

# The value of imperfect borehole information in mineral resource evaluation

Steinar L. Ellefmo and Jo Eidsvik

**Abstract** In mineral resource evaluation a careful analysis and assessments of the geology, assay data and structural data is performed. One critical question is where to position the exploration boreholes that render it possible to classify as much of the deposit as possible as a measured or indicated resource. Another important question is what method to use when analyzing the grade in the collected material. For the deposit we consider, a challenge is to assess whether one should analyze the collected core samples with accurate and expensive XRF equipment or the less accurate and less expensive XMET equipment. A dataset of 1,871 XMET and 103 XRF observations is available, along with relevant explanatory variables. At the 103 sites where XRF data is acquired, 103 XMET measurements are also available. We first derive estimates of the regression and covariance parameters of a Gaussian random field model for the log XMET and log XRF data. Next, the model is used to predict the decisive grade parameter on block support. To improve the predictions, the mining company has planned to drill and collect 265 core samples along new boreholes. The associated reduction in prediction variance, with XRF or XMET data collection, is studied. Moreover, we compute the value of the XRF or XMET information using the statistical model, the expected development costs and revenues. The value of information is a useful diagnostic here, comparing the actual price of the XRF or XMET data with its added value.

---

Steinar L. Ellefmo

Department of Geology and Mineral Resources Engineering, NTNU, Norway

Jo Eidsvik

Department of Mathematical Sciences, NTNU, Norway, e-mail: joeid@math.ntnu.no

Ninth International Geostatistics Congress, Oslo, Norway, June 11. – 15., 2012

## 1 Introduction

We analyze spatial data from a deposit in Norway. The main ore mineral is a particular oxide, and several exploration boreholes have been drilled to understand the spatial distribution of this mineral. The currently available data consist of about two thousand observations of the oxide along the boreholes. The deposit is still under consideration for mining, and the main purpose with this methodological paper is to evaluate different strategies for collecting more data. See also Eidsvik and Ellefmo (2012).

The oxide has been measured on crushed core samples using either a X-ray fluorescence (XRF) spectrometer in the laboratory or a portable X-ray meter (XMET). The XRF data are considered to be exact measurements of the oxide, providing perfect information at the locations where they are made. The analysis procedure is time consuming. The XMET data are considered to be a noisy observation of the true oxide level, providing imperfect information. These data are acquired more time efficiently and at a lower cost than the XRF data. We incorporate spatial dependence in the oxide by a Gaussian geostatistical model, and model the true oxide as a Gaussian random field.

In the mining industry resources are classified into measured, indicated or inferred, depending on the level of uncertainty. This is formalized through the JORC code. Since every deposit is unique, the assessment includes multiple levels of geological information, assay data, and structural data. From a more methodological viewpoint, we discuss several evaluation criteria in our context, computed at a set of resource blocks. The criteria are the reduction of the marginal variances, the increase in slope and correlation, the decrease in weight of the mean (Rivoirard, 1987) and the reduction of entropy. By today's standards the final classification is done by competent persons based on these criteria, and several other case-specific criteria.

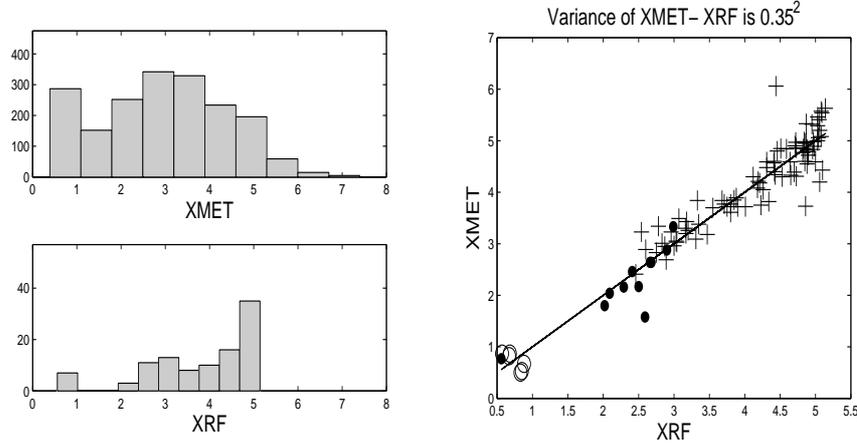
We also use the value of information (VOI) to study the potential of the different data collection schemes. The VOI relates the probabilistic model to the decision, using monetary units explicitly (Bhattacharjya et al., 2010).

