

Automatic mapping of the base of aquifer — A case study from Morrill, Nebraska

Mats Lundh Gulbrandsen¹, Lyndsay B. Ball², Burke J. Minsley², and Thomas Mejer Hansen¹

Abstract

When a geologist sets up a geologic model, various types of disparate information may be available, such as exposures, boreholes, and (or) geophysical data. In recent years, the amount of geophysical data available has been increasing, a trend that is only expected to continue. It is nontrivial (and often, in practice, impossible) for the geologist to take all the details of the geophysical data into account when setting up a geologic model. We have developed an approach that allows for the objective quantification of information from geophysical data and borehole observations in a way that is easy to integrate in the geologic modeling process. This will allow the geologist to make a geologic interpretation that is consistent with the geophysical information at hand. We have determined that automated interpretation of geologic layer boundaries using information from boreholes and geophysical data alone can provide a good geologic layer model, even before manual interpretation has begun. The workflow is implemented on a set of boreholes and airborne electromagnetic (AEM) data from Morrill, Nebraska. From the borehole logs, information about the depth to the base of aquifer (BOA) is extracted and used together with the AEM data to map a surface that represents this geologic contact. Finally, a comparison between our automated approach and a previous manual mapping of the BOA in the region validates the quality of the proposed method and suggests that this workflow will allow a much faster and objective geologic modeling process that is consistent with the available data.

Introduction

In order for a geologic model to be reliable, there has to be consistency between the model and the available sources of information about the earth structure, such as boreholes, geophysical data, and (or) geologic background knowledge. Regional geophysical data sets, such as airborne electromagnetic (AEM) surveys, are becoming increasingly common tools in the development of geologic and hydrological frameworks. For example, approximately 40% of Denmark has been mapped with the AEM data in support of their national groundwater program (Mielby, 2010); in Ireland and Northern Ireland, three different AEM studies have been merged to cover a total of 94,932 line-kilometers, or an area of 18,846 km² (Hodgson and Ture, 2015); in Australia, there have been several large AEM studies such as the Paterson Province survey (Cassidy et al., 2010/12) and the Muchison survey (Davis et al., 2016), with AEM data covering 47,600 and 51,284 km², respectively; and recent pilot projects have been conducted in India (Auken et al., 2013) and Thailand (Sanderson and Pjetursson, 2014), covering 3000 and 1000 km², respectively. These surveys can provide indirect geologic information that is spatially extensive

and challenging to interpret. Data-rich geophysical surveys make it progressively harder for the geologist to incorporate all available data when building a geologic model. A manual interpretation approach not only makes it difficult to evaluate the large amount of available data, but it also makes it practically impossible to develop geologic models consistent with all the available information in an objective fashion. In this paper, we demonstrate a workflow that quantitatively incorporates borehole lithologic logs and AEM data to create a map of a subsurface geologic contact. This map represents information about the contact geometry directly derived from the geophysical data and can be integrated with the geologic modeling process. In this way, geologists are provided with a highly informative and data-based map that can be thought of as an optimal starting point for manual updating of the geologic model based on their geologic knowledge. Following this workflow ensures that the final geologic model is consistent with geophysics, well logs, and geologic background knowledge, but it also allows for a much faster and more consistent modeling process.

In this study, we illustrate an automated workflow that maps the contact between a surficial aquifer and

¹The University of Copenhagen, Solid Earth Physics, Copenhagen, Denmark. E-mail: mats.lundh@nbi.ku.dk; tmeha@nbi.ku.dk.

²Crustal Geophysics and Geochemistry Science Center, United States Geological Survey, Denver, USA. E-mail: lball@usgs.gov; bminsley@usgs.gov.

Manuscript received by the Editor 31 October 2016; revised manuscript received 3 February 2017; published online 23 March 2017. This paper appears in *Interpretation*, Vol. 5, No. 2 (May 2017); p. T231–T241, 8 FIGS., 2 TABLES.

<http://dx.doi.org/10.1190/INT-2016-0195.1>. © 2017 Society of Exploration Geophysicists and American Association of Petroleum Geologists. All rights reserved.

its underlying aquitard by integrating an AEM data set and borehole data from western Nebraska, USA. This is done using an approach based on an attribute-guided regression technique called Smart Interpretation (SI) (Gulbrandsen et al., 2017). This method learns the relation between the depth to the base of aquifer (BOA) observed in borehole lithologic logs and a set of attributes from a deterministic inversion of the AEM data. Knowing this relation, the method can extrapolate the map of the BOA to the whole survey area, where AEM data are available. This result is compared with a manual interpretation of the same area. The SI method and the manual interpretation are based on the same deterministic inversion of the AEM data, allowing for a direct comparison of the SI with a more traditional and fully independent interpretation. Additionally, a Bayesian Markov chain Monte Carlo (McMC) inversion (Minsley, 2011) of the AEM data is considered as additional information. This probabilistic method explores the likelihood of vertical resistivity contrasts occurring throughout the depth of investigation (DOI), which can be used to identify the uncertainty in positions in which geologic contacts are likely to occur. Comparing these probabilistic results with the initial SI-interpreted BOA map gives an idea of where the SI interpretation can be improved through manual adjustment or addition of ground geophysical surveys. Information on likely geologic contact positions

from the McMC inversion is extracted and used in the SI approach to update the SI-interpreted BOA.

Regional and geologic setting

The study area lies in the panhandle of western Nebraska, a semiarid region consisting of bedrock tablelands bisected by the North Platte River Valley (NPRV). The NPRV is a relatively broad valley (locally ~15 km wide) with riparian bottomlands near the modern river channel and a series of low-relief terraces extending toward the tablelands. Our study focuses on the Morrill block, a rectangular region (5 × 30 km) of closely spaced (200 to 400 m) AEM flight lines that create a particularly high-resolution data set (Figure 1). The Morrill block encompasses the modern North Platte River and extends north through the NPRV and into the bedrock tablelands. The local shallow stratigraphy of the NPRV consists of Quaternary alluvial, colluvial, and eolian deposits of the modern and historical North Platte River ranging in grain-size distribution from silt and clay to sand and gravel (Table 1). These deposits form a productive unconfined aquifer that can be up to 60 m in total thickness in parts of the NPRV. Resistivity values vary widely within the unconsolidated surficial deposits, between 30 and >100 ohm-m (Ball et al., 2006; Abraham et al., 2012). These surficial deposits overlie the Tertiary Brule Formation, a massive siltstone unit rich in weathered vol-

canic glass that tends to be consistently conductive between 5 and 10 ohm-m (Ball et al., 2006; Abraham et al., 2012). Locally, the Brule Formation behaves as an aquitard and acts as the base of the surficial aquifer, although fractured zones and sand and gravel lenses within the Brule can serve as aquifers (Steele et al., 2002; Cannia et al., 2006). The Brule surface is incised by a few buried paleochannels. These paleochannels host localized groundwater storage zones and act as preferential flow paths between recharge locations along major canals that distribute irrigation water throughout the NPRV and the modern North Platte River. In the tablelands, the Tertiary sandstones of the Arikaree Group directly overlie the Brule Formation. These relatively high-resistivity units are typically considered as aquifers.

