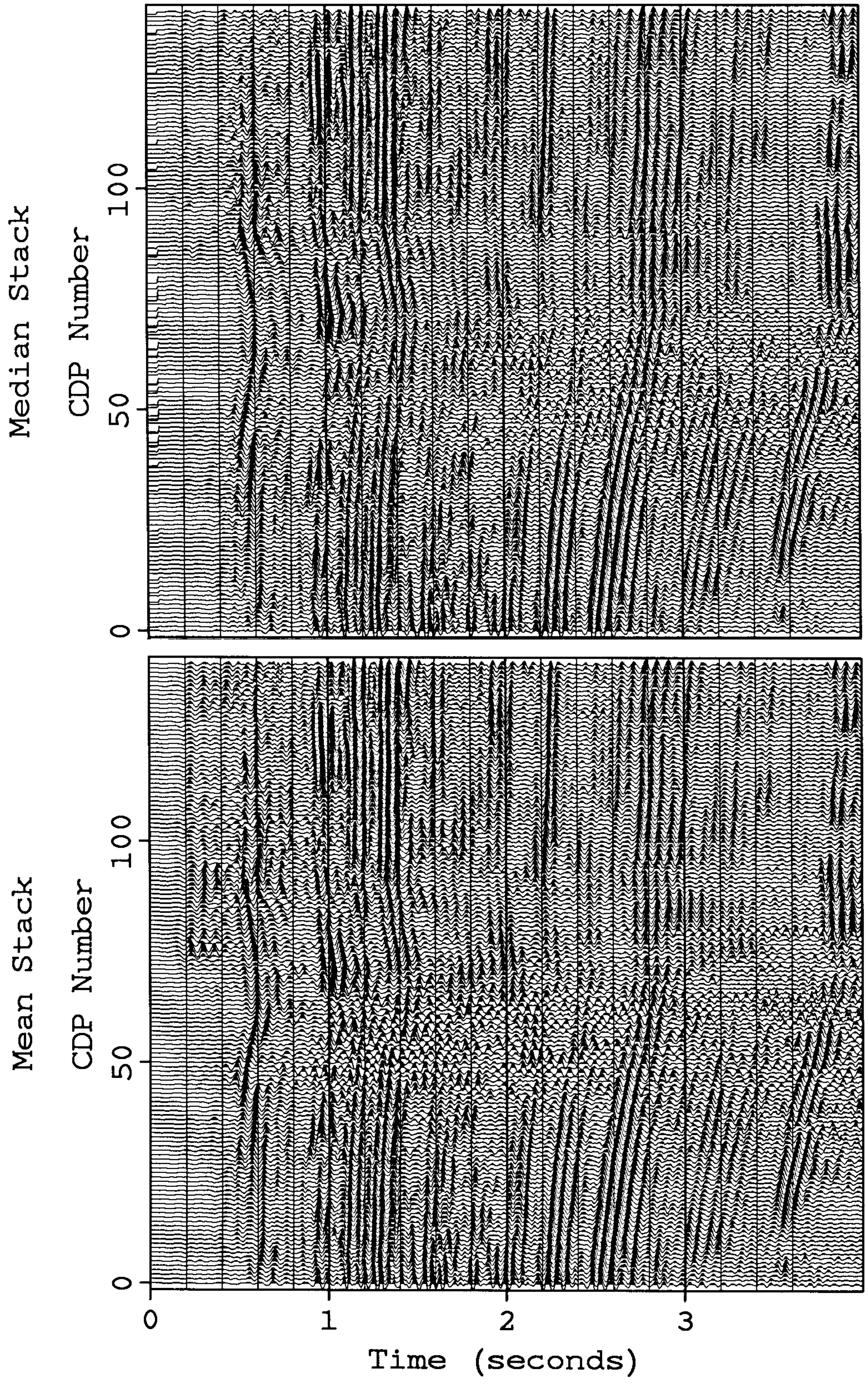

FIG. 2: Mean and median stacks of Vibroseis land data, plotted with the same clip. The sections were trace balanced after stack for display purposes. Data courtesy of Chevron.

