DRAGLINE COMPARISONS

BHP provided data on the performance of a dragline over a four-week period during which four different buckets were used. The Study Group examined this data and suggested a method of analysing data from such comparative studies.

1. Introduction

In surface mining of coal, draglines are used to expose coal by removing the overlying rock (overburden). A typical dragline might remove 10 million cubic metres of rock per year, exposing about 1 million tonnes of coal.

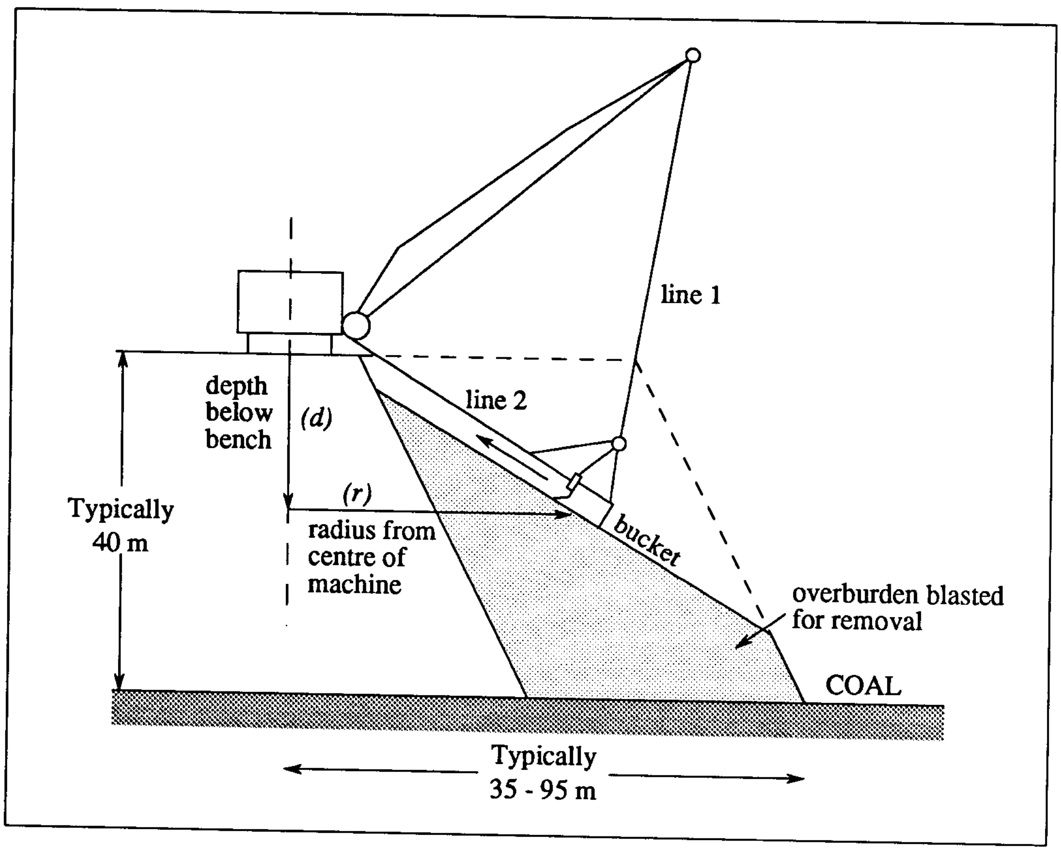

Figure 1: Side view of a dragline in operation

Figure 1 shows how the overburden is removed from above a coal seam. The diagram gives a side-on view of a strip being excavated by a dragline machine. The bucket is lowered onto the surface of the overburden and a drag rope is used to pull it up the slope. Once the bucket is filled, it is lifted into the air, swung through typically 90° and its contents dropped onto a spoil pile running along the side of the exposed coal seam.

One factor determining the performance of draglines is the design of the bucket used to remove the overburden. Small differences in bucket performance can have significant effects on the efficiency of the surface mining operation. It is estimated that an improvement of 1% in dragline performance over BHP Australia Coal's fleet of over 30 draglines would increase sales by $12 million. Therefore, the reliable detection of *small* performance changes is important.

When evaluating the performance of different bucket designs however, the engineers cannot control the field trials of buckets on real machines as much as they would like. Consequently they need reliable methods for analysing the data that they can get.

David Shanks of BHP Research asked that the Study Group consider statistical methodology which can be used for analysing field trials in which alternative equipment is evaluated. His interest was primarily in methodology, but he asked that the methodology be illustrated using data which he supplied.

He provided data from a trial in which four different buckets were used for a week each on a single dragline. He asked that we analyse this data set and indicate which bucket seemed to be best, giving some expression to the uncertainty in our conclusions.

2. The data

Data was provided on six quantities. These were as follows.

Dig length (metres):— the distance the drag rope is wound during the filling of the bucket.