## 2 Notation and data description

The deposit is about 2.5 km long, and is an intensively folded and lens formed body surrounded by mafic-felsic rocks. Geologists have defined three categories with increasing degree of mineralization. The classes are termed class 1, class 2 and class 3. Class 3 has the highest average oxide levels. The class 3 and class 2 categories are dominating in the central parts of the ore. The degree of mineralization is used as covariates in our analysis. We observe the oxide by XMET or XRF data acquired in boreholes. The laboratory XRF-data have been obtained from 10 meter long crushed sections (halves) of the core. The hand-held XMET data have been collected for every 25 cm of the core before crushing and aggregated into 10 meter long XMET-composites in correspondance with the XRF-analyses. There are 103 sites of XRF

data and 1871 sites of XMET data. The 103 locations with XRF data are also measured with XMET data. Therefore, this represents a partially heterotopic sampling scheme.

Figure 1 (left) displays histograms of the oxide data collected with XMET (top) and XRF (bottom), while Figure 1 (right) shows a crossplot of the 103 (XRF,



**Fig. 1** Left: Histogram of XMET (top) and XRF (bottom) observations. The oxide content ranges from 0.39 to 7.40 percent for XMET with a mean of 2.95. For XRF the range is from 0.56 to 5.14, with a mean of 3.80. Right: The XMET data (second axis) plotted against the XRF data (first axis) at 103 common sites.

XMET) measurement pairs. The degree of mineralization at the measurement location is also indicated in the figure: class 1 (o), class 2 (.) and class 3 (+). The highest measurements of the oxide are typically collected at locations with class 3 covariates. For the XRF histogram in Figure 1 (left, bottom) we might notice modes representing the different classes, but there is much variability within each class.

We next fit a joint geostatistical model for the XRF and XMET data. The XRF response ( $y_1(s)$ ) and XMET response ( $y_2(s)$ ) at a (north, east, depth) location  $s$  are modeled by

$$y_1(s) = x(s), \quad y_2(s) = x(s) + N(0, \tau^2). \quad (1)$$

The collection of XRF and XMET data is denoted by  $y = (y_1^t, y_2^t)$ , where

$$y_1 = (y_1(s_{1,1}), \dots, y_1(s_{1,103}))^t, \quad y_2 = (y_2(s_{2,1}), \dots, y_2(s_{2,1871}))^t.$$

Here, the true oxide at the (north, east, depth) location  $s$  is denoted  $x(s)$ . The XRF data provides perfect information about oxide at the location, while the XMET data is imperfect information of oxide with measurement noise variance  $\tau^2$ . The noise terms of the XMET observations are assumed to be independent from one location to another. We model the oxide as a Gaussian random field with expected

value  $\mu_x(s) = h^t(s)\beta$ , where  $h^t(s)$  includes a constant term and the mineralization covariate at site  $s$ , and  $\beta = (\beta_1, \beta_2)^t$  is a regression parameter. We choose a Matern covariance model to describe the spatial covariance structure of oxide:  $Cov(x(s), x(s')) = \sigma^2(1 + \phi h)\exp(-\phi h)$  for distance  $h = \|s - s'\|$ . Then the variance is  $\sigma^2$ , and  $\phi$  indicates the strength of spatial correlation.

Note that both datatypes are used together, in a joint Gaussian model, to predict the oxide grade at all spatial locations. Not dissimilar to what is typically done in cokriging, the current dataset and the joint modeling allows us to estimate the parameters and predict the grade in a unified way. The model parameters are assessed by maximum likelihood using the current XRF and XMET data.