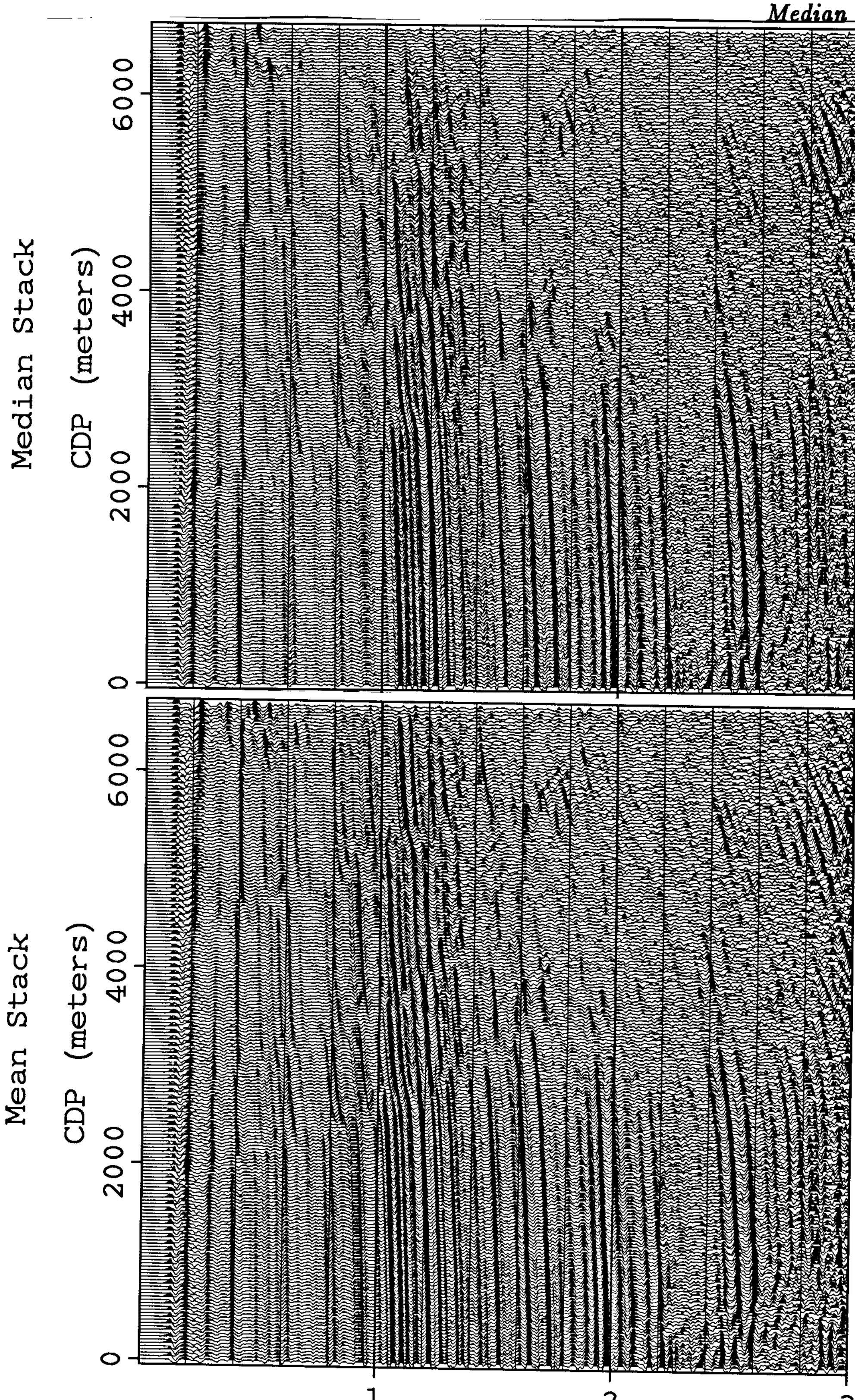

FIG. 3: Mean and median stacks of marine air-gun data, plotted with the same clip. Data courtesy of Western Geophysical.

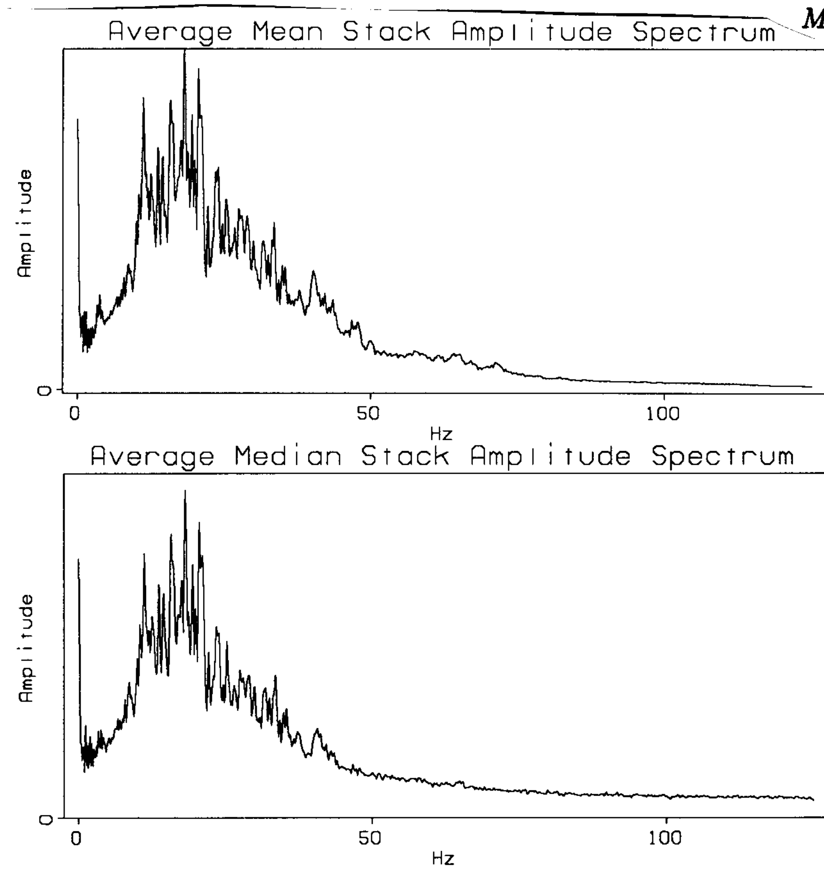

FIG. 4: Mean amplitude spectra of the mean and median stacked marine air-gun data, plotted at identical scales.

possess slightly asymmetric amplitude distribution functions. This conclusion is strengthened by examination of Figure 5. Showing the difference between the median and the mean to be a faint copy of the original sections, the plots confirm a generalized skewing of event amplitudes.

With trace balancing

In stacking data with the purpose of increasing the ratio of signal to noise, reflection amplitude is usually assumed to be constant across offset. Thus an alternative implementation of median stacking might follow an initial trace balancing. Figure 6 shows mean and median stacks of the marine data set where the traces were first normalized across offset to some power per number of nonzero samples. Reference to Figure 7 shows that this processing scheme significantly reduces the amplitude differences between the mean and median stacks for the marine data set; median stack power for this case is 3% greater than mean stack power. Furthermore, in contrast to Figure 5, most of the energy in Figure 7 appears in three places: in a random salt-and-pepper pattern perhaps related to the higher frequency content of the median; in the strong horizontal reflectors just after 1 second, and in a few scattered, diffuse patches.

