Ambiguous Models for Subsurface Prediction

Petter Abrahamsen *

Abstract

A frequently encountered problem when predicting the depth to subsurfaces is that alternative stochastic models for the depth to the subsurfaces exist. An approach to resolve this ambiguity, is to combine the predictors associated with each stochastic model. For a layered structure, the number of alternative stochastic models and associated predictors could be prohibitively large, so an alternative method is proposed that consists of constructing a single stochastic model from the alternative stochastic models. The result provides one consistent predictor for each subsurface which performs similarly to the approach combining several predictors while drastically reducing computational costs. The proposed method applies to layered geological structures using a combination of universal or Bayesian kriging and co-kriging.

KEY WORDS: ambiguous models, Bayesian kriging, combining models, combining predictors, depth conversion, Gaussian random fields, kriging.

INTRODUCTION

Consider the general problem of mapping the depth to subsurfaces separating different zones within a petroleum reservoir. The top and base of the reservoir are often visible on seismic lines so geophysicists can provide detailed depth maps from interpreted travel-time maps using seismic depth conversion. The internal subsurfaces separating different zones will only occasionally exhibit reliable seismic reflections, so geologists try to map the thickness of each reservoir zone based on bore-hole observations and an understanding of the depositional processes. These thickness maps are less detailed than the depth maps based on seismic data. The total mapped thickness of the internal reservoir zones will generally not add up to the thickness depicted in the detailed depth maps for the top and base of the reservoir. This ambiguity must be resolved to provide the final depth maps describing the depth to the internal subsurfaces. The specific problem considered here is to merge the detailed depth maps provided by the geophysicists with the cruder zone thickness maps provided by the geologists.

Two methods for resolving the ambiguity will be discussed. The first method is an adaption of an approach used in econometrics and forecasting (Bunn 1989, Granger 1989),

*Norwegian Computing Center, P.O.Box 114, N-0314 Oslo, Norway

Submitted to Mathematical Geology. Tables and illustrations at end of the paper.

and consists of predicting the depth to the subsurfaces based on different combinations of zone thickness maps and depth maps based on seismic data. The set of depth predictions are finally combined 'in an optimal manner' to come up with a single set of predicted depth maps for each subsurface. This approach works fine, but it is computationally expensive. An alternative new method is therefore proposed. Instead of combining the depth predictors for each subsurface, different stochastic models for each subsurface are combined. Each model correspond to a different combination of zone thickness maps and depth maps based on seismic data. The result is one stochastic model and one associated predictor for each subsurface. The predicted depth maps are very similar to the approach combining predictors, but the computer expenses are dramatically reduced.

The next two sections outline the problem in more detail and sketches the ideas underlying the two approaches. The subsequent three sections provide details on the stochastic models for depth and thicknesses and the technicalities of the two methods. Finally the properties of the proposed methods are explored by means of an example.

POSITION OF THE PROBLEM

The large scale geometry of petroleum reservoirs is defined by subsurfaces separating mainly homogeneous geological layers. A stochastic model for the depth to the subsurfaces is constructed by specifying stochastic models for the thickness of intervals between the subsurfaces. The stochastic model for the thickness of interval i includes a trend, $m_i(\mathbf{x})$, and a fully specified zero expectation Gaussian random field, $\epsilon_i(\mathbf{x})$, for the residual:

$$\Delta Z_i(\mathbf{x}) = m_i(\mathbf{x}) + \epsilon_i(\mathbf{x}); \qquad \mathbf{x} \in \mathbf{R}^2. \tag{1}$$

The stochastic model for the depth to subsurface l is obtained as $Z^l(\mathbf{x}) = \sum_{i=1}^{l} \Delta Z_i(\mathbf{x})$.

Fig. 1 illustrates a schematic cross-section of a petroleum reservoir where the subsurfaces 'Top reservoir' and 'Base reservoir' are assumed to be seismic reflectors. For an interval bounded by two seismic reflectors, the trend, $m_i(\mathbf{x})$, is constructed from seismic travel-times and seismic velocities. The trend will, in general, be detailed and accurate, so the variance of the residual error, $\epsilon_i(\mathbf{x})$, should be small. Travel-time-based models for the depth to 'Top reservoir' and 'Base reservoir' would be $Z^{\mathrm{TR}}(\mathbf{x}) = \Delta Z_{\mathrm{TR}}(\mathbf{x})$ and $Z^{\mathrm{BR}}(\mathbf{x}) = \Delta Z_{\mathrm{TR}}(\mathbf{x}) + \Delta Z_{\mathrm{R}}(\mathbf{x})$ respectively. (See Fig. 1 for notation.) At the same time, zonation within the reservoir formation is of great interest, and stochastic models of the form Eq. (1) can be provided for each reservoir zone. The trends for the zone thicknesses are usually contoured maps based on geological interpretation of bore-hole observations. The accuracy of these trends is generally low compared to trends based on seismic data, so the variance of $\epsilon_i(\mathbf{x})$ is large.

As an alternative to the travel-time based method for the depth to 'Base reservoir', the thickness of all the internal zones could be added to 'Top reservoir': $Z^{\mathrm{BR}}(\mathbf{x}) = \Delta Z_{\mathrm{TR}}(\mathbf{x}) + \Delta Z_{\mathrm{Z3}}(\mathbf{x}) + \Delta Z_{\mathrm{Z2}}(\mathbf{x}) + \Delta Z_{\mathrm{Z1}}(\mathbf{x})$. Now an ambiguity has occurred: Two stochastic models for the depth to 'Base reservoir' have been proposed. The travel-time based method will usually give more detailed and accurate depth maps so this is the preferred method in practical applications.

Consider the two intermediate zone boundaries, 'Top zone 1' and 'Top zone 2'. The depth to each of these subsurfaces can be obtained in two different ways:

$$Z^{\mathrm{T2}}(\mathbf{x}) = \Delta Z_{\mathrm{TR}}(\mathbf{x}) + \begin{cases} \Delta Z_{\mathrm{R}}(\mathbf{x}) - \Delta Z_{\mathrm{Z1}}(\mathbf{x}) - \Delta Z_{\mathrm{Z2}}(\mathbf{x}) & \text{Method (a) or} \\ \Delta Z_{\mathrm{Z3}}(\mathbf{x}) & \text{Method (b–d)} \end{cases}$$

$$Z^{\mathrm{T1}}(\mathbf{x}) = \Delta Z_{\mathrm{TR}}(\mathbf{x}) + \begin{cases} \Delta Z_{\mathrm{R}}(\mathbf{x}) - \Delta Z_{\mathrm{Z1}}(\mathbf{x}) & \text{Method (a,b) or} \\ \Delta Z_{\mathrm{Z2}}(\mathbf{x}) + \Delta Z_{\mathrm{Z3}}(\mathbf{x}) & \text{Method (c,d).} \end{cases}$$

The methods correspond to labels in Fig. 2 and Fig. 3. Once again ambiguities are encountered, but now it is far from obvious which alternative to choose. Since the seismic reflectors, 'Top reservoir' and 'Base reservoir', are assumed accurately determined, the figures suggest that Method (b–d) should be used for $Z^{\mathrm{T2}}(\mathbf{x})$ and Method (a,b) should be used for $Z^{\mathrm{T1}}(\mathbf{x})$. This choice corresponds to graph (b) in Fig. 2 and leaves a gap between the two subsurfaces so the trend for the thickness of 'Zone 2' is never considered. Also, this choice has a serious implication on the uncertainty of the thickness of 'Zone 2': Assuming the residuals to be independent implies that the variance of 'Zone 2' is

$$\mathrm{Var}\Big\{Z^{\mathrm{T2}}(\mathbf{x}) - Z^{\mathrm{T1}}(\mathbf{x})\Big\} = \mathrm{Var}\{\Delta Z_{\mathrm{R}}(\mathbf{x})\} + \mathrm{Var}\{\Delta Z_{\mathrm{Z3}}(\mathbf{x})\} + \mathrm{Var}\{\Delta Z_{\mathrm{Z1}}(\mathbf{x})\}\,.$$

This variance is usually significantly greater than $\mathrm{Var}\{\Delta Z_{\mathrm{Z2}}(\mathbf{x})\}$ and the potentially high correlation between the depth to subsurface 'Top zone 1' and 'Top zone 2' is lost.

The discussion has motivated the need for an approach where several methods can be used simultaneously, so the unpleasant need for choosing one particular method is obsolete. Two alternative solutions are suggested.

APPROACHES TO RESOLVING THE AMBIGUITIES

Combining Predictors

This approach is an adaption of a method used in time series analysis and forecasting and is reviewed by Bunn (1989) and Granger (1989). The idea is to make a linear combination of predictors which are based on different assumptions and/or methods. The linear combination is chosen such that the squared prediction error is minimized subject to a zero bias restriction. In the present context consider the four possible predictors for the depth to 'Top zone 1': $Z^{\mathrm{T1}\star}_{(\mathrm{a})}(\mathbf{x})$, $Z^{\mathrm{T1}\star}_{(\mathrm{b})}(\mathbf{x})$, $Z^{\mathrm{T1}\star}_{(\mathrm{c})}(\mathbf{x})$, and $Z^{\mathrm{T1}\star}_{(\mathrm{d})}(\mathbf{x})$. The combined predictor is a linear combination of these:

$$Z^{\mathrm{T1}\star}(\mathbf{x}) = w^{\mathrm{T1}}_{(\mathrm{a})}(\mathbf{x}) Z^{\mathrm{T1}\star}_{(\mathrm{a})}(\mathbf{x}) + w^{\mathrm{T1}}_{(\mathrm{b})}(\mathbf{x}) Z^{\mathrm{T1}\star}_{(\mathrm{b})}(\mathbf{x}) + w^{\mathrm{T1}}_{(\mathrm{c})}(\mathbf{x}) Z^{\mathrm{T1}\star}_{(\mathrm{c})}(\mathbf{x}) + w^{\mathrm{T1}}_{(\mathrm{d})}(\mathbf{x}) Z^{\mathrm{T1}\star}_{(\mathrm{d})}(\mathbf{x}). \quad (2)$$

The space dependent weights are chosen to minimize the combined prediction error. This requires the kriging prediction variances and covariances for all four predictors, so the drawback of this method is the necessity to evaluate several predictors and prediction variances and covariances.

