Appendix A

Sampling Protocols

Field collection

Phytoplankton monitoring in Lake Champlain was conducted according to a tiered monitoring and alert system. The tiered levels of Qualitative, Quantitative and Vigilance each entail an increasingly intense sampling regime (i.e. increased frequency and toxin analyses) and advancement through the tiers is triggered when potentially toxic cell densities surpass pre-determined limits. The plan was developed to best provide public health official with toxicity information (Watzin et al. 2006). Throughout the summer cyanobacteria abundances were measured with monthly in the spring and every other week in the summer. If counts of cyanobacteria exceeded levels indicating a bloom, weekly sampling was initiated. Nutrient and phytoplankton sampling were conducted between 10 am and 2 pm. Grab samples were collected just below the water surface, in duplicate, for nutrient content and phytoplankton cell counts. All phytoplankton samples were preserved in 1% Lugol's iodine solution. Nalgene high-density polyethylene bottles were used for all nutrient samples except total nitrogen, for which samples were collected in 50 mL polypropylene centrifuge tubes. Total phosphorus containers were cleaned with 20% hydrochloric acid solution prior to use. Nitrogen samples were preserved with sulfuric acid to a pH less than 2 and stored at 4°C until analysis. Total phosphorus samples were frozen until analysis.

Phytoplankton analysis

All net samples were evaluated using a rapid assessment protocol outlined in Watzin et al. (2006). A sub-sample of total samples collected each year was selected for settled whole water

counts. Whole water samples were examined using Utermöhl settling chambers following sedimentation for 1-4 days before enumeration (APHA 1998). Samples were counted using an Olympus IX70 or an Olympus IX71 inverted microscope with phase contrast at 400X. All taxa were identified to at least genus following Prescott (1982). Natural units (primarily colonies for cyanobacteria) were enumerated. Counting continued until 100 units of the most abundant taxa had been observed or 100 fields had been evaluated. For further statistical analysis, mean cell densities obtained from duplicate samples were used. Rapid analysis of net samples yields lower estimates of potentially toxic cell densities than settled counts collected at the same location but is an effective method for rapid notification of public health officials (Rogalus and Watzin 2008).

Water column nutrients

Samples for soluble reactive phosphorus (SRP) were filtered through 0.45 µm membrane filters and were analyzed using the ascorbic acid colorimetric method and a Shimadzu UV-1601 spectrophotometer (APHA 1998). Total P samples were thawed, mixed thoroughly, and an aliquot (generally 50 mL) was digested using ammonium persulfate and analyzed following Quikchem™ Method 10-115-01-1-F using a Lachat Quikchem™ 8000 Series Flow Injection Analyzer (APHA 1998). Total N samples were analyzed using persulfate digestion and cadmium reduction following Quikchem™ Method 10-107-06-2-H using a Lachat Quikchem™ 8000 Series Flow Injection Analyzer (APHA 1998).

