# Data Set 10: Water Motion Across a Kelp Bed

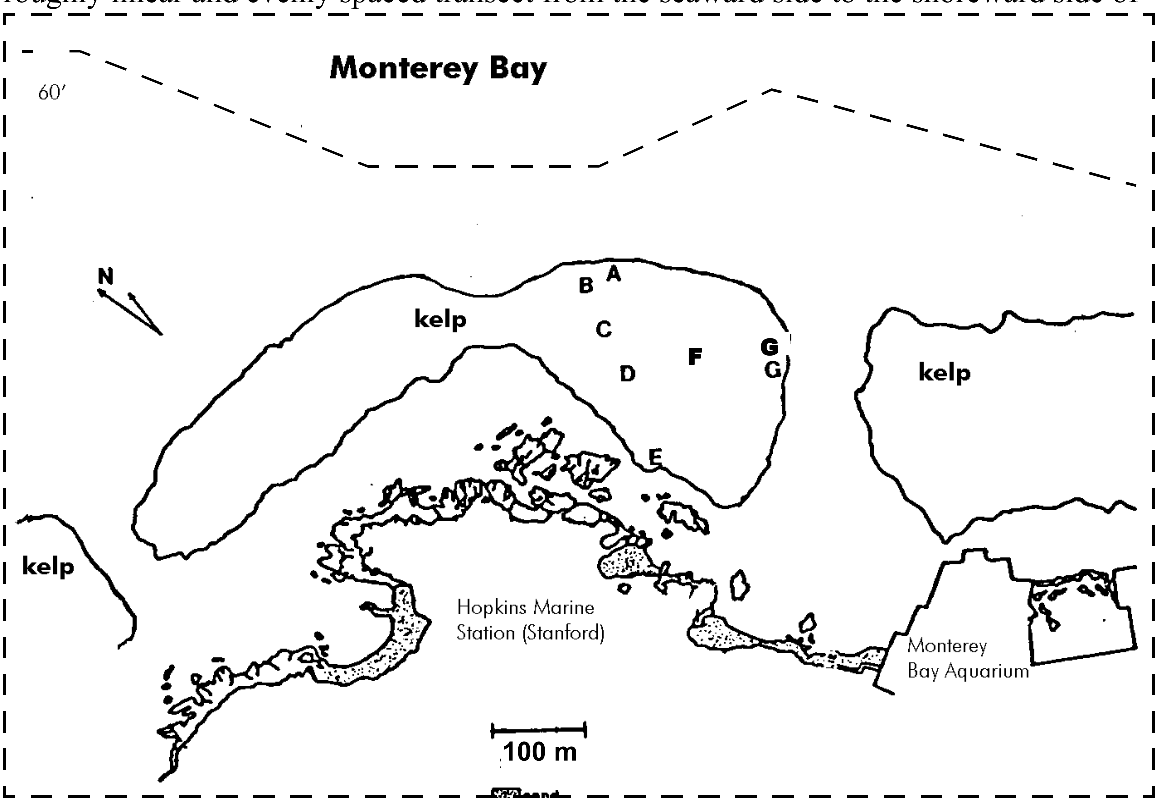
## Statistical Setting

This handout describes a nested ANOVA with subsampling within replicates of the factor of interest.

## Background

Shannon Fairres (ex-Botany grad student) wanted to determine how the amount of water flow varied in various parts of beds of giant kelp (*Macrocystis pyrifera*). Giant kelp is a marine brown alga found along the Pacific coast of North America. Buoyant fronds grow from an attachment ("holdfast") on the sea floor, reaching lengths of up to 60 m. The upper parts of the fronds often form a dense canopy at the sea surface.

Specifically, Shannon measured water motion at seven locations in one kelp bed in Monterey Bay (California). As shown on the map below, five locations (A-E) formed a roughly linear and evenly spaced transect from the seaward side to the shoreward side of




the kelp bed (about 250 m), and the other two locations (F, G) were on a line running parallel to shore from the transect to the side of the kelp bed. Shannon's expectation was that there would be less water motion further from the open ocean (closer to shore). She determined that there was enough water motion at all sites to cause considerable seasickness.

