Chemosphere 384 (2025) 144481

Contents lists available at ScienceDirect

Chemosphere

journal homepage: www.elsevier.com/locate/chemosphere

Research Paper

A Bayesian hierarchical modeling approach can improve measurement accuracy of microcystin concentrations☆

Sabrina Jaffe [a,*], Duane Gossiaux [b], Reagan M. Errera [b], Emanuela Gionfriddo [c,d,1], Song S. Qian [a]

[a] Department of Environmental Sciences, The University of Toledo, Toledo, OH 43606, USA
[b] Great Lakes Environmental Research Laboratory, National Oceanic and Atmospheric Administration, Ann Arbor, MI, USA
[c] Department of Chemistry and Biochemistry, The University of Toledo, Toledo, OH 43606, USA
[d] Department of Chemistry, University at Buffalo, The State University of New York, Buffalo, NY 14260-3000, USA

ARTICLE INFO

Handling Editor: Teresa Cutright

Dataset link: Lake Erie HABs monitoring data and R code (Original data)

Keywords:
Bayesian statistics
Calibration
Hierarchical modeling
Cyanobacteria
ELISA

ABSTRACT

The Bayesian hierarchical model (BHM) is a framework that improves parameter estimation by leveraging information from different sources. In an environmental monitoring program, we often measure important chemical concentrations using calibration-based methods. These methods require fitting a calibration curve repeatedly each time with a small number of standard solutions of known concentrations. This approach is often associated with large estimation uncertainty in the measured concentrations. BHM is a perfect method for reducing calibration curve uncertainty, thereby enhancing the accuracy and stability of the resulting concentration measurements. We demonstrate the effectiveness of a BHM approach by estimating microcystin concentrations from the Lake Erie harmful algal bloom (HAB) monitoring program operated by the Great Lakes Environmental Research Laboratory of the National Oceanic and Atmospheric Administration. We introduced a sequential updating algorithm to implement the BHM framework so that the BHM model can be fit and updated one test at a time. By comparing estimated quality control sample concentrations to their known values, we show that the BHM method yields the best accuracy compared to the currently used methods. Due to the sequential updating approach, the BHM can be readily incorporated into a lab without requiring additional changes to lab procedures, thus offering a key advantage over traditional calibration methods. This advancement could reduce health risks and false-positive shutdowns during HAB events.

1. Introduction

Cyanobacterial toxins threaten water quality worldwide (Huisman et al., 2018; Chorus and Welker, 2021). Major toxins of concern include microcystins (MC), which are a class of cyclic heptapeptides representing over 240 identified compounds (Svirčev et al., 2017; Babica et al., 2006; Gan et al., 2012) with varying toxicities. Toxic blooms of Microcystis spp., the major freshwater cyanobacteria producer of microcystins, occur annually within the western basin of Lake Erie (Boegehold et al., 2023) impacting both human and ecosystem health. The cyanobacteria harmful algal bloom (HAB) is extensively monitored to provide stakeholders (water managers, public, etc.) with data regarding bloom toxin concentrations, assisting their decision-making for public safety. Accurately measuring MC concentrations is not only a public safety concern, but also a social and economic concern. For instance, in 2014, high concentrations of MC in the western basin caused a "do not drink" water advisory in Toledo, Ohio, that lasted 3 days and left 500,000 residents without potable water (Toledo Department of Public Utilities, 2014). This event resulted in an estimated economic cost of 65 million dollars (Bingham et al., 2015).

The most common method for measuring MC concentrations is the Enzyme-Linked Immunosorbent Assay (ELISA). The ELISA is a calibration-based method shown to be highly variable (Qian et al., 2015). Calibration curves are regression models fit with instrument measured absorbances of standard solutions as the response and their corresponding known analyte concentrations as the predictor. Analyte

☆ Qian and Jaffe's work is partially supported by a grant from Ohio Sea Grant (2020-2023), Qian and Gionfriddo were partially supported by the University of Toledo Interdisciplinary grant (2020-2021). We thank two anonymous reviewers for their constructive feedback and suggestions.

* Correspondence to: The University of Toledo, 2801 Bancroft St, Toledo, OH 43606, USA.
E-mail addresses: sabrina.jaffe@rockets.utoledo.edu (S. Jaffe), duane.gossiaux@noaa.gov (D. Gossiaux), reagan.errera@noaa.gov (R.M. Errera), egionfri@buffalo.edu (E. Gionfriddo), song.qian@utoledo.edu (S.S. Qian).
[1] Current address: Department of Chemistry, University at Buffalo, The State University of New York, Buffalo, NY 14260-3000, USA.