An example of the latter—located between 5250 and 5500 meters at 1.6 seconds—is particu-

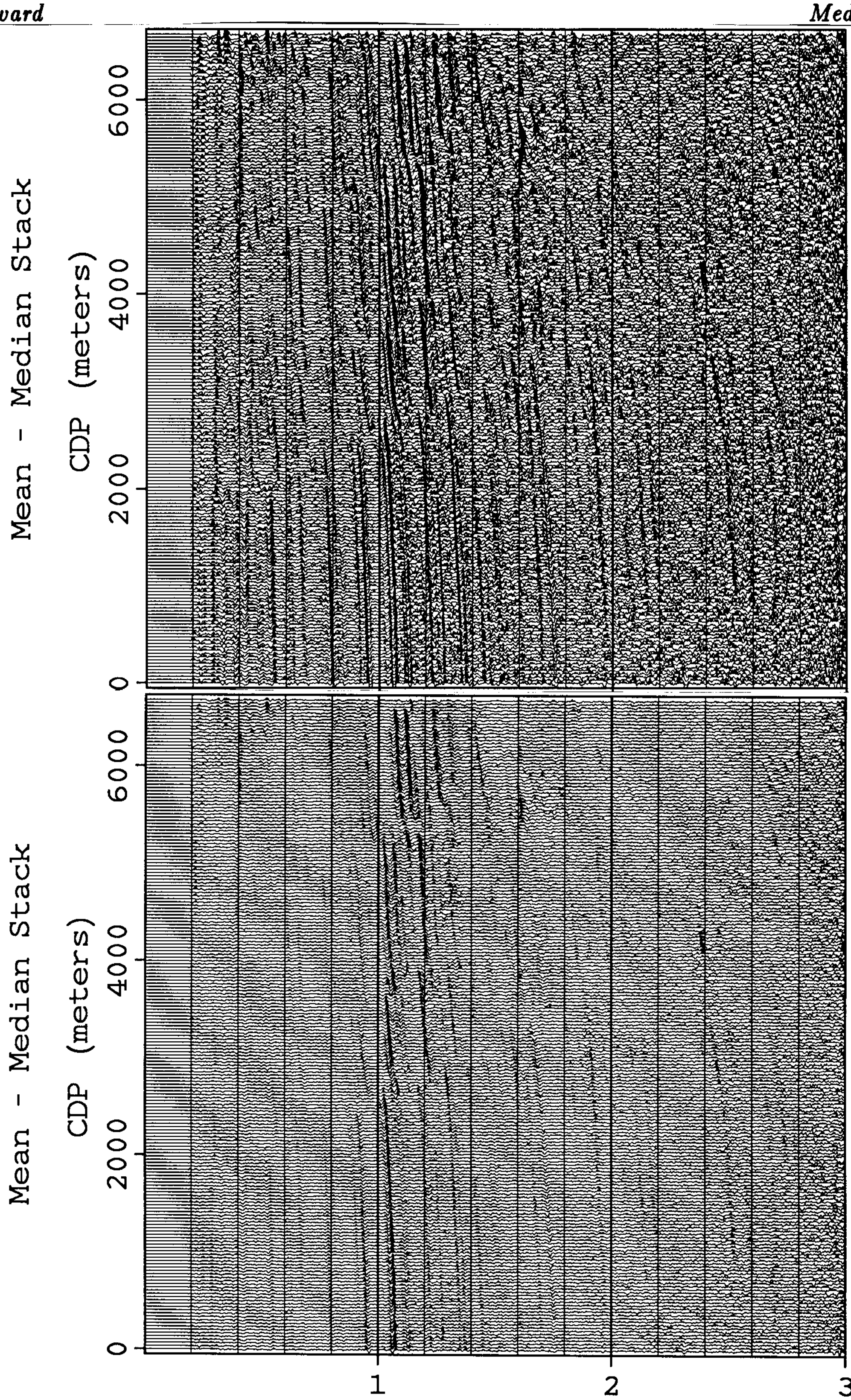

FIG. 5: The difference between the mean and the median for the marine data set, plotted on the left at the same clip as the stacked sections and gained on the right for display.