Water motion was measured by a simple and inexpensive method: buttons of Plaster of Paris were attached to kelp fronds, and the percent change in weight over 24 hours was used as an index of cumulative water motion.

At each location two kelp fronds were quasi-randomly selected. Ten Plaster of Paris buttons were attached to each frond. Unfortunately, some buttons were removed by seals or gulls or could not be relocated, producing a mildly unbalanced design. While this could easily be analyzed (and would give results very similar to those shown here), to simplify the discussion I have randomly selected 8 observations per frond to re-create balanced data set.

*Data*

Data are percent weight loss per button in 24 hours, by site and frond.

| A | | B | | C | | D | | E | |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 |
| 50.2 | 59.5 | 44.0 | 59.3 | 47.6 | 41.9 | 36.5 | 27.5 | 24.2 | 31.2 |
| 42.7 | 50.0 | 54.2 | 56.8 | 42.0 | 39.9 | 36.1 | 30.3 | 23.8 | 36.3 |
| 55.8 | 61.5 | 55.6 | 51.2 | 47.5 | 44.3 | 35.5 | 26.4 | 24.7 | 26.3 |
| 53.5 | 55.9 | 54.7 | 56.2 | 48.1 | 39.4 | 35.7 | 31.0 | 20.1 | 31.6 |
| 55.9 | 59.7 | 47.8 | 51.8 | 43.3 | 39.6 | 33.9 | 39.0 | 23.8 | 30.0 |
| 53.6 | 54.0 | 51.5 | 59.4 | 47.0 | 44.7 | 40.5 | 34.0 | 24.6 | 42.4 |
| 52.3 | 52.8 | 55.7 | 58.1 | 48.1 | 39.5 | 35.1 | 33.3 | 25.5 | 27.2 |
| 57.4 | 53.5 | 49.7 | 59.5 | 46.6 | 37.4 | 28.5 | 36.5 | 21.1 | 33.0 |

| F | | G | |
|---|---|---|---|
| 1 | 2 | 1 | 2 |
| 28.4 | 39.3 | 35.5 | 37.8 |
| 32.8 | 34.5 | 43.2 | 31.3 |
| 42.5 | 37.7 | 28.3 | 35.5 |
| 37.3 | 37.0 | 35.3 | 34.2 |
| 37.8 | 36.7 | 36.4 | 37.5 |
| 30.8 | 34.2 | 28.9 | 35.3 |
| 37.7 | 33.0 | 26.4 | 29.6 |
| 39.2 | 29.5 | 45.4 | 34.5 |

## Preliminary Data Exploration

*Descriptive Statistics*

| frond | | N | MEAN | MEDIAN | STDEV | site mean |
|---|---|---|---|---|---|---|
| A | 1 | 8 | 52.67 | 53.55 | 4.63 | 54.27 |
| | 2 | 8 | 55.86 | 54.95 | 4.01 | |
| B | 1 | 8 | 51.65 | 52.85 | 4.23 | 54.09 |
| | 2 | 8 | 56.54 | 57.45 | 3.34 | |
| C | 1 | 8 | 46.28 | 47.25 | 2.32 | 43.56 |
| | 2 | 8 | 40.84 | 39.75 | 2.57 | |
| D | 1 | 8 | 35.22 | 35.60 | 3.33 | 33.74 |
| | 2 | 8 | 32.25 | 32.15 | 4.31 | |
| E | 1 | 8 | 23.48 | 24.00 | 1.88 | 27.86 |
| | 2 | 8 | 32.25 | 31.40 | 5.18 | |
| F | 1 | 8 | 35.81 | 37.50 | 4.71 | 35.525 |
| | 2 | 8 | 35.24 | 35.60 | 3.10 | |
| G | 1 | 8 | 34.92 | 35.40 | 6.90 | 34.69 |

