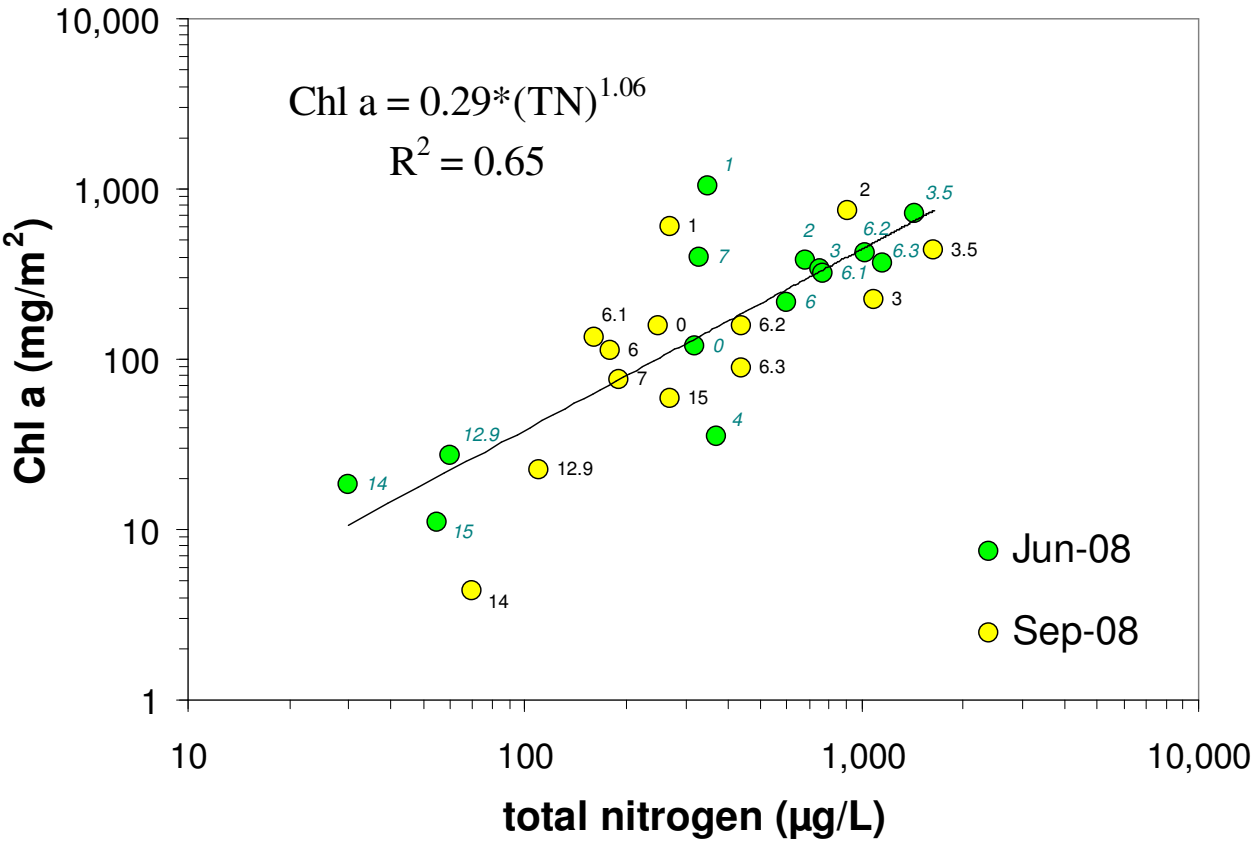
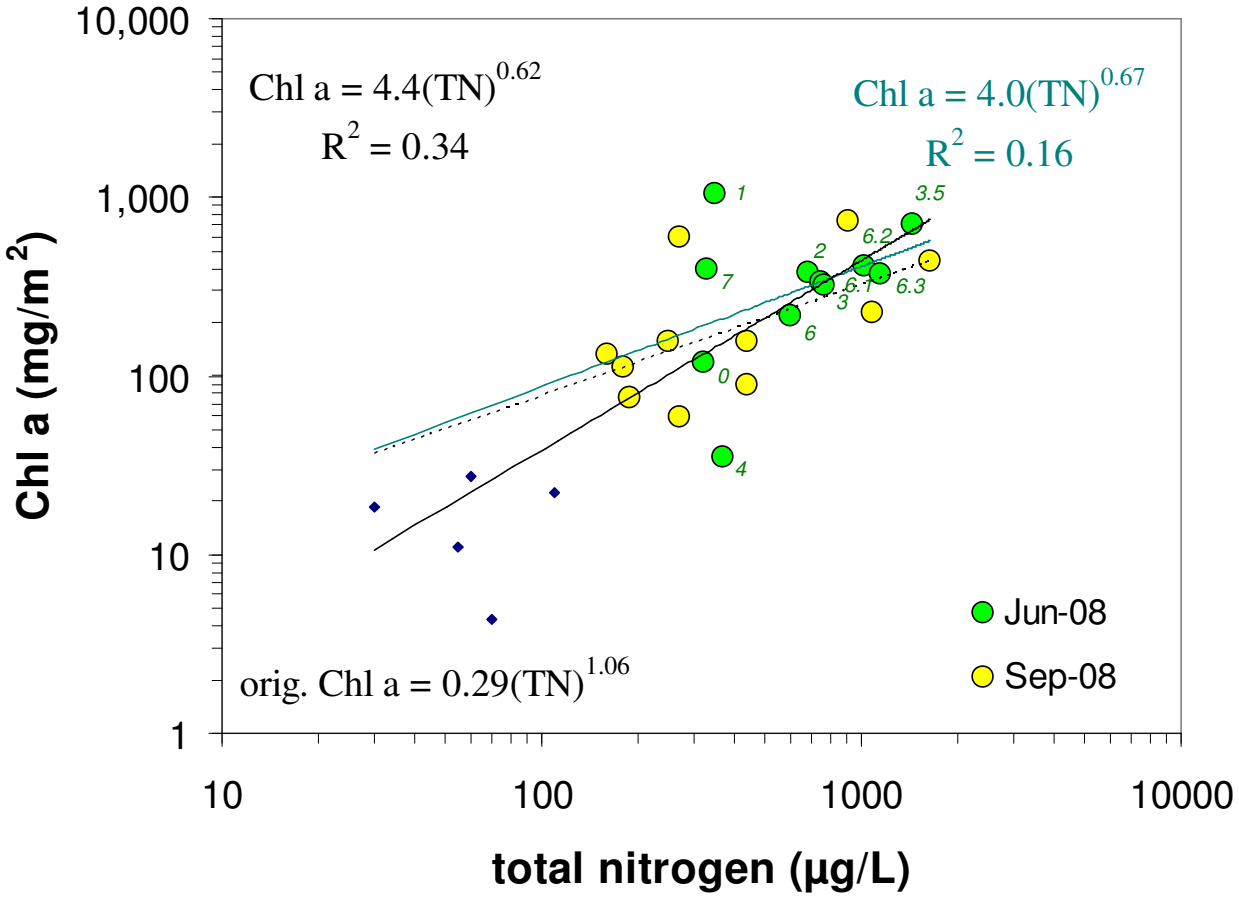


