P. Abrahamsen (1993), "Bayesian Kriging for Seismic Depth Conversion of a Multi-layer Reservoir" in A. Soares (ed.) "Geostatistics Troia '92". Kluwer Academic Publ., Dordrecht, pp. 385-398.

BAYESIAN KRIGING FOR SEISMIC DEPTH CONVERSION OF A MULTI-LAYER RESERVOIR

Petter Abrahamsen
Norwegian Computing Center
P.O.Box 114, Blindern
N-0314 Oslo
Norway

Abstract. A stochastic model for a petroleum reservoir with L seismic subsurfaces is presented. The seismic velocities within each layer is described by linear regression models and Gaussian random fields. Seismic interpretation errors are modeled as Gaussian random fields. Intercorrelations between all subsurfaces and all velocity fields are taken into consideration. This simplifies the handling of deviating wells and ensures consistent prediction and prediction variances for all L subsurfaces and L velocity fields. Bayesian kriging is used for prediction of subsurfaces and velocity fields.

In the limit corresponding to exact prior knowledge, the Bayesian method is equivalent to cokriging with $2L$ covariables. In the limit corresponding to no prior knowledge, the method is equivalent to a combination of universal kriging and cokriging with $2L$ dependent regression models and $2L$ covariables.

1 Introduction

The large scale geometry of a petroleum reservoir is usually described by a set of geological subsurfaces separating almost homogeneous layers. Available information for the depth to the subsurfaces are precise depth observations in wells, and less precise information from seismic travel times. The travel times are recordings of the time used by a sonic pulse from the surface to a reflecting subsurface. The importance of the seismic travel times are their lateral coverage. The travel times are usually available on a fine-meshed grid which allows an almost continuous — but inexact — description of the lateral depth trends. On the other hand, the almost exact well measurements are available in just a few locations possibly several kilometers apart. The challenge is therefore to combine the exact well measurements with the general trends from the travel times. This requires the simple kinematic relation between depth, the velocity of sound, and travel time: $z = v \cdot t$. In the next section, spatial stochastic models for v and t are established, and the resulting model for z is considered. Using kriging techniques, predictions for the velocity field and the depth

are available. The process of calculating the depth to a seismic reflector from the measured travel times is known as depth conversion.

In most applications several seismic reflectors are observed above the reservoir layers under study. The velocity of sound changes significantly at seismic reflectors. Thus, a separate model for the velocity field in every layer should be used to obtain the best results. The traditional approach to depth conversion is to predict the thickness of each layer independently. This paper proposes a method which considers all subsurfaces and velocity fields in one consistent model.

A general problem — at least in off-shore applications — is the very limited number of wells. The lack of data must be compensated by prior information, so Bayesian statistics is necessary. This ensures stable and reasonable results for any number of well observations — including zero.

The paper is organized as follows: the next section describes the stochastic models, Section 3 reviews Bayesian kriging and outlines how the current problem is organized to fit into the linear kriging machinery. Section 4 gives an example which illustrates the properties of the proposed method. Section 5 close the discussion by some final remarks.

2 Stochastic model for depth conversion

The fine-meshed grids containing the travel times are regarded as continuous 2-dimensional fields. The recorded travel times to subsurface l is denoted $t_l(\mathbf{x})$, where $\mathbf{x} \in \mathbf{R}^2$ is a lateral reference. The average vertical velocity of sound within a layer is called the interval velocity and is denoted $v_l(\mathbf{x})$. If both the travel times and the interval velocities are exactly known, the depth to subsurface L, $z_L(\mathbf{x})$, is found as the sum of the interval thicknesses, $\Delta z_l(\mathbf{x})$, above:

$$z_L(\mathbf{x}) = \sum_{l=1}^{L} \Delta z_l(\mathbf{x}), \tag{1}$$

where $\Delta z_l(\mathbf{x}) = v_l(\mathbf{x})\Delta t_l(\mathbf{x}) = v_l(\mathbf{x})\left(t_l(\mathbf{x}) - t_{l-1}(\mathbf{x})\right)$ The crucial notational convention is that properties associated with intervals are given the name of the subsurface below. The Δ-operator is used for the difference between a property at subsurface $l-1$ and subsurface l, i.e. associated with layer l. Note that by definition $z_0(\mathbf{x}) = t_0(\mathbf{x}) = 0$ for any $\mathbf{x}$.

2.1 Seismic travel times

The seismic travel times, $t_l(\mathbf{x})$ are usually collected on a regular grid sometimes as fine-meshed as 12.5m$\times$25m, so the lateral resolution appears very fine. However, the measured travel times are an averaged sum of reflections from a considerable area depending on the depth to the reflector. Thus the travel time map, $t_l(\mathbf{x})$, should be regarded as a smoothed indirect measurement of the depth to the reflector. There

are also physical limits to the vertical resolution of the seismic sound signals due to strong attenuation of short wavelengths. These two measuring errors are modeled as a space dependent random field, $R_l^t(\mathbf{x})$, The 'true' unsmoothed travel times to reflector l is therefore:

$$T_l(\mathbf{x}) = t_l(\mathbf{x}) + R_l^t(\mathbf{x}); \qquad \mathbf{x} \in \mathbf{R}^2. \tag{2}$$

2.2 Interval velocities

Within a homogeneous layer l the interval velocity usually varies laterally due to different average rock density and mineralogy. The lateral trend is a linear model:

$$V_l(\mathbf{x}) = \sum_{p=1}^{P_l} A_l^p g_l^p(\mathbf{x}) + R_l^v(\mathbf{x}); \qquad \mathbf{x} \in \mathbf{R}^2, \tag{3}$$

where A_l^p are partly known coefficient parameters, $g_l^p(\mathbf{x})$ are known space dependent regression functions, and $R_l^v(\mathbf{x})$ is modeled as a Gaussian random field with zero expectation. The regression functions are typically interval velocities from stacking velocities or functions of interpreted travel times. The part of Equation (3) excluding the residual is called the interval velocity trend, and is denoted by the symbol $\tilde{V}_l(\mathbf{x})$.

