

Abstract

We conducted a designed experiment to quantify sources of uncertainty in the expert interpretation of a geological cross-section. A group of 28 geologists participated in the experiment. Each interpreted borehole records which included three Palaeogene bedrock units, including the target unit for the experiment: the London Clay. The set of boreholes was divided into batches from which validation boreholes had been withheld; as a result we obtained 129 point comparisons between the interpreted elevation of the base of the London Clay and its observed elevation in a borehole not used for that particular interpretation. Analysis of the results showed good general agreement between the observed and interpreted elevations, with no evidence of systematic bias. Between-site variation of the interpretation error was spatially correlated, and the variance appeared to be stationary. The between-geologist component of variance was smaller overall, and depended on distance to the nearest borehole. There was also evidence that the between-geologist variance depends on the degree of experience of the individual. We used the statistical model of interpretation error to compute confidence intervals for any one interpretation of the base of the London Clay on the cross-section, and to provide uncertainty measures for decision support in a hypothetical route-planning process. The statistical model could also be used to quantify error-propagation in a full 3-D geological model produced from interpreted cross sections.

1 Introduction

Three dimensional (3-D) models are now the state of the art for presenting geologists' knowledge and interpretation of subsurface structure, and are supplied to varied users of geological information. Models can be produced by geostatistical interpolation (e.g., Lark and Webster, 2006) or by a combination of geostatistical methods with expert intervention to ensure geologically realistic results (e.g., Gunnink et al., 2013). How-

SED

6, 1687–1721, 2014

Uncertainty in cross section interpretations

R. M. Lark et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



ever, geostatistical interpolation requires relatively dense sets of observations of the subsurface, and is a largely mechanical process. For this reason there is interest in the development of methodologies for modelling which are able to make the fullest use of the geologist's experience and knowledge of the processes that underly the subsurface structure of interest.

Of particular interest here is an approach based on expert interpretation of boreholes along interlocking sets of cross-sections with subsequent interpolation from the interpreted sections to produce models of volumes in 3-D. This is exemplified by the GSI3D software (Kessler and Mathers, 2004; Kessler et al., 2009). Expert interpretation of a cross-section entails the interpretation of boreholes and the sequential construction of the basal contact of each geological unit in the stack. This process depends on the expert interpretation of boreholes in line with rules, explicit or tacit, which control the shapes of surfaces and the circumstances in which faults must be invoked to explain their observed positions. Because these rules encapsulate geological knowledge, they provide a sound basis for modelling particularly when limited observations are available. However, the interpretation of the cross sections inevitably has an attendant uncertainty, and this is propagated when the interpreted cross sections are combined to model volumes in 3-D by interpolation.

The uncertainty in a 3-D model is of interest to data users who will apply it for decision making. If information in 3-D is produced by geostatistical methods then the uncertainty can be quantified directly on the basis of the geostatistical model, but this is not the case with models produced from interpreted cross-sections.

Lark et al. (2013) studied empirically the uncertainty in a 3-D model, produced by the cross-section interpretation methodology. They compared the predicted heights of units with observed heights at a set of validation boreholes. This gave a quantitative measure of uncertainty. However, Lark et al. (2013) concluded that it is necessary to understand how error enters into the initial interpretation of cross-sections prior to interpolation, since the error in the cross-section may be predictable from factors such as distance to boreholes or crop line, but is propagated into the 3-D model by the

Uncertainty in cross section interpretations

R. M. Lark et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



interpolation step in a complex way. If we can understand and quantify the uncertainty in cross-section interpretation then it may be possible to develop quantitative models of the uncertainty for different “benchmark” geological settings, and use this to develop uncertainty measures for application to geological models.

To this end we undertook, and report here, an experiment to study the error in cross section interpretation, hypothesizing that the variability of the interpretation error varies along the section in ways that can be described by a statistical model. We consider, in particular, statistical models in which the distance to nearest borehole and the geologist’s experience are predictors of the variance of model error.

If our hypothesis is verified, then we could compute confidence intervals for the interpreted height of a contact along a cross-section, and model how this uncertainty may propagate in the subsequent interpolation from the interpreted cross-section into a 3-D geological model. If statistical models of the uncertainty in cross-section interpretation could be estimated for a variety of geological settings, then these could be used to compute uncertainty measures for new geological models, and so to calculate, for example, decision-theoretic measures of the value of the model information (Howard, 1966) or other criteria by which model-users can make rational decisions that account for model uncertainty.

2 Methods

2.1 The cross-section

This study is based on a 8-km section in London which roughly follows the A12 road from Hackney north-east to Wanstead. The local stratigraphy consists of Quaternary alluvium and river terrace deposits resting on four bedrock units the London Clay, the Lambeth Group, the Thanet Sand (all Palaeogene) and the Chalk (Cretaceous), the latter appears in approximately 10 % of the 143 available boreholes. In this study our unit of interest was the London Clay, an Eocene unit deposited in marine conditions in

Uncertainty in cross section interpretations

R. M. Lark et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Uncertainty in cross section interpretations

R. M. Lark et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



a succession of transgressive-regressive sequences which give rise to some textural variation across the formation. Most of the London Clay is silty clay to very silty clay, with some sandy clay and pockets of sand (Ellison et al., 2004). The London Clay is present over about 60% of the length of the section and the base of the unit, the surface of interest, is proven by 51 of the boreholes. Over the length of the section the elevation of the base of the London Clay varies from nearly 10 m to –13 m relative to Ordnance Datum.

2.2 Data subsetting, geologists' self-assessment and modelling

The key idea of the experiment was that each of a set of participating geologists would make an interpretation of the three Palaeogene bedrock units on the cross-section, drawing continuous (if occasionally interrupted) basal contacts of the units as interpretations of the information in a set of boreholes. Any one participant would use a subset of all available boreholes, so that their interpretation could be compared directly with each of a complementary validation subset. The difference between the interpreted and observed elevation of the base of the London Clay, the cross-section error, would then be treated as a variable for statistical analysis to identify important features of its variability. Note that, while we only examined the base of the London Clay, the participants interpreted this in the wider stratigraphical context by also drawing the bases of the other Palaeogene units.

The 51 available boreholes which prove the base of the London Clay were subdivided by independent random sampling without replacement into ten batches, each of five boreholes, with one remaining. One of these validation batches could be withdrawn from the total set to leave a particular subset of 46 boreholes, available for interpretation. In this way ten different although overlapping interpretation subsets, each of 46 boreholes, were prepared for use by geologists in the experiment. Any one participant would use just one interpretation subset. His or her interpretation of the cross-section could then be compared with the five boreholes in the corresponding validation subset, boreholes not used in the interpretation of the cross section.

Uncertainty in cross section interpretations

R. M. Lark et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



A total of 28 geologists participated in the experiment. Of these, 22 were delegates at the GSI3D workshop which took place at the British Geological Survey (BGS), Keyworth, from 17–18 October 2012 and the GSI3D software was used for the experiment. Some of the workshop participants were staff of BGS, others were geologists from a variety of organizations and countries, with varying levels of experience in geological modelling but all with some interest and experience, if rudimentary, in the use of the GSI3D software. The remaining 6 geologists were BGS staff who participated in the experiment after the workshop.

The key principle of the experiment was explained to all delegates, who were also provided with an explanation of the units in the cross section. Each participant in the experiment, on presenting at the workstations, was given a unique number, and an interpretation subset of boreholes. In addition to the boreholes a standard interpretation of the superficial material (as a single unit) was provided so that all participants were working to a common rockhead surface. The intersections of outcrops, as mapped in 2-D, with the cross-section were also provided to all participants. A set of guidance notes on the GSI3D software was available, and at all times a staff member experienced with the software was available to help. When the interpretation was complete it was saved with a code which indicated the participant's unique number and the number of the interpretation subset and complementary validation batch of boreholes which had been allocated. Batches were allocated to participants in order as they presented so that a more or less even distribution of participants over batches was achieved.