The times of start and finish of filling were estimated using data on rope tensions which were polled once per second by a computer on the dragline. The estimation is not completely reliable, as illustrated by the fact that a few very short dig lengths were recorded.

Dig time (seconds):— the time of the filling process as defined above. This is measured in whole seconds. No dig times were recorded as exceeding 60 seconds. Long dig times may have been removed from the data set, possibly in order to omit digging cycles during which operators were changed.

x (metres):— the horizontal displacement of the point where filling started from the centre of rotation of the dragline. This is also referred to as the radius.

y (metres):— the vertical displacement of the point where filling started from the base of the dragline.

The displacements x and y were calculated by the computer on board the dragline from measurements on the positions of the drums on which the drag and hoist ropes were wound.

Suspended load (tonnes):— estimated using measurements of the tensions in the drag and hoist ropes. This includes the weights of the bucket, rigging and payload.

Payload (tonnes):— the estimated weight of overburden in the bucket.

Some transformation of the data was done for commercial reasons. Details of data transformation were not provided. The four buckets were identified only as A, B, C and D, and details of bucket loads have not been included in this report.

Records were provided giving these six quantities for 2938 dig/swing/dump cycles of the dragline with bucket A, 2972 cycles with bucket B, 4788 cycles with bucket C and 4220 cycles with bucket D. The cycles were listed in chronological order, but it was noted that there were some periods during which the instrumentation was not in use and that only cycles having x in the range 30 to 90 and y in the range −35 to 5 were included.

3. Looking at the data

Several bivariate plots were drawn in order to gain a general appreciation of the data and to check for errors or anomalous results.

Figure 2 shows y against x for bucket A. From this figure it is possible to recognise the boundaries of the block of overburden being dug - compare to figure 1.

The dragline follows a set digging pattern which can lead to clusters of points in particular regions on such a figure. For instance there is a particularly dense cluster of points around $x = 65$, $y = 2$. In interpreting this particular cluster, it should be noted that the point on the bucket corresponding to (x, y) is at the top, just behind the "arch", so that small positive values of y can occur when digging close to bench level.

Additionally, the soft overburden was causing some sinkage at the front of the dragline so that the "y" values, calculated relative to the dragline's frame, were slightly but systematically higher than their true values.

Figure 3 shows the dig length against x for bucket A. A linear trend is clearly discernible in this figure.

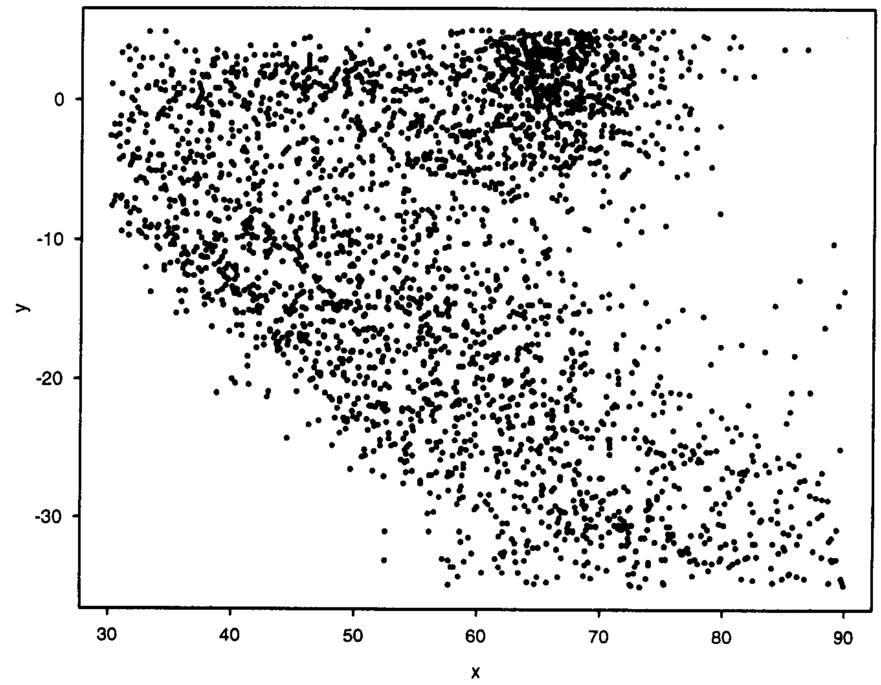

Figure 2: Plot of x against y for bucket A

David Shanks explained that when a dragline bucket is filled it is not in general hoisted immediately because the bucket would tilt at such an angle that some of the material in the bucket would fall out. This tendency to tilt can be alleviated by dragging the bucket closer to the dragline machine before hoisting it. Thus the dig length is not merely the length over which the bucket was filled. It may include some distance over which a full bucket was dragged until it could be hoisted without excessive loss of material.

