

Multiple realizations: Model variance and data uncertainty

Robert G. Clapp¹

ABSTRACT

Geophysicists typically produce a single model, without addressing the issue of model variability. By adding random noise to the model regularization goal, multiple equiprobable models can be generated that honor some *a priori* estimate of the model's second-order statistics. By adding random noise to the data, colored by the data's covariance, equiprobable models can be generated that give an estimate of model uncertainty resulting from data uncertainty. The methodology is applied to a simple velocity inversion problem with encouraging results.

INTRODUCTION

Risk assessment is a key component to any business decision. Geostatistics has recognized this need and has introduced methods, such as simulation, to attempt to assess uncertainty in their estimates of earth properties. Geophysics has been slower to recognize this need, as methods which produce a single solution have long been the norm.

The single solution approach, however, has a couple of significant drawbacks. First, because least-square estimates give invert for the minimum energy/variance solution, our models tend to have low spatial frequency than the true model. Second, it does not provide information on model variability or provide error bars on the model estimate. Geostatisticians have both of these abilities in their repertoire through what they refer to as "multiple realizations" or "stochastic simulations." They introduce a random component, based on properties of the data, such as variance, to their estimation procedure. Each realization's frequency content is more representative of the true model's and by comparing and contrasting the equiprobable realizations, model variability can be assessed.

In previous works (Clapp, 2000, 2001), I showed how we can modify standard geophysical inverse techniques by adding random noise into the model styling goal to obtain multiple realizations. Claerbout (1998) shows how an estimate of the correct scaling for the random noise can be obtained for the missing data problem. In this paper, I extend this early work. I show how we can modify the data fitting goal in a parallel manner achieving a potentially more significant result. By adding random noise colored by the inverse data covariance, we can obtain multiple model estimates that show how data uncertainty map to model uncertainty. In terms of velocity estimation we can rephrase this relation as how our uncertainty in semblance

¹email: bob@sep.stanford.edu

picks affects our estimate of interval velocity.

In this paper I begin by reviewing the methodology of adding noise to the model styling goal. I show why and when the methodology is effective. I then show how we can encode data uncertainty into our model estimation. I conclude by demonstrating both techniques on simple 1-D and 2-D tomography problems.

REVIEW: INCORPORATING MODEL VARIANCE

In Clapp (2001), I began with the standard geophysical problem, a linear relationship \mathbf{L} between a model \mathbf{m} and \mathbf{d} , with a regularization operator \mathbf{A} , written in terms of fitting goals as:

$$\begin{aligned}\mathbf{0} &\approx \mathbf{d} - \mathbf{Lm} \\ \mathbf{0} &\approx \epsilon \mathbf{Am}.\end{aligned}\tag{1}$$

To produce models with appropriate levels of variance we modify the second goal, replacing the zero vector $\mathbf{0}$ with standard normal noise vector \mathbf{n} , scaled by some scalar σ_m ,

$$\begin{aligned}\mathbf{0} &\approx \mathbf{d} - \mathbf{Lm} \\ \sigma_m \mathbf{n} &\approx \epsilon \mathbf{Am}.\end{aligned}\tag{2}$$

For the special case of missing data problem, Claerbout (1998) shows how to get an approximate value for σ_m . This paper suggests we first solve the simple problem described by

$$\mathbf{d} \approx \mathbf{Im},\tag{3}$$

where \mathbf{I} is the identity matrix and we do not allow the \mathbf{m} to change at known locations. For example, let's assume that as our input we have the data in Figure 1. We estimate a model using the fitting goal (3) and obtain Figure 2. We then look at the residuals \mathbf{r}_{data} at known locations, Figure 3. The variance of the residual should have the same variance as the random noise in (2). We can then estimate many different models by applying,

$$\begin{aligned}\mathbf{0} &\approx \mathbf{d} - \mathbf{Jm} \\ \sigma_m \mathbf{n} &\approx \epsilon \mathbf{Am},\end{aligned}\tag{4}$$

where \mathbf{J} is a selector matrix, 1 at known locations, 0 elsewhere, and \mathbf{A} is a Prediction Error Filter (PEF) estimated from the known data locations. Figure 4 shows three such realizations.

Comparison to geostatistics

The most common method to create models that honor both first and second order statistics is geostatistical simulation. A comparison between the method presented in this paper and the standard geostatistical technique is therefore warranted.

