Data Assimilation of Sensor Measurements to Improve Production Forecasts in Resource Extraction

T. WAMBEKE$^{1,2,*}$ and J. BENNDORF$^1$

$^1$Resource Engineering - Delft University of Technology, the Netherlands t.wambeke@tudelft.nl
$^2$MTI Holland - Royal IHC, the Netherlands
$^*$ presenting author

Abstract

One of the main challenges in the mining industry is to ensure that short-term production targets are continuously met. The executed mine plan, formulated in an attempt to comply with the targets, often turns out to be suboptimal due to unexpected deviations in the resource model.

Recent sensor technology enable the on-line characterization of these deviations at different points along the production process. There exists a potential to improve short-term planning if such on-line measurements can be integrated back into the resource model.

This contribution introduces a novel approach to real-time resource model updating. An algorithm was developed that specifically deals with the practical problems of the mineral industry. A case study demonstrates the added value.

1 Introduction

Traditionally, the mining industry has had mixed successes in achieving the production targets it has set out. Several projects have been identified where mineral grades are not as expected, schedules and plans are not met and recovery is lower than forecasted (McCarthy, 2004). Against this background, the industry is increasingly looking for effective methods for monitoring and reconciling the estimates and actual observations at different stages of the resource extraction process (Morley, 2014).

Recent developments in sensor technology enable the on-line characterization of production performance and raw material characteristics. To date, sensor measurements are mainly utilized in forward loops for downstream process control and material handling (Nienhaus et al., 2014). A backward integration of sensor information into the resource model to continuously improve the production forecasts and dispatch decisions does not yet occur.

The potential of real-time updating of mineral resource models is obvious. However, in order to apply it in practice, algorithms need to be developed which are computationally efficient, easy to use and can handle uncertainty. A general approach is needed to deal with differences in scale of support and measurement units. The objective of this paper is to present a practical methodology for assimilating sensor measurements of extracted material into the resource model.

Figure 1 illustrates the concept behind the designed algorithm. Its practicality mainly results from a design decision to exclude the forward observation model $A$ from the computer code. The design is adapted such that for each unique case an application specific observation model can be developed and used independently from the existing updating algorithm. Each run requires two types of inputs. The first input contains a collection of block model realizations. These realizations are generated via geostatistical simulations to characterize spatial heterogeneity. The second input consists of a collection of actual and predicted sensor measurements, characteristic for the blocks in a specific neighbourhood. The predicted sensor measurements are obtained by applying the observation model to each resource model realization.

The mathematical principles behind the updating framework are discussed later on in this paper. First, a brief overview of the most recent developments in sequential updating is presented. The paper concludes with a synthetic case study where truck loads measurements are assimilated into the resource model.
Figure 1: Closed-loop reconciliation framework to integrate on-line sensor measurements into the resource model. The design allows for the use of application specific observation models or process simulation software.

2 Recent developments in sequential updating

2.1 Updating of single best estimates

Several geostatistical algorithms for assimilating direct and indirect measurements have been proposed in the literature. Initial results were obtained in the field of inverse modelling. Consider the expression:

\[ d = A(z(x)), \]

where \( A \) is a forward observation model (linear or non-linear) that maps the spatial attributes \( z(x) \) into observations \( d \). Estimation of the attributes \( z(x) \) using the observation model \( A \) and the observed data \( d \) is referred to as solving the inverse problem (Oliver et al., 2008).

Kitanidis and Vomvoris (1983) introduced a geostatistical approach to the inverse problem in groundwater modelling. Both scarce indirect and more abundant direct measurements are used to estimate spatial attributes through a linear estimation procedure known in the geostatistical literature as cokriging (Chiles and Delfiner, 2012).