Figure 4 shows the dig time against x for bucket A. The bulk of points on this scatter graph are consistent with a linear relationship between x and the dig time. However there are a large number of dig times which are longer than the most typical dig times for given x. This is consistent with a view that the bucket was dragged until it was full and could be hoisted without excessive loss of material, with the drag speed being most commonly around 1.5 metres per second but slower than this when digging was difficult for any reason.

Some dig times are unrealistically small but they are consistent with the dig lengths. A histogram of dig speed calculated as dig length divided by dig time showed an approximately normal distribution with median 1.39 metres per second, quartiles of 1.01 and 1.73, and extremes of 0.02 and 2.86.

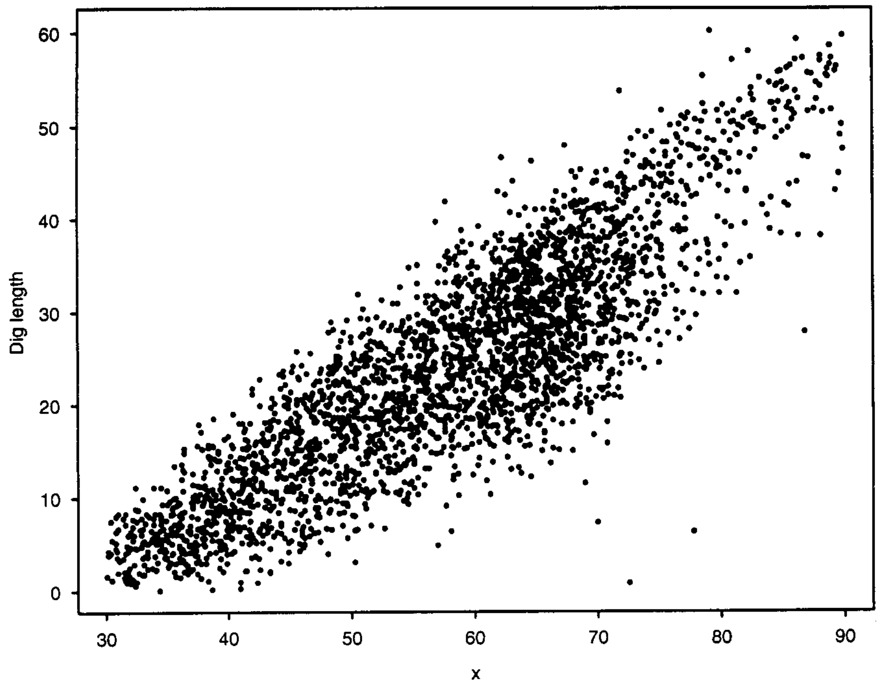

Figure 3: Plot of dig length against x for bucket A

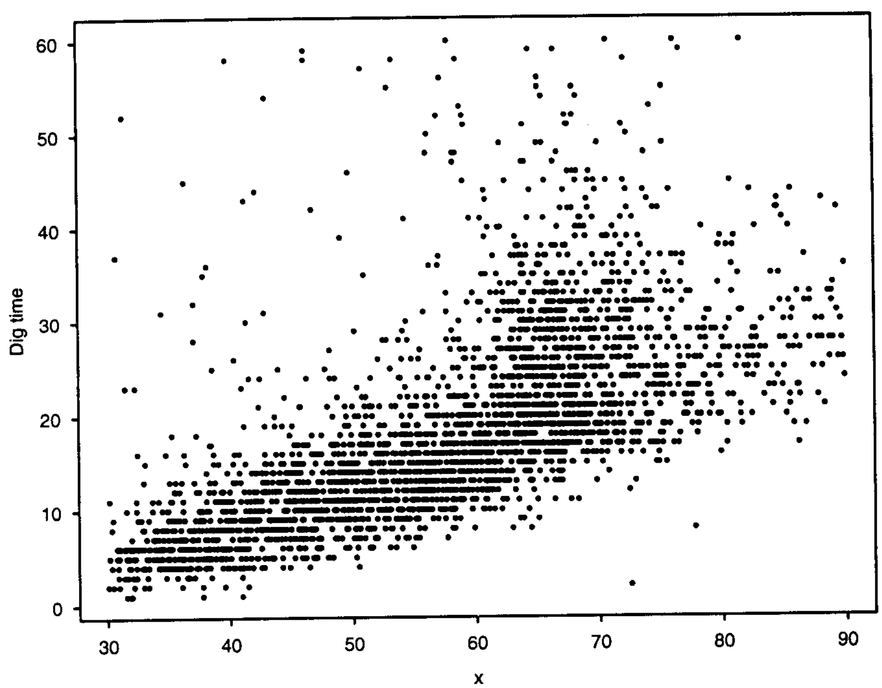

Figure 4: Plot of dig time against x for bucket A

Figure 5 shows the y values for bucket B. The horizontal axis for this graph is the index to the record within the file of records for bucket B. Thus the first digging cycle for bucket B is numbered 1, the second digging cycle is numbered 2, *etc.*