Figure 1: A wood texture from Claerbout (1998) with a hole removed to simulate a missing data problem. `bob1-wood.hole` [ER]

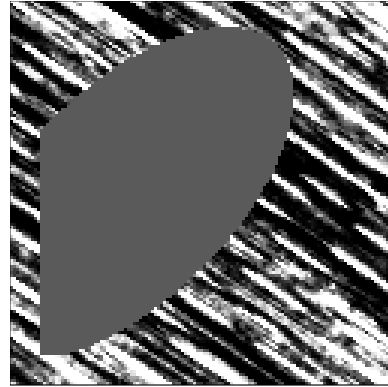


Figure 2: The result of solving a missing data problem using the data in Figure 1 as input, a PEF found from the known data locations, and fitting goal (3). `bob1-wood.pef` [ER,M]

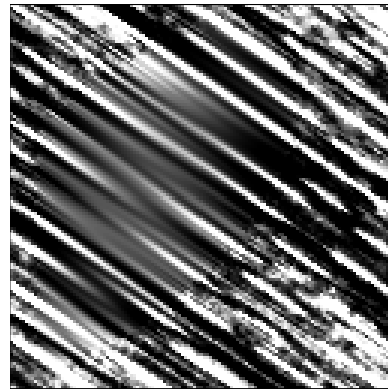


Figure 3: Left panel is the histogram of the known data locations shown in Figure 1. The right panel is the histogram of the residual at the same known locations after applying fitting goal (3). Note how the almost uniform distribution of the data approaches a normal distribution. `bob1-wood.histo` [ER,M]

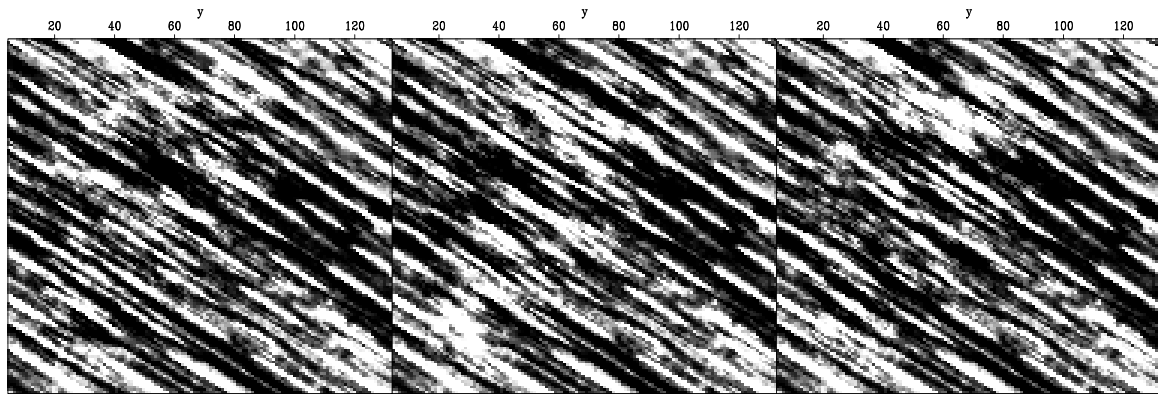
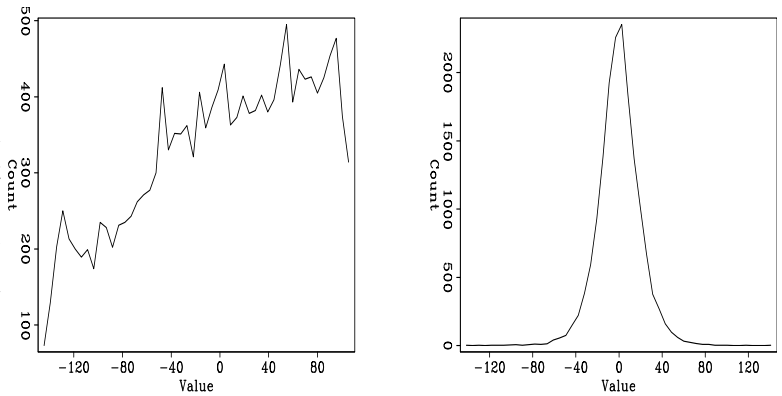


Figure 4: Three realizations using fitting goals (4) and the input data shown in Figure 1. `bob1-wood-multi` [ER,M]

For the comparison I will begin by describing the general procedure in sequential Gaussian simulation. First, some decision on stationarity is made. The known data is transformed to a new function that is standard normal. The covariance is described through variograms. Each point in the model space is randomly visited. At each point they find the best linear combination of known data, and points previously estimated through kriging (Isaaks and Srivastava, 1989), based on the predetermined variogram model. At the same time they calculate a kriging variance, which can be considered a measure of the uncertainty of the estimated value based on the surrounding points. A random number is then selected from a distribution function that has the calculated kriging mean and variance.