Data

Several AEM surveys have been conducted in the Nebraska panhandle to support regional groundwater modeling efforts and for the development of hydrogeologic frameworks (Smith et al., 2009, 2010). In 2008 and 2009, the Morrill block was surveyed using the Fugro Airborne Services RESOLVE (any use of trade, firm, or product names is for

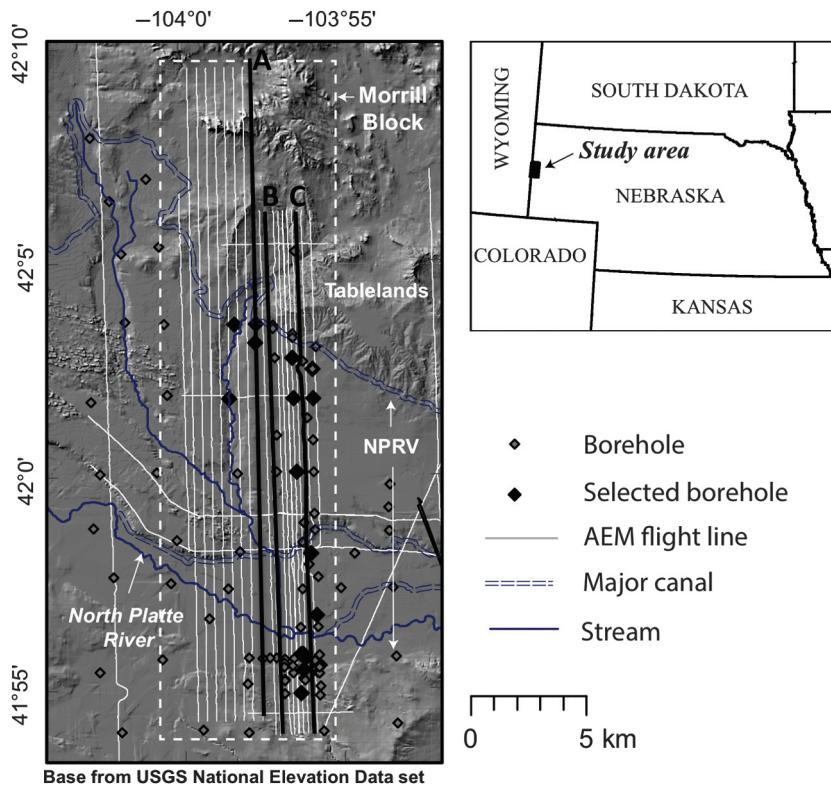


Figure 1. The Morrill block study area. The black diamonds show the selected boreholes among all available boreholes in the area (gray diamonds) used as input in the SI method. The highlighted lines A, B, and C represent the AEM flight lines 10080, 30020, and 30080, respectively, and are shown as cross sections in Figures 4, 6, and 7.

descriptive purposes only and does not imply endorsement by the U.S. government) frequency-domain electromagnetic (EM) helicopter-borne system. One-dimensional deterministic inversions of resistivity structure have been previously developed using EM1DFM (Farquharson et al., 2003) and are described in detail by Smith et al. (2010). These inversions used a DOI calculation similar to that described by Oldenburg and Li (1999) to mask poorly constrained portions of the resistivity models. This calculation compared multiple inversions using different reference models, in which the modeled resistivity values favor the reference model and substantially differ between inversions, and the models are considered to be relatively insensitive to the AEM data and therefore are removed from the final published inversion results (Smith et al., 2010).

Abraham et al. (2012) manually interpret the BOA throughout western Nebraska using the inverted AEM resistivity sections from multiple regional reconnaissance-style and block-style AEM surveys, including the Morrill block. This time-intensive manual process compared AEM sections with surface geologic maps, surface geophysical data, and borehole lithologic descriptions to determine the typical resistivity signature associated with the contact between the surficial aquifer and the underlying Brule Formation, a contact typically characterized as the transition from relatively resistive surficial materials to the approximately 10 ohm-m conductor typically found at depth. Using this information with expert geologic knowledge of the region and geophysical expertise, Abraham et al. (2012) manually pick the top of the relatively conductive Brule Formation surface along each inverted AEM section. Where the top of the conductive Brule Formation did not occur within the DOI, the interpretation was extended into the subsurface using the general geologic trends of the surrounding bedrock topography, with some guidance from the sub-DOI resistivity structure with the understanding that the resistivity models were poorly constrained in these areas. Through this process, Abraham et al. (2012) develop a data set defining the BOA along all flight lines with a typical spatial resolution of a BOA pick roughly every 200 m. This expert-interpreted BOA data set provides a robust independent test of the SI approach that is presented in

this paper. As part of this study, the frequency-domain EM data have also been inverted using a transdimensional Bayesian Monte Carlo approach (Minsley, 2011) that allows for a robust uncertainty analysis of the same data, and it will be used as an additional source of information about the BOA.

Smart interpretation

SI is an attribute-guided regression method that consists of two phases: the *learning* phase and the *predicting* phase (Gulbrandsen et al., 2017). In the first phase, we learn the relation between the geologic target and the available attribute data: in this case, the depth of the BOA and a set of attributes that are extracted from the inverted AEM data. In general, any kind of quantifiable information can be used as attributes. The attributes used in this study are listed in Table 2. The second phase of the method uses the relation learned in the first phase to predict the depth of the geologic target wherever attribute data exist, i.e., the depth of the BOA at every location where AEM data are present. The SI method aims at inferring a statistical model $f(\mathbf{d}_{\text{pred}} | \mathbf{M}_{\text{pred}})$, which is a probability distribution of a set of predicted interpretation points given a set of attributes. To infer such a model, we first need to learn the relation $h(\mathbf{M})$ between the interpretation points \mathbf{d} and

Table 2. For each 1D sounding from the deterministic inversion of the AEM data, the 54 attributes listed below are extracted and constitute the attribute vector \mathbf{M} .

Type of attributes	Number of attributes for each sounding
Terrain	1
Resistivity in each of the 17 layers	17
Depth to the top of each of the 17 layers	17
Depth to the base of each of the 17 layers	17
Geographical coordinates	2
Total number of attributes	54

Table 1. Summary of major hydrogeologic units within the Morrill block and the AEM DOI. Modified from Swinehart et al. (1985), Verstraeten et al. (2001), and Abraham et al. (2012).