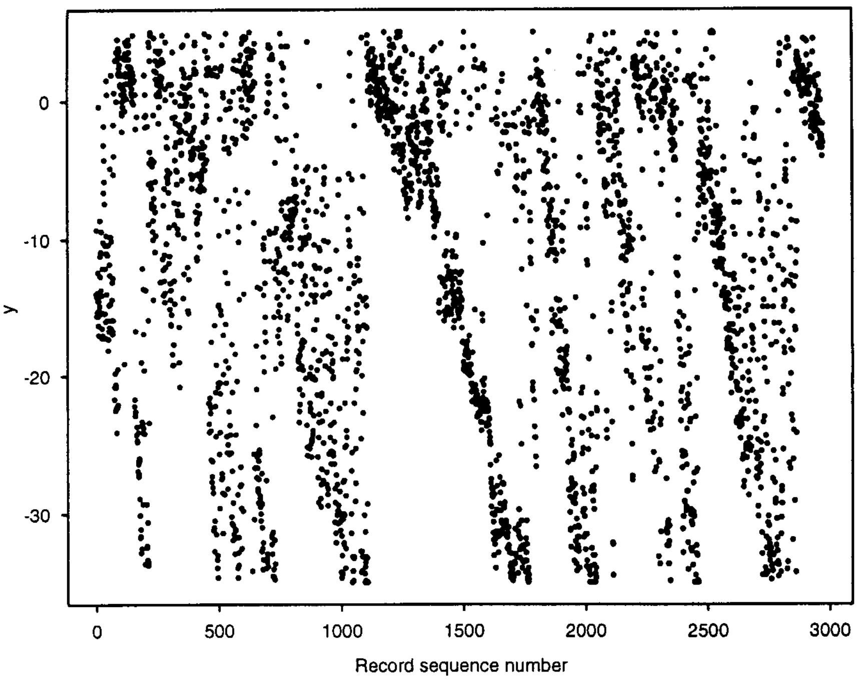

Figure 5: Plot of y against record sequence number for bucket B

Long-term cycles of digging behaviour can be clearly distinguished, with the most common value for y becoming larger negative over a period of a few hundred bucket-loads (which would be about one or two 8-hour shifts). The pattern near *index* = 1500 is particularly marked.

The pattern shown on figure 5 is typical of all four buckets, except that there was less digging below $y = -30$ for bucket D. The coal seam was slightly closer to the surface in the region of the pit where the dragline was working while bucket D was used. The pattern of there being 7–10 long-term cycles of digging behaviour during the week that a bucket was trialled was typical of all four buckets.

Scatter-plots of payload against x or y suggested that payload was not closely related to either of these variables. Multiple linear regressions also suggested that payload was not closely related to other variables. This is not unexpected, given that dragline operators would attempt to fill buckets under virtually all circumstances.

Bucket tilt

Some thought was given to the angle at which a bucket full of material would be held under the force of gravity and the tensions of the drag and hoist ropes. We would have liked to plot the payload against this tilt angle at the point where the bucket was hoisted. It is natural to speculate that very large payloads are less frequent where the tilt angle is such that material would be likely to fall out as the bucket is hoisted.

Figure 6 shows how the bucket payload is related to a simplistic estimate of the inclination of the hoist rope. It appears that very large payloads are slightly under-represented on the right hand side of the graph. The inclination of the hoist rope is itself only a crude measure of the tilt of the bucket as the bucket is lifted, so figure 6 should be interpreted as providing an indication that further work in this area would be likely to be worthwhile.

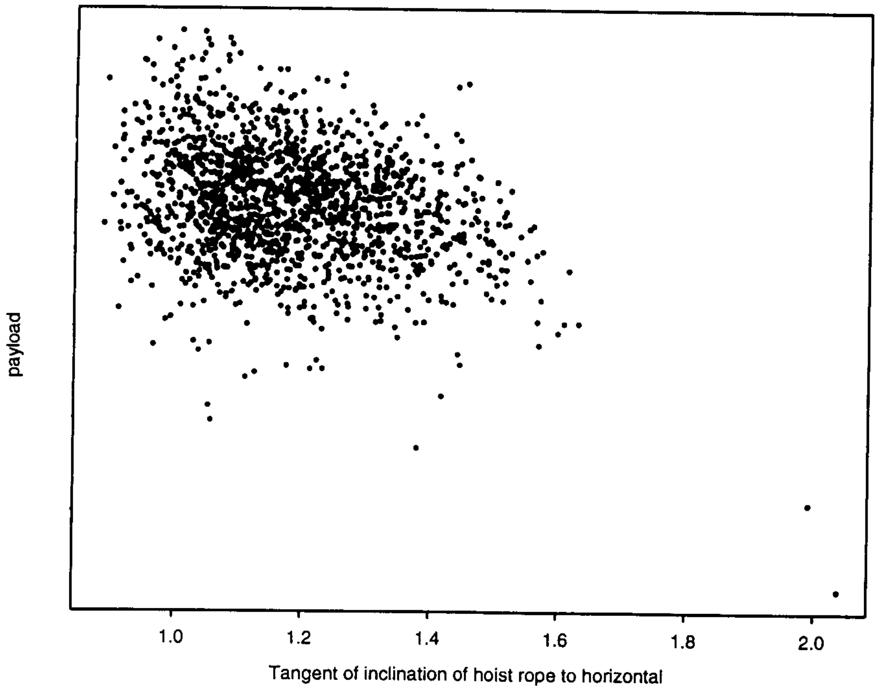

Figure 6: Plot of payload against inclination of hoist rope to horizontal at the time of hoisting for bucket *D*. (Cases where $y_h < -5$ only)