Figure A1

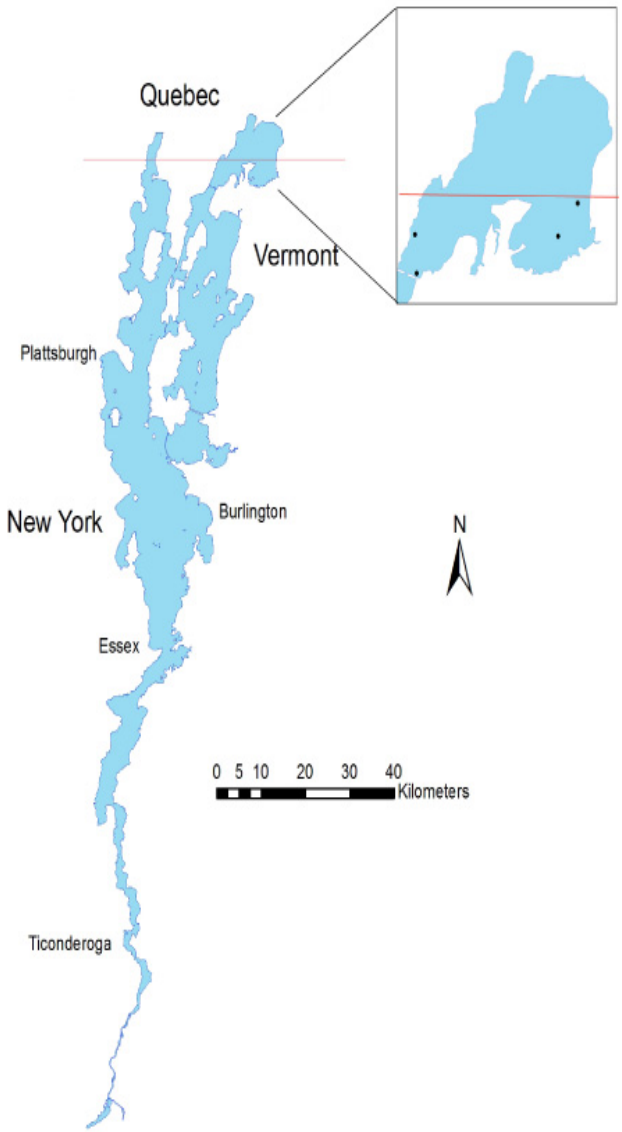

Figure A1: Map of study site: Lake Champlain, Vermont, USA. Inset shows sampling locations in the US waters of Missisquoi Bay.

70 **Appendix B**

71 **Analysis details**

72 *Time series analysis*

73 Three major taxa were sampled by the monitoring program, and we chose the numerically
74 dominant taxa from each year as the focus for our analysis (Figure B1). However dynamics
75 were generally similar among all taxonomic groups. In order to assess non-linearity, stationarity,
76 and feedback order in our time series (Figure B2) we followed the analytical methods of
77 Berryman (1999). We created phase plots of population growth rate (r_t) versus population size
78 (N_{t-1}) for the data from each year to assess non-linearity and look for evidence of feedbacks
79 (Berryman and Lima 2007). We tested process order by examining the partial rate correlation
80 function (PRCF;Berryman and Turchin 2001) and found that the data were all linear and
81 exhibited only first-order feedbacks (a single time-lag of one week). We plotted autocorrelation
82 functions for each time series to assess stationarity (Berryman 1999). Because none of the series
83 were stationary, we used the data splitting method outlined by Berryman (Berryman 1999,
84 Turchin 2003, Berryman and Lima 2006) and broke each series into a bloom phase and post
85 bloom phase. We chose the break point for each time series based on a visual assessment of the
86 raw data and the phase plots (Figure B3).

87

88 We assumed a linear functional form, $f(E_{t-d}) = \beta E_{t-d}$, for both nutrients and competitors ($f(E_{t-d})$
89 $= \beta S_{t-d}$) to minimize the numbers of parameters to be fit. Because the early series had only 19
90 points, we only fit models with a maximum of three total parameters. With these two composite
91 time series, we fit candidate models and assessed model fit using AICc (Aikake's Information
92 Criterion for small sample size; Burnham and Anderson 2002). We tested for the influence of

three different nutrient measurements (Figure B4) and one nutrient ratio: total nitrogen (TN), total phosphorus (TP), soluble reactive phosphorus (SRP) and the ratio of total nitrogen to total phosphorus (TN:TP) on r_0 and the strength of density dependence (e^c). Additionally, we tested for the effects on e^c of the density of four classes of potential algal competitors: Bacillariophyceae (diatoms), Chlorophyceae (green algae), Cryptophyceae (biflagellate cryptomonads), and Myxophyceae (non-toxin producing cyanobacteria) as well as the sum total of all competitors. We tested each nutrient measurement as a single effect with individual terms in equations 2 and 3 at lags 0 and 1 ($d = \{0,1\}$), and tested all four competitor taxa singly in equation 4 at lag 1 ($d = 1$). Because traditional models of competition model the effects of competitors on carrying capacity (Gotelli 2001) of the target species, we did not test models of direct effects of competitors on growth rate (r_0). We tested models in which growth rates merely tracked nutrients with no density dependence, a model in which there was simple density dependence with no effects of nutrients or competitors ($\theta = 1$ in equation 1), and null models of exponential growth ($r_0 > 0$) and a random walk ($r_0 = 0$). This resulted in 34 total that were tested (Table B1).

