AGU PUBLICATIONS

Water Resources Research

RESEARCH ARTICLE

10.1002/2013WR014372

Key Points:

- A Bayesian hurdle Poisson model predicted cyanobacteria abundance
- Temperature, flushing rate, and water column stability were key factors
- The model forecasted cyanobacteria watch and alert levels probabilistically

Correspondence to:
Y. Cha,
ykcha@umich.edu

Citation:
Cha, Y., S. S. Park, K. Kim, M. Byeon, and C. A. Stow (2014), Probabilistic prediction of cyanobacteria abundance in a Korean reservoir using a Bayesian Poisson model, *Water Resour. Res.*, *50*, 2518–2532, doi:10.1002/2013WR014372.

Received 2 JULY 2013
Accepted 28 FEB 2014
Accepted article online 4 MAR 2014
Published online 20 MAR 2014

Probabilistic prediction of cyanobacteria abundance in a Korean reservoir using a Bayesian Poisson model

YoonKyung Cha[1], Seok Soon Park[2], Kyunghyun Kim[3], Myeongseop Byeon[3], and Craig A. Stow[4]

[1]Cooperative Institute for Limnology and Ecosystems Research, University of Michigan, Ann Arbor, Michigan, USA, [2]Department of Environmental Science and Engineering, College of Engineering, Ewha Womans University, Seoul, South Korea, [3]National Institute of Environmental Research, Incheon, South Korea, [4]NOAA Great Lakes Environmental Research Laboratory, Ann Arbor, Michigan, USA

Abstract There have been increasing reports of harmful algal blooms (HABs) worldwide. However, the factors that influence cyanobacteria dominance and HAB formation can be site-specific and idiosyncratic, making prediction challenging. The drivers of cyanobacteria blooms in Lake Paldang, South Korea, the summer climate of which is strongly affected by the East Asian monsoon, may differ from those in well-studied North American lakes. Using the observational data sampled during the growing season in 2007–2011, a Bayesian hurdle Poisson model was developed to predict cyanobacteria abundance in the lake. The model allowed cyanobacteria absence (zero count) and nonzero cyanobacteria counts to be modeled as functions of different environmental factors. The model predictions demonstrated that the principal factor that determines the success of cyanobacteria was temperature. Combined with high temperature, increased residence time indicated by low outflow rates appeared to increase the probability of cyanobacteria occurrence. A stable water column, represented by low suspended solids, and high temperature were the requirements for high abundance of cyanobacteria. Our model results had management implications; the model can be used to forecast cyanobacteria watch or alert levels probabilistically and develop mitigation strategies of cyanobacteria blooms.

1. Introduction

Reports of harmful algal blooms (HABs) have increased worldwide [*Paerl and Huisman*, 2008]. HABs are a symptom of eutrophication, a state of high primary productivity resulting from nutrient overenrichment. They are often dominated by cyanobacteria [*Carey et al.*, 2012; *Downing et al.*, 2001; *Paerl and Huisman*, 2008] that can produce toxins, odors, and nuisance deposits, hampering ecosystem integrity and the designated uses of lakes [*Carmichael et al.*, 2001; *Codd et al.*, 2005; *Knoll et al.*, 2008].

Models to predict algal occurrence, development, and movement are important tools for HAB management [*Freeman*, 2011]. Predictive models can be useful both for developing strategies to reduce bloom frequency and severity, and for guiding actions to reduce the impact of a bloom that is underway. However, while the general factors that cause eutrophication are well understood [*Carpenter et al.*, 1998], the factors that influence cyanobacteria dominance and HAB formation can be site-specific and idiosyncratic, making prediction challenging [*Glibert et al.*, 2010].

Previous studies have documented climatic, hydrological, and nutritional conditions that contribute to a shift in the algal community composition toward cyanobacteria dominance. These conditions include high water temperature [*Beaulieu et al.*, 2013; *Elliott*, 2010; *Jöhnk et al.*, 2008; *McQueen and Lean*, 1987; *Paerl and Huisman*, 2008], high nutrient levels [*Phillips et al.*, 2005; *Wagner and Adrian*, 2009], increased water column stability [*Jones et al.*, 2005; *Romo et al.*, 2013; *Visser et al.*, 1996; *Wagner and Adrian*, 2009], low nitrogen to phosphorus ratios [*Elliott*, 2010; *Levine and Schindler*, 1999; *Smith*, 1983; *Wood et al.*, 2010], low light availability [*Scheffer et al.*, 1997; *Smith*, 1986], and decreased herbivore grazing [*Paerl*, 1988; *Vanderploeg et al.*, 2009]. However, primary predictors and specific ranges of those predictors associated with the proliferation of cyanobacteria may differ spatially due to location-dependent distribution of species, nutrient baselines, and hydrological and climatic conditions [*Carey et al.*, 2012; *Reynolds*, 2006]; for example, the relative importance between nutrient inputs and temperature in determining cyanobacteria abundance differs among lakes

[*Ahn et al*., 2011; *Brookes and Carey*, 2011; *Kosten et al*., 2012]. Thus, conditions that promote the cyanobacterial growth in one lake may not lead to blooms in other lakes [*Ahn et al*., 2011; *Kim et al*., 2005].

In oceanic and large-lake coastal systems, HAB prediction is often based on near real time, remotely sensed bloom detection [*Wynne et al*., 2013] coupled with hydrodynamic or process-based models to forecast bloom behavior over a period of days to weeks [*Zhang et al*., 2013]. These models are useful for management decisions such as beach or shellfish closures, or initiating enhanced treatment at drinking water intakes. Prediction in Lake Erie has received particular attention because of the severity and magnitude of recent blooms [*Michalak et al*., 2013]. In addition to near term forecasts of HAB movement [*Wynne et al*., 2011], seasonal projections using empirical models based on spring tributary inputs are also under development for Erie [*Stumpf et al*., 2012]. Empirical approaches have been widely used to capture the idiosyncratic nature of bloom formation [*Hamilton et al*., 2009; *Millie et al*., 2012, 2013], and *Graham et al*. [2004] emphasized the importance of empirical relationships for predicting microcystin concentration, a toxin produced by *Microsystis* under some circumstances [*Dyble et al*., 2008].

HAB development in Asian lakes can be distinctive as these lakes are influenced by climatic factors that differ markedly from those of the relatively well-studied North American and Western European lakes. In these lakes, the summer climate is governed by the Asian monsoon, which is characterized by hot, wet conditions, with rainfall during July–August constituting a major portion of the total annual precipitation [*Ahn et al*., 2002; *An and Jones*, 2000b]. *An and Jones* [2000a] reported cyanobacterial dominance in Taechung Reservoir, South Korea, during the summer of 1994 concurrent with the weak monsoon in contrast to infrequent cyanobacteria blooms in the summer of 1993 concurrent with the intense monsoon, indicating that rapid flushing and turbid inflows during the monsoon suppressed the growth of cyanobacteria.

Our study site, Lake Paldang, located close to the metropolitan area of the capital Seoul of South Korea, was formed through the construction of Paldang Dam in 1973. As a confluence of three rivers, the North Han, South Han, and Kyung-an, Lake Paldang is a typical river-run lake characterized by shallow depth (mean depth = 6.4 m) and a short hydraulic residence time ranging 2.6–9.0 days (mean residence time = 6.4 days) [*Na and Park*, 2005]. Lake Paldang supplies domestic and drinking water to 25 million residents of the capital area, and is used for flood control and electric power generation. The lake, which had been oligotrophic in the 1970s, became increasingly eutrophic with increases in phosphorus and nitrogen levels along with increased turbidity, and, since the 1990s, frequent dominance by cyanobacteria [*Na and Park*, 2006; *Park et al*., 2005]. The progression toward eutrophication was coincident with relaxed land-use regulations in the Paldang watershed, which was followed by urban and agricultural expansion, and increases in population density and nutrient loads [*Jung*, 2009]. A 4 year (1997–2000) study conducted between May and November found that the dominant summer (June–September) cyanobacterial genera included *Microcystis* and *Anabaena*, prone to colony and bloom formation [*Kim et al*., 2005]. Although toxin production by cyanobacteria was not substantial (microcystin level $< 1\ \mu g\ L^{-1}$), odor-generating substances often impaired drinking water quality [*Kim et al*., 2009]. In July 2008, unpleasant smelling tap water led to complaints from residents of the capital area and serious public concern. *Kim et al*. [2009] suggested that the odor originated from 2-Methylisoborneol (MIB) and especially high levels of geosmin, which were strongly associated with high levels of *Anabaena*.

For South Korean lakes supplying drinking water, such as Lake Paldang, a cyanobacteria watch is issued when cyanobacteria cell counts and chlorophyll *a* (Chla) jointly exceed 500 cells mL^{-1} and 15 mg m^{-3}, respectively. Further, an alert is issued when cyanobacteria cell counts and Chla jointly exceed 5000 cells mL^{-1} and 25 mg m^{-3}, respectively. Once a watch or alert is issued, then additional treatment processes are required at water treatment plants until the watch or alert is cleared. Additionally, when an alert is issued the movement of water intakes below where algae inhabit, and the analysis of cyanotoxin in the treated water are required.

