Geostatistical Modeling of McMurray Oil Sands Deposits

Oy Leuangthong[1], Emmanuel Schnetzler[2] and Clayton V. Deutsch[1,2]

[1]Department of Civil & Environmental Engineering,
University of Alberta

[2]Statios LLC

Abstract

The McMurray formation in the Athabasca oil sands deposits of Northern Alberta is part of the world's second largest proven crude oil reserves. The formation is characterized by stratigraphic layers that correspond to three different depositional environments: Marine, Estuarine and Fluvial facies. Resource estimation for oil sands has traditionally relied on polygonal and inverse distance schemes. These techniques are simple and straightforward in practice, but they do not permit reliable uncertainty assessment. This paper describes the modeling of the McMurray formation using modern geostatistical techniques.

Geostatistical modeling should be performed within homogeneous geological facies. The methodology for geostatistical simulation within each of the identified facies involves: (1) assessing the most appropriate stratigraphic transformation for optimisation of the correlation structure, (2) determining representative distributions with declustering and debiasing techniques, (3) modeling spatial continuity of the bitumen grade, fines grade, water saturation and other petrophysical variables, (4) performing estimation and cross validation as checks against simulation results, (5) performing simulation for uncertainty quantification of bitumen and fines grade, and (6) model checking of simulation results against the input data and comparisons against the kriged models. These steps are described in detail. The application of these models for uncertainty quantification in support of resource classification is also discussed.

Introduction

The Alberta oil sands deposits contain approximately 1.6 trillion barrels of bitumen, of which 174 billion barrels are proven reserves. In 2002, Canada's proven crude oil reserves comprised approximately 15% of the world's reserves, second only to Saudi Arabia (Alberta Energy, 2003). The oil sands deposits found in Alberta comprises the majority of Canada's crude oil reserves, with about 35% of Canada's crude oil output coming directly from the production of Alberta's oil sands, and this is expected to increase to 50% by the year 2005 (Alberta Energy, 2003). These statistics and economic trends clearly indicate the importance of the oil sands industry to Canada's current and future economy.

Crude bitumen can be found in three main regions in Alberta: Peace River, Athabasca and Cold Lake. Of these three, the Athabasca deposit is the largest with an areal extent of over 32,000 square kilometres (Steward and MacCallum, 1978). The Athabasca deposit has a Devonian limestone base and consists of three main formations (in ascending order): McMurray, Clearwater and Grand Rapids formations (Camp, 1976; McRory, 1982). The sands in the McMurray formation are host to the crude bitumen, in which three main lithofacies are recognized based on the depositional environments: Fluvial, Estuarine and Marine from the base to the top of the formation (Figure 1). These form the stratigraphic layers of interest for resource modeling.
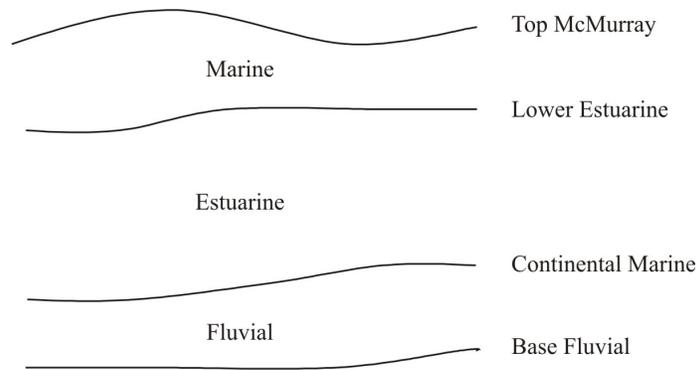
Figure 1 Schematic illustration of stratigraphic layers found within the McMurray formation of the Athabasca deposit.

Resource modeling of oil sands has traditionally employed fairly straightforward techniques, such as polygonal and inverse distance schemes. These resource models are usually combined with some type of geometric resource/reserve classification approach for public disclosure reporting. These methods, however, do not provide a measure of uncertainty for the resource and/or reserve estimates.

Instead, we propose to use a geostatistical simulation approach to model the McMurray formation that permits uncertainty assessment. Considerations, quality control measures, and the basic modeling framework specific to the McMurray formation are described. The methodology is illustrated for a small sanitized dataset, originally based on one of the current lease areas near Fort McMurray. This is followed by a discussion on the applications of this model, including resource classification.

