Review of the Chesapeake Bay Watershed Modeling Effort

By

Lawrence Band[1], Kenneth Campbell[2], Russell Kinerson[3],
Kenneth Reckhow[4], and Claire Welty[5]

[1]University of North Carolina, Chapel Hill, [2]University of Florida, [3]US Environmental Protection Agency, [4]Duke University, [5]University of Maryland, Baltimore County

STAC Publication 05-004

About the Scientific and Technical Advisory Committee

The Scientific and Technical Advisory Committee (STAC) provides scientific and technical guidance to the Chesapeake Bay Program on measures to restore and protect the Chesapeake Bay. As an advisory committee, STAC reports periodically to the Implementation Committee and annually to the Executive Council. Since it's creation in December 1984, STAC has worked to enhance scientific communication and outreach throughout the Chesapeake Bay watershed and beyond. STAC provides scientific and technical advice in various ways, including (1) technical reports and papers, (2) discussion groups, (3) assistance in organizing merit reviews of CBP programs and projects, (4) technical conferences and workshops, and (5) service by STAC members on CBP subcommittees and workgroups. In addition, STAC has the mechanisms in place that will allow STAC to hold meetings, workshops, and reviews in rapid response to CBP subcommittee and workgroup requests for scientific and technical input. This will allow STAC to provide the CBP subcommittees and workgroups with information and support needed as specific issues arise while working towards meeting the goals outlined in the *Chesapeake 2000* agreement. STAC also acts proactively to bring the most recent scientific information to the Bay Program and its partners. For additional information about STAC, please visit the STAC website at www.chesapeake.org/stac.

Publication Date:
June 2005

Publication Number:
05-004

Cover photo provided by the USGS.

To receive additional copies of this publication, contact STAC Staff at the Chesapeake Research Consortium and request the publication by title and number.

Chesapeake Research Consortium, Inc.
645 Contees Wharf Road
Edgewater, MD 21037
Telephone: 410-798-1283; 301-261-4500
Fax: 410-798-0816
http://www.chesapeake.org

Review of the Chesapeake Bay Watershed Modeling Effort

By

Lawrence Band[1], Kenneth Campbell[2], Russell Kinerson[3],
Kenneth Reckhow[4], and Claire Welty[5]

[1]University of North Carolina, Chapel Hill, [2]University of Florida, [3]US Environmental Protection Agency, [4]Duke University, [5]University of Maryland, Baltimore County

June 1, 2005

1. Introduction

Overview
In the spring of 2005 the Scientific and Technical Advisory Committee (STAC) of the Chesapeake Bay Program (CBP) recruited the authors as an independent panel of experts to review the Chesapeake Bay Watershed Model (CBWM) effort. The stated purpose of the review was to address the following broad questions:

(1) Does the current phase of the model use the most appropriate protocols for simulation of watershed processes and management impacts, based on the current state of the art in the HSPF model development?

(2) Looking forward to the future refinement of the model, where should the Bay Program look to increase the utility of the watershed model?

The authors met as a group on May 17 – 19, 2005 in Annapolis, MD. Handout materials were provided in advance and presentations were given to the review team by: Richard Batiuk, Gary Shenk, and Lewis Linker of the EPA Chesapeake Bay Program. The comments in this document summarize our assessment of work to date, and recommendations for future enhancements to the modeling effort.

It should be noted that in this review we have not seen any calibration or performance information for nutrient modeling for Phase 5, which is critical. We have limited information for Phase 4 which could be used by analogy, understanding that it was driven by different source loading as much as several years ago. **While the current Chesapeake Bay Model (CBM) may reproduce patterns of discharge and nutrient loads reasonably (although we have not received information on the latter), reproduction of nutrient concentrations is an important goal for diagnosing the model's performance.**

Current HSPF implementation and comparable programs
The CBWM team has done very good work in pulling together and integrating the range of information required to parameterize and operate the modeling system. Their activity is at the forefront and limits of the current technology available for this particular model

applied at the scale of the Chesapeake Bay. We point out that there are no templates for how this is best done. Watershed modeling for the scale and purposes envisioned by the CBP is the subject of considerable current research, while being recognized as a necessity for large-scale watershed management. The group has effectively partnered with other government and academic scientists to provide spatial data and GIS methods to aid in the parameterization and analysis of the model for the full Chesapeake Bay Watershed (CBW) and major tributaries. We commend the team for its work to date, and point out that our comments are geared towards refining methods and interpretation of the current CBWM and suggesting synergistic spatial data analysis and modeling approaches that can extend the utility of the current system with respect to the CBP goals.