Our objectives were to examine the environmental drivers that promote cyanobacterial growth, identify the conditions that affect the abundance of cyanobacteria relative to other phytoplankton, and develop a model for cyanobacteria prediction. To do this, we used a Bayesian hurdle Poisson model [*Mullahy*, 1986; *Neelon et al*., 2010]. Although the hurdle Poisson model does not distinguish between "true zeros" (real absence of organisms) and "false zeros" (failure to detect organisms that are present at low abundance) [*Martin et al*., 2005], the practical utility in our application is that it allows cyanobacteria presence/absence

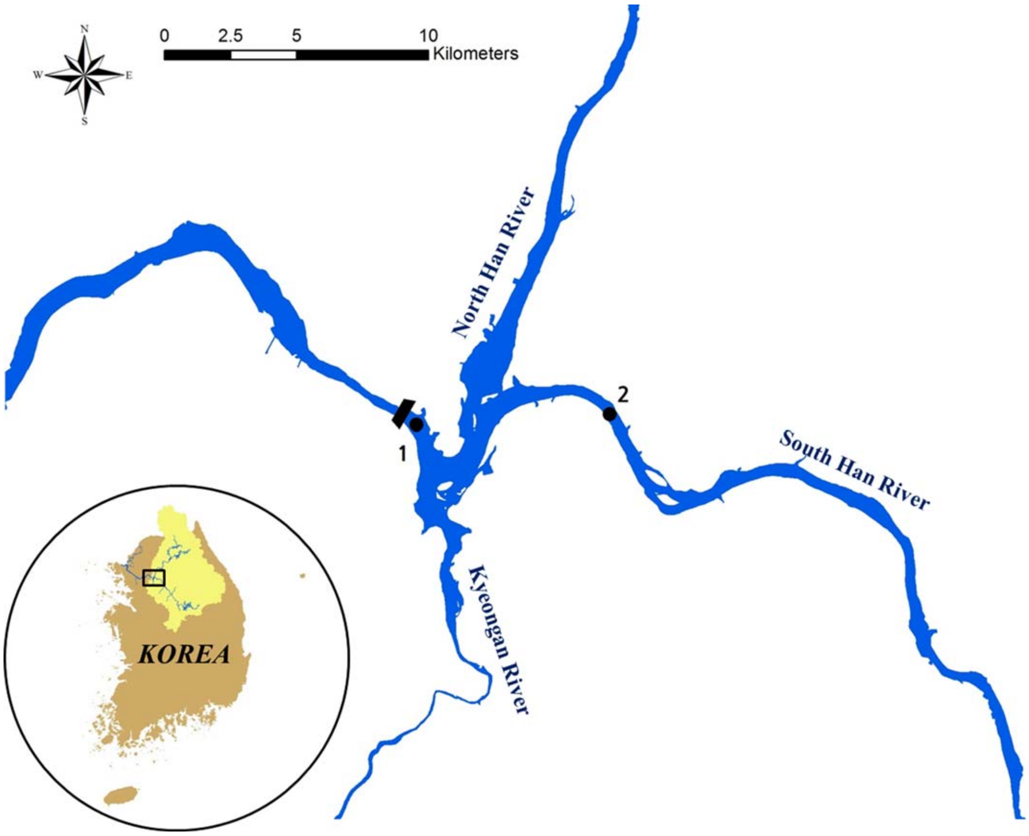

Figure 1. Map of Lake Paldang with the Paldang Dam (trapezoid). Sampling sites denoted as 1 and 2.

and cyanobacteria counts to be modeled as functions of different environmental factors. This capacity can be useful to differentiate conditions that initiate cyanobacterial proliferation from those that promote excessive concentrations. The hurdle Poisson model has been used in other fields of study [*Hu et al*., 2011; *Neelon et al*., 2013]; however, the application of the model integrated into a Bayesian framework to resolve water quality issues is novel. This model is appropriate to represent the distribution of cyanobacteria cells, characterized by discrete positive values with a distinct peak at zero. The model provides a scientific basis for lake management by predicting the level of cyanobacteria, and the probability of exceeding cyanobacteria criteria for gradients of combined predictor variables.

2. Methods

2.1 Data Description

For phytoplankton and water quality characteristics, we sampled at two locations (Figure 1) from April to November from 2007 to 2011, and in 2011 the sampling period was extended to December. During April–November in 2007–2011, sampling frequency was weekly at the station 1 and monthly at the station 2. In December 2011, sampling was daily in both stations. All samples were taken at 0.5 m below the water surface.

In our analysis, we included the water quality characteristics that are expected to be related to cyanobacteria blooms: water temperature, conductivity, secchi depth, suspended solids (SS), dissolved oxygen (DO), Chla, phosphorus species—total phosphorus (TP), total dissolved phosphorus (TDP), and phosphate (PO_4-P), and nitrogen species—total nitrogen (TN), total dissolved nitrogen (TDN), nitrate (NO_3-N), and ammonium (NH_4-N). Water temperature, conductivity, and DO were measured by YSI 6600 directly on the spot. SS was measured by filtering 1 L of water sample using a glass fiber filter (GF/F, 47 mm) and drying the filter at 105°C in an electric oven. For the Chla analysis, the same filter was used to filter 1 L of water sample. Then

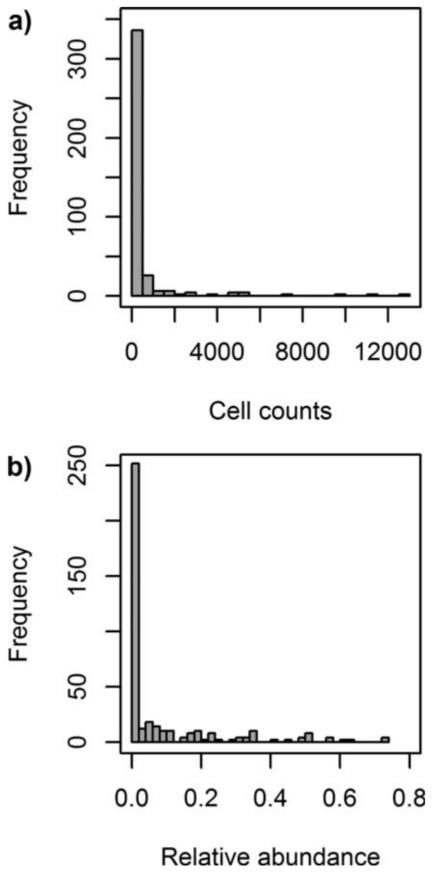

Figure 2. Distribution of cyanobacteria (a) cell counts (cells mL^{-1}) and (b) relative abundance based on individual observations sampled 2007–2011 in Lake Paldang.

the filter was soaked in 10 mL of 90% acetone solution to elute Chla. A spectrophotometer was used to measure the Chla concentration of the solution. TP, TDP, and PO_4-P were analyzed by the ascorbic acid method. TN was measured by the absorption photometry method and TDN, NO_3-N, and NH_4-N were analyzed by the Indo-Phenol method. The phytoplankton data consisted of cell counts (cells mL^{-1}) of diatom, green algae, cyanobacteria, and total phytoplankton. For the phytoplankton data, 1 mL sample taken from a well-mixed 2 L water sample was put into a Sedgwick-Rafter chamber and was analyzed by a phase-contrast microscope (Nikon Eclipse 80i) at 400–600× resolutions [*Komárek*, 1991].

Daily precipitation, inflow, and outflow data were obtained from online WAMIS of the Korean Ministry of Land, Transport, and Maritime Affairs (http://www.wamis.go.kr). In addition to the daily precipitation, inflow, and outflow, the average values of these characteristics over 6 days before the sampling date and including the sampling date of phytoplankton and water quality parameters were included as predictor candidates to account for hydraulic residence time. Mean and maximum daily wind speed, and daily irradiance were obtained from the Korea Meteorological Administration (http://www.kma.go.kr).

To compare the annual and seasonal patterns for a long-term period (1998–2006) and for the period of study (2007–2011), we obtained the data for water temperature, TP, TN, and Chla measured at the stations 1 and 2 with weekly to monthly frequency for the years 1998–2006 from the Korea National Institute of Environmental Research. Precipitation data for the corresponding period was obtained from the WAMIS.

2.2. Bayesian Poisson Hurdle Regression

Discrete data, such as cyanobacterial cell counts, are commonly modeled as either Poisson or negative binomial distributions. The negative binomial distribution can accommodate “overdispersion,” a situation that occurs when the variance of the data exceeds that which can be estimated using a Poisson model [*White and Bennetts*, 1996]. However, count data often have more zeros than either a Poisson or negative binomial model can represent (Figure 2). To address excessive zeros, and accommodate overdispersion, mixture models consider counts to arise from a combination of distributions. Zero-inflated mixture models represent zeros as having two distinct sources, an underlying Poisson or negative binomial process and a second process that generates extra zeros [*Lambert*, 1992]. In contrast, hurdle models consider all the zeros to arise from one process while the nonzero discrete values arise from a separate Poisson or negative binomial process that is truncated at zero [*Heilbron*, 1994; *Mullahy*, 1986]. The practical implication is that hurdle models can be used to reflect an underlying presence/absence state that is influenced by environmental drivers that may be distinct from the factors influencing the nonzero counts. In our case, the use of the hurdle model allowed us to explicitly characterize the environmental conditions that contribute to shift a lake from cyanobacteria-free to cyanobacteria-present states, while at the same time delineating the conditions that promote the proliferation of cyanobacteria once an outbreak occurs. We present the hurdle Poisson model herein; however, despite the conceptual differences, replacement with the negative binomial distribution or the zero-inflated modeling did not make substantial changes in the results. Further, a Bayesian framework was used to quantify the uncertainty associated with model parameters and predictions [*Ghosh et al*., 2006].