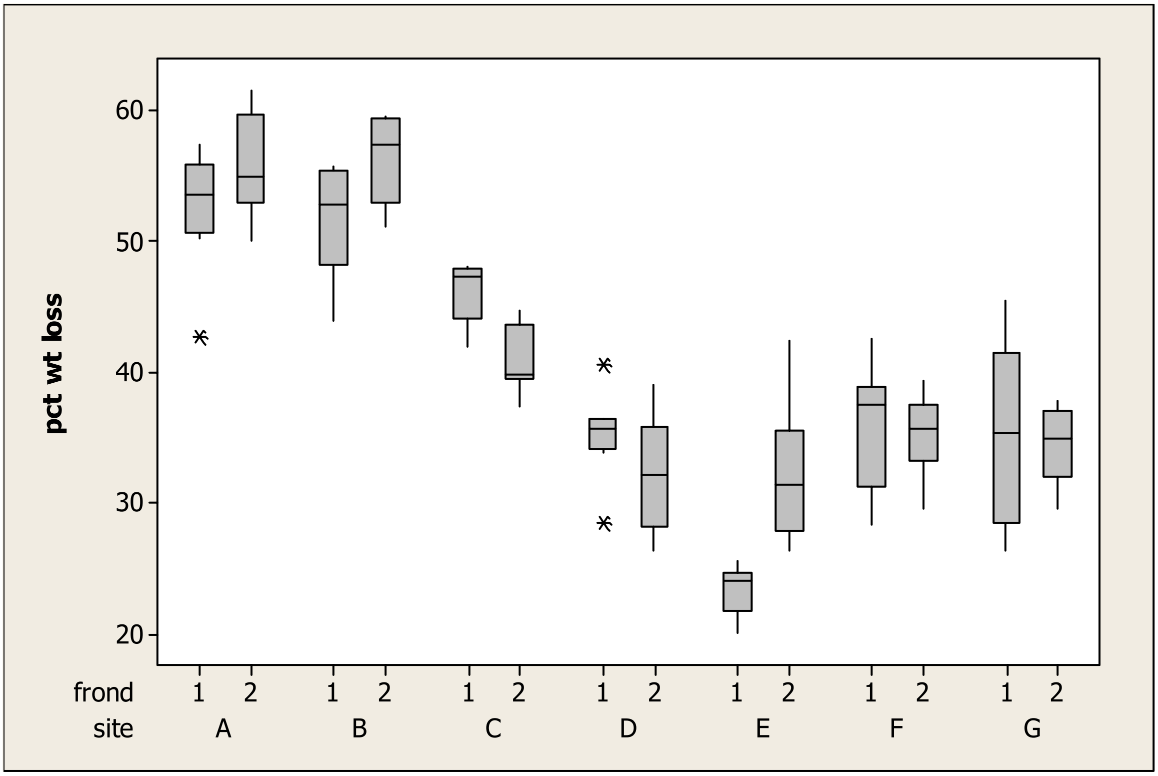
The means and medians indicate that there is variation in water motion across the kelp bed, with greater motion (loss from the buttons) at the seaward sites (A and B), progressively less towards shore (C through E, with F and G comparable to D). There also is variation between the two fronds at each site, but this is considerably less than the variation among sites.

The similarity of the means and medians indicates that the distributions of measurements within each frond generally are fairly symmetrical. The standard deviations show that the within-frond variability differs considerably among fronds: the largest standard deviation (frond 1 at site G) is nearly four times as large as the smallest (frond 1 at site E). The standard deviations do not appear to be systematically related to the mean levels or to location.

