Improving Water Quality Assessments through a Hierarchical Bayesian Analysis of Variability

ANDREW D. GRONEWOLD^{*,†} AND MARK E. BORSUK[‡]

NOAA, Great Lakes Environmental Research Laboratory, Ann Arbor, Michigan 48108, and Thayer School of Engineering, Dartmouth College, Hanover, New Hampshire 03755-8000

Received March 2, 2010. Revised manuscript received August 23, 2010. Accepted August 30, 2010.

Water quality measurement error and variability, while welldocumented in laboratory-scale studies, is rarely acknowledged or explicitly resolved in most model-based water body assessments, including those conducted in compliance with the United States Environmental Protection Agency (USEPA) Total Maximum Daily Load (TMDL) program. Consequently, proposed pollutant loading reductions in TMDLs and similar water quality management programs may be biased, resulting in either slower-than-expected rates of water quality restoration and designated use reinstatement or, in some cases, overly conservative management decisions. To address this problem, we present a hierarchical Bayesian approach for relating actual in situ or model-predicted pollutant concentrations to multiple sampling and analysis procedures, each with distinct sources of variability. We apply this method to recently approved TMDLs to investigate whether appropriate accounting for measurement error and variability will lead to different management decisions. We find that required pollutant loading reductions may in fact vary depending not only on how measurement variability is addressed but also on which water quality analysis procedure is used to assess standard compliance. As a general strategy, our Bayesian approach to quantifying variability may represent an alternative to the common practice of addressing all forms of uncertainty through an arbitrary margin of safety (MOS).

Introduction

Uncertainty and the TMDL Assessment Process. The United States Environmental Protection Agency (USEPA) Total Maximum Daily Load (TMDL) program is the nation's most comprehensive and far-reaching program governing protection and improvement of surface water quality (1-3). The TMDL program requires that states identify waters failing to meet water quality standards and then determine the maximum allowable pollutant load that can enter such waters and yet meet applicable water quality standards (4). Although the TMDL program was adopted into United States policy as part of the 1972 Amendments to the Federal Water Pollution Control Act (commonly referred to as the Clean Water Act), TMDL assessments were not completed on a

[‡] Dartmouth College.

large scale until the late 1990s (5). Since then, the number of TMDLs addressed by USEPA has increased almost every year (6).

The increased rate of TMDLs addressed by USEPA over the past ten years follows several federal initiatives, including development of final rules for the TMDL program and the publication of a National Research Council (NRC) report assessing the scientific basis for the program (1). The rate of TMDL assessments has also increased due to widespread application of computer-based mechanistic and empirical water quality models which link pollutant loading rates to receiving water body quality (7-9). Two frequently proposed improvements to TMDL assessment methodology, however, are noticeably absent in most, if not all, approved TMDLs. First, the NRC (1) recommended that TMDLs assess water quality impairment using innovative statistical methods of the type that have since been proposed by Borsuk et al., Gibbons, and Gronewold et al. (10-12). Second, the NRC (1), along with Reckhow and Shirmohammadi et al. (13, 14), recommended that TMDL models explicitly acknowledge uncertainty and discontinue the application of arbitrarily selected margins of safety (MOS). Despite these recommendations, uncertainty in TMDL assessments continues to be addressed primarily through an arbitrary MOS, based in most cases on a fixed percentage (e.g., 10%) of the recommended TMDL. Robust statistical approaches recommended by the NRC (such as probabilistic or Bayesian models) are rarely applied (1, 15).

One important source of uncertainty in water quality assessments is measurement variability, defined here as the extent to which observed values differ from true values of the quantity being assessed. Such measurement variability may result from spatial and temporal variability of the environment or variable and erroneous analytical methods. The goals of this paper are to demonstrate a Bayesian hierarchical approach for explicitly addressing measurement error and variability and to investigate whether adoption of this approach would change the pollutant loading reductions that have been recommended in some recent TMDL assessments.

Pathogen (Bacterial) TMDLs as an Example. Of the roughly 27,000 TMDLs approved by USEPA since October 1995, 5600 are pathogen TMDLs, the highest number of any single pollutant category. Despite their name, however, pathogen TMDLs typically assess compliance with water quality standards based on the concentration of nonpathogenic fecal indicator bacteria (FIB) including fecal coliforms, *Escherichia coli*, and *Enterococcus* species (*16–18*). Because they represent such a large portion of current and proposed TMDL assessments, and because of their multifaceted sources of error and variability, we focus here on pathogen TMDLs as an example.