The hurdle Poisson model can be expressed as [*Neelon et al*., 2010]:

$$\begin{aligned} &\Pr\ (Y_i=0)=1-p_i \ \text{ for } 0\le p_i\le 1 \\ &\Pr\ (Y_i=k)=p_i\mu^k e^{-\mu}/[k!(1\ -e^{-\mu})] \ \text{ for } k=1,\ \ldots,\ \infty,\ 0\ <\ \mu\ <\ \infty, \end{aligned} \tag{1}$$

where Y_i is the independent response (cyanobacteria cell count) for the sample $i = 1, \ldots, n$, k is the positive value measured, p is the probability of cyanobacteria presence in a sample, and μ is the mean for a truncated Poisson. The parameters, $\boldsymbol{p}$ and μ, are further modeled with predictor matrices $\mathbf{X} = (X_1, \ldots, X_m)$ and $\mathbf{W} = (W_1, \ldots, W_l)$,

$$\text{logit}\,(p_i)=\log\ [p_i/(1-p_i)]=\eta_0+\eta_1X_{i,1}+\ldots+\eta_mX_{i,m} \tag{2}$$

$$\log\,(\mu_i)=\beta_0+\beta_1W_{i,1}+\ldots+\beta_lW_{i,l}+\varepsilon_i,\varepsilon_i\sim\text{Normal}\,(0,\sigma^2) \tag{3}$$

where $\boldsymbol{\eta} = (\eta_0, \eta_1, \ldots, \eta_m)$ and $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_l)$ are vectors of regression parameters, and ε is an overdispersion parameter with the mean of zero and standard deviation of σ. Note that the predictors included in the logistic regression model may or may not be the same as the predictors included in the Poisson model.

Based on equation (3), if we put the logarithm of a baseline (in this case total phytoplankton cell counts) of counts as a predictor with its parameter fixed to one, then the counts can be interpreted as relative abundance (in proportion to total phytoplankton cell counts) of cyanobacteria. Abundance and relative abundance of cyanobacteria were separately modeled in our study.

We used noninformative prior distributions for all model parameters and posterior distributions were obtained using Markov chain Monte Carlo (MCMC) simulation procedures in the software program WinBUGS [*Lunn et al*., 2000]. To select predictor variables among a variety of environmental factors, first preliminary analyses, such as stepwise regression and regression trees, were performed with R packages [*R Development Core Team*, 2012]. During this process, variables that showed little relevance to the response (cyanobacteria presence/absence or abundance) were screened out. Among the remaining variables, inclusion of predictor variables and interaction between the variables was decided based on the deviation information criterion (DIC). The DIC is a Bayesian alternative of Akaike information criterion used for model comparison and selection [*Plummer*, 2008].

3. Results

Since the late 1990s, TP, TN, Chla, water temperature, and precipitation have varied from year to year without obvious trends (Figure 3). Based on ranges suggested by *Reckhow and Chapra* [1983], TP, TN, and Chla average concentrations (48 µg L^{-1}, 2.2 mg L^{-1}, and18 µg L^{-1}, respectively) indicate that Lake Paldang was eutrophic throughout the years 1998–2011 (Figure 3). TP levels were substantially higher during the summer monsoon (July–August) than annual average levels (Figure 3a), while the monsoon TN levels were comparable to annual TN levels in most of the years 1998–2011 (Figure 3b). Variability of Chla in the monsoon months exceeded the annual variability (Figure 3c), and, as expected, the monsoon water temperature ($\bar{x}=22.5°\text{C}$) was substantially higher than annual water temperature ($\bar{x}\ =13.6°\text{C}$) (Figure 3d). Precipitation during the monsoon largely determined the annual patterns, comprising 40–72% of total annual precipitation (Figure 3e).

In terms of cell counts diatoms were the most abundant phytoplankton group across the growing season (April–November) from 2007–2011 (Figure 4). Cyanobacteria was the least abundant group even in warm periods of the year, reaching its peak with the monthly average of 550–2600 cells mL^{-1} (0.07–0.27 as relative abundance) between June and September (Figure 4). However, instantaneous cyanobacteria cell counts reached ~13,000 cells mL^{-1} on 14 July 2010 and made up 74% of total phytoplankton cell counts on 14 July 2008 (Figure 2).

High water temperature and poor flushing were the conditions most strongly associated with the presence of cyanobacteria (Figures 5a and 5b) and are reflected in the model as:

$$\Pr\,(\text{cyanobacteria presence}) = \text{logit}^{-1}[-3.60\ (\pm 1 \text{ standard deviation} = 0.71) + 0.20\ (\pm\ 0.04) \cdot \text{WatTemp} - 0.14\ (\pm\ 0.05) \cdot 10^2 \cdot \text{Outflow}]$$

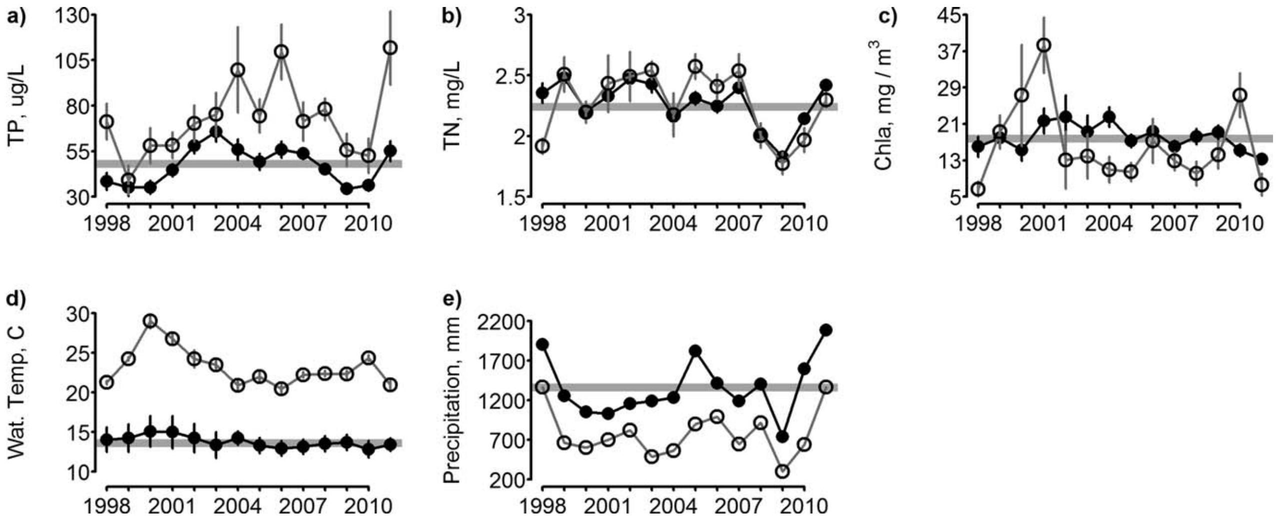

Figure 3. Annual and summer patterns of (a) mean TP, (b) mean TN, (c) mean Chla, (d) mean water temperature, and (e) total precipitation for the years 1998–2011 in Lake Paldang. Filled circles are mean values (or total value for precipitation) throughout the year, and empty circles are mean values (or total value for precipitation) during the summer monsoon period (July–August). Vertical bars indicate ±1 standard error of the mean estimate. Horizontal lines indicate overall mean for the years 1998–2011.

Our model indicates that at the overall mean weekly outflow (675 m^3 s^{-1}) cyanobacteria are more likely to be present than absent when the water temperature is above ∼23°C (Figure 5a). At a mean weekly outflow of 300 m^3 s^{-1}, an increase in water temperature from 15 (25th percentile) to 23°C (75th percentile) increased the probability of cyanobacteria occurrence (mean of posterior *p*) from 27 to 63%, while the same water temperature increase at 1000 m^3 s^{-1} of mean weekly outflow increased the probability of cyanobacteria occurrence from 12 to 40% (Figure 5c). With an increase in mean weekly outflow from 300 (25th percentile) to 1000 m^3 s^{-1} (75th percentile) at a temperature of 23°C, the probability of cyanobacteria

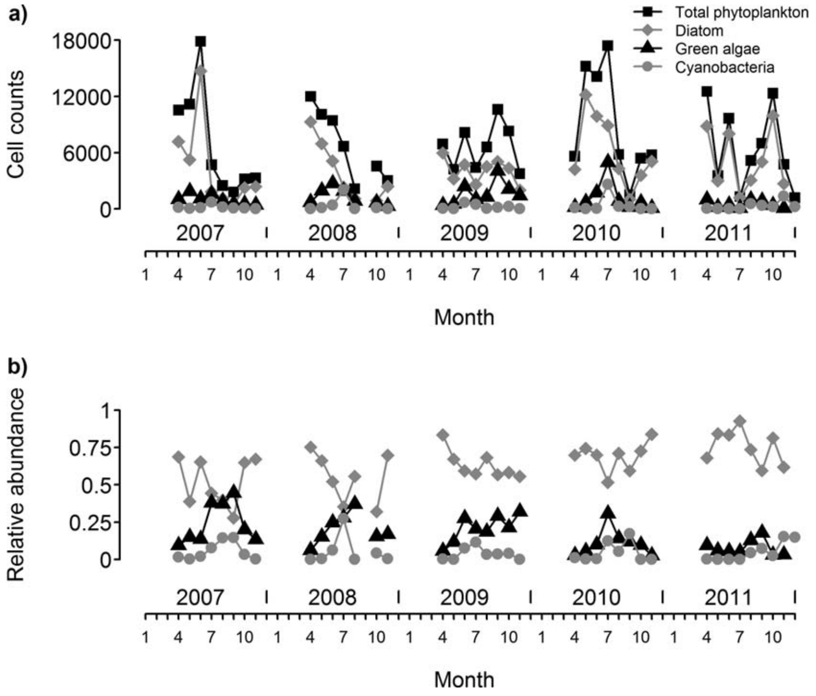

Figure 4. Monthly patterns of phytoplankton group composition with regard to (a) cell counts (cells mL^{-1}) and (b) relative abundance in Paldang Lake 2007–2011. Black square, gray diamond, black triangle, and gray circle indicate monthly means of total phytoplankton, diatom, green algae, and cyanobacteria, respectively.