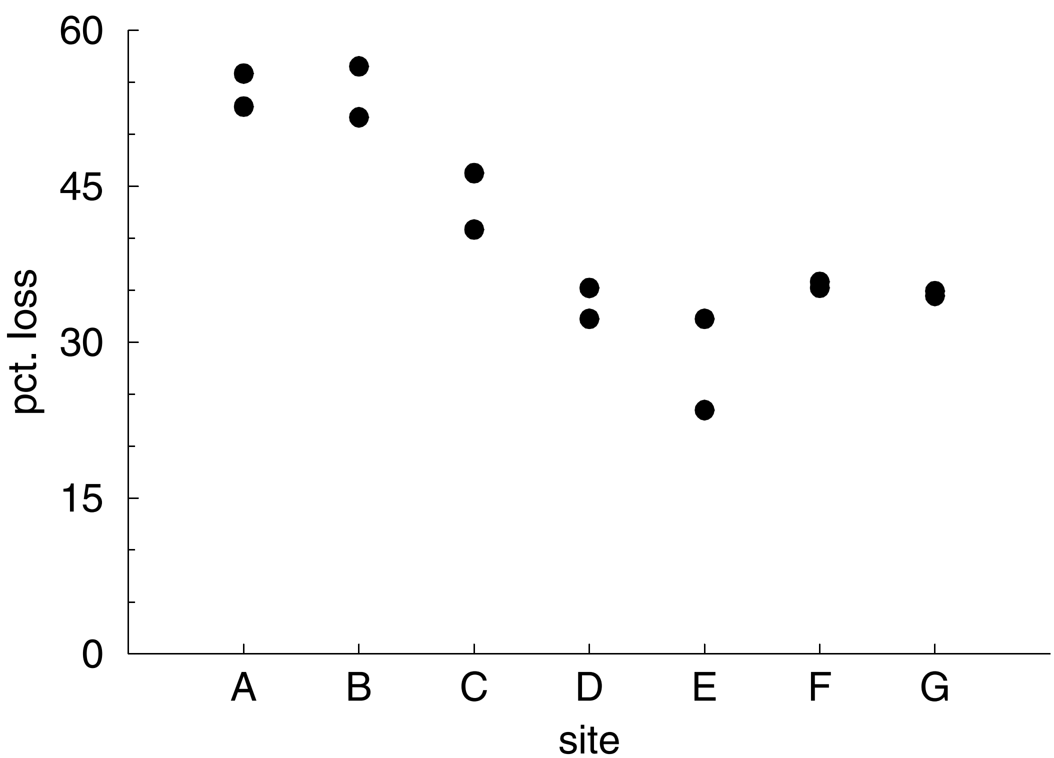
*Plots*

The boxplots and plot of means on the next page very clearly show the decrease in water motion from site A to site E, and the similarity of sites F and G to site D. The boxplots indicate a few possible outliers, but no clearly aberrant observations. They also show the differences in spread among the fronds, and the generally symmetric shape of most of the distributions.

Boxplots by frond:




Frond means *vs.* site:



## Analysis

The design of the study was a doubly nested one, with fronds nested within sites, and buttons nested within fronds. The site factor was a fixed effect: Shannon wanted to compare specific locations which differed in their proximity to the open ocean. The frond effect, however, was random: the selected fronds were of no particular interest but rather were chosen as representatives of fronds at a given site. The fronds were replicates, and the buttons on each frond were subsamples. In the jargon of subsampling designs, the among-frond (within-site) variability is "experimental error" and the among-button within-frond variability is "observation error."

*Nested ANOVA with Subsamples*

Using the individual observations (buttons), the data set would include columns identifying the site and the frond, along with one containing the response variable. The model would include site, and frond nested in site, which is specified as frond(site). The "error" term is the variation among buttons on a frond.

| Source | DF | SS | MS | F | P |
|---|---|---|---|---|---|
| site | 6 | 10361.63 | 1726.94 | 20.15 | 0.000 |
| frond(site) | 7 | 600.04 | 85.72 | 5.31 | 0.000 |
| Error | 98 | 1580.84 | 16.13 | | |
| Total | 111 | 12542.51 | | | |

| Source | Variance component | Error term | Expected Mean Square (using unrestricted model) |
|---|---|---|---|
| 1 site | | 2 | (3) + 8(2) + Q[1] |
| 2 frond(site) | 8.699 | 3 | (3) + 8(2) |
| 3 Error | 16.131 | | (3) |

The EMSes confirm that the frond-within-site term (the true replication) is the appropriate error term for testing the site main effect.

We conclude from these results that there is statistically very significant variation in water motion among sites, as well as between fronds within at least one site.

