Evaluating Components of a Multi-Source Output

J.E. Everett

Faculty of Economics and Commerce, University of Western Australia, Nedlands 6009

jim.everett@uwa.edu.au

ABSTRACT

In production problems with output from linear combination of multiple inputs, operators commonly believe single-input trials are required to evaluate individual contributions. Provided input proportions vary, zero-intercept linear multiple-regression can establish the component contributions. Originally applying the method to a laundry serving multiple hospitals, establishing unit costs for each type of clothing, the author has used it for several mining problems. In the iron ore industry, multiple ore types are crushed and blended before assay. The method allows the composition, and lump/fines split, to be estimated for each ore type. In another application, ores from multiple mines are combined at the port to form a product for shipping. Assays on loading are compared with the combined mine assays for each shipload, and the method provides estimates of the assay bias for each source mines. Further regression analysis applied to the residuals allows the error variances for the source mines to be estimated.

INTRODUCTION

A common occurrence in mining operations (and in other fields of endeavour) is a situation where a product is formed as a blend from multiple sources. Each source may have different properties, and the properties of the blend are the linear composite of the properties of the various sources. With mineral production, the properties may be the grades or behaviour of various ores. For example, in iron ore processing, the grade difference between lump and fines products may differ between various ore sources. The problem is not limited to mineral processing, but may occur in any situation where the total property is a linear combination of the source contributors.

The properties may be physical properties (such as mineral grade). They may equally be measurement errors: for example, ore sent from various mine sites is assayed at each mine site laboratory, stockpiled as a blend at the port and assayed again on loading each ship. The discrepancy (and error variance) between the grade predicted from the
mine assays and the grade assayed on shipment may vary between the mines, since the laboratories at some mines may be more reliable than the laboratories at other mines.

The properties being studied may also be cost (or profit) contributions.

The method of analysis to be described in this paper was first used by the author and a colleague in a cost analysis study done for the West Australian Linen and Laundry Service (Everett & McLeod, 1991). The West Australian Linen and Laundry Service provided a non-profit laundry service for a number of government institutions, including hospitals, old people’s homes and other locations. Each day a mix of items (such as bed sheets, nurse’s uniforms or operating theatre linen) was washed. It was desired to charge each institution with the cost of services provided, and this required the estimation of a unit cost for each type of laundry processed. The cost of each day’s operation was known. Since it was not feasible to run only one type of laundry through for any one day (and the operation would have been inefficient if this had been done) the unit costs had to be estimated from the records of what mix had in fact been laundered each day. Since each day’s total cost is a linear composite of the costs to be ascribed to the items laundered, and the mix varied from day to day, it was possible to carry out a linear regression, with zero intercept.

Many years later, in consulting to mining companies, the author began to meet similar situations. For example, in iron ore processing, low-grade ores are put through a “beneficiation” process, where gravity separation is used to enrich the ore. There are performance records for each day’s production, showing the beneficiation achieved for the mix of ores for that day. It is known that the beneficiation varies between ore types. The standard method had been to run the plant for a day treating only one ore type, and use the results to predict results for future mixes. This established method had a number of shortcomings:

• It was extremely disruptive and costly to run the beneficiation plant for a full day with only one ore type of feed.

• The feed might behave differently if only the one ore type of feed were running through the plant.

• At the end of the day’s trial, the results would be statistically inadequate, since there may be considerable random variation in performance.

Since there were copious records of the day-by-day performance of the beneficiation plant, with varying feed mixes, the zero-origin regression analysis was able to be applied to extract estimates (and estimates of standard error) for the performance of each ore type.
Iron ore is mined, then crushed and separated into lump and fines components. Before it is mined, its head grade (total composition) is estimated by taking samples while drilling blast holes. These blast hole samples have systematic and varying bias, because they tend to undersample the finer material and thus overestimate the iron content and underestimate the contaminants. The blast hole samples also give no estimate of how the ore grade will split between its lump and fines products (lump is usually richer in iron, and lower in contaminants). The final composition, for lump and for fines, is not known until trainloads or batches (containing a mixture of ore types) have been crushed and analysed. Both the systematic error and the lump-fines split differ between ore types. The zero-origin regression method was again used successfully to estimate both the required blast estimate adjustment and the lump-fines grade split for each of the ore types. Howard, James & Everett (2005) describe the lump-fines grade split estimation using this method.

