Regression Analyses of Vangorda Experimental Data

Given:

A Lotus type data file "data.wr1" containing whole rock specific gravities and assays on drill core from the Vangorda deposit. The file also records drill hole number, "from" and "to", sample number and rock type codes for each entry. The file has been sorted with respect to rock type. The total file was transformed to ASCII text and reformatted under the Pe2 editor to a BMDP readable file "data.prn".

Required:

To predict the physical property of in situ bulk density of dry rock from elemental mineral assays.

Solution:

The in situ bulk density of dry rock is, in principle, simply the inverse of the summation of its component dry mineral volume fractions adjusted for porosity. Let the solid portion of a rock volume have mineral mass fractions f_i ($i = 1, 2 ... n$) (tonnes mineral/tonne dry rock). The minerals are assumed to be homogeneous, discrete and nonporous. Each mineral has a fixed specific gravity (grain density) of ρ_i (tonnes/cubic metre). The in situ rock also has a certain homogeneous proportion of its volume occupied by void spaces. These spaces may be partially or wholly filled with fluids. However, when considered on a dry basis, the void fillings are assumed to be air and as such to have negligible mass. The amount of void volume is given as a fraction of the total volume using the symbol ϕ (cubic metres of voids/cubic metre of dry rock).

On the above basis it is easy to see that the in situ specific gravity of the dry rock is:

$$\rho = \frac{(1-\phi)}{\sum_{i=1}^{n}\left(\frac{f_i}{\rho_i}\right)}\left(\frac{t}{\mathrm{cu\ m}}\right) \tag{1}$$

The data available, however, does not give the mass fractions of <u>mineral</u> but that of some of the mineral elements. If the individual elemental assays infer, singly, a distinct mineral then the inverse of (1) could be written as:

$$vf = \frac{1}{(1-\phi)}\sum_{i=1}^{n}\left(\frac{vf_i \cdot a_i}{h_i}\right)\left(\frac{\mathrm{cu\ m}}{t}\right) \tag{2}$$

where a_i = elemental assay in mass %

h_i = mass % of element in pure mineral

vf_i = $1/\rho_i$

It is clear that (2) gives a linear function of the reciprocal of specific gravity in terms of elemental assays. Fitting a function to the given data then should be with respect to a mineralogical model and the reciprocal of the specific gravity or volume factor (vf).

The fitting process will have several sources of error.

These are:

(1) assumed mineralogy is incorrect
(2) mineral element content is variable
(3) mineral density is variable
(4) rock porosity is not homogeneous
(5) elemental assays are in error
(6) specific gravity measurements are in error

Items 1, 2, 3 and 4 can be termed model errors while 5 and 6 are of the measurement category.

In the absence of any further knowledge, the mineralogy of Vangorda rock is assumed to be that of simple sulphides with a homogeneous gangue. That is, all the lead is contained in galena (PbS), all the zinc in sphalerite (ZnS), all the iron in pyrite (FeS) and all the barium in barite (BaS).

The gangue mineral is then the remaining mass.

Assuming average atomic weights for the minerals, the mineral mass fractions are as follows:

Pb in pure galena = 86.6%
Zn in pure sphalerite = 67.1%
Fe in pure pyrite = 63.5%
Ba in pure barite = 81.1%

Using this simple mineralogical model with the given assay data, equation 2 can be written as:

$$\mathrm{vf} = \frac{1}{(1-\phi)}\left[\frac{\%\mathrm{Pb}}{86.6}\cdot \mathrm{vf}_1 + \frac{\%\mathrm{Zn}}{67.1}\cdot \mathrm{vf}_2 + \frac{\%\mathrm{Fe}}{63.5}\cdot \mathrm{vf}_3 + \frac{\%\mathrm{Ba}}{81.1}\cdot \mathrm{vf}_4 \right.$$

$$\left. + \left(1 - \frac{\%\mathrm{Pb}}{86.6} - \frac{\%\mathrm{Zn}}{67.1} - \frac{\%\mathrm{Fe}}{63.5} - \frac{\%\mathrm{Ba}}{81.1}\right)\cdot \mathrm{vf}_5\right] \qquad (3)$$

It is clear that, as equation 3 is linear in terms of the elemental mass assays, and may be used as a model to fit the reciprocal of the specific gravity measured. The fit will be with respect to a zero intercept. The regressors, $\beta_i = \frac{vf_i}{(1-\phi)}$ are then estimating the inverse of the mineral specific gravity adjusted for porosity and as such must be constrained as positive non zero. How close the estimators come to prediction of reasonable values gives <u>some</u> indication of the level of model error.