https://doi.org/10.1016/j.chemosphere.2025.144481
Received 25 October 2024; Received in revised form 2 May 2025; Accepted 6 May 2025
Available online 6 June 2025

concentrations are estimated using a two-step process of (1) fitting the calibration curve using the standard solutions and (2) estimating unknown concentrations using the inverse function of the fitted regression model. However, the calibration curve approach used to estimate MC concentrations in Lake Erie's western basin has a high level of uncertainty because of the small sample size (5 data points) used to establish the calibration curve, which results in highly variable estimated concentrations (Qian et al., 2015, 2025). This can be problematic when the estimated concentrations are used to make decisions that impact management and public safety. The small sample size used for fitting the calibration curve is likely due to the limited space on the ELISA plates used for processing samples. One way to improve the ELISA method is to increase the number of standard samples used to develop the calibration curve. However, dedicating more space to standard samples may not be practical or cost effective for labs. Alternatively, we can adopt a Bayesian regression approach and incorporate relevant information that already exists. Specifically, previous calibration curves developed in the lab can be used to create informative priors (Klauenberg et al., 2015). If these past calibration curves can be leveraged to help enhance the current calibration curve, then we would not need to change the lab procedures. This is a key advantage of the Bayesian statistical framework, summarizing all of the curves into prior distributions of calibration curve parameters thereby improving the quality of the next calibration curve. This can be achieved by using a Bayesian hierarchical modeling (BHM) approach (Qian et al., 2015, 2025).

The BHM is a flexible framework that can be used to leverage relevant information from similar sources to improve parameter estimation (Gelman et al., 2014). Intuitively, when a parameter is estimated with error, it means that the estimate is either too high or too low compared to the unknown underlying "true" value. If we are only estimating one parameter, then we prefer the unbiased estimator. Once we have multiple estimates of similar parameters, the multiple estimates can provide information on whether an estimate is likely over or underestimated. For example, if one estimated parameter is way above the overall average, we can improve the estimation uncertainty by moving it towards the overall average. By moving the individual estimates toward the overall average, we shrink the range of the multiple estimates. This is why this class of estimator is called the shrinkage estimator (Efron, 1975). The BHM is a shrinkage estimator which calculates the amount of shrinkage based on the relative magnitude of within and among group variances.

In the case of ELISA tests, information can be leveraged from within individual tests and across different tests to improve the overall estimation accuracy of MC concentrations. Within a single ELISA test, multiple water samples are estimated at the same time. According to Stein's paradox, when estimating three or more concentrations together, it is always better to shrink the estimates toward the overall average of the estimates (Stein, 1956; Efron, 1975). The BHM approach can be implemented to achieve estimation accuracy improvement predicted by Stein's paradox (Gelman et al., 2004). Across multiple tests, there is information available on the calibration curve parameters. The BHM is a modeling structure that can share information across all tests to develop prior distributions for the calibration curve estimated based on multiple curves. This would reduce the estimation uncertainty for individual MC estimates by shrinking curves toward the "average" curve. The BHM approach is mathematically predicted to improve the estimation accuracy of the individual samples within each test compared to the currently used two-step process.

Although BHM has been shown to be effective in improving estimation accuracy (Qian et al., 2025), data from multiple tests are needed when sharing information across tests. Therefore, running a BHM model can be impractical because it requires the availability of data from multiple tests. Even for labs with such data already available, using BHM for each additional test requires combining the most recent data with data from previous tests. As the number of tests increases, not only is the process of combining data cumbersome, but also the computational burden will inevitably become increasingly intolerable. To avoid these problems and to prevent changes to current lab procedures, we propose a sequential updating algorithm to easily implement the BHM approach to new and established labs. The algorithm is based on the idea that the posterior distribution from previous tests can serve as the priors for future tests (Efron, 1996). The subsequent (updated) posterior distribution can then be used as priors for the next test. This allows for new tests to be analyzed one at a time.

To assess the practical feasibility of the BHM approach, we applied it to data from a long-term water quality monitoring program conducted by the Great Lakes Environmental Research Laboratory (GLERL) of the National Oceanography and Atmospheric Administration (NOAA). This program has been regularly monitoring MC concentrations in the western basin of Lake Erie since 2012. Through our study, we illustrate how the BHM approach can be effectively implemented in real-world laboratory settings and how it significantly improves estimation accuracy when estimating unknown concentrations.

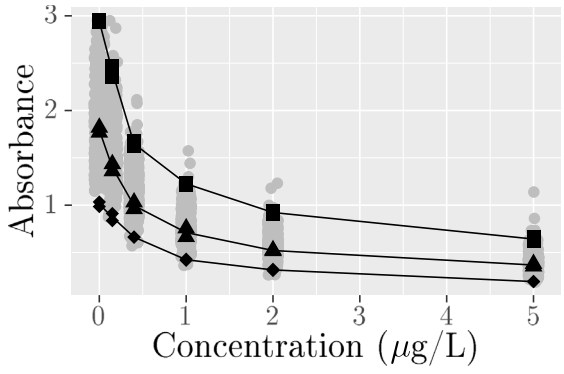

Fig. 1. Raw ELISA testing data from the GLERL western Lake Erie harmful algal bloom monitoring program. The gray, solid circles are the measured standard solution data points from the 214 tests. The black shapes are the results from three tests with the highest (squares) lowest (diamonds) and median response value (triangles) at the 0 concentration point.

