Short Note on Required Data for SAGD Reservoir Characterization

Clayton V. Deutsch, Oy Leuangthong, and Weishan Ren

Centre for Computational Geostatistics
Department of Civil & Environmental Engineering
University of Alberta

Abstract

Delineation requirements for mining and petroleum projects are often expressed in the units of data spacing or the number of drillholes/wells per unit area. The rationale for more data is to reduce uncertainty. There is a need to relate uncertainty to data density. Geostatistical procedures are used to establish this relationship; however, geological formations are not homogeneous and the same data density will lead to different uncertainty in different areas of the deposit. Some areas are homogenous and there will be lesser uncertainty than very heterogeneous areas for the same data density. This paper describes some thoughts on relating uncertainty to data density for different geological heterogeneity domains, with particular focus on SAGD reservoirs. Heterogeneous domains may warrant more drilling than homogenous domains. This note also discusses a procedure to establish the required data density as a function of geological heterogeneity.

Introduction

Mining and oil sands deposits are delineated by a certain number of drillholes. It is common to report the data spacing or the data density (number of data per unit area), e.g.,

the drillhole spacing is 100m or *there are 16 wells per section*

It is less common to report uncertainty targets, e.g.,

the true grades of $(15m)^3$ blocks are within ±15% of the predicted grades with 90% certainty

Although uncertainty targets are desirable from a statistical perspective, there are many reasons to retain the data spacing or the number of drillholes per unit area: (1) the units are easily understood – everyone knows how many data were collected, (2) the values can be easily compared between deposits/sites, (3) the requirements are directly related to cost ($/m of drilling), and (4) there is less room for statistical mischief to artificially increase or decrease drilling requirements.

The data spacing is often chosen arbitrarily based on experience. Professionals are increasingly uncomfortable with a simple geometric specification; there is a desire to relate a quantifiable measure of uncertainty to the drillhole spacing. This is an apparent contradiction with the preceding paragraph; however, it reflects a modern reality where uncertainty must be quantified and understood.

The aim of this short note is to comment on how the relationship between data spacing and uncertainty is established and to describe procedures for specifying conventional drilling requirements on the basis of a quantified measure of uncertainty. The importance of the bitumen resource in northern Alberta to the national and global crude oil reserves provides part of the motivation for this note. Recent studies in SAGD reservoir characterization and other mining and petroleum related studies confirm the strong industry and regulatory interest in this area of research.

Uncertainty

Application of geostatistical simulation tools permits a quantification of uncertainty for any location in the deposit. The uncertainty at an unsampled location is fully characterized by the conditional cumulative distribution function (ccdf): $F(z;\mathbf{u}|(n))$, where z represents the variable of interest, **u** a particular unsampled location and *(n)* the set of nearby relevant conditioning information. Standard practice consists of dividing the deposit into locations of equal area or equal volume. These equal areas/volumes may be a convenient block size or they may relate to a given time period of production.

Uncertainty may be predicted directly under a multiGaussian or indicator-based assumption. It is becoming increasingly common to apply geostatistical simulation to create realizations of the deposit. Geostatistical simulation is typically applied in a hierarchical fashion, that is, large scale structure is simulated, then facies or geological units, then continuous petrophysical properties. The values may have to be scaled up to the chosen block size. The details of geostatistical procedures to infer the uncertainty $\{F(z;\mathbf{u}_i|(n)), i=1,\ldots,N\}$ at N multiple locations is not the subject of this paper. The previous three papers on SAGD characterization of the Athabasca formation describe procedures for this very purpose (McLennan and Deutsch, 2004; Ren, Leuangthong and Deutsch, 2004; Ren and Deutsch, 2004). Our focus is to relate uncertainty to data spacing with the aim of specifying required data spacing instead of uncertainty, which would be more understandable and reasonable to most people.