It is instructive to rewrite equation 3 with assay values gathered together.

$$vf = \frac{1}{(1-\phi)}\left[vf_5 + \%Pb \cdot \left(\frac{vf_1 - vf_5}{86.6}\right) + \%Zn \cdot \left(\frac{vf_2 - vf_5}{67.1}\right)\right.$$

$$\left. + \%Fe \cdot \left(\frac{vf_3 - vf_5}{63.5}\right) + \%Ba \cdot \left(\frac{vf_4 - vf_5}{81.1}\right)\right] \qquad (4)$$

We can see that fitting the reciprocal of specific gravity to a non zero intercept linear regression on the assays can represent the simple mineralogical model. The regression coefficients however represent two or more physical parameters and are much more difficult to interpret. Algebraic inversion of equation 4 results in an infinite polynomial in the four assay variables whose coefficients are powers and cross products of those in equation 4. Interpretation here with respect to a mineralogical model is impossible.

At this point, the general basis for linear least squares regression deserves comment. The first and major assumption is that the input data are independent, identically distributed variates. The second and less restrictive assumption is that the residual resulting from lack of fit is independent and normally distributed with a mean of zero and a homogeneous variance. The assumption that the drill cores selected for specific gravity represent a random sample from an independent, identically distributed population is very difficult to substantiate. Firstly, if the population is to represent the complete orebody, the portions of broken and unrecovered core that were unsuitable for specific gravity analysis are certainly unrepresented. Secondly, unless the material (solid core sections) has been selected carefully with respect to the general ore mineralogy distribution, it is quite probable that specific mineralogies are over or under represented.

The least squares fitting procedure assumes all the lack of fit is due to differences between the model and the measured variable, i.e. specific gravity. The predictor variables, elemental assays, are assumed to be without error. This is certainly not so.

In summary, it is clear that the usual statistical criteria for regression may be poorly met with respect to the data. The results then should be treated with some caution.

Regression Methodology:

The regression were done using the BMDP software package with program 2R. This is a stepwise multiple linear regression program. That program selectively adds variates to the regression with respect to the relative increase in the ratio of regression sum of squares to the total sum of squares. When the relative increase in this ratio is below that which could happen by chance (approximately 1:20), addition of variates stops and the variates are examined as candidates for removal from the regression.

At each step of the regression an analysis of variance gives the regression coefficients and their standard errors. The standard error estimate of the regression is also given. On completion of the regression, many diagnostics of the regression on a case by case basis can be made to allow one to see if some observations unduly affect the results. It was found in the Vangorda data that residuals substantially beyond three standard deviations should be discarded. The amount of data discarded was never greater than 5%. It was also found that using the length of core as a weighting function improved the fitting process considerably.

Despite the preceding arguments in favour of volume factor regression on a mineralogical model, the direct empirical model of a specific gravity as a linear function of the assays will be useful for comparison with previous work. It was also found that, in statistical measures of fit, the volume factor mineralogical model was only marginally better than the empirical specific gravity model. Another correspondence between models was that data rejected in the empirical model was also rejected in the mineralogical model. Again the empirical and the mineralogical models when used to test for possible similarities in regressions between differing rock types gave equivalent results. Much of the exploratory analysis, therefore, was done using the simple empirical specific gravity model. The more computationally complex mineralogical model was used for final results.

Regression Results:

The BMDP input and output files of some 35 separate runs are enclosed on a 3-1/2 high density disk. The plots and residual tables have been deleted from these files as the mass of material would have taken up several more disks. The re-creation of the total output is perhaps best made by a BMDP rerun with the *.INP file or the echo output at the beginning of the *.OUT file.

Summary of the relevant statistics are given in Tables 1 and 2.