Combining Stochastic Models

This new method amounts to using a linear combination of the stochastic models for the depth to each subsurface deduced from the alternatives in Fig. 3. For the depth to 'Top zone 1', the combined stochastic model would be

$$\begin{aligned} Z^{\mathrm{T1}}(\mathbf{x}) = & \, w^{\mathrm{T1}}_{(\mathrm{a,b})}(\mathbf{x}) \left[\Delta Z_{\mathrm{TR}}(\mathbf{x}) + \Delta Z_{\mathrm{R}}(\mathbf{x}) - \Delta Z_{\mathrm{Z1}}(\mathbf{x})\right] \\ & + w^{\mathrm{T1}}_{(\mathrm{c,d})}(\mathbf{x}) \left[\Delta Z_{\mathrm{TR}}(\mathbf{x}) + \Delta Z_{\mathrm{Z2}}(\mathbf{x}) + \Delta Z_{\mathrm{Z3}}(\mathbf{x})\right]. \end{aligned}$$

The spatially dependent weights $w^{\mathrm{T1}}_{(\mathrm{a,b})}(\mathbf{x})$ and $w^{\mathrm{T1}}_{(\mathrm{c,d})}(\mathbf{x})$ are chosen to minimize the residual variance, $\mathrm{Var}\left\{Z^{\mathrm{T1}}(\mathbf{x})\right\}$, subject to the condition that the weights add to one. This leaves one stochastic model and a single associated predictor for the depth to each subsurface.

Combining predictors is based on the principle of minimizing the prediction error. Combining stochastic models however, is a heuristic approach which must be justified by comparing the results to the results obtained when combining predictors. An example will illustrate that the two approaches give almost identical results.

Further Alternatives

It is possible to consider Eq. (2) as a linear regression model for $Z^{l*}(\mathbf{x})$ with $Z^{l*}_a(\mathbf{x})$ as regressors and the weights as unknown parameters. An additional constant term accounting for possible bias can be added (Granger 1989). This approach, recently called "stacked regression" by Wolpert (1992) and Breiman (1992), either requires historical data or a large dataset allowing cross validation. LeBlanc and Tibshirani (1993) compares cross validation to bootstrapping, best subset regression, and regression trees.

PREDICTION

The best linear unbiased predictor for a random field with an unknown linear trend is the universal kriging predictor (Matheron 1963). Moreover, for a Gaussian random field, a linear predictor is the optimal predictor in the least squares sense. Introductions to universal kriging is found in Journel and Huijbregts (1978), Ripley (1981), and Cressie (1991). Therefore, assuming Eq. (1) is an appropriate stochastic model for the zone thicknesses leads to universal kriging as the optimal predictor for the zone thicknesses and the depth to subsurfaces (Olea 1974, Delfiner and Delhomme 1975). Universal kriging allows travel-time data to be included in the trend, and bore-hole observations are used to estimate trend parameters and tying the subsurfaces to observed depths. Since all subsurfaces are below the upper interval, all subsurface depths are correlated. Therefore, the universal kriging predictor for any particular subsurface depth should be conditioned on available bore-hole observations of all subsurfaces (Abrahamsen 1993). This requires that all covariances between depth observations from different subsurfaces must be known. These covariances depend on how zone thicknesses are added to obtain the depths to the entire set of subsurfaces. That is, choosing one method in Fig. 2, it is possible to deduce the covariances between subsurfaces based on the specified stochastic models for interval thicknesses.

The kriging predictors are linear predictors: $Z^*(\mathbf{x}) = \boldsymbol{\alpha}'(\mathbf{x})\,\mathbf{Z}$, where $\mathbf{Z}$ is a vector of depth observations from all subsurfaces. The kriging weights, $\boldsymbol{\alpha}(\mathbf{x})$, depend on the geometry

of the observations, the stochastic models for the zone thicknesses, and the method chosen for adding the zone thicknesses. Alternative methods therefore give different kriging weights for the same set of observations.

STOCHASTIC MODELS FOR SUBSURFACES

This section establish some basic notation for non-ambiguous models. The next section generalize the notation to include ambiguous models.

The stochastic model for the thickness of interval i includes a trend, $m_i(\mathbf{x})$, linear in unknown parameters, and a spatially correlated residual:

$$\Delta Z_i(\mathbf{x}) = \mathbf{g}_i'(\mathbf{x})\,\boldsymbol{\beta}_i + \epsilon_i(\mathbf{x}); \qquad \mathbf{x} \in \mathbf{R}^2, \tag{3}$$

where $\mathbf{g}_i(\mathbf{x})$ is a P_i-dimensional vector of known spatially dependent functions, $\boldsymbol{\beta}_i$ is a vector of P_i unknown coefficient parameters, and the residual, $\epsilon_i(\mathbf{x})$, is a zero mean Gaussian random field specified by the spatially varying standard error, $\sigma_i(\mathbf{x})$, and the correlation function $\rho_i(\mathbf{x}, \mathbf{y})$. Note that Eq. (3) is a linear regression model with a correlated error term, $\epsilon_i(\mathbf{x})$.

A typical isochore based model for interval i is

$$\Delta Z_i(\mathbf{x}) = g_{i1}(\mathbf{x})\,\beta_{i1} + \epsilon_i(\mathbf{x}), \tag{4a}$$

where $g_{i1}(\mathbf{x})$ is an interpreted isochore map for the interval supplied by geologists. The correlation function and standard error for the residual can occasionally be estimated from bore-hole data. A typical travel-time based model is

$$\Delta Z_i(\mathbf{x}) = [\beta_{i1} + \beta_{i2}\bar{t}_i(\mathbf{x})\,]\,\Delta t_i(\mathbf{x}) + \epsilon_i(\mathbf{x}), \tag{4b}$$

where $\bar{t}_i(\mathbf{x})$ is the interpreted seismic travel-time to the midpoint of interval i and $\Delta t_i(\mathbf{x})$ is the interpreted seismic travel-time for interval i. So $\mathbf{g}_i'(\mathbf{x}) = [\Delta t_i(\mathbf{x}),\, \bar{t}_i(\mathbf{x})\Delta t_i(\mathbf{x})]$. The expression in the square bracket in Eq. (4b) is the interval velocity trend for interval i. A positive value for β_{i2} gives the widely encountered velocity increase at larger depths due to compaction (Faust 1951, Acheson 1963). The residual must account for the uncertainty in the travel-times (Walden and White 1984, White 1984) and the uncertainty in the interval velocity field (Al-Chalabi 1974, Al-Chalabi 1979, Abrahamsen 1993).

Consider now a geological model including L intervals and subsurfaces. The depth to the lth subsurface is

$$Z^l(\mathbf{x}) = \sum_{i=1}^{l} \Delta Z_i(\mathbf{x}) = \mathbf{f}^{l\prime}(\mathbf{x})\,\boldsymbol{\beta} + \mathcal{E}^l(\mathbf{x}), \tag{5}$$

where

$$\mathbf{f}^{l\prime}(\mathbf{x}) = \begin{bmatrix} \mathbf{g}_1'(\mathbf{x}) & \cdots & \mathbf{g}_l'(\mathbf{x}) & \mathbf{0}' & \cdots & \mathbf{0}' \end{bmatrix} \qquad \text{and} \qquad \boldsymbol{\beta}' = \begin{bmatrix} \boldsymbol{\beta}_1' & \cdots & \boldsymbol{\beta}_L' \end{bmatrix}. \tag{6}$$

Here, $\mathbf{0}$ are zero vectors replacing $\mathbf{g}_m(\mathbf{x})$ for $m > l$, and $\mathcal{E}^l(\mathbf{x}) = \sum_{i=1}^{l} \epsilon_i(\mathbf{x})$

The covariance between two subsurfaces at arbitrary locations is

$$\mathrm{Cov}\Big\{Z^l(\mathbf{x}), Z^m(\mathbf{y})\Big\} = \mathrm{Cov}\Big\{\mathcal{E}^l(\mathbf{x}), \mathcal{E}^m(\mathbf{y})\Big\} = \sum_{i=1}^{l}\sum_{j=1}^{m} \sigma_i(\mathbf{x})\,\sigma_j(\mathbf{y})\,\rho_{ij}(\mathbf{x},\mathbf{y}), \qquad (7)$$

where $\rho_{ij}(\mathbf{x},\mathbf{y}) = \mathrm{Corr}\{\epsilon_i(\mathbf{x}), \epsilon_j(\mathbf{y})\}$. Kriging predictors for correlated subsurfaces are found in Abrahamsen (1993) and the most important equations are given in the Appendix.

AMBIGUOUS MODELS

The consecutive enumeration of the intervals used in the previous section will not do for ambiguous models. The summation of interval thicknesses must be replaced by a sum including the intervals needed to obtain the depth to a particular subsurface for a chosen method. To simplify later notation some sets are introduced:

$\mathcal{M}$	the set of methods for all subsurfaces,
$\mathcal{M}^l$	the set of possible methods for subsurface l.