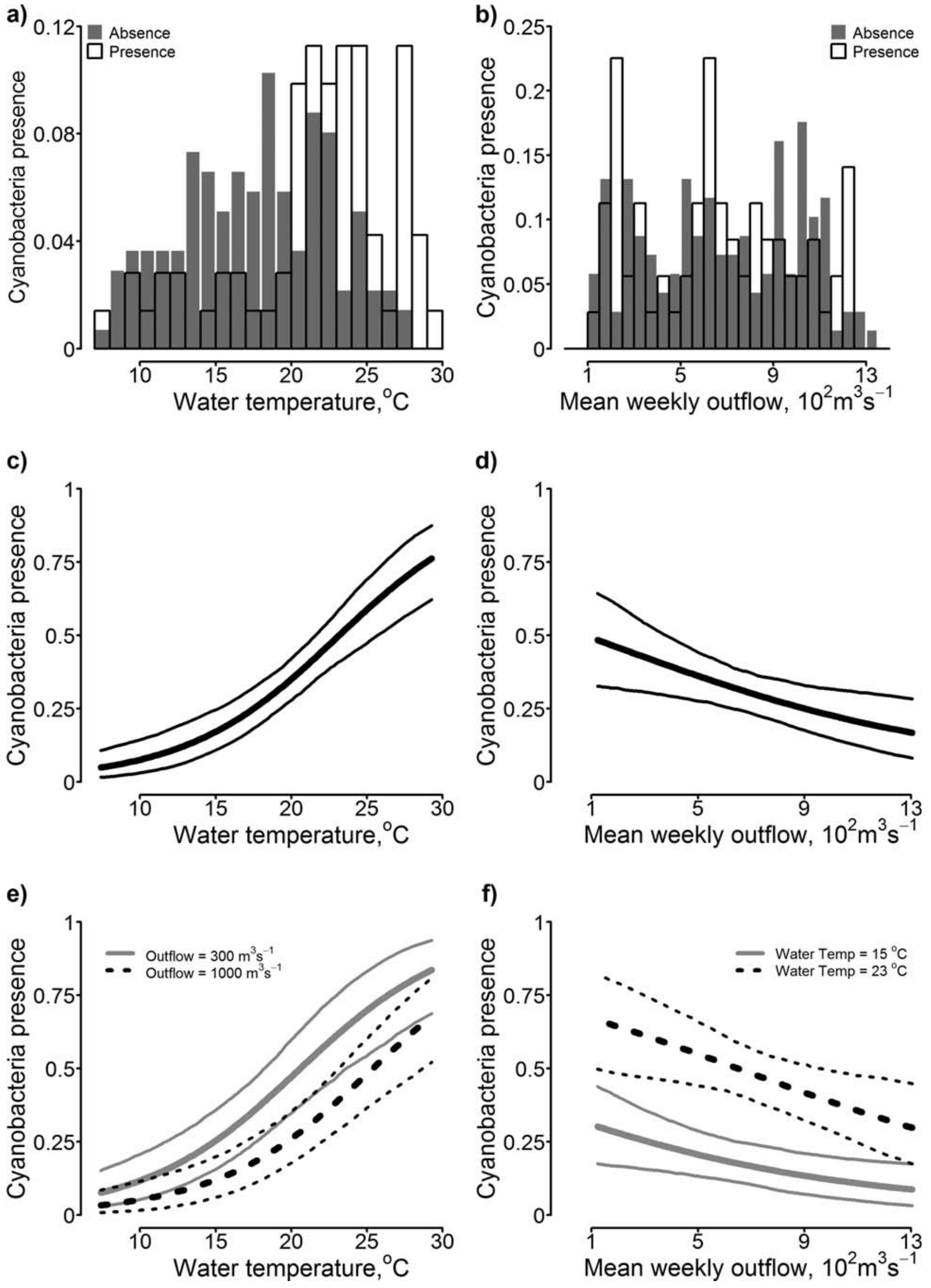

Figure 5. Probability of cyanobacteria presence as a function of (a) water temperature with outflow set at its mean (675 $m^3 s^{-1}$), (b) mean weekly outflow with water temperature set at its mean (19°C), (c) water temperature (gray lines with 25th percentile of outflow and black, dashed lines with 75th percentile of outflow), and (d) mean weekly outflow (gray lines with 25th percentile of water temperature and black, dashed lines with 75th percentile of water temperature). Circles indicate observed samples (1 when cyanobacteria = present; 0 = absent). Thick lines indicate mean predicted values, while thin lines indicate 95% predictive intervals.

occurrence would decrease from 63 to 40%, while at low water temperature the probability of cyanobacteria presence is consistently low regardless of outflow; the same outflow increase at 15°C would decrease the probability of cyanobacteria occurrence from 26 to 12% (Figure 5d). Therefore, water temperature was a more important factor than outflow in determining the probability of cyanobacteria presence.

Environmental factors that affect the Poisson mean (μ) of cyanobacteria cell counts were water temperature and SS (Figure 6) as indicated by the following model:

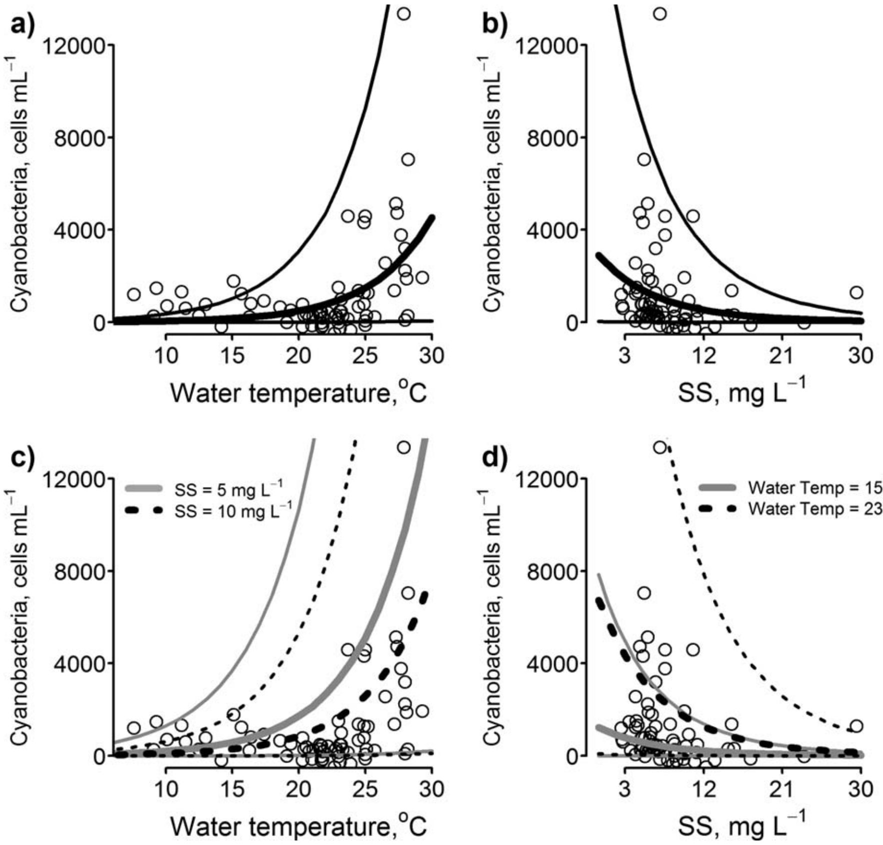

Figure 6. Changes in cyanobacteria cell counts in response to (a) water temperature with SS set at its mean (14 mg L^{-1}), (b) SS with water temperature set at its mean (19°C), (c) water temperature (gray lines with 25th percentile of SS and black, dashed lines with 75th percentile of SS), and (d) SS (gray lines with 25th percentile of water temperature and black, dashed lines with 75th percentile of water temperature). Circles indicate observed cell counts. Thick lines indicate mean predicted values, while thin lines indicate 95% predictive intervals.

$$\text{Log (cell counts)} = 2.94(\pm\ 0.55) + 0.20(\pm\ 0.03) \cdot \text{WatTemp} - 0.15(\pm\ 0.04) \cdot \text{SS} + \varepsilon,\ \varepsilon \sim \text{Normal}\ (0, 1.50^2)$$

These results indicate that predicted cyanobacteria cell counts increase as water temperature increases and SS decreases (Figures 6a and 6b). Water temperature effects on cyanobacteria growth differ depending on the SS level. At a low SS concentration (5 mg L^{-1}), a water temperature increase from 15 to 23°C increased predicted cyanobacteria cell counts (mean of posterior μ) from ~900 to ~4400 cells mL^{-1}, while at a higher SS concentration (10 mg L^{-1}), the same temperature increase raised the cell counts from ~400 to ~2000 cells mL^{-1} (Figure 6c). Water temperature played a critical role in determining the magnitude of cyanobacteria occurrence; at 23°C a SS increase from 5 (25th percentile) to 10 mg L^{-1} (75th percentile) corresponded to a decrease in cyanobacteria cell counts from ~4400 to ~2000 cells mL^{-1} (Figure 6d). In contrast, at low temperatures cyanobacteria levels were consistently low regardless of SS levels; at 15°C, an increase in SS from 5 to 10mg L^{-1} corresponded to cyanobacteria cell count decrease from ~850 to ~410 cells mL^{-1} (Figure 6d).