This is a plot of Kristie's Chl a vs. total nitrogen (TN) data. The regression line is for the combined dataset, but I've shown results from each approximate sampling date separately as well as indicating the SBCK site number for each point. My first comment is it looks like a reasonable regression (the original data, both Chl a and TN are not normally distributed and a log-log transformation not only yields the best regression coefficient, but is necessary to satisfy the normality requirement for regression validity). The relationship for June data is slightly better than for September, but both datasets yield similar equations. Taken at face value, the regression shows that total nitrogen concentrations can explain 65 % of the variation in Chl a.

My basic reservation is the same one I stated at the beginning of this research: that the validity of this relationship is partially dependent on the time chosen for field sampling. There was a bias built in to the data collection (inadvertent, I'll freely admit) due to delayed sampling that missed the first algal bloom – the peak of which occurred prior to late-May/early-June. Ordinarily, a May/June sampling window would not have been a problem, but in 2008 winter rains ended early, in February, giving way to an early spring and, consequently, an unusually early first bloom. By mid-May/early-June the bloom had not only ended at the Matilija and upper San Antonio sites, but much of the algal mass had already disappeared.

So my first point is that this graph may well exaggerate the importance of total nitrogen to algal density. Secondly, that the validity of the regression is heavily dependent on these very same suspect Matilija points. And third, that the failure to show confidence interval error bars for the mean Chl a values gives a mistaken impression of the relative precision of this measurement – which happens to be much lower than that of TN.