About a decade later, several researchers developed almost simultaneously and independently of each other a method to recursively incorporate subsets of data one at a time (Evensen, 1992; Harvey and Gorelick, 1995). The proposed sequential estimator improves previous subsurface models by using linearly weighted sums of differences between observations and model-based predictions:

\[ z_t = z_{t-1} + W_t(d_t - A_t(z_{t-1})), \]

where the vector \( z_t \) contains estimates of the spatial attributes after \( t \) updates; the vectors \( d_t \) and \( A_t(z_{t-1}) \) respectively hold actual observations and model-based predictions at time \( t \); and \( W_t \) is a matrix with weights defining the contribution of the detected deviations to the updated attributes. The sequential linear estimator bears a remarkable resemblance to Kalman filter techniques (Evensen, 1992; Bertino et al., 2002). The kriging weights are computed from the forecast and observation error covariance matrices:

\[ W_t = C_{(t-1, dz)}C_{(t-1, dd)}^{-1} = C_{(t-1, zz)}A_t^T(A_tC_{(t-1, zz)}A_t^T + R)^{-1}, \]

where \( C_{(t-1, dz)} \) and \( C_{(t-1, dd)} \) are the cross-covariance matrices between previous estimates and observations, and \( C_{(t-1, zz)} \) is the covariance matrix of previous estimates. The resultant matrix \( W_t \) is used to update the estimates and reduce uncertainties in the spatial attributes.
where $C_{t-1,zz}$ is the prior error covariance matrix of the attribute field; $R$ is a diagonal matrix which specifies the sensor precision and $A_t$ is a first order approximation of the non-linear observation model $A_t$ (Evensen, 1992).

Vargas-Guzman and Yeh (1999) provided the theoretical evidence that under a linear observation model sequential kriging and cokriging are equivalent to their traditional counterpart which includes all data simultaneously. In case of a nonlinear observation model, a sequential incorporation of data increases the accuracy of the first-order (linear) approximations $A_t$ since they are calculated around a progressively improving attribute field (Evensen, 1992; Harvey and Gorelick, 1995).

The linear estimator propagates the conditional mean and covariances from one update cycle to the next. The updated covariances now depend on the location and value of assimilated observations as well as on the spatial variability structure of the modelled field. The nonstationarity of the covariance results in perhaps the greatest limitation of the method, i.e. the necessity of storing large covariance matrices. Despite the promising results, the above mentioned techniques were not considered for application in mineral resource extraction for fast sequential updating of deposit models.

### 2.2 Updating of geostatistical simulation models

In order to assess the spatial variability and uncertainty of the attributes, it is common in geostatistics to simulate a set of realizations (Dimitrakopoulos, 1998). Vargas-Guzman and Dimitrakopoulos (2002) presented an approach that can facilitate fast-updating of generated realizations based on new data, without repeating the full simulation process. The approach is termed conditional simulation of successive residuals and was designed to overcome the size-limitations of the LU decomposition Davis (1987). The novel column partitioning requires the specification of a sequence of (future) data locations. A stored $L$ matrix can then facilitate the conditional updating of existing realizations if and only if the sequence of visited subsets and used production data are the same as the one used for generating the initial realizations (Jewbali and Dimitrakopoulos, 2011).

A major limitation of the techniques discussed thus far results from the necessity to store and propagate the conditional nonstationary covariances. To circumvent this limitation, the previously discussed sequential linear estimator can be integrated into a Monte Carlo framework (Evensen and van Leeuwen, 1996). At each time step, a finite set of realizations is updated on the basis of sensor measurements collected at time $t$:

$$\{z_i^t = z_i^{t-1} + W_i(d_t - A_t(z_i^{t-1}) + e_t) \quad i = 1 \ldots I\},$$

(4)

where the observation errors $e_t$ are randomly drawn from a normal distribution with a zero mean vector and a diagonal covariance matrix $R$. The conditional covariances, $C_{(t-1,zd)}$ and $C_{(t-1,dd)}$, are computed empirically from both sets of model-based predictions $A_t(z_i^{t-1})$ and field realizations $z_i^{t-1}$. Due to the applied Monte Carlo concept, the first-order approximation of the observation model becomes avertible (Burgers and van Leeuwen, 1998).

Benndorf (2014) was the first to recognize the potential of sequential updating in the field of solid resource extraction. Due to the nature of a mining operation, an observation is characteristic for blended material coming from neighbourhoods surrounding one or more extraction points. This contribution extends the work of Benndorf (2014) by fivefold.

1. The connection of the technique to the existing geostatistical literature is exemplified.
2. A sequences of Gaussian Anamorphoses is performed to deal with suboptimal conditions related to non-linear relationships and non-Gaussian distributions.
3. An interconnected parallel updating sequence (a double helix) is configured to reduce the number of necessary realizations.
4. A specific implementation of the empirical calculation of covariances overcomes problems with differences in scales of support.