Another iron ore application, which will be described here in greater detail, involved the comparison of laboratory errors. In one company’s operation, ore from each of a number of mines is railed to the port, and used to build blended stockpiles. The completed stockpiles are reclaimed and used to load ships for export. The ore is assayed as it is crushed and despatched from each mine, and again at the port when it is reclaimed from stockpiles for shipment. Each shipment thus has a port assay, and a mine assay. The mine assay is a combination of the assays from the multiple mines that contributed ore to the shipment. The corresponding port assays and mine assays differ. The discrepancy could be caused by errors in the mine assays or the port assays, or both. However it was accepted that the port assays were more reliable, and that the various mine laboratories might be more or less reliable. The zero-origin regression method was used to assess the reliability of the individual mine laboratories, for both systematic and random error. To demonstrate the effectiveness of the method, a simulated data set was also constructed. Hypothetical mine data and port data were generated, of known bias and random error. The method was applied to the simulated data set, to demonstrate that the bias and random error could be accurately recovered. The rest of this paper will describe the regression model, and its application to the simulated and real data set, to investigate the mine and port laboratory errors.

ANALYSIS

The Problem

Ships are loaded at the port with ore blended from seven source mines. Each stockpile at the port is built from ore that may be contributed by multiple mines. The grades for each mine contribution are reported, based on the mine assays. Each blended stockpile therefore has a ‘Mine’ assay, being the weighted composite of contributions from each mine. Each shipment is
sampled to give a ‘Port’ assay, and may be sourced from one or more blended stockpiles, each with its own ‘Mine’ assay, compounded to give a ‘Mine’ assay for the shipment. In this study the shipment Port assays are compared with the corresponding Mine assays to establish the mean bias and standard deviation of the discrepancy between the mine and port assays. It should be noted that both port and mine assays may be in error, but from the available data we can only estimate the discrepancy between them.

The Model

Assume the (Port-Mine) error for the \( k^{th} \) mine is normally distributed, \( N(m_k, s_k) \), for assays taken over a shipment tonnage. The discrepancy between the port and the \( k^{th} \) mine thus has mean \( m_k \) and variance \( s_k^2 \), both of which we would like to estimate.

For the \( j^{th} \) shipment, \( j = 1 \ldots J \), let \( D_j \) be the discrepancy between port and mine assays.

\[
D_j = \text{Port assay} - \text{Mine assay} = N(m_j, s_j)
\]  

(1)

Let \( p_{jk} \) be the proportion of the \( j^{th} \) shipment coming from mine \( k \), where \( k = 1 \ldots K \).

\[
m_j = \sum_k p_{jk} m_k
\]  

(2)

\[
s_j^2 = \sum_k p_{jk} s_k^2
\]  

(3)

So:

\[
D_j = \sum_k p_{jk} m_k + e_j
\]  

(4)

\[
\text{Variance}[e_j] = \text{Exp}[e_j^2] = s_j^2 = \sum_k p_{jk} s_k^2
\]

\[
e_j^2 = \sum_k p_{jk} s_k^2 + d_j
\]  

(5)

Equations (4) and (5) are linear regression models, with zero intercept, and can be solved to estimate \( m_k \) and \( s_k^2 \) for each mine \( k = 1 \) to \( K \).

The data for the regression of Equation (4) are the (Port-Mine) assay discrepancies \( D_j \).

The data for the regression of Equation (5) are the squared residuals \( e_j^2 \) from the regression of equation (4).

The regression analysis also provides standard errors of the estimates. For the tabulated and graphed results, the standard errors of the variance are used to compute upper and lower bounds of the standard deviation. The ranges (plus or minus one standard error) represent 67% confidence limits.

It should be emphasised that the parameters estimated by the regression models are the mean and standard deviations of the discrepancies between each mine and the port assays, over a shipment
tonnage. The estimates of the mean and standard deviation parameters each have a reported standard error, dependent on the variability and on the quantity of data available for each mine.

**Simulated Data**

To test the model, a program was written to generate a set of simulated data, of known $m_k$ and $s_k^2$, for seven mines, $k = 1...7$. These data were then subjected to the regression analysis, to find how accurately $m_k$ and $s_k^2$ were estimated. The simulation was designed to replicate the way trainloads of ore from the mines are railed to the port and used to build stockpiles that are reclaimed to fill shipments.

The specifications for one run of the simulation model are shown in Table 1. Trainloads of 24kt from each of six mines are railed to the port, and arrive at random in the proportions specified. Each 168kt stockpile is built from seven trainloads, and then reclaimed to fill 56kt shipments. Assays for trainloads from each mine have systematic biases and standard deviations of measurement around their “true” values, as shown in Table 1. The port assays could also be ascribed a bias (here set to zero) and a random standard deviation around the true values.