Each participant was asked to complete a questionnaire before undertaking the exercise. Their unique number was recorded on the form. They had the option of recording their name and contact details on the form, or of remaining anonymous. On the questionnaire each participant was asked to record a self-assessment of their experience of geological modelling in 3-D by identifying the most appropriate of four general descriptions. The descriptions and responses are presented in Table 1.

Once each geologist had completed and saved their interpretation, this was compared with the corresponding batch of validation boreholes, and the observed and

interpreted elevation of the base of the London Clay was extracted. One modeller's interpretation was not correctly saved, so this was lost, and in some cases the London Clay was not present in the interpretation at the location of a validation borehole. We therefore had a total of 129 comparisons of interpreted and observed elevations available for analysis with between 10 and 20 observations in any batch.

3 Data analysis

3.1 Linear mixed models

To analyse the results from this experiment we fitted and compared linear mixed models (LMM) (Verbeke and Molenberghs, 2000) for the cross-section errors. One observation of cross-section error corresponds to a particular geologist's interpretation at one of the validation sites withheld from the set of boreholes which that geologist had been allocated. This gives us a total of $N = 129$ observations of cross-section error. If the interpreted elevation of the base of the London Clay by geologist m at site k within batch i is $z^s(b_i, s_k, g_m)$ and the corresponding observed elevation in the validation borehole is $z^o(b_i, s_k)$ then the corresponding observation of cross-section error is defined as

$$\varepsilon(b_i, s_k, g_m) = z^o(b_i, s_k) - z^s(b_i, s_k, g_m). \quad (1)$$

A negative error therefore means that the geologist's interpretation is higher than the observed elevation of the base of the London Clay.

In a LMM a variable is treated as a linear combination of fixed effects (known or controlled factors) and random effects. The random effects represent sources of variation in the observations, and here account for differences between batches of validation boreholes, between the sites of the validation boreholes within batches, and between the geologists. We considered different models of these random effects, modelling their variance as functions of potential sources of uncertainty in the cross-section interpre-

SED

6, 1687–1721, 2014

Uncertainty in cross section interpretations

R. M. Lark et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



tation. In particular we considered the distance to the nearest borehole available for interpretation, and the geologist's self-identified experience.

In a LMM the random effects are modelled as Gaussian random variables with mean zero and a variance. The variance may be a constant, a parameter of the LMM, or it may be expressed as a parametric function of some covariate with parameters to be estimated (e.g., Nelder and Lee, 1998; Lark, 2009). The random effects and their parameters are of interest because they may be informative about sources of cross-section error, and allow us to predict cross-section error variance in similar settings. Once an appropriate model for the random effects has been selected, one may use generalized least squares to estimate the overall mean model error and test whether it appears to be significantly different from zero.

The models that we considered differ with respect to the random effects. The first model (*1a*) is the most general model for the data, and we examine two associated models (*1b* and *1c*) to see whether the general model can be simplified. These models include terms to account for spatial variations in model error (between sites), and variations among geologists. We describe the model-fitting procedure in the context of model *1a*, and the procedure for comparing models in the context of the evaluation of models *1b* and *1c*. We then examined a further variant, model *1d*, in which the variance of the between-site component of model error at any location is assumed to depend on the distance from that location to the nearest borehole available to the geologist when the corresponding model was made.

We then considered a further set of models in which the batch and site effects are taken from a model selected on statistical criteria from among *1a–1d*. These models differ with respect to the way the between-geologist component of variation is treated. In the first of these, model *2a*, it was assumed that the variance of the between-geologist effect at a location depends on on the distance from that location to the nearest borehole available to the geologist when the corresponding model was made. In the second, model *2b*, it was assumed that the the variance of the between-geologist effect at a location depends on the geologist's experience of 3-D geological modelling. In the third,

SED

6, 1687–1721, 2014

Uncertainty in cross section interpretations

R. M. Lark et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



model 2c, it was assumed that both distance to the nearest observation and geologist's experience determine the between-geologist component of variation in cross-section error.

3.2 The basic model (1a) and three variants

5 Model 1a takes the following form for a set of observations of cross-section error in a vector $\boldsymbol{\varepsilon}$ of length N :

$$\boldsymbol{\varepsilon} = \mathbf{M}\boldsymbol{\alpha} + \mathbf{X}_b\boldsymbol{\beta}_b + \mathbf{X}_s\mathbf{Z}_s + \boldsymbol{\eta}_g, \quad (2)$$

10 where \mathbf{M} is a $N \times p$ design matrix that associates each observation in $\boldsymbol{\varepsilon}$ with a value of a fixed effect variable, contained in the vector $\boldsymbol{\alpha}$ of length p . In all models considered in this paper the fixed effect is a constant, the mean cross-section error, so $p = 1$, $\boldsymbol{\alpha}$ contains the mean and \mathbf{M} is a $N \times 1$ vector of ones. Other terms in the model are explained in the following paragraphs.

15 The matrix \mathbf{X}_b is a $N \times N_b$ design matrix for the between-batch random effect where N_b is the number of batches. Row n of \mathbf{X}_b corresponds to the n th observation. If the n th observation belongs to the m th batch out of N_b then the element in column m of row n of \mathbf{X}_b is one and all other elements in the row are zero. The vector $\boldsymbol{\beta}_b$ is a $N_b \times 1$ vector which contains the mean errors for the batches, which are treated as random variables. One may write down an expression for the covariance matrix of the
20 N between-batch components of the observation, \mathbf{C}_b . Because the sites are randomly allocated to batches it is assumed that the batch effects are independent and so:

$$\begin{aligned} \mathbf{C}_b &= \sigma_b^2 \mathbf{R}_b \\ &= \sigma_b^2 \mathbf{X}_b \mathbf{X}_b^T, \end{aligned} \quad (3)$$

25 where σ_b^2 is the variance of the batch effect, \mathbf{R}_b denotes the correlation matrix of batch effects which is obtained, given the assumptions of independence, as the product of the batch design matrix with its transpose (denoted by the superscript T).

Uncertainty in cross section interpretations

R. M. Lark et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Uncertainty in cross section interpretations

R. M. Lark et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



The term \mathbf{X}_s is a $N \times N_s$ design matrix which associates each of the N observations to one of the N_s validation sites. These sites are not assumed to be independent of each other, since they were chosen by purposive sampling, and not by an independent random sampling design. Since the sampling design does not allow us to treat the between-site effect as an independent random variable, we rather invoke a random statistical model of the between-site effect (de Gruijter et al., 2006). The random variable, which is contained in the length- N_s random vector \mathbf{Z}_s is assumed to be N_s -variate Gaussian with mean zero and $N_s \times N_s$ covariance matrix \mathbf{S} :

$$\mathbf{Z}_s \sim \mathcal{N}(\mathbf{0}_{N_s}, \mathbf{S}), \quad (4)$$

where $\mathbf{0}_{N_s}$ is a vector length N_s of zeroes. We assume that \mathbf{Z}_s is a second-order stationary random variable so that the covariance of the values at any two sites depends only on the interval in space between those sites (Stein, 1999). Here we use a standard covariance function from geostatistics, the Matérn function (Matérn, 1986). Under this model the covariance between two locations separated by distance d is

$$C(d) = c_0 + c_1, \quad d = 0$$

$$= c_1 \left\{ 2^{\kappa-1} \Gamma(\kappa) \right\}^{-1} \left(\frac{d}{\phi} \right)^\kappa K_\kappa \left(\frac{d}{\phi} \right), \quad d > 0, \quad (5)$$

where $K_\kappa(\cdot)$ is a modified Bessel function of order κ , κ is a smoothness parameter – see Diggle and Ribeiro (2007) for a discussion, ϕ is a distance parameter and c_0 and c_1 are, respectively, the spatially uncorrelated and correlated components of variance of the between-site variable. Note that, while in principle, the covariance can be modelled as a function of the direction as well as the length of the separation vector between locations, when our observations are aligned on an almost-straight cross-section we consider distance only.