To illustrate these sets consider Fig. 1 with the four alternative methods illustrated in Fig. 2. Here $\mathcal{M}$ contains four methods, i.e., $\mathcal{M} = \{$(a), (b), (c), (d)$\}$. The possible methods for a particular subsurface is deduced from Fig. 2 are illustrated in Fig. 3. The set $\mathcal{M}^l$ for subsurfaces 'Top Reservoir' (TR) and 'Top zone 1' (T1) are $\mathcal{M}^{\mathrm{TR}} = \{$(a–d)$\}$ and $\mathcal{M}^{\mathrm{T1}} = \{$(a,b), (c,d)$\}$ respectively.

Three sets of intervals are also introduced:

$\mathcal{I}$	the set of all intervals,
$\mathcal{I}^l$	the set of intervals contributing to subsurface l,
$\mathcal{I}^l_a$	the set of intervals contributing to subsurface l using method $a \in \mathcal{M}$.

Note that $\mathcal{I}^l_a \subseteq \mathcal{I}^l \subseteq \mathcal{I}$. For the example in Fig. 1, $\mathcal{I}$ contains all five intervals, i.e. $\mathcal{I} = \{$TR, R, Z3, Z2, Z1$\}$. From Fig. 2 it is seen that $\mathcal{I}^{\mathrm{TR}} = \{TR\}$ and $\mathcal{I}^{\mathrm{T1}} = \{$TR, R, Z3, Z2, Z1$\}$. Finally, here are some examples of some $\mathcal{I}^l_a$ sets: $\mathcal{I}^{\mathrm{TR}}_{(\mathrm{a})} = \{TR\}$, $\mathcal{I}^{\mathrm{T1}}_{(\mathrm{a})} = \{$TR, R, Z1$\}$, $\mathcal{I}^{\mathrm{TR}}_{(\mathrm{c})} = \{TR\}$, and $\mathcal{I}^{\mathrm{T1}}_{(\mathrm{c})} = \{$TR, Z3, Z2$\}$.

Using this notation, Eq. (5) is generalized to

$$Z^l_a(\mathbf{x}) = \sum_{i\in\mathcal{I}^l_a} s_{ia}\,\Delta Z^i(\mathbf{x}) = \mathbf{f}^{l\prime}_a(\mathbf{x})\,\boldsymbol{\beta} + \mathcal{E}^l_a(\mathbf{x}); \qquad a \in \mathcal{M}, \qquad (8)$$

where s_{ia} is -1 if a thickness is subtracted and 1 otherwise. A negative s_{ia} corresponds to an arrow pointing upwards in Figs. 2 and 3. Now, $\boldsymbol{\beta}$ is a column vector including all β's from all intervals. Further, $\mathbf{f}^l_a(\mathbf{x})$ is a stacked vector of the $\mathbf{g}_i(\mathbf{x})$ vectors with a possible change of sign owing to s_{ia}. Similar to Eq. (6), $\mathbf{g}_i(\mathbf{x})$ vectors for $i \notin \mathcal{I}^l_a$ is replaced by zero vectors of corresponding length. Finally, $\mathcal{E}^l_a(\mathbf{x}) = \sum_{i\in\mathcal{I}^l_a} s_{ia}\epsilon_i(\mathbf{x})$.

The covariance between the depth to two subsurfaces at arbitrary locations, Eq. (7), must be replaced by

$$\mathrm{Cov}\Big\{Z^l_a(\mathbf{x}), Z^m_a(\mathbf{y})\Big\} = \sum_{i\in\mathcal{I}^l_a}\sum_{j\in\mathcal{I}^m_a} s_{ia}\,s_{ja}\,\sigma_i(\mathbf{x})\,\sigma_j(\mathbf{y})\,\rho_{ij}(\mathbf{x},\mathbf{y}); \qquad a \in \mathcal{M} \qquad (9)$$

Because these covariances depend on a, the predictors for the depth to all subsurfaces depend on which method in $\mathcal{M}$ has been chosen. The estimators and predictors in the Appendix apply directly to the stochastic model Eq. (8) for any method $a \in \mathcal{M}$.

COMBINING PREDICTORS

Consider the set of predictors $Z_a^{l*}(\mathbf{x})$ for the depth to subsurface l based on the different methods $a \in \mathcal{M}$. These are organized as a vector, $\mathbf{Z}^{l*}(\mathbf{x})$. To obtain a unique predictor a linear combination is used: $Z^{l*}(\mathbf{x}) = \boldsymbol{w}^{l\prime}(\mathbf{x})\,\mathbf{Z}^{l*}(\mathbf{x})$. Assuming each individual prediction, $Z_a^{l*}(\mathbf{x});\ a \in \mathcal{M}$, is unbiased, the combined prediction, $Z^{l*}(\mathbf{x})$, is unbiased provided $\boldsymbol{w}^{l\prime}(\mathbf{x})\,\mathbf{e} = 1$, where $\mathbf{e}$ is a vector of unit entries. The weights are chosen to minimize the combined prediction error. Assuming the covariance matrix, $\mathcal{C}_{ab}^{l}(\mathbf{x}) = \mathrm{Cov}\Big\{Z_a^{l*}(\mathbf{x}) - Z_a^{l}(\mathbf{x}), Z_b^{l*}(\mathbf{x}) - Z_b^{l}(\mathbf{x})\Big\}$, of the predictors is known, the minimum prediction variance is obtained using the location dependent weights

$$\boldsymbol{w}^{l}(\mathbf{x}) = \boldsymbol{\mathcal{C}}^{l^{-1}}(\mathbf{x})\,\mathbf{e} / \Big(\mathbf{e}'\boldsymbol{\mathcal{C}}^{l^{-1}}(\mathbf{x})\,\mathbf{e}\Big)\,. \tag{10}$$

This result is analogous to the weights obtained in ordinary kriging. The minimum prediction variance of the combined predictor is

$$\mathrm{Var}\Big\{Z^{l*}(\mathbf{x}) - Z^{l}(\mathbf{x})\Big\} = \Big(\mathbf{e}'\boldsymbol{\mathcal{C}}^{l^{-1}}(\mathbf{x})\,\mathbf{e}\Big)^{-1}\,, \tag{11}$$

which is always less than or equal to the individual prediction variances. The weighting will favor the predictor with smallest prediction variance at every location.

To calculate the off-diagonal elements of $\boldsymbol{\mathcal{C}}^{l}$ consider two predictors for the depth to subsurface l based on methods $a, b \in \mathcal{M}$: $Z_a^{l*}(\mathbf{x}) = \boldsymbol{\alpha}_a^{l\,\prime}(\mathbf{x})\,\mathbf{Z}_a$ and $Z_b^{l*}(\mathbf{x}) = \boldsymbol{\alpha}_b^{l\prime}(\mathbf{x})\,\mathbf{Z}_b$, where $\boldsymbol{\alpha}^{l}(\mathbf{x})$ are the so-called kriging weights. The prediction covariance is then given by (omitting coordinates)

$$\begin{aligned}\mathcal{C}_{ab}^{l} &= \mathrm{Cov}\Big\{Z_a^{l*} - Z_a^{l}, Z_b^{l*} - Z_b^{l}\Big\} \\ &= \boldsymbol{\alpha}_a^{l\,\prime}\,\mathrm{Cov}\{\mathbf{Z}_a, \mathbf{Z}_b\}\,\boldsymbol{\alpha}_b^{l} - \boldsymbol{\alpha}_a^{l\,\prime}\,\mathrm{Cov}\Big\{\mathbf{Z}_a, Z_b^{l}\Big\} - \mathrm{Cov}\Big\{Z_a^{l}, \mathbf{Z}_b\Big\}\,\boldsymbol{\alpha}_b^{l} + \mathrm{Cov}\Big\{Z_a^{l}, Z_b^{l}\Big\}\,. \end{aligned} \tag{12}$$

The covariances are calculated using a generalization of Eq. (9)

$$\mathrm{Cov}\Big\{Z_a^{l}(\mathbf{x}), Z_b^{m}(\mathbf{y})\Big\} = \sum_{i\in\mathcal{I}_a^{l}}\sum_{j\in\mathcal{I}_b^{m}} s_{ia}s_{jb}\,\sigma_i(\mathbf{x})\,\sigma_j(\mathbf{y})\,\rho_{ij}(\mathbf{x},\mathbf{y}); \qquad a, b \in \mathcal{M} \tag{13}$$

Note that $\boldsymbol{\mathcal{C}}^{l}(\mathbf{x})$ must be computed for all subsurfaces at every location where a prediction is required.

For the multi-layered structure of Fig. 1, $\mathcal{M}$ contains four methods so four predictors must be calculated for each of the subsurfaces. Introducing an additional interval thickness model for the interval from the reference to 'Base reservoir' introduces seven more methods for constructing the depth to the subsurfaces. Then, $\mathcal{M}$ will contain a total of eleven methods and the dimension of $\boldsymbol{\mathcal{C}}^{l}(\mathbf{x})$ becomes 11×11. For more complex multi-layered structures the dimension could become even larger and the computational demands required to obtain all the predictors and prediction covariances might become prohibitive.