*One-way ANOVA on Frond Means*

Alternatively, the averages of the observations on each frond could be used in a one-way ANOVA with site as the factor, and the two fronds per site as the replication.

| Source | DF | SS | MS | F | P |
|---|---|---|---|---|---|
| site | 6 | 1295.2 | 215.9 | 20.15 | 0.000 |
| Error | 7 | 75.0 | 10.7 | | |
| Total | 13 | 1370.2 | | | |

With respect to the analysis of differences among sites, these two analyses are exactly equivalent. The $F^*$ for site, and the *df* for site and frond (= error in the second analysis) are exactly the same, and the SSes and MSes differ by exactly a factor of 8 (within rounding error).

*Analysis of Site Effects*

Given the highly significant variation among site means, it is appropriate to examine the nature of these differences. Since the design is balanced, the mean percent weight loss of the 16 buttons at each site (i.e. $\bar{Y}_{i..}$) is an unbiased estimate of the true mean at that site (i.e. $\mu_{i.}$). Those means, in roughly the actual spatial pattern, were:

```
B 54.094   A 54.269
C 43.556
D 33.737        F 35.525      G 34.694
E 27.863
```

<u>All Pairwise Comparisons (Tukey):</u>

The simplest (most mindless) approach to analyzing these effects would be to ignore the spatial pattern and Shannon's expectation of greater water motion closer to the open ocean, and simply do all possible comparisons between means. For instance, the 95% Tukey confidence interval for the difference in means between sites $i$ and $i'$ would be

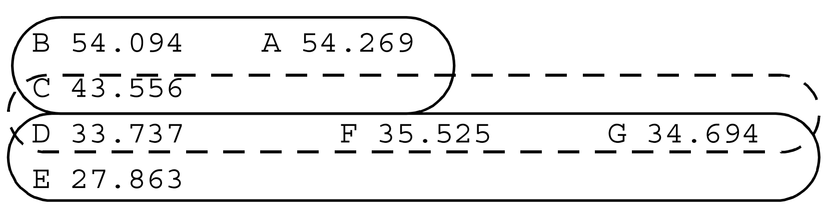
$$(\bar{Y}_{i..} - \bar{Y}_{i'..}) \pm \left(q_{0.95;7,7} \times \sqrt{\frac{\text{MSF(S)}}{16}}\right)$$

$$= (\bar{Y}_{i..} - \bar{Y}_{i'..}) \pm (5.61 \times 2.3146)$$

$$= (\bar{Y}_{i..} - \bar{Y}_{i'..}) \pm 12.9851$$

Equivalently, any two means differing by more than 12.9851 (the "honestly significant difference") would be significantly different at an overall Tukey significance level of $\alpha$=0.05. Of course the easiest way to do these comparisons would be as part of the one-way ANOVA on the frond means.

This approach yields the following conclusions, with ovals enclosed sets of means which are not significantly different:




Or in more traditional representation:

Comparisons of Adjacent Sites (Bonferroni):

A somewhat more structured analysis, taking into account the ordering Shannon expected, would be to compare only sites which are "adjacent" in terms of distance from the ocean: A–C, B–C, C–D, D–E, G–F, and F–D. With a total of six comparisons, the Bonferroni-adjusted minimum significant difference, for $\alpha$=0.05, is

$$t_{1-(0.05/(2\cdot 6)),\,7} \times \sqrt{\frac{2\mathrm{MSF(S)}}{16}} = 3.6358 \times 3.2734 = 11.9013$$

Since this minimum difference is only slightly less than for the Tukey comparisons above, despite the reduced number of comparisons (this reflects the difference in power between the Tukey and Bonferroni approaches), the results of the pairwise tests are identical to those shown above.

Linear Contrast:

The most structured possible approach would be to regress water motion (percent weight loss of the buttons) against actual distance from the seaward edge of the kelp bed. Since I do not have those distances I could not do this. The spacing of sites A/B through E, and G to D, however, is roughly even (A and B being similar in distance from the edge). Using sequence number along each of the transects as the explanatory variable therefore would give nearly the same result as using distance. This in turn is equivalent to doing linear contrasts on the means. Specifically, the effect of distance from the seaward edge could be tested by the contrast

$$C_1 = -3\frac{(\mu_A + \mu_B)}{2} - 1\mu_C + 1\mu_D + 3\mu_E$$

and the effect of the distance in from the side of the kelp bed could be tested by the contrast

$$C_2 = -1\mu_G + 1\mu_D$$

These contrasts are estimated simply by substituting in the appropriate site means, giving $\hat{C}_1 = -88.7745$ and $\hat{C}_2 = -0.957$