### 3 Information criteria

The deposit in question can potentially be mined in an open pit, possibly going underground at a later stage. The decisions about opening the mine and choosing mining strategies depend on many modifying factors. We focus on quantitative approaches based on the geostatistical modeling. In particular we study the following criteria with only current data and with additional borehole XRF or XMET data:

- Kriging variance
- Slope and correlation of grade
- Weight of the mean grade
- Entropy
- Value of information

A resource will be classified into the different categories by a competent person. She or he is a member of a recognized professional organization and has sufficient relevant experience. The classification will be done based on a detailed understanding of the mineralization and on uncertainty indicators. What indicators to use are deposit specific and in practice the choice of the competent person, and we will not try to draw any conclusions here.

We define 3740 resource blocks of size  $20^3 \text{ m}^3$  inside a possible pit where we predict the oxide grade. The oxide grade in a block is denoted  $x = (x(s_{0,1}), \dots, x(s_{0,n_b}))^t$ , where  $n_b = 64$  is the block discretization in our case. The final grade estimate in a block is the average of the  $n_b$  in-block estimates.

The mining company considers acquiring more XRF or XMET data. About 20 new boreholes have been planned, giving 265 additional measurements of either XRF or XMET data. We denote this new data by  $z = (z(s_{z,1}), \dots, z(s_{z,265}))^t$ .

### 3.1 Spatial prediction and prediction variance

The joint distribution of oxide  $x$ , planned data  $z$ , and current data  $y$  is Gaussian. From the joint model we can compute the conditional mean of  $x$  and  $z$  given current data  $y$  is

$$\mu_{x|y} = H_x \hat{\beta} + C_{x,y} C_y^{-1} (y - H_y \hat{\beta}), \quad \mu_{z|y} = H_z \hat{\beta} + C_{y,z}^t C_y^{-1} (y - H_y \hat{\beta}), \quad (2)$$

where we simply plug in the regression parameter estimate  $\hat{\beta}$ . Moreover,  $C$  with subscripts indicate the fitted covariance or cross-covariance between  $x$ ,  $y$  or  $z$ . When we write out the regression parameter  $\hat{\beta}$  as a function of the data  $y$ , we get

$$\mu_{x|y} = (M_{x,y} S_y H_y^t + C_{x,y}) C_y^{-1} y, \quad \mu_{z|y} = (M_{z,y} S_y H_y^t + C_{y,z}^t) C_y^{-1} y, \quad (3)$$

where  $M_{x,y} = H_x - C_{x,y} C_y^{-1} H_y$  and  $M_{z,y} = H_z - C_{y,z}^t C_y^{-1} H_y$ , and  $H$  with subscripts indicate the matrix of explanatory variables.

Under the Gaussian modeling assumptions, the kriging predictor in (3) is optimal, i.e. the unbiased predictor with minimum variance. We can easily account for the uncertainty in the regression parameters  $\hat{\beta}$ . The resulting conditional covariance expressions are

$$C_{x|y} = C_x - C_{x,y} C_y^{-1} C_{x,y}^t + M_{x,y} S_y M_{x,y}^t, \quad C_{z|y} = C_z - C_{y,z}^t C_y^{-1} C_{y,z} + M_{z,y} S_y M_{z,y}^t, \quad (4)$$

where the last terms with  $M_{x,y}$ ,  $M_{z,y}$  and the variance of  $\hat{\beta}$  given by  $S_y$  compensate for the increased variability caused by estimating  $\beta$ .

Let the average block grade be  $\bar{x} = \sum_i x_i / n_b$ . We denote the Gaussian density of  $\bar{x}$  given  $y$  by  $\pi(\bar{x}|y) = N(\mu_{\bar{x}|y}, C_{\bar{x}|y})$ . This is computed for every resource block, and we define  $\text{Std}_y$  as the length 3740 vector of Kriging standard errors ( $\sqrt{C_{\bar{x}|y}}$ ), conditional on the current data  $y$ .