2. Methods

2.1. Microcystin data and ELISA protocol

NOAA GLERL established eight monitoring sites in 2012 that are sampled weekly during the HAB season for a variety of water quality parameters, including microcystin concentrations (Boegehold et al., 2023). MC concentrations are analyzed using an ELISA kit (Abraxis) and data are distributed to stakeholders within 48-hours of collection. We obtained 214 sets of ELISA test results from NOAA-GLERL, encompassing data processed between 2012 and 2021. Per the manual of this specific ELISA kit, each test includes six standard solutions with known MC concentrations ranging from 0 to 5.00 μg/L and a quality control sample with an MC concentration of 0.75 μg/L. Throughout the 214 tests, the instrumental responses exhibited considerable variation (Fig. 1). This variation could be due to differences in the kits being used and the person performing the test. The absorbance variance across tests is higher at lower concentrations, but the variation actually has the opposite effect on the concentration variance. At higher concentrations, the curve becomes "flatter" which can make it more difficult to find the optimal concentration when estimating concentrations.

Up to 40 water samples with unknown MC concentrations could be tested with each test kit, and all samples were tested with a replicate. Before fitting the calibration curve and estimating unknown MC concentrations, the measured instrumental responses from each pair of replicates were averaged. The instrumental responses were fit as a

linear function of log MC concentrations for tests conducted until 18 July 2016 (tests 1-83) except for the test from 31 August 2015 which was fit with a nonlinear function (Fig. 2). The linear calibration curve followed the form:

$$\log(y) = \beta_0 + \beta_1 \log(x) + \varepsilon \quad (1)$$

where y represents the instrumental response, x is the known MC concentration, and ε is the residual assumed to follow a normal distribution $N(0, \sigma^2)$. The measured response values for standard solutions (each with two replicates) were averaged and divided by the average of the 0-concentration standard solution (relative absorbance). This log–log linear calibration curve is an approximation of the four-parameter logistic function (Eq. (4), see Supporting Materials from Nummer et al. (2018), https://github.com/songsqian/ELISA/blob/master/FPL/ELISAexpOnLine.pdf). Due to the log transformation of x, the calibration coefficients (β_0 and β_1) are estimated based on the five non-zero standard solutions (relative absorbances), resulting in a regression model fit with five data points. Consequently, the residual variance (σ^2) was estimated with degrees of freedom of 3. Starting from 25 July 2016 and the test conducted on 31 August 2015 (tests 84-214), the calibration curve was described by a nonlinear regression model (Fig. 2) in the form of the four-parameter logistic function:

$$y = f(x, \theta) = \theta_4 + \frac{\theta_1 - \theta_4}{1 + \left(\frac{x}{\theta_3}\right)^{\theta_2}} + \varepsilon \quad (2)$$

Where $\theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$. This model involves four unknown parameters, and the residual variance estimation has a degree of freedom of 1. The four-parameter logistic function is considered a more appropriate calibration curve because it characterizes the response better than the linear approach (Findlay and Dillard, 2007; Biosensis, 2024). In both the linear and nonlinear calibration models, the degrees of freedom are below 4, making it difficult to perform a reliable statistical assessment of predictive uncertainty for the fitted regression models because the sampling distribution of σ^2 is an inverse-Chi-square distribution and its variance does not exist when its degrees of freedom is less than 4. Alternatively, a Bayesian approach could help avoid the problems with the inverse-function method. However, priors are needed for the parameters when using a Bayesian approach (Box and Tiao, 1973). BHM is a form of empirical Bayes which derives priors using data from, in this case, the ELISA tests.

2.2. Bayesian hierarchical model as an alternative approach

The MC ELISA dataset is well suited for implementing a BHM approach because each test estimates multiple MC concentrations that can be leveraged to improve concentration estimation and there are multiple ELISA tests that can be leveraged to improve the estimation of calibration curve coefficients. Therefore, we proposed a BHM with two levels of information sharing.

Within an ELISA test, there are multiple MC concentrations to be estimated simultaneously. Stein's paradox suggests that the best estimator for each concentration is determined by shrinking all concentrations toward the overall average concentration of a test. This is because we know that each concentration is either over or underestimated, but we have no gauge for whether the estimate is over or underestimated if we only estimate one concentration. However, when multiple samples are estimated simultaneously, the overall average serves as a reference, hence, shrinking individual estimates can improve overall estimation accuracy. Accordingly, we can implement a hierarchical structure within each ELISA test to improve the estimates for individual concentration. The shrinkage effect is achieved in the BHM through a common prior for all unknown concentrations:

$$\log(x_0^j) \sim N\left(\mu_{x_0}, \sigma_{x_0}^2\right) \quad (3)$$

Where $\log(x_0)$ is the log transformed concentration. This transformation stabilizes the concentration error by fixing the uncertainty at a fixed percentage. $N\left(\mu_{x_0}, \sigma_{x_0}^2\right)$ is the normal prior distribution, with overall mean μ_{x_0} and between sample variance $\sigma_{x_0}^2$. Both parameters are estimated from data along with the unknown concentrations. We used weakly informative priors (Gelman et al., 2008) for μ_{x_0} and $\sigma_{x_0}^2$ for fitting the hierarchical model because we wanted the Bayesian structure to use the data to determine the level of shrinking for each test.