System	Geologic unit	Description	Location	Geophysical character
Quaternary	Alluvial, colluvial, and eolian deposits	Undifferentiated sand and gravel deposits with lenses of silt and clay	Surficial unit in NPRV, typically thin to absent in tablelands	Locally high resistivity (30 to >100 ohm-m)
Tertiary	Arikaree Group	Silty, very fine- to medium-grained sandstone rich in volcanic glass	Surficial unit in tablelands, frequently absent in NPRV	Locally high resistivity (30 to >100 ohm-m)
	Brule Formation of the White River Group	Massive siltstone and mudstone rich in volcanic glass and weathered ash	Locally acts as a confining unit and BOA throughout the study area	Locally low resistivity (<10 ohm-m)

the set of attributes \mathbf{M} , such that $\mathbf{d} = h(\mathbf{M})$. Estimating the function $h(\mathbf{M})$ given a set of \mathbf{M} and \mathbf{d} is essentially a supervised machine learning problem. There are several different techniques that can be used to estimate the function $h(\mathbf{M})$; however, in the SI method, $h(\mathbf{M})$ is computed based on the assumption that the potentially nonlinear relation between \mathbf{d} and \mathbf{M} for one interpretation point can be defined as

$$d_i = \sum_{j=1}^M \sum_{p=0}^P m_{ij}^p g_{jp}, \quad (1)$$

where p is the polynomial order of the relation between d_i and the M different attributes and g_{jp} is the regression coefficient weighting the p order polynomial of the j th attribute. Note that boldface letters denote vectors, boldface capital letters denote matrices, and nonboldface letters denote scalars. This way of computing the function $h(\mathbf{M})$ is known as a polynomial linear regression technique (Bishop, 2006). If assuming that the interpretation points \mathbf{d} and the prior expectation of the regression coefficients are associated with Gaussian uncertainty, $N(0, \mathbf{C}_d)$ and $N(\mathbf{g}_0, \mathbf{C}_g)$, respectively, the probability distribution over \mathbf{g} , i.e., $f(\mathbf{g}|\mathbf{d}, \mathbf{M})$, can be computed as

$$f(\mathbf{g}|\mathbf{d}, \mathbf{M}) = k * \exp\left(-\frac{1}{2}(\mathbf{g} - \tilde{\mathbf{g}})^T \tilde{\mathbf{C}}_g^{-1} (\mathbf{g} - \tilde{\mathbf{g}})\right), \quad (2)$$

where k is a constant, $\tilde{\mathbf{C}}_g^{-1}$ is the covariance matrix representing the Gaussian uncertainty on the regression coefficients and $\tilde{\mathbf{g}}$ is the mean and can be computed as

$$\tilde{\mathbf{g}} = \mathbf{g}_0 + (\mathbf{M}^T \mathbf{C}_d^{-1} \mathbf{M} + \mathbf{C}_g^{-1})^{-1} \mathbf{M}^T \mathbf{C}_d^{-1} (\mathbf{d} - \mathbf{M}\mathbf{g}_0). \quad (3)$$

Having estimated $\tilde{\mathbf{g}}$ (representing $h(\mathbf{M})$), the predicted interpretation points can be computed as

$$\mathbf{d}_{\text{pred}} = \mathbf{M}_{\text{pred}} \tilde{\mathbf{g}}, \quad (4)$$

where \mathbf{M}_{pred} represent the attributes where the prediction is made. For details on how to estimate \mathbf{C}_g , see Gulbrandsen et al. (2017).

SI in this study

SI was initially developed as a tool to be used by a geologist when making geologic layer models. A geologist would interpret some geologic contact by manually picking this contact based on their conceptual geologic understanding of the region, along with geophysical models, AEM data, and borehole data. The SI algorithm will then learn the relation between the geologist's manually picked points and the attribute data extracted from the inverted AEM data and will predict the geologic contact wherever these attribute data exist. The SI algorithm is computationally efficient, and the predicted geologic contact will be updated immediately as the geologist adds new interpretation points. In this

study, the SI method is used as a one-time automatic computation based on existing data, without any manual interpretation. A set of selected borehole logs is used as input points. The available borehole data within the Morrill block area consist of 73 wells with lithologic descriptions (Souders and Swinehart, 2000). A subset of these wells was selected to include only wells within 150 m of an AEM flight line that intersected the Brule Formation, which is defined by the shallowest depth where consolidated rocks (e.g., siltstone, shale, mudstone) were encountered (Figure 1). The resulting well subset provided 19 guide points to initiate/teach the SI-interpretation to interpret the BOA/Brule Formation contact. The final borehole data points show a spatial bias toward locations in the NPPR; no boreholes intersecting the Brule Formation in close proximity to the AEM data were available in the tablelands. As attributes, the resistivity of each inversion model layer above the minimum DOI for the Morrill block (17 layers), the upper and lower depth of each of these model layers, the terrain, and the geographic coordinates were also used. Mathematically, the SI method requires the same number of attributes everywhere, including a fixed number of resistivity model layers at every location, and as such, only resistivity values above the minimum DOI of the whole Morill block were used by the SI algorithm instead of the full but variable DOI in every location.

In this study, the resulting smart interpreted BOA map is automatically generated only using existing data. However, if this initial map does not agree with the expert geologic interpretation, the geologist can tune the map by using SI in a semiautomatic way by supplementing the borehole-defined guide points with manually interpreted points as proposed by Gulbrandsen et al. (2017).

Results

Comparison of manual and SI interpretations

Figures 2 and 3 show comparisons between the SI and manual interpretation approaches. The surface indicating the BOA resulting from the borehole-guided SI method (Figure 2b) is shown together with the previously published manual interpretation of the BOA (Abraham et al., 2012) (Figure 2a), and the difference between the manual- and SI-interpreted surfaces (Figure 2c). The automatic approach, as seen in Figure 2, clearly provides results that overall are similar to the manual picking. Figure 3 shows the BOA interpretation in the context of the deterministically inverted resistivity sections along three flight lines. The red dots represent the smart picks, whereas the blue crosses represent the BOA-contact picks manually interpreted by the geologist. The two black dots on the uppermost profile show the depth to the Brule Formation extracted from two well logs situated along this flight line. No other well logs are available for comparison along any of these three flight lines. Note that the uppermost plot is displayed from the opposite side as the two others. The upper, middle, and lower sections are highlighted as AEM flight lines A, B, and C in Figure 1, respectively. Only resistivity layers

above the minimum DOI are shown. The overall high agreement between the SI interpretations and the manually interpreted points is verified by Figure 4, which shows a histogram of the differences between the two maps of the BOA. The statistics tell that 68% (one standard deviation) of the picks are within 12 m of the manual picks, with a mean value equal to 0.275, indicating no bias in under- or over-prediction of the BOA. Alternately, it seems that the distribution of differences in BOA estimates is long tailed, indicating that a few automated picks are very far from the manual picks. The reason for this is, in part, that the resistivity data are not all of the same quality. Some resistivity models show poor data misfits and unreasonable changes in structure that indicate infrastructure coupling, such as near power-lines. These poor-quality data do not represent the geology, and, as such, SI picks using these data will be far off due to the misleading resistivity values. Examples of this are seen in Figure 3b and 3c (at distances of 0.9×10^4 and 1.8×10^4 m, respectively).