In the Bayesian approach the coefficient parameters are regarded as multi-Gaussian distributed random variables. The trained geophysicist should assign prior probability distributions to these variables, that is essentially, expectations and variances.

2.3 Depth to subsurfaces

By definition of interval velocity, the thickness of a layer l outlined by two seismic reflectors is:

$$\Delta Z_l(\mathbf{x}) = V_l(\mathbf{x})\Delta T_l(\mathbf{x}) = \left(\tilde{V}_l(\mathbf{x}) + R_l^v(\mathbf{x})\right)\left(\Delta t_l(\mathbf{x}) + \Delta R_l^t(\mathbf{x})\right). \tag{4}$$

The residuals are generally small compared to the trends so a reasonable simplification is to ignore products of residuals:

$$\Delta Z_l(\mathbf{x}) \approx \left(\tilde{V}_l(\mathbf{x}) + R_l^v(\mathbf{x})\right)\Delta t_l(\mathbf{x}) + \tilde{V}_l(\mathbf{x})\Delta R_l^t(\mathbf{x}).$$

Reorganizing by using $\Delta R_l^t(\mathbf{x}) = R_l^t(\mathbf{x}) - R_{l-1}^t(\mathbf{x})$, $R_0^t(\mathbf{x}) = 0$, and $\Delta\tilde{V}_{l+1}(\mathbf{x}) = (\tilde{V}_l(\mathbf{x}) - \tilde{V}_{l+1}(\mathbf{x}))$ gives the depth to subsurface L:

$$\begin{aligned} Z_L(\mathbf{x}) &= \sum_{l=1}^{L} \Delta Z_l(\mathbf{x}) \\ &= \sum_{l=1}^{L} \left(\tilde{V}_l(\mathbf{x}) + R_l^v(\mathbf{x})\right)\Delta t_l(\mathbf{x}) + \sum_{l=1}^{L-1} \Delta\tilde{V}_{l+1}(\mathbf{x})R_l^t(\mathbf{x}) + \tilde{V}_L(\mathbf{x})R_L^t(\mathbf{x}). \end{aligned}$$

The velocity changes, $\Delta\tilde{V}_{l+1}(\mathbf{x})$, are usually small compared to $\tilde{V}_L(\mathbf{x})$. Noting that these velocity changes are multiplied by residuals and using the fact that the sum of two residuals are dominated by the largest, suggests that simplifying by ignoring the 'velocity-change' terms has minor implications. This simplification reduce the stochastic model for depth to subsurface L to:

$$Z_L(\mathbf{x}) = \sum_{l=1}^{L} \left(\tilde{V}_l(\mathbf{x}) + R_l^v(\mathbf{x})\right) \Delta t_l(\mathbf{x}) + R_L^z(\mathbf{x}). \tag{5}$$

The time residual is replaced by the depth residual according to $R_L^z(\mathbf{x}) = \tilde{V}_L(\mathbf{x}) R_L^t(\mathbf{x})$. The depth residual is modeled as a Gaussian random field with expectation zero.

Two important aspects should be recognized from Equation (5): 1) The uncertainty in all interval velocities above subsurface L contributes to the uncertainty in the depth to subsurface L. 2) The uncertainty in the travel times to a subsurface above subsurface L does not contribute to the uncertainty in the depth to subsurface L. This is a consequence of removing the 'velocity-change' terms.

A full specification of the stochastic model for the depth to the subsurfaces requires a full specification of the residual Gaussian random fields and the prior multi-Gaussian distribution for the coefficient parameters. Independence among different residuals and the coefficient parameters is assumed:

$$\mathrm{Cov}\{R_l^z(\mathbf{x}), R_{l'}^z(\mathbf{x}')\} = 0 \qquad \text{for } l \neq l', \tag{6a}$$

$$\mathrm{Cov}\{R_l^v(\mathbf{x}), R_{l'}^v(\mathbf{x}')\} = 0 \qquad \text{for } l \neq l', \tag{6b}$$

$$\mathrm{Cov}\{R_l^z(\mathbf{x}), R_{l'}^v(\mathbf{x}')\} = 0 \qquad \text{for any } l \text{ and } l', \tag{6c}$$

$$\mathrm{Cov}\{A_l^p, R(\mathbf{x})\} = 0 \qquad \text{for any residual.} \tag{6d}$$

The expectations of the residuals are assumed zero everywhere. The variance and correlation function must be specified for every residual:

$$\mathrm{Var}\{R_l^{z/v}(\mathbf{x})\} = \left[\sigma_l^{z/v}(\mathbf{x})\right]^2 \qquad \text{for } l = 1, \ldots, L \tag{7a}$$

$$\mathrm{Corr}\{R_l^{z/v}(\mathbf{x}), R_l^{z/v}(\mathbf{x}')\} = \rho_l^{z/v}(\mathbf{x}, \mathbf{x}') \qquad \text{for } l = 1, \ldots, L. \tag{7b}$$

Finally consider a vector $\mathbf{A}$ containing all $P = \sum_{l=1}^{L} P_l$ coefficient parameters, A_l^p. The prior distribution for this multi-Gaussian vector must also be specified in terms of P expectations, $E\{\mathbf{A}\} = \boldsymbol{\mu}_0$, and a $P \times P$-dimensional covariance matrix, $\mathrm{Var}\{\mathbf{A}\} = \Sigma_0$.