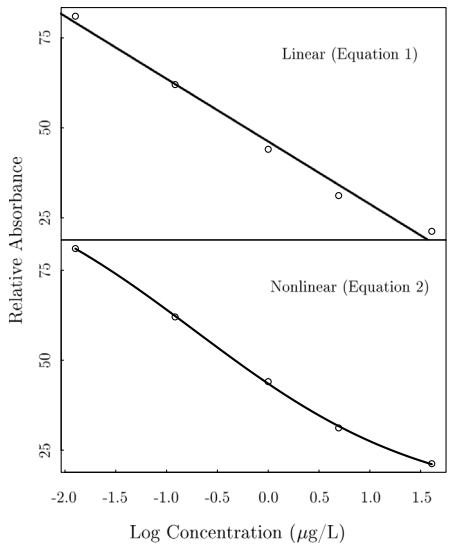

Fig. 2. Examples of the linear (top panel, Eq. (1)) and nonlinear (bottom panel, Eq. (2)) calibration curves fit using the standard samples from test 200. The circles are the relative absorbances (in percent) for the non-zero standard samples in log concentration scale and the black lines are the fitted curves.

Likewise, because we have multiple different ELISA tests that each fit their own calibration curve, we can consider each individual curve as a sample from a population of calibration curves. Therefore, we can leverage calibration curve information across different tests to further improve estimation accuracy. By imposing a common prior on calibration curve coefficients across multiple tests, the BHM induces the shrinkage effect on calibration curves:

$$\theta_k \sim N(\mu_\theta, \sigma_\theta^2) \quad (4)$$

Where θ_k is model coefficients and $N(\mu_\theta, \sigma_\theta^2)$ is the normal prior distribution with μ_θ as the mean for coefficients over multiple tests and σ_θ^2 being the among test variance. Again, we used weakly informative priors for μ_θ and σ_θ^2. We started the BHM modeling process by pooling data from the first nine tests to obtain posterior distributions of μ_θ and σ_θ^2. These posteriors are used as the priors for the subsequent sequential updating process.

2.3. Sequential updating

Sequential updating in the context of calibration curves is to fit each curve one at a time while taking into consideration the previous curves. This is possible under the BHM approach where we can use the hyperparameters of μ_θ and σ_θ^2 of Eq. (4) as priors for the next calibration curve because they are interpreted as the overall mean coefficient of multiple curves and the among curve variance. Efron (1996) suggested that the BHM approach is ideal for reducing the estimation uncertainty in an individual test (Fig. 3). Alternatively, Klauenberg et al. (2015) recommended to derive priors by summarizing existing data, a labor-intensive and highly technical process. Using sequential updating, we can use posterior distributions of μ_θ and σ_θ^2 from the initial nine tests as the prior for the next test (Eq. (4)), instead of non- or weakly-informative priors (Efron, 1996).

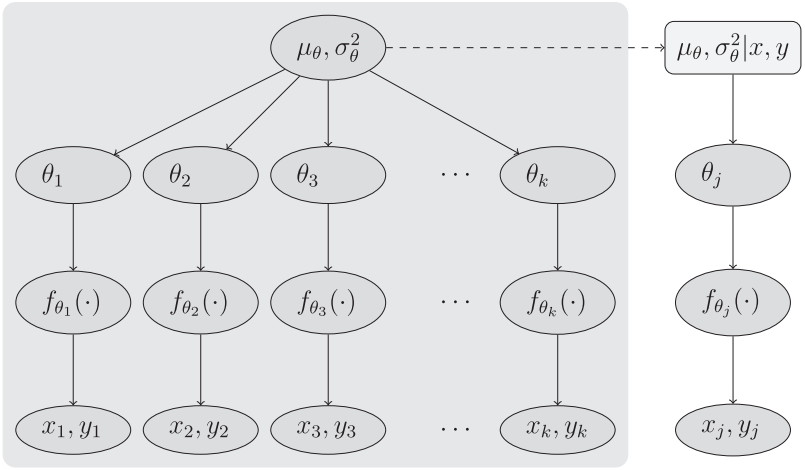

Fig. 3. Diagram showing the BHM framework and sequential updating process. In the gray shaded area, each column represents a different test, each with their own data (x_k, y_k) and test-specific parameters (θ_k). μ_θ and σ_θ^2 are the hyper-parameters (Eq. (4)) representing the overall mean and among test variance, respectively. When analyzing data from the next test, only the estimated μ_θ and σ_θ^2 are used (as priors, i.e., $\theta_j \sim N(\mu_\theta, \sigma_\theta^2)$). Sequential updating allows for tests to be evaluated individually while taking data from the previous tests into consideration. As more tests are evaluated, the hyperparameters can continue to be updated and used as priors for the next groups.