Two suggestions

Two possible improvements to the efficiency of dragline operations can be suggested at this point.

1. The calculation of the angle of tilt of the bucket if lifted could be performed every second by the computer on board the dragline. This information could be conveyed to dragline operators to help them make decisions about when to hoist the bucket.

2. Operators of draglines must choose the point at which the bucket is lifted. If they lift too soon—before the angle of tilt is satisfactory—then there is a risk that some material might fall out of the bucket as it is lifted. If they lift too late then they have wasted time. Given that loss of material is more easily noticed by their supervisors than is wasted time, it seems likely that they will tend to drag buckets for longer than the optimum time.

   BHP has confirmed that "overdigging" does indeed occur, and is the subject of research.

4. Recommended methodology

The specific methodological question to be addressed was to compare alternative dragline equipment. Practical decisions will generally involve many aspects of dragline performance. In this section we will illustrate the recommended methodology by looking only at the dig time.

A simple comparison of mean or median dig times would be inadequate because the distributions of x and y encountered by the various buckets were different. For the particular data set provided, bucket D was used less for y near -35 than were the other buckets. This was because the coal seam was nearer to the surface in the region where bucket D was used.

The recommended methodology is based on fitting linear regression models to the given dig times. Regressands x, x^2, y, xy, y^2 and y^3 were used on the basis that they were found to give a good fit to the data for bucket B. Separate regression models were fitted to the data for each of the buckets.

The choice of regressands does not affect the general ideas to be discussed below. In fact, the method could be applied even if different sets of regressands were used for different buckets.

Local comparisons

As a means of comparing buckets in various regions of their operation it is suggested that contour plots be drawn of the fitted surfaces. Figure 7 gives two such contour plots. Figure 7(a) is for bucket C and figure 7(b) is for bucket D. They show contours of the fitted dig time as a function of x and y.