2.4 Unification of depth and velocity models

Standard linear methods are used for predictions. The multi-layer model must therefore be formulated as a linear regression model with a spatial dependent Gaussian

residual:

$$Z(\mathbf{x}) = \sum_{p=1}^{P} A_p f_p(\mathbf{x}) + R(\mathbf{x}) = \mathbf{f}(\mathbf{x}) \cdot \mathbf{A} + R(\mathbf{x}). \tag{8}$$

All P coefficient parameters are organized in the column vector $\mathbf{A}$ and the corresponding regression functions are organized in the P-dimensional row vector $\mathbf{f}(\mathbf{x})$. First observe that Equation (5) — the model for depth to subsurfaces — has this form: the residual is $R(\mathbf{x}) = \sum_{l=1}^{L} R_l^v(\mathbf{x})\Delta t_l(\mathbf{x}) + R_L^z(\mathbf{x})$, and the sum of regression functions is: $\sum_{p=1}^{P} A_p(\mathbf{x}) f_p(\mathbf{x}) = \sum_{l=1}^{L} \sum_{p=1}^{P_l} A_l^p g_l^p(\mathbf{x}) \Delta t_l(\mathbf{x})$, where $P = \sum_{l=1}^{L} P_l$ is the total number of coefficient parameters. The interval velocity model, Equation (3), has exactly the form of Equation (8). The objective however, is to consider depth and velocity simultaneously so a common regression model must be used. The idea is to multiply Equation (3) by $\Delta t_l(\mathbf{x})$ to obtain a thickness having length units. Note that this thickness is not the true layer thickness since the depth residuals are ignored. Once the velocities are 'transformed' into thicknesses, a common regression model for both depths and velocities is possible. Although this is the key to simultaneous prediction of all subsurfaces and interval velocities, it is mainly a matter of notation. Details are therefore given in Appendix A.

3 Bayesian kriging

Two different kinds of data are considered: depth observations from every subsurface and velocity observations from every interval. The number of observations from subsurface and interval l is denoted N_l^z and N_l^v respectively. The lateral locations of the observations are in principal arbitrary but the position to observations from vertical wells obviously coincide. The observations are considered exact.

Travel times, $t_l(\mathbf{x})$, and regression functions, $g_l^p(\mathbf{x})$, are assumed known for every $\mathbf{x} \in \mathbf{R}^2$.

3.1 Posterior distribution for coefficient parameters

A posterior multi-Gaussian distribution for the coefficient parameters in the velocity models is assessed from a prior distribution and the available depth and velocity observations.

Consider a column vector, $\mathbf{Z}$, of all the $N = \sum_{l=1}^{L} (N_l^z + N_l^v)$ depth and velocity observations (rescaled by Δt), and a vector $\mathbf{R}$ of the corresponding N unobserved residuals: $\mathbf{Z}^T = [Z_1(\mathbf{x}_1), \ldots, Z_N(\mathbf{x}_N)]$, and $\mathbf{R}^T = [R_1(\mathbf{x}_1), \ldots, R_N(\mathbf{x}_N)]$, where '$T$' is used for transposed. In this notation, Equation (8) becomes:

$$\mathbf{Z} = F\mathbf{A} + \mathbf{R}. \tag{9}$$

The $N \times P$-dimensional 'design-matrix' F is constructed from the known space-dependent regression functions. Every row corresponds to an observation and every column corresponds to a particular regression function $f_p(\mathbf{x})$.

The prior $N \times N$-dimensional covariance matrix for all the observations are: $K_z = \text{Var}\{\mathbf{Z}\} = F\Sigma_0 F^T + K$, where $F\Sigma_0 F^T$ is the prior contribution from the trend, and K is the kriging matrix: $K = \text{Var}\{\mathbf{R}\}$. The Bayesian prediction and prediction variance for the coefficient parameters are the expectations and the covariances of the posterior multi-Gaussian distribution given by:

$$\hat{\boldsymbol{\mu}}_b = E\{\mathbf{A}|\mathbf{Z}\} = \boldsymbol{\mu}_0 + \Sigma_0 F^T K_z^{-1}(\mathbf{Z} - F\boldsymbol{\mu}_0) \tag{10}$$

$$\hat{\Sigma}_b = \text{Var}\{\mathbf{A}|\mathbf{Z}\} = \Sigma_0 - \Sigma_0 F^T K_z^{-1} F\Sigma_0. \tag{11}$$

This is a standard result from linear regression analysis found in most textbooks on Bayesian statistics such as Berger [2].

3.2 Kriging

Define the prior variance of $Z(\mathbf{x})$ at an arbitrary location, say $\mathbf{x}$, and the prior covariances between $Z(\mathbf{x})$ and the observation vector, $\mathbf{Z}$:

$$k_z(\mathbf{x}) = \text{Var}\{Z(\mathbf{x})\} = \mathbf{f}(\mathbf{x})\Sigma_0\mathbf{f}^T(\mathbf{x}) + k(\mathbf{x}) \tag{12}$$

$$\mathbf{k}_z(\mathbf{x}) = \text{Cov}\{Z(\mathbf{x}), \mathbf{Z}\} = F\Sigma_0\mathbf{f}^T(\mathbf{x}) + \mathbf{k}(\mathbf{x}), \tag{13}$$

where $k(\mathbf{x}) = \text{Var}\{R(\mathbf{x})\}$ and $\mathbf{k}(\mathbf{x}) = \text{Cov}\{R(\mathbf{x}), \mathbf{R}\}$. Using these definitions the Bayesian kriging predictor and the corresponding prediction variance are:

$$Z_b^*(\mathbf{x}) = E\{Z(\mathbf{x})|\mathbf{Z}\} = \mathbf{f}(\mathbf{x}) \cdot \boldsymbol{\mu}_0 + \mathbf{k}_z(\mathbf{x})K_z^{-1}(\mathbf{Z} - F\boldsymbol{\mu}_0) \tag{14}$$

$$\sigma_b^2(\mathbf{x}) = \text{Var}\{Z(\mathbf{x})|\mathbf{Z}\} = k_z(\mathbf{x}) - \mathbf{k}_z(\mathbf{x})K_z^{-1}\mathbf{k}_z^T(\mathbf{x}). \tag{15}$$

This result is found in Omre and Halvorsen [3] and in Omre, Halvorsen, and Berteig [4].