Specifically, we implemented sequential updating using Markov chain Monte Carlo (MCMC) as the computational method. The posterior random samples of μ_θ and σ_θ^2 are obtained from their joint distribution. Qian and Reckhow (2007) recommended that information represented by these random samples be summarized to quantify the conjugate prior (Bernardo and Smith, 2009) of a normal distribution with unknown mean and variance (inverse-gamma):

$$\begin{aligned} \mu_\theta \mid \sigma_\theta &\sim N(\mu_0, \sigma_\theta/\lambda) \\ \sigma_\theta &\sim IG(\alpha, \beta) \end{aligned}$$

This distribution can then be used as the prior distribution for the next calibration curve coefficient. Consequently, we can fit the next calibration curve using the Bayesian method with informative priors for model coefficients derived from the posterior from the previous BHM model, thereby avoiding fitting the BHM combining data from all tests. Because we used MCMC method, the posterior distribution of μ_θ and σ_θ^2 are represented by random samples and we can derive the four distribution parameters (μ_θ, λ, α, β) using the method of moments. That is, given random samples of μ_θ and σ_θ^2, we calculate their sample means and variances, and equate them to the theoretical mean and variance formulae of the two parameters. The joint prior distribution for the hyper-parameters is specified by four hyper-parameters: μ_θ, λ, α, and β. Assuming the joint posterior distribution of μ_θ and σ_θ^2 can be represented by a normal-inverse gamma distribution, the posterior parameters can be summarized using the method of moments:

$$E(\theta) = \mu, \qquad Var(\theta) = \frac{\beta_\theta}{(\alpha_\theta - 1)\lambda_\theta}$$
$$E(\sigma_\theta^2) = \frac{\beta_\theta}{\alpha_\theta - 1}, \quad Var(\sigma_\theta^2) = \frac{\beta_\theta^2}{(\alpha_\theta - 1)^2(\alpha_\theta - 2)}$$

Solving for the unknown parameters:

$$\mu_\theta^0 = E(\theta), \qquad \lambda_\theta = E(\sigma_\theta^2)/Var(\theta)$$
$$\alpha_\theta = 2 + E^2(\sigma_\theta^2)/Var(\sigma_\theta^2), \quad \beta_\theta = E(\sigma_\theta^2)(\alpha_\theta - 1)$$

These parameters are then used as prior parameters for analyzing data from the next test and their updated posterior distribution can be used to derive the prior for the next iteration. After a number of rounds of updating (nine in our case), the posterior parameters should converge and the subsequent updatings are essentially fitting a Bayesian linear/nonlinear regression model. The difference between BHM and current practices is that current practices fit each test individually while BHM fit individual tests using an informative prior summarized from previous tests (Fig. 3). Given that most labs conducting ELISA tests have data from previous tests, implementing the sequential updating process is feasible.

2.4. Model evaluation

To determine the effectiveness of the BHM approach, we used six different modeling methods for model evaluation (Table 1):

1. Standard Inverse-Function Estimator (IFE5): Following the protocol of the Abaxis ELISA test kits, we fitted calibration curves using the five relative absorbance observations from the standard solutions. Then, the unknown concentrations are estimated using the inverse function of the fitted regression model:

$$\log(x) = \log(\theta_3) - \frac{\log\left(\frac{\theta_1 - y}{y - \theta_4}\right)}{\theta_2} \tag{5}$$

2. Inverse-Function Estimator with All 12 Standard Sample Observations (IFE12): The same inverse-function method as IFE5, but instead the calibration curve was fit using the two replicates for all six standards (not using the relative absorbance approach), resulting in 12 standard samples data points for the model.
3. Bayesian Estimator (Bayes): This method combines the fitting and estimating processes, without leveraging information within and across tests. We used a non-informative prior for all concentration values (no hierarchical structure).
4. BHM with Information Shared Within Each Test (BHM1): Information from all unknown calibration samples within a test is shared to improve estimation accuracy (Eq. (3)). Across-test information is not shared.
5. BHM with Information Shared Across Each Test (BHM2): Information of calibration curve coefficients across all tests is shared to reduce estimation uncertainty. Within-test information is not shared and sequential updating was implemented to analyze one test at a time.
6. BHM with Information Shared Within and Across All Tests (BHM3). Information from unknown calibration samples and calibration curve coefficients are shared. Sequential updating was implemented to analyze one test at a time.

Both IFE5 and IFE12 used the maximum likelihood estimator (MLE).

For all six models, we employed the nonlinear calibration curve (Eq. (2)) for all 214 tests to estimate MC concentrations. We compared the absolute difference between the model estimated control concentrations and the known control concentration of 0.75 μg/L. This absolute difference is known as the accuracy. Additionally, we used an F-test to compare the variance of accuracy of IFE5 and IFE12 to assess the impact of sample size on calibration results. The outcome of IFE12 is

Table 1
Summary of the six models. We used the nonlinear calibration curve (Eq. (2)) for all models. Sample size refers to the number of standard samples used for fitting the calibration curve.

Model name	Sample size	Estimation method	BHM level
IFE5	5	MLE	N/A
IFE12	12	MLE	N/A
Bayes	12	Bayesian	N/A
BHM1	12	BHM	Within-test
BHM2	12	BHM	Across-test
BHM3	12	BHM	Within & across

then contrasted with the four Bayesian methods (Bayes, BHM1, BHM2, and BHM3), which also used the two replicates for all six standard samples (12 data points total) and instrumental responses from water samples with unknown MC concentrations to fit the standard curve and estimate the unknown MC concentrations. For BHM2 and BHM3, we implemented a sequential updating algorithm, where tests 1–9 were used to establish the initial priors. The algorithm allows us to evaluate tests 10-214 one at a time. Model fitting and estimation was completed using Stan (Stan Development Team, 2022b) through R (Core Team, 2022; Stan Development Team, 2022a). To show the improvements in accuracy of the models that used all 12 data points for fitting the curve (IFE12, Bayes, BHM1, BHM2, and BHM3), we evaluated the relative accuracies to directly compare these models to the best performing model.