While there are no templates for this effort, there are comparable projects with different models in the US and other countries. These include the modeling toolkit approach developed in Australia to simulate water, sediment, and water quality in large river basins (4000 – 150,000 km^2) (see http://www.catchment.crc.org.au), as well as a set of applications with SWAT and other large scale watershed models, that the CBWM team may wish to consider in a comparative mode.

Need for adaptive management framework

Based on previous experience with HSPF and other models of similar complexity and scope, prediction uncertainties may be large under certain conditions for some of the contaminants. In general, HSPF performs well for the simulation of river discharge, but is often less accurate for sediment and nutrient concentrations. Another way to state this point is that some predictions are likely to be wrong. **Given these circumstances, we recommend that assessments be adaptive; that is, "learning while doing" should occur during implementation of control measures (e.g., NRC, 2001).** This requires post-implementation monitoring (guided by the model) that might be used to assess compliance with the criterion, assess effectiveness of various BMPs, and suggest studies to improve the model. Risk of unanticipated outcomes can never be completely eliminated; this risk refers to both continued environmental degradation and/or excessive clean-up costs. As more knowledge is gained through monitoring/research, and this knowledge results in model (prediction) improvements, we can expect risk to be reduced. We believe that an adaptive implementation approach will most effectively lead to a reduction in risk and achievement (compliance) with program goals. **We emphasize that modeling and monitoring need to be effectively combined within this framework such that the modeling activity and results should be used to guide monitoring, while monitoring should be used to continuously test and refine the model structure and parameter sets.**

Need for formal uncertainty analysis

Prediction uncertainty can result from parameter uncertainty, model structural error, input errors, and unaccounted hydrologic variability. It is important that the current model be evaluated with respect to each source. Thus, the performance of the model should be specifically evaluated for hydrologic extremes (floods, droughts); in addition, seasonal effects should be assessed for wet versus dry conditions, and long-term trends in climate should be considered for assessment using the model. Other models should be run for comparison (e.g., SWAT), to assess model structural issues and parameter uncertainty should be evaluated with formal uncertainty analysis, making use of multiple realizations of the model parameter space. Another example may be the use of models such as SPARROW to evaluate different components of the system.

2. Calibration and parameter uncertainty Analysis

It has been suggested (e.g. Beven 2001) that large multi-parameter models are “overparameterized;” the result of this condition is that many “parameter sets” will lead to essentially equivalent good fits to the data. A particularly troubling aspect of this condition, called “equifinality,” is that individual parameters may vary greatly from one equivalently-fitting parameter set to another (since parameter covariance results in multi-parameter adjustments). As a consequence, one cannot be certain that the single parameter set chosen for the model on the basis of goodness-of-fit to a discharge time series correctly captures processes. **To address equifinality, and to estimate the impact of parameter uncertainty, we recommend that calibration results be presented as multiple parameter sets (all of which meet selected fitting criteria), and predictive application of the model involve Monte Carlo simulation (e.g. GLUE; generalized likelihood uncertainty estimation; Beven et al 2001, or other approaches) in order to produce a probabilistic range of feasible predictions.** This GLUE-based calibration (or similar approach) should reflect multiple system behaviors – from discharge and concentrations at the mouth to calibration at individual tributaries (to minimize compensating errors). In the current application, it is particularly important that model parameter sets be identified that can reproduce stream discharge, nutrient and sediment concentrations as well as their covariance structure. If available, additional internal state variables (e.g. soil moisture, groundwater levels) can be used as part of this procedure to further constrain the set of adequate parameter sets, and build confidence in the consistency of model predictions.

While calibration is typically based on goodness-of-fit of modeled and observed time series, **model performance evaluation should focus on a prediction/observation comparison using cumulative distribution functions (CDFs) instead of individual point-by-point fits.** The statistical distribution of outcomes is more important than fitting precise timing given uncertainty of exact loading of nutrient inputs, e.g., fertilizer application dates, sanitary system failures, spills, etc. In addition, regulatory instruments are typically geared toward exceedance frequencies.