COMBINING STOCHASTIC MODELS

This section gives the details of the new proposed method. Consider the set of stochastic models, Eq. (8), for the depth to subsurface l based on the different methods $a \in \mathcal{M}^l$. A unique stochastic model is obtained by forming a linear combination:

$$Z^l(\mathbf{x}) = \sum_{a \in \mathcal{M}^l} w_a^l(\mathbf{x})\, Z_a^l(\mathbf{x}) = \mathbf{f}^{l\prime}(\mathbf{x})\, \boldsymbol{\beta} + \mathcal{E}^l(\mathbf{x}). \tag{14}$$

The weights, $w_a^l(\mathbf{x})$, are the contribution to subsurface l from method $a \in \mathcal{M}^l$. The $\boldsymbol{\beta}$ vector contains the β's from all the interval thickness trends and the corresponding $\mathbf{f}^l(\mathbf{x})$ vector has the form

$$\mathbf{f}^{l\prime}(\mathbf{x}) = \begin{bmatrix} \cdots & W_i^l(\mathbf{x})\, \mathbf{g}_i'(\mathbf{x}) & \cdots \end{bmatrix}, \tag{15}$$

where

$$W_i^l(\mathbf{x}) = \begin{cases} \sum_{a \in \mathcal{M}^l} s_{ia}\, w_a^l(\mathbf{x}) & \text{if } i \in \mathcal{I}^l \\ 0 & \text{else.} \end{cases}$$

Here $s_{ia} = 1$ if $\Delta Z_i(\mathbf{x})$ adds to $Z_a^l(\mathbf{x})$ whereas $s_{ia} = -1$ if $\Delta Z_i(\mathbf{x})$ must be subtracted. If $i \notin \mathcal{I}^l$ the corresponding $W_i^l(\mathbf{x})\, \mathbf{g}_i(\mathbf{x})$ term in Eq. (15) is replaced by a zero vector of corresponding length. The weights $W_i^l(\mathbf{x})$ give the total contribution from interval i to subsurface l. The residual in Eq. (14) is

$$\mathcal{E}^l(\mathbf{x}) = \sum_{a \in \mathcal{M}^l} \sum_{i \in \mathcal{I}_a^l} s_{ia}\, w_a^l(\mathbf{x})\, \epsilon_i(\mathbf{x}) = \sum_{i \in \mathcal{I}^l} W_i^l(\mathbf{x})\, \epsilon_i(\mathbf{x}). \tag{16}$$

The weights, $w_a^l(\mathbf{x})$, are found by minimizing the residual variance $\mathrm{Var}\Big\{\mathcal{E}^l(\mathbf{x})\Big\}$ subjected to the restriction $\mathbf{w}^{l\prime}(\mathbf{x})\, \mathbf{e} = 1$. To justify this restriction consider the expectation of Eq. (14):

$$\mathrm{E}\Big\{Z^l(\mathbf{x})\Big\} = \sum_{a \in \mathcal{M}^l} w_a^l(\mathbf{x})\ \mathrm{E}\Big\{Z_a^l(\mathbf{x})\Big\}.$$

By assumption, all $\{Z_a^l(\mathbf{x});\, a \in \mathcal{M}^l\}$ are possible stochastic models for the depth to subsurface l. This means that the expected subsurface for all models should be approximately equal, giving $\sum_{a \in \mathcal{M}^l} w_a^l(\mathbf{x}) \approx 1$, or in vector notation $\mathbf{w}^{l\prime}(\mathbf{x})\, \mathbf{e} \approx 1$. Then, choosing to minimize the residual error of the combined model under this restriction gives

$$\mathbf{w}^l(\mathbf{x}) = \mathbf{C}^{l^{-1}}(\mathbf{x})\, \mathbf{e} / \Big(\mathbf{e}' \mathbf{C}^{l^{-1}}(\mathbf{x})\, \mathbf{e}\Big), \tag{17}$$

where $C_{ab}^l(\mathbf{x}) = \mathrm{Cov}\Big\{Z_a^l(\mathbf{x}), Z_b^l(\mathbf{x})\Big\}$ is given by Eq. (13) with $a, b \in \mathcal{M}^l$. The obtained weights, $w_a^l(\mathbf{x})$, will favor the most accurate models, that is, the models having smaller specified residual variance, $\mathrm{Var}\Big\{\mathcal{E}_a^l(\mathbf{x})\Big\}$. The combined residual variance is

$$\mathrm{Var}\Big\{Z^l(\mathbf{x})\Big\} = \Big(\mathbf{e}' \mathbf{C}^{l^{-1}}(\mathbf{x})\, \mathbf{e}\Big)^{-1} \tag{18}$$

which is always less than or equal to $\mathrm{Var}\Big\{Z_a^l(\mathbf{x})\Big\}$ for any method $a \in \mathcal{M}^l$.

The dimension of $\mathbf{C}^l(\mathbf{x})$ equals the number of methods in $\mathcal{M}^l$ which depends on l. This dimension is less than or equal to the dimension of $\boldsymbol{\mathcal{C}}^l(\mathbf{x})$ which is equal to the number of models in $\mathcal{M}$.

Subsurface Prediction for Combined Stochastic Models

The parameter estimators and predictors reviewed in the Appendix still hold, but the contents of some of the matrices and vectors must be slightly modified. The $\mathbf{f}^l(\mathbf{x})$ vectors will have the form given by Eq. (15) and the row in $\mathbf{F}$ corresponding to observation $Z^l(\mathbf{x}_i)$ is—as before—given by $\mathbf{f}^{l\prime}(\mathbf{x}_i)$. The covariances, $\mathbf{K} = \mathrm{Var}\{\mathbf{Z}\}$ and $\mathbf{k}^l(\mathbf{x}) = \mathrm{Cov}\Big\{Z^l(\mathbf{x}), \mathbf{Z}\Big\}$, are now deduced from Eq. (16) giving

$$\begin{aligned}\mathrm{Cov}\Big\{Z^l(\mathbf{x}), Z^m(\mathbf{y})\Big\} &= \sum_{\substack{a\in\mathcal{M}^l\\ b\in\mathcal{M}^m}} \sum_{\substack{i\in\mathcal{I}^l_a\\ j\in\mathcal{I}^m_b}} s_{ia}s_{jb}\, w^l_a(\mathbf{x})\, w^m_b(\mathbf{y})\, \sigma_i(\mathbf{x})\, \sigma_j(\mathbf{y})\, \rho_{ij}(\mathbf{x},\mathbf{y}) \\ &= \sum_{\substack{i\in\mathcal{I}^l\\ j\in\mathcal{I}^m}} W^l_i(\mathbf{x})\, W^m_j(\mathbf{y})\, \sigma_i(\mathbf{x})\, \sigma_j(\mathbf{y})\, \rho_{ij}(\mathbf{x},\mathbf{y}). \qquad (19)\end{aligned}$$

Co-linearities and Bayesian Kriging

A major problem with combining models is the large co-linearities in $\mathbf{F}$. The most striking consequence is the occurrence of extreme correlations in $\widehat{\mathbf{\Sigma}} = \mathrm{Var}\Big\{\widehat{\boldsymbol{\beta}}\Big\} = (\mathbf{F}'\mathbf{K}^{-1}\mathbf{F})^{-1}$. The large co-linearities are caused by the property that alternative methods can give almost identical predictions. One could argue that the estimates of the β parameters are of minor interest since prediction is the primary objective, but unfortunately, the example in the next section shows that even the predictor is vulnerable to this deficiency. In the most extreme cases, $\mathbf{F}$ becomes rank deficient so that the inverse of $\mathbf{F}'\mathbf{K}^{-1}\mathbf{F}$ is undefined and estimators and predictors fail to exist.

A simple remedy is to impose prior distributions on the β parameters. The β parameters will usually have a simple physical interpretation so prior knowledge is commonly available. The prior distribution will restrict the parameter space so that extreme correlations are reduced, and more importantly, the associated extreme estimation variances are reduced. If the prior distribution is assumed multinormal, the Bayesian kriging predictor (Kitanidis 1986, Omre 1987, Omre and Halvorsen 1989, Abrahamsen 1993) applies.

EXAMPLE

In this section combining models are compared to combining predictors for a synthetic example. The estimated trend will be used rather than the kriging predictor since:

- the local fitting to data points done by kriging is almost insensitive to the choice of method and therefore masks differences.
- kriging predictors coincide with estimated trends away from observations so results are valid outside the local influence of observations.
- the local fitting stabilizes the predictors so using the trends focus on worst-case behavior.

Using the trends will exaggerate potential problems and differences between the methods but the conclusions reached will carry over to the less sensitive kriging predictors. The estimated trend Eq. (A2) will be used instead of the kriging predictor Eq. (A4), and the unconditional prediction error (trend error) Eq. (A5) will be used for the prediction error. The corresponding estimated trend and trend error with Bayesian priors on $\boldsymbol{\beta}$ are Eq. (A6)

and Eq. (A8) respectively.

Stochastic Models

Consider the schematic cross section of a reservoir formation illustrated in Fig. 1. Simple isochore models assuming a constant thickness of each reservoir zone have been chosen [see Eq. (4a)]:

$$\Delta Z_{\mathrm{Z}i}(x) = \beta_{\mathrm{Z}i} + \epsilon_{\mathrm{Z}i}(x); \qquad i = 1, 2, 3. \tag{20a}$$

Moreover, $\Delta Z_{\mathrm{TR}}(x)$ and $\Delta Z_{\mathrm{R}}(x)$ are given as [see Eq. (4b)]:

$$\Delta Z_{\mathrm{TR}}(x) = \left[\beta_{\mathrm{TR1}} + \beta_{\mathrm{TR2}}\{t_{\mathrm{TR}}(x) - \mathrm{mean}(t_{\mathrm{TR}}(x))\}\right] t_{\mathrm{TR}}(x) + \epsilon_{\mathrm{TR}}(x) \tag{20b}$$

$$\Delta Z_{\mathrm{R}}(x) = \left[\beta_{\mathrm{R1}} + \beta_{\mathrm{R2}}\frac{5-x}{10}\right] \Delta t_{\mathrm{R}}(x) + \epsilon_{\mathrm{R}}(x) \tag{20c}$$

where $x \in \mathbf{R}$ since only a cross-section is considered. The expressions in the square brackets are the interval velocities. A cross-section of the travel-times appears in Fig. 4. A positive value for β_{TR2} implies the usual velocity increase at the flanks causing the subsurfaces to be more curved than the travel-times. Moreover, a positive value for β_{R2} leads to a reduced interval velocity for higher x values causing 'Base reservoir' to tilt upwards towards the right.