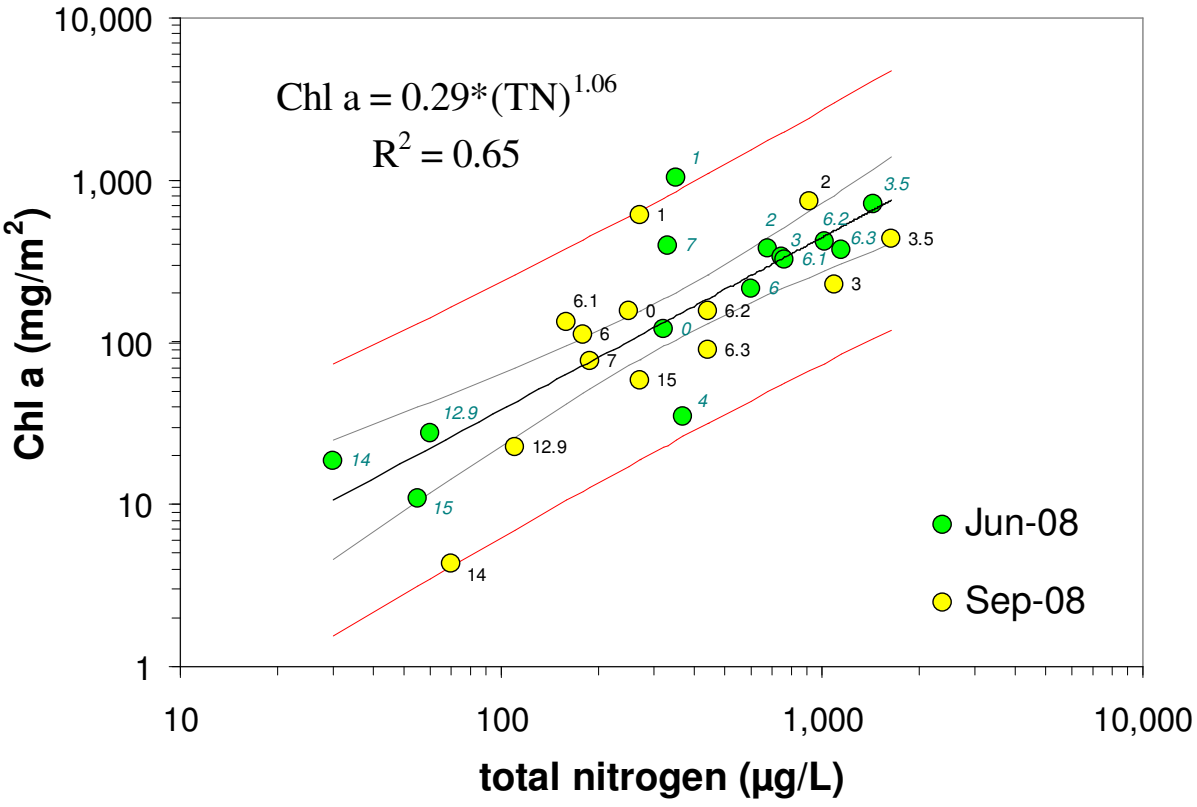


This graph addresses the heavy dependence of the regression on the Matilija surveys. (That a simple linear regression produces an r-squared of 0.21, i.e. TN can explain only about 20 percent of the Chl a variation, makes the same point.) I've excluded 5 of these 6 points and re-calculated the regression parameters. Exclusion of the Matilija data (I'm including VR12.9, the Ventura River at Camino Cielo, in the "Matilija" group since it's less than a mile below the Matilija confluence) produces a much less convincing relationship – especially for the earlier survey.

