Constrained inversion of reflection data using Gibbs' sampling

Danilo R. Velis*

Abstract

A procedure for the inversion of reflection data (zero offset section) is presented. The strategy is based on the generation of a 2-D Markov Random Field (MRF) with a certain spatial continuity that is appropriate to represent the underlying velocity model. The MRF, which is generated using the Gibbs' sampler, is forced to fit the data and to satisfy a set of well constraints during its (iterative) generation process. The spatial continuity of the MRF is implemented by defining a neighborhood system with associated potentials that favor the formation of regions with similar velocity values. Thus, the resulting MRF exhibits spatial continuity, fits the data and honors the well constraints. The algorithm is specially suited for obtaining high resolution velocity images dominated by horizontal layers or other constant velocity blocks. Moderate dipping layers are also tolerated. We tested the algorithm using various velocity models, including the hard Marmousi model. The results show that realistic high resolution blocky images can be recovered accurately. Since the generation of a MRF is a stochastic process, the uncertainty of the estimated models can be analyzed by performing several realizations.

KEYWORDS: inversion, Gibbs's sampler, Markov Random Field, Marmousi

*Facultad de Ciencias Astronómicas y Geofísicas, Universidad Nacional de La Plata, Argentina; and CONICET, Argentina

1 Introduction

The estimation of the earth model from measured reflection data represents a fundamental problem in seismic exploration. This inverse problem is usually non-unique because there are many models that fit the data equally well. Often, the non-uniqueness comes from the fact that the data is finite (and inaccurate) but the actual model space that represents the earth subsurface as a continuum is not. Fortunately, the non-uniqueness may be alleviated by selecting an appropriate parameterization so that the model space dimension is reduced significantly. Also it is possible to regularize the inversion within the numerical algorithm itself to guide the solution towards models with certain desired properties (see for example Tarantola (1997)). These regularization strategies are devised to discard implausible models and to select, among all plausible models that fit the data, those that are more likely to be correct. For this task it is usually necessary to assume certain a prior information that help us to impose within the mathematical framework, what type of model we would like to obtain (a smooth model, a layered model, etc.). The selected a prior information plays a key role in the solution of the inverse problem and constitutes the basic means of regularization.

The use of constraints contribute to reduce the non-uniqueness, too. If certain parameters are known a priori, these parameters may be fixed throughout the inversion process. Actually, this is a particular type of regularization. In the present work we use velocity constraints obtained from wells. The final model is constrained to honor the well data. In addition we regularize the inversion (in between wells) by modeling the velocity field via a discrete 2–D Markov Random Field (MRF) which exhibits a certain spatial continuity in the form of blocks that represent layers or other constant velocity regions (this constitutes the prior). For simplicity, the number of possible velocity values is fixed within a given range. This type of strategy has already been used for reflection tomography (Carrion and Pulleda, 1995; Carrion et al., 1997), where high resolution "blocky" models are obtained using Gibbs' sampling (implemented through Simulated Annealing). Gibbs' sampling has also been used for the inversion of marine reflection data (Sen and Stoffa, 1996),

where a 1–D layered model is derived using the so called heat bath algorithm with uniform prior.

In this paper we propose an algorithm based on the Gibbs' sampler to obtain 2–D "blocky" velocity images that satisfy the post-stack reflection data (zero-offset section). The solution is guided by the use of both well constraints and bounding constraints for the velocity values. The key relies in the selection of the appropriate potentials associated with the Gibbs' distribution to favor a desired spatial continuity to characterize the subsurface model in between the wells. We divide the earth into rectangular cells and assume that the velocity within each cell is constant. The size of the cells is small enough so that the discretization does not affect the final solution. Since the observations (reflection data) are naturally inaccurate, the final solution is forced to reproduce the data within a given threshold (expected misfit). Various numerical examples demonstrate the use of the Gibbs' sampler for the inversion of reflection data that yield high resolution "blocky" images of the subsurface velocity field. We test the algorithm in a simple 100×50 layered model, and in the complex 384×122 hard Marmousi model (Versteeg, 1994). Despite the fact that the procedure is rather demanding in terms of computational cost, the results are encouraging because complex realistic models can be recovered quite accurately.

2 Theory

The problem consists on finding the velocity model X that fits the data (zero offset section) and satisfies a set of constraints given by velocity values measured at a series of wells. Models with spatial continuity (specially lateral) are preferred. This is achieved by generating a Markov Random Field (MRF) with specified statistical properties.

The model may be assumed to be a MRF defined on a 2-D discrete regular lattice. In a 2-D MRF the full conditional probability

$$P(X_s = x_s | X_{S\backslash s} = x_{S\backslash s}) \tag{1}$$

where X_s is the velocity value at site s, $s \in S$, depends only on the neighborhood of s, N_s. That is

$$P(X_s = x_s | X_{S \setminus s} = x_{S \setminus s}) = P(X_s = x_s | X_r = x_r, r \neq s). \tag{2}$$

where $r \in N_s$. Usually, the neighborhood is defined in terms of *cliques* c. A clique c consists on any set of sites such that if $s_i, s_j \in c$, then s_j is in the neighborhood of $s_i, s_j \in N_{s_i}$. Also, s_i is in the neighborhood of $s_j, s_i \in N_{s_j}$.

The Hammersley-Clifford theorem states that the probability of a particular configuration obeys a Gibbs' distribution (Besag, 1974):

$$P(X) = \frac{1}{Z} \exp\left[-\frac{1}{T} U(X)\right], \tag{3}$$

where Z is a normalizing constant (partition function), T is a constant (temperature) and U is the energy function (typically defined in the neighborhood N_s). The importance of the Hammersley-Clifford theorem resides in the fact that the probability distribution $P(X)$ of any MRF is uniquely determined by the conditional probabilities. In other words, one can draw models from the Gibbs' distribution by sampling the conditional distributions. This is carried out by means of the Gibbs' sampler (Geman and Geman, 1984).

2.1 The Gibbs' sampler

The Gibbs' sampler is a technique for generating random variables from a (marginal) distribution indirectly (Geman and Geman, 1984; Casella and George, 1992). It consists on visiting all sites s, one site at a time, in a predefined order (any order is valid). Every time a site is visited, a random value is drawn from the conditional probability relative to the neighboring values (conditionals are much easier to calculate than marginals, specially in high dimensional spaces):

$$P(X_s = x_s | XX_r = x_r, r \neq s) = \frac{1}{Z_s} \exp\left[-\frac{1}{T} U(X_s)\right]. \tag{4}$$

Once all sites are visited, a sweep has been concluded. This process (sweep) is repeated a certain number of times, so that we end up with a Gibbs' sequence of random variables for every site $s \in S$ and a MRF sequence X_1, X_2, $\cdots$, X_k. It turns out that the distribution of X_k converges to $P(X)$ as $k \to \infty$. When k is large enough, $X_k = x_k$ is a sample from $P(X)$.

Usually

$$U(X_s) = \sum_{c \in C} V_c(X_s), \tag{5}$$

where V_c is the potential of clique c, and C is the set of all the cliques in the neighborhood N. The value of V_c depends on the local characteristics of the clique c; that is, its value depends only on those velocity values x_s for which $s \in C$ and it is selected so as to favor a desired behavior of the model (for example, in image de-noising and restoration, V_c will be such that images with regions of the same color are favored). So, the clique potentials describe the prior probability of a particular realization of the elements of the clique.