The standard errors of these contrasts are

$$s_{\{C\}} = \sqrt{\frac{\mathrm{MSAB}}{16}\sum_i c_i^2}$$

where the $c_i$ are the constants each sample mean is multiplied by in the contrast. The standard errors for the two contrasts are $s_1 = 9.1137$ and $s_2 = 3.2734$. Using the contrast estimates and standard errors to give *t*-tests gives

$t_1 = -88.7745 / 9.1137 = -9.7408$ and

$t_2 = -0.2924$.

With 7 *df*, these tests give *P*-values of $P_1 < 0.0001$ and $P_2 = 0.78$.

I would conclude thus that: (1) there is very strong (very highly significant) evidence of a linear decrease in water motion over the sequence of sites at increasing distances in towards shore (*i.e.* from A/B through E), but (2) there is no evidence of a trend with increasing distance into the bed from the side of the bed at a (roughly) constant distance in from the ocean (*i.e.* from G to D).

We might go further and look for curvilinear trends along the sea-shore transect, *e.g.* by a quadratic contrast. I did not do this, for two reasons. First, since we are not dealing with actual distances, I don't see the point in being so detailed. Second, the linear contrast accounts for a very large fraction of the total variation among sites. Specifically, the SS for the linear contrast is 8135.2, while the SS for site in the ANOVA is 10361.6: the linear trend accounts for over 78% of the variability among sites. The contrast SS can be gotten from SAS (see last section of this handout), or can be calculated from the *t* statistic above:

$$\begin{aligned} SS\{contrast\} &= MS\{contrast\} \text{ x } df\{contrast\} \\ &= MS\{contrast\} \text{ (since df=1)} \\ &= F\{contrast\} \text{ x } MSF(S) \\ &= t^2 \text{ x } MSF(S) \end{aligned}$$

*Variance Components*

The variances among fronds ("experimental units") and among buttons ("observation units") can be estimated in the usual way of taking differences between appropriate MSes and dividing by the appropriate constant. For example, $\sigma^2_F$ is estimated by

$$s_F^2 = \frac{MSB - MSF(S)}{8} = 8.699$$

The easier way to get the estimated variance components of course is to examine the EMS part of the ANOVA computer output (nested ANOVA results on page 5). From there we see that $s^2_B = 16.131$.

The conclusion here is that the variability on a very fine spatial scale (among buttons on a single frond) is nearly twice as great as that on a somewhat larger spatial scale (among fronds at a site). This to me is surprising. Perhaps this variability within a frond was due to buttons having to be placed on kelp blades along the length of the frond, so that those nearer the tip were more exposed than were those closer to the base of the frond and therefore more down inside the kelp bed?

