

Calculating the dilution between successive trainloads of iron ore during processing

J. E. Everett

*Faculty of Economics and Commerce,
The University of Western Australia, Australia*

Abstract

Iron ore is railed several hundred kilometres from inland mines to port, where it is crushed and separated into lump and fines products for shipment to customers. During crushing, each trainload is sampled and assayed to measure the grade, in iron, phosphorus, silica, alumina and a number of other minerals. Trainloads from different mines are of systematically different grade. Accurate product grades for each trainload, from each mine, are required for production planning and quality control. During crushing each trainload is contaminated with ore from the preceding and possibly from the following trainloads, so that the reported grade is biased by this dilution from the neighbouring trainloads. Although the dilution effect has long been recognised, it has not previously been quantified and therefore has not been corrected for. This paper describes the development, verification and use of a non-linear regression model enabling the amount of dilution to be estimated, so that the diluted grades can be corrected to undiluted grades.

Keywords: quality control, decision support, mining, non-linear regression, weighted least squares.

1 Introduction

The quality of the ore product depends upon it closely matching target composition, not only in iron, but also in phosphorus, silica and alumina. The four critically important composition elements will be referred to as the vector $\{\text{Fe}, \text{P}, \text{SiO}_2, \text{Al}_2\text{O}_3\}$. A number of other elements are also monitored. The customer blast furnaces are tuned to receive ore of the agreed target composition. Issues relating to quality control in iron ore production are discussed in [1,2].



Iron ore is railed from five mines of differing average grades. At the port, run of mine trainloads are processed sequentially and separated into lump and fines products. Each trainload is sampled during processing. The lump and fines products are stacked onto lump and fines stockpiles, blended from the five mines

The mine grades vary around consistently different means. For example, average Fe grade is considerably higher for Mine A than for the others. It has long been observed that trainload assays sampled at the port are biased according to the mine sources of trainloads processed before and after it. For example, a Mine A trainload followed and preceded by trainloads not from Mine A has lower Fe content than if its neighbours are other Mine A trainloads.

The effect is because ore from the neighbouring trainloads contaminates the beginning and end of each trainload sample, during dumping, screening and crushing. As each trainload is processed, some material mixes into the next trainload. Each trainload sample has a sharp beginning and end, but the sample contains some material from the following and/or preceding trainloads.

The dilution does not affect the estimated composition of the final product, because dilution of each trainload is compensated by dilution of its neighbours. But it does affect the reported composition of ore from each contributing mine, and this systematic error is important because it interferes with the planning and quality control for each mine. For each mine, each day's production is planned to satisfy composition targets, based on mine estimates of the currently available ore sources. Comparison between mine estimates and subsequent port assays are used to monitor and adjust the mine estimates: this adjustment is compromised if the port assays have error, depending on which type of trainload preceded or followed them.

A non-linear regression model is described, to estimate the average dilution from the preceding and following trainloads. It is shown that the lump and fines samples for 12 kilotonne trainloads are each contaminated by nearly one kilotonne from the preceding trainload and a median of about 450 tonnes from the following trainload.

The results of the analysis are now being routinely used to provide more accurate trainload grade data.

2 The model

Consider a variable X , with mean M ,

$$E[X_n] = M \quad (1)$$

$$X_n = M + d_n \quad (2)$$

where d_n is the deviation for the n th observation X_n .

M can be estimated as the average of X_n . But it can equally well be estimated as the value that minimises $\sum d_n^2$. The two formulations are equivalent, giving the same estimate for M , but the second formulation is a simple regression model.



Now consider the grade X_n for a trainload portion, which may come from any one of the five mines. It has total tonnage W_n kt, of which p kt comes from the previous trainload portion, and q kt comes from the following trainload portion. The expected value of X_n is now:

$$E[X_n] = (pM_p + (W_n - p - q)M_c + qM_q) / W_n \tag{3}$$

$$X_n = (pM_p + (W_n - p - q)M_c + qM_q) / W_n + d_n \tag{4}$$

where M_p, M_c, M_q are the mean values for the previous, current and following trainload portions, each being M_A, M_B, M_C, M_D or M_E , according to which of the five mines is its source.