Each step in the geostatistical simulation procedure has a corollary in an operator based simulation. We begin by making a decision on stationarity. If the problem is non-stationary, we can break the problem into patches (Claerbout, 1992) (what geostatisticians call “distinct sub-zones”) or we can use a non-stationary covariance description (Clapp et al., 1997; Crawley, 2000) (geostatisticians use kriging with external drift to allow some degree of non-stationarity). We transform into a space that is Gaussian (if we have an accurate description of the model’s covariance function, the residual space will have a Gaussian distribution). The operator approach solves a global, rather than a local, problem. The kriging estimate corresponds to our PEF estimated with a $\mathbf{0}$ in the model styling residual. Finally, the kriging variance corresponds to the variance of the residual in the PEF estimation.

Uses

Introducing model variance into the estimation process has several attractive properties. For missing data problems, we can produce models that have a more realistic texture and behave more accurately when used as inputs as for processes such as fluid flow. We can quickly test the accuracy of our covariance description by applying

$$\mathbf{m}_{\text{test}} = \mathbf{A}^{-1} \sigma_{\text{model}}(\mathbf{n} + \mathbf{r}_{\text{data}}), \quad (5)$$

where \mathbf{m}_{test} is our estimated model. Figure 5 shows the application of (5) to the Seabeam estimation problem (Claerbout, 1998; Crawley, 1995). With the correct covariance description, we get a believable estimate; the incorrect description gives a far less satisfactory answer.

When we have a more complex mapping operator, introducing model variability can help us understand the null space of our mapping operator. This can be potentially interesting for understanding Amplitude Versus Offset (AVO) effects from standard velocity estimation techniques (Mora and Biondi, 1999, 2000).

What the method doesn’t address is a potentially more interesting question: how does data uncertainty map to model uncertainty? In the next section, I will propose a methodology for addressing this issue.

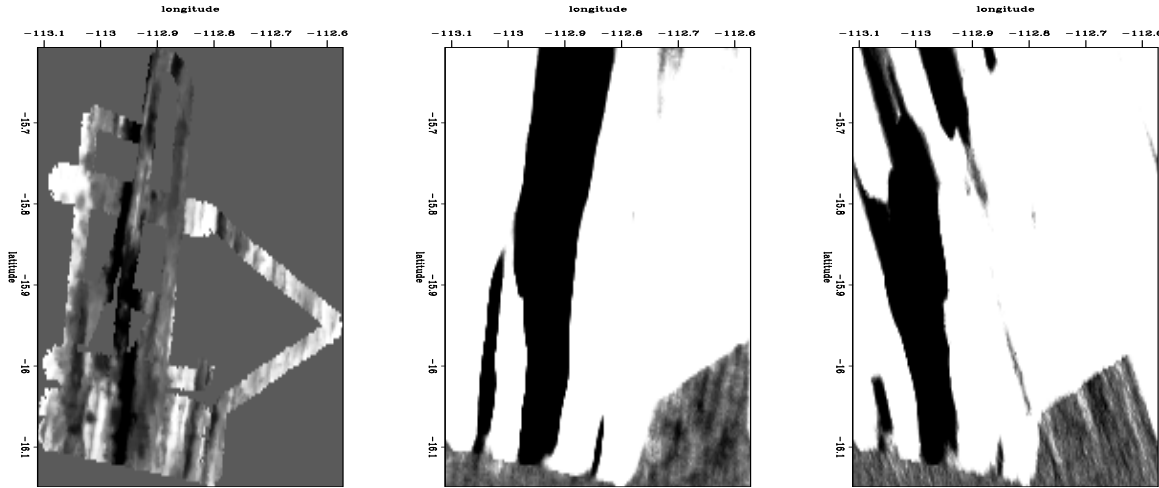


Figure 5: A fast way to check PEFs. The left is the input data. The center panel shows the result of applying fitting goals (5) with a reasonable covariance description. The right shows the result using an unrealistic covariance description. `bob1-fast` [ER,M]