4 Example and discussion

Sections 2.2 and 2.3 describes stochastic models for interval velocities and depth to subsurfaces. Section 2.4 shows how these models can be regarded as a single linear regression model with a correlated residual which is handled by standard estimation and prediction techniques as well as Bayesian prediction techniques. This 'unification' of the models determines correlations between separate subsurfaces and between subsurfaces and interval velocities. Consequences of these dependencies on predictions obtained by kriging will be illustrated.

4.1 Data

Data is taken from a Norwegian off-shore petroleum reservoir. The data are slightly manipulated to maintain confidentiality. Two subsurfaces — Top and Base — are considered.

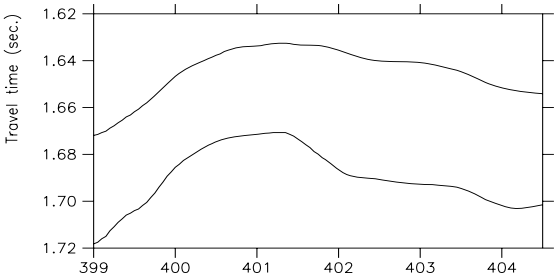


Figure 1: Recorded travel times, $t_i(\mathbf{x})$, to Top and Base.

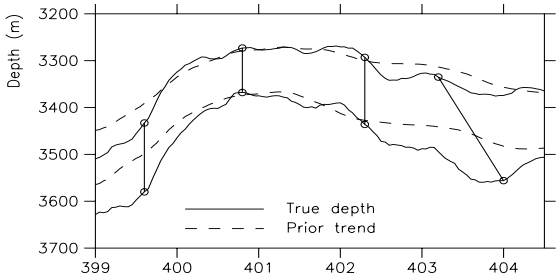
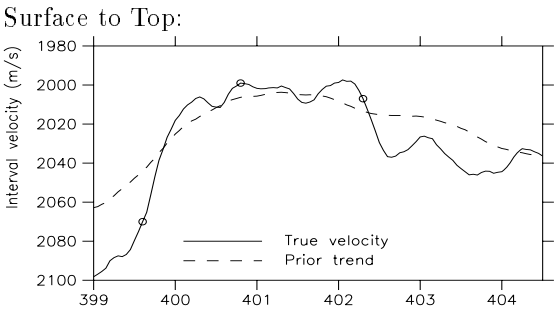


Figure 2: Possible true depth to Top and Base. Dashed lines are prior guesses. Depth observations are shown as small circles.



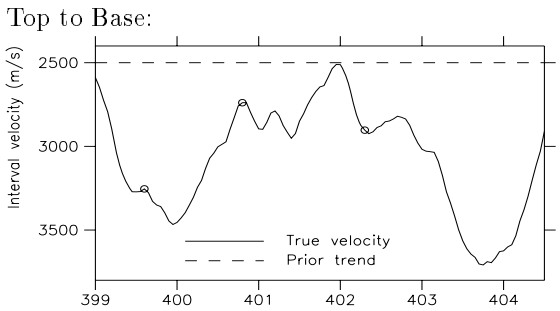


Figure 3: Possible true interval velocities. Dashed lines are prior guesses. Velocity observations are shown as small circles.

Figure 1 shows a 5.5 kilometer long cross-section of the travel time maps in the west-east direction. Figure 2 shows a cross section of the two subsurfaces. Four wells — three vertical and one deviating — are indicated by small circles connected by solid lines. To obtain depths from travel times, interval velocities must be specified. Figure 3 shows cross-sections of the possible true velocity fields and the six velocity observations from the vertical wells. Available observations are listed in Table 1.

4.2 Stochastic models

The stochastic model for Top is:

$$Z_{\text{Top}}(\mathbf{x}) = V_{\text{Top}}(\mathbf{x}) t_{\text{Top}}(\mathbf{x}) + R^z_{\text{Top}}(\mathbf{x}) \tag{16}$$

$$V_{\text{Top}}(\mathbf{x}) = A^1_{\text{Top}} + A^2_{\text{Top}} \left(t_{\text{Top}}(\mathbf{x}) - 1.63\text{s}\right) + R^v_{\text{Top}}(\mathbf{x}). \tag{17}$$

Note that $\min_{\mathbf{x}}\{t_{\text{Top}}(\mathbf{x})\} \approx 1.63$s. The two coefficient parameters have simple interpretations: A^1_{Top} is the average interval velocity to the crest of the geological structure and A^2_{Top} determines the increase (or decrease) in interval velocity at the flanks. So A^2_{Top} controls the curvature of the geological structure.

The stochastic model for Base is:

$$Z_{\text{Base}}(\mathbf{x}) = V_{\text{Top}}(\mathbf{x}) t_{\text{Top}}(\mathbf{x}) + V_{\text{Base}}(\mathbf{x}) \left(t_{\text{Base}}(\mathbf{x}) - t_{\text{Top}}(\mathbf{x})\right) + R^z_{\text{Base}}(\mathbf{x}) \tag{18}$$

$$V_{\text{Base}}(\mathbf{x}) = A^2_{\text{Base}} + R^v_{\text{Base}}(\mathbf{x}). \tag{19}$$

Well	Subsurface/ interval	X-coor. (km)	Depth (m)	Velocity (m/s)
1	Top	400.8	3273	1999
	Base		3368	2739
2	Top	402.3	3293	2007
	Base		3436	2903
3	Top	399.6	3433	2070
	Base		3580	3254
4	Top	403.2	3335	
	Base	404.0	3556	

Table 1: Available observations. Note that interval velocity observations are only available in vertical wells.