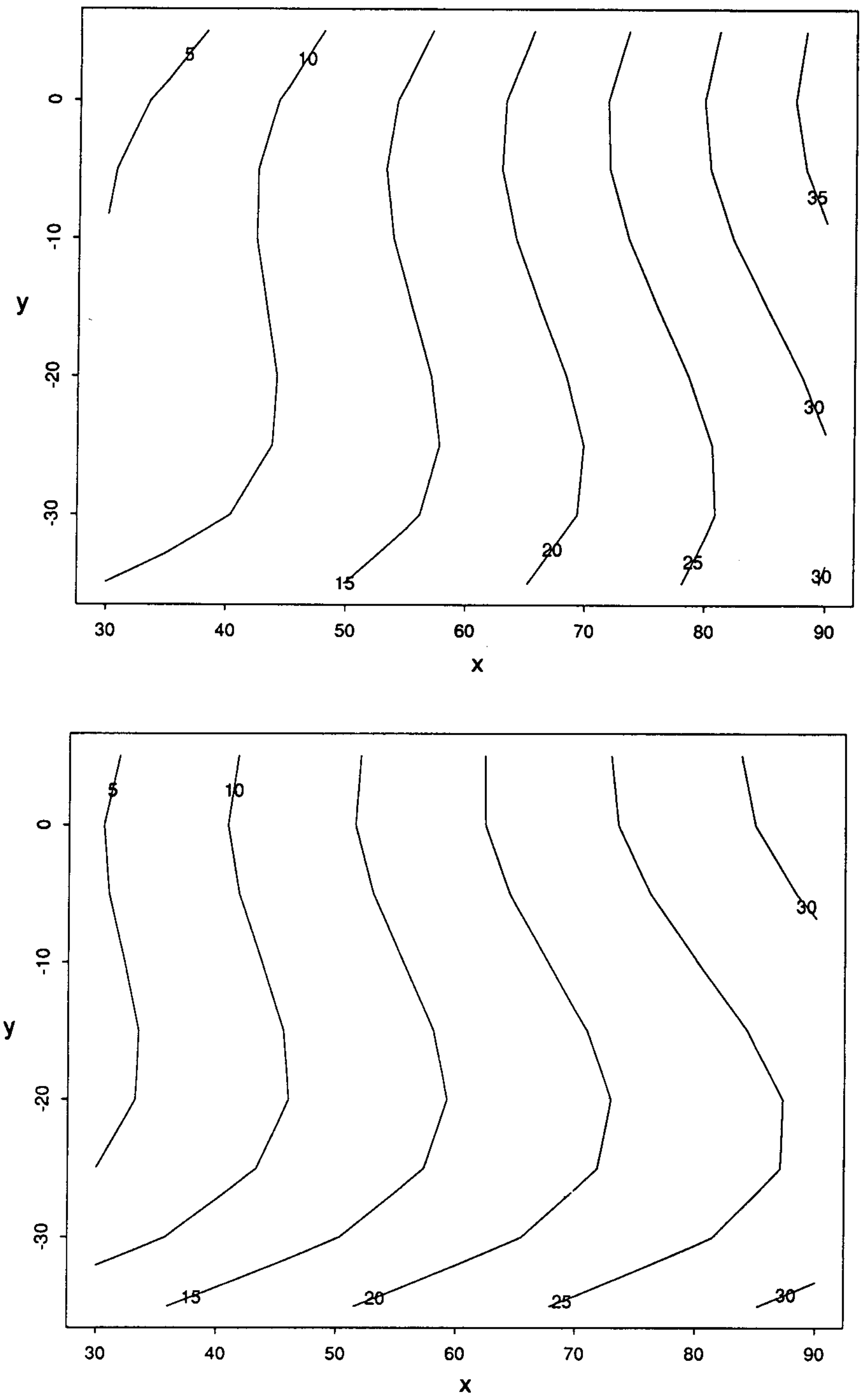

Figure 7: Contour plots of fitted average dig time in seconds as a function of x and y for buckets C and D. The upper figure is for bucket C and is referred to as figure 7(a)

In interpreting such contour plots it is important to remember that they are unreliable where there is no data. Scatter-plots like figure 2 can be looked at to see which regions of the contour plots are reliable and which regions are unreliable. The regions where the plots are unreliable are generally unimportant, because the dragline is seldom used for such combinations of x and y.

Comparison of buckets in the various regions is most easily made by making transparencies of the contour plots and looking at them simultaneously. Alternatively, contour plots of differences between fitted surfaces could be drawn.

Overall comparisons

An overall comparison of the dig times for various buckets can be made by calculating the average dig time for a benchmark distribution of values for x and y. Standard errors can also be calculated for such averages.

The benchmark distribution should in principle be selected carefully as corresponding to the distribution of x and y that is of greatest practical interest. It might be chosen to be illustrative of the distribution of x and y recently encountered, or it might be chosen to be an estimate of the distribution of x and y expected over the next few months of dragline operation.

We used the distribution of x and y for the data on bucket B as a benchmark distribution. This choice should be regarded as arbitrary.

The average dig time over the benchmark distribution could be calculated by finding the fitted dig time for each of the points in the benchmark distribution and averaging these fitted dig times. Computationally, it is much easier to first calculate the relevant moments of the benchmark distribution; and to calculate the average dig time over the benchmark distribution using these moments and the coefficients of the fitted model. This will give exactly the same answer.

For our chosen benchmark distribution the moments were as follows.

$$\begin{aligned} mean(x) &= 56.389 \\ mean(x^2) &= 3368.6 \\ mean(y) &= -12.109 \\ mean(xy) &= -747.98 \\ mean(y^2) &= 286.24 \\ mean(y^3) &= -7456.3 \end{aligned}$$

For bucket B the fitted response surface was

$$digtime = -1.214 - 0.0161x + 0.00588x^2 - 0.679y + 0.0151xy - 0.000639y^2 - 0.000213y^3$$

so its mean value over the benchmark distribution is

$$-1.214 - 0.0161 \times (56.389) + 0.00588 \times (3368.6) - 0.679 \times (-12.109)$$
$$+0.0151 \times (-747.98) - 0.000639 \times (286.24 - 0.000213 \times (-7456.3)$$
$$= 16.003$$

Some statistical computer packages allow this to be readily calculated as $m^T\hat{\beta}$ using a vector ($\hat{\beta}$) of fitted coefficients and a vector (m) of moments for the benchmark distribution, including a value 1 as the average value of 1 over the benchmark distribution since the constant term in the regression corresponds to a regressand which is 1 for all x and all y. (The package Splus was used by the group working on this problem.)

The true mean value over the benchmark distribution can be regarded as a linear combination of the parameters of the regression model, $m^T\beta$. The precision of estimating a linear combination of parameters of a regression model is discussed in many textbooks. The error incurred is $m^T\hat{\beta} - m^T\beta$ and the estimate is unbiased, so the expected mean squared error is

$$m^T E\left[(\hat{\beta} - \beta)(\hat{\beta} - \beta)^T\right] m$$

Multiple regression packages frequently provide an estimate (V) of the (unscaled) variance-covariance matrix $E\left[(\hat{\beta} - \beta)(\hat{\beta} - \beta)^T\right]$ so the expected mean square error in estimating $m^T\beta$ can be readily calculated as $m^T V m$. The standard error is the square root of this.