DATA UNCERTAINTY

To understand the method that I am proposing, let's begin by rewriting our standard fitting goals. We normally write

$$\begin{aligned} \mathbf{d} &\approx \mathbf{Lm} \\ \mathbf{0} &\approx \epsilon \mathbf{A} \mathbf{m}. \end{aligned} \tag{6}$$

We can rewrite these, adding our model variance information as

$$\begin{aligned} \mathbf{d} &\approx \mathbf{Lm} \\ \sigma_n \mathbf{n} &\approx \epsilon \mathbf{A} \mathbf{m}, \end{aligned} \tag{7}$$

or put another way

$$\begin{aligned} \mathbf{0} &\approx \mathbf{d} - \mathbf{Lm} \\ \mathbf{0} &\approx \epsilon \mathbf{A}(\mathbf{m} + \mathbf{m}_u), \end{aligned} \tag{8}$$

where $\mathbf{A} \mathbf{m}_u = -\sigma_n \mathbf{n}$ is the model variability not characterized by \mathbf{A} .

There is a corollary way to think about data variance. Normally we limit our characterization of data covariance to a diagonal matrix often referred to as the weight operator \mathbf{W} . A more appropriate choice is the data covariance matrix. We can rewrite our system of equations and add in an appropriate level of data variance:

$$\begin{aligned} \mathbf{0} &\approx \mathbf{W}(\mathbf{d} + \mathbf{d}_u - \mathbf{Lm}) \\ \mathbf{0} &\approx \epsilon \mathbf{A}(\mathbf{m} + \mathbf{m}_u). \end{aligned} \tag{9}$$

Or potentially more conveniently as a cascade of two operators

$$\mathbf{W} = \mathbf{D}\mathbf{H}, \quad (10)$$

where \mathbf{D} is a diagonal weighting matrix used before, and \mathbf{H} is the normalized data inverse covariance matrix, which we could characterize through a filtering operation.

Given this formulation we can define perturbed \mathbf{d}_{new} through

$$\mathbf{d}_{\text{new}} = \mathbf{d} + \mathbf{W}^{-1}\sigma_d\mathbf{n}, \quad (11)$$

where σ_d is our data variance. Assuming appropriate choices for \mathbf{D} and \mathbf{H} , the data for realization will have the correct structure, something that would happen if we simply added random noise to the data vector.

SUPER DIX

To test the methodology, I will return to the Super Dix example (Clapp et al., 1998; Rosales, 2000; Clapp, 2000). In 1-D we write the Super Dix fitting goals as

$$\begin{aligned} \mathbf{0} &\approx \mathbf{W}(\mathbf{d} - \mathbf{C}\mathbf{C}\mathbf{p}) \\ \mathbf{0} &\approx \epsilon\mathbf{p}, \end{aligned} \quad (12)$$

where \mathbf{C} is causal integration, the data is the $d(i)\sum_j^i v_{rms}^2(j)$, our model is $\mathbf{v}_{\text{int}}^2 = \mathbf{C}\mathbf{p}$, and we do not allow the model estimate to change at the first time sample. For this example, I approximated the inverse data covariance by a simple derivative in time. Figures 6 and 7 show the effect of introducing model and data variance. Note how, as expected, increasing model variance (Figure 6) produced higher frequency interval velocity estimates, but the general trends of the curves are preserved. When increasing data variance (Figure 7), we maintain approximately the same smoothness, but our estimate of approximate layer velocities change.