The standard errors of the interval residuals are specified as

$$\sigma_{\mathrm{TR}}(x) = \sigma_{\mathrm{R}}(x) = 0.1, \quad \sigma_{\mathrm{Z3}}(x) = \sigma_{\mathrm{Z2}}(x) = \sigma_{\mathrm{Z1}}(x) = 0.2,$$

and are assumed independent. The correlation lengths of the residuals are assumed to be less than the separation between observations so that spatial correlations do not interfere.

When combining models, the relevant subsurface descriptions are illustrated in Fig. 3, so that $\mathbf{C}^l$ has dimension one for 'Top reservoir' and dimension two for the three other subsurfaces. Note that $\mathbf{C}^l$ is independent of x since the standard error of the residuals are assumed constant. The resulting combined residual errors from Eq. (19), and weights from Eq. (17), are given in Table 1. The table shows that the weights favor the most accurate method and that residual errors are systematically reduced.

Assuming all β-parameters equal to one and combining the trends according to Eq. (14), using weights calculated by Eq. (17), give the depth trends illustrated in Fig. 5. It is seen how the shape of the intermediate subsurfaces reveals a transition from the shape of the upper seismic reflector to the base seismic reflector. This shows that the weights favor the shape of the closest reflector but also takes the shape of the other reflector into account.

Simulation Experiments

To compare the different approaches the x-axis has been divided into three segments of equal length. Within each segment, the location of a vertical well has been drawn from a uniform distribution. Then, the actual observation has been drawn with the expectations given by all trends in Fig. 5, and the covariances from Eq. (19). Ten sets of observations have been drawn and the β-parameters have been estimated. The resulting ten sets of trends are plotted in Fig. 6. Some of the trends are far off the 'true' trends in Fig. 5, an effect caused by the severe co-linearities making it almost impossible to estimate some of the β-parameters.

Imposing a prior distribution on the β-parameters with expectations 0.5 (a severe bias) and $\mathrm{Cov}\{\boldsymbol{\beta}\} = \mathrm{diag}(2)$ dramatically improves the estimated β-parameters. The corresponding ten trends are plotted in Fig. 7. The prior variances—although unrealistically large—effectively restrict the parameter space so that extreme β estimates are prohibited. Moreover, the prior expectations are effectively overruled by the data.

As a final comparison, estimated trends based on the four models illustrated in Fig. 2 have been combined using space dependent weights, $w^l(\mathbf{x})$, obtained from Eq. (10). The ten sets of combined trends are plotted in Fig. 8. The similarity to Fig. 7 is striking.

The model combination method is approximately ten times faster to calculate on a computer than the approach combining estimated trends.

Bias and Errors

An important concern is possible bias and the accuracy of the calculated prediction errors. To investigate this, one hundred sets of observations have been drawn using the algorithm described above. Fig. 9 displays the average bias of the resulting hundred estimated trends for subsurface 'Top zone 2'. It is seen that all three trend estimators behave well, in the sense that deviations from zero are small. This is expected since the model assumptions for the estimators agree with the model that generated the data.

Fig. 10 shows the empirical trend errors whereas Fig. 11 depicts the average theoretical trend errors. The two figures show that the model combination approach have difficulties caused by the co-linearities, but the two other approaches ensure good agreement between theoretical and empirical prediction errors. In particular the approach using Bayesian estimation gives an almost exact agreement.

The approach based on Bayesian estimation and model combination and the approach combining estimated trends are very similar. There is some bias at the edges for the combined estimated trends (see Fig. 9) making the empirical prediction error slightly larger.

The final test is to violate the model assumptions. By choosing $\sigma_{\mathrm{R}}(x) = 0.2$ rather than 0.1, the weights are modified in favor of models not including subsurface 'Base reservoir'. Data have been drawn from this model, but the model assumptions entering the trend estimators have been kept unchanged. Fig. 12 shows that the bias has been more than doubled for all three approaches so that the empirical errors given in Fig. 13 have increased. For the model combination approach the increase is significant, but for the two other approaches, the increase is less than 10%. The average theoretical prediction error in Fig. 11 now reveals underestimation of the true error. The approach using model combination with vague priors on the β parameters still has a slightly smaller empirical error. Comparing to the theoretical errors in Fig. 11 still gives agreement within approximately 10%.

CLOSING REMARKS

Two solutions to the fundamental problem of combining different methods for obtaining the depths to subsurfaces have been discussed. The first method combines alternative depth predictions and the second method merges alternative stochastic models. The latter method is less computer intensive but suffers from instabilities which are removed by imposing a prior distribution on the trend parameters.

When combining predictors a rigorous minimization criteria for the prediction error is used. The approach combining models however, uses a heuristic minimization criteria for the residual variance. The usefulness of this method is therefore justified by its properties. The two methods gave almost identical results for the synthetic example, so in this case it is possible to conclude that the model combination approach performs equally good using ten times less computer resources.

ACKNOWLEDGMENTS

The work was supported by a research fellowship from The Research Council of Norway and benefited from facilities at Norwegian Computing Center. The author is grateful to Erik Bølviken, Magne Aldrin, and Gudmund Høst for many valuable comments. I would also like to express my gratitude to Arne Skorstad and Geir Johansen who are currently implementing the methods in the commercially available STORM reservoir modeling package by Geomatic.

APPENDIX: PARAMETER ESTIMATION AND KRIGING

This appendix reviews predictors for a layered geological structure developed in Abrahamsen (1993).

Assume there exist observations of L subsurfaces at n distinct locations:

$$\mathbf{Z}' = \left[Z^1(\mathbf{x}_1), \ldots, Z^1(\mathbf{x}_n), \ldots, Z^L(\mathbf{x}_1), \ldots, Z^L(\mathbf{x}_n)\right].$$

Observations from each subsurface are assumed to be from identical locations for notational reasons, corresponding to observations from vertical wells. The linear model Eq. (5) yield

$$\mathbf{Z} = \mathbf{F}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \tag{A1}$$

where $\boldsymbol{\varepsilon}$ are the residuals corresponding to $\mathbf{Z}$ and the row in $\mathbf{F}$ corresponding to observation $Z^l(\mathbf{x}_i)$ is $\mathbf{f}^{l\prime}(\mathbf{x}_i)$.

Eq. (A1) is a linear regression model for all the β parameters with correlated Gaussian residual errors. The maximum likelihood estimates for $\boldsymbol{\beta}$ are the generalized least squares estimates:

$$\widehat{\boldsymbol{\beta}} = \left(\mathbf{F}'\mathbf{K}^{-1}\mathbf{F}\right)^{-1}\mathbf{F}'\mathbf{K}^{-1}\mathbf{Z}$$
$$\widehat{\boldsymbol{\Sigma}} = \mathrm{Var}\left\{\widehat{\boldsymbol{\beta}}\right\} = \left(\mathbf{F}'\mathbf{K}^{-1}\mathbf{F}\right)^{-1},$$

where the covariance matrix $\mathbf{K} = \mathrm{Var}\{\mathbf{Z}\} = \mathrm{Var}\{\boldsymbol{\varepsilon}\}$ is calculated using Eq. (7). The estimated trend for subsurface l is

$$\widehat{Z}^l(\mathbf{x}) = \mathbf{f}^{l\prime}(\mathbf{x})\,\widehat{\boldsymbol{\beta}}, \tag{A2}$$

with the corresponding estimation variance

$$\mathrm{Var}\left\{\widehat{Z}^l(\mathbf{x})\right\} = \mathbf{f}^{l\prime}(\mathbf{x})\,\widehat{\boldsymbol{\Sigma}}\mathbf{f}^l(\mathbf{x}). \tag{A3}$$

Universal Kriging

The best linear unbiased predictor for Gaussian random fields with a linear trend is the universal kriging predictor. In the present context all observations should be considered giving a universal co-kriging predictor for say subsurface l:

$$Z^{l*}(\mathbf{x}) = \widehat{Z}^l(\mathbf{x}) + \mathbf{k}^{l\prime}(\mathbf{x})\,\mathbf{K}^{-1}\left(\mathbf{Z} - \mathbf{F}\widehat{\boldsymbol{\beta}}\right), \tag{A4}$$

where $\mathbf{k}^l(\mathbf{x}) = \mathrm{Cov}\left\{\mathcal{E}^l(\mathbf{x}), \boldsymbol{\mathcal{E}}\right\}$. The corresponding prediction variance is

$$\begin{aligned}\mathrm{Var}\left\{Z^{l*}(\mathbf{x}) - Z^l(\mathbf{x})\right\} = \mathrm{Var}\left\{Z^l(\mathbf{x})\right\} - \mathbf{k}^{l\prime}(\mathbf{x})\,\mathbf{K}^{-1}\mathbf{k}^l(\mathbf{x}) \\ + \left[\mathbf{f}^l(\mathbf{x}) - \mathbf{F}'\mathbf{K}^{-1}\mathbf{k}^{l\prime}(\mathbf{x})\right]'\widehat{\boldsymbol{\Sigma}}\left[\mathbf{f}^l(\mathbf{x}) - \mathbf{F}'\mathbf{K}^{-1}\mathbf{k}^{l\prime}(\mathbf{x})\right].\end{aligned}$$

For $\mathbf{x}$ far away from all observations, $\mathbf{k}^l(\mathbf{x}) \to \mathbf{0}$, so that the prediction error approach

$$\mathrm{Var}\left\{Z^l(\mathbf{x})\right\} + \mathrm{Var}\left\{\widehat{Z}^l(\mathbf{x})\right\}. \tag{A5}$$

This is called the unconditional prediction variance or trend variance.