If the distance between site k in batch i and site l in batch j is denoted by $d_{\{i,k\},\{j,l\}}$ then one may compute a between-site covariance matrix \mathbf{S} , which is an $N_s \times N_s$ matrix.

If the r th out of N_s sites is site k in batch i , and the c th out of N_s sites is site l in batch j then

$$\mathbf{S}[r, c] = C(d_{\{i,k\},\{j,l\}}),$$

- 5 and the $N \times N$ between-sites effect covariance matrix for the LMM for all N observations is given by

$$\mathbf{C}_s = \mathbf{X}_s \mathbf{S} \mathbf{X}_s^T, \quad (6)$$

10 where \mathbf{X}_s is the $N \times N_s$ design matrix for sites. Given the site design matrix, and the distances among the observations, this covariance matrix is determined by the four parameters of the Matérn covariance function: c_0 , c_1 , κ and ϕ .

The geologist effect in model 1a, the term η_g in Eq. (2), is somewhat more complex. At each site within a batch a cross-section error is observed for each geologist who was allocated the corresponding batch of boreholes. The term η_g is the difference
 15 between the cross-section error for a particular geologist at a particular site, and the mean cross-section error at that site. It is therefore the between-geologist within-site effect, but we call it the geologist effect for brevity.

If each geologist had one and only one validation borehole then the geologist effect would be simply nested within sites as an independent random error (whether there
 20 was one or more observations of model error at each validation site). However, in the current experiment, each geologist appears at each site within the batch to which he or she was allocated. In model 1a we treat the geologist effects as correlated random variables within batches. If we denote by $\bar{\varepsilon}(b_i, s_k)$ the mean cross-section error at site k in batch i , the geologist effect for geologist m at the same site is

$$\eta(b_i, s_k, g_m) = \varepsilon(b_i, s_k, g_m) - \bar{\varepsilon}(b_i, s_k). \quad (7)$$

$$\eta(b_i, s_k, g_m) = \varepsilon(b_i, s_k, g_m) - \bar{\varepsilon}(b_i, s_k).$$

Uncertainty in cross section interpretations

R. M. Lark et al.

Title Page	
Abstract	Introduction
Conclusions	References
Tables	Figures
◀	▶
◀	▶
Back	Close
Full Screen / Esc	
Printer-friendly Version	
Interactive Discussion	



In the random effects component of model 1a, we assume that the correlation of the geologist effects is

$$\begin{aligned} \text{Corr}\{\eta(\mathbf{b}_i, \mathbf{s}_k, \mathbf{g}_m), \eta(\mathbf{b}_j, \mathbf{s}_l, \mathbf{g}_n)\} &= 1, \quad i = j, k = l, m = n \\ &= \rho, \quad i = j, k \neq l, m = n \\ &= 0, \quad \text{otherwise.} \end{aligned} \quad (8)$$

In words, the geologist effects for observations at two different sites are uncorrelated if the geologists are different (which includes all between-batch comparisons), and have a correlation of ρ if the geologist is the same. The covariance matrix for the geologist effect in model 1a is therefore

$$\mathbf{C}_g = \sigma_g^2 \mathbf{R}_g, \quad (9)$$

where \mathbf{R}_g is a $N \times N$ correlation matrix of geologist effects with values 1 on the main diagonal, ρ on off-diagonal elements which correspond to pairs of cross-section errors corresponding to the same geologist and zero in all other elements. The variance of the between-geologist effect is σ_g^2 .

The random effects of the model in Eq. (2) are characterized by the between-batch variance, σ_b^2 , the four parameters of the Matérn covariance model for the between-site variable (c_0 , c_1 , κ and ϕ), the between-geologist within site variance σ_g^2 and the correlation parameter ρ . We used residual maximum likelihood (REML) to estimate these parameters (Patterson and Thompson, 1971; Smyth and Verbyla, 1996). This proceeds on the assumption that $\boldsymbol{\varepsilon}$ in Eq. (2) is a realization of a multivariate Gaussian random variable, \mathbf{E} :

$$\mathbf{E} \sim \mathcal{N}(\mathbf{M}\boldsymbol{\alpha}, \mathbf{V}) \quad (10)$$

where \mathbf{V} is the covariance matrix given by

$$\mathbf{V} = \mathbf{C}_b + \mathbf{C}_s + \mathbf{C}_g. \quad (11)$$

Uncertainty in cross section interpretations

R. M. Lark et al.

Title Page	
Abstract	Introduction
Conclusions	References
Tables	Figures
◀	▶
◀	▶
Back	Close
Full Screen / Esc	
Printer-friendly Version	
Interactive Discussion	



Under this model the residual log-likelihood, ignoring constants, is given by

$$\ell_R = -\frac{1}{2} \left\{ \log |\mathbf{V}| + \log \left| \mathbf{M}^T \mathbf{V}^{-1} \mathbf{M} \right| + \boldsymbol{\varepsilon}^T \mathbf{P} \boldsymbol{\varepsilon} \right\}, \quad (12)$$

where

$$\mathbf{P} = \mathbf{V}^{-1} - \mathbf{V}^{-1} \mathbf{M} (\mathbf{M}^T \mathbf{V}^{-1} \mathbf{M})^{-1} \mathbf{M}^T \mathbf{V}^{-1}.$$

The Gaussian assumption can not be tested strictly because it is an assumption about a multivariate distribution of which we have a single realization. However, its plausibility can be tested by examining a histogram and summary statistics of the residuals of an ordinary least squares fit of the fixed effects model, equivalent to the statistics of the data in this case where a uniform mean is the only fixed effect. Where necessary data may be transformed to a new scale of measurement to make the assumption more plausible.

We used the `optim` procedure in the R package (R development core team, 2013) to find REML estimates of the random effects parameters, the values that maximize the likelihood as defined in Eq. (12). The `L-BFGS-B` optimization method was selected, a quasi-Newton optimizer in which upper and lower bounds are supplied for the parameters to be estimated (Byrd et al., 1995).

In the proposed model there are $P = 7$ random effects parameters (or variance parameters) to be estimated by REML. One may consider the “null hypothesis” that one of these parameters can be set at a fixed value, to simplify the model. For example, if one assumed that the cross-section errors for the same geologist at two sites within a batch are uncorrelated, then $\rho = 0$. In general a “null” model with $P - g$ parameters is said to be nested within a more complex “full” model with P parameters if the null model can be regarded as a particular case of the full model with the g additional parameters taking fixed values. The maximized residual likelihood for the full model $\ell_{R,F}$ is at least as large as that for the null model, $\ell_{R,N}$. To test whether the improvement of fit from the g additional parameters is large enough to justify their inclusion within the model one

Uncertainty in cross section interpretations

R. M. Lark et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



may compute the log-likelihood ratio statistic (Verbeke and Molenberghs, 2000)

$$L = 2(\ell_{R,F} - \ell_{R,N}). \quad (13)$$

In what is called the “standard case” where the g additional parameters all take definite values in the null model, and these parameter values are not on the boundary of the parameter space, then, in the case where the null model is the true model L is distributed as χ^2 with g degrees of freedom (Stram and Lee, 1994). Note that this procedure is valid for residual likelihoods only when the models have the same fixed effects structure.