Interestingly, water temperature was not a factor that affected the relative abundance (in proportion to total phytoplankton cell counts) of cyanobacteria (Figure 7a). Instead, SS and the TDN:TDP mass ratio were the predictors that showed strong relationships with relative abundance (Figures 7b and 7c) as reflected in the following model:

$$\text{Log (relative abundance)} = -1.04(\pm\ 0.32) - 0.07(\pm\ 0.03) \cdot \text{SS} - 0.01(\pm\ 0.00) \cdot \text{TDN:TDP} + \varepsilon,\ \varepsilon \sim \text{Normal}\ (0, 1.20^2)$$

At the mean TDN:TDP ratio (125), a SS increase from 5 to 10 mg L^{-1} was associated with a cyanobacteria relative abundance decrease from 0.34 to 0.24 (Figure 7b), while at the mean SS concentration a TDN:TDP

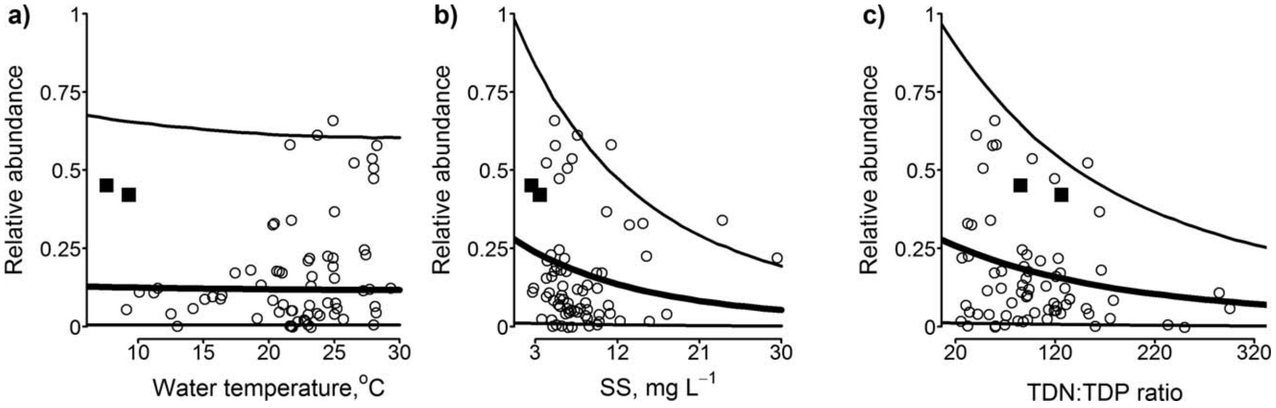

Figure 7. Changes in relative abundance of cyanobacteria in response to (a) water temperature, (b) SS, and (c) TDN:TDP mass ratio. Circles indicate observed samples. Solid squares indicate exceptional observations, high relative abundance associated with low water temperature, low SS, and low N:P ratio. Thick lines indicate mean predicted values, while thin lines indicate 95% predictive intervals.

increase from 65 (25th percentile) to 160 (75th percentile) was associated with a cyanobacteria relative abundance decrease from 0.28 to 0.14 (Figure 7c).

For South Korean lakes supplying drinking water, such as Lake Paldang, a cyanobacteria watch is issued when cyanobacteria cell counts and Chla jointly exceed 500 cells mL^{-1} and 15 mg m^{-3}, respectively. Further, an alert is issued when cyanobacteria cell counts and Chla jointly exceed 5000 cells mL^{-1} and 25 mg m^{-3}, respectively. However, our data indicate only a weak correlation between cyanobacteria counts and chlorophyll *a* (Figure 8a); the highest cyanobacteria levels occurred at intermediate Chla concentrations (~10.8–34.4 mg m^{-3}). Even when the cyanobacteria cell count was zero, Chla ranged from 0.6 to 63.8 mg m^{-3} with the highest observed Chla concentration occurring at a zero cyanobacteria count. Generally, Chla was much more strongly associated with diatoms (Figure 8b) and green algae (Figure 8c), than with cyanobacteria suggesting that high Chla concentrations may result from algal groups that pose a minimal drinking water concern. Consequently, the highest observed cyanobacteria counts were more likely to trigger a watch than an alert, and require only moderate response measures.

The weak correlation between cyanobacteria and Chla makes it difficult to characterize the antecedent conditions that would result in a watch or alert based on the joint Chla:cyanobacteria relationship. However, because cyanobacteria are of primary concern we can use our model to calculate the conditions that would result in high probabilities of exceeding the 500 cells mL^{-1} and 5000 cells mL^{-1} thresholds that, respectively, initiate a watch or alert. These probabilities are highest at high temperature, low SS, and low weekly outflow (Figure 9). With the weekly outflow at its mean values, the probability of exceeding the watch and

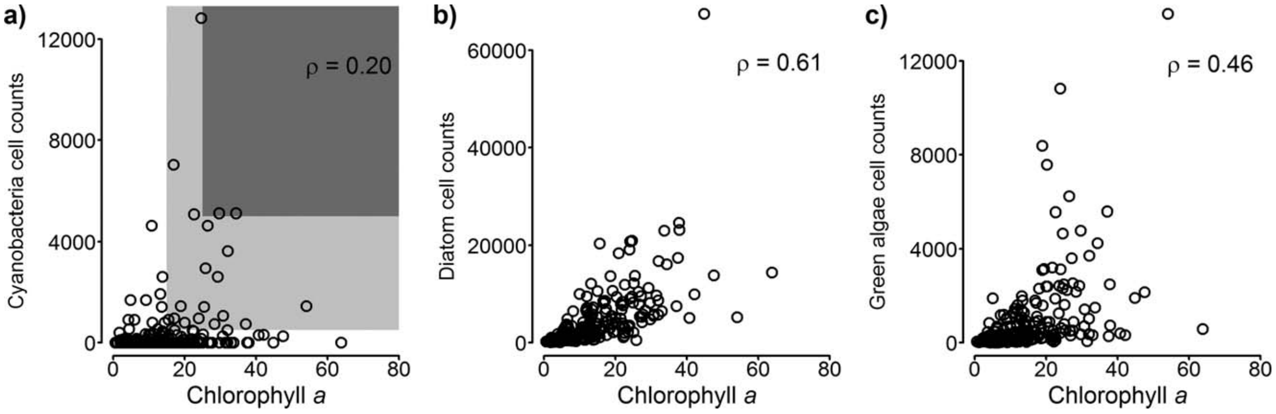

Figure 8. Relationship of chlorophyll *a* (mg m^{-3}) with (a) cyanobacteria cell counts (cells mL^{-1}), (b) diatom cell counts (cells mL^{-1}), and (c) green algae cell counts (cells mL^{-1}). ρ indicates Pearson's correlation coefficient. Light gray shading and dark gray shading represent the regions where the cell counts and chlorophyll *a* jointly exceed 500 cells mL^{-1} and 15 mg m^{-3}, and 5000 cells mL^{-1} and 25 mg m^{-3}, respectively.

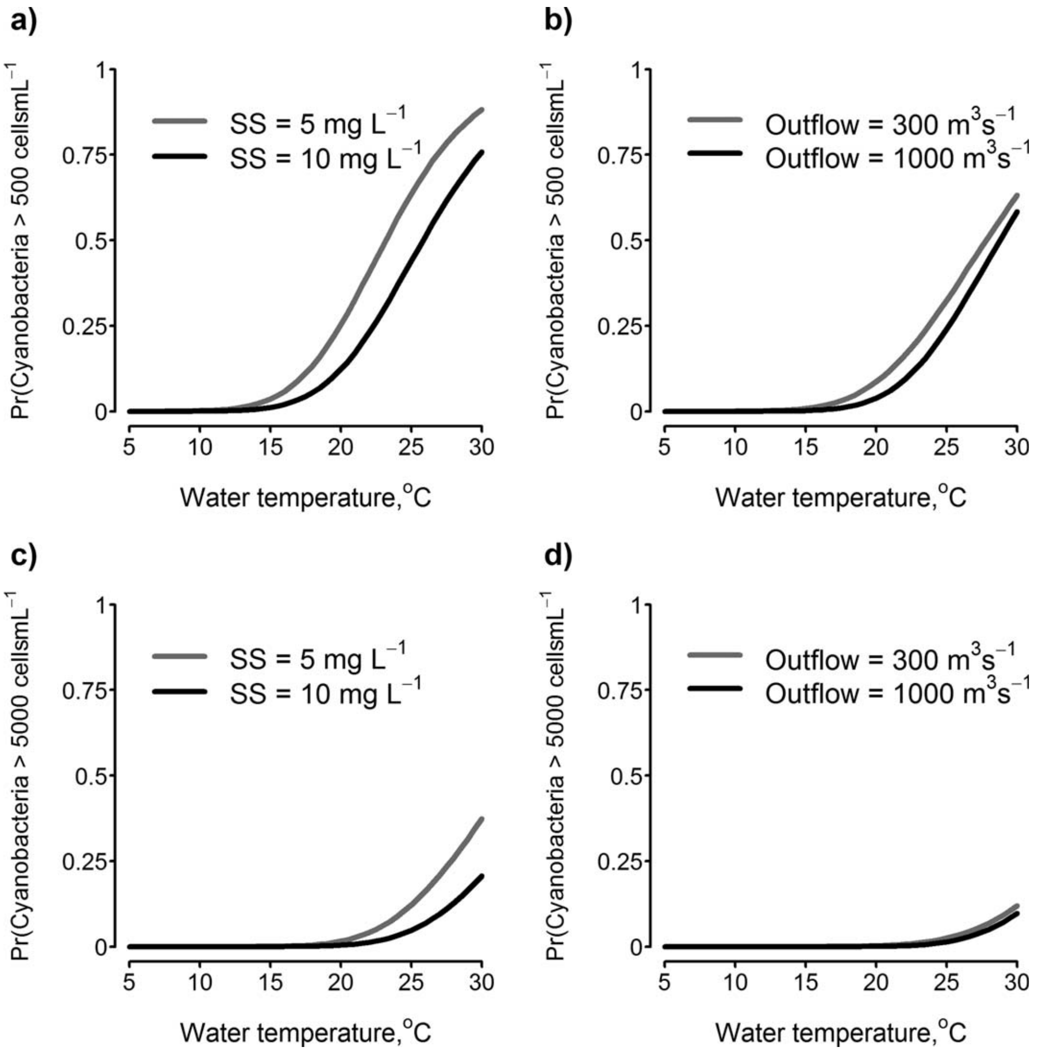

Figure 9. Probability of exceeding cyanobacteria watch level (500 cells mL^{-1}) and alert level (5000 cells mL^{-1}). The probability of exceeding the watch level given a gradient of (a) water temperature with 25th percentile of SS (gray line) and with 75th percentile of SS (black line), and (b) water temperature with 25th percentile of outflow (gray line) and with 75th percentile of outflow (black line). The probability of exceeding the alert level given a gradient of (c) water temperature with 25th percentile of SS (gray line) and with 75th percentile of SS (black line), and (d) water temperature with 25th percentile of outflow (gray line) and with 75th percentile of outflow (black line).