Even though Figure 2 illustrates an overall reasonable agreement between the two BOA maps, some differences are apparent as well. The largest differences not associated with cultural noise between the manual and the SI-picked BOA are located in three regions, namely, the northern part of the area correlating with the bedrock tablelands and in two channel structures within the NPrV: a narrow channel stretching from the northwest to the southeast parallel to the northernmost major canal, and a wider channel to the south close to the modern North Platte River (see Figure 2c). In the northern tablelands, the output from the SI indicates a deeper BOA than the geologist's picks, and in the two-channel structures, the geologist interpreted the BOA to be deeper than the smart interpreter. These differences are also illustrated in Figure 3. In the northern part of the study area (to the right in Figure 3a and to the left in Figure 3b and 3c), the BOA is interpreted to be deeper by SI than by the geologist. In contrast, SI picks the channel structures in the middle and to the

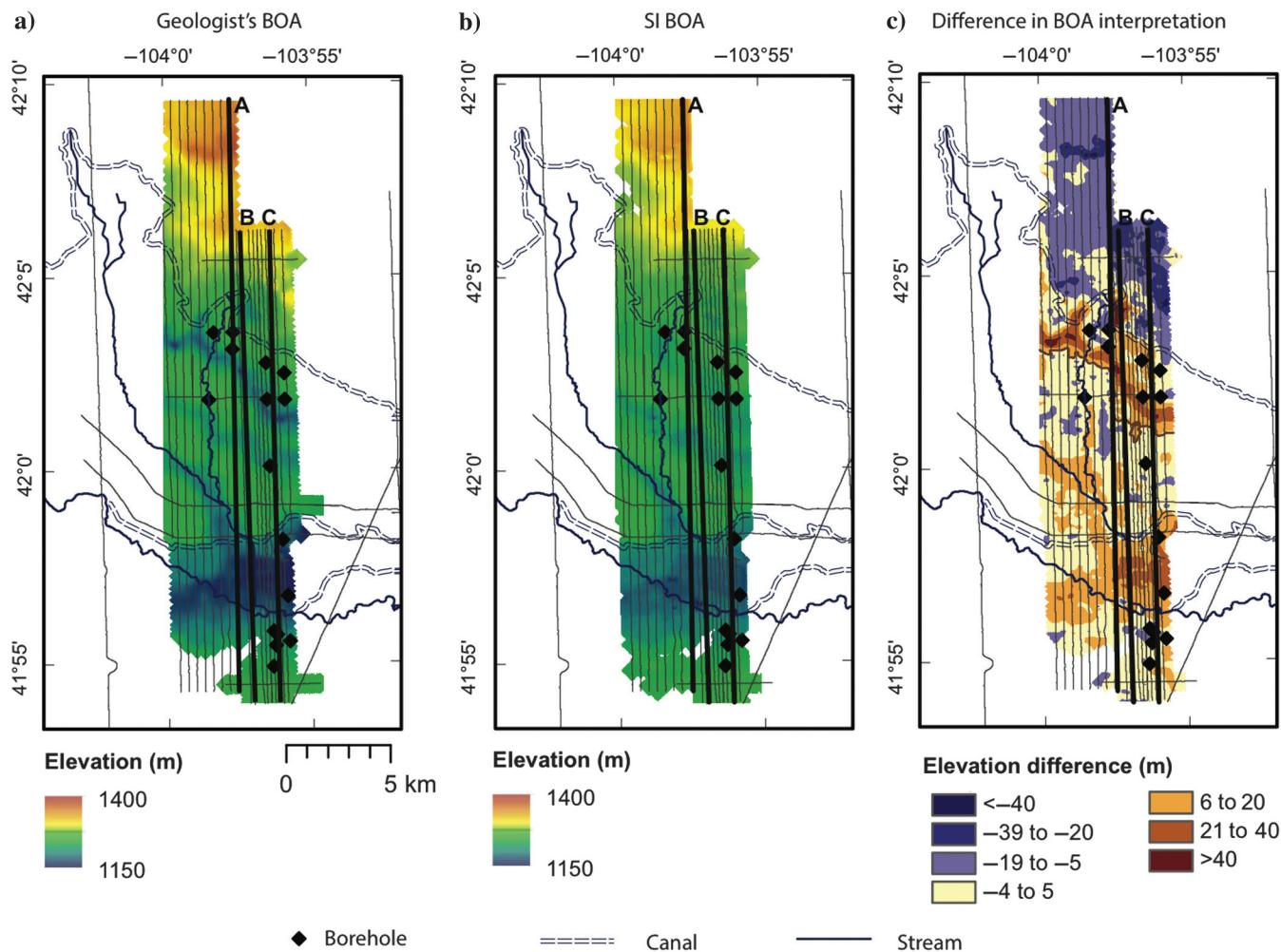


Figure 2. (a) The interpreted BOA resulting from a manual approach (Abraham et al., 2012) is plotted together with the results from (b) the SI method using selected well log data (black diamonds) and the difference between the two (c); positive values indicate that the SI interpretation is shallower than the geologist's interpretation. The highlighted lines A, B, and C represent the AEM flight lines 10080, 30020, and 30080, respectively.

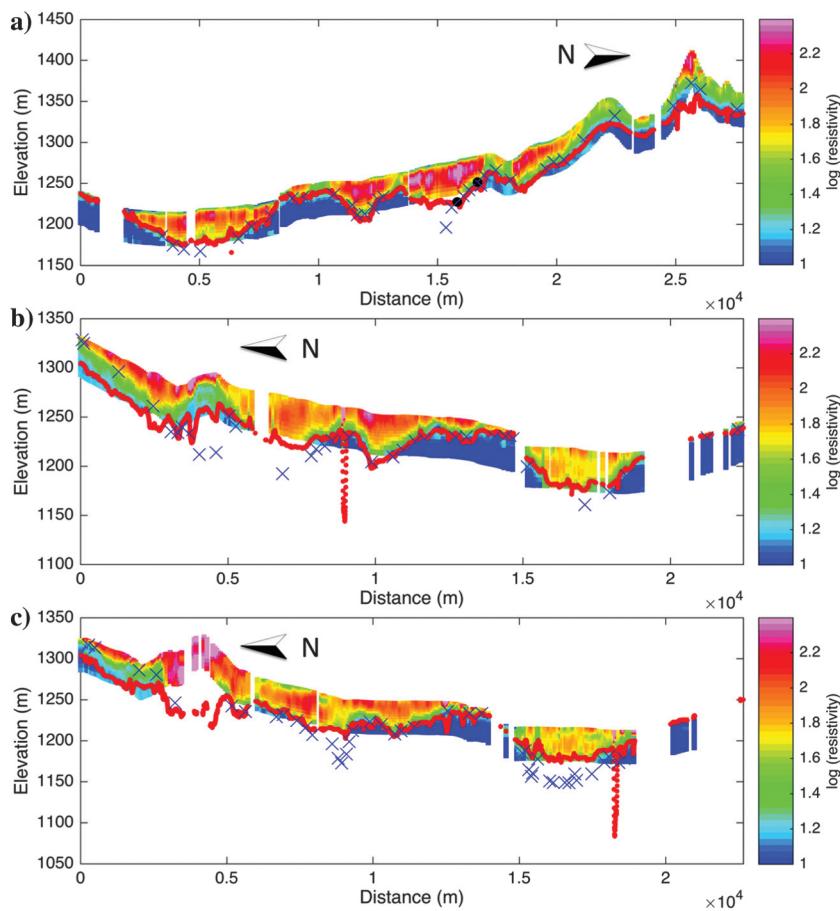


Figure 3. Three cross sections showing deterministically inverted resistivity values (represented by the color bar) from the Morrill survey. The red dots are the automated picks from the SI method, the blue crosses are the manually interpreted picks (Abraham et al., 2012), and the two black dots on the uppermost plot are the depth to the Brule Formation extracted from two well logs. (a-c) The cross sections represent the AEM flight lines A, B, and C in Figure 1, respectively.

south higher than the geologist chose to interpret these features. The middle channel seems to be “cut off” on all three profiles by the smart interpreter, whereas the differences on the channel to the south increase farther east.