One may use this procedure to compare the LMM in Eq. (2) with one in which the geologist effects are regarded as uncorrelated between sites within batches. In the full model $\rho \in [-1, 1]$, so the fixed value, $\rho = 0$, in the null model is not at a boundary. The comparison is therefore a standard case with $L \sim \chi^2(1)$ under the null hypothesis.

However, if we consider a null model in which the between-batch variance is zero this is not a standard case since zero is the lower bound for a variance. A more general criterion for comparing models of differing complexity, although not a formal test, is to compute for each model Akaike’s information criterion – AIC (Akaike, 1973):

$$A = -2\ell + 2P, \quad (14)$$

where ℓ is the maximized log likelihood (natural logarithms) and P is the number of parameters. That model is preferred for which A is smallest, so the term $2P$ is, in effect, a penalty for model complexity.

Model *1b* is a variant of *1a* in which the between-batch variance is dropped. Since the batches were formed at random one may expect that the mean error does not differ among the batches except for random sample variation. However, in a comparison between these two models the null (*1b*) is formed by fixing the between-batch variance at zero, which is a boundary in parameter space (variances cannot be negative). The models are therefore compared on the AIC.

Uncertainty in cross section interpretations

R. M. Lark et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Model *1c* is a variant of *1a* in which the correlation $\rho = 0$. As noted above, this comparison can be made by computing the log-likelihood ratio statistic L and testing it against $\chi^2(1)$.

Having selected one model from among *1a–1c*, a variant was considered in which the correlated variance of the between-site random variable, c_1 in Eq. (5), depends on the distance from that site to the nearest borehole available for interpretation (i.e. not in the validation set for the batch). We considered the possibility that this variance is a linear function of distance to the nearest borehole. The intercept and slope of this function, $\alpha_{s,0}$ and $\alpha_{s,1}$ respectively, are therefore substituted for c_1 in model *1d*. The comparison of between the null model selected from among *1a–1c* and the more complex variant *1d* can be made using the log-likelihood ratio, assumed to be distributed as $\chi^2(1)$ under the null model since model *1d* has one more parameter than the null.

3.3 Refining the model: explaining geologist variance (models *2a*, *2b* and *2c*)

Here we consider the possibility that the between-geologist variance can be replaced by a parametric function. In principle this is compatible with any variant of the models considered so far. The expression for the between-geologist covariance matrix in Eq. (9) is modified to

$$\mathbf{C}_g = \mathbf{\Sigma}_g \mathbf{R}_g \mathbf{\Sigma}_g, \quad (15)$$

where \mathbf{R}_g is defined as for Eq. (9), and

$$\mathbf{\Sigma}_g = \text{diag}(\boldsymbol{\sigma}_g) \quad (16)$$

where $\boldsymbol{\sigma}_g$ is a vector of length N which contains the standard deviation of the between-geologist effect for each observation, predicted from some parametric function. The operator “diag” denotes that the elements of this vector are put in order on the main diagonal of a $N \times N$ matrix, with off-diagonal elements equal to zero.

Three parametric functions were considered. In the first the between-geologist variance for the r th observation depends on the distance from the site which corresponds

Uncertainty in cross section interpretations

R. M. Lark et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Uncertainty in cross section interpretations

R. M. Lark et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



to the r th observed cross-section error and the nearest borehole available for interpretation to the corresponding site. Again a linear function was considered, so the parameter σ_g^2 in the first and second group of models was replaced by the intercept and slope of this predictive relationship, $\alpha_{g,0}$ and $\alpha_{g,1}$ respectively. These parameters, along with the remaining ones, were estimated by REML.

The second parametric model considered used the geologist's self-assessment of experience in 3-D geological modelling. There were four levels of experience to choose from, so the parameter σ_g^2 in the first and second group of models was replaced by four parameters, variances for each level of experience: $\sigma_{g,1}^2$, $\sigma_{g,2}^2$, $\sigma_{g,3}^2$, $\sigma_{g,4}^2$.

A final model was considered which combined the last two variants, with separate intercepts and slopes of the linear function for the geologist standard deviation being specified for each level of experience (i.e. eight new parameters replacing σ_g^2 in the first and second group of models).

Note that the parametric functions in these three models return variances, which may vary from one observation to another. The terms in σ_g are standard deviations, i.e. the square roots of the corresponding variances.

3.4 Simulating from the selected model to represent cross-section uncertainty

We used the selected model (model 2a as described in the results section below), to simulate realizations of the random component of cross-section error along a part of the cross-section (from 4000 m from the start of the section to the end). We considered a situation where all the boreholes along the cross-section were available to the geologist, we assumed that the cross-section error is zero at the location of a borehole, and simulated the components of the error under model 2a conditional on this at regularly-spaced locations along the cross-section. The between-site component was simulated as a multivariate normal random variate by Cholesky decomposition of the joint covariance matrix of the regularly-spaced sampling locations and the borehole locations. This is described in detail by Goovaerts (1997), we used the R procedure

Uncertainty in cross section interpretations

R. M. Lark et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



CHOL,(R development core team, 2013). To simulate the between-geologist component we evaluated the variance of this component at each regularly spaced location on the cross-section from the parameters of model 2a as a function of distance to the nearest borehole. A realization of the between-geologist component of model error at each location is then simulated as a normal random variable with mean zero and variance set to this computed value. We used the R procedure RNORM to do this, (R development core team, 2013). The overall cross-section error is then simulated by the sum of these two components. A total of 10 000 independent realizations of cross-section error were simulated this way.

By finding the 2.5th and 97.5th percentile of the simulated cross-section errors at any location we approximate the 95 % confidence interval for model error. This can be used to visualize the uncertainty. The simulations can also be used to answer other questions. Consider, for example, an engineer who wishes to dig a tunnel through the London Clay along the length of this part of the cross-section. We assume that the engineer wants to put the route of the tunnel as close as possible to the base of the London Clay, but wants to avoid intruding on the underlying Lambeth Group. The conditional simulations can be used to assess the risk of intruding on the Lambeth Group if the tunnel route is k m above the interpreted base of the London Clay everywhere along the route. Assume that the engineer specifies that the tunnel should enter the Lambeth Group over no more than 1 % of its length. What is the smallest value of k consistent with this? One could examine the 10 000 realizations of cross-section error and find, for increasing values of k , the number of realizations for which the engineer's specification is met. This number, out of 10 000, gives the approximate probability of meeting the specification given some k .

4 Results

4.1 Summary statistics on model error from all validation sites

Figure 1 shows a scatter plot of interpreted and observed height of the base of the London Clay for all observations. The points are scattered around the bisector (where observed and interpreted heights are equal) and there is no visual evidence of a systematic bias. Table 2 shows the summary statistics of cross-section error, and Fig. 2 shows the histogram of this variable. The symmetrical form of the histogram and the weak skewness and kurtosis values suggest that an assumption of normality is plausible for the analysis of these data. They also suggest that, if there is any systematic tendency for the base of the London Clay to be interpreted too high or too low, then this effect is small.