The CDF allows the modeler to focus on capturing the magnitude and frequency of concentrations/loads. Also, the modeler should continue to check for biases in model prediction – for example, does the model tend to over/under predict for high/low flows, or particular basins, or particular seasons?

3. Integration of monitoring and modeling

Integration of monitoring and modeling is a critical activity to the future of the CBWM effort. **The model can be used as a design tool to select monitoring locations, times, and frequencies, and the model should evolve and be revised as monitoring information yields new insights for model process components.** One could think of the model as the null hypothesis and ultimately it could be rejected as monitoring yields new information. Note that in this instance we do not suggest the model should be completely discarded, but that rejection would indicate the need to modify the model structure based on monitoring- generated information.

In assessing overall compliance with water quality criteria, compliance in individual tributaries, or effectiveness of particular BMPs, the model that was used to make the initial (pre-implementation) prediction and the post-implementation monitoring data each have something useful to contribute. The monitoring data reflect the actual system response (but may be less useful due to system response lags, under-sampling and natural variability), while the model forecast directly predicts the impact of the change (yet may be hampered by large prediction errors).

We recommend that both the pre-implementation model predictions, and the post-implementation monitoring data, be pooled for these post-implementation assessments. Methods such as Bayesian analysis and data assimilation (Draper et al., 1992) exist to do this pooling. Further, the mathematical model is the quantitative framework relating pollutant sources/controls, forcing functions, reactions, etc. to system responses of interest. Therefore, the model should be the analytic framework guiding the post-implementation monitoring design (Reckhow, 1999).

There are on the order of 284 flow gauging stations, 120 TSS stations, and 100 nutrient stations that are currently being monitored. Flow is for the most part measured continuously in time; nutrient and sediment are characterized largely by quarterly (or other periodic) grab samples at locations that are not all the same as those of flow measurements. To the extent possible, it would make sense to co-locate the sediment/nutrient sampling with the stream gauge monitoring. If resources become available, it would be desirable to take advantage of emerging sensor technologies to monitor nutrients and sediments continuously in time at selected gauge locations. Insofar as new monitoring stations are concerned, it would make sense to use the model to determine where new stations could be located.

Our understanding is that in Version 4.3 the CBM made use of more limited nutrient concentration data, and that these data have been significantly expanded for Phase 5. . We support this expansion and encourage the CBM team to make use of additional nutrient concentration data that exists for a set of smaller, research catchments in the CBW. We recommend that in Phase 5, nutrient concentration data be integrated with the modeling both by being used in the calibration steps and in the verification steps, in addition to load information. Combining discharge and concentration data to progressively constrain feasible model parameter sets will provide greater confidence in process representation and load predictions in response to development or control scenarios.

4. Scaling from representative smaller basins to the CBW

At the scale of the full CBW it is necessary to develop methods of producing uniform fields of meteorological, land use/landcover, soils, topography, hydrography and other critical system drivers and the CBWM team has effectively pursued and refined these approaches However, at the full CBW these procedures introduce some degree of error as the information base is sparse at this scale and the input parameters are necessarily spatially generalized. In addition, at this scale the ability to relate values of modeled state variables (e.g. soil moisture, groundwater levels) to observed variables are limited. This results in the need to "guesstimate" specific parameters representing small-scale processes that are difficult to evaluate at the CBW or large tributary scale. **The error structure, including uncertainty analysis, of CBM predictions should be quantified/evaluated using selected smaller basin studies that are representative of the range of subbasins within the CBW and for which more detailed input and monitoring information and modeling studies are available.**

Finer-scale work in representative smaller basins within the CBW would be valuable in providing more detailed information for the CBWM, and for more precise diagnosis of the model's performance. In order to carry out more detailed modeling, monitoring data will be needed at appropriate scales. Other entities are already conducting monitoring at smaller scales that the CBP may be able to take advantage of. Examples of smaller watersheds within the CBW where dense monitoring instrumentation arrays are currently deployed include the Baltimore LTER, the USDA OPE3 site in Beltsville, Smithsonian Environmental Research Center sites, the University of Virginia's Shenandoah Watershed Study, and the Virginia Trout Stream Sensitivity Study. In cases where additional instrumentation or monitoring information may be required beyond what is already in place to generate desired model input, the CBP could consider coordinating with the forthcoming efforts on large-scale environmental observatories (CUAHSI: www.cuahsi.org, CLEANER: www.cleaner.org, NEON: www.neon.org) that may have resources available for instrumentation.