This paper does not focus on modeling the stratigraphic controls; nevertheless, geostatistical modeling must always be performed within a stratigraphic framework accounting for all geological controls that are deemed representative. There is an important tradeoff that must be made in practice between respecting deterministic controls interpreted by experienced geological staff and respecting the inherent uncertainty in presence of widely spaced data relative to small scale heterogeneity. This paper describes the geostatistical procedures that are primarily aimed at modeling and understanding the stochastic uncertain aspects of oil sands deposits.

Geostatistical Modeling Methodology

The geostatistical approach can be described in six distinct steps:
1. Analyze of correlation structure. This investigates whether a transformation of the vertical coordinate system is required, in order to determine the true continuity structure of the deposit. There are a number of different correlation grids that should be considered: onlap, erosional, and proportional (Deutsch, 2003). Determination of the correct grid is dependent on the correlation grid that yields the maximum horizontal continuity.
2. Decluster drillhole data distribution. This is a critical stage of the modeling workflow; the relevant statistics must be deemed representative of the deposit prior to modeling. Any or a combination of cell, nearest neighbour and/or declustering by kriging weights may be employed to determine the summary statistics that are representative of the field.
3. Model spatial continuity of the bitumen grade, fines grade, water saturation and any other petrophysical variables of interest. A number of different spatial measures can be considered

here, among which, the semivariogram is the most common.

4. Perform estimation and cross validation using kriging as checks against simulation results. Kriging is widely accepted in the mining industry and yields a map of the large scale continuity of the deposit. Cross validation using kriging provides a quality control check on the estimation (and also simulation) parameters. This is considered good practice in order to ensure that the results from modeling are consistent with geological expectations prior to the next step of simulation.

5. Construct uncertainty models of bitumen grades, fines grades, etc., using geostatistical simulation. Sequential Gaussian simulation (Isaaks, 1990) is by far the most commonly applied geostatistical simulation algorithm applied in the natural resources sector. It has been extensively validated and provides a measure of local and global uncertainty, that is not afforded from kriging.

6. Check simulation results against the input data and compare results against the kriged models.

These uncertainty models can then be used for both long and short term decision making. For instance, resources/reserves can be estimated along with the uncertainty and hence risk associated to this estimate; grade control optimization can also be realized by way of optimizing dig limits, accurately predicting the head grade, identifying production periods of economic risk; and resources/reserves can be classified in accordance with National Instrument 43-101 for public disclosure. This last application is becoming more important as the modern classification approach is based on adopting a measure of uncertainty and then assessing the probability to be within this measure. Some of the world's leading mining companies have already begun to adopt this type of modeling workflow, realizing the advantages of these types of techniques over the more deterministic approaches currently used.

An application to an example oil sands dataset will illustrate each of the above steps in more detail. The data used in this application appear realistic, but are synthetic and do not belong to any particular site. The work flow and conclusions are reasonable and valid for a range of oil sands deposits. Moreover, the application is shown only for bitumen grades; however, the same workflow would apply to fines and any other variables of interest.

Application

To illustrate the proposed methodology, a small dataset consisting of 94 drillholes is used. This exercise will only be performed for the Estuarine facies; a similar approach and analytical workflow would be required for the Marine and Fluvial facies. The three facies are typically treated independently since they belong to separate environments. A total of 1609 3ft composites was available. Figure 1 shows the projection of the available DH data onto the horizontal and vertical planes, as well as the distribution of the bitumen grade. The smallest drillhole spacing is approximately 50m in plan view, and samples are 3m vertically spaced. A fairly fine scale model is chosen with a selected block size of 25m x 25m x 3m, covering a 7.5km x 7.5km x 100m volume. This yields a 3D model with a total of 2.57 million cells. It is common to construct numerical models with tens of millions of grid cells.

The cross sectional views of the data show a stratigraphic layer that appears fairly flat and tabular; however, undulations of the surface indicate that the different correlation structures should be evaluated for application. In addition to the direct elevation coordinates, stratigraphic coordinates corresponding to an onlapping, erosional/truncation, and proportional (Figure 3) correlation grids were considered (Deutsch, 2003). In each case, the data were transformed to normal scores and the variograms of these scores were calculated (Figure 4). The most continuous correlation structure is an appropriate choice for a stratigraphic grid. Figure 4 shows that the variogram corresponding to the original elevation coordinates is the most continuous in the short and long range structures. Therefore, no stratigraphic coordinate

transform is required for this particular dataset. This, however, does not preclude the need for a stratigraphic transform for different oilsands data. A similar study of continuity depending on correlation grid should be performed for each dataset. Knowledge of the geological depositional and tectonic history should also be taken into account when determining the stratigraphic framework.