Using McMC inversion results

Inverting the EM data using a transdimensional Bayesian McMC approach (Minsley, 2011) provides an estimate of the model a posteriori distribution, consisting of a large number (100,000) of 1D models with different numbers of layers, with a constant resistivity within each layer. There is one such a posteriori sample for every 10th observed sounding. From these model ensembles, a variety of statistics related to parameter uncertainty can be computed. In the present context, the BOA is expected to be associated with a high vertical change in resistivity (from the relatively resistive aquifer to the relatively conductive Brule Formation). One way to quantify the existence of such a high vertical change in resistivity is to compute the a posteriori probability that resistivity changes by more than 50% over a layer boundary. This interface probability is obtained by computing the frequency of layer interfaces in which a resistivity transition of greater than 50% occurs at every depth in the model. The results are presented as a grayscale image, in which light colors represent a low probability of locating a geologic contact with a high vertical resistivity contrast and darker colors represent higher probability.

Figure 5 shows the three profiles from Figure 3 displayed together with the SI predictions plotted on top of the grayscale plot of the McMC results for contact probability. There is high consistency among the output from the SI, manual interpretations derived from the deterministic inversion, and the contact probability from the McMC inversion. The largest inconsistencies are seen in the northern part of the survey and the deep valley structure in the southeast. The McMC results strongly indicate that there is a geologic contact with high certainty approximately 20 m above the SI predictions in the northern region, but it indicates with greater uncertainty (less sharp boundary) that the valley to the southeast is predicted too shallow by the SI (see the third section of Figure 5). In these regions, the McMC results agree with the interpretations made manually. Because the McMC results represent the probability of having an abrupt vertical resistivity contrast, and given that the BOA is defined as the contact between the relatively resistive aquifer and the more conductive Brule Forma-

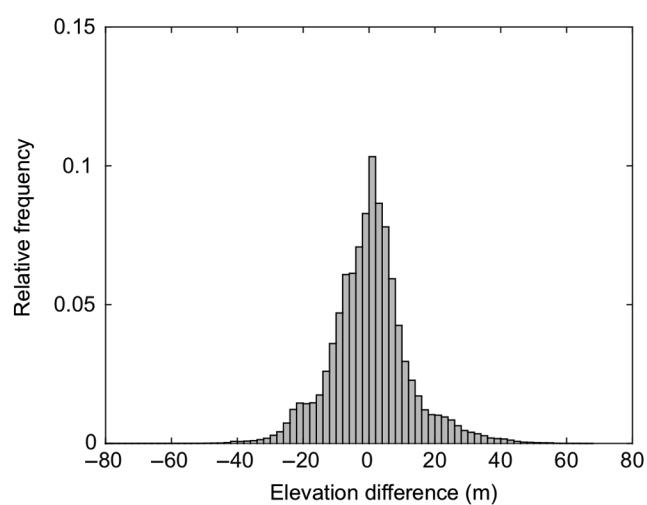


Figure 4. Histogram of the difference between the interpretations from the SI method and the manual interpretation approach.

tion, it is reasonable that the BOA contact will be represented in the McMC results. The correlation between expert interpretation and McMC results supports the use of McMC information and the geophysical uncertainty associated with the interfaces in interpreting the depth to the BOA. One way to do this is by adding the interface probability information directly into the SI

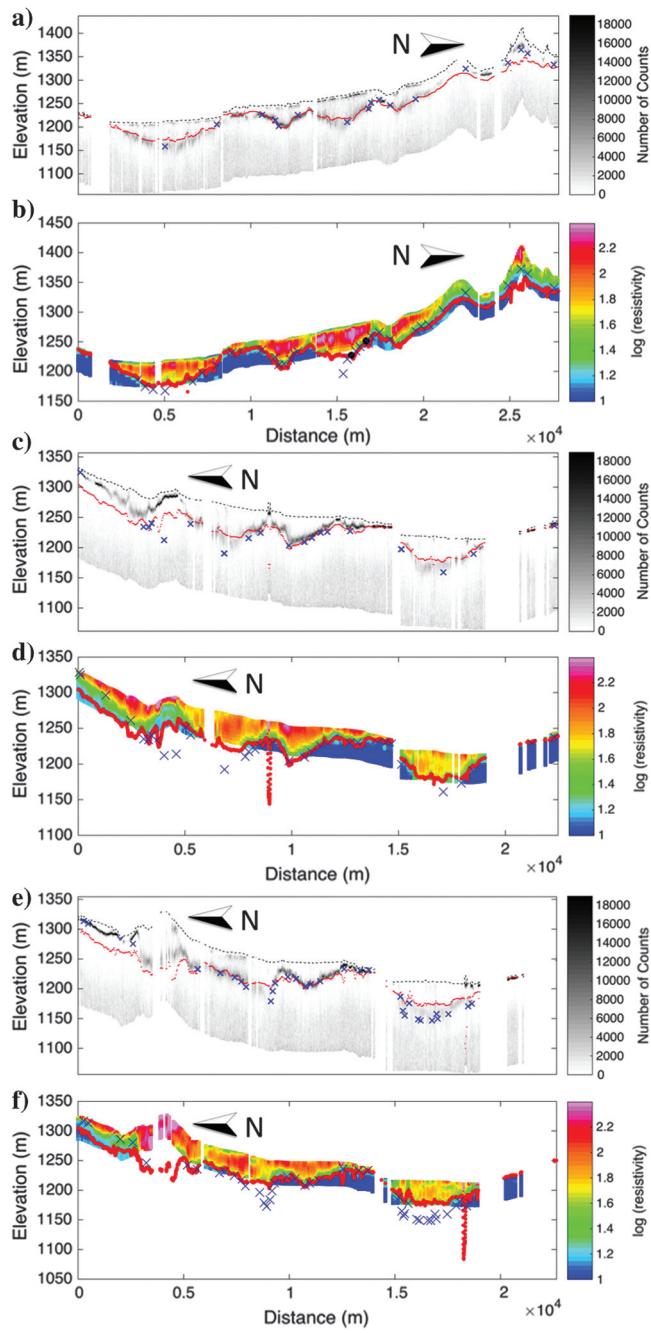


Figure 5. (a-f) The results from Figure 4 are displayed together with the SI predictions (red dots) plotted together with the grayscale plot resulting from the McMC inversions. The blue crosses represent the manual interpretations, the grayscale plots represent the number of times (of a total of 100,000 a posteriori realizations) resistivity changes more than 50% with depth, and the two black dots on the uppermost plot are the depth to the Brule Formation extracted from these two well logs.