If new, additional subbasin studies are needed, and in light of resource constraints, the program may wish to consider reallocation of resources from modeling and monitoring to fewer representative smaller basins for the purpose of diagnosing model behavior, including internal state variables other than discharge and nutrient/sediment concentrations at gauges.

Distributed models and special purpose models can be applied at a small scale to generate an understanding of system dynamics, including critical parameters, to feed into the larger scale model. Examples of these applications might include use of ANSWERS or a similar model to determine sediment and nutrients erosion and transport from agricultural fields and related BMP efficiencies, RHESSys (Tague and Band 2004) to evaluate nutrient cycling and delivery from forest and mixed land uses, or use of SWMM/EXTRAN or a similar model to determine runoff and sewer flows in urban areas. Key outputs from these simulations can be used to determine representative input values (e.g. delivery factors, BMP efficiencies, etc.) for the Bay model within appropriate landscape regions of the Bay watershed. A landscape classification scheme could be used to regionalize this information into the CBWM from detailed model studies to similar basins (see Winter (2001); Wolock et al. (2004); Brakebill et al. (2000)).

Evaluation of large-scale precipitation pattern estimation: Precipitation intensity and patterns are primary dynamic drivers of watershed hydrology. **The CBM team should assess the spatial pattern estimates of precipitation at the model time step (one hour) up through annual durations, specifically for distributional bias (both spatial and temporal) in selected, representative subbasins.** The regression method of estimating precipitation is an inexact interpolator (it does not reproduce measurements at the gauges). This may have the effect of smoothing precipitation surfaces and alteration (bias) of precipitation frequency distributions. The modeling team should consider choosing a set of precipitation gauges in different hydroclimate settings within the CBW, and compare interpolated and gauged precipitation frequency distributions for bias. **If significant differences in distributions are found, a check for residual propagation could be performed by simulating individual land segments with the two different time series.**

An additional test of the interpolation method can be gained with available, high quality NEXRAD derived precipitation data. Use of this information requires careful adjustment of the backscatter-rainfall (z-r) calibration. Existing 1-km resolution information may be gained from Jim Smith (Princeton) for areas in the Rapidan, Baltimore, and elsewhere. David Legates at University of Delaware may be an additional source.

Sediment and nutrient non-point sources, transport and remobilization: Non-point source loading to small streams and in-channel sediments from land disturbances such as historical agricultural and road building operations, are believed to be a major source of sediments and nutrients. Sediment and associated nutrient loading to these stream channels, and their subsequent contributions to the lower watershed, may arguably constitute the most important opportunities for improving water quality. This concept should be explored on selected sub-watersheds prior to possible incorporation into the full bay model. A full range of conditions should be explored: high and low nutrient areas, urban, agriculture, forest, etc.

Improved simulation of sediment and nutrients may require consideration of additional factors. These include representation of particle size distribution of mobilized and transported sediment, which may be important both in determining sediment loads to the Bay and associated nutrients. Incorporation of a model to better capture these types of sediment balance and dynamics might be considered (e.g. see comments regarding the use of ANSWERS, above).

At the scale of the full CBW, a threshold of 100 cfs as a mean annual flow is used for modeled river reaches. Processes (e.g. erosion, transport, retention) within the lower order streams and valley bottoms are not explicitly modeled but may require treatment by reducing the flow threshold modeled within the CBM or use of an alternative model. This area may be a large source for sediment and nutrients as stored alluvium accumulated in lower order streams over a long period of agricultural land use is scoured by upland generated runoff, particularly in urbanizing areas. This source may persist for an extended period, resulting in significant lags in achievement of sediment reduction targets. The current HSPF version does not simulate bank erosion, which is often the critical sediment source. **We suggest alternative river reach models, such as those developed at the National Sedimentation Lab, be considered.**

BMP dynamic behavior: Currently, BMPs are applied as constant percentages by land use category. It is known that their efficiencies are variable with storm size; this needs to be incorporated into the model. It may be advisable to test this storm-variable BMP effectiveness on selected subwatersheds to better understand their effect. An example of related work can be found in Emerson et al. (2005) which shows that stormwater detention basins designed for 2 - 100 year storms have essentially no impact on watershed-scale peak flow reduction for small storms (< 2 year), where small storms constitute 97% of the annual rainfall in the example application. BMP efficiency/effectiveness as a percentage reduction in load may be hard to defend in a regulatory situation. Additional research is needed to link smaller scale BMPs to large-scale effects. **The CBM team is currently compiling information on dynamic BMP efficiencies, and we encourage this activity as a critical component.**