5. A localization or neighbourhood strategy is applied to attain acceptable computation times.

The integration of these modifications results in an efficient and extremely practical algorithm capable of handling the specific challenges of the mineral industry. The following section further elaborates on the proposed modifications and integrates them into a general mathematical framework for sequential resource model updating.

3  Framework for sequential resource updating

In this section, the formal description of the proposed framework is presented. All statistical information is extracted from two provided ensembles with Monte Carlo realizations; the simulated attribute fields and the model-based predictions (Figure 1).

3.1 Gaussian Anamorphosis and treatment of measurement error

A series of local Gaussian anamorphoses (GA) is proposed to mitigate the effect of non-Gaussian data dependence on the performance of the sequential updating approach (Zhou et al., 2011). The objective of the transformation methodology is to handle the developing nonstationarity, implicitly model changes in support and correctly account for observation errors. Each individual Monte Carlo sample, either representative for a single grid node or observation, needs to be transformed according to its local statistics.

The nonparametric GA is suggested since this approach can extract the distributional shape from the Monte Carlo sample. Moreover, GA leads to an apparent pseudolinarisation of the dependence structure between the indirect sensor measurements and the resource attributes, which can be more accurately exploited by the linear updating equations (Schoniger et al., 2012).

The following paragraphs discuss the local transformation of individual grid nodes, predicted observations (with and without measurement error) and actual measurements.

A non-linear transformation $\Phi$ is applied to each Monte Carlo sample $z(x_n)$ to move the arbitrarily distributed grid nodes closer to Gaussianity:

$$\{u_i(x_n) = \Phi_{x_n}(z'(x_n)) = G^{-1}[F(z'(x_n))] \mid i = 1 \ldots I \} \ n = 1 \ldots N, \tag{5}$$

with the corresponding empirical cumulative distribution function (CDF) $F$ and the theoretical standard normal CDF $G$. As $G$ is, per definition, monotonously increasing, the inverse $G^{-1}$ exists. Note that for each grid location $x_n$ a new transformation function $\Phi_{x_n}$ needs to be constructed. The left side of Figure 2 illustrates the transformation of the resource attributes at the grid nodes.

Before discussing the transformation of observations, define the model-based predictions at time step $t$ as:

$$\{b_i^t(k) = A_{t,k}(z_{i-1}) \mid i = 1 \ldots I \} \ k = 1 \ldots K, \tag{6}$$

where $b_i^t(k)$ is a vector with $I$ predicted measurements taken at interval $k$ and period $t$ over resource attribute fields $i = 1 \ldots I$. The non-linear operator $A_{t,k}$ maps a single resource realization into the $k^{th}$ measurement.

In order to ensure compatibility with the real noisy observations $d_i$, random measurement errors $\epsilon_i N(0, R)$ are added to the model based predictions. The vector of $I$ transformed ensemble values $q_i^t(k)$ at each of the $k$ measurement locations or time intervals is obtained by evaluating the individual functions $\Phi_{t,k}$:

$$\{q_i^t(k) = \Phi_{t,k}(b_i^t(k) + \epsilon_i^t(k)) \mid i = 1 \ldots I \} \ k = 1 \ldots K. \tag{7}$$
Figure 2: Overview of Gaussian Anamorphosis applied to the Monte Carlo samples of the attribute fields, the model-based predictions and the real observations.

Note that the individual anamorphosis functions are based on the perturbed predictions. The correction of each attribute field by eq. 4 is based on the deviation of the simulated measurements from the observed ones. Therefore, the observed measurements have to be transformed according to the same anamorphosis function:

$$s_t(k) = \Phi_{t,k}(d_t(k)) \quad k = 1 \ldots K.$$  

(8)

The construction of the CDF and the subsequent transformation of real and perturbed predicted observations are depicted in Figure 2.