Cell declustering was then performed in order to determine the appropriate declustering weights. In this sense, data in densely clustered areas will receive less weight than data in sparsely sampled areas. Most times, the clustered data can be expected to be in the high grade regions as drillholes are preferentially drilled in these zones. Therefore, less weight can be expected to be given to the high grades, and higher weights can be expected for the low grade regions. The result of this type of weighting is a weighted mean (Figure 5) that is lower than the naïve mean (Figure 2). Of course, the reverse would be expected if we were declustering fines data, where the low fines regions are the areas that would have been preferentially sampled. Figure 5 is the % bitumen distribution that we should expect to reproduce in subsequent models. Note that declustering does not affect data values, but only assigns weights to each measurement.
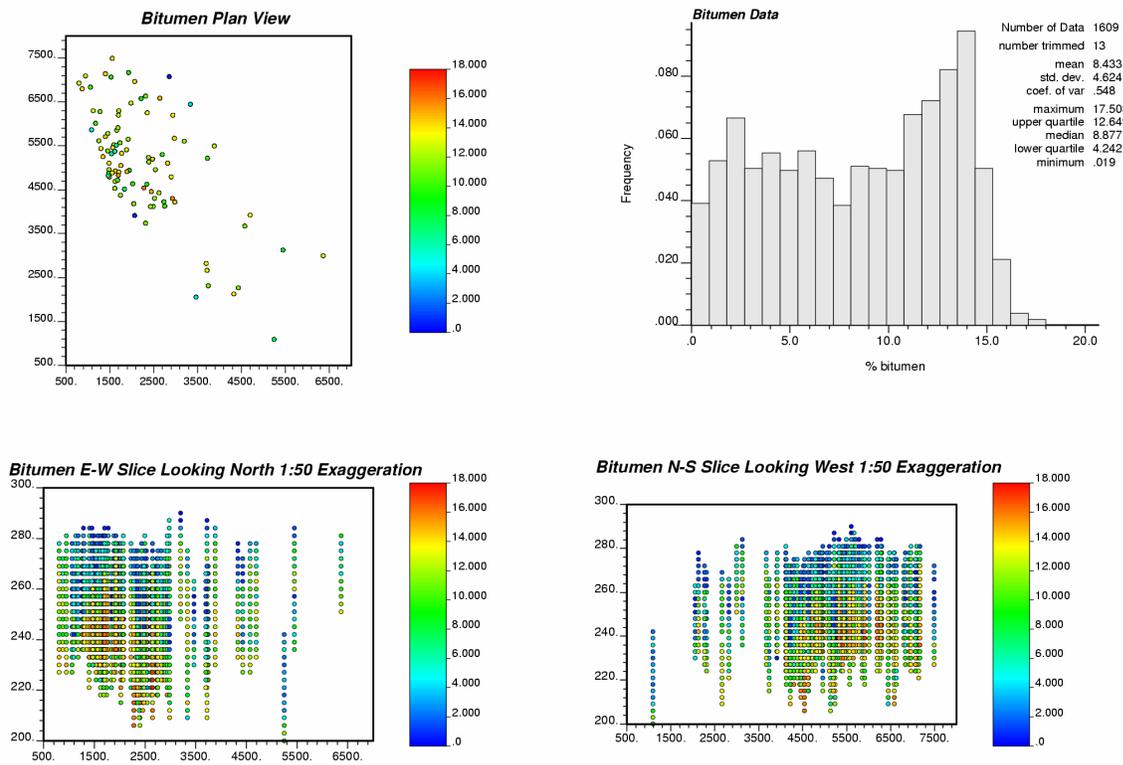
Figure 2 Drillhole data projected onto a horizontal plane (plan view (top left)), and two vertical planes (E-W sectional view (bottom left) and N-S sectional view (bottom right)). Distribution of % bitumen is also shown (top right).

