

GEOSPHERE, v. 14, no. 3

doi:10.1130/GES01510.1

6 figures; 8 tables; 1 supplemental file

CORRESPONDENCE: clare.bond@abdn.ac.uk

CITATION: Randle, C.H., Bond, C.E., Lark, R.M., and Monaghan, A.A., 2018, Can uncertainty in geological cross-section interpretations be quantified and predicted?: *Geosphere*, v. 14, no. 3, doi:10.1130/GES01510.1.

Science Editor: Shanaka de Silva  
Associate Editor: Jose Miguel Hurtado

Received 8 February 2017  
Revision received 9 July 2017  
Accepted 16 March 2018



This paper is published under the terms of the CC-BY license.

© 2018 The Authors

# Can uncertainty in geological cross-section interpretations be quantified and predicted?

Charles H. Randle<sup>1</sup>, Clare E. Bond<sup>1</sup>, R. Murray Lark<sup>2</sup>, and Alison A. Monaghan<sup>2</sup>

<sup>1</sup>Geology and Petroleum Geology, School of Geosciences, University of Aberdeen, Kings College, Aberdeen, AB24 3UE, UK

<sup>2</sup>British Geological Survey, Keyworth, Nottinghamshire, NG12 5GG, UK

<sup>3</sup>British Geological Survey, Edinburgh, EH14 4AP, UK

## ABSTRACT

Explicit interpretation of geological data by geologists forms the basis of many geological interpretations. However, quantitative, statistically valid research into how accurate and precise these interpretations of geological data are, and hence their uncertainty, is limited. As a result, the way that uncertainty differs between geological locations is poorly quantified and cannot be predicted. Here we show that uncertainty in cross-section interpretations varies significantly between different geological locations, and we examine the controls on this uncertainty. Using two cross-section interpretation experiments, through non-layer-cake superficial strata, we observe two distinct behaviors of uncertainty. In the first case, from Glasgow, errors in prediction of the depth of the rockhead from surrounding borehole data range from +18 m to –16 m of its actual position with a standard deviation of 6.03 m. The magnitude of this uncertainty, as measured by its mean squared value, is predictable from multiple factors relating to both the borehole data given to the geologist and their interpretation workflow. In the second case, from Manchester, the range in the predicted location of rockhead is +7 m to –9 m of the actual depth, with a standard deviation of the error of 2.69 m, and the uncertainty is only predictable from proximity to a cross-section intersection. We contrast these results with a previous experiment, on layer-cake strata from London, with an error range of  $\pm 7$  m (standard deviation 2.9 m) and where the mean square error was predictable from the experience of the interpreting geologist and the distance to the nearest borehole. Our results show that while one assumption for predicting uncertainty may be appropriate in one case, it cannot be generically applied to other cases. We conclude that care is needed when predicting uncertainty in geological cross sections from parameters associated with initial borehole data, such as data density, and further experiments are required to map out the differing behaviors of uncertainty in geological interpretation, if uncertainty is to be predicted from prior information.

## INTRODUCTION

Interpretation of borehole data allows the creation of 2D geological cross sections of subsurface strata. Borehole logs are the most common data set available for the interpretation of shallow superficial deposits. In a typical in-

terpretation workflow, the same horizons in adjacent boreholes are interpreted and linked across the intervening space to create a model of the geology in cross section. The resultant model outlines the main geological layers and discontinuities along the section line. Multiple cross sections are commonly interpreted and then linked to form a 3D geological model.

These geological models are often used as inputs into further models to make predictions. For example hydrological modeling of subsurface fluid-flow pathways (Hustoft et al., 2007), evaluating gas storage sites (Kaufmann and Martin, 2008), or to determine groundwater sources (Turner et al., 2015). But any geological model is uncertain, and these uncertainties are carried into further models. The uncertainties in geological models arise from the natural variability in any geological system, measurement uncertainty, sampling bias, and interpretation uncertainty resulting from the creation of the geological model from the data (Mann, 1993). Even if all available data are without error, the scarcity of data used to construct a geological model is such that the data will not absolutely constrain the geometry of the geological layers. Hence the accuracy of an expert's interpretation is uncertain. The error in a model may be systematic (indicating that there is a bias, e.g., a tendency to model certain contacts too low relative to datum). In the absence of bias there may be imprecision in a model, i.e., a scatter in the distribution of errors which may be narrow (low uncertainty) or wide (high uncertainty).

Geological models with significant uncertainty in the placement of geological layers and discontinuities are generally higher risk. In order to address these uncertainties, we must first map their behavior, both within specific geologies and between them. Doing so will allow us to better predict uncertainty in the future, resulting in the ability to determine uncertain regions of a 3D model, along with providing solutions to reduce that uncertainty. Here we describe two experiments that mapped out the behavior of uncertainty both within and across two geological settings.

## BACKGROUND

There are two general approaches in geological modeling. The first group is implicit, including data-driven methods that use algorithms such as kriging or other procedures to interpolate, more or less automatically, between observations. The second group of approaches is explicit, where the expert under-

taking the interpretation directly controls the interpretations on the basis of their knowledge and experience.

Many studies on model uncertainty have addressed implicit modeling methods. For example, Tacher et al. (2006) used the kriging variance as a measure of uncertainty, which is inherited directly from the statistical model that underlies kriging interpolation. Wellmann et al. (2010) used multiple conditioned simulations of the geology, and the variations between these simulations, to measure uncertainty at locations between data points.

In the case of explicit modeling, no statistical model of variability lies behind the output, and so there is no direct quantification of uncertainty, as in the case of kriging. However, there have been some attempts to quantify the uncertainty in outputs from interpretation-based geological modeling, in which a geologist's prior knowledge and experience is used to guide model construction. Lelliott et al. (2009) identified factors deemed to constrain the uncertainty in such models in a structured assessment using the fish-diagram approach of Cave and Wood (2002), and then they derived an algorithm to make local estimates of the magnitude of model uncertainty on the basis of factors such as distance to nearest borehole and automated quantification of local geological complexity.

There have been other studies aimed at determining the nature of uncertainty in interpretation of seismic image data. These studies focus on the observed differences in geological interpretations between individuals, with implications for hydrocarbon potential (Rankey and Mitchell, 2003), and on training and interpretation methodologies (Bond et al., 2007, 2012). Bond et al. (2012) show statistical relationships between factors associated with interpretation workflows and education and interpretation efficacy of a seismic image. Macrae et al. (2016) go further in a blind experiment to show a causative link between thinking about geological evolution and a positive interpretation outcome.

The experiments presented in this paper are based on the work of Lark et al. (2013, 2014), who used designed experiments to quantify uncertainty in models produced by geologists in workflows with explicit interpretation. In the study reported by Lark et al. (2014), a group of geologists independently interpreted superficial deposits in London from boreholes on a cross section. The investigators then directly measured the difference between the interpretations and withheld borehole data. The errors were examined statistically by using linear mixed modeling. The linear mixed model (LMM) allowed the authors to test for evidence of a systematic error in the interpretations (i.e., if the mean error differs from zero) and to examine whether the variance of interpretation error, which they treated as a quantitative measure of the uncertainty of the prediction, could be expressed as a function of some quantity (for example, distance to the nearest borehole). They found that the variance of the prediction errors increased with distance to the nearest borehole and was smaller for interpreters with greater experience of modeling.

The modeling of error variance (e.g., Lark et al., 2013, 2014) offers a basis to test hypotheses about the sources of uncertainty in geological interpretation from borehole data by designed experiments. An experiment of this style is

useful because the outcomes can be used to suggest improvements to data collection and model building workflows, with the aim of creating more accurate models with lower associated risk. They might also provide a basis for quantification of uncertainty in future interpretations on the basis of the factors identified as controlling the uncertainty. Conclusions drawn from a single experiment may however not be applicable to other geological locations. Interpretations using similar data sets at different localities may be affected by factors specific to the geology of an area, the presentation of data, or other unknown variables that may affect interpretation uncertainty.

We tested the hypothesis that both the magnitude of uncertainty and the factors that predict its magnitude will differ between different geologies. We quantified uncertainty in 2D vertical cross-section interpretations of borehole and geological map data at two localities with differing geologies. As in Lark et al. (2014), the standard deviation of interpretation error was used as a measure of the uncertainty in the interpretations. The results were analyzed using statistical models to determine whether properties of the data set presented to the geologists or factors relating to the interpretation itself affect the uncertainty of interpretations. The findings of these two experiments together with those of Lark et al. (2014) are used to assess whether common factors influence interpretation uncertainty at multiple geological locations and hence determine whether interpretation uncertainty can be predicted from these factors alone. If common factors are found, then they can be used for the prediction of uncertainty in the future, along with potentially suggesting where modeling workflows could be improved.