The distributions $\{F(z;\mathbf{u}_i|(n)), i=1,\ldots,N\}$ fully characterize uncertainty; however, different summary statistics are often used to quantify uncertainty. The local coefficient of variation is based on moments and is defined as:

$$CV(\mathbf{u}) = \frac{a \bullet \sigma(\mathbf{u})}{m(\mathbf{u})} \quad (1)$$

where $\sigma(\mathbf{u})$ is the local standard deviation, $m(\mathbf{u})$ is the local mean, and the value of a could be chosen as 1.28 so that $CV(\mathbf{u})$ could be interpreted as the deviation of the $P_{90}(\mathbf{u})$ or the $P_{10}(\mathbf{u})$ from the $P_{50}(\mathbf{u})$, if the distribution were standard normal. The measure $CV(\mathbf{u})$ does not account for the asymmetry of the local distributions and may be sensitive to extreme values. A non-parametric measure of local uncertainty based on selected quantiles could be defined:

$$U_1(\mathbf{u}) = \frac{1}{2}\left(\frac{P_{90}(\mathbf{u}) - P_{10}(\mathbf{u})}{P_{50}(\mathbf{u})}\right) \quad (2)$$

The statistic $U_1(\mathbf{u})$ is the average deviation of the $P_{90}(\mathbf{u})$ or the $P_{10}(\mathbf{u})$ from the $P_{50}(\mathbf{u})$, which would be the same as $CV(\mathbf{u})$ if a=1.28 and the distribution of uncertainty, $F(z;\mathbf{u}|(n))$, is non-standard Gaussian. The 80% P_{10} to P_{90} interval is arbitrary. Some practitioners prefer 90% or 95% probability intervals. The uncertain nature of geologic settings makes the 90/95% probability intervals quite large; the probability intervals can appear distressingly wide.

Another measure of uncertainty is the probability that the true value falls outside some specified fraction of the mean:

$$U_2(\mathbf{u}) = 1 - \Big(F\big((1+p)\bullet m(\mathbf{u});(n)\big) - F\big((1-p)\bullet m(\mathbf{u});(n)\big)\Big) \quad (3)$$

where p is the ± fraction from the local mean. This measure is consistent with the wording given above (…*true grade within 15% of the predicted*…). The probability of being outside the interval is used because it will have the same direction as $CV(\mathbf{u})$ and $U_1(\mathbf{u})$, that is, high $U_2(\mathbf{u})$ means high uncertainty and low $U_2(\mathbf{u})$ means low uncertainty. A low uncertainty of, say $U_2(\mathbf{u})<0.1$ entails that the probability of being inside the interval is above 0.9.

In general, absolute uncertainty measures such as $\sigma(\mathbf{u})$ or $P_{90}(\mathbf{u})$-$P_{10}(\mathbf{u})$ are not used because the uncertainty is a function of the predicted value, that is, there is a proportional effect where the uncertainty increases as the average value $m(\mathbf{u})$ increases.

Uncertainty is calculated by one of Equation (1), (2), or (3). The second non-parametric measure, $U_2(\mathbf{u})$, will be preferred if the wording of uncertainty target is expressed as a probability to be within a plus/minus tolerance of the predicted value. The first non-parametric measure of uncertainty $U_1(\mathbf{u})$ defined in (2) is robust and will be used as a default. Wherever possible, we will stay general and simply refer to *uncertainty*.

We will be relating uncertainty to data density. In many cases we will want an average or expected uncertainty over an area (or multiple areas) with similar data density and geologic heterogeneity. Multiple uncertainty values will be averaged when required.