alert levels at high SS conditions (10 mg/L) was higher than 0.1 when water temperature $>$20 and $>$28°C, respectively, and the probability of exceeding the watch level was higher than 0.5 when water temperature $>$26°C (Figures 9a and 9c). Under low SS conditions, the exceedance probabilities were higher at the same water temperature: at 5 mg L^{-1} SS combined with the mean level of weekly outflow, the probability of exceeding the watch and alert levels were higher than 0.1 when water temperature $>$17 and $>$24°C, respectively, and the probability of exceeding the watch level was higher than 0.5 when water temperature

Table 1. Summary of Variables Included in Models

Variable	Unit	Mean (±1 SD)	Range
Cyanobacteria cell count	cells mL^{-1}	514.44 (± 1648.98)	0–12840
Relative cyanobacteria abundance	-	0.08 (± 0.16)	0-0.74
Water temperature	°C	19.23 (± 5.50)	5.80–30.08
Mean weekly outflow	m^3 s^{-1}	592.13 (± 391.28)	1–1351
Suspended solids	mg L^{-1}	11.28 (± 41.60)	1–684
Total dissolved phosphorus	μg L^{-1}	21.04 (± 17.04)	3–101
Total dissolved nitrogen	mg L^{-1}	2.06 (± 0.40)	1.22–3.53

>23°C (Figures 9a and 9c). Weekly outflow had less impact than SS on the probabilities of the exceedance for a given level of water temperature: at 23°C, the probability of exceeding the watch level was 0.21 when outflow = 300 $m^3 s^{-1}$ and 0.13 when outflow = 1000 $m^3 s^{-1}$ (Figure 9b), and the probability of exceeding the alert level was much lower than 0.1 at both 300 and 1000 $m^3 s^{-1}$ of outflow (Figure 9d).

4. Discussion

Our model predicts cyanobacteria levels based on water temperature, outflow, and SS, three characteristics that are straightforward and relatively quick to measure (Table 1). Therefore, the model should be a useful tool for decision makers to anticipate imminent drinking water problems. While the model is empirically based and thus, specific to Lake Paldang, the factors incorporated are consistent with those reported in other systems as important to cyanobacteria dynamics.

Cyanobacteria absence/presence was influenced by outflow, an indicator of hydraulic residence time or flushing rate, and temperature (Figure 5). The temperature association may be related to the observation by *Visser et al*. [1995] that *Microcystis* colonies sank to the bottom due to a buoyancy loss caused by carbohydrate accumulation as temperature decreased from summer to autumn. Previous studies have also reported that vertical mixing associated with turbulent conditions minimized light exposure, inhibiting cyanobacteria growth [*Paerl*, 1988]. Additionally, at high outflows, the water residence time may be too low for cyanobacteria to proliferate [*Mitrovic et al*., 2006; *Paerl and Huisman*, 2008].

The cyanobacteria cell count in Lake Paldang was also affected by water temperature as well as SS levels (Figure 6). *Jöhnk et al*. [2008] found that the specific growth rate for *Microcystis* reached maximum at ~28°C, consistent with our results exhibiting accelerated increases in cyanobacteria cell counts as water temperature exceeds ~23°C (Figure 6). SS levels are closely associated with outflow and may largely represent light availability; the algal component of SS is minor. Note that the observed SS ranged from 2 to 684 mg L^{-1} but cyanobacteria never appeared when SS exceeded 30 mg L^{-1} (Figure 6b). Generally, our results are similar to those in another Korean lake, Taechung Reservoir, where warm temperature or a weak summer monsoon was required for cyanobacteria dominance or high cyanobacteria density [*Ahn et al*., 2002, 2011; *An and Jones*, 2000a].

Low SS was associated not only with high cyanobacteria abundance but also with a high relative abundance of cyanobacteria (Figure 7b). Previous studies have found that water column stability allows vertically motile (or buoyant) cyanobacteria to outcompete other nonmotile (or nonbuoyant) phytoplankton groups [*Jones et al*., 2005; *Serizawa et al*., 2008; *Visser et al*., 1996; *Wagner and Adrian*, 2009]. In stable, stratified waters cyanobacteria colonies are able to grow larger, and fast floatation by *Microcystis* and *Anabaena* allows them to migrate from the well-lit surface to the nutrient enriched bottom layer, yielding competitive advantages over other phytoplankton [*Carey et al*., 2012; O'*Brien et al*., 2004]. In contrast, in poorly mixed waters diatoms would be subject to sedimentation due to their heavier, larger cells compared to cyanobacteria [*Reynolds*, 2006].

The hypothesis that low N:P ratios, indicative of nitrogen limitation, promote nitrogen-fixing genera of cyanobacterias been a subject of long debate. Many studies have supported this idea [*Levine and Schindler*, 1999; *McQueen and Lean*, 1987; *Smith*, 1983; *Wood et al*., 2010], while others show no relationship between nitrogen deficiency and nitrogen fixers [*Downing et al*., 2001; *Kosten et al*., 2012; *Paerl et al*., 2011a]. In Lake Paldang, TN:TP ratio was not a good predictor of cyanobacteria proportion, but the TDN:TDP ratio and TDN (dissolved inorganic nitrogen) concentrations were negatively correlated with relative cyanobacteria abundance (Figure 7c). Note that the TN:TP and TDN:TDP ratios in Lake Paldang were generally higher (with the mean mass TN:TP of 62) than the range (mass TN:TP <29) that was suggested to indicate nitrogen limitation [*Smith*, 1983]. Even when the relative abundance of cyanobacteria exceeded 0.5, the mean TN:TP ratio was 47, ranging from 21 to 72, with a mean TN of 1.9 mg L^{-1}. Similar patterns, cyanobacteria dominance at TN:TP>29 with high nitrogen concentrations, were also observed in Taechung Reservoir, South Korea [*An and Jones*, 2000a].

Previous observational studies have reported the cyanobacterial dominance at high temperatures [*An and Jones*, 2000a; *Kosten et al*., 2012; *Soranno*, 1997], although a recent experimental study reported that optimal growth temperatures and mean growth rates at the optimum temperature do not differ significantly between cyanobacteria and chlorophytes [*Lürling et al*., 2013]. While water temperature was not a strong predictor of relative cyanobacteria abundance in our data (Figure 7a), there were general increases in relative cyanobacteria abundance with increasing temperature, with the exception of two observations (Figure 7a). In both cases, low SS (3 mg L^{-1} on average) and low flushing rates (mean weekly outflow = 224 m^3 s^{-1}), coincident with low N:P ratios (mass TDN:TDP = 106 compared to the overall mean of 150) may have contributed to high proportions of cyanobacteria to total phytoplankton cell counts under low temperature.

Our results corroborate temperature as a key factor associated with the initiation and proliferation of cyanobacteria. Numerous studies including experiments, field studies, modeling, and observational studies based on long-term data of both single lakes and cross-sectional data have indicated that temperature is an important factor influencing the success of cyanobacteria in freshwater ecosystems [*Davis et al*., 2009; *Downing et al*., 2001; *Jöhnk et al*., 2008; *Kosten et al*., 2012; *McQueen and Lean*, 1987; *Paerl*, 1988; *Visser et al*., 1995]. Further, recent publications addressing the effects of climate change on cyanobacteria blooms predict that the duration, intensity, or distribution of cyanobacteria blooms will increase with the climate warming [*Elliott*, 2012; *Kosten et al*., 2012; *Markensten et al*., 2010; *Paerl and Huisman*, 2008; *Paerl and Paul*, 2012; *Paerl and Scott*, 2010].

Temperate lakes generally experience warm summer temperatures accompanied by water column stratification and low flushing rates, conditions that are ideal for cyanobacteria growth. This is not often the case for lakes influenced by the Asian monsoon. Although across seasons, SS, outflow, and water temperature generally exhibit weak correlations, during the summer monsoon period, focused precipitation results in turbulent, well-mixed water columns, extremely high SS levels, and high outflow, conditions that do not favor algal growth. Thus, the intensity and duration of the summer monsoon play an important role in cyanobacteria bloom formation, which may be idiosyncratic to Asian lakes. An intense monsoon that occurs during the warmest period of the year can suppress the onset and magnitude of cyanobacteria blooms by increasing flushing rates, disturbing water column stability, and cooling the water [*An and Jones*, 2000a].

Contrary to our expectations, nutrients, in particular phosphorus concentrations did not show strong relationships with any types of responses (presence/absence, abundance, or relative abundance of cyanobacteria). One possible explanation is that the lake is overenriched with both phosphorus and nitrogen. During the study period, the average TP (47 (± 1.79; 1 standard error) $\mu g\ L^{-1}$) and average TN (2.2 (±0.02) mg L^{-1}) were consistent with eutrophic conditions [*Reckhow and Chapra*, 1983]. Many studies could have shown that high phosphorus and nitrogen levels interact with other physical factors, including temperature, to promote the growth of cyanobacteria [*Davis et al*., 2009; *Kosten et al*., 2012; *Paerl and Scott*, 2010; *Paerl et al*., 2011b]. Recently, *Brookes and Carey* [2011] reported nutrient input as the most important driver of cyanobacteria blooms, as evidenced by decreased blooms subsequent to decreases in nutrient inputs in a number of lakes, even under warming climate. Similarly, high phosphorus inputs, warm, quiescent conditions and a long residence time were all implicated as factors that promoted a record harmful algal bloom in Lake Erie in 2011 [*Michalak et al*., 2013].Thus, although loading was not explicitly examined herein, management of nutrient inputs to tackle cyanobacteria symptoms may also be important in Lake Paldang.