The partition function Z_s in (4) is given by

$$Z_s = \sum_{x_s \in \Gamma} \exp\left[-\frac{1}{T} \sum_{c \in C} V_c(X_s)\right]. \tag{6}$$

where Γ is the set of all possible values of the random variable X_s (e.g. pixel intensity in an image restoration problem). Temperature T controls the degree of "peaking" of the distribution. The smaller the T, the sharper the distribution. In general we set

$$T = \frac{T_0}{\log(1+k)}, \tag{7}$$

where T_0 is a constant and k stands for iteration (sweep).

2.2 Spatial continuity

To favor spatial continuity in the generation of models we chose neighborhood systems of first, second and third order, as depicted in Figure 1. The systems are conveniently adjusted at the boundaries.

Let $s = (i_s, j_s)$ be the current site with velocity x_s, and let $r = (i_r, j_r)$ be a neighbor site with velocity x_r. Also, let $\delta x = |x_r - x_s|$. We define the associated potentials to be

$$V_c = \alpha \frac{\delta x^p}{\delta x^p + \epsilon}. \tag{8}$$

where α, ϵ and p are predefined (positive) constants. Usually, $p = 2$ and ϵ is small compared to δx_{max}^p, where δx_{max} is the estimated maximum velocity difference. Clearly, when the neighbor velocity equals the current velocity (i.e. $x_r = x_s$), $V_c \rightarrow 0$, the corresponding term in the exponential (4) gets large and the probability increases. On the contrary, velocity values which do not contribute to the spatial continuity (i.e. $x_r \neq x_s$) are assigned low probabilities, since $V_c \rightarrow \alpha$ for large δx.

In the inversion of the reflection data, we also assign higher probabilities to those velocity values that contribute to spatial continuity in the horizontal direction (denoted by index j), relative to the spatial continuity in depth/time (denoted by index i). This is carried out by redefining (8):

$$V_c = \begin{cases} \alpha \frac{\delta x^p}{\delta x^p + \epsilon} & i_r = i_s \\ \rho\alpha \frac{\delta x^p}{\delta x^p + \epsilon} & i_r \neq i_s, \end{cases} \tag{9}$$

with $0 \leq \rho \leq 1$. For convenience we set $\alpha = (1/\Theta)/(2 + 4\rho)$, where Θ is the order of the selected neighborhood system. This ensures that $0 \leq U(x_s) \leq 1$.

Figure 2 depicts the energy function $U(x_s)$ defined in (5), which is the sum of the six clique potentials in a first order neighborhood system ($\Theta = 1$), with $\epsilon = 0.1$ and 0.5, $p = 2$, and $\rho = 1.0$ and 0.2. Only those values of x_s which exhibit a strong similarity with the neighbor values x_r will

be assigned low energy values (i.e. large probability). The neighbor values are shown in the same figure (left panel). For $x_s = 2.0$, the spatial continuity is largest, both for $\rho = 1.0$ and 0.2. For $x_s = 5.0$ the spatial continuity is assigned lower energy (even lower for $\rho = 0.2$, since this value is not at the same depth/time i_s). A maximum energy value of about 1.0 is assigned to those values for which there is no similarity with the neighbor values.

2.3 Misfit

Another important issue regarding the inversion of the reflection data is the misfit. Reflection data should be fitted within a given tolerance so that every time a site is visited, those velocity values which lead to a decrease in the misfit function are assigned higher probabilities. We define the new conditional probability

$$P(X_s = x_s | X_r = x_r, r \neq s) = \frac{1}{Z_s} \exp\left(-\frac{1}{T}\left[\beta_{j_s} U(X_s) + (1 - \beta_{j_s}) E_s\right]\right), \tag{10}$$

where E_s is the misfit (or error) function, and β_{j_s} is a constant used to balance the weight between spatial continuity and misfit at every offset. Usually $0 \leq \beta_{j_s} \leq 1$ so that low spatial continuity and large errors are discouraged. The partition function Z_s is adjusted accordingly to take into account this new term.

We define the misfit function as the normalized root mean square error

$$E_s = \sqrt{\frac{\sum_{i=1}^{N} (y_{ij_s} - y'_{ij_s})^2}{\sum_{i=1}^{N} y_{ij_s}^2}}, \tag{11}$$

where y_{ij_s} is the j_s-th trace obtained after convolving the reflectivity series derived from the velocity column vector x_{ij_s} and the known wavelet (data), and y'_{ij_s} is the trace with the velocity at depth/time i_s replaced by $x_s = x_{i_s j_s}$. N is the trace length. Note that the only difference between y and y' comes from the fact that a single velocity value has been modified. As a consequence, the

computation of the equation (11) can be simplified significantly (see Appendix).

2.4 Constraints and weighting factor

The inversion procedure honors any number of predefined well constraints by simply not changing the corresponding velocities at the appropriate sites. Assuming the well constraints are given by known velocity values at N_c offsets J_l, $l = 1, \cdots, N_c$, during every sweep of the sites in the rectangular lattice, whenever $j = J_l$, the velocity x_{ij} is kept constant. It is also possible not to honor the well constraints exactly. In this case, one can set a range of variation around the measured velocity values at these locations. In this work the constraints will be honored exactly.

The weighting factor β_{j_s} in equation (10) can be used to control how the well constraints are involved in the process of adding spatial continuity to the MRF. It seems reasonable to select a larger β_{j_s} for those offsets that are close to the known velocity values (which are honored at every sweep). On the contrary, a strong spatial continuity is not assigned to those offset which are far from the constraints because this may lead to the premature development of constant velocity regions (local minima) that may delay convergence (the MRF is poorly constrained here). However, the value of β_{j_s} at these offsets can be increased with the sweep number as the misfit decreases, so that a strong spatial continuity is also obtained at convergence.

For this purpose we define the following weighting factor, valid in the vicinity of each constraint J_l:

$$\beta_{j_s} = \beta_0 \frac{\eta}{\delta j_s^q + \eta}, \qquad \frac{J_{l-1} + J_l}{2} \leq j_s \leq \frac{J_l + J_{l+1}}{2}, \qquad l = 2, \cdots, N_c - 1, \tag{12}$$

where β_0, a user defined constant, is the maximum weight, η is a positive constant which depends on the sweep number only, $\delta j_s = |J_l - j_s|$ and $q = 2$. Clearly, when δj_s is small, $\beta_{j_s} \to \beta_0$. For offsets far from the constraint, δj_s^q becomes large compared to η and β_{j_s} is minimum. To control the way β_{j_s} decreases in between two constraints, we define

$$\eta = \beta_m \frac{(\delta_{max}/2)^q}{\beta_0 - \beta_m}, \qquad \text{with} \quad \beta_m = \beta_a + (\beta_b - \beta_a)\frac{k-1}{k_b - 1}. \tag{13}$$

Here δ_{max} is the maximum distance between any two consecutive constraints or between J_1 or J_{N_c} and the left ($j = 1$) or right ($j = M$) boundary of the model, that is

$$\delta_{max} = \max\left\{|J_l - J_{l+1}|, |J_1 - 1|, |J_{N_c} - M|\right\}, \qquad l = 1, \cdots N_c - 1 \tag{14}$$

β_a and β_b are the desired minimum and maximum values for $\beta_{j_s} = \beta_m$ at the middle point corresponding to δ_{max}, and k_b is the number of sweeps it takes β_m to go from β_a to β_b.