The previously constructed anamorphosis functions $\Phi_{t,k}$ are subsequently used to transform the simulated measurements, but now without their respective measurement errors:

$$\{y_i^t(k) = \Phi_{t,k}(b_i^t(k)) \quad i = 1 \ldots I \} \quad k = 1 \ldots K.$$  

(9)

The necessity of performing a second transformation on the original predicted observations (without error) will be exemplified later on in this section when calculating the empirical covariances. The transformation of predicted observations is displayed in Figure 2.

### 3.2 Double helix to reduce statistical sampling error

In order to avoid problems of inbreeding, a pair of sequential updating cycles is configured so that the assimilation of data into one set of realizations employs the weights calculated from the other one (Figure 3). The idea originates from Houtekamer and Mitchell (2001). They argue that the uncertainty in the updated realizations is likely to be underestimated if a single set of realizations was to be used to calculate both the weights and test their quality (posterior covariance).

The implementation of this concepts requires that all previously transformed Monte Carlo samples are split into two subsets of equal length. The necessary covariances can then be calculated empirically. Two observations error covariances can be directly calculated from the transformed perturbed measurements $q_i^j$:

$$C_{t-1,q}^{ij} = E[(q_i^j - E[q_i^j])(q_i^j - E[q_i^j])] \quad j = 1, 2,$$

(10)

where $i$ loops over all realizations in the respective set. Please note that the treatment of measurement error via a transformation overcomes the need to translate the diagonal matrix $R$ (Eq. 3) to the normal score domain. Likewise two cross-covariance matrices can be calculated for each subset:

$$C_{t-1,y}^{ij} = E[(u_i^j - E[u_i^j])(y_i^j - E[y_i^j])] \quad j = 1, 2$$

(11)
A significant speedup of the updating operation can be achieved by solving the assimilation equations first in the 'observation space' and then projecting the solutions into the attribute space (Evensen and van Leeuwen, 1996). A Cholesky decomposition of the two observation error matrices further results in considerable computational savings to solve following equations:

\[ C_{j'}^{(t-1,qq)} b_j^i = s_t - q_t^j \quad j = 1, 2. \]  

(12)

Here \( j' \) represents the ensemble that is complementary to ensemble \( j \). Both series of solution vectors \( b_j^i \) are subsequently combined with previously calculated cross-covariances to update the existing resource realizations:

\[ u_t^{ij}(x_n) = u_{t-1}^{ij} + C_{(t-1,uy)}^{j'} b_j^i \quad j = 1, 2. \]  

(13)

Finally, the resource attribute realizations are recombined and transformed back using the grid node specific anamorphosis functions which are stored in memory. To conclude this subsection, Figure 3 summarizes the mathematical formulation behind the interconnected parallel updating sequence (double helix).

### 3.3 Localization and change of support

The previous subsection already touched upon some of the issues related to the use of a finite Monte Carlo sample. From standard statistical theory, it further follows that the accurate estimation of weak correlations or covariances would require very large samples with thousands of realizations. When, due to practical considerations, the Monte Carlo samples are rather limited spurious correlations may arise between a sensor-based measurement and a remote grid cell. In order to avoid such a fallacy, a localization or neighbourhood strategy is applied, i.e. grid cells beyond a cut-off radius are not considered during an update. The previously outlined mathematical formulation thus only applies to grid cells \( x_n \) inside a local search neighbourhood. Furthermore, the use of a local neighbourhood greatly reduces the order the matrices in eq. 11 and 13 permitting a substantial computational economy to be realized.

Before concluding this section, one final practical concern needs to be addressed. Obviously sensor-based measurements are recorded on volumes substantially different from those of grid
cells or so-called resource blocks. The distributions of the Monte Carlo samples already reflect the support of the concerned random values, e.g., a distribution characterizing a resource block will generally be less variable than that of a predicted point measurement. As a result, differences in support are inherently accounted for through the empirical calculation of the covariances. Great flexibility is achieved since virtually every block or measurement involved can have a different support.

## 4 Updating of resource model based on measured truck loads

This experiment focuses on the reconciliation of 16 measured truckloads originating from a single mining block. Note that there are no limitations whatsoever to applying the algorithm when the truckloads do not originate from a single mining unit. The observations are simulated by blending nodes of the (known) true state. The resource realizations are obtained via a combination of unconditional simulations. Units are omitted to indicate the artificial nature of the resource models and observations.