Define \underline{M} as the vector $\{M_A, M_B, M_C, M_D, M_E\}$, and unit vectors $\underline{S}_p, \underline{S}_c, \underline{S}_q$ to specify the source mines of the preceding, current and following trainload portions. For example, if the preceding trainload portion comes from Mine B, then \underline{S}_p is $\{0, 1, 0, 0, 0\}$. Thus:

$$M_p = \underline{S}_p \cdot \underline{M}'; M_c = \underline{S}_c \cdot \underline{M}'; M_q = \underline{S}_q \cdot \underline{M}' \tag{5}$$

and equation (4) becomes:

$$X_n = (p\underline{S}_p + (W_n - p - q)\underline{S}_c + q\underline{S}_q) \cdot \underline{M}' / W_n + d_n \tag{6}$$

Equation (6) is a non-linear regression model.

Ordinary least squares (OLS) regression minimises $\sum d_n^2$. Weighted least squares regression (WLS) weights the variance terms to minimise $\sum W_n d_n^2$. This can be achieved by multiplying both sides of equation (6) by $\sqrt{W_n}$:

$$Y_n = X_n \sqrt{W_n} = (p\underline{S}_p + (W_n - p - q)\underline{S}_c + q\underline{S}_q) \cdot \underline{M}' / \sqrt{W_n} + e_n \tag{7}$$

The OLS solution to equation (7) is thus a WLS solution to equation (6), providing estimates of the dilution factors p and q and the mine means, \underline{M} .

3 Data used and its preparation

The analysis to be reported here is based upon a total of 9,403 ‘‘Port Actual’’ records, for ore crushed during the period of about ten months.

Each ‘‘Port Actual’’ record reported a trainload portion of lump or fines from a total of 2,468 trainloads, being 1,471 trainloads from Mine A, 548 from Mine B, 279 from Mine C, 108 from Mine D, and 62 from Mine E.

Each trainload portion was processed through one of two buildings, ‘‘TCB1’’ or ‘‘TCB2’’. Many trainloads were split between TCB1 and TCB2, with a trainload portion going to each. Successive trainload portions of the same



product (Lump or Fines) of the same trainload going through the same building were merged. For each building and product, the trainload portions were ordered by start time.

For each trainload portion record, five fields were added for the each of the unit vectors \underline{S}_p , \underline{S}_c and \underline{S}_q , to specify the preceding, current and following mine, as defined above. Each vector had four zero components, and a single unit component. Consider:

$$\underline{S}_p = \{0, 1, 0, 0, 0\}; \underline{S}_c = \{1, 0, 0, 0, 0\}; \underline{S}_q = \{0, 0, 0, 0, 1\} \quad (8)$$

In the example of equation (8) the preceding, current and following trainloads come from Mines B, A and E respectively.

Each trainload portion had a grade vector \underline{G} , defining its assayed composition. Assays were available not only for the four control minerals {Fe, P, SiO₂, Al₂O₃}, but also for {H₂O, LOI, Mn, TiO₂, CaO, MgO, S, K₂O}. LOI is “loss on ignition”, a measure of moisture content.

$$\underline{G} = \{\text{Fe, P, SiO}_2, \text{Al}_2\text{O}_3, \text{H}_2\text{O, LOI, Mn, TiO}_2, \text{CaO, MgO, S, K}_2\text{O}\}$$

The vector fields $\underline{\text{Prod}} = \{\text{L(ump), F(ines)}\}$ and $\underline{\text{Plant}} = \{\text{TCB1, TCB2}\}$ were added, specifying the product (lump or fines) and plant building for each trainload portion. For example, a lump trainload portion going through TCB2 has $\underline{\text{Prod}} = \{1, 0\}$ and $\underline{\text{Plant}} = \{0, 1\}$.

Fields T_p and T_q were calculated, as the time interval (in minutes) before the start and after the end of each trainload portion.