Bayesian Kriging

Assume a prior multinormal distribution on the β parameters is given: $\boldsymbol{\beta} \sim N_p(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$. The posterior expectation and covariances become:

$$\begin{aligned}\widehat{\boldsymbol{\mu}}_\mathrm{b} &= E\{\boldsymbol{\beta}|\mathbf{Z}\} = \boldsymbol{\mu}_0 + \boldsymbol{\Sigma}_0\mathbf{F}^T\mathbf{K}_Z^{-1}(\mathbf{Z} - F\boldsymbol{\mu}_0) \\ \widehat{\boldsymbol{\Sigma}}_\mathrm{b} &= \mathrm{Var}\{\boldsymbol{\beta}|\mathbf{Z}\} = \boldsymbol{\Sigma}_0 - \boldsymbol{\Sigma}_0\mathbf{F}^T\mathbf{K}_Z^{-1}\mathbf{F}\boldsymbol{\Sigma}_0,\end{aligned}$$

where $\mathbf{K}_Z = \mathrm{Var}\{\mathbf{Z}\} = \mathbf{F}\boldsymbol{\Sigma}_0\mathbf{F}' + \mathbf{K}$. The posterior trend and trend variance is given by replacing $\widehat{\boldsymbol{\mu}}$ with $\widehat{\boldsymbol{\mu}}_\mathrm{b}$ in Eq. (A2), and $\widehat{\boldsymbol{\Sigma}}$ with $\widehat{\boldsymbol{\Sigma}}_\mathrm{b}$ in Eq. (A3), i.e.

$$\widehat{Z}^l_\mathrm{b}(\mathbf{x}) = \mathbf{f}^{l\prime}(\mathbf{x})\,\widehat{\boldsymbol{\beta}}_\mathrm{b} \tag{A6}$$

$$\mathrm{Var}\left\{\widehat{Z}^l_\mathrm{b}(\mathbf{x})\right\} = \mathbf{f}^{l\prime}(\mathbf{x})\,\widehat{\boldsymbol{\Sigma}}_\mathrm{b}\mathbf{f}^l(\mathbf{x}). \tag{A7}$$

The Bayesian kriging predictor and prediction variance is given by:

$$\begin{aligned}Z^{l*}_\mathrm{b}(\mathbf{x}) &= \mathbf{f}^l(\mathbf{x}) \cdot \widehat{\boldsymbol{\mu}}_\mathrm{b} + \mathbf{k}^l(\mathbf{x})\mathbf{K}^{-1}(\mathbf{Z} - \mathbf{F}\widehat{\boldsymbol{\mu}}_\mathrm{b}) \\ \mathrm{Var}\left\{Z^{l*}_\mathrm{b}(\mathbf{x}) - Z^l(\mathbf{x})\right\} &= \mathrm{Var}\left\{\mathcal{E}^l(\mathbf{x})\right\} + \mathbf{f}^{l\prime}(\mathbf{x})\,\boldsymbol{\Sigma}_0\mathbf{f}^l(\mathbf{x}) - \mathbf{k}^l_Z(\mathbf{x})\,\mathbf{K}_Z^{-1}\mathbf{k}^l_Z{}'(\mathbf{x}),\end{aligned}$$

where $\mathbf{k}^l_Z(\mathbf{x}) = \mathrm{Cov}\left\{Z^l(\mathbf{x}), \mathbf{Z}\right\} = \mathbf{f}^{l\prime}(\mathbf{x})\,\boldsymbol{\Sigma}_0\mathbf{F}^T + \mathbf{k}^l(\mathbf{x})$. The 'unconditional' prediction variance or trend variance is:

$$\mathrm{Var}\left\{\mathcal{E}^l(\mathbf{x})\right\} + \mathrm{Var}\left\{\widehat{Z}^l_\mathrm{b}(\mathbf{x})\right\}. \tag{A8}$$

REFERENCES

Abrahamsen, P., 1993, Bayesian kriging for seismic depth conversion of a multi-layer reservoir, *in* A. Soares, ed., Geostatistics Tróia '92, Vol. 5 of Quantitative Geology and Geostatistics, proceedings from '4th International Geostatistics Congress', Tróia Portugal, 1992, Kluwer Academic Publishers, Dordrecht, p. 385–398.

Acheson, C. H., 1963, Time-depth and velocity-depth relations in western Canada, Geophysics v. 28, no. 5, p. 894–909.

Al-Chalabi, M., 1974, An analysis of stacking, RMS, average, and interval velocities over a horizontally layered ground, Geophysical Prospecting v. 22, p. 458–475.

Al-Chalabi, M., 1979, Velocity determination from seismic reflection data, *in* A. A. Fitch, ed., Developments in Geophysical Exploration Methods—1, Applied Science Publishers Ltd., London, chapter 1, p. 1–68.

Breiman, L., 1992, Stacked regression, Technical Report 367, Department of Statistics, University of California Berkeley, California 94720, 15 p.

Bunn, D., 1989, Forecasting with more than one model, J. Forecasting v. 8, no. 3, p. 161–166.

Cressie, N., 1991, Statistics for Spatial Data, John Wiley & Sons, New York, 900 p.

Delfiner, P., and Delhomme, J. P., 1975, Optimum interpolation by kriging, *in* J. C. Davis and M. J. McCullagh, eds, Display and Analysis of Spatial Data, John Wiley & Sons, New York.

Faust, I. Y., 1951, Seismic velocity as a function of depth and geological time, Geophysics v. 16, p. 192–206.

Granger, C. W. J., 1989, Invited review: Combining forecasts—twenty years later, J. Forecasting v. 8, no. 3, p. 167–173.

Journel, A. G., and Huijbregts, C. J., 1978, Mining Geostatistics, Academic Press Inc., London, 600 p.

Kitanidis, P. K., 1986, Parameter uncertainty in estimation of spatial functions: Bayesian analysis, Water Resour. Res. v. 22, no. 4, p. 499–507.

LeBlanc, M., and Tibshirani, R., 1993, Combining estimates in regression and classification, Technical report, Univ. of Toronto, 21 p.

Matheron, G., 1963, Principles of geostatistics, Economic Geol. v. 58, p. 1246–1266.

Olea, R. A., 1974, Optimal contour mapping using universal kriging, J. Geophys. Res. v. 79, no. 5, p. 695–702.

Omre, H., 1987, Bayesian kriging—merging observations and qualified guesses in kriging, Math. Geol. v. 19, no. 1, p. 25–39.

Omre, H., and Halvorsen, K. B., 1989, The Bayesian bridge between simple and universal kriging, Math. Geol. v. 21, no. 7, p. 767–786.

Ripley, B. D., 1981, Spatial Statistics, John Wiley & Sons, New York, 252 p.

Walden, A. T., and White, R. E., 1984, On errors of fit and accuracy in matching synthetic-seismograms and seismic traces, Geophysical Prospecting v. 32, no. 5, p. 871–891.

White, R. E., 1984, Signal and noise estimation from seismic reflection data using spectral coherence methods, Proc. IEEE v. 72, no. 10, p. 1340–1356.

Wolpert, D., 1992, Stacked generalization, Neural Networks v. 5, p. 241–259.

Table. 1: Specified residual errors σ_a^l, calculated weights w_a^l, and combined residual errors σ^l. The methods correspond to labels in Fig. 3. Note how the weights favor the assumed most accurate models.

Subsurface l	Method a	Res. error $\mathrm{Var}\{Z_a^l\}^{1/2}$	Weight w_a^l	Comb. res. error $\mathrm{Var}\{Z^l\}^{1/2}$
TR	(a–d)	0.1	1	0.1
T2	(a)	0.316	0.31	0.194
	(b–d)	0.224	0.69	
T1	(a,b)	0.245	0.62	0.202
	(c,d)	0.300	0.38	
BR	(a–c)	0.141	0.92	0.139
	(d)	0.361	0.08	

List of Figures

Fig. 11 Average theoretical error from 100 simulations. (- - -) combined model. (—) combined model with priors on β parameters. $(\cdots)$ combined predictors.

Fig. 12 Empirical bias from 100 simulations. Data drawn from model with $\sigma_{\mathrm{R}}(x) = 0.2$. (- - -) combined model. (—) combined model with priors on β parameters. $(\cdots)$ combined predictors.

Fig. 13 Empirical error from 100 simulations. Data drawn from model with $\sigma_{\mathrm{R}}(x) = 0.2$. (- - -) combined model. (—) combined model with priors on β parameters. $(\cdots)$ combined predictors.

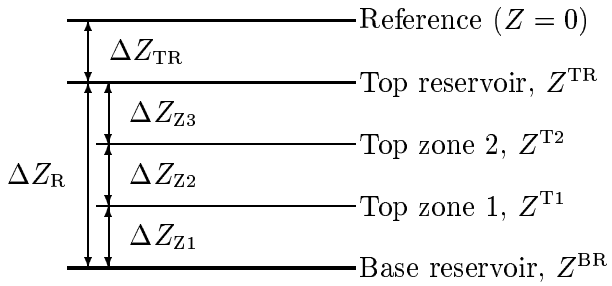

Fig. 1: Schematic illustration of a reservoir formation. The double arrows indicate a stochastic model for the thickness of the corresponding interval ΔZ_i. 'Top reservoir' and 'Base reservoir' are seismic reflectors.

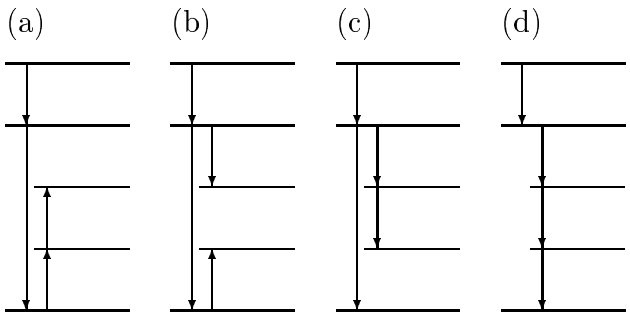

Fig. 2: Four alternative methods for determining the depth to subsurfaces. The horizontal lines are the subsurfaces given in Fig. 1. The depth to a particular subsurface is found by following the arrows; an arrow pointing downwards means that the corresponding thickness is added whereas an arrow pointing upwards means that the corresponding thickness is subtracted.