Data Density

The data density is the number of data per unit area or volume. The complexity of the number of data per unit volume will be avoided since most deposits are sampled by vertical drilling through the entire extent. The data spacing d_s is often in the units of meters or feet. The data density d_d is often in the units of number per km^2 or per mi^2 (per section). The relationship between d_s in m and d_d in $\#/km^2$ is given by:

$$d_d = \left(\frac{1000}{d_s}\right)^2 \quad \text{and} \quad d_s = \frac{1000}{\sqrt{d_d}} \quad (4)$$

The relationship when d_s is in ft and d_d in $\#/mi^2$ is similar; just replace the 1000 m/km with 5280 ft/mile in (4). In practice the data spacing is often irregular. The `dhdens` program can be used to calculate the data density. The number of data within some specified radius is counted and the data density calculated directly. The results may be smoothed to avoid artifacts of irregular spacing and the search radius.

Uncertainty versus Data Density

Provided the statistical model does not change, uncertainty will decrease with additional data. The sketch on the left of Figure 1 shows how uncertainty decreases with increasing data density. The three curves indicate the behavior for three different models of geologic heterogeneity. It is disconcerting when additional data is collected and uncertainty increases, but that is the expected result when the deposit is found to be more complicated/heterogeneous than first thought. The sketch on the right of Figure 1 shows how the apparent uncertainty can increase when the available data (or the interpretation of the available staff) indicates that the deposit is more geologically complex. These jumps in uncertainty are less likely when the geological heterogeneity is realistically appraised early in delineation. It is difficult to plan for such increases. An important assumption of this work is that the uncertainty decreases monotonically with increasing data spacing.

In practice, there are two different ways to establish the quantitative relationship between uncertainty and data density:

1. When the area is sampled to varying degrees, the values of uncertainty from a detailed geostatistical study can be plotted against the drillhole density. This is ideal since no resampling and strong assumptions of stationarity are required; however, it does require a site that has areas of dense sampling and areas of sparse sampling.
2. In presence of few data or in presence of uniform data density, geostatistical simulation can be used. Data are sampled from a reference geostatistical realization, multiple realizations are constructed, and then uncertainty is calculated. This is repeated for different drillhole densities.

A number of examples of the first scenario have been considered where both the absolute and relative uncertainty is plotted against the well density. The relationship between relative uncertainty and well density appears better behaved and easily understandable due to the unitless measure. The Centre for Computational Geostatistics is committed to providing the best technical guidance on this subject; thorough case studies will be presented next year.

Geologic Heterogeneity

In presence of different facies or geological domains, there will be different relationships. More heterogeneous regions will require more drillholes for delineation within some acceptable uncertainty level; less heterogeneous areas will require fewer drillholes to achieve similar uncertainty targets.

One idea is to propose distinct regions that clearly have different relationships; these regions can be defined by a geological heterogeneity index (GHI). In the case of a single variable of interest, these regions can be based on that particular property. Note that in this case, the uncertainty measure and the GHI would be based on the same variable. In the multiple variables scenario, the easiest approach is to identify a property similar to that chosen for the uncertainty measure. For instance, in the case of an oil sands deposit, the uncertainty measure may be based on the net continuous bitumen while the GHI can be based on the trend in this same variable. Clearly, the two variables are closely related.

Recent experience with the concept of identifying GHI regions has produced mixed results. Since the idea of the GHI is to identify regions of variability, these should conform to our expert understanding of the deposition. However, in recent applications, this conformity has not been clear, thus making interpretability of a GHI map difficult. Future work is required to determine an appropriate measure of GHI and to interpret such a result.

Uncertainty Threshold Curves

In presence of multiple geological domains within the same deposit the uncertainty threshold is unlikely to be constant. Figure 2 shows a schematic illustration of an uncertainty threshold curve.