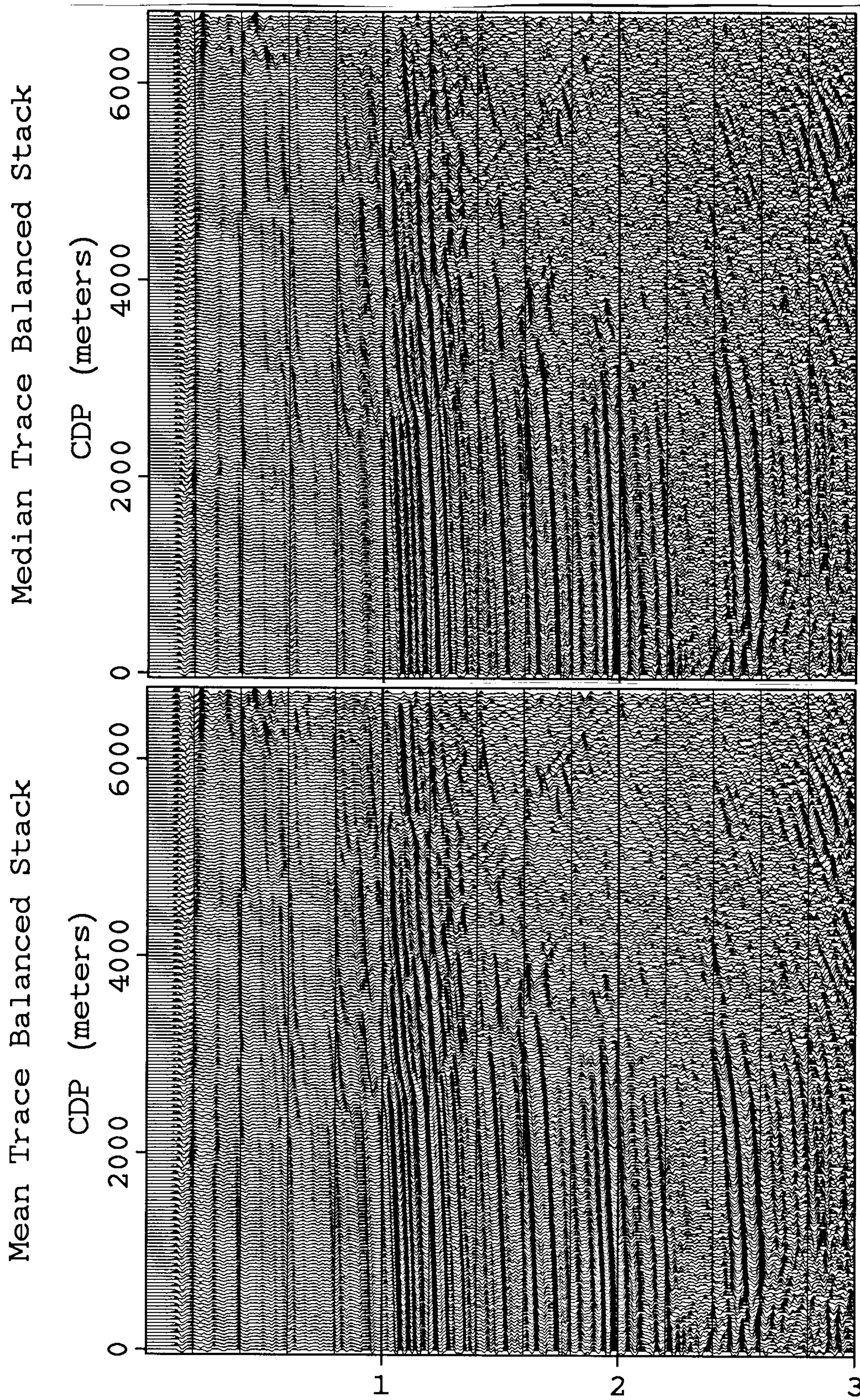

FIG. 6: Mean and median stacks of trace balanced marine data.