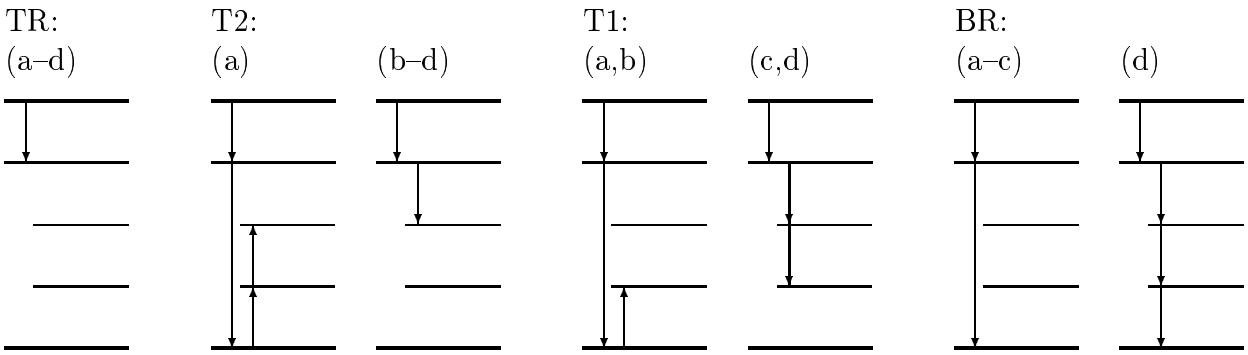

Fig. 3: Alternative methods for obtaining the depth to 'Top reservoir' (TR), 'Top zone 2' (T2), 'Top zone 1' (T1), and 'Base reservoir' (BR). Labels correspond to the illustrations shown in Fig. 2. The horizontal lines and the arrows have the same meaning as in Fig. 2.

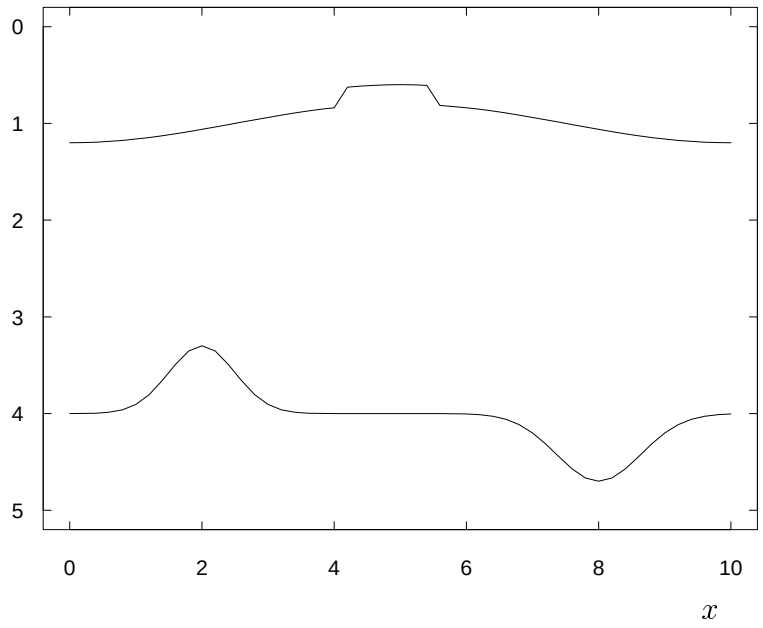

Fig. 4: Travel times to 'Top reservoir' and 'Base reservoir'.

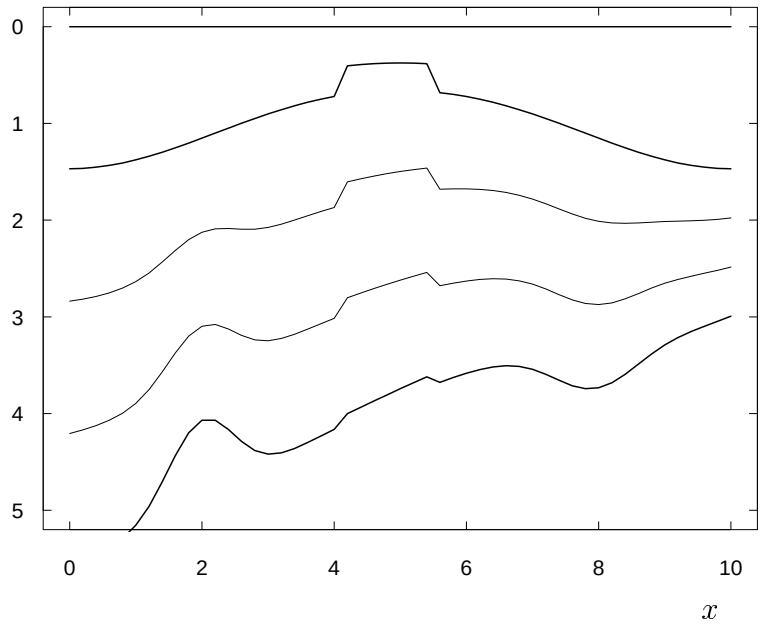

Fig. 5: Depth trends obtained when choosing all β's equal to one.

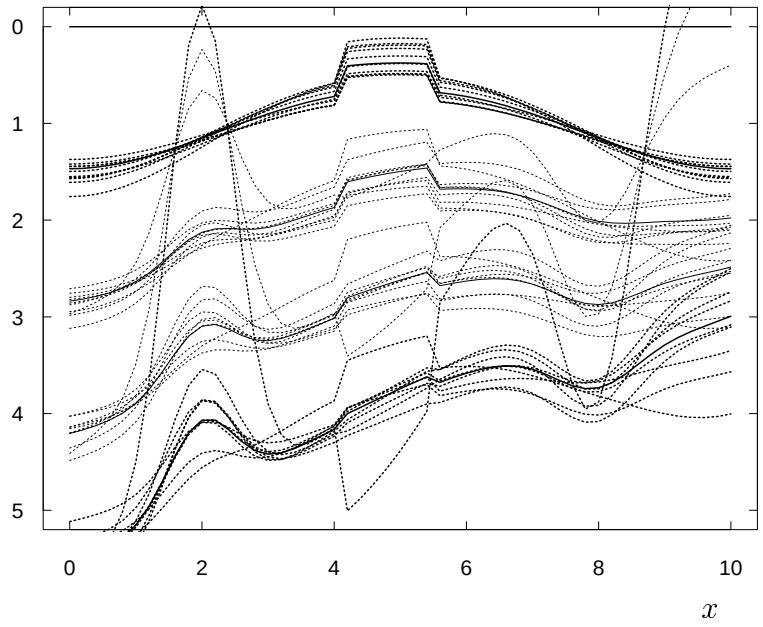

Fig. 6: Ten sets of trends obtained using a combined model. Each set is conditioned on data drawn from the model.

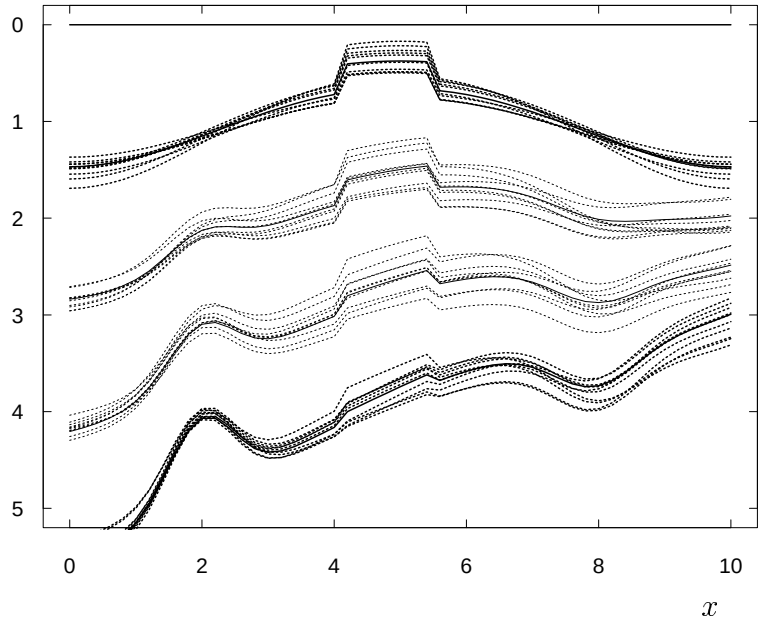

Fig. 7: Ten sets of trends obtained using a combined model and a Bayesian prior on the β parameters. Each set is conditioned on data drawn from the model.

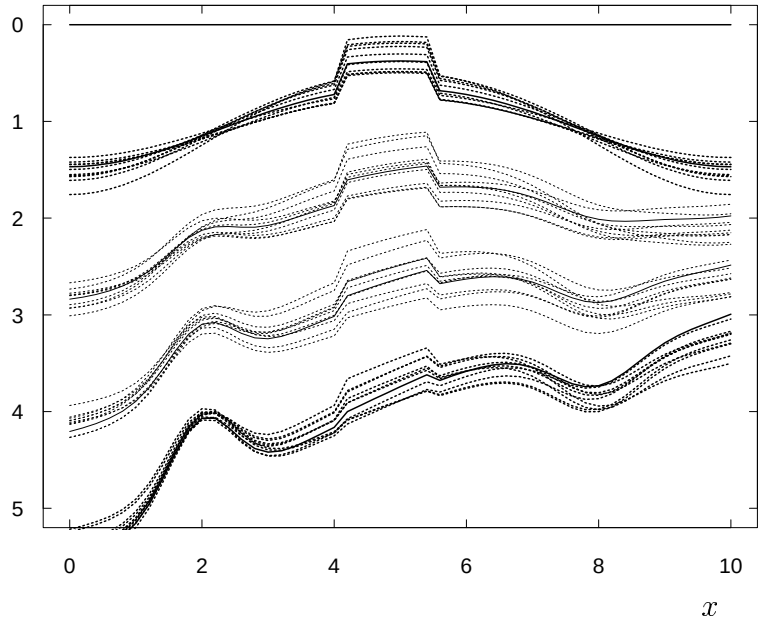

Fig. 8: Ten sets of trends obtained by combining estimated trends based on different models. Each set is conditioned on data drawn from the model.