## METHODOLOGY

The objective of the experiment was to measure the uncertainty of professional geologists' interpretations of borehole data and to identify factors that contribute to the uncertainty. The methodology followed was similar to that of Lark et al. (2013) in which geologists are presented with a set of real borehole and geological map data to interpret. The interpretations of the geologists are then tested through comparison to withheld borehole data (Fig. 1). The methodology results in a direct measure of the error, rather than an estimation. Statistical modeling is then used to test the evidence for systematic errors in the interpretations (mean error is not zero) and to test hypotheses that the mean square error of the interpretation depends on specified factors. The interpretation workflow used in the experiment is similar to that used by the professional geologists, ensuring that the results are representative of a typical interpretation project.

Professional geologists from the modeling team at the British Geological Survey were asked to volunteer for the experiment. While the cohort of geologists that took part may not be representative of all geologists who carry out 3D modeling, we felt that the advantage of using a single software package that would be familiar to all participants, along with the improved logistical arrangements, outweigh the potential biases. Likewise, asking for volunteers

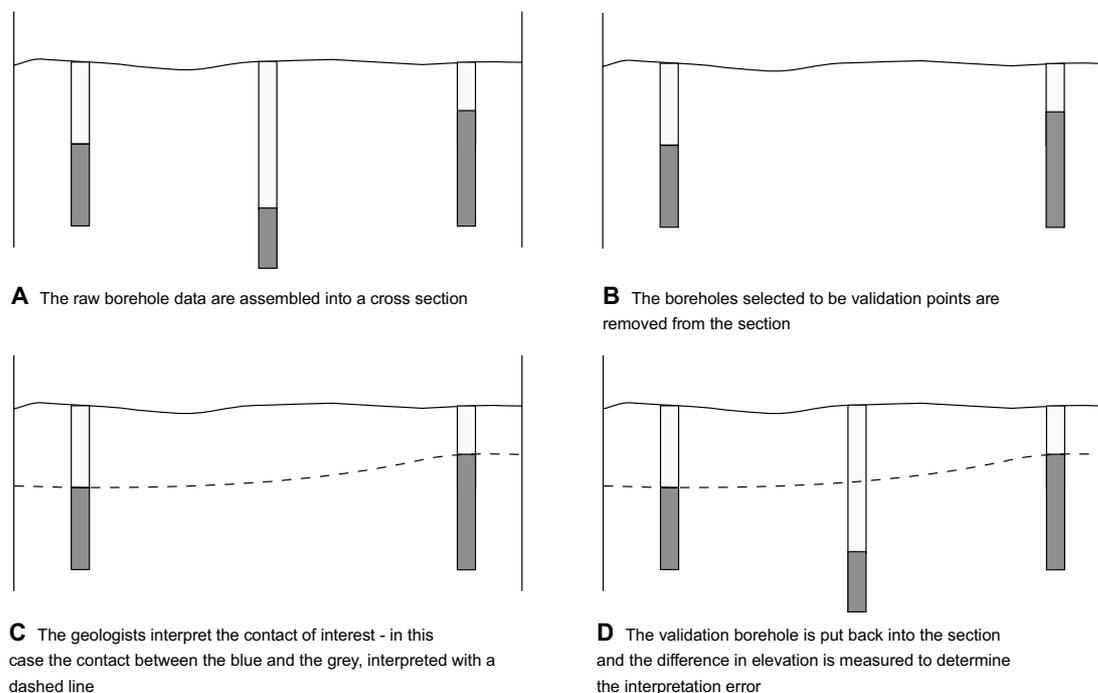


Figure 1. Diagram of the methodology followed in this experiment.

may have introduced a bias by oversampling those that are typically more confident in their interpretation abilities; however, again, the logistical benefits of using volunteers outweigh the potential drawbacks.

### Experiment Locations

For these experiments, boreholes through the glacial and postglacial superficial deposits of Glasgow and Manchester were chosen. The geological contact used to test the geologist's interpretations was the rockhead—the base of the superficial deposits and/or top of the bedrock. The geometry of the rockhead is more variable than the conformable contacts of London's superficial deposits previously investigated by Lark et al. (2014); this is due to the erosional nature of the rockhead. The structures present in Glasgow and Manchester are also distinct from each other. While both have fluvial deposits overlying till, which overlies glaciofluvial sands and gravels, in Glasgow, these sands and gravels are found in multiple narrow channels that are manifested in the rockhead as short-frequency changes in dip. In Manchester, however, the channels are much wider, creating gentle changes in the dip geometry of the rockhead. These distinct but similar geologies allow comparison to determine if differences in geology affect interpretation uncertainty.

At both locations, there is a large collated data set of boreholes (Price et al., 2012; Monaghan et al., 2014). These collated boreholes provided the key resource for the experiments, from which section lines connecting multiple boreholes were constructed. Borehole combinations were chosen that recorded enough data for an interpreter to build an understanding of the superficial geology and create a 2D vertical cross section. For the Glasgow experiment, the result was a set of 112 boreholes running west to east, in a slight arc, over a distance of 11 km. For the Manchester experiment, the result was 115 boreholes arranged in a set of seven sections covering an area of 4 km<sup>2</sup>, the data replicates that that would typically be presented to an interpreter in a standard cross-section interpretation workflow (Fig. 2).

The superficial geology of Glasgow and Manchester is the result of glacial and postglacial processes. In Glasgow (Fig. 2A), Scottish Coal Measures Group bedrock consisting of mudstone, siltstone, sandstone, and coal is overlain by a localized lower till, channelized fluvial outwash sands, and gravels, which are in turn overlain by a thick glacial till that covers the majority of the area. This is overlain by a complex succession of spatially restricted glaciofluvial, glaciolacustrine, and glacio-marine sand, gravel, and mud followed by postglacial fluvial and lacustrine deposits and artificial (man-made) ground (Browne and McMillan, 1989; Hall et al., 1998; Monaghan et al., 2014). In the Manchester

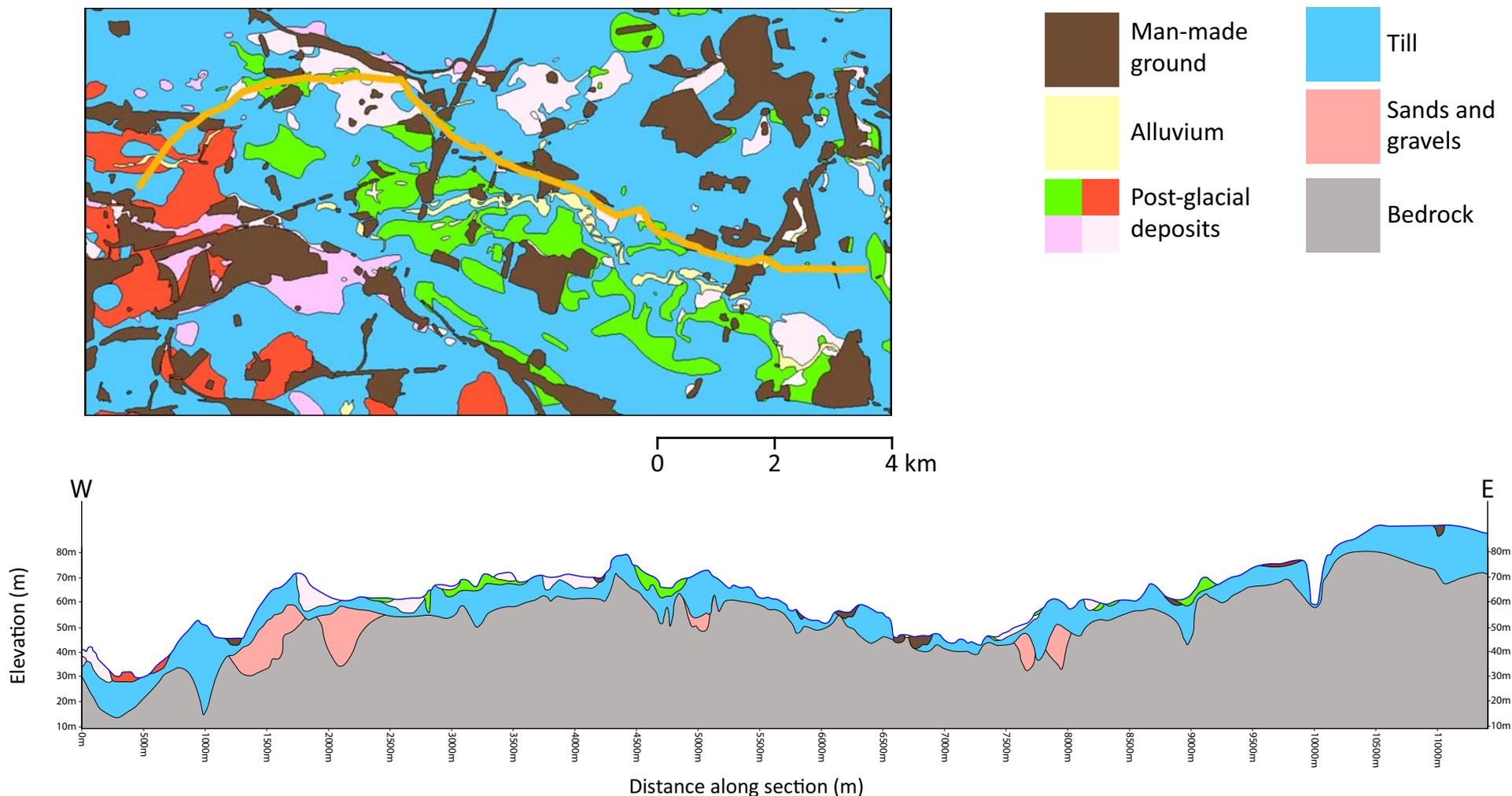


Figure 2 (on this and following page). (A) The cross section used for the Glasgow portion of this experiment along with the associated geological map. The section trace is marked by the yellow line.