The procedures can be extended to include both current data  $y$  and  $z$ . The conditional mean of  $x$ , given both data, is

$$\mu_{x|yz} = \mu_{x|y} + C_{x,z|y} C_{z|y}^{-1} (z - \mu_{z|y}), \quad (5)$$

where  $C_{x,z|y}$  is the covariance of  $x$  and  $z$ , given  $y$ . Just like in the situation with only current data  $y$ , we define  $\text{Std}_{yz}$  as the Kriging standard errors of average block grades, now given both  $y$  and  $z$  data. The reduction in prediction variance depends on the locations of the new observations  $z$ , relative to each other, and to the current data  $y$  and the resource blocks  $x$ . Collecting XRF data in the planned boreholes provides a larger reduction in prediction variance than with XMET, but in general the uncertainty reduction is a complicated function of the covariance parameters.

### 3.2 Slope, correlation and weight of the mean

Rivoirard (1987) uses the regression between the predicted and true block grades, called the slope, to assess the effects of different Kriging neighborhoods. This criterion has also been used to quantify the degree of measured or indicated resources in mining. For each resource block we have

$$\text{Slope}_y = \text{Cov}(\bar{x}, \mu_{\bar{x}|y}) / \text{Var}(\mu_{\bar{x}|y}) = (w^t G_{x,y} C_y^{-1} C_{x,y}^t w) / (w^t G_{x,y} C_y^{-1} G_{x,y}^t w), \quad (6)$$

where  $w^t = 1^t / n_b$  and 1 is a vector of ones. Moreover,  $G_{x,y} = M_{x,y} S_y H_y^t + C_{x,y}$  is recognized in (3). The correlation is a normalized version of the slope;

$$\text{Corr}_y = \text{Corr}(\bar{x}, \mu_{\bar{x}|y}) = \text{Slope}_y \cdot \sqrt{\text{Var}(\mu_{\bar{x}|y}) / \text{Var}(\bar{x})}. \quad (7)$$

The weight of the mean (Rivoirard, 1987) is another useful quality indicator in kriging. Given data  $y$ , the weight of the mean is interpreted as the relative impact of the regression, compared with that of the simple Kriging predictor  $C_{x,y} C_y^{-1} y$ . From the prediction formula in (3) we recognize the simple Kriging predictor as the last term, and the regression effect in the first part. When the deposit is more densely sampled, the second term will dominate over the first term. We have

$$\text{Weight}_y = (w^t M_{x,y} S_y H_y^t C_y^{-1} 1) / [(w^t M_{x,y} S_y H_y^t C_y^{-1} 1) + (w^t C_{x,y} C_y^{-1} 1)]. \quad (8)$$

The slope, correlation and weight of the mean are computed for each resource block. In total, they can be represented as length 3740 vectors, with one value for each resource block. They can be defined similarly conditioning on both current data  $y$  and the new data  $z$ . When we get more accurate predictions of the grade, the slope is closer to 1, the correlation is closer to 1, while the weight of the mean is closer to 0. The effect is expected to be clearer with perfect information (XRF) than with imperfect data (XMET).

### 3.3 Reduction of entropy

We compare the information content using the entropy given current data  $y$  and prospective data  $z$ . Now let  $x^*$  denote the oxide variable at the center of each of the 3740 resource blocks. The entropy (disorder) decreases with more information. It is defined as the negative expected value of the log density. For a Gaussian  $\pi(x^*) = N(\mu, \Sigma)$  we have entropy

$$\text{Ent}(x^*) = - \int \pi(x^*) \log \pi(x^*) dx = \frac{n}{2} \log(1 + 2\pi) + \frac{1}{2} \log |\Sigma|. \quad (9)$$

The entropy reduction when acquiring the new data  $z$  becomes  $\delta \text{Ent} = \text{Ent}(x^*|y) - \text{Ent}(x^*|y, z) = \frac{1}{2} (\log |C_{x^*|y}| - \log |C_{x^*|yz}|)$ , where  $C_{x^*|y}$  is the covariance at all resource

blocks given  $y$ , while  $C_{x^*|yz}$  is conditional on both  $y$  and  $z$  data. The determinant expressions can be evaluated before the actual data are collected.