2-D example

We can extend the basic 1-D formulation into 2-D. We redefine our data by

$$\mathbf{d}_1 = \mathbf{d} - \mathbf{C}\mathbf{C}\mathbf{v}_0, \quad (13)$$

where \mathbf{d}_1 is our new data array and \mathbf{v}_0 is an array containing our zero time velocity at each Common Midpoint (CMP) location. The end result of equation (13) is to remove the zero frequency component of our model. Our fitting goals change to

$$\begin{aligned} \mathbf{d}_1 &\approx \mathbf{CMA}^{-1}\mathbf{p} \\ \mathbf{0} &\approx \epsilon\mathbf{p}, \end{aligned} \quad (14)$$

Figure 6: Effect of changing the model variance on interval velocity.
bob1-model-var [ER,M]

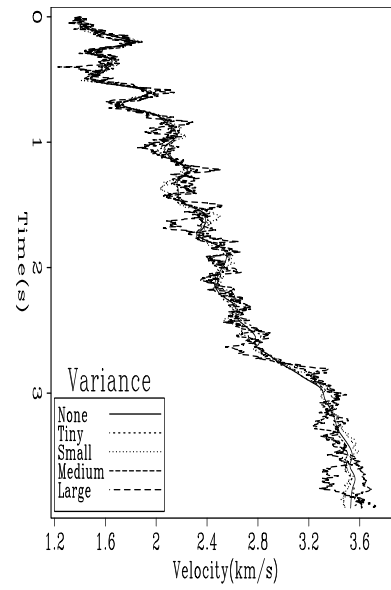
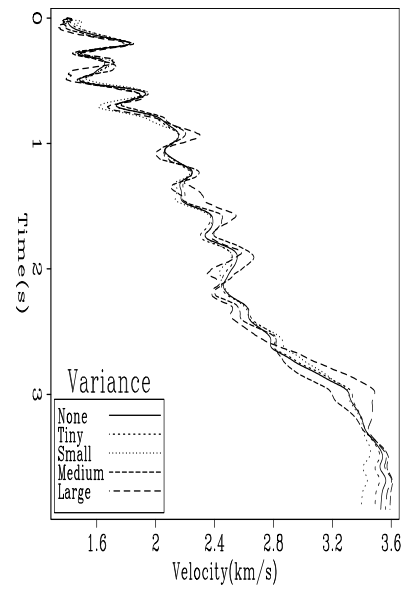


Figure 7: Effect of changing the data variance on interval velocity.
bob1-data-var [ER,M]



where \mathbf{M} is a masking operator that doesn't allow our velocity estimate to change at zero time and \mathbf{A}^{-1} is our model covariance estimate. After estimating \mathbf{p} we can convert back to interval velocity through

$$\mathbf{v}_{\text{int}} = \sqrt{\mathbf{C}\mathbf{v}_0 + \mathbf{A}^{-1}\mathbf{p}}. \quad (15)$$

The modifications to the original fitting goals is necessary because of the constraint that we do not change our velocity at $t = 0$. In the 1-D case, we didn't have to worry about our change of variable from $\mathbf{v}_{\text{int}}^2$ to \mathbf{p} because our preconditioned operator did not modify the zero time. In the 2-D case our \mathbf{A}^{-1} operator can introduce smoothness laterally as well as vertically, forcing us to modify our inversion scheme in order to take advantage of the preconditioning speed up.

To test the methodology I chose the relatively simple Gulf of Mexico dataset provided by Western and used in Claerbout (1995). Figure 8 shows a near offset section from the data. I performed semblance analysis and chose a fairway within which all valid \mathbf{v}_{rms} picks would fall (Figure 9). For each CMP, I automatically picked the maximum semblance at each time within the fairway (Figure 10). For my diagonal operator \mathbf{D} , I used the amplitude of the semblance at the picked v_{rms} (left panel of Figure 11) and then scaled by $1/t$ to correct for \mathbf{d} being the result of summing operation. Figure 12 shows the result of estimating an interval