experiment area (Fig. 2B), the bedrock consisting of the Sherwood Sandstone Group is overlain by channelized sands and gravels resulting from glaciofluvial outwash. This is overlain by glacial till across the majority of the section lines, but in some places the channelized glaciofluvial deposits are overlain by glaciofluvial sheet deposits of sand and gravel. In places, glaciofluvial sheet sands also occasionally overlie the till. Artificial ground is found on all section lines (Price et al., 2012).

### Allocation of Data

At each location, the entire borehole data set (112 boreholes for Glasgow; 115 for Manchester) was divided by simple random sampling, without replacement, into a number of validation batches equal to the number of geologists taking part in each experiment (ten for the Glasgow experiment; five for the Manchester experiment). The validation batches comprised the same number

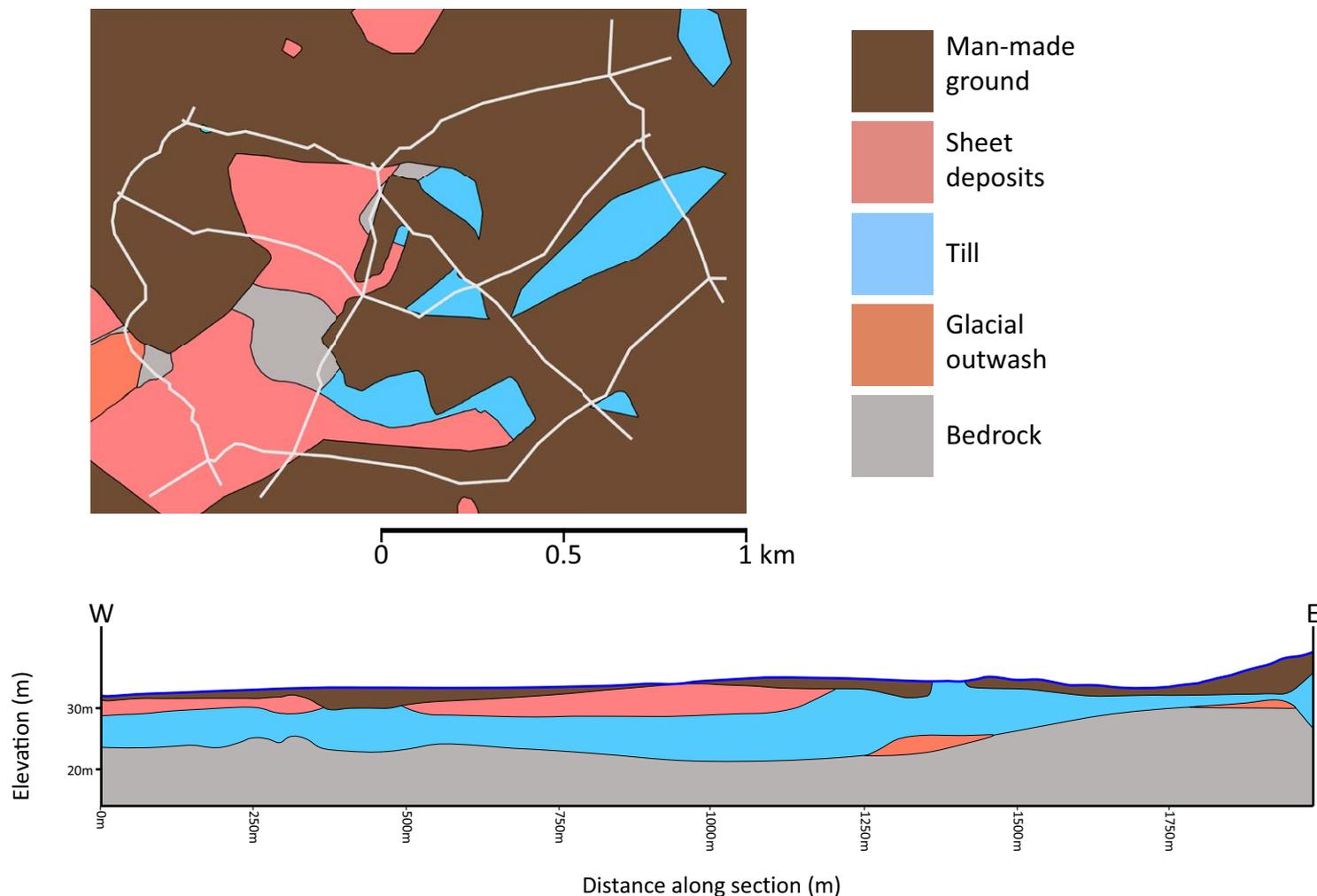


Figure 2 (continued). (B) A geological map of the region used in Manchester showing the traces of the seven sections (white) and the southernmost west-east section as an example of the geological structure.

of observations (11 for Glasgow; 23 for Manchester). Each geologist was then given a unique set of boreholes to interpret (101 for Glasgow; 92 for Manchester), from which one of the batches of boreholes had been withheld for validation. This batching process means that although each geologist's set of boreholes is unique, a borehole is validated only once, and some boreholes are common across different geologists' sets. This allows testing of error across the whole data set while minimizing loss of borehole data density.

### Interpretation Process

The modeling software package GSI3D (Kessler and Mathers, 2004) was used as the interpretation environment for the experiments. The software was chosen because it is a commonly used package for the explicit interpretation of superficial deposits within the British Geological Survey. This ensured that the interpreters had a good working knowledge of the software. In addition to

### Briefing document for Glasgow superficial cross section interpretation

#### Project outline

This experiment is part of a wider PhD study of uncertainties in two dimensions and their propagation into a 3D framework model. This is the first of several similar experiments exploring uncertainty within a variety of geological situations.

The objective of this part of the project is to determine the overall uncertainty in cross section interpretation, from that the reasons behind the uncertainty can be determined through statistical modelling.

The surfaces that you will be interpreting are the base of the Wilderness Till Formation (WTF-DMTN) and every unit below it. Some of the boreholes in the section have been withheld. These will be used to compare your interpretations to the proven geology. The results of which will enable the uncertainty to be estimated.

All responses will be confidential and participation in this experiment is entirely voluntary and you may stop participating at any time. Furthermore, the sampling design will not allow us to make meaningful estimates of the quality of your interpretations, but only general measures of interpretation uncertainty.

#### Experiment objectives

The objectives of this experiment can be summarised as:

1. To determine the overall range in cross section interpretation by comparing your interpretations to a set of withheld boreholes
2. To determine the origins of the uncertainty through the use of statistical modelling
3. To compare the uncertainty in this geological situation with that present in others
4. To suggest methods for reducing the uncertainty in cross section interpretation

<sup>1</sup>Supplemental Material. Briefing document for Glasgow superficial cross section interpretation. Please visit <http://doi.org/10.1130/GES01510.S1> or the full-text article on [www.gsapubs.org](http://www.gsapubs.org) to view the Supplemental Material.

the borehole data sets, a map of the superficial geology was included in the released data set showing the location of the boreholes. The GSI3D files were placed onto a shared drive accessible to each geologist, with their specific data set for interpretation entitled with a code assigned to each individual geologist. For both experiments, the geologists were given two weeks to complete their interpretations, allowing them time to fit in the experiment with other work commitments.

The geologists were given a set of guidelines to follow. These guidelines covered technical aspects of how to use GSI3D as well as a statement that the data presented to them should be considered to be perfectly accurate. The statement on the accuracy of the data provided was included to ensure that interpreters did not selectively ignore any of the raw data, believing it to be erroneous. The briefing document also covered a description of the experiment and the experimental process along with a brief geological summary. These are available in this paper's Supplemental Material<sup>1</sup>.