4 Discriminant analysis

Solution of the non-linear regression equation (7) requires that the mean values of X differ significantly between the five mine sources.

The model could be run repeatedly, for each mineral, giving repeat estimates for the dilution. It is more efficient to let X be the linear composite of the minerals that best separates the sources. The “best” separation is defined as maximising the between-group variance divided by the within-group variance. The linear composite satisfying this condition is known as the first discriminant function. The second discriminant function again maximises the between-group variance divided by the within group variance, subject to the constraint that the second discriminant function is orthogonal (uncorrelated) to the first discriminant function. A useful description of discriminant analysis is supplied in the SPSS™ reference guide [3].

With five groups, four orthogonal discriminant functions can be extracted. We are interested only in the first discriminant function, since it has the greatest possible power to discriminate between the five groups. The statistical package

SPSS™ was used to apply discriminant analysis. The first discriminant function scores were added to the data file. The discriminant analysis was run separately for lump and for fines.

The correlations between the mineral grades and a discriminant function are referred to as the “loadings” of the mineral grades on the discriminant function. The loadings for the first discriminant function, for lump and for fines, are tabulated below.

Table 1: Loadings on the first discriminant function.

Mineral	Fe	P	SiO ₂	Al ₂ O ₃	H ₂ O	LOI	Mn	TiO ₂	CaO	MgO	S	K ₂ O
Lump	-.53	.54	.09	.58	-.01	.87	-.04	.33	-.26	-.28	.29	-.17
Fines	-.77	.47	.57	.60	-.01	.72	-.17	.22	-.29	-.33	.38	-.13

5 Model solution

The non-linear regression equation (7) can be solved to estimate p and q and the mine means M. The dilutants p and q are not necessarily constants, but may themselves be expressed as functions, of the trainload portion weight W_n, the crusher plant (TCB1 or TCB2), the product (Lump or Fines), time, et cetera.

The model of equation (7) could be run separately for lump and fines, and it is to be expected that the mean value vector will be different for lump and fines. However, the two analyses can be combined. Letting the vector of mean values M be M_L and M_F respectively for lump and fines, equation (7) becomes the combined equation:

$$Y_n = X_n \sqrt{W_n} = (pS_p + (W_n - p - q)S_c + qS_q) \cdot (L \cdot \underline{M}_L' + F \cdot \underline{M}_F') / \sqrt{W_n} + e_n \tag{9}$$

The statistical package SPSS™ was again used, to solve the non-linear regression model for the unknown parameters p, q, ML and MF, which could be expressed as functions.

A range of physically believable models were tested, rejecting any terms not significant at the 5% significance level. The following model was found significant in all its terms:

$$p = P_1 * W_n \tag{10}$$

$$q = Q_1 * W_n * \exp(-H * T_q) \tag{11}$$

$$M_L = \{M_{AL}, M_{BL}, M_{CL} + G_{CL} * \text{Time}, M_{DL}, M_{EL}\} \tag{12}$$

$$M_F = \{M_{AF} + G_{AF} * \text{Time}, M_{BF} + G_{BF} * \text{Time}, M_{CF}, M_{DF}, M_{EF}\} \tag{13}$$



Parameters defining p and q are of operational interest. All are significant at the 0.0001% level, as tabulated below.

Parameters for M_L and M_F represent discriminant function score means and are of no further interest. The coefficients G_{CL} , G_{AF} and G_{BF} , were all significantly negative, indicating that the means for Lump from Mine C Lump and for Fines from Mines A and B decrease with time. Slopes against time for the other means were tested, and none were significant at the 5% level.

Table 2: Dilution parameter estimates.

Parameter	StdErr	Sig.
P_1	.083	< 0.0001%
Q_1	.119	< 0.0001%
H/minute	.060	< 0.0001%

6 Testing for further dilution models

The dilution model of equations (10) and (11) support the model for which:

- 1) The tonnage dilution from the previous trainload portion is proportional to the tonnage of the current trainload portion.
- 2) The tonnage dilution from the following trainload portion is proportional to the tonnage of the current trainload portion, declining exponentially according to the time gap between the current and following trainload portions.