Figure 3 shows a typical selection of the weighting factor β_{j_s} for three well constraints at offsets $J_1 = 10$, $J_2 = 40$ and $J_3 = 90$, respectively (the velocity model is defined in the range 1-100, arbitrary units). In this example, $\beta_0 = 0.2$, $\beta_a = \beta_0/4$ and $\beta_b = \beta_0/2$. At the first stages of the MRF generation (first sweeps), a strong spatial continuity is forced around the well constraints. Far from the well, the misfit term prevails. On the other hand, as the number of sweeps increases (and the overall misfit decreases), the spatial continuity term becomes large for those offsets which are not so close to the constraints, too. At these stages, one can assume that the misfit has already been significantly reduced. This will become more evident in the Examples section.

3 Examples

In all the examples we describe the velocity model using a $M \times N$ rectangular lattice with discrete velocities in the range 1-10 (arbitrary units). Here M is the number of traces (columns) in the offset direction, and N is the number of samples (rows) in the vertical direction. The data is the zero-offset section (primary reflections only) obtained after convolving the derived reflectivity series with a known Ricker wavelet. For simplicity, the reflection coefficients are calculated assuming a constant density. Gaussian noise with $SNR = 10$ is added to the data.

The initial velocity model used to start the iteration is set to uniform random values. The inversion process stops after a maximum of number of sweeps or when the expected misfit is reached. The expected misfit, χ, can be estimated easily by inspecting formula (11) at convergence. We define the expected misfit as the mean rms error of all the traces:

$$\chi^2 = \frac{1}{M} E_s^2 = \frac{1}{M} \sum_{j=1}^{M} \frac{\sum_{i=1}^{N} (y_{ij} - y'_{ij})^2}{\sum_{i=1}^{N} y_{ij}^2} = \frac{1}{M} \sum_{j=1}^{M} \frac{\sum_{i=1}^{N} n_{ij}^2}{\sum_{i=1}^{N} y_{ij}^2} \simeq \frac{1}{SNR}, \tag{15}$$

where n_{ij} is the added noise. For $SNR = 10$, $\chi \simeq 0.316$. In all cases we set $p = 2$ and $\epsilon = 0.5$ for the clique potentials in equation (9).

3.1 Layered model

Figure 4 (first two panels) shows a layered velocity model (Example 1) and the data, with $M = 100$ and $N = 50$. To evaluate the behavior of the algorithm under different number of constraints, the inversion was repeated three times assuming that

a) the velocity was known at three wells located at offsets $J_1 = 10$, $J_2 = 40$ and $J_3 = 90$;

b) the velocity was known at two wells located at offsets $J_1 = 10$ and $J_2 = 90$;

c) the velocity was unknown everywhere.

In all the cases we set $T_0 = 0.1$, $\rho = 0.2$, $\beta_0 = 0.2$, $\beta_a = 0.25$, $\beta_b = 0.75$ and $k_b = 1500$. A maximum of 3000 sweeps were performed, and the third order neighborhood system was used.

The results of the inversion for case (a) are shown in Figure 4 (rows 2 to 6). The figure shows the evolution of the estimated velocity model for various sweeps along with the corresponding calculated data and error images. The resulting misfit in each case are displayed in the figure. The true model was recovered quite accurately after 781 sweeps, when the expected misfit was reached and the iteration stopped. At about 500 sweeps, the recovered model is also very accurate. Note how

the model is progressively recovered (as the number of sweeps increases) from the wells towards regions less constrained. For sweep number 100, for example, the model is recovered only in the neighborhood of the three wells. This behavior is controlled by the selection of the appropriate parameters in equations (12) to (14). In this particular case, β_{j_s} is that shown in Figure 3.

In Example 1b and 1c, when the velocity is known at two locations for the 1b case and nowhere for the 1c case, the results are shown in Figure 5. The final model and data after 3000 sweeps for the two well constraints case are shown in the first row of the figure. Although the expected misfit has not been reached, all details of the true model are recovered quite well. The results are not so accurate in the unconstrained case (second row), where the main features of the model were recovered but the final misfit was too large (0.351). After 3000 sweeps the solution does not improve significantly.

3.2 Layered model with constant velocity block

Here we used exactly the same settings (constraints, parameters, etc.) than in Example 1. The only difference resides in the model, where we have added a constant velocity block in the center of the rectangular lattice, as depicted in Figure 6, row 1. The results for Examples 2a, 2b and 2c, are shown in rows 2, 3 and 4, respectively. Example 2a yields the best results. In this case, the expected misfit was reached after 939 sweeps, and the recovered model is very accurate, except at the vertical boundaries of the block.

Row 3 of the same figure displays the results for the two well constraints case. The expected misfit was not reached after 3000 sweeps. However, the main features of the true model were recovered quite well, including the constant velocity block. Some zones with wrong velocity values can be distinguished in the regions far away from the well constraints. The results are less accurate, as expected, in the unconstrained case (row 4). The final solution (after 3000 sweeps) clearly shows the layering and the constant velocity block, but there are many regions with wrong velocity values.

3.3 Marmousi model

Now we test the algorithm with a more realistic model. For the test we selected the 2–D hard Marmousi model (Versteeg, 1994), which is defined in a 384×122 rectangular lattice and is displayed in Figure 7, first row. The hard Marmousi model exhibits many normal faults, reflectors, steep dips and strong lateral and vertical velocity variations. For simplicity, we eliminated the first two rows of the original model to get rid of the water layer. Also, we scaled and re-discretized the velocity model so that we end up with a 384×120 grid with 10 velocity values in the range 1–10 (arbitrary units). Actually, this is the model which is shown in the first row of Figure 7. The data (zero offset section) was generated by convolving a Ricker wavelet with the reflectivity sequences derived from the velocity column vectors assuming constant density. Gaussian noise with a signal-to-noise ratio of 10 was added to the data. Again, the expected misfit was 0.316.

Because of the different character of the velocity model (thinner layers, etc.), we selected a different set of parameters for the inversion of the data. In particular we set $T_0 = 0.1$, $\rho = 0.1$, $\beta_0 = 0.1$, $\beta_a = 0.25$, $\beta_b = 0.75$ and $k_b = 10000$. A maximum of 10000 sweeps were performed, and the second order neighborhood system was used. As for the constraints, we assumed we have ten wells at offsets $10, 50, 90, \cdots, 370$. The results are shown in Figure 7. The final misfit after 10000 iterations was 0.385. The main features of the model were recovered quite well. Steep dip velocity layers were not recovered as accurately as those which are essentially horizontal, specially at the bottom of the model. Figure 8 shows the results when only five well constraints were used (at locations 10, 100, 190, 280, and 370). The results are not as accurate as in the previous case, but they are satisfactory, considering the complexity of the model. In this case, the final misfit after 10000 sweeps was 0.399.

3.4 Uncertainty

It is possible to check the uncertainty of the estimated models by generating several MRFs using different seeds. This was carried out for the examples 1a and 2a. The inversion was repeated using 20 different seeds with a maximum of 3000 sweeps. Figure 9 shows the convergence curves for the 20 runs in the two models. Note that convergence is achieved in most runs: the average misfits are 0.3147 ± 0.0002 and 0.3149 ± 0.0006 for the examples 1a and 2a respectively (the expected misfit is 0.3162 in both cases). In general, the convergence in the layered model is faster than in the model with the constant velocity block.

Figure 10 displays the average velocity models (and data) obtained after the 20 independent runs, along with their standard deviations. As expected, the uncertainty is smaller around the constraints. Larger uncertainties take place at isolated locations, specially in the second model. In general, both models are recovered consistently, except for some minor artifacts at the vertical boundaries of the block in the second model.