### 3.4 Value of information

The VOI is the maximum monetary amount a decision maker should pay to collect data. In our context there are two levels of decisions. The downstream decision is whether to open the mine or not. This question is incorporated to solve for the second level of decisions; whether one should collect XMET or XRF data in the planned boreholes, or no further data.

The VOI is defined as the difference between prior and posterior value:

$$\text{VOI} = \text{PoV} - \text{PV}. \quad (10)$$

We decide to purchase the data  $z$ , either XRF or XMET in planned boreholes, only if the VOI is larger than the price of data acquisition. The XRF data is perfect information, and the VOI of XRF is always larger than the VOI of the imperfect XMET data. However, the XRF data has a higher price than the imperfect XMET data. For the Gaussian model, the prior and posterior value can be computed analytically (Eidsvik and Ellefmo, 2012).

## 4 Results and discussion

We now evaluate the planned boreholes using the different information criteria .

Table 1 shows the values of different evaluation criteria. The XRF data are of

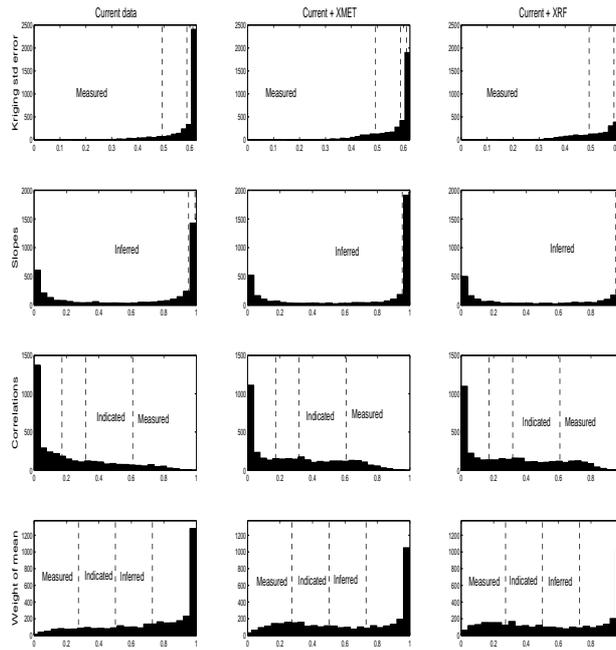
**Table 1** Various evaluation criteria: Distance to nearest borehole, Kriging std, slope, correlation, weight of mean and entropy, using current data, and with planned XRF and XMET borehole data. The distance, standard error, slope and weight of the mean are vectors for all resource blocks. Here, we display the averages over all blocks.

	distance	std	slope	corr	weight of mean	entropy
Current data	55.1	0.59	0.62	0.21	0.73	Ent=-2930
XMET data	48.0	0.57	0.69	0.27	0.63	$\delta\text{Ent}=43.0$
XRF data	48.0	0.57	0.70	0.29	0.62	$\delta\text{Ent}=72.4$

course more informative than XMET data, and for some criteria we clearly gain some by acquiring XRF instead of XMET data. For instance, the reduction in entropy is almost twice as large when collecting XRF. Of course, a pure distance criterion does not separate between XRF and XMET in the new boreholes. For the Kriging std the average difference between XRF and XMET collection is minis-

cule. There is a slight improvement in the slope, correlation and weight of the mean criteria, but the added value of XRF, compared with XMET, is small considering the reduction from the current data.