As with most types of TMDL assessments, there is a broad range of factors which may contribute to discrepancies between modeled and observed water quality conditions in pathogen TMDLs. Intrinsic sources of variability, such as spatial or temporal variability in the environment or in a sample aliquot, are often quite significant in common measures of FIB concentration, such as the most probable number (MPN) and colony forming unit (CFU). Both the MPN and CFU can have relatively broad and potentially biased probability distributions (depending on characteristics of the analytical procedure, such as sample aliquot volume), and the MPN is known to be a biased estimator for any given *in situ* concentration (19-22). Therefore, relating model-based FIB concentration predictions to the

^{*} Corresponding author phone: (734)741-2444; fax: (734)741-2055; e-mail: drew.gronewold@noaa.gov.

[†] NOAA. [‡] Dartmouth Coll

equivalent sampling distribution of MPN or CFU values requires consideration of not only spatial and temporal sampling patterns but also the water quality analysis method. We find that pathogen TMDLs rarely account for these sources of variability.

Using a hierarchical Bayesian approach, MPN and CFU measurements can be graphically and probabilistically related to the statistics on which standard compliance assessments are based as well as to FIB concentration parameters. This allows for assessments that proceed seamlessly from field data (possibly collected using multiple methods) to integrated inferences of *in situ* FIB concentrations or from model-based concentration predictions to probabilistic standard compliance evaluations - all while appropriately accounting for measurement error and variability.

We demonstrate the hierarchical Bayesian approach by applying it to the assessment of two tidal embayments in Chesapeake Bay determined to be out of compliance with water quality standards for shellfish harvesting areas. We are interested in investigating whether an explicit and detailed accounting for variability in estimates of the in situ FIB concentration might affect the pollutant loading reduction recommendations made as part of the TMDL program. We explore this issue by generating Bayesian posterior probability distributions of the quantities used to assess standard compliance in these two water bodies, under current conditions as well as under conditions expected to result from the recently approved TMDLs for these water bodies. Finally, in a novel set of calculations, we infer the full probability distribution of the reduction in the fecal coliform loading rate which, given intrinsic variability in both the in situ concentration and water quality analysis procedures, will lead to full compliance with standards.

Methods

Case Study: Oyster and Mosquito Creeks, Virginia. Oyster and Mosquito Creeks are two tidal embayments along the western shore of Chesapeake Bay in Lancaster County, Virginia. They have been found to be out of compliance with state water quality standards (see Table S1 in the Supporting Information for a summary of pertinent standards), and TMDLs were recently developed and approved (23). We selected these two embayments for our example not only because they address the widespread problem of fecal contamination in coastal shellfish harvesting waters but also because (along with many other FIB-based TMDLs in coastal waters) their TMDLs were developed using the simple volumetric model. This model (also referred to as the modified tidal prism model) provides an ideal contrast to our proposed approach because it is widely used, it is less complex than other TMDL-support models including (among others) the well-known HSPF (Hydrological Simulation Program-Fortran) and WARMF (Watershed Analysis Risk Management Framework) models (24, 25), and it allows us to explore how different approaches to quantifying intrinsic water quality measurement variability alone might lead to conflicting and, perhaps, misleading TMDL loading reduction recommendations.

Summary of Conventional TMDL Model and Assessment Results. In the conventional tidal volumetric models used to develop the Oyster Creek and Mosquito Creek TMDLs, water quality is first characterized by the maximum rolling 30-sample geometric mean (μ_{geo}^m) and maximum rolling 30sample 90th percentile (q_{90}^m) of monthly MPN and CFU values (in organisms per 100 mL) collected roughly once per month between May 1995 and July 2008 in Oyster Creek (resulting in 158 samples for this assessment) and between June 1998 and July 2008 in Mosquito Creek (resulting in 122 samples for this assessment). The use of rolling statistics in the conventional approach, rather than data from mutually exclusive time periods, is a precautionary step (implicitly representing an MOS) that results in conservative assessments relative to NSSP (National Shellfish Sanitation Program)-based standards (*26*). In the simple tidal volumetric modeling paradigm, μ_{geo}^m and q_{90}^m are next multiplied by the average estuary volume to obtain what are commonly interpreted as corresponding statistics of the FIB loading rate. Recommended loading rate reductions are then estimated based on the fractional reduction in the loading rate statistics which bring μ_{geo}^m and q_{90}^m into compliance with the water quality standard for μ_{geo} and q_{90} . Summary statistics and current loading reduction requirements from the TMDL reports for Oyster and Mosquito Creeks (*23*) are presented in Table S2 in the Supporting Information.