4 Summary and conclusions

The inversion of reflection data can be carried out by generating a MRF that favors certain spatial continuity given by a set of potentials defined over a neighborhood system of cliques. This MRF, which is generated using Gibbs' sampling, is guided to honor the data and satisfies a set a well constraints at every stage of the iterative process. Well constraints are important to improve the convergence of the algorithm in a reasonable number of sweeps (iterations). The defined potentials are specially suited for obtaining high resolution velocity images dominated by horizontal layers or other constant velocity blocks. Moderate dipping layers are tolerated. The selection of an appropriate scheme to progressively include the well constraints into the inversion is very important to avoid the premature formation of constant velocity regions that may delay the convergence. This strategy is used to control how the spatial continuity and the misfit terms are weighted.

The procedure allows one to estimate the uncertainty of the inversion by generating several MRFs using different seeds. The results in this regards show that the obtained velocity models exhibit low variability except for some small regions which were poorly constrained.

In this work we selected first, second and third order neighborhood systems with associated potentials that favor the generation of constant velocity regions which are essentially flat or moderately deep (when $\rho \ll 1$). Though the neighborhood system and the associated potentials can be adjusted for models with different spatial continuity behavior (for example selecting $\rho = 1$ would favor both horizontal and vertical continuity), it is clear that alternative potentials could be easily incorporated into the inversion procedure.

For simplicity we defined the models using 10 discrete values within a given range of arbitrary units. A larger number of velocity values can be used at the expense of an increase in the computational cost. The computational cost is completely dominated by the calculation of the reflection data and misfit. Though the replacement of a single velocity value at a given offset affects only locally the trace at that offset and efficient formula have been derived to update the misfit accordingly, about 80–90% of the computational cost of the whole inversion process is accounted by the evaluation of the misfit term.

5 References

Besag, J., 1974, Spatial interaction and the statistical analysis of lattice systems: Journal of the Royal Statistical Society B, **36**, no. 2, 192–236.

Carrion, P. M., and Pulleda, S., 1995, Bayesian inference and Gibbs' sampler in non-linear inversion: Journal of Seismic Exploration, **4**, 73–79.

Carrion, P. M., Pulleda, S., and Comelli, P., 1997, Gibbs' statistics in crosswell and reflection tomography: Inversion of noisy data: Geophysics, **62**, no. 4, 1208–1213.

Casella, G., and George, E. I., 1992, Explaining the Gibbs' sampler: The American Statistician, **46**, no. 3, 167–174.

Geman, S., and Geman, D., 1984, Stochastic relaxation, Gibbs' distributions, and the Bayesian restoration of images: IEEE Trans. Pattern Analysis and Machine Intelligence, **PAMI-6**, 721–741.

Sen, M. K., and Stoffa, P. L., 1996, Bayesian inference, Gibbs' sampler and uncertainty estimation in geophysical inversion: Geophysical Prospecting, **44**, 313–350.

Tarantola, A., 1997, Inverse problem theory, methods and data fitting and model parameter estimation: Elsevier Science Publishing Co.

Versteeg, R., 1994, The Marmousi experience: velocity model determination on a synthetic complex data set: The Leading Edge, **13**, 927–936.

A Appendix

Let

$$y_i = \sum_k h_k r_{i-k+1} \tag{16}$$

be the seismic trace coefficients at offset j_s, where h_i and r_i are the corresponding seismic wavelet and reflectivity coefficients. Assuming a constant density, the reflection coefficient is

$$r_i = (x_i - x_{i-1})/(x_i + x_{i-1}), \quad i = 2, 3, \cdots, \tag{17}$$

with $r_1 = 0$, where x_i is the velocity at depth/time i. At any given stage of the inversion procedure, let $x'_{i_s} = x_{i_s} + \delta x_{i_s}$ be the new velocity value at offset j_s and depth/time $i = i_s$. Then

$$r'_i = \begin{cases} (x_{i_s} - x_{i_s-1} + \delta x_{i_s})/(x_{i_s} + x_{i_s-1} + \delta x_{i_s}), & i = i_s \\ (x_{i_s+1} - x_{i_s} - \delta x_{i_s})/(x_{i_s+1} + x_{i_s} + \delta x_{i_s}), & i = i_s + 1 \\ r_i. & \text{otherwise} \end{cases} \tag{18}$$

is the new reflectivity sequence. The new trace is

$$\begin{aligned} y'_i &= \sum_k h_k r'_{i-k+1} \\ &= h_{i-i_s+1} r'_{i_s} + h_{i-i_s} r'_{i_s+1} + \sum_{\substack{k,\, k \neq i-i_s \\ k \neq i-i_s+1}} h_k r_{i-k+1} \\ &= h_{i-i_s+1} r'_{i_s} + h_{i-i_s} r'_{i_s+1} + \sum_k h_k r_{i-k+1} - h_{i-i_s+1} r_{i_s} - h_{i-i_S} r_{i_s+1} \\ &= y_i + h_{i-i_s}(r'_{i_s+1} - r_{i_s+1}) + h_{i-i_s+1}(r'_{i_s} - r_{i_s}) \\ &= y_i + h_{i-i_s}\gamma_1 + h_{i-i_s+1}\gamma_2 \end{aligned} \tag{19}$$

where

$$\gamma_1 = \frac{-2\delta x_{i_s} x_{i_s+1}}{(x_{i_s} + x_{i_s+1})(x_{i_s} + x_{i_s+1} + \delta x_{i_s})} \tag{20}$$

$$\gamma_2 = \frac{2\delta x_{i_s} x_{i_s-1}}{(x_{i_s} + x_{i_s-1})(x_{i_s} + x_{i_s-1} + \delta x_{i_s})} \tag{21}$$

Note that the trace needs to be updated using equation (19) only for $i_s \leq i \leq L_h + i_s$, where L_h is the wavelet length. This procedure reduces significantly the computational cost of evaluating the misfit function.

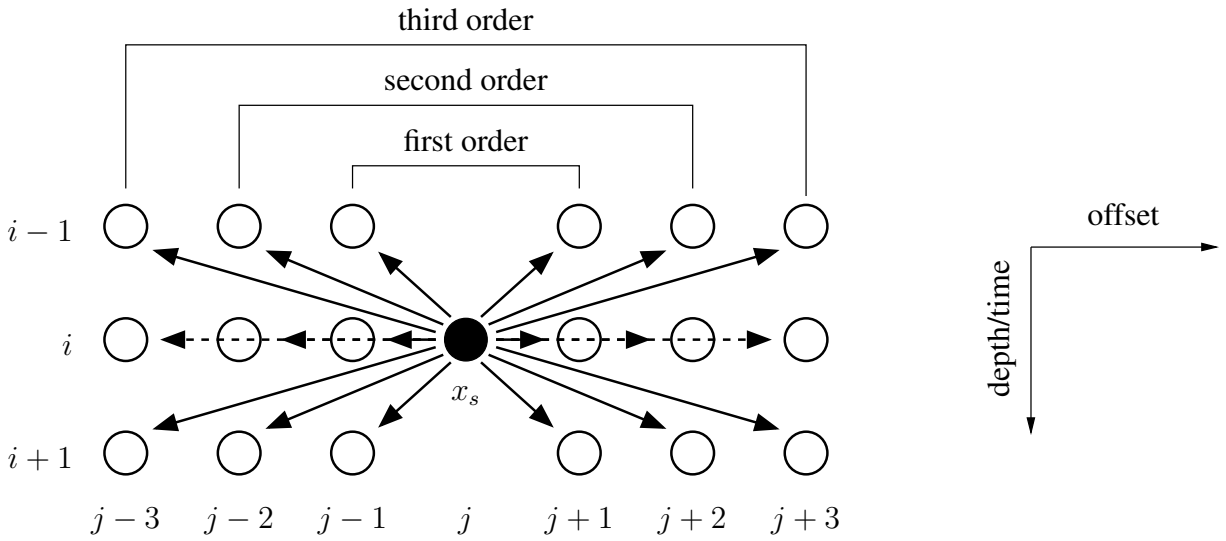

Figure 1: First, second and third order neighborhood system. Each circle represents a cell. Each arrow a clique.