The geologists were asked to interpret all superficial deposits present in the boreholes as if they were part of a typical cross-section interpretation project. The geologists were told that they should not edit any of the map or preexisting cross-section line work that was provided to them. This was to ensure that every participant was working under the same set of assumptions and data. In the software package, the geologists can view the geological map and the borehole data along a section line. Using the borehole information, the tops of each lithology above the rockhead, and with the possibility of co-visualizing the geological map, each interpreter explicitly interprets the top of each unit, by digitizing a line between the boreholes.

## Statistical Modeling

The basic variable analyzed is the model error. The model error can be computed for each location in each of the validation batches. Recall that a validation batch is a set of boreholes withheld from the set supplied to one of the participating geologists. That geologist's interpretation is compared with each borehole in the validation batch at its corresponding location in space. The model error for the selected surface (i.e., rockhead) is the difference between the observed elevation of the surface at the location of a validation borehole and the elevation in the interpretation of the cross section by the geologist from whom that borehole was withheld. An error of zero therefore indicates that the interpretation and observation coincide. A positive error indicates that the interpreted height of the rockhead was below the observation at the validation site.

The set of errors obtained this way from all validation batches was then modeled statistically to test our hypotheses; a linear mixed model (LMM) framework was used for the modeling (Verbeke and Lesaffre, 1997) carried out using the R statistical package (R Core Team, 2015). The LMM framework was needed because the validation boreholes, while selected at random from the available boreholes, cannot be regarded as independently drawn along

the cross section since boreholes were selected to form a continuous section instead. It is therefore necessary to be able to model the spatial dependence between observed errors in any statistical model of them. If the interpreters are unbiased, the observed errors can be treated as a random variable of mean zero. If some factor, which varies between validation sites (e.g., distance to the nearest borehole or experience of the interpreter), determines the uncertainty of the interpretation, it is possible to model the variance of the observed errors as a function of these factors in the LMM framework (Nelder and Lee, 1991).

In the linear mixed model, we treat a variable as the sum of a fixed effect (here a constant mean value) and random effects. The random effects are Gaussian random variables with particular statistical properties. Because the boreholes were not chosen according to an independent random sampling scheme, we start by investigating whether the model error is spatially dependent (i.e., whether errors are correlated according to their proximity in space). The simplest linear mixed model treats the variation around the mean model error as a simple uniform Gaussian random variable with independent values with mean zero and some unknown standard deviation or variance.

The parameters of this random variable are estimated by residual maximum likelihood (REML), maximizing the log likelihood of a linear combination of the data determined by the fixed effects structure (here a constant mean). We may then fit a more complex model in which the random variation comprises two components, an independent random variable and one in which the values at any two locations are spatially correlated. This introduces two new parameters, a distance parameter that defines how the spatial correlation between two locations depends on the separation between them in space and the variance of the additional random component. Because this new model has two extra parameters, it should have at least as large a log likelihood as the simpler model. The question is whether the improvement in fit is large enough to justify the additional complexity. There are two ways to address this question. The first, which can be used in any such comparison, is to compute the Akaike information criterion (AIC; Akaike, 1974), determined from:

$$AIC = 2k - 2l, \quad (1)$$

where  $k$  is the number of parameters—coefficients that represent the factors we are testing—in the model, and  $l$  is the maximized log likelihood. We select the model for which AIC is smallest, which shows that the number of parameters to be estimated in the model is a penalty factor.

In cases where the models are nested (i.e., we can treat the simpler one as a special case of the more complex model in which certain parameters take fixed values) and for a standard comparison (i.e., the fixed values of the parameters for the simpler model take definite values not at the limit of their range), we can test the evidence that the extra terms in the more complex model are significant by computing the log-likelihood ratio statistic,  $L$ , from:

$$L = 2(l_c - l_n), \quad (2)$$

where  $l_c$  and  $l_s$  denote, respectively, the maximized log likelihood for the more complex model and the simpler model. If there are  $P$  extra parameters in the more complex model, than in the null case (i.e., where the simpler model is correct), then  $L$  is distributed asymptotically as chi-squared with  $P$  degrees of freedom. One can therefore evaluate a  $p$  value for the null hypothesis and the evidence, i.e., the probability that, if the null hypothesis were true, one would obtain a value of  $L$  as large or larger than the one observed. A small value of  $p$  is evidence against this null hypothesis.

We used AIC to evaluate the evidence for spatial correlation in the random component of the model. We also use this approach to evaluate evidence that the standard deviation (square root of the variance) of the random component depends on some factor, rather than being constant (i.e., the model uncertainty varies). For example, one might compare a simple model in which the standard deviation is a constant with a more complex alternative in which it is determined by the expression:

$$\sigma = \beta_0 + \beta_1 d, \quad (3)$$

where  $\sigma$  is the standard deviation of the error, and  $d$  is the corresponding distance to the nearest borehole. The two parameters  $\beta_0$  and  $\beta_1$  are a constant and a distance coefficient. One may compare this model to the simpler alternative on the log-likelihood ratio and so evaluate the evidence that the standard deviation of the model error depends on distance to the nearest borehole.

## Hypotheses

We tested two hypotheses through our statistical analysis. First, as in the previous work (Lark et al., 2013, 2014), that the standard deviation (and equivalently the variance) of the errors in the interpretation at a location depends on their proximity to neighboring boreholes. Second, that the standard deviation of the error was either related to or could be predicted by other factors recorded in the borehole data, the interpretations, and the interpreted sections. These additional factors are presented in Table 1.

TABLE 1. LIST OF THE PARAMETERS CHOSEN TO TEST FOR RELATIONSHIPS TO UNCERTAINTY ALONG WITH DESCRIPTIONS OF EACH

Parameter	Description
Mean depth to rockhead	The mean* of the interpreted and actual depth of the rockhead at the validation site
Number of lithological units	The number of lithological units present above the rockhead in the validation borehole (Fig. 3)
Change in number of lithological units	The difference in the number of lithological units between the two boreholes adjacent to the validation borehole
Dip angle	The angle of an interpretation created by a straight line between the two boreholes adjacent to the validation point
Experience of 3D software	The time (in years) that the geologist had been regularly using 3D geological software
Experience of GSI3D	The time (in years) that the geologist had been regularly using GSI3D
Time spent on the interpretation	The time spent (in hours) interpreting the provided dataset

\*Mean depth to rockhead was used over either the interpreted or observed depths to minimize the potential influence of regression to the mean, where exceptional differences between the interpretations and reality would skew the statistical analysis.

## RESULTS

### General Statistics

The base of the superficial deposits (also known as the rockhead) was used as the surface of interest in both Glasgow and Manchester. The elevations of interpretations of the rockhead were manually extracted from GSI3D and compared to the actual elevation observed within the withheld boreholes. Figure 3 shows the distribution of errors measured in both experiments.

The near-zero mean and symmetrical distributions indicate that there was no bias in the geologists' interpretations (Glasgow, mean error = 0.02 m; Wald test to reject the null hypothesis of zero mean = 0.001,  $p = 0.97$ ; no evidence against the null hypothesis; Manchester, mean error = -0.21 m; Wald test to reject the null hypothesis of zero mean = 0.53,  $p = 0.46$ ; no evidence against the null hypothesis); i.e., the interpretations were just as likely to be above the observed rockhead elevation as below. The clearest difference between the two distributions is that the range in predicted depth of the Glasgow rockhead is almost twice that of the Manchester data set. The results suggest that the uncertainty of an interpretation with geology similar to Glasgow is higher than an interpretation of somewhere with geology similar to Manchester.

### Statistical Modeling Results

#### Initial Analyses

The starting point of the analysis is the construction of a null model. This model is the simplest used and is fitted using only the mean error and random variable, with no additional parameters included.

This model serves as a starting point to which single-parameter-based models are compared. Next a test for spatial correlation was carried out to determine if the observed errors in the models were spatially dependent. To do this, a model was constructed using separation between boreholes as a parameter to determine the covariance of errors at those boreholes. This

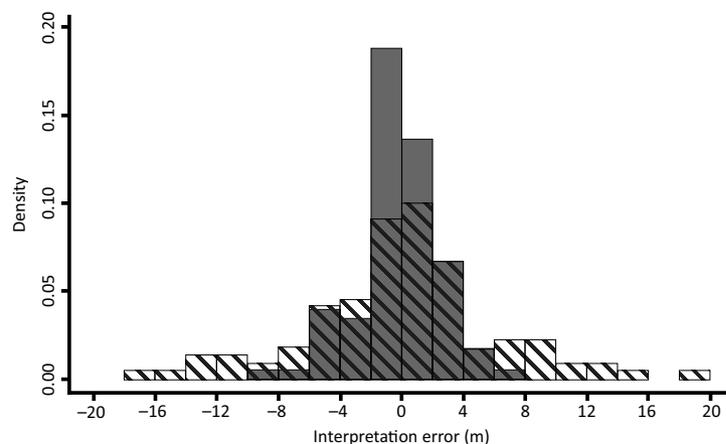


Figure 3. Stacked density plot of the measured errors in interpretations in Glasgow (hatched) and Manchester (gray). Note the narrower and more peaked distribution of the Manchester results compared to those in Glasgow and that both distributions are largely symmetrical.

was an important test because the borehole data were not randomly and independently selected. The results are presented in Table 2. The model with spatially dependent covariance can be compared with a null model without spatial dependence on the AIC because the log-likelihood ratio does not have a simple distribution under the null hypothesis of no spatial dependence (Lark, 2012).