$$z_{strat} = \frac{zrel - zbase}{ztop - zbase} \cdot thickness$$
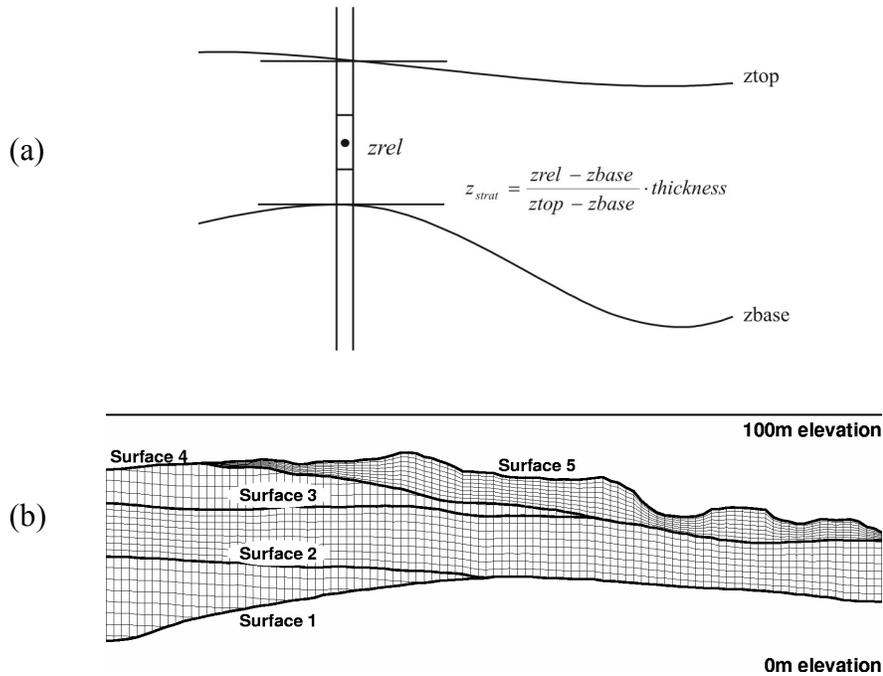
Figure 3 (a) schematic illustration of stratigraphic transform for a proportional grid, and (b) outcrop example with four stratigraphic layers bounded by five major surfaces: the style from bottom to top: onlap, proportional, erosional, and proportional.
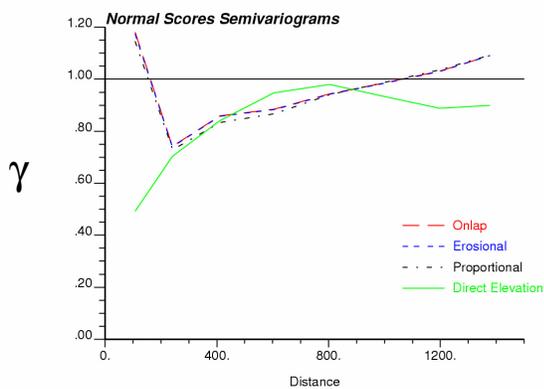
Figure 4 Variograms calculated based on different correlation structures: elevation, proportional, erosional and onlap.
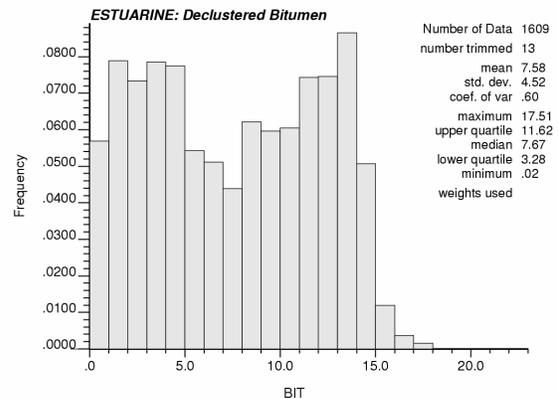
Figure 5 Histogram of declustered bitumen data.

A normal scores transform was then performed. Using these normal scores, directional variograms were calculated and modeled. In the calculation, a number of tools were used to determine the directions of anisotropy including variogram maps and calculating the variogram in a series of multiple directions. From this analysis, the major continuity direction was determined to be N-S and the minor direction was E-W. The variograms and fitted models are shown in Figure 6. The fitted model is

$$\gamma(\mathbf{h}) = 0.05 + 0.25 Sph_{ah\max=100}(\mathbf{h}) + 0.50 Sph_{ah\max=1500}(\mathbf{h}) + 0.20 Sph_{ah\max=2200}(\mathbf{h})$$

$$\begin{array}{ccc} ah\min=120 & ah\min=800 & ah\min=800 \\ ahvert=33 & ahvert=33 & ahvert=33 \end{array}$$

where *ahmax* indicates the range in the horizontal maximum continuity direction (N-S), *ahmin* denotes the range in the horizontal minimum continuity direction (E-W), and *ahvert* is the range in the vertical direction.