6. Bigger picture issues and model simplification

The modeling team is in a good position to develop assessments of "emergent behavior" of the CBW suggested by the numerous model runs, sensitivity analyses and scenarios tested, in addition to monitoring data. What are the repeated patterns that are persistent in different runs in terms of dominant controls of CB water quality changes? This requires stepping back from the details of the models and examining and summarizing major model output. Are there dominant processes that can be retained in a simpler model or set of models that can be applied to specific parts of the CBW? Can dominant processes among the different basins in the watershed be regionalized in a way that would point to different management strategies? This may already be forthcoming, but would be useful for a review team or managers to see.

If a set of dominant drivers for the different areas can be determined, the CBM team should assess whether simpler models, based on these dominant drivers, can be produced for the different regions of the CBW. This approach is based on the premise that the same model structure may, in fact, not be suitable for all areas, or that the comprehensiveness and complexity of a fully general model may not allow the use of Monte Carlo methods for formal uncertainty analysis. This recommendation is not based on the assumption that a general model is less physically realistic, but on the assumption that the availability of required data to adequately parameterize such a model is the limiting factor determining model reliability. Therefore, simpler models that can be demonstrated to be applicable, or to yield as high a level of explanation of watershed system response (in this case river discharge, nutrient and sediment concentrations) can be more reliably parameterized and assessed for uncertainty. Note that one of the main advantages of the simpler models is to operate them in parallel with the general model to better assess uncertainty, not necessarily to replace the full CBWM.

7. Concluding Thought

We applaud the Chesapeake Bay Modeling team; their modeling efforts and their openness during our review significantly facilitated our task. The team has accomplished a great deal with models that exceed the scale of any previous work. We believe that their continued modeling activity, in consideration of the recommendations raised in this review, can lead to a modeling-monitoring effort on the Chesapeake that will both effectively guide management and advance the science.

8. References

Beven, K.J. 2001. *Rainfall-Runoff Modeling, The Primer*. John Wiley and Sons, Ltd., Chichester, U.K. 360p.

Beven, K.J., and J. Freer. 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modeling of complex environmental systems, *J. Hydrology* 249:11-29.

Brakebill, J.W., and S.K. Kelley. 2000. Hydrogeomorphic Regions of the Chesapeake Bay. U.S. Geological Survey Open-File Report OFR-00-424, digital data set accessed at http://water.usgs.gov/lookup/getspatial?hgmr

Draper, D., D.P. Gaver, Jr., P.K. Goel, J.B. Greenhouse, L.V. Hedges, C.N. Morris, J.R. Tucker, and C.M. Waternaux. 1992. *Combining Information - Statistical Issues and Opportunities for Research*. Washington, DC: National Academy Press.

Emerson, C.H., C.Welty, and R.G. Traver. 2005. A Watershed-Scale Evaluation of a System of Stormwater Detention Basins. *ASCE J. of Hydrologic Engineering*. 10(3):237-242.

National Research Council (NRC). 2001. *Assessing the TMDL Approach to Water Quality Management.* Committee to Assess the Scientific Basis of the Total Maximum Daily Load Approach to Water Pollution Reduction, Water Science and Technology Board. National Academy of Sciences. Washington, DC.

Reckhow, K.H. 1999. Water quality prediction and probability network models. *Canadian Journal of Fisheries and Aquatic Sciences*. 56:1150-1158.

Tague, C.L., and L.E. Band. 2004. RHESSys: Regional Hydro-Ecologic Simulation System—An object-oriented approach to spatially distributed modeling of carbon, water, and nutrient cycling. *Earth Interactions* 2004(8):1-42.

Winter, T.C. 2001. The concept of hydrologic landscapes. *JAWRA*. 37(2):335-348.

Wolock, D.M., T.C. Winter, and G. McMahon. (in press) Delineation and evaluation of hydrologic landscape regions of the United States using geographic information systems tools and multivariate statistical analyses. *J. Environmental Management.* http://water.usgs.gov/lookup/getspatial?hlrus