approach. Figure 6 shows the same results as in Figure 5; however, now the predictions of the BOA from the SI method are made with two additional input points derived from the McMC contact probability information. These points are manually added by analyzing the McMC results and used by the SI as additional input

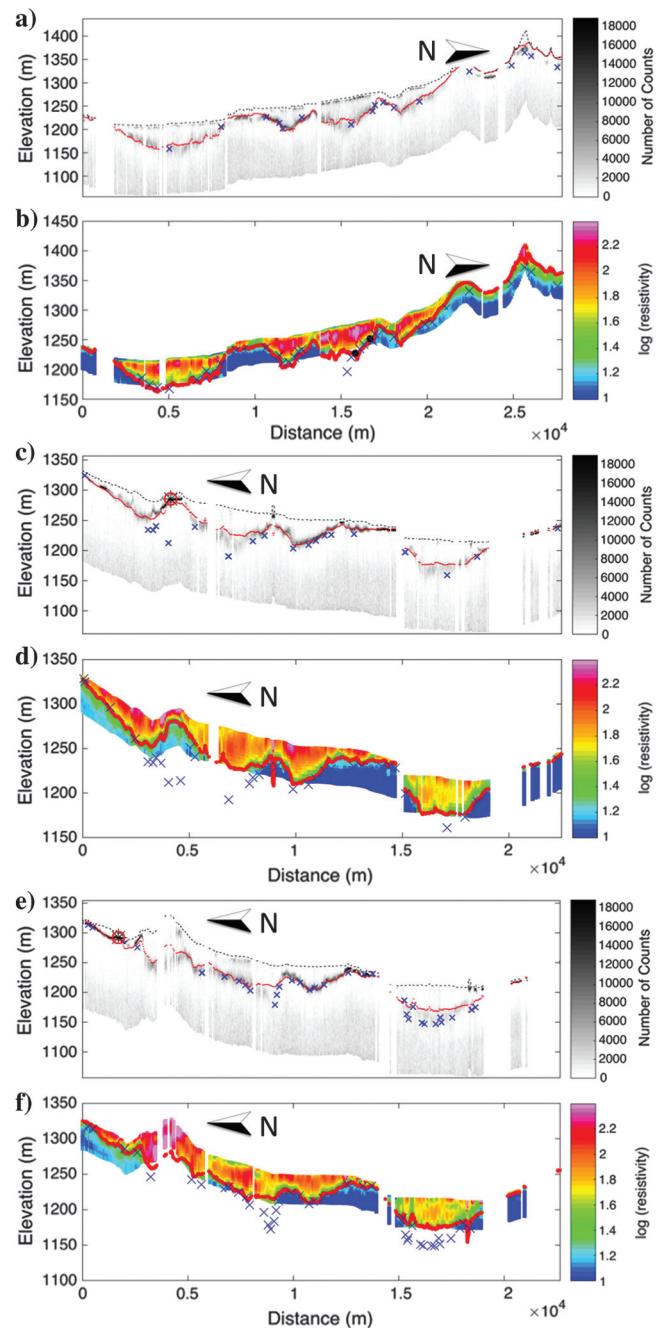


Figure 6. (a-f) The predictions from the SI method (red dots) are computed here with an additional two input points in the learning phase of the method. The black asterisks with a red circle indicate these additional points. The blue crosses represent the manual interpretations, the grayscale plots represent the number of times (of a total of 100,000 a posteriori realizations) resistivity changes more than 50% with depth, and the two black dots on the uppermost plot are the depth to the Brule Formation extracted from these two well logs.

points that augment the original 19 borehole-defined points. The added input points are marked by black asterisks and are located on the second and third cross sections in Figure 6. These points were chosen because they represent parts of the cross sections in which the disagreement between the SI predictions and the contact estimations from the McMC method is high, in addition to the high probability that these points represent a geologic contact (i.e., a well-defined boundary in the grayscale plot). Figure 6 shows that by adding the two extra input points to the smart interpreter, the output predictions become even more consistent with the McMC results and ideally become a better representation of the BOA. The final map of BOA of the Morrill block is displayed in Figure 7b, with the geologically interpreted BOA based on the manual approach as in Figure 7a, and the difference (with respect to the SI BOA) in Figure 7c. The final SI map of BOA after adding the extra information about the geologic contact extracted from the McMC inversion and the manual inter-

pretation of the BOA are more similar. This is also illustrated by the statistics of the differences between the maps (Figure 8). However, there are still some inconsistencies, the most prominent of which is the paleochannel in the middle of the survey, in which the SI approach still interprets the channel to be shallower than the geologist does. Although incorporation of the McMC results has helped to reduce overestimates of the BOA depth in the northern tablelands, some significant underestimates remain in which the depth of the manual pick is greater than the DOI in the resistivity section. This result is manifested as an asymmetry in the differences shown in Figure 8, compared with the original result in Figure 4.

These results demonstrate that AEM data inverted using the probabilistic McMC approach provide results that can be visualized in a way that is natural for use as part of modeling geologic contacts. Even though such an McMC inversion is more computationally expensive than a deterministic inversion, it is manageable for a 3D