The lack of spatial correlation suggests that error at any one validation point is independent of error at any other position along the sections. In further modeling, errors were therefore treated as independent, conditional on any factors included in the model of error variance.

### Distance to Nearest Borehole

To allow comparison with the previous work of Lark et al. (2013, 2014) a model was constructed to test whether distance to nearest borehole (DNB) was related to the standard deviation of error. This was a measurement of the horizontal distance between the validation and its nearest neighbor along the same section. The results are presented in Table 3.

The log-likelihood ratio test and the improvement (reduction) in the Akaike information criterion in the case of the Glasgow model indicate that including a linear function of distance to the nearest borehole to compute the standard deviation of the interpretation error at a site is justified. As shown in Table 3 (column F), this is a positive effect: the closer the nearest borehole is to a validation site, the smaller the error standard deviation at that site, as would be expected. For the Manchester experiment, however, there is no such effect, indicating that the improvement in the model resulting from factoring in distance to nearest borehole is not significant enough to justify the increased complexity of the statistical model. Figure 4 shows the distances to nearest borehole for the Glasgow and Manchester experiment data sets.

The spacings of the validation boreholes to the surrounding boreholes at the two localities are similar. The mean spacing for Glasgow is 76.55 m and the

TABLE 2. RESULTS OF THE NULL MODEL AND DISTANCE TO THE NEAREST BOREHOLE STATISTICAL TESTS FOR BOTH GLASGOW AND MANCHESTER

(A) Location	(B) Model	(C) Log likelihood	(D) AIC
Glasgow	Null	-252.68	509.36
	Spatial correlation	-252.68	511.36
Manchester	Null	-131.71	267.43
	Spatial correlation	-131.55	273.02

Notes: The table shows columns (A) the location, (B) the model tested, (C) log likelihood, and (D) AIC—Akaike information criterion. The increase in AIC with the addition of spatial correlation means that the model is no better with spatial correlation than without.

TABLE 3. RESULTS OF THE NULL MODEL AND DISTANCE TO THE NEAREST BOREHOLE STATISTICAL TESTS FOR BOTH GLASGOW AND MANCHESTER

(A) Model number	(B) Parameter(s)	Glasgow				Manchester			
		(C) Log likelihood	(D) AIC	(E) Correlation with error standard deviation	(F) $p$ value	(C) Log likelihood	(D) AIC	(E) Correlation with error standard deviation	(F) $p$ value
0	Null	-252.68	509.36	—	—	-155.64	315.27	—	—
1	DNB	-247.36	500.73	Positive	$10^{-3}$	-154.16	314.32	None	0.09

Notes: The table shows columns (A) the model number, (B) the parameters tested, (C) log likelihood, (D) AIC—Akaike information criterion, (E) the correlation with standard deviation, and (F) the  $p$  value of the improvement over the null model. DNB—distance to nearest borehole.

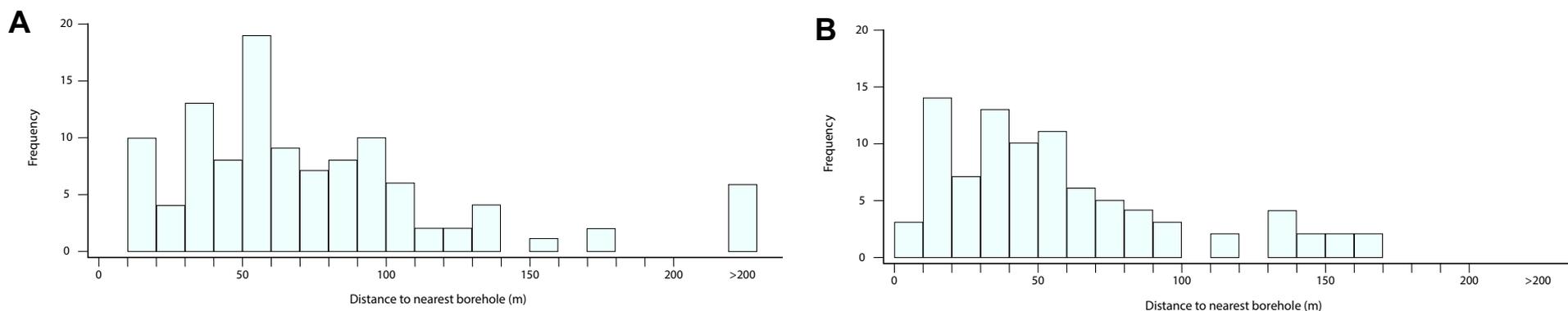


Figure 4. Histograms of the measured distances between the validation boreholes and their nearest neighbor for (A) Glasgow and (B) Manchester.

median 64 m, along an 11 km section line; for Manchester, the mean spacing is 56.5 m and the median spacing 48.5 m, and the section lines are between 1 and 2 km long. A close spacing of the boreholes may be more important for Glasgow given the shorter length frequency of the change in rockhead geometry for Glasgow, as compared to Manchester.

**Other Data Set Parameters**

To investigate whether distance to nearest borehole was the best predictor of error for the Manchester and Glasgow locations, other data-related parameters, described in Table 1, were compared to error standard deviation using the method previously described. The results of the statistical analysis for the significant parameters are presented in Table 4.

For Glasgow, the significant improvement in AIC of the depth model compared to the distance to nearest borehole and number of units in the borehole suggests that uncertainty is more significantly related to depth than either distance to the nearest borehole or the number of lithological units in the borehole above the rockhead. This means that for an interpretation in Glasgow,

the best predictor of uncertainty is the depth of the rockhead from the surface. To test this, further combinations of all three parameters were considered (Table 5). The multi-parameter models were built in a similar manner to the single-parameter models but with the addition of a second and third random variable, each with standard deviations of the additional parameters. The depth model (model 3) replaced the null model as the basis for comparison because it was the best of the single-parameter models.

The addition of the number of lithological units in the borehole above the rockhead to a model considering depth (model 4) results in the best-fitting Glasgow model so far (the model with lowest AIC). The addition of distance to nearest borehole (model 6) does not provide a significant improvement ( $p = 1$ ); neither does the combination of depth and distance to nearest borehole (model 5,  $p = 0.5$ ). This does not necessarily indicate that distance to the nearest borehole is not a significant predictor of uncertainty, although it could suggest that the apparent effect is a coincidence resulting from the validation points rather than a true effect. Notably none of these predictors or combinations of predictors were found to be significant in the Manchester experiment, suggesting that uncertainty cannot be predicted using these parameters at that location.

TABLE 4. RESULTS FOR THE THREE PARAMETERS THAT WERE FOUND TO CORRELATE WITH UNCERTAINTY IN GLASGOW AND MANCHESTER

(A) Model number	(B) Parameter(s)	Glasgow				Manchester			
		(C) Log likelihood	(D) AIC	(E) Correlation with error standard deviation	(F) $p$ value	(C) Log likelihood	(D) AIC	(E) Correlation with error standard deviation	(F) $p$ value
0	Null	-252.68	509.36	-	-	-155.64	315.27	-	-
1	DNB	-247.36	500.73	Positive	$10^{-3}$	-154.16	314.32	None	0.09
2	Lithological units	-247.69	501.39	Positive	$10^{-3}$	-155.56	317.12	None	1
3	Mean depth to rockhead	-232.17	470.33	Positive	$10^{-10}$	-155.63	317.26	None	0.9

Notes: The null model result is presented again for reference. The table shows columns (A) the model number, (B) the parameters tested, (C) log likelihood, (D) AIC—Akaike information criterion, (E) the direction of the correlation between error standard deviation and that parameter, and (F) the  $p$  value of any improvement over the null model. DNB—distance to nearest borehole.