Figure 2 illustrates the variability in the Kriging standard error, slope, correlation and the weight of mean at the 3740 resource blocks. The histograms show current values (left), with XMET data (middle) and with XRF data (right). Clearly, more data pushes the histogram of the standard errors (top) towards smaller values, the slope and correlation (middle) to higher values, and the weight of the mean (bottom) to smaller values. Thus, at many resource blocks there is clearly added information



**Fig. 2** Histogram of evaluation criteria at all resource blocks. Kriging standard errors (top), slope (middle, top), correlation (middle, bottom) and weight of the mean (bottom). The left displays are based on current data, middle displays on current data and XMET in planned boreholes and right displays on current data and XRF in planned boreholes. The vertical dashed lines are the criteria-based separation of measured, indicated and inferred resources.

in the planned borehole data. The improvement going from XMET to XRF is visible for resource blocks close to the planned boreholes, but not far away from these locations. In fact, the Kriging prediction errors have larger variability after conditioning on more information. Of course, the planned data acquisition is guided to the spatial

domains of most interest, and the reduction of uncertainty is highest where we want to predict the grade accurately.

A resource classification is based on multiple criteria and experience of the local geology, usually evaluated by a so-called competent person. Here, we simply compare the presented geostatistical criteria and classify based on thresholding. It is not an attempt to do a resource classification in compliance with the JORC-code. The categorization limits are obtained from the currently available data using geometric considerations as follows: For each resource block we compute the azimuth angles and distances to the five nearest borehole measurement locations. These are used to group the resource blocks in four categories: Category 1: The fifth closest point is within 30m and the standard deviation of the azimuth angles to data locations within 100 m is between 80 and 130 degrees. Category 2: The fifth closest point is between 30m and 60m and the standard deviation of the azimuth angles to data locations within 100 m is between 80 and 130 degrees. Category 3: The fifth closest point is within 60m and 200m and the standard deviation of the azimuth angles to data locations within 100 m is between 80 and 130 degrees. Category 4 is defined by the remaining resource blocks. The azimuth variability condition ensures that there are proximal measurements in more directions, not only one borehole. Given this categorization of resource blocks, the 75 percentiles of all criteria are computed from the Kriging errors, slopes, correlation and weight of the mean in resource blocks belonging to each category. These values define the thresholding values for measured, indicated and inferred. They are displayed by vertical dashed lines in Figure 2. Recall that this is based on the current boreholes. The same thresholds are next applied for the planned data as well.

The categorization we have done here is used to study the information content in the new data and allows us to compare the methodologies. In particular, we aim to study the effects of XMET and XRF data acquisition in the planned boreholes. The geometric criteria based on distances and angles is easy to understand, but it is not useful to compare the XRF and XMET data, since they are equally informative in terms of distances and angles.

In Table 2 we show the resulting tonnages in the measured, indicated and inferred categories. Here, the resource blocks falling in the measured, indicated and inferred categories are converted to tonnes of resource. The block volumes outside the ore are not included in the calculation. We use a cut-off value of 2.5 % (based on current data) to separate waste from ore. With the current data, using the Kriging standard error as criterion, there are about 9 million tonnes of measured resource and 20 million tonnes indicated. There are only slight variations between the criteria using our thresholding method. Obviously, with more data, there are more resource blocks in the measured category. When we collect XMET data in the planned boreholes, the measured category in Table 2 has around 14-15 million tonnes. The indicated category is around 21 million tonnes. Some blocks have gone from indicated to measured, while others have gone from inferred to indicated. The sum of measured and indicated resources is close to 40 million tonnes. Collecting XRF data in the planned boreholes gives only slightly larger numbers in the measured category: 15-17 million tonnes. Note that the pure geometric distance criterion has the same number as

**Table 2** Resource classification (in million tonnes) based on current data, XMET data in planned boreholes and XRF data in planned boreholes. The measured, indicated and inferred classification is done from thresholds in different evaluation criteria: Distances, kriging standard deviations, slope, correlation and weight of mean.