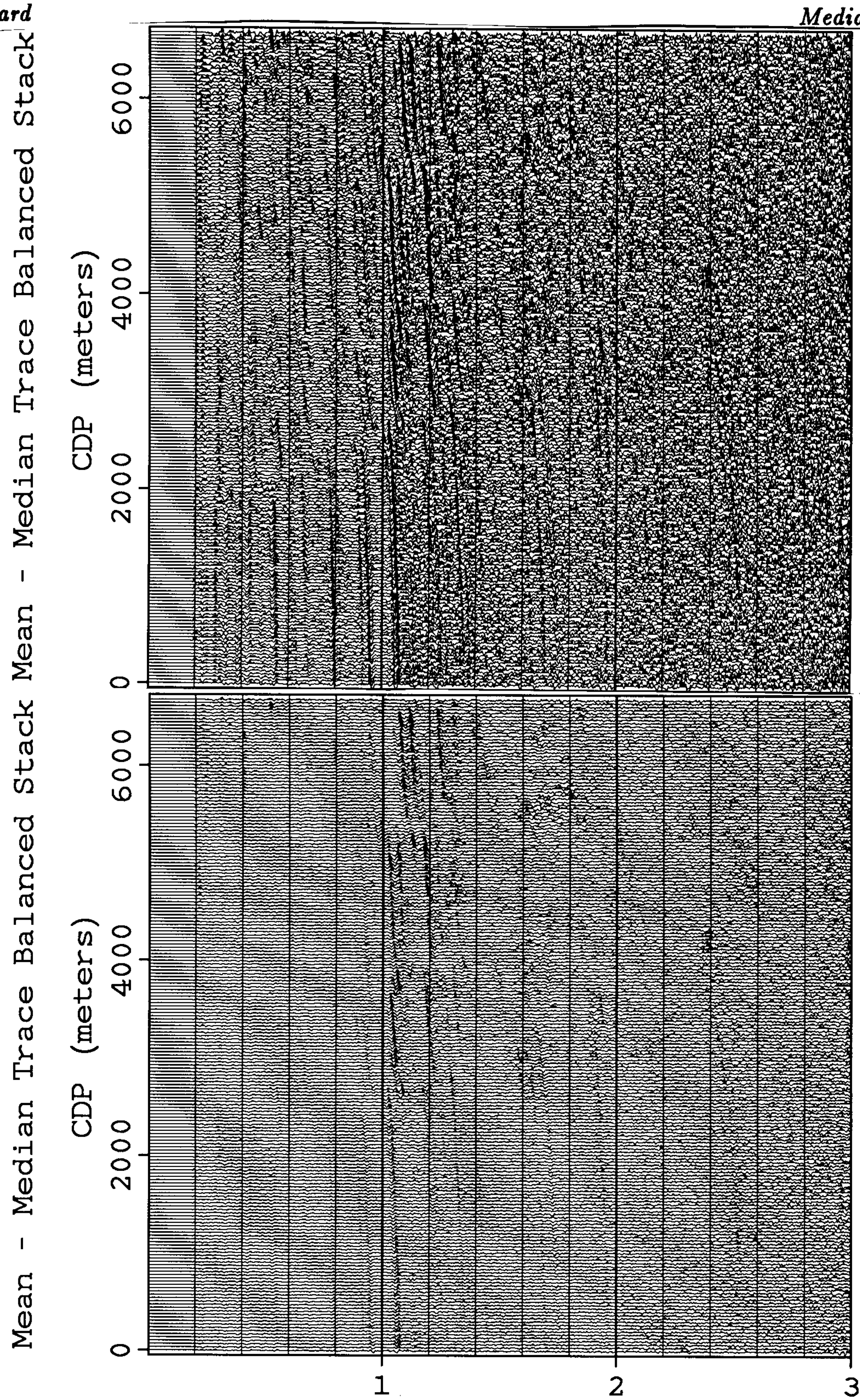

FIG. 7: The difference between the mean and the median for the trace balanced marine data set, plotted on the left at the same clip as the stacked sections and gained on the right for display.

larly interesting because it corresponds to an event present on both the balanced and unbalanced mean stacks and absent from both of the median stacks. Because the gathers containing this disappearing event generally manifest some residual normal moveout, it is reasonable to hypothesize that this fourth contrast may be due to the differing responses of means and medians to RNMO.

Median vs. Mean Velocity Panels

The variation in the response of the two statistical averages to RNMO is best examined through observations of mean and median velocity analyses—formed from constant velocity mean and median stacks. Figures 8 (unnormalized) and 9 (trace equalized) show mean and median velocity analyses for the marine data set averaged over five neighboring CDP's, with velocity sampled evenly in *slowness* **2. As predicted above, the median stack appears more sensitive to RNMO than the mean for the unbalanced trace analyses: for Figure 8, a given error in velocity would reduce the stacked amplitude of an event more for the median than for the mean. Unfortunately, for the balanced traces in Figure 9 the contrast between median and mean velocity resolution is less clear, and other explanations become possible.

Horizontal smearing of mean velocity analyses results from the flatness of the tops of normal moveout hyperbolas; because the events on inner offsets fail to respond to changes in moveout velocity, they contribute to mean stack power at a wide range of velocities and velocity resolution is lost (Thorson, 1984). If the contrast between the mean and median unnormalized trace velocity analyses reflects real differences in the response of the statistics to RNMO, then the median would have to respond differently than the mean to the inner offsets. This difference is hard to conceptualize; perhaps in selecting a single, middle value across offset, the median is less influenced by inner offsets—smearing is reduced and resolution is increased. Alternatively, if the contrast between the mean and the median is due instead to the presence of high coherent event amplitudes on the inner offsets of this particular data set, then the difference between the unnormalized mean and median velocity analyses would have to be explained as an example of the spike-rejecting habit of the median. Given the great jump in resolution between the mean unnormalized and normalized trace velocity analyses, the latter explanation appears most likely. Note resolution for both mean and median velocity analyses decreases at late times when truncation effects dominate.

Thorson (1984) suggests velocity smearing may be reduced by using weighted mean stacks to diminish the influence of inner offsets. The rightmost plots in Figures 8 and 9 show weighted mean constant velocity stacks formed by weighting each trace in a gather by the magnitude of its offset:

$$weighted\ mean = \frac{\sum_{i=1}^{N} h_i a_i}{\sum_{i=1}^{N} h_i}$$

(a_i represents signal amplitude across $i = 1 \cdots N$ offsets of magnitude h_i). The results display