Linear discriminant analysis of the separatrix

To evaluate the existence of multiple population equilibrium points, Berryman (1999) recommends a phase plot with the x-axis as log (N_{t-1}) and visual assessment of whether or not a line could be drawn cleanly between points from the different phases (a separatrix). We wanted to quantify how well our visual separatrix worked by testing it against the null hypothesis that an arbitrary set of breakpoints could be chosen and would separate into two different equilibrium points equally well. To do this we used a quadratic discriminant function (Gotelli and Ellison

116 2004) to create a linear combination of variables to predict membership in the bloom or decline
117 phase.

$$Z_i = a_{i1}Y_1^2 + a_{i2}Y_2^2$$

118 In this equation, $Y_1$and Y_2 are growth rate and population size at *t-1* and the model assesses how
119 well a linear or quadratic transformation of these variables predict bloom or no bloom status? In
120 the randomization, we tested every possible break point, with the restriction that every series had
121 to have at least one point in either the bloom phase or decline phase. Random group assignments
122 were constrained because a point within a series at *t=3* could not be assigned to the decline phase
123 if at *t=4* it was in the bloom phase (Figure B2). Even with this restriction, there were more than
124 250,000 possible break point combinations. For each breakpoint combination, we then fit a
125 quadratic discriminant function and asked how well the function predicted bloom and decline
126 phase. If a given partition can separate every point perfectly, it will have a 0% misclassification.
127 We created a distribution of all possible proportions of misclassified points. If the visually
128 assessed break points represented an arbitrary pattern, the classification would be no better or
129 worse than a random set of break points. The discriminant function for the visually assessed
130 breakpoints (Figure B5) misclassified only 9% of the points, and performed better than 99% of
131 all other possible splits.
132
133
134
135
136
137

Literature Cited

Berryman, A. A. 1999. Principles of Population Dynamics and Their Application. Garland Science, London.

Berryman, A., and M. Lima. 2006. Deciphering the effects of climate on animal populations: Diagnostic analysis provides new interpretation of soay sheep dynamics. American Naturalist 168:784–795.

Berryman, A., and M. Lima. 2007. Detecting the order of population dynamics from time series: Nonlinearity causes spurious diagnosis. Ecology 88:2121–2123

Berryman, A., and P. Turchin. 2001. Identifying the density-dependent structure underlying ecological time series. Oikos 92:265–270.

Gotelli, N. J. 2001. A primer of ecology. Sinauer Associates, Sunderland, MA.

Gotelli, N. J., and A. M. Ellison. 2004. A Primer Of Ecological Statistics. Sinauer Associates, Sunderland, MA.

Prescott, G. W. 1982. Algae of the Western Great Lakes Area. Page 997. W.C. Brown, Dubuque, IA.

Rogalus, M. K., and M. C. Watzin. 2008. Evaluation of sampling and screening techniques for tiered monitoring of toxic cyanobacteria in lakes. Harmful Algae 7:504–514.

Turchin, P. 2003. Complex population dynamics: a theoretical/empirical synthesis. Monographs in Population Biology. Princeton University Press.

Watzin, M. C., E. B. Miller, A. D. Shambaugh, and M. A. Kreider. 2006. Application of the WHO alert level framework to cyanobacterial monitoring of Lake Champlain, Vermont. Environmental Toxicology 21:278–288.