Residual	σ	a (m)	ρ
$R^z_{\text{Top}}(\mathbf{x})$	4m	1800	Spherical
$R^v_{\text{Top}}(\mathbf{x})$	12m/s	600	Gaussian
$R^z_{\text{Base}}(\mathbf{x})$	8m	2000	Spherical
$R^v_{\text{Base}}(\mathbf{x})$	300m/s	600	Gaussian

Table 2: Specifications for the Gaussian random fields. The covariance functions are: $\text{Cov}\{R(\mathbf{x}), R(\mathbf{x}')\} = \sigma^2\rho(|\mathbf{x}-\mathbf{x}'|; a)$.

The stochastic models for Top and Base include four residuals. They are modeled as Gaussian random fields with zero expectation. Table 2 gives the specifications used in the computations. The spherical correlation function is defined by $\rho(r;a) = 1 - \frac{3}{2}\frac{r}{a} - \frac{1}{2}\left(\frac{r}{a}\right)^3$ for $r < a$ and 0 else. The Gaussian (2. order exponential) correlation function is defined by $\rho(r;a) = \exp(-3(r/a)^2)$.

A full specification of the stochastic model requires prior guesses on the coefficient parameters. The expectations and standard deviations are found in Table 3.

General figure caption. Cross-sections of predicted depths and velocities are shown in Figures 4 to 8. Predictions, $Z^*(\mathbf{x})$, are shown as solid lines and prediction variances, $\sigma^2(\mathbf{x})$, are shown as dotted lines of $Z^*(\mathbf{x}) \pm \sigma(\mathbf{x})$. Well observations are shown as small circles. Observations from the same well is connected by a solid line. The horizontal axis are in the west-east direction and spans approximately 5.5 kilometers. The vertical scale are meters and meters per second for depth and velocity predictions respectively.

4.3 "Traditional" versus proposed method

The traditional approach to predicting the depth to subsurfaces in a layered media is to start by predicting the uppermost subsurface. Prediction of the second subsurface is done by adding the predicted thickness of the intermediate layer to the uppermost subsurface. Prediction variances are obtained for the first subsurface and the intermediate layer. The problem however, is to evaluate the prediction variance for the second subsurface. Simple addition of the two prediction variances would give the correct answer if the first subsurface and the thickness of the intermediate layer where uncorrelated. This is not so: the thickness of the intermediate layer between

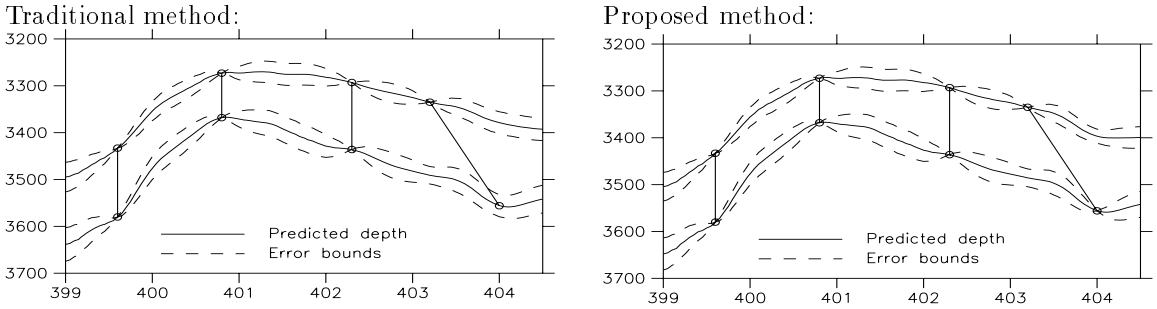


Figure 4: Predicted depth to Top and Base. Notice the failure of the traditional approach at the deviating well.

Table 3: Expectation μ, standard deviation σ, and correlations between the coefficient parameters for 0, 1, 2, 3, and 4 wells.

	Bayes estimates										GLS estimates (universal kriging)					
# of wells	0 (prior)		1		2		3		4		2		3		4	
# of obs.	0		2+2		4+4		6+4		8+6		4+4		6+4		8+6	
Parameter	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
A^1_{Top} m/s	2000	50	1993	12	1993	10	1997	9	1996	9	1991	25	1978	14	1987	11
A^2_{Top} m/s^2	1500	750	1493	749	1499	731	2269	625	2402	469	1701	3609	4050	1169	2992	610
A^1_{Base} m/s	2500	500	2675	257	2772	195	2864	186	2963	157	2821	212	2826	208	3000	166
Corr$\{A^1_{\text{T}}, A^2_{\text{T}}\}$	0		−0.25		−0.49		−0.70		−0.77		−0.94		−0.88		−0.84	
Corr$\{A^1_{\text{T}}, A^1_{\text{B}}\}$	0		0		0		0.01		−0.02		0		0.20		0.01	
Corr$\{A^2_{\text{T}}, A^1_{\text{B}}\}$	0		0		0		−0.16		−0.06		0		−0.30		−0.09	

Top and Base is:

$$\begin{aligned}\Delta Z_{\text{Base}}(\mathbf{x}) &= Z_{\text{Base}}(\mathbf{x}) - Z_{\text{Top}}(\mathbf{x}) \\ &= V_{\text{Base}}(\mathbf{x})\,(t_{\text{Base}}(\mathbf{x}) - t_{\text{Top}}(\mathbf{x})) + R^z_{\text{Base}}(\mathbf{x}) - R^z_{\text{Top}}(\mathbf{x}),\end{aligned}$$

which is (anti-)correlated to $Z_{\text{Top}}(\mathbf{x})$ due to the common residual $R^z_{\text{Top}}(\mathbf{x})$. Ignoring negative correlations causes to large prediction variances. Figure 4 compares the traditional approach to an approach where all correlations are considered. Both are produced by Bayesian kriging. The results are almost identical except from at the right flank. The prediction variance for Base fails to be zero at the well observation. The very strong intercorrelations between $\Delta Z_{\text{Base}}(\mathbf{x})$ and $Z_{\text{Top}}(\mathbf{x})$ in the vicinity of the well is ignored in the traditional approach leading to wrong results.