By basing loading reduction estimates on sample-based characteristics, the conventional tidal volumetric model (like most FIB TMDL-support models) implicitly equates in situ FIB concentrations with the MPN and CFU values on which water quality standards are based. However, MPN and CFU values have significant and distinct sources of bias and variability relative to in situ concentrations (22, 27). Put differently, for a given set of conditions, a probability distribution of measured MPN values will very likely differ from the probability distribution for CFU values as well as from the true in situ FIB concentration distribution. While this is apparent to most water resource management agencies, few assessments formally acknowledge this discrepancy. In fact, it is notable that, from 1984 to 2007, the Virginia Department of Environmental Quality (VDEQ) used the 3-tube ×3 series multiple tube fermentation (MTF) procedure (28-31) to estimate MPN values and then in 2007 switched to using standard membrane filtration (MF) procedures (20) to derive CFU values. The measurements, although made with different procedures, are currently treated as equivalent for the purposes of TMDL determination. Our approach will allow us to directly test this assumption.

Proposed Bayesian Approach. We propose a Bayesian procedure for inferring both current conditions and the FIB loading reduction required to bring a waterbody into compliance with standards. This procedure allows variability in estimates of the *in situ* FIB concentration arising from single or multiple analysis procedures to be reflected explicitly in the underyling FIB concentration assessment. The method is based on the recognition that, within the framework of the simplified tidal prism model, MPN and CFU values (along with the associated statistics which serve as the basis for water quality standards) are related to FIB loading rates through an implied, but rarely acknowledged, hierarchical relationship. This hierarchy (presented graphically in Figure 1) begins with the observation that variability and uncertainty in the *in situ* FIB concentration at a monitoring site can be represented by assuming that concentration values c (in organisms per 100 mL) arise from a log-normal $LN(c \mid \mu, \sigma)$ probability distribution with log-concentration mean μ and log-concentration standard deviation σ , both in log-organisms per 100 mL (32, 33).

The relationship between the *in situ* FIB concentration *c* and corresponding MPN and CFU values can then be expressed through a series of commonly employed and empirically validated statistical models (for details, see the Supporting Information and refs *19*, *22*, *28*, and *33*–*36*). By "reversing" the logic of these statistical models (*12*, *22*), and of the hierarchical relationship shown in Figure 1, we use the pattern of positive tubes from MTF measurements, or the number of colonies counted on a growth plate in an MF procedure, to infer the Bayesian posterior distribution of the underlying *in situ* FIB *c* as well as the parameters μ and σ . This approach represents a logical and defensible framework for combining FIB concentration analysis results from a variety of different testing procedures (each with distinct



FIGURE 1. Graphical representation of two approaches to estimating the FIB loading reductions required to achieve standard compliance in Oyster and Mosquito Creeks. Ovals represent variables and parameters, while rectangles represent constants. Shaded nodes represent measured quantities. The conventional approach (bounded by dashed lines) is based on maximum rolling 30-sample statistics (μ_{geo}^m , q_{g0}^m) of combined MPN and CFU values. Our proposed approach (bounded by solid lines) is based on inferring the FIB log-concentration mean (μ) and standard deviation (σ) from "raw" data on the total number (n), volume (v_1 , v_2 , v_3), and number of *positive* tubes (t_1 , t_2 , t_3) in each dilution series of MTF experiments as well as the number of colonies counted on a growth plate (y) and volume filtered (V) in MF experiments (for additional details, see the Supporting Information).

sources of variability) into a comprehensive assessment of uncertainty in the *in situ* FIB concentration. We know of no other water quality assessment procedure which provides a comparable approach to resolving this issue.

Using these inferred parameter values, we follow the logic of Figure 1 (and eqs 1 and 2 in the Supporting Information) to simulate 1000 size-30 FIB concentration samples and associated MPN and CFU values. For each size-30 sample we calculate the geometric mean (μ_{geo}) and 90th percentile (q_{90}), leading to 1000 samples from the posterior distribution for each of these quantities.