Fitted dig times averaged over the benchmark distribution and their standard errors were as follows.

Bucket identifier	Benchmark average dig time (secs.)	Standard error assuming independence (secs.)
A	17.82	0.019
B	16.00	0.018
C	15.70	0.016
D	15.83	0.018

The mean dig times for the four buckets (calculated as simple averages and therefore based on different distributions of x and y for each bucket) were 18.61, 16.00, 17.01 and 16.64 for buckets A, B, C and D, respectively. The mean for bucket B is the same as the fitted dig time averaged over the benchmark distribution, because the benchmark distribution was taken (arbitrarily) to be that experienced by bucket B. For the other three buckets the averages over the benchmark distribution differ from the simple averages. Of particular note is the fact that the order of preference of buckets would be B, D, C, A if simple averages were used, but the order of preference would be C, D, B, A if averages over the benchmark distribution were used.

There are some assumptions behind the calculation of standard errors of such linear combinations of parameters. People tend to worry about the fact that the departures from the model are not normally distributed, having a skew distribution with tails heavier than those of a normal distribution. Figure 8 shows the departures from the multiple linear regression model for the data for bucket B. However the departure from normality is not a major cause for concern since the procedures being used are based on least-squares and do not require normality.

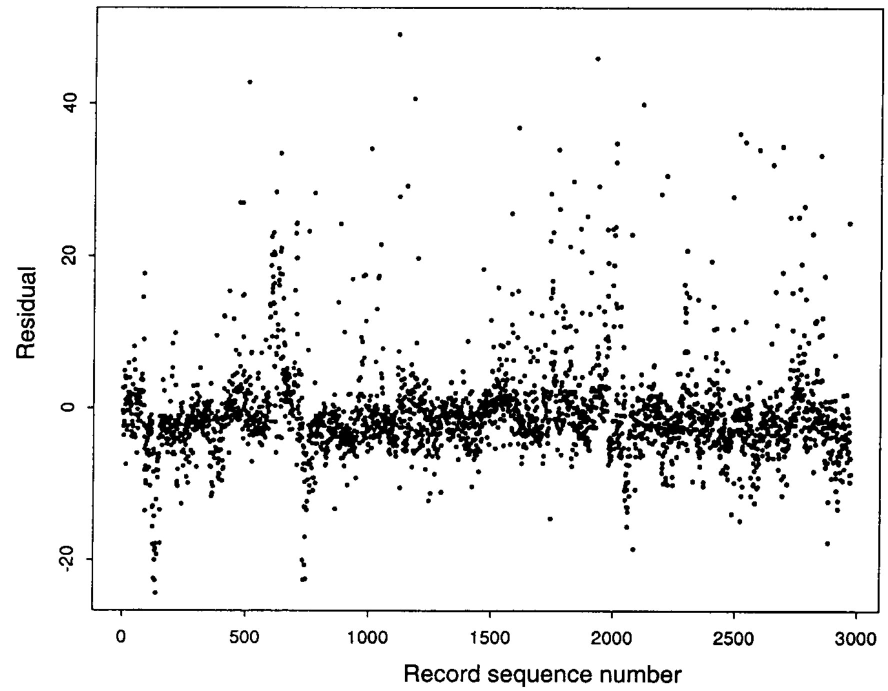

Figure 8: Residuals from multiple linear regression for the data on bucket B

A much more serious concern is that the departures from the model are correlated. Batches of mostly negative residuals and batches of mostly positive residuals are discernible on figure 8. It is to be expected that residuals will not be independent, since operator-to-operator variation and geological variation are both expected to affect dig times and to vary slowly compared to the digging cycle time.

One way to display the extent of departure from the assumption of independence is to plot a semi-variogram for the residuals. This is done in figure 9 for bucket B.