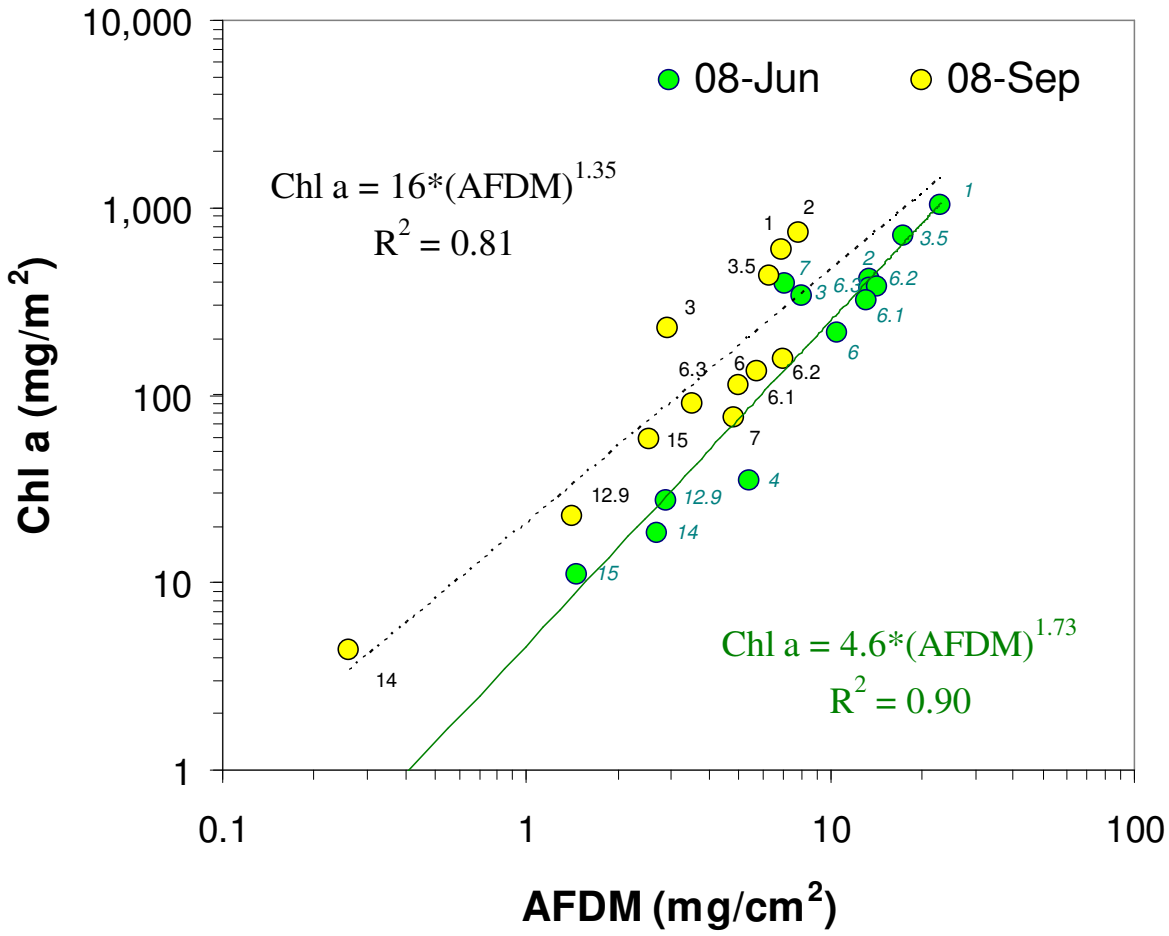
Perhaps the most obvious feature of these June data are the three points with almost equal TN concentrations, but vastly diverging Chl a (4, 7 and 1). (I exclude the lagoon data from this observation for many reasons, but mainly because there is no way of knowing what exactly it represents.) There is also a major grouping of points with what I might term "centrist values," characterized by relatively close Chl a means (215-380) but more widely diverging TN concentrations (600-1150). This leaves only the data point for VR03.5, the site immediately below the treatment plant on which the regression is also highly dependent, unaccounted for. Exclusion of this point would further reduce the r-square (June data) to 0.08, i.e. TN now explaining less than 10 of the variation.

I'm struck by the arbitrariness of the regression, by how dependent it is on site selection and timing. Had the initial survey been done in April results might have been very different (the placement of September VR15 data, the one Matilija point I didn't exclude – thanks to a late second bloom of spirogyra – gives some indication on how tightly grouped the Chl a data might have been).

Similarly, had the second survey been delayed a month or so, when watercress made its resurgence on the lower river, results would also have been different – especially at 03.5 where macroalgae were no longer in evidence, replaced by a massive expansion of watercress (with diatoms playing a subsidiary role). By May algae had so diminished on San Antonio Creek that only the 07 site was included in the survey. The inclusion of sites like VR10 (where algae were noticeably prevalent in April with TN circa 4,000 µgm/L) would have also changed the impression.