Finally, we infer the required reduction (η) in the FIB organism loading rate that achieves compliance with water quality standards by calculating the reduction in the FIB log-concentration mean μ that will bring the distribution of c into a range that is compliant with current MPN- and CFU-based standards. In this way, the posterior distribution of η will account for uncertainty at all levels of the assessment procedure and provide an alternative to arbitrary selection of an MOS in the TMDL. Following ref *12*, we define the compliance range as an inequality in terms of μ and σ (see the Supporting Information for details)

$$\mu \le \frac{2.34 - \sigma^{1.30}}{0.88} \tag{1}$$

We inferred posterior probability distributions of η and other unknown parameters for each creek by encoding the

hierarchical relationship (Figure 1) in the Bayesian analysis software package WinBUGS (37) using diffuse prior probability distributions for each model parameter (38, 39). Data consisted of the pattern of positive tubes from 3 tube \times 3 series MTF procedures *and* the number of colonies counted on each growth plate in MF procedures conducted between May 1995 and July 2008 in Oyster Creek and between June 1998 and July 2008 in Mosquito Creek. We confirmed the assumed serial independence of each time series through visual inspection of the autocorrelation function. Code and additional details of our WinBUGS procedure are included in the Supporting Information.

Results

The posterior distributions of the MPN and CFU geometric means and 90th percentiles are shown for Mosquito Creek in Figure 2 and, for Oyster Creek, in Figure S1 (in the Supporting Information). These represent the uncertainty in the actual values of these parameters as inferred from all available raw data from both MTF and MF procedures. We find that the observed maximum rolling 30-sample statistics (thick black vertical line in each figure) are indeed conservative, based on the fact that they lie at the upper tails of the posterior distributions. However, their degree of conservatism is inconsistent between the two creeks; the sample statistics are located much further out in the tails in the posterior distributions for Oyster Creek than for Mosquito Creek.

Mosquito Creek (existing conditions)



FIGURE 2. Histograms of simulated geometric mean (left-hand column) and 90th percentile (right-hand column) values from a size-30 sample of MPN (top row) or CFU (bottom row) values based on historic water quality data from Mosquito Creek, Virginia. Vertical black lines indicate maximum rolling 30-sample values from the Mosquito Creek TMDL assessment (see Table S2, Supporting Information), and vertical gray lines indicate NSSP-recommended numeric water quality standards (see Table S1, Supporting Information).

The posterior distributions shown in Figure 2 and Figure S1 also confirm that Oyster Creek and Mosquito Creek are both highly unlikely to comply with NSSP-recommended numeric standards (thick vertical gray line in each figure) under existing conditions. This is evidenced by the fact that the creeks have posterior distributions for the 90th percentile that have a significant probability of exceeding the relevant standards (measured as the proportion of the histogram to the right of the vertical gray line). In fact, Mosquito Creek has a high probability of exceeding the geometric mean standard as well.

Additionally, our results indicate that the 30-sample CFU geometric mean (lower left-hand panel of Figure 2 and Figure S1) is negatively biased relative to the 30-sample MPN geometric mean (upper left-hand panel in each figure), yet the corresponding water quality standard for the geometric mean is 14 organisms per 100 mL regardless of whether CFU or MPN values are used in the assessment (see Table S1, Supporting Information). There also appears to be a somewhat less significant bias in the CFU 90th percentile relative to that of the MPN (right-hand panels of each figure). While the water quality standards for the 90th percentile are adjusted to each testing procedure (Table S1, Supporting Information), they still appear to lead to quantitatively different assessments.

Our hierarchical Bayesian analysis integrates raw data from both MTF and MF procedures to yield posterior distributions on the underlying FIB concentration parameters μ and σ (Figure 3). These confirm that Oyster Creek and Mosquito Creek are both highly unlikely to comply with water quality standards under existing conditions (Figure 3, left column) with none of the joint distribution lying in the compliance range for either water body. These plots also indicate that Mosquito Creek is further from compliance than Oyster Creek, since the joint probability distribution for Mosquito Creek falls further to the right of the compliance line than the probability distribution for Oyster Creek.

The loading reductions recommended by the VDEQ translate into a shift of the joint distribution of μ and σ to the left for each creek. However, even after these reductions,

only approximately 5% of each distribution lies in the compliance region of the graph (Figure 3, right column). Therefore, we can say that the *confidence of compliance* (CC) with relevant standards after the currently approved loading reduction is approximately 5% (*10, 12, 21*).