To issue a cyanobacteria watch, the joint exceedance of cyanobacteria and Chla threshold levels is required. However, Chla concentration was not a good indicator of high cyanobacteria concentrations as indicated by the week correlation between cyanobacteria cell counts and Chla (Figure 8a). Also, many zeros for the cell counts made it difficult to characterize the cell counts-Chla relationship (Figure 8a). The poor relationship probably arises because Chla is a component of all types of phytoplankton. The weak cyanobacteria-Chla correlation has important management implications with regard to issuing cyanobacteria watch or alert. Our data demonstrate that by requiring the levels of Chla to be exceeding 15 mg m^{-3} and 25 mg m^{-3} jointly with cyanobacteria cell counts, we miss 12 among 31 cyanobacteria levels $>$ 500 cells mL^{-1} and miss three among five cyanobacteria levels $>$ 5000 cells mL^{-1} (Figure 8a). Thus, the cyanobacteria-Chla relationship and using the joint exceedance as a requirement to issue the watch/alert should be revisited in other lakes that supply drinking water.

5. Conclusions

The hurdle Poisson model provided a framework to address excessive zeros, capturing the factors attributable to the shift from cyanobacteria-free to cyanobacteria-existent conditions. Our results indicate that low temperatures inhibit cyanobacteria bloom initiation and rapid flushing may remove cyanobacteria before they can proliferate. Also, low temperatures along with low suspended solids were conditions that promoted low cyanobacterial abundance. Further, our modeling approach deals with differing measures of cyanobacteria, presence/absence, and (relative) abundance of cyanobacteria in a single model, while previous studies have focused on analyzing either cyanobacteria abundance or relative abundance, not explicitly taking into account the state of cyanobacteria absence. However, mechanisms underlying the differing responses and the degree of overlap between the mechanisms are unclear, and invite further investigation.

As shown in predictions, large uncertainty is associated with high cyanobacteria levels, which would be of interest to lake managers. Using Bayesian inference quantifies the uncertainty in parameters and predictions. Recognizing that the sample size of cyanobacteria cell counts exceeding the watch level (500 cells mL^{-1}) and the alert level (5000 cells mL^{-1}) is merely 31 and 5 out of 207 observations sampled for 5 years, continued monitoring efforts especially at high cyanobacteria abundance would reduce the uncertainty associated with parameter estimates and ensure the better delineation of the relationships with predictors.

Acknowledgments
GLERL contribution number 1702. We would like to thank Tara Stow and anonymous reviewers for their insightful comments. This study was partly supported by the "Early Warning System of Algal Bloom" project of NIER(National Institute of Environmental Research), Ministry of Environment, Korea.

References

Ahn, C.-Y., A.-S. Chung, and H.-M. Oh (2002), Rainfall, phycocyanin, and N:P ratios related to cyanobacterial blooms in a Korean large reservoir, *Hydrobiologia*, *474*(1–3), 117–124.

Ahn, C.-Y., H.-M. Oh, and Y.-S. Park (2011), Evaluation of environmental factors on cyanobacteria bloom in eutrophic reservoir using artificial neural networks, *J. Phycol.*, *47*(3), 495–504.

An, K.-G., and J. R. Jones (2000a), Factors regulating bluegreen dominance in a reservoir directly influenced by the Asian monsoon, *Hydrobiologia*, *432*, 37–48.

An, K.-G., and J. R. Jones (2000b), Temporal and spatial patterns in salinity and suspended solids in a reservoir influenced by the Asian monsoon, *Hydrobiologia*, *436*(1–3), 179–189.

Beaulieu, M., F. Pick, and I. Gregory-Eaves (2013), Nutrients and water temperature are significant predictors of cyanobacterial biomass in a 1147 lakes data set, *Limnol. Oceanogr.*, *58*(5), 1736–1746.

Brookes, J. D., and C. C. Carey (2011), Resilience to blooms, *Science*, *334*(6052), 46–47.

Carey, C. C., B. W. Ibelings, E. P. Hoffmann, D. P. Hamilton, and J. D. Brookes (2012), Eco-physiological adaptations that favour freshwater cyanobacteria in a changing climate, *Water Res.*, *46*(5), 1394–1407.

Carmichael, W. W., S. Azevedo, J. S. An, R. J. R. Molica, E. M. Jochimsen, S. Lau, K. L. Rinehart, G. R. Shaw, and G. K. Eaglesham (2001), Human fatalities from cyanobacteria: Chemical and biological evidence for cyanotoxins, *Environ. Health Perspect.*, *109*(7), 663–668.

Carpenter, S. R., N. F. Caraco, D. L. Correll, R. W. Howarth, A. N. Sharpley, and V. H. Smith (1998), Nonpoint pollution of surface waters with phosphorus and nitrogen, *Ecol. Appl.*, *8*(3), 559–568.

Codd, G. A., L. F. Morrison, and J. S. Metcalf (2005), Cyanobacterial toxins: Risk management for health protection, *Toxicol. Appl. Pharmacol.*, *203*, 264–272.

Davis, T. W., D. L. Berry, G. L. Boyer, and C. J. Gobler (2009), The effects of temperature and nutrients on the growth and dynamics of toxic and non-toxic strains of *Microcystis* during cyanobacteria blooms, *Harmful Algae*, *8*(5), 715–725.

Downing, J. A., S. B. Watson, and E. McCauley (2001), Predicting cyanobacteria dominance in lakes, *Can. J. Fish. Aquat. Sci.*, *58*(10), 1905–1908.

Dyble, J., G. L. Fahnenstiel, R. W. Litaker, D. F. Millie, and P. A. Tester (2008), Microcystin concentrations and genetic diversity of *Microcystis* in the lower Great Lakes, *Environ. Toxicol.*, *23*(4), 507–516.

Elliott, J. A. (2010), The seasonal sensitivity of cyanobacteria and other phytoplankton to changes in flushing rate and water temperature, *Global Change Biol.*, *16*(2), 864–876.

Elliott, J. A. (2012), Is the future blue-green? A review of the current model predictions of how climate change could affect pelagic freshwater cyanobacteria, *Water Res.*, *46*(5), 1364–1371.

Freeman, K. S. (2011), Forecasts aid HABs response, *Environ. Health Perspect.*, *119*(12), a510.

Ghosh, S. K., P. Mukhopadhyay, and J.-C. Lu (2006), Bayesian analysis of zero-inflated regression models, *J. Stat. Plann. Inference*, *136*(4), 1360–1375.

Glibert, P. M., J. Allen, A. Bouwman, C. W. Brown, K. J. Flynn, A. J. Lewitus, and C. J. Madden (2010), Modeling of HABs and eutrophication: Status, advances, challenges, *J. Mar. Syst.*, *83*(3), 262–275.

Graham, J. L., J. R. Jones, S. B. Jones, J. A. Downing, and T. E. Clevenger (2004), Environmental factors influencing microcystin distribution and concentration in the Midwestern United States, *Water Res.*, *38*(20), 4395–4404.

Hamilton, G., R. McVinish, and K. Mengersen (2009), Bayesian model averaging for harmful algal bloom prediction, *Ecol. Appl.*, *19*(7), 1805–1814.

Heilbron, D. C. (1994), Zero-altered and other regression models for count data with added zeros, *Biometrical. J.*, *36*(5), 531–547.

Hu, M.-C., M. Pavlicova, and E. V. Nunes (2011), Zero-inflated and hurdle models of count data with extra zeros: Examples from an HIV-risk reduction intervention trial, *Am. J. Drug Alcohol Abuse*, *37*(5), 367–375.

Jöhnk, K. D., J. E. F. Huisman, J. Sharples, B. E. N. Sommeijer, P. M. Visser, and J. M. Stroom (2008), Summer heatwaves promote blooms of harmful cyanobacteria, *Global Change Biol.*, *14*(3), 495–512.

Jones, I. A. N., G. George, and C. Reynolds (2005), Quantifying effects of phytoplankton on the heat budgets of two large limnetic enclosures, *Freshwater Biol.*, *50*(7), 1239–1247.

Jung, Y. (2009), Management plan for the control of eutrophication in Lake Paldang, South Korea, Policy Res. Rep. 2009-43, Gyeonggi Res. Inst., Suwon, South Korea.

Kim, J., M.-S. Han, W. Jheong, and J. Park (2005), Correlation between phytoplankton dynamics and water quality in Paldang Reservoir, *Korean J. Limnol.*, *38*(2), 217–224.

Kim, J., S. Lee, H. Bang, and S. Hwang (2009), Characteristics of algae occurrence in Lake Paldang, *J. Korean Soc. Environ. Eng.*, *31*(5), 325–331.

Knoll, L. B., O. Sarnelle, S. K. Hamilton, C. E. H. Kissman, A. E. Wilson, J. B. Rose, and M. R. Morgan (2008), Invasive zebra mussels (Dreissena polymorpha) increase cyanobacterial toxin concentrations in low-nutrient lakes, *Can. J. Fish. Aquat. Sci.*, *65*(3), 448–455.

Komárek, J. (1991), A review of water-bloom forming *Microcystis* species, with regard to populations from Japan, *Algol. Stud./Archiv für Hydrobiologie*, *64*, 115–127.

Kosten, S., et al. (2012), Warmer climates boost cyanobacterial dominance in shallow lakes, *Global Change Biol.*, *18*(1), 118–126.

Lambert, D. (1992), Zero-inflated Poisson regression, with an application to defects in manufacturing, *Technometrics*, *34*(1), 1–14.

Levine, S. N., and D. W. Schindler (1999), Influence of nitrogen to phosphorus supply ratios and physicochemical conditions on cyanobacteria and phytoplankton species composition in the Experimental Lakes Area, Canada, *Can. J. Fish. Aquat. Sci.*, *56*(3), 451–466.