TABLE 5. RESULTS OF MULTI-PARAMETER ANALYSIS FOR GLASGOW

(A) Model number	(B) Parameter(s)	(C) Log likelihood	(D) AIC	(E) Correlation with error standard deviation	(F) <i>p</i> value
3	Mean depth to rockhead	-232.17	470.33	Positive	10 <sup>-10</sup>
4	Depth + lithological units	-230.07	468.14	Positive/positive	0.04
5	Depth + DNB	-231.90	471.79	Positive/positive	0.5
6	Depth + DNB + lithological units	-229.81	469.62	Positive/positive/positive	1.0

Notes: Model 3 is repeated from Table 3 because it is used as a baseline model to compare to models 4 and 5. Model 6 was compared to model 4 because it was the best-fitting model of the pair of two-parameter models. The columns are: (A) the model number, (B) the parameters tested, (C) log likelihood, (D) AIC—Akaike information criterion, (E) the direction of the correlation between error standard deviation and the parameters, and (F) the *p* value of any improvement over the previously best model. DNB—distance to nearest borehole.

**Interpreter Effects**

Next we tested parameters that related to the geologists themselves or their interpretations. The results for the Glasgow experiment are presented in Table 6. As with the previous tests, none of these factors were found to be significant for the Manchester experiment.

We found that a geologist’s experience of any 3D geological software is in general unrelated to the uncertainty in their interpretation, for both Glasgow and Manchester experiments. This is contrary to the findings of Lark et al. (2014) for their London experiment. However, the geologists’ experience in GSI3D (the specific software used) was found to be a significant factor in the Glasgow interpretation. We observed improvement in the accuracy of interpretations with increasing time spent on the interpretation. While the result is not conclusive, due to the low number of participants (ten), it suggests that when interpreting superficial deposit data for Glasgow, ample time should be given to the geologists to complete their interpretations.

**Cross-Section Effects**

In a further test, we consider the effect of the presentation of the data to the interpreters. For Glasgow, this was the distance along the section from the left (westernmost) side of the section. It was found to be a significant parameter (Table 7; *p* value of improvement over the null model = 10<sup>-6</sup>, AIC 488.41), though not as significant as depth (Table 4, model 3). The addition of depth to this model results in the best-fitting model of all of the Glasgow models (*p* = 0.02, AIC 466.49) despite being dependent on fewer parameters. This means that the best prediction of the uncertainty in an interpretation of the Glasgow data is given by a combination of the depth and distance along the section. Distance along section was not a significant parameter in Manchester, although the 3D arrangement of the multiple sections may have obscured a relationship.

Why the variance of errors depends on the distance along section in Glasgow is not clear. One hypothesis is that it may be a result of the interpreters becoming more accurate with increased time spent interpreting, as-

TABLE 6. RESULTS OF THE GEOLOGIST-BASED STATISTICAL TESTS FOR BOTH GLASGOW AND MANCHESTER

(A) Model number	(B) Parameter(s)	Glasgow				Manchester			
		(C) Log likelihood	(D) AIC	(E) Correlation with error standard deviation	(F) <i>p</i> value	(C) Log likelihood	(D) AIC	(E) Correlation with error standard deviation	(F) <i>p</i> value
7	Experience of 3D software	-252.01	510.01	Negative	0.24	-155.63	317.26	None	0.8
8	Experience of GSI3D	-245.01	496.02	Negative	10 <sup>-5</sup>	-155.19	316.39	None	0.34
9	Time spent doing the interpretation	-246.00	499.20	Negative	10 <sup>-4</sup>	-155.56	317.13	None	0.7

Notes: The table shows columns (A) the geological location, (B) the model tested, (C) log likelihood, (D) AIC—Akaike information criterion, (E) the correlation with standard deviation, and (F) the *p* value of any improvement over the null model.

TABLE 7. THE RESULTS FOR MODELS THAT INCLUDED DISTANCE ALONG SECTION AS A PARAMETER

(A) Model number	(B) Parameter(s)	(C) Log likelihood	(D) AIC	(E) Correlation with error standard deviation	(F) <i>p</i> value
10	Distance along section	-241.20	488.41	Negative	10 <sup>-6</sup>
11	Distance + depth	-229.25	466.49	Negative/positive	0.02

Notes: Model 10 was compared to the null model; model 11 was compared to model 10. The columns are: (A) the model number, (B) the parameters tested, (C) log likelihood, (D) AIC—Akaike information criterion, (E) the direction of the correlation between error standard deviation and the parameters, and (F) the *p* value of that improvement over the compared model.

TABLE 8. THE RESULTS OF THE ONLY SIGNIFICANT PARAMETER FOUND IN MANCHESTER, PROXIMITY TO A CROSSING SECTION LINE

(A) Model number	(B) Parameter(s)	(C) Log likelihood	(D) AIC	(E) Correlation with error standard deviation	(F) <i>p</i> value
0	Null	-155.64	315.27	-	-
12	Proximity to a section intersection	-129.60	265.2	Positive	0.04

Notes: With the null model result repeated from Table 3. The columns are: (A) the model number, (B) the parameters tested, (C) log likelihood, (D) AIC—Akaike information criterion, (E) the direction of the correlation between error standard deviation and that proximity, and (F) the *p* value of the improvement over the null model.

suming a left to right interpretation direction. Alternatively, there may be some aspect of the geology not picked up by any of the tested parameters. Discussion with the geologists who participated in the experiment suggested that there is a change along section, with the westernmost end being dominated by fluvial and glaciofluvial deposits that introduce more conceptual possibilities in the interpretation and likely a greater range in top rockhead morphology as compared to more simple glacial deposits at the eastern end.

For the Manchester experiment, the presentation of the data is more complex than Glasgow. Rather than a single cross-section line (Fig. 2A), the borehole data are arranged in a set of seven intersecting cross sections (Fig. 2B). Given the network of sections, testing simply for distance along section did not seem appropriate, especially given that no data were captured on whether interpreters worked on each section line in turn and in what order, or if they jumped between sections during their interpretation. Instead an additional test of proximity to a crossing section line was made. The distance between the validation and the nearest of the 11 section intersections was used as the parameter of interest, and when tested against the null model, was found to be significant (Table 8). This result suggests that intersecting cross-section lines reduce interpretation uncertainty. We hypothesize that this improvement is caused by the addition of data from the third dimension giving better control of what the geology is doing, thus allowing better decision making.

## DISCUSSION

### Can We Predict Uncertainty?

The two experiments show two distinct behaviors of error and, hence, two distinct behaviors of uncertainty. The Glasgow experiment shows that the uncertainty in interpretation can be predicted using several parameters, including both factors associated with the data used for the interpretation and factors relating to the interpreting geologists. In contrast, the results of the Manchester experiment suggest that uncertainty is almost entirely unpredictable, with the only factor found that can be used to predict the uncertainty being proximity to a crossing section line.

To further investigate these differing behaviors, we compare the results of our Glasgow and Manchester experiments to the results of Lark et al. (2014) for the “layer-cake” stratigraphy of the London Clay. The distribution of interpretation errors for the London experiment was similar to those found in the

Manchester experiment (Fig. 5) with a symmetrical distribution about a mean of zero. Statistical modeling found that error variance and, hence, uncertainty, increased with distance to the nearest borehole and the experience in using 3D geological modeling software of the interpreting geologist, similar to the findings of the Glasgow experiment.

Across the three experiments, three distinct behaviors of uncertainty are documented: (1) predictable and high uncertainty in the variable superficial geology of Glasgow; (2) unpredictable and low uncertainty in the moderately variable superficial geology of Manchester; and (3) predictable high uncertainty in the simple layer-cake geology of London. The differing behaviors of uncertainty allow us to classify them based on the predictability and magnitude of uncertainty (Fig. 6).

In the classification scheme (Fig. 6), each observed behavior is treated as an end member. The experiment localities are placed in the corners of the diagram, with the diagram effectively normalized to them. Further studies in other localities would be likely to result in movement of the current localities toward the center of the diagram as they are replaced with localities (e.g., with even higher predictability and higher magnitude uncertainty). Further experiments may also prove the existence of the fourth as yet unobserved behavior of uncertainty—high in magnitude and low predictability.