Table B1

Model	*Hypothesis tested*
$r_t = 0$	Random walk / Null model
$r_t = r_0$	Exponential growth or decline (depending on sign)
$r_t = r_0 - N_{t-1}\mathrm{e}^c$	Density dependence
$r_t = r_0 - X_t$	Effect of nutrient *X* on population growth rate and no density dependence with no time lag, *X={TN, TP, TN:TP, SRP}*
$r_t = r_0 - X_{t-1}$	Effect of nutrient *X* on population growth rate and no density dependence with a 1 week time lag, *X={TN, TP, TN:TP, SRP}*
$r_t = r_0 - N_{t-1}\mathrm{e}^c + \beta_1 X_t$	Effect of nutrient *X* on population growth rate parameter with no time lag, *X={TN, TP, TN:TP, SRP}*
$r_t = r_0 - N_{t-1}\mathrm{e}^c + \beta_1 X_{t-1}$	Effect of nutrient *X* on population growth rate parameter with a 1 week time lag, *X={TN, TP, TN:TP, SRP}*
$r_t = -r_0 - N_{t-1}\mathrm{e}^{(c-\beta_1 X_t)}$	Effect of nutrient *X* on the strength of density dependence with no time lag, *X={TN, TP, TN:TP, SRP}*
$r_t = -r_0 - N_{t-1}\mathrm{e}^{(c-\beta_1 X_{t-1})}$	Effect of nutrient *X* on the strength of density dependence with a 1 week time lag, *X={TN, TP, TN:TP, SRP}*
$r_t = -r_0 - N_{t-1}\mathrm{e}^{(c-\beta_1 S_{t-1})}$	Effect of species *S* on the strength of density dependence with a 1week time lag, *S = { Bacillariophyceae, Chlorophyceae, Cryptophyceae, Myxophyceae, SUM(All taxa)}*

A table of all 34 of the models that we tested for each phase.