Further terms were added to the model of equations (10) to (13), one by one, to test if the model required significant augmentation.

6.1 Potential additions to the “p” model

$$p = P_0 + P_1 \cdot W_n \quad P_0 = .034 \text{ +/- } .023 \\ \text{Sig.}(P_0=0) = 14\% \quad (14)$$

$$p = P_1 \cdot W_n \cdot \exp(-K \cdot T_p) \quad K = 1.92 \text{ +/- } 1.11 \\ \text{Sig.}(K=0) = 8.4\% \quad (15)$$

$$p = P_1 \cdot W_n^K \quad K = .808 \text{ +/- } .091 \\ \text{Sig.}(K=1) = 3.5\% \quad (16)$$

$$p = (P_B \cdot \text{TCB2} + P_1) \cdot W_n \quad P_B = -.0005 \text{ +/- } .008 \\ \text{Sig.}(P_B=0) = 94\% \quad (17)$$

$$p = (P_F \cdot \text{Fines} + P_1) \cdot W_n \quad P_F = -.015 \text{ +/- } .010 \\ \text{Sig.}(P_F=0) = 15\% \quad (18)$$



6.2 Potential additions to the “q” model

$$q = Q_0 + Q_1 \cdot W_n \cdot \exp(-H \cdot T_q) \quad Q_0 = -.016 \text{ +/- } .019$$

$$\text{Sig.}(Q_0=0) = 40\% \quad (19)$$

$$q = Q_1 \cdot W_n \cdot K \cdot \exp(-H \cdot T_q) \quad K = .895 \text{ +/- } .175$$

$$\text{Sig.}(K=1) = 55\% \quad (20)$$

$$q = (Q_B \cdot \text{TCB2} + Q_1) \cdot W_n \quad Q_B = -.042 \text{ +/- } .020$$

$$\text{Sig.}(Q_B=0) = 3.7\% \quad (21)$$

$$q = (Q_F \cdot \text{Fines} + Q_1) \cdot W_n \quad Q_F = -.019 \text{ +/- } .024$$

$$\text{Sig.}(Q_F=0) = 43\% \quad (22)$$

7 Results

None of the examined extensions to the model of equations (14) to (22) are significant at the 3% level. Two extensions, (16) and (21), significant at the 5% level, provide weak evidence that:

- 1) p may increase a little less than linearly with W_n .
- 2) q may be smaller for the second plant TCB2.

It was concluded that the evidence for these effects was not strong enough to justify further complicating the model.

8 Undiluting the data

Having established that each trainload is diluted with a tonnage p from the previous rake, and q from the following rake, we can write the formula for calculating the trainload's undiluted composition vector X_U from its diluted composition X_D , and the compositions X_P and X_Q of the preceding and following trainloads.

$$X_U = (X_D - pX_P - qX_Q)/(1-p-q) \quad (23)$$

where $p = 0.083$, and $q = 0.119 \cdot \exp(-.06T_q)$

The dilution model of equations (10) and (11) was used to undilute the diluted data, in accord with equation (23). The compositions X_P and X_Q of the preceding and following trainloads should really be the undiluted values. For convenience, the diluted values were used instead, introducing a negligible second-order error.

As an example of the undilution process, cumulative distributions for diluted and undiluted lump Fe values are plotted in Figure 1, and for lump Al_2O_3 values are plotted in Figure 2. The graphs show the results only for Mines A and B. Fe is consistently higher for Mine A, and undilution makes it higher still. Al_2O_3 is consistently lower for Mine A, and undilution makes it lower still. The reverse applies to Mine B.