A different, but less dramatic effect, is that the rightmost observation of Base should have an impact on the prediction of Top. This is obviously not true for the traditional approach since Top is predicted independently of Base. The proposed method however, gives a deformation of the prediction for Top directly above the observation. Notice also a small local reduction in the prediction variance.

4.4 Bayesian kriging

Bayesian kriging is a necessity to obtain reasonable results in some applications; the lack of well observations in off-shore applications make universal kriging unreliable.

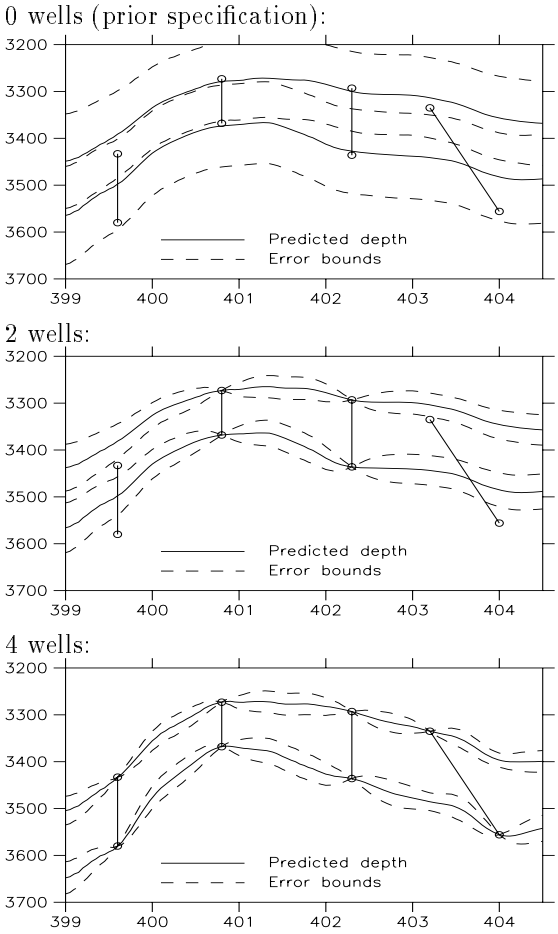


Figure 5: Predicted depth to Top and Base by Bayesian kriging.

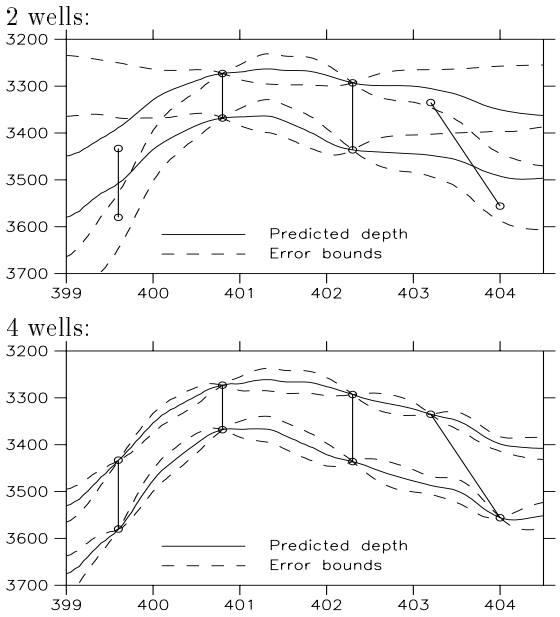


Figure 6: Predicted depth to Top and Base by universal kriging.

Figures 5 and 6 show predictions by Bayesian and universal kriging respectively. The predictions are conditioned on different numbers of wells.

The difference between Bayesian and universal kriging is the method for obtaining estimates for the coefficient parameters. Both methods use simple kriging for local adaptions in the vicinity of observations. Table 3 contains Bayesian and GLS estimates for the coefficient parameters. The estimation of A^1_{Top} and A^1_{Base} are very successful even when a single well is used. The reason is that the specified variance for the residuals are quite small compared to the prior variances for the coefficient parameters. The acceleration parameter, A^2_{Top}, controlling the curvature, is not successfully estimated before any of the deep observations in well 3 or 4 are included.

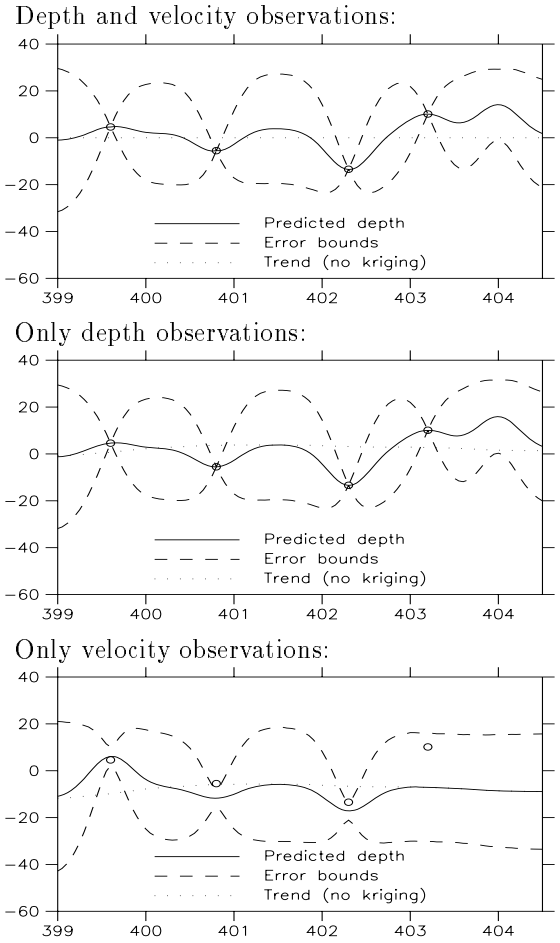


Figure 7: Bayesian depth predictions for Top. The trend for predictions with both depth and velocity data has been subtracted to exaggerate the local influence of the data.