Primary screening showed some anomalies in the data file. Firstly, there were a set of six records that had been entered twice. These were of hole 90V-2 from 29.6 to 35.9. Visual examination of an early regression residual record pointed to a duplication of records and a short fortran program was written to do a thorough search for duplicates. Five more, all under the same hole code, were discovered and deleted from the source ASCII file data.prn. Further screening via BMDP controls showed that hole number 90V-38 sample number 6027 has a "to" equal to the "from" giving it zero length and excluding it from analysis by the length weighted regression. In addition, two further records had one or more assays at zero less and were excluded from analyses. All these identified invalid records were deleted from the data.prn data base. There is still one reject record that in rock type E has not been identified and removed. This record appears under the column heading "No. rej." in both tables.

Examination of Table 1, the emperical model regression on specific gravity indicates, to me, the following with respect to specific gravity prediction:

(1) There are definitely two types of rock, the quartzite based material A, C and D and the massive sulphide material E, G and H. (The remaining rock categories, 3GO, J, K and L and 5C*, have too small a sample size for analysis.)

(2) Within the quartzite group (A, C and D), one might consider group A as distinct from C and D. However when separate regressions are done an subgroups of A (A7, A4 and A0), we can see there is as much variability in parameters between subgroups of A as there is between the overall A and C. Also, subgroups under C (C8, C7, C78, C3, C38 and C0) appear to be homogeneous and do not significantly differ from C or A + C + D.

(3) Within the massive sulphide rock types (E, G and H), I do not see differences between groups.

In general then my conclusion is that for specific gravity prediction purposes, two regressions, one for quartzite type rocks and one for massive sulphides, should be sufficient.

Conclusions

The "best" regression functions for Vangorda "rock", as represented by the experimental data, are of volume factors (cubic metres/tonne) for two general rock types. These functions are in terms of an assumed mineralogy of galena, sphalerite, pyrite and barite. The mineral mass fractions are related to the elemental assay through molecular proportion as follows:

Galena (mass fraction) = % Pb/86.6
Sphalerite (mass fraction) = Zn/67.1
Pyrite (mass fraction) = % Fe/63.5
Barite (mass fraction) = % Ba/81.1
Gangue (mass fraction) = 1 – galena – sphalerite – pyrite – barite

The volume factor predictor (cubic metres/tonne) for quartzite rock (types A, C and D, line 6, Table 2) is as follows:

$$\begin{aligned} \text{vf (A, C and D)} = {} & 0.3759 \cdot \text{gangue} + 0.1210 \cdot \text{galena} \\ & + 0.3598 \cdot \text{sphalerite} + 0.1489 \cdot \text{pyrite} \\ & + 0.0848 \cdot \text{barite} \end{aligned} \tag{5}$$

The standard error of this function is 0.0233, and the number of samples used in computation was 809.

In terms of element assays, this is:

$$\begin{aligned} & \text{vf (A, C and D; Pb, Zn, Fe, Ba)} = \\ & \quad 0.3759 - 2.943 \cdot 10^{-3} \cdot \%\ \text{Pb} - 2.399 \cdot 10^{-4} \cdot \%\ \text{Zn} \\ & \quad - 3.575 \cdot 10^{-3} \cdot \%\ \text{Fe} - 3.589 \cdot 10^{-3} \cdot \%\ \text{Ba} \end{aligned} \tag{6}$$

The volume factor (cubic metres/tonne) for massive sulphide rock (types E, G and H, line 2, Table 2) is:

$$\text{vf (E, G and H)} = 0.3148 \cdot \text{gangue} + 0.0986 \cdot \text{galena} + 0.3198 \cdot \text{sphalerite} + 0.1970 \cdot \text{pyrite} + 0.1669 \cdot \text{barite} \qquad (7)$$

The standard error of estimate is 0.0231 and 858 samples were used in estimation.

In terms of elemental assays this is:

$$\text{vf (E, G and H; Pb, Zn, Fe, Ba)} = 0.3148 - 2.497 \cdot 10^{-3} \cdot \%\ \text{Pb} + 7.452 \cdot 10^{-5} \cdot \%\ \text{Zn} - 1.855 \cdot 10^{-3} \cdot \%\ \text{Fe} - 1.824 \cdot 10^{-3} \cdot \%\ \text{Ba} \qquad (8)$$