Our inferred distributions of the parameter η (Figure 4) translate the joint distributions shown in Figure 3 (left column) into the FIB loading reduction required to comply with standards. The proportions of the histograms to the left of the gray vertical lines in Figure 4 indicate the CC associated with the current VDEQ recommended loading reductions. Consistent with Figure 3, at both locations there is only approximately a 5% CC attached to these reductions. To achieve a CC of 50% or greater, a reduction upward of 60% is required at Oyster Creek. Similarly, at Mosquito Creek, a reduction greater than 80% is necessary to achieve greater confidence in compliance than in noncompliance.

Discussion

Our Bayesian hierarchical approach acknowledges that MPN and CFU values are distinctly different measures of the actual in situ FIB concentration and therefore explicitly accounts for the variability and biases of each. This is important because many water quality assessments continue to be based on long (e.g., 3-4 year) historic records, which increasingly include both MPN and CFU values and, in some cases, MPN values derived from differing serial dilution analysis-based procedures. It is erroneous to consider these methods as equivalent, yet the older data should not be ignored. The approach we introduce here has the distinct advantage of being able to accommodate all FIB concentration analysis procedures simultaneously because they are each represented mathematically by an appropriate likelihood function for the "true" FIB concentration c. Each likelihood function reflects the sources of variability and uncertainty inherent to each procedure. This means that our method also has the advantage of being able to easily incorporate novel microbial analysis procedures that have recently been developed and may soon be employed in water



FIGURE 3. Contour plots of the joint posterior probability density function of fecal coliform log-concentration mean (μ) and standard deviation (σ), both in log-organisms per 100 mL, for Oyster and Mosquito Creeks under existing conditions (left column) and approved loading reduction (right column) as described in ref 23. The diagonal dashed line in each panel represents the compliance boundary: combinations of μ and σ to the left violate NSSP CFU-based standards with less than 0.5% probability. Confidences of compliance (CC) with such standards are given in the upper right of each panel.



FIGURE 4. Histograms of samples from the posterior probability distribution for η (the required fractional FIB loading reduction) for Oyster and Mosquito Creeks. Vertical gray lines indicate the loading reduction required based on the VDEQ TMDL report (23).

quality assessment programs (40-42). It should also be noted, however, that the proper exposition of the likelihood function for the true FIB concentration requires "raw" data on the actual pattern of positive tubes (in an MPN experiment) or the number of colonies counted on a growth plate and the dilution factor (in a CFU experiment). Unfortunately, these raw data may not always be available for performing retrospective analyses (43).

When using our method to compare Bayesian posterior distributions of sample-based statistics μ_{geo} and q_{90} against rolling statistics and NSSP-recommended standards at the two case study locations, Oyster and Mosquito Creeks, we found that there were significant inconsistencies (Figure 2 and Figure S1). Point values based on maximum rolling 30-

sample statistics showed a varying degree of conservatism relative to the full posterior probability distributions. As the degree of conservatism associated with the use of such rolling statistics is not typically quantified, the loading reduction requirements implied by these statistics (Table S2, Supporting Information) may be either too conservative or too tolerant. Additionally, the determined loading reduction will depend critically on which metric (CFU or MPN) is used in the analysis. Although the standards differ slightly between the two methods, the difference is apparently not sufficient, as evidenced by the difference in probability of a violation at Oyster and Mosquito Creeks when CFU, rather than MPNbased, standards are assumed (see Figure 2). This is consistent with the general relations between the various standards noted by ref 12. The fact that TMDL implementation success may depend on which analytical procedure is used to assess future compliance has not, to our knowledge, been ac-knowledged in any TMDL assessments.

The TMDLs developed by VDEQ for Oyster and Mosquito Creeks are based on determinations of the difference between the observed maximum rolling 30-sample values of the geometric mean (μ_{geo}^m) and 90th percentile (q_{90}^m) and the NSSPrecommended standards regarding these statistics. This approach has two drawbacks. First, it assumes that the required reduction should be based entirely on the statistic that is furthest from compliance. Second, it assumes that the reduction required to *exactly* bring that single statistic into compliance with standards would provide just the right degree of confidence that the water body actually will be in compliance in future monitoring. The problem with these assumptions is that the full range of the FIB concentration distribution is unlikely to respond homogeneously to a reduction in FIB loading, so that barely achieving compliance with one uncertain statistic such as the maximum rolling value of q_{90} may or may not achieve adequate compliance for either that statistic itself or for the other statistic being monitored.