4.2 Model comparisons

The results for model *1a* and its variants are shown in Table 3. Note that the estimated between-batch variance is zero. When a REML estimate of a parameter is at the boundary of parameter space, as here, it is advisable to examine the likelihood profile in the vicinity of the estimate. To compute the likelihood profile for a model parameter that parameter is fixed at a series of values and, for each, the remaining parameters are estimated by maximum (residual) likelihood. The maximized likelihoods are then plotted against the values of the parameter of interest. The profile likelihood should increase smoothly towards the estimated value. The profile likelihood for the batch variance satisfied this requirement. This is not unreasonable; because the batches were formed at random we would hope that the between-batch variation is purely explicable in terms of sampling error. The comparison of models *1a* and *1b* can be done by examining the AIC, which is smaller for the latter model in which the batch effect is dropped. Model *1b* is therefore selected over *1a*. The profile likelihood for the uncorrelated between-site variance, c_0 , in these models also approached the estimated value, 0.0, smoothly.

SED

6, 1687–1721, 2014

Uncertainty in cross section interpretations

R. M. Lark et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Uncertainty in cross section interpretations

R. M. Lark et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



In model *1c* the correlation between within-site effects for particular geologists is dropped (set to zero). The maximum likelihood is slightly smaller than for model *1b* in which this parameter is estimated. However, the log-likelihood ratio statistic, L , for the comparison of (null) model *1c* with (full) model *1b* is small, the probability of obtaining a value of L this large or larger under the null model is large, and so the more complex model is rejected in favour of the null. This is also consistent with the small estimated value of this correlation, -0.09 .

In model *1d* a stationary correlated variance for the between-site effect (as in model *1c*) is replaced by two parameters for a linear function which expresses this variance as a function of distance to the nearest borehole available for interpretation. This (full) model can be compared with a (null) model (*1c*) with a stationary variance by the log-likelihood ratio test. Once again, L is too small to support a choice of the more complex model.

In summary, the consideration of model *1a* and its variants in Table 3 leads us to the selection of Model *1c* (smallest AIC in the Table) in which the batch effect and the correlation parameter ρ for geologist effects are dropped, and the between-site variation is modelled as a stationary correlated random variable.

Table 4 shows results for model *2a* and its variants. These models are based on *1c* but differ in that, rather than assuming a stationary geologist effect, the between-geologist within-site variance is modelled as a function of covariates. In model *2a* the geologist variance is modelled as a linear function of distance to the nearest borehole available to the geologist for interpretation. The zero value of the intercept, $\alpha_{g,0}$ is plausible, under the assumption implicit in our analysis that the borehole data are correct, the cross-section error should be zero at the location of a borehole. The positive value of $\alpha_{g,1}$ implies that the geologist variance increases with increasing distance from a borehole, which is also plausible. Model *1c* can be regarded as nested within *2a*, a null model with $\alpha_{g,0}$ equivalent to σ_g^2 and $\alpha_{g,1} = 0$. The models can be tested by the log-likelihood ratio statistic, Table 4 shows that the null model (*1c*) can be decisively rejected in favour of the full model *2a*.

**Uncertainty in cross
section
interpretations**R. M. Lark et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

Model *2b* is an alternative to *2a* in which the geologist variance depends on the self-identified experience of the geologist at 3-D modelling. The estimated parameters in Table 4 are plausible in that the variance is largest for geologists who identified themselves as having “no experience of modelling in 3-D” and smallest for those who identified themselves as having “substantial experience more than 2 years of modelling independently.” Once again this model could be compared with *1c* by a log-likelihood ratio test, and the null model (*1c*) can be rejected, indicating that there is significant evidence for differences in geologist variance, related to geologists experience. However, the evidence for this model is weaker than for *2a*.

In model *2c* different relationships between geologist variance and distance to nearest borehole were fitted for the four levels of geological experience. In the fitted model the intercepts were all zero, the smallest slope is for the geologists with the highest experience level. However, while the log-likelihood ratio test shows that model *2c* is significantly better than model *2b*, i.e. adding the information on distance to nearest borehole to a model with geologist experience gives a significant improvement, the comparison of model *2c* with *2a* leads to the conclusion that adding geologist experience to a model which already has the distance to nearest borehole incorporated does not give a significant improvement. On the basis of the AIC model *2a* is preferred among all those considered in this study.

Table 5 shows the estimated mean cross-section error and its standard error, under model *2a*. The Wald statistic (e.g., Dobson, 1990) is a test of the null hypothesis that the mean error is zero, and the large *p* value shows that this cannot be rejected, so the data provide no evidence for systematic bias in the interpretation. Note that the estimate of mean error is rather less than the average for all observations reported in Table 2. That is because (i) the mean reported in Table 5 is the model mean, the fitted effect in the underlying statistical model for cross-section error and (ii) the original data were not a random sample in space, and show some local clustering which is likely to bias the arithmetic average as an estimate of the spatial mean of cross-section error.

Uncertainty in cross section interpretations

R. M. Lark et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Figure 3 shows the 95 % probability interval for cross-section errors along the section, approximated by the 2.5th and 97.5th percentile of the conditionally simulated errors. The red symbols show the location of the boreholes. There are two features of the interval. First, there is a rapid narrowing near the boreholes (the interval is zero at the boreholes but this is only seen if the borehole coincides with a point where the error is sampled). This arises from the spatial correlation of the between-site component of cross-section error. The second feature is a gradual widening of the interval to a local maximum at the midpoint between successive boreholes. This is particularly apparent in the second half of the plot. This arises from the dependence of the between-modeller effect on distance to nearest borehole, showing how the constraint of the borehole on model error decays with distance. In Fig. 4 the confidence intervals are added to the interpretation of the base of the London Clay by one of the modellers.

Figure 5 shows a plot of the estimated probability that a tunnel built k m above the interpreted base of the London Clay will intrude on the underlying Lambeth Group over no more than 1 % of its length for different values of k . This shows that the engineer can be 90 % confident that this specification will be met if the route is a little less than 8 m above the interpreted base.

5 Conclusions

Both the summary statistics and the scatter plot (Fig. 1) and the estimate of the mean cross-section error from the selected model 2a (Table 5) show that the data obtained in this study provide no evidence that there is any bias in the interpretation of the base of the London Clay by the geologists in this study, i.e. the mean error is not significantly different from zero.

We established this experiment to test the hypothesis that the variability of the error of interpretations of cross-sections varies spatially. This hypothesis has been supported. First, we found that there is spatial dependence in the variability of the between-site component of cross-section error. That is to say the cross-section error at one location

is likely to be more strongly correlated with the error at a nearby location than at one further away. This is reasonable since if, for example, a surface tends to be interpreted too high above Ordnance Datum at a site, perhaps because of faulting, then it is likely that a similar error will occur at nearby sites. There was no evidence, however, that the between-site variance depends on the distance to the nearest borehole.

The between-geologist variance is rather smaller than the between-site variance (compare c_0 with σ_g^2 in model 1c). However, there was evidence that the variance of this error depends on geologist experience and also on the distance to the nearest borehole available for interpretation. The results for these two models are consistent with our hypothesis, and also make intuitive sense in that the variance of cross-section error declines with the geologist's experience, and increases with increasing distance from the borehole. However, the preferred model for the data, given a penalty on model complexity, considers only the distance to nearest borehole.

The fitted model can be used to simulate cross-section errors conditional on a distribution of boreholes. One may use this procedure to compute confidence intervals around the interpreted cross-section which quantifies uncertainty in this interpretation and shows how this changes in space. One could also use this simulation method to study the propagation of cross-section error in further processing to interpolate the surface into 2-D and so produce 3-D volumes.

The methodology presented in this paper could be deployed in a wider range of geological settings in order to generate statistical models of cross-section error for those settings. These could then be used to compute confidence intervals for new models or measures of uncertainty specific to the requirements of particular data users, such as the example for the London Clay illustrated in Fig. 5.