Before using these variogram models and representative statistics to construct simulation models, some quality control tasks should be performed to prevent errors later in the workflow. Amongst the quality control checks are standard cross validation tools. The idea of cross validation is to estimate at known sample locations as if it was unknown, using the surrounding data. The estimate is then compared to the true value and the error in the estimation can be assessed directly. A look at the distribution of errors shows that it is approximately normally distributed with a mean of zero, and the correlation between the estimate and the true values is fairly high at 0.707 (Figure 7). A mean of zero shows that the errors are unbiased and a high correlation between the estimate and data is encouraging for subsequent modeling.

Simulation was then performed using sequential Gaussian simulation (Isaaks, 1990). In all, 100 realizations of the 2.5 million cell model were generated in 6.75 hours (approximately 4 minutes per realization). Figure 8 shows four arbitrarily chosen realizations at elevation 248m. All the maps are clearly different, yet there are some similar features. For instance, in the northwest quadrant of the maps, there is a very clear region that is characterized by a high % bitumen grade. Just southeast of this high grade region, there is a small pocket that is clearly low grade. These areas were verified against the data and the kriged models, both of which confirmed that the available drillhole data in these regions reflected the same grade trends. Further examination shows that in the northeast and southwest areas of the map where no drillholes are available (see Figure 2), there is greater variability from one realization to the next. Realizations 25 and 75 show a medium to high grade region in the northeast, while realizations 50 and 100 show that this region could actually be a low grade region. This variability in the realizations reflects the uncertainty in these regions where there are no drillhole and/or blasthole to suitably condition the models.