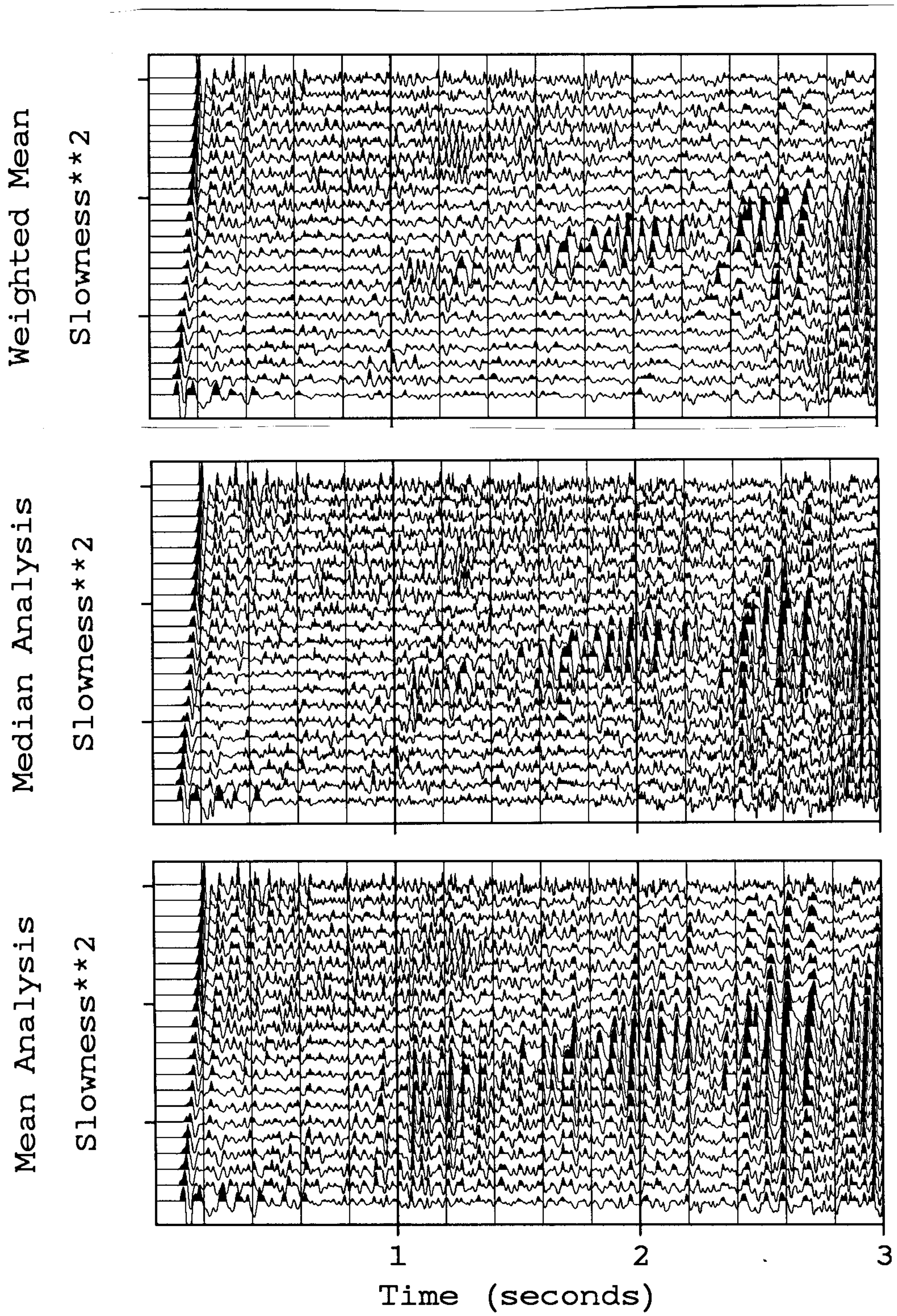

FIG. 8: Mean, median and weighted mean velocity analyses for the marine data set, sampled evenly in slowness**2 from $(1/1500)^2$ sec^2/m^2 on the left to 0 on the right and plotted at the same clip.

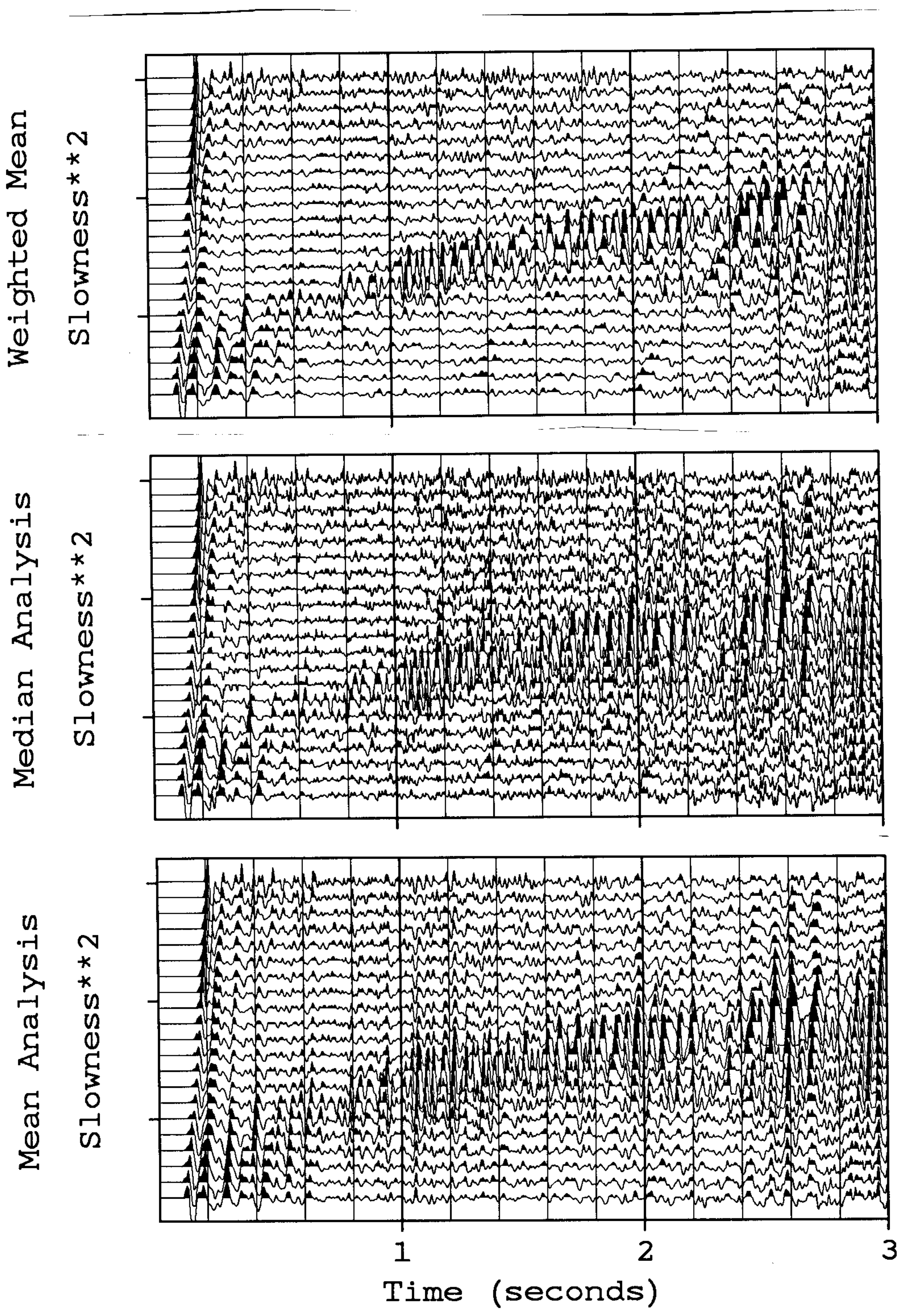

FIG. 9: Mean, median and weighted mean velocity analyses for the trace balanced marine data set, sampled evenly in slowness**2 from $(1/1500)^2$ sec^2/m^2) on the left to 0 on the right and plotted at the same clip.

smearing on the order of the median stack examples, with a strong reduction in the level of background noise. Data dependent optimum-weight averaging schemes (Larner, 1979) are beyond the scope of this paper.