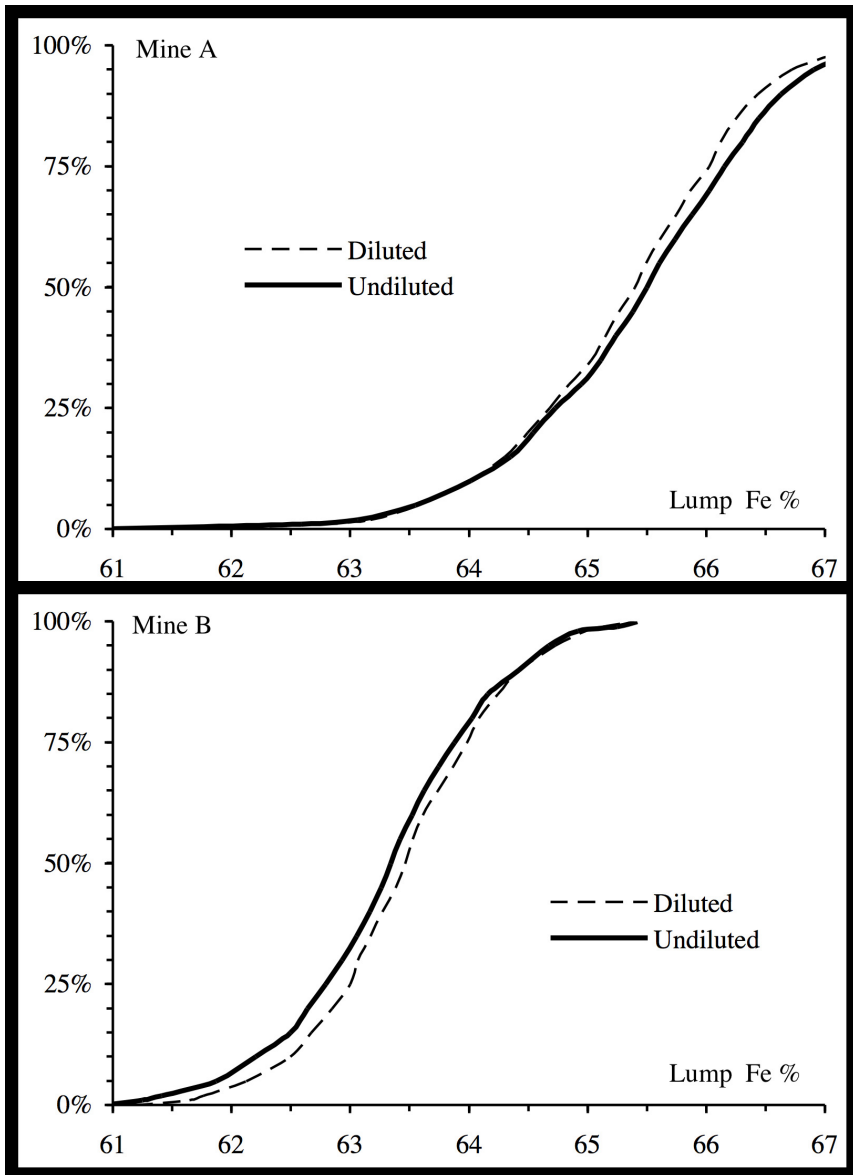


Figure 1: Cumulative distributions for lump Fe, Mines A and B.

9 Conclusion

The results plotted in Figures 1 and 2 show examples of the undilution process. Similar comparisons can be made for the other three mines, and for the other minerals of interest, P and SiO₂.