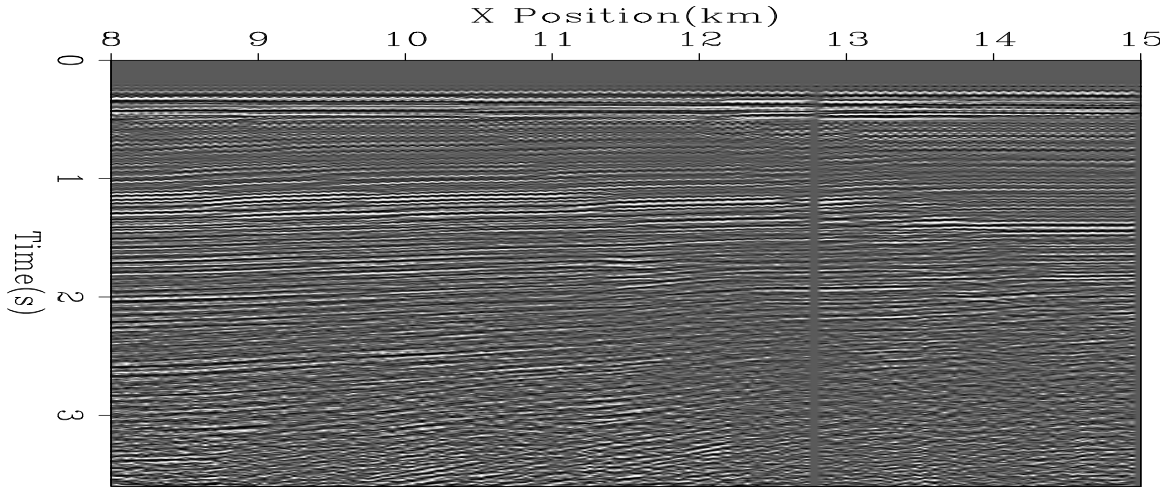


Figure 8: Near offset section of a Gulf of Mexico dataset. `bob1-beidata` [ER]

velocity without adding any model variance or data uncertainty. Figure 13 shows two different realizations for the interval velocity adding model variability. I estimated a PEF from my v_{rms} picks and used filtering with that PEF for \mathbf{H} and polynomial division with it for \mathbf{H}^{-1} . For my data variance I used the width of the picked semblance block (the right panel of Figure 11). Note how, as expected, the variance generally increases with depth. For my different data realizations I applied equation (14). Figure 14 shows two different realizations of the interval velocity using the data with the added uncertainty.

Figure 9: Semblance analysis for a CMP from the data shown in Figure 8. Overlaid is the fairway used by the automatic picker. `bob1-fairway` [ER]

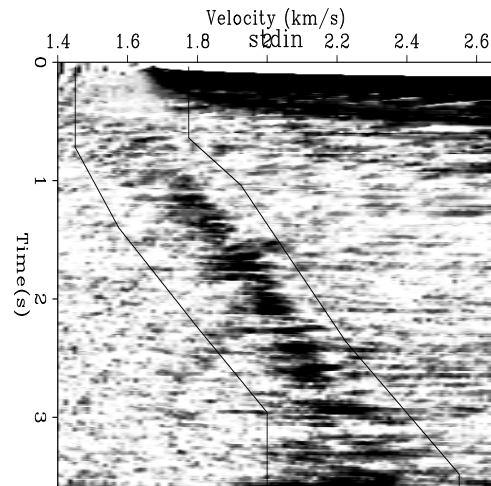
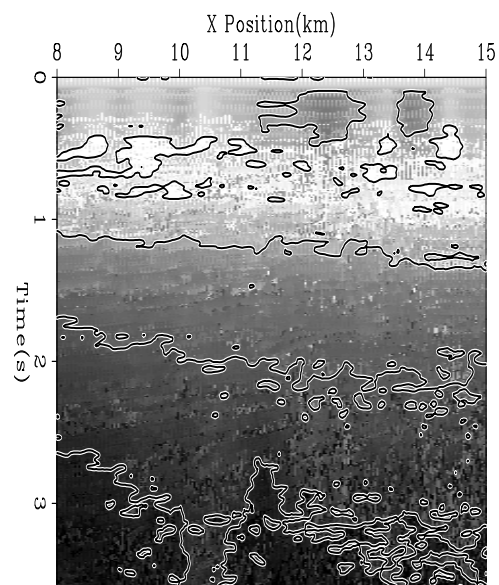


Figure 10: The automatically picked velocities for each CMP within the fairway shown in Figure 9. Overlaid are smoothed contours for the same field. `bob1-rms` [ER]