Lunn, D. J., A. Thomas, N. Best, and D. Spiegelhalter (2000), WinBUS—A Bayesian modelling framework: Concepts, structure, and extensibility G, *Stat. Comput.*, *10*, 325–337.

Lürling, M., F. Eshetu, E. J. Faassen, S. Kosten, and V. L. M. Huszar (2013), Comparison of cyanobacterial and green algal growth rates at different temperatures, *Freshwater Biol.*, *58*(3), 552–559.

Markensten, H., K. Moore, and I. Persson (2010), Simulated lake phytoplankton composition shifts toward cyanobacteria dominance in a future warmer climate, *Ecol. Appl.*, *20*(3), 752–767.

Martin, T. G., B. A. Wintle, J. R. Rhodes, P. M. Kuhnert, S. A. Field, S. J. Low-Choy, A. J. Tyre, and H. P. Possingham (2005), Zero tolerance ecology: improving ecological inference by modelling the source of zero observations, *Ecol. Lett.*, *8*(11), 1235–1246.

McQueen, D. J., and D. R. S. Lean (1987), Influence of water temperature and nitrogen to phosphorus ratios on the dominance of blue-green algae in Lake St. George, Ontario, *Can. J. Fish. Aquat. Sci.*, *44*(3), 598–604.

Michalak, A. M., et al. (2013), Record-setting algal bloom in Lake Erie caused by agricultural and meteorological trends consistent with expected future conditions, *Proc. Natl. Acad. Sci. U. S. A.*, *110*(16), 6448–6452.

Millie, D. F., G. R. Weckman, W. A. Young II, J. E. Ivey, H. J. Carrick, and G. L. Fahnenstiel (2012), Modeling microalgal abundance with artificial neural networks: Demonstration of a heuristic 'Grey-Box' to deconvolve and quantify environmental influences, *Environ. Modell. Software*, *38*, 27–39.

Millie, D. F., G. R. Weckman, W. A. Young II, J. E. Ivey, D. P. Fries, E. Ardjmand, and G. L. Fahnenstiel (2013), Coastal 'Big Data' and nature-inspired computation: Prediction potentials, uncertainties, and knowledge derivation of neural networks for an algal metric, *Estuarine Coastal Shelf Sci.*, *125*, 57–67.

Mitrovic, S. M., B. C. Chessman, L. C. Bowling, and R. H. Cooke (2006), Modelling suppression of cyanobacterial blooms by flow management in a lowland river, *River Res. Appl.*, *22*(1), 109–114.

Mullahy, J. (1986), Specification and testing of some modified count data models, *J. Econometrics*, *33*(3), 341–365.

Na, E. H., and S. S. Park (2005), A hydrodynamic modeling study to determine the optimum water intake location in Lake Paldang, Korea, *J. Am. Water Resour. Assoc.*, *41*(6), 1315–1332.

Na, E. H., and S. S. Park (2006), A hydrodynamic and water quality modeling study for spatial and temporal patterns of phytoplankton growth in a stratified lake with buoyant incoming flow, *Ecol. Modell.*, *199*, 298–314.

Neelon, B. H., A. J. O'Malley, and S.-L. T. Normand (2010), A Bayesian model for repeated measures zero-inflated count data with application to outpatient psychiatric service use, *Stat. Modell.*, *10*(4), 421–439.

Neelon, B. H., P. Ghosh, and P. F. Loebs (2013), A spatial Poisson hurdle model for exploring geographic variation in emergency department visits, *J. R. Stat. Soc.: Ser. A*, *176*(2), 389–413.

O'Brien, K. R., D. L. Meyer, A. M. Waite, G. N. Ivey, and D. P. Hamilton (2004), Disaggregation of *Microcystis aeruginosa* colonies under turbulent mixing: Laboratory experiments in a grid-stirred tank, *Hydrobiologia*, *519*(1–3), 143–152.

Paerl, H. W. (1988), Nuisance phytoplankton blooms in coastal, estuarine, and inland waters, *Limnol. Oceanogr.*, *33*(4), 823–847.

Paerl, H. W., and J. Huisman (2008), Blooms like it hot, *Science*, *320*(5872), 57–58.

Paerl, H. W., and J. T. Scott (2010), Throwing fuel on the fire: Synergistic effects of excessive nitrogen inputs and global warming on harmful algal blooms, *Environ. Sci. Technol.*, *44*(20), 7756–7758.

Paerl, H. W., and V. J. Paul (2012), Climate change: Links to global expansion of harmful cyanobacteria, *Water Res.*, *46*(5), 1349–1363.

Paerl, H. W., N. S. Hall, and E. S. Calandrino (2011a), Controlling harmful cyanobacterial blooms in a world experiencing anthropogenic and climatic-induced change, *Sci. Total Environ.*, *409*(10), 1739–1745.

Paerl, H. W., H. Xu, M. J. McCarthy, G. Zhu, B. Qin, Y. Li, and W. S. Gardner (2011b), Controlling harmful cyanobacterial blooms in a hyper-eutrophic lake (Lake Taihu, China): The need for a dual nutrient (N & P) management strategy, *Water Res.*, *45*(5), 1973–1983.

Park, H.-K., H.-J. Lee, E.-K. Kim, and D.-I. Jung (2005), Characteristics of algal abundance and statistical analysis of environmental factors in Lake Paldang, *Korean Soc. Water Qual.*, *21*(6), 584–594.

Phillips, G., A. Kelly, J.-A. Pitt, R. Sanderson, and E. Taylor (2005), The recovery of a very shallow eutrophic lake, 20 years after the control of effluent derived phosphorus, *Freshwater Biol.*, *50*(10), 1628–1638.

Plummer, M. (2008), Penalized loss functions for Bayesian model comparison, *Biostatistics*, *9*(3), 523–539.

R Development Core Team (2012), *R: A Language and Environment for Statistical Computing*, R Found. for Stat. Comput., Vienna.

Reckhow, K. H., and S. C. Chapra (1983), *Engineering Approaches for Lake Management, Volume 1: Data Analysis and Empirical Modeling*, Butterworths, Boston, Mass.

Reynolds, C. S. (2006), *Ecology of Phytoplankton*, Cambridge Univ. Press, Cambridge, U. K.

Romo, S., J. Soria, F. FernÁNdez, Y. Ouahid, and Á. BarÓN-SolÁ (2013), Water residence time and the dynamics of toxic cyanobacteria, *Freshwater Biol.*, *58*(3), 513–522.

Scheffer, M., S. Rinaldi, A. Gragnani, L. R. Mur, and E. H. v. Nes (1997), On the dominance of filamentous cyanobacteria in shallow, turbid Lakes, *Ecology*, *78*(1), 272–282.

Serizawa, H., T. Amemiya, A. Rossberg, and K. Itoh (2008), Computer simulations of seasonal outbreak and diurnal vertical migration of cyanobacteria, *Limnology*, *9*(3), 185–194.

Smith, V. H. (1983), Low nitrogen to phosphorus ratios favor dominance by blue-green algae in lake phytoplankton, *Science*, *221*(4611), 669–671.

Smith, V. H. (1986), Light and nutrient effects on the relative biomass of blue-green algae in lake phytoplankton, *Can. J. Fish. Aquat. Sci.*, *43*(1), 148–153.

Soranno, P. A. (1997), Factors affecting the timing of surface scums and epilimnetic blooms of blue-green algae in a eutrophic lake, *Can. J. Fish. Aquat. Sci.*, *54*, 1965–1975.

Stumpf, R. P., T. T. Wynne, D. B. Baker, and G. L. Fahnenstiel (2012), Interannual variability of cyanobacterial blooms in Lake Erie, *PloS One*, *7*(8), e42444.

Vanderploeg, H. A., T. H. Johengen, and J. R. Liebig (2009), Feedback between zebra mussel selective feeding and algal composition affects mussel condition: Did the regime changer pay a price for its success?, *Freshwater Biol.*, *54*(1), 47–63.

Visser, P., B. A. S. Ibelings, B. Van Der Veer, J. A. N. Koedood, and R. Mur (1996), Artificial mixing prevents nuisance blooms of the cyanobacterium *Microcystis* in Lake Nieuwe Meer, the Netherlands, *Freshwater Biol.*, *36*(2), 435–450.

Visser, P. M., B. W. Ibelings, and L. R. Mur (1995), Autunmal sedimentation of *Microcystis* spp. as result of an increase in carbohydrate ballast at reduced temperature, *J. Plankton Res.*, *17*(5), 919–933.

Wagner, C., and R. Adrian (2009), Cyanobacteria dominance: Quantifying the effects of climate change, *Limnol. Oceanogr.*, *54*(6), 2460–2468.

White, G. C., and R. E. Bennetts (1996), Analysis of frequency count data using the negative binomial distribution, *Ecology*, *77*(8), 2549–2557.

Wood, S. A., M. J. Prentice, K. Smith, and D. P. Hamilton (2010), Low dissolved inorganic nitrogen and increased heterocyte frequency: Precursors to *Anabaena* planktonica blooms in a temperate, eutrophic reservoir, *J. Plankton Res.*, *32*(9), 1315–1325.

Wynne, T. T., R. P. Stumpf, M. C. Tomlinson, D. J. Schwab, G. Y. Watabayashi, and J. D. Christensen (2011), Estimating cyanobacterial bloom transport by coupling remotely sensed imagery and a hydrodynamic model, *Ecol. Appl.*, *21*(7), 2709–2721.

Wynne, T. T., R. P. Stumpf, and T. O. Briggs (2013), Comparing MODIS and MERIS spectral shapes for cyanobacterial bloom detection, *Int. J. Remote Sens.*, *34*, 6668–6678.

Zhang, H., W. Hu, K. Gu, Q. Li, D. Zheng, and S. Zhai (2013), An improved ecological model and software for short-term algal bloom forecasting, *Environ. Modell. Software*, *48*, 152–162.