What appears to be the case, in the examples so far, is that parameters based purely on the “complexity” of the geology cannot be used to predict uncertainty, or the magnitude of uncertainty, in a cross-section interpretation. If geological “complexity” was a good predictor of interpretation uncertainty, we would expect the London example with simple layer-cake stratigraphy to

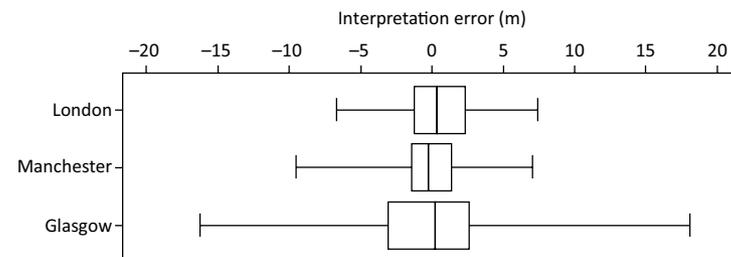
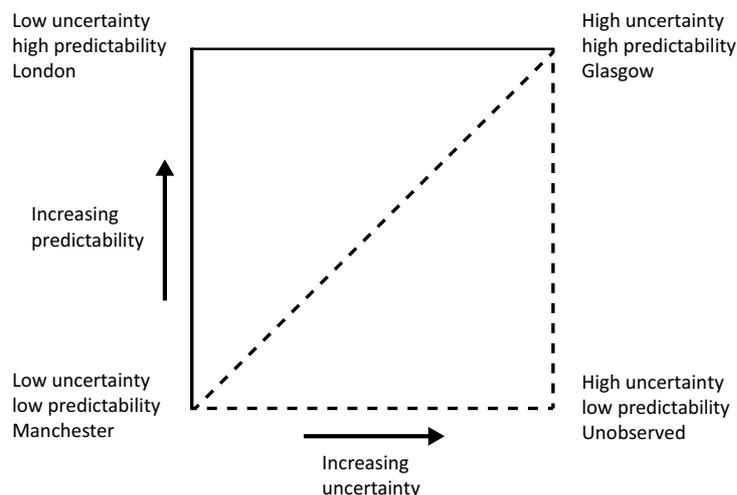


Figure 5. Boxplots of the measured error interpretations for London from Lark et al. (2014), and Manchester and Glasgow (this study). Note the similarity between the ranges and interquartile range of the London and Manchester results and the considerably wider distribution of the Glasgow results.



**Figure 6.** Uncertainty classification diagram—a graphical representation of the three behaviors of uncertainty found in this experiment and in Lark et al. (2014).

have a highly predictable, low uncertainty; the Manchester example to have a moderately predictable, moderate uncertainty; and the Glasgow example, with the apparently most “complex” glaciofluvial stratigraphy, to be least predictable and have high uncertainty. The Manchester experiment results do not fit with the hypothesis that geological complexity and interpretational uncertainty have a simple predictable linear relationship. The results raise several questions that are discussed below.

### Defining Geological Complexity

Various authors have attempted to integrate geological complexity into their consideration of how best to determine uncertainty in geological models. For example, many authors attempt to evoke a hierarchical set of features and relationships to build models (Deutsch and Wang, 1996; Wu et al., 2005), combine geological reasoning with statistics (Lelliott et al., 2009), or use the number of geological units or faults within a defined cell to determine model entropy or uncertainty in a model (Wellmann and Regenauer-Leib, 2012; Richards et al., 2015). These latter examples equate complexity (defined by the number of “options” within a cell) to uncertainty. Such attempts to define complexity in a quantitative manner can be used as a comparator between sites. The term “geological complexity” when used in geology is often used in a qualitative manner, and what may be complex to one geologist (working in polydeformed metamorphic terrains) may be quite different to another geologist (working on sedimentary basins). In situations where the geological strata

are conformable and undisrupted by faults or folds, as for example in the Lark et al. (2014) London experiment, the lack of “complexity” should make the uncertainty in a cross-section interpretation predictable. Indeed this is what Lark et al. (2014) found: a small-magnitude uncertainty that is unbiased (symmetrical around the real location of the test geological boundary) and predictable based on data density (distance to the nearest borehole) and with some effect on the magnitude of uncertainty from the geologist’s experience in using 3D interpretation software. In such examples, where the geology is simple, stochastic modeling and interpolation using implicit modeling (e.g., Wellmann et al., 2010) to create a cross section may give as good a result as explicit interpreter-led cross-section construction, providing a realistic prediction of the uncertainty in the placement of a geological boundary along the section line.

In geological terrains that include any one of the following aspects—unconformities, folds, or faults—geological reasoning (Frodeman, 1995) and rules of superposition must be employed in an interpretation (Bond, 2015). The superficial deposits of Manchester and Glasgow both contain numerous unconformable surfaces where the glacial deposits have eroded each other and the rockhead. For Glasgow, the frequency of the individual erosional features is high, resulting in high-frequency changes in the elevation of the rockhead. Despite visually determining the Glasgow geology as more complex, describing this quantitatively to model statistically is a challenge. Here we used: (1) the dip angle (the angle created by joining the two boreholes adjacent to the validation point with a straight line) as a proxy for the rate of change in slope around the test borehole; (2) the number of lithological units above the rockhead in the surrounding boreholes to give a picture of the lithological variety; and (3) the depth below the surface of the rockhead as proxies for “complexity.” All but the dip angle were found to relate to uncertainty for the Glasgow experiment, but there was no relationship for the Manchester experiment. This suggests that our proxies for geological complexity are poor, lacking the sophistication required to truly define complexity. Alternatively, beyond a threshold point, complexity becomes less important than other factors in determining uncertainty. Such a threshold point may actually not be far from simple layer-cake stratigraphy, with uncertainty in sections containing any features (e.g., unconformity, fault, or fold) that require an element of reasoning being better determined from other factors.

### General Predictors of Uncertainty

Our results support contentions that conclusions drawn from single experiments and therefore single localities may not provide robust general predictors of interpretational uncertainty. In most previous studies (Lelliott et al., 2009; Lark et al., 2014), only a single geological location has been used to test reasons for interpretational uncertainty. An exception to this is a seismic interpretation exercise by Bond et al. (2012); their findings, which were tested in a further seismic interpretation experiment (Macrae et al., 2016), independently confirmed that both education level and thinking about the geological evolu-

tion of the interpreted geology were significant factors in reducing interpretation uncertainty and/or improving interpretation efficacy.

The experiments presented here are based on cross-section interpretations of dense borehole data in superficial geology. In more “complex” solid geology locations, particularly in metamorphic and igneous terrains, borehole density, apart from in actively mined areas, is likely to be much lower. In these types of geological terrains, the requirement for robust geological reasoning is high, and conceptual uncertainty (the possibility that more than one conceptual model is probable from the given data; see Bond et al., 2007) is great. In such terrains, we would hypothesize that the potential magnitude of uncertainty is likely to be high and its predictability low. As a result of both low-density borehole data and “complex” geology, our workflow and the ability to draw conclusions from it may not be easily transferable to solid geological situations, particularly in igneous and metamorphic terrains. Simply, there might be too little constraint on the geological possibilities between boreholes in some areas.

The issue of transferring the workflow presented here to other localities (and geologies) is further complicated by the different data sets available to aid in the interpretation of subsurface geology (e.g., gravity, magnetics, seismic image data, and borehole logs, including wireline and formation imaging tools) and the fact that more than one of these data sets is often combined in a subsurface interpretation workflow. Understanding the interplay and controls on uncertainty behavior in situations where multiple data sets are invoked in an interpretation workflow is a further challenge not considered here. Consideration has been given to these issues by Lindsay et al. (2013), who combine geophysical and geological data to assess uncertainty in the Ashanti Greenstone Belt; they found that some geophysical techniques highlight geological complexities in the prospective gold layer and that combining data sets can aid in the understanding of uncertainty.

The lack of predictability in being able to determine uncertainty from our tested parameters, despite using identical workflows at multiple geological locations, suggests that the conclusions drawn from previous studies based on observations at one locality are not likely to be applicable to other locations. Analysis of our own work, and that of others, suggests that significant further work is required to test the validity of suggested reasons for interpretational uncertainty, if they are to be used as predictors in other locations. Challenges include designing tests that can be used on a variety of multiple subsurface data sets and defining geological complexity.

## CONCLUSIONS

The experiments presented here show that the nature of interpretational uncertainty changes across different locations. Our two experiments for Glasgow and Manchester, in combination with the results of Lark et al. (2014) for London, demonstrate three distinct behaviors of uncertainty. Each experiment varies in either the predictability (i.e., predictable by many

parameters versus few parameters) or the magnitude of uncertainty in interpretation of the data.

Our results show that the behavior of uncertainty is not reliably predictable from factors relating to the geology, the geologists, or the interpretation. Hence, the behavior of interpretation uncertainty at one geological location is not necessarily transferrable to other locations. This is the case even when the geologies are broadly similar; therefore, care must be taken when using factors from one site to predict uncertainties in another.

To truly determine if specific parameters, or combinations of parameters, can be used to predict interpretation uncertainty across multiple geological locations, a greater number of locations with similar data sets need to be tested. These further experiments would serve to populate the uncertainty behavior chart (Fig. 6) and in doing so may provide insight into the root causes of the behavior of uncertainty in borehole interpretation of superficial deposits. The optimum outcome would be to allow the behavior of uncertainty to be predicted ahead of interpretation. This would have value through both making the geologist more aware of where the interpretation may be more challenging and allowing better modeling of uncertainty in preexisting and new 3D geological models.