As previously demonstrated in Chapter 9 of Qian (2016), we used posterior simulation for IFE5 and IFE12 to estimate their estimation uncertainty. Within each Abraxis kit, a control sample with a known MC concentration of 0.75 μg/L was provided. To assess estimation accuracy, we compared the posterior distribution of the estimated control sample MC concentration (represented by 5000 random samples from Monte Carlo simulations) with the known value. To quantify the estimation uncertainty for each test, we define the accuracy as the absolute values of the differences between the 5000 random samples and known value of 0.75 μg/L of the control sample. The median of these absolute differences represents the deviance of the estimated values from the true value, where a smaller deviance represents better accuracy.

3. Results

As expected, increasing the sample size for fitting the calibration curve from 5 (degrees of freedom or df = 1) to 12 (df = 8) significantly improved the estimation accuracy (Fig. 4). This reduction is evidenced by the accuracy of 0.261 and 0.145 for IFE5 and IFE12, respectively. The F-test between IFE5 and IFE12 to compare the variance of accuracies yielded a p-value of 1.176e−12, suggesting that using all 12 standard samples for the standard curve (IFE12) is the better approach. When employing the Bayesian estimator (Bayes), the accuracy is 0.127 μg/L (Fig. 4). This value is smaller than that of the inverse function estimator (IFE12), a result of using the weakly informative prior on all unknown concentrations, which prevented extreme values of the estimated concentrations. As such, we expect that the Bayes estimator will consistently outperform the inverse-function estimator. By implementing the BHM approach to utilize relevant information, accuracy was further improved (Fig. 4). Leveraging multiple water samples within a test only (BHM1) resulted in an accuracy of 0.114 μg/L, while leveraging multiple calibration curves to improve estimation accuracy of calibration curve coefficients only (BHM2) resulted in an accuracy of 0.121. Leveraging both within and across tests (BHM3) resulted in an accuracy of 0.109, which is similar to BHM1. The similarity between the accuracy of BHM3 and BHM1 is because of the high across-test variance (Fig. 1), as the across-test overall mean is downweighed, thereby the results are dominated by the within-test level (Gelman et al., 2014). In other words, within-test variation dominates, and across-test pooling contributes little due to heterogeneity in calibration curve shapes. Ultimately, BHM3 has the greatest accuracy followed by BHM1, BHM2, Bayes, and IFE12, respectively (Fig. 5).

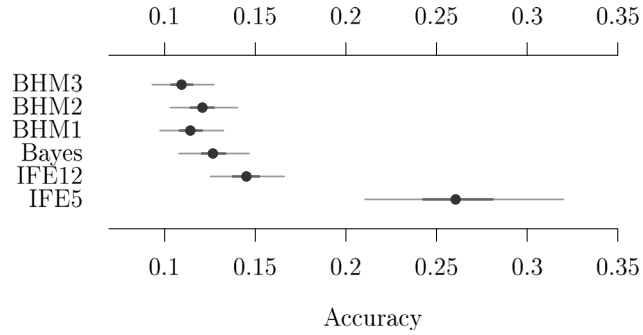

Fig. 4. Estimation accuracy of the quality control sample from the six methods (listed on the vertical axis) are compared. The accuracy is the absolute value of the difference between the estimated and the true concentration value (0.75 μg/L), labeled on the horizontal axis. The solid circles are the medians, the black (thick) lines are the 50% credible intervals, and the gray (thin) lines are the 95% credible intervals. The smaller the accuracy, the better.

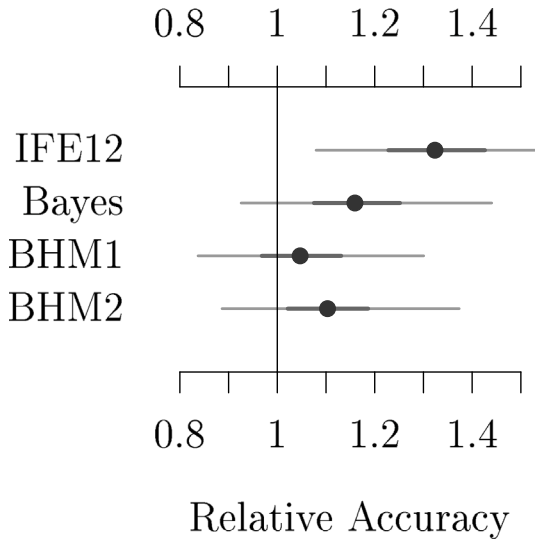

Fig. 5. The relative accuracy of the four models that used all 12 standard samples for estimating the curve (listed on the vertical axis) are compared to the best performing model (BHM3). BHM3 is represented as the vertical line with all other models being compared to it. The solid circles are the medians, the black (thick) lines are the 50% credible intervals, and the gray (thin) lines are the 95% credible intervals. The closer the relative accuracy of the four models to 1, the closer the model's accuracy is to the best performing model (BHM3).