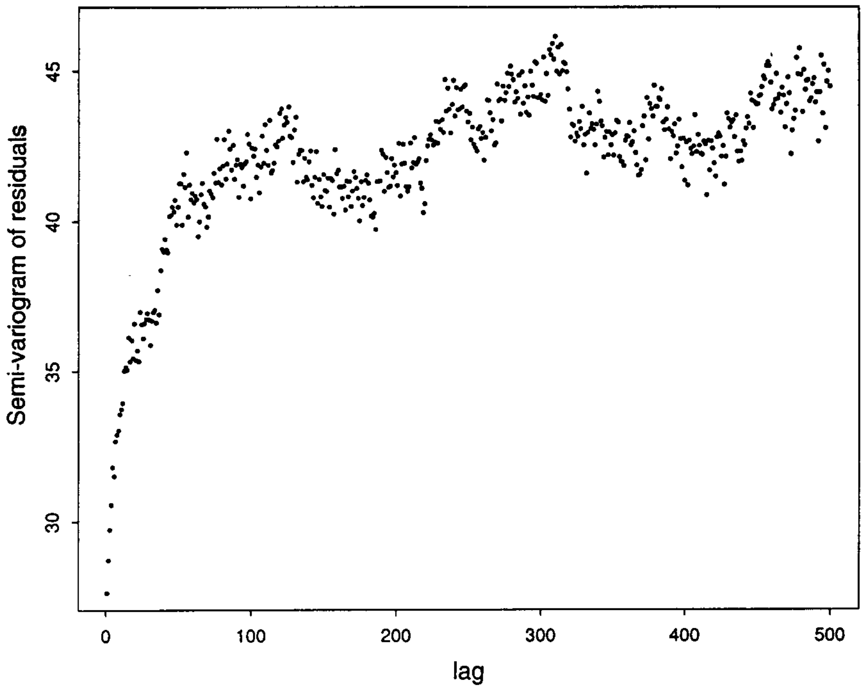

Figure 9: Semi-variogram of residuals from multiple linear regression for the data on bucket B

The semi-variogram for lag h is half the average squared difference between residuals which are separated by lag h. If the residuals were independent then the expected value of the semi-variogram would be equal to the variance of the residuals (about 42) for all lags. The sample semi-variogram would be expected to show some variation about that value.

Figure 9 shows that the semi-variogram is smaller for lags up to about 40 than it is for larger lags. Estimation of the nature and extent of any non-independent components of variance which are present is quite unreliable. A very crude interpretation of figure 9 is that the residuals seem to have a component of variance which has variance about 25 and is independent from one residual to the next, plus a component which has variance about 17 and is highly correlated. This correlation takes a lag of about 50 before it drops off to small values.

A model for a component of variance such that correlations become negligible at a lag of 50 and which is easy to use in calculations is that sets of 25 adjacent data points are affected in the same way. While this is not expected to be precisely true, it should give an indication of the extent to which the assumption of independent errors gives an inflated idea of the precision of comparisons.

The average of 3000 independent observations with variance 42 would have a variance of

$$42/3000 = 0.014$$

Assuming a component of variation with variance 25 acting on individual observations and a component of variation with variance 17 acting on each of the 120 groups of 25 observations, the average of 3000 observations would have a variance of

$$\frac{25}{3000} + \frac{17}{120} = 0.15$$

The variance of the average is larger in the second case by a factor of 0.15/0.014 = 10.7. Hence it seems reasonable to inflate the variances associated with the estimation by a factor of about 10.7. Thus the standard errors should be inflated by a factor of $\sqrt{10.7}$. The least statistically significant difference between benchmark mean dig times is approximately $1.96 \times \sqrt{2}$ times the largest standard error. This is approximately 0.17.

Thus the difference between buckets *C* and *D* seems to be just statistically significant at the 95% level, and the other pairwise comparisons are all statistically significant.

Readers should note that the estimation of the standard error of the estimates has been rather crude. The method used would not be found in any textbook. The textbook answer to the question of how reliable the estimates are is that the question cannot be answered using the available information.

The objective assessment of reliability requires that real replication be included in the experiment. This means that each bucket must be used over at least two separate periods, and that there must be enough information for estimation of the amount of variation between operators and the amount of variation between areas of blasted overburden.

Commenting of the analysis, David Shanks observed that reasons for the correlation amongst residuals could include:

- Poor fit between the regression surface and the "true" surface at particular regions. Since the dragline tends to dig in a limited region for periods in the order of an hour, a series of correlated residuals can result.

- Operator differences, as suggested. At least one operator was noted to be "different".

- Related to the previous point, slight differences in technique may have led to some systematic error in detection of the start-of-dig point by the computer instrumentation. (The software for this has been much improved since the example data were gathered.)

While completely controlled trials would probably never be possible, recent improvements in both the quality and the quantity of data available would help ease the problem of making reliable field performance assessments.

5. Alternative methodology

David Shanks had requested that some comments be made about another methodology for comparing buckets, discussed in his original statement of the problem. This involves defining subsets of the data as regions of values for x and y. The average dig times are compared for each subset separately. This provides some insight provided that the regions are few in number and can be readily interpreted.

By taking a benchmark distribution specifying the relative frequencies of the various subsets it is possible to amalgamate these comparisons across subsets and obtain an overall comparison between buckets.

Difficulties with this methodology are as follows.

1. The method becomes cumbersome if more than two variables (here x and y) are used to define the subsets.
2. Within a region, the (x, y) pairs for one bucket may have a sufficiently different distribution from the (x, y) pairs for another bucket to bias comparisons between buckets.
3. It is difficult to consider the lack of independence of the data points using this method.

Acknowledgements

The moderator (Geoff Robinson) would like to thank David Shanks for providing the data and for having sensibly delimited the problem. Contributors to the discussion of this problem included David Gates, Darryl Grieg, Michael Hasofer, Patrick Howden and Tony O'Connor.