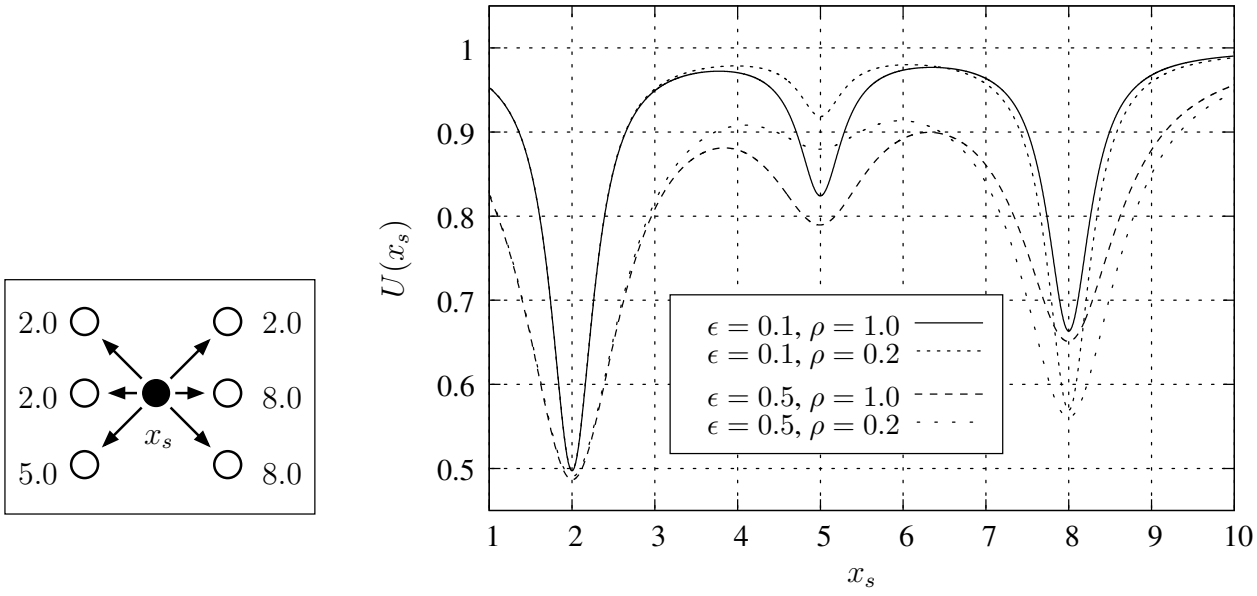

Figure 2: Energy function $U(x_s)$ for the given first order clique (left panel). Note how low energy peaks are assigned to those values of x_s that yield greater spatial continuity (right panel).

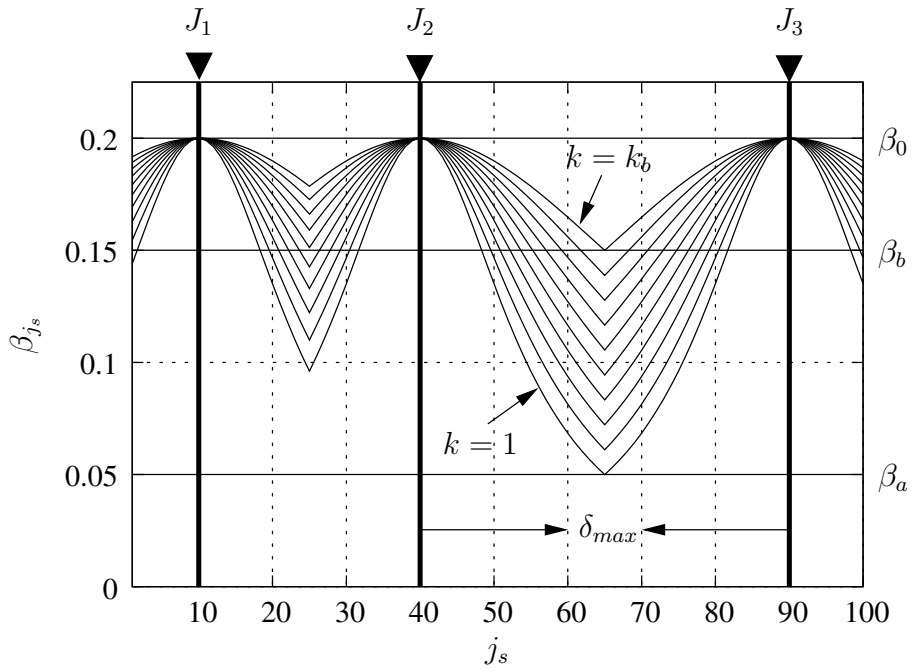

Figure 3: Weighting factor β_{j_s} for sweep number varying from 1 to k_b. The maximum weight, β_0, is assigned to the offsets at the well constraint locations (marked with small arrows). In between the constraints, the weight is smaller to avoid premature development of constant velocity regions at the early stages of the MRF generation process.