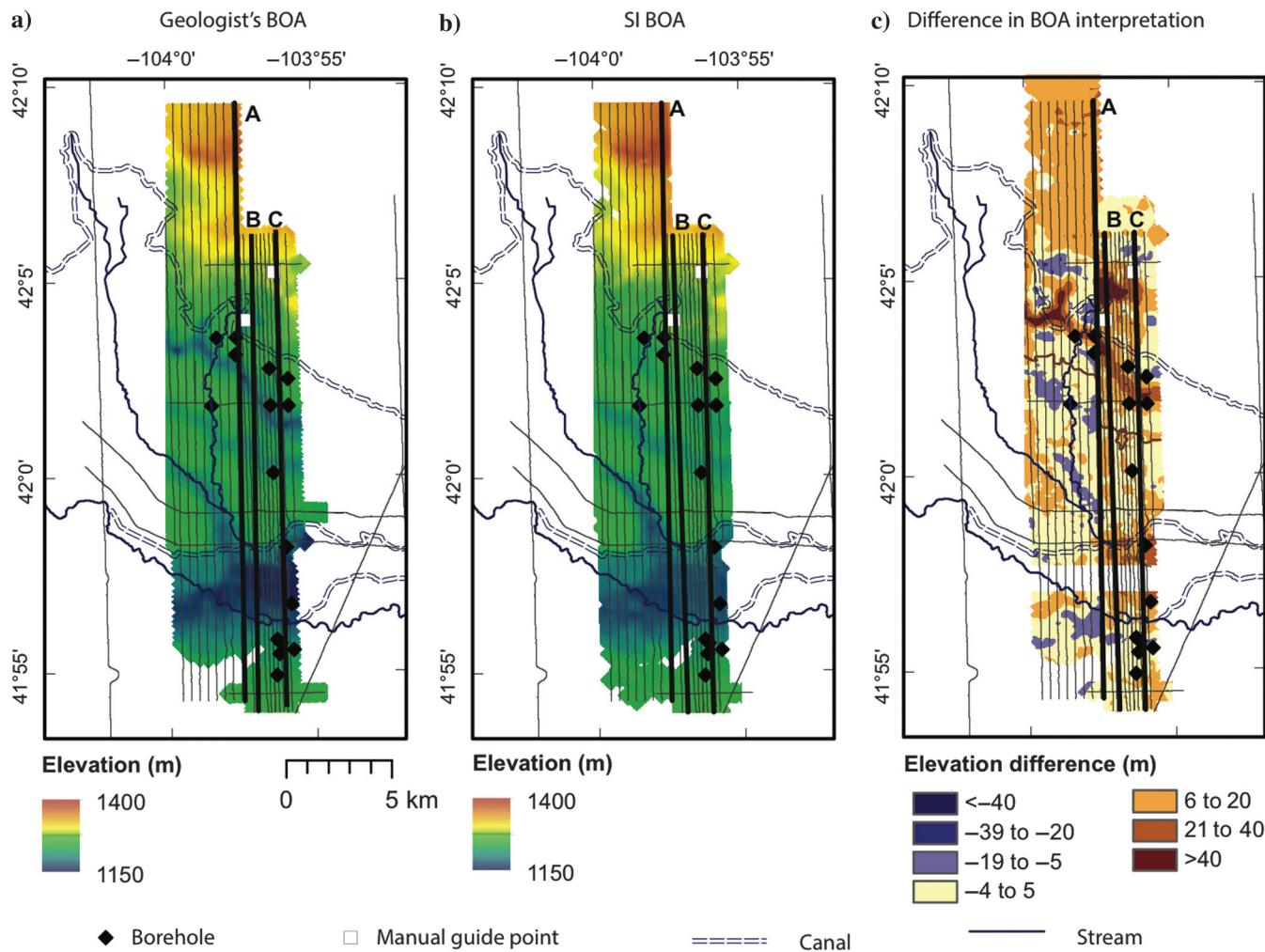


Figure 7. (a) The interpreted BOA resulting from the SI method using well log data (black diamonds) and (b) two additional points extracted from the McMC inversion (white squares) is plotted together with the results from a manual approach, and (c) the difference between the two. The highlighted lines A, B, and C represent the AEM flight lines 10080, 30020, and 30080, respectively.

hydrogeologic setting considered here. Given the utility of the McMC information, one could consider applying such McMC methods on an even coarser grid than the one presented here. Based on these results, inverting the AEM data probabilistically to include in final interpretation of geologic contacts is highly recommended.

Discussion

There might be several reasons why the manual interpretation and the automatic predictions differ. However, two main categories of why these two methods differ and stand out are (1) the training data, i.e., the borehole logs, do not consistently sample all geologic settings and (2) the DOI does not always extend to the depth necessary to fully capture the BOA. This potential lack of information from the two sources of data (well logs and geophysics) will be reflected in the outputs of the two different approaches. SI can be seen as an unbiased approach to quantify the relation between the AEM data and BOA position. It performs only as well as the information available. On the other hand, the geologist has conceptual knowledge about geologic structures based on the information they have beyond the local study area, information which is not explicitly implemented in the SI method, and hence might be absent if not directly available in the data. In this study, these differences might be the reason why the two approaches differ in certain areas of the survey. If the geologist has additional information, which is not represented in the boreholes or the geophysical data, such as the typical shape or depth of buried valleys throughout the region, then the two maps can differ. For instance, the deep valley structure in the middle of the survey might not have been picked up by the SI approach due to the lack of data. Neither well logs nor the AEM data (above DOI) provide information at these depths. However, the geologist may have knowledge

about the depth of typical paleovalleys in these regions and might be able to interpret how these valleys are most likely to be configured in the absence of local data. For example, this can be interpreted based on how the geologist interprets the beginning of the valley structure, and hence has knowledge of how it should continue, to keep the most plausible geologic shape. An example seen in this study is that the geologist interpreted the paleovalley in the southern part of the survey to get deeper farther east, a trend that is not picked by the SI approach.

The overall accuracy of the interpretations from SI, using the manual interpretation as a reference, is in general good. A geologic expert should be able to deduce whether the SI model is accurate enough as is, or whether it should be further adjusted. In any case, based on the accuracy of this automatically generated starting point, the geologist should be able to reach a better geologic model, much faster than without using this workflow.

The current discussion is comparing the SI-interpreted BOA with respect to a previous study in which the BOA has been manually interpreted. Even though this study has been used as a reference to discuss the interpretation from the automatic approach, the manual picks do not necessarily represent the true BOA. If the automated map was given to the geologist before he/she started picking, he or she might have chosen the BOA differently. For instance, the geologist might have picked the BOA a bit deeper in the northern region if the SI results would have been available. This is, however, a region of high uncertainty of the results and is difficult both for the geologist and for SI due to the absence of boreholes. In the study area, the majority of well drilling has occurred in the unconsolidated Quaternary sediments underlying the NDRV. In the tablelands north of the canals, the Arikaree Group forms the local aquifer above the Brule Formation. The resistivity signature of the BOA changes between these two settings, but there are no borehole data to train the SI algorithm to recognize this contact. This presents a problem for the SI algorithm as well as for the geologist because both lack ground control for their picks. However, the geologist could use available supplemental information such as distant borehole logs, general stratigraphic thicknesses, or outcrop data to help make a decision.

Conclusions

In this paper, we show that the smart interpretation method provides an automated, fast, and objective map of the BOA by combining borehole information with a set of geophysical attributes extracted from deterministically inverted AEM data. Additional information about the BOA is also extracted from an McMC inversion of the same AEM data that are used in the SI method to update the initial BOA map. The accuracy and quality of these results are verified by comparing them with an earlier manual mapping of the same geologic contact. The automatically generated map of the BOA is intended to assist the geologist in the interpretation process by providing

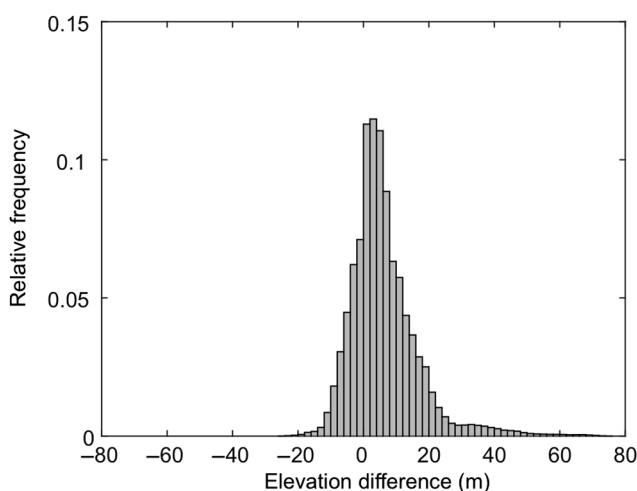


Figure 8. Histogram of the difference between the interpretations from the SI method, including the two additional points from extracted from the McMC inversion, and the manual interpretation approach.

as much information from the geophysical data as possible and through consistently interpreting the large quantities of geophysical information. The strong correlations between the SI and expert-interpreted surface in the Morill block illustrate the general utility of this approach. The results of the SI approach are intended to be integrated into the geologic modeling process by providing a high-resolution starting model for the geologist that can be evaluated and edited, instead of manually developing a new model with limited guidance. The efficiency of the method and the accurate results suggest that the geologist should be able to more quickly generate geologic models that are consistent with the borehole lithologic logs, geophysical data, and geologic expert knowledge.