In addition to providing a framework for integrating FIB concentration measures from multiple and variable analytical methods, we feel that our inferred concentrationbased analysis (represented by Figure 3) addresses the two problems with common practice described in the previous paragraph. This is because the full FIB concentration distribution as well as the uncertainty in its location and spread are represented by the joint posterior distribution of μ and σ . Assumptions about how the characteristics of this distribution are likely to change as a result of loading reductions can then be made explicit, with the results "tracked" graphically in the (μ, σ) parameter space. In our analysis, we assumed that loading reductions would shift the log-concentration mean μ to lower values, while the logconcentration standard deviation σ would remain constant. This is reasonable if we consider the variability in FIB concentration to be the result of multiplicative, rather than additive, departures from the expected value. We have confirmed this assumption in other settings experiencing pollutant load reductions (44). Certainly, other assumptions could be made and defended on a case-specific basis.

Once current conditions are located as a joint distribution in the (μ, σ) parameter space (Figure 3, left panels) and an assumption is made regarding the effect of a loading reduction on μ and/or σ , then a distribution for the reduction required to achieve compliance (which we refer to as η) can be directly inferred (Figure 4). The cumulative probability of values of η following this distribution represents the associated confidence we can have that the value of η will lead to compliance with applicable standards, i.e., the *confidence of compliance*, or CC. A CC assessment acknowledges that, in an impacted natural system, it is not possible to ensure that samples will *never* be found to be in violation with water quality standards (i.e., there is always some nonzero probability of noncompliance) (10, 12, 21).

Directly addressing the question of "How confident do we need to be?" that a water body will be in compliance with standards is a practical and rational alternative to arbitrarily adding an MOS to TMDL determinations. In fact, for Oyster and Mosquito Creeks, we found that approved TMDLs yield only about a 5% CC, even with an implicit MOS (through the use of maximum rolling statistics). This suggests that the MOS often employed in TMDLs is not adequate to address the typically unrecognized sources of error and variability we acknowledge here, let alone more widely acknowledged uncertainties, such as those related to pollutant source, transport, and fate processes. The question then becomes the following: "If a TMDL assessment were to adopt the specific procedure we propose, what additional MOS would be necessary to account for these other uncertainties?". Of course, our suggestion would be to employ a probabilistic modeling approach to modeling these other processes in a manner similar to the one we describe here. Some recent examples in the literature include refs *10, 11,* and *45.*

We recognize that most water quality models used to support pathogen TMDLs currently yield only deterministic predictions of pollutant concentrations rather than probabilistic assessments that account for sampling variability. We also realize that the probabilistic models necessary to relate actual (or model-simulated) in situ concentrations to sample-based measures are relatively complex and specialized. Therefore, to encourage practical implementation, we have created a free software tool, ProVAsT-FIB (Probabilistic Violation Assessment Tool), which translates in situ FIB concentrations into the expected frequency of sample-based standard violations and provides an estimate of the confidence of compliance in the assessment (46). Users can input their own actual or model-simulated concentration values as well as specify the particular analytic methods employed and numeric limits that pertain to local water quality standards. It is our hope that such a tool will encourage rational consideration of measurement error and variability in water quality assessments by reducing the burden of complex statistical calculations.

Acknowledgments

The authors thank Jeffrey Chanat of the Virginia Department of Environmental Quality (VDEQ) and Daniel Powell of the Virginia Department of Health (Shellfish Sanitation Division) for providing valuable guidance on Virginia's water quality monitoring programs and for sharing their water quality monitoring data. The United States Environmental Protection Agency, through its Office of Research and Development, partially funded and collaborated in the research described here. It has been subjected to Agency review and approved for publication. This paper is GLERL contribution No. 1575. The comments of three anonymous reviewers helped improve the technical quality of this paper.

Supporting Information Available

Existing water quality standards, current TMDL assessment results, our hierarchical Bayesian analysis results for existing conditions in Oyster Creek, and the WinBUGS code used in our analysis. This material is available free of charge via the Internet at http://pubs.acs.org.