## ACKNOWLEDGMENTS

This work was undertaken while C.H. Randle held a joint British Geological Survey University Funding Initiative (BUFI) and University of Aberdeen, College of Physical Sciences Ph.D. Studentship at Aberdeen University. The contributions by C.H. Randle, R.M. Lark, and A.A. Monaghan are published with the permission of the Executive Director of the British Geological Survey Natural Environment Research Council. We would also like to thank all those who took part in both experiments as well as the many people who have given input on our results.

## REFERENCES CITED

- Akaike, H., 1974, A new look at the statistical model identification: *IEEE Transactions on Automatic Control*, v. 19, no. 6, p. 716–723, <https://doi.org/10.1109/TAC.1974.1100705>.
- Bond, C.E., 2015, Uncertainty in structural interpretation: Lessons to be learnt: *Journal of Structural Geology*, v. 74, p. 185–200, <https://doi.org/10.1016/j.jsg.2015.03.003>.
- Bond, C.E., Gibbs, A.D., Shipton, Z.K., and Jones, S., 2007, What do you think this is?: “Conceptual uncertainty” in geoscience interpretation: *GSA Today*, v. 17, p. 4–10, <https://doi.org/10.1130/GSAT01711A.1>.
- Bond, C.E., Lunn, R.J., Shipton, Z.K., and Lunn, A.D., 2012, What makes an expert effective at interpreting seismic images?: *Geology*, v. 40, p. 75–78, <https://doi.org/10.1130/G32375.1>.
- Browne, M.A.E., and McMillan, A.A., 1989, Quaternary geology of the Clyde Valley: British Geological Survey Research Report SA/89/1, Onshore Geology Series, 63 p.
- Cave, M.R., and Wood, B., 2002, Approaches to the measurement of uncertainty in geoscience data modelling: British Geological Survey Internal Report IR/02/068, 42 p.
- Deutsch, C.V., and Wang, L., 1996, Hierarchical object-based stochastic modelling of fluvial reservoirs: *Mathematical Geology*, v. 28, no. 7, p. 857–880, <https://doi.org/10.1007/BF02066005>.
- Frodeman, R., 1995, Geological reasoning: Geology as an interpretive and historical science: *Geological Society of America Bulletin*, v. 107, no. 8, p. 960–968, [https://doi.org/10.1130/0016-7606\(1995\)107<0960:GRGAAL>2.3.CO;2](https://doi.org/10.1130/0016-7606(1995)107<0960:GRGAAL>2.3.CO;2).
- Hall, I.H.S., Browne, M.A.E., and Forsyth, I.H., 1998, Geology of the Glasgow District: Memoir of the British Geological Survey, Sheet 30E (Scotland).
- Hustoft, S., Mienert, J., Bünz, S., and Nouzé, H., 2007, High-resolution 3D-seismic data indicate focussed fluid migration pathways above polygonal fault systems of the mid-Norwegian margin: *Marine Geology*, v. 245, no. 1–4, p. 89–106, <https://doi.org/10.1016/j.margeo.2007.07.004>.

- Kaufmann, T., and Martin, T., 2008, 3D geological modelling from boreholes, cross-sections and geological maps, application over former natural gas storages in coal mines: *Computers & Geosciences*, v. 34, no. 3, p. 278–290, <https://doi.org/10.1016/j.cageo.2007.09.005>.
- Kessler, H., and Mathers, S.J., 2004, Maps to models—Finally capturing the geologists' vision: *Geoscientist*, v. 14, p. 4–6.
- Lark, R.M., 2012, Distinguishing spatially correlated random variation in soil from a 'pure nugget' process: *Geoderma*, v. 185–186, p. 102–109.
- Lark, R.M., Mathers, S.J., Thorpe, S., Arkley, S.L.B., Morgan, D.J., and Lawrence, D.J.D., 2013, A statistical assessment of the uncertainty in a 3-D geological framework model: *Proceedings of the Geologists' Association*, v. 124, no. 6, p. 946–958, <https://doi.org/10.1016/j.pgeola.2013.01.005>.
- Lark, R.M., Thorpe, S., Kessler, H., and Mathers, S.J., 2014, Interpretative modelling of a geological cross section from boreholes: Sources of uncertainty and their quantification: *Solid Earth*, v. 5, p. 1189–1203, <https://doi.org/10.5194/se-5-1189-2014>.
- Lelliott, M.R., Cave, M.R., and Walthall, G.P., 2009, A structured approach to the measurement of uncertainty in 3D geological models: *Quarterly Journal of Engineering Geology and Hydrogeology*, v. 42, p. 95–105, <https://doi.org/10.1144/1470-9236/07-081>.
- Lindsay, M.D., Perrouty, S., Jessell, M.W., and Aillères, L., 2013, Making the link between geological and geophysical uncertainty: Geodiversity in the Ashanti Greenstone Belt: *Geophysical Journal International*, v. 195, p. 903–922, <https://doi.org/10.1093/gji/ggt311>.
- Macrae, E.J., Bond, C.E., Shipton, Z.K., and Lunn, R.J., 2016, Increasing the quality of seismic interpretation: *Tulsa, Interpretation*, v. 4, no. 3, p. T395–T402, <https://doi.org/10.1190/INT-2015-0218.1>.
- Mann, C.J., 1993, Uncertainty in geology, *in* Davis, J.C., and Herzfeld, U.C., eds., *Computers in Geology—25 Years of Progress*: Oxford, UK, Oxford University Press, p. 241–254.
- Monaghan, A.A., Arkley, S.L.B., Whitbread, K., and McCormac, M., 2014, Clyde superficial deposits and bedrock models released to the ASK Network 2014: A guide for users: Version 3: *British Geological Survey Open Report OR/14/013*.
- Nelder, J.A., and Lee, Y., 1991, Generalized Linear Models for the Analysis of Taguchi-Type Experiments: *Applied Stochastic Models in Business and Industry*, v. 7, p. 107–120, <https://doi.org/10.1002/asm.3150070110>.
- Price, S.J., Kessler, H., Burke, H.F., Hough, E., and Reeves, H.J., 2012, Model metadata report for Manchester and Salford, NW England: *British Geological Survey Open Report OR/12/068*.
- R Core Team, 2015, *R, A language and environment for statistical computing*: Vienna, Austria, R Foundation for Statistical Computing: URL <http://www.R-project.org/>.
- Rankey, E.C., and Mitchell, J.C., 2003, That's why it's called interpretation: Impact of horizon uncertainty on seismic attribute analysis: *The Leading Edge*, v. 22, p. 820–828, <https://doi.org/10.1190/1.1614152>.
- Richards, F.L., Richardson, N.J., Bond, C.E., and Cowgill, M., 2015, Interpretational variability of structural traps: Implications for exploration risk and volume uncertainty, *in* Richards, F.L., et al., eds., *Industrial Structural Geology: Principles, Techniques and Integration*: Geological Society of London, Special Publications, v. 421, no. 1, p. 7–27, <https://doi.org/10.1144/SP421.13>.
- Tacher, L., Pomian-Srednicki, I., and Parriaux, A., 2006, Geological uncertainties associated with 3-D subsurface models: *Computers & Geosciences*, v. 32, p. 212–221, <https://doi.org/10.1016/j.cageo.2005.06.010>.
- Turner, R.J., Mansour, M.M., Dearden, R.O., Dochartaigh, B.E., and Hughes, A.G., 2015, Improved understanding of groundwater flow in complex superficial deposits using three-dimensional geological-framework and groundwater models: An example from Glasgow, Scotland (UK): *Hydrogeology Journal*, v. 23, no. 3, p. 493–506, <https://doi.org/10.1007/s10040-014-1207-0>.
- Verbeke, G., and Lesaffre, E., 1997, The linear mixed model. A critical investigation in the context of longitudinal data, *in* Gregoire, T.G., Brillinger, D.R., Diggle, P.J., Russek-Cohen, E., Warren, W.G., and Wolfinger, R.D., eds., *Modelling Longitudinal and Spatially Correlated Data*: New York, Springer, p. 89–99.
- Wellmann, J.F., and Regenauer-Lieb, K., 2012, Uncertainties have a meaning: Information entropy as a quality measure for 3-D geological models: *Tectonophysics*, v. 526, p. 207–216, <https://doi.org/10.1016/j.tecto.2011.05.001>.
- Wellmann, J.F., Horowitz, F.G., Schill, E., and Regenauer-Lieb, K., 2010, Towards incorporating uncertainty of structural data in 3D geological inversion: *Tectonophysics*, v. 490, p. 141–151, <https://doi.org/10.1016/j.tecto.2010.04.022>.
- Wu, Q., Xu, H., and Zou, X., 2005, An effective method for 3D geological modeling with multi-source data integration: *Computers & Geosciences*, v. 31, no. 1, p. 35–43, <https://doi.org/10.1016/j.cageo.2004.09.005>.