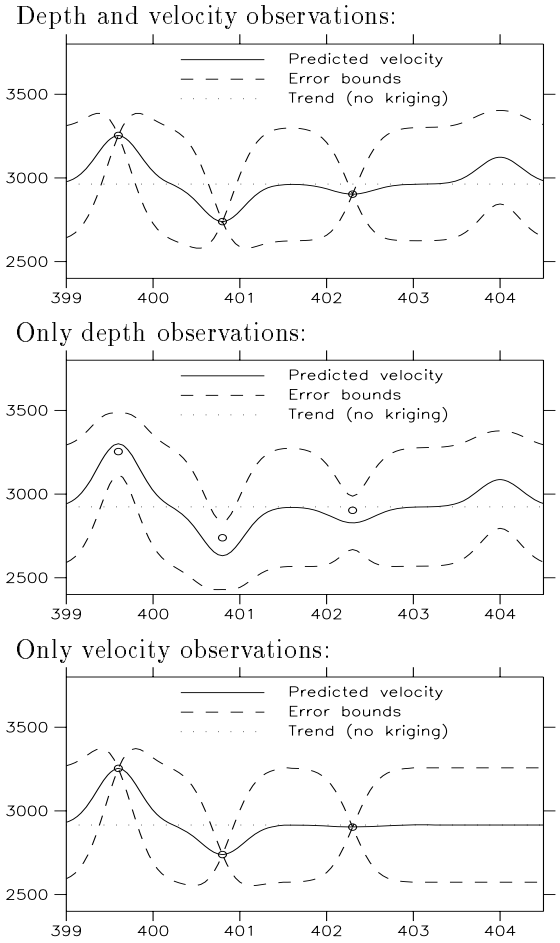


Figure 8: Bayesian velocity predictions for interval between Top and Base.

4.5 Use of velocity information

The impact of velocity observations on depth predictions are quite small due to strong intercorrelations. To visualize this Figure 7 shows depth predictions for Top conditioned on 1) all observations 2) depth observations alone 3) velocity observations alone. To exaggerate the variations the predicted trend obtained by using all data are subtracted from all the predictions. The two upper figures are almost identical indicating minor influence from velocity observations. Notice especially the 'hump' on the right side due to the depth observation of Base below. The reduction in uncertainty at this location is seen quite clearly. The prediction in the lower figure do not interpolate the depth observation since it is conditioned on velocity observations alone.

Figure 8 shows the predicted interval velocity in the layer between Top and Base.

Table 4: Generalized least squares estimates for expectation μ, standard deviation σ, and correlations of the coefficient parameters for all observations, depth observations, and velocity observations.

GLS estimates (universal kriging)						
Observations:	Depth+velocity		Depth		Velocity	
# of observations	8+6		8		6	
Parameter	μ	(σ)	μ	(σ)	μ	(σ)
A^1_{Top} (m/s)	1987	(11)	1990	(11)	1984	(11)
A^2_{Top} (m/s^2)	2992	(610)	2880	(621)	2915	(638)
A^1_{Base} (m/s)	3000	(166)	2975	(204)	2965	(173)
Corr$\{A^1_{\text{Top}}, A^2_{\text{Top}}\}$	-0.84		-0.84		-0.80	
Corr$\{A^1_{\text{Top}}, A^1_{\text{Base}}\}$	0.01		-0.01		0	
Corr$\{A^2_{\text{Top}}, A^1_{\text{Base}}\}$	-0.09		-0.10		0	

The influence from the depth observation of Base in the deviating well is seen as a hump and a small reduction in prediction variance.

Table 4 shows the corresponding estimated coefficient parameters. The more sensitive GLS estimates are shown. The general observation is that adding velocity data has minor implications since depth and velocity data are highly correlated.

Note however that A^1_{Base} is determined more precisely using the 3 relevant velocity observation than the 4 relevant depth observations. The reason is that the 6 velocity observations splits into two sets of uncorrelated observations while all the depth observations are strongly correlated.

5 Final remarks

A new method for depth conversion of seismic travel times to L reflecting subsurfaces has been presented. The method has the following characteristics:

- it is a combination of cokriging with $2L$ covariables and L dependent linear regression models for velocity and depth trends,
- predictions are possible by Bayesian or universal kriging,
- the predicted depths and velocities are conditioned on all observations from the N subsurfaces and N velocity fields,
- depth predictions interpolate the relevant depth observations, and velocity predictions interpolate the relevant velocity observations,
- prediction variances for all predictions are available,
- using deviating wells are trivial,
- estimates of coefficient parameters in the velocity models are based on all observations.

Conditional simulation based on the same model is described in Abrahamsen and Omre [1]. Simulation is necessary for the assessment of uncertainty in volumetric predictions.

References

[1] P. Abrahamsen and H. Omre. Horizon 4.0 — geometrical module, user guide and theory. NR-note SAND/01/1992, Norwegian Computing Center, P.O.Box 114 Blindern, N-0314 Oslo, Norway, 1992.

[2] J. O. Berger. *Statistical Decision Theory and Bayesian Analysis.* Springer series in statistics. Springer-Verlag Inc., New York, 2. edition, 1985.

[3] H. Omre and K. B. Halvorsen. The bayesian bridge between simple and universal kriging. *Math. Geol.*, 21(7):767–785, 1989.

[4] H. Omre, K. B. Halvorsen, and V. Berteig. A bayesian approach to kriging. In M. Armstrong, editor, *Geostatistics*, volume 1, pages 109–126. Kluwer Academic Publishers, 1989.