Literature Cited

- (1) National Research Council. *Assessing the TMDL approach to water quality management*; National Academy Press: Washington, DC, 2001.
- (2) Houck, O. A. The Clean Water Act TMDL program: Law, Policy, and Implementation; 2nd ed.;Environmental Law Institute: Washington, DC, 2002.
- (3) Cooter, W. S. Clean Water Act assessment processes in relation to changing U.S. Environmental Protection Agency management strategies. *Environ. Sci. Technol.* 2004, *38*, 5265–5273.
- (4) U.S. Environmental Protection Agency. National Water Quality Inventory: Report to Congress (2002 Reporting Cycle), EPA 841-R-07-001; 2002.
- (5) U.S. Environmental Protection Agency. Guidance for 2006 Assessment, Listing and Reporting Requirements Pursuant to Sections 303(d), 305(b) and 314 of the Clean Water Act; 2005.
- (6) U.S. Environmental Protection Agency. National Section 303(d) List Fact Sheet; 2007.
- (7) Chapra, S. C. Engineering water quality models and TMDLs. J. Water Resour. Plann. Manage.-ASCE 2003, 129, 247–256.
- (8) Benham, B. L.; Baffaut, C.; Zeckoski, R. W.; Mankin, K. R.; Pachepsky, Y. A.; Sadeghi, A. M.; Brannan, K. M.; Soupir, M. L.; Habersack, M. J. Modeling bacteria fate and transport in

watersheds to support TMDLs. Trans. ASABE 2006, 49, 987-1002.

- (9) Muñoz-Carpena, R.; Vellidis, G.; Shirmohammadi, A.; Wallender, W. Evaluation of modeling tools for TMDL development and implementation. *Trans. ASABE* **2006**, *49*, 961–965.
- (10) Borsuk, M. E.; Stow, C. A.; Reckhow, K. H. Predicting the frequency of water quality standard violations: A probabilistic approach for TMDL development. *Environ. Sci. Technol.* 2002, 36, 2109–2115.
- (11) Gibbons, R. D. A statistical approach for performing water quality impairment assessments. J. Am. Water Resour. Assoc. 2003, 39, 841–849.
- (12) Gronewold, A. D.; Borsuk, M. E.; Wolpert, R. L.; Reckhow, K. H. An assessment of fecal indicator bacteria-based water quality standards. *Environ. Sci. Technol.* **2008**, *42*, 4676–4682.
- (13) Reckhow, K. H. On the need for uncertainty assessment in TMDL modeling and implementation. J. Water Resour. Plann. Manage.-ASCE 2003, 129, 245–246.
- (14) Shirmohammadi, A.; Chaubey, I.; Harmel, R. D.; Bosch, D. D.; Muñoz-Carpena, R.; Dharmasri, C.; Sexton, A.; Arabi, M.; Wolfe, M. L.; Frankenberger, J.; Graff, C.; Sohrabi, T. M. Uncertainty in TMDL models. *Trans. ASABE* **2006**, *49*, 1033–1049.
- (15) Freedman, P. L. Navigating the TMDL Process: Evaluation and Improvements; Water Environment Research Foundation and IWA Publishing: 2003.
- (16) Noble, R. T.; Fuhrman, J. A. Virus decay and its causes in coastal waters. *Appl. Environ. Microbiol.* **1997**, *63*, 77–83.
- (17) U.S. Environmental Protection Agency. *Protocol for Developing Pathogen TMDLs*; Technical Report EPA 841-R-00-002; Office of Water (4503F), United States Environmental Protection Agency: Washington, DC, 2001.
- (18) Boehm, A. Enterococci concentrations in diverse coastal environments exhibit extreme variability. *Environ. Sci. Technol.* 2007, 41, 8227–8232.
- (19) El-Shaarawi, A. H.; Esterby, S. R.; Dutka, B. J. Bacterial density in water determined by Poisson or negative binomial distributions. *Appl. Environ. Microbiol.* **1981**, *41*, 107–116.
- (20) Dufour, A. P.; Strickland, E. R.; Cabelli, V. J. Membrane-filter method for enumerating Escherichia coli. *Appl. Environ. Microbiol.* **1981**, *41*, 1152–1158.
- (21) McBride, G. B.; Ellis, J. C. Confidence of compliance: a Bayesian approach for percentile standards. *Water Res.* 2001, 35, 1117– 1124.
- (22) Gronewold, A. D.; Wolpert, R. L. Modeling the relationship between most probable number (MPN) and colony-forming unit (CFU) estimates of fecal coliform concentration. *Water Res.* 2008, 42, 3327–3334.
- (23) Virginia Department of Environmental Quality. Oyster and Mosquito Creeks total maximum daily load (TMDL) report for shellfish condemnation areas listed due to bacteria contamination, 2009.
- (24) Donigian, A.; Imhoff, J.; Bicknell, B.; Kittle, J. Application guide for the hydrologic simulation program - FORTRAN1984.
- (25) Chen, C.; Herr, J.; Ziemelis, L. Watershed Analysis Risk Management Framework: A decision support system for watershed approach and total maximum daily load calculation; Report TR-110709; Electric Power Research Institute; Palo Alto, CA, 1998.
- (26) Food and Drug Administration and Interstate Shellfish Sanitation Conference. National Shellfish Sanitation Program - Guide for the Control of Molluscan Shellfish; 2005.