4. Discussion

Given the importance of effectively monitoring HABs, our study demonstrates that the estimation accuracy of MC concentrations can be greatly improved by implementing a BHM approach. The four Bayesian estimators produced accuracy distributions with comparable and low variances, indicating improved consistency (Figs. 4 and 5). In contrast, the standard inverse estimator (IFE5) has much larger variances (Fig.

4). The inverse estimator using 12 data points (IFE12) offers improved accuracy and variance compared to IFE5, but does not perform as well as the Bayesian estimators. In short, Bayesian estimators demonstrate improved accuracy and greater consistency compared to the inverse-function estimators, with the within-test (BHM1) and two-level BHM (BHM3) having the largest improvement. The city of Toledo's "do not drink" advisory in 2014 was executed based on the MC concentration of one water sample. However, this estimated concentration was later shown to have high uncertainty (Qian et al., 2015). Had this information about the concentration uncertainty been available at the time, the next step may have been to redo the analysis before making a decision. Our BHM approach reduces the estimation uncertainty which would be beneficial for real-world circumstances such as those like the Toledo water crisis. Because calibration-based methods are commonly used in analytical chemistry (Miller and Miller, 2010), this approach can be used for a wide range of calibration-based problems (Qian et al., 2025).

The sequential updating algorithm is necessary for implementing the BHM approach in a typical lab setting. Without sequential updating, new tests need to be analyzed with previous data. This can be computationally burdensome as the number of tests increases. With sequential updating, the algorithm stores the prior information accumulated from the previous tests which allows for new tests to be analyzed one at a time. Computationally, analyzing one test would be a quick process. As tests are constantly improving, information from older tests can be irrelevant. Accordingly, we can modify the sequential updating process to allow systematic discounting of information from older tests by inflating the variance (σ^2_θ, Eq. (4)) from the previous iteration, similar to the discount factor used in Bayesian time series analysis (West and Harrison, 1997). For example, Qian and Reckhow (2007) suggested that "we can set the sum of prior parameters α_θ and β_θ to be similar to the data sample size if we want to give the prior information a weight similar to that given to the data". Both the BHM and sequential updating algorithm can be easily automated through the development of a computer application such as a Shiny app. A Shiny app would allow non-technical users to use these methods without needing the statistical or coding knowledge to run the analyses in its current form. This would allow for these methods to be easily integrated into the current lab setting where the only change to protocols would be to replace the currently used Excel software with the new app.

5. Conclusions

We provided an alternative modeling approach that improves the measurement accuracy of algal toxin concentrations. Based on our results, we recommend always using all 12 standard measurements and applying shrinkage models when feasible. All of the models that used all 12 standard samples performed better than the model that used only five samples (IFE5). Furthermore, a shrinkage estimator should always be considered as a way to improve estimation accuracy. Even the across-test only hierarchical model (BHM2), which was the weaker of the three hierarchical models, provided improvements compared to the non-hierarchical models. This is because pooling data across tests provides guardrails for unusual curve results. Ultimately, when measuring toxins, it is always better to include stronger mathematical regulations against extreme results, which is what the BHM offers (Efron, 1975). Given that our approach does not require any changes to the current lab procedure, there would be no monetary costs to implement this method in a lab. Additionally, sequential updating allows labs to analyze new tests efficiently.

CRediT authorship contribution statement

Sabrina Jaffe: Writing – original draft, Methodology, Formal analysis, Conceptualization. **Duane Gossiaux:** Writing – review & editing, Data curation. **Reagan M. Errera:** Writing – review & editing, Data curation. **Emanuela Gionfriddo:** Writing – review & editing, Methodology, Funding acquisition, Data curation. **Song S. Qian:** Writing – original draft, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Song S. Qian reports financial support was provided by Ohio Sea Grant. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data and computer code (in R) is publicly available at

Lake Erie HABs monitoring data and R code (Original data) (GitHub)

References

Babica, P., Bláha, L., Maršálek, B., 2006. Exploring the natural role of microcystins—a review of effects on photoautotrophic organisms 1. J. Phycol. (ISSN: 0022-3646) 42 (1), 9–20.

Bernardo, J.M., Smith, A.F.M., 2009. Bayesian Theory. volume 405, John Wiley & Sons.

Bingham, M., Sinha, S.K., Lupi, F., 2015. Economic Benefits of Reducing Harmful Algal Blooms in Lake Erie. Technical report, Environmental Consulting and Technology, Inc., URL https://legacyfiles.ijc.org/tinymce/uploaded/Publications/Economic-Benefits-Due-to-Reduction-in-HABs-October-2015.pdf.

Biosensis, 2024. Technical Note 2, Data Analysis for Elisa Assays. Technical report, biosensis, URL https://www.biosensis.com/documents/enhancedinfo/Technical-Note-2-ELISA-Data-Analysis.pdf.