The experimental design used in this study allowed us to make best use of somewhat sparse boreholes by examining multiple geologist interpretations at each validation site. However, if there had been a significant correlation between within-site effects for the same geologist, then subsequent modelling of the geologist variance would have been complicated. Alternatively one might use an experimental design in which

SED

6, 1687–1721, 2014

Uncertainty in cross section interpretations

R. M. Lark et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



validation sites are nested within modellers (so each modeller has a unique subset of validation sites). This requires that there are many boreholes available, however, since each validation borehole is compared with just one interpretation. It also reduces the information that we obtain on between-modeller differences.

5 One way to get around the problem of insufficient validation observations is to generate synthetic cross-sections, perhaps conditioned on geophysical data such as interpretations from seismic lines. These synthetic cross-sections can then be notionally sampled at as many locations as we want to provide synthetic borehole data for interpretation and validation. In such an experiment the synthetic validation boreholes
10 should be sampled according to an optimized design (e.g., Lark, 2002) design to ensure good estimation of the spatial variance parameters and to give good coverage of possible covariates: e.g. spanning a range of distances to the nearest borehole available for interpretation.

15 *Acknowledgements.* We are grateful to our colleague Luz Ramos Cabrera for her contribution to setting up the trial, and to those delegates to the 2012 GSI3D workshop, and other colleagues at the British Geological Survey who participated. This paper is published with the permission of the executive director of the British Geological Survey (NERC).

References

- 20 Akaike, H.: Information theory and an extension of the maximum likelihood principle, in: Second International Symposium on Information Theory, edited by: Petov, B. N. and Csaki, F., Akademia Kiado, Budapest, 267–281, 1973. 1700
- Byrd, R. H., Lu, P., Nocedal, J., and Zhu, C.: A limited memory algorithm for bound constrained optimization, SIAM J. Sci. Comput., 16, 1190–1208, 1995. 1699
- Diggle, P. J. and Ribeiro, P. J.: Model-Based Geostatistics, Springer, New York, 2007. 1696
- 25 Dobson, A. J.: An Introduction to Generalized Linear Models, Chapman & Hall, London, 1990. 1706
- de Gruijter, J. J., Brus, D. J., Biekens, M. F. P., and Knotters, M.: Sampling for Natural Resource Monitoring, Springer, Berlin, 2006. 1696

Uncertainty in cross section interpretations

R. M. Lark et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Uncertainty in cross section interpretations

R. M. Lark et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Ellison, R. A., Woods, M. A., Allen, D. J., Forster, A., Pharaoh, T. C., and King, C.: Geology of London. Memoir of the British Geological Survey, Sheets 256 (North London), 257 (Romford), 270 (South London) and 271 (Dartford) (England and Wales), British Geological Survey, Keyworth, 2004. 1691

5 Gunnink, J. L., Maljers, D., van Gessel, S. F., Menkovic, A., and Hummelman, H. J.: Digital Geological Model (DGM): a 3-D raster model of the subsurface of the Netherlands, *Netherlands Journal of Geosciences – Geologie en Mijnbouw*, 93, 33–46, 2013. 1688

Goovaerts, P.: *Geostatistics for Natural Resources Evaluation*, OUP, New York, 1997. 1702

Howard, R. A.: Information value theory, *IEEE T. Syst. Sci. Cyb.*, 2, 22–26, 1966. 1690

10 Kessler, H. and Mathers, S. J.: Maps to models – finally capturing the geologists' vision, *Geoscientist*, 14, 4–6, 2004 1689

Kessler, H., Mathers, S. J., and Sobisch, H.-G.: The capture and dissemination of integrated 3-D geospatial knowledge at the British Geological Survey using GSI3D software and methodology, *Comput. Geosci.*, 35, 1311–1321, 2009. 1689

15 Lark, R. M.: Optimized spatial sampling of soil for estimation of the variogram by maximum likelihood, *Geoderma*, 105, 49–80, 2002. 1709

Lark, R. M.: Kriging a soil variable with a simple non-stationary variance model, *J. Agr. Biol. Envir. St.*, 14, 301–321, 2009. 1694

20 Lark, R. M. and Webster, R.: Geostatistical mapping of geomorphic variables in the presence of trend, *Earth Surf. Proc. Land.*, 31, 862–874, 2006. 1688

Lark, R. M., Mathers, S. J., Thorpe, S., Arkley, S. L. B., Morgan, D. J., and Lawrence, D. J. D.: A statistical assessment of the uncertainty in a 3-D geological framework model, *P. Geologist. Assoc.*, 124, 946–958, 2013. 1689

25 Matérn, B.: *Spatial Variation*, Lecture Notes in Statistics, No. 36, Springer, New York, 1986. 1696

Nelder, J. A. and Lee, Y.: Joint modelling of mean and dispersion, *Technometrics*, 40, 168–171, 1998. 1694

Patterson, H. D. and Thompson, R.: Recovery of inter-block information when block sizes are unequal, *Biometrika*, 58, 545–554, 1971. 1698

30 R Development Core Team: R: a language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0, <http://www.R-project.org/>, 2013. 1699, 1703

Uncertainty in cross section interpretations

R. M. Lark et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Table 1. Questionnaire on modelling experience and responses received.

Description	Number of participants selecting this description
I have no experience of geological modelling in 3-D	2
I have some experience of geological modelling in 3-D (perhaps through a training course) but little (up to 6 months) or no experience of modelling independently	8
I have moderate experience of geological modelling in 3-D (six months to 2 years of modelling independently)	8
I have substantial experience of geological modelling in 3-D (more than 2 years of modelling independently)	10

SED

6, 1687–1721, 2014

Uncertainty in cross section interpretations

R. M. Lark et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Table 2. Summary statistics of cross-section error.

Mean	0.70
Median	0.38
SD	2.90
Min	-6.67
Max	7.44
Skewness	0.28
Kurtosis	-0.14

Uncertainty in cross section interpretations

R. M. Lark et al.

Title Page	
Abstract	Introduction
Conclusions	References
Tables	Figures
⏪	⏩
◀	▶
Back	Close
Full Screen / Esc	
Printer-friendly Version	
Interactive Discussion	



Table 3. Model 1 and variants.

Model	Random effects parameters					Geologist		ℓ_R	AIC	Contrast*	L	ρ	
	Batch	Site		κ	ϕ	σ_g^2	ρ						
	σ_b^2	c_0	c_1										
1a	0.0	0.0	6.84	2.5	4.36	1.45	-0.093	-148.89	311.78				
1b	**↓	0.0	6.84	2.5	4.36	1.45	-0.093	-148.89	309.78				
1c	↓	0.0	6.86	2.5	4.38	1.45	↓	-149.75	309.50	1c vs. 1b	1.72	0.19	
1d	↓	0.0	$\alpha_{s,0}$ 6.03	$\alpha_{s,1}$ 0.01	2.5	4.38	1.45	↓	-149.48	310.96	1c vs. 1d	0.54	0.46

* The first-named model is the null model.

** A ↓ indicates that a term has been dropped from the model.

Uncertainty in cross section interpretations

R. M. Lark et al.

Table 4. Model 2 and variants.

Model	Site				Random effects parameters Geologist				ℓ_R	AIC	Contrast	L	ρ
	c_0	c_1	κ	ϕ	$\alpha_{g,0}$	$\alpha_{g,1}$	$\sigma_{g,1}^2$	$\sigma_{g,2}^2$					
2a	0.0	6.63	2.5	4.73	0.0	0.0217			-117.18	246.36	1c vs. 2a	65.1	$< 10^{-15}$
2b	0.0	7.53	2.5	4.59	4.44	2.25	1.32	0.46	-144.16	304.31	1c vs. 2b	11.2	0.01
2c*	0.0	6.72	2.5	4.58	0.022	0.021	0.030	0.018	-116.63	257.26	2a vs. 2c 2b vs. 2c	1.1 55.1	0.98 $< 10^{-10}$

* In this model a separate slope and intercept to compute the between-geologist variance as a function of distance to nearest borehole was computed for each level of experience. All estimates of the intercept were zero exactly, so the slopes only are reported here.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



SED

6, 1687–1721, 2014

Uncertainty in cross section interpretations

R. M. Lark et al.

Table 5. Estimate of the mean cross-section error conditional on Model 2a.