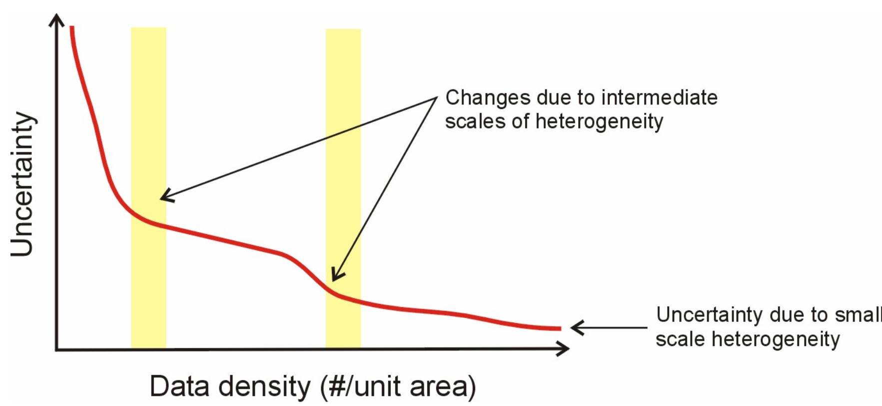
The dashed line on the left indicates a constant uncertainty threshold. In practice, we are likely to expect more knowledge in homogeneous areas and tolerate more uncertainty in heterogeneous areas (the solid line on the left). The sketch on the right shows the required data density to achieve an uncertainty threshold curve.

Given the concept of different geologically heterogeneous zones, we can imagine the construction of similar plots for each zone. Two maps could then be used for any location of interest to determine (1) the zone of geologic heterogeneity, and (2) the corresponding data spacing required to achieve the required confidence target for that particular phase of development.

Comments on Case Studies

A number of case studies were undertaken by CCG researchers during the 2003/2004 reporting period. Our intent is to establish and develop general results from these case studies and (1) present general methodological results with well-understood contrived examples, and (2) where appropriate and with permission present specific case studies.

A research focus of CCG has been characterizing the McMurray formation for planning and managing both mining and in-situ (SAGD) projects. Three projects were considered in detail. The results were consistent and fit nicely within the context presented above. Uncertainty decreases monotonically with increasing number of wells per section. There are interesting and important breakpoints that are related to the scale of geological variability, see the schematic below. There is good reason to target the shaded data densities for delineation.



We believe that some general results can be established to permit stakeholders (including companies and regulators) to set the requirements on the basis of technical criteria. Additional research is required to understand the geological reasons for the breakpoints and to propose general results for the McMurray.

The application of these concepts (relating uncertainty to geometric measures of data density or spacing) was also applied to a large base metal mining project. The relationship was not so clear. In fact, the best correlation between uncertainty measures and geometric measures was no greater than 0.3. In the end, we chose to use the uncertainty measures from the simulation directly. Additional research is required to understand the settings in which it is not practical to relate uncertainty to data spacing.

Future Work

One goal of drilling is to reduce uncertainty. Two primary factors affect uncertainty: (1) number of data available, and (2) geological complexity of the formation. These two factors are related; the more complicated the geology, the more data are required to understand the true complexity. Traditionally, specification of drillhole or well spacing requirements has relied on simple geometric measures that are easily comprehensible.

Uncertainty depends on the geologic heterogeneity of the region and the phenomena of interest. We have geostatistical tools that permit uncertainty quantification, yet we do not have the tools to relate this assessment to the data spacing required to achieve specific uncertainty targets for resource delineation and/or classification. The concept of identifying regions of geological heterogeneity and determining the corresponding uncertainty threshold curves provides a good start towards achieving this goal.

Preliminary results show there is much work to be done in this area. To proceed down the path of utilizing a geological heterogeneity index, we must first consider the criteria that make certain parameters an ideal measure of heterogeneity. Once established, robustness of this defined measure must be tested. Of course, we expect that this measure will be project-specific, so some practical guidelines would be helpful for general consumption.