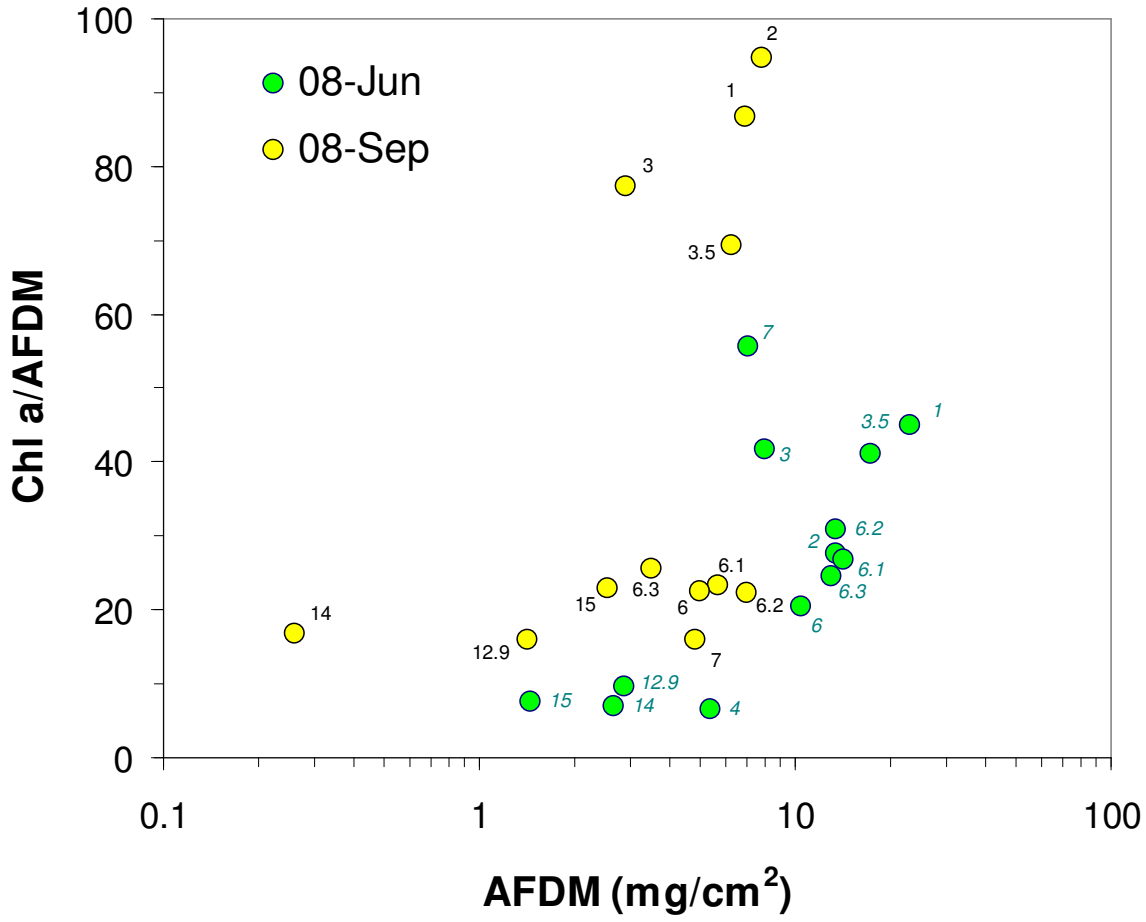
I wanted to point out the low level of predictability of the Chl a vs. TN relationship. I've added to the graph the 95 % confidence intervals for both a mean prediction (grey) and an individual prediction (red). Knowing the total nitrogen concentration will allow you to predict (granted prior acceptability of the regression model), with this level of confidence, Chl a within roughly two magnitudes, i.e. only slightly better than the 4 to 1000 mg/sq-meter total range of values actually measured (the confidence interval for a single estimate extends from roughly 80 % lower to 500 % higher than the predicted value). As an example, a TN of 400 µg/L yields an estimated single measurement of 166 mg/sq-meter for Chl a (c.i. of 28 to 978); the confidence interval for a mean measurement is better (118 to 234), but still straddles the generally assumed boundary of 150-200 for acceptable Chl a. I should point out that 400 µg/L meets both the California and EPA TN criteria for a water body in good condition.



In view of an earlier assertion, that while Chl a may have appreciably decreased at sites surveyed passed the peak of the algal bloom, ash-free-dry-mass (AFDM) would not be greatly changed (retaining, so to speak, "the ghost of algae past"), I've been interested in how the relationship between AFDM and Chl a would turn out. In the above graph I've plotted Kristie's data, again separating the two surveys and showing SBCK site numbers. I've also plotted separate log-log lines-of-best-fit for each survey.