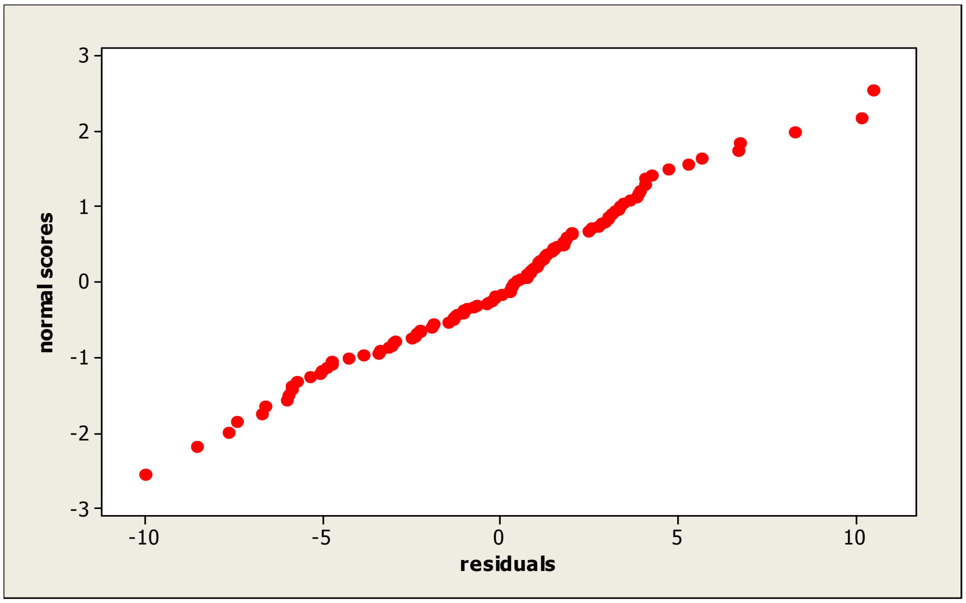
**Aptness of Model**

The residuals (deviations of observations from their respective "cell" [frond within site] means) were obtained for use in assessing the validity of the ANOVA assumptions.

Transformation?

It is common practice to arcsine-square-root transform percentage (or proportion) data such as these. This transformation has very little effect on percentages between 20% and 80%, which all those in this study were. I did repeat the analysis and diagnostics on transformed data, but as expected there was essentially no difference.

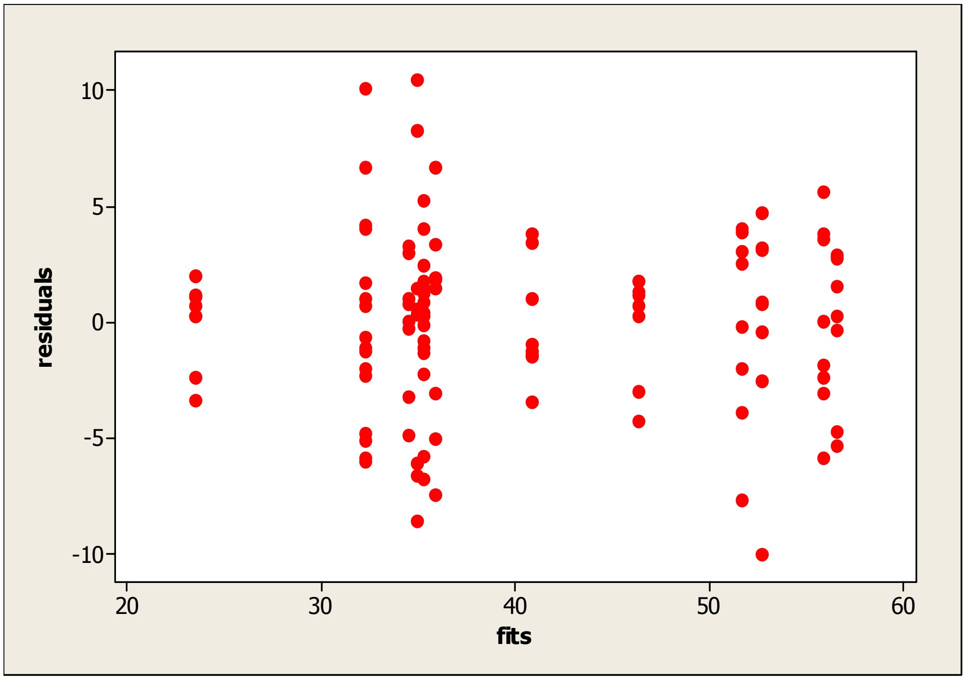
Normality:




The normal scores plot (above) and histogram of the residuals (not shown) show the distribution of the residuals to be very close to a normal distribution. The correlation between the residuals and their normal scores was 0.991, further indicating very good agreement between this distribution and a normal distribution. These conclusions are consistent of those concerning symmetry, in the preliminary data exploration.

Constancy of Error Variance:

As noted during the preliminary data analysis, the variability among subsamples (buttons on a frond) differed considerably among fronds: the largest within-frond standard deviation was about 3.6 times as large as the smallest. Given that this is a balanced design (giving maximum robustness to unequal variances), I do not consider this a serious threat to the validity of the ANOVA. Furthermore, the variability in residuals showed no apparent relationship to the fitted value (frond mean), so that it is unlikely any transformation could substantially reduced the heterogeneity of variances; this lack of relationship between means and standard deviations was noted in the descriptive statistics, and is confirmed by plotting residuals against means, as below.