![Figure 4: Synthetic ore body configuration and construction of the centre field for the initial set of realizations. The left panel shows the true field with semivariogram model 0 (range 120 m). The middle panel shows a realization of a random field characterized by model 1 (range 90 m). The right hand panel displays the centre field and is the sum of the left and middle panel.](image)

### 4.1 Generation of prior realizations and measurements

The synthetic experiment is performed on a two-dimensional 300 m x 300 m grid containing 90,000 nodes. The true configuration of the ore body is denoted by the symbol $z_t$ and is shown in the left hand side of Figure 4. The true state is characterized by an isotropic spherical semivariogram model with a range of 120 m (model 0).

A field $z_c$, representing the best available estimate of the true field, is generated by adding a random perturbation field to the true state: $z_c = z_t + \text{random field (model 1)}$. This perturbation field is characterized by a semivariogram model with a range of 90 m (model 1). The right hand side of Figure 4 displays the resulting central field $z_c$. Note that the central field deviates substantially from the true state.

The initial set of resource realizations are constructed according to the following equation:

$$\{z^i = z_c + \text{random field (model 1)} \mid i = 1 \ldots I\}, \quad (14)$$

where the random fields are generated using the same semivariogram structure as for the construction of the central field.

Subsequently, a reconciliation based on measurements from a single mining block is performed. The mining block is a 20 m x 20 m area, defined by (140, 90) x (160, 110), and is
further subdivided into 16 smaller 5 m x 5 m subblocks representing individual truckloads. The average value of each truckload is measured and recorded.

Figure 5 displays in red the real truckload observations. The 150 sets of predicted measurements are shown as blue dots. For completeness, the mean of the predicted measurements is computed for each truckload and plotted as well. Figure 5 illustrates that the best estimate calculated from the initial set continuously underestimates the true state in the selected region.

4.2 Reduction of uncertainty in the neighbourhood

To appreciate the improvements and the reduction of uncertainty after the data assimilation of the average truck load measurements, the prior and posterior attribute values along a cross section at $y = 100$ m are plotted. For all selected grid nodes, the mean and 95% confidence intervals are calculated. The resulting numbers together with the true values are plotted in Figure 6.

The region between 140 m and 160 m refers to the nodes which are located inside the...
sampled mining unit. After updating, the best estimate in this zone almost perfectly represents the true state. A perfect match is not to be expected since averages over larger volumes (5 x 5 nodes) are used to condition single nodes. Hence, the remaining uncertainty is not entirely reduced to zero. This observation can be attributed to a dispersion of many smaller nodes in a larger 5 m x 5 m truckload. The average value of the 16 sets of 25 nodes however do exactly match the real observed values after updating, while the corresponding uncertainty is reduced to zero (not shown).

In the 40 m long area east of the sampled region, the best estimate is considerably improved and matches the true state reasonably well. The reduction in uncertainty seems to be dependent on the distance away from the measured subvolumes. In the 40 m long area west of the sampled region, the best estimate improves as well but less significant than in the eastern region. However, a similar reduction in uncertainty can be observed.

5 Conclusion

This paper presents a practical methodology for assimilating sensor measurement of extracted material into the resource model to improve production forecasts. The methodology is based on a sequential application of cokriging which improves prior resource models by using linearly weighted sums of differences between sensor-based measurements and model-based predictions. The sequential application of cokriging results in a loss of stationarity in the covariance structure. To circumvent the explicit propagation and storage of the conditional covariances in time, the cokriging system was integrated into a Monte Carlo framework.

A novel algorithm was developed that deals specifically with the practical problems of the mineral industry, i.e. very large grids, non-linear measurements models, differences in scales of support, non-Gaussian data and observations resulting from a blend of several mining blocks. A Gaussian anamorphosis of grid nodes, sensor-based measurements and model-based predictions was implemented to deal with suboptimal conditions related to the absence of multivariate Gaussian data and non-linear relationships. The treatment of measurement error has been addressed. Further, an integrated parallel updating sequence was configured to reduce the statistical sampling error without the need of increasing the number of realizations. A neighbourhood search strategy is applied to constrain computation time and to avoid the emergence of spurious correlations. The integration of these modifications results in an efficient and extremely practical algorithm capable of handling the specific challenges of the mineral industry.

References