Figure B1

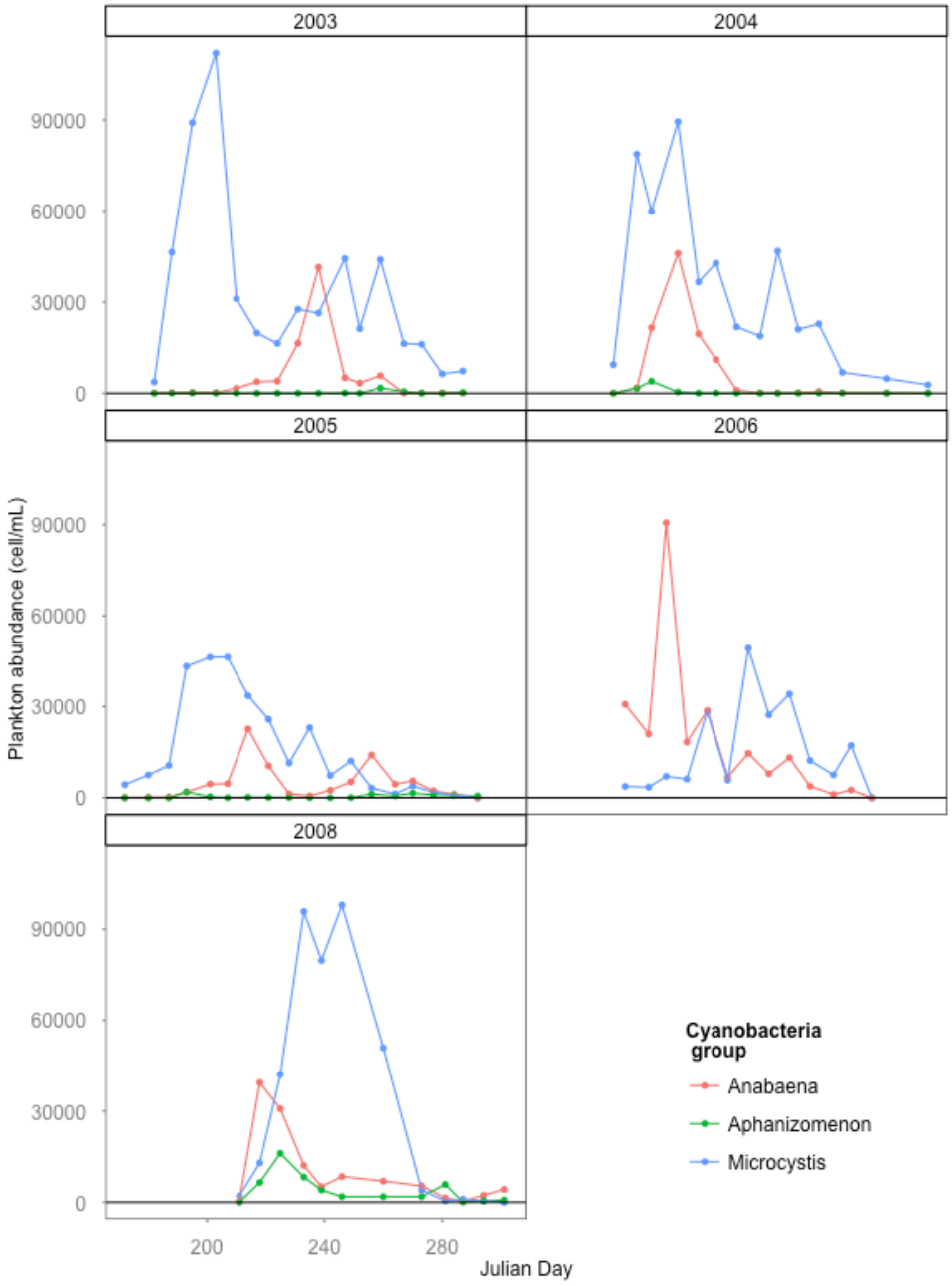

Figure B1. Plots of raw weekly plankton data by taxonomic group for each year. The dominant taxa in each year was chosen for our combined model.