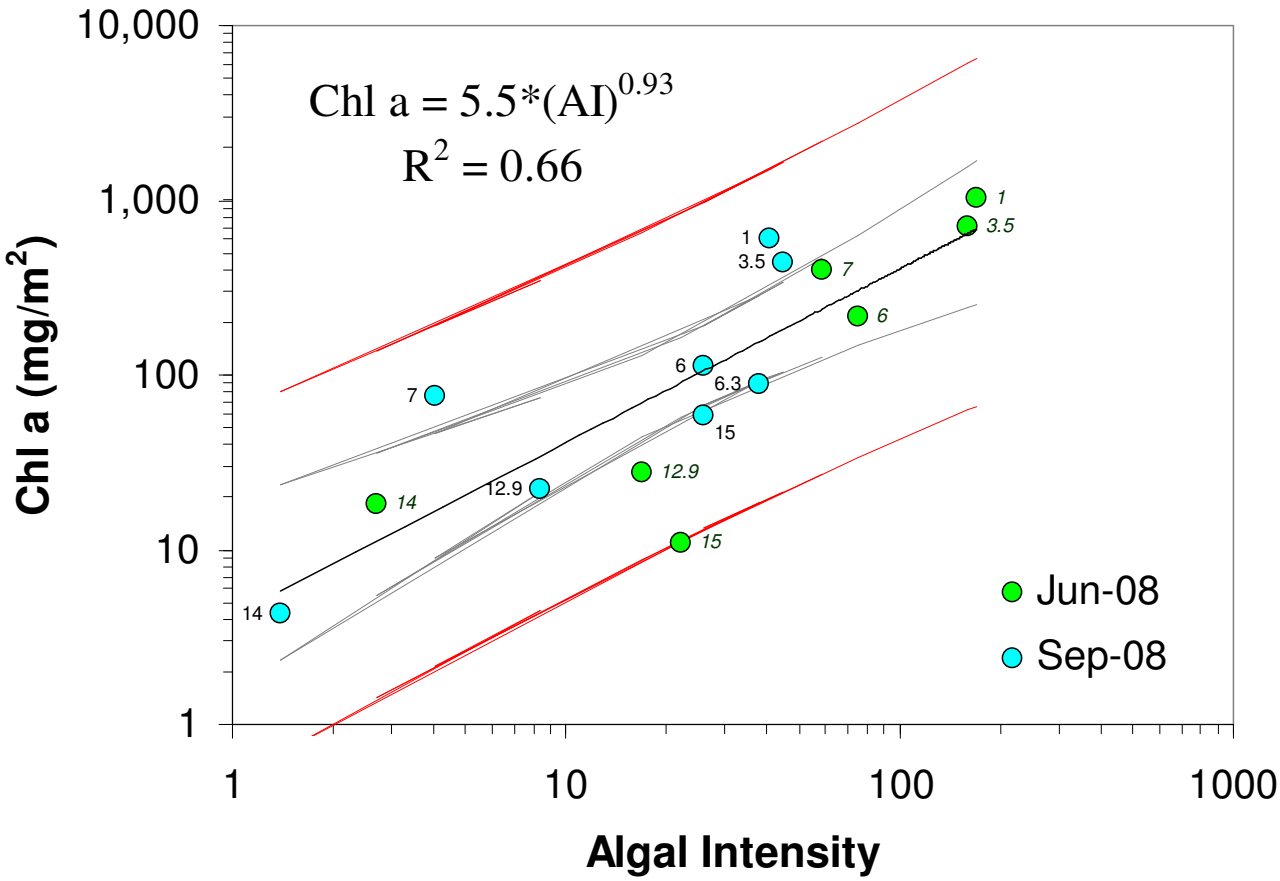
I find a number of curious things in the data: mainly very different relationships for the two surveys, with September data showing a greater amount of Chl a per unit AFDM and an increased absence of linearity (r-squared of around 0.5 vs. 0.8 for the June data). But I'm more interested in why there is any relationship at all. Linear or not. Or, even more basically, what might these wide ranging values, both between locations and at different times, imply?

I'll assume that as an algal bloom develops the relationship between Chl a and AFDM will vary in some defined pattern. My best guess would be a gradual increase in the Chl a to AFDM ratio (as competition for light increases with increasing algal density) followed by a gradual decrease (as shaded out interior algae begin to die even as the bloom is still advancing. This would be followed by a much steeper decrease as the bloom declines and the relative proportion of dead to active algae increases, and as the algal mat becomes enriched with organic detritus, heterotrophic critters and other non-chlorophyllous organisms. I would also assume that different species will exhibit different (perhaps even characteristic) indices.



It turns out there is a formal expression for the Chl a to AFDM ratio: the autotrophic index. Or as it is also alternately expressed: the AFDM to Chl a ratio. The EPA describes periphyton in surface water relatively free of organic matter as having a Chl a/AFDM ratio of 1 to 2%; circa 0.4 or 0.5% as growing in inorganically enriched waters; and < 0.25% as indicative of organically polluted waters. I've shown Kristie's autotrophic indices values in the graph. I've kept her original units (divide by 100 to convert to percent Chl a) and I've again added SBCK site identifiers to each point. The ratios vary from about 0.06 up to 1%.

The data do show the ghost of algae past (the very low June ratio values in the lower left-hand corner – 4, 12.9, 14 and 15, sites where the peak of the algal bloom had been long gone – but the question remains of how much of a ghost might have been left. Another good question might be: why should the lower river sites (1-3.5) have higher percentages of Chl a in Sept. than in June, when any algae present should have been more contaminated with misc. organic matter? And why should these sites be so different from all others? Disregarding these two "islands" of differences, we are left with reasonably consistent values for each of the surveys – with the later survey having a lower proportion of Chl a, as we might expect. Except that the Sept. data are more tightly grouped, even though the assemblages at each site were quite different, while there was greater variation in June, when the blooms were almost uniformly cladophora.