- (27) McBride, G. B. Preparing exact Most Probable Number (MPN) tables using occupancy theory, and accompanying measures of uncertainty. *NIWA Technical Report* **2003**, *121*, 62.
- (28) McCrady, M. H. The numerical interpretation of fermentation tube results. *J. Infect. Dis.* **1915**, *17*, 183–212.
- (29) Cochran, W. G. Estimation of bacterial densities by means of the 'most probable bumber'. *Biometrics* **1950**, *6*, 105–116.
- (30) Woodward, R. L. How probable is the most probable number. J. Am. Water Works Assoc. 1957, 49, 1060–1068.
- (31) Edberg, S. C.; Allen, M. J.; Smith, D. B. National field-evaluation of a defined substrate method for the simultaneous enumeration of total coliforms and escherichia coli from drinking water comparison with the standard multiple tube fermentation method. *Appl. Environ. Microbiol.* **1988**, *54*, 1595–1601.
- (32) Ott, W. *Environmental statistics and data analysis*; Lewis Publishers: Boca Raton, FL, 1995.
- (33) McBride, G. B. Using statistical methods for water quality management. Issues, problems and solutions; John Wiley & Sons Ltd.: Chichester, UK, 2005.
- (34) Eisenhart, C.; Wilson, P. W. Statistical Methods and Control in Bacteriology. *Bacteriol. Rev.* **1943**, *7*, 57–137.
- (35) Christian, R. R.; Pipes, W. O. Frequency distributions of coliforms in water distribution systems. *Appl. Environ. Microbiol.* **1983**, 45, 603–609.
- (36) Beliaeff, B.; Mary, J. Y. The most probable number estimate and its confidence-limits. *Water Res.* 1993, *27*, 799–805.
- (37) Lunn, D. J.; Thomas, A.; Best, N.; Spiegelhalter, D. WinBUGS-A Bayesian modelling framework: Concepts, structure, and extensibility. *Stat. Comput.* 2000, *10*, 325–337.
- (38) Gelman, A. Prior distributions for variance parameters in hierarchical models (Comment on Article by Browne and Draper). *Bayesian Anal.* 2006, 1, 515–534.
- (39) Gelman, A.; Hill, J. Data Analysis Using Regression and Multilevel/ Hierarchical Models; Cambridge University Press: New York, NY, 2007.
- (40) Boehm, A.; Ashbolt, N.; Colford, J. M.; Dunbar, L.; Fleming, L.; Gold, M.; Hansel, J.; Hunter, P.; Ichida, A.; McGee, C.; Soller, J.; Weisberg, S. A sea change ahead for recreational water quality criteria. *J. Water Health* **2009**, *7*, 9–20.
- (41) Griffith, J.; Cao, Y.; McGee, C.; Weisberg, S. Evaluation of rapid methods and novel indicators for assessing microbiological beach water quality. *Water Res.* 2009, *43*, 4900–4907.
- (42) Converse, R.; Blackwood, A.; Kirs, M.; Griffith, J.; Noble, R. Rapid QPCR-based assay for fecal Bacteroides spp. as a tool for assessing fecal contamination in recreational waters. *Water Res.* 2009, 43, 4828–4837.
- (43) Gronewold, A.; Qian, S.; Wolpert, R.; Reckhow, K. Calibrating and validating bacterial water quality models: A Bayesian approach. *Water Res.* **2009**, *43*, 2688–2698.
- (44) Stow, C. A.; Borsuk, M. E. Assessing TMDL effectiveness using flow-adjusted concentrations: A case study of the Neuse River, North Carolina. *Environ. Sci. Technol.* **2003**, *37*, 2043–2050.
- (45) Smith, E. P.; Ye, K. Y.; Hughes, C.; Shabman, L. A. Statistical assessment of violations of water quality standards under section 303(d) of the Clean Water Act. *Environ. Sci. Technol.* **2001**, *35*, 606–612.
- (46) Gronewold, A. D.; Borsuk, M. A software tool for translating deterministic model results into probabilistic assessments of water quality standard compliance. *Environ. Model. Soft.* 2009, 24, 1257–1262.

ES100657P