Figure B2

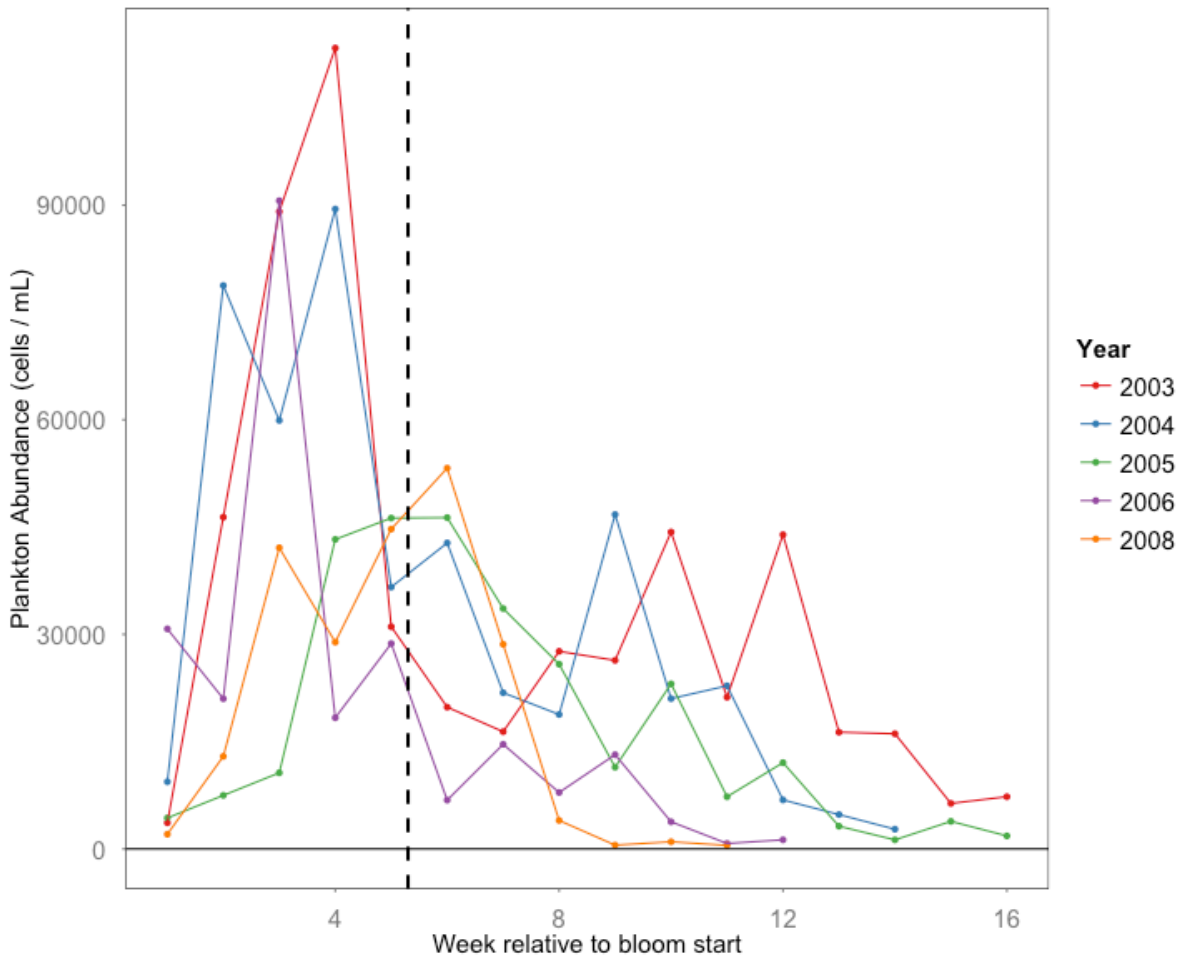

Figure B2. Plot of raw data for all time series with relative week since bloom initiation. The black vertical line represents the division between "bloom" and "non-bloom" phases that were used in the discriminant analysis.

Figure B3

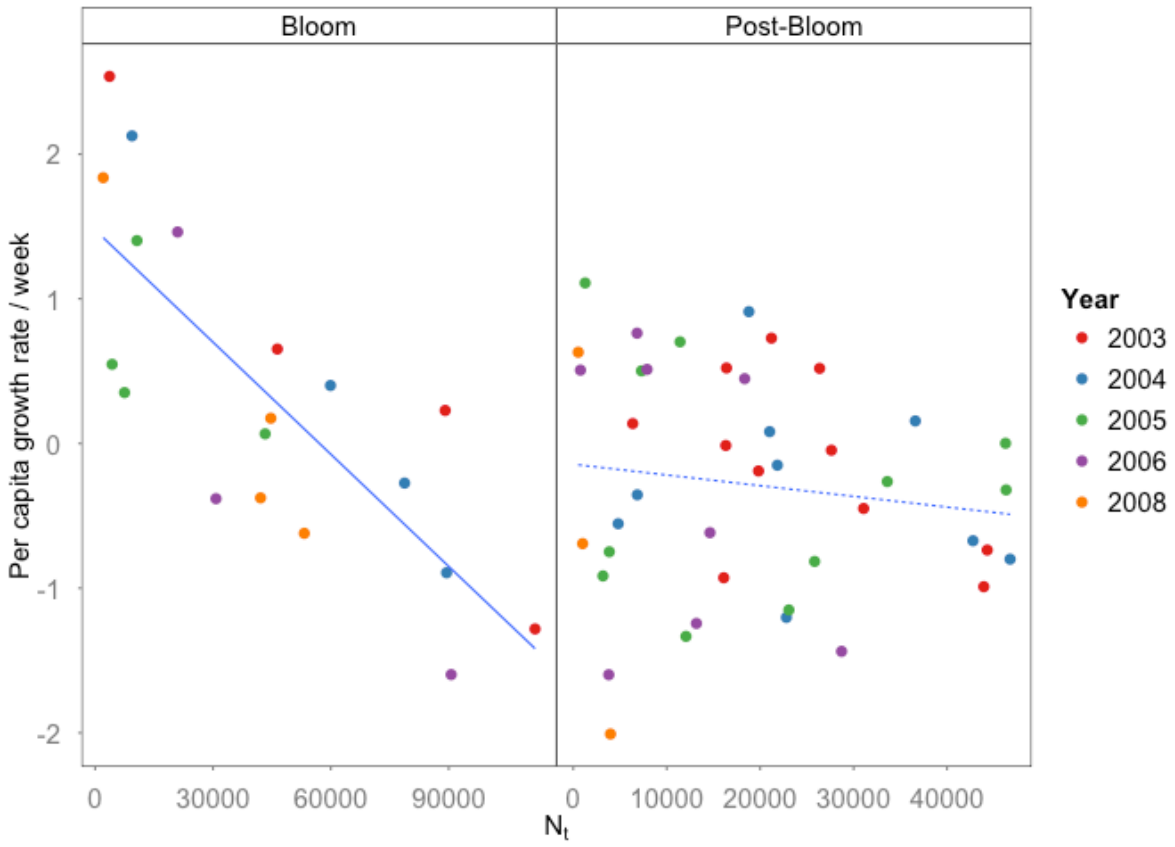

Figure B3. Bloom and post-bloom points plotted out by ear. N_t is the population size at time *t*. Years are dispersed evenly along the line of best fit, demonstrating that aggregation across years increases power only and that the signal we observe is not an artifact of our data aggregation method.