Median vs. Mean Semblance

The problem of horizontal smearing in velocity analysis is generally avoided by using semblance instead of constant velocity stacks. Defined as output energy divided by input energy, semblance is effectively a measure of signal variance over offset (Neidell and Taner, 1971)—and consequently it is far less sensitive to locally flat hyperbola tops. For a mean stack, output energy is the mean squared and semblance is described by

$$\frac{\left(\sum_{i=1}^{N} a_i\right)^2}{N \sum_{i=1}^{N} a_i^2};$$

for a median stack, output energy is the median and semblance is described by

$$\frac{N\,(median\{a_i\})^2}{\sum_{i=1}^{N} a_i^2}$$

(where a_i is defined as above). Contour plots of mean and median semblance corresponding to the velocity analyses of Figures 8 and 9 are shown in Figures 10 and 11, for normalized and unnormalized traces, respectively. Here the unnormalized trace median shows limited improvement of large scale velocity resolution and little or no effect on small scale resolution; the trace normalized mean and median semblances are almost equivalent. Velocity models constructed from the four contour plots would be identical. This result is not surprising, since comparing mean and median semblance is like comparing mean and median absolute values. For this data set: trace balanced input produces median and mean stacked powers that are roughly equivalent; unnormalized trace input produces median stacked powers that are less than their mean equivalents.

Conclusion

The differences between the mean and the median as statistical averages for velocity analysis and stack may be exploited to achieve differing processing goals. For data sets with random bad traces, the robustness of the median serves as an automatic editing routine, rejecting high amplitude noise. Similarly, for poorly balanced gathers, the median rejects high amplitude coherent events from overly powerful traces—yielding a potentially weaker stack. If the high power traces were on inner offsets, the median would then evidence greater velocity resolution than the mean; if the traces were on outer offsets, the median would be less sensitive to NMO errors. Although there was some suggestion of the possibility of the median being more sensitive to RNMO than the mean, no final conclusions were formed.

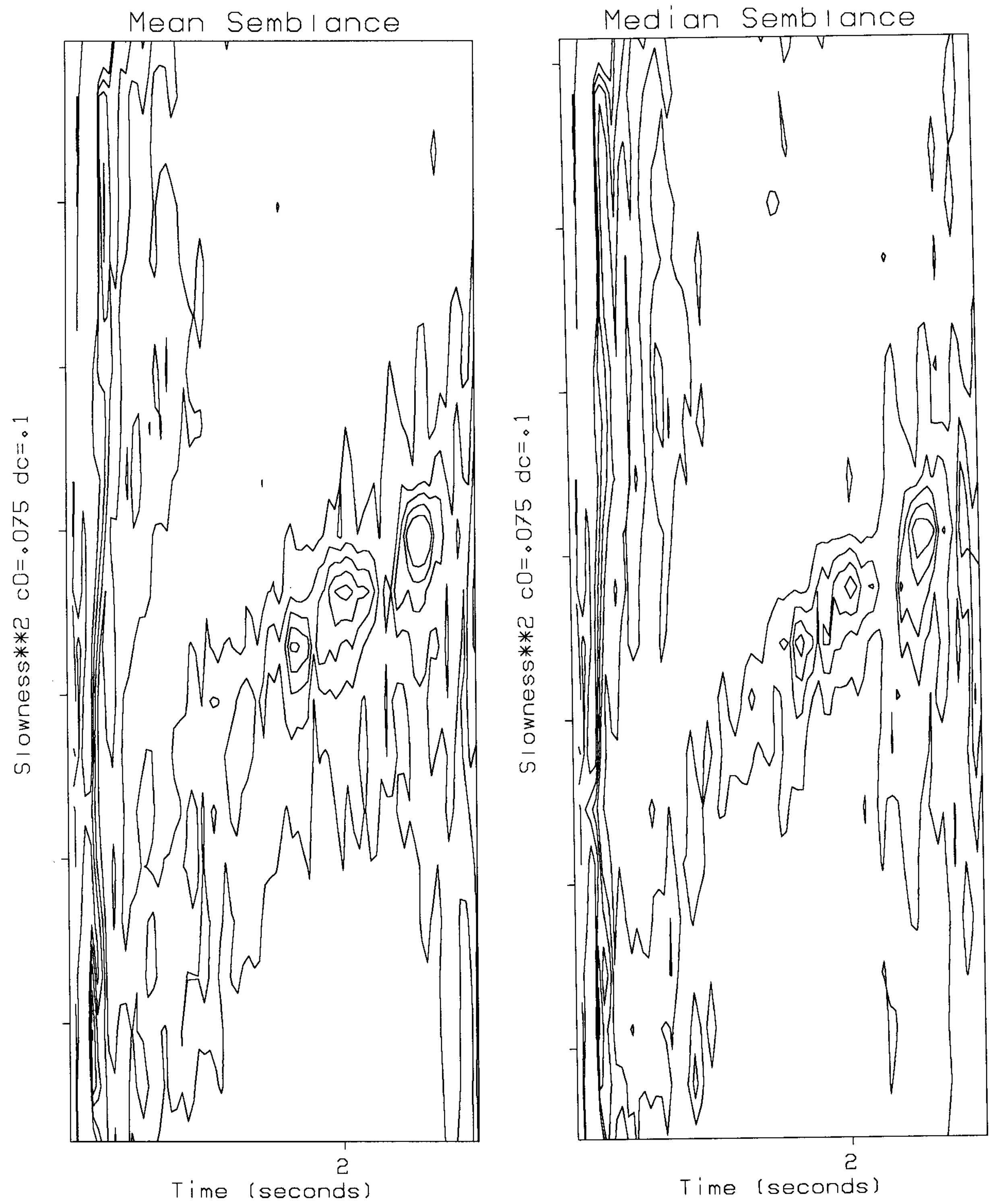

FIG. 10: Mean and median semblance for the marine data set, unnormalized traces. Semblance was calculated over 20 point (80 msec) overlapping time windows, with slowness ranging from $(1/1500)^2$ sec^2/m^2 on the left to 0 on the right