Boegehold, A.G., Burtner, A.M., Camilleri, A.C., Carter, G., DenUyl, P., Fanslow, D., Fyffe Semenyuk, D., Godwin, C.M., Gossiaux, D., Johengen, T.H., 2023. Routine monitoring of western lake erie to track water quality changes associated with cyanobacterial harmful algal blooms. Earth Syst. Sci. Data Discuss. (ISSN: 1866-3591) 2023, 1–39.

Box, G.E.P., Tiao, G.C., 1973. Bayesian Inference in Statistical Analysis. Addison-Wesley, Reading, MA.

Chorus, I., Welker, M., 2021. Toxic Cyanobacteria in Water–a Guide To their Public Health Consequences, Monitoring and Man-Agement: The Totally Revised Second Edition Is now Available. Taylor and Francis.

Core Team, R, 2022. R: A language and environment for statistical computing. R foundation for statistical computing, vienna, Austria. URL https://www.R-project.org/.

Efron, B., 1975. Biased versus unbiased estimation. Adv. Math. (ISSN: 0001-8708) 16 (3), 259–277.

Efron, B., 1996. Empirical bayes methods for combining likelihoods. J. Amer. Statist. Assoc. (ISSN: 0162-1459) 91 (434), 538–550.

Findlay, J.W., Dillard, R.F., 2007. Appropriate calibration curve fitting in ligand binding assays. AAPS J 9, 29.

Gan, N., Xiao, Y., Zhu, L., Wu, Z., Liu, J., Hu, C., Song, L., 2012. The role of microcystins in maintaining colonies of bloom-forming microcystis spp. Environ. Microbiol. (ISSN: 1462-2912) 14 (3), 730–742.

Gelman, A., Carlin, J.B., Stern, H.S., Dunson, David B., Vehtari, Aki, Rubin, D.B., 2014. Bayesian Data Analysis, third ed. CRC Press, Boca Raton, Florida.

Gelman, A., Chew, G.L., Shnaidman, M., 2004. Bayesian analysis of serial dilution assays. Biometrics (ISSN: 0006-341X) 60 (2), 407–417.

Gelman, A., Jakulin, A., Pittau, M.G., Su, Y.S., 2008. A weakly informative default prior distribution for logistic and other regression models. Ann. Appl. Stat. 2 (4), 1360–1383.

Huisman, J., Codd, G.A., Paerl, H.W., Ibelings, B.W., Verspagen, J.M.H., Visser, P.M., 2018. Cyanobacterial blooms. Nat. Rev. Microbiol. (ISSN: 1740-1526) 16 (8), 471–483.

Klauenberg, K., Walzel, M., Ebert, B., Elster, C., 2015. Informative prior distributions for elisa analyses. Biostatistics (ISSN: 1465-4644) 16 (3), 454–464, 01.

Miller, J.N., Miller, J.C., 2010. Statistics and Chemometrics for Analytical Chemistry, Sixth ed. Pearson/Prentice Hall, Essex, UK.

Nummer, S.A., Weeden, A.J., Shaw, C., Snyder, B.K., Bridgeman, T.B., Qian, S.S., 2018. Updating the elisa standard curve fitting process to reduce uncertainty in estimated microcystin concentrations. MethodsX (ISSN: 2215-0161) 5, 304–311.

Qian, S.S., 2016. Environmental and Ecological Statistics with R. Chapman and Hall/CRC.

Qian, S.S., Chaffin, J.D., DuFour, M.R., Sherman, J.J., Golnick, P.C., Collier, C.D., Nummer, S.A., M.G., Margida, 2015. Quantifying and reducing uncertainty in estimated microcystin concentrations from the ELISA method. Environ. Sci. Tech. 49, 14221–14229.

Qian, S.S., Jaffe, S., Gionfriddo, E., Wang, H., Richardson, C.J., Godage, N.H., 2025. Rethinking calibration as a statistical estimation problem to improve measurement accuracy. Analitica Chimica Acta, in review URL www.github.com/songsqian/calibration.

Qian, S.S., Reckhow, K.H., 2007. Combining model results and monitoring data for water quality assessment. Environ. Sci. Tech. 41, 5008–5013.

Stan Development Team, 2022a. Rstan: the r interface to stan, version 2.21.5. URL http://mc-stan.org/rstan.html.

Stan Development Team, 2022b. Stan modeling language user's guide and reference manual, version 2.30. URL http://mc-stan.org/manual.html.

Stein, C., 1956. Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. In: Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability. vol. 1, Berkeley, California, pp. 197–206, University of California Press.

Svirčev, Z., Drobac, D., Tokodi, N., Mijović, B., Codd, G.A., Meriluoto, J., 2017. Toxicology of microcystins with reference to cases of human intoxications and epidemiological investigations of exposures to cyanobacteria and cyanotoxins. Arch. Toxicol. 91 (0340–5761), 621–650.

Toledo Department of Public Utilities, 2014. Microcystin event preliminary summary. http://goo.gl/m4m0ky.

West, M., Harrison, J., 1997. Bayesian Forecasting and Dynamic Models, second ed. Springer.