This is the original point of this exercise: seeing how Kristie's measures of algal density match up with my "Algal Intensity" parameter. Algal Intensity is the maximum range of the daily dissolved oxygen cycle (delta-DO, in mg/L) multiplied by flow (in cfs). From now on I'll abbreviate it as AgI (since a simple AI would match the usual shorthand used for autotrophic index). Unfortunately, there were only a limited number of data points where measurements of both Chl a and AFDM, on one hand, and AgI, on the other, were taken within more-or-less the same time frame (mainly because a number of Kristie's locations could not be easily reached by SBCK in the dark). In the graph I've plotted Chl a against AgI, showing SBCK site identifiers and the approximate time-frame of Kristie's surveys. The regression line is for the entire dataset, and the 95 % confidence limits for both a mean and individual prediction are shown.

The correlation is pretty decent: AgI explaining 67% of the Chl a variation. This is roughly the same explanatory power as the TN equation I started with without, of course, the drawbacks discussed – a simultaneity of the algal cycle at each survey location not being required, there is no apples-with-oranges comparison drawback. Unfortunately, however, the prediction confidence intervals are also roughly the same. Still, AgI is a measurement easily made, and in my opinion, offers promise. The regression fit for June data, when algal intensity was at its maximum was better (r-square of 0.74, 0.86 if the anomalous site 15 data point were removed) than in September – as it should be since the parameter has less utility at lower values.

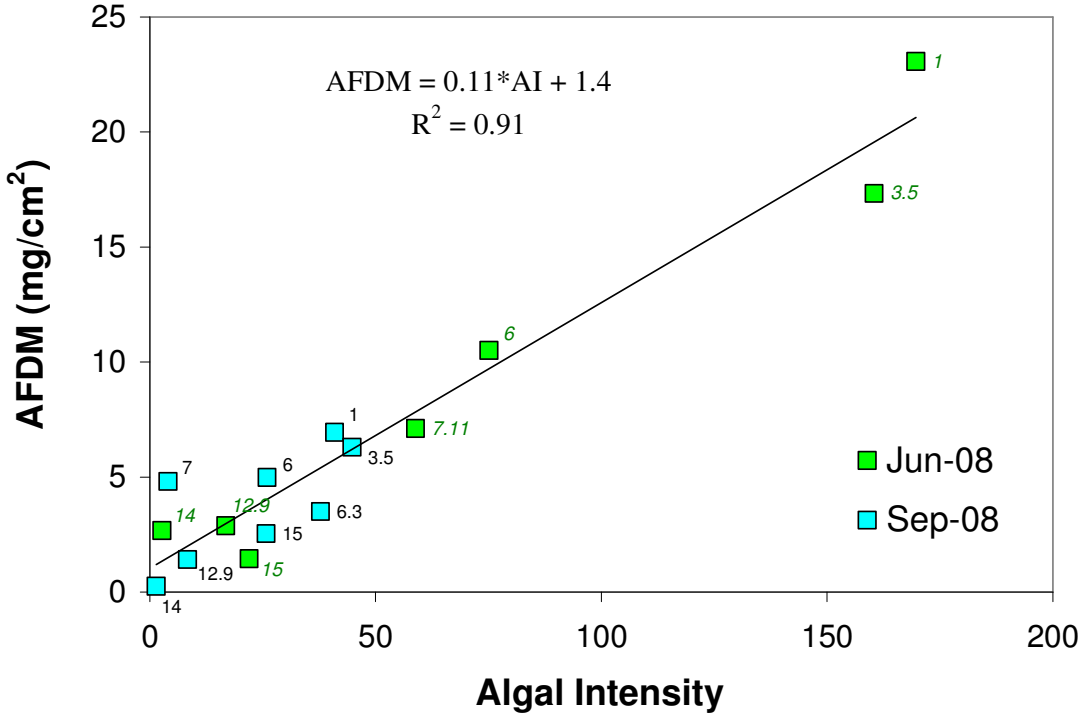
AgI better describes the algal impact on streamflow as flow increases and/or as algal density also increases. At low flows and low algal densities the relationship breaks down, since it is a simplification of the true situation. AgI assumes that the primary impact on dissolved oxygen will be caused by the extremes of algal photosynthesis and respiration. However, other processes depress oxygen (e.g. aerobic decay) or increase oxygen (e.g. physical re-aeration). AgI ignores all other factors, attributing delta-DO to algae alone, a reasonable assumption only as long as the magnitude of algal productivity dwarfs other processes. As the amount of algae increases, and as flow increases (reducing the relative amount of oxygen gain and loss per unit flow via physical processes), this becomes increasingly true. As algal biomass and/or flow decrease, other factors become increasingly important and the utility of AgI as a measure decreases (which is probably why there is an increased dispersion of low-value data points, but the relationship becomes increasingly linear at higher AgI numbers).