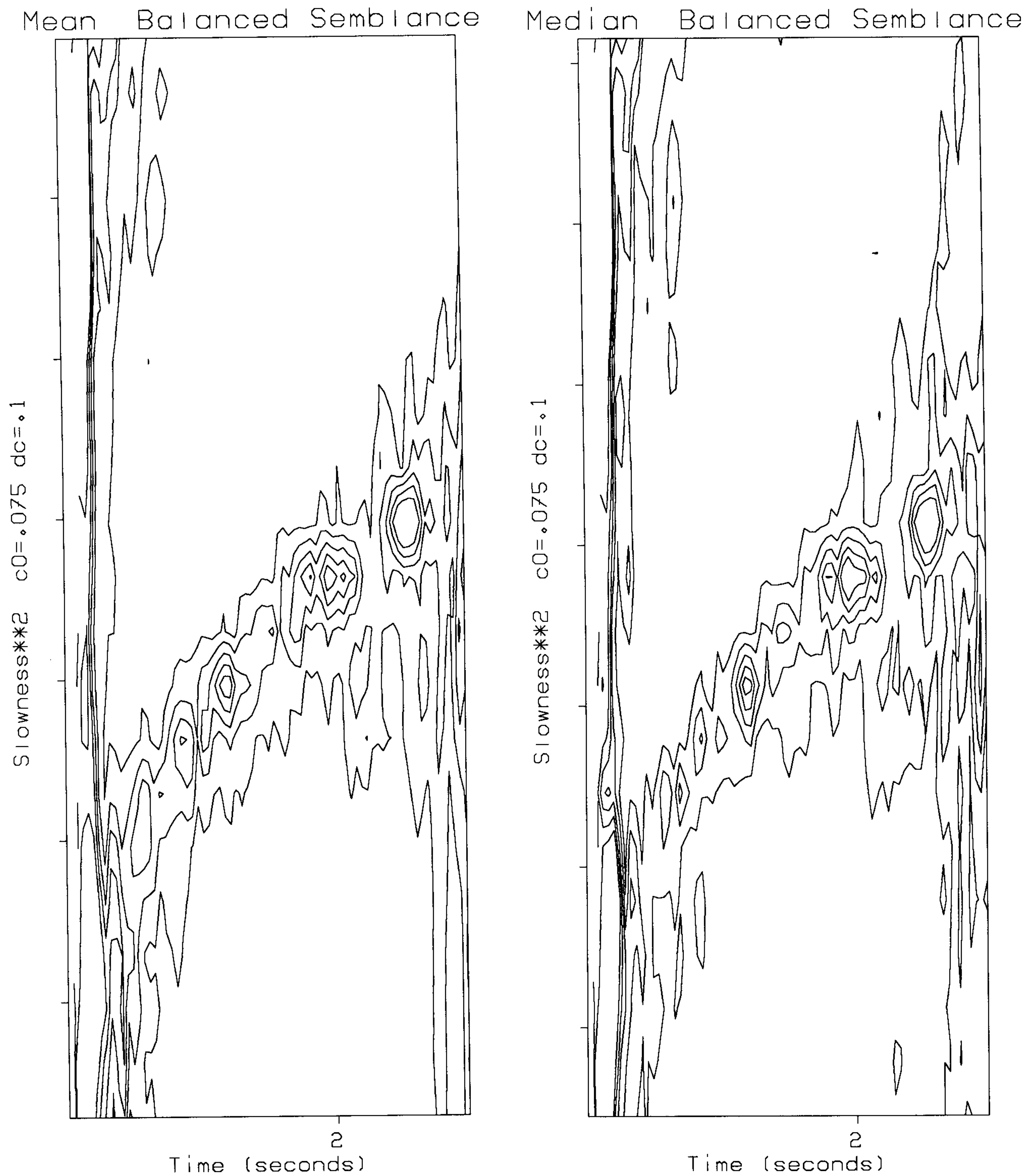

FIG. 11: Mean and median semblance for the marine data set, trace balanced prior to stack. Semblance was calculated over 20 point (80 msec) overlapping windows, with slowness ranging from $(1/1500)^2$ sec^2/m^2 on the left to 0 on the right

References

Claerbout, J.F., and Muir, F., 1973, Robust modeling with erratic data: Geophysics, v. 38, p. 826-847.

Gray, W., 1979, Variable norm deconvolution, SEP-19.

Larner, K., 1979, Optimum-weight CDP stacking, unpublished notes, Western Geophysical Co.

Neidell, N.S., and Taner, T., 1971, Semblance and other coherency measures for multichannel data: Geophysics, v. 36, p. 482-497.

Thorson, J.R., 1984, Velocity and slant stack inversion, SEP-39

c13=258.00 zet=480.00
c11=745.00 c33=592.00 eps=523.00
c44=222.00 c66=300.00 del=370.00

Cotton Valley shale