	Distance	Kriging Std	Slope	Corr	Weight
	Current data				
Measured	9.0	8.8	8.8	8.8	8.5
Indicated	20.8	21.3	21.9	21.2	21.4
Inferred	10.6	10.4	9.8	10.5	10.6
	Current data and XMET in planned boreholes				
Measured	13.8	14.2	15.1	14.2	15.2
Indicated	21.6	21.5	21.0	21.4	20.0
Inferred	5.1	4.8	4.5	5.0	5.4
	Current data and XRF in planned boreholes				
Measured	13.8	15.0	16.0	15.0	16.7
Indicated	21.6	21.0	20.4	20.8	18.7
Inferred	5.1	4.6	4.2	4.8	5.1

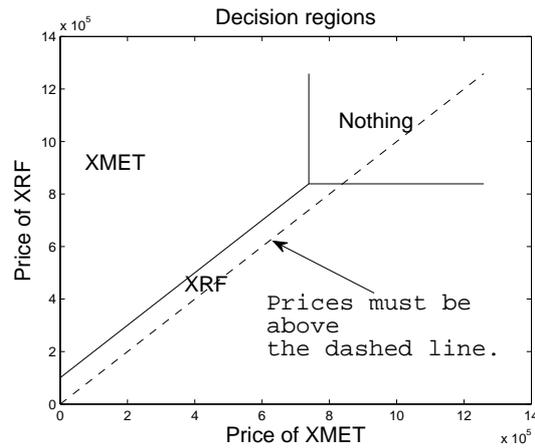
for XMET (13.8 million tonnes measured), since it uses no uncertainty modeling. The indicated resource blocks are about 19-21 million tonnes. In summary, there is a clear increase in measured tonnages going from current to XMET, but not such an improvement when collecting XRF data instead of XMET.

Recall that these numbers are based on our subjective criteria. A real-life resource classification would have been based performed by a competent person in compliance with the JORC-code or other similar codes.

In order to assess the value of XMET and XRF information, we specify revenues, costs, processing parameters and tonnages. This allows us to compute the prior and posterior value, and the VOI (Eidsvik and Ellefmo, 2012). We do this for XMET and XRF in the planned boreholes. In Figure 3 we show the decision regions as a function of XMET and XRF data acquisition prices. The decision regions are computed by selecting the data type that gives the largest added value, compared with the price of data. This entails a selection rule of:

$$\text{Decision} = \operatorname{argmax} \{ \text{VOI}_{\text{XRF}} - \text{Price}_{\text{XRF}}, \text{VOI}_{\text{XMET}} - \text{Price}_{\text{XMET}}, 0 \}, \quad (11)$$

where we decide to purchase XRF if  $\text{VOI}_{\text{XRF}} - \text{Price}_{\text{XRF}}$  is the highest element in the length three vector in (11). If none of the first two entries are positive, we decide to purchase no more data. In our situation, the actual prices of XRF and XMET mean that we are just within the 'Nothing' region. Given that the already defined tonnages were enough to open the mine, we would decide not to purchase more data. Recall that the price of XRF is always higher than the price of XMET, and the relevant price ranges are above the straight line in Figure 3. XRF data is the most lucrative data type for very low laboratory prices. For more expensive laboratory analysis, XMET data is preferable.



**Fig. 3** Decision regions. Whether to purchase full XMET or XRF data, or nothing. The first axis represents the price of XMET data. The second axis is the price of XRF data.

## 5 Closing remarks

We have presented a unified geostatistical model for XMET and XRF data used in mining exploration. Several criteria for uncertainty reduction are discussed, and we apply them for resource evaluation on a case study from Norway.

## 6 Acknowledgments

We acknowledge the mining company for letting us use their data.

## 7 References

- Bhattacharjya D, Eidsvik J, Mukerji T (2010) The value of information in spatial decision making. *Mathematical Geosciences* 42: 141-163.
- Eidsvik J, Ellefmo, SL (2012) The value of information in mineral exploration. Technical report, 2, Department of Mathematical Sciences, NTNU, Norway.
- JORC (2004) Australasian Code for Reporting of Exploration Results, Mineral Resources and Ore Reserves. 2004 Edn.
- Rivoirard J (1987) Two key parameters when choosing the Kriging neighborhood. *Mathematical Geology* 8: 851-856.