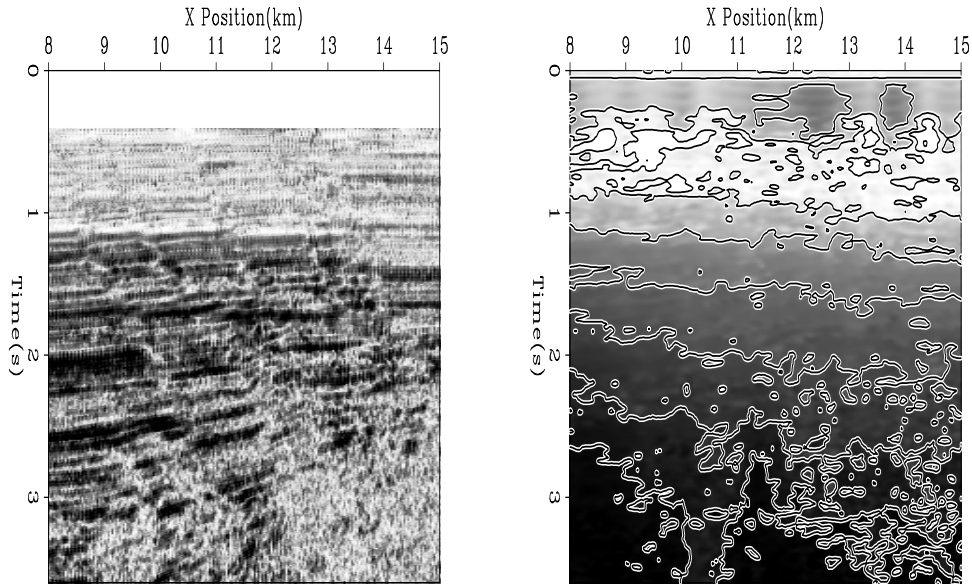


Figure 11: The left panel is the amplitude of the semblance at each v_{rms} shown in Figure 10. The right panel is the approximate width of the corresponding semblance blob at the same location. Overlaid on the right plot are contours for the same field. `bob1-range-wt` [ER,M]

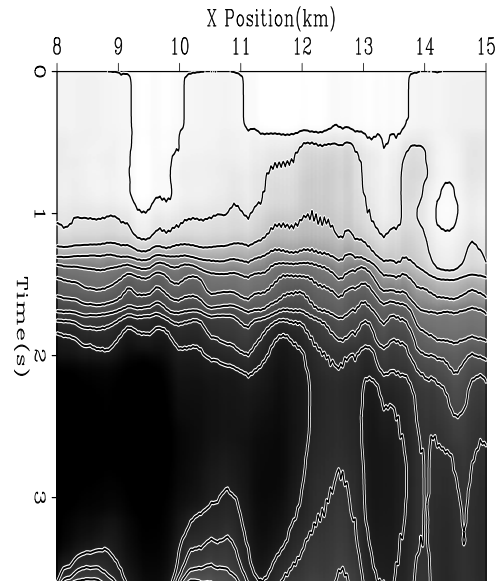


Figure 12: The result of estimating for an interval velocity from the v_{rms} picks in Figure 10 without adding any model variance or data uncertainty. Overlaid are contours for the same field. `bob1-int` [ER]

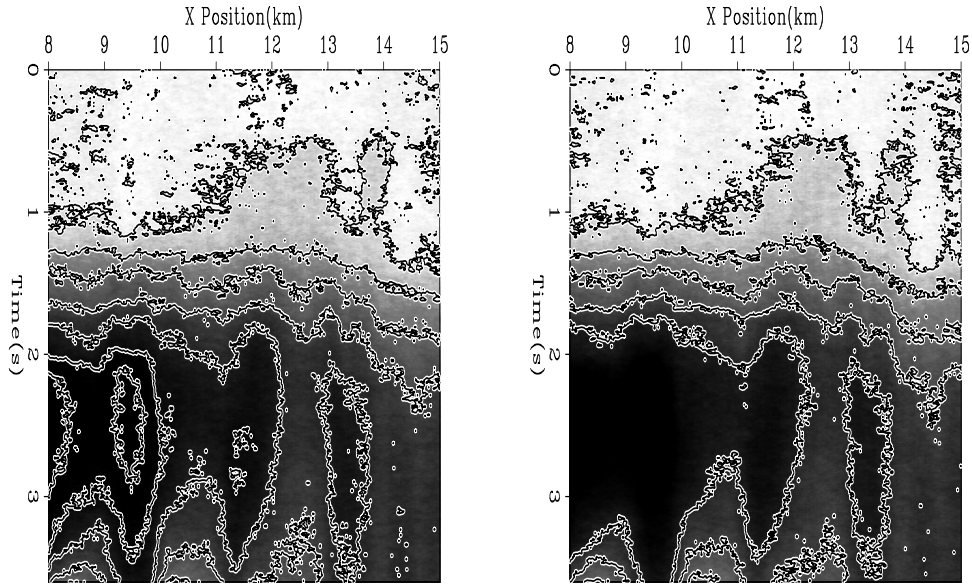


Figure 13: Two different realizations for the interval velocity adding model variability. Overlaid are contours for the same fields. `bob1-model-var-dix2d` [ER,M]