<table>
<thead>
<tr>
<th>Mine Source</th>
<th>Assay Bias</th>
<th>Standard Deviation</th>
<th>Proportion of Trains</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.05</td>
<td>0.10</td>
<td>15%</td>
</tr>
<tr>
<td>2</td>
<td>0.30</td>
<td>0.20</td>
<td>15%</td>
</tr>
<tr>
<td>3</td>
<td>0.15</td>
<td>0.50</td>
<td>15%</td>
</tr>
<tr>
<td>4</td>
<td>0.00</td>
<td>0.10</td>
<td>15%</td>
</tr>
<tr>
<td>5</td>
<td>0.20</td>
<td>0.20</td>
<td>20%</td>
</tr>
<tr>
<td>6</td>
<td>0.40</td>
<td>0.60</td>
<td>20%</td>
</tr>
</tbody>
</table>

**Table 1: Specification of the Simulation Model**

The simulation was run to send a total of 21,000 trainloads to the port, to build 3,000 stockpiles, which were reclaimed to fill 9,000 shipments. For each shipment, (Port-Mine) assay discrepancies were calculated using the statistical parameters as specified (assuming normal distributions), and the proportion coming from each mine was recorded.

The zero-intercept regression models of equations (4) and (5) were applied to the 9,000 (Port-Mine) assay discrepancies, to estimate the bias and variance parameters for each mine, and the standard errors for each parameter.

Figure 1 shows the simulation results for the (Port-Mine) bias. It demonstrates that the regression procedure does a good job of recovering the (Port-Mine) bias for each mine. It should be emphasised that the method is unable to sort out the port bias from the mine bias, but merely shows the bias for each mine relative to the port assays.
To bring the unit of estimation back to trainload scale, the generating standard deviation is:

$$\text{Generating Standard Deviation} = \sqrt{\text{Mine Variance} + 7*\text{Port Variance}}$$  \hspace{1cm} (6)

Figure 2 shows the results after applying the regression of equation (5) to the squared residuals of the regression (4). To enable comparison with the generating standard deviation of equation (6), the discrepancy variances for shipments had to be multiplied by seven to bring them back to trainload scale.

Figure 2 shows that the regression method does a reasonably good job of recovering the mine variances, although the standard errors of the estimated standard deviations are considerably greater than for the bias estimates.

It should be noted from equation (6) that, unless the port variance is much smaller than the mine variance (as in this simulation example), it would swamp the mine variance.
Applying the Model to Real Data

Having established that the method is reliable, at least for estimating (Port-Mine) bias, it was applied to a set of real data where we had Port and Mine measurements for 2,600 shipments, with estimates of what proportion of each shipment came from each of nine source mines.

The estimated biases for Fe grade are plotted in Figure 3, along with the standard errors for each estimate. If standard errors do not overlap, we can say that the estimated biases are significantly different.

Statistically significant differences exist between the biases for the nine mines. In particular, mine B is very much out of line. The results prompted an investigation into the methods used in that mine’s laboratories. A sample preparation problem was identified.

Similar analyses for the other minerals provided results that enabled the various mine laboratory methods to be evaluated and in some cases revised.
The standard deviations of the (Port-Mine) discrepancies were also estimated using the regression model of equation (5). Since stockpiles averaged 130 kt, the standard deviation estimates (based on shiploads taken from stockpiles) were brought back to the 24 kt trainload scale by multiplying by \(2.33 = \sqrt{130/24}\).

Figure 4 shows the results for the estimates of the (Port-Mine) standard deviation for trainloads from each mine.

It is clear that mine D assays best conform to the port assays. The bias is very small (-.08) and the standard deviation of the discrepancies is also less than 0.1, considerably less than for any of the other mines.

Similar regression analyses were carried out, not only for Fe but also for the contaminant minerals silica, alumina and phosphorus. The results identified the mines having larger biases and larger standard deviations in their discrepancies between their assays and the port assays, and focussed attention on several areas where the sampling preparation and assaying could be improved.
CONCLUSION

The examples considered have illustrated how zero-intercept regression techniques can be used to estimate individual contributions to a blended or mixed result. The results are statistically much more powerful than single trials. The regression method described is generally more feasible and far less costly than running individual trials for each contributor. The examples have shown how the method is equally applicable to cost analysis, to measurement of physical properties, and to error analysis.

REFERENCES