**Study Design: Effect of Subsampling**

*Optimal Allocation*

If we knew the "costs" (presumably in time) of additional fronds *vs*. additional buttons per frond, these variance estimates could be used to determine the optimal amount of subsampling (i.e. number of buttons per frond). Without knowing these relative costs, all we can say (assuming a model of additive costs applies) is that the estimated optimal number of subsamples is

$$m_{opt} = \frac{s_B}{s_F}\sqrt{\frac{c_F}{c_B}} = 1.362\sqrt{\frac{c_F}{c_B}}$$

where $c_F$ is the cost of sampling an additional frond (at each site) and $c_B$ is the cost of using an additional button on each frond. I doubt that it would take much more time to move the boat to a different frond than to add another button on the same frond, in which case $c_F$ would be quite small relative to $c_B$, and only one button per frond should be used.

If adding a new frond required not only moving the boat but also getting in and out of it with diving gear, and if sticking a button on a frond is very fast and easy, this would cause the ratio of costs (*i.e.* $c_F/c_B$) to be considerably larger than one, so that at least some subsampling (*i.e.* multiple buttons per frond) would be advantageous. For the 8 subsamples which Shannon used to have been optimal, $c_F/c_B$ would have to be about 35: it would have to take 35 times as long to move to a new frond as to put another button on a frond. This seems unlikely to me, but then what do I know about small boats, kelp beds, and big waves?

*Value of Subsamples*

The subsample observations (buttons) are not used directly in comparing sites; even when they are included in the analysis, in the nested design, it is the variation among frond means which is used as the "error" for comparing sites. And as discussed above, it is unlikely that this degree of subsampling was optimal.

This does not mean, though, that the subsampling was not useful. If only one button had been used per frond, and still only two fronds per site, the degrees of freedom in the ANOVA of sites would be the same as in this handout. The variance of the single observations, however, would have been eight times the variance of the means of the eight subsamples. This difference in study design, though, would not affect the site means, and thus the `Q[1]` term contributing to the site MS would be unchanged. The expected means squares for the design without subsampling therefore would be

EMS(site) $= 8\sigma^2_f + Q[1]$

EMS(error) $= 8\sigma^2_f$

where $\sigma^2_f$ is the variance among frond means in the design with subsampling. Using the current data, $\sigma^2_f$ is estimated to be 10.7, and Q[1] can be estimated analogously to how variance components are estimated, as MS(site) − MSE = 205.1. If there had not been subsamples, we would expect MSE to be around 85.6, and MS(site) to be around 85.6 + 205.1 = 290.7, yielding an $F^*$ around 3.4 rather than 20, and a *P*-value between 0.1 and 0.05.

In short, putting only one button per frond on only two fronds per site, while certainly faster and easier than using 8 buttons, would have provided much less precise information on water motion at each frond, and thus a much less powerful analysis.

**Conclusions**

The ANOVA is a valid analysis, whether done as a nested design with subsamples or as a one-way analysis of the frond means, and yields clear and sensible conclusions:

(1) There are differences in water motion among the seven sites ($P < 0.0005$).

(1) The differences are primarily a roughly linear decrease in water motion with increasing distance from the seaward edge of the kelp bed ($P < 0.0001$).

(1) There is no evidence that water motion varies with distance in from the side of the kelp bed ($P = 0.78$).

(1) There is more variation among buttons on a single kelp frond than among nearby fronds (perhaps reflecting differences in depth into the bed), so that a sampling design using more fronds and fewer (probably only one) subsamples per frond would have been more efficient.

An important qualification to these conclusions (especially (2) and (3)) is that there is no replication of distances into the bed, so interpreting the differences seen among sites as being due in fact to distance from the edge is less certain than if the same pattern had also been seen at one or more other kelp beds. The nice, roughly linear relationship of water motion to distance from the edge, however, does strengthen the argument for interpreting the "site" effect in this way.