Estimated mean	0.52
Standard error	0.42
Wald statistic*	1.56
p value	0.21

* The Wald statistic tests the null hypothesis that the true mean error is zero. The p value is the probability of obtaining a Wald statistic this large or larger under the null hypothesis.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



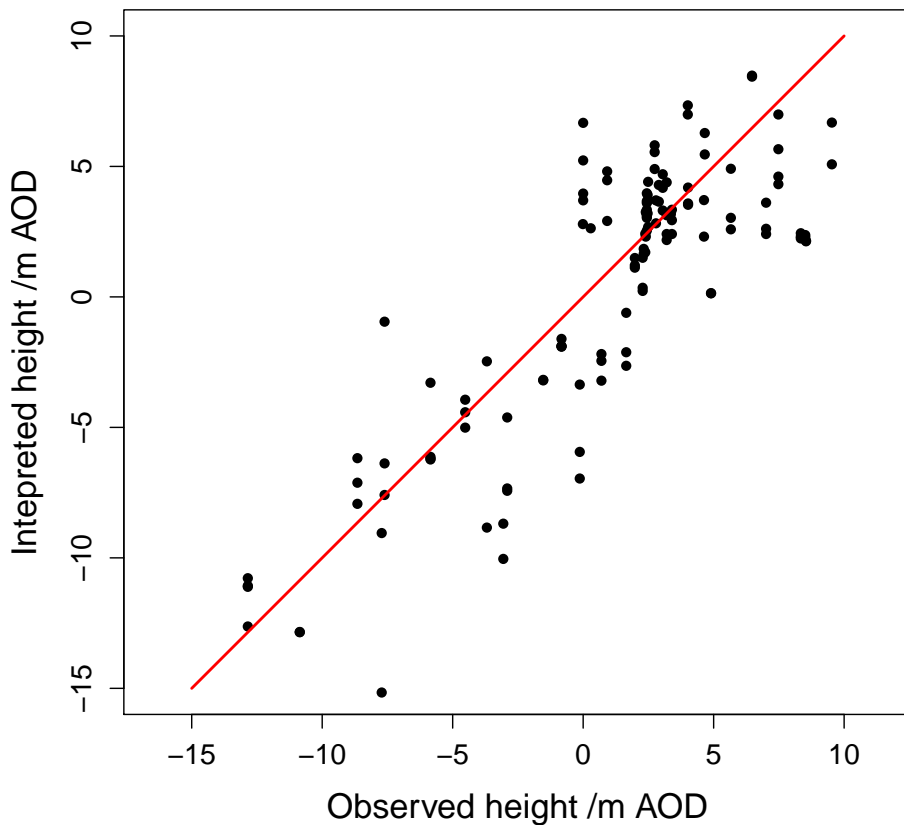


Figure 1. All validation observations of the interpreted and observed height of the base of the London Clay AOD. The red line is the bisector.

Uncertainty in cross section interpretations

R. M. Lark et al.

Title Page	
Abstract	Introduction
Conclusions	References
Tables	Figures
◀	▶
◀	▶
Back	Close
Full Screen / Esc	
Printer-friendly Version	
Interactive Discussion	



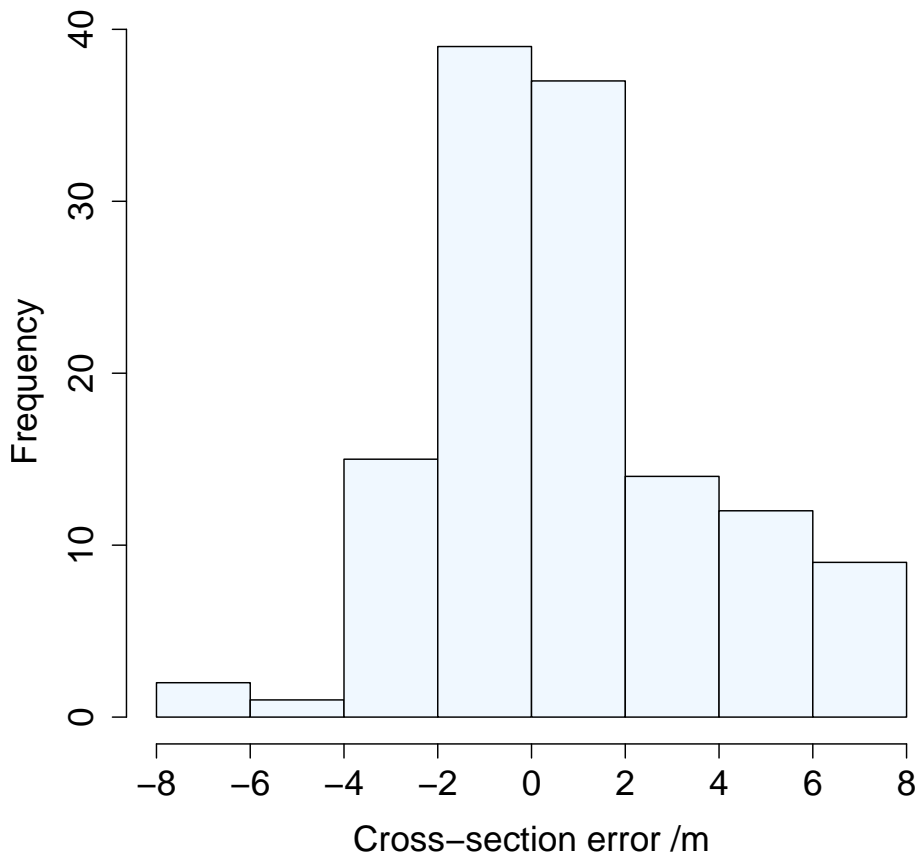


Figure 2. Histogram of cross-section errors.

SED

6, 1687–1721, 2014

Uncertainty in cross section interpretations

R. M. Lark et al.

Title Page	
Abstract	Introduction
Conclusions	References
Tables	Figures
◀	▶
◀	▶
Back	Close
Full Screen / Esc	
Printer-friendly Version	
Interactive Discussion	