Acknowledgments

This study is partly financed by the Danish High Technology foundation (journal number: 113-2013-1), the University of Copenhagen, and the U.S. Geological Survey Frontier Geophysical Methods and Applications project. As far as we understand, there is no potential conflict of interest.

References

- Abraham, J. D., J. C. Cannia, P. A. Bedrosian, M. R. Johnson, L. B. Ball, and S. S. Sibray, 2012, Airborne electromagnetic mapping of the base of aquifer in areas of western Nebraska: U.S. Geological Survey Scientific Investigations Report 2011-5219.
- Auken, E., S. Chandra, G. Vignoli, A. Shankeel, M. K. Sen, and S. Gupta, 2013, Large scale mapping of groundwater resources in India with results from test sites in different geological terrain: 13th SAGA Biennial Conference & Exhibition October 2013, Conference Contribution.
- Ball, L. B., H. K. Wade, V. S. Gregory, J. C. Cannia, and M. J. Andersen, 2006, Determination of canal leakage potential using continuous resistivity profiling techniques, interstate and tri-state canals, Western Nebraska and Eastern Wyoming, 2004: U.S. Geological Survey Scientific Investigations Report 2006-5032.
- Bishop, C. M., 2006, Pattern recognition and machine learning: Springer.
- Cannia, J. C., D. Woodward, and L. D. Cast, 2006, Cooperative hydrologic study cohyst, hydrostratigraphic units and aquifer characterization report, http://cohyst.dnr.ne.gov/document/dc012hydro_aquifer_022406.pdf, accessed 2 June 2009.
- Cassidy, K. F., M. T. Costelloe, K. Czarnota, P. Duerden, D. K. Hutchinson, D. Huston, S. Jaireth, S. F. Liu, D. Maidment, D. Miggins, I. C. Roach, C. Sorensen, A. J. Whitaker, J. R. Wilford, and N. C. Williams, 2010/12, Geological and energy implications of the Paterson Province airborne electromagnetic (AEM) survey, Western Australia: Record-Geoscience Australia.
- Davis, A., S. Macaulay, T. Munday, C. Sorensen, J. Shudra, and T. Ibrahim, 2016, Uncovering the groundwater resource potential of Murchison Region in Western Australia through targeted application of airborne electromagnetics: 25th International Geophysical Conference and Exhibition, ASEG-PESA-AIG, 459–464.
- Farquharson, C. G., D. W. Oldenburg, and P. S. Routh, 2003, Simultaneous 1D Inversion of loop-loop electromagnetic data for magnetic susceptibility and electrical conductivity: *Geophysics*, **68**, 1857–1869, doi: [10.1190/1.1635038](https://doi.org/10.1190/1.1635038).
- Gulbrandsen, M. L., K. S. Cordua, T. Bach, and T. M. Hansen, 2017, Smart Interpretation — Automatic geological interpretations based on supervised statistical models: *Computational Geosciences*, doi: [10.1007/s10596-017-9621-8](https://doi.org/10.1007/s10596-017-9621-8).
- Hodgson, J. A., and M. D. Ture, 2015, Processing and data merging report of Tellus airborne geophysical datasets, 2012–2015: Technical Report, Geological Survey of Ireland.
- Mielby, S., 2010, Den nationale grundvandskortlægning i Danmark — faglige resultater fra 2010: Technical Report, Geological survey of Denmark and Greenland.
- Minsley, B. J., 2011, A trans-dimensional Bayesian Markov chain Monte Carlo algorithm for model assessment using frequency-domain electromagnetic data: *Geophysical Journal International*, **187**, 252–272, doi: [10.1111/j.1365-246X.2011.05165.x](https://doi.org/10.1111/j.1365-246X.2011.05165.x).
- Oldenburg, D. W., and Y. Li, 1999, Estimating depth of investigation in DC resistivity and IP surveys: *Geophysics*, **64**, 403–416, doi: [10.1190/1.1444545](https://doi.org/10.1190/1.1444545).
- Sandersen, R. T. P., and B. Pjetursson, 2014, The integrated ground water mapping concept of Denmark transferred to Thailand: Presented at the 8th Annual Meeting DW RIP.
- Smith, B. D., J. D. Abraham, J. C. Cannia, and P. Hill, 2009, Helicopter electromagnetic and magnetic geophysical survey data for portions of the North Platte River and Lodgepole Creek, Nebraska, June 2008: U.S. Geological Survey Open-File Report 2009-1110.
- Smith, B. D., J. D. Abraham, J. C. Cannia, B. J. Minsley, M. Deszcz-Pan, and L. B. Ball, 2010, Helicopter electromagnetic and magnetic geophysical survey data, portions of the North Platte and South Platte natural resources districts, Western Nebraska, May 2009: U.S. Geological Survey Open-File Report 2010-1259.
- Souders, V. L., and J. B. Swinehart, 2000, Morrill County test-hole logs: Nebraska Water Survey Test — Hole Report No. 62, Conservation and Survey Division, Institute of Agriculture and Natural Resources, University of Nebraska-Lincoln, 106.
- Steele, G. V., C. J. Cannia, and K. G. Scripter, 2002, Hydrologic characteristics of selected alluvial aquifers in the North Platte natural resources district, Western Nebraska: U.S. Geological Survey Water-Resources Investigations Report 2001-4241.
- Swinehart, J. B., V. L. Souders, H. M. DeGraw, and R. F. Diffendal Jr., 1985, Cenozoic paleogeography of the western Nebraska, in R. M. Flores, and S. S. Kaplin, eds., Ceno-

zoic Paleogeography of West Central United States: Rocky Mountain Section, SPEM, 209–229.

Vertstraeten, I. M., G. V. Steele, J. C. Cannia, D. E. Hitch, K. G. Scripter, J. K. Bohlke, T. F. Kraemer, and J. S. Stanton, 2001, Interaction of surface water and ground water in the Dutch Flats area, western Nebraska,

1995–1999: U.S. Geological Survey Water-Resources Investigations Report 01-4070.

Biographies and photographs of the authors are not available.