Table 1

SPECIFIC GRAVITY REGRESSION BY ELEMENTAL CONTENT

Rock Type	Std. Error of Est.	No. in File	No. Out.	No. Rej.	Const. Est.	Lead		Zinc		Iron		Barium	
						Coef. $\cdot 10^2$	$\sigma \cdot 10^3$	Coef. $\cdot 10^2$	$\sigma \cdot 10^3$	Coef. $\cdot 10^2$	$\sigma \cdot 10^3$	Coef. $\cdot 10^2$	$\sigma \cdot 10^3$
ALL	0.3414	1701	17	1	2.5478	5.13	3.5	-	-	4.08	0.88	4.01	1.3
H	0.2842	30			3.6142	-	-	-	-	-	-	7.10	23.3
G	0.3409	356	2		3.3948	-	-	-	-	2.44	3.4	2.37	2.8
E	0.3495	472	9	1	2.8628	5.54	7.0	-	-	3.37	2.3	2.78	6.5
D	0.2733	101	3		2.5165	4.83	11.1	-	-	3.94	4.3	4.42	8.4
C	0.2588	476	8		2.7594	5.77	17.1	-	-	2.95	2.0	-	-
A	0.1910	232	8		2.5569	-	-	2.42	4.6	3.18	2.1	5.76	5.7
E, G, H	0.3578	858	11	1	3.0536	3.80	5.0	-	-	2.79	1.8	2.78	1.9
G8	0.3055	23			4.2722	-	-	-	-	-	-	-	-
G4	0.3274	111	2		2.9651	-	-	-	-	3.93	7.3	3.37	5.2
G48	0.3005	78			4.2668	-	-	-	-	-	-	-	-
G0	0.3666	34	1		4.4834	-	-	–8.25	38.0	-	-	-	-
E8	0.3214	35	2		2.8948	-	-	-	-	3.88	11.1		
E64	0.3933	27	1		2.5760	7.70	31.7	-	-	3.62	9.5	4.76	18.2
E1	0.3069	66			2.5248	-	-	-	-	4.51	6.4	-	-
E18	0.2747	35			2.9100	-	-	-	-	3.15	7.2	9.43	35.8
E14	0.2718	26			2.7779	8.38	24.7	-	-	2.78	7.0	-	-
E0	0.1980	30	2		2.8978	–19.33	70.7	-	-	4.50	7.0	-	-
C8	0.2485	24			3.0195	-	-	-	-	2.00	8.4	-	-
C7	0.2299	103			2.8045	-	-	-	-	2.48	3.6	-	-
C78	0.2208	29			2.7418	-	-	-	-	2.86	5.8	15.72	49.3
C3	0.2307	113	2		2.8285	-	-	13.85	37.6	2.57	4.2	-	-
C38	0.2407	43	1		2.6136	-	-	-	-	4.10	7.1	-	-
C0	0.1488	49	4		2.5396	9.46	33.8	-	-	3.92	4.7	-	-
A7	0.0977	29			2.7286	-	-	6.77	25.2	1.43	4.9	-	-
A4	0.2624	94	1		2.4857	4.97	19.9	-	-	3.58	4.0	5.65	8.8
A0	0.1585	70	4		2.5297	7.07	23.2	-	-	3.60	3.1	-	-

Table 2

VOLUME FACTOR REGRESSION BY MINERAL CONTENT

Rock Type	Std. Error of Est.	No. in File	No. Out.	No. Rej.	Gangue		Galena		Sphalerite		Pyrite		Barite	
					Coef. $\cdot 10^2$	$\sigma \cdot 10^3$	Coef. $\cdot 10^2$	$\sigma \cdot 10^3$	Coef. $\cdot 10^2$	$\sigma \cdot 10^3$	Coef. $\cdot 10^2$	$\sigma \cdot 10^3$	Coef. $\cdot 10^2$	$\sigma \cdot 10^3$
ALL	0.0253	1701	18	1	36.87	1.8	9.71	33.7	33.32	23.2	15.35	2.8	12.35	8.5
E, G, & H	0.0231	858	11	1	31.48	4.0	9.86	39.2	31.98	28.7	19.70	4.3	16.69	9.3
D	0.0215	101	3		37.08	9.9	-	-	35.70	49.1	16.50	13.8	12.93	53.3
C	0.0215	476	9		35.24	3.9	-	-	26.38	102.1	19.26	6.9	-	-
A	0.0199	232	8		38.07	2.7	25.26	112.3	25.38	49.7	17.62	12.2	-	-
A, C & D	0.0233	809	10		37.59	2.2	12.10	61.3	35.98	34.7	14.89	4.1	8.48	38.9