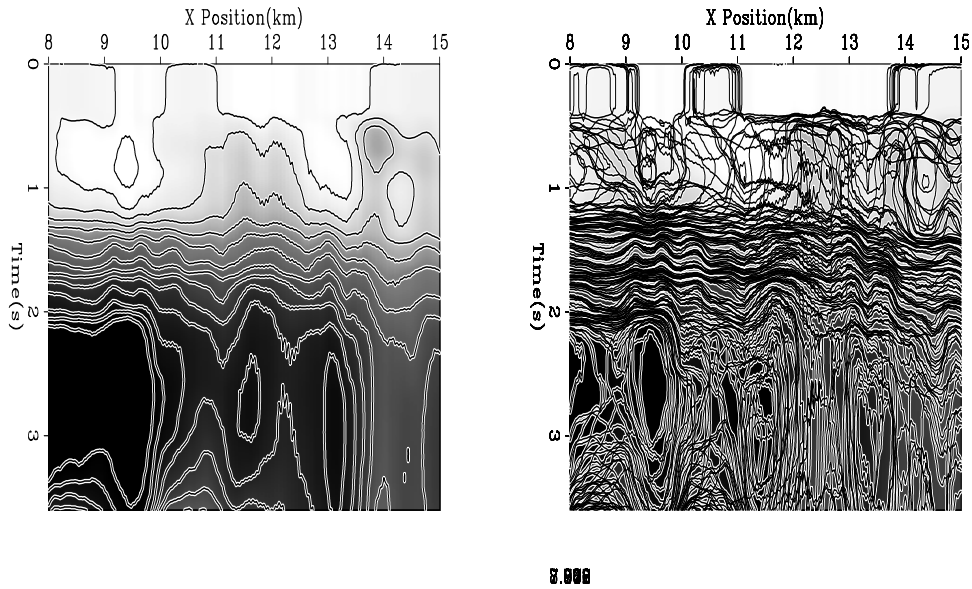


Figure 14: Two different realizations for the interval velocity adding data variability. Overlaid are contours for the same fields. `bob1-data-var-dix2d` [ER,M]

CONCLUSIONS

I have demonstrated a method for creating equiprobable realizations that shows how data uncertainty affects model uncertainty.

ACKNOWLEDGMENTS

I would like to thank Morgan Brown, Jon Claerbout, and Sergey Fomel for many useful conversations.

REFERENCES

- Claerbout, J. F., 1992, Nonstationarity and conjugacy: Utilities for data patch work: SEP-73, 391–400.
- Claerbout, J. F., 1995, Basic Earth Imaging: Stanford Exploration Project.
- Claerbout, J. Geophysical Estimation by Example: Environmental soundings image enhancement: <http://sepwww.stanford.edu/sep/prof/>, 1998.
- Clapp, R. G., Fomel, S., and Claerbout, J., 1997, Solution steering with space-variant filters: SEP-95, 27–42.
- Clapp, R. G., Sava, P., and Claerbout, J. F., 1998, Interval velocity estimation with a null-space: SEP-97, 147–156.
- Clapp, R., 2000, Multiple realizations using standard inversion techniques: SEP-105, 67–78.
- Clapp, R. G., 2001, Geologically constrained migration velocity analysis: Ph.D. thesis, Stanford University.
- Crawley, S., 1995, Multigrid nonlinear SeaBeam interpolation: SEP-84, 279–288.
- Crawley, S., 2000, Seismic trace interpolation with nonstationary prediction-error filters: Ph.D. thesis, Stanford University.
- Isaaks, E. H., and Srivastava, R. M., 1989, An Introduction to Applied Geostatistics: Oxford University Press.
- Mora, C., and Biondi, B., 1999, Seismic attribute sensitivity to velocity uncertainty: SEP-102, 157–176.
- Mora, C., and Biondi, B., 2000, Estimation of AVO attributes sensitivity to velocity uncertainty using forward modeling: A progress report: SEP-103, 349–366.
- Rosales, D. A., 2000, Research interest: Improving the velocity model: SEP-103, 69–78.