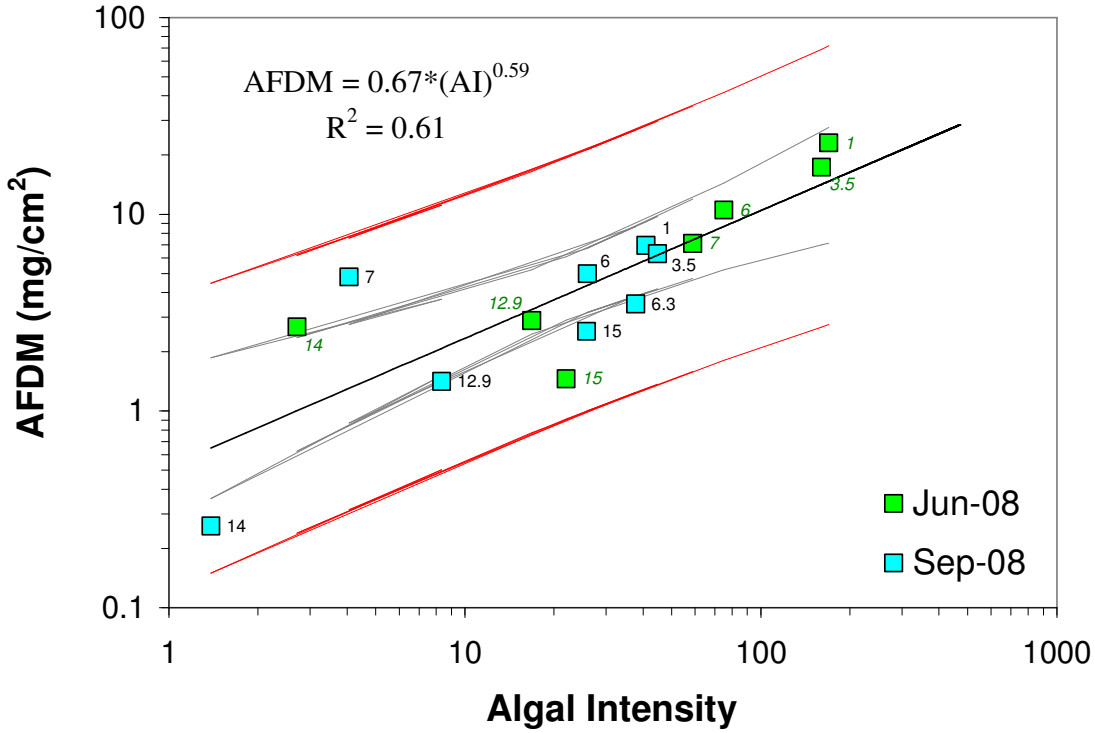
So AgI is an imperfect measure of algal productivity, and algal productivity is not always a good measure of algal mass (e.g. an overcast cold day producing a diminished DO cycle) which is what we are probably trying to measure in the first place. (I came up with the idea by applying a "beats a jab in the eye with a sharp stick" measure of scientific utility. I've never said it was perfect. Easy determination is its chief attraction.)

(There are, however, ways of improving the measurement of Ag. Measuring minimum and maximum DO concentrations over the course of a single 12 hour period by estimating the approximate times of occurrence was, admittedly, rather casual. Hourly measurements over several days using sondes would avoid dependence on haphazard measurements and the vagaries of a single day's weather conditions (which might, given abnormal temperatures or cloud conditions, be uncharacteristic). It would also allow the application of more accurate methods of estimating primary production.)

However, estimating Chl a or AFDM may be even more problematical, not to mention further removed from evaluating the actual consequences of algal growth. My chief complaint about the methodology used in this study (10 transects, 3 "quarter" sized samples per transect) is its time-consuming difficulty and expense, not to mention the high level of expertise required. Although I have not seen the actual data, I have to assume that the variation between transects was wide, and the standard error of the mean values appreciable.. If, as a rough rule, the 95 % confidence interval for a mean is twice the standard error on either side of the mean , the fact that one mean (say at 06.3) is less than another (say at 03.5) doesn't mean (no pun intended) that the two measurements were really different. If the confidence intervals overlap we have no way of being sure that values at the two location were different -- statistically different, that is. It's entirely possible that had we gone out there the next day, or even later the same day, and repeated the exercise the new results would be exactly opposite the first set i.e. 03.5 < 06.3.

Kristie's results may show real differences between sites, or they may not. Only the actual transect data will indicate that. This assumes that the methodology was suitable and precise enough in the first place, and accurately applied in the second. I still have reservations about the repeatability of these measurements; something I have never seen evaluated. Quantifying algae at a location is not an easy or simple problem. And unless the differences between sites are gross and easily distinguished, we may not be able to compare some sites with each other with any degree of confidence.





Finally, here's the plot of AFDM against Algal Intensity (AgI). I've shown it as both a linear and log-log plot since the regular correlation is the one that first caught my eye; but of course AgI is not normally distributed so the r-square value of the power relationship is the more appropriate measure.