The simulated model was then checked for histogram and variogram reproduction. Although one would expect that these two key inputs should be reproduced by any representative model, only a simulated model will reproduce these statistics. An estimated model, obtained from either kriging, inverse distance or any other conventional method, should produce the same large scale features, but these will not be able to reproduce either the histogram or the variability of the true data. Figure 9 shows the histogram of one realization of bitumen; this histogram is comparable to the declustered distribution of Figure 5, in terms of its shape and the magnitude of the summary statistics. Further, the reproduction of the summary statistics was also checked over the 100 realizations (Figure 10).
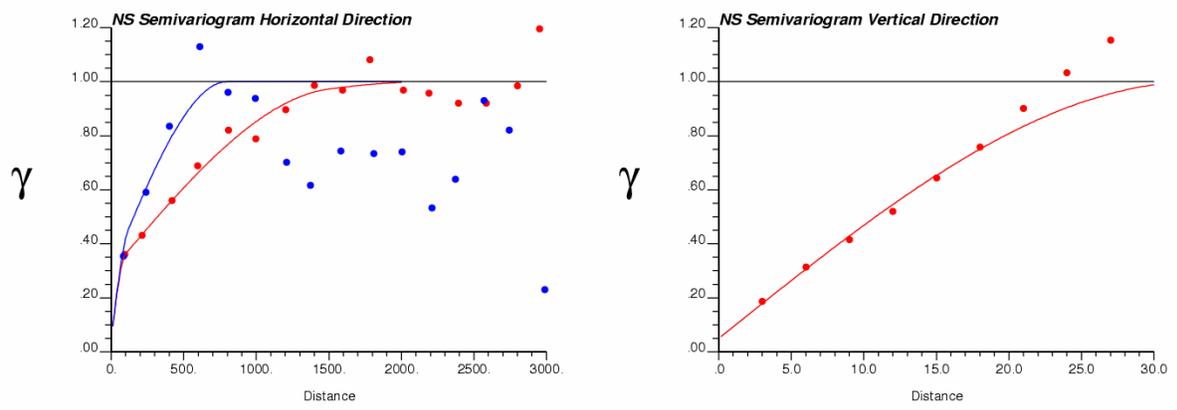
Figure 6 Directional variograms for normal scores bitumen: horizontal (left) with blue corresponding to the E-W direction, and red corresponding to the N-S direction, and vertical (right) direction.
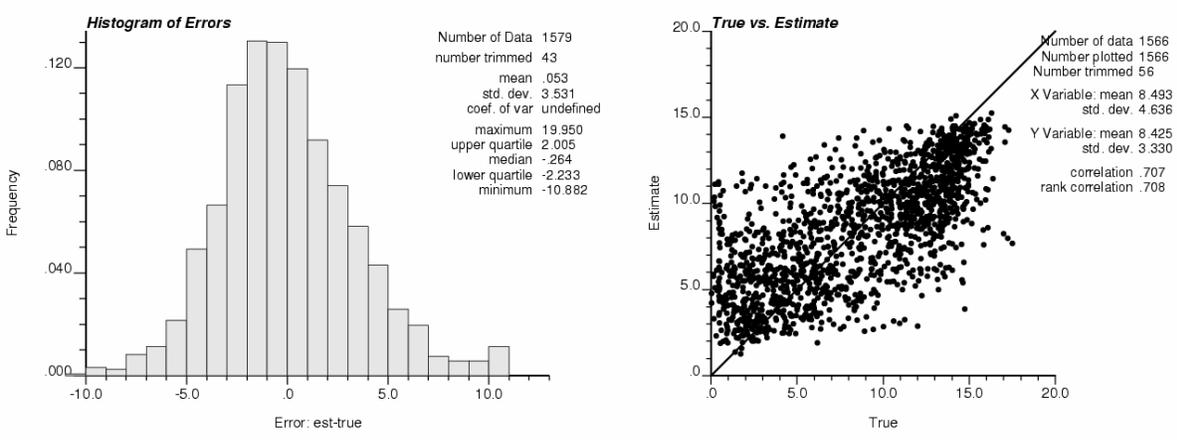
Figure 7 Cross validation results: histogram of errors (left) and crossplot of true vs. estimate (right).
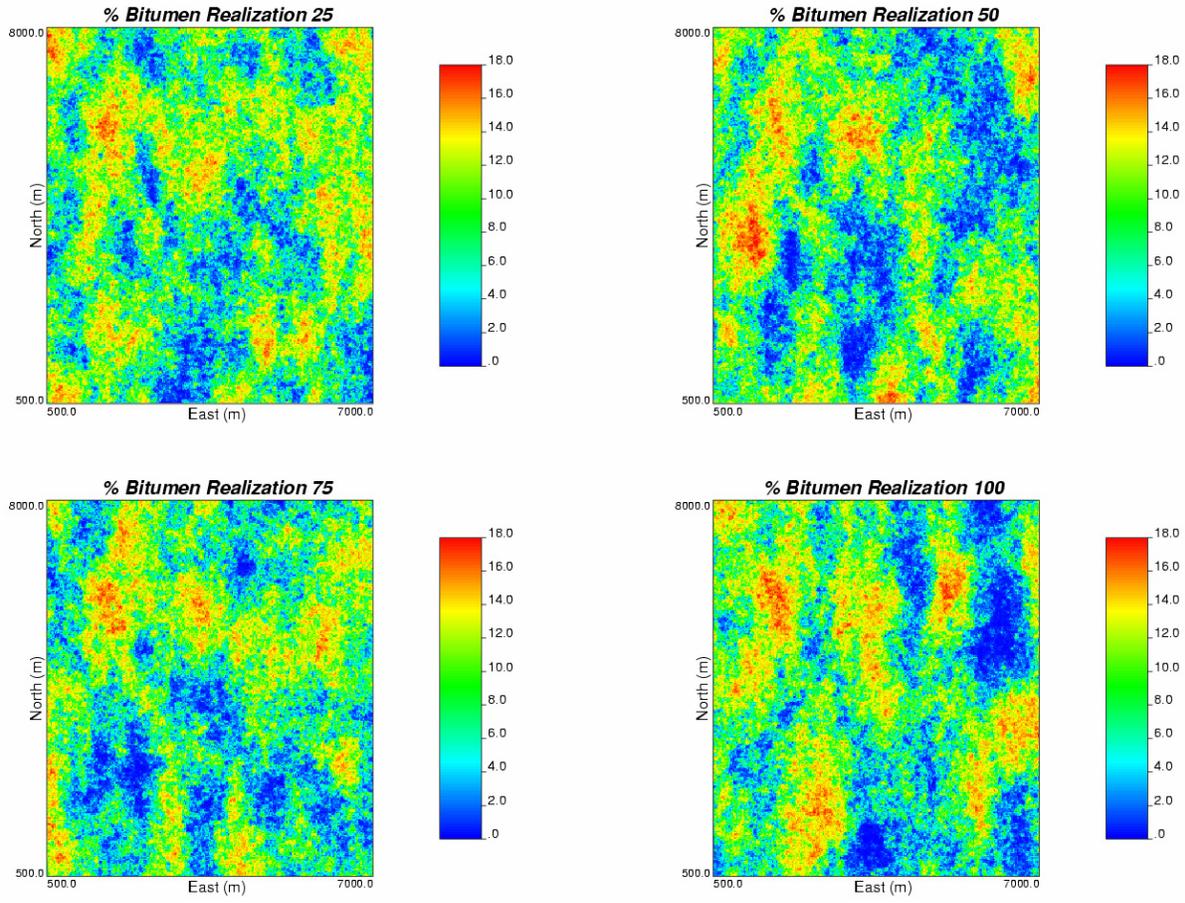
Figure 8 Four realizations of % Bitumen obtained from sequential Gaussian simulation.
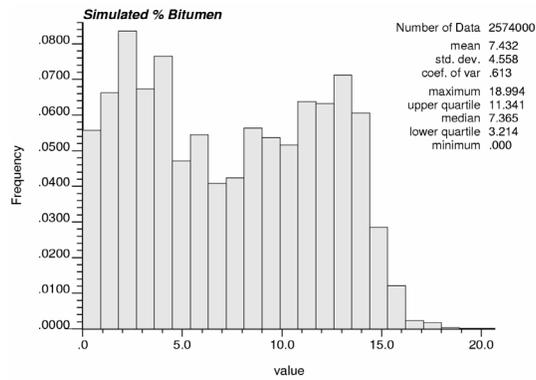
Figure 9 Histogram of simulated % bitumen from one realization.
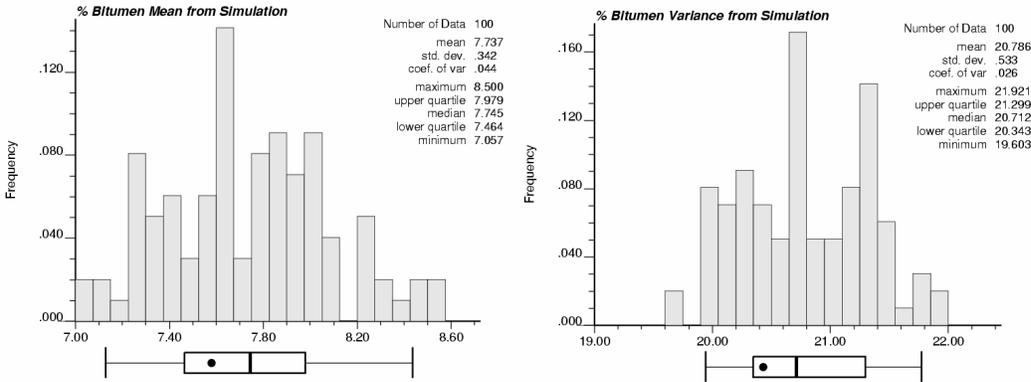
Figure 10 Box plots show the reproduction of summary statistics: mean (left) and variance (right). In both cases, the representative statistic falls within the interquartile range, and shows reasonable reproduction.