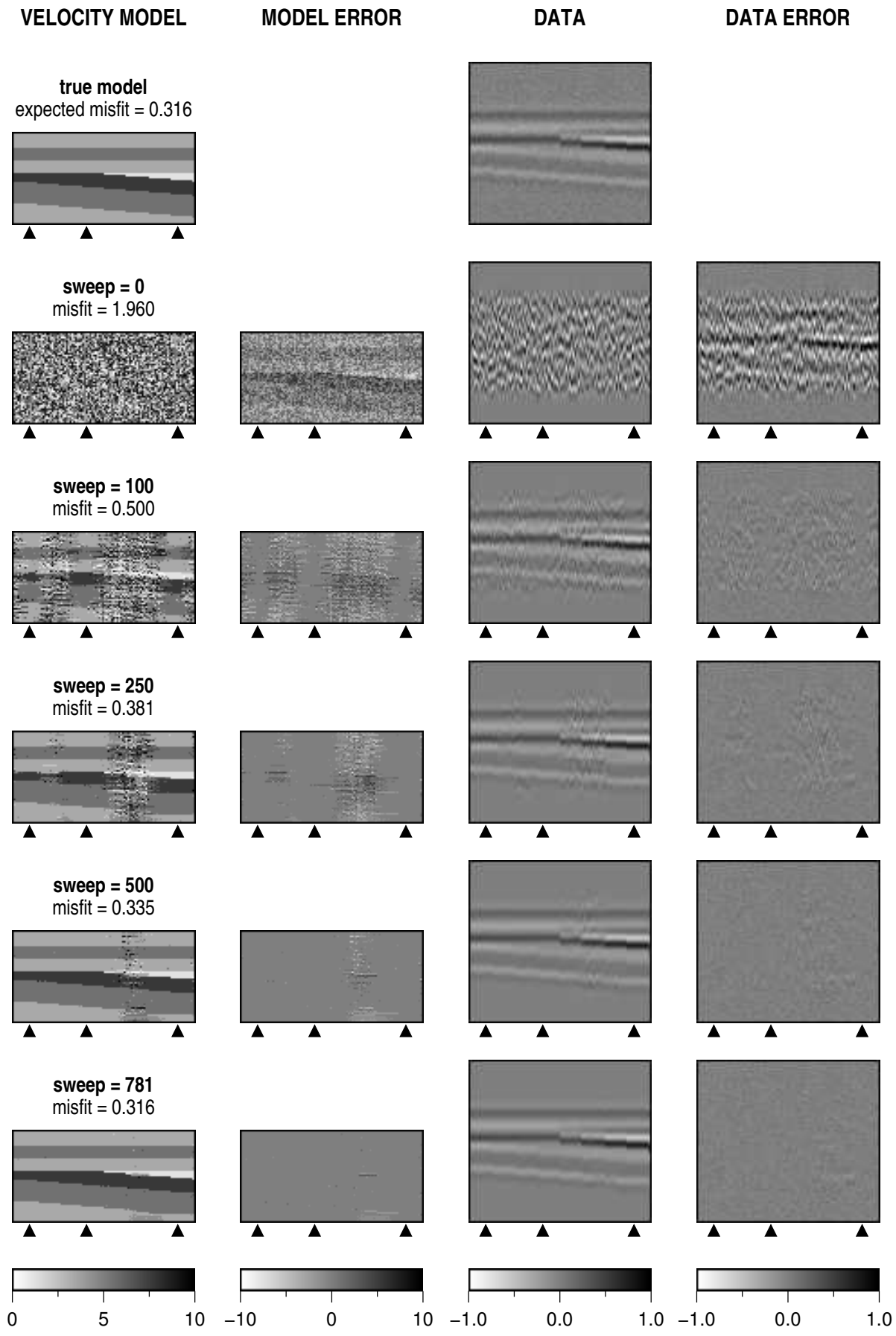

Figure 4: Example 1a: layered model. First row: true model and data. Rows 2 to 6: estimated model, data and error images for various sweeps. The small arrows show the location of the constraints.

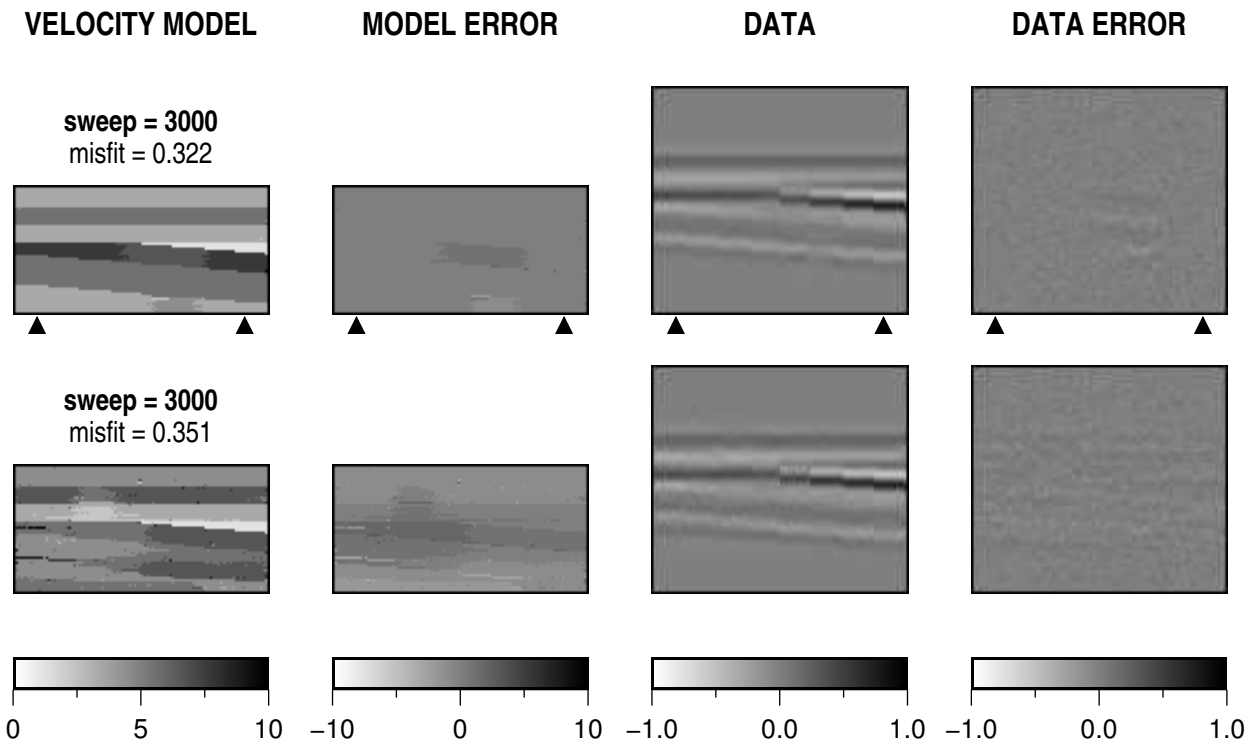

Figure 5: Examples 1b and 1c: layered model. Estimated model, data and error images after 3000 sweeps in the layered model with two well constraints (first row) and no well constraints (second row). The true model and data are shown in Figure 4, first row.