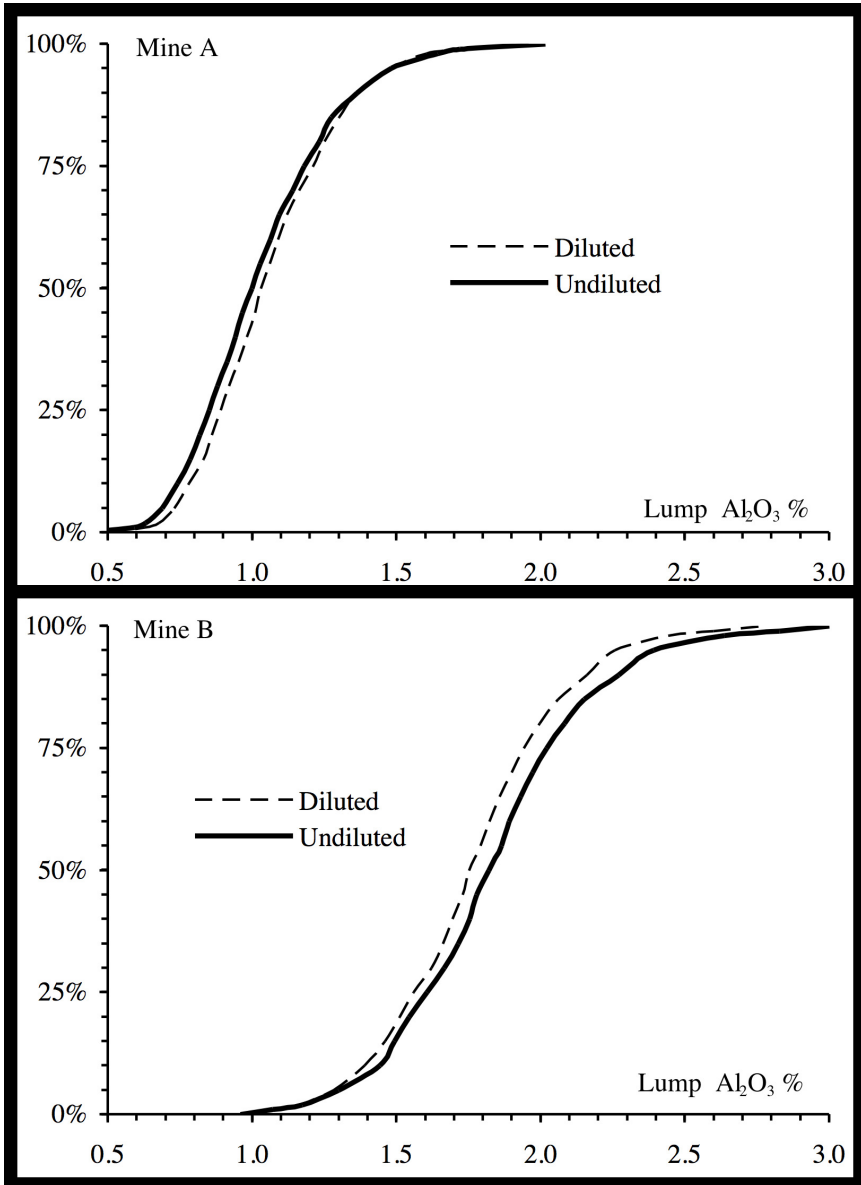


Figure 2: Cumulative distributions for lump Al_2O_3 , Mines A and B.

It should be noted that the undilution applied to a particular trainload takes account of the particular trainloads that precede and follow it. If a trainload from a particular mine is preceded and followed by trainloads also from the same mine, then the undiluted composition will be close to the diluted composition.



The procedure is now used routinely to improve the estimates of composition for trainloads from each mine. The greater accuracy has helped in improving the tracking of mine estimate errors and thus has improved the mine planning capability.

The model of equations (9)–(13), leading to the undilution formula of equation (23) is somewhat counter-intuitive. It finds that the tonnage contaminated from the previous and following trainloads is proportional to the tonnage of the trainload being contaminated. Our expectation had been that the contamination should be a tonnage amount independent of the tonnage of the contaminated trainload. The significances of the model parameters in equations (10), (11), (14) and (19) support the counter-intuitive model.

The model finds that the contamination from the following trainload decays exponentially, according to the time interval between the trainload being contaminated and the following trainload. The decay has a half-life of about twelve minutes. This makes sense. The sampling from a trainload stops when the operator judges that the trainload processing has finished. If the following trainload starts being fed in straight away, some of it will be picked up in the sample, but not if there is an appreciable time interval between the trainloads.

References

- [1] Everett, J.E. Iron ore handling procedures enhance export quality, *Interfaces*, **26(6)**, pp. 82-94, 1996.
- [2] Everett J.E. Iron ore production scheduling to improve product quality, *European Journal of Operational Research*, **129**, pp. 355-361, 2001.
- [3] SPSS, *Reference Guide*, SPSS Inc: Chicago, 1990.