Discussion

Model checking is important. Quality control steps were integrated into the proposed methodology through declustering, cross validation and model checking after simulation. These are straightforward, basic checks that should be performed as part of any resource modeling procedure; unfortunately, these are often neglected in practice (Leuangthong, McLennan and Deutsch, 2004). Aside from checking that the implementation of the modeling method was correct, validation using production data remains the true test of the model's ultimate value in resource estimation. In this aspect, blasthole data validation is crucial. This can only performed once production begins, but even at the early stages of production, the performance of the simulations can be directly assessed. Comparisons between blasthole samples and simulated values at the same location will indicate how well the simulated model is able to predict the grade at previously unsampled locations.

Once validation using production data is performed, there is a number of interesting applications for the multiple realizations of a simulated model. The uncertainty assessment afforded by the model can be used to assess dilution at different scales, simulate production over different periods, forecast recovery, and resource/reserve classification. With the passing of National Instrument 43-101 in 2001 (Canadian Securities Administrators, 2001), this latter application has come to the forefront of useful simulation applications. With the use of simulation, resources/reserves classification accounts for more than just the geometry of the data configuration; it also accounts for the sample grades and the spatial continuity of the grade(s) of interest. Further, classification is based on a measure of uncertainty that inherently captures the state of knowledge and information at the time of model construction.