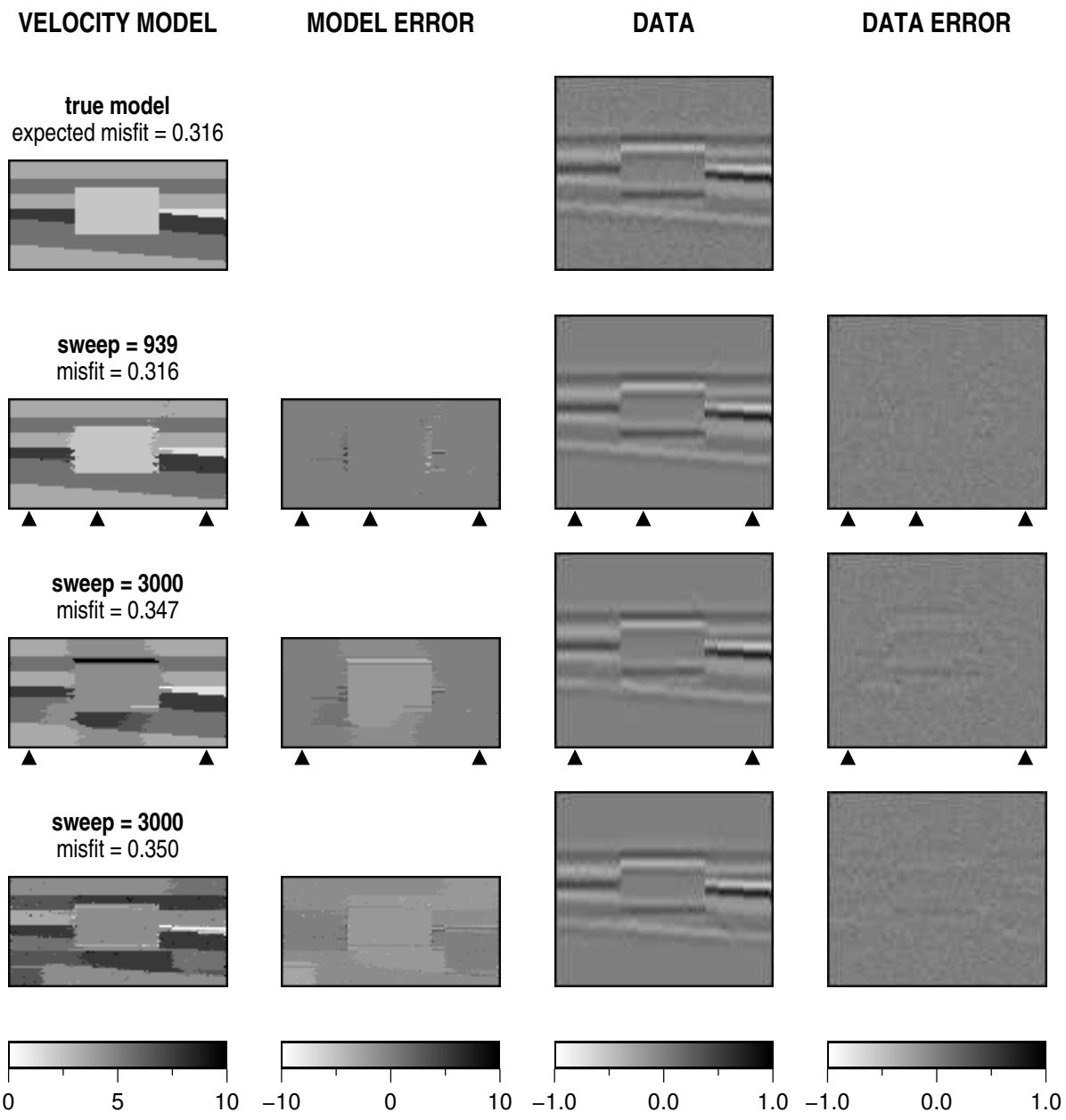

Figure 6: Example 2: layered model with constant velocity block. First row: true model and data. Rows 2, 3 and 4: estimated model, data and error images for the three, two and none well constraints cases, respectively.

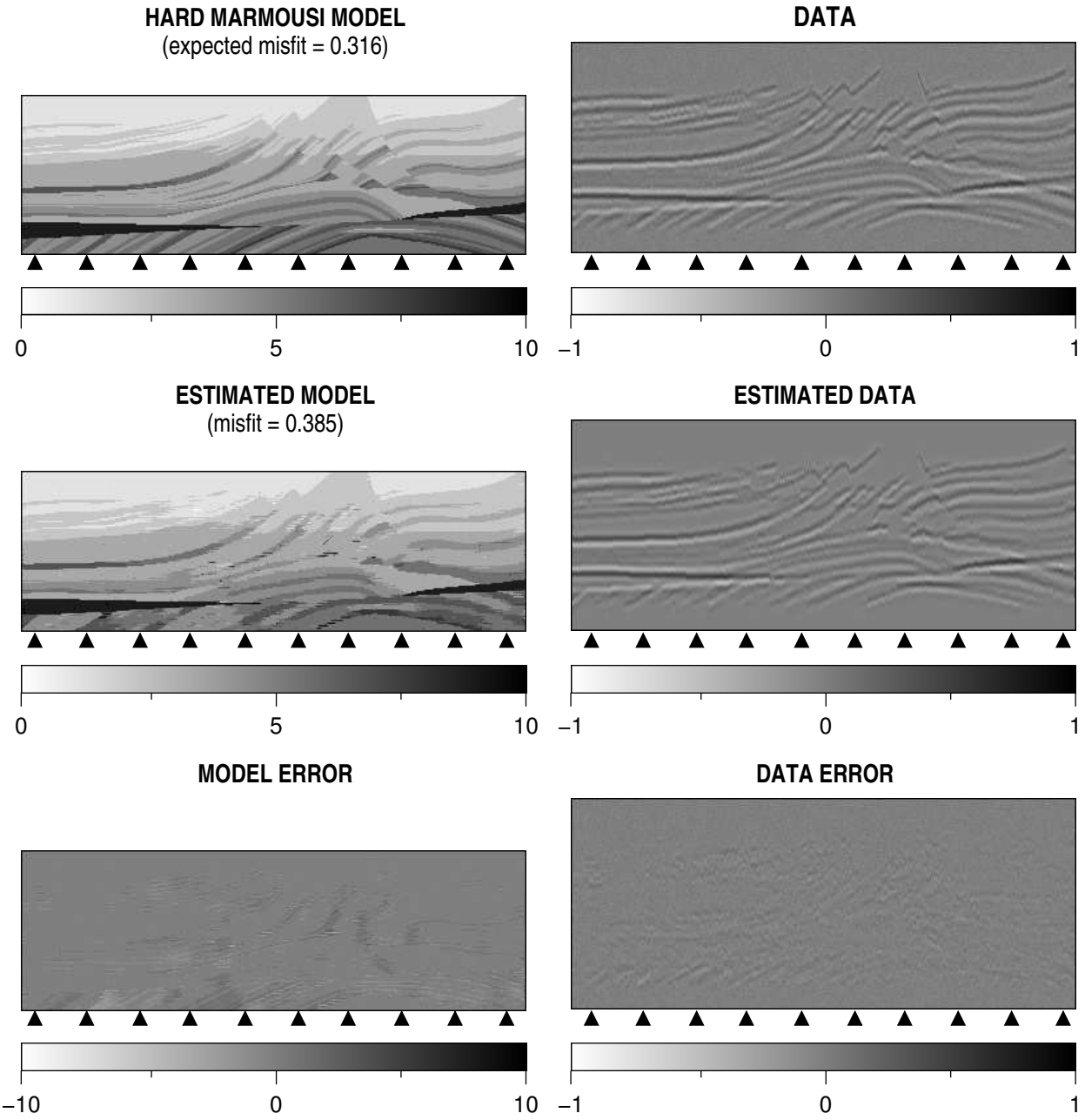

Figure 7: Example 3: hard Marmousi model. First row: true model and data. Second row: estimated model and data after 10000 sweeps. Third row: model and data error images after 10000 sweeps. The small arrows show the location of the constraints.