References

- Deutsch, C.V. 2002, *Geostatistical Reservoir Modeling*. Oxford University Press, New York.
- Deutsch, C.V. and Journel, A.G., 1998. *GSLIB: Geostatistical Software Library: and User's Guide*. Oxford University Press, New York, 2nd Ed.
- McLennan, J.A. and Deutsch, C.V., SAGD Reservoir Characterization Using Geostatistics: Application to the Athabasca Oil Sands, Alberta, Canada, *CCG Technical Report 6*, Edmonton, AB, 2004.
- Ren, W., Leuangthong, O. and Deutsch, C.V., Geostatistical Reservoir Characterization of McMurray Formation by 2-D Modeling, *CCG Technical Report 6*, Edmonton, AB, 2004.
- Ren, W. and Deutsch, C.V., Geostatistical Reservoir Characterization of McMurray Formation by 3-D Modeling, *CCG Technical Report 6*, Edmonton, AB, 2004.

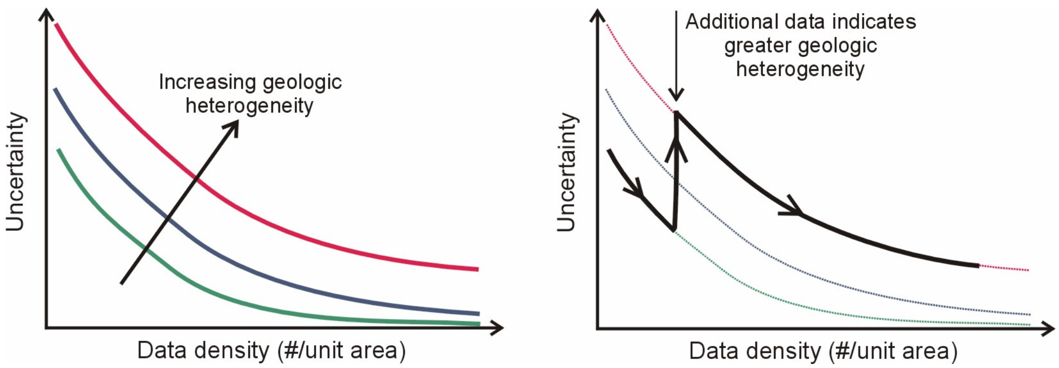


Figure 1: Schematic illustration of how uncertainty decreases with increasing data density. The sketch on the left shows three curves with increasing geologic heterogeneity. The sketch on the left shows how the apparent uncertainty increases when sufficient data is collected to indicate that the deposit is more geologically complex than first thought

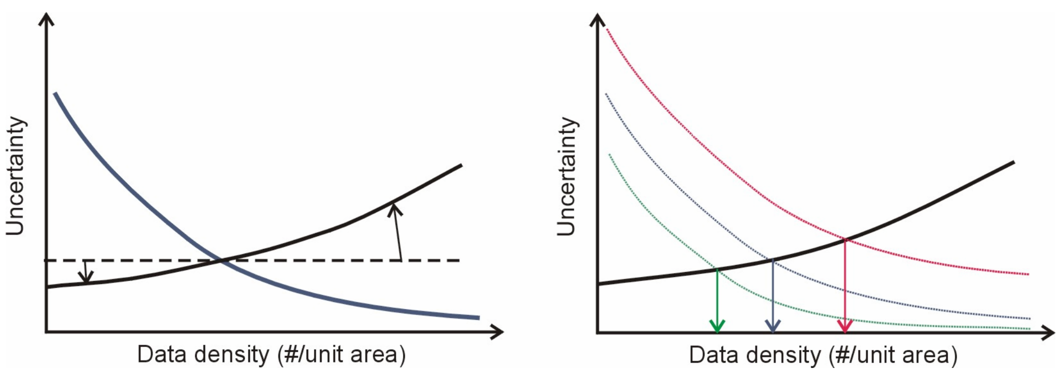


Figure 2: Schematic illustration of an uncertainty threshold curve. The dashed line on the left indicates a constant uncertainty threshold. In practice, we are likely to expect more knowledge in homogeneous areas and tolerate more uncertainty in heterogeneous areas (the solid line on the left). The sketch on the right shows the required data density to achieve an uncertainty threshold curve.