Figure B4

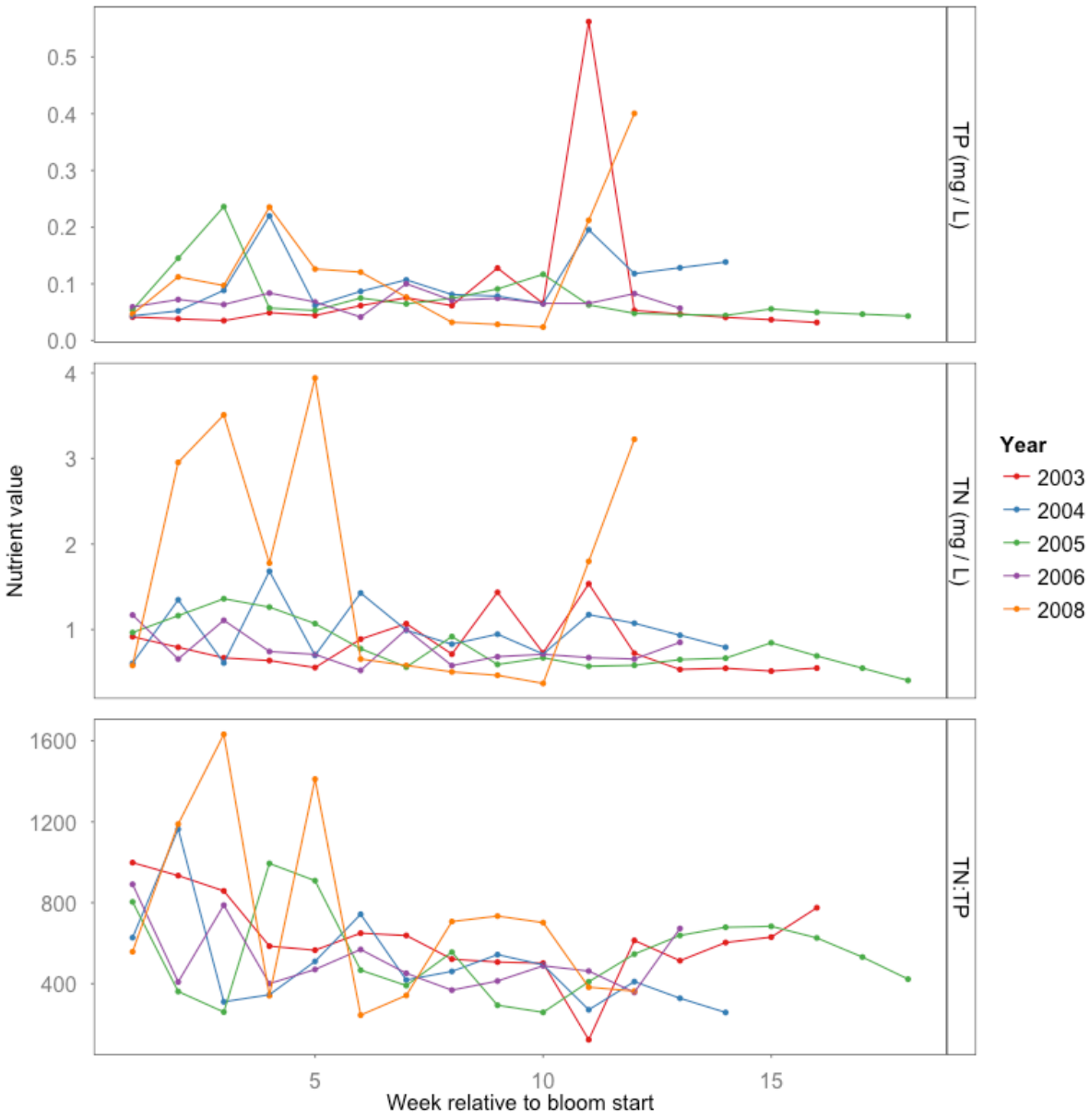
TP (mg / L)
TN (mg / L)
TN:TP
Nutrient value
Week relative to bloom start
Year
2003
2004
2005
2006
2008

Figure B4: Time series plots of Total Phosphorus (TP), Total Nitrogen (TN) and the ratio of the two by year. Plots show that besides a spike of TN in early 2008, there is very little between year variability.