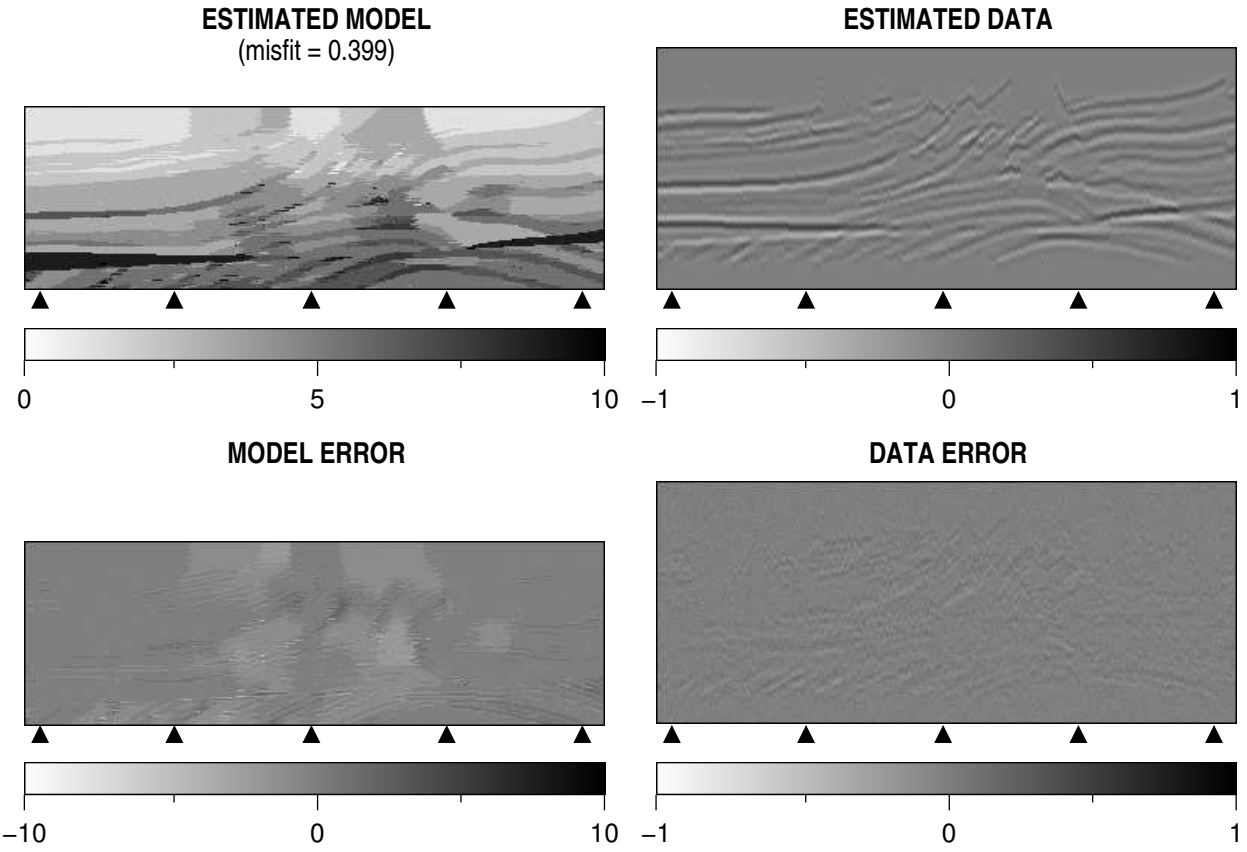

Figure 8: Example 3: hard Marmousi model. First row: estimated model and data after 10000 sweeps. Third row: model and data error images after 10000 sweeps. The small arrows show the location of the constraints.

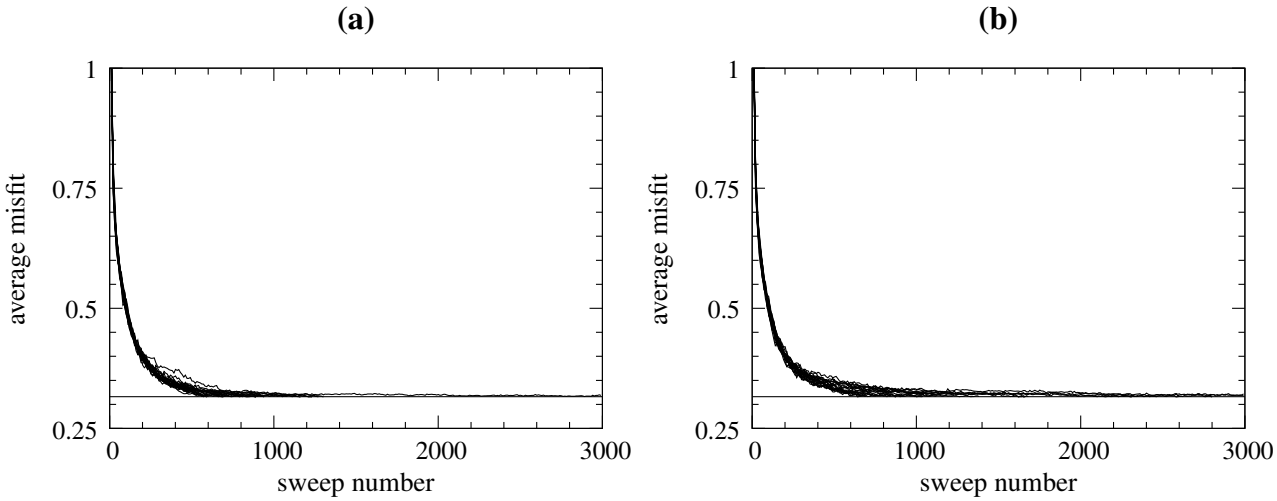

Figure 9: Average data misfit versus sweep number for 20 independent runs: (a) layered model; (b) layered model with a constant velocity block. The horizontal line shows the expected misfit, which is equal to 0.316 in both cases. Three well constraints have been used.

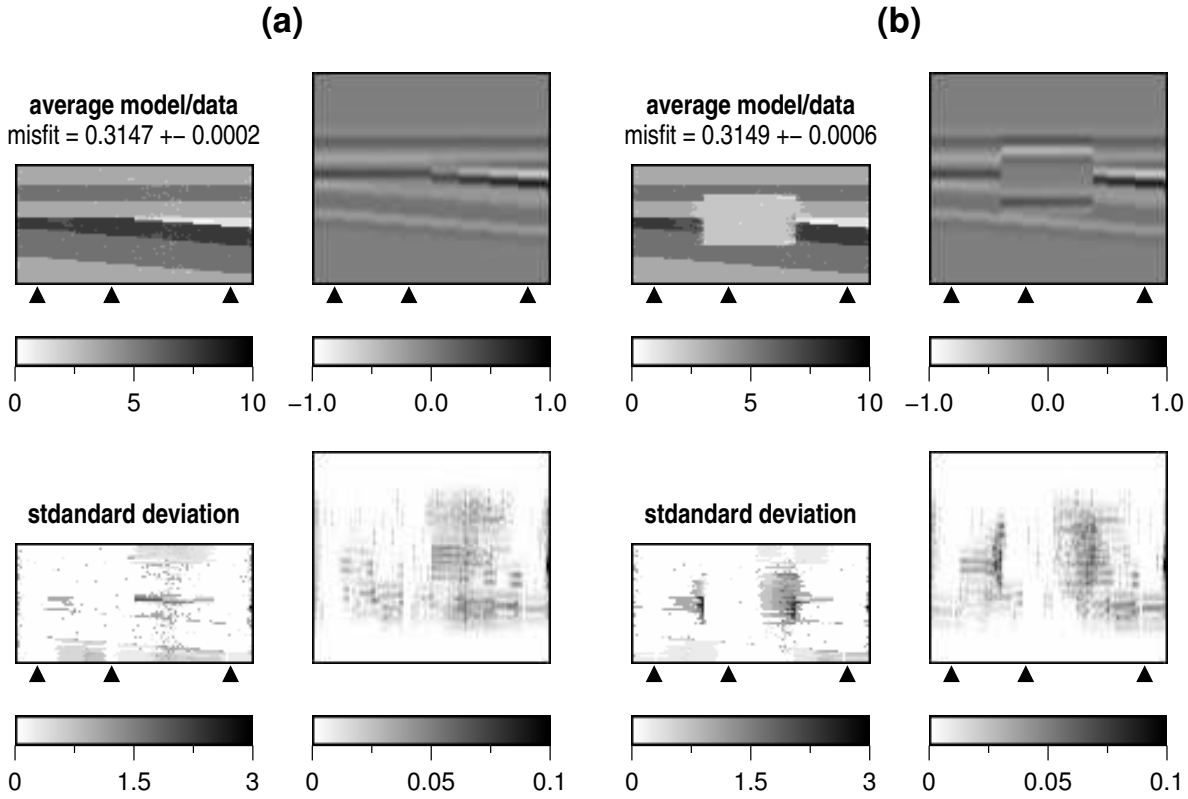

Figure 10: True (first row), average (second row) and standard deviation (third row) of the velocity model and data obtained after 20 independent runs. (a) Layered model; (b) layered model with a constant velocity block. The small arrows show the location of the constraints.