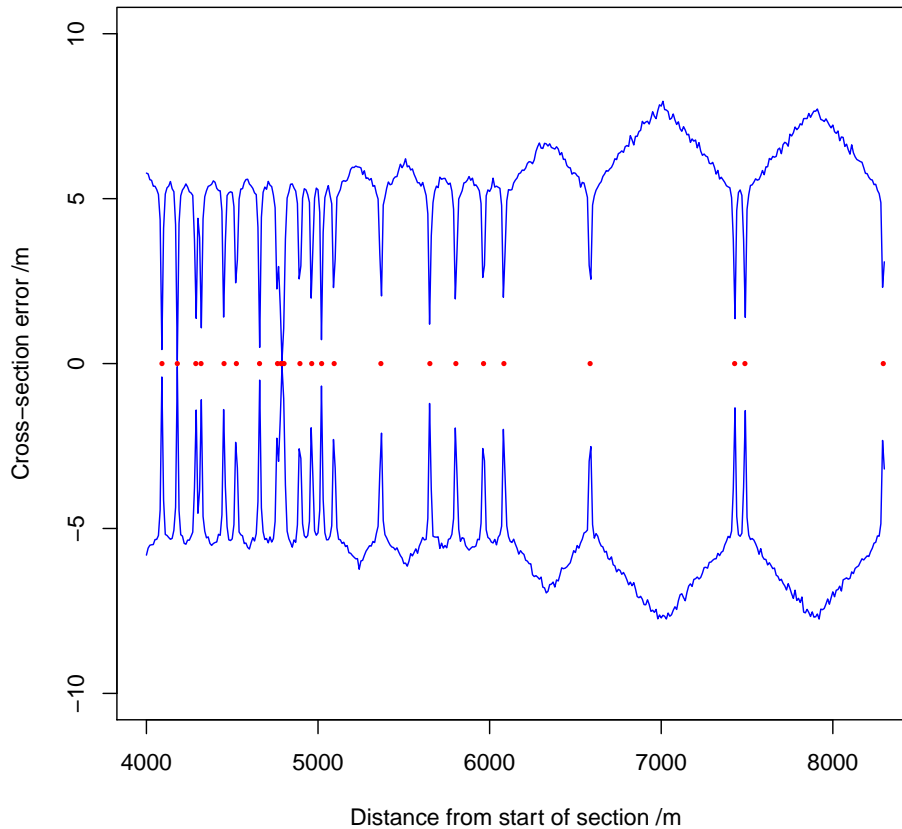


Figure 3. 95% probability interval for simulated cross-section errors conditional on location of nearest borehole (red symbol) and model 2a. Note that these are evaluated at discrete locations.

Uncertainty in cross section interpretations

R. M. Lark et al.

Title Page	
Abstract	Introduction
Conclusions	References
Tables	Figures
◀	▶
◀	▶
Back	Close
Full Screen / Esc	
Printer-friendly Version	
Interactive Discussion	



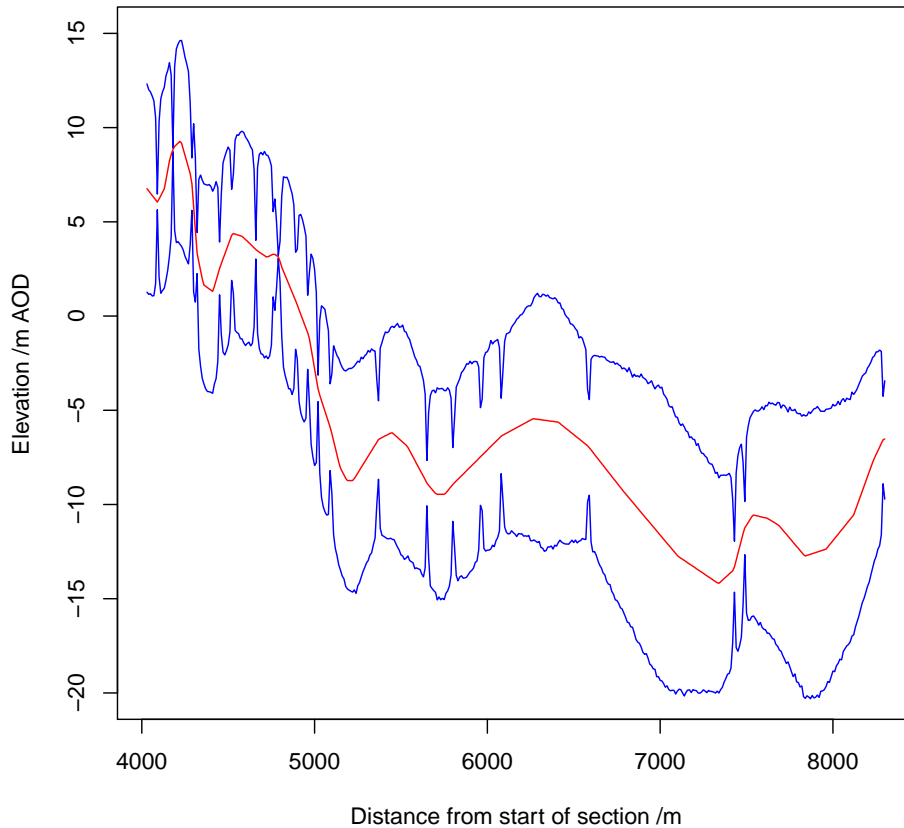


Figure 4. One geologist's interpretation of the base of the London Clay (red) with 95 % confidence intervals (blue).

Uncertainty in cross section interpretations

R. M. Lark et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



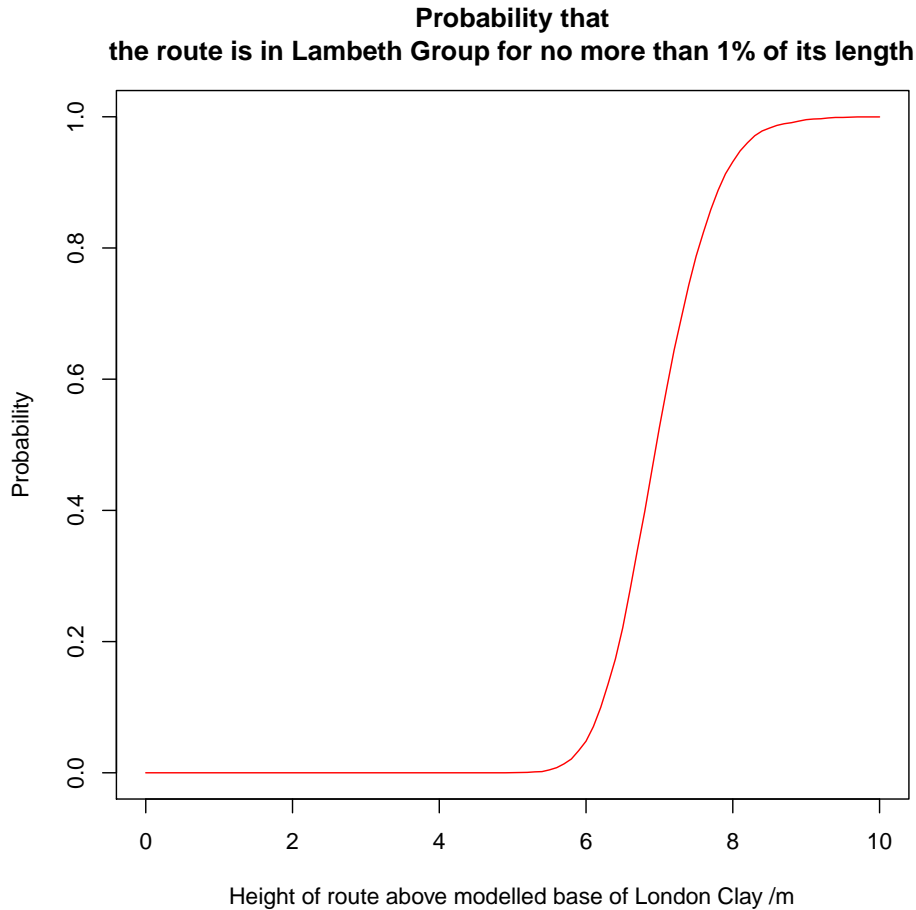


Figure 5. How close to the modelled base of the London Clay could you build a tunnel (over the last 4 km of the cross section) and have a specified probability (ordinate) that the tunnel strays into the underlying Lambeth Group for no more than 1% of its length?

Uncertainty in cross section interpretations

R. M. Lark et al.

Title Page	
Abstract	Introduction
Conclusions	References
Tables	Figures
◀	▶
◀	▶
Back	Close
Full Screen / Esc	
Printer-friendly Version	
Interactive Discussion	