Figure B5

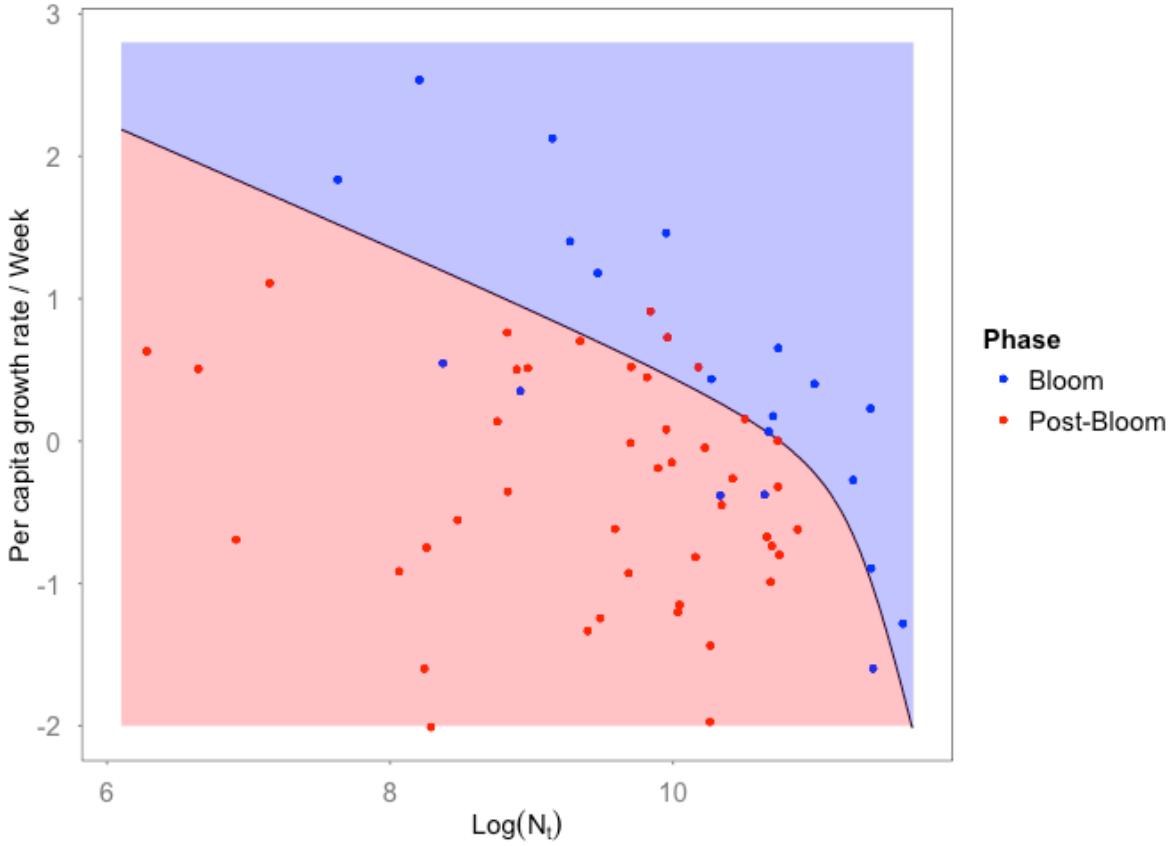

Figure B5. A plot of natural of log population size vs per capita growth rate separated by bloom phase (blue) and decline phase (red) points. The line represents the best-fitting discriminant function for the set of break points that properly classifies 91% of the data into the bloom or the

287 decline phase. Red shaded areas represent the predicted decline phase and blue shaded areas the
288 predicted bloom phase.
289
290