Quality control measures taken throughout any resource modeling task must also be addressed in reporting for public disclosure. The qualified person is required to document specific measures that were duly undertaken in an effort to validate the reported reserves. These measures must be an integrated part of the workflow and should be taken throughout the project, in all aspects of resource modeling.

Conclusions

There is little doubt that the future of Canada's energy economy will depend largely on the oil sands resources found in Alberta. Of the three main crude bitumen deposits, the Athabasca formation is by far the largest and most amenable to production using current surface mining technologies. Estimation of these resources is incredibly important to forecasting its availability and determining the feasibility of current mining practice for certain deposits. In this assessment, uncertainty is a key issue, which cannot

be addressed using conventional estimation tools that are prevalent in the industry. Instead, a geostatistical approach is proposed that is fairly straightforward and is well positioned to satisfy the conditions and requirements set in classification guidelines.

The proposed geostatistical methodology for modeling the oil sands deposits of the McMurray formation can be summarized into six steps: (1) determine an appropriate stratigraphic coordinate system that maximizes the correlation structure, (2) obtain a distribution of grades that is representative of the deposit by employing declustering techniques, (3) model the spatial continuity measure that shows the variability of grades with distance, (4) perform cross validation and kriging as quality control measures to detect any potential blunders, (5) generate multiple realizations using sequential Gaussian simulation, and (6) check the realizations to ensure that input data and statistics are reproduced. Although this is a fairly classical geostatistical simulation methodology, an important feature for the oil sands setting is the analysis of correlation at the very beginning of the modeling workflow. Given the known depositional environments found within the McMurray formation, it is imperative that this is reflected in the stratigraphic controls imposed on the models.

Uncertainty is inevitable when dealing with incomplete data. The paradigm of statistics is to quantify the uncertainty and permit risk qualified decision making in presence of uncertainty. The advantage of geostatistics is that spatial correlations between the data are modeled and used to improve local grade estimates and uncertainty estimates. There has been some controversy regarding the use of geostatistics. A claim is made that classical statistics is more straightforward and that geostatistics is needless. The use of the correlation between multiple data is an established principle of statistics. Correlation between data has been accounted for since the beginning of statistics. de Moivre, Laplace, Gauss, Bayes and others luminaries of statistics pioneered these techniques since the late 1700s. Geostatistics is the modern application of very classical statistical principles.

References

ALBERTA DEPARTMENT OF ENERGY, September 9, 2003, "Alberta's Oil Sands", accessed on March 26, 2004 at http://www.energy.gov.ab.ca/osd/docs/osgenbrf.pdf.

CANADIAN SECURITIES ADMINISTRATORS, 2001, National Instrument 43-101 Standards of Disclosure for Mineral Projects, Form 43-101F1, Technical Report, and Companion Policy 43-101CP.

CAMP, D.F., 1976, The Tar Sands of Alberta, Canada. 3rd Edition, Cameron Engineers Inc., Denver, CO, 81 p.

DEUTSCH, C.V., 1996. Direct assessment of local accuracy and precision. In E.Y. Baafi and N.A. Schofield, ed., Geostatistics Wollongong '96, vol. 1, Holland, Kluwer Academic Publishers, pp. 115-125.

DEUTSCH, C.V., 2003. Geostatistical Reservoir Modeling. Oxford University Press, New York, 376 p.

DEUTSCH, C.V. and JOURNEL, A.G., 1998. GSLIB: Geostatistical Software Library and User's Guide. 2nd Edition, Oxford University Press, New York, 369 p.

ISAAKS, E.H., 1990. The Application of Monte Carlo Methods to the Analysis of Spatially Correlated Data, Ph.D. Thesis, Stanford University, Stanford, CA, 213 p.

LEUANGTHONG, O., McLENNAN, J.A., and DEUTSCH, C.V., 2004. Minimum Acceptance Criteria for Geostatistical Realizations, accepted by Natural Resources Research, 20 p.

McRORY, R.E., 1982, Oil Sands and Heavy Oils of Alberta, Alberta Energy and Natural Resources, 94 p.

STEWART, G.A. and MACCULLUM, G.T., 1978. Athabasca Oil Sands Guide Book, Canadian Society of Petroleum Geologists, Calgary, Canada, 33 p.