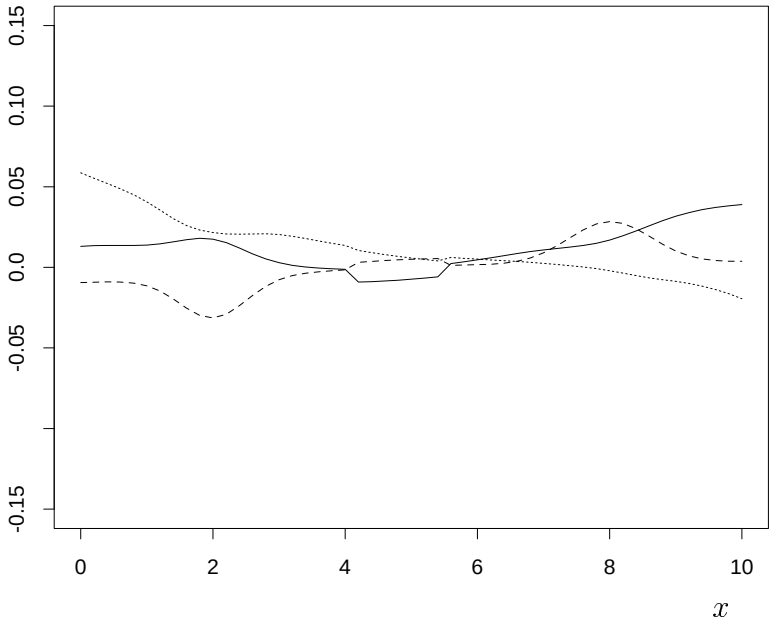

Fig. 9: Empirical bias from 100 simulations. (- - -) combined model. (—) combined model with priors on β parameters. $(\cdots)$ combined predictors.

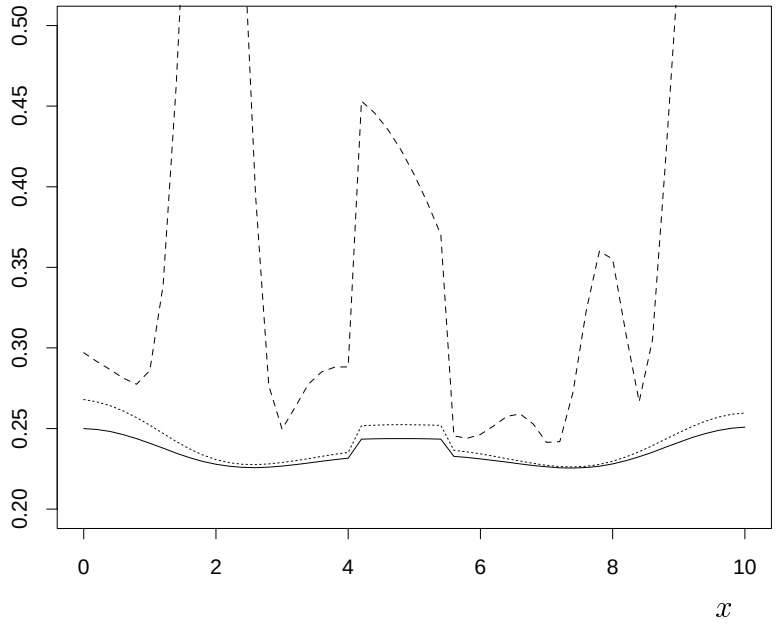

Fig. 10: Empirical error from 100 simulations. (- - -) combined model. (—) combined model with priors on β parameters. $(\cdots)$ combined predictors.

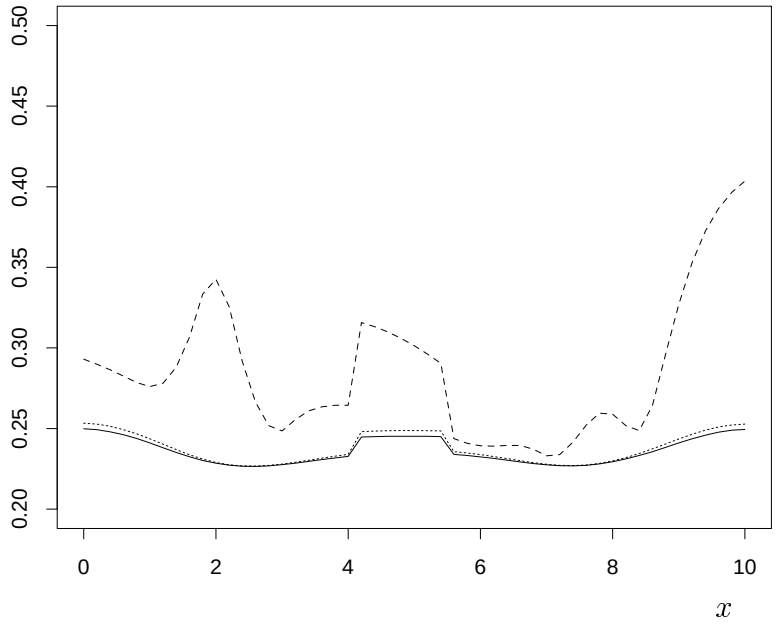

Fig. 11: Average theoretical error from 100 simulations. (- - -) combined model. (—) combined model with priors on β parameters. $(\cdots)$ combined predictors.

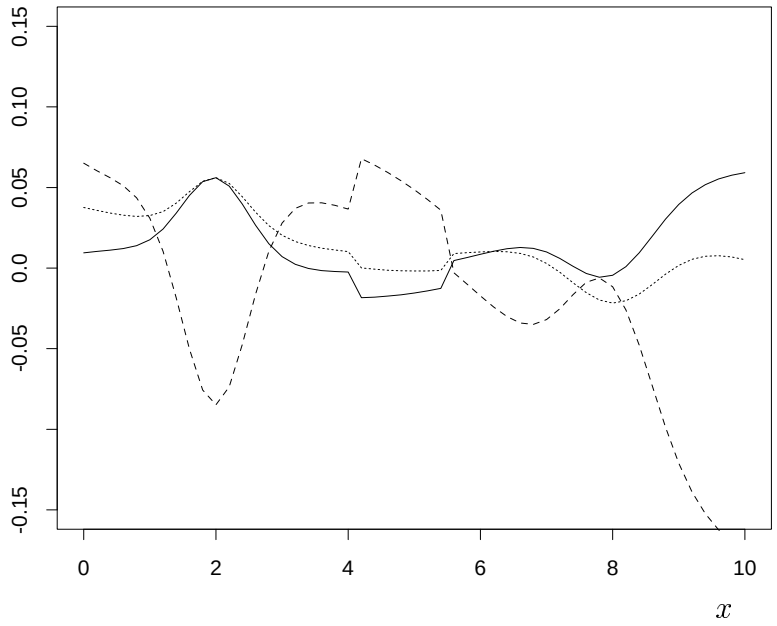

Fig. 12: Empirical bias from 100 simulations. Data drawn from model with $\sigma_{\mathrm{R}}(x) = 0.2$. (- - -) combined model. (—) combined model with priors on β parameters. $(\cdots)$ combined predictors.

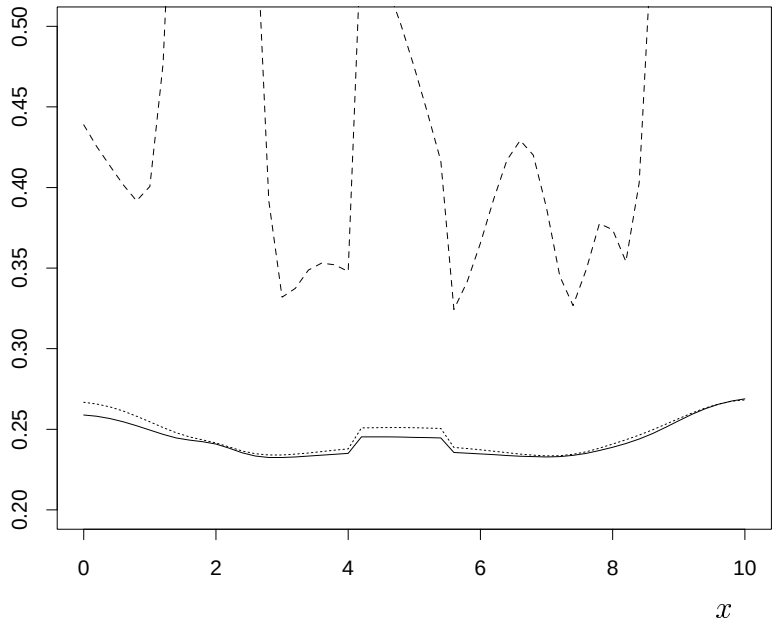

Fig. 13: Empirical error from 100 simulations. Data drawn from model with $\sigma_{\mathrm{R}}(x) = 0.2$. (- - -) combined model. (—) combined model with priors on β parameters. $(\cdots)$ combined predictors.