A Unified notation for depth to subsurfaces and interval velocities

We intend to show that the regression models:

$$V_l(\mathbf{x}) = \sum_{p=1}^{P_l} A_l^p g_l^p(\mathbf{x}) + R_l^v(\mathbf{x}) \tag{20}$$

$$Z_L(\mathbf{x}) = \sum_{l=1}^{L} \left(\tilde{V}_l(\mathbf{x}) + R_l^v(\mathbf{x})\right) \Delta t_l(\mathbf{x}) + R_L^z(\mathbf{x}) \tag{21}$$

can be written as a single regression model:

$$Z(\mathbf{x}) = \mathbf{f}(\mathbf{x}) \cdot \mathbf{A} + R(\mathbf{x}).$$

The exercise is mainly a change in notation.

It will prove convenient to use a new equivalent set of regression functions:

$$f_l^p(\mathbf{x}) = \Delta t_l(\mathbf{x}) g_l^p(\mathbf{x}).$$

Multiplying Equation (20) by $\Delta t_l(\mathbf{x})$ and introducing new regression functions gives:

$$V_l(\mathbf{x})\Delta t_l(\mathbf{x}) = \sum_{p=1}^{P_l} A_l^p f_l^p(\mathbf{x}) + R_l^v(\mathbf{x})\Delta t_l(\mathbf{x}) \tag{22}$$

$$Z_L(\mathbf{x}) = \sum_{l=1}^{L} V_l(\mathbf{x})\Delta t_l(\mathbf{x}) + R_L^z(\mathbf{x}). \tag{23}$$

Note that Equation (22) is equivalent to Equation (20) since $\Delta t_l(\mathbf{x})$ is considered a known function.

All coefficient parameters will be organized in a $P = \sum_{l=1}^{L} P_l$-dimensional column vector starting with the P_1 parameters for the first interval and so on:

$$\begin{aligned}\mathbf{A}^T &= (A_1, \ldots, A_P) \\ &= \left([A_1^1, \ldots, A_1^{P_1}], \ldots, [A_L^1, \ldots, A_L^{P_L}]\right).\end{aligned}$$

The square brackets simply emphasize the initial grouping of parameters. The space dependent regression functions $f_l^p(\mathbf{x})$ are organized accordingly in a P-dimensional row vector such that the trends can be written

$$\begin{aligned}\tilde{V}_l(\mathbf{x})\Delta t_l(\mathbf{x}) &= \mathbf{f}_l^{[v]}(\mathbf{x}) \cdot \mathbf{A} \\ \tilde{Z}_l(\mathbf{x}) &= \mathbf{f}_l^{[z]}(\mathbf{x}) \cdot \mathbf{A}.\end{aligned}$$

The crucial point is to replace all $f_l^p(\mathbf{x})$'s irrelevant for a certain surface by zeroes:

$$\begin{aligned}\mathbf{f}_l^{[v]}(\mathbf{x}) &= \left(0, \ldots, 0, [f_l^1(\mathbf{x}), \ldots, f_l^{P_l}(\mathbf{x})], 0, \ldots, 0\right) \\ \mathbf{f}_l^{[z]}(\mathbf{x}) &= \left([f_1^1(\mathbf{x}), \ldots, f_1^{P_1}(\mathbf{x})], \ldots, [f_l^1(\mathbf{x}), \ldots, f_l^{P_l}(\mathbf{x})], 0, \ldots, 0\right).\end{aligned}$$

Introducing generic forms:

$$Z(\mathbf{x}) = \begin{cases} Z_L(\mathbf{x}); & \text{for depth to subsurface } L \\ V_L(\mathbf{x})\Delta t_L(\mathbf{x}); & \text{for velocities in } L-1 \text{ to } L \end{cases} \tag{24}$$

$$\mathbf{f}(\mathbf{x}) = \begin{cases} \mathbf{f}_L^{[v]}(\mathbf{x}); & \text{for } Z(\mathbf{x}) = Z_L(\mathbf{x}) \\ \mathbf{f}_L^{[z]}(\mathbf{x}); & \text{for } Z(\mathbf{x}) = V_L(\mathbf{x})\Delta t_l(\mathbf{x}) \end{cases} \tag{25}$$

$$R(\mathbf{x}) = \begin{cases} R_L^z(\mathbf{x}) + \sum_{l=1}^{L} R_L^v(\mathbf{x})\Delta t_L(\mathbf{x}); & \text{for } Z(\mathbf{x}) = Z_L(\mathbf{x}) \\ R_L^v(\mathbf{x})\Delta t_L(\mathbf{x}); & \text{for } Z(\mathbf{x}) = V_l(\mathbf{x})\Delta t_l(\mathbf{x}) \end{cases} \tag{26}$$

Using these gives the single regression model

$$Z(\mathbf{x}) = \mathbf{f}(\mathbf{x}) \cdot \mathbf{A} + R(\mathbf{x}), \tag{27}$$

for both depths to subsurfaces and velocities. The parameter vector, $\mathbf{A}$, is independent of the interpretation of Z, $\mathbf{f}$, and R.

Equation (27) describes $V_l(\mathbf{x})\Delta t_l(\mathbf{x})$ rather than $V_l(\mathbf{x})$ and the same applies to the Bayesian kriging equation. This implies that the velocity observations must be multiplied by $\Delta t_l(\mathbf{x})$ when used in the kriging equation. The result — the predicted surface — must therefore be divided by $\Delta t_l(\mathbf{x})$ to obtain the interval velocity field.

For the N observations of the surfaces Equation (27) can be written

$$\mathbf{Z} = F\mathbf{A} + \mathbf{R}. \tag{28}$$

where $\mathbf{Z}$ is the N observations organized as a column vector and $\mathbf{R}$ is the corresponding unknown residuals. The $N \times P$-dimensional matrix F contains the corresponding regression functions. Every row of F is the P-dimensional $\mathbf{f}(\mathbf{x})$ vector for the corresponding observation.